# A Hierarchical Deep Learning Framework for Nuclei 3D Reconstruction from Microscopic Stack-Images of 3D Cancer Cell Culture

Tarek Maylaa, Feryal Windal, Halim Benhabiles, Gregory Maubon, Nathalie Maubon, Elodie Vandenhaute, Dominique Collard

# A hierarchical deep learning framework for nuclei 3D reconstruction from microscopic stack-images of 3D cancer cell culture

Tarek Maylaa[1], Feryal Windal[1], Halim Benhabiles[1], Gregory Maubon[2], Nathalie Maubon[2], Elodie Vandenhaute[2], and Dominique Collard[1]

[1] University of Lille, CNRS, Central Lille, Junia, Univ. Polytechnic Hauts-de-France
UMR 8520 - IEMN - Institute of Electronics, Microelectronics and Nanotechnology
F-59000 Lille, France,
I.tarek.maylaa@junia.com,
[2] HCS Pharma, Lille, France,

**Abstract.** In this article, we propose a hierarchical deep learning framework for the nuclei 3D reconstruction from a stack of microscopic images representing 3D cancer cell culture. The framework goes through three successive stages namely: at the slice level of the stack i) the spheroid detection and ii) their nuclei segmentation then at the stack level iii) nuclei 3D reconstruction. For this purpose, we prepared a dataset of bright-field microscopic images acquired from 3D cultures of HeLa cells and manually annotated by the experts for both tasks (spheroids detection and nuclei segmentation). Two CNN models namely YOLOv5x and U-Net-VGG19 have been trained and validated on our dataset for the detection and the segmentation tasks respectively. For the 3D reconstruction task, the delaunay triangulation technique has been adopted by exploiting point cloud clusters that represent the segmented nuclei in the stack. Our framework offers to the biologists an efficient assisting tool for quantifying the number of spheroids and analyzing the morphology of their nuclei. The conducted experiments on our generated dataset show the promising results obtained by our framework with notably an average precision of 0,892 and 0,76 on the spheroids detection and nuclei segmentation respectively. Moreover, our 3D reconstruction technique shows visually a consistant representation of nuclei in term of volumetery and shape.

**Keywords:** 3D cell culture, confocal microscopy, z-stack images, deep learning, object detection, segmentation, 3D reconstruction

## 1 Introduction

In drug screening process against cancer, in-vitro 2D cell culture technique is widely used due to its simplicity and its low-cost [1]. However, this technique does not fully represent the physiological characteristics of the cells leading to a limited biological interpretation of the effectiveness of a candidate drug [2]. To

overcome this limit, 3D cell culture technique has been introduced [3]. Indeed, this technique has shown its ability to ensure cell to cell communication and interaction which permit to better mimic the tissue structure as well as function [2]. Neverthless, due to the translucent environment of this type of culture, the acquired microscopic images are affected by several artefacts impacting the visual appearance of the cell environment objects. Figure 1 illustrates the degraded visual quality of such type of images (bright-field images) while emphasizing blurred objects namely spheroids (see on the left the first slice of the z-stack images). The figure shows also the complexity of distinguishing the nuclei due to their confusing contours (see on the right, the nuclei of one spheroid). In addition, the 3D information generated by the depth propagation along the z-axis of the confocal microscopy remains insufficient to guarantee a faithful representation of the whole environment notably with regards to occluded objects (spheroids hidden by other spheroids located at the same 2D position with different z-depth). All these observations make the automatic analysis of these images a great challenge.

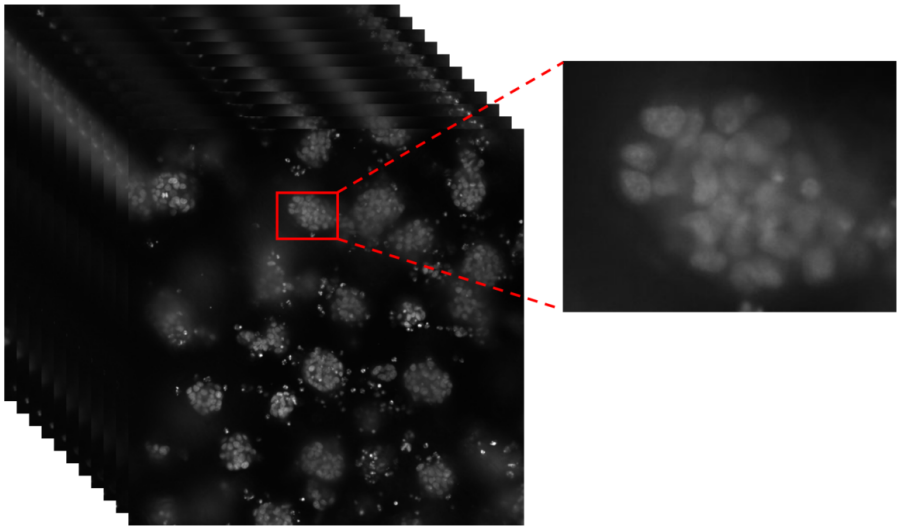


Fig. 1: A z-stack images of 3D culture HeLa cells showing blurred spheroids on the left and a focus on the nuclei of one spheroid with confusing contours on the right.

In this context, nuclei segmentation is a fundemental step for the 3D reconstruction [4]. To address the segmentation task, the deep learning based segmentation algorithms such as U-Net [5] [6] could be exploited. Indeed, these algorithms have demonstrated a high efficiency in the field of medical imaging segmentation [7]. However, this type of approach is well adapted for analyzing:

1) data caracterized by a single level of semantic information (for instance a population of cells or a population of nuclei) which is not the case of our 3D culture-based images. More precisely, the images contain two levels of semantic information, namely a population of spheroids, each of which representing a population of nuclei. 2) a population of separated objects with regular shapes which is also not the case of our 3D culture-based images as highlighted in Figure 1. This also raises the problem of preparing a training dataset with manual segmentations of all the nuclei.

In this article we propose a hierarchical deep learning framework for the nuclei 3D reconstruction from a stack of microscopic images. The framework goes through three successive stages namely: at the slice level of the stack i) the spheroid detection and ii) their nuclei segmentation then at the stack level iii) nuclei 3D reconstruction. In fact, our observation on the multi-level semantic information that characterize our data has motivated our strategy of hierarchical analysis where spheroids are first detected and then their nuclei are segmented from each slice in the stack. We precise that we chose to detect spheroids instead of segmenting them because they represent an agregation of small objects (nuclei) without a continuous contour. To deal with the confusing contours of the nuclei, we propose to segment only the most visible nuclei having explicit contours. Indeed, we make the assumption that the fact to detect all the spheroids, to partially segment their nuclei from each slice in the stack and then apply a 3D reconstruction on the segmented nuclei should permit a relevant quantitative data analysis. More specifically, having an assisting tool for quantifying the number of spheroids and measuring their sizes offers to the biologists a first level of interpreting the effectiveness of a drug. Furthermore, the identification of some nuclei inside each spheroid will offer a second level of analysis with respect to their morphology. To demonstrate the efficiency of our framework, we prepared a dataset of bright-field microscopic images acquired from in-vitro 3D culture of HeLa cells (cervical cancer cells cultured in our laboratory). The images have been manually labeled by experts for both tasks namely spheroids detection and nuclei segmentation. It is worth mentioning that the partial manual segmentation of nuclei still allowed to train a segmentation model on a weakly-labeled dataset. The conducted experiments on our dataset shows the promising results of our framework notably outperforming direct nuclei segmentation methods. Our dataset and source code will be made available upon request.

## 2 Related Work

Several works have been proposed in the literature to address nuclei segmentation from microscopic images [4]. Nevethless, as raised in [4] most of the proposed methods deal with microscopic images acquired from 2D cell culture. In this context, the 2018 Data Science Bowl contest [5], revealed the high interest shown by the scientific community on this task. Indeed, as indicated in [5], the contest has attracted 3,891 teams worldwide challenged on a dataset of 841 2D images containing 37333 nuclei manually annotated. The evaluation protocol established

in the frame of this contest has shown that the best performing solution was based on a variant of the U-Net deep learning architecture. This latter result has motivated the community to pursue the efforts on the development of new variants of U-Net architecture such as Cellpose [8] and StarDist [9]. However, these methods are not adapted to perform an efficient nuclei segmentation from 3D culture-based images in reason of their strong hypothesis with regard to the analyzed patterns. More specifically, they are designed and trained to segment a population of patterns sharing all the same shape (a repetetive shape) and having explicit contours.

In the case of 3D cell culture, Błażej et al. [10] have proposed a 3D reconstruction method of the nuclei surface from a z-stack images. To this end, they adopted a 2D image analysis approach that aims to delimit the contours of the nuclei from one image slice and to track their continuity over the stack to reconstruct the final surface. For the contour delimitation (nuclei segmentation) they exploited hand crafted geometric features that permit to generate a set of boundary points. As raised by the authors, the proposed method performs well for nuclei with convex shapes and with the assumption that for each reconstructed 3D nucleus object there exist at least one 2D slice where the considered object is well separated. Wu et al. [11] proposed a deep learning method to detect and quantify the 3D nuclei centers from 3D fluorescence microscopic images. They based their calculations on the estimation of 2D centroids of these objects from 2D slices. The authors raised that the performance of their detection method is affected by nuclei with non ellipsoidal shapes. Maylaa et al. [12], proposed a comparative study of several hand crafted machine learning methods for the segmentation of nuclei from a z-stack image. For this purpose, they trained several classifiers that allow to identify nuclei objects pixels from one slice image and reconstruct final objects. As raised by the authors, although the trained classifiers succeeded to detect the nuclei pixels they failed to delimit the nuclei contours resulting in a blob-like objects. The obtained results in the afformentioned works on 3D culture cells show that nuclei segmentation and reconstruction tasks are still challenging and require the development of new efficient methods. This observation has been also confirmed in the recent survey [4]. Moreover, the survey highlighted also the lack of public datasets as well as manual annotations.

Regarding spheroids detection from 3D cell culture images, recently Grexa et al. [13] proposed rather to segment the spheroids by exploiting several techniques namely classical ones (Otsu threshold and watershed techniques) and deep learning ones (U-Net and R-CNN models). Neverthless as raised by the authors the delimitation of the spheroids contours remains challenging specifically in the case of close or adjacent spheroids. To the best of our knowledge, no work has been proposed in the frame of 3D cell culture on the detection of spheroids issued from the growth of cancer cells [14]. However, in the field of computer vision, object detection is a well established axis. Indeed, several deep learning based architectures have been proposed to address this task such as the YOLOv series [15] and detection transformers [16].

# 3 Materials and Methods

## 3.1 Dataset

***Image generation*** – To generate the image dataset, we first prepared in our laboratory 3D cultures of HeLa cells (cervical cancer cells) following a standard cell proliferation protocol based on BIOMIMESYS technology [17]. Indeed, cervical cancer is currently ranked at the top 4 worldwide cancers within women [18]. For the microscopic image acquisition, we used the ImageXpress system from molecular device company illustrated in Figure 2. For this purpose, the microscope confocal magnification has been set to 20-x. In total, 600 z-stack (z=0 to z=50$\mu$m) bright-field images with a resolution of 2048 x 2048 pixels have been acquired from a 24-well plate.


Fig. 2: Our microscopic image acquisition system.

***Dataset organization*** – Due to the high amount of generated images (30 000 images), the manual labelling process required to build our training and validation sets for both spheroids detection and nuclei segmentation rapidly became a tedious task for the experts. For this reason, we considered only a subset of z-stack images namely 50 stacks for the detection and among them 29 stacks

for the segmentation. The remaining stacks have been kept for qualitative tests. To ensure a good generalization during the training process of our models, the selected stacks for the labelling have been randomly picked from several wells. Furthermore, from each stack, the experts have labeled one slice image that they have selected along the z-depth according to their own visual perception. More specifically, they have been requested to select the most informative slice in term of spheroid objects for the detection and the nuclei for the segmentation. Following these steps, we created two datasets: 1) a spheroid dataset composed of 50 labeled images with 854 bounding boxes of spheroids. The set has been split into 40 images for training and 10 images for validation. 2) a nuclei dataset composed of 326 spheroid patches (issued from the 29 images) with 1996 delineated nuclei. In this latter case, the set has been split into 249 patches for training and 77 patches for validation. We precise that we made sure that the training patches and validation patches do not come from the same slices. Additionnaly, only separated nuclei with explicit contours have been delineated by the experts.

### 3.2 Methods

***Workflow process*** – As illustrated in Figure 3, our framework takes in input a stack of images (Figure 3(A)) for which a set of spheroids are detected at the slice level (see traced bounding boxes in Figure 3(B)) based on a deep learning model. Each detected spheroid is then cropped and placed into the center of a black background square patch of 512 x 512 pixels (see Figure 3(C)). The choice of this latter resolution has been established empirically based on the largest width and height of the spheroids of our dataset. Each spheroid patch is then passed into our deep learning segmentation model which generates a binary mask of identified nuclei (see Figure 3(D)). Each binary mask is denoised by removing tiny surfaces according to an empirical threshold. A global binary mask (see Figure 3(E)) of the original input image is then generated based on the set of previous masks as well as the spheroid bounding boxes. The resulting spheroid bounding boxes and their nuclei delimited by contours are visualized in Figure 3(F). The 3D surface of each segmented nuclei over the stack is then reconstructed as shown in Figure 3(G).

***Deep learning based spheroids detection model*** – To train our detection model we considered two types of deep architectures namely YOLOv5x [19] [20] and DETR-Resnet50 [16]. Indeed both of these recent architectures have demonstrated a high efficiency over several datasets such as COCO [21] and Pascal [22]. We used the Adam optimizer for both architectures which is the recommended one in the referenced articles. We also used the recommended loss functions namely IOU and hungarian for YOLOv5x and DETR-Resnet50 respectively. The two architectures have been trained following a transfer learning strategy. More specifically, they have been pretrained on the COCO dataset and then trained and validated on our spheroid image dataset producing two detection models.

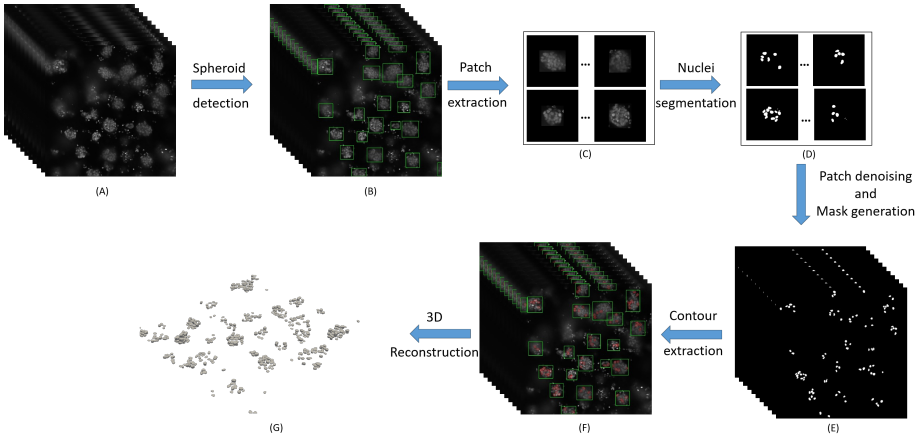



Fig. 3: Workflow process of our hierarchical deep learning framework for nuclei 3D reconstruction.

***Deep learning based nuclei segmentation model*** – To train our segmentation model we considered the U-Net deep architecture [7]. Indeed, this architecture is widely exploited in the field of medical image segmentation [23]. In our case, we customized the architecture by replacing its orginal CNN backbone by a VGG19 [24] [25] one which offers a good tradeoff between the architecture depth and its learning capacity. The model has been trained and validated on our nuclei dataset using Adam optimizer and Jaccard index loss function which is the complement of IOU metric [26].

***Unsupervised learning based nuclei 3D reconstruction model*** – Our built model goes through three main steps (See Figure 4). First, it takes in input the stack of binary masks representing all segmented nuclei and convert it into a 3D point cloud representation. Then, a density-based spatial clustering technique (DBSCAN) [27] is applied in order to identify clusters of 3D points that represent separated nuclei. Finally, the resulting clusters are then used to reconstruct the underlying surfaces of the nuclei using the Delaunay triangulation technique [28].

***Evaluation metrics*** – To evaluate the performance of our detection and segmentation models we adopted standard metrics namely the Precision (Prec), Recall (Rec) and Average Precision (AP) defined as follow:

$$Prec = \frac{TP}{TP + FP} \tag{1}$$

$$Rec = \frac{TP}{TP + FN} \tag{2}$$

$$AP = \frac{TP}{TP + FP + FN} \tag{3}$$

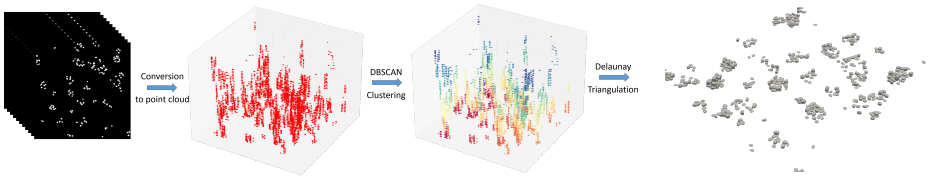



Fig. 4: 3D stack reconstruction pipeline

where TP, FP and FN correspond to True Positive, False Positive and False Negative objects. In the case of spheroid objects TP/FP/FN rates are calculated based on an IOU (intersection over union) metric between ground-truth spheroids (i.e. manual annotations of bounding boxes made by the experts) and predicted ones according to several thersholds set in the range [0,5 to 0,9]. The same calculation methodology is applied for estimating the TP/FP/FN rates of nuclei objects. In this latter case, the ground-truth segmentation of nuclei correspond to binary masks of these objects issued from the manual delineation of their contours by the experts. For the 3D reconstruction quality evaluation, we limit our experiments in this study on a visual analysis of the shape and the volume of the generated nuclei.

## 4 Experimental study

To evaluate the performance of our framework, we organized our experimental study into 4 sections: the first section is dedicated to the results obtained by our detection models, the second concerns the segmentation model, the third is related to a performance comparison with the direct nuclei segmentation methods from the state of the art and the last section presents some qualitative results.

### 4.1 Spheroids detection performance

Figure 5 shows the AP (Average Precision) curves of the two detection models namely YOLOv5x and DETR-Resnet50 obtained on the validation image set (10 images). One can notice that the YOLOv5x have reached the highest performance with notably an AP of 0,848 at an IOU thresholod of 0,5. Nevethless, the DETR model reached also competitive performance. To further improve the performance of the best model namely YOLO we trained it on augmented dataset. More specifically, based on the original training set (40 images) we generated 3 augmented sets: 1) spatial-based augmented set (240 images), 2) texture-based augmented set (200 images) and 3) combined augmented set (400 images). For the spatial augmentation, we applied 3 rotations (90, 180 and 270 degrees) and 2 flips (horizontal and vertical). For the texture augmentation, we applied 2 levels

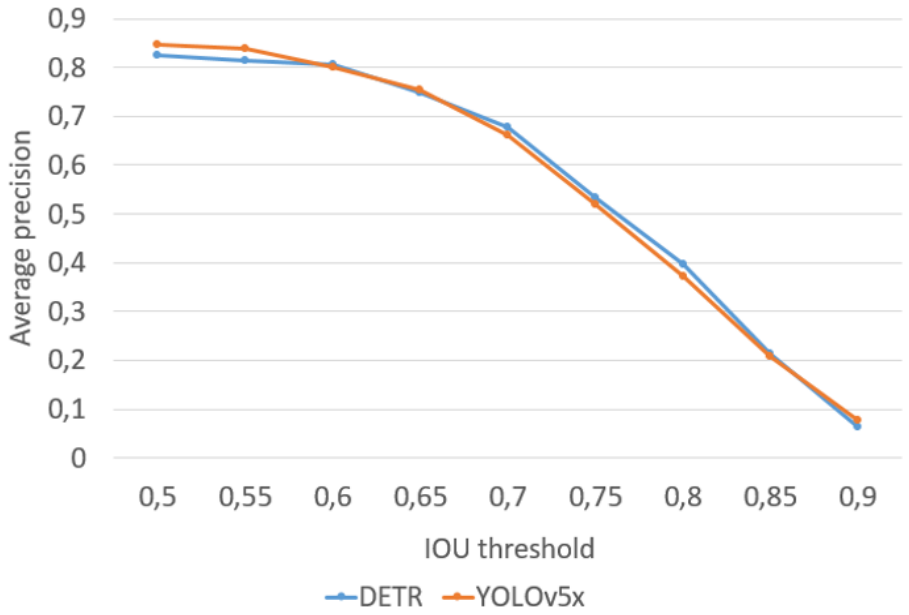



Fig. 5: Performance of the spheroids detection models obtained on the validation set.

of contrast and brightness variations. The combined augmentation corresponds to the merge of all the afformentioned augmentations. Figure 6 shows the obtained AP curves on the validation set for each augmentation scenario together with the original train set (without augmentation). The figure shows that the combined augmentation has permitted to gain 4,4% more in term of AP reaching a score of 0,892 at IOU threshold equal to 0,5. Table 1 sumarizes for this latter scenario the TP/FP/FN rates as well as Prec and Rec metrics calculated on the basis of an IOU threshold set to 0,5. The table shows the high performance of the trained detection model.

Table 1: Performance of the YOLOv5x-based spheroids detection model obtained on the validation set with $IOU_{@0,5}$. Model trained on the combined augmented set (400 images).

| TP | FP | FN | Prec | Rec |
|---|---|---|---|---|
| 157 | 012 | 007 | 0,929 | 0,957 |

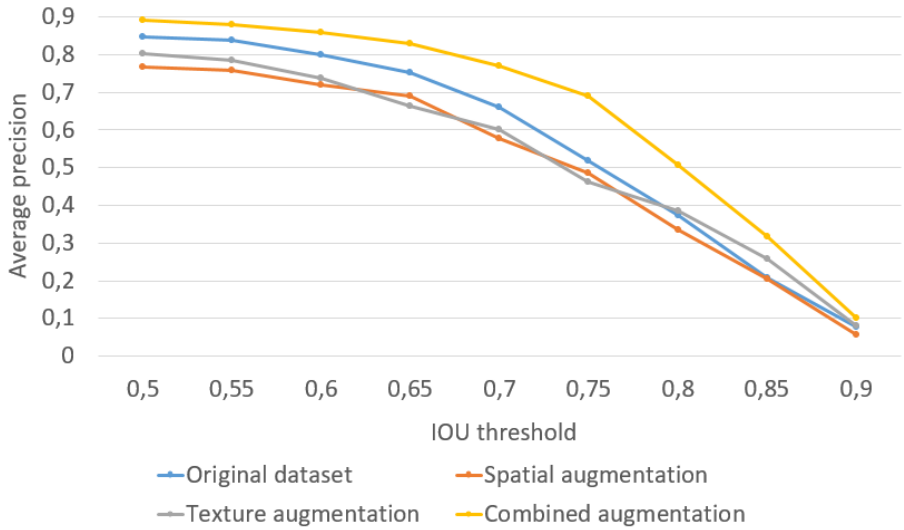


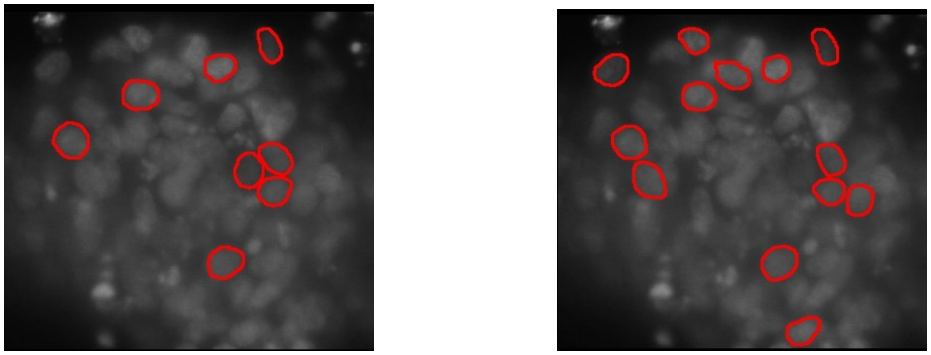

Fig. 6: Performance of four YOLOv5x-based spheroid detection models obtained on the validation set. The models are trained following four scenarios of data augmentation.

### 4.2 Nuclei segmentation performance

Similarly to the detection model, we have trained our segmentation model on 4 training sets including the original training set (249 patches) and 3 augmented sets following the same strategies indicated in the previous section, i.e. spatial (1494 patches), texture (1245 patches) and combined (2490 patches) augmentations. Table 2 summarizes the obtained performance by our models on the validation set (77 patches) for each training scenario and with an IOU threshold set to 0,5. The table shows that the highest performance are obtained by the model trained on the spatial-based augmentation set. The model has reached a Rec score of 0,723 indicating that a weak number of ground-truth nuclei have been missed by it. However, the Prec score of the model is low indicating the high number of detected FP nuclei. In fact, after visually analyzing these FP by the experts it was found that several objects among them are corresponding to true nuclei. Figure 7 highlights this observation through a compartive example between a ground-truth nuclei segmentation vs. predicted one on the same spheroid patch. Hence, the nuclei ground-truth of the validation set has been updated by adding correctly predicted nuclei by our model. Following this update, the number of ground-truth nuclei has increased from 301 to 519. In addition, the Prec and the AP have increased to 0,911 and 0,760 respectively.

Table 2: Performance of nuclei segmentation models obtained on the validation set of spheroid patches with $IOU_{@0,5}$.

| Augmentation | Prec | Rec | AP |
|---|---|---|---|
| None | 0,535 | 0,616 | 0,401 |
| Texture | **0,649** | 0,369 | 0,307 |
| Spatial | 0,594 | **0,723** | **0,484** |
| Combined | 0,620 | 0,608 | 0,443 |



(a) (b)

Fig. 7: Manually labeled nuclei (a) vs Predicted nuclei (b).

### 4.3 Segmentation performance comparison: our framework vs. state of the art direct segmentation methods

We considered two methods from the state of the art: Cellpose [8] and StarDist [9]. The Cellpose is a generalized segmentation method that is designed and trained to efficiently segment a population of objects with repetitive shapes whatever their natures (cells, neurons, etc.) and their geometric forms. The StarDist is a nuclei segmentation method that aims to delimit contours of objects (nuclei cells) having a convex shape. Both of the methods are built on a U-Net architecture. To evaluate these two methods on our dataset: i) The StarDist model pretrained on the Bowl 2018 nuclei dataset has been fine-tuned on the slice images of our training set. To this end, ground-truth global masks have been generated from shperoid patches. ii) The Cellpose method has not been trained since it offers a generalized nuclei segmentation model trained on several combined microscopy datasets including the Bowl 2018 nuclei dataset. However, the model has been parametrized by setting up its nuclei diameter option according to our dataset nuclei.The two models have been tested on our updated nuclei validation set (519 nuclei). In addition, we trained a U-Net with a VGG19 backbone directly on the slice images and validated it on the same set (519 nuclei). Figure 8 shows the AP curves of the 3 models compared to the AP curve of our framework. One may observe that our framework outperformed the three methods. More specifically, it has reached an AP of 0,76 at an IOU set to 0,5.

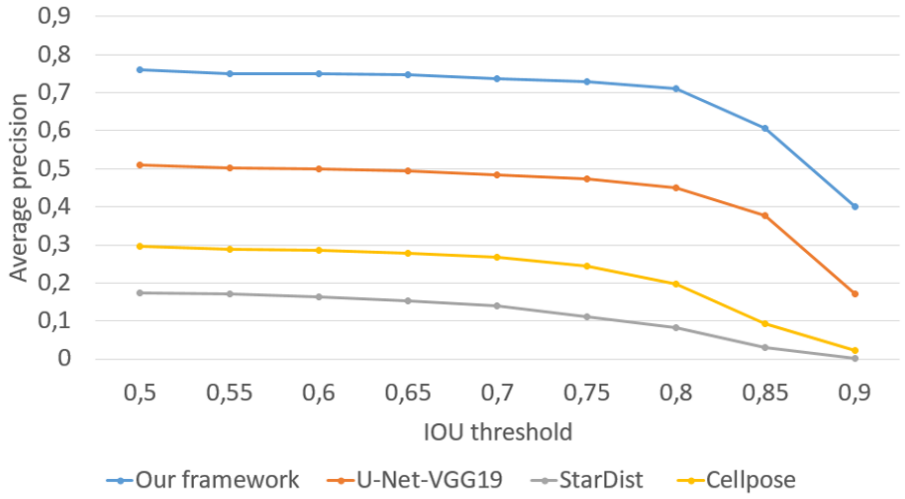
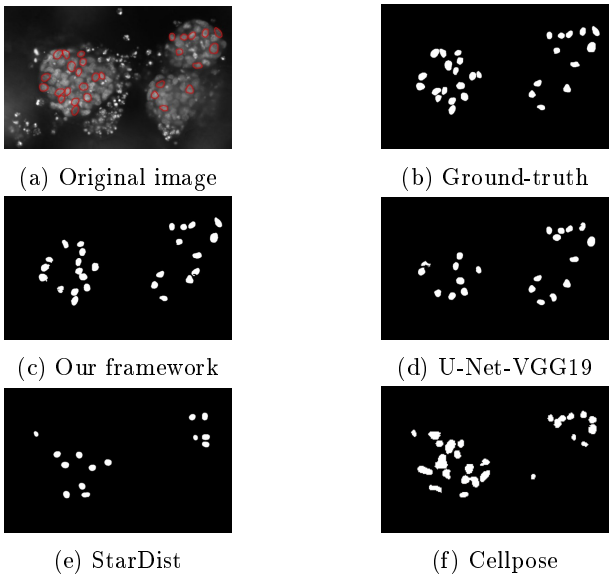



Fig. 8: Segmentation performance comparison between our framework and state of the art methods obtained on the validation set.

We can observe also that the StarDist model has the lowest AP (0,174) although it has been trained (fine-tuned) on our nuclei training set. This result was expected due to the unadapted criterion to our data on which the architecture was designed namely the convexity of the nuclei shape. The Cellpose reached a better AP (0,295) than StarDist although it has been tested directly on the validation set without training. However its score remains low in comparison with our framework showing again the diffculty of segmenting the nuclei of our data by exploiting a direct segmentation approach. Finally, this latter observation is confirmed with the U-Net-VGG19 model that has been trained on the slice images and reached an AP of 0,509. These results demonstrate the contribution of our analysis strategy which first detect the spheroids and then segment their nuclei. Moreover, as illustrated in Figure 9 our strategy permits to cover all the spheroids and thus detect nuclei in each of them which is not guaranteed with the direct segmentation methods.

### 4.4 Qualitative results

***Generalization of detection and segmentation models*** – To analyze the generalization potential of our framework as well as its z-depth sensitivity, we have tested it on several blind z-stacks (stacks that haven't been exploited in the training/validation sets). Figure 10, illustrates some qualitative results obtained on one z-stack with z= 3, 12, 26 and 48. One can observe that our framework has succeeded to detect almost all the spheroids and segment some of their nuclei whatever the slice depth. We also observe that the framework offers a good robustness against noise present in the slices. Indeed, all the identified objects



(a) Original image (b) Ground-truth

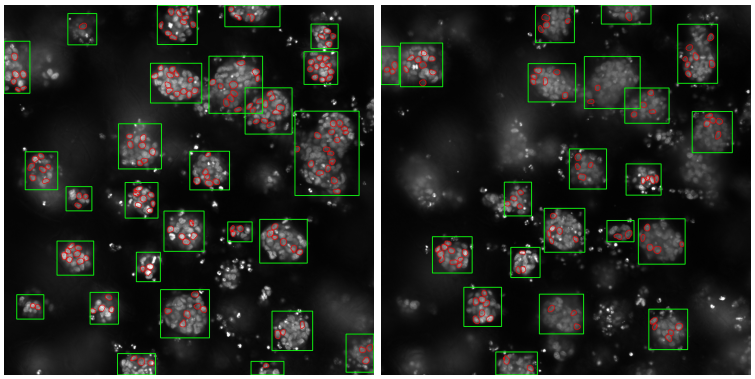
(c) Our framework (d) U-Net-VGG19

(e) StarDist (f) Cellpose

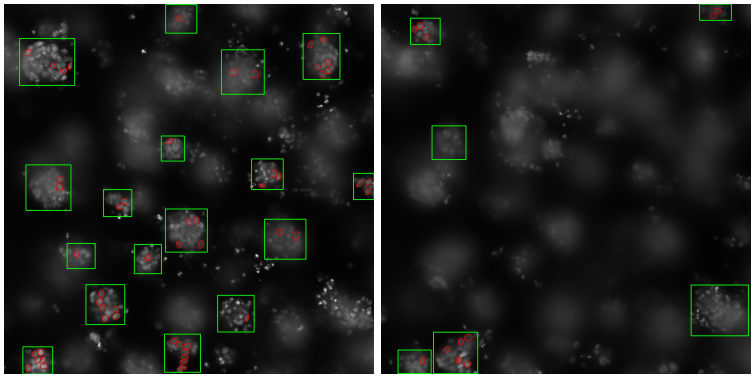
Fig. 9: Qualitative results of segmentation obtained by our framework and state of the art methods.

either from the detection model or from the segmentation model are relevant objects.

***3D Reconstruction*** – In Figure 11(A) we show the result of our reconstruction model on a blind stack (not used in the training and validation process). In order to analyse the impact of the clustering technique on the resulting reconstruction, we have replaced DBSCAN by the Mean Shift Clustering (MSC) [29]. The reconstruction results obtained on the same blind stack is shown in Figure 11(B). One can notice that from a global point of view the reconstruction model is able to generate surfaces of several objects whatever is the exploited clustering technique. Nevertheless, as illustrated in the Figure, a special focus on a small area of the 3D view permits to highlight the efficiency of the DBSCAN compared to the MSC. More specifically, we can observe that the 3D nuclei surfaces generated from MSC are not well separated, are cracked, and include noise. To understand the reasons of this result, we have investigated the correlation between the identified clusters and the binary masks of the segmented nuclei. Figure 12 illustrates an example of this correlation on a small patch extracted from one slice image. One may observe that, contrary to the MSC, the DBSCAN is able to correctly identify the nuclei objects since each one of them is represented by one separated cluster.



(a) z=2 (b) z=12



(c) z=26 (d) z=48

Fig. 10: Qualitative result obtained by our framework on a blind stack (not exploited in the training and validation sets) along several z-depth.

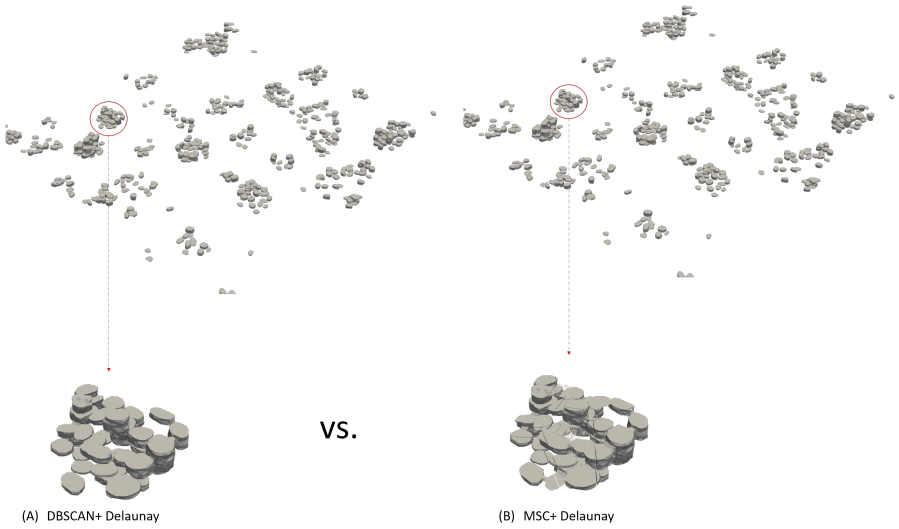



Fig. 11: Comparison between two 3D reconstruction mesh results formed with two different two clustering methods (DBSCAN and MSC) with the Delaunay triangulation

## 5 Conclusion

In this article, we have presented a hierarchical deep learning framework for the nuclei 3D reconstruction. For this purpose we prepared a dataset of bright-field microscopic images composed of 854 manually anotated spheroids and 1996 delineated nuclei by the experts. The two trained CNN models namely yolov5x for the detection and Unet-VGG19 for the segmentation have shown high performance on our validation set by reaching an AP (Average Precision) scores of 0,892 and 0,76 respectively. Several findings have been revealed in our study: 1) combined augmentations based on spatial and texture transformations have permitted to improve the precision of the detection model while only spatial transformations permitted to improve the precision of the segmentation model. 2) segmentation model trained on weakly labeled data succeeded to segment nuclei that have not been annotated by the experts. 3) hierarchical analysis (spheroids detection then nuclei segmentation) permitted to improve the segmentation quality compared to direct segmentation methods from the state of the art. 4) training the models on the most informative slices (only one slice picked from each stack) permitted to ensure a good generalization potential of these models, a result that has been confirmed qualitatively on several stacks along the z-depth. 5) The Delaunay triangulation applied on segmented nuclei achieved a better 3D visual representation when using DBSCAN clustering.

(a) MSC (b) DBSCAN

Fig. 12: MSC vs DBSCAN clustering results

## References


1. M. Kapałczyńska, T. Kolenda, W. Przybyła, M. Zajączkowska, A. Teresiak, V. Filas, M. Ibbs, R. Bliźniak, Ł. Łuczewski, K. Lamperska, Archives of Medical Science : AMS **14**(4), 910 (2018). DOI 10.5114/aoms.2016.63743. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6040128/
2. O. Habanjar, M. Diab-Assaf, F. Caldefie-Chezet, L. Delort, International Journal of Molecular Sciences **22**(22), 12200 (2021). DOI 10.3390/ijms222212200. URL https://www.mdpi.com/1422-0067/22/22/12200
3. B.W. Graf, S.A. Boppart, in *Live Cell Imaging*, vol. 591, ed. by D.B. Papkovsky (Humana Press, Totowa, NJ, 2010), pp. 211–227. DOI 10.1007/978-1-60761-404-3_13. URL http://link.springer.com/10.1007/978-1-60761-404-3_13. Series Title: Methods in Molecular Biology
4. R. Hollandi, N. Moshkov, L. Paavolainen, E. Tasnadi, F. Piccinini, P. Horvath, Trends in Cell Biology p. S0962892421002518 (2022). DOI 10.1016/j.tcb.2021.12.004. URL https://linkinghub.elsevier.com/retrieve/pii/S0962892421002518
5. J.C. Caicedo, A. Goodman, K.W. Karhohs, B.A. Cimini, J. Ackerman, M. Haghighi, C. Heng, T. Becker, M. Doan, C. McQuin, M. Rohban, S. Singh, A.E. Carpenter, Nature Methods **16**(12), 1247 (2019). DOI 10.1038/s41592-019-0612-7. URL http://www.nature.com/articles/s41592-019-0612-7
6. C. Ling, M. Halter, A. Plant, M. Majurski, J. Stinson, J. Chalfoun, in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (IEEE, Seattle, WA, USA, 2020), pp. 4157–4163. DOI 10.1109/CVPRW50498.2020.00491. URL https://ieeexplore.ieee.org/document/9150701/
7. O. Ronneberger, P. Fischer, T. Brox, in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, vol. 9351, ed. by N. Navab, J. Hornegger, W.M. Wells, A.F. Frangi (Springer International Publishing, Cham, 2015), pp. 234–241. DOI 10.1007/978-3-319-24574-4_28. URL http://link.springer.com/10.1007/978-3-319-24574-4_28. Series Title: Lecture Notes in Computer Science
8. C. Stringer, T. Wang, M. Michaelos, M. Pachitariu, Nature Methods **18**(1), 100 (2021). DOI 10.1038/s41592-020-01018-x. URL http://www.nature.com/articles/s41592-020-01018-x

9. U. Schmidt, M. Weigert, C. Broaddus, G. Myers, in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*, vol. 11071, ed. by A.F. Frangi, J.A. Schnabel, C. Davatzikos, C. Alberola-López, G. Fichtinger (Springer International Publishing, Cham, 2018), pp. 265–273. DOI 10.1007/978-3-030-00934-2_30. URL http://link.springer.com/10.1007/978-3-030-00934-2_30. Series Title: Lecture Notes in Computer Science
10. B. Ruszczycki, K.K. Pels, A. Walczak, K. Zamłyńska, M. Such, A.A. Szczepankiewicz, M.H. Hall, A. Magalska, M. Magnowska, A. Wolny, G. Bokota, S. Basu, A. Pal, D. Plewczynski, G.M. Wilczyński, Frontiers in Neuroanatomy **13**, 81 (2019). DOI 10.3389/fnana.2019.00081. URL https://www.frontiersin.org/article/10.3389/fnana.2019.00081/full
11. L. Wu, S. Han, A. Chen, P. Salama, K.W. Dunn, E.J. Delp, in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (IEEE, Nashville, TN, USA, 2021), pp. 3750–3760. DOI 10.1109/CVPRW53098.2021.00416. URL https://ieeexplore.ieee.org/document/9522908/
12. T. Maylaa, F. Windal, H. Benhabiles, G. Maubon, N. Maubon, E. Vandenhaute, D. Collard, Current Computer-Aided Drug Design **18** (2022). DOI 10.2174/1573409918666220208120756. URL https://www.eurekaselect.com/200924/article
13. I. Grexa, A. Diosdi, M. Harmati, A. Kriston, N. Moshkov, K. Buzas, V. Pietiäinen, K. Koos, P. Horvath, Scientific Reports **11**(1), 14813 (2021). DOI 10.1038/s41598-021-94217-1. URL http://www.nature.com/articles/s41598-021-94217-1
14. F. Bubba, C. Pouchol, N. Ferrand, G. Vidal, L. Almeida, B. Perthame, M. Sabbah, Journal of Theoretical Biology **479**, 73 (2019). DOI 10.1016/j.jtbi.2019.07.002. URL https://linkinghub.elsevier.com/retrieve/pii/S0022519319302796
15. U. Nepal, H. Eslamiat, Sensors **22**(2), 464 (2022). DOI 10.3390/s22020464. URL https://www.mdpi.com/1424-8220/22/2/464
16. N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, S. Zagoruyko, in *Computer Vision – ECCV 2020*, vol. 12346, ed. by A. Vedaldi, H. Bischof, T. Brox, J.M. Frahm (Springer International Publishing, Cham, 2020), pp. 213–229. DOI 10.1007/978-3-030-58452-8_13. URL https://link.springer.com/10.1007/978-3-030-58452-8_13. Series Title: Lecture Notes in Computer Science
17. E. Vitali, I. Boemi, G. Tarantola, S. Piccini, A. Zerbi, G. Veronesi, R. Baldelli, G. Mazziotti, V. Smiroldo, E. Lavezzi, A. Spada, G. Mantovani, A.G. Lania, Cancers **12**(8), 2143 (2020). DOI 10.3390/cancers12082143. URL https://www.mdpi.com/2072-6694/12/8/2143
18. World Health Organization. Cervical cancer. URL https://www.who.int/health-topics/cervical-cancer#tab=tab_1
19. G. Jocher. Yolov5 ultralytics. URL https://github.com/ultralytics/yolov5
20. W. Zhan, C. Sun, M. Wang, J. She, Y. Zhang, Z. Zhang, Y. Sun, Soft Computing **26**(1), 361 (2022). DOI 10.1007/s00500-021-06407-8. URL https://link.springer.com/10.1007/s00500-021-06407-8
21. T.Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, C.L. Zitnick, in *Computer Vision – ECCV 2014*, ed. by D. Fleet, T. Pajdla, B. Schiele, T. Tuytelaars (Springer International Publishing, Cham, 2014), pp. 740–755
22. M. Everingham, L. Van Gool, C.K.I. Williams, J. Winn, A. Zisserman, International Journal of Computer Vision **88**(2), 303 (2010). DOI 10.1007/s11263-009-0275-4. URL http://link.springer.com/10.1007/s11263-009-0275-4
23. N. Siddique, S. Paheding, C.P. Elkin, V. Devabhaktuni, IEEE Access **9**, 82031 (2021). DOI 10.1109/ACCESS.2021.3086020. URL https://ieeexplore.ieee.org/document/9446143/

24. F. Xu, X. Li, H. Yang, Y. Wang, W. Xiang, Biomedical Signal Processing and Control **73**, 103416 (2022). DOI 10.1016/j.bspc.2021.103416. URL https://linkinghub.elsevier.com/retrieve/pii/S1746809421010132
25. F. Liu, L. Wang, Construction and Building Materials **322**, 126265 (2022). DOI 10.1016/j.conbuildmat.2021.126265. URL https://linkinghub.elsevier.com/retrieve/pii/S0950061821039957
26. T. Eelbode, J. Bertels, M. Berman, D. Vandermeulen, F. Maes, R. Bisschops, M.B. Blaschko, IEEE Transactions on Medical Imaging **39**(11), 3679 (2020). DOI 10.1109/TMI.2020.3002417. URL https://ieeexplore.ieee.org/document/9116807/
27. M. Ester, H.P. Kriegel, X. Xu, p. 6
28. Boris Delaunay, Bull. Acad. Science USSR VII: Class. Sci. Mat **17**(8), 793 (1934)
29. Yizong Cheng, IEEE Transactions on Pattern Analysis and Machine Intelligence **17**(8), 790 (1995). DOI 10.1109/34.400568. URL http://ieeexplore.ieee.org/document/400568/