Generation extraction and characterisation on brain organoid culture images with small datasets constraints.

Clara Brémond Martin

PhD in Sciences and Technologies of Information and Communication

ED EM2PSI - Economy, Management, Mathematics, Physic and Informatic Sciences CY
Cergy Paris University

Manuscript

Generation, Extraction and Characterisation on Brain Organoid Culture Images With Small Dataset Constraints

---

by

Clara Brémond Martin

ETIS Laboratory UMR 8051 (CY Cergy Paris University, ENSEA, CNRS)
6 avenue du Ponceau, 95014 Cergy-Pontoise Cedex, France

Witsee 33 Av. des Champs-Élysées, 75008 Paris, France

Defense on June the 22th, 2022

With the jury composed of :

Pr	Caroline Petitjean	LITIS, University of Rouen Normandie	Rapporteure
Dr	Yannick Benezeth	ImViA, University of Bourgogne Franche-Comté	Rapporteur
Dr	Pierre Jacob	LaBRI, University of Bordeaux	Examiner
Pr	Emmanuel Pauthe	ERRMECe, CY Cergy Paris University	Examiner and President
Pr	Aymeric Histace	ENSEA, CY Cergy Paris University	Thesis Director
Dr	Camille Simon Chane	ENSEA, CY Cergy Paris University	Academic advisor
Dr	Cédric Clouchoux	Witsee, Neoxia	Industrial advisor

Short abstract

Brain organoids (BO) are promising recent brain models, but tools are lacking to follow their development, with very few images available. This thesis aims to automate the extraction and appropriately characterize their morphologies on small datasets.

A small dataset was augmented by various generative adversarial networks (GANs) and the best one was optimized. Most optimisations are suitable to enrich the dataset. The effect of a noise variation by GAN to generate images closer to the exceptions has been studied, and these images have also been validated automatically, statistically and psychovisually.

The shape of the original BOs was then extracted by deep learning and a better accuracy was obtained if the training is carried out by an optimized GAN. The architecture of the extraction has been reduced to adapt it to small datasets.

Thanks to the segmented images, morphology of the organoids was characterized by calculating morphological indices, but also by finding their imprint by analyzing topological data. Characterizing an index of their development (neuroepithelia) in association with another laboratory would be a future step, such as validating this approach on other brain cultures and for various microscopic acquisition methods.

Résumé

Les organoides cérébraux (OC) sont des modèles récents et prometteurs pour le suivi du développement cérébral. Cependant, les biologistes manquent d'outils d'analyse automatique ou semi-automatique pour suivre leur développement in vitro, du fait de la faible quantité d'images disponibles actuellement. Dans ce contexte, cette thèse vise à automatiser l'extraction et à caractériser de façon adaptée la morphologie des organoïdes cérébraux sur de petits datasets d'images acquises en microscopie à divers stades de la croissance de ces cultures.

Dans cet objectif, un petit dataset a été augmenté par divers réseaux adverses génératifs (GAN) et le meilleur d'entre eux au regard de critères psychovisuels, statistiques et, quantitatifs a été optimisé. Quel que soit l'optimisation utilisée pour générer des images, elle enrichit le dataset de manière constructive, excepté pour un groupe. De manière complémentaire, l'effet d'une variation du bruit par GAN pour générer des images plus proches des exceptions a été étudié, et les images générées ont été validées également automatiquement, statistiquement et psychovisuellement.

Dans la continuité de cette première contribution, la forme des OC originels a été extraite par deep learning et nous montrons qu'une meilleure précision peut être obtenue si l'entraînement est réalisé par un GAN optimisé. L'architecture de l'extraction a été réduite pour l'adapter aux petits datasets.

Enfin, grâce aux images segmentées, la morphologie des OC a pu être caractérisée par calcul d'indices morphologiques, mais également en déterminant leur empreinte par analyse de données topologiques. Caractériser un indice de leur développement (neuroépithéliums) en association avec un autre laboratoire serait une étape future comme la validation de cette approche sur d'autres cultures cérébrales et pour diverses modalités d'acquisition microscopiques.

Abbreviations

AAE: Adversarial autoencoder
BCE: Binary cross entropy
BCEL1: Binary cross entropy with a L1 normalisation
BO: Brain organoids
cGAN: conditional generative adversarial network
CM: center of mass
CNN: Convolutional neural network
CO: Cerebral organoids
CP: cortical plate
DL: deep learning
ELISA: enzyme-linked immunosorbent assay
FCN: Fully convolutional neural network
FID: Frechet inception distance
FGF: Fibroblast growth factors
FN: false negative
FP: false positive
GAN: Generative adversarial network
GT: Ground truth
hESCs: human embryonic stem cells
HIAS: high-refractive index aqueous solutions
HS: hyperhydrating solutions
iPSCs: induced pluripotent stem cells
KL: Kullback-Leibler
L1: mean absolute error
L2: mean square error
LS: Least square
MEA: micro-electrode array
MI: mutual information
ML: machine learning
MNIST: mixed national institute of standards and technology
MSE: mean square error
OS: organic solvent
PCA: principal component analysis
PSNR: peak signal to noise ratio
P. Wass: Perceptual Wasserstein
RNAseq: ribonucleic acid sequencing
ROI: regions of interest
RTqPCR: quantitative retrotranscriptase-polymerase chain reaction
SSIM: Structural similarity
SSH: Sonic hedgehog protein
SVM: Support vector machine
SVZ: subventricular zone
TDA: topological data analysis
t-SNE: t-distributed stochastic neighbor embedding
TT: tissue transformation
TN: true negative
TP: true positive

UQI: Universal quality index
VAE: Variational autoencoder
VZ: ventricular zone
WGAN: Wasserstein generative adversarial network
WNT: Wnt protein

Acknowledgements

La thèse solitaire n'existe pas, comme sur les sentiers on rencontre des repères pour ne pas se perdre en chemin.

Je voulais ici remercier ici ceux qui ont été les repères de ma thèse.s Tout d'abord je voulais remercier les membres du jury : à Caroline Petitjean rapportrice, à Yannick Benezeth rapporteur et, à Emmanuel Pauthe et Pierre Jacob qui ont accepté d'évaluer mon travail. Je voulais aussi remercier les membres de mes comités de mi-parcours également qui ont su m'indiquer de nouvelles pistes un nouveau merci à Caroline Petitjean et Fanny Mann.

Puis j'en viens à mes encadrants Camille Simon-Châne et Cédric Clouchoux et mon directeur de thèse Aymeric Histace. Tous les trois vous avez su me donner de précieux conseils sur mon sujet de thèse, sur des méthodes, des axes de recherche, ou encore sur les figures en latex... Vous m'avez apporté vos différents points de vue tant académique qu'industriel, mais aussi beaucoup de soutien dans mes diverses démarches : un très très très (je voudrais écrire pleins de très pour à chaque fois que vous m'avez aidé mais ça ne rentrerait pas) grand merci.

Je voulais remercier les membres du laboratoire ERRMECe pour avoir accepté de participer aux études psychovisuelles, Michele Bertacchi de l'institut Valrose et, Madeline Lancaster et Ilaria Chiaradia du Laboratory of Molecular Biology pour m'avoir transmis de nouvelles données.

Je voulais remercier l'Anrt qui a attendu un 1er avril en pleine pandémie pour me laisser entrer en thèse (la plus belle blague de ma vie). Mais surtout l'équipe Cell et les membres du laboratoire Etis et, l'équipe Witsee (ancienne aisee-Mvp-3dns) de l'entreprise Neoxia pour m'avoir apporté un cadre bienveillant pour faire ma thèse. Je voulais dire d'abord un merci à la direction de Neoxia (Gilles Mergoil) et Etis (Olivier Romain) pour m'avoir laisser entrer en thèse. J'aimerai remercier l'ancienne équipe (3dns-MVP-aisee et de Witsee) avec qui j'ai pu travailler deux ans avant de commencer cette thèse. Toutes les personnes que j'ai pu y rencontrer et qui sont parties, celles qui sont encore là et avec qui j'ai pu apprendre et évoluer.

Parmi les personnes du laboratoire Etis qui m'ont soutenu je voulais en remercier plus particulièrement certaines sans qui cette thèse n'aurait certainement pas été possible (dans tous les sens du terme).

Les premiers que je dois remercier sont Mohammed Amine Khelif et Pierre Jacob pour m'avoir déjà expliqué qu'on ne vouvoyait pas les doctorants et, présenté le club des doctorants qui a été un lien social durant toute la période pandémie. Puis mes co-responsables du club des doctorants, philosophes de tous les jours, Théo Serru et Louis Annabi : je ne pensais pas avoir un jour des amis avec qui pouvoir argumenter de tout, mais vous êtes là tous les deux alors merci pour ces trois années près de vous, sans presque jamais vous avoir vu, et qui m'ont tant apporté. Louis, tu nous as bien devancé depuis la fin de ta thèse, et je suis toujours autant admirative des conseils et de ta façon de percevoir les choses, alors merci pour tout et ne change rien (je vois bien la petite chèvre s'afficher). Théo, tu seras le dernier mousquetaire qui devra soutenir, je compte bien être là comme tu l'as été là pour nous, et à te redonner le sourire que tu m'as tant prêté quelque soit la situation (et cela quitte à te faire un gâteau

à chaque fois). Laure Acin et Claire Béranger, j'ai été bien plus heureuse à partir du moment où vous êtes arrivées toutes les deux (enfin des doctorantes dans l'équipe!). N'oubliez pas de rire même dans les moments où on a envie de pleurer et, de continuer à manger des goûters toutes les deux heures pour nourrir vos neurones qui travaillent! Je voulais te remercier avec les filles, mais j'ai évité de justesse, Alexandre Bordat, pour avoir avec Laure repris la gestion du club (on sait la charge qu'on vous a laissé...), participé à animer la vie des doctorants, organisé tout sous forme de listes interminables, ne jamais finir tes phrases et... pour avoir voulu me prêter un déambulateur ces derniers mois. Pour avoir été le relais des doctorants du laboratoire Etis entre Saint-Martin et l'Ensea, m'avoir entrainée sur la glace pour penser à autre chose que l'écriture et la fin de ma thèse, jouté avec moi au badminton ou hockey, et pour être d'une extrême gentillesse et aide à toute épreuve avec toutes et tous les doctorants, je voulais remercier Sylvain Colomer. Mes anciens et, nouveaux colocataires de bureau : Louis Desportes, l'homme qui murmure aux oreilles des canards de debuggages de code et, Etienne Vareilles notre dessinateur de bds de la vie au labo des vendredi aprem. Un bureau sans vous deux cela n'est pas le bureau 360. Puis tous les doctorants qui ont participé aux journal club, à la vie scientifique et des rabbit signal du 4h que je n'ai pas encore pu citer: Tarek Elouaret (et dire qu'on aurait pu soutenir le même jour), Arnau Dillen (toujours le sourire, ne change pas), Juliette Courson et Pauline Rousseau (accrochez vous toutes les deux)... J'aimerai citer tous les doctorants là mais je crois que je ne pourrai pas sans dépasser la page. Je voulais remercier aussi les permanents du laboratoire que je n'ai pas encore mentionné de nous avoir souvent aiguillé et, participé aux débats avec nous (Emmanuelle, Inbar, Laurent, Son, Guillaume,...).

Enfin je voulais remercier ma famille qui m'a suivi de très loin (900km). Je n'ai pas été là pour vous ces quelques années alors que j'aurais tant aimé pouvoir aider plus à la maison. Merci maman et papa de m'avoir permis de faire toutes mes études, de m'avoir appelée au téléphone quasiment tous les soirs à 20h... Merci à Chris et Coraline, pour les fous rires, les sorties entre frères et soeurs, les fruits et légumes, j'aurais aimé partager plus de moments avec vous et je sais combien mes études me font partir bien loin. Merci à mes deux grands-mères et mon grand-père pour le soutien par machine à coudre ou madeleine. A Viviane pour avoir été là quelques soirées au téléphone, initié quelques sorties sur Paris ou dans notre sud et, permis de beaux voyages. Un merci à mes amis d'enfance qui ont tenté de garder contact en m'envoyant un message tous les 1ers du mois, Morgane, Antonin, promis je me rattrape dès que j'ai fini la thèse! Merci à Laurent, tu as brillé par ton absence ces trois années, mais tu as apporté des étoiles à mon chemin. Un merci spécial à Rumbo, promis tu n'as plus besoin de macher les cables de l'ordinateur pour que je joue plus avec toi, à Pumba pour s'escaper et courir avec moi dans les collines, aux chemins de randonnées non indiqués, à la patinoire de Cergy, aux auteurs présents dans ma bibliothèque et aux paysages enneigés, d'automne, désertiques ou fleuris sur nuvole bianche.

Un merci enfin à ceux que j'ai perdu en chemin, à chaque pas que je fais je pense à vous.

Contents

Chapter 1

Introduction

The Human brain is one of the most complex organ to study in terms of functions and structures. Thus, researchers in Neuroscience use models to study and attempt to understand our brain such as animals models (in vivo), post-mortem brain slices (ex vivo) and cellular cultures (in vitro) (Kelava et al., 2016a). However, each one of these models have limitations: In vivo, the cell diversity is different between the Human and animal models, and some cognitive processes are peculiar to Humans; Ex vivo, ethical considerations limit some studies when the quick cell death does not allow it (Arlotta et al., 2019); In vitro, 2D cell cultures do not reproduce the brain architecture, and suffer from the lack of connections between cells (Poli et al., 2019).

A promising alternative lies on brain cultures. To understand our brain development or disease, different research teams have been trying to cultivate brain cultures for over a century (Madeline A Lancaster and Knoblich, 2014). In 2013, this goal was reached by the Lancaster team: Brain Organoids, better known as Minibrains, became a reality (Madeline A. Lancaster et al., 2013). Brain organoids are 3D brain-like structures mimicking some brain regions measuring around 4 mm of diameter and suffering from Batch syndrome: despite being grown in the same environment, they develop differently. For example instead of developing two lateral ventricles, one can innately develop none, although its neighbor can have ten.

The lack of knowledge about the physiological development of brain organoids, precludes biologists to use them as alternative solutions to others models (animal, post-mortem). To accomplish this, morphology is analyzed by acquiring images by various methods (Brémond Martin et al., 2021b). However, at this moment, no adapted tool exists to extract and characterize the complex shape of cerebral organoids automatically.

In the literature, solutions are often constituted with software assembly, or implies the use of pre-processing step by a human operator to segment (extract the shape from the background inside the image) or characterize (calculate some morphological indexes on these shape). The machine learning solutions scarcely developed cannot been applied on other datasets. The only automatic solution dedicated to extract the shape of brain organoids concerns only ventricular cavities and uses deep learning (Albanese et al., 2020). However, deep learning strategies requires an amount of data to their training phase, while only a few image datasets are available, consisting of a couple of images. My PhD aims to address the following needs :

- The automation of processing step on image analysis for brain organoids.
- A dedicated shape extraction tool for the overall morphology of brain organoids and facilitate the following of their growth.

- An increase of the amount of data in tiny datasets in order to use deep learning strategies.
- Morphological characterisation tools dedicated to the complex shape of brain organoids.

1.1 Objectives

This thesis aims at presenting an insightful approach to extract and characterize brain organoid culture images. This work is centered on the automatic shape extraction and characterisation, with tiny dataset constraint, which is an issue for both the industrial and academic point of view.

Neoxia, the project's partner company, has a strong expertise in process industrialisation from academic projets. Recently, they created a spin-off company, Witsee, from a consortium between academics and industrial partners during a "Projet d'investissement d'Avenir" (PIA2). The proof of concept was based on preclininal research, including neurological (Alzheimer) and oncological (cancer) disease using various models (animal, 3D culture). The output of this project was the creation of a collaborative platform of large biomedical data processing, including microscopic images processing and analysis. Similarly to academic laboratories, Witsee runs into the lack of large available datasets issue. To solve this problem, data augmentation strategies are used so deep learning methods can be implemented (Hamdi, Bouvier, Delzescaux, et al., 2021; Hamdi, Bouvier, Delzesceaux, et al., 2021). In the context of collaborative projects, the platform developed by Witsee relies on classical morphological tools to help customers characterize the shape of biological contents in images. However, morphological indexes (described in Section 2) are not adapted to the particular growing shape of brain organoids.

The CELL team from ETIS laboratory, UMR 8051 (CY Cergy Paris University, ENSEA, CNRS), is specialized in smart embedded systems for bio-medical applications. This team has a strong expertise in automated learning in constraint spaces (limited database, embedded calculation...) specially for real-time detection of polypes in video-coloscopy (Bernal et al., 2017; Leenhardt et al., 2019), a project financed by the PIA2 (Initiative d'Excellence Paris Seine, I-Site 2017) and by the SATT IdfInnov (creation of the startup "Augmented-Endoscopy"). The lack of data and the variability of biomedical tiny datasets force them to use machine learning strategies before deploying deep based systems in capsule image endoscopy (Houdeville et al., 2022; Romain et al., 2022). This team also brings solutions in optimisations for signature representations in deep learning strategies (P. Jacob et al., 2019), and other solutions focus on the spectral analysis of cell cultures (De Roux, Terosiet, Kolbl, et al., 2017; De Roux, Terosiet, Kölbl, et al., 2017).

In this thesis, we pave the way to analyse tiny datasets of brain organoid images, in an attempt to answer the following points (also see Figure 1.1):

- Generating images: in a pandemic situation, acquiring our own brain organoid images at different developmental stages became a real issue. Another issue is to find a database with enough images to train deep-learning based solutions. In such tiny databases constraint, synthetic learning images must be generated. The choice of the type of image generation and the kind of architecture and optimisation needed to get the most qualitative images, as well as similar to the original dataset. Does a comparative automatic validation of the generated images corroborate with a human expert validation? Can we use these results to construct a new automatic tool which could replace Human evaluation of generated images and gain time?

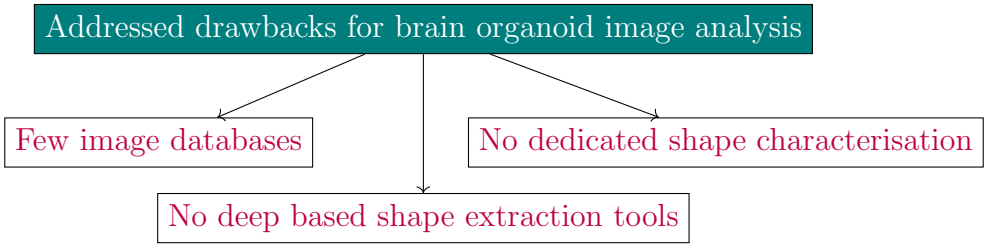


Figure 1.1: Summary of the three drawbacks we attempt to answer in this thesis

- Segmentation of brain organoid images: this topic raises the question of using a machine learning or a deep-based strategy. While machine learning solutions are not suitable for all the datasets, in the second case the requirement of amount of images is a break. Which strategy will best fit the segmentation of brain organoids? How to validate a small segmentation architecture designed for tiny datasets? Another question raised linked to the first item is: is a validated data augmentation strategy more useful to improve a deep learning segmentation than no validated data augmentation? Or classical machine learning segmentation strategies are they sufficient to extract the shape of brain organoids tiny datasets?

- Morphological characterisation of particular developmental or pathological shape: does a classical morphological index characterize the best the biological state or developmental stage of a brain organoid culture? Are new topological methodologies solutions to accurately classify BO according to developmental stage, or physiopathological state? Are the neuroepithelial formations detected with such tools? If some topological or morphological indexes are given during the segmentation training, does this improve its accuracy?

1.2 Development and contributions

Development and contributions are summarized in Figure 1.2. The first months of the thesis were dedicated to a literature review on the image acquisition and analysis of brain organoids cultures, presented and updated in section 2 The lack of appropriate morphological analysis tools is depicted in this review and used by other researcher in the field. This work helps identify all available brain organoid databases. Our work and that of Chakradhar, 2016 shows the difficulties to develop this kind of culture in normal and more specifically in pandemic restrictions with no laboratory access.

The access of databases containing a certain amount of images could help to develop automated image analysis tools.

Among the few database available, we needed to choose one allowing to study the development of brain organoid, and thus with an imaging strategy bearing it. Once we select a bright-field brain organoid image database from Gomez-Giro et al., 2019, the augmentation of number of images is a requirement, to use and develop appropriate deep based segmentation tools. In Brémond Martin et al., 2021a, we compare various generative architectures to classical transformation of brain organoid dataset, we verify the effect of loss optimisation thanks to automated metrics and observe the contribution of each loss in both clustering and in segmentation training stages. This original contribution highlights the normalised Wasserstein loss, originally proposed in this study, which outperforms other losses. As the metric validation is still contentious for generative evaluation, we compare it with a Human

validation and describe possible alternative metric combinations, for the first time to our knowledge (article under submission). The images generated are in the same statistical space than the original ones, except for some with a particular gradient background. Thus we study the effect of adding different noises inside the generative process and proceed as for the loss optimisations. We do not highlight a particular noise to naturalize the generation in Brémond Martin et al., 2022 and in Section 3.6 (unpublished).

Regarding organoid image segmentation methods, we first verified the contribution of each loss optimized generated image on the segmentation accuracy by using U-Net (Brémond Martin et al., 2021a). We then checked if a training performed using psycho-visually validated generated images could help to increase the accuracy of the segmentation, see Section 4.2 (original research, not published). Due to time and huge computational requirements, we search to implement a small architecture with few input image requirements, which has never been implemented before, and compared it with two others architectures. This work was presented at the "Towards pragmatic learning in a context of limited labeled visual data" GDR-ISIS communication.

We also produce an ablation study, which consists in suppressing parts from the original segmentation architecture one by one obtaining a new one. This study contributes to validate this latest architecture as the best one to segment bright-field brain organoid images (see Section 4.3). On this small segmentation architecture, we verify the contribution of all the optimized loss and the adding of psychovisual validated images to the training step and discuss the results compared to machine learning approach for segmentation.

Once the organoid shape is extracted, we perform morphological characterisation. We first calculate morphological indexes documented in the literature and summarized in Brémond Martin et al., 2021b. Then we implement a topological data analysis tool (increasingly used in the biomedical field), and mix it with a clustering approach in order to classify the organoids developmental stage (Brémond-Martin et al., 2022).

We also characterize their state (physiological or pathological) (see Section 5.4) and, the cross-factor analysis of their developmental stage according to their state (see Section 5.5), which are not published parts. We compare morphological indexes clustering with TDA, to determine the most fitted analysis for brain organoids characterisation.

We finally create a pipeline dedicated to brain organoid bright-field images with the best optimisations and the best architectures for each kind of the three steps summarized in Figure 1.3. This pipeline is validated for other datasets for the steps highlighted.

1.3 Outline

The manuscript is written in the five subsequent chapters summarized in Figure 1.2:

- The Chapter 2 describes the recent trends and perspectives for the brain organoid imagings.
- The Chapter 3 presents a brief review of data augmentation strategies for biomedical images, and the different architectures, loss and noise optimisations we use, along with validation. In this part we also describe evaluation of synthetic images, and discuss the results.
- The Chapter 4 focuses on segmentation: we present the different algorithms used in the biomedical field, and describe the architecture we choose and its drawback for tiny datasets. We describe then an ablation study and the Mu-Net architecture we create.

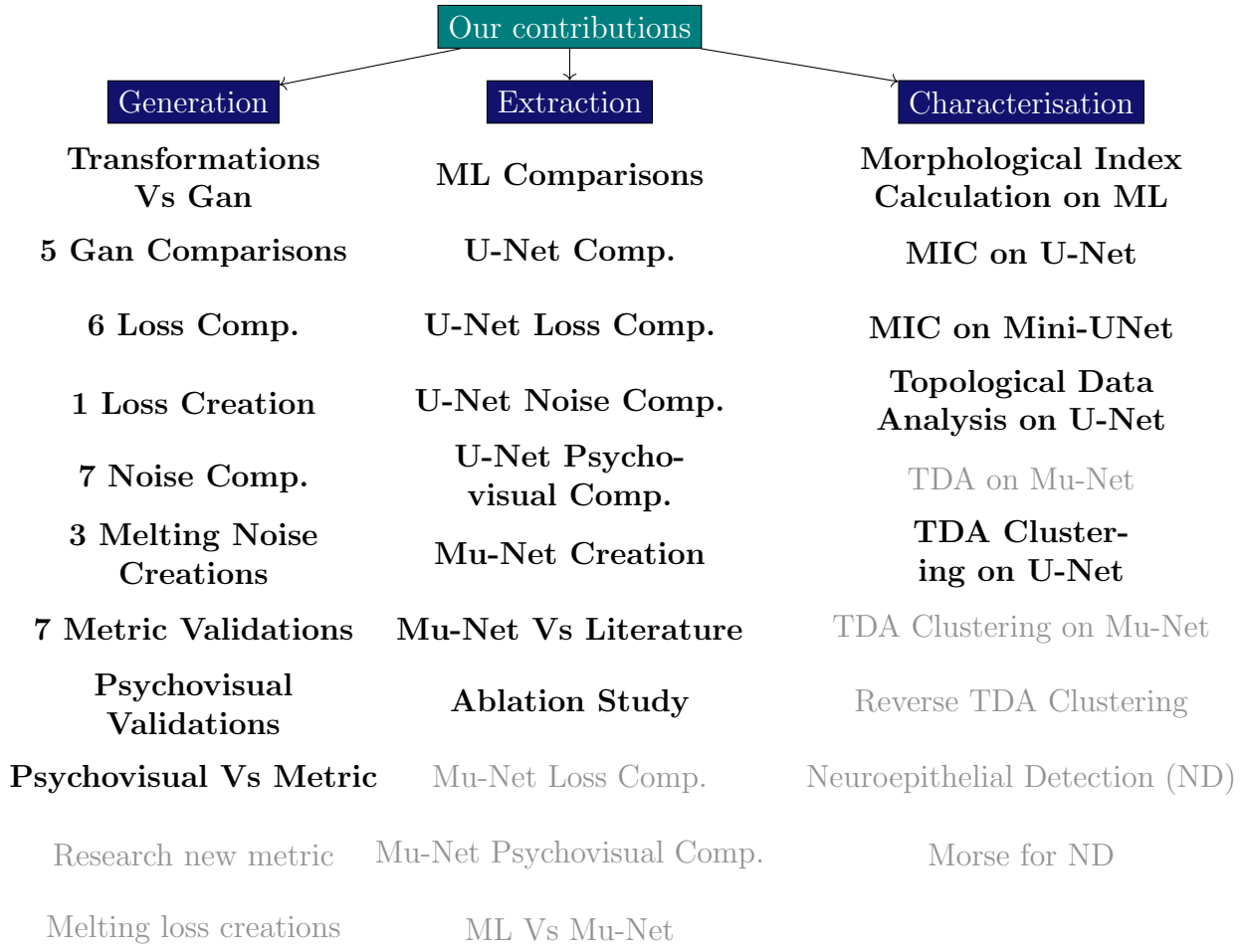


Figure 1.2: Summary of developed contributions on brain organoid image analysis. In bold: completed work; in lightgray: in progress.

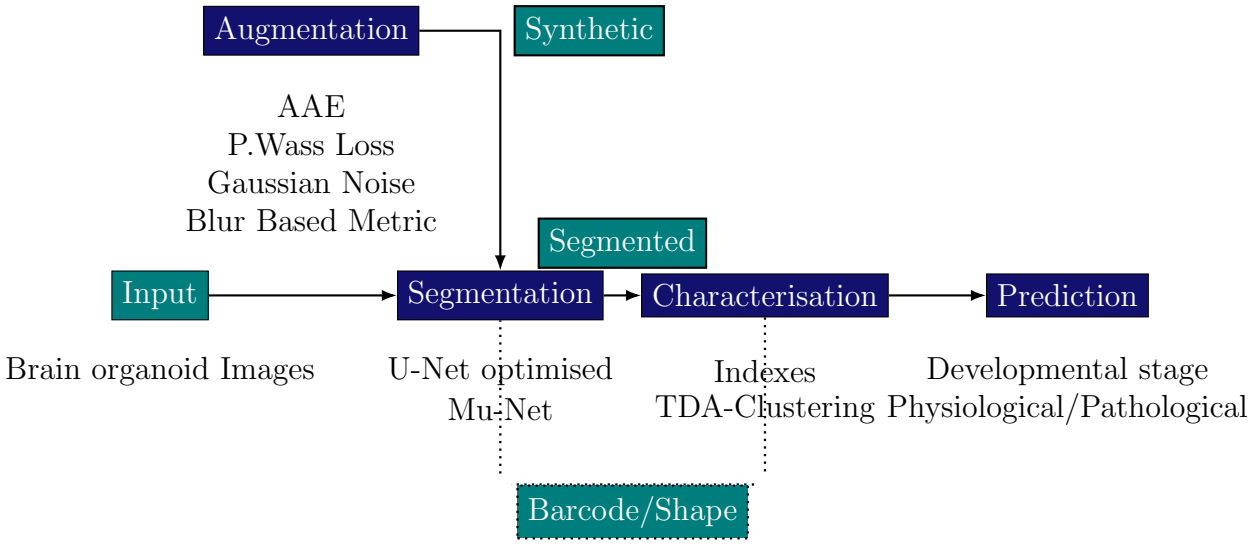


Figure 1.3: Pipeline for brain organoid image analysis. In blue: the process, in green: the image state, the dotted feedback line corresponds to futur implementations.

We show the inflow of generated images validated by metrics of psychovisual evaluation on the segmentation accuracy on our new architecture compared to previously described approaches.

- The Chapter 5 concerns shape characterisation, based on morphological indexes and on topological data analysis. We show the results obtained using our TDA clustering, reverse it on original images, and compare it with the characterisation using the classical morphological indexes.
- The last Chapter 6 summarizes our work and the perspectives on brain organoid tiny datasets analysis.

Chapter 2

State of the art

This section is the update of our review paper published in July 2021 at Frontiers in Neuroscience (Brémond Martin et al., 2021b).

2.1 Historical context

Experimental cerebral models are used to observe and analyze structure and function, both of which are complex to identify in human brain tissues (Stan et al., 2006). These models are often classified in three categories: in vivo; post-mortem; in vitro. However, in vivo and post-mortem brain animal models are often prone to controversy due to ethical considerations added to technical impairments due to divergences with the Human brain structures (Lodato et al., 2015; Kelava et al., 2016b). Key benefits of in vitro models are these cultures can be derivatives from human cells on the one hand, and on the other hand, be more relevant to replicate its physiology. Despite this great asset, standard 2D neuronal cultures lack of tissue structures, diversity of self-patterning cells and some disease patterns, presenting then with strong limitations for in vitro study. Three-dimensional (3D) brain cultures (Kapalczynska et al., 2016; Bolognin et al., 2019; Cederquist et al., 2019) have become in the last years a very promising alternative to overcome these limitations.

In this context, recently, cerebral organoids (CO) have emerged by the differentiation of induced pluripotent stem cells (iPSCs), or human embryonic stem cells (hESCs) (Madeline A. Lancaster et al., 2013). Such 3D cultures are no larger than 4 mm in diameter and they develop some structures similar to those developed by the brain during the second semester at numerous random locations (Kelava et al., 2016a). To study these cerebral organoids, researchers use methods originally developed to analyze other post-mortem and in vitro models: enzyme-linked immunosorbent assay (ELISA (Raja et al., 2016)), quantitative retrotranscriptase-polymerase chain reaction (RTqPCR (Sakaguchi, Taisuke Kadoshima, et al., 2015)), ribonucleic acid sequencing (RNAseq (Quadrato et al., 2017)), micro-electrode array (MEA (Monzel et al., 2017)) and others techniques focused on for example proteins or metabolites. Because these techniques can lead to complex and costly experimental set-up, in addition to them, imaging techniques are now used in almost every study focusing on cerebral organoids both to complete and to validate other molecular analysis. It can also be used to observe features that are unavailable with other methods, for example to quantify the growth of such cerebral organoids (Iefremova et al., 2017).

The commercialisation of cerebral organoids since 2016 (Chakradhar, 2016) has resulted in the widespread generalisation of their use by laboratories (see Figure Consequently, the microscope technique, analysis methods and tools must be tailored for the issue at hand.

Given this increase and the importance of the image analysis in this field, it has become essential to identify the methods employed to study cerebral organoids, as well as the improvements that can be performed and the challenges that need to be overcome to handle image analysis on cerebral organoids at a larger scale.

2.2 Scope and positioning

This review summarizes the recent advances in 3D brain cultures imaging and analysis, and particularly for cerebral organoids. We performed statistical analysis on the 457 articles on cerebral organoids referenced by Pubmed between January 2018 and June 2020. We chose to perform this review study starting from 2018 because the number of articles per year was less than 100 before this date. Of note, 670 articles on key words "cerebral organoids" have been published since 2013 according to Pubmed. Among these 457 articles, 63 mentioned these key words but are not on this topic, and 46 % of the remaining articles are reviews.

Most of these reviews focus on brain diseases (Donegan et al., 2020), comparison of cultures (Chhibber et al., 2020), ethical considerations (Bayne et al., 2020), challenges in vascularisation (Hong et al., 2019) and connectivity between cells (Marton and Pașca, 2020). A few reviews focus on evolution (Stefano L Giandomenico et al., 2017), biobanking (S. Li et al., 2020), or summarizing a single article (Madeline A. Lancaster, 2018). This includes a review on the possible emergence of cerebral organoids connected to other organ models (Chukwurah et al., 2019).

Most of these reviews addressed brain diseases, cultures comparisons including a review on the possible emergence of cerebral organoids connected to other organ models (Chukwurah et al., 2019), and development Figure Less than 3% of the reviews addressed 3D brain cultures images analysis. Among them, only three about image analysis applied to cerebral organoids data have been published. (Poli et al., 2019) reviewed computational models of formation and organisation of these cultures, and also reviewed protocols and other experimental methods (in electrophysiologic field) applied on cerebral organoids. For these authors, even if cerebral organoids are promising in terms of in vitro models of human brain, the generation protocols and procedures characterisation still need refinement. (Booij et al., 2019) analyzed imaging techniques, image analysis methods and high-content images in 3D cultures but not particularly focused on cerebral organoid cultures. They concluded on the requirement to "validate these technologies and to demonstrate clearly that using biologically relevant in vitro systems actually improves the efficiency of early drug discovery. A direct comparison of the predictive value of 2D and 3D models for in vivo efficacy is required." (Grenier et al., 2020) mentioned in a diagram the perspective of generating a high-throughtput platform for drug testing including image analysis on cleared cerebral organoids with deep learning to identify functional and architectural markers. The authors also discussed the challenges allowing integration of additional variables and risk factors (toxic agents, vasculature) in order to make cerebral organoids a formidable and scalable system to improve our understanding, provide precision to diagnostic and prognostic predictions and personalize drug discovery efforts for neurodegenerative diseases. Of note, in another field, (M. E. Boutin et al., 2019) studied retinal organoids to summarize perspectives on drug testing. One of their expectations was also to apply machine learning on both high-content cell imaging and others chemical methods for their retinal model. They expected work was "being done to apply machine learning approaches to score and predict control versus disease phenotypes from cell imaging assays, including work on photoreceptor outer segment formation. Most of this work has so far been done in 2D systems, and the

hope is that with the development of techniques that allow HT cell imaging in 3D, those will be applied to this more complex systems."

In the time range considered for our paper, we did not find any review focused on image analysis tools dedicated to cerebral organoids. However, a very recent study was published by Albanese et al., 2020 (December), creating for the first time a pipeline named Scout including deep-learning methods to segment the ventricular zone of 3D images of cleared cerebral organoids. They gave a first attempt to an holistic approach to characterize the content and structure of cerebral organoids in 3D.

The current review focuses on the recent trends in acquisition and image analysis methods on cerebral organoids to highlight the specific needs of the field. For all 214 included articles published on cerebral organoids, between January 2018 and July 2020, we identified: their scope; the kind of organoid generated; the acquisition method of images presented in the figures; the analysis methods used specifically, the software and algorithms developed or used; and finally advantages and limitations of the proposed approaches.

The following section gives an overview of the emergence of 3D brain cultures. Then we describe the sample preparation and the image acquisition methods. Three-dimensional imaging is particularly detailed in this paragraph because it captures better the shape and allows quantification for the full brain culture. In the third section, image analysis methodologies are described in two parts: quantification and morphological analysis. Software used to that aim are presented in the fourth section. They remains for the most part semi-automatic due to the recent generation of this model. Following this methodology section, we discuss the pros and cons of each described method, as well as the potentially insightful image analysis tools to implement in order to handle the increasing amount of generated data.

2.3 Three dimensional brain cultures

Advent of cerebral organoids

One of the first 2D neural models were neural rosettes stabilized from iPSCs or hESCs (Chambers et al., 2009). These rosettes are structures composed by neural cells surrounding a lumen. Cerebral organoids were derived from these rosettes.

Over the past ten years, a considerable increase in the use of 3D cultures has been observed. Figure 2.1 shows an exponential growth in the number of articles citing spheroids, organoids and cerebral organoids.

Between 2013 and June 2020, 671 out of 4509 published articles on organoid cultures were treating about cerebral organoids. Before explaining how imaging cerebral organoids, we summarized in this section what are cerebral organoids and how their generation has evolved in the last decade.

Organoids mimic organs: they contain multiple organ-specific cell types, are spatially organized, and simulate organ-specific functions (M. A. Lancaster et al., 2014).

The first 3D neural organoid was a self-organized optic cup made of retinal epithelium (Eiraku et al., 2011). Two years later, (T. Kadoshima et al., 2013) created guided forebrain organoids and (Madeline A. Lancaster et al., 2013) the first self-patterned cerebral organoids. These organoids replicate human fetal brain growth during the second semester (Kelava et al., 2016a). The discrepancy between these two cultures is mainly due to the growth pattern and both methods are currently used for cerebral organoid generation.

(Paşca et al., 2015) created cortical spheroids, also called dorsal forebrain organoid (Arlotta et al., 2019), an assembly of differentiated cells producing deep and superficial

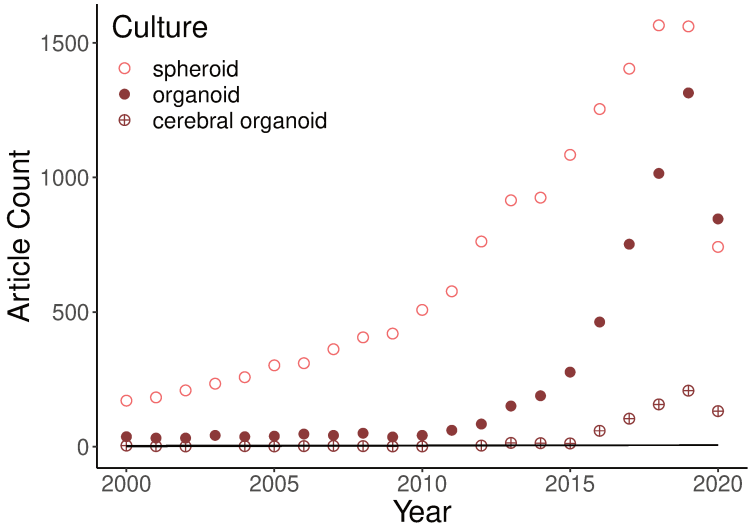


Figure 2.1: Graphic representation of the soaring of three-dimensional cultures based on a pubmed search of the following keywords: "spheroid", "organoid", and "cerebral organoid". Cerebral organoid articles are a subset of the keyword "organoid" research. The expanding of published articles is explained by an exponential model at 89 % (Rsquared: 0.8876: Growth Model = log(Count) Year:Culture). The first generation of cerebral organoids was in 2013, so the previous few articles identified by pubmed contain the two key words but do not talk about these 3D brain cultures inside the body of the text. The points in 2020 are not on the curves due to the fact that the year was not over the time of counting.

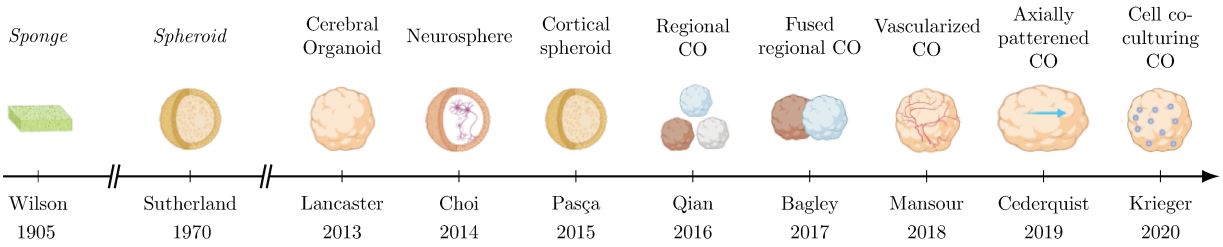


Figure 2.2: Evolution of 3D brain cultures over time. Non-brain cultures which led the way are labeled in italics. The abbreviation CO is used for "cerebral organoid".

layers around ventricular zones. Then, 3D bio-printing bioreactors allowing the generation of cerebral region-specific organoids (forebrain, midbrain and hyppocampic) have emerged (Qian, H. N. Nguyen, et al., 2016). While these region-specific have been created, some authors proposed to fuse them to reproduce the connectivity observed between structures in the human brain (Birey et al., 2017; Bagley et al., 2017). One of the remaining weaknesses of this system is the absence of vasculature, later (Mansour et al., 2018) transplant cerebral organoids inside in vivo model to vascularize the culture. Others teams observed that human organoid transplantation inside injured in vivo mice brains helped lost functions recovering (S.-N. Wang et al., 2019). Nevertheless, the inter-organoid heterogeneity and their cell diversity, failing to reproduce the topological organisation of the human brain, conduct others authors to axially pattern cerebral organoids as occurring during the fetal growth (Cederquist et al., 2019).Only recently, cerebral organoids have been co-cultured with others cell type (tumoral for example), to model disease progression (Krieger et al., 2020). Figure 2.2 summarizes the evolution of 3D cultures from sponges to modern cerebral organoids.

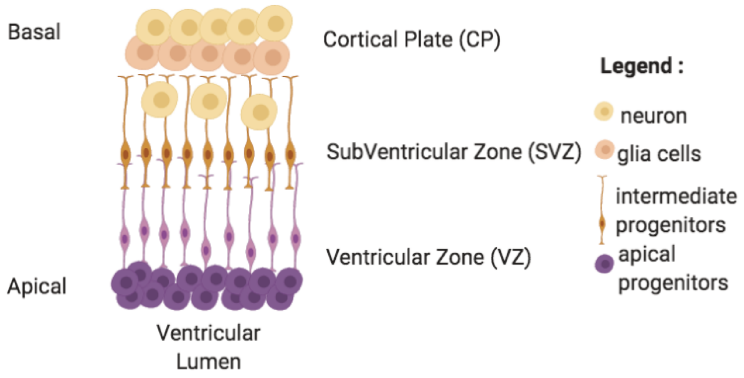


Figure 2.3: Neuroepithelium formation inside cerebral organoids. This formation is present at discrete random locations around lumen ventricles in cerebral organoids. It grows from the apical perimeter thanks to progenitor cells to the basal zone. Cells migrate and differentiate along these axes. Cerebral organoids' neuroepithelium are made of 3 zones: a ventricular zone (VZ), a subventricular zone (SVZ) and a cortical plate (CP) each composed of specific cell types.

Variability in 3D brain cultures

The importance of imaging cerebral organoids is linked to their particular constitution. The cyto-architectural complexity of cultures mimicking brain formation (see Figure 2.2) greatly depends on the culture protocol (Sidhaye et al., 2020). Cerebral organoids containing self-patterned regions are larger and more complex than cortical spheroids showing rosette patterns. In turn, these are more complex than an assembly of different cell-types in a neurosphere (Kelava et al., 2016a). However, differentiating a regional cerebral organoid (i.e. dorsal or ventral forebrain) is more tedious than letting a cerebral organoid self pattern, as such differentiation requires various factors additions to the media at specific times (Madeline A. Lancaster et al., 2013; Bagley et al., 2017).

During the cerebral organoid generation process, model complexity increases with time. First, iPSCs are derived and aggregated in an embryoid body, which undergoes a neural induction (containing a core and a peripheral zone). It is then embedded in a matrix for maturation (Kelava et al., 2016a). During the maturation phase, cerebral organoids innately almost mimic second semester fetal brain growth by developing neuroepithelium regions (Figure 2.3 (Madeline A Lancaster and Knoblich, 2014)). Similarly to human development, neuroepithelium are constituted by a ventricular zone surrounding lumen, a subventricular zone (more recenlty, both inner and outer subventricular zone were generated (Qian, Su, et al., 2020)) and a cortical plate constituted by various cell populations with neurons producing action potentials and synapses (Madeline A. Lancaster et al., 2013). Moreover, comparative studies between fetal human brain developmental stage and cerebral organoids showed some similar transcriptome even if few genes are down or up regulated (Qian, H. N. Nguyen, et al., 2016). However, there are more complex signatures in the human case due in part to vascularisation, to radial glia frequency, and to consequent neuron generation in later fetal stages (Bershteyn et al., 2017; Qian, H. N. Nguyen, et al., 2016). Despite these differences and different growth conditions, parallels can potentially be made between human brain and cerebral organoid tissues development, as investigated in some studies, using histological images (LaMonica et al., 2013; Ostrem et al., 2015; Kostović et al., 2019).

Table 2.1: Percentage of articles studying diseases on cerebral organoids over 2019 and 2020. Neurodegenerative diseases include articles on Alzheimer and Parkinson. Neurodevelopmental diseases include autism, lissencephaly, microcephaly, skizophrenia and various syndromes. Tumors include glioblastoma invasion in cerebral organoids. Infections correspond to viral infections; injury to brain lesions.

Organoid model	Articles (%)
Healthy	52.34
Neurodevelopmental	14.95
Neurodegenerative	11.68
Tumor	8.41
Infection	8.41
Injury	3.74

Time and growth are also quite important parameters, since they can lead to necroses at the core of cultures, mostly due to shortage in nutrients and oxygenation. An answer to this problem consists in slicing cerebral organoids during their growth (Qian, Su, et al., 2020). Such a process increases the number of neuroepithelium layers and the culture longevity.

An important morphological variability between cultures of the different batches exists ("batch to batch syndrome"), as well as variability within a given batch ("batch syndrome"), although not as important as the former. Such variability consists in regions developing in various locations and in an undetermined number (Madeline A. Lancaster et al., 2013). One explanation lies in the non homogenisation between pluripotent stem cells at the origin of the cerebral organoid colonies in term of morphology and pluripotency. Another reason is the thickness of media culturing (Poli et al., 2019). Such variability precludes atlas creation for cerebral organoids (Zaslavsky et al., 2014). In order to reproduce the brain cyto-architectural development with a higher reproducibility, some studies investigated the addition of specific factor to the media (WNT, SSH, FGF) (Krefft et al., 2018; Cederquist et al., 2019; H. Kim et al., 2019; Sivitilli et al., 2020) whereas others used bioreactors (Qian, H. N. Nguyen, et al., 2016; Velasco et al., 2019; Eremeev et al., 2019) or changed the type of culture (Berger et al., 2018; Nickels et al., 2020).

Many authors also chose to study cerebral organoids replicating various diseases (neurodevelopmental, neuro-degenerative, tumoral, infectious or injury models) originating from patient biopsies (A. Tian et al., 2020). Indeed, almost the half of the reviewed articles studies cerebral organoids model disease (Table 2.1). Cerebral organoids are complex to produce and to standardize, but they are already used in pathological cases. The complexity of studying cerebral organoids is also related to protocols and imaging methods described in the following section.

Microscopic studies of 3D brain cultures

Cerebral organoids and other 3D brain cultures are studied both as a whole and at the molecular, cellular or regional level. The frequencies, aims and major disadvantages are summarised in Table 2.2.

Cerebral organoids are most often studied by microscopic observation and analysis. A small fraction of articles do not use microscopy: these either propose a new model or they only rely on RNAseq for the analysis. Cerebral organoids are generally studied first intact

Table 2.2: Methodology to study cerebral organoids

	Percentages	Advantages	Inconvenient
Microscopy	95%	Visualize proteins	sliced in 2D
RNAseq	50.5%	Full Transcriptome	Only on thousand cells located
RTqPCR	34.1%	Detect/identify proteins	Localisation lost
WB	26.2%	Detect/identify proteins	Localisation lost
ELISA	7.2%	Detect antigens	Localisation lost

and then sliced, as shown Figure 2.4.

A few studies (4.33 %) study whole clarified organoids. Most studies produce fluorescent images from confocal microscopy (54.53 %). The two main analysis performed on these images of 3D cultures are quantifications (counting cells and their components, measuring marker intensity or advanced quantifications in particular regions) and morphological measurements (size, shape, etc.) (details in section 2.5). The great majority of studies rely on software or lab-developed scripts for image analysis maybe due to the quicker accessibility of results by automation and the accessibility to reproducible results. The remaining 4.21 % realized only image observations or manual analysis such as cell counting with 1.05 %. One can argue that observation does not allow quantification but contrary to manual counting, it is far less time consuming.

2.4 Preparation and imaging

Sample preparation

Immunohistochemistry Using a microscope may require the preparation of the 3D brain culture through fixation, slicing and immunolabeling.

- Fixation. The fixation step allows the preservation and the long term storage of tissues, by stopping enzymatic reactions (Stanly et al., 2016). In our search, paraformaldehyde was the most commonly used fixation method for cerebral organoids.

- Slicing. Most of the protocols generating cerebral organoids and spheroids cut the samples in slices to facilitate imaging. In the 214 articles analyzed for this review, slices are cut between 5 µm and 50 µm. Slices are realized with different apparatus depending of culture conservation method: cryostat or microtome for frozen samples in the major cases(Mansour et al., 2018); microdissection laser microscopes when only a region is used (Buchsbaum et al., 2020); and a few use vibratome for cultures stored in PBS and agarose (Logan et al., 2020; Gomez-Giro et al., 2019; Nickels et al., 2020; Monzel et al., 2017; Berger et al., 2018; Lisa Maria Smits et al., 2020). Paraffin embedded methods are rarely used on cerebral organoids, due to the size of these cultures (less than a few millimeters).
  In order to avoid slicing and to image a full cerebral organoid in a single acquisition, (Durens et al., 2020) created a protocol aiming at reducing the organoid thickness to around 100 µm. This protocol enables imaging by a single acquisition with high throughput imaging systems, such as confocal microscopes.

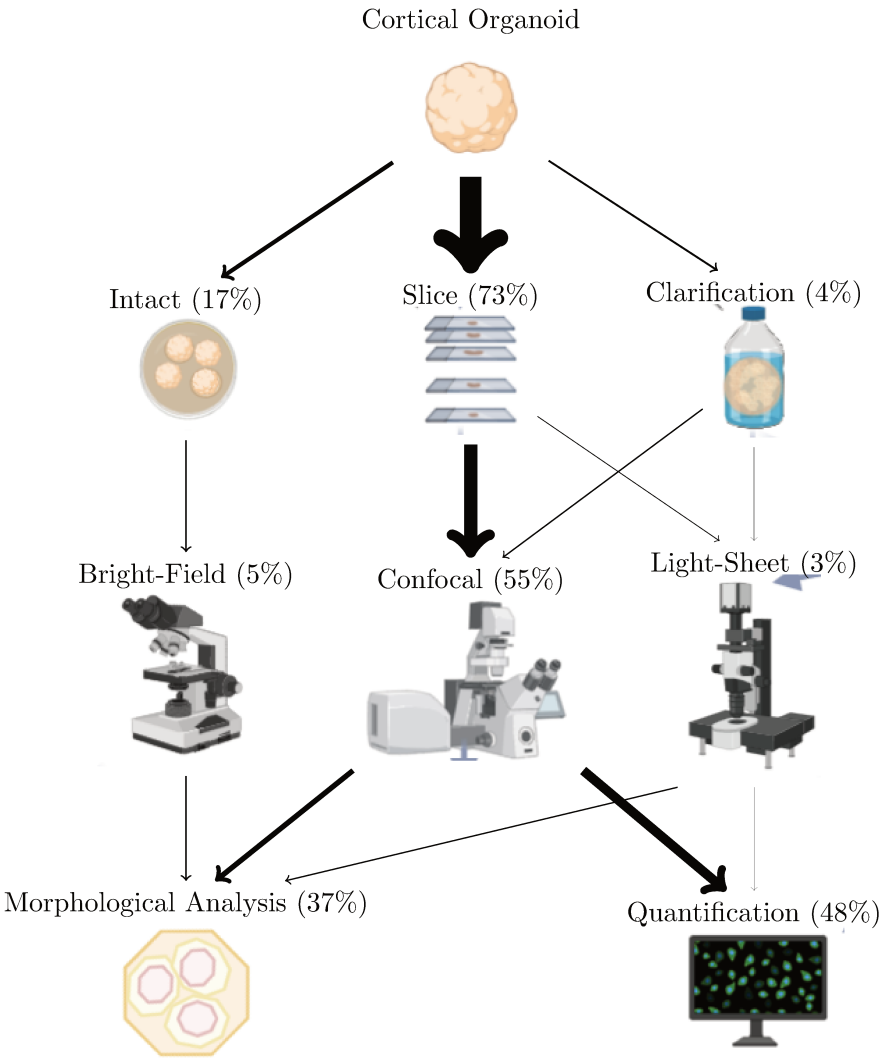


Figure 2.4: Imaging techniques most used to study cerebral organoids. Arrows width is proportional to use of the methods over the reviewed articles. Totals are less than 100 % since only the most used methods are included in this figure. Also, a given article can describe multiple analyze types.

Table 2.3: References of articles using clarification on 3D brain cultures and corresponding image analysis between January 2018 and June 2020. Protocol abbreviations are: hyperhydrating solutions (HS), tissue transformation (TT), high-refractive index aqueous solutions (HIAS), organic solvent (OS).

Reference	Clarification Category	Image analysis
(Sloan, Andersen, et al., 2018)	OS	cell migration
(Masselink et al., 2019)	HIAS	fluorescence intensity and regional marker observation
(Rakotoson et al., 2019)	HIAS or HS	nuclear detection and intensity
(Sakaguchi, Ozaki, et al., 2019)	TT	observation of markers
(Krieger et al., 2020)	Hybrid HIAS and HS	tumor invasion
(Buchsbaum et al., 2020)	OS	cell migration
(Wilpert et al., 2020)	HIAS	observation of marker intensity

- Immunolabeling. Immunolabeling is a crucial biochemical step to prepare samples for the detection and the localisation of an antigen — often a protein — inside a cell, a tissue or an organ. To detect these antigens, a complex of antibodies targeting them are tagged. Fluorescent tags are used for confocal microscopy but an enzyme that catalyzes a colored reaction can be used for other microscopic methods, less used to study 3D brain cultures.
  Immunolabeling is used in 3D brain cultures to detect a cell components such as nuclei (Gomez-Giro et al., 2019), microtubules (Buchsbaum et al., 2020), or mitochondria (Daviaud et al., 2018); a given cell type (neurons (Lisa M. Smits et al., 2019) dopaminergic ones (Bolognin et al., 2019), microglia (Ormel et al., 2018), oligodendrocytes (Marton, Miura, et al., 2019), astrocytes (Watanabe et al., 2017)); or an extracellular marker (Y.-T. Lin et al., 2018). Regions are also identified thanks to immunolabeling, with the combination of different cells markers (Anastasaki et al., 2020; R. Li et al., 2017). Marked cells allow to monitor the tumor invasion inside cerebral cultures (Liu et al., 2020).

Organoids clarification To study a whole 3D sample without cutting, an old practice from the early 1900s consists in rendering it transparent: this method is called clarification. There are 4 main clarification protocols: based on organic solvents (OS), high-refractive index aqueous solutions (HIAS), hyperhydrating solutions (HS) and tissue transformation (TT). To find out more about each of the cited protocols, you can find more information in Matryba et al., 2019. Clarification is not commonly used for cerebral organoids: only 4% of articles use it (Table 2.3).

The major drawback of this method is the time required by the the protocols; the transparency varies over time and is tissue-dependent; protocols can modify the morphological aspect of the culture, inducing over-sizing or shrinking; and some reagents are not compatible with the use of some immunolabelings. Nevertheless, clarification protocols are widely developed for the study of other organs models and even tumoral spheroids (Costa et al., 2019; Molly E. Boutin et al., 2018; Schmitz et al., 2017; Nürnberg et al., 2020). Clarified 3D brain cultures are acquired with confocal (mono-photon), multiphoton or light-sheet microscopy.

Table 2.4: Percentage of articles per microscopy and per task performed for the analysis of cerebral organoids.

Task	Bright-field	Confocal	Light-sheet	Not mentioned	Other / None	Total
Observation	0.84	3.79	0.42	0.42	4.1	9.47
Morphology	3.79	19.16	2.11	2.32	9.4	36.84
Quantification	0.42	30.53	0.00	4.21	12.5	47.58
None	–	–	–	–	–	6.1
Total	5	53.5	2.5	6.9	32.1	100

Imaging Techniques

High quality images are necessary to perform reliable analyses on 3D brain cultures. Bright-field, confocal and light-sheet microscopy are the most often used modalities in this context (Table 2.4). We do not further describe microscopic methods not reaching 2% of use, such as inverted and phase contrast microscopy; those are grouped in the "others" category. The microscope used to acquire an images is chosen based on brain culture type, more specifically the thickness and preparation (Thorn, 2016), as well as the desired analysis to be performed.

Bright-field microscopy Bright-field microscopy is used to observe shape (Monzel et al., 2017) and surface parameters (Iefremova et al., 2017) of 3D brain cultures. On other 3D organ cultures, these images are also used to measure the overall size with automatic methods (Borten et al., 2018; Kassis et al., 2019; Hasnain et al., 2020). In such cases, samples do not require any particular preparation. Cultures can be examined without staining and the illumination does not alter the true colors of the sample. This system is simple and practical to use.

The light source is emitted below the sample and contrasts are created by the absorption of light in the sample. The in-plane resolution does not exceed 2 µm.

The issue often met using Bright-field microscopy is its 2-dimensional nature: although very useful for length and areas measures, only partial shape measures can be realized as the 3-dimensional information is not captured. Another problem is that the quality of the observation is reduced when the contrast is too high, creating distortions in the image. At low contrast, most of the cells are not observable as they are not stained. Confocal microscopy, for example, is better suited for cell observation.

Confocal microscopy The most commonly used fluorescence microscope for 3D brain cultures is the confocal microscope (Table 2.4). The acquired images are analyzed to measure various parameters at the sub-cellular level such as: intensity (Raja et al., 2016); shape (Cullen et al., 2019); surface (Karzbrun et al., 2018); cell distribution (Qian, H. N. Nguyen, et al., 2016) or for 3D reconstruction (Monzel et al., 2017). Confocal microscopy allows the study of samples in the third dimension, which is impossible in bright-field. This optical microscope acquires images at low depth of field (around 500 nm). A laser sweeps the objective via a reflecting mirror. The beam goes through the sample to be imaged and a diaphragm reduces the light received by the sensor to the desired field of view. The whole image is acquired as a mosaic, making possible leveling down the sample plate of a increment of z to image the depth of the culture, and sweep another image. As a result,

these stacked images can be used to reconstruct the 3D volume, enabling measures of 3D parameters characterizing culture structural properties. Immunolabeling via fluorescent tags is necessary to observe confocal images, contrary to bright-field, which conserves the natural color of samples.

One of the principal issues of confocal microscopy is the long acquisition time, particularly for in-depth imaging (in the z plane) where several hours per slice can be necessary. Moreover, only the first few slices produce a sharp signal. For these reasons, some teams prefer to use light-sheet microscopy for 3D culture imaging even though it requires a longer and more complex sample preparation protocol.

Light-sheet microscopy Light-sheet is commonly used to observe 3D samples. However, only 3% of cerebral organoid studies rely on this imaging method, mainly because of the high cost of the device and samples preparation. The illuminating laser source is in the acquisition plane, forming a light-sheet between 4 µm and 10 µm of depth, and of the sample width. The light-sheet is divided in 3 sub-beams (to limit artifacts) which converge toward the sample.

Light-sheet microscopy can acquire organoid images but the in-plane resolution and the light depth penetration are not sufficient to reconstruct a connectivity map according to (Poli et al., 2019). For spheroids, which are 4 times smaller than cerebral organoids, the imaging of clarified data is feasible by light-sheet or confocal microscopy (Costa et al., 2019; Molly E. Boutin et al., 2018).

Other imaging methods Others methods are sometimes used to study cerebral organoids for live imaging (Madeline A. Lancaster et al., 2013), to acquire Ca^{++} activity (Sakaguchi, Ozaki, et al., 2019), or to monitor permeability to certain molecules (Bergmann et al., 2018).

2.5 Image analysis

The aim of cerebral organoids image analysis is to quantify and characterize cell types (stem or proliferative cells, neuronal populations, oligodendrocytes, astrocytes, microglia or epitheliums), cells components (nucleus, neurites as dendrites or axons, mitochondria, synapses), pathological markers of specific disease, cell migration, permeability of tissues to specific molecules, necrosis and structure formations inside the core of culture. In case of group studies, analysis is used to compare size, shape, and dimensions between cerebral organoid groups. In some cases, these results are used to complete and validate information obtained with another method (RTqPCR, ELISA, etc).
Pre-analysis stages are sometimes required to prepare data for future investigations. For example, 3D-reconstruction from acquired slices avoids counting cells multiple times when they appear in multiple z planes (Kartasalo et al., 2018). 3D-reconstruction also allows the visualisation of the multi-view images acquired from light-sheet microscopy (Dobosz et al., 2014). Reconstruction methods from histological slices are based on different features: Fourier, blob or high level features. Validation methods are based on observation, landmark detection, or measures of overlaps (Pichat et al., 2018). After pre-processing, cerebral organoid images are processed with different methods described in this section. As previously mentioned, the two main tasks performed on these images are quantification and

Table 2.5: Quantifications performed on cerebral organoid images, given in percentage of the reviewed articles.

Quantification	Type	Percent
Counting	Cell	20.84
	Protein	8.21
	Nuclei	6.74
	Synapses	1.89
	Pathological	1.05
	Mitochondria	0.21
Density	Various markers	7.16
Total		47.58

morphology (Figure 2.4).

Quantification

Quantification is the main analysis realized on cerebral organoid images (occurring in more than 47% of the reviewed studies, see Table 2.5). Quantification includes markers detection and identification, counting, calculation of intensity and advanced methods for studying cerebral organoid regions.

Counting Counting is performed on specific cells or cell components. In this section, after describing the different quantified structures, we detail some of the counting methods described in the literature.

- Biological structures.
  - Neurons and glial cells: cell counting constitutes 20% of image analysis performed on cerebral organoids (Table 2.5). Brain growth can be tracked by counting markers of neural stem cells (Lisa M. Smits et al., 2019), proliferative cells (Cullen et al., 2019) or differentiated neurons (Berger et al., 2018). In addition to neurons, the brain is constituted of glial cells. Astrocytes are responsible for nutrition and neuronal communication while oligodendrocytes constitute the neuronal myelin gain. Both cell types have been quantified in previous studies (Cullen et al., 2019; Nickels et al., 2020; X. Zhong et al., 2020; H. Kim et al., 2019). Counting of microglia — another kind of glial cell responsible for immunity — has also been investigated (Brownjohn et al., 2018; Ormel et al., 2018). Quantifying organoids microglial cells can help study both their development and their interaction with neurons in case of disease. The last kind of glial cell, constituting the epithelium barrier of brain cavities, is also quantified in choroid plexus organoid models (Pellegrini et al., 2020). Their function of secretion is measured in this previous article by quantifying a typical molecule of transport (transthyretin) only expressed in choroid plexus.
  - Nuclei: The nuclear compartment present in eukaryotic cells contains its genetic information. Brain culture development is assessed by counting the total nuclei

number (and therefore the total cell number) in a slice, a particular region or an entire brain culture (Berger et al., 2018; F. Jacob et al., 2020; Bagley et al., 2017; Qian, Su, et al., 2020; Park et al., 2018; Kielkowski et al., 2020). Identifying nuclei also allows identifying the proportion of apoptosis (cell death), helpful to quantify organoid viability (Lisa M. Smits et al., 2019; Zheng et al., 2020; Pedrosa et al., 2020; Nickels et al., 2020). A similar process with a counter-stain permits the characterisation of the neuronal population density. For example, (Lisa M. Smits et al., 2019) and (Berger et al., 2018) segment nuclei and dopaminergic neuronal markers in midbrain organoids to determine the neurons proportion of their models.

  - Synapses: connective zones between neurites of neurons where the information is transmitted. Number of synapses and their functionalities are altered in case of organoid models of various diseases (Gomez-Giro et al., 2019; Ghatak et al., 2019).
  - Pathological and physiological proteins: Proteins constitute cells and play various roles in transmitting information or regulating factors. In cerebral organoids, proteins are quantified to identify for example a particular cell component such as regulating factors of transcription or tubulin markers (Madeline A. Lancaster et al., 2013). To quantify diseases markers, a key is to count any excessive or insufficient amount of physiological marker, or identifying a pathological marker. For example, the number of A*beta* puncta is used to identify Alzheimer markers in cerebral organoids (Y.-T. Lin et al., 2018).
  - Mitochondria: involved in energy conversion resulting from cellular respiration. Mitochondrial abnormalities caused by genetic mutations in some diseases like in parkinson organoid models (midbrain organoids) can result in cell death (Bolognin et al., 2019).

- Counting methods. Counting cell markers relies on many different procedures. For example, different studies use the following steps: first, images are denoised using median filtering. Second, a Gaussian filter is applied in order to obtain a mask for the marker. Then a median filtering is used on masks, and connectivity is search to remove small connected components (Berger et al., 2018; Lisa M. Smits et al., 2019; Bolognin et al., 2019; Nickels et al., 2020). Finally, expression levels of markers are expressed in pixels or percentage, and sometimes are normalized by the expression level of nuclear markers.
  Another way to count cells consists in binarizing each channel using Otsu thresholding (Otsu, 1979a), and separating overlay cells using watershed (Meyer, 1994). Images are then denoised and channels are overlayed to count cells and calculate ratios (Cullen et al., 2019).
  Most nuclei identification methods use a foreground and background image which are first convolved with a Gaussian filter, then substracted from one another to obtain segmented nuclei (Berger et al., 2018; Bolognin et al., 2019; Nickels et al., 2020). In some cases, the Gaussian filtering is applied directly on the Hoechst channel (Lisa M. Smits et al., 2019).
  A way to quantify synapses is to manually segment them using a specific software (Quadrato et al., 2017; Gomez-Giro et al., 2019). Others choose to co-localize pre-synaptic and post-synaptic punta inside a population of neuronal cells by semi-automatic tools and quantify them per micrometer of neurite length (Ghatak et al., 2019).
  In order to quantify mitochondria, (Bolognin et al., 2019) segmented the plate of

organoid culture, cell nuclei, cell and then a mitochondrial mask was defined via a difference of Gaussians. Masks were refined using a sequence of operations (connected component removal, erosion and skeletonisation).

Intensity In order to quantify the proportion of cell components or molecules inside brain cultures, marker intensity measure has been proposed (around 7% of the image analysis). Different markers can then be measured: neurotransmitters (Jorfi et al., 2018; Sartore et al., 2017), molecule transporters (Wilpert et al., 2020), infiltration of tumors (Liu et al., 2020), nuclei (Rakotoson et al., 2019) or pathological markers (Y.-T. Lin et al., 2018).
To measure the neurotransmitter intensity, the mean grey value of this specific marker is measured in three points of each cerebral organoid border, delimited by a rectangular selection. This fluorescence intensity is then normalized for the tissue background (Jorfi et al., 2018). To assess the neurotransmitter intensity per particular neurons, this parameter is normalized to total neuronal intensity (Ghatak et al., 2019). To quantify the tumoral infiltration regions the fluorescence intensity is thresholded (Liu et al., 2020). For intensity of nuclear markers, background image was subtracted from stained one, the image (originally in 16bits) is converted in 8bit gray-scale and, the intensity of this marker is measured (Stachowiak et al., 2017).

Advanced regional quantification When (Madeline A. Lancaster et al., 2013) generated the first cerebral organoid, they discovered the presence of various brain regions, similar to the ones already described in human brain. It is possible to identify regions using a combination of different markers, marker density or marker location. (Paşca et al., 2015) were the first to quantify different types of cells inside cortical spheroid regions: a ventricular zone (VZ), a deep layer and a superficial layer. One year later (Raja et al., 2016) counted nuclei expressing a caspase to determine the cell death from the center to the external cortex of a cerebral organoid. Indeed, markers of cell death and proliferation are often measured in VZ and SVZ regions (Qian, Su, et al., 2020; Anastasaki et al., 2020; F. Jacob et al., 2020; W. Zhang, Ma, et al., 2020). Other articles also calculate the percentage of particular neurons in VZ, SVZ, outer SVZ (R. Li et al., 2017) or CP (W. Zhang, S.-L. Yang, et al., 2019). With the emergence of fused specific region organoid, (Bagley et al., 2017) expressed the percentage of various fluorescent markers in dorsal and ventral forebrain organoids.
As of today, regional quantification mostly remains on a semi-automatic process (Albanese et al., 2020). All of the articles cited use imageJ after a manual extraction of the region of interest. Regional organisation is also scored manually by three authors in "no organisation", "geographic segregation" and "laminar structures" to determine the degree of differentiation (Cullen et al., 2019).
Between January 2018 and June 2020, we only found classic segmentation methods to identify cell components. It would be interesting to test various segmentation methods to identify the most adapted to accurately identify cellular components.

Morphological analysis

Morphological analysis represent approximately 37% of the studies of 3D cerebral organoids images and are summarized in Table 2.6.

Upon these morphological parameters harvested, two categories are further detailed in this section: basic and advanced metrics containing 2D (diameter, perimeter, length, area,

Table 2.6: Morphological analysis performed on cerebral organoid images, given in percentage of the reviewed articles.

Type	Analysis	Percent	Dimension
Basic	Diameter	4.84	2D
	Perimeter	0.84	2D
	Unspecified size	2.95	2D
	Distances	4.21	Mix
	Neurite	2.95	2D
	Radialisation	0.42	2D
	Ventricles	1.68	2D
	Nuclear Morphology	0.21	2D
	Area	11.58	2D
	Volume	1.26	3D
Advanced	Thickness	4.63	2D
	Folding	0.63	2D
	Tortuosity	0.21	2D
	Curvature and Wrinkling	0.21	2D
	Circularity	1.05	2D
	Sphericity	0.42	3D
Total		36.84	

folding, wrinkling, curvature and circularity) and 3D analysis (volume, sphericity and distances) see Figure 2.5.

Basic metrics: two dimensional analysis Some studies investigated organoid global growth, by measuring size indices to identify the state of growth and well being of the culture, as well as to compare methods of culturing or disease models of cerebral organoids.

- Diameter, Perimeter. Diameter and perimeter are measured in (6%) of cerebral organoid articles. They are part of tools to measure the size of cerebral organoid to evaluate their growth or to compare different groups of culture (healthy and disease models for example). Indeed, their size is evaluated by their diameter (Monzel et al., 2017; Sartore et al., 2017; Sivitilli et al., 2020) or perimeter (Buchsbaum et al., 2020) on bright-field images. Others authors use confocal microscopy to measure the size in term of perimeter (Iefremova et al., 2017).
  To measure these parameters, semi-automatic tools are provided in some software. One of the method is to sample diameter twice in a perpendicular angle using the line tool of FIJI, on maximum z-projections made from image stacks acquired by confocal microsocopy (Schindelin et al., 2012). In bright-field microscopy, perimeter of an element inside an image is measured on boundaries of manual or semi-automated selected regions. For diameter, the longest distance between two points of a selected region is measured. These measures had to remain in early stage of development due to heterogeneous shape in later stage in this culturing model.
  Such metrics could become an indicator of cerebral organoid shape only in early stages.

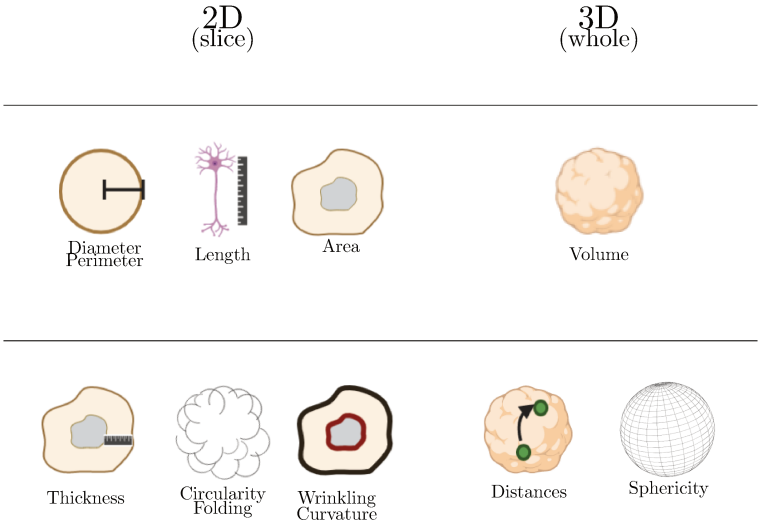


Figure 2.5: Summary of major morphological analysis performed on cerebral organoids. The first line in this table corresponds to basic morphological analysis and the second one to advanced. Basic parameters are used to calculate the advanced ones.

Nevertheless in other kind of organoid models, diameters are an indicator of their shape all along their growth as most of them stay spherical and a few are elliptical (Kassis et al., 2019). In this article, intestinal organoids are identified, and their diameters are measured thanks to deep neural network based on anchor boxes and features pyramidal network from (T.-Y. Lin et al., 2018). Some software are developed solely to measure the spheroid perimeter (W. Chen et al., 2014).

- Length. The first form of measuring distances is measuring it at cell scale. Measuring cell component allow to identify characteristic of growth culture. Researchers measure for example neurites (Xiang et al., 2019; Cullen et al., 2019; Ao et al., 2020; Durens et al., 2020) or cilium length (W. Zhang, S.-L. Yang, et al., 2019). Using lengths, ratio can be calculated to compare neurites in different regions inside cerebral organoids (Xiang et al., 2019), or to evaluate the direction of growing of this cell component (Durens et al., 2020).
  To measure the length of cell components, semi-automatic tools are used to define the boundaries of each of neurites or cilium, and distances between the two boundaries are calculated.

- Area. The surface area better represents cerebral organoids shape in latter stages than other 2D parameters, being more acute on non-spherical shape. Area is the most calculated morphological parameter on 3D brain culturing (12%), and is used to compare various cerebral organoids growth in different conditions or groups (Watanabe et al., 2017). At the sample level, the surface area of 3D reconstruction of light-sheet microscopic images can be performed (Qinying Wang et al., 2020; Yun Li et al., 2017). Nevertheless, some authors prefer measuring regions (VZ and SVZ and CP) in term of areas on slices to follow there growth (Watanabe et al., 2017; Iefremova et al., 2017). Particularly ventricular lumen area are measured in order to know the state of 3D brain culture or to test a pharmacological component (Qian, H. N. Nguyen, et al., 2016; Iefremova et al., 2017; Sartore et al., 2017; Di Matteo et al., 2020). Area can also

serve to determine the culture viability. The necrotic and viable areas are measured in the case of comparisons of two kind of culturing (Berger et al., 2018). The calculation of an expressed marker area without considering the regional segregation can also be done (in entire slices). The area of a kind of neurons or glial cells (Park et al., 2018) and the area of all the nucleus (Cullen et al., 2019) are another example of this kind of measurement. Of note, measuring area enables the evaluation of co-localisation of some markers (Ao et al., 2020) like presynaptic and postsynaptic ones.
To identify the growth of cerebral organoid culture, authors calculate the total surface area of the whole organoid. Regions of interest are surrounded manually around the entire organoid from a bright-field microscopy image, and thanks to an imageJ module ("Area Measurements of a Complex Object"), the surface area is calculated (in pixels) (W. Zhang, Ma, et al., 2020; Gomez-Giro et al., 2019). Viability of cultures can be assessed both at regional or cellular level. (Berger et al., 2018) choose a typical fluorescent marker not expressed in necrotic core region and measure its surface expression related to the total surface area. This parameter is measured as the minimum area in pixels that an object must have after its selection, thanks to semi-automatic tools (Zen software). For cell viability, areas of some cell component markers (such as plasmic membrane or enzyme) are also calculated. To that aim, (Cullen et al., 2019) convert the two channel corresponding to plasmic membrane and enzyme in 8 bits images, then binarize images to obtain cell shape regions. The area of these two markers is then quantified, and their ratio is calculated. Synapses quantification can be achieved using marker areas co-localisation. Synapse areas are for example calculated by overlapping Homer (post-synaptic) and Bassoon (pre-synaptic) channel signal in the case of assembloid of organoids using a lab-developed tool (Sloan, Darmanis, et al., 2017). For particular cerebral organoids, areas are even calculated. For example in fused ventral and dorsal forebrain organoids, areas of typical expressed markers are also calculated (Bagley et al., 2017). For the blood brain barrier organoids, areas are equally measured, particularly the core area by measuring it at 50 µm from the surface. A scale bar is used as a reference to correctly assess the distance (Bergmann et al., 2018). In mammalian, colon and intestinal organoids, the whole area of the entire organoid digitized after bright-field imaging is calculated (Borten et al., 2018; Ren et al., 2018; Hasnain et al., 2020). As an example, for (Borten et al., 2018), after a segmentation of colon organoids (by a conversion, opening-closure, thresholding, filtering to denoise, filling holes, denoising and removing debris), the surface area of identified region of interest is measured.

Basic metrics: three-dimensional analysis Measuring the cerebral organoid size in 3D is also possible in light-sheet images, where the volume of this 3D brain culture is assessed (Qinying Wang et al., 2020; Yun Li et al., 2017). Only few authors calculated this parameter, possibly because this imaging modality is poorly used. Indeed 2.5% of articles use light-sheet, and 1.3 % calculate the volume of cerebral organoids (Table 2.5).
The outline of the cerebral organoid is delineated and used to compute both volume and surface area, with semi-automatic tools (Yun Li et al., 2017). Such metric could be use to indicate if an antitumoral treatment works like it was made for spheroids. However, the number of spheroids is too important to semi-automatically or manually measure volume when performing drug testing. (Wojaczek et al., 2019; Kalaydina et al., 2019) use deep learning method based on the YOLOv2 architecture (using anchor boxes instead of fully connected layers) (Redmon et al., 2016) to identify spheroids and calculate their volume V

from the radius r, assuming a perfect sphere. Manual calculation of the radius r was made by measuring the diameter twice for each spheroid, then averaged and divided to obtain r, using a scale bar as a reference. After automated identification, coordinates of predicted bounding box enable the measurement of radius and the volume calculation (Kalaydina et al., 2019).

Advanced metrics

- Length and thickness. Advanced specific distances are calculated in Cederquist et al., 2019 article to identify the cerebral organoid patterning. First, the center of mass (CM) of a factor-organizing cells is computed, inside a grid applied on the image. The CM is a function of its mean gray value intensity and the total intensity. After choosing a marker of a typical protein, intensity is thresholded and regions of interest (ROI) are identified. Finally the Euclidean distance between each ROI and the CM is obtained.

  The second kind of distance is the neuroepithelium thickness. In cerebral organoids, this thickness is used to characterize an organoid model (Watanabe et al., 2017; Sakaguchi, Ozaki, et al., 2019; W. Zhang, Ma, et al., 2020; Di Matteo et al., 2020; Buchsbaum et al., 2020) and to follow the patterning of the culture (Cederquist et al., 2019) or the effect of various culturing on the growth of the regions contained in it (Qian, Su, et al., 2020).

  A specific feature of the neuroepithelium thickness is the relative thickness R_{thick}, which is the ratio of the total layer thickness TL_{thick} over the VZ region thickness VZ_{thick} (W. Zhang, S.-L. Yang, et al., 2019):

$$R_{thick} = \frac{TL_{thick} - VZ_{thick}}{TL_{thick}} \tag{2.1}$$

  Another way to calculate the relative VZ thickness is defined as the ratio of VZ thickness to VZ plus outer layer thickness (Qian, H. N. Nguyen, et al., 2016).

- Circularity and Folding. The shape of the cerebral organoid is one of the parameter used to distinguish it from spheroids, and a marker of later stage of the cerebral organoid growth. Circularity (C) is a shape parameter measured in the early stage (day 6) of development in intact cerebral organoids, and is defined by Yoon et al., 2019 as:

$$C = 4\pi \cdot \frac{A^2}{P} \tag{2.2}$$

  where A is the object area and P is the perimeter. An index of 1 reflects a perfect circle.

  Human cortical surface is characterized by folding (gyri and sulci), which is not always present in mammalian models (Kelava et al., 2016b). To determine if a cerebral organoid model reproduces gyrification, (S. Li et al., 2020; Qinying Wang et al., 2020) quantify folding. On bright-field or in higher magnification view images, the Canny edge detector is used to extract edges. Once edges are found, their total length is used to compute a folding index (Wojaczek et al., 2019).

- Wrinkling and Curvature. Wrinkling occurs at two brain formation stages: during the emergence of folds along the neural tube, and during the expansion of surface area. Measuring wrinkling is a relevant index to characterize diseases such as lissencephaly. (Karzbrun et al., 2018) calculate the wrinkling and the curvature inside cerebral organoid. 2D wrinkling is the measure of the real perimeter of the organoid divided by the total maximal perimeter as a circle containing the organoid. The curvature is defined as the average of the tangent angle $\theta(r)$ derivative along the surface of inner and outer neuroepithelium perimeter contour $\eta r\theta(r)$.

- 3D: Sphericity and Distances. For 3D images, circularity cannot be characterized, hence the identification of brain gyrification uses the sphericity (how spherical an object is) on light-sheet images (S. Li et al., 2020; Qinying Wang et al., 2020). The calculation of sphericity, ϕ originally generated by Wadell in 1932, is defined as the ratio of the cell surface area of a sphere over the cell surface area of a particle, with V the volume of the particle and A the surface area of the particle:

$$\phi = \frac{\pi^{1/3}(6V)^{2/3}}{A} \tag{2.3}$$

  The latest measure performed on cerebral organoids evaluates the tumor propagation in some models. The distances between tumoral cells or between them and the center of the cerebral organoid is computed. From binarized images, several steps are then performed: exclusion of single cells (using by connected components), holes filling, organoid surface approximation (by a Delauney triangulation). Normal distances between tumoral cell voxels is then calculated (Krieger et al., 2020).

Summary on morphological parameter extraction Over the considered time range (January 2018 to June 2020), we only found methods focusing on classic extraction of shapes. More recently ((Albanese et al., 2020)), authors extracted ventricular region of cleared organoïds using a Deep-Learning approach (U-net architecture (Ronneberger et al., 2015a)). This original work paves the way to Deep-based approaches and clearly shows the potential of such methods. Similar methods could potentially be used for all types of cerebral organoids structures.

2.6 Software

Pre-analysis software

Most imaging platforms include a software able to perform pre-analysis. For example, the tiles module and the position module of the Zen software can be used to reconstruct multi-view images in 3D (Watanabe et al., 2017); while the NIS imaging software (Nikon) can measure the size of cerebral organoids (Berger et al., 2018).

However, these software packages are generally not adapted to perform the tasks variability required by researchers who want to analyze cerebral organoid imagings. To analyze images, neuroscientists choose dedicated software depending on the study topic, imaging type, ease of use, source code flexibility, their computing knowledge, and budgets. ImageJ, Matlab, CellProfiler and Imaris, are the most used software solutions in this context, as shown in table 2.7.

Table 2.7: Software used to analyze cerebral organoid images.

Software	Open source	Automatism	percent
imageJ / Fiji	yes	semi-automatic	51.37
Matlab	no	automatic	4.42
CellProfiler	yes	semi-automatic	0.84
Vast	yes	semi-automatic	0.42
Imaris	no	semi-automatic	3.16
Visiopharm	no	automatic	≤ 0.2
ImageScope	yes	automatic	≤ 0.2

ImageJ / Fiji

ImageJ is an open-source software which can run on all the main operating systems (Windows, macOS, Linux/Unix). It does not require knowledge in coding and the interface is somewhat user friendly; for example it supports "drag and drop" of the image to analyze. ImageJ is the most widely used software for the analysis of 3D brain cultures (used in over half the articles surveyed, see table 2.7). The most popular modules include the "cell counter" plugin, the "particle" counter, the "length" and "area" measurement functions, the "ROI tool", and the "Canny edge detection" to measure folding density.

For those who need further analysis, the advantage of this software is the possibility to code macros in Java to automate analysis or to create new tools (Raja et al., 2016; Ormel et al., 2018). One drawback is that some file extensions require additional plugins to be handled (for example bioformat files) while in-house extensions are not handled at all. Also, ImageJ performances are impacted when used with large images and may require increasing memory allocation.

Theoretically, it is possible to perform 3D analysis with the "ImageJ3Dviewer" plugin. However, to our knowledge, these tools have not been used for the analysis of 3D brain cultures.

Matlab

Matlab is a numerical computing environment and proprietary programming language widely used by the scientific community, for example for image and data processing or simulations [1]. Matlab can also run on the main operating systems. Many toolboxes exist and can be used to develop new tools. Matlab is more versatile and faster than the other software on large amounts of data, but it requires specialized knowledge to develop and validate new tools. Matlab is the second most used software (with 5% of use), and has been used for a wide range analysis tasks: nuclei segmentation (Lisa M. Smits et al., 2019); cell segmentation (Bolognin et al., 2019); puncta co-localisation (Sloan, Darmanis, et al., 2017); curvature, folding and surface measurement (Karzbrun et al., 2018); and tumoral cell dispersion evaluation (Krieger et al., 2020).

CellProfiler CellProfiler is an open source software developed in Matlab; it thus requires a Matlab license. Many plugins are available and used by different teams analyzing 3D brain cultures (Park et al., 2018; Pedrosa et al., 2020). The major inconvenient is that not all

[1] MathWorks, Matlab

image formats are currently accepted. Specific scripts must be developed, but new plugins can be coded in Matlab as mentioned before.

Vast Vast [2] is a Matlab-based semi-automatic segmentation tool for 2D and 3D images and is used to segment images from transmission electron microscopy, including segmentation of synaptic compartment in cerebral organoids (Quadrato et al., 2017).

Imaris

Imaris is a commercial software which allows 3D and 4D (along the time) analysis of cell cultures, but it remains a semi-automatic tool. User selects objects inside images to detect and process them. Imaris is used in 3% of the articles surveyed for this review, and is particularly used to reconstruct images in 3D (T. Kadoshima et al., 2013; M. Renner et al., 2017), to count cells (R. Li et al., 2017) and to quantify volumes, surface area and sphericity (Yun Li et al., 2017).

Others solutions

Visiopharm[3] is a commercial solution composed of a range of AI-based image analysis and tissue mining tools. It has been used on fluorescent cerebral organoid images to count cells (Stachowiak et al., 2017). ImageScope[4] is a commercial automatic quantitative software for widefield microscopy, used to count pixels labeled with a specific marker for prion in a Creutzfield-jacob model of cerebral organoid images (Groveman et al., 2019).

Others methods have been validated for the study of non-cerebral organoids: Cytocensus for retinal organoids (Hailstone et al., 2020); OrgDyn for widefield images of mammalian organoids (Hasnain et al., 2020); OrganoSeg for 3D bright-field images of colon organoids (Borten et al., 2018). Most of these tools are based on image filtering and segmentation. Notably, OrgaQuant locates and quantifies the size distribution of human intestinal organoids in bright-field images based on a deep learning network (Kassis et al., 2019). Only recently a software was created to characterize the cytoarchitectures of cerebral organoids imaged by light-sheet microscopy (Albanese et al., 2020).

2.7 Available datasets

The number of available dataset in the organoid/spheroid field is limited see Table 2.8, and only few of them could be used to characterize the development of brain organoids. We choose to use the bright-field brain organoid dataset to our work and ask to other authors if they could let us some others bright-field brain organoid images. Even if this dataset contains only a few samples, these images represents some advantages: open access, and acquired at three different developmental stages; The imaging method of this dataset is the one usually used to follow brain organoids growth, and organoids are centered within the image. Therefore, such dataset allows a developmental morphological characterisation.

About the ground truth The first work before segmenting images is to prepare the dataset. One of the preparations consists in creating the ground truth (GT) of images. The ground truth is the manual annotation of images which allow to compare fairly automatic

[2] https://software.rc.fas.harvard.edu/lichtman/vast/
[3] https://www.visiopharm.com/
[4] https://www.leicabiosystems.com/fr/imagerie-pathologique/analyser/

Table 2.8: Available organoid/spheroid image datasets found in the literature.

Reference	Culture type	Acquisition	Healthy/Total	Size
Gomez-Giro et al., 2019	Brain org.	Brighfield	20/40	1388×1038
Nickels et al., 2020	Midbrain org.	Brightfield	300/300	1388×1038)
Kassis et al., 2019	Intestinal org.	Brightfield	756/756	450×450
Borten et al., 2018	Colon and breast cancer spheroid	Brightfield	120/120	864×648
Stachowiak et al., 2017	Brain org.	Confocal	16/40	1389×1040
Gomez-Giro et al., 2019	Brain org.	Confocal	30/60	1907×2355
Lisa M. Smits et al., 2019	Midbrain org.	Confocal	76/76	1388×1040
Nickels et al., 2020	Midbrain org.	Confocal	10/10	1024×1024
Schmitz et al., 2017	Melanoma spheroid	Light-sheet	2/2	690×726

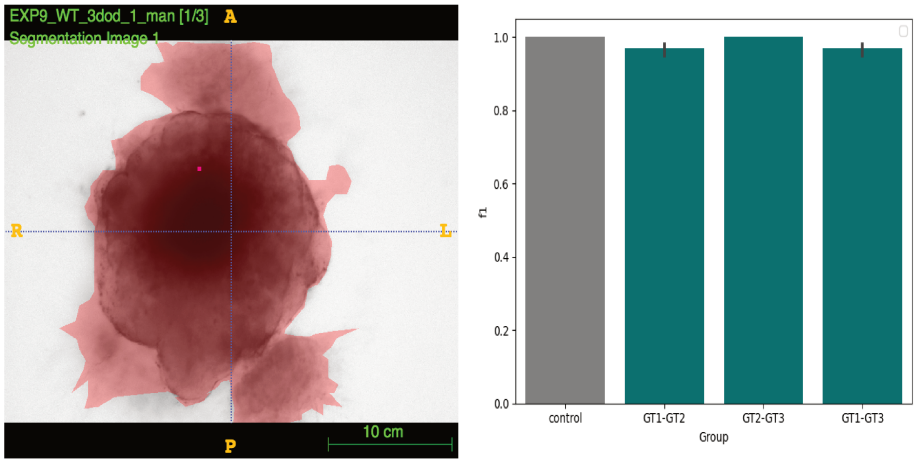


Figure 2.6: Ground truth annotation on itksnap and variations in ground truth annotation by a biological expert. Three manual segmentation have been made by the experts, and F1 score has been calculated. Kruskall-wallis test performed on the dice score between control and GT comparison group does not show significant differences (p-value > 0.05).

with a biological expert annotation. The ground truth has been created by a superimposition of 3 annotations of experts witch selected the contour of images with the ITK-SNAP software www.itksnap.org (Yushkevich et al., 2006a). To verify the three manual segmentation are quite similar we compare the annotation variations Figure 2.6.

Resources We develop all the scripts in Python 3.6 with an Anaconda framework containing Keras 2.3.1 and Tensorflow 2.1 and run them on an Intel Core i7-9850HCPU with 2.60 GHz and a NVIDIA Quadro RTX 3000s GPUdevice.

2.8 Discussion

This section gives an overview of the current limits in: cerebral organoids generation, existing imaging solutions and analysis methods and tools. We also present expectations for new image and volume analysis tools. Indeed, one of the keypoint in the context of image analysis

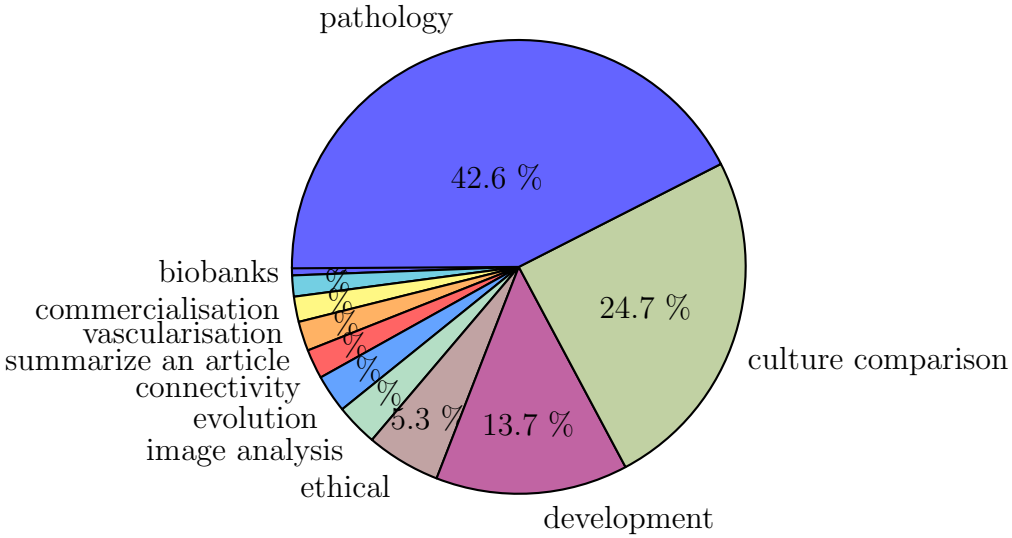


Figure 2.7: Main domains of reviews published. Percentages less than 5% are not mentioned.

of cerebral organoids is the faisability of the analysis supported by the quality of generated images and on their imaging.

Cerebral organoid generation limitations

Some limitations remain in the generation of cerebral organoids. The main limitation is the necrosis occurring during the growth of cerebral organoids due to lack of nutrients and oxygenation. Slicing the organoid and optimizing the culture medium have reduced this necrosis (Qian, Su, et al., 2020; Berger et al., 2018). However, the lack of vasculature of cerebral organoids remains the root of the problem. In some protocols cerebral organoids are transplanted in mice brains for irrigation (Mansour et al., 2018; Pham et al., 2018; Shi et al., 2020) while others generate blood-brain barrier organoids (Cho et al., 2017; Nzou et al., 2018; Bergmann et al., 2018) but these solutions lack the self-patterning of vessel generation. Recently, the theoretical elucidation of this problem has been exposed based on two models of gradient diffusion of the vascular endothelial growth factor (Hong et al., 2019). A recent study also documents the generation of telencephalic and choroid plexus organoids allowing the production of cerebrospinal fluid (Pellegrini et al., 2020). A combination of these barriers in a cerebral organoid model could potentially increase its lifespan.

"Batch syndrome" and batch-to-batch variability as previously described are a major inconvenient for the commercialisation and robust analysis of cerebral organoids. A pre-requisite for commercialisation consists in measuring size and morphological complexity (cf 2.3) to validate the model (Choudhury et al., 2020). However, existing tools to measure the overall size present drawbacks like time consumption as they are manual or semi-automatic, making them unsuitable for mass production. Though the generation of this model is less than a decade old and not well stabilized, growth monitoring of cerebral organoids neglected for the benefit of articles comparing pathological and physiolocal cerebral organoid models. Almost half of the related articles and reviews included in this review are about pathological organoids (Table 2.1, Figure 2.7). In others organ models, automatic tools have emerged to measure the size or to classify the morphology of others organ models (Hasnain et al., 2020; Borten et al., 2018; Kassis et al., 2019). We think a similar tool for cerebral organoids could help to measure and identify the growing step of cerebral organoids.

The large amount of cells to handle in generated cerebral organoids, in addition to their

variability in numbers, also increases the difficulty in analyzing images (from 3000 to 11000 cells at 6 months depending on the protocol). Nethertheless, similar problems have already been addressed. For example, connectome has already been investigated for larger central nervous systems like drosophilae (25000 neurons and their projection), but also in part for the human brain (containing around 86 billions of neurons and their projections) (Scheffer et al., 2020; Rosen et al., 2021; Maller, 2019). The Human Connectome Project requires to create collaborations between laboratories and a large storage capacity, as terabytes of storage are required in computing resources for a whole human brain. In order to investigate the development of cerebral organoid connectome under various protocols, we think it could be necessary to create a similar initiative collaboration, and biobanks dedicated to cerebral organoids images. Another review discussed about the benefit and limitations of conserving cerebral organoid generated or their cell contents inside biobanks (S. Li et al., 2020). It could help also to investigate for instance each neuropathological model created in cerebral organoids as it has been done for glioblastoma (brain tumor) cultures of patients (F. Jacob et al., 2020).

Generation of cerebral organoids is not the only limitation of these models, to an image analysis point of view, the imaging remains an issue.

Preparation and imaging method limitations

Preparation Sectioning during preparation restricts the efficiency and throughput of organoids and spheroids (Paşca et al., 2015). The loss of bio-material is critical for these small cultures that do not exceed 4 mm in diameter for cerebral organoids and 0.5 mm for spheroids. Moreover 3D reconstructions computed from these altered images can introduce a bias (Richardson et al., 2015). To avoid slicing and to obtain a full cerebral organoid image in a single acquisition, (Durens et al., 2020) generate an organoid with a thickness of 100 µm. Another problem of classic immunohistochemistry methods is the poor diffusion of markers in the depth of cultures. A possible solution is to use clarification. Nowadays, only a few teams use this expensive solution on cerebral organoids (see Table 2.3). The aim in the near future is to use clarification in high-throughput platforms (Poli et al., 2019; Grenier et al., 2020). Very recently, out of the time scope of this review, cleared cerebral organoids, were analyzed in one of this expected platform called SCOUT (Albanese et al., 2020). Others authors tried also others clarified methods to analyze in 3D their cerebral organoids (Adhya et al., 2021; H. Renner et al., 2020). However, contrarily to the spheroids field, to our knowledge there is no article comparing existing clarification methods to find the most accurate one, allowing better image analysis on 3D brain cultures (Nürnberg et al., 2020). An appropriate clarification method applied on cerebral organoids could help to acquire images of quality and allow the most accurate 3D analysis.

Imaging Each of the various acquisition methods used on cerebral organoids has specific limitations. Bright-field microscopy is only used to analyze intact samples and is a powerful and simple acquisition modality to identify 2D morphology and follow the growth, however not suited for inner cells study. The resolution of confocal microscopy is satisfactory only for the superficial sweeps while only a halo of markers are visible in the deepest views (Lisa M. Smits et al., 2019). Accordingly, only cell counting in a single acquisition plane is possible (Qian, Su, et al., 2020). Light-sheet microscopy is only used by few teams (See Table 2.4). This method is an expensive solution that requires to be tested on cerebral organoids clarified by various protocols before obtaining good quality data. This imaging method has been used in the recently published articles on only one clearing method (Albanese et al., 2020;

Adhya et al., 2021). A comparison of images resulting from various clarification protocols in light-sheet and confocal modality, not already provided to our knowledge, could be an important step to identify the best methodology for the observation and analysis of 3D cerebral organoids. Such image acquisition methods diversity yields additional complexity in the automated analysis tools standardisation (Table 2.4).

Analysis

Some authors chose to develop their own algorithm rather than using already available software modules (Stachowiak et al., 2017; Berger et al., 2018; Lisa M. Smits et al., 2019; Cullen et al., 2019; Bolognin et al., 2019; Krieger et al., 2020). In addition to software imaging and updates high costs, these are usually not optimized for their specific imaging modalities. Also, commercialized software source code is not always available, to be modified to fit custom needs. With regard to clarified samples images, only a few software are allowing 3D-data analysis (See section 2.6). Noteworthy is the fact that most of the existing solutions remain semi-automatic. In the actual context of data expansion and increase of organoids models (Ashok et al., 2020; Choudhury et al., 2020), the development of fast and automated tools is mandatory.

Indeed, manual characterisation of spheroids, smaller than cerebral organoids, is time consuming (Soetje et al., 2020). In contrast, automated processing based on computational neural network (CNN) can provide real time measures (Kalaydina et al., 2019; Anagnostidis et al., 2020; Wojaczek et al., 2019). In other imaging disciplines such as MRI brain tumor detection, similar methods are already widely developed (Gordillo et al., 2013).

Aside from quantification speed optimisation, another benefit of CNNs is that they are not subject to human error (except from the manual annotation process). Nowadays, CNNs are used to measure size parameters from 3D intestinal organ models, (Kassis et al., 2019) or to count cells in retinal organoids (Hailstone et al., 2020).

Despite the fact that is widely developed for others 3D cultures, to our knowledge, only one article included deep learning methods in order to segment ventricules of their cerebral organoids (Albanese et al., 2020). However, the comparison of machine learning methods applied to cerebral organoids would bring precise information on analysis precision and reproducibility. Nevertheless, the lack of shared images databases precludes such a comparison (Chakradhar, 2016).

Need of analysis tools

Automatic monitoring during cerebral organoids development, although essential for their commercialisation (Chakradhar, 2016) and management of the increase culture amount (Figure 2.1), is still lacking.

In others organ models — i.e. mammary organoids (Hasnain et al., 2020) —, automated tools allowed the discovery of various groups of morphology. Such classification would be interesting to highlight in cerebral organoids.

Studying the morphology and measuring the size of a cerebral organoid in 2D images can help to compare groups inside a study (Watanabe et al., 2017; Monzel et al., 2017; Iefremova et al., 2017). However, the tools used to that aim are still semi-automatic or manual. A possible answer lies in the use of CNNs, which can help identifying and characterizing cerebral organoids in the culture (Kassis et al., 2019; Wojaczek et al., 2019; Kalaydina et al., 2019; Anagnostidis et al., 2020; Soetje et al., 2020). These tools completed with transcriptome analysis in various locations inside some cerebral organoids blindly selected in a batch, could help to automatically validate the growing step of a cerebral organoid, in a

productivity chain. Such a tool used in research would improve the speed of organoid groups comparison. Additionally, automated size and growth measurement could be helpful in other 3D cultures (organ models or spheroids), less complex in term of morphology (Friedrich et al., 2009).

Another interesting feature of cerebral organoids is the presence of regions mimicking similar human brain regions, but at random location, with shape variability and in random numbers (Madeline A. Lancaster et al., 2013). Regional quantification has already been performed in two dimensions with semi-automatic tools (Qian, Su, et al., 2020; Anastasaki et al., 2020; F. Jacob et al., 2020; W. Zhang, Ma, et al., 2020). Conversely, automatic 3D structures extraction has not been done yet, except for ventricular regions (Albanese et al., 2020). To observe or quantify molecules in specific brain regions, researchers use atlases on the assumption that structures localisation and shapes are identical to the ones found in a healthy subject. Such assumptions are not valid for cerebral organoids, because of the previously mentioned variability. Moreover, atlas creation process is a complex task, even in the case of in vivo models or human brain (Bazin et al., 2020; Johnson et al., 2010; Quanxin Wang et al., 2020). Recently, some authors developed a brain atlas based on deep learning in order to automate the segmentation of mice brain regions, which are variable in size and shape (Iqbal et al., 2019). This study demonstrates the feasibility of localizing brain structures despite mild brain variability, and could be translated to cerebral organoid study.

Additionally, the minimal density of markers defining a region in 3D would be interesting to highlight. Such characterisation could help identifying unknown functional and architectural markers, as mentioned in Grenier et al., 2020, with the perspective of generating a high-throughput deep learning-based image analysis platform for drug testing.

Such platform could benefit many other applications. Defining regions with a reduced number of markers on a single sample could leave room for another marker, more relevant for a specific study. Moreover, organoids structures are manually extracted to count markers, or are cut to analyze in RNAseq (Buchsbaum et al., 2020; Sloan, Andersen, et al., 2018). Nevertheless, structures are microscopic, and the tools enabling the selection of regions depend on the accuracy of the operator. This becomes particularly critical when regions are cut with laser microscopes. Precision in cerebral organoid cutting could be increased using automatic region identification, or error correction through a dedicated analysis tool.

Automated quantification of cells and their components would be of great interest, as such measures remain the main analysis realized on cerebral organoids Figure 2.4. While cell counting is the principal quantification realized on cerebral organoids (cf Table 2.5), authors only use classical segmentation (thresholding, watershed for example) (Cullen et al., 2019). Similarly quantifying physiological or pathological markers inside cerebral organoid regions has been performed only with semi-automatic tools (Qian, Su, et al., 2020; Anastasaki et al., 2020; F. Jacob et al., 2020; W. Zhang, Ma, et al., 2020). Automatic tools developed for other culture models could potentially be used to achieve such quantification (Piccinini et al., 2020). Development of new methods could also be inspired by approaches already used for in vivo brain models (C. Zhang, 2017), however with some limitations regarding methods used for real human brain tissue study. In this specific case, cell counting is based on three different approaches: histological or stereological approaches, DNA extraction and isotropic fractionnating. Only the first method keeps the localisation of the cells (Von Bartheld et al., 2017) and would therefore be suited for cerebral organoids.

Another interesting project to develop is the creation of cerebral organoid connectomes. We think connectivity mapping has to be developed at various scales, between two organoids of an assembloid, between regions inside an organoid, but also between the constituting

cells. In assembloids, the connectivity could help to explain neurodevelopmental defects using pluripotent stem cells derived from neurological diseases patients and to test potential therapeutic compounds (Bagley et al., 2017). Another review addresses the challenge of connecting an organ culture with cerebral organoids, in order to reproduce important axes in the human body, although this raises major ethical questions (Chukwurah et al., 2019).

New computational methods identifying connections could help to understand organoid inner structure. For instance, regional connectivity could be helpful to identify a pathological formation inside the neuroepithelium, and help to understand the neurodevelopemental formation (Seto et al., 2019). Finally, characterizing the full connectivity of the whole organoid, or inside a particular region, could help distinguishing relations between different cell types, relevant to identify neurodegenerative diseases (Marotta et al., 2020).

The identification of cell interactions has been described in another review (Poli et al., 2019), and is based on a connectivity map realized after segmentation of clarified tissues and visualized with virtual reality. This method could also be applied for fused regional cerebral organoids or for connected organ culture with the brain one. A unified analysis tools platform would benefit simultaneously to the manufacturing process standardisation and 3D cultures research (summarized in Figure 2.8).

2.9 Conclusion

The use of cerebral organoids in laboratories has increased exponentially since their first creation in 2013. However, we observe in this review that actual tools to study images from these 3D brain cultures in all their dimensions suffer from some limitations. The structural variability occurring during maturation needs to be limited by improved protocols or by computational analysis solutions. The best combination of "clarification protocol - microscopic device" remains to be highlighted to acquire images from cerebral organoids that could be analyzed in all their dimensions. Specific tools need to be developed to improve the speed and the accuracy of their identification and quantification, but also to better understand their physiology and their entire 3D cyto-architecture. However, such an approach implies access to very large image datasets, which seems only possible when they will be stored in the "Organobanks".

As already mentioned by two other teams, and once the current limitations are overcome, the ideal platform would combine molecular/transcriptome and high throughput image analysis tools (Grenier et al., 2020; Poli et al., 2019). The first milestone of this kind of research was very recently published (Albanese et al., 2020; H. Renner et al., 2020). However, cerebral organoids dedicated image analysis tools remain to be developed, as summarized in Figure 2.8.

We are convinced that cerebral organoids coupled with high performance image analysis tools have the potential to highlight features others brain models are not able to show yet, and will help evaluating theories in the neuroscience field.

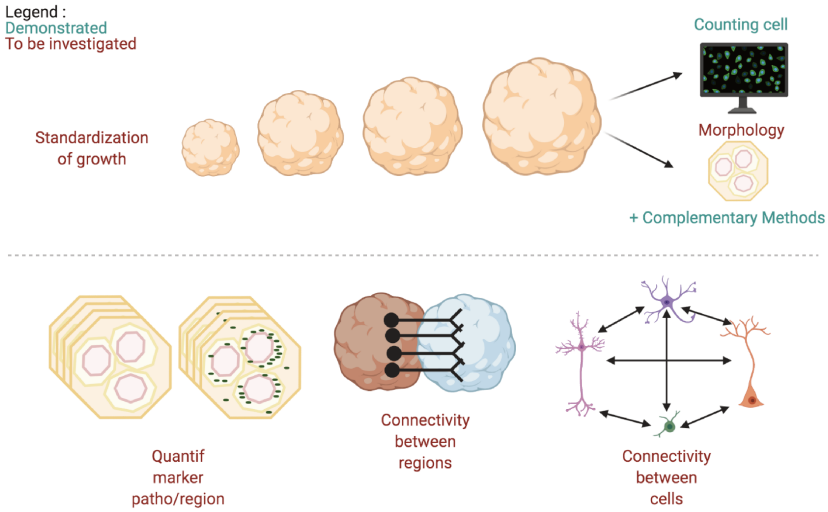


Figure 2.8: Quantitative and morphological expected tools to analyze images of 3D brain cultures.

Chapter 3

Generation

The scarcity of public datasets of annotated biomedical images remains an unresolved bottleneck to develop specialized and robust automatic analysis tools. Research groups do not usually share experimental data for privacy reasons. The high costs of equipment, long acquisition times, and necessary in-depth expertise can be a brake to acquisitions (Chakradhar, 2016). To benefit from the advances in deep-learning (DL) for automated image analysis, large training datasets are necessary. Moreover, original dataset constraints create a problem of class imbalance with deep learning training procedures. These problems are emphasized with small sets, reduced to a few images (Tajbakhsh et al., 2016).

A solution widely used in various domain is data augmentation (N. K. Singh et al., 2021). This concept gathers techniques used to increase the amount of data by adding slightly modified copies of already existing data or newly created synthetic data from existing data. Data augmentation aims to reduce over-fitting when training a machine or deep learning model. The main current applications in the biomedical field are in detection and classification,but has been used before for synthesis and reconstruction or detection and segmentations (Yi et al., 2019; Y. Chen et al., 2022). Usually three data augmentation strategies are described: classical, deep learning-based, or a combination of both (Chlap et al., 2021).

3.1 State of the art on data augmentation

3.1.1 Classical data augmentation

Classical data augmentations gather the oldest, and the most often used strategies, mostly because of high processing speed. These strategies are divided in six methodologies: geometric, color space, noise, filter, mixture (or patch), and erasing based.

Color augmentation gathers contrast transformation (increase, reduction, linear, gamma...), histogram equalisation or whitening of images (Mikolajczyk et al., 2018), while color space augmentation is based on RGB pixel values manipulation or hue transformation, or color jittering (Wu et al., 2015). Affine geometric transformations contain shear, scale, translation, reflection, rotation, flip-flop, skew, random crop, occlusion, edge enhancement (Mikolajczyk et al., 2018; Taylor et al., 2018). Others transformations include kernel filtration, patch or noise injection (such as Gaussian, Uniform, Salt and Pepper, sharpening, blur, or emboss) (Hussain et al., 2017; S.-H. Wang et al., 2021; Chlap et al., 2021).

Most of these transformations have been compared by Taylor et al., 2018 experiment on CIFAR10, showing a not so strong accuracy in classification tasks, with a maximal accuracy of 79%, obtained by cropping. A possible explanation of this result is the lack of variation in

transformed images, in comparison with the original ones. Other transformations have been proposed to overcome this limitation. Displacement based on mapping (a reference image and a second one which is deformed, to become similar to the first) (Krivov et al., 2018), diffeomorphic image registrations combined with affine augmentation (Nalepa et al., 2019) or atlas based (Tustison et al., 2019) are also in the spectrum of possible transformations, although they require a reference images or an atlas of the object. Spline interpolation is another mathematical transformation which smoothly deforms the image content. Two interpolation methods, thin and B spline, are common. The thin is the most use, however the content is sometimes unrecognizable (Chlap et al., 2021). Concerning the latest kind of classical transformations, they mix patches of images of various content or, erase randomly including important parts which do not allow to well train algorithms (Summers et al., 2019; Takahashi et al., 2020; Z. Zhong et al., 2020).

Recently, other classical approaches has emerged such as statistical shape models, which describe the shape variability inside a dataset with a model, used to generate deformations within the range of original parameters, with a dimensional reduction approach for instance. However, the variability in image intensities between protocols and acquisitions, and the absence of shape references of biological contents lead to errors (Bhalodia et al., 2018; Corral Acero J, 2019).

In summary, classical data augmentation are the most used and the fastest methods, but should be used with caution due to severe drawbacks such as: clone images with no variability, unrecognizable generated data, non natural biological shape in the content, need of a reference or a template and mathematically limited number of transformations.

3.1.2 Deep learning based data augmentation

Deep learning (DL) data augmentation approach automatically learn the representation of images and generate realistic but not clone images, and is generally used to increase the model generalities, synthesizing variable images, and reduce over-fitting during training by the increase number of images. DL strategies are in majority composed of four groups described below: feature space augmentation, adversarial training, generative adversarial networks (and variations), and neural style transfer.

Feature space augmentation

The principle of feature space augmentation consists in increasing not the input database, but the learned space by adding noise, interpolating, or extrapolating by a neighborhood search (Chawla et al., 2002; Nacereddine et al., 2012). An example of this data augmentation strategy is illustrated in Figure 3.2.

The deep learning architecture used for this approach is a sequential autoencoder. The autoencoder is constituted by an encoder which receive the input and transform it by applying one or more nonlinear parameters into a new representation of a lower dimension. The decoder also uses nonlinear transformations to reconstruct the original image with the lower dimensional representation. An illustration of the classical auto-encoder architecture is given Figure 3.1. Autoencoder are usually divided in many fields: sparse, denoising, contractive, etc.

During the forward step, the hidden states of the recurrent layers are propagated through the layer stack. The encoder hidden states at the final time step, called the context vector, seeds the hidden state of the decoder at its first time step. The sequential auto-encoder works similarly; however, the encoder and decoder use recurrent layers which can encode

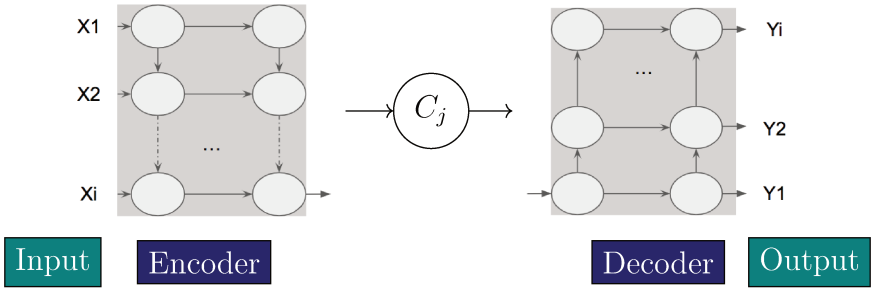


Figure 3.1: Auto-encoder architecture.(X_1,X_i) corresponds to input information, circle non linear transformed resulting vectors, C_j represents the context vector also called latent space and (Y_i,Y_A) the output information.

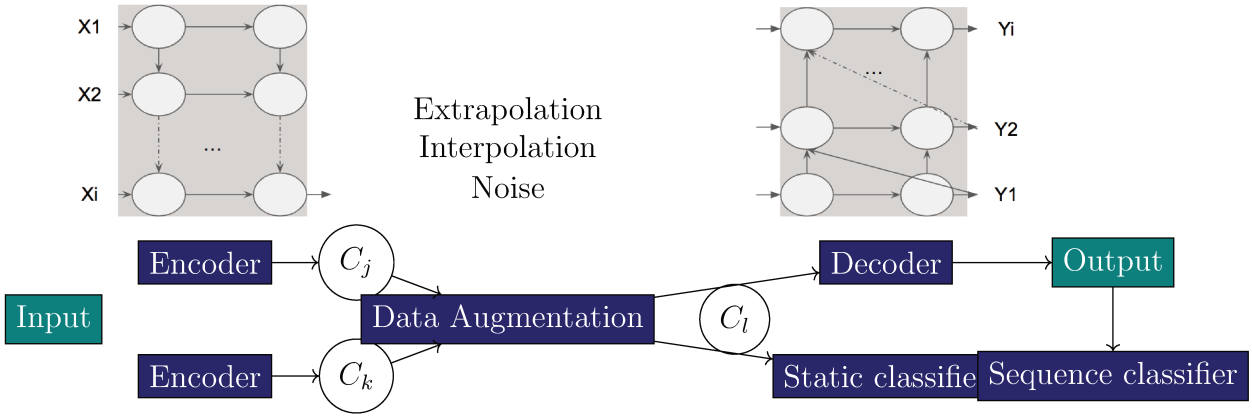


Figure 3.2: An example of feature space augmentation lying on a sequential auto-encoder illustration based on (DeVries et al., 2017). C_j, C_k and C_l represent context vectors, in blue are the process and in green images. Encoder architecture and the sequential decoder are illustrated in grey.

and decode variable length sequences. Indeed, the decoder learns on the context vector at each time step and not on the first step, which improves the reconstruction (Kyunghyun et al., 2014).

A limitation of this method comes from the proposed data augmentation, based on simple transformations: only a limited number of transformations can be calculated on the contextual vector.

Adversarial training: origin of Generative Adversarial Networks

Generative adversarial networks (GANs) are unsupervised deep learning-based architectures composed of generator (G) and a discriminator (D). The generator aims at creating visually realistic and natural images while the discriminator tries to decipher whether the result is generated or belongs to the original dataset. Created for the first time by Goodfellow et al., 2014, its surname is Vanilla due to its architecture, directly drawing samples from the original image distribution, without needing to model the underlying probability density function; An illustration is given in Figure 3.3.

The generative model (GM) maps the images into the space (z) by an objective function (F) while the discriminative model (DM) determines the probability for which a point from z belongs to the original dataset (o) or to the generated dataset (g). Training F increases the probability that the data synthesized is attributed to o. The probability of correct sample

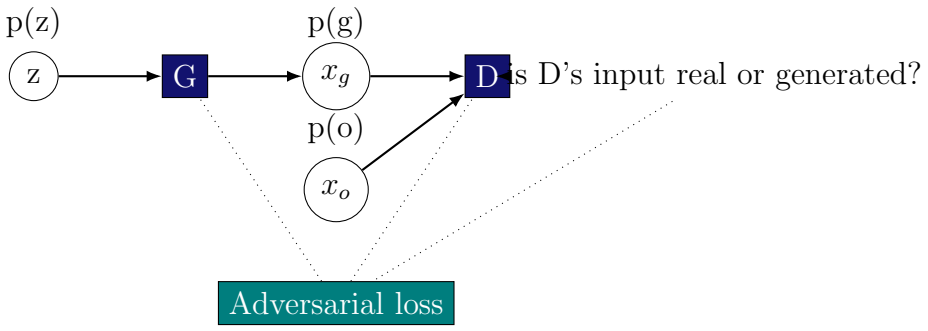


Figure 3.3: Vanilla Generative Adversarial Network from (Goodfellow et al., 2014)

labeling (belonging to generated g or original o) is maximised by D. Simultaneously, GM is trained to leverage the discriminator function expressed by:

$$\min_{GM} \max_{DM} F(DM, GM) = E_x \sim p_o(x)[logD_o] + E_x \sim p_z(z)[log(1 - D(G_z))] \tag{3.1}$$

Both networks are trained simultaneously with the same loss joint function, usually using binary cross entropy loss function. To explain equation 3.1, we define: H the cross entropy and x a sample which have 50% to belong to the probability to belongs to the original dataset p_o or, the generated dataset p_g and, suppose $y = 1$ if x is real or $y = 0$ if x is generated. Joint entropy for the discriminator is described as:

$$H(y, y') = H(y, D(x)) \tag{3.2}$$

$D(x)$ is the probability for which D is confident about x belongs to p_o. With the cross entropy formula, we obtain $H(y, D(x))$ with N the size of the dataset:

$$H(y, D(x)) = E_y[-logD(x)] = \frac{1}{N} \sum_{i=1}^{N} yilog(D(xi)) \tag{3.3}$$

We then split the sum in equation 3.3 in two parts, as each class possesses $N/2$ samples:

$$H(y, D(x)) = -(\frac{1}{N} \sum_{i=1}^{N/2} yilog(D(xi)) + \frac{1}{N} \sum_{i=N/2}^{N} (1 - yi)log(1 - D(xi))) \tag{3.4}$$

The first part of equation 3.4 represents the distribution of p_o and the second part the distribution of p_g. We then convert all the sums in expectations as all the yi could be validated, and 2 is a constant:

$$H(y, D(x)) = -(\frac{1}{2} E_x p_o(log(D(xi))) + \frac{1}{2} E_x p_g(log(1 - D(xi)))) \tag{3.5}$$

The p_g sample data (equation 3.5) being outputs from G, it can be replaced with D(G(z)), with z explained by p_z to obtain the discriminator loss function L_D which goal is to minimize itself:

$$L_D = \max_{DM}(E_x \sim p_o(x)(log(D(x_o))) + E_z \sim p_z(z)(log(1 - D(G(z))))) \tag{3.6}$$

The generator goal is to maximizes incertitude, proven to be equivalent to minimizing the Jensen Shannon divergence between $p_o(x)$ and $p_g(x)$:

$$L_G = \min_{GM}(E_z \sim p_z(z)(log(1 - D(G(z))))) \tag{3.7}$$

Table 3.1: Limitations of Generative adversarial networks and possible causes.

Drawback	Description
Collapse	generation of identical images as the input (fail in a local minima)
Noise/Blur	generation of noisy images or blurry when the training is not enough or when the generator is incertain when receiving the feedback
Overfitting	generated model contains more characteristics that the initial dataset
Vanishing gradient	the discriminator won, it determines better the false images and does not give enough feedback to the generator, so the generator fails to train due to the vanishing gradient
Divergence	model learns a pattern that provokes the generation of images more and more distant from the input

Coupling these two loss functions, we obtain the adversarial loss written in equation 3.1.

Compared to feature space augmentation, there is no restriction for the latent space size generation with GAN. However several drawbacks have been highlighted. Table 3.1 describes the most often observed.

To overcome these limitations, others GAN architectures have been proposed over the last years.

GANs architecture variations

Since its first creation, multiple GAN architecture variations have been proposed to generate and extend datasets (Y. Chen et al., 2022; Fernández et al., 2021; Lan et al., 2020). Among the various architecture variations, we only describe here the most known which allow to overcome some Vanilla's limitations. We also classify them in generator and discriminator objective variations.

Generator objective variations To stabilize the training of Vanilla GAN, one solution is to replace residual blocks architectures using fully connected layers by fully convolutional downsampling/upsampling layers such as DCGAN (Radford et al., 2015) (Figure 3.4). Pooling layers are replaced by strided convolutions for the discriminator and by fractional strided convolution for the generator. It contains BatchNorm for regulating the extracted feature scale (Yoffe et al., 2015), and LeakyRelu for preventing dead gradients (Maas et al., 2013). ReLU activation is used in generator for all layers except for the output, which uses Tanh and LeakyReLU activation in the discriminator for all layers (Radford et al., 2015).

Another generator architecture variation used in the literature consists in replacing the fully connected layers by an auto-encoder (Larsen et al., 2016; Kingma et al., 2013; Makhzani et al., 2015; Isola et al., 2017; Zhu et al., 2017) (Figure 3.5). Where in the Vanilla GAN, the generator transforms the z sample into g belonging to $G(g)$, the VAEGAN uses pixel-wise reconstruction loss to enforce the decoder part of VAE to generate structures to match the real images (Larsen et al., 2016). In this architecture, the GAN becomes conditional (cGAN) and the generation process drives the output to have certain properties, expressed as $xg = G(z, c)$ where c is an image. Another VAE from (Kingma et al., 2013)uses a KL divergence penalty to impose a prior distribution on the hidden code vector of the autoencoder. In Makhzani et al., 2015, the proposed AEGAN uses an adversarial training procedure by matching the aggregated posterior of the hidden code vector with the prior distribution and outperforms VAE architectures on test likelihoods of real-valued MNIST

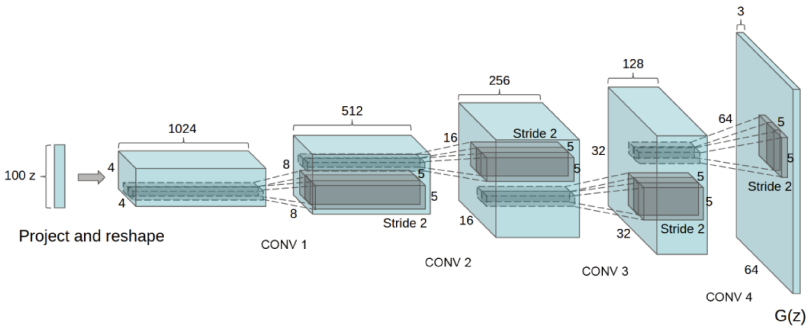


Figure 3.4: "DCGAN generator used for LSUN scene modeling. A 100 dimensional uniform distribution Z is projected to a small spatial extent convolutional representation with many feature maps. A series of four fractionally-strided convolutions (in some recent papers, these are wrongly called deconvolutions) then convert this high level representation into a 64 × 64 pixel image. Notably, no fully connected or pooling layers are used" from (Radford et al., 2015)

datasets. Concerning image-to-image translation neural network, pix-2-pix is the most known and presents generator updates (Isola et al., 2017). Indeed, this generator is a "U-Net"-based architecture to avoid vanishing gradient (Ronneberger et al., 2015b), which differs from the encoder-decoder original architecture, by applying skip connections between the mirror layers from the two networks.

However, pix2pix requires aligned training data. This is not the case with CycleGAN (Zhu et al., 2017) , which is made of two generators and discriminators. Synthetic result from a generation is given as an input to the second generator, and contains cycle consistency loss. However, this architectures is limited by the performance of the discriminative loss.

Discriminator objective variations The discriminator is also a structure which could spread limitations in the generative network. Two main strategies are generally chosen to optimize this part: varying the loss or the latent vector. A few teams start to modify the discriminator architecture.

To stabilize the training and avoid mode collapse, various losses for the discriminative network are proposed. In Nowozin et al., 2016, the interest of using variational divergence functions on the complexity and quality of the generative model is discussed. Kullback Leibler (KL), reverse LM, Pearson X2, Squared Heligner, Jensen-Shannon and GAN loss functions are compared, and the kind of activation layer best for each case, depending on the application, is discussed. As previously mentioned, the binary cross entropy loss (BCE), originally used by Goodfellow et al., 2014 is the most used function, but since 2016 an amount of new loss based upon regularisation loss or type loss has emerged. Among these losses, we will discuss the most popular such as the least square loss optimisation, which avoids vanishing gradient (Mao et al., 2017):

$$LD = \frac{1}{2}E_x \sim p_o(x)((D(xo) - l_o)^2) + \frac{1}{2}E_z \sim p_z((D(G(z)) - l_g)^2) \tag{3.8}$$

$$LG = \frac{1}{2}E_z \sim p_z(D(G(z)) - h)^2) \tag{3.9}$$

Where l_o and l_g are labels from original and generated images and h the hyperparameter G wants D to recognize. When l_o, l_g and h responds to $lo - h = 1$ and $lo - lg = 2$, the Pearson

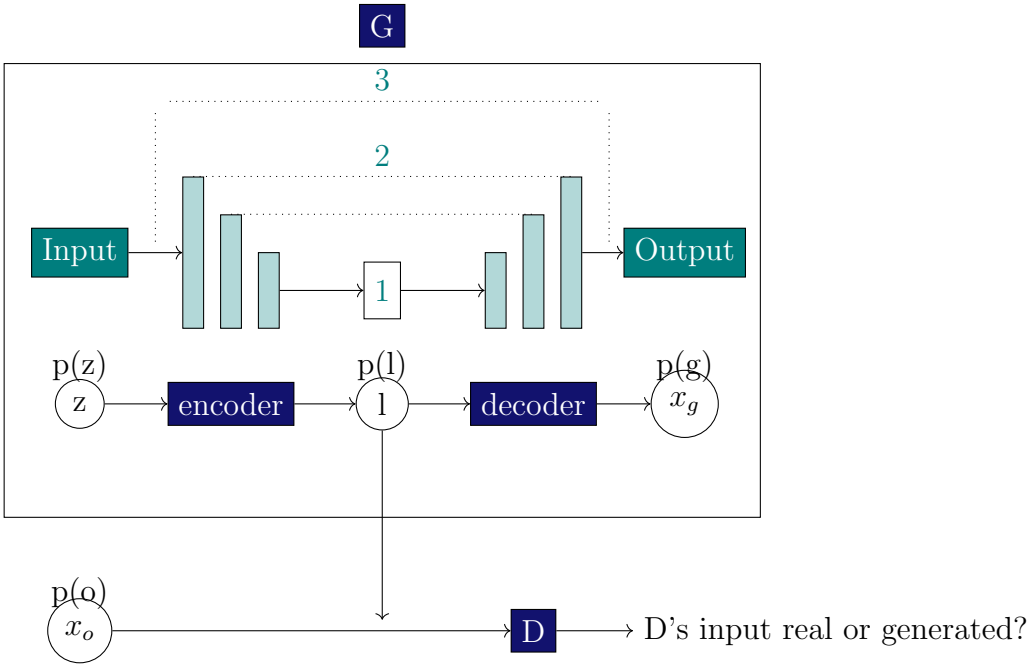


Figure 3.5: Classical Gans architectures based on auto-encoders. The basic architecture is given by the AAE from (Makhzani et al., 2015). Variations are given by numbers where: 1. KL divergence penalty to impose a prior distribution (Kingma et al., 2013); 2. U-Net architecture with skip connections instead of a classical auto-encoder architecture (Isola et al., 2017); 3. Reconstruction loss based upon pixel-wise (Larsen et al., 2016)

divergence between $p_o + p_g$ and $2pg$ is minimized. Thus large errors, far from decision boundaries, are penalized by D, and this provides sufficient gradient when G is updated, contrary to the original GAN (Goodfellow et al., 2014).

The earth-mover's (or Wasserstein) distance is also one of the most used loss functions (Arjovsky et al., 2017):

$$LD = E_x \sim p_o(x)(log(D(xo)) - E_z \sim p_z(log(1 - D(G(z))) \tag{3.10}$$

Wasserstein reflects distance even when p_o and p_g do not overlap, is continuous and is able to provide a valid gradient for training G. Thus, instead of using a binary classifier, D fits a Wasserstein distance and behaves as a regression task, so the terminal sigmoid layer is removed. This solves the vanishing gradient problem and the mode collapse.

Despite the great asset of these loss functions on the generative process, they still present limitations. Some articles try updating the content loss, which is particularly used in super-resolution generation of images or style transfer (Cheon et al., 2018; Kupyn, Budzan, et al., 2018; Y. Gu et al., 2020). The perceptual content loss is a L2 content loss instead of the L1 content loss.

$$ContL1 = |y - y'| \tag{3.11}$$

$$ContL2 = (y - y')^2 \tag{3.12}$$

Where y is an original tensor information and y' its prediction. One of the WGAN improvement consists in adding a perceptual (L2) content loss such as in Kupyn, Budzan, et al., 2018. Others use a penalty gradient applied on the Wasserstein loss (Gulrajani et al., 2017):

$$LD_{grad} = LD + \lambda E_x \sim p_x(x)(||\Delta_x D(G(z))||_2 - 1)^2 \tag{3.13}$$

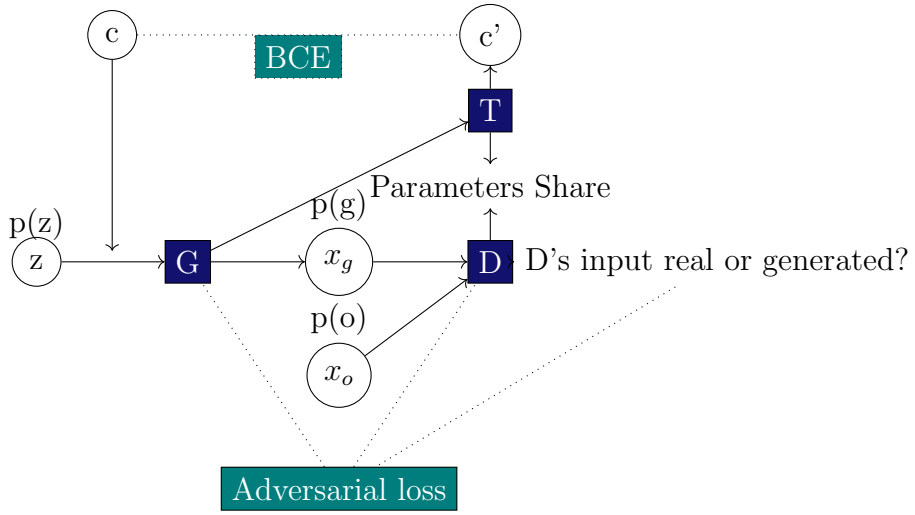


Figure 3.6: InfoGAN architecture. G and T acts like an autoencoder architecture which aim is to find the embedding G(z) with c conditions and minimizing the binary cross entropy between c and c'.

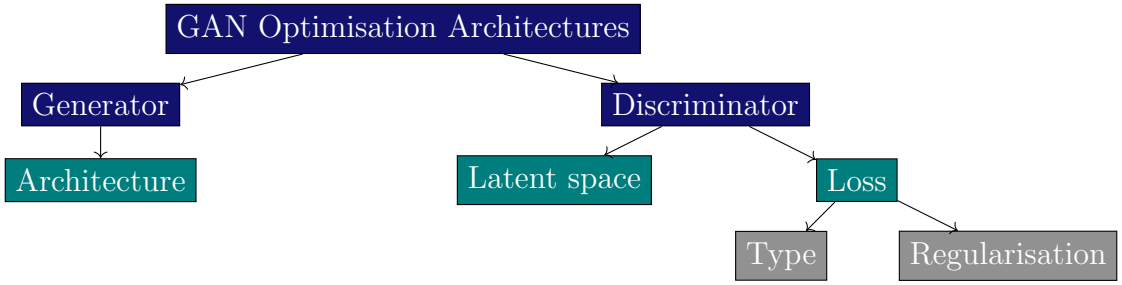


Figure 3.7: Global overview of GAN optimisations, split in discriminator and generator optimisation. However, other classifications exist.

Where λ equals 10 and p_x is a uniform sampling between paired point samples from the p_o and p_g distributions. In pix2pix discriminator they use a convolutional "PatchGAN" classifier, which only penalizes structure at the scale of image patches to capture local style statistics by applying a L1 instead of a L2 loss (which aim is to normalize) and avoids blur results (Isola et al., 2017).

Another discriminator optimisation consists in updating the latent space. Another conditional architecture with only one generator and discriminator, called InfoGAN, presents the advantage of being lighter than CycleGAN, while the output of the generator is given to feed the input at the subsequent iteration (B. Hu et al., 2019; Zhu et al., 2017). This architecture maximizes the mutual information between conditional variables and the synthetic data by introducing a transformer T to predict c' with a certain $G(z)$ fed with c conditions, see Figure 3.6.

Possible GAN optimisations detailed in this manuscript are summarized in Figure 3.7.Other GAN optimisations based on discriminator or generator networks are available; however, the efficiency of a particular GAN optimisation depends on the application it is used on. An efficient GAN architecture is characterized by a good image quality, dataset diversity and avoiding vanishing gradients (Zhengwei Wang et al., 2021).

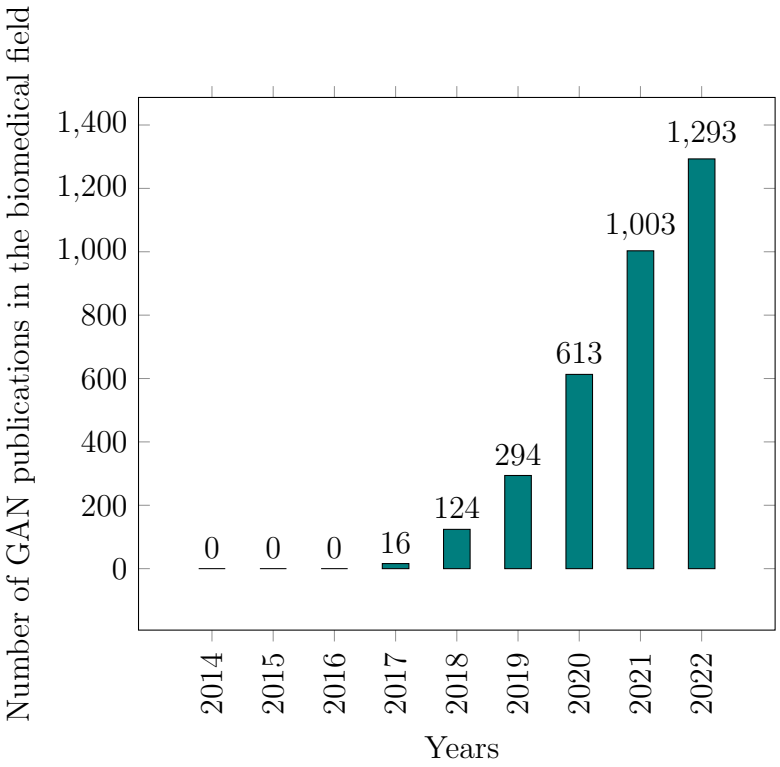


Figure 3.8: Number of publications on GAN's applications per year in the biomedical field since 2014.

3.1.3 Data Augmentation in the biomedical field

Generalities on data augmentation for biomedical datasets

The first GAN application on a biomedical dataset is from (Shang et al., 2017). The approach is validated based on a combination of GAN and multimodal autoencoder to implement missing values of two genetic databases, and uses the MNIST dataset (a dataset of digits written by hand and largely used to train, developed and test deep learning models). Since this proof of concept, the number of articles using GAN for biomedical applications has kept increasing. In total, 3079 related articles are published as of today, according to PUBMED see Figure 3.8.

Two kinds of uses are mentioned in the biomedical literature, linked with the generator or discriminator objective: the first aims to study the structure of trained biomedical data and avoid patient/biomedical image scarcity; the second aims to regularize or detect abnormal data. A summary of the major GAN applications in the biomedical field, is given in the next section.

GAN's biomedical applications

The biomedical applications of GANs covers several domain. The main one is the detection of biological contents in images (more than 20% of published applications), followed by classification tasks (20%), segmentation (17%) and reconstruction (15%). Registration and synthesis share almost the same proportion of GAN use (12%), see Figure 3.9.

The major image acquisition domain using GAN is X-rays followed by ultrasound, MRI and CT, and histopathology (with a little more than 5%). The other microscopic or medical image acquisitions consist in less than 5% (Figure 3.10).

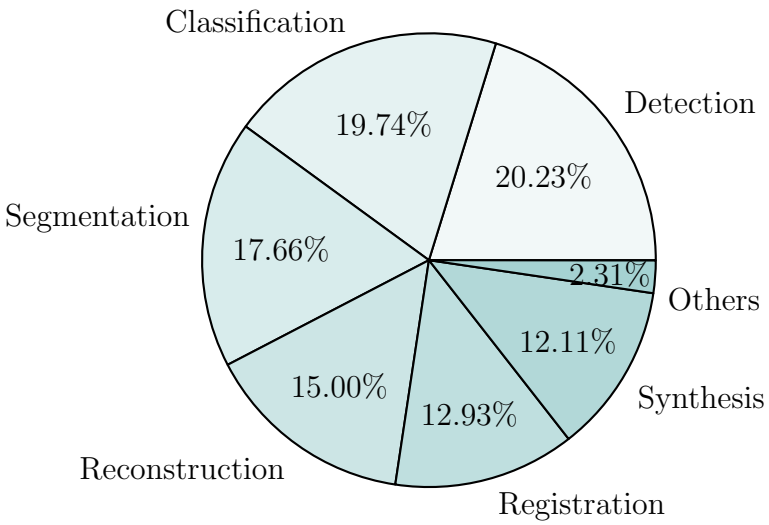


Figure 3.9: Publication on GAN's biomedical field applications. We only count the number of publications in the main domain and show it in percentages. Registration is also called mapping in certain publications.

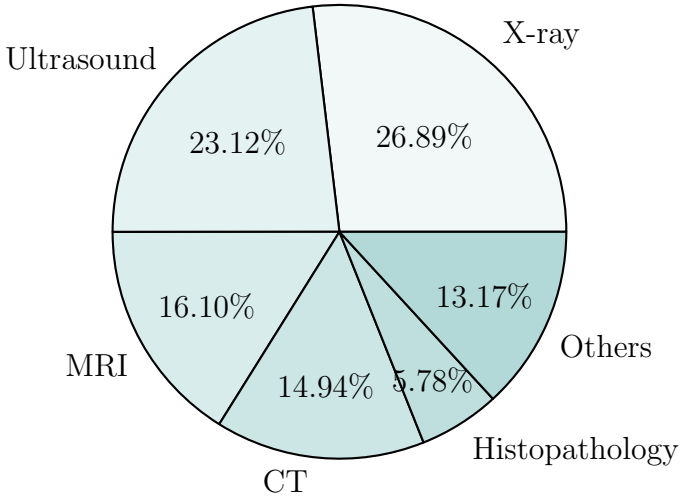


Figure 3.10: Publication on GAN's image acquisition applications. We only count the number of publications in the main domain and show it in percentages. Others contains PET (3.25%), Clarification (1.78%), Fundus (1.75%), Fluorescence (1,33%), Mammogram (0,78%), Bright-Field (0.68%), Dermoscopy (0,26%) and others (3.34%).

Limitations of GANs for biomedical applications

The use of GAN's in the biomedical domain suffers from two major inconveniences: the inter-image from patient or biological model variability, and the few examples of images inside each database. These issues render GAN's use difficult, notably for their efficiency. Indeed, the efficiency of GANs increases with the network depth, which requires more data than lighter architectures. This constraint, not always applicable for biomedical dataset, forces researchers to optimize architectures and develop application-based solutions: for a biological domain or for a particular image acquisition (Zhengwei Wang et al., 2021).

The second major problem lies in the validation of synthetic biological images.A particular problem encountered with GANs is the mislabeling of results to another category due to the added noise at the input, which is a major problem for preclinical or clinical data (Shorten et al., 2019). The validation of GAN synthetic images are another emerging issue. Manual validation consists in validation of each image by an expert, while similitude or qualitative metrics are used for automatic solutions. The first one is time consuming whereas the second one is controversial (Borji, 2018): which kind of metric can highlight similitude between biological contents and realistic image quality? Data augmentation for brain organoids does not derogate to these limitations.

3.1.4 GAN for brain organoid culture images: our positioning and strategy

The first question addressed in this part is: which kind of natural images have to be augmented in the brain organoid field to extract their shape and characterize their development? As previously discussed in part 2.7, the developmental bright-field images of brain organoid from (Gomez-Giro et al., 2019) could be the best compromise between advantages and drawbacks see Paragraph 2.7. As mentioned in part 3.1.3 and Figure 3.9, segmentation of biological contents are the third GAN application in the biomedical field. However, bright-field data augmentation by GAN's are not well known (less than 1% of GAN applications in the biomedical field, see Figure 3.10). Thus, attempting to identify a strategy to augment brain organoid bright-field images in order to segment them and characterize their developmental shape should be validated on other bright-field brain organoid datasets (and not on midbrain or cortical organoids, see Annexes) to allow a generalisation of GAN use for this particular microscopic acquisition.

The second question addressed is: what kind of strategy should be used to increase image datasets, in order to generate the most natural synthetic brain organoid image? This question can be divided in four sub-questions:

- Which GAN architecture fits the best such data augmentation application?
- Which kind of loss optimisation can improve brain organoid natural generation?
- Which kind of noise input generates the best microscopic features?
- How to validate these synthetic biological images?

The strategies adopted to answer these questions are summarized in the Figure 3.11.

To answer these questions, we compare GAN architectures, then update the loss function on the best one, input various noises types, and we finally compare automatic (metric) versus psychovisual (Human expert) evaluation of each image. The methodology, results and discussion are summarized in the four following subsections.

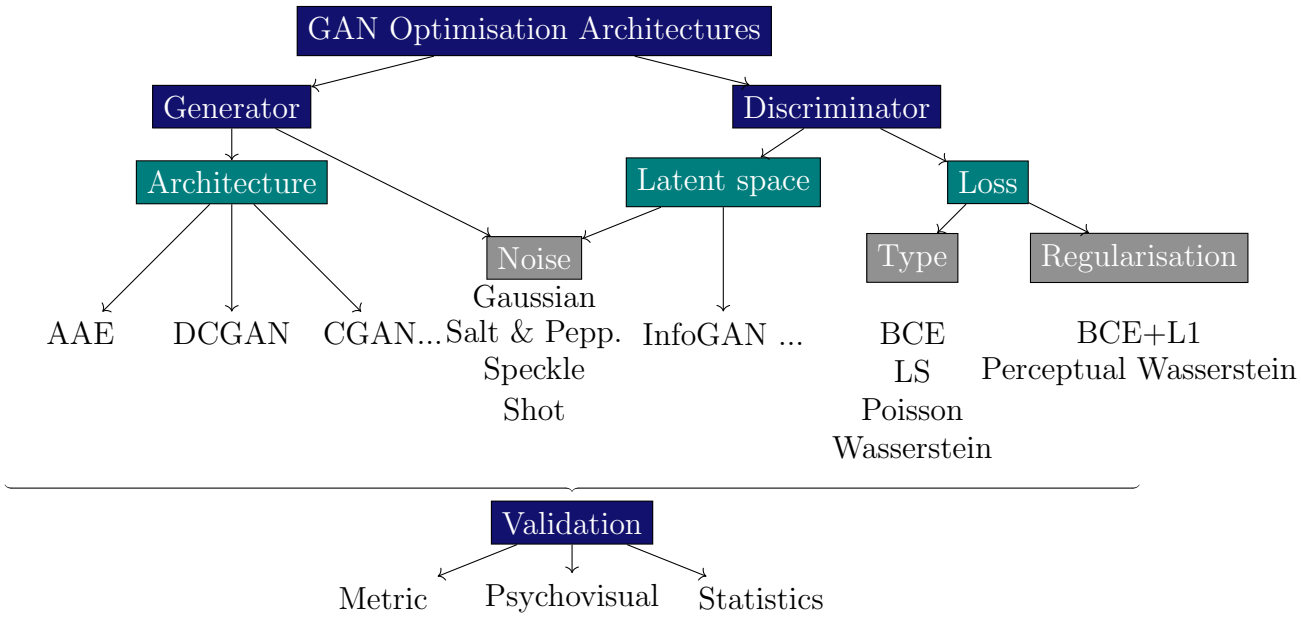


Figure 3.11: Global overview of GAN optimisation strategies we choose to develop. This figure is an extension of the Figure 3.7.

3.2 GAN architectures for bright-field brain organoid image generation

This section is almost entirely published in Brémond Martin et al., 2021a in VISAPP conference in 2022. The unpublished parts address dimensional reduction methods, results and discussion for GAN architecture comparisons and collapse figure.

The contribution of this section is to qualitatively and quantitatively investigate the influence of various GAN-based approaches in the specific case of bright-field brain organoid image generation using quantitative metrics from the literature and a dimensional reduction of parameters.

3.2.1 Methods

Resources Our dataset is composed of 40 images from an open access database (Gomez-Giro et al., 2019). Pathological (N=20) and healthy (N=20) CO were digitized using a bright-field microscope, over 3 days. The grayscale images are 1088×1388 pixels. However, to compare several networks within a reasonable time, the input images are cropped and resized to 250×250 pixels, maintaining the original proportions.

Resources are described in part 2.7.

Generative Adversarial Networks To find the most suited GAN (described in section 3.1.2), we consider five of the most used architectures to increase the dataset: **GAN** (Goodfellow et al., 2014) is the original implementation; **CGAN** (Yi et al., 2019)gives the correct label (physiological or pathological) as a generator input; **DCGAN** (Yi et al., 2019)is constituted by a convolutional neural network instead of the generator; **INFOGAN** (Yi et al., 2019)uses the generated images at an epoch to train the subsequent; **AAE** (Makhzani et al., 2015) uses an autoencoder as a generator.

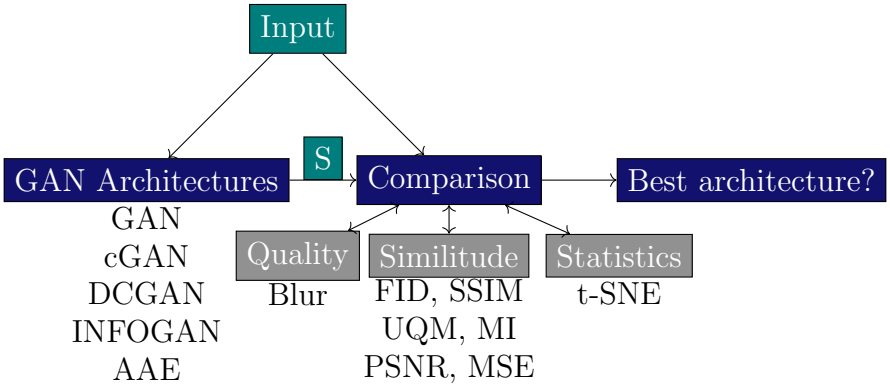


Figure 3.12: Procedure of GAN architecture comparison for the generation of bright-field brain organoid images. S corresponds to synthesised images.

During a 1000 epoch training, input images (size of 250×250 pixels) are used to generate synthetic images. In this work, the original 40 images of the dataset are used to generate 40 synthetic images for a better follow-up by each architecture. The number of images generated are chosen to guarantee no mode collapse. We follow the procedure explained in Figure 3.12.

Metric evaluation Six metrics are used to compare the similitude of the synthetic images generated by the AAE to the original dataset. A blur metric is used to evaluate the quality of these synthetic images.

The Frechet Inception Distance (FID) is calculated between two groups of images (Heusel et al., 2017).This score tends towards low values when the two groups (original O or generated G images) are similar, with μ the average value of the pixels of all images of a group, and Σ the covariance matrix of a group:

$$\mathrm{FID}(O,G) = |\mu_O - \mu_G|^2 +^T (\Sigma_O + \Sigma_G - 2(\Sigma_O \Sigma_G)^{\frac{1}{2}}) \tag{3.14}$$

The Structural Similarity Index (SSIM) is calculated using luminescence, contrast and structure between two images o and g belonging respectively to O and G (Zhou Wang, Bovik, et al., 2004).

$$\mathrm{SSIM}(o,g) = \frac{(2\mu_o\mu_g + c_1)(2\sigma_{og} + c_2)}{(\mu_o^2 + \mu_g^2 + c_1)(\sigma_o^2 + \sigma_g^2 + c_2)} \tag{3.15}$$

where σ represents the standard deviation, c_1 is a constant that ensures the luminance ratio is always positive when the denominator is equal to 0, and c_2 is another constant for the contrast stability. The SSIM ranges between 0 (no similitude) and 1 (high similitude).

The Universal Quality Metric (UQM) is based on the calculation of the same parameters as SSIM (Zhou Wang and Bovik, 2002). UQM ranges between 0 and 1 (1 being the highest quality):

$$\mathrm{UQM}(o,g) = \frac{4\mu_o\mu_g\mu_{og}}{(\mu_o^2 + \mu_g^2)(\sigma_o^2 + \sigma_g^2)} \tag{3.16}$$

Entropy-based Mutual Information (MI) measures the correlation between original and generated images and ranges between 0 (no correlation) and 1 (high correlation) (Pluim et al., 2003):

$$\mathrm{MI}(o,g) = \sum_{o \in O} \sum_{g \in G} P(o,g) log \frac{P(o,g)}{P(o)P(g)} \tag{3.17}$$

where $P(o, g)$ is the joint distribution of o belonging to O and g from G.

The Mean Square Error (MSE) between an original image and a synthetic image is calculated as:

$$\text{MSE}(o, g) = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} (o(i, j) - g(i, j))^2 \tag{3.18}$$

The Peak Signal to Noise Ratio (PSNR) indicates a high signal power against noise, as used in M. Jiang et al., 2021. High values correspond to qualitative images. Pixels in images are ranked between 0 and 255, so the maximum pixel value of an image is noted $max(o)$ and equals at most 255.

$$\text{PSNR}(o, g) = 20 \log \max(o) - 20 \log \text{MSE}(o, g) \tag{3.19}$$

To evaluate image quality, we calculate the blur index based on local image variance (Tsomko et al., 2008), where low score stands for a sharp image. In the following equation, the size of the image is (m,n), the predictive residues for a given image pixel are $(p(i, j))$ and their median $(p'(i, j))$:

$$\text{Blur} = \frac{1}{m(n-1)} \sum_{i=1}^{m} \sum_{j=1}^{n-1} [p(i, j) - p'(i, j)]^2 \tag{3.20}$$

The FID is designed to compare groups of images. We thus successively compare each group of synthetic images with the original input images. The FID reference range is calculated on the original image developmental stages. The SSIM, UQM, PSNR, MI and MSE are designed to compare two images. For each group of synthetic images (all 6 losses), we successively compare every image with each original image and then compute the average of these 40×40 values. We also calculate these values on all pairs of original images to compare the results to the original range. The Blur index is calculated on individual images. We store the minimum and maximum value of this index for the original images and the average value per loss for the synthetic images.

Dimensional Reduction The dimensional reduction goal is to observe in the same statistical space if, for each optimisation, generated image representations are close or far from the original image representations. We choose to perform a t-distributed Stochastic Neighboor Enbending (t-SNE) dimensional reduction. Contrary to others dimensional reduction methods, t-SNE preserves the local dataset structure by minimizing the divergence between the two distributions with respect to the locations of the points in the map. To avoid subjective or calculated indexes, we perform t-SNE directly on images features extracted from the GAN networks. t-SNE is constituted with Stochastic Neighbor Embedding where first an asymmetric probability (p) based on dissimilarities (symmetric) is calculated between each object (x_i), and its probably neighborhood (x_j) (Hinton et al., 2003). The effective number of local neighbors called perplexity (k) is chosen manually:

$$p_{i,j} = \frac{\exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma_i^2}\right)}{\sum_{k \neq i} \exp\left(-\frac{\|x_i - x_k\|^2}{2\sigma_i^2}\right)} \tag{3.21}$$

The larger the perplexity, the larger the variance of a Gaussian kernel used to have an uniform induced distribution. We therefore choose the maximal value possible which is 80, the number of individuals in our dataset. To match the original $(p_{i,j})$ and induced

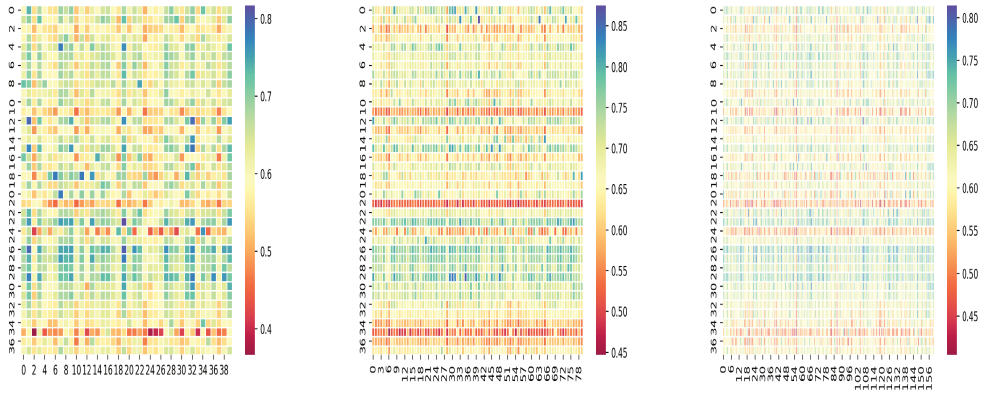


Figure 3.13: Mode collapse for each generation with a GAN architecture. From left to right: with 40, 80 and 160 image generation. Mode collapse occurs when the model only fit a few data distribution samples, and ignore the rest of them. All the SSIM values up to 90 are considered as mode collapse (in blue).

distributions ($p'_{i,j}$) in a low dimensional space (the enbending aim), the objective is to minimize the Kullback–Leibler (KL) cost function:

$$C = \sum_i \sum_j p_{i,j} \log \frac{p_{i,j}}{p'_{i,j}} \tag{3.22}$$

This minimisation allows t-SNE to preserve the dataset structure contrary to other dimensional reduction methods (as Principal Component Analysis). Then, Student t-distribution with one degree of freedom is used to avoid the crowding problem (Maaten et al., 2008).

We use a momentum term to reduce the number of iterations required (set at 1000 iterations at the beginning)The map points have become organized at 450 iterations in a scatterplot. Each point in the map corresponds to the feature vector while the axes are the embedding following the similarity properties i.e. the neighborhood of points. Each run of the t-SNE algorithm generates a different setting of the scatterplot. Even if point locations are different, the grouping remains similar. We launch the t-SNE between original and all the generated features 10 times to validate the similar grouping. We retrieve the KL divergence values between original and each generated distributions which could indicate a degree of similitude. A low KL divergence means that the two distributions are close.

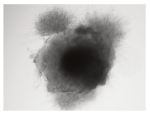
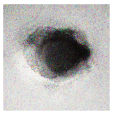
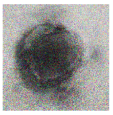
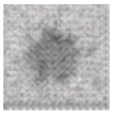
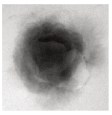
3.2.2 Results

We aim at generating qualitative images of cerebral organoid by GAN strategies to increase the open-source dataset (Gomez-Giro et al., 2019).

To determine the maximum number of images generated without collapse, we calculate the SSIM between original and generated images. The summary of GAN collapse generation with 40, 80 or 160 image generations can be observed in Figure 3.13.

We observe that the maximal similitude without creating a twin content is 0.90 (the maximum of similitude between two original images is only of 0.87). Generated images are at the minimum 45 % similar to original images. When we double the generative process, some identical images appear. We choose to generate only 40 images for each case in the testing phase to avoid these duplicates. To verify the best suited GAN architecture for cerebral organoid bright field images, we first compare the original images and the ones generated using the five architectures.

Table 3.2: Image quality and similitude of cerebral organoids generated by various GAN architectures. Scores within the original range are in gray, and best values are displayed in bold.

metric	best	Original 	GAN 	CGAN 	DCGAN 	INFOGAN 	AAE 
FID	low	0.47 – 0.80	2.02	2.13	≥ 4	2.89	**1.20**
SSIM	high	0.65 – 0.71	0.12	0.12	0.27	0.10	**0.63**
UQM	high	0.63 – 0.87	0.79	0.80	0.69	0.66	**0.81**
MI	high	0.21 – 0.47	0.17	0.24	0.25	0.19	**0.37**
BLUR	low	0.10 – 86.28	2504.24	7561.47	704.47	724.38	**135.93**
PSNR	low	11.9 – 16.6	**12.16**	12.64	28.35	28.35	12.89
MSE	low	93.25 – 106.23	105.41	103.07	107.12	107.14	**102.72**

In Table 3.2, sample images produced with GAN, CGAN and AAE are the most resembling images compared to the originals. Mode collapse is the most seen in GAN and CGAN architectures as reported in the literature. We can observe also a strong noise for these two architectures results with a white imprint around the shape of the organoid in the GAN case. While AAE generated images are characterized with blurry contours, DCGAN and INFOGAN generate a divergent background making the images difficult to exploit. To verify these observations, we calculate qualitative and similitude metrics, introduced in the section 3.2.1, by pairing first original images and then original and generated images.

Table 3.2, presents these results. For the output images, we underline the metric values within range of the original images. We observe only a low proportion of architecture metric within the original range. Indeed, only the AAE and the CGAN answer to only four metrics (UQM, MI, PSNR, MSE) on the seven calculated and only the UQM is within the range of original ones for all the architectures. Regarding FID, SSIM, UQM, and MI scores, AAE generate the most comparable images to the original ones.

In term of quality, this architecture generate the sharpest images, even if the blur index is higher than original images indexes. All the architectures express MSE of between the minimal and maximal values of this metric calculated for the original images. However, regarding the PSNR only the GAN, CGAN and AAE produce images with a score of between the original images limits. To summarize, according to the metric values, AAE is the most suited architecture to generate cerebral organoid images.

Dimensional Reduction To analyze all at once the similitude and the variability of the generated images with various GAN architectures, we study images in the same statistical space. Dimensional reduction is implemented on the extracted features of 20 synthetic images from each architecture (the minimal number of images without mode collapse for one architecture) in the generative process with the dimensional reduction methodology called t-SNE see Figure 3.14.

Feature space representation of original images (in black) and generated (others colors) are in the same range. Particular original images are grouped in the bottom of the representation and only ponctual generation are in their space. cGAN and GAN representations seem as spread as the original images, while AAE representations are at the center of the group. We observe that point representations of two generative networks are distant from

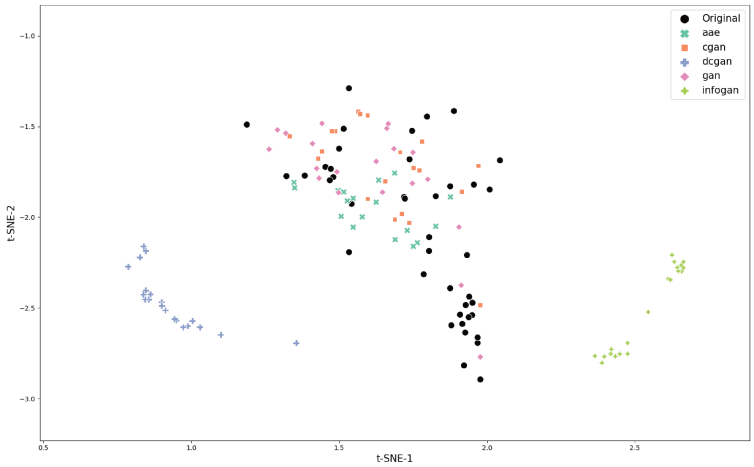


Figure 3.14: t-SNE representation of GAN architecture comparisons.

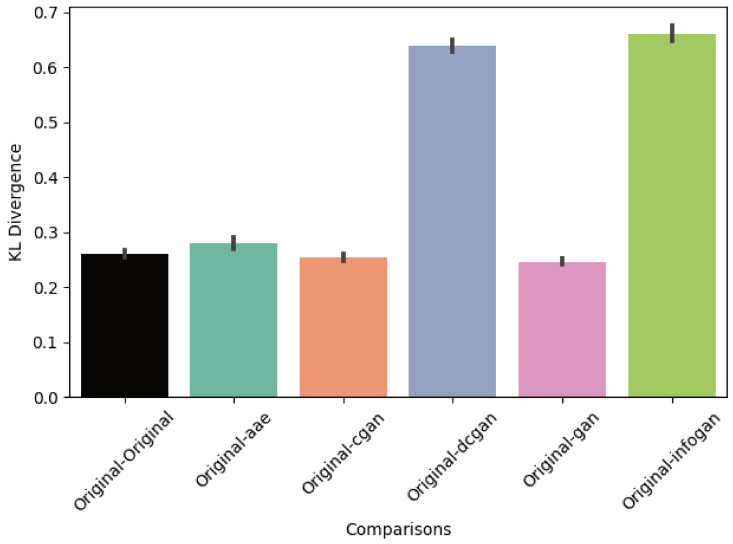


Figure 3.15: Kullback Lieber divergences calculated on t-SNE representations.

the others representations: InfoGAN and DCGAN, which is consistent with qualitative and similitude calculations.

We compare Kl divergences between original and generated representations from 10 t-SNE launches and plot the results in Figure 3.15. This bar-graph suggests DCGAN and INFOGAN from point cloud t-SNE representation diverge from Original representation. These results highlight cGAN, GAN and AAE architectures can bring various and maybe complementary information during the generative process.

3.2.3 Discussion

This work presents for the first time to our knowledge data augmentation of cerebral organoids bright-field images, using GAN and comparisons between different networks architectures. We avoid mode collapse by stopping the generation with a limited number of qualitative images. The AAE architecture outperforms other GANs architectures chosen in this work qualitatively and quantitatively based upon metrics. Dimensional reduction strategy shows two architectures generate foreigner synthetic images, not represented in the

same statistical space, while three are within the original range representations. Statistically, cGAN, GAN and AAE feature space are in the space of original image and suggest these images could be used in other tasks.

Mode collapse is avoided by generating at each step a few sampling of images. We observe since the doubling of the original dataset a lot of mode collapse with GAN and, and we decide to compare every architecture only on the same effective as the original dataset. In other published works, the generation still give a considerable amount of images (hundreds to thousands of images (Sandfort et al., 2019)). However, their relevance is to discuss, as they are almost identical and do not provide useful information for resulting model if used for segmentation. While the consensus is to use amounts of data to train deep learning based tasks, we thought, maybe for the first time of GAN generation, generating only few examples more representatives to the original dataset, but not identical, could help to these tasks.

Synthetic images generated with AAE are coherent with original dataset quality contrary to other architectures, with almost no collapse mode but also adding a sought diversity similar to the acquisition of the original dataset. This architecture has already show its advantages in data augmentation (Makhzani et al., 2015). The particularity to update the generator in an auto-encoder structure helps to impose a specific distribution on the latent variable in its encoding layer (Makhzani et al., 2015). An update of this architecture (for example replacing the autoencoder part by a U-Net) may improve these results(Isola et al., 2017). However, the major inconvenient with this architecture is the blurry content, which has already been described in Goodfellow et al., 2014; Larsen et al., 2016; Zemouri, 2020. Second, we chose to focus only on the use of five architectures to generate our images, we could compare these results with other types of GANs such as CycleGAN or PixtoPix (T. Zhou et al., 2017; Isola et al., 2017).

The synthetic images from an AAE are not in the range of original Blur metric, however, they give the best result for this metric. This inconvenient render the biological image not natural, or seems like the microscopic device acquisition has to be transformed to be analyzed (Jasim, 2019), so not useful in the biomedical field. We think a discriminative loss update could maybe help with this de-bluring task (Kupyn, Budzan, et al., 2018; Gulrajani et al., 2017; Isola et al., 2017; Mao et al., 2017).

Statistically feature generated images from three architectures are in the space of features from original images encoded. Statistical results from Figures 3.14 and 3.15 are consistent with qualitative and similitude metrics calculated in Table 3.2. Indeed, GAN, cGAN, and AAE, which have similitude metrics in the range of the original are also the three architectures containing their feature representations in the range of original. However, as mentioned before, if cGAN and GAN are spreading in the original feature spaces, AAE are grouped in the center of the representation which could be a problem for the variety of the synthesis: some original space are not represented by the AAE. We think a variation of noise given in the generative process could help to generate some diversity with this architecture (Sonderby et al., 2016; Salimans et al., 2016; Feng et al., 2021). Indeed, the microscopic acquisition of bright-field images are characterized by shot noise, while the usually noise used in a GAN is a Gaussian one (Boyat et al., 2015; Goodfellow et al., 2014) The other idea could be giving to the input a label representing the physiological or pathological state of input images such as in cGAN (Yi et al., 2019).

Two architectures provides non natural images, worst results in term of similitude and quality, and all features representations are totally out of the original ones. For DCGAN, we could explain the results by the fact with have a small dataset, while it requires a lot of images and deep architecture to render better results than those obtained in our work

(Radford et al., 2015). The only optimisation could be to increase the number of images inside the dataset or increase the generative steps. For INFOGAN, these results could be explained by the learning of a small part of the image, and learning only this pattern (which feed the subsequent step) (B. Hu et al., 2019), the generation became biologically not sense. We think these two architectures should not be applied on small datasets of biological bright-field images in the future.

There is still room for improvement in the proposed GAN architecture strategy evaluation. First, to propose a quantitative evaluation of the generalisation of the results obtained on cerebral organoid bright-field images: we would like to use this methodology on others bright-field biomedical images (Nickels et al., 2020) and Table 2.8. The best way to evaluate synthetic images and, compare various results of GAN optimisations is still a challenge (Alqahtani et al., 2019). Indeed there is no commonly approved specific metric to evaluate whether GAN-generated synthetic images can be considered as natural (Borji, 2018) and statistical space comparisons with a dimensional reduction strategy do not allow us to elect a particular architecture as the best in this work (it highlight three of them). Biological experts aim at psychovisually evaluating the generated images, and strengthen the quantitative evaluation proposed. In particular, we project to validate the suitability of the metrics we use and observe the training effect on the segmentation task with only validated images by biological experts and maybe highlight a metric combination which could replace the time consuming psychovisual evaluation.

For the first time to our knowledge, we compare GAN architectures for brain organoid Results remains exploratory with the mentioned small dataset we used. Improvements could be augmenting the number of input images for all of the architectures, increasing the training time, update the auto-encoder architecture in the generative part, or update the discriminator loss of an AAE, which provide the best results in term of quality, similitude and a same feature space with the original dataset. Indeed the major inconvenient highlighted by this work is the blurry content in brain organoid synthetic images generated with an AAE. The effect of the loss variations on this architecture could be of interest.

3.3 AAE loss optimisations

This part is published in Brémond Martin et al., 2021a for the VISAPP conference of 2022. Discriminative loss Figure is a supplementary work not published.

In this work, we improve the best GAN architecture (AAE, already described in the previous section 3.2) to generate cerebral organoid bright-field images. If the loss effect has already been explored for others biological models in MRI (Lv et al., 2021), to our knowledge, there is no systematic comparative study proposed in the specific context of CO bright-field image generation that gives a quantitative appreciation of this effect. In particular, we are interested in choosing a loss while guaranteeing good quality of the generated data, as well as a good variability of images obtained compared with the inputs in order to improve characterisation tasks.

3.3.1 Methods

We proceed as in the previous section, with the same resources described in part 3.2.1 and the AAE previously selected by the architecture comparison in Section 3.2. We use the same

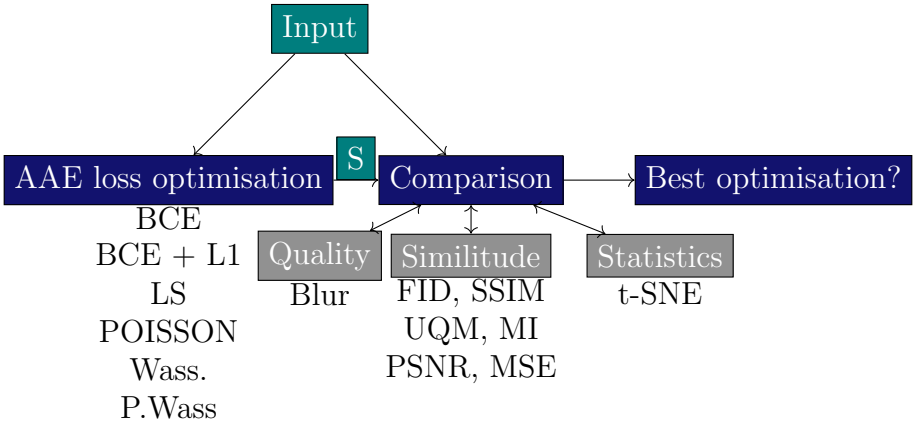


Figure 3.16: Procedure of AAE loss optimisation comparison for the generation of bright-field brain organoid images.

metric evaluation and dimensional reduction described in respectively in paragraphs 3.2.1 and 3.2.1. The procedure of this work is described in Figure 3.16.

Loss Optimisations

The images generated by the AAE architecture we use are somewhat blurry. To overcome this phenomenon we study how the discriminator loss can influence the quality of the image generation. We consider six losses: the Binary Cross Entropy (BCE) which is most commonly used in GANs and five other losses which are specifically known to improve the contrast or sharpness of the generated images.

BCE is the most commonly used loss for GANs and the baseline of this work. It is calculated by:

$$\text{BCE} = -\frac{1}{n}\sum_{i=1}^{n}(y_i(\log(y_i'))) - ((1-y_i)(\log(1-y_i'))) \tag{3.23}$$

with y the real image tensor and y' the predicted ones (Makhzani et al., 2015) and n the number of training.

Summing the L1 norm to the BCE is reported to reduce over-fitting (Wargnier-Dauchelle et al., 2019). We hypothesize this norm could improve the quality of the generation as reported in image restoration tasks which does not over-penalize large errors (H. Zhao et al., 2017).

$$\text{L1} = \frac{1}{n}\sum_{i=1}^{n}|y_i - y_i'|_1 \tag{3.24}$$

$$\text{BCE}_{\text{L1}} = \text{BCE} + \alpha\text{L1}. \tag{3.25}$$

α is set to 10^{-4}, as in the original paper.

The least square loss (LS) is reported by Mao et al., 2017 to avoid gradient vanishing in the learning process step resulting in better quality images:

$$\text{LS} = \frac{1}{n}\sum_{i=1}^{n}(y_i - y_i')^2 \tag{3.26}$$

A Poisson loss is used in Wargnier-Dauchelle et al., 2019 to improve the sensitivity of a segmentation task:

$$L_{Poisson} = \frac{1}{n}\sum_{i=1}^{n}(y_i' - y_i)\log(y_i' + \epsilon) \tag{3.27}$$

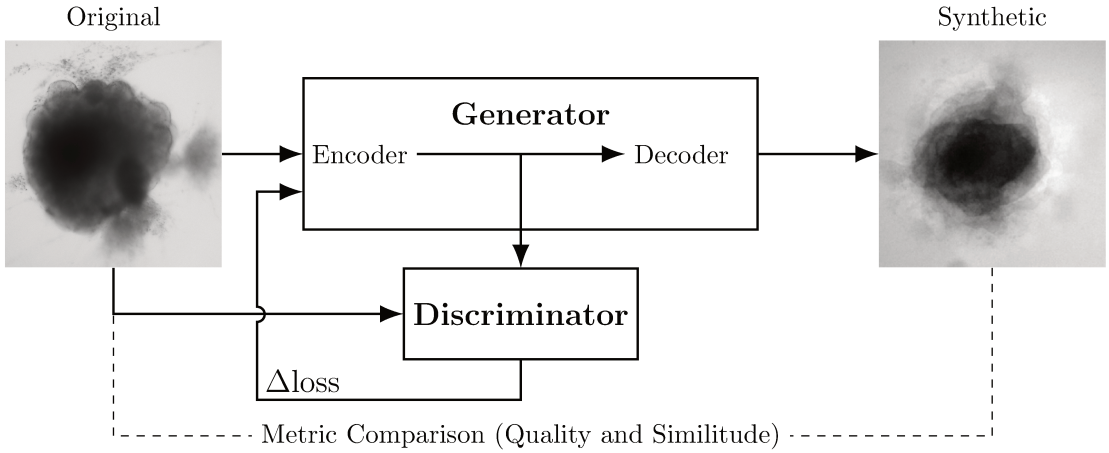


Figure 3.17: Experimental scheme of AAE supporting data augmentation of cerebral organoids bright-field images. The generator tries to persuade the discriminator that it has generated a true and slightly variable image of input dataset. The discriminator tries to find the true ones. They improve each other by backpropagation, formulated by an objective function based on a loss. Losses variations implemented in this article are symbolized by Δ. Input image is from (Gomez-Giro et al., 2019).

where ϵ is a regularisation term set to 0.25.

The DeblurGAN was developed to unblur images using the Wasserstein loss (Kupyn, Budzan, et al., 2018). Since we are also interested in deblurring the output images, we have tested this loss with the proposed AAE where $(P(y, y'))$ is the joint distributions of y and y' for which the distributions are equal to Py and Py', and p(y,y') the proportion of y or y' to move to have $Py = py'$:

$$Wass(P(y, y')) = \sum_{i=1}^{n} \inf_{p(P(y,y'))} E_{yi,yi'} \delta p(||yi - yi'||) \tag{3.28}$$

However, we do not apply a l2 content loss such as in Kupyn, Budzan, et al., 2018 added to the Wasserstein loss, or add a penalty gradient to the Wasserstein loss such as (Gulrajani et al., 2017). Since the dataset contains various subclasses (physiological and pathological brain organoid images acquired at three developmental stages), we aim at creating a Normalized Wasserstein loss to avoid imbalanced mixture proportions (Balaji et al., 2019). We apply the L2 normalisation on the Wasserstein loss, producing a new loss we call the Perceptual Wasserstein loss for the first time to our knowledge, applied on an AAE architecture:

$$P.Wass(P(y, y')) = \sum_{i=1}^{n} \sqrt{\inf_{p(P(yi,yi'))} E_{yi,yi'} \delta p(||yi - yi'||)^2} \tag{3.29}$$

Training

Figure 3.17 shows the global training setup. The 40 original images are used to generate 40 synthetic images for each architecture (each loss). Input and output images measure 250×250 pixels. Training lasts 2000 epochs for each optimisation; this corresponds to the plateau before over-fitting for all loss optimisations, see Figure 3.18.

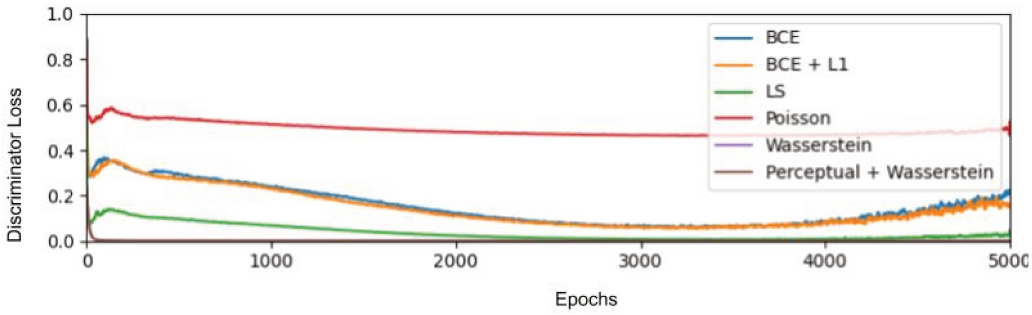
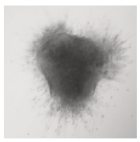
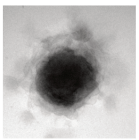
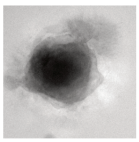
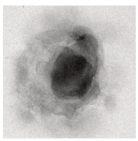
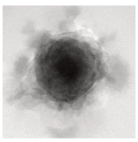
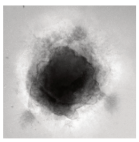
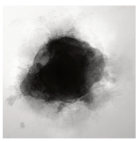
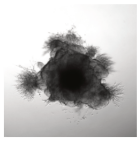
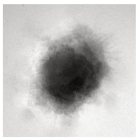
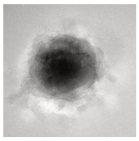
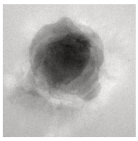
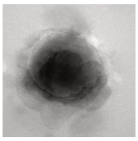
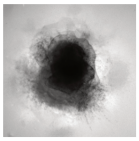
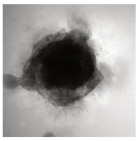
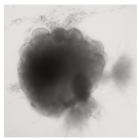
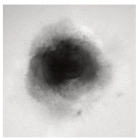
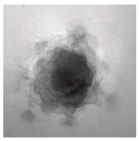
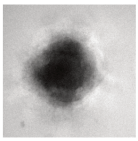
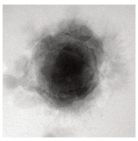
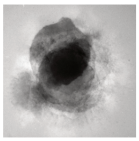
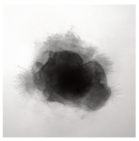


Figure 3.18: Training curves comparisons of discriminator loss optimisations.

Table 3.3: Samples of original and synthetic images generated by the AAE. We show 3 of the 40 images for each group: original and per AAE loss variation.

Original	BCE	BCE + L1	LS	Poisson	Wasserstein	P. Wass.
						
						
						

3.3.2 Results

We first present samples from each loss optimisation, the metric evaluation of the synthetic images, and then dimensional reductions.

Qualitative evaluation

Sample images produced by each AAE variation are shown in Table 3.3 more particularly three of the 40 images generated for each of the six AAE variations. While some of the generated samples are blurry and present a white imprint (BCE, BCE + L1, LS). Others show sharper edges and less visible imprints (Poisson, Wasserstein and P. Wass.). For this group of three losses, only a few of the generated images seem to be identical to the input images: these networks do not suffer from mode collapse.

Metric evaluation

To quantitatively confirm the visual analysis of the generated images, we calculate several metrics on both the original and synthetic images. Results are shown Table 3.4. The AAE loss optimisations allow generated images to be within the range on five metrics with the Wass. and P. Wass. loss (against four for the other loss). Indeed, the Blur index is with these two optimisations within the range of original images. Of the 7 metrics calculated,

Table 3.4: Metric evaluation of AAE brain organoid bright-field generated images. We have calculated metrics on generated images from each AAE loss variations, with the BCE loss as the baseline. Scores outside the original range are in gray, best values are displayed in bold.

Metric	Best	Original	BCE	BCE + L1	LS	Poisson	Wass.	P. Wass.
FID	low	0.47 – 0.80	1.20	1.41	1.33	1.41	1.10	**0.82**
SSIM	high	0.65 – 0.71	**0.63**	0.62	0.60	**0.63**	0.62	0.50
UQM	high	0.63 – 0.87	0.83	0.83	**0.84**	**0.84**	0.83	0.82
MI	high	0.21 – 0.47	0.37	0.39	0.36	0.41	**0.46**	0.42
Blur	low	0.10 – 86.28	135.93	116.30	135.01	106.71	59.84	**59.00**
PSNR	low	11.9 – 16.6	13.47	13.74	13.53	13.74	13.17	**12.86**
MSE	low	93.25 – 106.23	103.13	103.35	104.01	103.33	103.11	**102.93**

only FID and SSIM are not in the range for all the optimisations. However, the FID for the P. Wass. loss is quite close to the upper bound of the input range (0.82 vs. 0.80).

Quantitatively, Wass. and P. Wass. networks generate better image quality than other networks, based on lower PSNR, MSE score and Blur index.

Otherwise, the similarity seems to depend on the architecture. Based upon the FID, the P. Wass. loss generates the most similar images. MI is higher for the images generated using Wass. networks. However, compared to images produced with BCE and Poisson, images produced with P. Wass. loss are not similar to the original dataset regarding the SSIM. Images generated with a Poisson or LS loss have the best UQM index.

The P. Wass. loss is the most appropriate loss optimisation for generating cerebral organoid images with the AAE. It performs best for four metrics and is within the original range for five metrics.

Dimensional reduction

To analyze all at once the similitude and the variability of the generated images with AAE loss optimisation, we study images in the same statistical space. We implement a dimensional reduction on the features extracted on images in the generative process with t-SNE, see Figure 3.19. We first observe that, independantly from the loss optimization, similar positions in the map are kept between original and generated images. Some original images constitute a cluster and are distant to the generated ones. This could be explained by the incapacity of generated images to replicate a background similar to the bright-field acquisition with a light gradient. While at the exterior of the map images generated with a Poisson a Wasserstein or a Perceptual Wasserstein loss are represented, inside the map BCE, BCE+L1, and LS losses are. This observation suggests that each loss optimisation could bring different information during the generative process. We compare the KL divergences between original and generated images which remains similar (all results are approximating the null: inferior at 0.3). To summarize, loss optimisations generate similar contents to original images keeping its variability and creating intermediate shapes not seen in the original population.

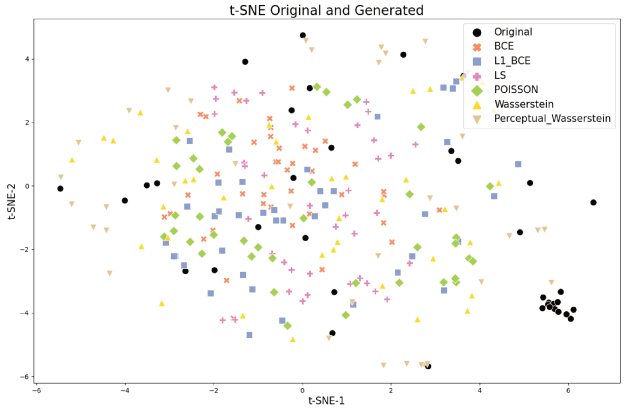


Figure 3.19: t-SNE representation of original and generated images with optimized AAE

3.3.3 Discussion

We present data augmentation of cerebral organoid bright-field images, using various loss optimisation of an Adversarial Autoencoder (AAE). The Perceptual Wasserstein discriminator loss optimisation outperforms the state-of-the-art original loss article to generate these images, according to the metrics used. Nevertheless, our dimensional reduction experiments suggest that any loss optimisations can bring some variability to the generative process while remaining similar to the original dataset. Each loss present a particular interest for specific tasks (segmentation, classification, detection etc.)

As in the previous Section 3.2, synthetic images generated with AAE are still coherent with original dataset in terms of image quality, with almost no collapse mode but also adding a sought diversity similar to the acquisition of the original dataset. As previously discussed, generative architecture can potentially be optimized by implementing a U-Net generator instead of a simple Auto-encoder (Isola et al., 2017).

The Perceptual Wasserstein loss optimisation of AAE performs best according to metrics see Table 3.4. Other loss optimisations also show high similitude, though with a lower quality. Poisson and Wasserstein loss bring the best results after the Perceptual Wasserstein. It could be interesting to observe if a regularisation term such as in the Perceptual Wasserstein for the Poisson loss can outperforms such result. Indeed, the Poisson loss increases when the prediction is a generated example (Wargnier-Dauchelle et al., 2019). Another potential improvement is to update the perceptual part with a L1 normalisation, as it might avoid blur Isola et al., 2017 and therefore avoid the white imprints around the shape. We also aim at trying others embedded losses (already used for segmentation tasks) during the generative process based upon high level prior like object shape, size topology or inter-regions constraints (El Jurdi et al., 2021; Kingma et al., 2013). These losses could be used on the condition that the morphological development of CO is better characterized.

However, the dimensional reduction experiment suggests that some loss functions can be used to generate more images and a good diversity enriching the original training set. Such observation raises the following question: which role can each loss optimisation results have for biomedical applications? The GAN optimisations dedicated to a particular biomedical application have already been discussed in Zhengwei Wang et al., 2021. In this context, we plan to explore what type of information each loss brings during the image generation: such as comparing the effect of each loss variation in each segmentation task. However the Perceptual Wasserstein and Wasserstein losses seem to be the nearest to extreme spaces of

the original range and provide the best diversity while BCE losses, LS and Poisson are more in the heart of original feature representations. This could be explained by the Wasserstein loss aim, which is to transform a distribution into another at a minimal cost (Arjovsky et al., 2017). Thus the point feature distributions of Wasserstein and Perceptual Wasserstein images are near all the feature original limit distributions.

Noteworthy is the the similarity of original images in the t-SNE right corner. Generated images were not able to generate a similar background such as the lightning gradient of some bright-field acquisition in white light (Boyat et al., 2015). To resolve this issue, we aim at studying the effect of a similar bright-field background injection during the generative process.

There is still room for improvement in the proposed optimized AAE network strategy evaluation. A first possibility is to quantitatively evaluate the generalisation of the results obtained on cerebral organoid bright-field images: we would like to use this methodology on others bright-field biomedical images ((Nickels et al., 2020) and Table 2.8). The best way to evaluate synthetic images and, compare various results of GAN optimisations is still a challenge (Alqahtani et al., 2019). Indeed there is no commonly approved specific metric to evaluate whether GAN-generated synthetic images can be considered as natural (Borji, 2018) and statistical space comparisons with a dimensional reduction strategy do not allow us to elect a particular loss optimisation as the best in this work. Biological experts aim at psychovisually evaluating the generated images, and strengthen the quantitative evaluation proposed. In particular, we project to validate the suitability of the metrics we use and observe the training effect on the segmentation task with only validated images by biological experts and maybe highlight a metric combination which could replace the time consuming psychovisual evaluation.

This study answer to the first emerging issue in the cerebral organoid field, highlighted in Brémond Martin et al., 2021b and updated in Section 2. These results show that small databases augmentation of cerebral organoids bright-field images is possible using GANs. Particularly the AAE Perceptual Wasserstein loss optimisation generates the most qualitative images, while being similar to the original dataset. However it remains to discover what kind of information each loss optimisations can bring during the generative process, with coherent diversity in the initial dataset. For these applications, we need to psychovisually validate the synthetic images. We also aim to generate particular backgrounds by injecting noise during the generation, similar to bright-field acquisition noise.

3.4 AAE optimisation by noise injection

This section is published in Brémond Martin et al., 2022 for the IPTA conference in 2022.

In the previous Section 3.3, we highlight that Perceptual Wasserstein loss contributes to generate the most qualitative and similar results as the original images, with a certain amount of variability. However, the background of the generated images remains different from original bright-field images background, usually characterized by a shot noise (Boyat et al., 2015; Gilroy, 2019).

In recent years, noise addition has been largely investigated for improving data augmentation strategy. In the literature, Salt and pepper or Gaussian noise are usually used to increase datasets (S.-H. Wang et al., 2021; Hussain et al., 2017; Chlap et al., 2021). For feature space augmentation, others authors use Gaussian noise to increase the dataset inside

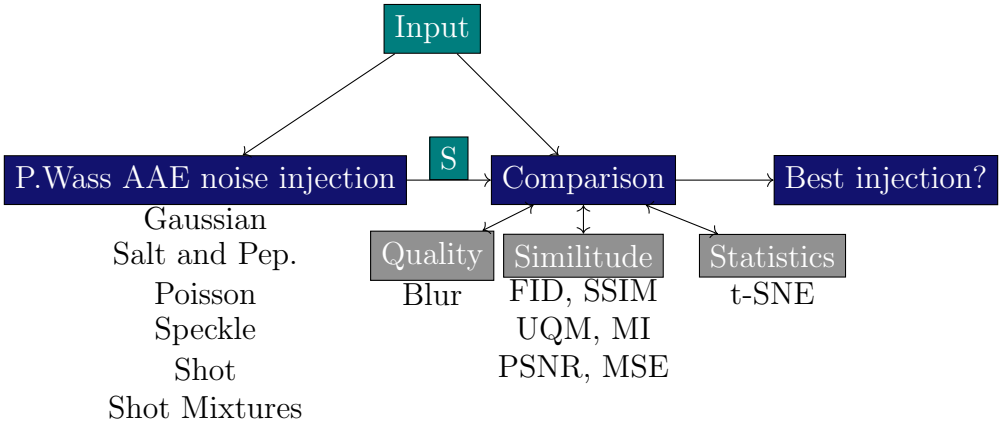


Figure 3.20: Procedure of noise injection comparison in AAE with P.Wass loss optimisation for the generation of bright-field brain organoid images more similar to microscopic natural acquisition. S corresponds to synthesized images.

the learned space (Nacereddine et al., 2012). In the GAN field, some authors already use Gaussian noise but do not compare the effect of different noise types in the generation to synthesize more natural images, and not try adding a microscopic natural noise such as the shot one (Salimans et al., 2016; Sonderby et al., 2016; Feng et al., 2021). To our knowledge, no study have already used GAN to produce at the same time a realistic microscopic image background, and a similar content of the object acquired in microscopy.

The contribution of this part lies in the quantitative and feature statistical space investigation of the influence of a noise injection during the generative process using AAE. The goal is to find the noise type generating images with both satisfying quality and a good variability, so that they can be incorporated in the learning dataset and improve automated characterisation tasks. We validate data variability by a dimensional reduction of the feature space related to generated images and compared to real ones. We summarize the procedure in Figure 3.20.

3.4.1 Methods

Resources, metric evaluation and dimensional reduction are the same as in Section 3.2. As in the previous section we use an AAE with a Perceptual Wassertstein loss optimisation as it render the best qualitative and similitude images to the original bright-field brain organoid images (Brémond Martin et al., 2021a) and see Section 3.2. In the original implementation, Gaussian noise is added to the generative part. However in our case, generated images have a different background from natural microscopic images and occasionally presents a white imprint (Brémond Martin et al., 2021a). Usually, bright-field images background is characterized with a shot noise (Boyat et al., 2015; Gilroy, 2019). We aim at quantifying and visualizing the effect of various specific noises or mixture noises injected on the decoding part of the generative process. The summary of experimental scheme is described in Figure 3.21.

Noise injection

The overall goal is to generate synthetic images as close as possible to real bright-field ones, including realistic background. In the literature, shot noise is the signal emitted in this acquisition modality (Boyat et al., 2015; Gilroy, 2019). We therefore inject shot noise *Shot*. in the generative step instead of the usual Gaussian noise *Gauss*. (a variation applied to

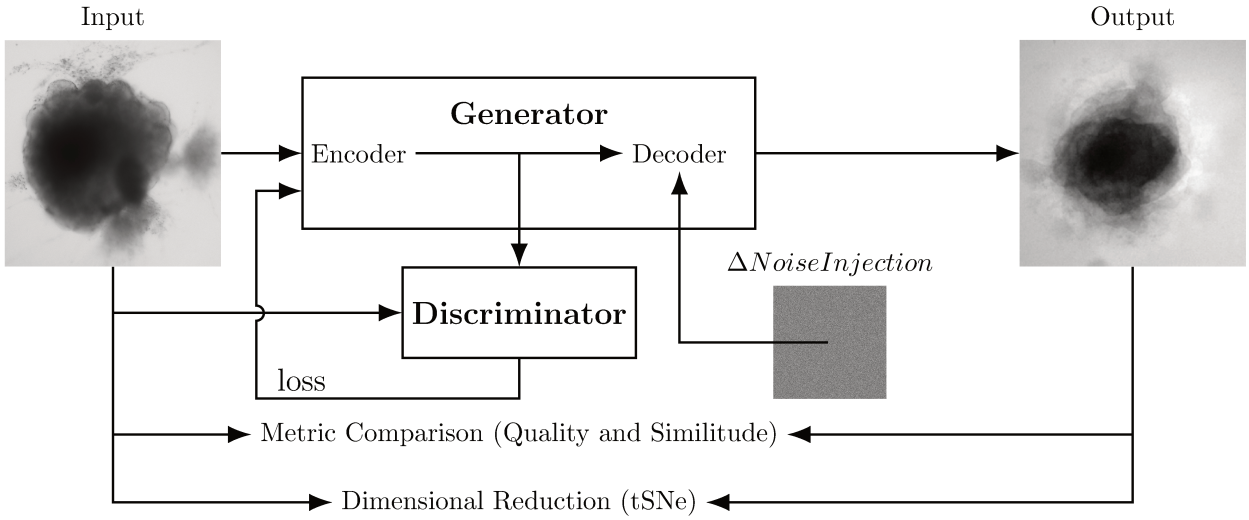


Figure 3.21: AAE supporting data augmentation of brain organoids bright-field images with noise injection procedure.

each pixel value between 0.1 and 0.05).

We calculate the pixel noise estimation from (Immerkær, 1996) in original images to obtain an estimated value of shot noise σ to apply on generated images. For each original image (I of dimensions X=Y=250) we first convolve an approximation of the Laplacian matrix (L of 3X3 pixels) of I at N positions (x,y) to obtain the c term.

$$c = \sum_{i,j=0}^{N} \sum_{x,y=0}^{X,Y} |(I, L)| \tag{3.30}$$

This term is used to found σ of an original image I calculated such as:

$$\sigma = c * \sqrt{\frac{\pi}{2}} \frac{1}{(X-2)(Y-2)} \tag{3.31}$$

Then we calculate the median of the σ for all the original images to obtain the value of noise to reach (represented by the blue line at 50 in Figure 3.22).

To find a satisfying shot noise (pixel variance=50), we produced 500 images of shot noise with a variation of the peak parameter (ranging from 0 to 500). Then we calculate each image's noise variation and selected the shot noise image equal to σ. To visualize the selection, we plot the noise variation in function of the shot noise peak value used to produce the shot noise image with the line representing the baseline σ (see Figure 3.22): the intersection of the two curves corresponds to the selected value to produce a shot noise image with an estimated noise variation similar to the original dataset. We found the best shot noise peak to be 200.

To verify if and how another type of noise can modulate the generated image, we implement both Salt and Pepper noise *SaltPepp.* (25% of the pixels, randomly selected, are set to 255 or 0) and Speckle noise *Spe.*. Moreover, to verify the effect of a mixed noise on the generative process, we inject *Shot* with one of the three other noises.

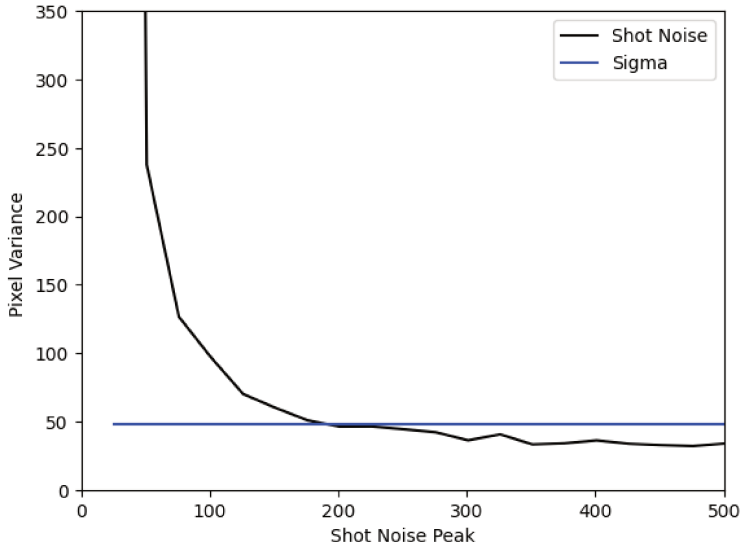


Figure 3.22: Peak variation used to Shot Noise regarding the pixel noise variance. The median of original image pixel estimated noise variance σ (baseline) is represented by the blue line. The Shot noise image is represented by the black curve.

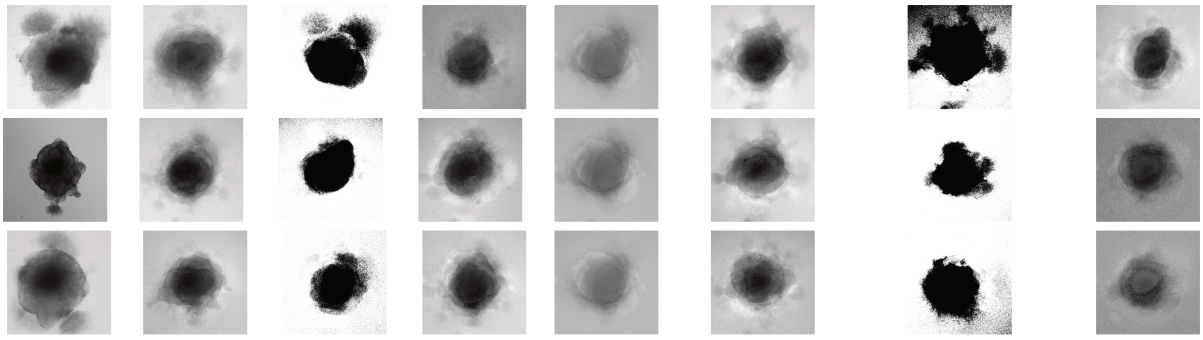
Original	Gaussian	Salt&Pepp.	Speckle	Shot	Shot+Gaussian	Shot+Salt&Pepp.	Shot+Speckle
							

Table 3.5: Sample of generated images from each AAE with noise injection optimisation. Gaussian loss is the baseline. We add two other samples for each case compared to the published article.

Table 3.6: Metrics calculated on generated images from each AAE noise injection, with Gaussian loss as the baseline.

	FID	SSIM	UQM	MI	BLUR	PSNR	MSE
Original	0.47 to 0.80	0.65 to 0.71	0.63 to 0.87	0.21 to 0.47	0.10 to 86.28	11.9 to 16.6	93.25 to 106.23
Gaussian	1.19	0.60	**0.84**	**0.83**	169.70	14.2	**2 998**
Salt & Pepper	1.18	0.38	**0.84**	0.64	822.58	14.2	281 639.8
Speckel	1.28	0.51	0.76	0.74	791.24	12.3	29 410.33
Shot	1.40	**0.67**	0.82	0.70	**76.21**	13.8	430 051.75
Shot + Gaussian	**0.99**	0.56	**0.84**	0.81	231.98	14.0	3216.23
Shot + Salt & Pepper	1.11	0.46	0.73	0.71	1273.29	**11.5**	1 148 012.25
Shot + Speckle	1.05	0.59	**0.84**	0.82	172.88	14.0	319 333.2

3.4.2 Results

Influence of noise injection

To reproduce the effect of bright-field acquisition, we inject various noises (Gaussian, Salt and Pepper, Speckle and Shot, and mixed Shot noise) to the decoding part of the generative process. Sampling of the generated images is represented in Figure 3.5. Original images of brain organoids are characterized with a core and peripheral zones with neuroepitheliums (projections around the core) on a diffuse background with a shot noise and gradients of light. With Gaussian noise, a slight white imprint is generated around the organoid shape. The Salt and Pepper injection results in more contrasted brain organoid contours than with Gaussian noise. However, the generated images are black and white, which is different from grey-level original bright-field images. Speckle noise injection adds a dense zone in the middle of the core and a white imprint around its shape, whereas the shot noise injection renders the entire image pale. Mixed shot noise renders all the images less contrasted than a single noise does. The generated image background (whatever the noise injected) seems modified respectively to the original AAE loss optimisation from (Brémond Martin et al., 2021a). Neuroepithelial zones are more observable in images generated with a salt and pepper injection, but remains visible with other noise injections.

To strengthen the previous observations, we choose to calculate a set of metrics (Table 3.6), indicating the quality of generated images and their similitude with original images. Even though shot noise is related with bright-field acquisitions in the literature, only the $BLUR$ shows the lowest score and the highest $SSIM$. According to others metrics (UQM, MI, MSE), Gaussian noise seems the best suited to generate COs bright field images. However, only the shot noise produces images with four metric scores in the range of original scores ($SSIM$, UQM, $BLUR$ and $PSNR$). Mixing the shot noise with another noise does not significantly improve nor the quality nor the similitude with the original dataset according to the computed metrics (except for the UQM). Of note, the MSE score is higher for generated images than for original ones, independently from the used noise.

Dimensional Reduction

To verify that the image generation with noise injection respects the original image variability, we study image representation in the same optimized statistical space. We implement a dimensional reduction on the feature extracted on images during the generative process with t-SNE, see Figure 3.23.

To observe the t-SNE representation we remove three outlying points corresponding to three images generated with a speckle noise. Keeping these outliers completely skews the output representation. The remaining 316 image points are represented the figures plotted with t-SNE. One of the 10 representations is shown in Figure 3.23. All the image points (original and generated) are in the same representation space and not clustered per noise. Only three images (one original, and one generated with a Salt and Pepper injection or a Gaussian injection) are in the center of the circle represented in the plot and some generated are not on the circle.

However when we compare KL divergence between original and generated images with these optimisations, no significant difference exists (less than 0.3). To summarize Figure 3.23, noise injection in the optimized AAE allows to generate both similar content and background when compared to original ones. It also add a certain amount of statistical variability: the generated images using purposely noise injection fill the gaps between original images as illustrated in Figure 3.23.

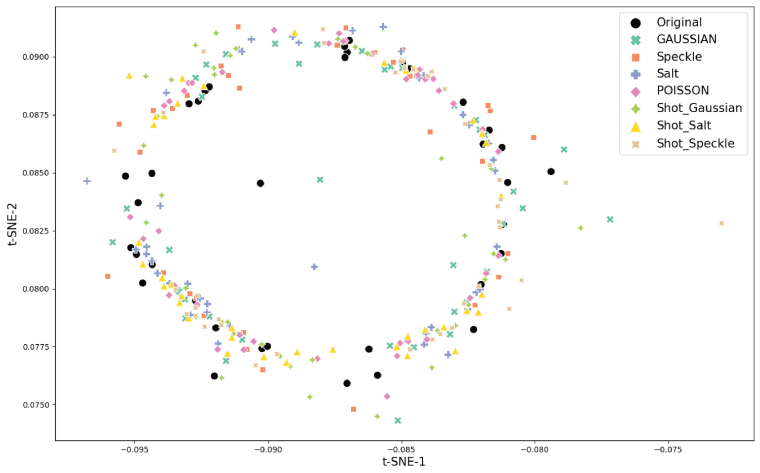


Figure 3.23: t-SNE representation of original and generated images with various noise injected in AAE.

3.4.3 Discussion

For the first time to our knowledge, we propose in this work an enriched generation of brain organoid image using a purposely injection of noise in an AAE architecture. We show that shot noise injection provides an overall better quality and realistic images than our gaussian baseline while the Gaussian noise injection provides the best scores. The dimensional reduction experiment shows that noise modulations fill the gaps between each original image representation.

Our experiment demonstrates noise injection modulates AAE-generated images. In the original implementation, a Gaussian noise is used (Makhzani et al., 2015; Goodfellow et al., 2014). However, this noise is not suited to generate bright-field organoid images, as suggested in Brémond Martin et al., 2021a. The current study also shows that other type of noise can be used to change the content and background of AAE generated images, to make them more realistic.

The main idea of our work is to generate images with similar background than an actual microscopic acquisition. A shot noise injection seems to produce the better range of quality and similitude metrics compare to original images. This noise injection outperforms the others noise injections including the classic one used in all others submission (Gaussian noise) (Goodfellow et al., 2014). Shot noise characterize the bright-field acquisition (Gilroy, 2019; Boyat et al., 2015) this could explain why in this application it render the best results. To our knowledge, the noise in GAN is only Gaussian, and no implementation has concerned various noise injections comparisons (Salimans et al., 2016; Sonderby et al., 2016; Feng et al., 2021). It could be interesting to assess if such noise injection is appropriate to augment brain organoids image databases taken with other microscopy modalities (i.e. confocal, light-sheet) or for other bright-field biomedical images (see part 2.8).

Only four score metrics are in the range of the original images scores with the shot noise, and some metrics calculated are never in the range of the original metrics and thus whatever the noise injection used such as MI or the FID while others are always in the range (UQM, $PSNR$). If UQM and $PSNR$ characterize the best the similitude and the quality estimation on natural and generated images, they cannot be used as a discriminating metric with the used optimisation strategy. Regarding the mutual information, this similitude index is based upon the relationship calculations between pixels and does not take into account pixel's respective neighborhood. Given that the generative process creates regions

similar to original ones, including neuroepithelial formation, it could be interesting to test another metric more suited for regions such as the regional mutual information used in registration (Russakoff et al., 2004). The FID remains too high in our generative process, however its calculation, based on covariance and mean, (Borji, 2018) is the precursor of $SSIM$ calculation.

A first interpretation lies in the assumption that the defined metrics are not well suited to the current problem. To validate the use of metrics, Human psycho-visual evaluation of generated images could be used as a complementary approach.

The second interpretation is that the generative process needs to be improved to generate images with more similar scores to the original range. To tackle this issue, the proposed architecture can be optimized. For example, it could be split in two, as it has been done in others studies, to provide a better global precision: one architecture part will be implicate for local precision (a discriminator which take in input only patches); the second one will be implicate for global precision (a discriminator which take in input global images) (Kupyn, Martyniuk, et al., 2019). It could be interesting to use noise injection with the split architecture in order to observe if similitude increases.

The dimensional reduction strategy shows that generated images are in the same statistical space as original images, independently from the type of injected noise during the generative process (except for generated images from a speckle AAE optimisation). The generated images obtained with other noise injection better cover the statistical space than the loss optimisations (Brémond Martin et al., 2021a). This optimisation reaches the generation of gradient light background that could not been achieved with just loss optimisation, and thus avoid an original image cluster formations. These images could be used in tasks requiring a larger dataset due to their variability covering original images space, after validation by biological experts. Speckle noise is not implicated in bright-field acquisitions, explaining its part remoteness in the t-SNE graph (Boyat et al., 2015; Gilroy, 2019). However, its implication in other microscopic acquisition could be interesting, such as in fluorescence or light-sheet microscopy (2D or 3D).

In our case, we study noise injection in the AAE architecture used in Brémond Martin et al., 2021a. Future works will test noise injection in the generative process of others GAN architectures, to assess the increase in quality and similitude of the generated images. Noise optimisations will also be tested to increase the original dataset of brain organoid images in others microscopic acquisition modalities. Another perspective is the potential use of other noise mixture (Gaussian, Speckle or Salt based noise). All the results must also be validated on larger and more diverse datasets, as the original data can have an effect on the generation.

We use various noise injections during the generative process of an AAE in order to increase the size of a brain organoid bright-field image dataset. We observed the modulation of generated images, and quality and similitude metrics in the range of original image ones calculated with a shot noise injection, observed in real bright-field images. Dimensional reduction points out the generated images are in the same statistical space as the original images. After validation, generated images can be incorporated in the dataset used to train a model using deep learning strategies, and then used for segmentation to characterize brain organoid morphological development.

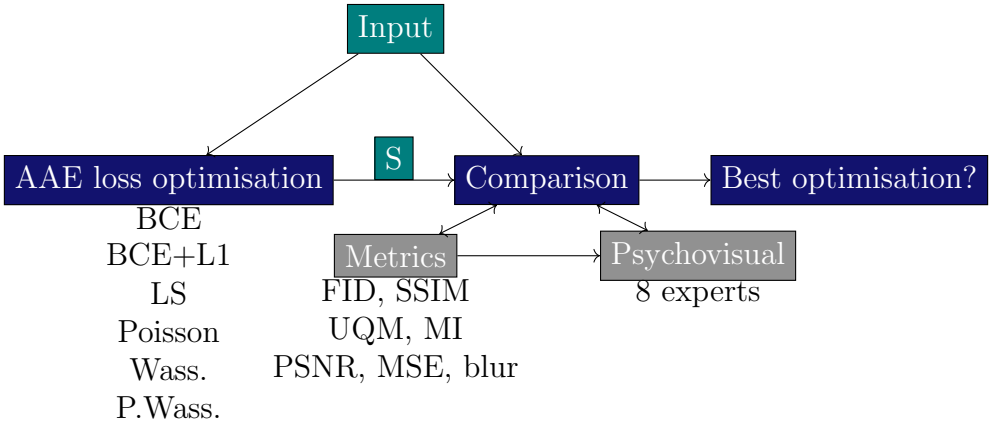


Figure 3.24: Procedure of loss optimisation psychovisual evaluation in AAE for the generation of bright-field brain organoid images more similar to microscopic natural acquisition. S corresponds to synthesized images.

3.5 Evaluation of loss-optimized synthetic brain organoid images

This section takes part of a work submitted to Computers in Biology and Medecine, and entitled "Brain Organoid Data Synthesis and Evaluation".

The validation of synthetic images remains a challenge (Alqahtani et al., 2019). Psychovisual evaluation is a time consuming gold standard which requires many subjects to reduce its' intrinsic subjectivity. On the other hand, there is no commonly approved specific metric to evaluate whether GAN-generated synthetic images can be considered as natural. The use of common metrics is also controversial (Borji, 2018). To improve the sharpness during the generation of brain organoid bright-field images, we test various loss functions to improve the adversarial network (Brémond Martin et al., 2021a). However, these results are based upon metric calculation and a dimensional reduction to compare all feature images (original and generated with each optimisation) in the same statistical space. Indeed some metrics may not been suited to identify the naturality of an image as they are originally created to test the similitude or the quality of images (Brémond Martin et al., 2021a; Borji, 2018). If these images still deserve to train a DL segmentation algorithm, another fundamental issue remains unresolved: do these synthetic images seem natural to a biological expert point of views as for metric(s)? Psychovisual evaluations have been already made on others bright-field cell synthetic cell generation (Malm et al., 2015). This evaluation is an important step for the validation of a particular generative model of images. Thus the selected images as natural by Human Biological experts could maybe help to train deep based segmentation methods and characterize their development but with now a double psychovisual-metric validation.

We propose to evaluate the synthetic images generated an AAE (Brémond Martin et al., 2021a) using both with similarity metrics and biological experts. The other part of this work attempt to dechipher a metric or a combination of metrics which could replace the psychovisual validation. The procedure is summarized in Figure 3.24.

3.5.1 Methods

Resources, AAE architecture and metric evaluations are the same than in article 3.2. Loss optimisations and trainings are the same detailed in article 3.3. For the psychovisual evaluation methodologies and the comparisons between metrics and psychovisual evaluations see the following method parts.

Psychovisual evaluation

Dataset of original and synthetic images The dataset to evaluate contains two classes of images the 40 microscopic acquisitions mentioned in section 2.7 and the 240 synthetic images created by the previous mentioned AAE loss optimisations: binary cross entropy (BCE), binary cross entropy with a L1 normalisation (BCE + L1), least squares (LS), Poisson, Wasserstein (Wass.), perceptual Wassertein (P. Wass). The dimension of these 280 images is 250×250 pixels.

Randomisation To perform a double blind test the images are not labelled during the visualisation so neither the experts or the team can be biased by the images information (real, generated, nor its kind of generation). Real and generated images are randomized at each test run. Each biological expert evaluates the complete randomized dataset (280 images). The randomisation and corresponding labels are stored in a .csv file which is only accessible for result analysis.

Experts The experts who evaluated the database are biologists from ERRMECe laboratory, EA1391, CY Cergy Paris University. The group of 8 experts is composed of 3 men and 5 women who are either PhD students, research engineers or researchers. They all have an expertise in neuronal culture and microscopy acquisition. We do not allow duplicate evaluators across evaluation procedure. During each evaluation session the evaluator is physically isolated from the other participants, without knowledge of other experts responses or images labels.

Evaluation software To help experts in their evaluation and to ensure consistency throughout the entire experiment, we built a dedicated software using Python 3.6, as shown Figure 3.25. The interface consists of a displayed image (250×250 pixels), a cursor with 3 buttons: *Real*, *Generated*, *Next* and the number of remaining images to classify. Keyboards shortcuts are available for *Real*, *Generated* and *Next* buttons (respectively A, P and Tab keys on an AZERTY keyboard) to facilitate the process. Images on the screen are updated each time *Next* is hit (or the corresponding keyboard shortcut. Clicks on *Next* are counted as a pass if not preceded by a *Real* or *Generated* click.

Experimental protocol An operator enters the expert name and the date and hour of the recording. All eight experts chose to use their own mouse with the experiment laptop. The protocol, consisting of a single session and including all 280 images (real and synthetic), is described in Figure 3.25. A pass is consider as an answer. The decision and answer time are saved at each click in a .csv file only accessible to analyze the results. The operator is present nearby to verify the smooth functioning of the experimental process and to capture any comments made by the experts. The list of questions the expert has to answer after the process is listed below:

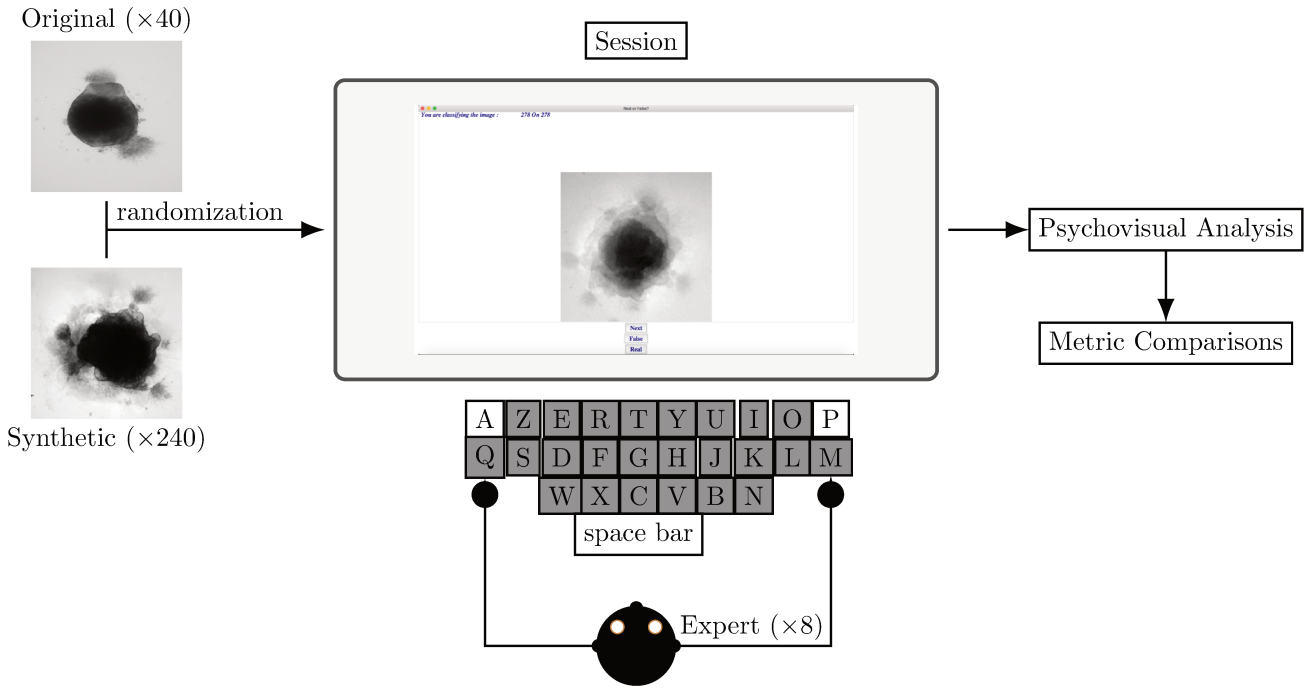


Figure 3.25: Experimental procedure of psychovisual evaluation. a) represents a sample from original dataset and b) a specific loss optimised sample generated with an AAE. Screenshot of the software helping biological experts to decipher natural to non natural content of brain organoid culture images. Eight biological experts test this procedure.

- Why do you classify this generated image as a false? The operator show a generated image with the longest hesitation.
- Why do you classify this original image as a false? The operator show an original image with the longest hesitation if this situation exists.
- What should we improve in future sessions?

We summarize the answers of these questions in the result part.

Each evaluation run produces two .csv files: One stores the randomisation i.e. the order and label of each image presented. The second stores the experts' name and for each image present the answer time and the decision

Analysis

The first analysis step consists in associating the randomisation and the results files. We obtain, for each expert and for each image the decision and the decision time. The decision is then labeled as true positive (TP) or false negative (FN) for the original images and false positive (FP) or true negative (TN) for the synthetic images.

Parameter Calculations It is then standard to calculate the error rate (ER), defined as the number of false decisions divided by the total number of decisions.

$$\mathrm{ER} = \frac{\mathrm{FP} + \mathrm{FN}}{\mathrm{FP} + \mathrm{FN} + \mathrm{TP} + \mathrm{TN}} \tag{3.32}$$

For the original images, this becomes:

$$\mathrm{ER_O} = \frac{\mathrm{FN}}{\mathrm{FN} + \mathrm{TP}} \tag{3.33}$$

and for the synthetic images:

$$\mathrm{ER_G} = \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TN}}. \tag{3.34}$$

However, we wish to compare the proportion of synthetic images falsely labeled as true, with the proportion of original images label as true. We thus calculate the Positive Rate (PR) of the original images:

$$\mathrm{PR_O} = 1 - \mathrm{ER_O} = \frac{\mathrm{TP}}{\mathrm{FN} + \mathrm{TP}}. \tag{3.35}$$

For the synthetic images $\mathrm{PR_G} = \mathrm{ER_G}$.

As a second parameter, we calculate the decision occurrence for each modality for each subgroup by a simple counting and render it in a % according to the total effective of a group of images.

We also evaluate the number of positive answers given by each expert as a count and the number of images given as a positive by zero expert, one expert,2 experts etc. or the 8 experts. Time decision and all these parameters are calculated between original and generated images, or between original and each modality of loss generation (BCE, $BCE{+}L1$, LS, $Poisson$, $Wass.$, P. Wass.), globally or by each decision subgroups (FP, FN, TP, TN). All results are rendered as bargraphs representing variables (Time Decision in seconds or occurrence in % or error rates) according to one or many factors (group and subgroups of decision).

Metric versus human decision To verify if some metrics highlight the same loss as producing the most natural images as experts, we plot each metric values for each loss group by each decision factor modality (FP, FN, TP, TN). In the dot representations for each loss group, each metric is plotted according to the Normalized Error Rate NER with individuals decision time t for the FP and for FN modality:

$$\mathrm{NER} = \frac{\mathrm{FP} \times t_{\mathrm{FP}} + \mathrm{FN} \times t_{\mathrm{FN}}}{\mathrm{FP} + \mathrm{FN} + \mathrm{TP} + \mathrm{TN}} \tag{3.36}$$

where t_{FP} and t_{FN} are the average decision time respectively for FP and FN. In practice, this becomes, for the original images:

$$\mathrm{NER_O} = \frac{\mathrm{FN} \times t_{\mathrm{FN}}}{\mathrm{FN} + \mathrm{TP}} \tag{3.37}$$

and for the synthetic images:

$$\mathrm{NER_G} = \frac{\mathrm{FP} \times t_{\mathrm{FP}}}{\mathrm{FP} + \mathrm{TN}}. \tag{3.38}$$

To verify if a relation exists between a metric or a particular combination of metrics (M) and the decision or the time decision (DT), we calculate KL divergences on dimensional reduction results (Joyce, 2011).

$$KL(M||DT) = \sum_{x \in X} M(x) \times \log\left(\frac{M(x)}{DT(x)}\right) \tag{3.39}$$

We only represent here KL divergence correlation plots for each individual metric (and not metric combinations) for space considerations. We consider all possible metric combinations C :

$$C = nk | k \in \mathbb{N}^*, k \leq n \tag{3.40}$$

Table 3.7: Metric combinations explored as alternative to psychovisual evaluations.

Factorial	1	2	3	4	5	6	Total
	Blur SSIM PSNR MSE MI UQI	Blur,SSIM Blur,PSNR Blur,MSE Blur,MI Blur,UQI SSIM,PSNR SSIM,MSE SSIM,MI SSIM,UQI PSNR,MSE PSNR,MI PSNR,UQI MSE,MI MSE,UQI MI,UQI	Blur,SSIM,PSNR Blur,SSIM,MSE Blur,SSIM,MI Blur,SSIM,UQI Blur,PSNR,MSE Blur,PSNR,MI Blur,PSNR,UQI Blur,MSE,MI Blur,MSE,UQI Blur,MI,UQI SSIM,PSNR,MSE SSIM,PSNR,MI SSIM,PSNR,UQI SSIM,MSE,MI SSIM,MSE,UQI SSIM,MI,UQI PSNR,MSE,MI PSNR,MSE,UQI PSNR,MI,F MSE,MI,F	Blur,SSIM,PSNR,MSE Blur,SSIM,PSNR,MI Blur,SSIM,PSNR,UQI Blur,SSIM,MSE,MI Blur,SSIM,MSE,UQI Blur,SSIM,MI,UQI Blur,PSNR,MSE,MI Blur,PSNR,MSE,UQI Blur,PSNR,MI,UQI Blur,MSE,MI,F SSIM,PSNR,MSE,E SSIM,PSNR,MSE,F SSIM,PSNR,MI,F SSIM,MSE,MI,UQI PSNR,MSE,MI,UQI	Blur,SSIM,PSNR,MSE,MI Blur,SSIM,PSNR,MSE,UQI Blur,SSIM,PSNR,MI,UQI Blur,SSIM,MSE,MI,UQI Blur,PSNR,MSE,MI,UQI SSIM,PSNR,MSE,MI,UQI	Blur,SSIM,PSNR,MSE,MI,UQI	
Effective	6	15	20	15	6	1	63

where $n = 6$ is the total number of metrics. The total number of metric combinations considered is thus 63. In total we consider 63 metric combinations calculated using a binomial low with the subsequent binomial coefficient B for k among M metrics, with k belonging to (0:L): A summary of each metric combination is given in the following Table 3.7.

We then calculate Pearson and Kendall correlations between metric combinations and KL divergence results (for error rate and time decison) for original and synthetic groups. We show the Pearson correlation for the ten best metric combinations. The best representation of these results (error rate or time by KL divergence) are represented as a scatter plot.

Statistical analysis The normality is verified by a Shapiro and quantile to quantile graphics. We verify the homoscedasticity in normal cases by a Bartlett test and in the case of non-normality by a Levene test. In case of a normality and homoscedasticity, parametric tests are used (Anova), and non-parametric tests otherwise (Kruskall-Wallis with a Tuckey post-hoc test). Regression models are implemented to verify the interaction of factors (group and decision) on a specific variable (time or error rate or occurrence for instance). We use a post-hoc Holm test to compare two by two the effect of factors on variables after the regression. We take an alpha risk at 5%. Correlation matrices are based on Pearson correlation tests.

3.5.2 Results

Psychovisual evaluation of synthetic images

We compare the mean positive rate between the original and synthetic images in Figure 3.26 left and observe that it is significantly higher for the original images. In Figure 3.26 right, we compare the occurrence of each decision in percentage for *original* and *generated* groups. A misleading corresponds to a false positive answer. There is less misleading in original and generated group than right decisions. However, 40% of misleading are observed in the generated group.

We found that the number of false positive selected images by all the participants is small (less than 20), and 30 images are selected by 5 participants (Figure 3.27, left). Almost 70 synthetic images are not selected at all by experts as being natural (first column). To observe

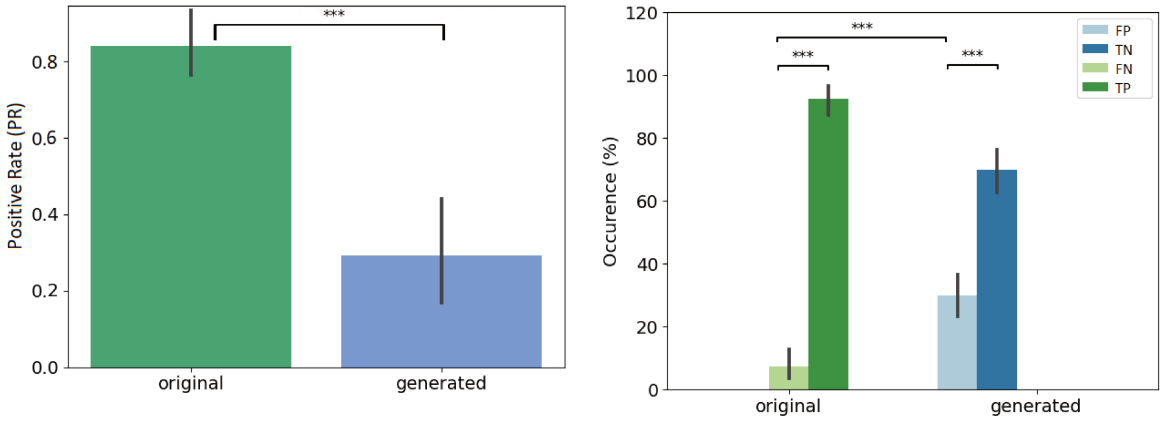


Figure 3.26: Overall view of decision per original and generated group. Error bars indicate the variability per expert. Left: Error rate per original and generated group. The baseline (original) corresponds to Positive rate PR. Right: Occurrence of answers per original and generated group. (* indicates p-values ≤ 0.05, ** p-value ≤ 0.01 and *** p-value ≤ 0.001.

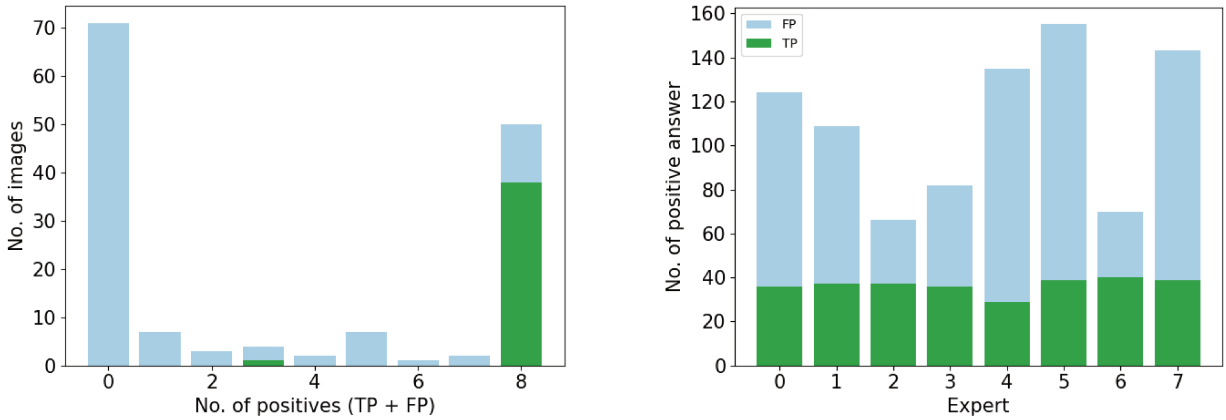


Figure 3.27: Study of positive answers. Right: Number of images per number of positive answers. Left: Number of Positive answers by experts.

the number of false positive answer by each experts, see Figure 3.27 right. Three experts answer less positive answers than the others (less than the half of the visualized dataset). Two of these expert are experts in microscopic acquisitions and not only biologists. One expert labelled around 150 images as being natural.

We retrieve the decision time before the expert give an answer whatever the kind of generation (Figure 3.28) left. Biological experts answer in the same time for generated and original images. This is confirmed also when the kind of generation is considered see Figure 3.28 left. However, the time is increased when they are misled.

Feedback on the psychovisual procedure

When we ask experts why they classify a generated image as a synthetic with hesitation, they point out the background (imprint, superimposed contours or artifact) or the image high noise, but they hesitate longer due to the credible organoid content. When they classify an original image as synthetic, it is because of a microscopic acquisition artefact, learned by an architecture, and reproduced on the worst synthetic images. Feedbacks on the experiment include having less images to label in a session, and a larger image on the screen.

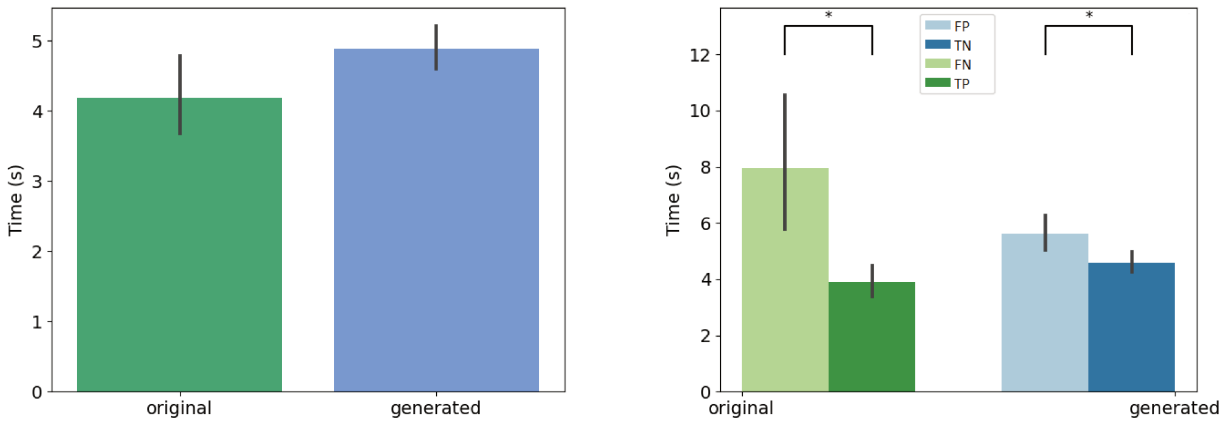


Figure 3.28: Mean decision time for all the experts. a) Per original and generated group b) per kind of decision

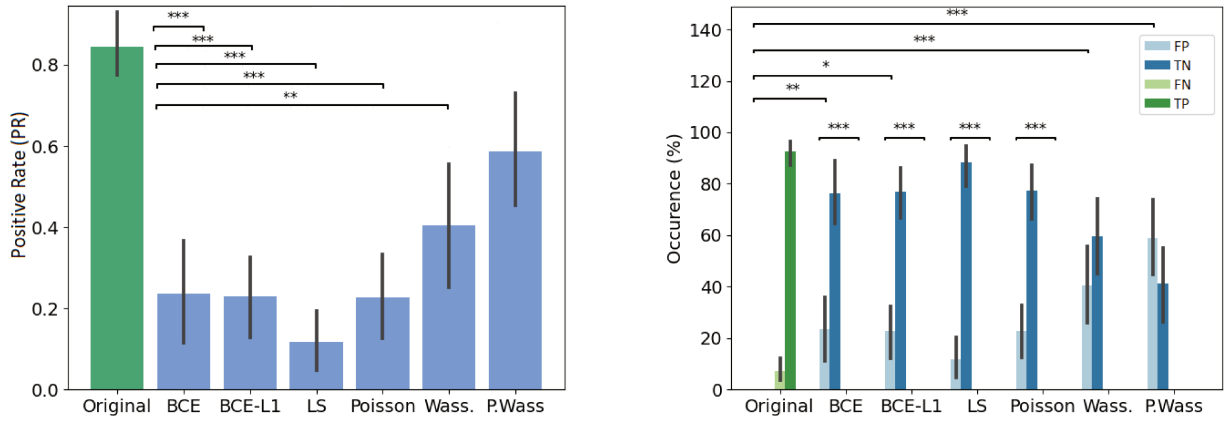


Figure 3.29: Overall results of decision per loss optimisation groups. Left: Effect of loss optimisations on the error rate. The baseline (original) corresponds to Positive rate (PR). Right: Occurrence of answer per loss optimisations.

Psychovisual evaluation of loss optimisation

The error rate is particularly higher in the Wass. and P. Wass groups than for the original one, see Figure 3.29 left. These particular loss optimisations drive the experts to mislead and consider the images from these two groups as natural. In Figure 3.29 (right), there is a difference between false positives of original and generated images from BCE, $BCE + L1$, Wass. and P. Wass. loss optimisation. However if we consider the intra-factor loss comparison, we can observe statistical differences between F_P and T_N of each for the BCE, BCE + L1, LS and Poisson loss rendering too small the proportion of misleading. There is no differences between these two occurrences decision for images generated by a Wass. or a P. Wass. loss showing 42 % of F_P and almost 60 % of F_P.

To observe which group of images is the most selected as positive, we observe the number of images selected by group in the Figure. 3.30 left. P. Wass. and Wass. images are selected as natural by the most of experts (a few Poisson, and a few BCE by 7 experts and L1 + BCE). The same three experts as in 3.27 only answer positive in most of the case for P. Wass. images see Figure 3.30 b). Four experts answer more synthetic images but more even for P. Wass. Only one expert seems to answer identically for all synthetic group of images.

If we do not consider the kind of decision, there is no difference of decision time per group

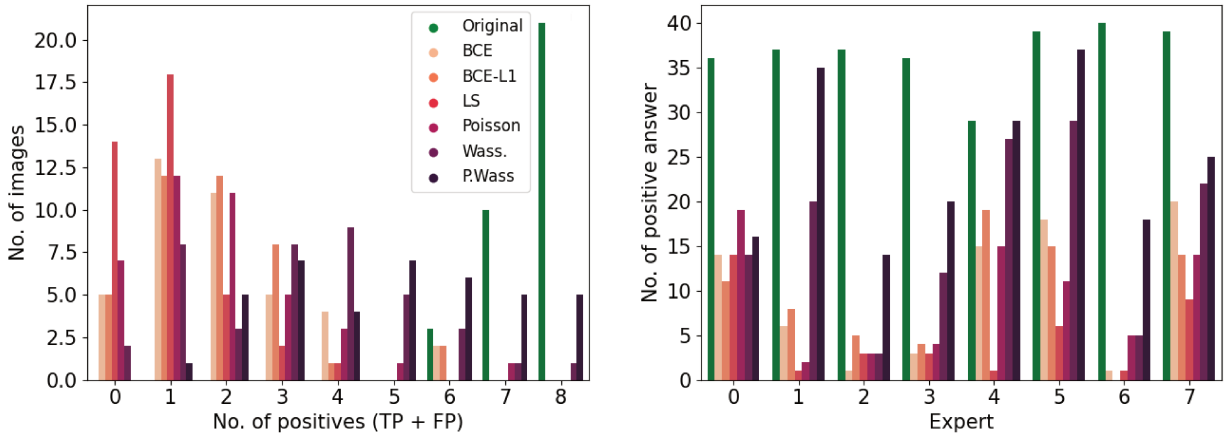


Figure 3.30: Study of positive answers per loss optimisation. Left: Number of images per loss optimisation per number of positive answer. Right: Number of positive per loss optimisation for each expert

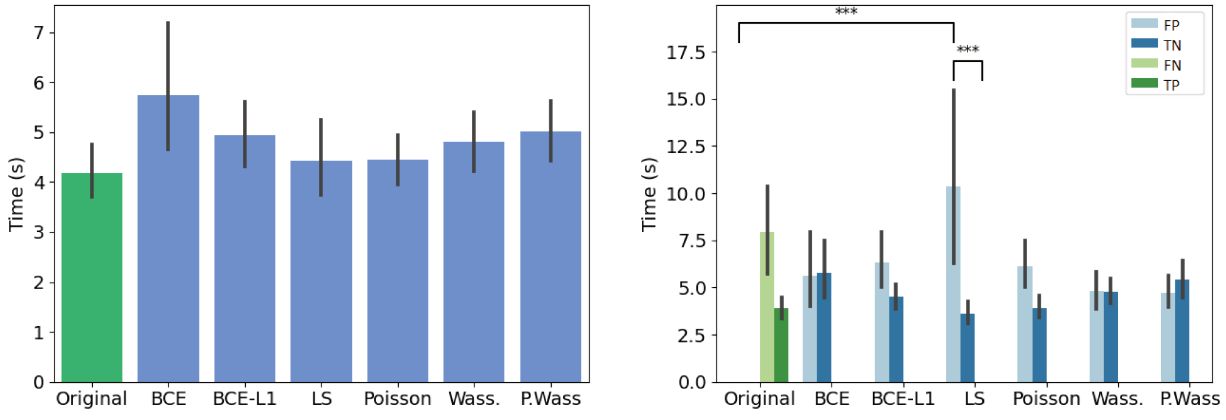


Figure 3.31: Mean decision time for per loss optimisations compared to the original input images. Left: Globally. Right: Per decision.

Table 3.8: Average and standard deviation of five of metrics on original and synthetic images per psychovisual decision (true or false) and per loss optimisation.

	Decision	Count	Blur	SSIM	PNSR	MSE	MI	UQI
			avg. – σ	avg. – σ	avg. – σ	avg. – σ	avg. – σ	avg. – σ
POSITIVE								
Original	TP	297	64 – 50	**0.78** – 0.17	**42.13** – 41.49	2974 – 3164	0.90 – 0.46	**0.86** – 0.12
BCE	FP	79	100 – 15	0.62 – 0.06	13.52 – 3.20	3713 – 2611	0.86 – 0.14	0.83 – 0.10
BCEL1	FP	77	104 – 39	0.61 – 0.07	13.54 – 2.94	3549 – 2320	0.82 – 0.13	0.83 – 0.09
LS	FP	39	**156** – 45	0.59 – 0.07	13.77 – 3.00	3404 – 2276	0.78 – 0.11	**0.84** – 0.09
Poisson	FP	74	93 – 23	**0.63** – 0.06	**13.86** – 2.84	3275 – 2140	0.86 – 0.14	**0.84** – 0.09
Wass.	FP	134	61 – 18	0.61 – 0.06	13.33 – 2.38	3499 – 1980	0.90 – 0.14	0.83 – 0.07
P. Wass.	FP	197	47 – 17	**0.63** – 0.07	12.48 – 2.44	**4272** – 2389	**0.95** – 0.17	0.80 – 0.07
NEGATIVE								
Original	FN	23	45 – 25	0.71 – 0.01	10.10 – 0.90	6418 – 1326	0.88 – 0.23	0.78 – 0.04
BCE	TN	241	105 – 21	0.62 – 0.06	13.57 – 3.16	3647 – 2555	0.85 – 0.13	0.83 – 0.10
BCEL1	TN	243	110 – 39	0.60 – 0.07	13.60 – 2.87	3478 – 2248	0.81 – 0.13	0.83 – 0.09
LS	TN	281	**157** – 44	0.58 – 0.06	13.79 – 2.80	3307 – 2119	0.78 – 0.11	**0.84** – 0.09
Poisson	TN	246	106 – 25	0.62 – 0.06	**13.86** – 2.77	3249 – 2090	0.84 – 0.13	**0.84** – 0.09
Wass.	TN	186	57 – 17	**0.63** – 0.07	13.41 – 2.58	3510 – 2116	0.90 – 0.13	0.83 – 0.07
P. Wass.	TN	123	41 – 14	**0.63** – 0.07	12.69 – 2.55	**4124** – 2415	**0.95** – 0.16	0.80 – 0.08

of synthesis, (Figure 3.31 left. When we study the decision time per group, the experts take more time to answer only when they are confronted to synthetic images generated with a least square optimisation (Figure 3.31 right.

Concordance of metrics and psychovisual evaluations

After observing the psychovisual decisions by generated groups, we compare qualitative and similitude metrics to the previous results in order to verify if the same groups are selected, but also to verify if some metrics or metric combinations can be used as a proxy to human psychovisual evaluation.

An overview of these results is given in Table 3.8 shows no differences between decision whatever the group of loss optimisation or the calculated metric. *Fid* is the highest for *Poisson* loss than for others groups whatever the kind of decision. BLUR metric is the highest for the decision with LS generated images. In term of SSIM, UQI indexes and PSNR, the decision is the highest for original images and no improvement is visible with generative methods. For MSE and MI the decision rate are the highest similarly for original and P. Wass. generated images. No differences are visible in term of decision with UQI.

Correlations To identify the metrics which corresponds the best the psychovisual evaluation, we plot metrics against error rate, the divergences between points inside these representations in Figure 3.32. We only show here the Blur scatterplot as an example of point representations (others scatterplot for single metrics are represented in Figure 7.1). In this graphic, we observe that the lighter color points (*Original*) are near the darker-violet ones (P. Wass.). This is represented in green in the successive heatmap, while purple represents a high KL divergence between two groups of points (such as between P. Wass. and *LS*). BLUR, SSIM, MI, *Uqi* maps represents the best the psychovisual evaluation decision (green cases for the last line). PSNR, MSE, and *FID* does not highlight the same results as psychovisual evaluation for each group.

The Figure 3.33 summarizes the correlations between the psychovisual assertion and the time or error rate according to each metric or metric combination. The main result is the

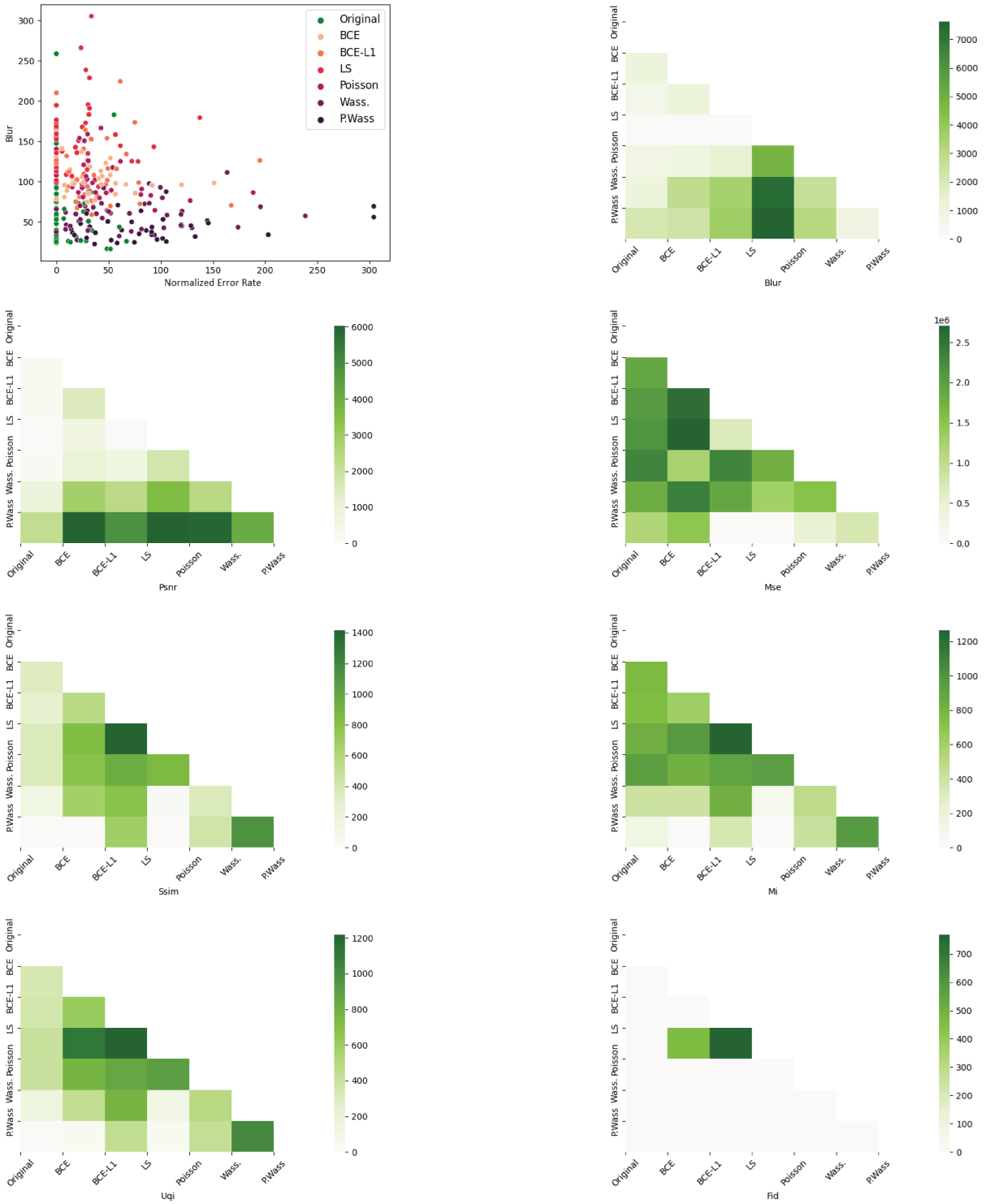


Figure 3.32: Part II: Comparisons of Metrics and various pondered decision time of error rate according to the group of images.

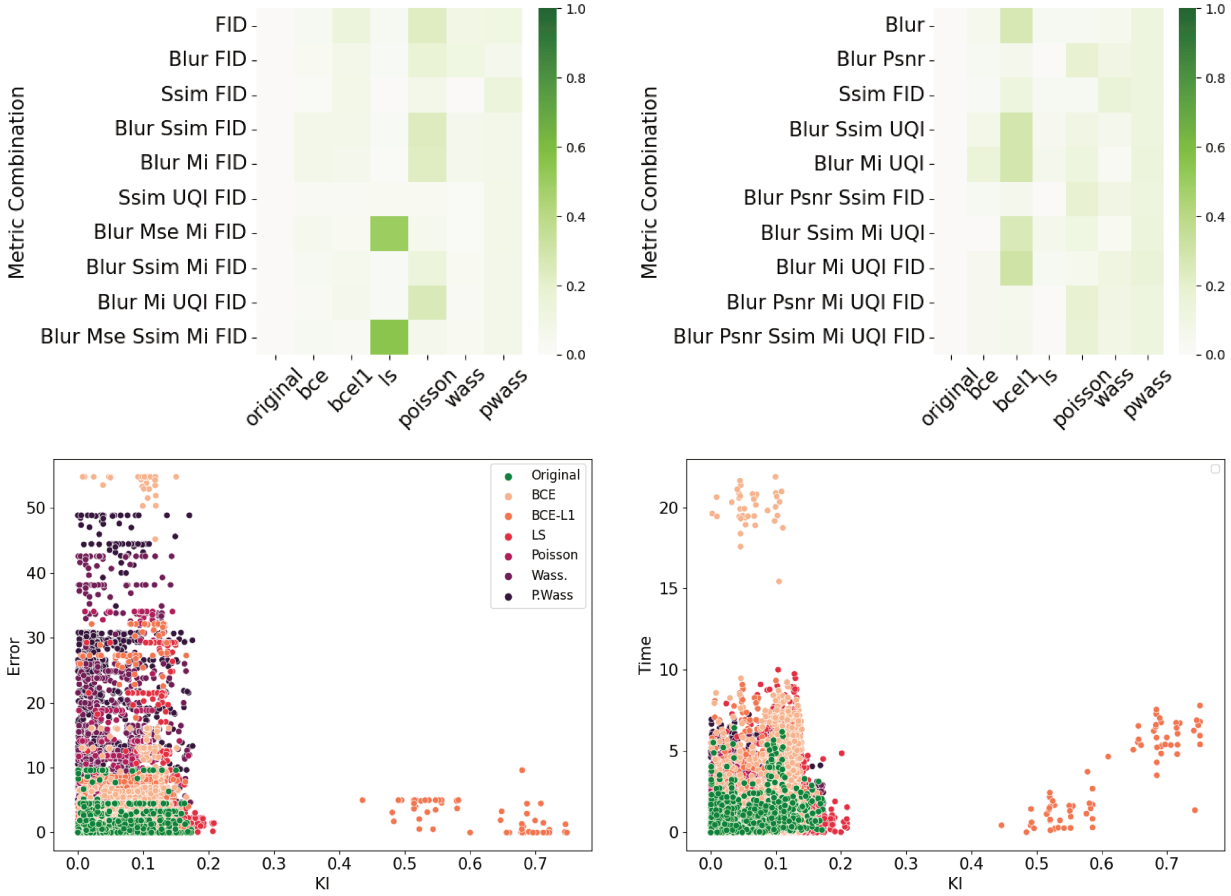


Figure 3.33: Correlations between metrics and psychovisual assessment (left error rate and right time) on various synthetic images. a) Correlation matrix of the ten best combinations b) Left Blur and FID representation and Right Blur Ssim MI UQI representation.

absence of correlation for single metrics, however, combination with a Blur metric, FID (for the error rate correlations) and Ssim-FID-Mi (for the time correlation) render the best results in Figure 3.33 top. The same result is represented in Figure 3.33 bottom. To observe the group representation between the Error rate or the Time and the KL divergence of points represented for these two combinations (Blur-FID and Blur-SSIM-FID-MI), see Figure 3.33 bottom left. The LS and BCE group are far from the others point representations. P. Wass. group is superimposing the original one with Wass. Others groups are not distinguishable, however, there are at the peripheral zone of the perceptual-original amount. The t-sne representation of these two metric combinations are Figure 3.33 bottom right.

3.5.3 Discussion

In this part we present to our knowledge the first psychovisual and metric evaluation comparison of Loss optimized generative adversarial network of brain organoid bright-field images. This study helps at validating most natural images generated by various AAE loss optimisations. We also contribute to strengthen metric evaluation by highlighting some images from optimized generated adversarial network to be percepted by Human biological expert as natural microscopic images: with a P. Wass. loss perception neurobiologists are misled 60% of the time and 40% by Wass.. They take more time to answer when they are misled. We compare human and metric evaluation and found mutual information to be the most

related to their decision metric, although no correlation appeared in our experiment for single metrics, but only for combinations including blur.

Biomedical experts select around 40% of synthetic images as natural compared to the original dataset. Thus, the generation by AAE networks generate a large part of realistic images such as the background of bright-field acquisition or, their content. The non selected images where considered sometimes as non natural due to some artefacts reproduced in some of them, or by a superposition of contours. Nethertheless the selected images could help in the future at training task for DL segmentation algorithms on brain organoid images for instance.

A first argument of the strong validation of the selected images as natural is the time to take a decision (Shaffrey et al., 2002). If the time to answer natural for a generated group corresponds to the time to answer natural or original images, we could consider these two groups are perceived as similar. We found no differences between original and generated images and thus whatever the kind of loss optimisation used to produce them. So they are not doubting when they classify an image as natural or generated. However psychovisual evaluation shows an increase of decision time before answering when they answer as false positive (depending of the loss optimisation) or false negative. This behavior is specifically shown from a Least Square Loss Optimisation generated images considered as natural. When we ask to participants why they have doubt on a particular image, they answer us it was linked with some acquisition artefact learned by the generated process and found on a lot of images (a bunch of cells) or, by a blurry contour which could may be due to the acquisition in the case of original images(Ali et al., 2022). For false negative answers, they only said that the artefact acquisition is also present (and they thought it was a generated). In the future, we think a pre-process image treatment has to be done on images to correct the acquisition artefact before the generative process, to avoid these false negative in the psychovisual evaluation or, to add a component in the generative network to avoid these artifacts (Ali et al., 2022; Galteri et al., 2017).

However only 15 images are selected by all the experts as natural. Five experts select a huge amount of images as natural and three less. This study raises a question: could we use images considered by only five experts as natural in a training step? Thus we thought we need to increase the number of biological experts to overcome future studies in order to be more precise on the number of images considered as natural. We could also analyze the answers by the field of expertise of biological experts too (separate those who made only culture or only microscopic acquisition from those working in both fields).

Nonetheless, human Psychovisual experts choose in majority images from generative adversarial network as natural if they are from Wass. and P. Wass. loss and, a few BCE, BCE-L1. These two first kinds of generation are also highlighted by most of the metrics in another study to have the better quality and to be the most similar to original images (Brémond Martin et al., 2021a) and see part 3.3. Thus, the psychovisual evaluation strengthen the choice of the use of these two and particularly the perceptual one in generative process. We can now confirm the idea that the regulation term of the Wass. distance between two images (Kupyn, Budzan, et al., 2018) could improve the learning of the pattern or characteristics of brain organoids in images and contribute to generate more natural images in term of content and aspects. We have also to consider to remove in a future study the non selected images by human psychovisual experts in the training session, in order to observe if it could increase the segmentation accuracy and see the impact on their morphological characterisation.

However the few BCE and BCE-L1 images selected as natural by psychovisual experts could maybe have also a great interest whereas the metric are not pointing them as natural

images (Brémond Martin et al., 2021a) and see part 3.3. As we know the use of metric is still controversial for the GAN evaluation as they are measuring similitude and quality (Borji, 2018).Here, we could not highlight a strong correlation between the use of certain metrics and the decision to reject or not a generated image as natural. To correlate a metric and psychovisual evaluation instead of a binary answer 'natural' or 'not natural', some authors use a graduation scale (Pedersen, 2015; Pedersen and Hardeberg, 2012). This could explain our results, we use only a binary answer (natural or not) and not a graduation scale. In future studies we could use a graduation scales instead of a binary answer.

In consequence, no metrics used in this study could replace a human perceptual evaluation to decipher the naturality of an image generated. There is a certain link with FID, BLUR or MI and the group and MI with the mean decision but it remains weak. We could only say that similitude and referenced-bases metrics are more linked to the decision than qualitative metrics and non reference-based metrics. And when we compare metrics with decisions some patterns appear according to the kind of loss optimisation. The use or not of a measure to decipher natural generated examples is an issue recently discussed (Borji, 2018). To compare fairly images generated by various optimised models, there is no consensus for a use of a particular metric. In other fields such metric comparisons highlight a wsi and SSIM metric in term of quality assessment(Pedersen, 2015). However, we do not want an identical image but one just resembling as a natural one. This could explain these metrics are not well designed for the GAN specific evaluation when they are lonely considered. This study comparing the overall psycho-visual evaluation and 7 metrics is a first step to help at pointing a metric of "natural".

Based upon our KL divergence maps, we suggest that a combination of metric (BLUR, SSIM, MI, UQI which represents the best the psychovisual evaluation decision) could maybe help at replacing one day the human psychovisual evaluation which is time consuming. Nevertheless, this work on metric combinations replacing a psychovisual evaluations need to be further studied. In other fields the combination of metrics help at pointing out some results in term of quality or similitude (Susu Yao et al., 2005; Pedersen and Hardeberg, 2012; Okarma et al., 2021). Another idea could also to use non reference quality metrics combinations (Rubel et al., 2022). Some authors also try to implement directly a discriminator of generative adversarial networks based upon human perception, this could be a solution if not time consuming (Arnout et al., 2021; Fujii et al., 2020). It is not the case in this study, as we for us, an important task is to found an appropriate metric for highlighting "the naturality" of the image and replacing the psycho-visual evaluation. An idea is to test psycho-metrics instead of classical similitude or quality metrics such as Hype from (S. Zhou et al., 2019) which is an alternative of FID from (Heusel et al., 2017), or implementing the GFI quality assessment created by Y. Tian et al., 2022.

We could also applied this psychovisual evaluation on others datasets to attempt to answer more specifically to the metric replacement. We thought about noise optimised generated images of brain organoids with an AAE for the same aim (Brémond Martin et al., 2022). It could be interesting to observe if with a noise injection, similar to the bright-field acquisition images, generated images are more perceive as natural even if metrics are not pointing a particular kind of noise. Indeed, qualitative and similitude metrics point out Gaussian noise and shot noise injection. But as said previously, this could maybe only due to the metric choice (Borji, 2018). An analysis of Psychovisual evaluation could maybe help at highlighting a combination of metrics. In this future study, it could be also interesting to observe for example the microscopic experience of the Biological expert as a new criterion. A larger application of this methodology could be made on others kind of generation (such as on GAN (Goodfellow et al., 2014), DCGAN (Radford et al., 2015)...) and maybe help

at pointing out the better GAN model for brain organoid generation used during a training segmentation task.

In this study we use an unbalanced dataset with more synthetic images than original ones. Nethertheless, biological experts does not know the number of real or synthetic images, so the test is unbiased. In future studies, we need to obtain and use more original images in order to balance the set. We use an in-house software created specifically for the psychovisual task for brain organoid images, and identified a few limitations. We have to run batch process with pauses to limit the tiredness of biological experts, in the same way thant others psychovisual evaluations (Shaffrey et al., 2002). We also have to add a cursor with a score instead of a button to estimate a natural range, to facilitate correlations studies (Y. Tian et al., 2022). The size of the image of the screen has to be increased, but not for all the participants. Depsite the identified limitations, the use of the software is simple and practical according to users feedbacks. Concerning the Biological experts, we have to increase their number to strengthen our statistical analysis. Another amelioration would consist in increasing the number of biological experts. However, it could include biases of the experts: it requires knowledge of which image is considered as natural and which one is not for the target domain, so the number of experts available in the field is diminished while our task required more experts. The performance of human judges is not fixed and can improve over time, other articles choose a validation by 15 experts for instance which is not possible in our biomedical context which requires experts in the field (Salimans et al., 2016; Denton et al., 2015). More over, psychovisual evaluation is limited to the number of images that can be reviewed in a reasonable time (Borji, 2018).

In this study psycho-visual evaluations allow us to:

- Validate some synthetic image generated from loss optimisation of generative brain organoid images with an AAE in term of decision time and decision.
- Describe the quality and similitude of the synthetic images with the original dataset by a metric validation.
- Verify if some synthetic images could be considered as natural by psychovisual expert decision.
- Compare psychovisual and metric evaluations.
- Paves the way to finding a metric or a metric combination that mimics psychovisual evaluations.

This selected images could be use in the training phase of a segmentation task in order to help at their morphological development characterisation.

We also need to evaluate psychovisually noise injected optimized synthesized images.

In future studies we suggest a combination of metrics or a perceptual metric could maybe help at replacing the psycho-visual assessment which is time consuming. Such methodology could be used for others brain organoid data-sets generated with a generative adversarial network.

3.6 Evaluation of noise-injected synthetic brain organoid images

This section describes unpublished works.

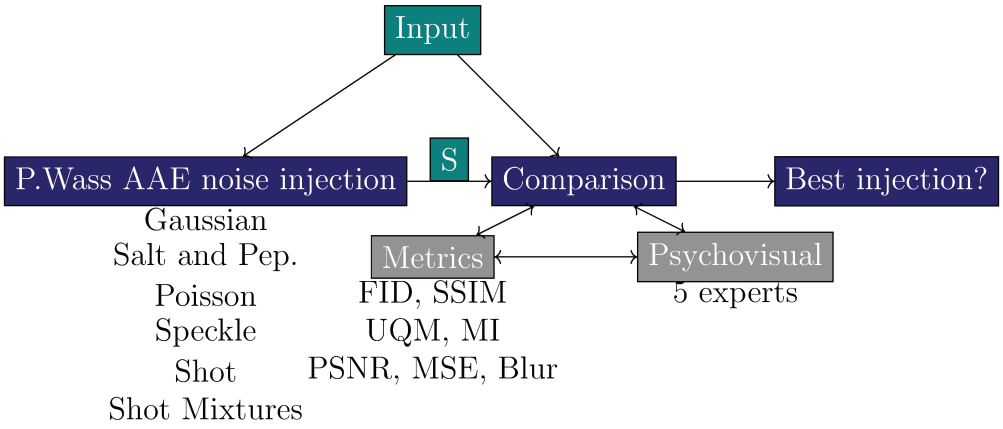


Figure 3.34: Procedure of noise injection psychovisual evaluation in AAE with P.Wass loss optimisation for the generation of bright-field brain organoid images more similar to microscopic natural acquisition. S corresponds to synthesized images.

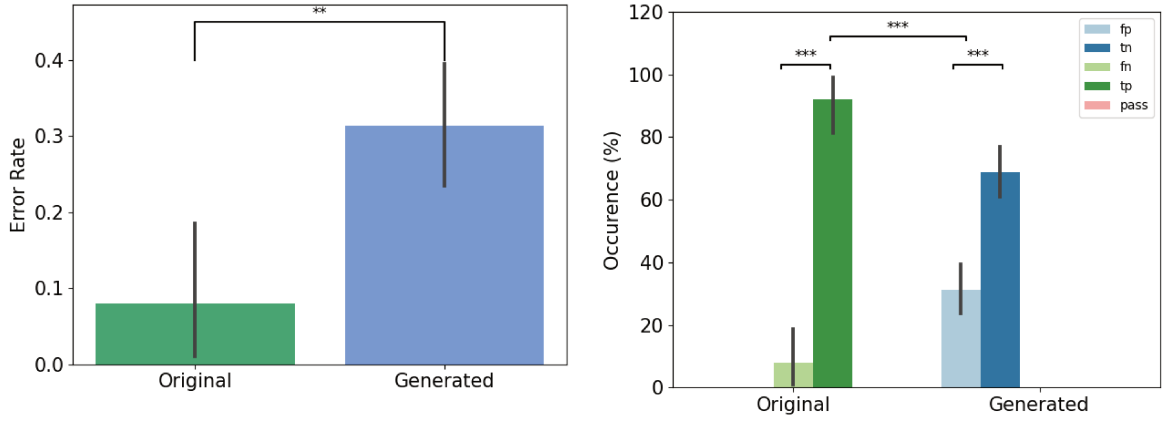


Figure 3.35: Decision per original and generated group. Left: Error rate per original and generated group. Right: Occurrence of answers per original and generated group. * if p-values ≤ 0.05, ** if p-value ≤ 0.01 and with *** if p-value ≤ 0.001.

Previously we validate psychovisually the effect of loss optimisation on generative network synthesizing brain organoid bright-field images. In this part, we attempt the same procedure on noise optimized generative network results which is summarized in Figure 3.34.

We process as similarly to 3.5. The only difference consist in the number of Biological experts: only five of them could conduct this evaluation for time consideration.

3.6.1 Results

Psychovisual evaluation of original and synthetic images

To observe differences between the original and synthetic groups, we first compare the error rates in Figure 3.35. The Error rate ratio is increased in the case of synthetic images by 3 times. If we compare inside generated group, TN answers are significantly higher than FP, but there is a difference between FP and TN answers.

Individual answers show most of images are classified in non natural and only 5 images are classified as natural by all the experts see Figure 3.36 left. Three experts choose more than 100 images as natural and the others two 75 and 50 images only see Figure 3.36 right.

There is no differences of decision time between the synthetic and the original group,

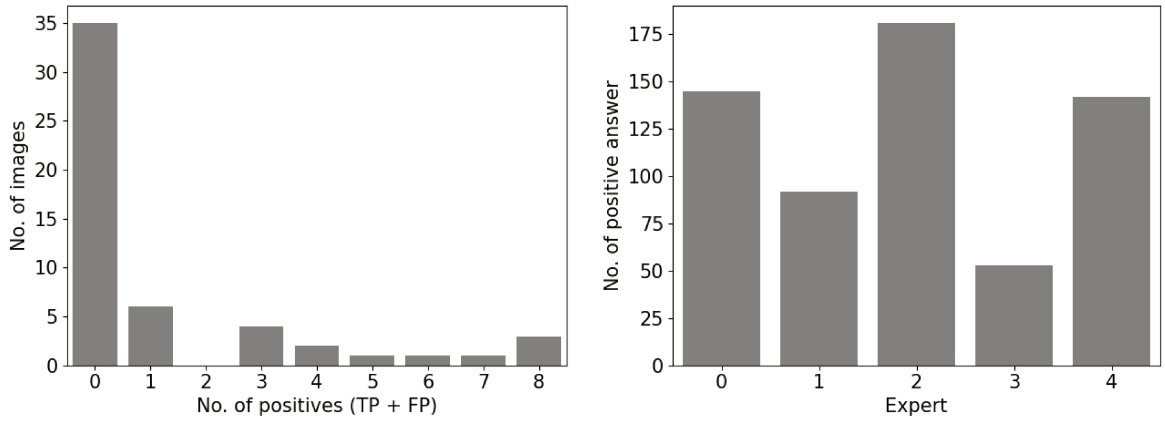


Figure 3.36: Study of positive answers. Right: Number of images per number of positive answers. Left: Number of Positive answers by expers.

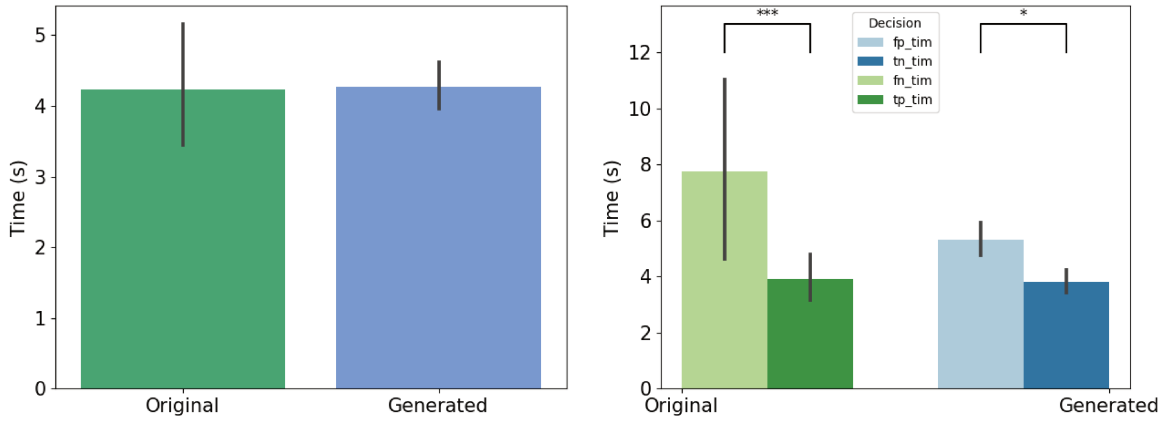


Figure 3.37: Mean decision time for all the experts. a) Per original and generated group b) And per kind of decision

however, when they are mislead, experts took statistically more time to answer within the synthetic or the original group, but not between the groups see Figure 3.37.

Feedback on the psychovisual procedure

When we ask to experts why they classify a generated image as a synthetic with hesitation, they answer the background contains an imprint, or there is an artefact, or the image is too noisy, but they hesitate longer due to the possible content. When they classify an original image as a synthetic is because of a microscopic acquisition artefact, learned by an architecture, and reproduced on the worst synthetic images. They would like to have less images in a session, and a larger image on the screen.

Psychovisual evaluation of noise optimized synthetic images

The error rate and occurrence of each decision per group of synthesis are summarized in Figure 3.38. There is no difference between original and the various synthetic groups due to variations. However, while looking at the decision occurrence, a statistical difference is observed in Poisson, Speckle, Poisson Gaussian and Poisson Speckle such as in the original group. Only the Gaussian, Salt and Poisson Salt groups presents no differences between the

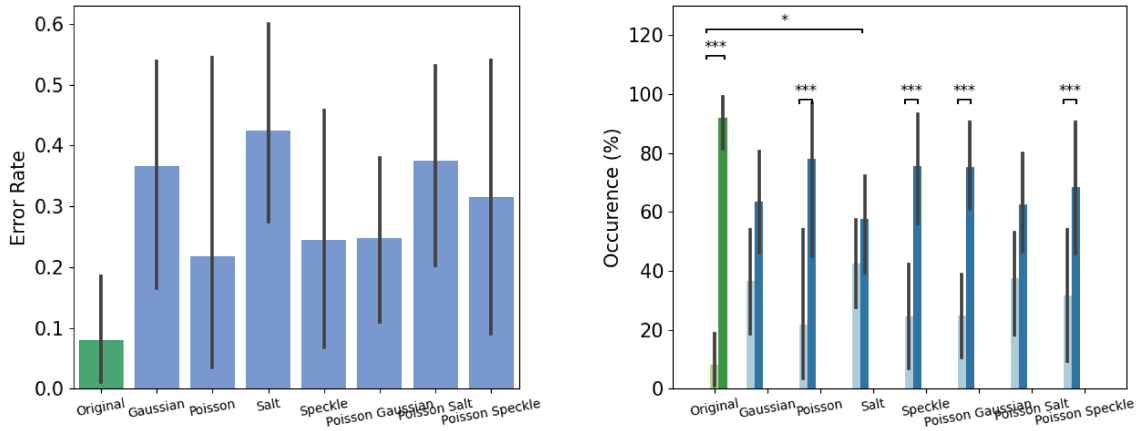


Figure 3.38: Overall results of decision per loss optimisation groups. Left: Effect of loss optimisations on the error rate. Right: Occurrence of answer per loss optimisations.

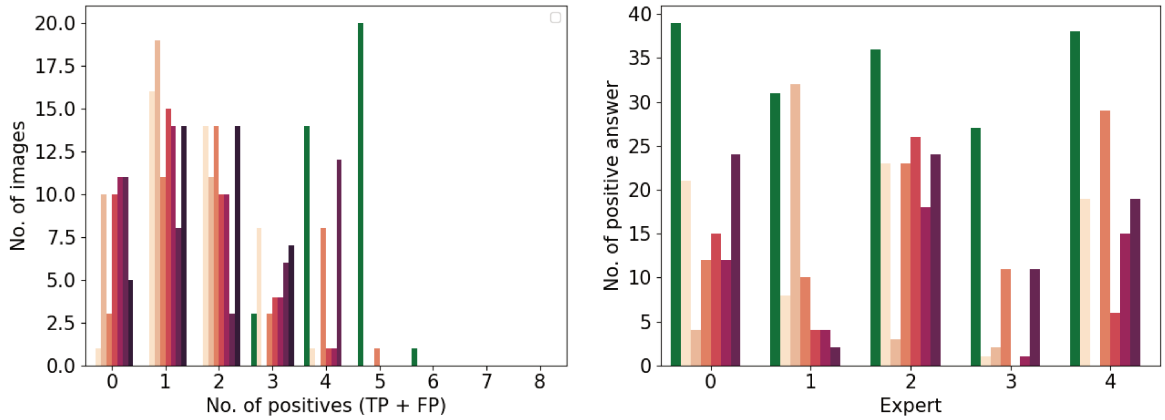


Figure 3.39: Study of positive answers per NOISE optimisation. Left: Number of images per loss optimisation per number of positive answer. Right: Number of positive per loss optimisation for each expert

TN and FP. More over, a difference is observed between FP in the Salt group confronted to the original, showing, this group mislead particularly the experts.

The individual answers per groups show interestingly original images are only deciphered by 4 or 5 experts, in Figure 3.39. If we compare experts positive answers, three experts answer more for positive for the salt group, one for the Poisson and one for the Poisson.

We do not observe differences for decision time between the original and various noise synthesis Figure 3.40 and also whatever the kind of decision.

Concordance of metrics and psychovisual evaluations

After the observation of psychovisual evaluation, we compare metrics evaluation with the psychovisual one. The aim of this part is to verify if they highlight the same results and in a second time if some metric combination could replace the psychovisual evaluation.

An overview of these results is given in Table 3.9. For positive answers, only the UQI and PSNR show best results for Gaussian and almost for Speckle and Poisson Gaussian mixture. All the other metrics have the best value for the original group. Concerning the Poisson Salt it reaches the worst values for almost all metrics, and Salt, the second worst values.

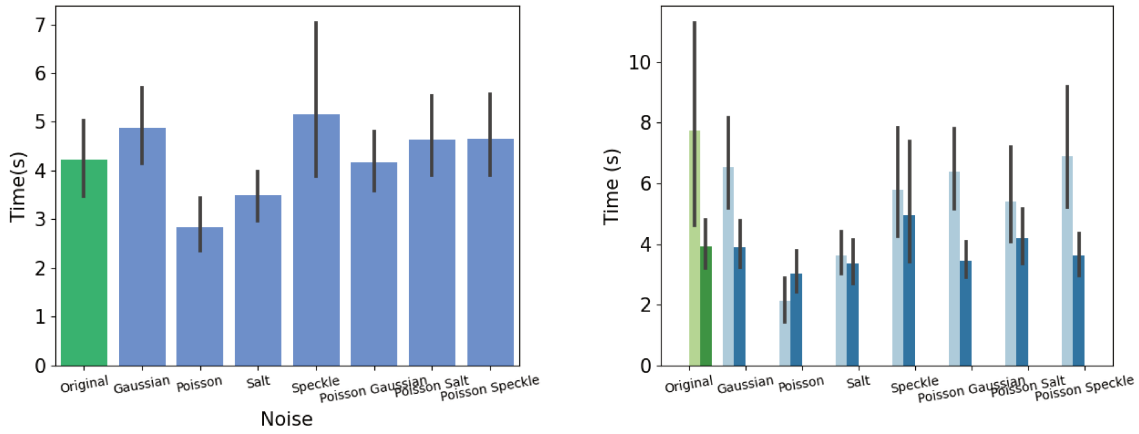


Figure 3.40: Mean decision time for all the experts per loss optimisations. a) Per original and generated group b) And per kind of decision.

Table 3.9: Average and standard deviation of five of metrics on original and synthetic images per psychovisual decision (true or false) and per noise injection.

	Decision	Count	Blur	SSIM	PNSR	MSE	MI	UQI
			avg. – σ	avg. – σ	avg. – σ	avg. – σ	avg. – σ	avg. – σ
POSITIVE								
Original	TP		**61.95** – 48	**0,68** – 0.08	**14.87** – 14.28	4298 – 2819	**0.93** – 0.56	0.81 – 0.09
Gaussian	FP		166 – 53	0.60 – 0.06	14.26 – 2.82	1975 – **1935.35**	0.83 – 0.13	**0.85** – 0.09
Poisson	FP		76 – 6	0.67 – 0.06	13.82 – 2.06	3018 – 1484	0.70 – 0.10	0.82 – 0.08
Salt	FP		1231 1260	0.48 – 0.13	11.32 – 3.24	6027 – 3694	0.70 – 0.14	0.82 – 0.13
Speckle	FP		731 –305	0.39 – 0.08	14.19 – 2.30	2844 – 1549	0.64 – 0.09	0.84 – 0.07
Poisson Gaussian	FP		167 – 49	0.59 – 0.06	14.01 – 3.13	3289 – 2312	0.83 – 0.12	0.84 – 0.10
Poisson Salt	FP		2099 1414	0.38 – 0.16	10.01 – 2.62	7636 – 4266	0.67 – 0.13	0.66 – 0.14
Poisson Speckle	FP		251 – 83	0.54 – 0.07	14.08 – 2.95	3157 – 2131	0.80 – 0.12	0.84 – 0.09
NEGATIVE								
Original	FN		66 – 46	0.67 – 0.08	14.80 – 14.27	4308 – 2801	0.91 – 0.57	0.81 – 0.09
Gaussian	TN		173 – 46	0.59 – 0.06	14.23 – 2.77	2982 – 1924	0.82 – 0.12	**0.85** – 0.09
Poisson	TN		76 – 5	0.66 – 0.06	13.80 – 2.06	3026 – 1487	0.69 – 0.09	0.82 – 0.08
Salt	TN		541 –799	0.53 – 0.09	12.94 – 3.33	4287 – 3032	0.75 – 0.12	0.79 – 0.12
Speckle	TN		805 –314	0.37 – 0.07	14.12 – 2.29	2886 – 1564	0.63 – 0.09	0.84 – 0.07
Poisson Gaussian	TN		175 – 59	0.58 – 0.06	14.03 – 3.03	3231 – 2221	0.81 – 0.12	0.84 – 0.09
Poisson Salt	TN		901 1295	0.49 – 0.15	12.29 – 3.16	4943 – 3677	0.72 – 0.13	0.77 – 0.14
Poisson Speckle	TN		224 – 68	0.55 – 0.06	14.01 – 2.95	3202 – 2149	0.80 – 0.12	0.84 – 0.09

For negative answers, they do not reach highest or worst values for a metric. The Gaussian noise reaches the best value in UQI, and speckle the worst in MI.

The Figure 3.41 summarize correlations between the pondered error rate and metric values. The highest Blur, UQI, PSNR MSE, MI and SSIM correlations are observed for Salt injected optimized generation. Then Poisson-Speckle mixtures reaches the highest values for Blur, SSIM, MSE, MI and UQI. The third one is Poisson Salt for PSNR.

Original and Gaussian show similar correlations values for MI, PSNR and Blur. Original shows similar correlations values to Poisson also for the Blur, Uqi, Psnr and Ssim (and a few for MI). Original correlations are the most different with Salt ones except for the MSE.

The ten best metric combinations recapitulating the best the relations between the psychovisual evaluation and the kind of synthesis are shown in Figure 3.42. Best correlations are observed for Salt and Poisson-Salt in term of decision, but no strong correlations are observed concerning the time. The ten best combinations for the decision are composed by Blur (and for almost by PSNR) while the ten best combinations for the time are composed by MSE (and for the almost by MI). The Original correlations are similar to those from

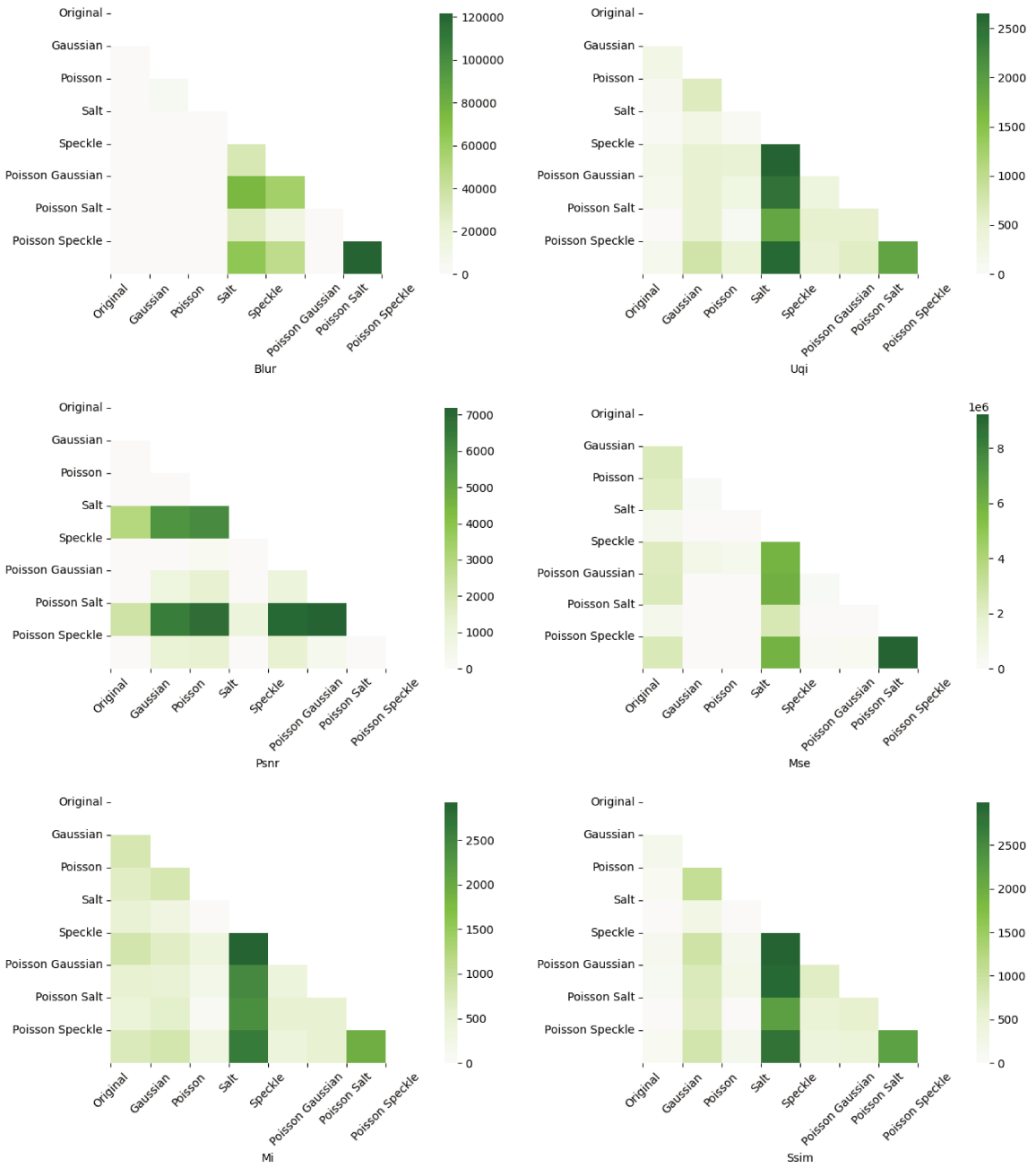


Figure 3.41: Comparisons of Metrics and various pondered error rate according to the noise injected.

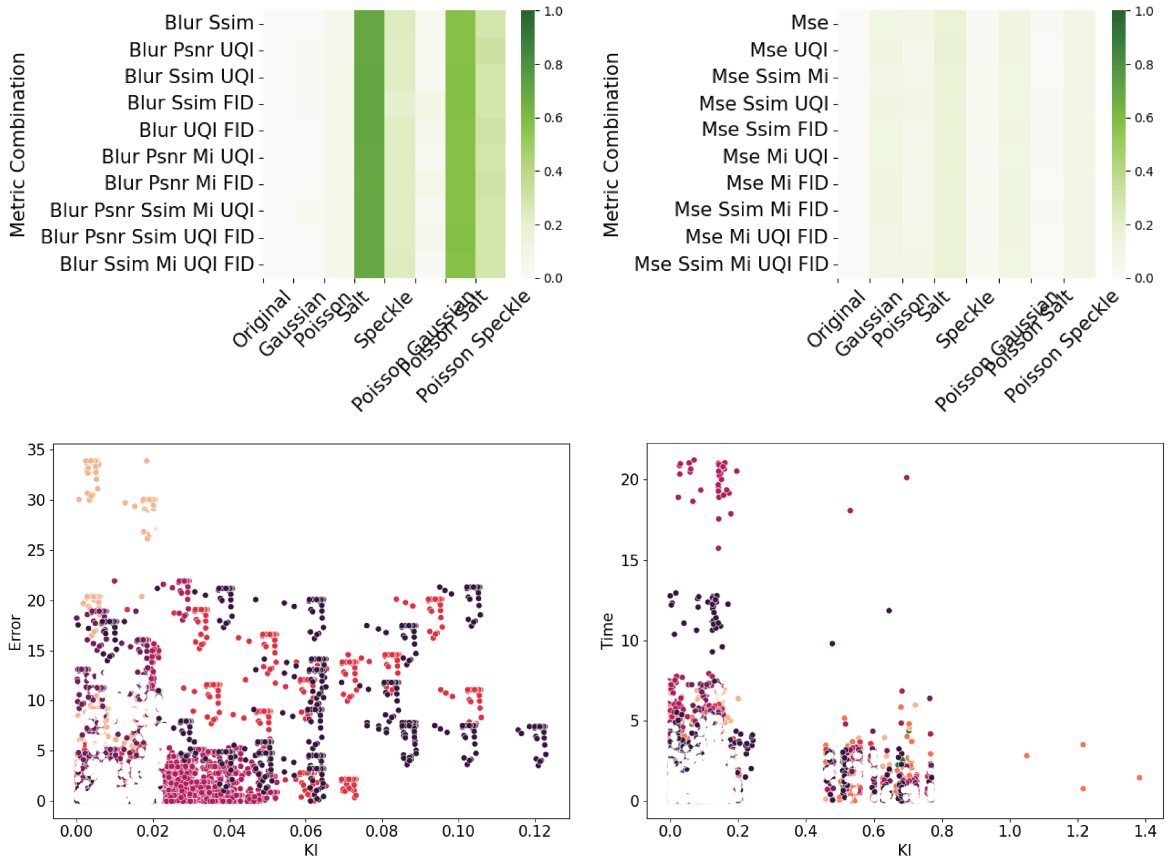


Figure 3.42: Correlations between metrics and psychovisual assessment (left error rate and right time) on various noise injected synthetic images. Top: Correlation matrix of the ten best combinations. Bottom: representation and representation

Gaussian, Poisson and Poisson-Gaussian mixture in decision or time. The two figures at the bottom of Figure 3.42 recapitulate the two best combination of metrics

3.6.2 Discussion

In this part we evaluate psychovisually a synthetic dataset of brain organoid bright field images generated with various AAE noise injections, and compare it with automatic evaluation. Results highlight generated images with noise injection could be considered as natural by Biological experts. They take more time to answer when they are misled except for Poisson noise injection. Gaussian, Salt, and Poisson-Salt variations images are the most considered as natural while in a metric point of view, Salt and Poisson-Salt give the worst results. However we have to consider the few number of experts in this experiment and there is some uncertainty to conclude on the best noise optimisation according to these results. Gaussian seems the best in term of metric and decision comparisons, and Poisson gives similar profile of correlations (of metric and pondered error rate) to the original images. The combination of metrics which could replace the psychovisual decision contains Blur, and for the time MSE. A discussion on all of these results follows in this subsection.

Thirty percents of AAE generated images with noise injection are considered as natural by biological experts. This percent is the same with previous reported results on Loss optimisations with this same architecture and highlight some consistency see Section 3.3. As previously discussed the non natural aspect of synthetic images is due in majority to

reproduced artifacts (white imprint, bluriness, or a noise too strong on the image), already discussed in previous GAN studies with a biological purpose (Salimans et al., 2016; Goodfellow et al., 2014). These selected images could help to train a DL task such as classification or segmentation of brain organoide images.

The validation of Poisson images is strengthen by the fact they take less time to answer to this kind of generation which is not reproduced with other noise injected synthetic images. Natural bright-field acquired images are based upon this noise also called shot noise (Boyat et al., 2015; Gilroy, 2019). However, only a few of Poisson synthetic images (around 20%) are considered as natural, which result in the fact not only the noise injection could allow to consider these images as natural see Figure 3.38. We suppose, biological experts, misled by those images displaying sibling background of natural images with a high contrast content, answer directly they are natural.

The noise injected synthetic images producing the most misleading are Gaussian, Salt and Gaussian Salt mixture images with almost 40% of images from each group selected as natural. This is less than for our previous observation on loss optimisation (with the only Perceptual Wasserstein optimisation around 60% of image generated are considered as natural) see Figure 3.29 in the part 3.3. The addition of noise during the generation could maybe attenuate the natural of images. It could be interesting to ask to experts if an image only optimized by a perceptual wasserstein loss is most natural for them than a perceptual wasserstein with noise in another experiment. The Gaussian noise is the noise the most used for GAN synthesis (Salimans et al., 2016; Sonderby et al., 2016; Feng et al., 2021). Interestingly Poisson noise images are not the one considered as the most natural while the background of original microscopic images are exhibiting this noise(Boyat et al., 2015; Gilroy, 2019). Only the Poisson Salt mixture and the Salt synthetic images are similarly to the Gaussian images considered as natural. The other noise mixtures are not producing satisfying results (Poisson - Gaussian and Poisson-Speckle), while our first thought is to hypothesize the most used and most similar background mixture noise could be more natural (Salimans et al., 2016; Boyat et al., 2015; Tanh et al., 2015). The explanation could be the mixture of Poisson and Gaussian noise characterize the electronic and IRM acquisitions and not the bright-field microscopic one. More over, the Poisson Salt mixture and Salt render the worst results according to metric considerations, which render the strong validation by Biological experts surprising see Section 3.4. The Salt noise is not considered in a biological point of view to be natural because in biomedical images is created during the acquisition in the data transmission step (Salimans et al., 2016). We think the combination of Poisson and salt and pepper noise added during the acquisition is better recognize by Biological experts than a single Poisson noise added in synthetic images. And thus Biological expert validate more synthetic images producing some features of acquisition issues. Looking at the particular images considered as natural for each individual, we observe almost the same 10 salt images are selected by the 5 experts, which could highlight those images could maybe also used in future tasks. It raises a question what bring some noise optimisation for some images which are not given in others from the same group of generation and are so not validated? Furthers studies need to be engaged on this domain to answer this question.

However, these results have to be analyzed with caution. Only five experts could answer to this experiment contrary to the one from 3.3 and thus, these results are less robust than the previous study. Maybe with more Biological expert we will highlight others noise to be considered by individual as the best. Here the Gaussian noise seems the one preferred by four experts, but one answer positively 30 times on 40 for Salt synthetic images. The doubling of experimental experts could help us to conclude on this part.

Comparing metric and psychovisual evaluation highlight the Gaussian noise as the best

noise optimisation, while correlation highlight the Poisson one. Metrics could not allow us to choose a particular noise optimisation as the best one, nor the statistical analysis with dimensional reduction in the previous part see Section 3.4, nor an expert validation. It seems only a combination of Human and automatic evaluations could decipher a particular noise optimisation as the best. The question which is raised until then is: Do we have to choose a comparison or a correlation to highlight this best noise image generation, or could we choose the both noise?

Such as in previous experiment on loss optimisation see part 3.5, a Blur metric combination could maybe replace a psychovisual decision but not the time consideration. Similarly as previously discussed, we may not conclude as we only have five point of views, and more should help to strengthen this result.

In future articles, we should thus increase the number of biological as previously discussed, but not only. The analysis of Biological expert background need to be further analyzed: These experts are only experts in biological cultures, or also in microscopic acquisition? In this study it may influence the final decision. We could also modify the binary button with a graduation scale to calculate, such as others articles, directly the correlation between the certainty of natural decision and numeric metric values (Pedersen, 2015; Pedersen and Hardeberg, 2012). Another experiment could be to ask to the biological expert to compare and select the most natural image between a not and a noise injected synthetic image. We also have to consider to incrise the size of image inside the screen to facilitate the decision of Biological experts. Finally another psychovisual experimentation could be conducted on other noise mixture, we do not consider here for time considerations, such as Gaussian or Salt based.

In this part we:

- Validate noise injected synthetic images as natural by Biological experts.
- Consider various group of noise injection to be the most validated as natural, and that depends at the same time from the noise of the background similar to a natural biological image and from the sharpness of the content.
- Conclude on the requirement to increase the Biological expert number in future experiments.
- Gaussian and Poisson noise are highlighted by Metric and Psychovisual comparison and correlation.
- We show for the second time a combination of metric based upon the Blur may replace psychovisual evaluation.

We expect these validated noise injected synthetic bright-field brain organoid images could help to train DL algorithm, for a classification or a segmentation purpose.

3.7 Discussions and conclusions on generation

The aim of this generative part is to increase the number of a tiny dataset of bright-field brain organoid images, in a qualitative manner in order to segmente and characterize latter their shape with DL algorithms. By comparing various architecture comparisons, loss optimisation and noise injection, we produce certain synthetic images which may impersonate natural images from three point of view: automatically by metrics, statistically by dimensional reduction and psychovisually by Biological experts. In this discussion we will

highlight the benefits of these results, some drawbacks, the initial work on other datasets and, academic/ industrial future perspectives of research of this work.

Discussion on our work In our work we validate an AAE architecture with a Perceptual Wasserstein Loss optimisation with certain noise injections to synthesize the most natural synthetic images of bright-field brain organoid images. Thanks to this work the brain organoid field could now compensate the lack of bright-field images for their developmental characterisation (Chakradhar, 2016) by applying this data augmentation strategy. This strategy could be improve by comparing others architectures or other optimisations, however, it gives yet satisfying results in three validation ways and we think could be use in future works such as segmentation or classification tasks as given. It could be interesting for others microscopic acquisition of brain organoid images, which are also suffering from limited number, such as confocal two-photons or light-sheet microscopy ((Brémond Martin et al., 2021b) or see part 2 to follow the same procedure of data augmentation comparison strategy, we propose here). Similarly, we need to compare various GAN architectures, loss and noise optimisations on these others microscopic acquisition of brain organoid cultures. It could be interesting particularly to study the effect of noise injections in this context: such as light buzzing for confocal (Tanh et al., 2015), or bi-photon (Scott et al., 2021), or light-sheet microscopy Gaussian beam effect (Hillman et al., 2019). The validation on other images of bright-field biological models could help to generalize our procedure. For instance, a validation step could be made on developmental blastocytes bright-field image dataset from this link https://bbbc.broadinstitute.org/image_sets.

Initial work on perspectives The main developed annexed work concerns the validation of our procedure with another bright-field cortical organoid image datasets from the Nice Sofia-Antipolis Valrose Biological Institute with Physiological and Pathological mechanisms of brain development team (thanks to Michèle Studer and Michele Bertacchi).In this dataset only 30 images could be used for the data augmentation process we increase the dataset as previously with only metrics and statistical evaluations. Some results are summarized in the annexes part 7.1.1. Additional works concern the effect of combined loss instead of a single loss on the generative process, and the effect of updating loss during the learning step, developed by two master students.

Academic and industrial perspectives Industrial and academic perspectives from this work are various:

- It confirms for biological culture type, on a certain acquisition modality, for a tiny dataset, a comparison sheme has to be established. The most use architecture, and optimisations are not suited to augment this dataset.
- The interest of this comparison scheme is it could be developed and propose to augment others biological models.
- We nonetheless have to compare others architectures such as double Auto-encoder generative architecture to augment the background and content simultaneously (Kupyn, Martyniuk, et al., 2019)and observe the impact of each noise on this kind of generation.
- We also have to test a DifAugment architecture which augment real and false samples to produce a best generation of image according to (S. Zhao et al., 2020). With the same idea we also have to test or improve architectures based upon transformers such

as InfoGAN which render bad synthetic images (B. Hu et al., 2019) but which seems to be promising for the future.

- The effect of input number variation during the generation has also to be studied.
- Concerning the loss optimisation, we could apply the perceptual component or a regularisation term on all the other available loss tested, to observe its impact (Isola et al., 2017).
- For other microscopic acquisition and for other biological models, a loss/noise comparison and other losses/noise should be tested.
- The noise injection impact should be tested on the input of the generative architecture, or/and on the feature space.
- The architecture developped is not adapted to generate specific stages of culture, or a particular pathological organoid state, given the very few input images available (approx. 3 per category). In the future, instead of using a cGAN (Yi et al., 2019), we could try an AAE (Makhzani et al., 2015) which take in input the label of each image. An idea could be to complete it with a shape prior based loss such as in El Jurdi et al., 2021, a topological loss (Clough et al., 2020), or a constraint loss (Bateson et al., 2021) originally produced for segmentation tasks, to synthesize a more accurate generation per group.
- The development of each of these ideas could open new issues to answer and to develop. The main idea would be developing a scalable comparative methodology which augment in the best way each biological tiny dataset whatever its kind of acquisition.

Conclusion This work is the first to:

- Augment brain organoid images with GANs,
- To compare architectures, loss and noise optimisation for this aim.
- To compare automatic and psychovisual evaluation
- To attempt replacing psychovisual evaluation by a combination of metrics.
- To be validated on another dataset (except yet for the psychovisual validation part).

This work and approach should be tested on others models in order to generalize it. It opens various perspectives of research and development for the futur. An interesting work would be to verify the interest of the generation of synthetic brain organoid images on a segmentation task.

Chapter 4

Segmentation

Image segmentation aims at representing a given input image in meaningful form in order to identify a region of interest, measure the volume or size of a particular object and, help at characterizing its content. For a while, segmentation was made by hand, but this task is time consuming for biomedical experts. Numbers of biomedical segmentation algorithms have been created to relieve the experts, among them traditional and, more recently deep learning-based segmentation algorithms.

Nowadays, brain organoid shape extraction is almost not automated. Most of the time, neurobiologists use a software combination or semi-automatic solutions such as ImageJ or Imaris (Brémond Martin et al., 2021b) and see Paragraph 2. A few teams constitute their own solution by creating some macros in these software. One team uses the automatic U-Net a deep learning algorithm to segment ventricular zones (Albanese et al., 2020). However, automated and accurate segmentation has been already well studied in other biomedical fields and some architectures seem more suited to particular datasets (Maier et al., 2019). The solutions for biomedical content segmentation of small datasets has been also poorly described inside the literature and, segmentation could be more accurate with machine learning solutions (Qi et al., 2020). The need of implementing a dedicated shape extraction algorithm for small brain organoid datasets is important for neurobiologists who want to use them as a model.

In this section, we describe the various segmentation methodologies used in biomedical field and for brain organoids, the drawbacks and limitations of these architectures, as well as our strategy and contribution to segment them properly.

4.1 State of the art on segmentation

4.1.1 Classical segmentation

Classical segmentation strategies are constituted by various group of methodologies classified in: threshold, region, clustering, graph, shape and morphological based methods.

A threshold segmentation consists in classifying a pixel into an object or a non object class, depending on a level of a pixel value called threshold. The most known threshold segmentation is called Otsu which is automatic, in other segmentation methods, the threshold is set by hand (Otsu, 1979b). In this particular threshold segmentation method, the intra-class variance is minimized in order to set the optimal threshold which would best separate the two classes. The first step is to characterize the histogram and probabilities of each intensity levels and, initializing variables. Next, mean and variance probability are calculated for all the possible thresholds. Finally, the minimal intra-class variance is retrieved

to be used as the threshold. The various step of a threshold by Otsu are summarized in the pseudo-code described in **??**.

The most known regional segmentation is called the region growing algorithm (Adams et al., 1994). The principle of this segmentation is to initialize seed points and explore their neighborhood. The neighborhood corresponds to a certain pixel intensity, or a particular texture or color, and, is categorized as an object or a non object class. The process is achieved when all the pixels of an image are classified. The pseudo-code in **??** illustrates the region growing algorithm segmentation.

A widespread clustering strategy for segmenting an image is the Kmean (Vora et al., 2013). The initialisation is made by fuzzy or by random partition, then a pixel is compared to a cluster which contains a centroid. Thus the pixel is associated to the cluster with the nearest mean, the centroid is updated, and the process is iterated. The process is achieved when the centroids could not be updated due to the well clusterisation of each pixel. The pseudo-code of this procedure is described in **??** with randomisation.

One graph based method is for instance the graph cut algorithm (X. Chen et al., 2018). This segmentation is based upon Kmean and energy functions.

Two shape based methods are for instance the active contour and level set algorithms (Menet et al., 1990; Meziou et al., 2012). The active contour is compared as the snake game, it uses energy minimizing. The principle is to apply a distortion spline influenced by constraint and forces, that pull the spline towards object contours and internal forces that resist deformation. The pseudo-code is represented in **??**.

The morphological segmentation is represented most of the time by a watershed algorithm (Sharma et al., 2022). This name takes part of the process of filling water inside a recipe, which is compared as geographical valley and, mountains obtains are the shapes of the segmented object. Technically, this method uses local minima of the gradient of the image as markers, which are determined often manually by user.

Other classic segmentations exist and, most of the time optimisations are proposed according to certain applications to improve them.

4.1.2 Deep learning based segmentation

Deep neural networks are used in many applications, among them the segmentation requires often this use. The pre-processing step on data is an important step to its use. Generally, the dataset is divided in three parts: training, validation and test datasets. Then, various architectures are used for segmentation purposes.

Convolutional Neural Networks

Convolutional Neural Networks (CNN) (Z. Li et al., 2021), see Figure 4.1, are constituted by three main types of layers: convolution, pooling and, fully connected layers. The convolution consists in a multiplication of local neighbours of an image pixel with various kernels, in order to produce feature maps. The pooling is a subsampling, it reduces the width and height of the given input and computational requirements. Contrary to a single convolutional layer, each input vector influences every output of the output vector in the fully connected layer to give the result.

Fully convolutional neural network (FCN) consists in transforming the latest fully connected layer, into a fully convolutional one. The inconvenient is the high time computation, and, low resolution of images obtained with fuzzy boundaries, due to propagation through alternate convolution and pooling layers (Milletari et al., 2016).

Figure 4.1: Convolutional Neural Network

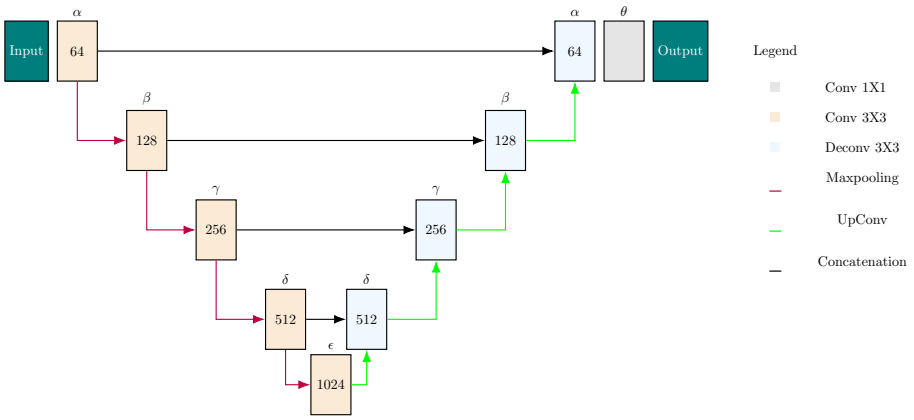


Figure 4.2: U-Net architecture inspired by the original from (Ronneberger et al., 2015a)

Recurrent convolutional networks, works with feed-backs. After the convolutions and the activation, a filtration is produced on the feature map which allow a better accuracy and performance thanks to its update (Qiu et al., 2021).

The 'atrous' convolution drives the resolution at which feature responses are computed. DeepLab is the first network to propose these convolutions, with atrous spatial pyramid pooling (L.-C. Chen et al., 2017). Convolutional layers use filters with multiple sampling rates and, enlarged the field of view which allow to capture image context and objects at multiple scales, without increasing the number of parameters or the computation time. Since then others DeepLab versions have been developed.

U-Net and its derivatives

A common segmentation model based upon encoder-decoder is called U-Net created by Ronneberger et al., 2015a, see Figure 4.2. This network uses the FCN architecture in the downsampling part to extract the features. In this part, each successive block contains two successive 3×3 convolutions, followed by a ReLU activation unit and a max pooling layer. The upsampling part uses deconvolutions to decrease the number of feature maps and obtains an accurate localisation: each bloc upsamples the feature map using a 2×2 up-convolution. The feature map is reduced by a 1×1 convolution to the required number of channels and produces the segmented image. Between the contractive and upsampling part, skip connections (crop and concatenation) are given to avoid pixel information at the edge, which contains less information. It allows to propagate contextual information. An issue found for U-Net is the slow down computation time, due to accumulations in the same scale feature map while skipping connections.

Since its creation a lot of U-Net variants have been proposed, to overcome the limitations of this network.

In 3D U-Net, all the operations are replaced by 3D in order to segment volume in three-dimensional images (Çiçek et al., 2016). In attention U-Net, some gate are used to avoid the propagation of non useful features in the expansive path before concatenation (Oktay et al., 2018). A gate accelerates the procedure and gives an accurate segmentation. In inception U-

Net, various filter sizes are tried on each layer, and, render the network available to segment images with various heterogeneity (Rad et al., 2020). Residual U-Net is a deeper network. With this framework based on the ResNet architecture, the input to a convolutional layer is added (by an element-wise addition) to the output from the following convolutional layer using a skip connection (Kerfoot et al., 2019).

However, these networks are not appropriate, to segment multimodal images (from various microscopic scales). MultiResNet proposes to combine an Inception and Residual U-Net. In this architecture, they replace the convolutional layer sequence by a block constituted from an inception block filter with concatenation, to preserve the features with different context size. Then, these filters are factorized to create successive layers with a residual connection, to preserve the dimensions (Ibtehaz et al., 2020). Also for multimodal images, MRN employs multiple encoders corresponding to different resolutions, that are structurally identical for downsampling, and one single decoder for upsampling(F. Gu et al., 2018). Hook-Net proposes to combine two U-Net structures: one for the target (in a high focus point of view) and one for the context (for low focus) in histopathological images (Van Rijthoven et al., 2021), and, outperforms U-Net and MRN for this task but has not been compared with MultiResNet architecture.

In all these architectures, a limitation persists: the residual skip connection does not allow to preserve from vanishing gradient. The solution is to use Dense U-Net which consists in replacing convolutional layers by DenseNet architecture in deep networks. Compared with ResNet, each layer receives the features from all the previous layers, and, the maps are combined by concatenation into tensors(Dong et al., 2019).In one hundred layer Tiramisu network, they use DenseNet to work as a FCN by upsampling path to recover the full input resolution. The upsampling path results in a number of feature maps with very high resolution prior to the softmax layer (Jégou et al., 2017). U-Net++ avoids the loss of semantic information between the upsampling and downsampling part by transforming each level by a dense block. For this aim, skip connections are used between mirrored layers (Z. Zhou et al., 2018). Unet3+ variation uses full-scale skip connections and deep supervisions (Huang et al., 2020).

Attention based models

CNN networks have the limitation to produce segmentation mask without accurate boundaries. This particularity come from the consecutive pooling and convolutional information, which does not allow the extraction of a right context and necessary features. This becomes a challenge in biomedical image datasets, which are composed by intra-class variations, inter-class non distinct features and, noise. In order to improve the accuracy of the segmentation, some models propose to use attention mechanisms for semantic segmentation.

Attention Gated Networks (AG) use a 3D-Unet, in which two attention modules have been added in the contractive path. The aim of these modules is to suppress irrelevant regions in an input image and, highlight salient features useful for the segmentation (Schlemper et al., 2019). CE-Net is similar to U-Net architecture, with two additional blocks at the U-shape basement: a dense atrous convolution and, a residual multi-kernel pooling to preserve more high-level features and more spatial information (Z. Gu et al., 2019). For real time images, Bilateral Segmentation Network (BiSeNet) has been propose. To preserve the spatial information and keep high-resolution features, they create a spatial path with a small stride, and conserve the context by increasing the speed of the downsampling path. They use a feature fusion module to combine these two paths (Yu et al., 2018). In BA-Net (for boundary aware), they add some modules to retrieve the meaningful contextual information

to drive the decoder part. Between the encoder and decoder architecture, they use an atrous spatial pyramide pooling module, and between each mirrored layer they successively process a pyramide edge extraction, a mini multitask learning module and a cross feature fusion (R. Wang et al., 2022). BA-Unet gives better results than others architectures such as U-Net or CE-Net, multiresUnet or DeepLabV3. In the same way, SA-Net proposes scale attention modules between the encoder and decoder mirrored layer before the concatenation(J. Hu et al., 2021). Sau-Net uses for textural information a U-Net architecture. However, the encoder is replaced with dense blocks from DenseNet-121 similar to the Tiramisu and the decoder block used is a dual attention decoder (Sun et al., 2020). Instead of using a U-Net architecture, the Multi-scale dual guided attention module applies a ResNet architecture, with four focus attention modules for each dense local features(Sinha et al., 2020).

Instead of using attention modules, some architectures use directly the shape, or topological information to guide the segmentation. In BB-Unet the encoder layers are fed with the image for contextual feature extraction, similarly to U-Net. Bounding filter, a binary map which indicates the position of the object, is fed independently to the additional BB-ConV layer for shape and location feature extraction (El Jurdi et al., 2020). Others prefer to update the loss functions with shape or topological information. Instead or complementary to fitting loss (Dice or cross entropy based), some authors apply regularisation prior loss based upon size, shape, inter-region or topology of contents inside the image to improve the segmentation (El Jurdi et al., 2021).

Light architectures

All these architectures presented previously, are designed for big datasets, contrary to the original U-Net and, and most of the time requires a lot of computation time. Moreover, the attention gate limitations have a high complexity and, are relatively less interpretable. To perform a segmentation on very small datasets, reducing the U-Net architecture seems to be a solution which start to appear.

In U-Net ++ the structure for instance is lightened by giving dense nested strip pathways (Z. Zhou et al., 2018). The stripped-down U-Net (SD-Unet) is constituted by depth-wise separable convolutions with a weighted standardisation, in order to avoid the slow performance and high computing time. This model contains less parameters to learn, a smaller size architecture and lower computation time than U-Net (Gadosey et al., 2020). Contrary to SD-Unet, SmaAT-Unet proposes an attention module added to the use of depth-wise separable convolutions (Trebing et al., 2021). In the CNL-Unet structure, a pre-trained encoder is proposed to allow the segmentation of small datasets. By using the transfer learning properties of the encoder, the modified skip connections to reduce the gaps between mirrored layers and, the classifier and localiser module CNL, they learn efficiently the small amount of data(Shuvo et al., 2021). Another update of the encoder architecture consists in using transformers after convolutional layers, to lighten the architecture such as in LeViT-Unet(Xu et al., 2021).

However, these procedures modify the simple aspect of the U-Net architecture, the encoder and decoder structure. Recently a Unet-Mini architecture has been created to segment small datasets, which reduce the number of parameters drastically only by updating three elements (Jirik et al., 2020): they use element-wise skip connection (addition instead of a concatenation), reduce the number of layers and the kernel size.

Another strategy consists in transforming the learning session for a small dataset, particularly when they are partially or not annotated. Some solutions work also with weakly supervised (or semi-supervised such as few or one shot learning), and unsupervised (also

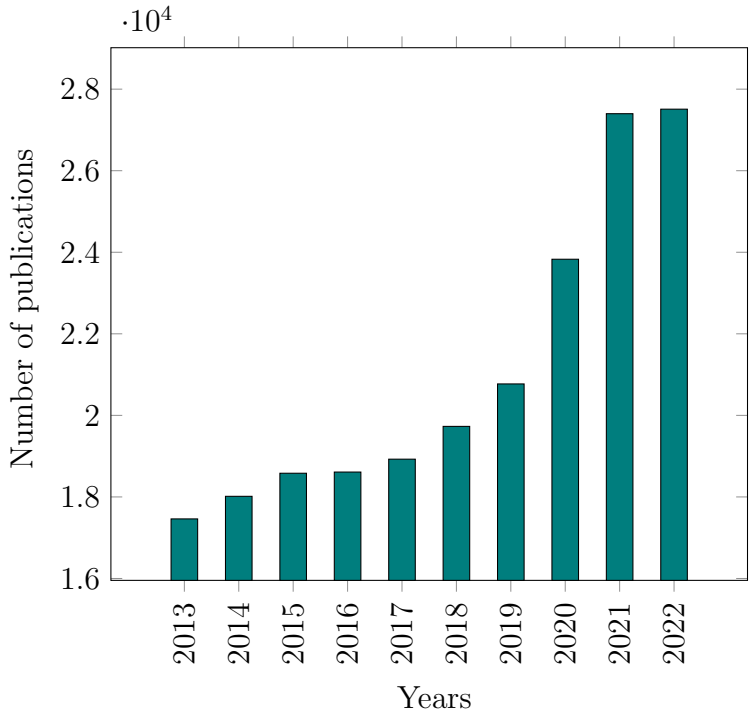


Figure 4.3: Number of publications on segmentation per year in the biomedical field since 2000.

called zero-shot learning or self training) (Pastore et al., 2021).

Recent segmentation architectures

GAN based Other particular U-Nets exist, such as adversarial U-Net. This adversarial structure uses a U-Net inside the generative part of a GAN, which takes in input an image, and produces its segmentation. The discriminator has the aim to decipher synthetic segmentations from the given ground truth (Mirza et al., 2014).A cGAN, using a Markovian-GAN inside the discriminator part, aims at segmenting better the images by calculating errors and optimizing them (Rezaei et al., 2018). The aim is to outperforms the GT segmentation applied on certain images.

Multi-Adaptative U-Net and pipelines An existing pipeline containing an automatic segmentation (after a manual split of the dataset) is the InstantDL software. The segmentation process is based upon a Mask-RCNN with a ResNet50 architecture. However, this architecture has been validated only on some cells (Waibel et al., 2021) and for a kind of microscopic acquisition. Scalable architectures are necessary to be adaptative to all microscopic acquisitions or biomedical datasets. (Isensee, Jaeger, Kohl, Petersen, and Maier-Hein, 2021), proposes nn-Unet, an adaptative U-Net architecture, which automatically optimizes itself according to the given database.

4.1.3 Segmentation in the biomedical field

Biomedical image segmentation has increased over the last decade see Figure 4.3. Segmentation is one of the main operation realized on biomedical images, and its automation seems necessary to save time for biomedical experts.

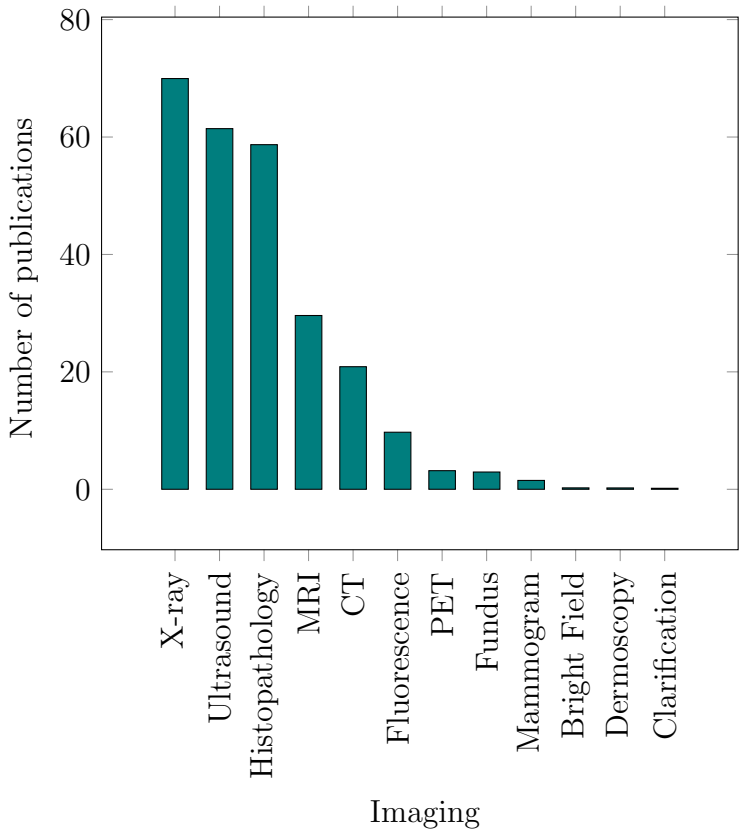


Figure 4.4: Publication on segmentation per imaging applications. We only count the number of publications in the main domain and show it in percentages (on 109307 inputs).

Among the imaging the most often segmented, we found X-ray, ultrasound and histopathological images with more than 50% of the use see Figure 4.4.

And the domain of application concerns most of the time the segmentation of cell components (more than 50% of use) tumor, heart, brain and eyes. The segmentation of biomedical cultures concerns only 11% of the cases, see Figure 4.5.

The most use architecture remains U-Net and its derivatives (almost the half of the use) see Figure 4.6. The second are CNN based architectures. With less than 10 %, the FCN, atrous and light architectures are only emerging architectures for segmentation purpose.

4.1.4 Segmentation of brain organoid culture image: our positioning and strategy

The segmentation of brain organoid images has been made in majority with semi-automatic tools see part 2. Only one team segmented brain organoid ventricles with U-Net trained with classic data augmentation strategies (Albanese et al., 2020). No segmentation tool has been compared and developed to extract accurately the shape of brain organoid images to characterize their development. The constraint is to accurately and automatically segment a small dataset.

The summary of segmentation optimisation strategies are given in Figure 4.7. The developed strategy in this work is to test various classical segmentation methodologies and deep learning based methodologies. The second part proposed is to validate a data augmentation strategy to better segment with deep learning tools. Then the last part of this work consists in reducing the size of the architecture to be appropriate to small datasets.

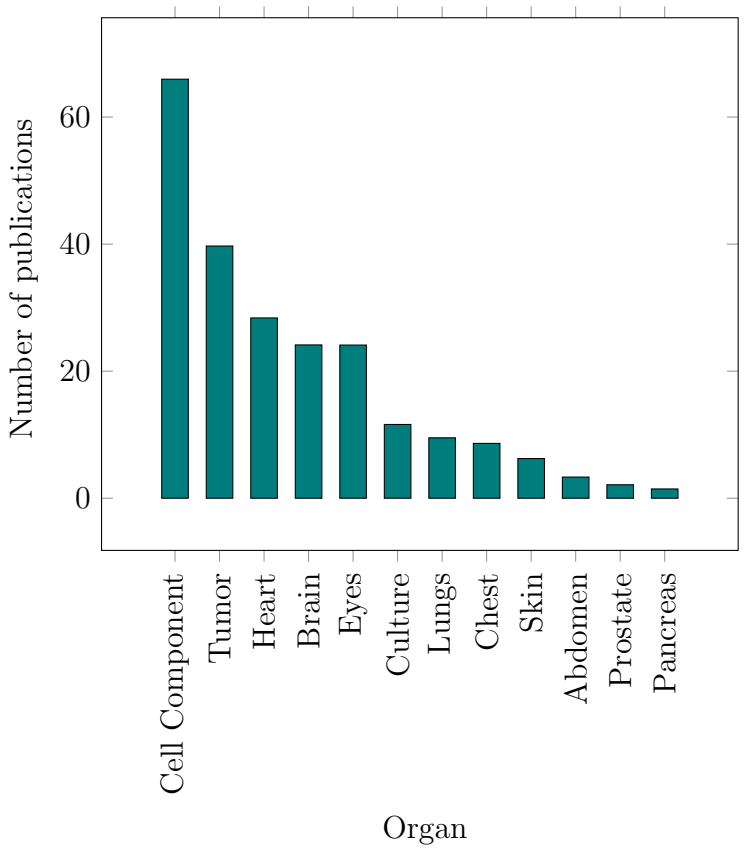


Figure 4.5: Publication on segmentation per organ. We only count the number of publications in the main domain and show it in percentages (on 109307 inputs).

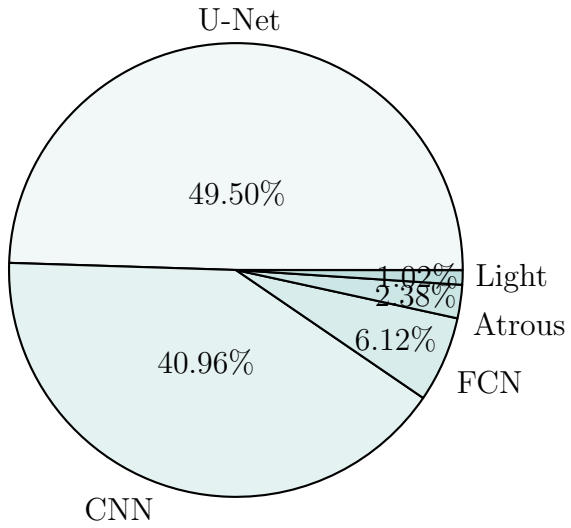


Figure 4.6: Architecture segmentations most use in the Biomedical field. We only count the number of publications in the main domain and show it in percentages (on 125461 inputs since 2017).

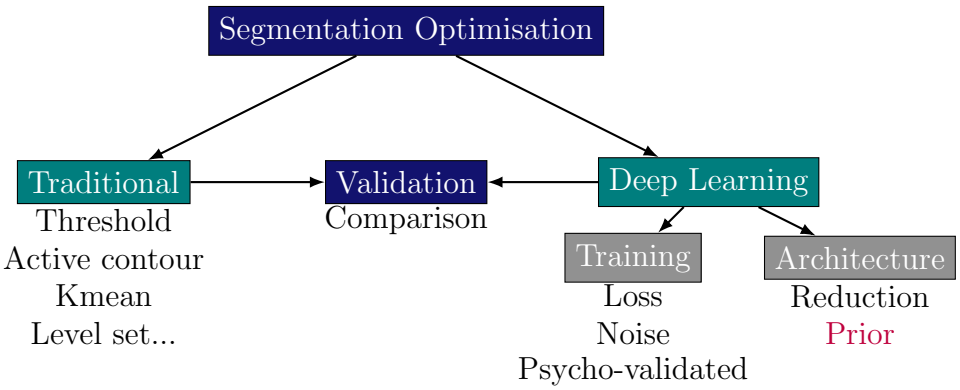


Figure 4.7: Global overview of Segmentation optimisation strategies we choose to develop.

4.2 Validation of GAN optimisations by a deep learning-based segmentation

This section is in part published in Brémond Martin et al., 2021a in VISAPP conference in 2022. Noise optimized and psychovisually validated images used during the training are not published parts.

Segmentation allows the extraction of an image content from its background. Various segmentation procedures exists see part 4.1. In the biomedical field, U-Net architecture is the most used see Figure 4.6 but requires a certain amount of images (Ronneberger et al., 2015a).

However, in the case of tiny dataset, such as for brain organoid cultures, which kind of data augmentation could contribute to accurate the segmentation when they are used during the training step?

To determine the effect of data augmentation strategies on a segmentation task, we consider several training scenarios using various synthetic datasets:

- The effect of synthetic images resulting from various loss optimisation of AAE (see Section 3.3)
- The effect of synthetic images validated by none to 8 biomedical experts (see Section 3.5)
- The effect of synthetic images resulting from various noise optimisation of AAE (see Section 3.4)

4.2.1 Methods

Datasets

We aim to segment images from (Gomez-Giro et al., 2019) which is composed of 40 images in open access. Twenty pathological and twenty healthy brain organoids were numerized with a bright-field microscope over 3 days. The grayscale images are 1088×1388 pixels. However, to compare several networks within a reasonable time, the input images are cropped and resized to 250×250 pixels, maintaining the original proportions.

The first "classical" training dataset for comparisons purpose is composed by 80 classical augmentations involving flip-flops, rotations, whitenings, or crops. The "loss" training

datasets are constituted by the generated images from each AAE loss optimisation see Section 3.3: BCE, BCE+L1, LS, Poisson, Wass. P.Wass. The "noise" training datasets are composed by the generated images from each AAE noise injection 3.4: Gaussian, Poisson, Salt, Speckle, Poisson Gaussian, Poisson Salt, Poisson Speckle. The "psychovisual" training dataset is composed by the psychovalidated images from AAE loss optimisation see Section 3.5: 0 expert validations, 1 expert validation, 2 expert validation, until the 8 expert validation. While the "loss" and "noise" training are compared with the "classical" training, in the "psychovisual" training case, if an image is validated by 0, 1, 2, 4, 6 or 8 experts, it replaces classical augmented images among the 80 to conserve the number of images inside the training dataset for each of the mentioned group.

Ground truth as been generated with the ITK-SNAP software see Paragraph 2.7 and, resources are the same as written in previous sections and Paragraph 2.7.

U-Net

To segment these images we use the U-Net architecture, widely used in the biomedical field (Ronneberger et al., 2015a). Segmentation allows the extraction of an image content from its background. Various segmentation procedures exist but we have chosen U-Net for its advantages to work well for small training sets with data augmentation strategies, and to have already been used for the ventricle segmentation of cleared brain organoids (Albanese et al., 2020).

Training

To make the performance evaluation more robust, a "leave-one-out" strategy is used, resulting in 40 training sessions (numbers of images in our original dataset). For the "loss" and "noise" case each training is performed on 79 images. We stop the training at 1000 epochs with an average time of training of more than 1 hour for each leave one out loop (7 cases of augmentations $\times$ 40 images = 280 hours almost for the total training step for noise and 240 hours for loss). For the "psychovisual" case each training is performed on 119 images. We stop the training at 1000 epochs with an average time of training of more than 1 hour for each leave-one-out loop (6 cases of augmentations $\times$ 40 images = 240 hours almost for the total training step).

The summary of the leave-one-out strategy for every tested case is summarised in Figure 4.8.

Comparison of segmentations

To compare ground truth cerebral organoid content segmentation (GT) and U-Net (u) ones in various conditions, mean Dice scores are calculated as:

$$Dice_{(GT,u)} = \frac{2|GT \cap u|}{|GT| + |u|} \tag{4.1}$$

Thanks to the TP, FP, TN and FN we could calculate the Accuracy, the Specificity, the Sensitivity, and the F1-score. The Accuracy is the ratio of true on the positives labels:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{4.2}$$

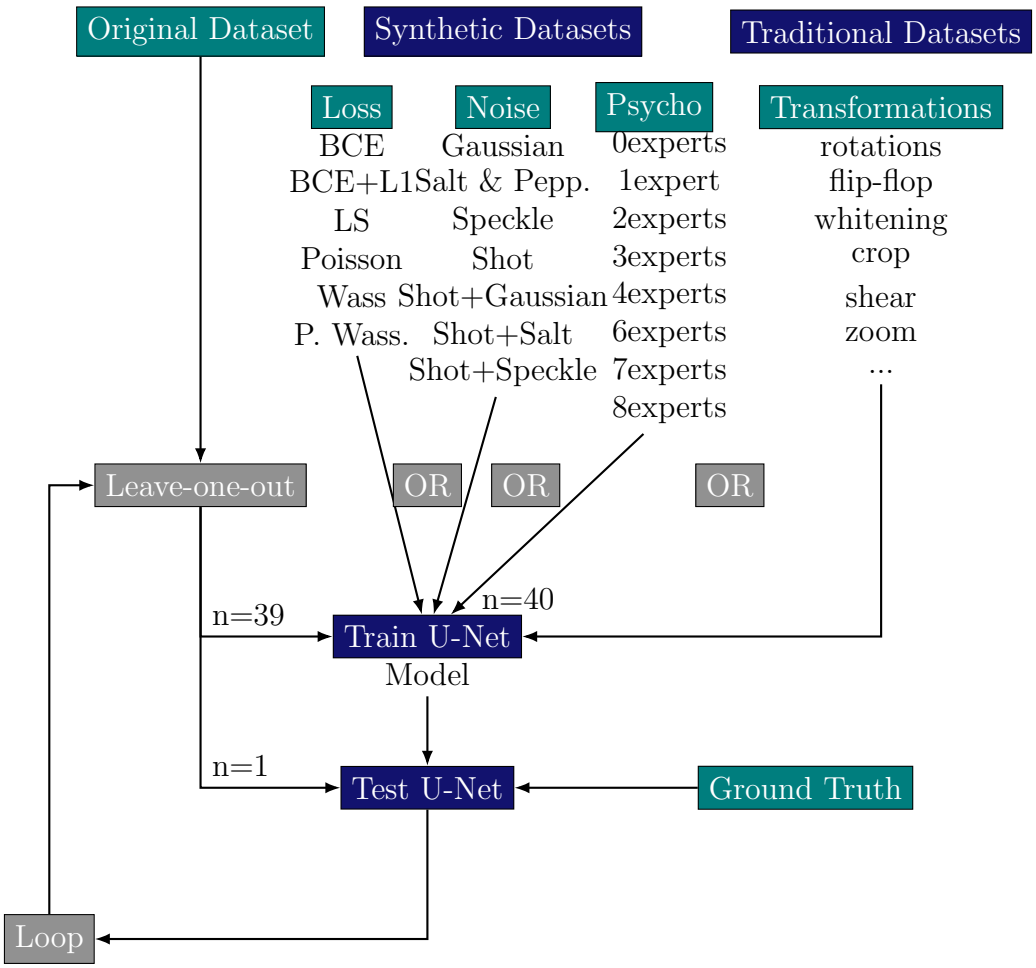


Figure 4.8: Experimental scheme of the leave-one-out strategy to test the effect of various data augmentation on the segmentation quality.

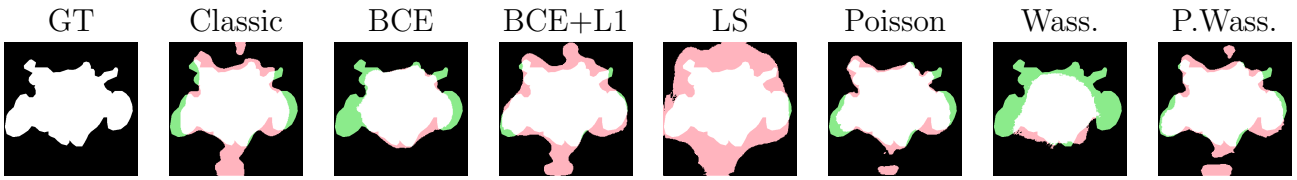


Table 4.1: Sample of each segmentation with a particular loop optimisation training.

The Sensitivity is the ratio between how much were correctly identified as positive to how much were actually positive:

$$Sensitivity = \frac{TP}{TP + FN} \tag{4.3}$$

The Specificity is the ratio between how much were correctly identified as negative to how much were actually negatives:

$$Specificity = \frac{TN}{TN + FP} \tag{4.4}$$

The Precision is the ratio between how much were correctly identified as positives to how much were actually labeled as positives:

$$Precision = \frac{TP}{TP + FP} \tag{4.5}$$

The F1-Score allow to summarize the precision and the recall (Sensitivity) in an unique metric:

$$F1 - Score = 2 * \frac{Precision * Sensitivity}{Precision + Sensitivity} \tag{4.6}$$

Visualisation

To highlight real/false positive/negative segmentation we create a superimposed image composed by the ground truth and a sample of each segmentation resulting from the various trainings. We update the pixels values in lightpink the FP cerebral organoid segmentations and, in lightgreen the FN.

4.2.2 Results

Influence of loss optimised synthetic images in training step of a segmentation task

To illustrate the influence of generated images by an optimised AAE against classical data augmentation, we tackle a segmentation task in a leave-one-out strategy (n=79 for training and n=1 for testing). We choose the classic U-Net architecture and we consider the different losses to compare the segmentation performance for each data augmentation.

Psychovisually, samples in Figure 4.1 are the best segmented with a training involving images resulting from the AAE Perceptual Wasserstein optimisation. They show the less false positive and negative segmentations compared to others AAE optimisations and to classical data augmentation.

Quantitatively the mean Dice index highlight the segmentation performance. Results are summarized in Table 4.2. Mean Dice index is higher for segmented images with Perceptual

Table 4.2: Sample Cerebral Organoid image (left) with ground truth (GT) segmentation (our baseline), compared to Classical and AAE-based data augmentation, and the corresponding Mean Dice index and Standard Deviation (SD). Pixels are colored according to the following legend: Black and white represent respectively true negatives and true positives while magenta highlights false positives and cyan a false negatives. The best Dice index is displayed in bold.

	Classic	BCE	BCE + L1	LS	Poisson	Wass.	Per. + Wass.
Mean Dice	0.87	0.85	0.87	0.86	0.87	0.88	**0.90**
SD	**0.03**	0.05	0.05	0.14	0.09	0.05	0.04
Accuracy	0.77	**0.84**	0.82	0.77	0.75	0.83	**0.84**
Sensitivity	0.95	0.95	0.90	0.92	0.90	**0.97**	0.90
Specificity	0.92	0.51	0.75	0.93	0.75	0.37	**0.96**
F1-score	**0.88**	0.61	0.73	0.87	0.75	0.57	**0.88**

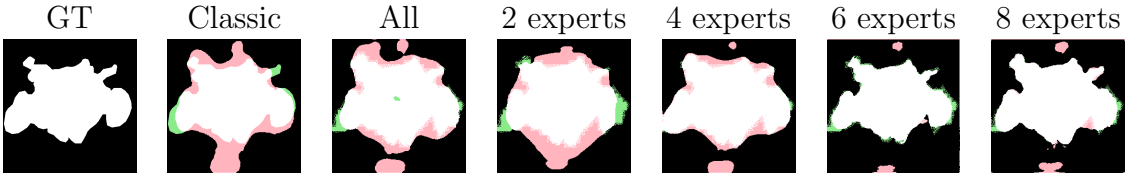


Table 4.3: Sample of each segmentation with various data augmentation strategies validated by a certain number of experts.

Wasserstein augmentation, in accordance with the selected visual illustration. The standard error, accuracy, and f1-score are the best for classic and near followed by the Perceptual Wasserstein group.

In conclusion, images generated from Perceptual Wasserstein AAE allow a more accurate segmentation than other AAE loss, in accordance with previous results on quality. The influence of the Perceptual loss combined with Wasserstein distance, such as a data attachment term based on the difference of generated and images features maps, improve their sharpness and textural information, making it a viable strategy for data augmentation in this context.

Influence of psychovisually-validated images

Then the second task is to verify the interest of using synthetic images in a training segmentation task which have been validated by 0, 1...or 8 biological experts.

Qualitative results To observe the quality of the segmentation, we observe a ground truth segmentation performed with the ITK-SNAP software, and automatic segmentations performed with a U-Net architecture with various data augmentation strategies see Figure 4.3. In the 0 expert group: the synthetic images used during the training are selected by none of the experts. In the others training images are previously selected by 2, 4, 6 and 8 experts. We observe less false positive (in pink) and false negative (in green) region, if the segmentation is realized with a training step constituted with a dataset containing images validated by 8 experts or 6 experts. If the synthetic images are selected by 6 or more experts, we could observe almost no errors.

	GT	classical	0	2	4	6	8
Nbr of synthetic images	-	0	33	41	22	16	14
Dice	1.0	0.80	0.62	0.59	0.61	0.67	**0,82**
Accuracy	1.0	0.77	0.63	0.61	0.68	0.62	**0.93**
Sensitivity	1.0	0.94	0.96	**0.97**	0.87	0.83	0.91
Specificity	1.0	0.92	0.50	0.48	0.60	0.56	**0.94**
F1-score	1.0	0.84	0.64	0.61	0.64	0.59	**0.87**

Table 4.4: Comparison of segmentation according to the number of expert validations. All metric results are the mean calculated on a group of segmented images: GT, training with classical data augmentation, non validated images, 2, 4, 6, or 8 experts validating some images.

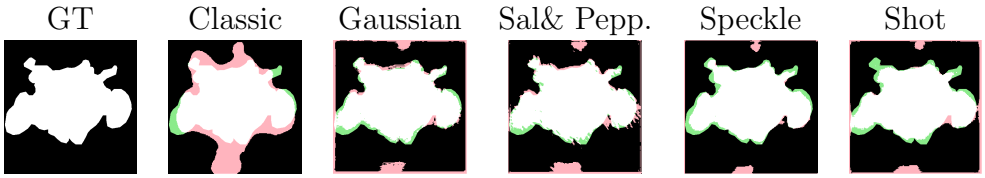


Table 4.5: Sample of each segmentation with various noise optimised synthetic images used during the training.

Quantitative results Some parameters calculations between the ground truth segmentation and classical or automatic with various data augmentation strategies are summarized in Table 4.4. The segmentation is better with 8 experts selected images for the training (with higher levels of Dice, Accuracy, Sensibility and F1-score). The sensitivity is higher for segmentations with 0 or 2 experts groups training, however the Specificity is lower than the ground truth group and the segmentation is worst than the classical one.

Influence of noise optimised synthetic images

In this part, we study the influence of various noise optimisations used during the generation of synthetic images used during the training step of the segmentation. We compare four kind of noise optimisation : Gaussian, Salt & Pepper, Speckle and Shot.

Qualitative results Similarly as previously, we obtain the segmentation with U-Net from various training with noise optimized AAGAN synthesized images see Figure 4.5. These segmentations are compared with a segmentation from a classical training, and with the ground truth. Only a few external shape of neuroepithelial regions are not well segmented whatever the noise optimisation. With shot noise two neuroepithelial regions seems not well segmented for this image. However, false positive segmentation are bigger in Gaussian and Salt & Pepper noise optimisation training than for images segmented with Speckle or Shot optimized images for an AAE. These False positive regions represent cells conglomerates and not the brain culture.

Quantitative Results The quantification of the effect of various GAN noise optimisation used during the training of brain organoid segmentation is summarized in Table 4.6.

	classical	Gaussian	Salt & Pepp.	Speckle	Shot
Nbr of synthetic images	0	40	40	40	40
Dice	**0.83**	0.76	0.77	0.77	0.74
Accuracy	0.77	0.80	0.80	**0.81**	0.73
Sensitivity	**0.95**	0.79	0.80	0.79	0.87
Specificity	0.92	0.95	0.96	**0.97**	0.84
F1-score	0.88	0.80	0.81	**0.82**	0.78

Table 4.6: Comparison of segmentation according to noise optimisation.

The best Dice score is obtained with Salt and Pepper or Speckle training, while the best Accuracy is obtained with the shot noise. The optimum F1-score, Specificity and Accuracy are obtained with the Speckle. In summary none of the noise injected produce better results than a segmentation produced with a classical data augmentation strategy except for the Speckle noise.

4.2.3 Discussion

Approaching a segmentation experiment, images generated with a Perceptual Wasserstein loss could bring a better precision to a segmentation task. Other losses may be interesting for different tasks. The psychovisual validation of these trained images highlight accurate segmentation with more experts validation. The 8 expert validation of GAN images used during the training allow to improve all the segmentation score compared to a classical transformation used during the training. Speckle noise optimized augmented images used during the training of the segmentation improve the segmentation quality.

Attempting to distinguish the contribution of each loss optimisation, this strategy can potentially bring better pixel-wise precision for segmentation tasks. Shown here as a proof of concept, using a U-Net architecture, we demonstrate that the Perceptual Wasserstein loss can fruitfully enrich the original dataset. This may also show a kind of regularisation achieved by the Perceptual loss leading to a good variability of generated data without being too generic. The contribution of others loss could not been highlighted in this task. Nevertheless, segmentation could be even more appropriate with algorithms suited for small datasets or by increasing the number of training iterations. In future parts we extract morphological parameters, such as areas, perimeters or higher-order statistics needed for the growth follow-up of cerebral organoid cultures on segmented images, see Section 5. In this work, we only segmented organoid vs. non organoid regions. Another objective could be reproducing the same work differentiating the peripheral and the core zones of the cerebral organoid in these images.

The more experts validate the dataset, the best is the segmentation quality. This suggests that more experts are available to select images, and strenghen the 'naturality' of the synthetic dataset used during the training, it could improve the accuracy of the segmentation results. However, even if the psychovalidation of certain synthetic images allow us to improve the segmentation, this method is still subjective. It could include biases of the experts about the model, its configuration, and the project objective. It requires knowledge of which images are considered as natural and which ones are not for the target domain, so the number of experts available in the field diminished while our task required more experts. It is limited to the number of images that can be reviewed in a reasonable time (Booij et al.,

2019).

Validation by six experts is the minimum to improve the segmentation, but quantitative analysis show us, 8 experts validation is the minimum due to the equilibrium state between the Specificity and the sensibility. The performance of human judges is not fixed and can improve over time, other articles choose a validation by 15 experts for instance which is not possible in our biomedical context which requires experts in the field (Salimans et al., 2016; Denton et al., 2015).

Segmented images with a training made upon Speckle optimized augmented image are the most accurate according to the metrics we choose. Speckle is a granular textural noise which does not takes part of bright-field image acquisition which makes its contribution to the improvement of the segmentation process more difficult to interpret. Indeed, the natural noise contained in bright-field images (shot noise), which render the best data augmentation strategies can not render similar improvement of the segmentation (Gilroy, 2019; Brémond Martin et al., 2022). Thus the explanation of this result should be in another domain. Speckle is present in other biomedical acquisition modalities such as radar tomography or ultrasound images and has to be eliminated in order to highlight relevant biological structures (Liba et al., 2017; P. Singh et al., 2021; Karaoğlu et al., 2022). Edge extraction on images containing this noise is particularly difficult due to the direction of contents edges while, the speckle is randomly distributed. The segmentation render best quantitative results with a Gabor-based anisotropic diffusion pre-processing denoising system on ultrasound images for instance (Haobo Chen et al., 2021). However, in our case, the noise is injected in the feature space of images during the generative process of the GAN and not on the image, see Section 3.4 and (Brémond Martin et al., 2022). Considering our results, an explanation could be the power of generalisation on a segmentation task trained with data augmentation with GAN noise optimized image, is the best with a speckle noise injection. Indeed, speckle noise is the only noise which follows a gamma distribution (Arulpandy et al., 2020), this property could explain its particularity to generate generic images that will allow a better training. The other noise optimisations could maybe improve the result of other tasks which has to be studied.

In future studies, a metric based method or a qualitative based approach which could replace the validation of 8 or more experts could help to automate or, win some time during the segmentation procedure, could be useful. It could be interesting to segment others images from others GAN variations such as one from noise optimisation (Brémond Martin et al., 2022). In the training step, we only use psychovalidated images without verifying the kind of optimisation. In another part, we only use a specific GAN variation during the training. As in 8 expert validation images contained in the training set are in majority from the P.Wass or Wass loss optimisation considered as the best optimisation to produce the most realistic images and to accurate the segmentation, it strengthen this result 3.3 and 3.5.3. To be fair, it could be interesting to test a similar 8 expert validation set with others optimisations in the same number. The 0 to 4 expert dataset contains almost all BCE, BCE+l1, LS psychovalidated images, however, these images are specifically blur with white imprint and others issues see Section 3.3. Maybe, the blur and artefacts of these generations provoke the huge amount of FP around the segmented shape of the organoids for these groups due to the increase of incertitude.

Then, other biomedical datasets could be segmented with this methodology to estimate the range of experts needed at minimum and maximum for validating GAN images.

In this part we aim at studying the effect of various GAN data augmentation strategies on the segmentation quality. We bring new insights on which kind of optimisation could benefit to the best learning process on a segmentation procedure for brain organoid bright-

field images. We verify the Perceptual Wasserstein loss optimisation on AAE is the best loss optimisation to give the best segmentation results. We verify the most biological expert validate a synthetic image used during the training step, the better is the segmentation. Finally we highlight the speckle noise injection during the generative process of synthetic images used during the training of segmentation give the best accurate segmentations.

4.3 Mu-Net a light architecture for small dataset segmentation

This article will be soon submitted for publication consideration. It has been presented in a special session of the GDR-ISIS called "Towards pragmatic learning in a context of limited labeled visual data" at this link .

In the previous Section 4.2 we observe the effect of various data augmentation strategies and optimisation on a deep learning segmentation based upon U-Net. We observe a huge training time and a certain amount of data are required. Recently some reduced architectures for segmenting small datasets have appeared see Paragraph 4.1.2.

In this part, we will study the effect of various data augmentation strategies on the reduction of the segmentation architecture.

4.3.1 Methods

Resources

Original 40 images are from the dataset (Gomez-Giro et al., 2019). Twenty pathological and physiological cultures are captured with a bright-field microscope over 3 days. The grayscale images are 1088×1388 pixels. However, to execute our script in a reasonable time, the input images are cropped and resized to 250×250 pixels, maintaining the original proportions. Transform groups are original images transformed by linear data augmentation strategies chosen randomly (flip flop, rotation and whitening) in order to compare AAE with a classical approach. We created 40 transformed images. Generated groups are constituted by: 240 images in 6 loss group optimisation of an Adversarial Auto-encoder (AAE) (Brémond Martin et al., 2021a) and 160 in 4 noise groups optimized in AAE from (Brémond Martin et al., 2022). The each 6 loss groups are composed by 40 images generated with a binary cross entropy (BCE), a BCE with a normalisation (BCE +L1), a least square (LS), a poisson (POISSON), a wassertein (Wass.), and a perceptual wasserstein loss (P.Wass.) whereas each of the 4 noise groups are composed by 40 images generated with a gaussian, a speckle, a salt and pepper and a shot noise injection. Bright-field image acquisitions are characterized with shot noise (Gilroy, 2019). Resources are listed in paragraph 2.7.

Segmentation architectures

We choose to realize an architecture based on the U-Net architecture as we want to reduce the time execution but conserving its precision.

U-Net Originally created in 2015 by Ronneberger et al., 2015a, U-Net creates detailed segmentation maps that are particularly well used in the biomedical domain thanks to its two part (encoder and decoder) u-shape architecture which spread the contextual information of

Dataset	Effectives	Sample
1) Original	40	
2) Transformed	40	
3) Loss optimisation AAE	240	
4) Noise injection AAE	160	

Table 4.7: Sample view for each dataset. 1) is from (Gomez-Giro et al., 2019), 3) and 4) samples are respectfully from (Brémond Martin et al., 2021a; Brémond Martin et al., 2022). Only samples from 2) are purposely transformed from the original dataset 1).

images inside the network. The encoder part use a CNN architecture: two 3×3 convolutions followed by a ReLU activation unit and a max-pooling layer repeated several times. The difference between U-Net and others previous CNN is the second part called decoder part where at each stage, the feature map is up-sampled using 2×2 up-convolution. In the encoder part, the feature map from the corresponding layer is cropped and concatenated onto the up-sampling map which is a long skip connection. The ultimate stage consists in a 1×1 convolution to reduce the feature map and produce the segmented image.

UNet-Mini Segmenting a small dataset is an actual challenge in the domain, in 2020, (Jirik et al., 2020) realize a new architecture based on U-Net with a few changes: they update the skip connections with addition as an element-wise product instead of concatenations; they reduce the layer number at four to reduce over-fitting situations and also the number of kernels used in each convolution (ks $= 3 \times 3$ and stride s $= 1$ and they use a soft-max activation task for their non binary). This last update reduce the number of hyper-parameters to learn (and the execution time) at 128K instead of 17M with the original implementation.

Mu-Net Our reduced U-Net architecture contains as similarly with the UNet-Mini 4 encoder-decoder layers with 16, 32, 64 and 64 number of kernels as our objective is to reduce the execution time. However we conserve the original skip connections. Indeed, we choose the concatenation for the feature re-usability and compaction. We also update size strides (the amount of steps between filter applications to the input image) when we down-sample in convolution layers while stride stay at 1 whatever the convolution layer in UNet-Mini. We used an Adam optimizer, with a binary cross entropy loss and a sigmoid activation (because of the binary segmentation purpose). Summarized model is recapitulated on the subsequent Table 4.8

To observe the differences between the three models see Table 4.9.

Training

To segment original images, we train the chosen architectures in a leave one out strategy: all the original and an augmentation group are used during the training stage except one original image which is segmented during the test phase thanks to the model generated. We process 40 leave one out training models * 6 loss groups + 40 leave one out training models * 4 noise groups + 40 leave one out training models with a classical augmentation dataset constituted by flip flop, rotation, or whitening of original images: which results in

name	filter	parameters
Conv2D	16X16	[k=3,s=1,a=relu]
Conv2D	16X16	[k=3,s=2,a=relu]
BatchNorm		
Conv2D	32X32	[k=3,s=1,a=relu]
Conv2D	32X32	[k=3,s=2,a=relu]
BatchNorm		
Conv2D	64X64	[k=3,s=1,a=relu]
Conv2D	64X64	[k=3,s=2,a=relu]
BatchNorm		
Conv2D	64X64	[k=3,s=1,a=relu]
Conv2D	64X64	[k=3,s=2,a=relu]
BatchNorm		
Dropout		[d=0.5]
Maxpooling		[p=2,2]
DConv2D	64X64	[k=3,s=1,a=relu]
DConv2D	64X64	[k=3,s=2,a=relu]
Upsampling	64X64	[k=2,a=relu,sz=2,2]
Concatenation		
DConv2D	64X64	[k=3,a=relu]
DConv2D	64X64	[k=3,a=relu]
Upsampling	32X32	[k=2,a=relu,sz=2,2]
Concatenation		
DConv2D	32X32	[k=3,a=relu]
DConv2D	32X32	[k=3,a=relu]
Upsampling	16X16	[k=2,a=relu,sz=2,2]
Concatenation		
DConv2D	16X16	[k=3,a=relu]
DConv2D	16X16	[k=3,a=relu]
Conv2D	2X2	[k=3,a=relu]
Conv2D	1X1	[k=1,a=sigmoid]

Table 4.8: Mu-Net Model. k corresponds to the kernel size, s to the stride, a to activation, d to dropout, p to pool size, sz to upsampling size

Update Name	Unet	UNet-Mini	Mu-Net
Layer Number	5	4	4
Filter	[64, 128, 256, 512, 1024]	[16,32,64,64]	[16,32,64,64]
Stride	default(1)	default(1)	1 or 2
Activation	Sigmoid	Softmax	Sigmoid
Skip connections	Concatenation	Addition	Concatenation

Table 4.9: Summary of differences

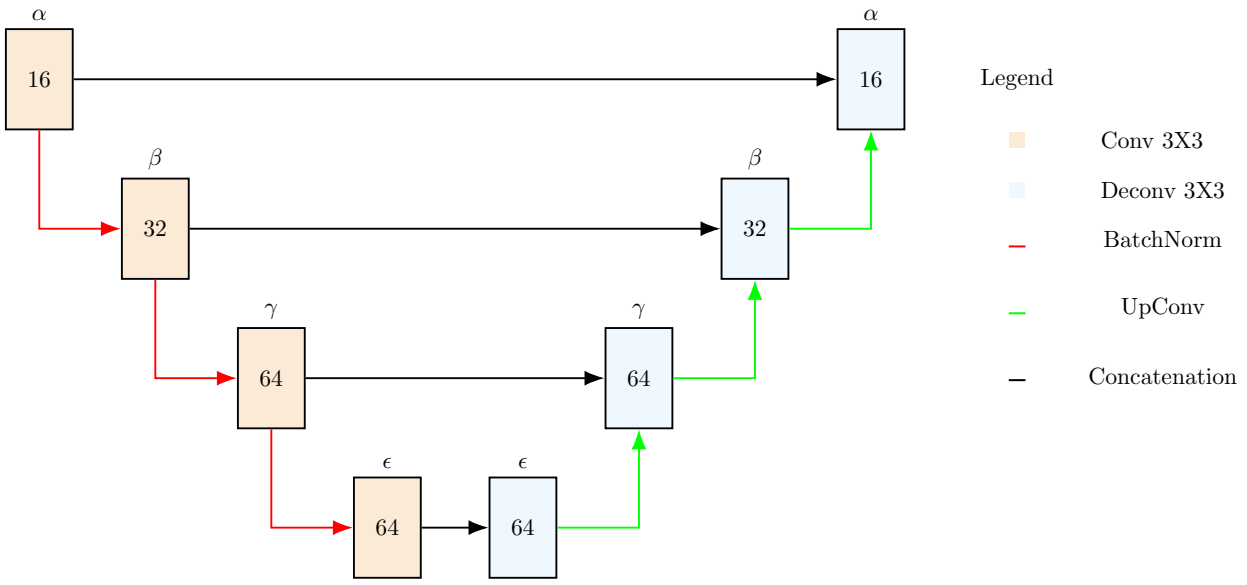


Figure 4.9: Mu-Net Architecture. Arrows are representing operations and blocs are representing convolutional or deconvolutional layers constituted by two 3×3 convolutions (except for the latest blue bloc which is constituted by two 3×3 and one 1×1 deconvolution).

440 training launch for each model. We retrieve the time of training of models. To stop the training and avoid over-fitting, we retrieve the models before the loss value hits a 'plateau'. Original, transformed and generated images are manually segmented with the ITK-SNAP software to obtain a ground truth purpose (Yushkevich et al., 2006b).

Comparison of segmentations

To compare all the strategies we compare median scores between groups with all the metrics already described in Paragraph 4.2.1. We retrieve execution time for each image segmented by a network, and for an entire leave-one-out process (the procedure of the leave-one-out strategy is summarized in Figure 4.10).

The visualisation of segmented images is similar to the one described in Paragraph 4.2.1.

Ablation study

In order to validate the benefit and inconvenience of our methodology, we perform an ablation study between Unet, UNet-Mini and our architecture in term of scores. It results we tested N models with only:

- Unet,
- "Layer",
- "filter",
- "kernel",
- "Ewise",
- "layer + filter",
- "layer + kernel",

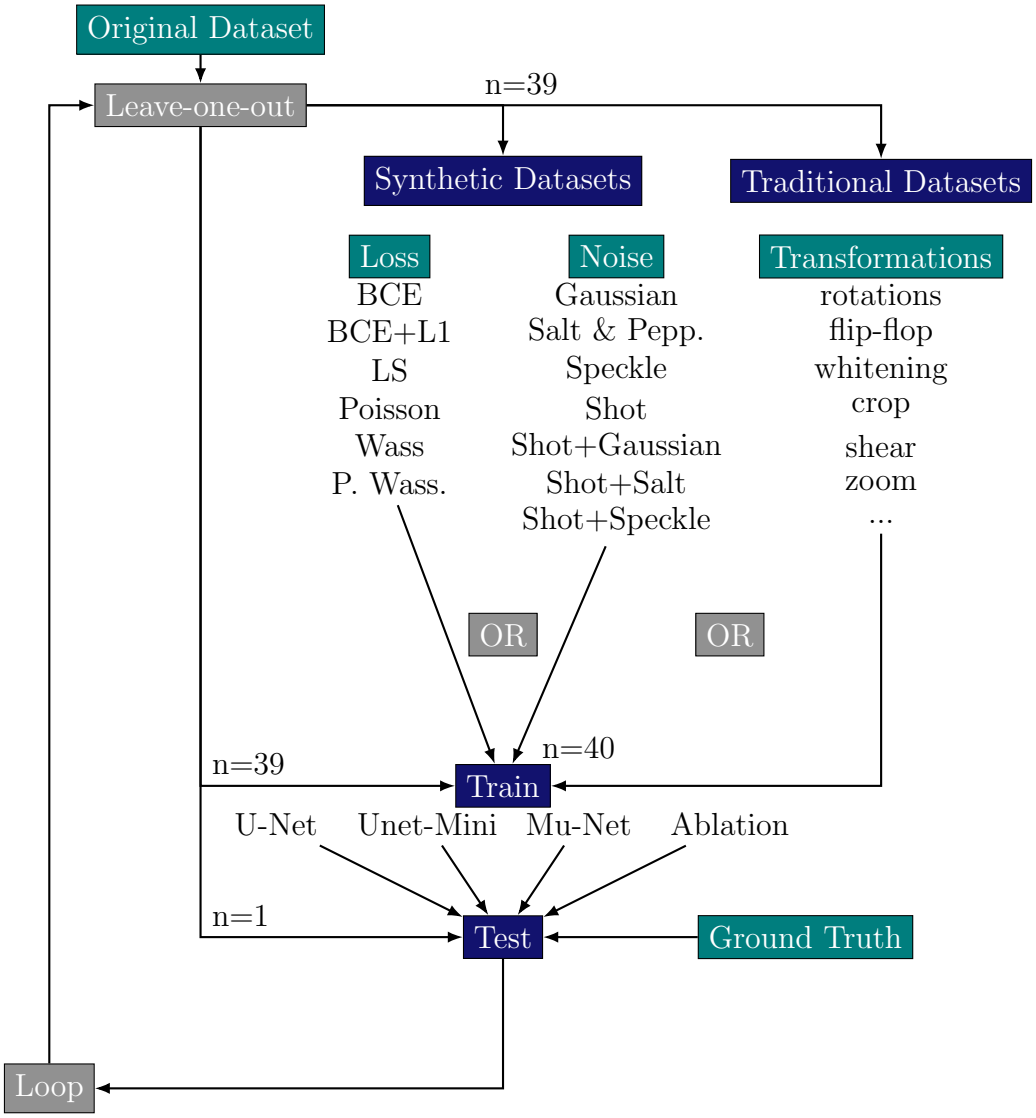


Figure 4.10: Experimental scheme of the leave-one-out strategy to test the effect of various data augmentation on segmentation. The synthetic and traditional datasets are generated at each loop without the leaved one out image.

- “layer + ewise”,
- “Filter + kernel”,
- “Filter + ewise”,
- “Kernel + ewise”,
- “layer + filter + kernel”,
- “layer + kernel + ewise”,
- “layer + filter + ewise”,
- “filter + kernel + ewise”,
- “layer + filter + kernel + ewise”,
- “layer + filter + kernel + stride” (Mu-Net),
- “layer + filter + kernel + ewise + stride + activation”(UNet-Mini).

We decide to determine a binary content and always use a sigmoid activation.

4.3.2 Results

We suggest to tackle a segmentation task in a leave-one-out strategy (n=79 for training and n=1 for testing for each data augmentation strategy). We choose the classic U-Net architecture and its two lighten architectures as UNet-Mini and the purpose Mu-Net.

Qualitative

We suppose that the data augmentation strategy used during each architecture training time could modify the segmentation results. Table 4.10 represents segmented samples from U-Net, UNet-Mini and Mu-Net according to the AAE loss optimisations against classical data augmentation strategies while Table 4.10 represents segmented samples according to the AAE noise injection. Concerning the Table 4.10, U-Net architecture seems to be the most accurate while UNet-Mini and particularly Mu-Net segmentations show high amounts of false positive pixels except for Wassertein loss optimisations. Concerning the Table 4.11, U-Net and Mu-Net show almost similar result with again a small amount of false positives around the shape of the cerebral organoid. For UNet-Mini, a particular false positive around the contour of images are seen, and the salt of paper is characterized by a huge amount of false negative.

Quantitative

As we cannot decipher the real precision of segmentations, with human observation on a single image from a dataset, we calculate median scores for each case summarized for loss optimisations in Table 4.12 and for noise injections in Table 4.13. Dice scores are for all the architectures more elevated for each AAE loss optimisation strategy respectfully to classic data augmentation. The highest Dice score is reached by U-Net architecture when AAE Perceptual Wassertein loss optimisation is used during training. This loss optimisation used in UNet-Mini training allow also to reach the highest score however lower than for the U-Net

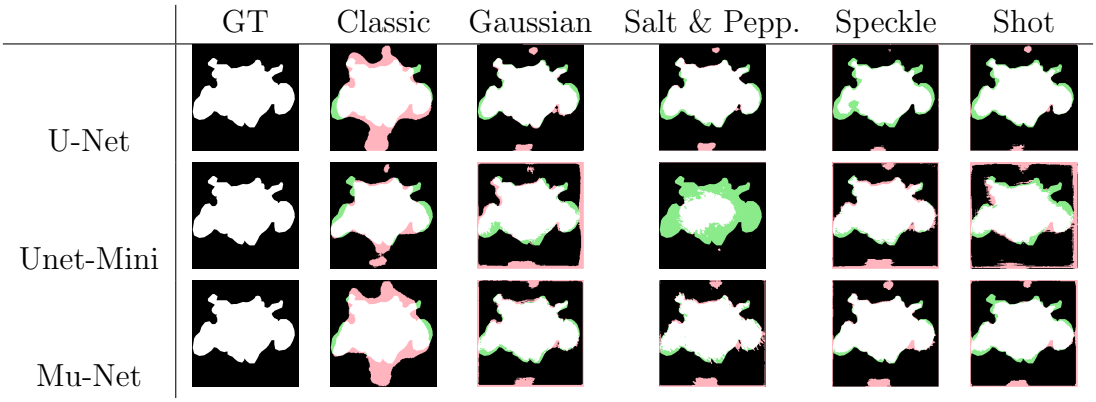


Table 4.10: Observation of Segmentation realized with several loss optimisation data augmentation strategy during the training for U-Net and its derivatives architectures. GT corresponds to the ground truth segmentation. False positive pixels are colored in purple and false negatives in cyan, real negatives in black and real positives in white.

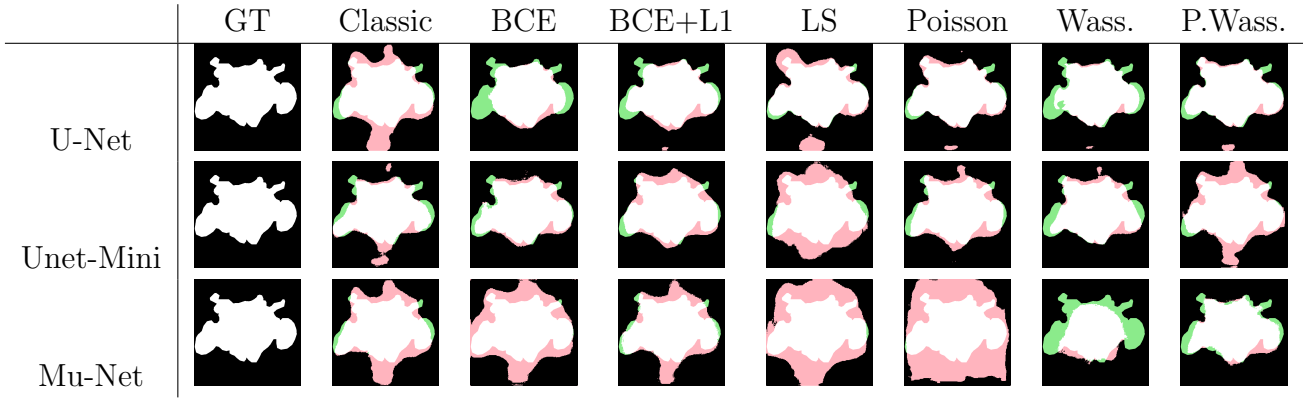


Table 4.11: Observation of segmentations realized with several noise injections for data augmentation strategy during the training for U-Net and its derivatives architectures. GT corresponds to the ground truth segmentation. False positive pixels are colored in purple and false negatives in cyan, real negatives in black and real positives in white.

	Classic	BCE	BCE + L1	LS	POISSON	Wass.	P. Wass.
U-Net	0.68	0.87	0.87	0.86	0.87	0.88	**0.90**
Layer	0.83	0.89	0.88	0.87	0.90	0.88	**0.91**
Filter	0.85	0.56	0.84	0.82	0.83	**0.85**	0.83
Kernel	**0.85**	0.17	0.84	**0.85**	0.68	**0.85**	0.64
Ewise	0.51	0.69	0.72	**0.80**	0.66	0.68	0.73
Layer + Filter	0.84	0.87	0.87	0.85	**0.88**	0.87	**0.88**
Layer + Kernel	0.34	0.35	0.68	0.68	0.35	0.50	**0.70**
Layer + Ewise	0.33	0.21	0.28	0.28	0.25	0.32	0.37
Filter + Kernel	0.71	**0.78**	0.72	0.59	0.70	0.71	0.74
Filter + Ewise	0.67	0.39	**0.77**	0.76	0.59	0.31	0.65
Kernel + Ewise	0.17	**0.65**	0.55	0.59	0.64	**0.65**	0.56
Layer + Filter + Kernel	0.83	**0.91**	0.88	0.86	0.87	0.88	0.87
Layer + Kernel + Ewise	0.35	0.33	0.16	0.16	0.21	0.31	0.37
Layer + Filter + Ewise	0.42	0.28	0.19	0.36	0.13	0.29	0.41
Filter + Kernel + Ewise	0.50	0.34	0.29	0.29	0.45	0.37	0.21
Layer + Filter + Kernel + Ewise (Mu-net)	0.85	0.88	0.87	**0.89**	0.87	**0.89**	0.88
Layer + Filter + Kernel + Stride	0.84	0.83	**0.85**	0.78	0.84	0.82	0.81
Layer + Filter + Kernel + Stride + Activation + Ewise (Mini-Unet)	0.86	0.84	0.87	**0.88**	0.85	**0.88**	0.82

Table 4.12: Dice scores for segmentation with several loss optimisation of data augmentation strategies during training

architecture. According to Table 4.13, Mu-Net seems the most accurate in term of median Dice scores, followed by U-Net and by UNet-Mini. Images from an AAE optimized with a Gaussian noise injection used during the training of Mu-Net show the highest Dice scores. However, others noise injections in the generative process do not improve the segmentation task.

To summarize, U-Net gives the most accurate results in term of segmentation of bright-field images from brain organoids. However, Mu-Net is a promising alternative with its almost same accuracy with classical data augmentation strategy and particularly execution time.

To verify the effect of each component of the architecture on the segmentation precision of bright-field image in a specific data augmentation strategy, we realize an ablation study. It consists in suppressing or adding an architecture component (one by one) and calculating then the median Dice scores for each intermediate architecture. Results of this study are summarized in the lowest part of Tables 4.12 and 4.13. At first sight, the layer reduction allows an increase in Dice score whatever the data augmentation strategy realized, whereas the filtration does not improve the results. The kernel reduction allows the highest Dice score for the binary cross entropy. Layer reduction of the U-Net architecture allow to increase the segmentation accuracy particularly when images from a generative process with perceptual Wasserstein loss are used during the training time of this model. Similar results are obtained with a layer, a filter and a kernel reduction combined with generated images from BCE loss used during the training. Best results are obtained with P.Wass loss used for generated images used during the training time. Regarding noise injections, best results are obtained with a Gaussian noise whatever the reduction made.

	Classic	Gaussian	Salt Pepp.	Speckle	Shot
U-Net	0.68	**0.83**	0.82	0.74	0.74
Layer	0.83	**0.91**	0.83	0.83	0.85
Filter	0.85	**0.83**	0.52	0.73	0.74
Kernel	0.85	0.64	0.72	0.58	**0.75**
Ewise	0.51	**0.73**	0.22	0.18	0.38
Layer + Filter	0.84	**0.88**	0.83	0.69	0.80
Layer + Kernel	0.34	0.70	**0.77**	0.49	0.40
Layer + Ewise	0.33	0.20	0.54	0.20	0.32
Filter + Kernel	0.71	0.74	0.59	0.55	0.62
Filter + Ewise	0.67	**0.65**	0.45	0.10	0.21
Kernel + Ewise	0.17	0.56	0.45	0.11	0.26
Layer + Filter + Kernel	0.83	**0.87**	0.81	0.77	0.74
Layer + Kernel + Ewise	0.35	0.42	0.37	0.26	0.32
Layer + Filter + Ewise	0.42	0.18	0.14	0.08	0.14
Filter + Kernel + Ewise	0.50	0.21	0.29	0.26	0.16
Layer + Filter + Kernel + Ewise (Mu-Net)	0.85	**0.88**	0.80	0.81	0.83
Layer + Filter + kernel + stride	0.84	0.81	0.59	0.75	0.81
Layer + Filter + Kernel + Stride + Activation + Ewise (Mini-Unet)	0.86	**0.82**	0.74	0.48	0.77

Table 4.13: Dice scores for segmentation with several noise optimisation of data augmentation strategies during training

Architecture	Time (1 launch s=h)	Time (LOO s=h)
U-Net	990=16.5	39613=660
UNet-Mini	157=2	6282=104
Mu-Net	**151=2**	**6044=100**

Table 4.14: Training time execution for each segmentation architecture in seconds for one image and for the entire leave one out process (LOO on 40 images). The time is written in seconds=hours (s=h).

Computational comparisons

We also verify our supposition measuring the average execution time for a training and a test phase for a single image and for the full leave one out strategy (40 passages) for each architecture, whatever the data augmentation strategy used while training, which are summarized in Table 4.14. As supposed, UNet-Mini and Mu-Net realize the quickest execution compare to the original implementation.

4.3.3 Discussion

The main outcomes of this study are: we presents a new light segmentation methodology (Mu-Net), compared with architectures from literature, according to various data augmentation strategies, to segment small bright-field image datasets of brain organoids, which have particular formations essential to extract. The presented framework based on U-Net network runs the quickest for a single launch or for an entire leave-one-out process. Mu-Net reaches almost the highest scores obtained with U-Net in term of precision with the particularity of reaching it with classical data augmentation strategies used during its training

time, whereas similar scores are obtained with optimized AAE strategies with U-Net. The ablation study confirms that another intermediate architecture, based upon a single layer reduction could be used to obtain higher scores.

To our knowledge, this is the first automated segmentation dedicated to small dataset of two dimensional bright-field images of brain organoids. Segmentation of small datasets (not always labeled) remains an actual issue in the domain of biomedical images. In the case of brain organoid growth following, segmentation is a crucial tool to extract their morphological shape and characterize their physiological or pathological development, in order one day to make them a brain in vitro model which could fill the gaps of in vivo models (Madeline A. Lancaster et al., 2013). Only U-Net has been used to segment automatically ventricular zones in the Scout software for clarified brain organoid images and to extract the shape of two dimensional brain organoid images in order to validate the accurate data augmentation strategy (Albanese et al., 2020; Brémond Martin et al., 2021a). However, training time consideration, the amount of labelled images required and, the approximate segmentation of neuro-epithelial regions (essential to determine if the culture is well developed) convinced us to found an appropriate solution for these particular cultures.

Among light U-Net architectures, Mu-Net allows a reduction of training time execution. This could be explained by the drastic reduction of learned hyper-parameters, and also by the layer reduction in these networks. Usually, the sophistication of the architectures aims at producing highly accurate segmentation, however it involves a cost in term of proceeding and computation. Contrary to others U-Net derivatives applied on others biomedical images (3D, attention, inception, residual, current neural network, and adversarial) reducing the architecture remains a few update proceeding task (Siddique et al., 2020) and could be quickly implemented on platform rendering image segmentation.

Mu-Net is almost as accurate as U-Net according to the data augmentation strategy used during the training. In contrast, UNet-Mini never reaches higher Dice scores than 0.85 though they shared an important part of similitude in their implementation. This is explained by the transfer learning method used during the training time: while U-Net and our Mu-Net architecture use concatenations to long skip connections, UNet-Mini uses an additional long skip connection based upon an element-wise product (Jirik et al., 2020; Ronneberger et al., 2015a). A skip connection skips layers in the neural network and feeds the output of one layer as the input to the next layers. In addition, data captured in the initial layers are features corresponding to lower semantic information that are extracted from the input and shared to output. In concatenation, low-level information is shared between the input and output, and this information is directly shared across the network. This method allow the feature compaction and high standardisation which is not possible with addition.

The particularity of Mu-Net is to reach highest Dice scores in case of a classical data augmentation strategy is used during the training time. The ablation study confirmed that the combination of layer reduction and kernel size allows to conserve a high Dice score segmentation and not only the layer reduction. The supposition we propose would be the combination of the reduction layers + kernel size and the concatenation allows a transfer learning on essential features. Classical data augmentation chosen here is supported by simple transformations such as flip-flop, rotations of 90 degrees or whitening of images by applying a filtration. GAN solutions are based on two part generative and discriminative networks which have to be train and optimised. If this behaviour is observed with another dataset of bright-fields images, it could render this architecture more attractive (to avoid using data augmentation strategies based upon GAN architectures which is time consuming).

We notice a high rate of false positives in samples chosen to illustrate the architecture

results in the case of classical data augmentation. A false positive is defined as an instance of an organoid region that is present in the predicted mask has zero intersection with the ground truth mask region. Surroundings of brain organoids are constituted by spreading cells (Kelava et al., 2016a) that could be interpreted by these network as organoid regions. The ground truth could have an effect on this result. This may be due to the small amount of dataset used to be trained combined with the less deeper architecture which could cause the approximate learning of surrounding details. It could be interesting to realize a segmentation with all data augmentation strategies in the same training dataset in order to observe its segmentation effect.

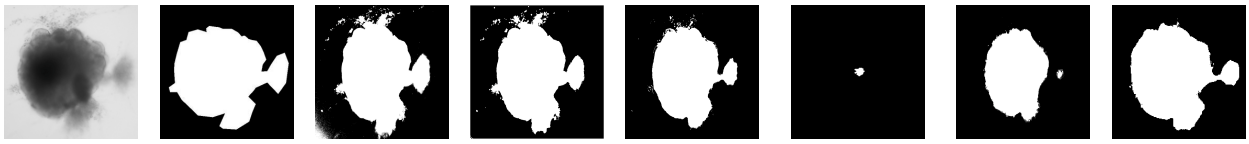
A comparison that could be done is to compare this Mu-Net architecture with others U-Net variation such as described in Siddique et al., 2020. We observe the fundamental role of concatenation in this architecture, however others U-Net variations are characterized by dense connections (Hoang et al., 2021). Such update could be interesting to combine with the reduced architecture in order to observe its contributions to the bright-field image segmentation of these cultures. We choose also a supervised version (with manually labeled datasets as ground truth) but in future comparisons it could be interesting to realize a zero shot learning solution in order to avoid the neurobiologist time vanishing in manual segmentation (Qi et al., 2020). Another possible comparison could be also made with machine learning tools (active contours, Kmean etc.). Concerning the metrics comparison, we could calculate others known scores (F1, accuracy...) in future studies. Augmentation strategies could be also realized with non linear transformations or others GAN architectures (Yi et al., 2019). We mention the noise injection optimisation benefit during the training need to be thorough. We verify the contribution on ablation study which implies only a few configuration. In the future, it will be interesting to realize an ablation study on filtering, kernel reduction, or even others skip connections. The main future objective is particularly to reproduce all these procedures on another bright-field data-set of brain organoid images, or of another biomedical model if there is no available.

We propose our first results on Mu-Net a framework for bright-field brain organoid image small data-set segmentation. Apart from validating our methodology on a second data-set, we need also to compare the particular contribution of the combination of a lighten architecture with others specific transfer learning methodologies. Once validated, this segmentation architecture could help at deciphering the particular growth formation on these cultures in two dimensions in a quickest time than the original U-Net implementation and maybe help at their morphological characterisation. Transferring this framework to others biomedical models bright-field images could also be useful.

4.4 Machine learning segmentation of brain organoids

This part has not been published.

In recent articles the advent of deep learning methodologies has been exposed, however, for some use, classical tools could maybe more accurate. Particularly when the training set is too small, the model cannot learn well the task and the segmentation became less accurate than others. To verify if the previous developed deep tools render better results than classical machine learning ones, we compare some classical segmentation tools with the best deep methodology.



(a) (b) (c) (d) (e) (f) (g) (h)

Figure 4.11: Qualitative comparisons of ML segmentations with Mu-Net. a) Real, b) GT, c) Threshold, d) Active contour, e) Kmean, f) Region, g) Watershed, h) Mu-Net.

4.4.1 Methods

The protocol is the same as developed previously in part 4.3 and we use the same resources and metric comparisons as previously. The classical segmentation algorithm does not require training time.

Machine learning tools

We decide to compare the segmentation results by testing one of each machine learning methods described in 4.1.1. We test an Otsu thresholding from (Otsu, 1979b), a region growing algorithm (Adams et al., 1994), a Kmean clustering with a particle swarm optimisation from (Vora et al., 2013), an active contour from (Menet et al., 1990), a level set from (Meziou et al., 2012), and a watershed method from (Sharma et al., 2022). The threshold has been set to 185 pixel value. For region growing algorithm, we set the regional threshold to 0.2, and the size of growth to 1.

Deep learning tool

We compare U-Net, Mu-Net and the previous detailed classic segmentation algorithms with the Dice, F1-score, Accuracy, Sensitivity and Specificity as previously described.

4.4.2 Results

The interest of this work is to compare various machine learning algorithm with Mu-Net, our deep-learning lighten architecture, first by observing a sample of images obtained, then quantitatively by using metrics.

Qualitative analysis

A sample of each segmentation is presented in Figure 4.11. Qualitatively, the threshold, active contour, Kmean and Mu-Net seems to render the best results. The region growing algorithm and threshold render the worst result with the watershed maybe due to their lower precision linked to the manual setting of some parameters.

Quantitative analysis

The metric comparisons between GT and each segmentation in Table 4.15 are consistent with the previous observations in Figure 4.11. Comparisons of indices between Mu-Net results with classical ml show it render best results with the Kmean algorithm. But does not outperform this method in term of Dice and Specificity scores but in F1-score, Sensitivity and Accuracy a see. Others segmentation methods does not reach high scores, and are the worst for the threshold and region growing algorithms.

	Threshold	Active	Kmean	Region	Watershed	Mu-Net
Dice	< 0.1	0.62	**0.89**	< 0.1	0.61	0.88
F1-score	0.40	0.66	0.72	0.20	0.61	**0.77**
Sensitivity	0.39	0.69	0.77	0.19	0.53	**0.90**
Specificity	**0.96**	0.84	0.89	0.86	0.94	0.86
Accuracy	0.83	0.90	0.90	0.7	0.88	**0.92**

Table 4.15: Mu-Net with classical segmentation comparisons

4.4.3 Discussion

In this part, we originally compare our original segmentation contribution Mu-Net with classical machine learning (ML) algorithms. Some ML segmentation algorithms seem not well suited to extract the shape of brain organoid without the optimisation of a pre-processing step and, are only semi-automated. Mu-Net also outperforms almost all the ML procedures, except the Kmean clustering algorithm for some scores.

ML strategies do not give strong results of segmentation. The main drawback observed in this study is the use of pre-setting of some parameters in ML solutions. Set most of the time manually, and despite the fact of testing various combinations, these parameters render the segmentation process semi-automated and, especially not expected scores. The non expected scores are also linked with the fact ML algorithms are not appropriate for datasets with strong variations (Malhotra et al., 2022). However, in early developmental stage of brain organoids there is no neuroepithelial formation while after 14 days the shape is transforms due to this structure (Gomez-Giro et al., 2019). They are also influenced by the batch syndrome: brain organoids are growing differently in the same environment producing one or a multiple of these neuroepithelial zones (Madeline A. Lancaster et al., 2013). Thus instead of furnishing a pre-setting parameter for each image contained in particular dataset, ML solution are not dedicated for these cultures.

The Kmean machine learning algorithm is an exception and reaches the best segmentation scores. Taking appart the number of parts to segment (here the background and the brain organoid), this algorithm do not require another setting which will influence the score, and this setting could be given automatically. The particularity of the Kmean clustering algorithm is to compare the cluster containing a centroid with a pixel (Vora et al., 2013). The addition of the particle swarm optimisation allow to get a more precise clustering efficiency and thus a better Accuracy than a single Kmean algorithm. Taken together, this could explain why the Kmean render the best results and should be used to segmente brain organoid tiny datasets images.

Mu-Net outperforms the ML strategies except the Kmean. Our reduced model of segmentation for brain organoid tiny dataset of images, is automated and do not require presetting. The light architecture allow to perform the segmentation of an image almost as quickly a Kmean algorithm and produce similar results. However, for now, this architecture need to be improve to be the one to choose to segment brain organoid tiny datasets. An improvement could be to use prior information during the Mu-Net training (such as giving shape or topological information on the kind of image segmented) (El Jurdi et al., 2021), but this will require a better knowledge of the morphological characterisation of these cultures.

To conclude in this part, the comparative study helps highlighting Mu-Net architecture can render almost as accurate segmentation as the Kmean and, should be proffered to replace others ML strategies as Mu-Net outperforms their scores. We are convinced an optimisation by shape/topological information given during the data augmentation step to train better

our model could improve the segmentation results of brain organoid tiny datasets and, may outperforms also the Kmean in the future.

4.5 Discussions and conclusions on segmentation

The objective of this segmentation part is to extract the shape of brain organoid tiny datasets of microscopic bright-field acquisition in order to characterize their morphology. We first compare various training procedure with the golden standard deep learning segmentation methodology which is U-Net, in order to decipher the most usefull data augmentation methodology for a tiny dataset. Then, we reduce this architecture which requires a certain amount of data and trianing time in order to simplify it and we obtain particular accurate segmentation which are not far from a ground truth (manual segmentation). In the last part we show how a reduced deep learning architecture with a particular data augmentation strategy for tiny datasets could produce the most precise segmentation compared to ML architectures. In this discussion, we will highlight the benefits of these results, some drawbacks, the pioneer work we attempts, and the industrial and future perspectives of research in the segmentation domain for these tiny datasets.

4.5.1 Discussion on our work

In our work we propose to verify the influence of data augmentation during the learning step of a deep learning segmentation process. We propose a light U-Net architecture dedicated to the extraction of brain organoid shape in bright-field acquisition and, verify if it succeed to be better than ml processes. Thanks to this work the brain organoid field can now count into an automatic way to extract their shape which is fast, and does not require a lot of training steps. This architecture can be improved by shape prior, topological or contextual information. It could be also interesting to test the segmentation with this architecture on other microscopic acquisition of this culture. Similarly, we need to compare this architecture with others non U-Net variation architectures such as CNN, FCN, or attention based models detailed in Paragrap 4.1.2. It could be also interesting to test the effect of Few, One or Zero shot learning on this kind of architecture such as proposed in Pastore et al., 2021. The interest of learning a small database in an unsupervised way will be a challenge to raise in the future. Then, in order to generalize its use on others bright-field images, it could be interesting to test this architecture on others biological models such as the already cited blastocystes bright-field image dataset from this link https://bbbc.broadinstitute.org/image_sets.

4.5.2 Initial work on perspectives

The main future work concerns the test on others architectures such as described in part 4.1.2. Concerning the comparisons with ML architecture, we need to observe the influence of some optimisations particularly on the initialisation step. The validation on other bright-field biomedical microscopic images has steel to be tested in order to generalize the use of Mu-Net.

4.5.3 Academic and industrial perspectives

Industrial and academic perspectives from this segmentation work on segmentation are various:

- The Mu-Net architecture could help to segment accurately tiny datasets of bright-field brain organoid images.
- It confirms the data augmentation used during the training step of a deep learning based segmentation algorithm influence the accuracy of the shape extraction. In an automatic procedure to extract every shape accurately, on a tiny dataset, the test of various data augmentation strategies is required.
- The interest of this comparison scheme is to propose to segment other biological models, or other microscopic acquisitions on brain organoid images.
- We nonetheless have to compare others optimisation of the Mu-Net architecture to outperform a Kmean optimized machine learning segmentation.
- The effect of few-, one- and zero-shot learning needs to be tested.
- The test of this architecture on others biomedical models or on the other microscopic acquisitions of brain organoids could help to verify the generality of this algorithm and spread its use in other applications.
- Once brain organoid is characterized, an interesting perspective could be to add a prior shape/topological/channel/cnstraint knowledge to extract the shape of others brain organoid datasets according to their development more accurately.
- The development of these ideas could open new issues to answer and develop. Developing a scalable algorithm which could be optimized or reduced according to the tiny dataset given in the biomedical field could be also of interest.

4.5.4 Conclusion

The main contributions of this chapter are:

- The effect on the segmentation accuracy of brain organoid bright-field images of various data augmentation strategies based upon GAN optimized used during the training session, has been verified.
- Perceptual Wasserstein loss and Speckle noise optimisation, or psychovalidated images by 8 experts, for a training synthetic dataset render the best qualitative segmentations.
- We propose a lighten U-Net architecture dedicated to the brain organoid shape extraction.
- We validated this architecture on another dataset, we propose an ablation strategy, and compare it with other deep or ml based algorithms.

This work and approach should be tested on others models in order to generalize it. An interesting work could be to to use, and thus verify, our segmentation strategy to characterize the brain organoid development.

Chapter 5

Characterisation

The study of the morphology of biological structures, is essential to characterize the physiological development. Morphology originally describes a branch of biology that deals with the form and structure of animals and plants. Also called set theory in other domains, is used to extract features representing a regional shape in images (Hausdorff, 2005).

In the previous sections we generated synthetic images of brain organoid cultures in order to segment their shape see Sections 3 and 4. The binary images obtained are used to characterize the morphology of brain organoid cultures. The interest of morphological characterisation is to better understand these structures in order to predict a developmental stage, or, to distin a physiological and pathological development, or to follow the growth of a culture.

In this section we describe morphological calculation which describe biological contents in images: first the traditional operations, then the topological based analysis, and the deep learning methods, the use of these tools in the biomedical field, our strategy and our contributions.

5.1 State of the art on characterisation

5.1.1 Classical morphological characterisation

Traditional morphological characterisation contains various types of index calculations depending on boundary or regional descriptors (Gonzales et al., 1987). Here, we describe the brain organoid literature classical descriptors without taking into account those only describe in Section 2.

Boundary descriptors

Also called simple descriptors, boundary descriptors contain the length, diameter, curvature the shape number, Fourier descriptors and moments. The length is the simplest descriptor, it consists in counting the number of pixel constituting an axis of a shape. While the diameter of a boundary consists in creating a line connecting the two extreme points of the boundary, and is also called the major axis. The curvature is a rate of slope changes, and is more difficult to calculate than the length and diameter due to no continuous boundaries, and the various length of individual segments. It consists in decomposing the boundaries into straight segments which simplify the description process and then calculating slopes. At a point p, the curvature descriptor is described by ranges of change slope: p takes part of the nearest segment if the change is less than 10°, p takes part of a corner if the change

exceed 90°. Thus the boundary vertices can also be interpreted as curvature descriptors: a vertex point p takes part of a convex segment if the change of slope at p is non negative and, p takes part of a concave segment in the contrary.

Another boundary descriptor takes its origin from the chain coded boundary and is called the shape order. A chain code boundary represents the boundary by a connected sequence of straight lines and each direction of a segment is encoded by a number. The difference between chain code boundaries are the starting point. While the shape order takes into account only the number of values constituting the representations. For instance a square s with a starting point at the top left corner has a shape order of 4, a chain code of 0321 and a difference of shape number of 3333.

Fourier descriptors distinguish boundary shapes, are not sensitives to the starting point, and the level of details can be modified by increasing or decreasing the number of elements. The Cartesian descriptor, also called elliptical Fourier descriptor, can describe some invariants. Indeed, to compare shapes, a Fourier descriptor is invariant to geometric transformations, which is at minimum: for scale changes, translation, rotation and curvature for 2D and for 3D. For an arbitrary closed 2D polygon $Poly$ of V from $(v_{p0}, \ldots, v_{p-1})$, a sequence of p points representing the vertices of $Poly$ we set a number of Fourier coefficient pairs Np and the length of Fourier descriptors will be $2Mp + 1$ (see Appendix pseudocode **??**).

Moments describe the shape of boundary segments or signatures. Signatures are 1D representations of boundaries by plotting the distance from the centroid to the boundary as an angle. This method allow to reduce the boundary to a 1D representation, which is simpler to describe, and translation invariant. However, they are rotation and scale-dependant. Thus if a shape is noisy, the dependencies generate errors. A segment is represented as a 1D function $g(v)$ of a variable v, and its amplitude as a random variable a. The amplitude histogram (AH), with i increments of a $(1 : K)$ is then calculated, with K the total number of amplitude increments, or in a normalisation case of $g(v)$ with v $(1 : B)$ where B is the number of boundary points. The n^{th} moment is then defined by the mean and variance of a or v such as described in equation 5.1 and 5.2, 5.3 and 5.4.

$$m_a = \sum_{i=1}^{K} a_i AH(a_i) \tag{5.1}$$

$$\mu_{n(a)} = \sum_{i=1}^{K} (a_i - m_a)^n AH(a_i) \tag{5.2}$$

$$m_v = \sum_{i=1}^{B} v_i g(v_i) \tag{5.3}$$

$$\mu_{n(v)} = \sum_{i=1}^{B} (v_i - m_v)^n g(v_i) \tag{5.4}$$

Usually, to differentiate signature of distinct shapes, only the first moments are used. For instance $\mu_{n(v)}$ is related to the shape of g(v): the second moment measures the spread of the curve of m_v; the third moment measures the symmetry of m_v.

Regional descriptors

Boundary descriptors are not sufficient to describe the morphology of images. They are completed by regional descriptors such as: perimeters, areas, compactness, and axes or more complex as topological, or textural descriptors.

The perimeter is defined as the length of the boundary while the area is defined as the number of pixels contained inside the boundary. These two morphological indexes compose the compactness, which is insensitive to scale changes and orientation:

$$Compactness = \frac{perimeter^2}{area} \tag{5.5}$$

The direction of a region inside a shape corresponds to the biggest eigenvalue and its eigenvector. Matrix covariance eigenvectors are also called principal axes and obtained from pixels inside a region described by a vector from a population of random vectors with n dimensions $x = [x1, x2...xn]$ and its mean as $m_x = Ex$ where E is the expected value:

$$Covar_x = E(x - m_x)(x - m_x)^T \tag{5.6}$$

These descriptors are insensitive to rotations, but depends on the scale changes. To overcome this issue, the ratio of the largest to the smallest eigenvalue is taken as the descriptor.

A more complex regional descriptor is to quantify the texture which measures the regularity or smoothness for instance of images. To calculate these descriptors, statistical, structural and spectral approaches are used. The statistical approach usually uses moments of the gray-level histogram from a region and, the methodology is optimized with the position of pixels. The structural approach yield with regular parallel space lines while the spectral approach deals with detect the periodicity in images by highlighting the high energy peaks in Fourier spectrums.

5.1.2 Topological based characterisation

Traditional morphological analysis may miss some information when the analysis concerns some extreme datasets such as tiny or big datasets in the biomedical field. Topology is the study of the figure properties which are not affected by any transformation (Zomorodian et al., 2012). For instance, a topological descriptor can be defined by the number of holes H in a region, which are not affected by rotations. Second topological descriptor is the number of connected components C inside a region. Connected component is a subsample of maximal size, with a connected curve joining two points of its points. The number of Euler ϵ, the third topological property is defined then by:

$$\epsilon = C - H \tag{5.7}$$

For a polygon P, regions are represented with straight line segments, and the region interior with a number of faces F, holes H, vertices V and edges E and the Euler number is then defined by:

$$C - H = V - E + F \tag{5.8}$$

The principle of topological data analysis (TDA) is to convert complex datasets due to the huge amount or the lack of information into a simple summary of their features and render may be a solution to this issue. TDA is split in two strategies: persistence homology and mapper (Skaf et al., 2022; Iniesta et al., 2022).

Mapper is dedicated to the visualisation, it transforms a high dimensional data into a skeleton which form a simplicial complex. The skeleton of a region is obtained by a thinning and most of the time is defined by the medial axis transformation MAT. In MAT, the nearest neighbor of the border B is found for each point p in a region R. When p has more than one neighbor, it belongs to the medial axis, also called the skeleton of R. The neighborhood depends on the distance chosen.

In the mapper case, a data set is a point cloud expressed by p in P and a filter as a function $f(p)$ is applied on each p to obtain its x coordinate in the P space. These values are then displayed vertically to obtain a map with overlapping intervals in colors which form boxes. These boxes represent the pullback of $f(p)$: a box contains p when the values are in the range of the same colors. Points inside a pullback are finally clustered as a node (0-simplex), and if two clusters have a point in common, their nodes are linked to form the simplicial complex.

The persistence homology lies on homology and on persistence properties. The homology counts the topological features for each dimension. The number of connected components in 0 dimension is called $\beta 0$ and points belonging to the same connected component are described as homologous. The unique 1-dimensional hole in 1 dimension is called $\beta 1$ while, 2-dimensional hole such as a cavity or a void is called $\beta 2$. To summarize, in homology, β_n is called the $n^{th} Bettinumber$ and describes the useful information inside a shape, but must be use with persistence. The persistence aims at connecting points close together and proceed as a Vietoris Rips complex formation by creating simplicial complexes and barcodes for instance. At each step (range values), the number of connections added depends on a parameter (radius) which rule the distance at which each point should be connected. First each connected component is link to its data point p, and only p. Thus $\beta 0$ equals to the number of connected components and the number of points which is the number of birth (departure of lines) of persistence barcode. During the second step, around each point a ball of radius r is set and if two balls are overlapping the points $p1$ and $p2$ are connected and takes part of the same connected component. This is represented by a death in a persistence barcode (one of the line which has been birth die), and the balls and radius are representing simplicial complexes. If there is no overlapping the line is continued in the persistence barcode. While the number of $\beta 0$ is superior than 1 (and thus the lines constituting the barcode are not dead), updates of homology are collected and, the radius is increased.

5.1.3 Deep learning for morphological characterisation

Thanks to morphological characterisation, the images can be classified or used to predict the developmental stage or pathological state of its content. Classification is the main task in deep learning use (almost 30%) and is followed by the prediction (27.53%)in the biomedical field see Figure 5.1. To compare these results segmentation is around 27% of the use and GAN and detection are minor applications. However, we have to mention, classification do not use only deep learning strategies.

Three kind of features can be extracted (W. Wang et al., 2020). Low-level features are extracted upon grayscale density, color, texture shape or position inside the image, or hand-crafted by an expert. Mid-level features are known as word descriptors. These extracted features can be made upon support vector machine(SVM), clustering strategies such as Kmean or fuzzy C-means, or scale-invariant feature transform (SIFT). And high level features are extracted by deep learning solutions such as CNN, FCN or inception networks. The issue with high level features extraction is based on the high amount of image required during the training step of classification based on deep learning strategy.

The disease evolution, or a physiological developmental stage of a particular structure are the two main applications of the prediction. Thus, the goal of classification is to assign one or more labels to an image. After extracting features, the label are assigned to various contents in images by a classifier(Cai et al., 2020).

In the case of small datasets, two strategies are preferred to use deep learning methods for classifications. The first consists in using GANs to increase the original data set number

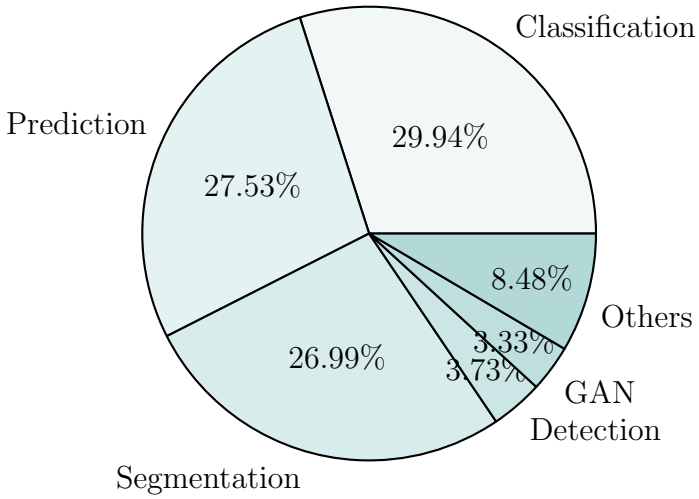


Figure 5.1: Deep learning use in the Biomedical field. We only count the number of publications in the main domain and show it in percentages (on 27617 inputs since 2013).

of images and its variability (J. Wang et al., 2017), and the second is based upon transfer or active learning. The principle of GAN have been already described in chapter 3 so here we will focus on learning strategies.

A requirement for these large datasets, is on top of that to be labeled. Sometimes only a small part of a large dataset is labeled, particularly for diagnosis applications. A way to allow deep learning prediction of small non labeled datasets is to use active learning (Beluch et al., 2018). The use of active learning is in majority used in non deep learning strategy (such as SVM). Recently certain active learning strategies for deep-learning classification tools has emerged such as: ensembles to estimate the prediction uncertainty of deep neural networks in the context of outlier detection and reinforcement learning; uncertainty estimation by measuring the variance. Transfer learning for medical image classification is also used in case of few labeled data. The transfer of information takes birth from the parameters of an original image. Two kind of transfers are existing: by feature extraction when the convolutional layers are frozen or by fine tuning when the model is fit (H. E. Kim et al., 2022).

5.1.4 Morphological characterisation in the biomedical field

Morphological characterisation in the biomedical field has increase since the last decade see Figure 5.2. Among the imaging modalities, the most often characterized morphology, we found histopathological images with 40% of the use and X-ray with 23%, see Figure 5.3.

The domain of application concerns most of the time the morphological characterisation of tumors and cultures (around 20%) see Figure 5.4.

Since 2013, the number of articles in the biomedical field studying the topology are 187.

5.1.5 Positioning and strategies on morphological characterisation

Morphological characterisation has been used for various biomedical images and microscopic acquisitions. Morphological characterisation on cultures concerns almost 20% of the studies and histopathological use are the most studied images. However, the main question raised is if the morphological indexes are adapted to the complex growing of brain organoid

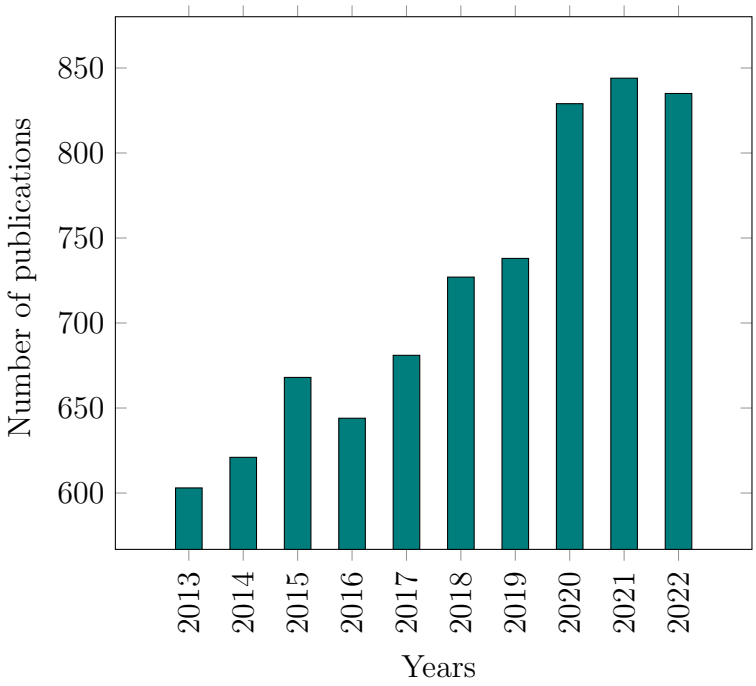


Figure 5.2: Number of publications on morphological characterisation per year in the biomedical field for microscopic acquisition since 2013.

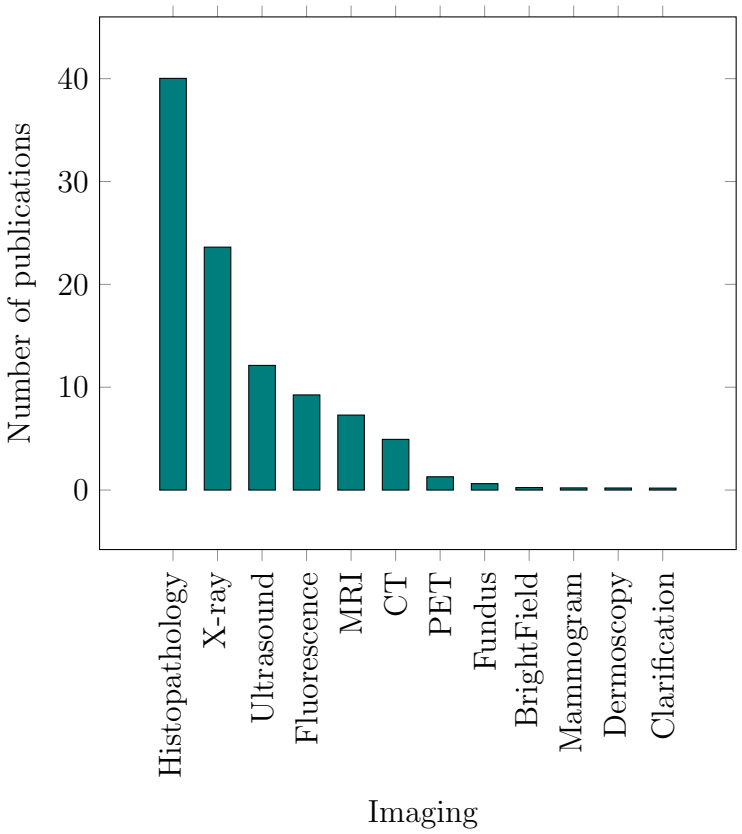


Figure 5.3: Publication on morphology per image acquisition applications. We count the number of publications in the main domain and show it in percentages (on 18223 inputs).

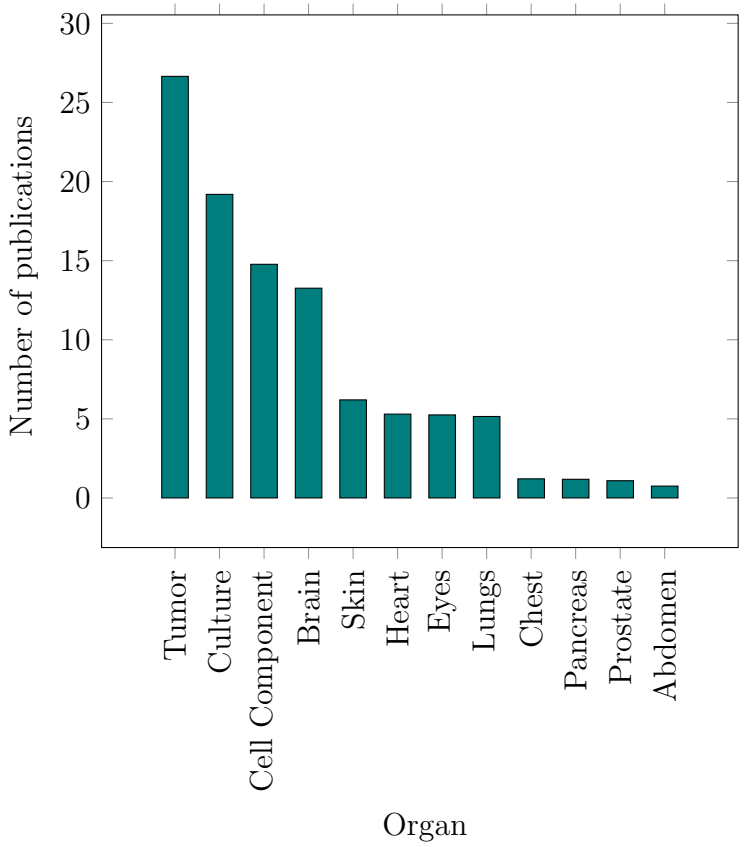


Figure 5.4: Publication on morphological characterisation per organ of study. We count the number of publications in the main domain and show it in percentages (on 18204 inputs).

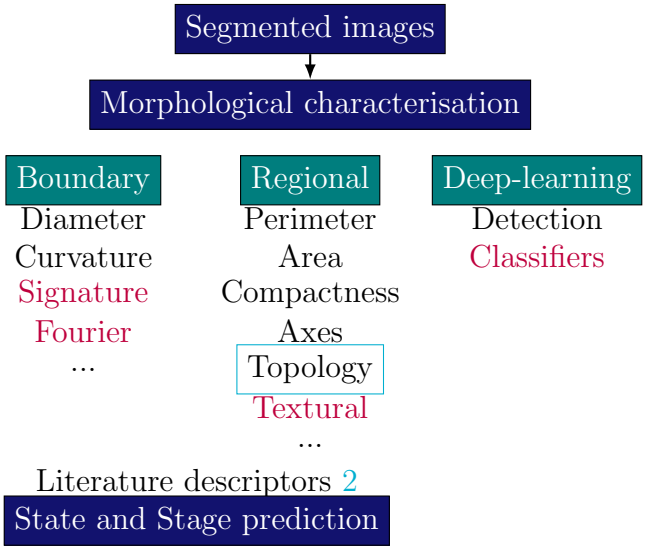


Figure 5.5: Morphological characterisation strategy. Topological analysis is the main strategy developed in this thesis. Non tested strategies are in red.

and their physiological characterisation? We theoretically explain the reasons why such characterisation could not be used alone in Section 2.

We hypothesize a topological data analysis could help to characterize the shape of brain organoid, due to its particularity to reduce the complexity of high dimensional information contained in images. We characterize the morphology of brain organoids by two strategies: the traditional index calculation and the topological analysis in order to compare their contribution to their characterisation, and attempt it also on physiological and pathological images. We attempt to strenghen the analysis by a TDA-clustering strategy. In a second part, we attempt to predict the physiological state by using these topological analysis and we try also to detect directly the neuroepitheliall formation (a clue of their correct development). We are convinced that a deep learning based strategy is not the accurate solution to predict nowadays the brain organoid development with a small dataset and we develop a new solution based upon Morse Theory compared with deep learning-based strategy.

Our strategy is in summary based upon morphological characterisation by TDA and indices. We think these tools could then be used in few set learning strategies in deep learning-based solutions. This order seems to us a logical way to understand the developmental growth of brain organoid and allowing in the future an accurate prediction of its physiological or pathological developmental stage. The summary of the strategy realized is presented in Figure 5.5.

5.2 Evaluation of morphological indexes and segmentation strategies for organoid shape characterisation

This section describes unpublished work.

The morphological index calculation is the main used approach nowadays to characterize the brain organoid morphology. Indeed, morphological characterisation distinguish a physiological from a pathological state or various stages of their culture. However, no article has

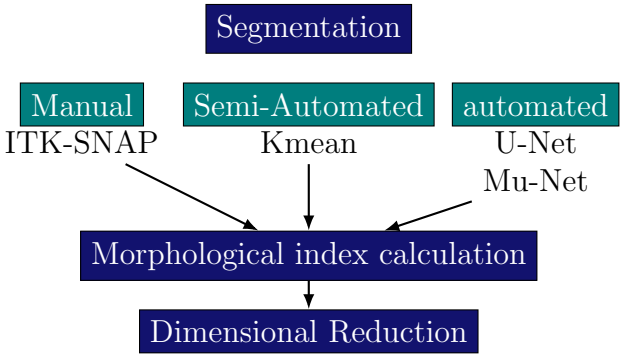


Figure 5.6: Procedure of morphological characterisation of brain organoid with index calculations.

searched which morphological index is dedicated to best characterize a brain organoid at a particular step: all morphological parameters are used according to the subjectivity of the authors.

The second issue resides in the manual or the semi-automated solutions which are used to calculate their morphology; at this moment no article uses an automatic morphological index calculation. Indeed, to calculate morphological parameters, they use one or more software which realize manually or semi-automatically acontour of the shape. This approach is time consuming for biological experts which have to set some parameters before using these tools.

In this part we propose to automatically characterize the shape of brain organoids at particular states or stages by morphological indexes used in others articles. The particularity of our index calculation approach is based upon the comparison of manual, semi-automatic and our automated optimized training segmentation of shape extraction. This methodology is also designed to decipher which kind of morphological parameters could be dedicated to the brain organoid characterisation and to avoid calculate useless parameters in future. The summary of the procedure is presented in Figure 5.6.

5.2.1 Methods

Datasets

We consider 40 binary images (250×250 pixels), resulting from the segmentation of the 40 bright-field images of brain organoids (Gomez-Giro et al., 2019) (see Chapter 4). Images are acquired at 9 days (7 images), 14 days (6 images) and 15 days of culture (6 images) for physiological group and, at 9 days (7 images), 14 days(6 images) and 15 days(6 images) for the pathological group.

We use a second dataset constituted by images of cortical organoids given by the Valrose Institute of Nice. Cortical organoid images are acquired 3 by 3 images every 3 days between day 1 and 55 days with a bright-field microscope with 2 experimental protocol of cortical organoid formations. On the 77 images acquired, only 29 could be used in this data augmentation protocol as we need to use entire images. This cortical organoid dataset is constituted by 512×341 pixel images and centered and focused on the entire cortical organoid. The acquisition render RGB images with no canal alpha. These images are resized for time consideration of algorithms and homogeneisation of the dataset in 250 by 250 pixels.

We constitute three datasets based upon our previous works with the Gomez-Giro et al., 2019 dataset, the second has not been split. The interest is to compare morphological index

calculation on a manual (by ITK-SNAP software), semi-automated (by Kmeans clustering) and automated segmentation (U-Net Ronneberger et al., 2015b and Mu-Net trained with generated images from a GAN optimized with a Perceptual Wassertein loss optimisation) see Chapter 4. All scripts are written in python.

Morphological index calculation

The morphological index calculation is based upon parameters described in Chapter 2 and in Brémond Martin et al., 2021b. Among parameters listed, we choose the ones described in the literature and not only the morphological parameters originally calculated on the original images from the article (Gomez-Giro et al., 2019).

The parameters calculated are: the area, aspect ratio, bounding box area, barycenters (x and y), compactness, circularity, convexity, elongation, folding, convex hull, perimeter, specificity, solidity, wrinkling.

Dimensional reduction strategies

To verify if each morphological index contributes to characterize the morphological state of brain organoid we conduct a dimensional reduction analysis based upon the Principal Component Analysis (PCA) and compare it with t-SNE explained in previous sections see Paragraph 3.2.1.

Statistical analysis

The group we aim to compare are the stages (with the 9, 14 and 15 days modalities) and, the state of culture (with physiological or pathological modalities). To verify if a morphological index is different between two states or the three stages, we verify the significance of each difference with statistics. First we verify the homocedasticity and normality of the responses variables for each group and then apply a T-test or Mann-Whitney test in the case of two groups. We use the analysis of variance or the Kruskall Wallis test in the case of more than two groups comparisons.

5.2.2 Results

We first compare the classical morphological indexes found in the literature to characterize the brain organoid morphology for various stage of culture from various extractive methods: manual (by ITK-SNAP software), semi-automatic (Kmean), and automated (optimized U-Net and Mini-Unet) strategies. We first observe in the manual case 6 morphological index which are significantly different between three developmental stages, see Table 5.2. The area, circularity, convexity, hull and solidity are the highest while the wrinkling is the lowest in the earliest stage. The 14^{th} day is characterized by the lowest circularity and solidity and the highest wrinkling. The latest stage by the lowest area, convexity, and hull. The semi-automatic method highlights other measures such as the aspect ratio and the elongation which are the highest at later stages, and the bounding box at the 14^{th} day. They share the hull but not in the same way: the later stages have highest scores. Automated solutions share the circularity differences with the manual and the convexity, but only for the Mini-Unet extraction.

If we pay attention to the physio-pathological comparison, see Table 5.1, first in a manual way. Healthy brain organoid have highest bounding box, folding, hull, perimeter and wrinkling than pathological brain organoi models and a smaller compactness, and circularity

Table 5.1: Physiological and pathological comparison of morphological indexes calculated on various segmentation strategies on brain organoids.

Index	State	Manual	Kmean	U-Net	Mu-Net
area	Physiological	17177.06	22160.26	19046.5**	18678.67
	Pathological	13249.26	26731.11	14118.68**	20337.21
aspect ratio	Physiological	1.10	0.84***	0.95	0.82
	Pathological	1.03	0.88***	0.93	0.85
bounding box	Physiological	31901.11***	44816	33887.56***	36570.94
	Pathological	22044.63***	42600.17	23106.79***	34723.05
barycentersx	Physiological	128.35	127.98***	129.12	122.45
	Pathological	127.40	122.72***	129.49	130.66
barycentersy	Physiological	122.38	123.87	121.97	113.42
	Pathological	125.47	122.13	124.82	120.79
compactness	Physiological	0.06***	0.16	0.41	0.26
	Pathological	0.09***	0.17	0.47	0.34
circularity	Physiological	1115.80***	4635.98	7817.46	6382.73
	Pathological	1140.68***	7963.78	7454.26	7168.54
convexity	Physiological	14.29	25.59***	29.67	28.46
	Pathological	13.78	23.05***	28.19	28.78
elongation	Physiological	0.81	0.83	0.83	0.77
	Pathological	0.85	0.86	0.86	0.83
folding	Physiological	803.28***	17536.21	622.94***	703.61
	Pathological	611.21***	21702.67	466.63***	711.37
hull	Physiological	26523.5***	32487.26	22963.11***	25998.28
	Pathological	19469.26***	35668.56	16372.58***	24941.53
perimeter	Physiological	1836.11***	1295.79	768.06***	863.83
	Pathological	1407.84***	1752.78	577.74***	860.53
sphericity	Physiological	0.38	0.27	0.45**	0.36
	Pathological	0.43	0.32	0.55**	0.40
solidity	Physiological	0.64	0.68	0.84***	0.72
	Pathological	0.68	0.77	0.91***	0.83
wrinkling	Physiological	3.35***	1.65	1.18	1.24
	Pathological	2.94***	2.15	1.18	1.24

index. Similarly as previously, the semi-automated approach highlights other parameters such as a lowest aspect ratio, or highest x coordinate of the barycenter and convexity index. However, the U-Net automated approach only show similar differences with the manual solution but with other indexes such as the lowest area, and the highest sphericity and solidity in the pathological case.

In cross factor comparisons see Tables 5.3 and 5.4, we observe at the same time the influence of the day of culture and the physio-pathological state. Manual and automatic algorithms based upon U-Net show significant differences between the day and state for the area, bounding box, and hull. However, the U-Net highlight more stages differences and also in aspect ratio and convexity. Semi-automated does not show as much difference as manual solution and automated based upon Micro-Unet does not highlight any differences.

Table 5.2: Developmental stage comparison of morphological indexes calculated on various segmentation strategies on brain organoids.0.05: *, 0.01:**, below = ***

Index	Developmental	Manual	Kmean	U-Net	Mu-Net
area	9 days	17893.21*	20632.8	18811.57	17829.71
	14 days	14431.67*	27631.45	16343.75	22004
	15 days	12476.18*	26251.55	13782.27	19246.92
aspect ratio	9 days	1.02	0.80*	0.92	0.80
	14 days	1.12	0.90*	0.91	0.87
	15 days	1.06	0.91*	0.98	0.85
bounding box	9 days	30286.21	35439.53*	28406.71	33365.93
	14 days	28072.25	50739.82*	30787.92	40202.82
	15 days	21108.55	48052.36*	25623.27	34055.08
barycentersx	9 days	125.86	119.83	126.53	116.79
	14 days	131.52	129.17	133.40	132.47
	15 days	126.41	129.30	128.39	132.87
barycentersy	9 days	126.35	130.75	124.07	115.90
	14 days	122.93	118.26	122.74	116.21
	15 days	122.07	117.26	123.37	119.64
compactness	9 days	0.08	0.15	0.47	0.31
	14 days	0.07	0.17	0.39	0.30
	15 days	0.089	0.17	0.45	0.29
circularity	9 days	1430.85***	3419.90	9048.06*	6752.28*
	14 days	891.45***	11147.47	6719.16*	7199.96*
	15 days	1002.55***	5228.28	6822.04*	6446.66*
convexity	9 days	15.14***	20.92	28.47	27.58*
	14 days	13.40***	30.00	29.76	31.84*
	15 days	13.31***	23.39	28.54	26.89*
elongation	9 days	0.85	0.79*	0.84	0.76
	14 days	0.80	0.87*	0.84	0.83
	15 days	0.84	0.89*	0.86	0.82
folding	9 days	700.79	17291.33	554.93	602.21
	14 days	784.42	17862.91	581.08	754.18
	15 days	622.55	24361.27	485.18	787.83
hull	9 days	25509.36*	27033.6*	20205.79	22524.07
	14 days	23654.25*	40778.82*	21251.75	30071.27
	15 days	18759.82*	36838.27*	16955.73	24644.75
perimeter	9 days	1658.71	1374.67	691.71	745.71
	14 days	1752.83	1559.36	711.25	919.45
	15 days	1413	1672.45	598.45	945.42
sphericity	9 days	0.41	0.31	0.56	0.37
	14 days	0.38	0.32	0.47	0.38
	15 days	0.41	0.26	0.46	0.40
solidity	9 days	0.70**	0.77	0.94	0.76
	14 days	0.62**	0.64	0.82	0.76
	15 days	0.67**	0.73	0.84	0.80
wrinkling	9 days	2.95*	1.76	1.22	1.17
	14 days	3.49*	2.11	1.27	1.24
	15 days	2.99*	1.85	1.041	1.32

Table 5.3: Cross factor comparison of morphological indexes calculated on various segmentation strategies on brain organoids.

Index	Developmental	State	Manual	Kmean	U-Net	Mu-Net
area	9 days	physiological	20416.29***	19902.75	22434**	18381.57
		pathological	15370.15***	21467.14	15189.15**	17277.86
	14 days	physiological	16825.17*	15350.83**	18970.83**	18940.5
		pathological	12038.16667*	42368.2**	13716.66667**	25680.2
	15 days	physiological	13064.4	33943.6	14394.8*	18780.4
		pathological	11986	19841.5	13271.83*	19580.14
aspect ratio	9 days	physiological	1.06	0.78	0.94	0.75
		pathological	0.98	0.82	0.91	0.86
	14 days	physiological	1.16	0.80*	0.91	0.90
		pathological	1.08	1.01*	0.92	0.83
	15 days	physiological	1.08	1.00	1.014*	0.83
		pathological	1.05	0.83	0.96*	0.86
bounding box	9 days	physiological	35161.57**	38368.88	34886.14***	34062.57*
		pathological	25410.86**	32091.71	21927.29***	32669.29*
	14 days	physiological	35050.17**	42823*	36424**	41152.5
		pathological	21094.33**	60240*	25151.83**	39063.2
	15 days	physiological	23557.6	57523	29445.8***	34584.8
		pathological	19067.67	40160.17	22437.83***	33676.71
barycentersx	9 days	physiological	127.87	120.20	125.59	104.39
		pathological	123.85	119.40	127.46	129.19
	14 days	physiological	128.35	131.31	129.89	131.20
		pathological	134.69	126.60	136.90	133.99
	15 days	physiological	129.01	136.42	133.13	137.24
		pathological	124.25	123.37	124.44	129.75
barycentersy	9 days	physiological	123.73	128.51	122.71	103.98
		pathological	128.97	133.32	125.43	127.82
	14 days	physiological	120.65	117.98	120.58	118.29
		pathological	125.21	118.58	124.90	113.72
	15 days	physiological	122.57	123.52	122.60	120.79
		pathological	121.65	112.04	124.01	118.82
compactness	9 days	physiological	0.04	0.15	0.46	0.23
		pathological	0.11	0.15	0.50	0.39
	14 days	physiological	0.05**	0.11	0.37	0.26
		pathological	0.08**	0.25	0.40	0.34
	15 days	physiological	0.08	0.23	0.39	0.30
		pathological	0.09	0.13	0.49	0.28
circularity	9 days	physiological	1430.10	3924.39	9935.04	6224.23
		pathological	1431.60	2843.35	8161.07	7280.34
	14 days	physiological	896.72	2398.97	6616.28	5761.44
		pathological	886.17	21645.66	6822.04	8926.19
	15 days	physiological	938.69	8458.95	6294.26	7350.19
		pathological	1055.77	2536.05	7261.86	5801.28
convexity	9 days	physiological	15.10	23.88	30.46	26.04
		pathological	15.17	17.55	26.48	29.12
	14 days	physiological	14.17	24.74	28.76	32.42
		pathological	12.63	36.32	30.76	31.14
	15 days	physiological	13.30	29.36**	29.65*	27.07
		pathological	13.31	18.41**	27.61*	26.77

Table 5.4: Cross factor comparison of morphological indexes calculated on various segmentation strategies on brain organoids.

Index	Developmental	State	Manual	Kmean	U-Net	Mu-Net
elongation	9 days	physiological	0.82	0.78	0.80	0.69
		pathological	0.88	0.80	0.88	0.83
	14 days	physiological	0.80	0.79 *	0.87	0.83
		pathological	0.80	0.97*	0.82	0.83
	15 days	physiological	0.82	0.97	0.83	0.81
		pathological	0.86	0.83	0.88	0.82
folding	9 days	physiological	802.57	13674	634.57	605
		pathological	599	21425.43	475.29	599.43
	14 days	physiological	898.83	14601.33	686.5	727.17
		pathological	670	21776.8	475.67	786.6
	15 days	physiological	689.6	27237.6	530.4	813.4
		pathological	566.67	21964.33	447.5	769.57
hull	9 days	physiological	29059.71 **	27427.25	24534.14**	23610.57
		pathological	21959**	26583.71	15877.43**	21437.57
	14 days	physiological	28742.33**	29953.83*	24326.17*	29513.83
		pathological	18566.17**	53768.8*	18177.33*	30740.2
	15 days	physiological	20310.4*	43623.4	19128**	25122.4
		pathological	17467.67*	31184	15145.5**	24303.57
perimeter	9 days	physiological	1884	1157.88	788.43	761.29
		pathological	1433.43	1622.43	595	730.14
	14 days	physiological	2031.83	1231	837.67	907.83
		pathological	1473.83	1953.4	584.83	933.4
	15 days	physiological	1534.2	1594.2	656	954.6
		pathological	1312	1737.67	550.5	938.86
sphericity	9 days	physiological	0.34	0.26	0.48	0.29
		pathological	0.49	0.36	0.65	0.45
	14 days	physiological	0.36	0.30	0.46	0.38
		pathological	0.41	0.35	0.48	0.38
	15 days	physiological	0.46	0.26	0.40	0.44
		pathological	0.37	0.26	0.52	0.37
solidity	9 days	physiological	0.69	0.72	0.92	0.68
		pathological	0.70	0.83	0.96	0.83
	14 days	physiological	0.59	0.56	0.79	0.70
		pathological	0.65	0.74	0.85	0.84
	15 days	physiological	0.64	0.75	0.77	0.79
		pathological	0.68	0.71	0.90	0.81
wrinkling	9 days	physiological	3.13	1.52	1.22	1.14
		pathological	2.77	2.04	1.22	1.20
	14 days	physiological	3.87	1.90	1.29	1.17
		pathological	3.12	2.35	1.25	1.33
	15 days	physiological	3.04	1.54	1.01	1.48
		pathological	2.95	2.12	1.06	1.21

5.2.3 Dimensional Reduction

We reduce the dimension of all the calculated variables to observe which one could explain and characterize the dataset, we made a PCA based upon the morphological indexes calculated from shape extracted with an optimized U-Net algorithm. The summary of this PCA is shown in Figure 5.7.

The two first dimensions of the top left graph explain around 55% of the variance which is calculated with eigenvalues. The eigenvalues measure the explained variance quantity by each axis and allow to determine the number of principal components. If we pay attention to the first three columns in the top right correlation plot we observe most of the morphological indexes contribute to these dimensions, which is observed with in strong blue color the highly correlated variables.

These variables are then represented in the circle plot below. These variables are represented according to their correlation (the length of the arrow tendering toward the limit of the circle represents the most explainable variables) and a $cos2$ value (the square cosine of variables which show the quality of variable representation in heat colors). Thus, the solidity, area, hull and bounding box are the most explainable variables, then the sphericity, folding, perimeter, and convexity. We consider not orange to red values to not well characterize the shape of brain organoids.

In the lowest graph are represented each brain organoid position according to the two first dimensional axis. The day of culture is represented by the dotted size, and the pathophysiological state in color. The space point is homogeneously distributed, but, early stages are concentrate in the top of the graph and pathological state at the left and physiological states at the right. However, due to the small amount of data, we could not predict the point distribution of another assertion. We could not classify a new point at some location into a physiological or pathological group (at the top of the graph between -2 and 1) without uncertainty, and we cannot say nothing of the day of culture.

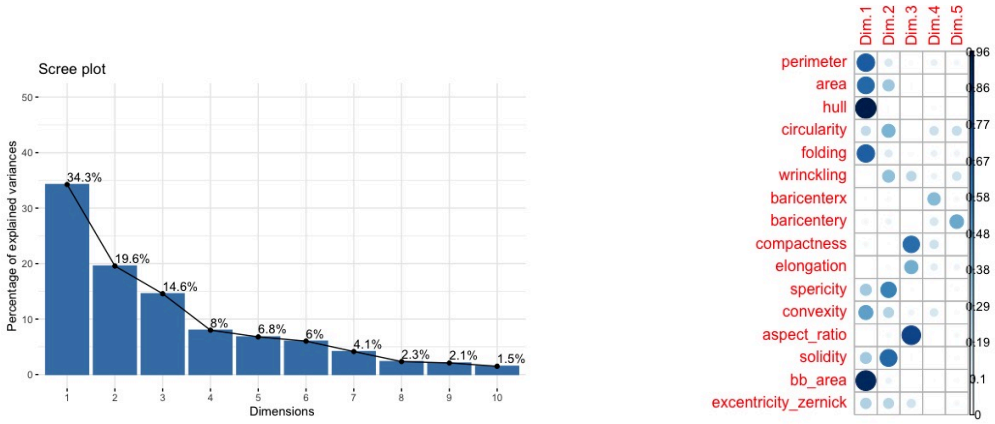
If we observe the dimensional reduction with another strategy (t-SNE), the clustering is similarly not observed, even if a grouping is almost visible by state, but not with the stage see Figure 5.8

Validation on the cortical dataset

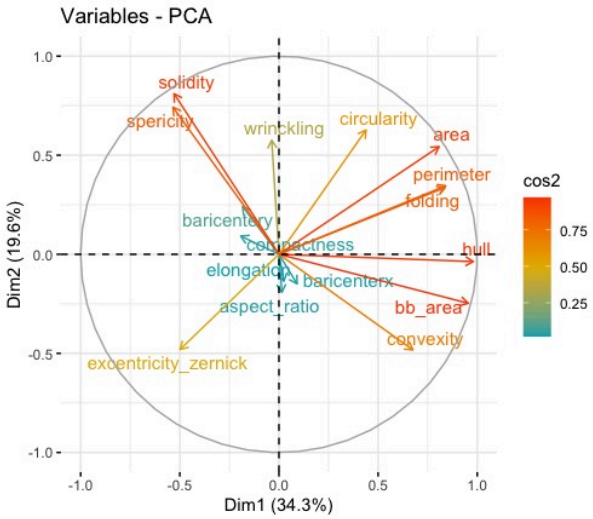
In this part, we validate the morphological index characterisation on a second dataset based on cortical organoids. As previously we calculate morphological indexes on optimized U-Net extracted shapes and then realize a dimensional reduction strategy based upon PCA summarized in see Figure 5.7.

The two first dimensions of the top left graph explain around 50% of the variance which is calculated with eigenvalues. This explainability is a bit lower than for the first dataset. The correlation plot show a few variations with the previous subsection. The most explainable variables are near to the limit of the circle, and in orange colors, and thus are: the perimeter, folding, bounding box, hull, area, circularity, convexity, and wrinkling (which are almost the same as the first dataset). The only two exceptions concern the wrinkling which is here below 0.50 cos2 for the first dataset, and the sphericity, which here does not characterize the dataset, but represents more than 0.50 in cos2.

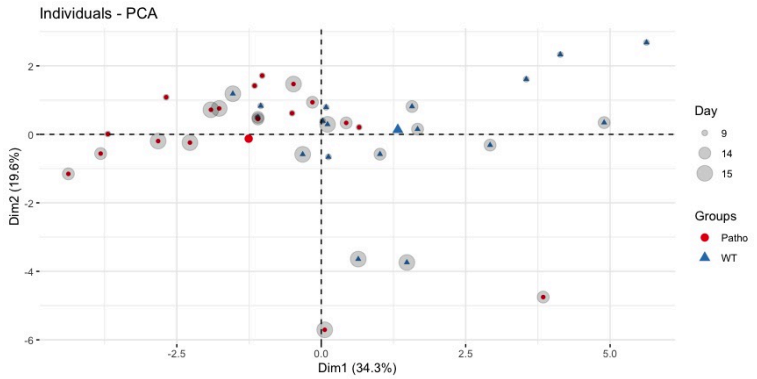
We also observed the distribution of the images according to these morphological indexes scores in dimensional reduction as black points in the bottom right graphic. There is no cluster formation and the data seems homogeneously distributed. However, if we superimpose the distribution of the morphological index calculated on the synthetic images deserved



(a) (b)



(c)



(d)

Figure 5.7: PCA on morphological index calculated on extracted images with an optimized U-Net on a brain organoid dataset. a) Variance of each axis of the PCA. b) Correlation of those dimensions with the variables measured. In darkblue are the strongest correlation. c) Variables explaining the PCA according to the first two dimensions. The best qualitative variable are in red cos2 values. d) Projection of data on the PCA dimension. The stage is represented by the size and state by the color.

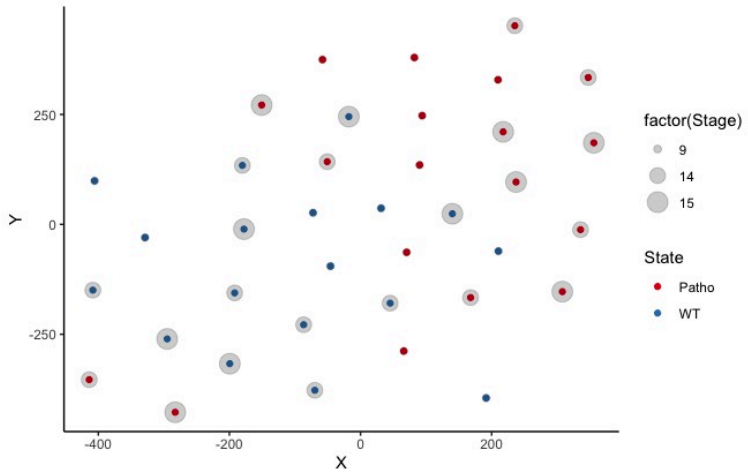


Figure 5.8: Dimensional reduction strategies made upon morphological index calculation with t-SNE for the characterisation of brain organoids.

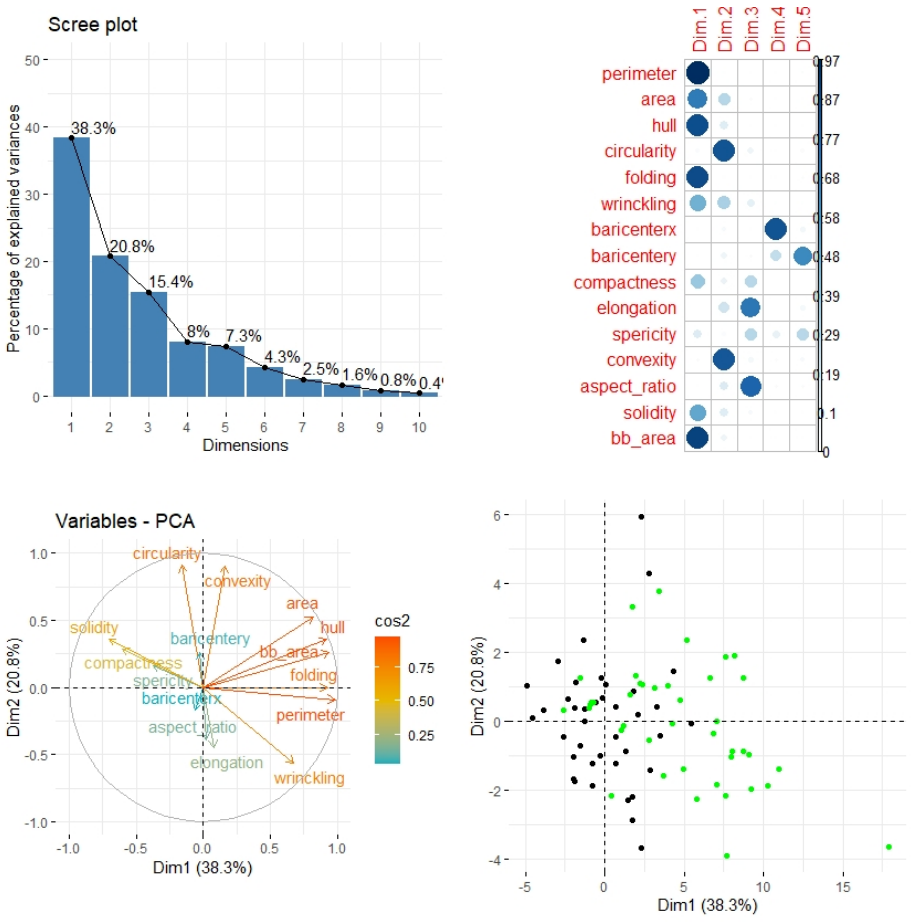


Figure 5.9: Dimensional reduction on morphological index characterisation based upon a U-Net segmentation with an optimized training on a cortical organoid dataset. From left to right: top line represents the correlation plot with the variables explaining each dimension, the second the variances per dimension, the third the visual representation of PCA with in black original images and in green generated which have been used to train U-Net optimized algorithm, and then the variable plot of the PCA.

to train the U-Net algorithm, they are not well superimposing, thus the synthetic images are creating contents with maybe not usual brain organoid shapes.

5.2.4 Discussion

In this part, we conduct morphological index calculation on brain organoid cultures in order to better characterize their morphology, and realize a dimensional reduction strategy to observe if a classification could be made based upon morphological indexes and finally select better the morphological index to characterize these cultures. Some morphological indexes seems better to characterize the morphology of brain organoid than others parameters according to their developmental stage (circularity, convexity, hull, solidity, wrinkling), their physiopathological state (bounding box, folding, hull, perimeter, sphericity, solidity) or the both(area, bounding box, hull, convexity). These indexes best characterize the morphology with an automatic optimized U-Net extraction compared to other manual, semi-automated or another automatic strategy. However, the dimensional reduction strategy shows: no clustering formation appear clearly with a PCA. We could not classify a new point into a developmental stage or a physiopathological state; the synthetic images which deserve to train the extraction have a slightly modified morphology .

The morphological index which best characterize a brain organoid culture is the hull. Indeed, this parameter is a strong tool to decompose the boundary by computing the smallest convex set contained inside the set. This particularity could allow to this parameter an index of a region inside the shape or for the full shape and explain why is found to characterize at the same time a developmental stage, a physiopathological stage and the combination of the both.

Others parameters seem dedicated to a single task: the physio-pathological characterisation or the developmental stage, but not for the both such as the solidity and the wrinkling. The solidity is also link to the convexity a shape, its calculation is a fraction between the area and the convexity. However, the wrinkling is measured as the contour length, normalized by the length of an outer convex contour averaged over all brain organoids. The notion of convexity to characterize the brain organoid state or stage seems of the main importance. If we look at parameters for the both state and stage characterisation, the convexity calculation appear too, with the area, and bounding box.

Some indexes are particular to the developmental stage such as the circularity which decrease with the growth and the formation of neuroepithelial formation, or the folding and perimeter for the physiopathological state which is dedicated to the particular pathology studied here. Thus, it will be of interest to see, in other pathological models of brain organoids, if these parameters appear to characterize their developmental stage.

Another point to discuss is the particularity of the optimized U-Net extraction methodology to extract almost as similarly to the manual segmentation the shape of brain organoids but not the others semi or automated solutions. We think using the best synthetic images, according to their similitude, quality triple-validation, during the training of the extraction, allows to perform an accurate segmentation which renders possible to highlight the same parameters as a manual extraction and more. The difference with the other automatic extraction solutions resides in the smallest architecture, and the slightly less accuracy of the segmentation, particularly for the neuro-eptihelial formations. We think the extractive automated methods could be improved for instance by giving a shape or topological prior information during the synthesis of images used during the training step for instance (El Jurdi et al., 2021). Indeed, we observe during the dimensional reduction the morphological index calculation of these synthetic images are not well superimposed for a cortical organoid

dataset on the original, and thus the generation need to be improved.

Concerning the dimensional reduction, the PCA allows to observe some particular spreading of groups but not sufficient to allow a good classification of a new image. To allow a better clustering strategy we think first of using others parameters not mentioned in the literature review of brain organoid (Brémond Martin et al., 2021b), but described in the Paragraph 5.1.1, such as using signatures, or topological data analysis. By coupling methods, maybe the dimensional reduction will allow to segregate some clustering formations. Also the clustering strategy we used may not be the most suited. Indeed, recently U-MAP or t-SNE based dimensional reduction strategy provide more accurate clusterisation (Hinton et al., 2003; Zu et al., 2022; Devassy et al., 2020). Another solution could be to use classifiers based upon deep learning solutions, as described in Paragraph 5.1.2.

Another question raised by this study concerns the availability of morphological index calculation to describe the shape and stage of brain organoids. Indeed, in the literature of brain organoids, only certain morphological indexes has been used by the neurobiological community. However, we observe previously in Section 5.1, others morphological descriptors could be used in term of boundary (such as signatures or Fourier descriptors), or regional descriptors (textural or topological). Indeed, with a completion of other approaches not used in the neurobiological field, the characterisation could be more accurate and a prediction could be attempted. Another characterisation could be also attempted, with deep learning-based strategies with classifiers (see Section 5.1.3).

In conclusion, the morphological index calculation is an accurate solution to characterize the morphology of brain organoids. However, it cannot help to predict a physiological or pathological state and has to be taken with caution. We now want to verify, if a topological approach could help to characterize and predict the state and stage of these cultures contrary to the morphological index calculation.

5.3 TDA-clustering for brain organoid characterisation

This subsubsection is published in the 2nd Topological Data Analysis MICCAI conference in 2022 Brémond-Martin et al., 2022.

Our aim is to develop a clustering solution to help neurobiologists automatically characterize their images and recognize their developmental stage. We use Topological Data Analysis (TDA) to infer high dimensional structures from low dimensional representations to assemble discrete points into global structures such as persistence diagrams. TDA, increasingly used in the biomedical domain, converts the original data into global topological objects (Skaf et al., 2022).

Previously, brain organoid shapes were extracted from bright-field images using a U-Net segmentation algorithm with various data augmentation strategies (Brémond Martin et al., 2022) (see Section 5.2). Here we propose to characterize the developmental stage of segmented brain organoids images with TDA and TDA-clustering. As a new analysis method for these biological images, we combined TDA with Kmean and SVM to group features within each diagram. We pave the way to identify the developmental stage of these cultures by using TDA as a predictive solution.

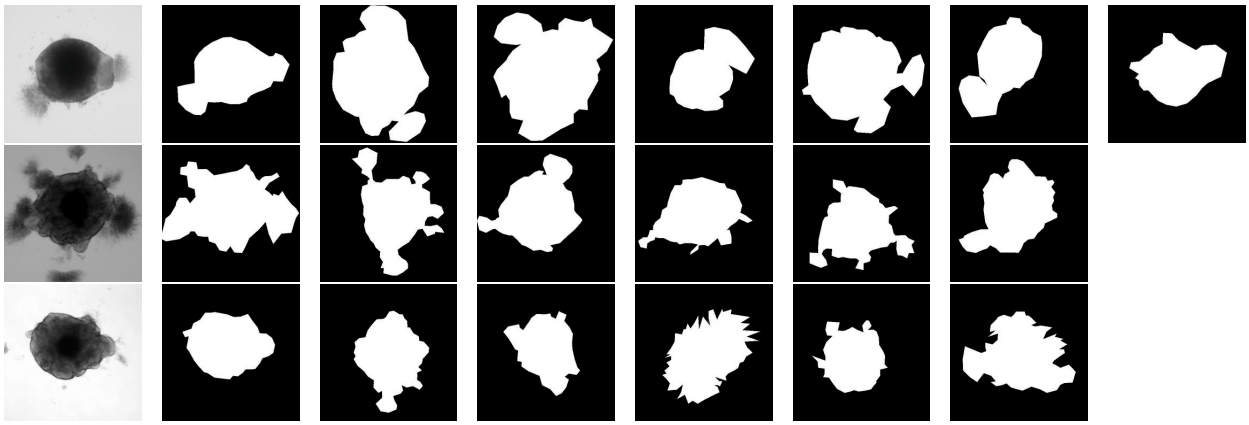


Figure 5.10: Dataset of binary images on which the TDA is performed. A sample organoid from (Gomez-Giro et al., 2019) is given for each developmental stage. All input segmentation masks are shown. First row: 9 days. Second row: 14 days. Third row: 15 days.

5.3.1 Methods

Resources We consider 19 binary images (250 × 250 pixels), shown Fig. 5.10. They result from the segmentation of the 19 bright-field images of healthy brain organoids from the largest public dataset available (Gomez-Giro et al., 2019). Images are acquired at 9 days (7 images), 14 days (6 images) and 15 days of culture (6 images). These images were segmented with a U-Net framework trained with a classical data augmentation strategy (Ronneberger et al., 2015a; Brémond Martin et al., 2022). All scripts are written in python.

TDA We based our implementation on the Ripser algorithms as it outperforms others codes in computation times and memory efficiency and proposes a Vietoris–Rips filtration step (Bauer, 2021; Tralie et al., 2018). We perform TDA on each image by first computing the Persistence Homology (PH). PH maps a set of points P with a distance function F to retrieve topological features in segmented images.

To summarize this approach, a $S_d(P)$ space is defined by all points within the distance d (a disc centered in P with a d radius) from a point p in P. A connection is equal to two discs overlapping in the same topological space. If d is small, the number of connected components (C) in the space is equal to p in P, while if d increase, C is equal to one. Thus, for a given d, the topological feature vector associated contains shape information and the aim is to study them (Edelsbrunner et al., 2002; H. Edelsbrunner et al., 2008; H. Edelsbrunner et al., 2010).

A barcode encodes the topological feature evolution. The beginning of a bar represents the "birth" of a feature. All zero-dimensional features are "born" at zero and their length represents the non-static nature of the clustering. The largest code is the most persistent feature. The connected components are sorted by increasing d value. Lower d correspond to high components similarities (H. Edelsbrunner et al., 2010; Ghrist, 2007).

Persistence diagrams describing data structures can also be used to compare PH. A point in this diagram represents a topological feature according to its' birth (abscesses) and death (ordinate). We plot zero-dimensional and one-dimensional topological features, respectively $H0$ and $H1$. H0 represents the connected components of the point cloud; H1 corresponds to holes and informs about dynamics (Otter et al., 2017). We compare the PH of brain organoid images at different developmental stages of these cultures.

Clustering feature vectors We cluster the feature vector points from the persistence diagrams (B. Yang et al., 2017a). Our goal is to observe if H0 and H1 feature clustering are identical; if some clusters, representing the structure of brain organoids are linked to each other; if the clustering varies according to the developmental stage. We implement clustering both with Kmean (Panagopoulos, 2022) and with a SVM (García et al., 2007) and obtained identical clustering in all cases. We thus refer to the algorithm generically as "the clustering algorithm". To automatically select the number of clusters k, we calculate elbow points on inertia curves (N.-K.-K. Nguyen et al., 2021): we retrieve the distance between each data point and its centroid, squaring this distance, and summing these squares across a cluster for all the dataset. k equals to 4 for H0 and H1 at each developmental stage.

Quantitative comparisons To compare the persistence diagrams we first perform a linear regression on the H1 point clouds (Baas et al., 2020). Then we compute the average point cloud entropy (Bouleux et al., 2019), dispersion (Turner, 2020), and persistence (N.-K.-K. Nguyen et al., 2021) for each developmental stage and for each cluster within each developmental stage. We also compute the pairwise Wasserstein distance between developmental stages and between clusters (Berwald et al., 2018).

We compare each feature vector in the same statistical space with a t-distributed Stochastic Neighbor Embedding (t-SNE). t-SNE preserves the local dataset structure by minimizing the divergence between two distributions respective to the map point locations (Hinton et al., 2003). An asymmetric probability based on dissimilarities is calculated between an object and its probable neighbor (Hinton et al., 2003). We choose the effective number of local neighbors as the number of clusters points. To match distributions in a low dimensional space, we aim at minimizing a Kullback-Leibler (KL) cost function such as in Maaten et al., 2008. One-degree freedom Student t-distribution avoids the crowding issue (Maaten et al., 2008). We reduce iterations with momentum. To compare feature vectors representation, we calculate the KL divergence between groups.

5.3.2 Results

Qualitative Global barcodes and persistence diagrams for each stage of brain organoid cultures are shown Fig. 5.11. For each developmental stage we show a global representation, instead of sample barcodes and persistence diagrams corresponding to an individual input image. Barcode patterns differ between each developmental stage. Bars at 9 days are shorter than from 14 and 15 days. This reveals a most persistent H1 feature of the point cloud at later stages, whereas H0 features are similar between groups. H0 and H1 features of persistence diagrams seems less dispersed at 9 days than for later stages.

To further compare feature vector distributions between groups we combine TDA with clustering. H0 and H1 clustering seems different in term of point proportions and positions according to the stage of culture. To quantify these observations, we calculate several metrics on H0 and H1 features.

Quantitative In this section all calculations are made on the individual feature vectors for all images and rendered as a mean. We calculate the linear regression on the H1 feature diagrams. Tab. 5.5 shows the slope and intercept of these linear regressions, as well as the entropy, dispersion, and average persistence for each developmental stage. We also compute the pairwise average Wasserstein distance between features for each developmental stage.

The linear regression for H1 at 9 days is markedly different from that at 14 and 15 days. The entropy for H1 at 14 days is slightly higher than for 9 and 15 days. All other

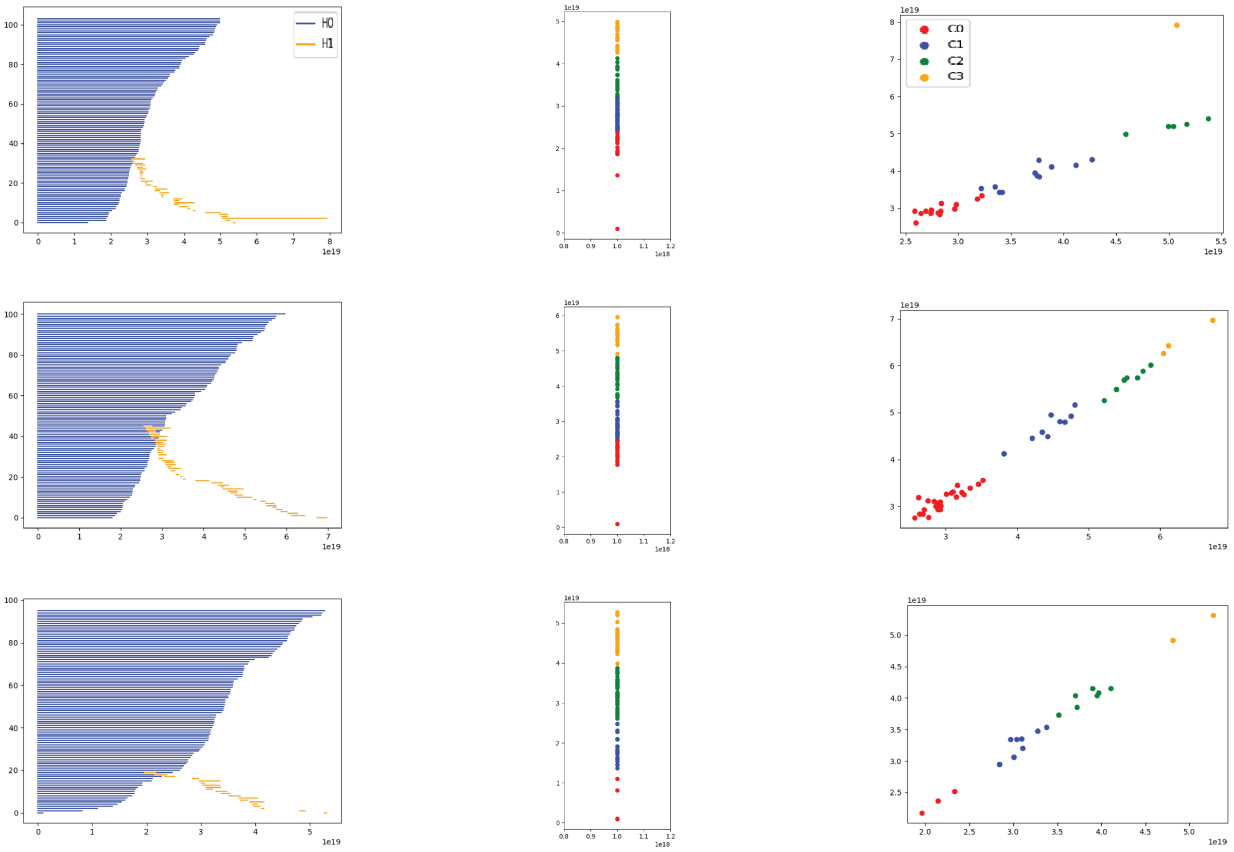


Figure 5.11: Average representation for each developmental stage of barcode and H0 and H1 persistence diagrams. First row: 9 days. Second row: 14 days. Third row: 15 days. For this figure, average barcode and persistence clustering diagrams are calculated from the individual representations for each input image.

Table 5.5: Slope and intercept from the linear regression; entropy, dispersion and average persistence for each developmental stage; pairwise average Wasserstein distance between each developmental stage.

		Linear reg.		Entropy	Dispersion	Average	Wasserstein distance		
		Slope	Intercept		($\times 10^{34}$)	persistence	9	14	15
H0	9	–	–	47.92	9.82	5.38×10^{17}	561 222	579 195	609 972
H0	14	–	–	47.93	10.07	5.44×10^{17}	–	597 156	627 910
H0	15	–	–	47.93	10.23	5.48×10^{17}	–	–	658 625
H1	9	0.6	704	10.46	9.82	8.70×10^{2}	633	656	632
H1	14	1.1	304	12.26	10.69	8.66×10^{2}	–	672	651
H1	15	1.4	208	9.31	10.23	8.26×10^{2}	–	–	617

Table 5.6: Point distribution across clusters and entropy within clusters for each developmental stage for H0 and H1.

	Points (%)						Entropy					
	H0			H1			H0			H1		
	9	14	15	9	14	15	9	14	15	9	14	15
C0	50	51	46	42	39	27	10.33	10.34	10.23	9.72	9.39	10.03
C1	1	1	1	13	13	24	44.21	44.21	44.21	4.40	2.56	4.21
C2	18	21	20	27	27	17	9.41	9.89	9.51	5.78	5.41	4.64
C3	31	27	33	18	21	32	9.98	10.01	9.96	6.18	3.20	4.83

parameters are quite close within H0 and H1 for all three developmental stages. We calculate the average Wasserstein distance between all points in every pair of developmental stage. For H0, distances are greater for 15 days than for 9 or 14 days. For H1, 14 days has the biggest distance with others groups and within itself.

Table 5.6 shows the point distribution across clusters (in percentage) and the entropy within clusters for each developmental stage for H0 and H1. For H0, the entropy is greater for C1 clusters and lower for clusters further from the origin, for all three groups. For H1, the entropy is greater for C0 clusters and lower for C1 clusters. For both H0 and H1 features, the percentage of points in a persistence diagram constituting a cluster is largest at C0 for the earliest stages and the lowest for C1. The differences are attenuated at the latest stage for H1.

We calculate the average pairwise point to point Wasserstein distance between clusters to them within a persistence diagram, see Tab. 5.7. For H0 features, there is no pair of feature that is consistently the furthest or the closest over time. However, the average distance between clusters is greatest at 14 days. For H1, the distance between clusters is strikingly bigger than that within clusters. The distances are also greater at 14 days, especially for C3.

Table 5.8 shows the Wasserstein distances within each cluster for each developmental stage. The distance is greatest at 15 days for H0 and at 14 days for H1. To analyze the variability and potential prediction of an image, we compare persistence feature vectors in the same statistical space using t-SNE, see Tab. 5.9. For H0, distances between and within groups are near. For H1, the distance between 14 and 15 days is markedly greater than all other distances.

Figure 5.12 shows the KL divergence between developmental stages and clusters. For H0, the divergence between clusters of different groups is much bigger than that within a

H0	C0	C1	C2	C3	H1	C0	C1	C2	C3
9 - C0	43.54	44.99	54.94	52.31	9 - C0	45.49	183.15	561.63	870.68
9 - C1	–	38.16	47.61	50.55	9 - C1	–	55.78	478.48	687.54
9 - C2	–	–	54.75	58.94	9 - C2	–	–	31.31	488.46
9 - C3	–	–	–	58.73	9 - C3	–	–	–	53.20
14 - C0	71.84	74.85	70.74	74.91	14 - C0	53.86	235.34	520.37	965.67
14 - C1	–	73.80	70.14	72.68	14 - C1	–	51.85	401.07	750.29
14 - C2	–	–	66.15	69.20	14 - C2	–	–	48.66	531.16
14 - C3	–	–	–	71.28	14 - C3	–	–	–	71.50
15 - C0	50.09	59.48	30.65	44.81	15 - C0	47.43	137.96	546.72	706.30
15 - C1	–	61.85	37.19	53.85	15 - C1	–	43.98	408.77	568.35
15 - C2	–	–	18.03	26.50	15 - C2	–	–	35.09	439.85
15 - C3	–	–	–	37.97	15 - C3	–	–	–	55.16

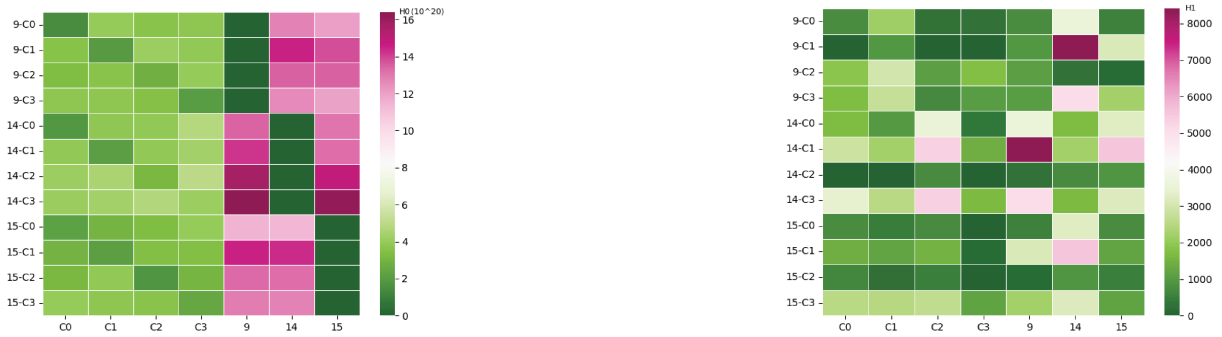
Table 5.7: Average Wasserstein distance between clusters.

H0 ($\times 10^{19}$)	9	14	15	H1 ($\times 10^{19}$)	9	14	15
C0	0.990	0.263	3.038	C0	1.056	2.892	2.249
C1	1.395	0.929	4.577	C1	1.471	3.780	3.186
C2	1.838	1.646	5.628	C2	2.146	6.497	3.922
C3	2.285	2.271	6.426	C3	2.671	5.123	5.080

Table 5.8: Wasserstein distances within each cluster for each developmental stage.

H0 ($\times 10^{19}$)	9	14	15	H1	9	14	15
9	2.83	2.81	2.75	9	3886	1215	2648
14	–	3.23	2.85	14	–	3682	5741
15	–	–	2.76	15	–	–	2556

Table 5.9: KL divergence between developmental stages for H0 and H1 feature vectors reduced by t-SNE.



(a) (b)

Figure 5.12: KL divergences between developmental stages and clusters for H0 (a) and H1 (b) feature vectors reduced by t-SNE.

group. For H1 the biggest divergences are for certain clusters at 14 days.

5.3.3 Discussion

Our aim is to automatically characterize brain organoid morphology according to the step of culture using TDA-clustering. The TDA pattern of brain organoids and their feature clusters evolve according to the developmental stage. The persistence diagram of early cultures approaches a right line characterized by a lower slope and higher intercept than that of later cultures. The clustering strategy strengthens the characterisation of developmental groups by showing differences in distances and entropy in inter-intra groups. Dimensional reduction on clustered feature vectors confirms the particularity of the 14th day.

There is an evolution in the persistence diagram over the culture development which could be linked to neuroepithelial formation, a crucial event that occurs before 14 days of culture (Stefano L. Giandomenico et al., 2021). Our goal is to find a single representation of images by persistence diagram that characterize this formation which is usually determined by the calculation of multiple morphological indexes (Brémond Martin et al., 2021b).

To verify these qualitative results, we calculate parameters currently used to compare persistence diagram representations (N.-K.-K. Nguyen et al., 2021; Turner, 2020; Bouleux et al., 2019; Baas et al., 2020; Berwald et al., 2018). All parameters, except the dispersion and persistence average, seem to support the qualitative results. The slope and entropy increase with later stage while the intercept decreases. The Wasserstein distance between feature points representations evolves between early and later stages. Another calculation of dispersion and persistence average, based upon the difference of distribution inside cluster proposed in N.-K.-K. Nguyen et al., 2021, could be used to see if such difference persist in future studies.

To further analyze the feature point distributions of persistence diagrams from brain organoid images, we observe TDA with two clustering methodologies (B. Yang et al., 2017b; García et al., 2007). Instead of applying clustering on the overall feature vectors (N.-K.-K. Nguyen et al., 2021), we segregate H0 and H1 vectors. As SVM render same clustering as Kmean, we calculate the parameters on clusters from each persistence diagrams once. Analysis of feature cluster representations at various days, allows to further characterize the brain organoid morphology pattern. Some clusters are characterized by extreme metric values. The relation between the clusters and the morphological characteristics of the segmented images must still be explored.

The groups clusters comparison, based upon feature vectors constituting the diagrams, needs to be further studied. We could also try other dimensional reduction methodologies or, propose a better classification of feature clustering, based on a Random forest classifier on persistence images (Frahi et al., 2020). Another topic to explore is the comparison of TDA-clustering strategies with classical shape analysis on brain organoid cultures (index calculation such as perimeter, surfaces...) or others strategies usually calculated on biomedical images (Brémond Martin et al., 2021b; Mingqiang et al., 2008; Thibault et al., 2013). When TDA-clustering strategy will be validated on other brain organoid image datasets, such automatic methodologies could help follow the growth of these cultures, predict their developmental stage, or compare physio-pathological models (Brémond Martin et al., 2021b).

We propose using TDA-clustering to characterize the morphological development of brain organoids. TDA-clustering allows us to highlight some patterns verified by the calculation of parameters on persistence diagrams and this both with Kmean and SVM clustering. The evolution of the TDA representation seems to appear before 14 days of culture and, could be linked to neuroepithelial formations. t-SNE analysis applied on TDA-clustering feature vectors suggests such strategy could help the developmental stage prediction of segmented images. Future work will include putting these clusters in relation with the corresponding morphological characteristics of the segmented images. We need to verify the physio-pathological characteristics before exploring this perspective.

5.4 TDA-clustering for the physio-pathological classification of brain organoids

This section describes unpublished work.

The objective of the work presented in this section is to pave the way to compare physiological and pathological brain organoids images using TDA and TDA-Clustering. We propose to ally TDA with clustering in order to intent classify features into low representation as coming from a physiological or a pathological content. We propose to use SVM or Kmean clustering strategies.

Contributions are many-folded: we propose TDA for the characterisation of physiological and pathological brain organoid models; We observe if the TDA-Clustering by a Kmean methodology allow an accurate characterisation of brain organoid shape; We propose to classify images into physiological or pathological content by a TDA-Clustering-Predictive novel strategy.

5.4.1 Methods

Resources

Our dataset contains 38 binary images (16 physiological and 16 pathological) of 250×250 pixels. We process a U-Net with a classical data augmentation strategy on the (Gomez-Giro et al., 2019) to obtain this dataset, such as described in Section 4.2.

Topological Data Analysis (TDA)

The TDA strategy is the same as described in the previous Section 5.3.1.

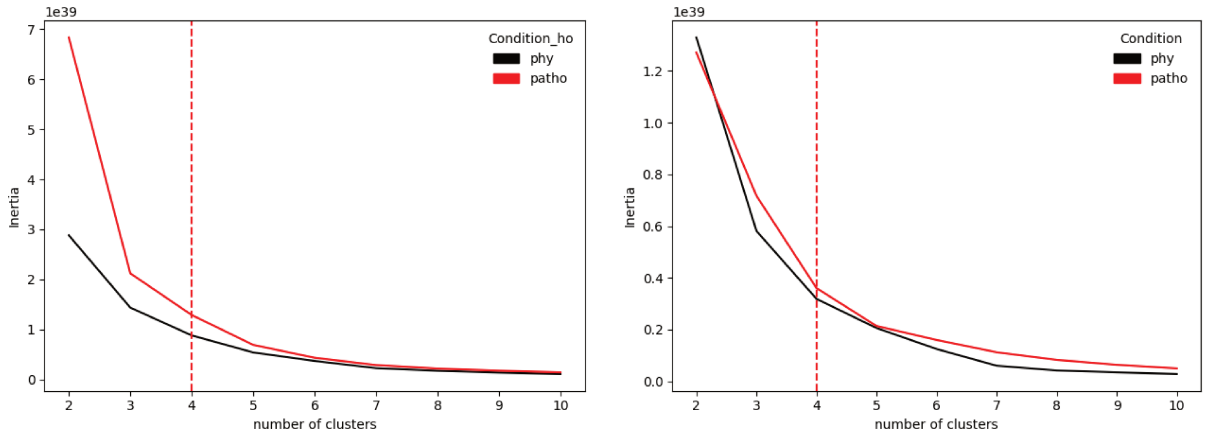


Figure 5.13: Inertia according to the number of clusters and their elbow point highlighted in dotted red line.

Clustering

We cluster feature vector points in persistence diagram (B. Yang et al., 2017a) and observe: if H0 or H1 feature clustering are identical (if some C, structuring the content of brain organoids, are linked to each others); if between the two states clustering are identical (if their diagrams change with the model). To do so we propose to combine the TDA with a Kmean algorithm (Panagopoulos, 2022). To initialize the number of clusters k automatically, we calculate the elbow points (the point where the inertia begins to slow) (N.-K.-K. Nguyen et al., 2021). We retrieve the distance between each data point and its centroid, squaring this distance, and summing these squares across one cluster for all the dataset D: $inertia = \sum_{i=1} D(x_i - C_k)^2$ (N.-K.-K. Nguyen et al., 2021). We plot inertia relative to the number of clusters for each feature vectors. Thanks to the curves in Figure 5.13, we find k = 4 for Kmean.

Characterisation

We compare the diagrams, by calculating five parameters within each cluster (by SVM or Kmean) or within each entire feature vector: the average persistence (N.-K.-K. Nguyen et al., 2021), the point dispersion (Turner, 2020), the entropy (Bouleux et al., 2019), the slope and intercept (Baas et al., 2020), the Wasserstein Distance (Berwald et al., 2018). Results are rendered as mean for each parameter per group or cluster per group.

Prediction

To predict the belonging group of persistence diagram of brain organoid images, we compare each feature vector in the same statistical space thanks to t-distributed Stochastic Neighbor Embedding (t-SNE). This method preserves the local dataset structure by minimizing the divergence between the two distributions respectively to the map point locations (Hinton et al., 2003). An asymmetric probability based on dissimilarities is calculated between an object and its probable neighbor (Hinton et al., 2003). We choose the effective number of local neighbors, as the maximal number (the number of clusters points for each groups in our dataset). To match the distributions, in a low dimensional space, we aim at minimizing a Kullback-Leibler cost function. One-degree freedom Student t-distribution avoids the crowding issue (Maaten et al., 2008). We use only 250 iterations thanks to a momentum.

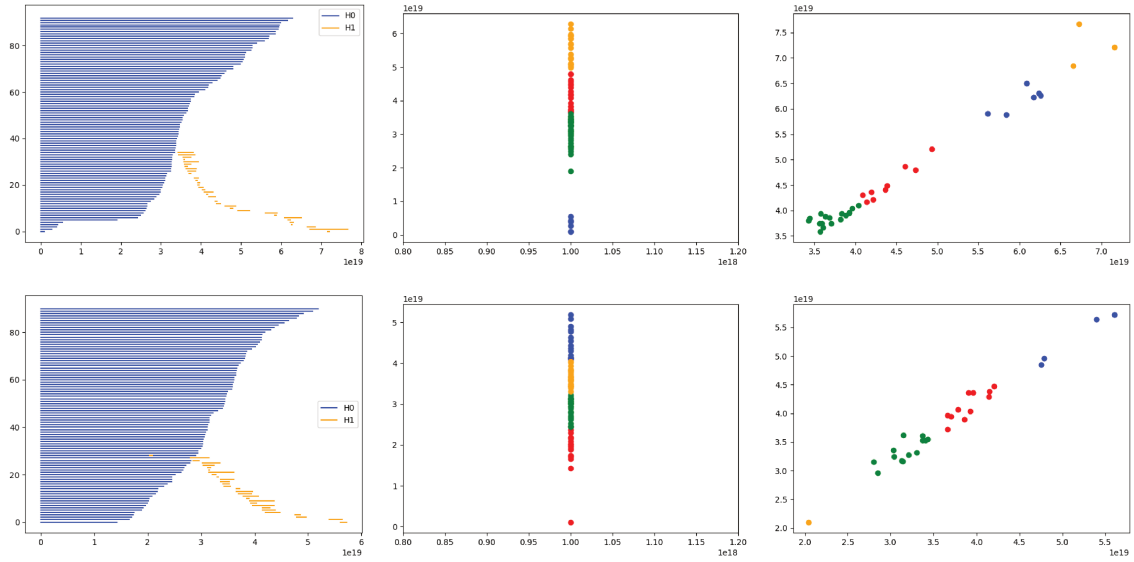


Table 5.10: Barcode, Persistence Diagrams and TDA-Clustering for Physiological and Pathological Images

We compare feature vectors representation from groups calculating KL divergences.

5.4.2 Results

To compare physological and pathological contents of segmented images, we first observe qualitatively the barcode and persistece diagrams, then we quantitatively compare parameters calculated on those diagrams, and then predict the state of images.

Qualitative

Average barcodes and persistence diagrams representations of each barcode and persistence diagrams from H0 and H1 feature vectors are represented in Figure 5.10. We could observe average barcodes for physiological images are more narrowed than for pathological images. H0 and H1 feature vectors representations are more dispersed for physiological than for pathological images. Point proportions seem identical for H0 and H1 vectors in the case of physiological image, however, C0 clusters seems less populated than other clusters and more dispersed in the case of pathological images.

To verify the qualitative observations on barcodes and persistence diagrams from H0 and H1 feature vectors, we calculate metrics on persistence diagrams.

Quantitative

Summary of calculated metrics on the overall persistence diagrams are shown in 5.11. We first calculate a linear regression on H1 persistence diagrams of each image from the two conditions to measure the slope and the intercept. Slope is smallest for pathological images and the intercept is higher. Persistence of H1 feature vectors is similarly lower for pathological images. For others parameters, there is no differences observable.

To go further in comparisons, we calculate metrics in intra-cluster persistence diagrams, see Table 5.12. Point proportions in each cluster seem higher in C0. Only a few points are clustered in C1 and C2 and C3 have similar proportion of points whatever the considered feature vector and group. Dispersion are similar between physiological and pathological

	Physiological	Pathological
Slope	1.3	1.3
Intercept	−121.1	−80.6
Dispersion-H0 (10^{17})	5.0	5.0
Entropy-H0	47.9	47.9
Persistence-H0 (10^{18})	1.0	1.0
Dispersion-H1	543.3	484.6
Entropy-H1	14.3	14.1
Persistence-H1	1086.6	969.2

Table 5.11: Phy and Pathological metric calculation on persistance diagrams.

	Physiological				Pathological			
	C0	C1	C2	C3	C0	C1	C2	C3
Points-H0 (%)	0.5	0.0	0.2	0.3	0.5	0.0	0.3	0.3
Dispersion-H0 (10^{18})	8.3	3.3	8.3	8.3	8.3	3.3	8.3	8.3
Entropy-H0	10.3	44.2	9.9	10.0	10.3	44.2	9.6	10.0
Persistence-H0 (10^{18})	1.0	2.0	1.0	1.0	1.0	2.0	1.0	1.0
Wasserstein-H0 (10^{19})	1.5	1.6	2.1	2.7	0.9	1.5	1.9	2.3
Points-H1 (%)	0.4	0.3	0.1	0.2	0.4	0.2	0.2	0.2
Dispersion-H1 (10^{3})	5.5	22.3	6.5	20.0	4.8	8.5	14.8	22.4
Entropy-H1	7.1	1.4	5.2	1.7	7.3	4.5	1.6	7.2
Persistence-H1	1247.7	1768.8	1542.6	1481.3	1034.2	1951.5	1488.2	1325.4
Wasserstein-H1 (10^{19})	3.8	4.5	6.1	7.0	2.1	3.3	4.0	5.2

Table 5.12: Intra-cluster physiological and pathological calculations

groups except for H1 where C1 and C2 dispersion are different. For the pathological group, dispersion is higher in clusters further from the origin. Entropy is similar for H0 features, while for H1 foreign clusters are different between physiological and pathological groups. Persistence average is similar between clusters of the 2 groups from H0 feature vectors. The H1-C1 persistence is higher to others clusters whatever the group consider. Others clusters persistence values are higher in the physiological group than for the pathological group. In the physiological group, Wasserstein distance intra-cluster is lower in nearest clusters than further clusters. Intra-cluster Wasserstein distance seems higher in the physiological group than for the pathological group, whatever the feature vector considered.

To observe more particularly the inter-cluster Wasserstein distance in same persistence diagrams and between clusters from the other group, see Tab. 5.13. For H0, the inter-cluter Wasserstein distance is greater for C0 between two diagrams from physiological images. For pathological images, is between C3 and C0. Between two identical clusters of physiological and pathological, C3 is the cluster with a maximal distance.

For H1 inter-cluster Wasserstein distance, C3 with all others clusters has the biggest distance for physiological diagrams and C0 for pathologicals. Physiological and pathological distances from same type of cluster are biggest than if clusters are compared from the same group, except for C2.

To observe the grouping of features, to pave the way of a tool to predict the group, we realize a dimensional reduction and calculate KL divergences on the rendered representation, the summary of these calculations are shown in Table 5.14 and in Figure 5.14.For H0, the

H0	C0	C1	C2	C3
phy-C0	81.3	73.4	75.0	76.7
phy-C1	–	47.5	61.9	67.0
phy-C2	–	–	66.2	69.2
phy-C3	–	–	–	71.3
path-C0	65.0	69.8	73.3	68.6
path-C1	–	60.8	54.7	60.4
path-C2	–	–	38.5	47.1
path-C3	–	–	–	54.3
H1	C0	C1	C2	C3
phy-C0	479.54	561.03	506.39	687.67
phy-C1	–	572.13	582.54	680.24
phy-C2	–	–	432.02	700.30
phy-C3	–	–	–	693.56
path-C0	488.27	412.71	408.20	458.53
path-C1	–	260.59	254.23	330.78
path-C2	–	–	172.91	319.40
path-C3	–	–	–	379.70

Table 5.13: Inter-cluster and inter-group Wasserstein distance comparisons.

H0 (10^{19})	Physiological	Pathological
Physiological	2.70	2.83
Pathological	3.28	3.12
H1	Physiological	Pathological
Physiological	3317.9	3628.1
Pathological	2830.3	2568.4

Table 5.14: KL divergences between groups on overall representations after a t-SNE reduction

KL divergence in intra-cluster is higher between foreign clusters, and particularly for the C2. For H1, intra-cluster KL divergence is greater for C2 clusters and for C0 in the case of physiological diagrams only. If we compare Physiological and Pathological clusters, the greatest difference appears between phy-C1 and patho-C1 as well as between phy-C3 and patho-C3.

5.4.3 Discussion

An important issue in the brain organoid field is the morphological comparison of two of these models knowing they are suffering from batch syndrom, and there is no dedicated tools (Kelava et al., 2016a). Here we propose to use TDA-clustering methodology to answer this problem. Physiological TDA brain organoid morphology is characterized by the largest dispersion and persistence in H1 feature vectors compared to a pathological model and some metric values are higher in specific clusters and particularly the persistence, dispersion

(a) (b)

Figure 5.14: Intra-cluster KL divergence and between groups

and Wasserstein distance (for the foreigner cluster). Dimensional reduction applied on the resultant feature vectors shows also a particular signature in the physiological group.

Resulting persistence diagrams are different between physiological and pathological brain organoid models particularly marked with a different dispersion of point clouds but not for slope and intercept. However, previous work show the slope and intercept evolve according to the culture stage (Brémond-Martin et al., 2022). It would be interesting to observe the TDA pattern of a concomitant culture condition and the culture stage to strengthen results.

Physiological images have a higher dispersion and persistence value which could be linked with they higher size area and development such as described in Gomez-Giro et al., 2019. The benefit of TDA methodology contrary to morphological index calculation on the overall segmented image are: its non subjective aspect, we do not choose the metric to compare the two models, because it is based upon features directly; a single representation can highlight the difference, which is not the effect with morphological index calculation; we do not need a combination of morphological indexes to characterize the two models; information missed by traditional morphological calculation is uncovered by TDA (Skaf et al., 2022). However, the TDA rendering time is higher to process than a morphological index calculation. The comparison of each diagrams and feature vectors need to be lightened.

The clustering approach allow us to more specifically compare pathological and physiological segmented brain organoid images. Indeed, physiological segmented image feature topological vectors clusters are represented by a higher dispersion, persistence and distance than for pathologicals. Particularly the nearest origin cluster and the furthest have specific values in the two groups. In future studies, we need to retrieve the part of the segmented image with each feature cluster in order to link which kind of part is specific in these groups, such as suggested in Peláez et al., 2021.

The clustering methodology seems to not have an impact on the clustering representations, groupings are identical with Kmean and SVM. Another clustering methodology could be used to make our approach more robust. We could also try a Random forest classifier on persistence cluster images such as (Frahi et al., 2020).

The t-SNE KL divergence representation, show us physiological representations are more variables than for pathological, but physiological and pathological representation are diverging. Another idea is to try to predict the belonging of an image at a group by making a predictive analysis on t-SNE dimensional reduction (Bois et al., 2022). To further study physiological and pathological images we could also combine a morphological development study with the physiological and pathological condition. Maybe during the development, their pattern are also evolving.

Based on persistence diagrams, we can segregate physiological from pathological brain organoid segmented images. Clustering approach helps at strengthening their differences and a dimensional reduction show physiological images pattern are more diversified. This study paves the way for an automatic morphological analysis of brain organoids based on indexes. We also aim to analyse at the same time the state and stage of brain organoids with TDA.

5.5 TDA cross-factor clustering for brain organoid characterisation

This section describes unpublished work.

In the previous sections we characterize the topology of developmental stage or the physiological or pathological state of a brain organoid culture. However, the topology may be influenced by the two factors as the same time which as not been studied until then.

In order to verify the effect of the developmental stage (9,14 or 15 days) and the state (physiological or pathological) of brain organoid cultures, we aim at producing a cross factor analysis of their topology.

The methods are the same as described in the two previous Sections 5.3 and 5.4.

5.5.1 Results

Cross factoring of the state (physiological or pathological) and the developmental day (9, 14 or 15 days) of culture are confronted with TDA strategies.

The barcode for each combination are represented in Figure 5.15. For physiological H0 and H1 features, there is a physiological spreading with the developmental evolution. In H0 the spreading seems curved. While for pathological stage a tightening is observed with the development.

On the persistence diagrams, these tightening and spreading are observed too Figure 5.16. Points representation are spreading with the developed stages in physiological cases while the centering of points representations are observed in the pathological case.

To verify the assumptions observed previously we calculate parameters on the overall persistence diagrams for each image and average it for each combination of conditions which are summarized in Table 5.15. We first describe the H1 diagram parameters. The slope increase with the developmental stage whatever the state. However, the slope value is higher in a pathological state. Concerning the intercept, entropy and average they are at their maximum for the earliest physiological state, and the maximum is reached by the pathological state for the Wasserstein. The 14 day whatever the state obtain their lowest values for these four parameters. For the H0, the physiological state obtains the maximal values particularly for the 14 days, even if the differences remains slight.

In Table 5.16, we pay attention to intra-cluster parameters calculation. The earliest developmental stage obtains the higher values and more particularly the 2nd cluster whatever the state. The entropy increase with the cluster number also whatever the developmental stage and state. The average and H1 entropy values are higher in the pathological groups than in physiological. This is the only difference between the two states.

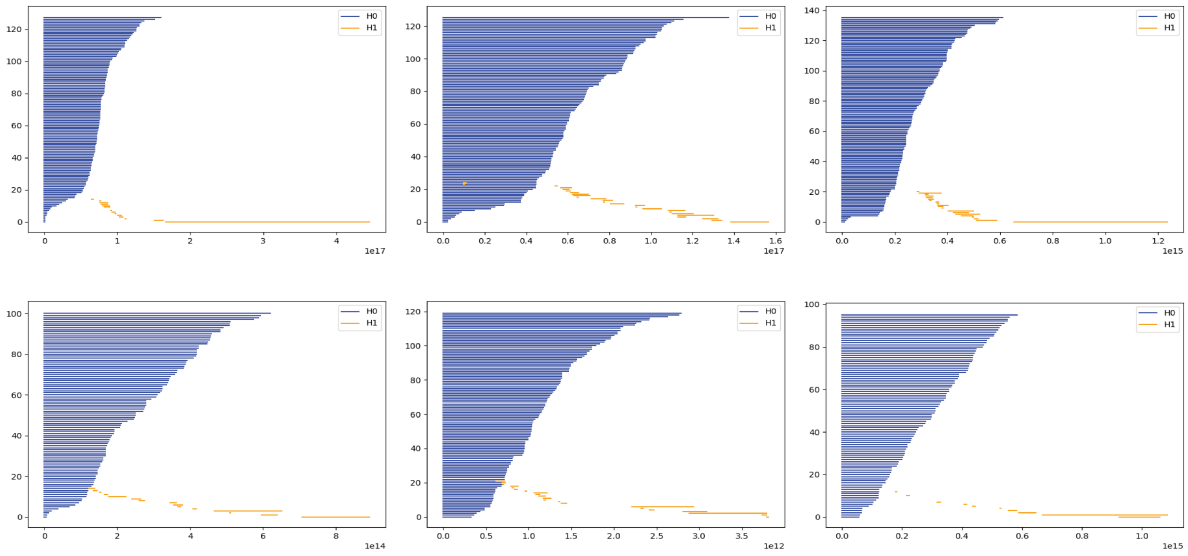


Figure 5.15: Average barcode representation of brain organoid images at each developmental stage (from left to right 9, 14 and 15 days) and for each state(physiological is the top line and pathological the bottom line).

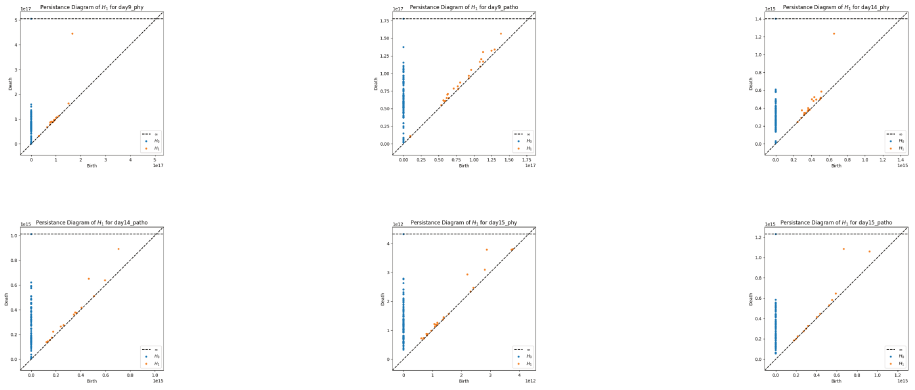


Figure 5.16: Average persistence diagrams representation of brain organoid images at each developmental stage (from left to right 9, 14 and 15 days)and for each state(physiological is the top line and pathological the bottom line).

state	day	slope	intercept	entropy	average	dispersion	Wasserstein
H1							
phy	9	0.23	**1027.52**	**28.12**	**8.55**	44.70	185182.7
phy	14	0.76	655.62	2.22	1.02	30.82	568.05
phy	15	1.38	1.25	16.71	6.37	13.94	3.32
patho	9	0.99	381.12	13.80	7.76	11.03	**107904.3**
patho	14	1.44	47.83	2.29	6.47	**168.14**	487.77
patho	15	**1.43**	380.17	26.26	7.71	49.08	8.43
H0							
phy	9	-	-	9.85	421.31	5.0	**33.4**
phy	14	-	-	**9.93**	**497.86**	5.0	6.40
phy	15	-	-	9.84	409.58	5.0	7.24
patho	9	-	-	9.69	427.06	5.0	30.7
patho	14	-	-	9.76	374.55	5.0	8.25
patho	15	-	-	9.83	373.92	5.0	8.63

Table 5.15: Parameters calculated on the entire persistence diagrams.

			H0				H1			
state	day	cluster	entropy	average	dispersion	Wasserstein	entropy	average	dispersion	Wasserstein
phy	9	0	10.28	-	5.0	97311.16	10.20	1059.23	82.79	45.49
phy	9	1	**44.22**	1.0	< 0.01	312539.3	3.21	1112.77	81.38	55.78
phy	9	2	9.92	**496.02**	5.0	**608212.7**	9.73	**1518.36**	87.86	31.31
phy	9	3	9.37	94.61	5.0	< 0.01	9.70	1324.95	81.78	53.20
phy	14	0	10.39	-	5.0	430.07	**11.96**	1248.69	25.78	53.86
phy	14	1	**44.21**	1.0	< 0.01	1346.77	6.72	1264.37	38.36	51.85
phy	14	2	10.00	447.80	5.0	2392.92	2.93	1441.45	131.86	48.66
phy	14	3	9.87	110.91	< 0.01	< 0.01	9.877726	1329.96	42.93	71.50
phy	15	0	10.35	-	5.0	1.47	10.27	680.83	64.44	5.26
phy	15	1	**44.21**	1.0	< 0.01	4.54	1.50	1187.49	56.71	3.48
phy	15	2	9.79	340.68	5.0	9.28	9.93	674.63	18.54	3.36
phy	15	3	8.25	72.45	< 0.01	< 0.01	8.27	889.44	3.63	4.53
patho	9	0	10.17	-	5.0	111075.2	10.20	836.66	40.69	16933.86
patho	9	1	**44.21**	1.0	< 0.01	27592.32	21.06	1149.61	34.17	1016632.0
patho	9	2	9.27	**553.61**	5.0	**490494.8**	6.48	894.27	8.46	0.0
patho	9	3	9.40	115.31	< 0.01	0.0	**36.50**	1039.40	34.72	2242588.0
patho	14	0	10.22	-	5.0	261.05	11.99	815.67	161.33	204.56
patho	14	1	**44.21**	1.0	< 0.01	1142.68	3.83	916.29	201.68	4119.82
patho	14	2	9.86	380.93	5.0	2701.52	9.71	1135.42	200.35	6031.74
patho	14	3	9.75	193.18	< 0.01	0.0	12.01	742.62	**217.16**	9690.95
patho	15	0	10.13	-	4.1	275.48	26.11	1130.37	48.14	1136.95
patho	15	1	**44.21**	1.0	8.3	1236.27	35.18	**1471.51**	172.80	4592.89
patho	15	2	9.27	540.69	5.0	2217.90	9.12	1216.01	66.51	4404.67
patho	15	3	9.88	115.31	< 0.01	0.0	32.68	887.13	47.10	9538.67

Table 5.16: Intra cluster parameters calculation according the state and stage of persistence diagrams.

5.5.2 Discussion

In this part, we evaluate the influence of the physiopathological state and the developmental stage on the topology of brain organoid in order to characterize them. The pattern of barcodes are different for the state and the developmental stage which is also observed in the persistence diagrams. The quantitative analysis show a difference in slope, intercept, entropy and dispersion with a root at 14 days of culture. The clustering analysis show the entropy, average and Wasserstein distance analysis are particular.

We show for the first time the state and stage of a brain organoid culture is characterized by particular barcode and persistence diagrams. Similarly as previous study in 5.3, and 5.4 we observe the topology can be used to characterize these brain cultures. The conclusions are similar to these two parts. The clustering approach help to compare more specifically the state and stage of brain organoids.

Similarly all the parameters calculated on the persistence diagrams are not highlighting the state and stage differences between the culture topology. The question raised could be which kind of parameters are dedicated to decipher similitude and differences between two persistence diagrams due to the absence of consensus in the domain and standardize methodology of comparison. A strategy could be to compare all the contribution of these parameters in a dimensional reduction strategy.

Another contribution in the field could be to compare morphological index calculations with the topological analysis made upon segmented brain organoid images in order to observe which one could contribute to help the most to characterize these cultures. The other interest could be to use these solutions in order to predict the state and stage of the culture, or to help to follow their growth automatically.

This third study on the topological aspect of brain organoids reinforce the interest of using this strategy to distinguish the state and stage of these cultures. Clustering approach help to differentiate the developmental and physiopathological approaches at the same time. Futur studies need to be set up to observe the contribution of topology versus the morphological index based calculation to predict the state and stage of these culture, or others on the validation of metrics used to compare two topological approaches.

5.6 General discussion on characterisation

The goal of this characterisation part is to describe the shape of brain organoid in order to understand their physiopathological development and help to the prediction of new cultures. By comparing these various procedures based upon boundary, regional or topological descriptors, we highlight some strategies are more accurate than others. In this discussion we will highlight the benefits of these results, some drawbacks, the first work we attempts such as the Morse theory for the neuroepithelial detection, and finally the industrial and future perspectives of research for the characterisation of these tiny datasets.

Discussion on our work In this work we propose to verify first the manual, semi or automated extraction in order to highlight the morphological index which seems to characterize the most these cultures. This work highlight all the indexes related to the convexity of the shape have a particular relation with the state and stage of brain organoids. We validate it on two datasets, but if the results seems validated for an automated solution (U-Net optimized with in the training step synthetic images triple validated), we do not reach a similar results with the other automatic solution (Mini-Unet). Progress need to be

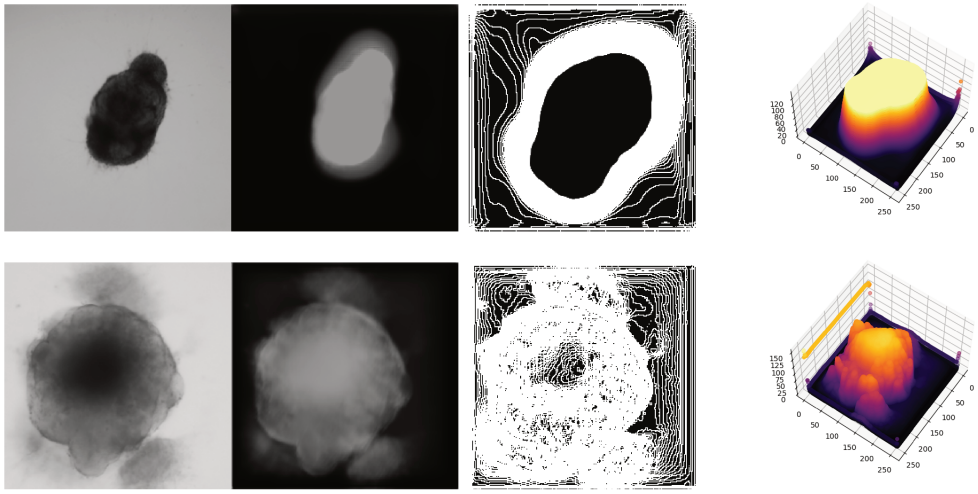


Figure 5.17: Neuroepithelial detection in brain organoids: by Morse theory (top lines from left to right original image, segmented map by the optimized U-Net, map of saddle points and 3D density map: the first corresponds to an early developmental stage and the second a later stage) and Deep learning methods (bottom line bounding box of neuroepithelial detections obtained by Yolo

.

further investigated to reach similar results with this lower computational cost architecture, and particularly a study on shape prior synthesis.

The second issue with a simple morphological index characterisation is this methodology does not allow a prediction: maybe due to the not well accurate extraction on neuroepithelial formation, or maybe due to the only boundary or region related link. For this reason we propose to use topological data analysis which allow us first to characterize the stage, the state and both for brain organoids. These results need to be verified on a second dataset.

Taken together, boundary, regional and topological characterisation could help to improve the synthesis of supplementary images, which could accurate the segmentation and thus improve also the morphological characterisation of tiny datasets.

Initial work on perspectives Concerning the future directions they are many folded. First we attempt to detect the neuroepithelial formations by using morse theory and a deep learning tool (Delgado-Friedrichs et al., 2014; Robins et al., 2011; Redmon et al., 2017). The results are summarized in Figure 5.17. The idea behind the prediction of neuro-epithelial formation is to better characterize the state of development of these culture, and the second objective is to verify if the culture grow well. Thus a brain organoid which do not develop these neuroepithelial formation need to be left apart from all the future analysis.

Academic and industrial perspectives Industrial and academic perspectives from this characterisation work are various:

- First we provide an automatic solution to calculate morphological index parameter usualy used to characterize brain organoid morphology which has been validated on two datasets.
- This solution could be applied on other tiny datasets to characterize the morphology of particular segmented structures.

- We observe the convexity based calculated parameters are the one which describes the best the state and stages of these cultures, it could be interesting to study the case in others organoids or other biomedical datasets.
- We produce a new clustering-TDA strategy to complete the morphological index calculation which could help to predict the state and stage of these cultures.
- Similarly this strategy could be used on other datasets to characterize the morphology.
- We observe however a particular drawback, if these solution are automated, they could be accurated by improving the automatic extraction by giving synthetic images used during the training step more similar to the reality by adding a prior shape of topological knowledge information during the generation (F. Wang et al., 2020; Moor et al., 2020).
- We propose new perspective based upon the neuroepithelial detection in order to better characterize the state of these brain organoid or their well growth.
- The development of this latest axis could open new issues to answer and develop.

Conclusion This work contains major contributions such as:

- We automate morphological index calculation based upon all the literature use.
- The comparison of manual, semi-automated and automated extraction tools to highlight morphological index calculation to characterize the shape of particular state and stage of tiny datasets of brain cultures.
- We highlight the interest of using particularly convex based tools for the brain culture characterisation of state and shape alone or both and has been validated on other datasets.
- We show the interest of coupling region, boundary and topological based tools to characterize these cultures, and particularly for predictive purpose.
- We create an original clustering-TDA strategy which accurate the morphological characterisation.
- We propose two future strategies for the detection of neuroepithelial formation inside these cultures and show on tiny datasets, the deep learning solution are maybe not the solution.

This work and approach should be tested on others models in order to generalize this strategy.

Chapter 6

Conclusion and perspectives

The objective of this PhD thesis was to study and propose new automatic tools for the segmentation and the characterisation of brain organoid images with the constraint of tiny datasets.

The first obstacle resided in the fact of augmenting the initial tiny datasets with a natural but not identical content in image in order to train deep learning extracting tools. The associated issue was to validate the synthetic images which was not under the control of a consensus.

The second objective, once new synthetic natural images are created, was to segment the shape of organoid from their background with automated tools and compare the machine learning segmentation with deep learning. The main part was to propose a segmentation architecture adapted to tiny datasets.

The third objective, once the shapes of brain organoid were well extracted was to characterize them and compare morphological index found in literature with new methodologies developed during the thesis. The latest objective is to predict their patho-physiological state or their developmental stage.

The objective sub-latent is to validate each proposition by verifying each implementation and result on minimum two tiny datasets of these brain culture in a pandemic situation and almost no open source database available. The validation on other datasets could help to generalize the results, and maybe use them in other situations in the case of the platform proposed by the company involved in this CIFRE.

6.1 Results and contributions

The first contribution of this thesis was to realize a state of the art about the image analysis made with brain organoid in order to highlight which kind of image analysis were needed in this field. This study help us to highlight the lack of datasets and automatic extractive and morphological characterisation tools adapted to the shape of these cultures. This contribution has been published in the review Frontiers in Neuroscience.

To increase the few image datasets found in the literature, we tested various data augmentation strategies based on generative adversarial networks and optimized the loss function for the best architecture, which is based on an adversarial auto-encoder. We highlight, a particular loss function (Perceptual Wasserstein) used during the generation to create the most natural image. This work has been published in the International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (Visap). We validate the results by a triple validation due to the lack of a consensus on the use of validation tools for synthetic images: quantitatively by metrics, statistically by a dimensional

reduction strategy, and manually by a psychovisual validation made by biological experts in culture and microscopy. This comparison of generative tools and loss optimisations has been published in the Visapp international conference, while the validation study has been sent to the Frontiers in Neuroscience journal.

While doing the statistical comparison, we observe synthetic images do not have a specific gradient background found in some original images. We thus study the effect of adding various noise during the generative step and reproduce the same three step validation. Two kinds of noise, Gaussian and Shot noise, help to produce more natural images according to qualitative and psychovisual validation. Looking at the statistical space, we found each noise contributes to generate images in the space of original images. The comparison of various noise injection during the generative step has been published in the International Conference on Image Processing Theory, Tools and Applications (Ipta), while the validation has not been yet published.

Once these images were validated, we use them to train and compare the various data augmentation strategies for the shape extraction by deep learning tools which requires a huge amount of data. We found a U-Net deep learning extractive network produces the best extraction if trained with Perceptual Wasserstein loss optimized synthetic images, or, images selected by the most experts as natural, or by Speckle noise. These studies has not been published as they require a second validation on a second dataset. We observe the extraction was time consuming (the training requires time), are not well accurate due to the small amount of data even augmented and in the literature, we found some reduction of architectures could be more adapted to tiny datasets. We made an ablation study of U-Net in order to produce a segmentation architecture Mu-Net more robust to various data-augmentation strategies, which require less time of training and dedicated to tiny datasets. This contribution has been presented in the GDR-ISIS on "Towards pragmatic learning in a context of limited labeled visual data", and the article is still unpublished. We also segmented the shape by semi-automatic tools in order to observe if the best automatic extraction by our optimized segmentation strategies outperform the non deep-learning tools used in most of the case in the literature.

Once segmented, we characterized the shape of organoids of various state and stage of culture. First we compare a manual, semi-automated, and automated extraction by calculating on their shape only morphological index that we observed in our state of the art. While some morphological parameters based upon convexity seems to correctly characterize the shape of brain organoid at various state and stage, these parameters based upon regional or boundary descriptors do not allow to predict the state or stage of a culture. Thus, we also analyze the topology of their shape by topological data analysis (TDA) and propose a clustering-TDA strategy which could allow the prediction of an image. The TDA-clustering strategy to characterize and predict the developmental stage of brain organoids has been published in the 2nd workshop of TDA of the Miccai international conference. We also detect the neuroepithelial formations on two datasets by deep learning and machine learning strategies. However, further analysis need to be furnish in the future to strenghen our results.

We are conscious the analysis on two datasets is not enough to provide a generalisation, thus, we ask datasets to some laboratories as we could not acquired images due to pandemic restriction. We obtain two new datasets from the Valrose institute from Nice and the MRC Laboratory of Molecular Biology from Cambridge, which will allow to validate our approach in the near future.

6.2 Research perspectives

This work is a pioneer in the field of automated shape extraction and characterisation of brain organoid, and we think many other lines of research could be studied both in academia and industria. Here, we will describe some perspective of research highlighted on this thesis (Figure 6.1).

Shape prior constraints for the synthesis of brain organoid images One of the main perspective of research could be to use the morphological index calculation made upon the boundary, regional and topological descriptors to help synthesize more natural images. Indeed, we have observed the morphological index calculated on synthetic images are not always in the range of morphological index calculated on original images see Section 3.3 and 3.4. This issue could contribute to explain why the tri-validation do not perform high scores in some cases (Section 3.5 and 3.6), why some segmentation using these images during the training do not reach high accurate score (Section 4.2), and the automatic extractive solution have not exactly the same results to the manual(Section 4.4). Thus, giving a shape-topological prior information during the generation could help to synthesize maybe more natural images and improve all the chain of procedures. The strategy adopted could be to first verify boundary, then the regional, and the topological and, in a third try to add the three kind of information. This strategy may help to observe if a kind of morphological characteristic helps better to naturalize the synthetic images than another, or if all or the most information improve the results.

Few-shot synthesis of brain organoid images The second improvement in the generative process is trying to generate images with only two, one or unlabeled images. In the case of tiny datasets, the partitioning of the dataset in batches to create a model of generation is particularly a difficulty because the samples have to reach an equilibrium while we do not have so much possibilities of combinations see Section 3.2. Using only two images by two, or one, or generating images by a zero shot strategy could help maybe to reduce the time required to analyze the dataset before using it during the generation. The strategy could be to use a cascaded generative architecture in order to produce images in the range of few images. We could then compare the variability and naturality of these generated images, with a non few-shot strategy and combine it, if it render best results with the proposed loss noise or shape-topological prior information.

Metric equivalent to the tri-validation of synthetic images Nowadays there is no consensus on a metric which could select images has the most natural when they are generated by a GAN, and the psychovisual evaludation is time and ressource consuming. In this thesis we observe a combination of metrics could maybe replace the psychovisual evaluation see Section 3.5 and 3.6. Instead of using various metric calculations, the idea is to create a single metric. The strategy to create such metric is to combine the objective of the metric which indicates synthetic images in the range of original.

Tri-validation extended on other biomedical datasets We validate synthetic images by tri-validation for brain organoid tiny datasets, but the main objective could be to generalize this approach on other biomedical datasets or on other microscopic images see Section 2.7. The idea could be to synthesize images by GAN on multiple contents and then to validate by metric, statistic and psychovisual tools these images.

Shape prior constraints for the segmentation of brain organoid images The objective is similar to the shape prior constraint for the generation but applied to the segmentation, in order to improve it. Here, we aim at adding the shape or topological information to the training step of the segmentation process. This kind of process is still in research and more difficult to implement than for classification or generation. The idea could be to generalize it for other tiny biomedical datasets.

Scalable reduction of segmentation architecture for small biomedical datasets A reduction of the segmentation architecture could be proposed according to the size of the dataset. Indeed, proposing a deep and reduced architecture according to the kind of dataset could help to improve the shape segmentation and strengthen the analysis. The algorithm scalability is a recent opportunity to proposed dedicated solution for each application, and thus for each dataset containing more or less content and more or less variability.

Explainability of convex shape descriptors for the characterisation of brain cultures The convexity and its variations seems the most important morphological index to characterize the morphology of brain organoids see Section 5.2. The idea is to better understand the growth of brain organoids by calculating the convexity at each step of culture in order to identify in the future the step of brain organoid culture automatically. The second idea is to modele the convexity of the physiological growth of brain organoid in order to understand the pathological development and, better understand why this measure is so important.

Reverse TDA for the explainability of the morphological characterisation of small biomedical datasets In this thesis we realize a TDA-clustering to characterize and predict the state and stage of brain organoid cultures (see Sections 5.3, 5.4, and, 5.5). However, which kind of the part is linked to each clustering formation in the persistence diagram? During the thesis we made the supposition that some neuro-epithelial formations are contained in some cluster and axes of boundaries others. By doing a reverse TDA-clustering strategy, we could maybe highlight which part of the shape are evolving or staying static during the growth or a state variation.

Classification of original brain organoid images The prediction of brain organoid state and stage has been attempted in this thesis by the creation of a morse theory to detect for instance neuro-epithelial zones, or using TDA-clustering strategies see Chapter 5. However, these propositions has been made upon a segmented image, the idea is to classify the images in groups by directly taking the original images not segmented. We suggest to use deep learning tools and compare it with classical approaches.The use of data augmentation strategies which will create only a kind of state or stage of brain organoid culture (for instance based on few-shot generation) could help to this task.

Prediction of brain organoid culture stage and state We observe the prediction is still tedious with morphological index and topological data analysis see Chapter 5. We propose finally to use all the previous studies detailed in this perspective subsection to combine them in a strong tool to predict the state and stage of brain organoids. We are convinced the effort provided in all these perspectives could help to determine a predictive solution to detect if a culture is well growing correctly.

6.3 The last word

The last words aims at giving in a macroscopic point of view the interest of this thesis. First, this thesis has shown even in a context of tiny dataset, brain organoid could be analyzed with automated image analysis tools. Then, that a pandemic situation in which the access of laboratory are forbidden to acquire images, some solutions such as data augmentation and the literature research of open source database could help to furnish enough content to develop some image analysis tools. We also could observe the generative process by GAN solutions can be adapted and improve by some updates to produce more natural images particularly the loss update and noise injections. A tri-validation of synthetic images help to strengthen the accuracy of the segmentation step when they are used during the training, even in a case of small amount of images used during the training. The validation of synthetic images by metrics are not yet under a consensus and the psychovisual approach is time consuming for biological experts, thus this thesis work proposes by this tri-validation a strategy which could be a solution until finding a dedicated metric for this purpose of synthetic image validation. The segmentation architecture for these tiny datasets can be lightened in order to reduce the amount of data required and the training time. In a context of energy cost diminution, the scalability of architectures adapted to the amount of data and furnishing a similar accuracy is raised. A morphological characterization of brain organoid lies on convex information, but only topological and topological clustering approach render a prediction of the state and stage of culture. Another perspective could be to use these shape-topological information in earliest steps in order to improve all the procedure and thus the analysis of brain organoid tiny datasets of images.

We expect these tools associated with the development of brain organoids more similar to the human brain structures could help to complete other brain model such as 2D cultures or animal models, and maybe replace them in the future. The second expectation could be to generalize these approach for other kind of datasets and propose this solution in an industrial platform of biomedical image analysis services.

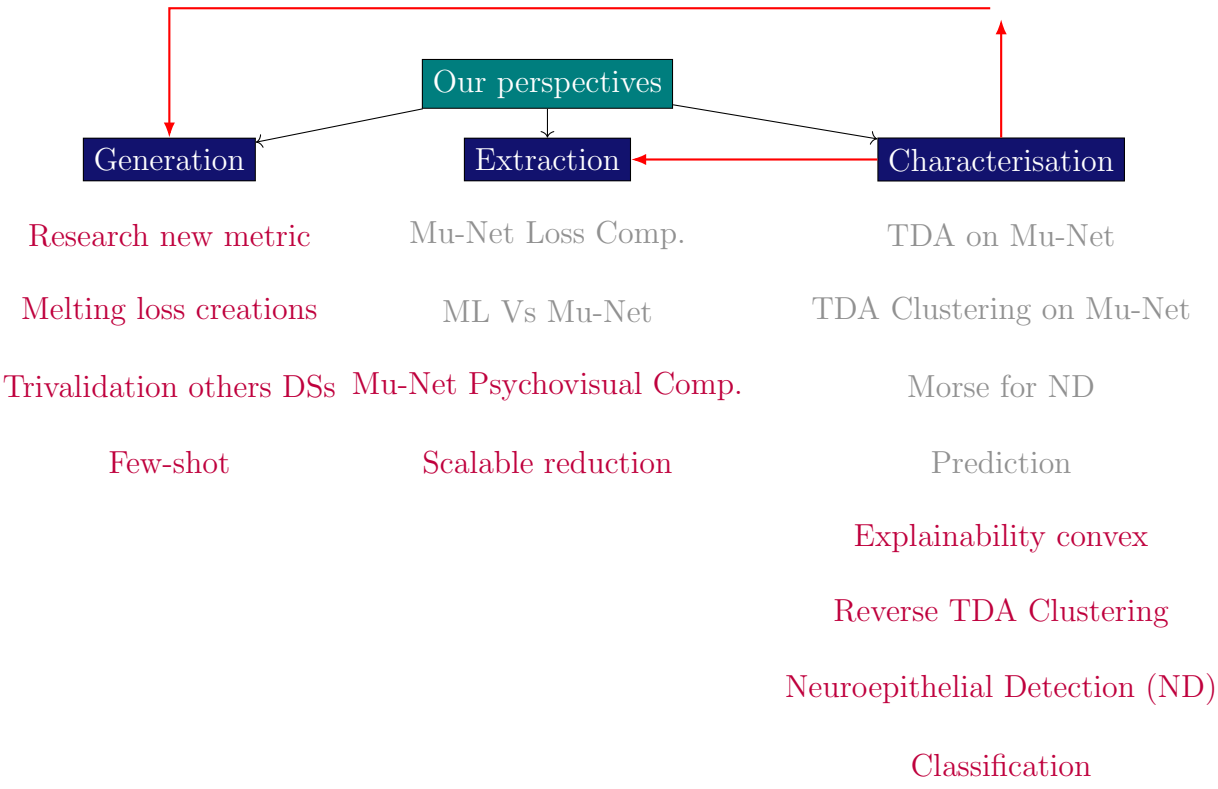


Figure 6.1: Summary of perspectives on brain organoid image analysis. In lightgray: in progress; in red: work we would like to implement; arrows: shape prior information given to other tasks.

Publications

6.4 Published review

Clara Brémond Martin et al. (July 2021b). "Recent Trends and Perspectives in Cerebral Organoids Imaging and Analysis". en. In: *Frontiers in Neuroscience* 15, p. 629067

6.5 Published international conference papers

Clara Brémond Martin et al. (2021a). "AAEGAN Loss Optimizations Supporting Data Augmentation on Cerebral Organoid Bright-Field Images". en. In: *Proceedings of the 17th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications VISAPPGRAPP*, p. 8

Clara Brémond Martin et al. (2022). "AAEGAN Optimization by Purposeful Noise Injection for the Generation of Bright-Field Brain Organoid Images". In: *Internal Conference On Image Proceccing Theory Tools and Application : Special Session Biological & Medical Image Analysis*, p. 6

Clara Brémond-Martin et al. (2022). "TDA-Clustering Strategies for the Characterization of Brain Organoids". In: *Ethical and Philosophical Issues in Medical Imaging, Multimodal Learning and Fusion Across Scales for Clinical Decision Support, and Topological Data Analysis for Biomedical Imaging: 1st International Workshop, EPIMI 2022, 12th International Workshop, ML-CDS 2022, 2nd International Workshop, TDA4BiomedicalImaging, Held in Conjunction with MICCAI 2022, Singapore, September 18–22, 2022, Proceedings*. Springer, pp. 113–122

6.6 Unpublished papers

Brémond Martin, C., Simon Chane, C., Clouchoux, C., and Histace, A. TDA-Clustering Strategies for the Pathological Characterization of Brain Organoids.

Brémond Martin, C., Simon Chane, C., Clouchoux, C., and Histace, A. MU-net: a light architecture for small dataset segmentation applied to Brain Organoid bright-field images.

Brémond Martin, C., Simon Chane, C., Clouchoux, C., and Histace, A. Brain Organoid Data Synthesis and Evaluation. Computers in biology and medicine.(send : CIBM-D-23-02255)

Bibliography

Adams, R. and L. Bischof (June 1994). "Seeded region growing". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 16.6, pp. 641–647.

Adhya, Dwaipayan, George Chennell, James A. Crowe, Eva P. Valencia-Alarcón, James Seyforth, Neveen A. Hosny, Marina V. Yasvoina, Robert Forster, Simon Baron-Cohen, Anthony C. Vernon, and Deepak P. Srivastava (Dec. 2021). "Application of Airy beam light sheet microscopy to examine early neurodevelopmental structures in 3D hiPSC-derived human cortical spheroids". en. In: *Molecular Autism* 12.1, p. 4.

Albanese, Alexandre, Justin M. Swaney, Dae Hee Yun, Nicholas B. Evans, Jenna M. Antonucci, Silvia Velasco, Chang Ho Sohn, Paola Arlotta, Lee Gehrke, and Kwanghun Chung (Dec. 2020). "Multiscale 3D phenotyping of human cerebral organoids". en. In: *Scientific Reports* 10.1, p. 21487.

Ali, Mohammed A. S., Kaspar Hollo, Tõnis Laasfeld, Jane Torp, Maris-Johanna Tahk, Ago Rinken, Kaupo Palo, Leopold Parts, and Dmytro Fishman (Dec. 2022). "ArtSeg—Artifact segmentation and removal in brightfield cell microscopy images without manual pixel-level annotations". In: *Scientific Reports* 12.1, p. 11404.

Alqahtani, Hamed, Manolya Kavakli, and Dr. Gulshan Kumar Ahuja (July 2019). "An Analysis of Evaluation Metrics of GANs". In: *International Conference on Information Technology and Applications* 7.

Anagnostidis, Vasileios, Benjamin Sherlock, Jeremy Metz, Philip Mair, Florian Hollfelder, and Fabrice Gielen (2020). "Deep learning guided image-based droplet sorting for on-demand selection and analysis of single cells and 3D cell cultures". In: *Lab on a Chip* 20.5, pp. 889–900.

Anastasaki, Corina, Michelle L. Wegscheid, Kelly Hartigan, Jason B. Papke, Nathan D. Kopp, Jiayang Chen, Olivia Cobb, Joseph D. Dougherty, and David H. Gutmann (Apr. 2020). "Human iPSC-Derived Neurons and Cerebral Organoids Establish Differential Effects of Germline NF1 Gene Mutations". In: *Stem Cell Reports* 14.4, pp. 541–550.

Ao, Zheng, Hongwei Cai, Daniel J. Havert, Zhuhao Wu, Zhiyi Gong, John M. Beggs, Ken Mackie, and Feng Guo (Mar. 2020). "One-Stop Microfluidic Assembly of Human Brain Organoids To Model Prenatal Cannabis Exposure". In: *Analytical Chemistry* 92.6, pp. 4630–4638.

Arjovsky, Martin, Soumith Chintala, and Léon Bottou (2017). "Wasserstein generative adversarial networks". In: *International conference on machine learning*. PMLR, pp. 214–223.

Arlotta, Paola and Sergiu P Paşca (June 2019). "Cell diversity in the human cerebral cortex: from the embryo to brain organoids". In: *Current Opinion in Neurobiology* 56, pp. 194–198.

Arnout, Hiba, Johanna Bronner, and Thomas Runkler (2021). "Evaluation of Generative Adversarial Networks for Time Series Data". In: *2021 International Joint Conference on Neural Networks*. Vol. -. -, pp. 1–7.

Arulpandy, P and M Trinita Pricilla (2020). "Speckle noise reduction and image segmentation based on a modified mean filter". In: *Computer Assisted Methods in Engineering and Science* 27.4, pp. 221–239.

Ashok, Aswathi, Deepak Choudhury, Yu Fang, and Walter Hunziker (Mar. 2020). "Towards manufacturing of human organoids". In: *Biotechnology Advances* 39, p. 107460.

Baas, Nils A., Gunnar E. Carlsson, Gereon Quick, Markus Szymik, and Marius Thaule, eds. (2020). *Topological Data Analysis: The Abel Symposium 2018*. Vol. 15. Abel Symposia. Cham: Springer International Publishing.

Bagley, Joshua A, Daniel Reumann, Shan Bian, Julie Lévi-Strauss, and Juergen A Knoblich (July 2017). "Fused cerebral organoids model interactions between brain regions". In: *Nature Methods* 14.7, pp. 743–751.

Balaji, Yogesh, Rama Chellappa, and Soheil Feizi (2019). "Normalized Wasserstein for Mixture Distributions With Applications in Adversarial Learning and Domain Adaptation". In: *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 6499–6507.

Bateson, Mathilde, José Dolz, Hoel Kervadec, Hervé Lombaert, and I Ben Ayed (2021). "Constrained domain adaptation for image segmentation". In: *IEEE Transactions on Medical Imaging* 40.7, pp. 1875–1887.

Bauer, Ulrich (2021). "Ripser: efficient computation of Vietoris-Rips persistence barcodes". In: *Journal of Applied and Computational Topology*.

Bayne, Tim, Anil K. Seth, and Marcello Massimini (Jan. 2020). "Are There Islands of Awareness?" In: *Trends in Neurosciences* 43.1, pp. 6–16.

Bazin, Pierre-Louis, Anneke Alkemade, Martijn J Mulder, Amanda G Henry, and Birte U Forstmann (2020). "Multi-contrast anatomical subcortical structures parcellation". en. In: *Elife* 9, e59430.

Beluch, William H, Tim Genewein, Andreas Nürnberger, and Jan M Köhler (2018). "The power of ensembles for active learning in image classification". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 9368–9377.

Berger, Emanuel, Chiara Magliaro, Nicole Paczia, Anna S. Monzel, Paul Antony, Carole L. Linster, Silvia Bolognin, Arti Ahluwalia, and Jens C. Schwamborn (2018). "Millifluidic culture improves human midbrain organoid vitality and differentiation". In: *Lab on a Chip* 18.20, pp. 3172–3183.

Bergmann, Sonja, Sean E. Lawler, Yuan Qu, Colin M. Fadzen, Justin M. Wolfe, Michael S. Regan, Bradley L. Pentelute, Nathalie Y. R. Agar, and Choi-Fong Cho (Dec. 2018). "Blood–brain-barrier organoids for investigating the permeability of CNS therapeutics". In: *Nature Protocols* 13.12, pp. 2827–2843.

Bernal, Jorge, Nima Tajkbaksh, Francisco Javier Sanchez, Bogdan J. Matuszewski, Hao Chen, Lequan Yu, Quentin Angermann, Olivier Romain, Bjorn Rustad, Ilangko Balasingham, Konstantin Pogorelov, Sungbin Choi, Quentin Debard, Lena Maier-Hein, Stefanie Speidel, Danail Stoyanov, Patrick Brandao, Henry Cordova, Cristina Sanchez-Montes, Suryakanth R. Gurudu, Gloria Fernandez-Esparrach, Xavier Dray, Jianming Liang, and Aymeric Histace (June 2017). "Comparative Validation of Polyp Detection Methods in Video Colonoscopy: Results From the MICCAI 2015 Endoscopic Vision Challenge". In: *IEEE Transactions on Medical Imaging* 36.6, pp. 1231–1249.

Bershteyn, Marina, Tomasz J. Nowakowski, Alex A. Pollen, Elizabeth Di Lullo, Aishwarya Nene, Anthony Wynshaw-Boris, and Arnold R. Kriegstein (Apr. 2017). "Human iPSC-Derived Cerebral Organoids Model Cellular Features of Lissencephaly and Reveal Prolonged Mitosis of Outer Radial Glia". In: *Cell Stem Cell* 20.4, 435–449.e4.

Berwald, Jesse J., Joel M. Gottlieb, and Elizabeth Munch (Nov. 2018). "Computing Wasserstein Distance for Persistence Diagrams on a Quantum Computer". In: *arXiv:1809.06433 [quant-ph]*. arXiv: 1809.06433.

Bhalodia, Riddhish, Anupama Goparaju, Tim Sodergren, Alan Morris, Evgueni Kholmovski, Nassir Marrouche, Joshua Cates, Ross Whitaker, and Shireen Elhabian (Sept. 2018). *Deep Learning for End-to-End Atrial Fibrillation Recurrence Estimation*. en.

Birey, Fikri, Jimena Andersen, Christopher D. Makinson, Saiful Islam, Wu Wei, Nina Huber, H. Christina Fan, Kimberly R. Cordes Metzler, Georgia Panagiotakos, Nicholas Thom, Nancy A. O'Rourke, Lars M. Steinmetz, Jonathan A. Bernstein, Joachim Hallmayer, John R. Huguenard, and Sergiu P. Paşca (May 2017). "Assembly of functionally integrated human forebrain spheroids". In: *Nature* 545.7652, pp. 54–59.

Bois, Alexandre, Brian Tervil, Albane Moreau, Aliénor Vienne-Jumeau, Damien Ricard, and Laurent Oudre (May 2022). "A topological data analysis-based method for gait signals with an application to the study of multiple sclerosis". In: *PLOS ONE* 17.5. Ed. by Chan Hwang See, e0268475.

Bolognin, Silvia, Marie Fossépré, Xiaobing Qing, Javier Jarazo, Janez Ščančar, Edinson Lucumi Moreno, Sarah L. Nickels, Kobi Wasner, Nassima Ouzren, Jonas Walter, Anne Grünewald, Enrico Glaab, Luis Salamanca, Ronan M. T. Fleming, Paul M. A. Antony, and Jens C. Schwamborn (Jan. 2019). "3D Cultures of Parkinson's Disease-Specific Dopaminergic Neurons for High Content Phenotyping and Drug Testing". In: *Advanced Science* 6.1, p. 1800927.

Booij, Tijmen H., Leo S. Price, and Erik H. J. Danen (July 2019). "3D Cell-Based Assays for Drug Screens: Challenges in Imaging, Image Analysis, and High-Content Analysis". In: *SLAS DISCOVERY: Advancing the Science of Drug Discovery* 24.6, pp. 615–627.

Borji, Ali (Oct. 2018). "Pros and Cons of GAN Evaluation Measures". In: *Computer Vision and Image Understanding*.

Borten, Michael A., Sameer S. Bajikar, Nobuo Sasaki, Hans Clevers, and Kevin A. Janes (Dec. 2018). "Automated brightfield morphometry of 3D organoid populations by Organo Seg". In: *Scientific Reports* 8.1.

Bouleux, Guillaume, Mael Dugast, and Eric Marcon (Oct. 2019). "Information Topological Characterization of Periodically Correlated Processes by Dilation Operators". In: *IEEE Transactions on Information Theory* 65.10, pp. 6484–6495.

Boutin, M. E., C. Hampton, R. Quinn, M. Ferrer, and M. J. Song (2019). "3D Engineering of Ocular Tissues for Disease Modeling and Drug Testing". In: *Pluripotent Stem Cells in Eye Disease Therapy*. Ed. by Kapil Bharti. Vol. 1186. Series Title: Advances in Experimental Medicine and Biology. Cham: Springer International Publishing, pp. 171–193.

Boutin, Molly E., Ty C. Voss, Steven A. Titus, Kennie Cruz-Gutierrez, Sam Michael, and Marc Ferrer (Dec. 2018). "A high-throughput imaging and nuclear segmentation analysis protocol for cleared 3D culture models". In: *Scientific Reports* 8.1.

Boyat, Ajay Kumar and Brijendra Kumar Joshi (Apr. 2015). "A Review Paper : Noise Models in Digital Image Processing". en. In: *Signal & Image Processing : An International Journal* 6.2, pp. 63–75.

Brémond Martin, Clara, Camille Simon Chane, Cédric Clouchoux, and Aymeric Histace (2021a). "AAEGAN Loss Optimizations Supporting Data Augmentation on Cerebral Organoid Bright-Field Images". en. In: *Proceedings of the 17th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications VISAPPGRAPP*, p. 8.

Brémond Martin, Clara, Camille Simon Chane, Cédric Clouchoux, and Aymeric Histace (July 2021b). "Recent Trends and Perspectives in Cerebral Organoids Imaging and Analysis". en. In: *Frontiers in Neuroscience* 15, p. 629067.

— (2022). "AAEGAN Optimization by Purposeful Noise Injection for the Generation of Bright-Field Brain Organoid Images". In: *Internal Conference On Image Proceccing Theory Tools and Application : Special Session Biological & Medical Image Analysis*, p. 6.

Brémond-Martin, Clara, Camille Simon-Chane, Cédric Clouchoux, and Aymeric Histace (2022). "TDA-Clustering Strategies for the Characterization of Brain Organoids". In: *Ethical and Philosophical Issues in Medical Imaging, Multimodal Learning and Fusion Across Scales for Clinical Decision Support, and Topological Data Analysis for Biomedical Imaging: 1st International Workshop, EPIMI 2022, 12th International Workshop, ML-CDS 2022, 2nd International Workshop, TDA4BiomedicalImaging, Held in Conjunction with MICCAI 2022, Singapore, September 18–22, 2022, Proceedings*. Springer, pp. 113–122.

Brownjohn, Philip W., James Smith, Ravi Solanki, Ebba Lohmann, Henry Houlden, John Hardy, Sabine Dietmann, and Frederick J. Livesey (Apr. 2018). "Functional Studies of Missense TREM2 Mutations in Human Stem Cell-Derived Microglia". In: *Stem Cell Reports* 10.4, pp. 1294–1307.

Buchsbaum, Isabel Y, Pavel Kielkowski, Grazia Giorgio, Adam C O'Neill, Rossella Di Giaimo, Christina Kyrousi, Shahryar Khattak, Stephan A Sieber, Stephen P Robertson, and Silvia Cappello (May 2020). "ECE2 regulates neurogenesis and neuronal migration during human cortical development". In: *EMBO reports* 21.5.

Cai, Lei, Jingyang Gao, and Di Zhao (2020). "A review of the application of deep learning in medical image classification and segmentation". In: *Annals of translational medicine* 8.11.

Camp, J. Gray, Farhath Badsha, Marta Florio, Sabina Kanton, Tobias Gerber, Michaela Wilsch-Bräuninger, Eric Lewitus, Alex Sykes, Wulf Hevers, Madeline Lancaster, Juergen A. Knoblich, Robert Lachmann, Svante Pääbo, Wieland B. Huttner, and Barbara Treutlein (Dec. 2015). "Human cerebral organoids recapitulate gene expression programs of fetal neocortex development". In: *Proceedings of the National Academy of Sciences*, p. 201520760.

Cederquist, Gustav Y., James J. Asciolla, Jason Tchieu, Ryan M. Walsh, Daniela Cornacchia, Marilyn D. Resh, and Lorenz Studer (Apr. 2019). "Specification of positional identity in forebrain organoids". In: *Nature Biotechnology* 37.4, pp. 436–444.

Chakradhar, Shraddha (Apr. 2016). "New company aims to broaden researchers' access to organoids". In: *Nature Medicine* 22.4, pp. 338–338.

Chambers, Stuart M, Christopher A Fasano, Eirini P Papapetrou, Mark Tomishima, Michel Sadelain, and Lorenz Studer (Mar. 2009). "Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling". In: *Nature Biotechnology* 27.3, pp. 275–280.

Chawla, Nitesh V., Kevin W. Bowyer, Lawrence O. Hall, and W. Philip Kegelmeyer (2002). "SMOTE: Synthetic Minority over-Sampling Technique". In: *J. Artif. Int. Res.* 16.1, pp. 321–357.

Chen, Haobo, Yuqun Wang, Jie Shi, Jingyu Xiong, Jianwei Jiang, Wanying Chang, Man Chen, and Qi Zhang (2021). "Segmentation of lymph nodes in ultrasound images using u-net convolutional neural networks and Gabor-based anisotropic diffusion". In: *Journal of Medical and Biological Engineering* 41.6, pp. 942–952.

Chen, Liang-Chieh, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L Yuille (2017). "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs". In: *IEEE transactions on pattern analysis and machine intelligence* 40.4, pp. 834–848.

Chen, Wenjin, Chung Wong, Evan Vosburgh, Arnold J. Levine, David J. Foran, and Eugenia Y. Xu (July 2014). "High-throughput Image Analysis of Tumor Spheroids: A User-friendly Software Application to Measure the Size of Spheroids Automatically and Accurately". In: *Journal of Visualized Experiments* 89.89.

Chen, Xinjian and Lingjiao Pan (2018). "A survey of graph cuts/graph search based medical image segmentation". In: *IEEE reviews in biomedical engineering* 11, pp. 112–124.

Chen, Yizhou, Xu-Hua Yang, Zihan Wei, Ali Asghar Heidari, Nenggan Zheng, Zhicheng Li, Huiling Chen, Haigen Hu, Qianwei Zhou, and Qiu Guan (May 2022). "Generative Adversarial Networks in Medical Image augmentation: A review". en. In: *Computers in Biology and Medicine* 144, p. 105382.

Cheon, Manri, Jun-Hyuk Kim, Jun-Ho Choi, and Jong-Seok Lee (2018). "Generative adversarial network-based image super-resolution using perceptual content losses". In: *Proceedings of the European Conference on Computer Vision (ECCV) Workshops*.

Chhibber, Tanya, Sounak Bagchi, Behnaz Lahooti, Angela Verma, Abraham Al-Ahmad, Manash K. Paul, Gurudutt Pendyala, and Rahul Dev Jayant (Feb. 2020). "CNS organoids: an innovative tool for neurological disease modeling and drug neurotoxicity screening". In: *Drug Discovery Today* 25.2, pp. 456–465.

Chlap, Phillip, Hang Min, Nym Vandenberg, Jason Dowling, Lois Holloway, and Annette Haworth (Aug. 2021). "A review of medical image data augmentation techniques for deep learning applications". en. In: *Journal of Medical Imaging and Radiation Oncology* 65.5, pp. 545–563.

Cho, Choi-Fong, Justin M. Wolfe, Colin M. Fadzen, David Calligaris, Kalvis Hornburg, E. Antonio Chiocca, Nathalie Y. R. Agar, Bradley L. Pentelute, and Sean E. Lawler (Aug. 2017). "Blood-brain-barrier spheroids as an in vitro screening platform for brain-penetrating agents". In: *Nature Communications* 8.1.

Choudhury, Deepak, Aswathi Ashok, and May Win Naing (Mar. 2020). "Commercialization of Organoids". In: *Trends in Molecular Medicine* 26.3, pp. 245–249.

Chukwurah, Evelyn, Allison Osmundsen, Shannon W. Davis, and Sofia B. Lizarraga (June 2019). "All Together Now: Modeling the Interaction of Neural With Non-neural Systems Using Organoid Models". In: *Frontiers in Neuroscience* 13.

Çiçek, Özgün, Ahmed Abdulkadir, Soeren S Lienkamp, Thomas Brox, and Olaf Ronneberger (2016). "3D U-Net: learning dense volumetric segmentation from sparse annotation". In: *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part II 19*. Springer, pp. 424–432.

Clough, James R, Nicholas Byrne, Ilkay Oksuz, Veronika A Zimmer, Julia A Schnabel, and Andrew P King (2020). "A topological loss function for deep-learning based image segmentation using persistent homology". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 44.12, pp. 8766–8778.

Corral Acero J Zacur E, Xu H et al (2019). "Data Augmentation Based on Statistical Models of Deformation to Enhance Segmentation in 2D Cine Cardiac MRI." In: *Functional Imaging and Modeling of the Heart. Lecture Notes in Computer Science* 361, p. 9.

Costa, Elisabete C., Daniel N. Silva, André F. Moreira, and Ilídio J. Correia (Oct. 2019). "Optical clearing methods: An overview of the techniques used for the imaging of 3D spheroids". In: *Biotechnology and Bioengineering* 116.10, pp. 2742–2763.

Cullen, D. Kacy, Wisberty J. Gordián-Vélez, Laura A. Struzyna, Dennis Jgamadze, James Lim, Kathryn L. Wofford, Kevin D. Browne, and H. Isaac Chen (Nov. 2019). "Bundled Three-Dimensional Human Axon Tracts Derived from Brain Organoids". In: *iScience* 21, pp. 57–67.

Daviaud, Nicolas, Roland H. Friedel, and Hongyan Zou (Nov. 2018). "Vascularization and Engraftment of Transplanted Human Cerebral Organoids in Mouse Cortex". In: *eneuro* 5.6, ENEURO.0219–18.2018.

De Roux, Edwin, Mehdi Terosiet, Florian Kolbl, Michel Boissière, Philippe Banet, Johnatan Chrun, Pierre-Henry Aubert, Emmanuel Pauthe, Olivier Romain, and Aymeric Histace (2017). "Prototype of a Wireless System for Electrical Impedance Spectroscopy: Demonstration of Real Time Measurements on Cells and Biological Medium". In: *IEEE International Workshop on Signal Processing Systems*, accepted.

De Roux, Edwin, Mehdi Terosiet, Florian Kölbl, Johnatan Chrun, Pierre-Henry Aubert, Philippe Banet, Michel Boissière, Emmanuel Pauthe, Aymeric Histace, and Olivier Romain (2017). "Wireless and portable system for the study of in-vitro cell culture impedance spectrum by electrical impedance spectroscopy". In: *2017 Euromicro Conference on Digital System Design (DSD)*. IEEE, pp. 456–461.

Delgado-Friedrichs, Olaf, Vanessa Robins, and Adrian Sheppard (2014). "Morse theory and persistent homology for topological analysis of 3d images of complex materials". In: *2014 IEEE International Conference on Image Processing (ICIP)*. IEEE, pp. 4872–4876.

Denton, Emily, Soumith Chintala, Arthur Szlam, and Rob Fergus (2015). "Deep Generative Image Models Using a Laplacian Pyramid of Adversarial Networks". In: *Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 1*. NIPS'15. Montreal, Canada: MIT Press, pp. 1486–1494.

Devassy, Binu Melit and Sony George (2020). "Dimensionality reduction and visualisation of hyperspectral ink data using t-SNE". In: *Forensic science international* 311, p. 110194.

DeVries, Terrance and Graham Taylor (Feb. 2017). "Dataset Augmentation in Feature Space". In: *Arxiv*.

Di Matteo, Francesco, Fabrizia Pipicelli, Christina Kyrousi, Isabella Tovecci, Eduardo Penna, Marianna Crispino, Angela Chambery, Rosita Russo, Ane Cristina Ayo-Martin, Martina Giordano, Anke Hoffmann, Emilio Ciusani, Laura Canafoglia, Magdalena Götz, Rossella Di Giaimo, and Silvia Cappello (May 2020). "Cystatin B is essential for proliferation and interneuron migration in individuals with EPM1 epilepsy". In: *EMBO Molecular Medicine*.

Dobosz, Michael, Vasilis Ntziachristos, Werner Scheuer, and Steffen Strobel (Jan. 2014). "Multispectral Fluorescence Ultramicroscopy: Three-Dimensional Visualization and Automatic Quantification of Tumor Morphology, Drug Penetration, and Antiangiogenic Treatment Response". In: *Neoplasia* 16.1, 1–W7.

Donegan, Jennifer J. and Daniel J. Lodge (Mar. 2020). "Stem Cells for Improving the Treatment of Neurodevelopmental Disorders". In: *Stem Cells and Development*.

Dong, Rongsheng, Xiaoquan Pan, and Fengying Li (2019). "DenseU-net-based semantic segmentation of small objects in urban remote sensing images". In: *IEEE Access* 7, pp. 65347–65356.

Durens, Madel, Jonathan Nestor, Madeline Williams, Kevin Herold, Robert F. Niescier, Jason W. Lunden, Andre W. Phillips, Yu-Chih Lin, Derek M. Dykxhoorn, and Michael W. Nestor (Apr. 2020). "High-throughput screening of human induced pluripotent stem cell-derived brain organoids". In: *Journal of Neuroscience Methods* 335, p. 108627.

Edelsbrunner, Herbert and John Harer (2008). "Persistent homology—a survey". In: *Contemporary Mathematics*. Ed. by Jacob E. Goodman, János Pach, and Richard Pollack. Vol. 453. Providence, Rhode Island: American Mathematical Society, pp. 257–282.

— (2010). "Computational Topology". en. In: *American Mathematical Society*, p. 294.

Edelsbrunner, Letscher, and Zomorodian (Nov. 2002). "Topological Persistence and Simplification". en. In: *Discrete & Computational Geometry* 28.4, pp. 511–533.

Eiraku, Mototsugu, Nozomu Takata, Hiroki Ishibashi, Masako Kawada, Eriko Sakakura, Satoru Okuda, Kiyotoshi Sekiguchi, Taiji Adachi, and Yoshiki Sasai (Apr. 2011). "Self-organizing optic-cup morphogenesis in three-dimensional culture". In: *Nature* 472.7341, pp. 51–56.

El Jurdi, Rosana, Caroline Petitjean, Paul Honeine, and Fahed Abdallah (2020). "Bb-unet: U-net with bounding box prior". In: *IEEE Journal of Selected Topics in Signal Processing* 14.6, pp. 1189–1198.

El Jurdi, Rosana, Caroline Petitjean, Paul Honeine, Veronika Cheplygina, and Fahed Abdallah (2021). "High-level prior-based loss functions for medical image segmentation: A survey". In: *Computer Vision and Image Understanding* 210, p. 103248.

Eremeev, A. V., E. A. Volovikov, L. D. Shuvalova, A. V. Davidenko, E. A. Khomyakova, M. E. Bogomiakova, O. S. Lebedeva, O. A. Zubkova, and M. A. Lagarkova (Mar. 2019). ""Necessity Is the Mother of Invention" or Inexpensive, Reliable, and Reproducible Protocol for Generating Organoids". In: *Biochemistry (Moscow)* 84.3, pp. 321–328.

Feng, Ruili, Deli Zhao, and Zheng-Jun Zha (2021). "Understanding noise injection in gans". In: *International Conference on Machine Learning*. PMLR, pp. 3284–3293.

Fernández, Rubén, Pilar Rosado, Esteban Vegas, and Ferran Reverter (2021). "Medical image editing in the latent space of Generative Adversarial Networks". en. In: *Intelligence-Based Medicine* 5, p. 100040.

Frahi, Tarek, Clara Argerich, Minyoung Yun, Antonio Falco, Anais Barasinski, and Franscesco Chinesta (2020). "Tape surfaces characterization with persistence images". In: *AIMS Materials Science* 7.4, pp. 364–380.

Friedrich, Juergen, Claudia Seidel, Reinhard Ebner, and Leoni A Kunz-Schughart (Mar. 2009). "Spheroid-based drug screen: considerations and practical approach". In: *Nature Protocols* 4.3, pp. 309–324.

Fujii, Kazuki, Yuki Saito, Shinnosuke Takamichi, Yukino Baba, and Hiroshi Saruwatari (May 2020). "Humangan: Generative Adversarial Network With Human-Based Discriminator And Its Evaluation In Speech Perception Modeling". In: *IEEE International Conference on Acoustics, Speech and Signal Processing*. Barcelona, Spain: IEEE, pp. 6239–6243.

Gadosey, Pius Kwao, Yujian Li, Enock Adjei Agyekum, Ting Zhang, Zhaoying Liu, Peter T Yamak, and Firdaous Essaf (2020). "SD-UNET: Stripping down U-net for segmentation of biomedical images on platforms with low computational budgets". In: *Diagnostics* 10.2, p. 110.

Galteri, Leonardo, Lorenzo Seidenari, Marco Bertini, and Alberto Del Bimbo (Oct. 2017). "Deep Generative Adversarial Compression Artifact Removal". en. In: *IEEE International Conference on Computer Vision*. Venice: IEEE, pp. 4836–4845.

García, Elkin and Fernando Lozano (2007). "Boosting Support Vector Machines". In: *Mldm posters*, pp. 153–167.

Ghatak, Swagata, Nima Dolatabadi, Dorit Trudler, XiaoTong Zhang, Yin Wu, Madhav Mohata, Rajesh Ambasudhan, Maria Talantova, and Stuart A Lipton (Nov. 2019). "Mechanisms of hyperexcitability in Alzheimer's disease hiPSC-derived neurons and cerebral organoids vs isogenic controls". In: *eLife* 8.

Ghrist, Robert (Oct. 2007). "Barcodes: The persistent topology of data". In: *Bulletin of the American Mathematical Society* 45.01, pp. 61–76.

Giandomenico, Stefano L and Madeline A Lancaster (Feb. 2017). "Probing human brain evolution and development in organoids". In: *Current Opinion in Cell Biology* 44, pp. 36–43.

Giandomenico, Stefano L., Magdalena Sutcliffe, and Madeline A. Lancaster (Feb. 2021). "Generation and long-term culture of advanced cerebral organoids for studying later stages of neural development". en. In: *Nature Protocols* 16.2, pp. 579–602.

Gilroy, Kyle D (2019). "Overcoming Shot noise Limitations with Bright Field Mode". en. In: *Vision Research*, p. 5.

Gomez-Giro, Gemma, Jonathan Arias-Fuenzalida, Javier Jarazo, Dagmar Zeuschner, Muhammad Ali, Nina Possemis, Silvia Bolognin, Rashi Halder, Christian Jäger, Willemijn F. E. Kuper, Peter M. van Hasselt, Holm Zaehres, Antonio del Sol, Herman van der Putten, Hans R. Schöler, and Jens C. Schwamborn (Dec. 2019). "Synapse alterations precede neuronal damage and storage pathology in a human cerebral organoid model of CLN3-juvenile neuronal ceroid lipofuscinosis". In: *Acta Neuropathologica Communications* 7.1.

Gonzales, Rafael C and Paul Wintz (1987). *Digital image processing*. Addison-Wesley Longman Publishing Co., Inc.

Goodfellow, Ian, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Y. Bengio (June 2014). "Generative Adversarial Networks". In: *Advances in Neural Information Processing Systems* 3.

Gordillo, Nelly, Eduard Montseny, and Pilar Sobrevilla (Oct. 2013). "State of the art survey on MRI brain tumor segmentation". In: *Magnetic Resonance Imaging* 31.8, pp. 1426–1438.

Grenier, Karl, Jennifer Kao, and Phedias Diamandis (Feb. 2020). "Three-dimensional modeling of human neurodegeneration: brain organoids coming of age". In: *Molecular Psychiatry* 25.2, pp. 254–274.

Groveman, Bradley R., Simote T. Foliaki, Christina D. Orru, Gianluigi Zanusso, James A. Carroll, Brent Race, and Cathryn L. Haigh (Dec. 2019). "Sporadic Creutzfeldt-Jakob disease prion infection of human cerebral organoids". In: *Acta Neuropathologica Communications* 7.1.

Gu, Feng, Nikolay Burlutskiy, Mats Andersson, and Lena Kajland Wilén (2018). "Multi-resolution networks for semantic segmentation in whole slide images". In: *Computational Pathology and Ophthalmic Medical Image Analysis: First International Workshop, COMPAY 2018, and 5th International Workshop, OMIA 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16-20, 2018, Proceedings 5*. Springer, pp. 11–18.

Gu, Yuchong, Zitao Zeng, Haibin Chen, Jun Wei, Yaqin Zhang, Binghui Chen, Yingqin Li, Yujuan Qin, Qing Xie, Zhuoren Jiang, et al. (2020). "MedSRGAN: medical images super-resolution using generative adversarial networks". In: *Multimedia Tools and Applications* 79, pp. 21815–21840.

Gu, Zaiwang, Jun Cheng, Huazhu Fu, Kang Zhou, Huaying Hao, Yitian Zhao, Tianyang Zhang, Shenghua Gao, and Jiang Liu (2019). "Ce-net: Context encoder network for 2d medical image segmentation". In: *IEEE transactions on medical imaging* 38.10, pp. 2281–2292.

Gulrajani, Ishaan, Faruk Ahmed, Martin Arjovsky, Vincent Dumoulin, and Aaron Courville (2017). "Improved Training of Wasserstein GANs". In: *Proceedings of the 31st International Conference on Neural Information Processing Systems*. NIPS'17. Long Beach, California, USA, pp. 5769–5779.

Hailstone, Martin, Dominic Waithe, Tamsin J Samuels, Lu Yang, Ita Costello, Yoav Arava, Elizabeth Robertson, Richard M Parton, and Ilan Davis (May 2020). "CytoCensus, mapping cell identity and division in tissues and organs using machine learning". In: *eLife* 9.

Hamdi, Ramzi, Clément Bouvier, Thierry Delzescaux, and Cédric Clouchoux (2021). "Histopathology learning dataset augmentation: a hybrid approach". In: *MIDL*.

Hamdi, Ramzi, Clément Bouvier, Thierry Delzesceaux, and Cédric Clouchoux (2021). "Evaluation of data augmentation for GAN with limited histological datasets". In: *miua*.

Hasnain, Zaki, Andrew K Fraser, Dan Georgess, Alex Choi, Paul Macklin, Joel S Bader, Shelly R Peyton, Andrew J Ewald, and Paul K Newton (May 2020). "OrgDyn: feature- and model-based characterization of spatial and temporal organoid dynamics". In: *Bioinformatics* 36.10. Ed. by Alfonso Valencia, pp. 3292–3294.

Hausdorff, Felix (2005). *Set theory*. Vol. 119. American Mathematical Soc.

Heusel, Martin, Hubert Ramsauer, Thomas Unterthiner, Bernhard Nessler, and Sepp Hochreiter (2017). "GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium". In: *Proceedings of the 31st International Conference on Neural Information Processing Systems*. NIPS'17. Long Beach, California, USA: Curran Associates Inc., pp. 6629–6640. ISBN: 9781510860964.

Hillman, Elizabeth MC, Venkatakaushik Voleti, Wenze Li, and Hang Yu (2019). "Light-sheet microscopy in neuroscience". In: *Annual review of neuroscience* 42, p. 295.

Hinton, Geoffrey and Sam Roweis (Jan. 2003). "Stochastic Neighbor Embedding". en. In: *Advances in neural information processing systems* 0.no, pp. 857–864.

Hoang, Hong Hai and Hoang Hieu Trinh (Feb. 2021). "Improvement for Convolutional Neural Networks in Image Classification Using Long Skip Connection". en. In: *Applied Sciences* 11.5, p. 2092.

Hong, Yean Ju and Jeong Tae Do (Dec. 2019). "Neural Lineage Differentiation From Pluripotent Stem Cells to Mimic Human Brain Tissues". In: *Frontiers in Bioengineering and Biotechnology* 7.

Houdeville, Charles, Romain Leenhardt, Marc Souchaud, Guillaume Velut, Nicolas Carbonell, Isabelle Nion-Larmurier, Alexandre Nuzzo, Aymeric Histace, Philippe Marteau, and Xavier Dray (May 2022). "Evaluation by a Machine Learning System of Two Preparations for Small Bowel Capsule Endoscopy: The BUBS (Burst Unpleasant Bubbles with Simethicone) Study". en. In: *Journal of Clinical Medicine* 11.10, p. 2822.

Hu, Bo, Ye Tang, Eric I-Chao Chang, Yubo Fan, Maode Lai, and Yan Xu (2019). "Unsupervised Learning for Cell-Level Visual Representation in Histopathology Images With Generative Adversarial Networks". In: *IEEE Journal of Biomedical and Health Informatics* 23.3, pp. 1316–1328.

Hu, Jingfei, Hua Wang, Jie Wang, Yunqi Wang, Fang He, and Jicong Zhang (2021). "SA-Net: A scale-attention network for medical image segmentation". In: *PloS one* 16.4, e0247388.

Huang, Huimin, Lanfen Lin, Ruofeng Tong, Hongjie Hu, Qiaowei Zhang, Yutaro Iwamoto, Xianhua Han, Yen-Wei Chen, and Jian Wu (2020). "Unet 3+: A full-scale connected unet for medical image segmentation". In: *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, pp. 1055–1059.

Hussain, Zeshan, Francisco Gimenez, Darvin Yi, and Daniel Rubin (2017). "Differential data augmentation techniques for medical imaging classification tasks". In: *AMIA annual symposium proceedings* 2017, p. 979.

Ibtehaz, Nabil and M Sohel Rahman (2020). "MultiResUNet: Rethinking the U-Net architecture for multimodal biomedical image segmentation". In: *Neural networks* 121, pp. 74–87.

Iefremova, Vira, George Manikakis, Olivia Krefft, Ammar Jabali, Kevin Weynans, Ruven Wilkens, Fabio Marsoner, Björn Brändl, Franz-Josef Müller, Philipp Koch, and Julia Ladewig (Apr. 2017). "An Organoid-Based Model of Cortical Development Identifies Non-Cell-Autonomous Defects in Wnt Signaling Contributing to Miller-Dieker Syndrome". In: *Cell Reports* 19.1, pp. 50–59.

Immerkær, John (Sept. 1996). "Fast Noise Variance Estimation". en. In: *Computer Vision and Image Understanding* 64.2, pp. 300–302.

Iniesta, Raquel, Ewan Carr, Mathieu Carriere, Naya Yerolemou, Bertrand Michel, and Frédéric Chazal (2022). "Topological Data Analysis and its usefulness for precision medicine studies". In: *Sort: Statistics and Operations Research Transactions*.

Iqbal, Asim, Romesa Khan, and Theofanis Karayannis (June 2019). "Developing a brain atlas through deep learning". In: *Nature Machine Intelligence* 1.6, pp. 277–287.

Isensee, Fabian, Paul F Jaeger, Simon AA Kohl, Jens Petersen, and Klaus H Maier-Hein (2021). "nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation". In: *Nature methods* 18.2, pp. 203–211.

Isola, Phillip, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros (2017). "Image-to-image translation with conditional adversarial networks". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1125–1134.

Jacob, Fadi, Ryan D. Salinas, Daniel Y. Zhang, Phuong T.T. Nguyen, Jordan G. Schnoll, Samuel Zheng Hao Wong, Radhika Thokala, Saad Sheikh, Deeksha Saxena, Stefan Prokop, Di-ao Liu, Xuyu Qian, Dmitriy Petrov, Timothy Lucas, H. Isaac Chen, Jay F. Dorsey, Kimberly M. Christian, Zev A. Binder, MacLean Nasrallah, Steven Brem, Donald M. O'Rourke, Guo-li Ming, and Hongjun Song (Jan. 2020). "A Patient-Derived Glioblastoma Organoid Model and Biobank Recapitulates Inter- and Intra-tumoral Heterogeneity". In: *Cell* 180.1, 188–204.e22.

Jacob, Pierre, David Picard, Aymeric Histace, and Edouard Klein (Aug. 2019). "Metric Learning With HORDE: High-Order Regularizer for Deep Embeddings". en. In: *arXiv:1908.02735 [cs]*. arXiv: 1908.02735.

Jasim, Muna Khalid (Dec. 2019). "Image Noise Removal Techniques : A Comparative Analysis". en. In: *International Journal of Science and Applied Information Technology* 8.6, pp. 24–29.

Jégou, Simon, Michal Drozdzal, David Vazquez, Adriana Romero, and Yoshua Bengio (2017). "The one hundred layers tiramisu: Fully convolutional densenets for semantic segmentation". In: *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pp. 11–19.

Jiang, Mingfeng, Minghao Zhi, Liying Wei, Xiaocheng Yang, Jucheng Zhang, Yongming Li, Pin Wang, Jiahao Huang, and Guang Yang (2021). "FA-GAN: Fused attentive generative adversarial networks for MRI image super-resolution". In: *Computerized Medical Imaging and Graphics* 92, p. 101969.

Jirik, Miroslav, Ivan Gruber, Vladimira Moulisova, Claudia Schindler, Lenka Cervenkova, Richard Palek, Jachym Rosendorf, Janine Arlt, Lukas Bolek, Jiri Dejmek, et al. (2020). "Semantic segmentation of intralobular and extralobular tissue from liver scaffold H&E images". In: *Sensors* 20.24, p. 7063.

Johnson, G. Allan, Alexandra Badea, Jeffrey Brandenburg, Gary Cofer, Boma Fubara, Song Liu, and Jonathan Nissanov (Nov. 2010). "Waxholm Space: An image-based reference for coordinating mouse brain research". en. In: *NeuroImage* 53.2, pp. 365–372.

Jorfi, Mehdi, Carla D'Avanzo, Rudolph E. Tanzi, Doo Yeon Kim, and Daniel Irimia (Dec. 2018). "Human Neurospheroid Arrays for In Vitro Studies of Alzheimer's Disease". In: *Scientific Reports* 8.1.

Joyce, James M. (2011). "Kullback-Leibler Divergence". In: *International Encyclopedia of Statistical Science*. Berlin, Heidelberg: Springer Berlin Heidelberg. Chap. 2, pp. 720–722.

Kadoshima, T., H. Sakaguchi, T. Nakano, M. Soen, S. Ando, M. Eiraku, and Y. Sasai (Dec. 2013). "Self-organization of axial polarity, inside-out layer pattern, and species-specific progenitor dynamics in human ES cell-derived neocortex". In: *Proceedings of the National Academy of Sciences* 110.50, pp. 20284–20289.

Kalaydina, Regina-Veronicka, Hedi Zhou, Elena Markvicheva, Sergey Burov, Farhana Zulkernine, and Myron R Szewczuk (Dec. 2019). "Impact of Fucosylation on Self-Assembly of Prostate and Breast Tumor Spheroids by Using Cyclo-RGDfK(TPP) Peptide and Image Object Detection". In: *OncoTargets and Therapy* Volume 12, pp. 11153–11173.

Kapalczynska, Marta, Tomasz Kolenda, Weronika Przybyla, Maria Zajaczkowska, Anna Teresiak, Violetta Filas, Matthew Ibbs, Renata Blizniak, Lukasz Luczewski, and Katarzyna Lamperska (2016). "2D and 3D cell cultures – a comparison of different types of cancer cell cultures". In: *Archives of Medical Science*.

Karaoğlu, Onur, Hasan Şakir Bilge, and İhsan Uluer (2022). "Removal of speckle noises from ultrasound images using five different deep learning networks". In: *Engineering Science and Technology, an International Journal* 29, p. 101030.

Kartasalo, Kimmo, Leena Latonen, Jorma Vihinen, Tapio Visakorpi, Matti Nykter, and Pekka Ruusuvuori (Sept. 2018). "Comparative analysis of tissue reconstruction algorithms for 3D histology". In: *Bioinformatics* 34.17. Ed. by Robert Murphy, pp. 3013–3021.

Karzbrun, Eyal, Aditya Kshirsagar, Sidney R. Cohen, Jacob H. Hanna, and Orly Reiner (May 2018). "Human brain organoids on a chip reveal the physics of folding". In: *Nature Physics* 14.5, pp. 515–522.

Kassis, Timothy, Victor Hernandez-Gordillo, Ronit Langer, and Linda G. Griffith (Dec. 2019). "OrgaQuant: Human Intestinal Organoid Localization and Quantification Using Deep Convolutional Neural Networks". In: *Scientific Reports* 9.1.

Kelava, Iva and Madeline A. Lancaster (Dec. 2016a). "Dishing out mini-brains: Current progress and future prospects in brain organoid research". In: *Developmental Biology* 420.2, pp. 199–209.

— (June 2016b). "Stem Cell Models of Human Brain Development". In: *Cell Stem Cell* 18.6, pp. 736–748.

Kerfoot, Eric, James Clough, Ilkay Oksuz, Jack Lee, Andrew P King, and Julia A Schnabel (2019). "Left-ventricle quantification using residual U-Net". In: *Statistical Atlases and Computational Models of the Heart. Atrial Segmentation and LV Quantification Challenges: 9th International Workshop, STACOM 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16, 2018, Revised Selected Papers 9*. Springer, pp. 371–380.

Kielkowski, Pavel, Isabel Y. Buchsbaum, Volker C. Kirsch, Nina C. Bach, Micha Drukker, Silvia Cappello, and Stephan A. Sieber (Dec. 2020). "FICD activity and AMPylation remodelling modulate human neurogenesis". In: *Nature Communications* 11.1.

Kim, Hee E, Alejandro Cosa-Linan, Nandhini Santhanam, Mahboubeh Jannesari, Mate E Maros, and Thomas Ganslandt (2022). "Transfer learning for medical image classification: A literature review". In: *BMC medical imaging* 22.1, p. 69.

Kim, Hyosung, Ranjie Xu, Ragunathan Padmashri, Anna Dunaevsky, Ying Liu, Cheryl F. Dreyfus, and Peng Jiang (May 2019). "Pluripotent Stem Cell-Derived Cerebral Organoids

Reveal Human Oligodendrogenesis with Dorsal and Ventral Origins". In: *Stem Cell Reports* 12.5, pp. 890–905.

Kingma, Diederik and Max Welling (2013). "Auto-encoding variational bayes". In: *arXiv preprint arXiv:1312.6114*.

Kostović, I., G. Sedmak, and M. Judaš (2019). "Neural histology and neurogenesis of the human fetal and infant brain". In: *NeuroImage* 188, pp. 743–773.

Krefft, Olivia, Ammar Jabali, Vira Iefremova, Philipp Koch, and Julia Ladewig (Jan. 2018). "Generation of Standardized and Reproducible Forebrain-type Cerebral Organoids from Human Induced Pluripotent Stem Cells". In: *Journal of Visualized Experiments* 131.131.

Krieger, Teresa G, Stephan M Tirier, Jeongbin Park, Katharina Jechow, Tanja Eisemann, Peter Angel, Roland Eils, and Christian Conrad (2020). "Modeling glioblastoma invasion using human brain organoids and single-cell transcriptomics". In: *Neuro Oncol*, p. 24.

Krivov, Egor, Maxim Pisov, and Mikhail Belyaev (2018). "MRI Augmentation via Elastic Registration for Brain Lesions Segmentation". en. In: *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*. Ed. by Alessandro Crimi, Spyridon Bakas, Hugo Kuijf, Bjoern Menze, and Mauricio Reyes. Vol. 10670. Series Title: Lecture Notes in Computer Science. Cham: Springer International Publishing, pp. 369–380.

Kupyn, Orest, Volodymyr Budzan, Mykola Mykhailych, Dmytro Mishkin, and Jiri Matas (Apr. 2018). "DeblurGAN: Blind Motion Deblurring Using Conditional Adversarial Networks". en. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 8183–8192.

Kupyn, Orest, Tetiana Martyniuk, Junru Wu, and Zhangyang Wang (Oct. 2019). "DeblurGAN-v2: Deblurring (Orders-of-Magnitude) Faster and Better". In: *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*. Seoul, Korea (South): IEEE, pp. 8877–8886.

Kyunghyun, Cho, van Merrienboer Bart, Bahdanau Dzmitry, Bougares Fethi, Schwenk Holger, and Bengio Yoshua (2014). "Learning phrase representations using rnn encoder–decoder for statistical machine translation". In: *In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*, pp. 1724–1734.

LaMonica, Bridget E, Jan H Lui, David V Hansen, and Arnold R Kriegstein (2013). "Mitotic spindle orientation predicts outer radial glial cell generation in human neocortex". en. In: *Nature communications*, p. 21.

Lan, Lan, Lei You, Zeyang Zhang, Zhiwei Fan, Weiling Zhao, Nianyin Zeng, Yidong Chen, and Xiaobo Zhou (May 2020). "Generative Adversarial Networks and Its Applications in Biomedical Informatics". en. In: *Frontiers in Public Health* 8.

Lancaster, M. A. and J. A. Knoblich (July 2014). "Organogenesis in a dish: Modeling development and disease using organoid technologies". In: *Science* 345.6194, pp. 1247125–1247125.

Lancaster, Madeline A, Nina S Corsini, Simone Wolfinger, E Hilary Gustafson, Alex W Phillips, Thomas R Burkard, Tomoki Otani, Frederick J Livesey, and Juergen A Knoblich (May 2017). "Guided self-organization and cortical plate formation in human brain organoids". In: *Nature Biotechnology* 35.7, pp. 659–666.

Lancaster, Madeline A and Juergen A Knoblich (Sept. 2014). "Generation of cerebral organoids from human pluripotent stem cells". In: *Nature Protocols* 9.10, pp. 2329–2340.

Lancaster, Madeline A. (May 2018). "Crinkle-Cut Brain Organoids". In: *Cell Stem Cell* 22.5, pp. 616–618.

Lancaster, Madeline A., Magdalena Renner, Carol-Anne Martin, Daniel Wenzel, Louise S. Bicknell, Matthew E. Hurles, Tessa Homfray, Josef M. Penninger, Andrew P. Jackson,

and Juergen A. Knoblich (Sept. 2013). "Cerebral organoids model human brain development and microcephaly". In: *Nature* 501.7467, pp. 373–379.

Larsen, Anders Boesen Lindbo, Soren Kaae Sonderby, Hugo Larochelle, and Ole Winther (2016). *Autoencoding beyond pixels using a learned similarity metric*, pp. 1558–1566.

Leenhardt, Romain, Pauline Vasseur, Cynthia Li, Jean Christophe Saurin, Gabriel Rahmi, Franck Cholet, Aymeric Becq, Philippe Marteau, Aymeric Histace, Xavier Dray, Sylvie Sacher-Huvelin, Farida Mesli, Chloé Leandri, Isabelle Nion-Larmurier, Stéphane Lecleire, Romain Gerard, Clotilde Duburque, Geoffroy Vanbiervliet, Xavier Amiot, Jean Philippe Le Mouel, Michel Delvaux, Pierre Jacob, Camille Simon-Shane, and Olivier Romain (Jan. 2019). "A neural network algorithm for detection of GI angiectasia during small-bowel capsule endoscopy". In: *Gastrointestinal Endoscopy* 89.1, pp. 189–194.

Li, Rui, Le Sun, Ai Fang, Peng Li, Qian Wu, and Xiaoqun Wang (Nov. 2017). "Recapitulating cortical development with organoid culture in vitro and modeling abnormal spindle-like (ASPM related primary) microcephaly disease". In: *Protein & Cell* 8.11, pp. 823–833.

Li, Shuang, Min Wang, and Junmei Zhou (Apr. 2020). "Brain Organoids: A Promising Living Biobank Resource for Neuroscience Research". In: *Biopreservation and Biobanking* 18.2, pp. 136–143.

Li, Weizhe, Ronald N. Germain, and Michael Y. Gerner (Aug. 2017). "Multiplex, quantitative cellular analysis in large tissue volumes with clearing-enhanced 3D microscopy (C $_{e}$ 3D)". In: *Proceedings of the National Academy of Sciences* 114.35, E7321–E7330.

Li, Yun, Julien Muffat, Attya Omer, Irene Bosch, Madeline A. Lancaster, Mriganka Sur, Lee Gehrke, Juergen A. Knoblich, and Rudolf Jaenisch (Mar. 2017). "Induction of Expansion and Folding in Human Cerebral Organoids". In: *Cell Stem Cell* 20.3, 385–396.e3.

Li, Zewen, Fan Liu, Wenjie Yang, Shouheng Peng, and Jun Zhou (2021). "A survey of convolutional neural networks: analysis, applications, and prospects". In: *IEEE transactions on neural networks and learning systems*.

Liba, Orly, Matthew D Lew, Elliott D SoRelle, Rebecca Dutta, Debasish Sen, Darius M Moshfeghi, Steven Chu, and Adam de La Zerda (2017). "Speckle-modulating optical coherence tomography in living mice and humans". In: *Nature communications* 8.1, p. 15845.

Lin, Tsung-Yi, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár (Feb. 2018). "Focal Loss for Dense Object Detection". In: *arXiv:1708.02002 [cs]*. arXiv: 1708.02002.

Lin, Yuan-Ta, Jinsoo Seo, Fan Gao, Heather M. Feldman, Hsin-Lan Wen, Jay Penney, Hugh P. Cam, Elizabeta Gjoneska, Waseem K. Raja, Jemmie Cheng, Richard Rueda, Oleg Kritskiy, Fatema Abdurrob, Zhuyu Peng, Blerta Milo, Chung Jong Yu, Sara Elmsaouri, Dilip Dey, Tak Ko, Bruce A. Yankner, and Li-Huei Tsai (June 2018). "APOE4 Causes Widespread Molecular and Cellular Alterations Associated with Alzheimer's Disease Phenotypes in Human iPSC-Derived Brain Cell Types". In: *Neuron* 98.6, 1141–1154.e7.

Liu, S. John, Martina Malatesta, Brian V. Lien, Parna Saha, Shivani S. Thombare, Sung Jun Hong, Leslie Pedraza, Mark Koontz, Kyounghee Seo, Max A. Horlbeck, Daniel He, Harjus S. Birk, Miten Jain, Hugh E. Olsen, Mark Akeson, Jonathan S. Weissman, Michelle Monje, Nalin Gupta, David R. Raleigh, Erik M. Ullian, and Daniel A. Lim (Dec. 2020). "CRISPRi-based radiation modifier screen identifies long non-coding RNA therapeutic targets in glioma". In: *Genome Biology* 21.1.

Lodato, Simona, Ashwin S. Shetty, and Paola Arlotta (Feb. 2015). "Cerebral cortex assembly: generating and reprogramming projection neuron diversity". In: *Trends in Neurosciences* 38.2, pp. 117–125.

Logan, Sarah, Thiago Arzua, Yasheng Yan, Congshan Jiang, Xiaojie Liu, Lai-Kang Yu, Qing-Song Liu, and Xiaowen Bai (May 2020). "Dynamic Characterization of Structural, Molecular, and Electrophysiological Phenotypes of Human-Induced Pluripotent Stem Cell-Derived Cerebral Organoids, and Comparison with Fetal and Adult Gene Profiles". In: *Cells* 9.5, p. 1301.

Lv, Jun, Jin Zhu, and Guang Yang (June 2021). "Which GAN? A comparative study of generative adversarial network-based fast MRI reconstruction". en. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 379.2200, p. 20200203.

Maas, AL, AY Hannun, and AY Ng (2013). "Rectifier nonlinearities improve neural network acoustic models". In: *Proceedings icml*, p. 3.

Maaten, Laurens van der and Geoffrey Hinton (2008). "Visualizing Data using tSNE". en. In: *Journal of Machine Learning Research* 9.1, pp. 2579–2605.

Mahmoud, Amira A, S EL Rabaie, TE Taha, O Zahran, FE Abd El-Samie, and W Al-Nauimy (2013). "Comparative study of different denoising filters for speckle noise reduction in ultrasonic b-mode images". In: *International Journal of Image, Graphics and Signal Processing* 5.2, p. 1.

Maier, Andreas, Christopher Syben, Tobias Lasser, and Christian Riess (2019). "A gentle introduction to deep learning in medical image processing". In: *Zeitschrift für Medizinische Physik* 29.2, pp. 86–101.

Makhzani, Alireza, Jonathon Shlens, Navdeep Jaitly, Ian Goodfellow, and Brendan Frey (2015). "Adversarial autoencoders". In: *arXiv preprint arXiv:1511.05644*.

Malhotra, Priyanka, Sheifali Gupta, Deepika Koundal, Atef Zaguia, and Wegayehu Enbeyle (2022). "Deep neural networks for medical image segmentation". In: *Journal of Healthcare Engineering* 2022.

Maller, Jerome J (2019). "Revealing the Hippocampal Connectome through Super-Resolution 1150-Direction Diffusion MRI". en. In: *Scientific reports*, p. 13.

Malm, Patrik, Anders Brun, and Ewert Bengtsson (2015). "Simulation of bright-field microscopy images depicting pap-smear specimen". In: *Cytometry Part A* 87.3, pp. 212–226.

Mansour, Abed AlFatah, J Tiago Gonçalves, Cooper W Bloyd, Hao Li, Sarah Fernandes, Daphne Quang, Stephen Johnston, Sarah L Parylak, Xin Jin, and Fred H Gage (May 2018). "An in vivo model of functional and vascularized human brain organoids". In: *Nature Biotechnology* 36.5, pp. 432–441.

Mao, Xudong, Qing Li, Haoran Xie, Raymond Y. K. Lau, Zhen Wang, and Stephen Paul Smolley (Apr. 2017). "Least Squares Generative Adversarial Networks". en. In: *IEEE International Conference on Computer Vision*, pp. 2813–2821.

Marotta, Nick, Soojin Kim, and Dimitri Krainc (Apr. 2020). "Organoid and pluripotent stem cells in Parkinson's disease modeling: an expert view on their value to drug discovery". In: *Expert Opinion on Drug Discovery* 15.4, pp. 427–441.

Marton, Rebecca M., Yuki Miura, Steven A. Sloan, Qingyun Li, Omer Revah, Rebecca J. Levy, John R. Huguenard, and Sergiu P. Pașca (Mar. 2019). "Differentiation and maturation of oligodendrocytes in human three-dimensional neural cultures". In: *Nature Neuroscience* 22.3, pp. 484–491.

Marton, Rebecca M. and Sergiu P. Pașca (Feb. 2020). "Organoid and Assembloid Technologies for Investigating Cellular Crosstalk in Human Brain Development and Disease". In: *Trends in Cell Biology* 30.2, pp. 133–143.

Masselink, Wouter, Daniel Reumann, Prayag Murawala, Pawel Pasierbek, Yuka Taniguchi, François Bonnay, Katharina Meixner, Jürgen A. Knoblich, and Elly M. Tanaka (Feb.

2019). “Broad applicability of a streamlined ethyl cinnamate-based clearing procedure”. In: *Development* 146.3, dev166884.

Matryba, Pawel, Leszek Kaczmarek, and Jakub Golab (Aug. 2019). “Advances in Ex Situ Tissue Optical Clearing”. In: *Laser & Photonics Reviews* 13.8, p. 1800292.

Menet, Sylvie, Philippe Saint-Marc, and Gerard Medioni (1990). “Active contour models: Overview, implementation and applications”. In: *1990 IEEE International Conference on Systems, Man, and Cybernetics Conference Proceedings*. IEEE, pp. 194–199.

Meyer, Fernand (July 1994). “Topographic distance and watershed lines”. In: *Signal Processing* 38.1, pp. 113–125.

Meziou, Leila, Aymeric Histace, Frédéric Precioso, Bogdan Matuszewski, and Franck Carreiras (2012). “3D Confocal Microscopy data analysis using level-set segmentation with alpha-divergence similarity measure”. In: *International Conference on Computer Vision Theory and Applications*, pp. 861–864.

Mikolajczyk, Agnieszka and Michal Grochowski (May 2018). “Data augmentation for improving deep learning in image classification problem”. en. In: *2018 International Interdisciplinary PhD Workshop (IIPhDW)*. Swinoujście: IEEE, pp. 117–122. ISBN: 978-1-5386-6143-7.

Milletari, Fausto, Nassir Navab, and Seyed-Ahmad Ahmadi (2016). “V-net: Fully convolutional neural networks for volumetric medical image segmentation”. In: *2016 fourth international conference on 3D vision (3DV)*. Ieee, pp. 565–571.

Mingqiang, Yang, Kpalma Kidiyo, and Ronsin Joseph (Nov. 2008). “A Survey of Shape Feature Extraction Techniques”. In: *Pattern Recognition Techniques, Technology and Applications*. Ed. by Peng-Yeng Yin. InTech.

Mirza, Mehdi and Simon Osindero (2014). “Conditional generative adversarial nets”. In: *arXiv preprint arXiv:1411.1784*.

Monzel, Anna S., Lisa M. Smits, Kathrin Hemmer, Siham Hachi, Edinson Lucumi Moreno, Thea van Wuellen, Javier Jarazo, Jonas Walter, Inga Brüggemann, Ibrahim Boussaad, Emanuel Berger, Ronan M.T. Fleming, Silvia Bolognin, and Jens C. Schwamborn (May 2017). “Derivation of Human Midbrain-Specific Organoids from Neuroepithelial Stem Cells”. In: *Stem Cell Reports* 8.5, pp. 1144–1154.

Moor, Michael, Max Horn, Bastian Rieck, and Karsten Borgwardt (2020). “Topological autoencoders”. In: *International conference on machine learning*. PMLR, pp. 7045–7054.

Nacereddine, Hammami, Bedda Mouldi, and Farah Nadir (2012). “Spoken Arabic digits recognition using MFCC based on GMM”. In: *In Sustainable Utilization and Development in Engineering and Technology, IEEE Conference*, pp. 160–163.

Nalepa, Jakub, Grzegorz Mrukwa, Szymon Piechaczek, Pablo Ribalta Lorenzo, Michal Marcinkiewicz, Barbara Bobek-Billewicz, Pawel Wawrzyniak, Pawel Ulrych, Janusz Szymanek, Marcin Cwiek, Wojciech Dudzik, Michal Kawulok, and Michael P. Hayball (2019). “Data Augmentation via Image Registration”. In: *2019 IEEE International Conference on Image Processing (ICIP)*, pp. 4250–4254.

Nguyen, Ngoc-Kim-Khanh and Marc Bui (2021). “Detecting Anomalies in the Dynamics of a Market Index with Topological Data Analysis”. In: *International Journal of Systematic Innovation* 6.6, p. 14.

Nickels, Sarah Louise, Jennifer Modamio, Bárbara Mendes-Pinheiro, Anna Sophia Monzel, Fay Betsou, and Jens Christian Schwamborn (July 2020). “Reproducible generation of human midbrain organoids for in vitro modeling of Parkinson’s disease”. In: *Stem Cell Research* 46, p. 101870.

Nowozin, Sebastian, Botond Cseke, and Ryota Tomioka (2016). "f-gan: Training generative neural samplers using variational divergence minimization". In: *Advances in neural information processing systems* 29.

Nürnberg, Elina, Mario Vitacolonna, Julia Klicks, Elena von Molitor, Tiziana Cesetti, Florian Keller, Roman Bruch, Torsten Ertongur-Fauth, Katja Riedel, Paul Scholz, Thorsten Lau, Richard Schneider, Julia Meier, Mathias Hafner, and Rüdiger Rudolf (Feb. 2020). "Routine Optical Clearing of 3D-Cell Cultures: Simplicity Forward". In: *Frontiers in Molecular Biosciences* 7.

Nzou, Goodwell, R. T. Wicks, E. E. Wicks, S. A. Seale, C. H. Sane, A. Chen, S. V. Murphy, J. D. Jackson, and A. J. Atala (Dec. 2018). "Human Cortex Spheroid with a Functional Blood Brain Barrier for High-Throughput Neurotoxicity Screening and Disease Modeling". In: *Scientific Reports* 8.1.

Okarma, Krzysztof, Piotr Lech, and Vladimir V. Lukin (Sept. 2021). "Combined Full-Reference Image Quality Metrics for Objective Assessment of Multiply Distorted Images". en. In: *Electronics* 10.18, p. 2256.

Oktay, Ozan, Jo Schlemper, Loic Le Folgoc, Matthew Lee, Mattias Heinrich, Kazunari Misawa, Kensaku Mori, Steven McDonagh, Nils Y Hammerla, Bernhard Kainz, et al. (2018). "Attention u-net: Learning where to look for the pancreas". In: *arXiv preprint arXiv:1804.03999*.

Ormel, Paul R., Renata Vieira de Sá, Emma J. van Bodegraven, Henk Karst, Oliver Harschnitz, Marjolein A. M. Sneeboer, Lill Eva Johansen, Roland E. van Dijk, Nicky Scheefhals, Amber Berdenis van Berlekom, Eduardo Ribes Martínez, Sandra Kling, Harold D. MacGillavry, Leonard H. van den Berg, René S. Kahn, Elly M. Hol, Lot D. de Witte, and R. Jeroen Pasterkamp (Dec. 2018). "Microglia innately develop within cerebral organoids". In: *Nature Communications* 9.1.

Ostrem, Bridget E L, Jan H Lui, Caitlyn C Gertz, and Arnold R Kriegstein (2015). "Control of outer radial glial stem cell mitosis in the human brain". en. In: *Cell reports*, p. 18.

Otsu, Nobuyuki (Jan. 1979a). "A Threshold Selection Method from Gray-Level Histograms". In: *IEEE Transactions on Systems, Man, and Cybernetics* 9.1, pp. 62–66.

— (1979b). "A threshold selection method from gray-level histograms". In: *IEEE transactions on systems, man, and cybernetics* 9.1, pp. 62–66.

Otter, Nina, Mason A Porter, Ulrike Tillmann, Peter Grindrod, and Heather A Harrington (Dec. 2017). "A roadmap for the computation of persistent homology". en. In: *EPJ Data Science* 6.1, p. 17.

Panagopoulos, Dimitrios (Jan. 2022). "Topological data analysis and clustering". en. In: *arXiv:2201.09054 [cs, math]*. arXiv: 2201.09054.

Park, Joseph, Isaac Wetzel, Ian Marriott, Didier Dréau, Carla D'Avanzo, Doo Yeon Kim, Rudolph E. Tanzi, and Hansang Cho (July 2018). "A 3D human triculture system modeling neurodegeneration and neuroinflammation in Alzheimer's disease". In: *Nature Neuroscience* 21.7, pp. 941–951.

Paşca, Anca M, Steven A Sloan, Laura E Clarke, Yuan Tian, Christopher D Makinson, Nina Huber, Chul Hoon Kim, Jin-Young Park, Nancy A O'Rourke, Khoa D Nguyen, Stephen J Smith, John R Huguenard, Daniel H Geschwind, Ben A Barres, and Sergiu P Paşca (July 2015). "Functional cortical neurons and astrocytes from human pluripotent stem cells in 3D culture". In: *Nature Methods* 12.7, pp. 671–678.

Pastore, Giuseppe, Fabio Cermelli, Yongqin Xian, Massimiliano Mancini, Zeynep Akata, and Barbara Caputo (2021). "A closer look at self-training for zero-label semantic segmentation". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 2693–2702.

Pedersen, Marius (2015). "Evaluation of 60 Full-Reference Image Quality Metrics on the CID:IQ". en. In: *IEEE International Conference on Image Processing*, p. 5.

Pedersen, Marius and Jon Hardeberg (Jan. 2012). "Full-Reference Image Quality Metrics: Classification and Evaluation". In: *Foundations and Trends® in Computer Graphics and Vision* 7, pp. 1–80.

Pedrosa, Carolina da S. G., Leticia R. Q. Souza, Tiago A. Gomes, Caroline V. F. de Lima, Pitia F. Ledur, Karina Karmirian, Jimena Barbeito-Andres, Marcelo do N. Costa, Luiza M. Higa, Atila D. Rossi, Maria Bellio, Amilcar Tanuri, Arnaldo Prata-Barbosa, Fernanda Tovar-Moll, Patricia P. Garcez, Flavio A. Lara, Renato J. R. Molica, and Stevens K. Rehen (Mar. 2020). "The cyanobacterial saxitoxin exacerbates neural cell death and brain malformations induced by Zika virus". In: *PLOS Neglected Tropical Diseases* 14.3. Ed. by Paulo F. P. Pimenta, e0008060.

Peláez, Rebeca, Ricardo Cao, and Juan M. Vilar (2021). "Nonparametric estimation of the probability of default with double smoothing". In: *SORT. Statistics and Operations Research Transactions* 45, pp. 93–120.

Pellegrini, Laura, Claudia Bonfio, Jessica Chadwick, Farida Begum, Mark Skehel, and Madeline A. Lancaster (June 2020). "Human CNS barrier-forming organoids with cerebrospinal fluid production". In: *Science*, eaaz5626.

Pham, Missy T., Kari M. Pollock, Melanie D. Rose, Whitney A. Cary, Heather R. Stewart, Ping Zhou, Jan A. Nolta, and Ben Waldau (May 2018). "Generation of human vascularized brain organoids:" in: *NeuroReport* 29.7, pp. 588–593.

Piccinini, Filippo, Tamas Balassa, Antonella Carbonaro, Akos Diosdi, Timea Toth, Nikita Moshkov, Ervin A. Tasnadi, and Peter Horvath (2020). "Software tools for 3D nuclei segmentation and quantitative analysis in multicellular aggregates". en. In: *Computational and Structural Biotechnology Journal* 18, pp. 1287–1300.

Pichat, Jonas, Juan Eugenio Iglesias, Tarek Yousry, Sébastien Ourselin, and Marc Modat (May 2018). "A Survey of Methods for 3D Histology Reconstruction". In: *Medical Image Analysis* 46, pp. 73–105.

Pluim, J.P.W., J.B.A. Maintz, and M.A. Viergever (2003). "Mutual-information-based registration of medical images: a survey". In: *IEEE Transactions on Medical Imaging* 22.8, pp. 986–1004.

Poli, Daniele, Chiara Magliaro, and Arti Ahluwalia (Mar. 2019). "Experimental and Computational Methods for the Study of Cerebral Organoids: A Review". In: *Frontiers in Neuroscience* 13.

Qi, Guo-Jun and Jiebo Luo (2020). "Small data challenges in big data era: A survey of recent progress on unsupervised and semi-supervised methods". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 44.4, pp. 2168–2187.

Qian, Xuyu, Ha Nam Nguyen, Mingxi M. Song, Christopher Hadiono, Sarah C. Ogden, Christy Hammack, Bing Yao, Gregory R. Hamersky, Fadi Jacob, Chun Zhong, Ki-jun Yoon, William Jeang, Li Lin, Yujing Li, Jai Thakor, Daniel A. Berg, Ce Zhang, Eunchai Kang, Michael Chickering, David Nauen, Cheng-Ying Ho, Zhexing Wen, Kimberly M. Christian, Pei-Yong Shi, Brady J. Maher, Hao Wu, Peng Jin, Hengli Tang, Hongjun Song, and Guo-li Ming (May 2016). "Brain-Region-Specific Organoids Using Mini-bioreactors for Modeling ZIKV Exposure". In: *Cell* 165.5, pp. 1238–1254.

Qian, Xuyu, Yijing Su, Christopher D. Adam, Andre U. Deutschmann, Sarshan R. Pather, Ethan M. Goldberg, Kenong Su, Shiying Li, Lu Lu, Fadi Jacob, Phuong T.T. Nguyen, Sooyoung Huh, Ahmet Hoke, Sarah E. Swinford-Jackson, Zhexing Wen, Xiaosong Gu, R. Christopher Pierce, Hao Wu, Lisa A. Briand, H. Isaac Chen, John A. Wolf, Hongjun

Song, and Guo-li Ming (Mar. 2020). “Sliced Human Cortical Organoids for Modeling Distinct Cortical Layer Formation”. In: *Cell Stem Cell*.

Qiu, Bingjiang, Jiapan Guo, Joep Kraeima, Haye Hendrik Glas, Weichuan Zhang, Ronald JH Borra, Max Johannes Hendrikus Witjes, and Peter MA van Ooijen (2021). “Recurrent convolutional neural networks for 3D mandible segmentation in computed tomography”. In: *Journal of personalized medicine* 11.6, p. 492.

Quadrato, Giorgia, Tuan Nguyen, Evan Z. Macosko, John L. Sherwood, Sung Min Yang, Daniel R. Berger, Natalie Maria, Jorg Scholvin, Melissa Goldman, Justin P. Kinney, Edward S. Boyden, Jeff W. Lichtman, Ziv M. Williams, Steven A. McCarroll, and Paola Arlotta (Apr. 2017). “Cell diversity and network dynamics in photosensitive human brain organoids”. In: *Nature* 545.7652, pp. 48–53.

Rad, Reza Moradi, Parvaneh Saeedi, Jason Au, and Jon Havelock (2020). “Trophectoderm segmentation in human embryo images via inceptioned U-Net”. In: *Medical Image Analysis* 62, p. 101612.

Radford, Alec, Luke Metz, and Soumith Chintala (2015). “Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks”. In: *CoRR*.

Raja, Waseem K., Alison E. Mungenast, Yuan-Ta Lin, Tak Ko, Fatema Abdurrob, Jinsoo Seo, and Li-Huei Tsai (Sept. 2016). “Self-Organizing 3D Human Neural Tissue Derived from Induced Pluripotent Stem Cells Recapitulate Alzheimer’s Disease Phenotypes”. In: *PLOS ONE* 11.9. Ed. by Jaya Padmanabhan, e0161969.

Rakotoson, Irina, Brigitte Delhomme, Philippe Djian, Andreas Deeg, Maia Brunstein, Christian Seebacher, Rainer Uhl, Clément Ricard, and Martin Oheim (Aug. 2019). “Fast 3-D Imaging of Brain Organoids With a New Single-Objective Planar-Illumination Two-Photon Microscope”. In: *Frontiers in Neuroanatomy* 13.

Redmon, Joseph and Ali Farhadi (Dec. 2016). “YOLO9000: Better, Faster, Stronger”. In: *arXiv:1612.08242 [cs]*. arXiv: 1612.08242.

— (2017). “YOLO9000: better, faster, stronger”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 7263–7271.

Ren, Jiong, Zhibin Niu, Xiaoqin Li, Jie Yang, Meijiao Gao, Xudong Li, Tao Zhang, Lei Fang, Boyang Zhang, Junping Wang, Yongping Su, and Fengchao Wang (Dec. 2018). “A novel morphometry system automatically assessing the growth and regeneration of intestinal organoids”. In: *Biochemical and Biophysical Research Communications* 506.4, pp. 1052–1058.

Renner, Henrik, Martha Grabos, Katharina J Becker, Theresa E Kagermeier, Jie Wu, Mandy Otto, Stefan Peischard, Dagmar Zeuschner, Yaroslav TsyTsyura, Paul Disse, Jürgen Klingauf, Sebastian A Leidel, Guiscard Seebohm, Hans R Schöler, and Jan M Bruder (Nov. 2020). “A fully automated high-throughput workflow for 3D-based chemical screening in human midbrain organoids”. en. In: *eLife* 9, e52904.

Renner, Magdalena, Madeline A Lancaster, Shan Bian, Heejin Choi, Taeyun Ku, Angela Peer, Kwanghun Chung, and Juergen A Knoblich (May 2017). “Self-organized developmental patterning and differentiation in cerebral organoids”. In: *The EMBO Journal* 36.10, pp. 1316–1329.

Rezaei, Mina, Konstantin Harmuth, Willi Gierke, Thomas Kellermeier, Martin Fischer, Haojin Yang, and Christoph Meinel (2018). “A conditional adversarial network for semantic segmentation of brain tumor”. In: *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: Third International Workshop, BrainLes 2017, Held in Conjunction with MICCAI 2017, Quebec City, QC, Canada, September 14, 2017, Revised Selected Papers 3*. Springer, pp. 241–252.

Richardson, Douglas S. and Jeff W. Lichtman (July 2015). "Clarifying Tissue Clearing". In: *Cell* 162.2, pp. 246–257.

Robins, Vanessa, Peter John Wood, and Adrian P Sheppard (2011). "Theory and algorithms for constructing discrete Morse complexes from grayscale digital images". In: *IEEE Transactions on pattern analysis and machine intelligence* 33.8, pp. 1646–1658.

Romain, Leenhardt, Koulaouzidis Anastasios, Histace Aymeric, Baatrup Gunnar, Beg Sabina, Bourreille Arnaud, de Lange Thomas, Eliakim Rami, Iakovidis Dimitris, Dam Jensen Michael, Keuchel Martin, Margalit Yehuda Reuma, McNamara Deirdre, Mascarenhas Miguel, Spada Cristiano, Segui Santi, Smedsrud Pia, Toth Ervin, Eugenio Gian, Eyal Klang Tontini, Dray Xavier, and Kopylov Uri (2022). "Key research questions for implementation of artificial intelligence in capsule endoscopy". In: *Therapeutic Advances in Gastroenterology*.

Ronneberger, Olaf, Philipp Fischer, and Thomas Brox (May 2015a). "U-Net: Convolutional Networks for Biomedical Image Segmentation". In: *arXiv:1505.04597 [cs]*. arXiv: 1505.04597.

— (2015b). "U-net: Convolutional networks for biomedical image segmentation". In: *International Conference on Medical image computing and computer-assisted intervention*. Springer, pp. 234–241.

Rosen, Burke Q. and Eric Halgren (Jan. 2021). "A Whole-Cortex Probabilistic Diffusion Tractography Connectome". en. In: *eneuro* 8.1, ENEURO.0416–20.2020.

Rubel, Andrii, Oleg Ieremeiev, Vladimir Lukin, Jarosław Fastowicz, and Krzysztof Okarma (Feb. 2022). "Combined No-Reference Image Quality Metrics for Visual Quality Assessment Optimized for Remote Sensing Images". en. In: *Applied Sciences* 12.4, p. 1986.

Russakoff, Daniel B., Carlo Tomasi, Torsten Rohlfing, and Calvin R. Maurer (2004). "Image Similarity Using Mutual Information of Regions". In: *Computer Vision - ECCV 2004*. Ed. by Takeo Kanade, Josef Kittler, Jon M. Kleinberg, Friedemann Mattern, John C. Mitchell, Oscar Nierstrasz, C. Pandu Rangan, Bernhard Steffen, Madhu Sudan, Demetri Terzopoulos, Dough Tygar, Moshe Y. Vardi, Gerhard Weikum, Tomáš Pajdla, and Jiří Matas. Vol. 3023. Series Title: Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 596–607.

Sakaguchi, Hideya, Taisuke Kadoshima, Mika Soen, Nobuhiro Narii, Yoshihito Ishida, Masatoshi Ohgushi, Jun Takahashi, Mototsugu Eiraku, and Yoshiki Sasai (Dec. 2015). "Generation of functional hippocampal neurons from self-organizing human embryonic stem cell-derived dorsomedial telencephalic tissue". In: *Nature Communications* 6.1.

Sakaguchi, Hideya, Yuki Ozaki, Tomoka Ashida, Takayoshi Matsubara, Naotaka Oishi, Shunsuke Kihara, and Jun Takahashi (June 2019). "Self-Organized Synchronous Calcium Transients in a Cultured Human Neural Network Derived from Cerebral Organoids". In: *Stem Cell Reports*.

Salimans, Tim, Ian Goodfellow, Wojciech Zaremba, Vicki Cheung, Alec Radford, and Xi Chen (2016). "Improved techniques for training gans". In: *Advances in neural information processing systems* 29.

Sandfort, Veit, Ke Yan, Perry J Pickhardt, and Ronald M Summers (2019). "Data augmentation using generative adversarial networks (CycleGAN) to improve generalizability in CT segmentation tasks". In: *Scientific reports* 9.1, p. 16884.

Sartore, Rafaela C., Simone C. Cardoso, Yury V.M. Lages, Julia M. Paraguassu, Mariana P. Stelling, Rodrigo F. Madeiro da Costa, Marilia Z. Guimaraes, Carlos A. Pérez, and Stevens K. Rehen (Feb. 2017). "Trace elements during primordial plexiform network formation in human cerebral organoids". In: *PeerJ* 5, e2927.

Scheffer, Louis K et al. (Sept. 2020). "A connectome and analysis of the adult Drosophila central brain". en. In: *eLife* 9, e57443.

Schindelin, Johannes, Ignacio Arganda-Carreras, Erwin Frise, Verena Kaynig, Mark Longair, Tobias Pietzsch, Stephan Preibisch, Curtis Rueden, Stephan Saalfeld, Benjamin Schmid, Jean-Yves Tinevez, Daniel James White, Volker Hartenstein, Kevin Eliceiri, Pavel Tomancak, and Albert Cardona (July 2012). "Fiji: an open-source platform for biological-image analysis". In: *Nature Methods* 9.7, pp. 676–682.

Schlemper, Jo, Ozan Oktay, Michiel Schaap, Mattias Heinrich, Bernhard Kainz, Ben Glocker, and Daniel Rueckert (2019). "Attention gated networks: Learning to leverage salient regions in medical images". In: *Medical image analysis* 53, pp. 197–207.

Schmitz, Alexander, Sabine C. Fischer, Christian Mattheyer, Francesco Pampaloni, and Ernst H. K. Stelzer (May 2017). "Multiscale image analysis reveals structural heterogeneity of the cell microenvironment in homotypic spheroids". In: *Scientific Reports* 7.1.

Scott, Hannah, Dominic Branford, Niclas Westerberg, Jonathan Leach, and Erik M Gauger (2021). "Noise limits on two-photon interferometric sensing". In: *Physical Review A* 104.5, p. 053704.

Seto, Yusuke and Mototsugu Eiraku (Dec. 2019). "Toward the formation of neural circuits in human brain organoids". In: *Current Opinion in Cell Biology* 61, pp. 86–91.

Shaffrey, Cian W, Ian H Jermyn, and Nick G Kingsbury (2002). "Psychovisual Evaluation of Image Segmentation Algorithms". In: *Advanced Concepts for Intelligent Vision Systems*, p. 7.

Shang, Chao, Aaron Palmer, Jiangwen Sun, Ko-Shin Chen, Jin Lu, and Jinbo Bi (Dec. 2017). "VIGAN: Missing view imputation with generative adversarial networks". en. In: pp. 766–775.

Sharma, Arpit Kumar, Amita Nandal, Arvind Dhaka, Deepika Koundal, Dijana Capeska Bogatinoska, and Hashem Alyami (2022). "Enhanced watershed segmentation algorithm-based modified ResNet50 model for brain tumor detection". In: *BioMed Research International* 2022.

Shi, Yingchao, Le Sun, Mengdi Wang, Jianwei Liu, Suijuan Zhong, Rui Li, Peng Li, Lijie Guo, Ai Fang, Ruiguo Chen, Woo-Ping Ge, Qian Wu, and Xiaoqun Wang (May 2020). "Vascularized human cortical organoids (vOrganoids) model cortical development in vivo". In: *PLOS Biology* 18.5. Ed. by Bing Ye, e3000705.

Shorten, Connor and Taghi M Khoshgoftaar (2019). "A survey on image data augmentation for deep learning". In: *Journal of big data* 6.1, pp. 1–48.

Shuvo, Md Badiuzzaman, Rifat Ahommed, Sakib Reza, and MMA Hashem (2021). "CNL-UNet: a novel lightweight deep learning architecture for multimodal biomedical image segmentation with false output suppression". In: *Biomedical Signal Processing and Control* 70, p. 102959.

Siddique, Nahian, Paheding Sidike, Colin Elkin, and Vijay Devabhaktuni (2020). "U-Net and its variants for medical image segmentation: theory and applications". In: *arXiv preprint arXiv:2011.01118*.

Sidhaye, Jaydeep and Jürgen A. Knoblich (June 2020). "Brain organoids: an ensemble of bioassays to investigate human neurodevelopment and disease". In: *Cell Death & Differentiation*.

Singh, Nripendra Kumar and Khalid Raza (2021). "Medical Image Generation using Generative Adversarial Networks". en. In: *Health Informatics: A Computational Perspective in Healthcare*, p. 19.

Singh, Prabhishek, Manoj Diwakar, Achyut Shankar, Raj Shree, and Manoj Kumar (2021). "A Review on SAR Image and its Despeckling". In: *Archives of Computational Methods in Engineering* 28, pp. 4633–4653.

Sinha, Ashish and Jose Dolz (2020). "Multi-scale self-guided attention for medical image segmentation". In: *IEEE journal of biomedical and health informatics* 25.1, pp. 121–130.

Sivitilli, A., J.T. Gosio, B. Ghoshal, A. Evstratova, D. Trcka, P. Ghiasiighorveh, J. J. Hernandez, J. M. Beaulieu, J. L. Wrana, and L. Attisano (Mar. 2020). *Robust Production of Uniform Human Cerebral Organoids from Pluripotent Stem Cells*. preprint. Neuroscience.

Skaf, Yara and Reinhard Laubenbacher (May 2022). "Topological Data Analysis in Biomedicine: A Review". In: *Journal of Biomedical Informatics*, p. 104082.

Sloan, Steven A., Jimena Andersen, Anca M. Pașca, Fikri Birey, and Sergiu P. Pașca (Sept. 2018). "Generation and assembly of human brain region–specific three-dimensional cultures". In: *Nature Protocols* 13.9, pp. 2062–2085.

Sloan, Steven A., Spyros Darmanis, Nina Huber, Themasap A. Khan, Fikri Birey, Christine Caneda, Richard Reimer, Stephen R. Quake, Ben A. Barres, and Sergiu P. Paşca (Aug. 2017). "Human Astrocyte Maturation Captured in 3D Cerebral Cortical Spheroids Derived from Pluripotent Stem Cells". In: *Neuron* 95.4, 779–790.e6.

Smits, Lisa M., Lydia Reinhardt, Peter Reinhardt, Michael Glatza, Anna S. Monzel, Nancy Stanslowsky, Marcelo D. Rosato-Siri, Alessandra Zanon, Paul M. Antony, Jessica Bellmann, Sarah M. Nicklas, Kathrin Hemmer, Xiaobing Qing, Emanuel Berger, Norman Kalmbach, Marc Ehrlich, Silvia Bolognin, Andrew A. Hicks, Florian Wegner, Jared L. Sterneckert, and Jens C. Schwamborn (Dec. 2019). "Modeling Parkinson's disease in midbrain-like organoids". In: *npj Parkinson's Disease* 5.1.

Smits, Lisa Maria and Jens Christian Schwamborn (May 2020). "Midbrain Organoids: A New Tool to Investigate Parkinson's Disease". In: *Frontiers in Cell and Developmental Biology* 8.

Soetje, Birga, Joachim Fuellekrug, Dieter Haffner, and Wolfgang H. Ziegler (Mar. 2020). "Application and Comparison of Supervised Learning Strategies to Classify Polarity of Epithelial Cell Spheroids in 3D Culture". In: *Frontiers in Genetics* 11.

Sonderby, Casper Kaae, Jose Caballero, Lucas Theis, Wenzhe Shi, and Ferenc Huszar (2016). "Amortised map inference for image super-resolution". In: *arXiv preprint arXiv:1610.04490*.

Stachowiak, E. K., C. A. Benson, S. T. Narla, A. Dimitri, L. E. Bayona Chuye, S. Dhiman, K. Harikrishnan, S. Elahi, D. Freedman, K. J. Brennand, P. Sarder, and M. K. Stachowiak (Nov. 2017). "Cerebral organoids reveal early cortical maldevelopment in schizophrenia—computational anatomy and genomics, role of FGFR1". In: *Translational Psychiatry* 7.11.

Stan, Ana D., Subroto Ghose, Xue-Min Gao, Rosalinda C. Roberts, Kelly Lewis-Amezcua, Kimmo J. Hatanpaa, and Carol A. Tamminga (Dec. 2006). "Human postmortem tissue: What quality markers matter?" In: *Brain Research* 1123.1, pp. 1–11.

Stanly, Tess A., Marco Fritzsche, Suneale Banerji, Esther García, Jorge Bernardino de la Serna, David G. Jackson, and Christian Eggeling (Sept. 2016). "Critical importance of appropriate fixation conditions for faithful imaging of receptor microclusters". In: *Biology Open* 5.9, pp. 1343–1350.

Summers, Cecilia and Michael J. Dinneen (2019). "Improved Mixed-Example Data Augmentation". In: *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*. Vol. -. -, pp. 1262–1270.

Sun, Jesse, Fatemeh Darbehani, Mark Zaidi, and Bo Wang (2020). "Saunet: Shape attentive u-net for interpretable medical image segmentation". In: *Medical Image Computing and*

Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part IV 23. Springer, pp. 797–806.

Susu Yao, Weisi Lin, EePing Ong, and Zhongkang Lu (2005). "Contrast signal-to-noise ratio for image quality assessment". In: *IEEE International Conference on Image Processing 2005*. Genova, Italy: IEEE, pp. I–397.

Tajbakhsh, Nima, Jae Y. Shin, Suryakanth R. Gurudu, R. Todd Hurst, Christopher B. Kendall, Michael B. Gotway, and Jianming Liang (2016). "Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?" In: *IEEE Transactions on Medical Imaging*.

Takahashi, Ryo, Takashi Matsubara, and Kuniaki Uehara (2020). "Data Augmentation Using Random Image Cropping and Patching for Deep CNNs". In: *IEEE Transactions on Circuits and Systems for Video Technology* 30.9, pp. 2917–2931.

Tanh, D. N. H and S. D. Dvoenko (2015). "A Variational Method to Remove the Combination of Poisson and Gaussian Noises:" en. In: *Proceedings of the 5th International Workshop on Image Mining. Theory and Applications*. Berlin, Germany: SCITEPRESS - Science, pp. 38–45. ISBN: 978-989-758-094-9.

Taylor, Luke and Geoff Nitschke (Nov. 2018). "Improving Deep Learning with Generic Data Augmentation". en. In: *2018 IEEE Symposium Series on Computational Intelligence (SSCI)*. Bangalore, India: IEEE, pp. 1542–1547.

Thibault, Guillaume, Bernard Fertil, Claire Navarro, Sandrine Pereira, Pierre Cau, Nicolas Levy, Jean Sequeira, and Jean-Luc Mari (Feb. 2013). "Shape and Texture Indexes Application to Cell Nuclei Classification". In: *International Journal of Pattern Recognition and Artificial Intelligence* 27.01, p. 1357002.

Thorn, Kurt (Jan. 2016). "A quick guide to light microscopy in cell biology". In: *Molecular Biology of the Cell* 27.2. Ed. by Doug Kellogg, pp. 219–222.

Tian, Ai, Julien Muffat, and Yun Li (Feb. 2020). "Studying Human Neurodevelopment and Diseases Using 3D Brain Organoids". In: *The Journal of Neuroscience* 40.6, pp. 1186–1193.

Tian, Yu, Zhangkai Ni, Baoliang Chen, Shiqi Wang, Hanli Wang, and Sam Kwong (Jan. 2022). "Generalized Visual Quality Assessment of GAN-Generated Face Images". In: *arXiv:2201.11975 [cs, eess]*. arXiv: 2201.11975.

Tralie, Christopher, Nathaniel Saul, and Rann Bar-On (Sept. 2018). "Ripser.py: A Lean Persistent Homology Library for Python". In: *Journal of Open Source Software* 3.29, p. 925.

Trebing, Kevin, Tomasz Stanczyk, and Siamak Mehrkanoon (2021). "SmaAt-UNet: Precipitation nowcasting using a small attention-UNet architecture". In: *Pattern Recognition Letters* 145, pp. 178–186.

Tsomko, Elena, Hyoung Joong Kim, Joonki Paik, and In-Kwon Yeo (2008). "Efficient Method of Detecting Blurry Images". en. In: *Journal of Ubiquitous Convergence Technology* 2.1, p. 14.

Turner, Katharine (2020). "Medians of populations of persistence diagrams". In: *Homology, Homotopy and Applications* 22.1, pp. 255–282.

Tustison, Nicholas J., Brian B. Avants, Zixuan Lin, Xue Feng, Nicholas Cullen, Jaime F. Mata, Lucia Flors, James C. Gee, Talissa A. Altes, John P. Mugler III, and Kun Qing (Mar. 2019). "Convolutional Neural Networks with Template-Based Data Augmentation for Functional Lung Image Quantification". en. In: *Academic Radiology* 26.3, pp. 412–423.

Van Rijthoven, Mart, Maschenka Balkenhol, Karina Siliņa, Jeroen Van Der Laak, and Francesco Ciompi (2021). "HookNet: Multi-resolution convolutional neural networks for

semantic segmentation in histopathology whole-slide images". In: *Medical image analysis* 68, p. 101890.

Velasco, Silvia, Amanda J. Kedaigle, Sean K. Simmons, Allison Nash, Marina Rocha, Giorgia Quadrato, Bruna Paulsen, Lan Nguyen, Xian Adiconis, Aviv Regev, Joshua Z. Levin, and Paola Arlotta (June 2019). "Individual brain organoids reproducibly form cell diversity of the human cerebral cortex". In: *Nature*.

Von Bartheld, CS, J Bahney, and S Herculano-Houzel (2017). "The Search for True Numbers of Neurons and Glial Cells in the Human Brain: A Review of 150 Years of Cell Counting". en. In: *Journal of Comparative Neurology*, p. 56.

Vora, Pritesh, Bhavesh Oza, et al. (2013). "A survey on k-mean clustering and particle swarm optimization". In: *International Journal of Science and Modern Engineering* 1.3, pp. 24–26.

Waibel, Dominik Jens Elias, Sayedali Shetab Boushehri, and Carsten Marr (2021). "InstantDL: an easy-to-use deep learning pipeline for image segmentation and classification". In: *BMC bioinformatics* 22, pp. 1–15.

Wang, Fan, Huidong Liu, Dimitris Samaras, and Chao Chen (2020). "Topogan: A topology-aware generative adversarial network". In: *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part III 16*. Springer, pp. 118–136.

Wang, Jason, Luis Perez, et al. (2017). "The effectiveness of data augmentation in image classification using deep learning". In: *Convolutional Neural Networks Vis. Recognit* 11.2017, pp. 1–8.

Wang, Qinying, Xiaoxu Dong, Jing Lu, Tingting Hu, and Gang Pei (Feb. 2020). "Constitutive activity of a G protein-coupled receptor, DRD1, contributes to human cerebral organoid formation: DRD1 contribute to neurogenesis in brain organoids". In: *STEM CELLS*.

Wang, Quanxin, Song-Lin Ding, Yang Li, Josh Royall, David Feng, Phil Lesnar, Nile Graddis, Maitham Naeemi, Benjamin Facer, Anh Ho, Tim Dolbeare, Brandon Blanchard, Nick Dee, Wayne Wakeman, Karla E. Hirokawa, Aaron Szafer, Susan M. Sunkin, Seung Wook Oh, Amy Bernard, John W. Phillips, Michael Hawrylycz, Christof Koch, Hongkui Zeng, Julie A. Harris, and Lydia Ng (May 2020). "The Allen Mouse Brain Common Coordinate Framework: A 3D Reference Atlas". en. In: *Cell*.

Wang, Ruxin, Shuyuan Chen, Chaojie Ji, Jianping Fan, and Ye Li (2022). "Boundary-aware context neural network for medical image segmentation". In: *Medical Image Analysis* 78, p. 102395.

Wang, Shu-Na, Zhi Wang, Tian-Ying Xu, Ming-He Cheng, Wen-Lin Li, and Chao-Yu Miao (Dec. 2019). "Cerebral Organoids Repair Ischemic Stroke Brain Injury". In: *Translational Stroke Research*.

Wang, Shui-Hua, Yin Zhang, Xiaochun Cheng, Xin Zhang, and Yu-Dong Zhang (Mar. 2021). "PSSPNN: PatchShuffle Stochastic Pooling Neural Network for an Explainable Diagnosis of COVID-19 with Multiple-Way Data Augmentation". en. In: *Computational and Mathematical Methods in Medicine* 2021. Ed. by Martti Juhola, pp. 1–18.

Wang, Weibin, Dong Liang, Qingqing Chen, Yutaro Iwamoto, Xian-Hua Han, Qiaowei Zhang, Hongjie Hu, Lanfen Lin, and Yen-Wei Chen (2020). "Medical image classification using deep learning". In: *Deep learning in healthcare: paradigms and applications*, pp. 33–51.

Wang, Zhengwei, Qi She, and Tomas E Ward (2021). "Generative adversarial networks in computer vision: A survey and taxonomy". In: *ACM Computing Surveys (CSUR)* 54.2, pp. 1–38.

Wang, Zhou and A.C. Bovik (2002). "A universal image quality index". In: *IEEE Signal Processing Letters* 9.3, pp. 81–84.

Wang, Zhou, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli (2004). "Image quality assessment: from error visibility to structural similarity". In: *IEEE Transactions on Image Processing* 13.4, pp. 600–612.

Wargnier-Dauchelle, Valentine, Camille Simon-Chane, and Aymeric Histace (2019). "Retinal Blood Vessels Segmentation: Improving State-of-the-Art Deep Methods". en. In: *Computer Analysis of Images and Patterns*. Vol. 1089. Cham: -, pp. 5–16.

Watanabe, Momoko, Jessie E. Buth, Neda Vishlaghi, Luis de la Torre-Ubieta, Jiannis Taxidis, Baljit S. Khakh, Giovanni Coppola, Caroline A. Pearson, Ken Yamauchi, Danyang Gong, Xinghong Dai, Robert Damoiseaux, Roghiyh Aliyari, Simone Liebscher, Katja Schenke-Layland, Christine Caneda, Eric J. Huang, Ye Zhang, Genhong Cheng, Daniel H. Geschwind, Peyman Golshani, Ren Sun, and Bennett G. Novitch (Oct. 2017). "Self-Organized Cerebral Organoids with Human-Specific Features Predict Effective Drugs to Combat Zika Virus Infection". In: *Cell Reports* 21.2, pp. 517–532.

Wilpert, Nina-Maria, Martin Krueger, Robert Opitz, David Sebinger, Sarah Paisdzior, Bianca Mages, Angela Schulz, Joachim Spranger, Eva K. Wirth, Harald Stachelscheid, Philipp Mergenthaler, Peter Vajkoczy, Heiko Krude, Peter Kühnen, Ingo Bechmann, and Heike Biebermann (Apr. 2020). "Spatiotemporal Changes of Cerebral Monocarboxylate Transporter 8 Expression". In: *Thyroid*.

Wojaczek, Alex, Regina-Veronicka Kalaydina, Mohammed Gasmallah, Myron R. Szewczuk, and Farhana Zulkernine (Dec. 2019). "Computer Vision for Detecting and Measuring Multicellular Tumor Shperoids of Prostate Cancer". In: *2019 IEEE Symposium Series on Computational Intelligence (SSCI)*. Xiamen, China: IEEE, pp. 563–569. ISBN: 978-1-72812-485-8.

Wu, Ren, Shengen Yan, Yi Shan, Qingqing Dang, and Gang Sun (2015). *Deep Image: Scaling up Image Recognition*.

Xiang, Yangfei, Yoshiaki Tanaka, Bilal Cakir, Benjamin Patterson, Kun-Yong Kim, Pingnan Sun, Young-Jin Kang, Mei Zhong, Xinran Liu, Prabir Patra, Sang-Hun Lee, Sherman M. Weissman, and In-Hyun Park (Mar. 2019). "hESC-Derived Thalamic Organoids Form Reciprocal Projections When Fused with Cortical Organoids". In: *Cell Stem Cell* 24.3, 487–497.e7.

Xu, Guoping, Xingrong Wu, Xuan Zhang, and Xinwei He (2021). "Levit-unet: Make faster encoders with transformer for medical image segmentation". In: *arXiv preprint arXiv:2107.08623*.

Yang, Bo, Xiao Fu, Nicholas D Sidiropoulos, and Mingyi Hong (2017a). "Towards k-means-friendly spaces: Simultaneous deep learning and clustering". In: *international conference on machine learning*. PMLR, pp. 3861–3870.

— (2017b). "Towards k-means-friendly spaces: Simultaneous deep learning and clustering". In: *international conference on machine learning*. PMLR, pp. 3861–3870.

Yi, Xin, Ekta Walia, and Paul Babyn (2019). "Generative adversarial network in medical imaging: A review". In: *Medical Image Analysis* 58, p. 101552.

Yoffe, S and C Szegedy (2015). "Batch normalization: Accelerating deep network training by reducing internal covariate shift". In: *Proceedings of the 32nd International Conference on International Conference on Machine Learning* 37, pp. 448–456.

Yoon, Se-Jin, Lubayna S. Elahi, Anca M. Pașca, Rebecca M. Marton, Aaron Gordon, Omer Revah, Yuki Miura, Elisabeth M. Walczak, Gwendolyn M. Holdgate, H. Christina Fan, John R. Huguenard, Daniel H. Geschwind, and Sergiu P. Pașca (Jan. 2019). "Reliability of human cortical organoid generation". In: *Nature Methods* 16.1, pp. 75–78.

Yu, Changqian, Jingbo Wang, Chao Peng, Changxin Gao, Gang Yu, and Nong Sang (2018). "Bisenet: Bilateral segmentation network for real-time semantic segmentation". In: *Proceedings of the European conference on computer vision (ECCV)*, pp. 325–341.

Yushkevich, Paul A, Joseph Piven, Heather Cody Hazlett, Rachel Gimpel Smith, Sean Ho, James C Gee, and Guido Gerig (2006a). "User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability". In: *Neuroimage* 31.3, pp. 1116–1128.

— (2006b). "User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability". In: *Neuroimage* 31.3, pp. 1116–1128.

Zaslavsky, Ilya, Richard A. Baldock, and Jyl Boline (Sept. 2014). "Cyberinfrastructure for the digital brain: spatial standards for integrating rodent brain atlases". In: *Frontiers in Neuroinformatics* 8.

Zemouri, Ryad (2020). "Semi supervised adversarial variational autoencoder". en. In: *Machine Learning and Knowledge Extraction* 2.3, pp. 361–378.

Zhang, Chen (2017). "A platform for stereological quantitative analysis of the brain-wide distribution of type-specific neurons". en. In: *Scientific Reports*, p. 12.

Zhang, Wei, Li Ma, Mei Yang, Qiang Shao, Jian Xu, Zhipeng Lu, Zhen Zhao, Rong Chen, Yang Chai, and Jian-Fu Chen (Apr. 2020). "Cerebral organoid and mouse models reveal a RAB39b–PI3K–mTOR pathway-dependent dysregulation of cortical development leading to macrocephaly/autism phenotypes". In: *Genes & Development* 34.7-8, pp. 580–597.

Zhang, Wei, Si-Lu Yang, Mei Yang, Stephanie Herrlinger, Qiang Shao, John L. Collar, Edgar Fierro, Yanhong Shi, Aimin Liu, Hui Lu, Bruce E. Herring, Ming-Lei Guo, Shilpa Buch, Zhen Zhao, Jian Xu, Zhipeng Lu, and Jian-Fu Chen (Dec. 2019). "Modeling microcephaly with cerebral organoids reveals a WDR62–CEP170–KIF2A pathway promoting cilium disassembly in neural progenitors". In: *Nature Communications* 10.1.

Zhao, Hang, Orazio Gallo, Iuri Frosio, and Jan Kautz (2017). "Loss Functions for Image Restoration With Neural Networks". In: *IEEE Transactions on Computational Imaging* 3.1, pp. 47–57.

Zhao, Shengyu, Zhijian Liu, Ji Lin, Jun-Yan Zhu, and Song Han (2020). "Differentiable augmentation for data-efficient gan training". In: *Advances in Neural Information Processing Systems* 33, pp. 7559–7570.

Zheng, X., L. Zhang, Y. Kuang, V. Venkataramani, F. Jin, K. Hein, M. P. Zafeiriou, C. Lenz, W. Moebius, E. Kilic, D. M. Hermann, M. S. Weber, H. Urlaub, W.-H. Zimmermann, M. Bähr, and Thorsten R. Doeppner (May 2020). "Extracellular Vesicles Derived from Neural Progenitor Cells—a Preclinical Evaluation for Stroke Treatment in Mice". In: *Translational Stroke Research*.

Zhong, Xiali, Georgina Harris, Lena Smirnova, Valentin Zufferey, Rita de Cássia da Silveira e Sá, Fabiele Baldino Russo, Patricia Cristina Baleeiro Beltrao Braga, Megan Chesnut, Marie-Gabrielle Zurich, Helena T. Hogberg, Thomas Hartung, and David Pamies (Feb. 2020). "Antidepressant Paroxetine Exerts Developmental Neurotoxicity in an iPSC-Derived 3D Human Brain Model". In: *Frontiers in Cellular Neuroscience* 14.

Zhong, Zhun, Liang Zheng, Guoliang Kang, Shaozi Li, and Yi Yang (2020). "Random Erasing Data Augmentation". In: *Proceedings of the AAAI Conference on Artificial Intelligence* 34, pp. 13001–13008.

Zhou, Sharon, Mitchell L Gordon, Ranjay Krishna, Austin Narcomey, Durim Morina, and Michael S Bernstein (2019). "Hype: Human Eye Perceptual Evaluation Of Generative Models". In: *lctr*, p. 15.

Zhou, Ting, Lei Tan, Gustav Y. Cederquist, Yujie Fan, Brigham J. Hartley, Suranjit Mukherjee, Mark Tomishima, Kristen J. Brennand, Qisheng Zhang, Robert E. Schwartz, Todd Evans, Lorenz Studer, and Shuibing Chen (Aug. 2017). "High-Content Screening in hPSC-Neural Progenitors Identifies Drug Candidates that Inhibit Zika Virus Infection in Fetal-like Organoids and Adult Brain". In: *Cell Stem Cell* 21.2, 274–283.e5.

Zhou, Zongwei, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, and Jianming Liang (2018). "Unet++: A nested u-net architecture for medical image segmentation". In: *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4*. Springer, pp. 3–11.

Zhu, Jun-Yan, Taesung Park, Phillip Isola, and Alexei A Efros (2017). "Unpaired image-to-image translation using cycle-consistent adversarial networks". In: *Proceedings of the IEEE international conference on computer vision*, pp. 2223–2232.

Zomorodian, Afra J et al. (2012). *Advances in Applied and Computational Topology: American Mathematical Society Short Course on Computational Topology, January 4-5, 2011, New Orleans, Louisiana*. Vol. 70. American Mathematical Soc.

Zu, Xinrui and Qian Tao (2022). "SpaceMAP: Visualizing High-dimensional Data by Space Expansion". In: *International Conference on Machine Learning*. PMLR, pp. 27707–27723.