



HAL
open science

Segmentation 3D de cellules par prédiction de surfaces actives

Quentin Rapilly, Anaïs Badoual, Charles Kervrann

► **To cite this version:**

Quentin Rapilly, Anaïs Badoual, Charles Kervrann. Segmentation 3D de cellules par prédiction de surfaces actives. RFIAP 2024 - Congrès Reconnaissance des Formes, Image, Apprentissage et Perception, Université de Lille, Cité Scientifique Campus; IMT Nord Europe, Jul 2024, Lille, France. pp.1-3. hal-04611350

HAL Id: hal-04611350

<https://hal.science/hal-04611350>

Submitted on 13 Jun 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Segmentation 3D de cellules par prédiction de surfaces actives

Quentin RAPILLY^{1,*}

Anaïs BADOUAL¹

Charles KERVRANN¹

¹ Centre Inria de l'Université de Rennes, Equipe SAIRPICO, Unité Chimie et Biologie de la Cellule, U1143 INSERM, UMR3666 CNRS, Institut Curie, Université PSL

*Email : quentin.rapilly@inria.fr

1 Introduction

En microscopie, les grandes quantités de données acquises nécessitent le développement de méthodes de segmentation automatiques multi-instances. La révolution de l'apprentissage profond a conduit au développement de méthodes très performantes en segmentation de cellules. Pour des images 2D, Cellpose [8] ou StarDist [7] sont les algorithmes les plus performants et les plus communément utilisés par les biologistes pour la segmentation multi-objets. Ils sont basés sur des architectures de réseaux de neurones convolutifs (type UNet [6]) et prédisent les paramètres d'un modèle de forme dans le cas de StarDist, ou des cartes de gradients utilisées pour construire les formes pour Cellpose. SplineDist [4] est un algorithme moins répandu, qui tente de corriger les faiblesses de StarDist en généralisant les formes prédites, sans se restreindre à des formes "étoilées". Les images biologiques 3D sont difficiles à analyser manuellement. Une adaptation de StarDist a été proposée pour analyser de telles images [9]. En revanche, Cellpose traite une image 3D comme une série d'images 2D indépendantes. Notre méthode est une généralisation de la version 3D de StarDist. Son principal avantage est de prédire des surfaces actives non restreintes à la segmentation de formes étoilées et de garantir une représentation continue des objets segmentés.

2 Méthode

On représente les objets biologiques par des surfaces paramétriques fermées, à topologie sphérique, comme présentées dans [2]. La fonction représentant la surface est décrite comme suit :

$$\forall (u, v) \in [0, 1], \quad \mathbf{s}(u, v) = \sum_{k=0}^{2M-1} \sum_{l=-1}^{M+1} \mathbf{c}_{k,l} \varphi_{\text{per}}(2Mu - k) \varphi(Mv - l), \quad (1)$$

où les vecteurs $\{\mathbf{c}_{k,l}\}_{\substack{0 \leq k \leq 2M-1 \\ -1 \leq l \leq M+1}}$ sont les points de contrôle 3D de la surface, et $M \geq 2$ est un entier décrivant le degré de complexité de la surface; plus M est grand, plus la surface peut varier localement. La fonction φ détermine les formes possibles de la surface paramétrique. Dans notre cas, φ correspond à la B-spline exponentielle, étudiée dans [2], tout comme sa version $2M$ -périodique φ_{per} . Afin d'assurer la C^1 -continuité de la surface au niveau de ses pôles, des conditions sur certains points de contrôle doivent être imposées [2]. Au final, la surface est entièrement décrite par $I = [2M(M-1) + 6]$ paramètres : $\{\mathbf{f}_i\}_{1 \leq i \leq I} = \{\mathbf{c}_N, \mathbf{c}_S, \mathbf{T}_{1,N}, \mathbf{T}_{2,N}, \mathbf{T}_{1,S}, \mathbf{T}_{2,S}, \mathbf{c}_{0,1}, \dots, \mathbf{c}_{2M-1, M-1}\}$ où $\mathbf{c}_S, \mathbf{c}_N$ sont les pôles nord et sud de la surface et les vecteurs $\mathbf{T}_{1,N}, \mathbf{T}_{2,N} \in \mathbb{R}^3$, $(\mathbf{T}_{1,S}, \mathbf{T}_{2,S} \in \mathbb{R}^3$, respectivement) déterminent le plan tangent au pôle nord (sud, respectivement). La correspondance entre les $\{\mathbf{f}_i\}$ et les $\{\mathbf{c}_{k,l}\}$ a été établie dans [2].

Les surfaces paramétriques actives sont des méthodes de segmentation qui consistent à estimer les paramètres $\{\mathbf{f}_i\}$ permettant d'ajuster au mieux la surface (1) aux frontières de l'objet cible. Cet ajustement est habituellement effectué en minimisant une fonction d'énergie. Notre méthode procède différemment. Elle vise à prédire directement, à l'aide d'un réseau de neurones, pour chaque objet d'intérêt de l'image, les paramètres $\{\mathbf{f}_i\}$ d'une surface s'ajustant avec précision aux contours, évitant ainsi des initialisations coûteuses car nombreuses.

Architecture. L'architecture de notre réseau est un UNet [6]. À partir d'une image de dimension (C_{in}, H, W, D) le réseau prédit un volume de dimension (C_{out}, H, W, D) , où C_{in} vaut 1 (niveau de gris) et $C_{out} = 1 + 3I$. En effet, pour chaque voxel $\mathcal{V} = (x, y, z)$, on applique la fonction sigmoïde au premier canal de sortie afin d'obtenir la probabilité $p_{\mathcal{V}}$ de \mathcal{V} d'être le barycentre d'un objet. Les $3I$ autres canaux correspondent aux 3 coordonnées sphériques des I paramètres $\{\mathbf{f}_{\mathcal{V},i}\}_{1 \leq i \leq I}$ caractérisant la surface entourant l'objet dont \mathcal{V} fait partie, dans un repère dont \mathcal{V} est l'origine.

Entraînement. Soit $\hat{p}_{\mathcal{V}}$ la probabilité prédite que \mathcal{V} soit au centre d'un objet d'intérêt et $p_{\mathcal{V}}$ la vérité-terrain correspondante (qui vaut 0 dans le fond de l'image, et évolue de 0 sur les contours des objets, à 1 pour les voxels les plus éloignés des contours). Soit $\hat{S}_{\mathcal{V}} = \{\hat{s}_{\mathcal{V},p}^p\}_{1 \leq p \leq P}$ un échantillon composé de P points, obtenus par l'intermédiaire d'une grille de Fibonacci

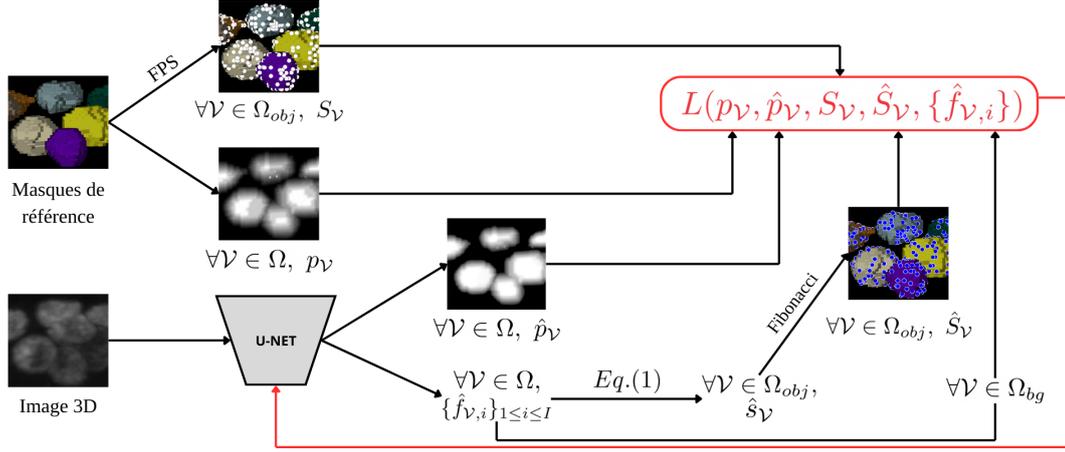
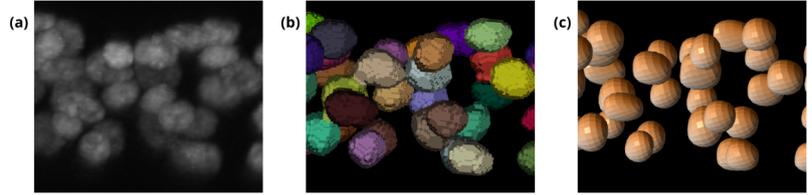


FIGURE 1 – **Processus d’entraînement** : en haut, les masques vérité terrain sont utilisés pour obtenir les échantillonnages S_V des objets (FPS) ainsi que la carte de probabilité p_V d’être au centre d’un objet. En bas, l’image est fournie au réseau qui infère la probabilité prédite \hat{p}_V et les paramètres $\{f_{V,i}\}$ permettant de reconstruire la surface \hat{S}_V (voir (1)), que l’on échantillonne en \hat{S}_V via la grille de Fibonacci. L’ensemble $\Omega = \Omega_{bg} \cup \Omega_{obj}$ englobe les voxels du fond et des objets.

FIGURE 2 – **Inference**. (a) : image issue de l’ensemble de test de "WORM"; (b) : segmentation de référence; (c) : surfaces prédites par notre algorithme. Sur les 7 images de test, la détection donne une précision de 0.76 et un rappel de 0.94. L’IoU moyen entre les surfaces de référence et les vrais positifs vaut 0.64, l’IoU médian vaut 0.68.



[3], régulièrement distribués sur la surface générée par les $\{\hat{f}_{V,i}\}$ prédits. Soit $S_V = \{s_V^p\}_{1 \leq p \leq P}$ un échantillonnage de la surface vérité terrain obtenu en utilisant un algorithme de "Farthest Point Sampling" (FPS) [5]. Pour chaque voxel de l’image, on définit la fonction de coût d’apprentissage comme suit :

$$L(\hat{p}_V, p_V, \hat{S}_V, S_V, \{\hat{f}_{V,i}\}) = BCE(\hat{p}_V, p_V) + \lambda_s L_s(p_V, \hat{S}_V, S_V, \{\hat{f}_{V,i}\}), \quad (2)$$

où le paramètre λ_s pondère l’importance accordée aux surfaces par rapport aux probabilités estimées au cours de l’apprentissage. Le premier terme BCE correspond à la fonction de coût standard "Binary Cross Entropy" permettant d’évaluer la probabilité prédite \hat{p}_V . Le terme L_s compare la surface prédite à la surface de vérité terrain à partir de leur échantillonnage \hat{S}_V et S_V . Pour l’ensemble des voxels appartenant à un objet d’intérêt Ω_{obj} (i.e., $\{V, p_V > 0\}$), il correspond à la "Sliced Wasserstein Distance" (SWD) [1], moins coûteuse qu’une distance de Wasserstein, pondérée par p_V . Pour l’ensemble des voxels du fond de l’image Ω_{bg} (i.e., $\{V, p_V = 0\}$) c’est un terme de régularisation qui est appliqué à la place :

$$L_s(p_V, \hat{S}_V, S_V, \{\hat{f}_{V,i}\}) = p_V \cdot \mathbb{1}_{\Omega_{obj}}(V) \cdot SWD(\hat{S}_V, S_V) + \lambda_{reg} \cdot \mathbb{1}_{\Omega_{bg}}(V) \cdot \sum_i \|\hat{f}_{V,i}\|. \quad (3)$$

Le paramètre λ_{reg} contrôle l’importance du terme de régularisation, dont le rôle est de contraindre le réseau à prédire des surfaces de volume nul en l’absence d’objet. Ce terme limite aussi l’explosion des normes des paramètres. Une illustration représentant une époque d’entraînement est sur la figure 1.

Inference. Pour analyser une nouvelle image, le réseau prédit $(\hat{p}_V, \{\hat{f}_{V,i}\})$ pour chaque voxel V . On extrait les centres potentiels $\{V_q\}_{1 \leq q \leq Q}$, dont la probabilité \hat{p}_{V_q} est supérieure à celle de tous leurs voisins (i.e. les maxima locaux) et à un seuil donné. Pour chaque voxel V_q , la surface \hat{S}_{V_q} est générée à partir des $\{\hat{f}_{V_q,i}\}$ prédits. Une étape de "Non Maximum Suppression" basée sur "l’Intersection-over-Union" est finalement opérée afin d’éliminer les potentiels doublons parmi les centres sélectionnés.

3 Résultats expérimentaux

Afin d’évaluer notre méthode, nous avons entraîné notre réseau pour $M = 2$ (i.e., $I = 10$), avec les paramètres $P = 101$, $\lambda_s = 1$, $\lambda_{reg} = 10^{-3}$, sur le jeu de données "WORM" [9], représentant les noyaux des cellules des larves de vers *C. Elegans*, observés en microscopie de fluorescence. Le jeu comprends 28 images (18 d’entraînement, 3 de validation, 7 de test), de taille $(140 \times 140 \times 1100)$ voxels, peu bruitées mais mal contrastées et denses (contenant chacune environ 550 objets). Le réseau est entraîné pendant 10000 époques ($\sim 20h$), sur un GPU Nvidia A100. La détection des objets (nombre, position) durant l’inférence est de qualité similaire et a été obtenue dans un temps comparable ($\sim 8s$) à StarDist. La précision des surfaces est prometteuse mais devra être améliorée. La figure 2 illustre typiquement les surfaces prédites lors de l’inférence.

Références

- [1] Nicolas Bonneel, Julien Rabin, Gabriel Peyré, and Hanspeter Pfister. Sliced and radon wasserstein barycenters of measures. *Journal of Mathematical Imaging and Vision*, 51 :22–45, 2015.
- [2] Ricard Delgado-Gonzalo, Nicolas Chenouard, and Michael Unser. Spline-based deforming ellipsoids for interactive 3d bioimage segmentation. *IEEE Transactions on Image Processing*, 22(10) :3926–3940, 2013.
- [3] Álvaro González. Measurement of areas on a sphere using fibonacci and latitude–longitude lattices. *Mathematical Geosciences*, 42 :49–64, 2010.
- [4] Soham Mandal and Virginie Uhlmann. Splinedist : Automated cell segmentation with spline curves. In *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, pages 1082–1086. IEEE, 2021.
- [5] Gabriel Peyré and Laurent D. Cohen. *Geodesic Methods for Shape and Surface Processing*, pages 29–56. Springer Netherlands, Dordrecht, 2009.
- [6] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net : Convolutional networks for biomedical image segmentation. In Nassir Navab, Joachim Hornegger, William M. Wells, and Alejandro F. Frangi, editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, pages 234–241, Cham, 2015. Springer International Publishing.
- [7] Uwe Schmidt, Martin Weigert, Coleman Broaddus, and Gene Myers. Cell detection with star-convex polygons. In *Medical Image Computing and Computer Assisted Intervention - MICCAI 2018 - 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part II*, pages 265–273, 2018.
- [8] Carsen Stringer, Tim Wang, Michalis Michaelos, and Marius Pachitariu. Cellpose : a generalist algorithm for cellular segmentation. *Nature methods*, 18(1) :100–106, 2021.
- [9] Martin Weigert, Uwe Schmidt, Robert Haase, Ko Sugawara, and Gene Myers. Star-convex polyhedra for 3d object detection and segmentation in microscopy. In *The IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 3655–3662, March 2020.