

Sujet de thèse : Analyse d'images de microscopie électronique par apprentissage profond pour le diagnostic de la leucémie lymphoïde chronique

Encadrement

- Directeur : Benoît NAEGEL - ICube - b.naegel@unistra.fr
- Co-directeur : Patrick SCHULTZ - IGBMC - patrick.schultz@igbmc.fr

Date de début

1er septembre 2019

Financement

Contrat doctoral interdisciplinaire : financement IDEX de l'Université de Strasbourg.

Contexte

L'objectif de cette thèse est le développement de méthodes de segmentation robustes dédiées à l'imagerie biologique couplant apprentissage profond et opérateurs connexes dans un but de description quantitative de l'organisation tridimensionnelle des cellules. Ce sujet, à l'interface de l'imagerie biologique (équipe Schultz, IGBMC) et de l'analyse numérique des images (équipe IMAGEs, ICube) se prolonge à travers une collaboration avec une équipe de cliniciens de l'hôpital de Hautepierre (Dr L. Mauvois, UMR S-1113) sur l'étude morphologique de différentes populations de cellules lymphoïdes B normales et provenant de patients atteints de leucémie lymphoïde chronique, l'objectif étant de pouvoir améliorer le diagnostic des différentes formes ou stades de la maladie.

La morphologie des cellules, analysée par microscopie électronique, fournit une grande quantité d'informations sur la taille, la forme et la distribution des composants cellulaires tels que le noyau, les pores nucléaires, le nucléole ou l'état de condensation de la chromatine. Dans le cytoplasme, le réticulum endoplasmique, l'appareil de Golgi, le cytosquelette ou encore le réseau de mitochondries sont résolus sans difficultés. Un atlas des paramètres de morphologie cellulaire normale et pathologiques pourrait, à l'avenir, améliorer le diagnostic ou guider les chercheurs qui étudient une maladie inconnue vers des voies déjà décrites pour lesquelles un traitement existe. Jusqu'à récemment, ces informations étaient essentiellement descriptives mais le développement de nouveaux modes d'imagerie, tels que l'imagerie en microscopie électronique à balayage combinée à un faisceau ionique focalisé (FIB/SEM), ouvrent la perspective d'extraire des informations 3D quantitatives sur tout type cellulaire dans différentes conditions physiologiques [Kizilyaprak et al. 2014]. Dans sa version la plus récente, un faisceau ionique focalisé (FIB) découpe une couche mince à la surface de l'échantillon, tandis qu'un faisceau d'électrons à balayage (SEM) image la face du bloc nouvellement créée. L'itération de ces deux processus offre la possibilité d'imager des zones de 40 x 40 μm avec une résolution spatiale isotrope de 3 nm sur une profondeur de plusieurs dizaines de micromètres. L'interprétation des images et l'extraction de données quantitatives se heurtent le plus souvent à des problèmes de segmentation d'image : la technologie FIB/SEM, avec sa capacité à imager des cellules entières de haute résolution spatiale, fournit des images de très grande taille (de l'ordre de 2048 x 1536 x 500, soit plus de 1.5 milliard de pixels), et contenant de nombreux objets biologiques complexes (voir figure 1).

L'étape de segmentation est nécessaire pour annoter les compartiments cellulaires et les organelles afin que des informations quantitatives telles que la taille, la distribution, la morphologie puissent être extraites de ces régions délimitées. Un expert humain peut identifier et annoter ces caractéristiques cellulaires de

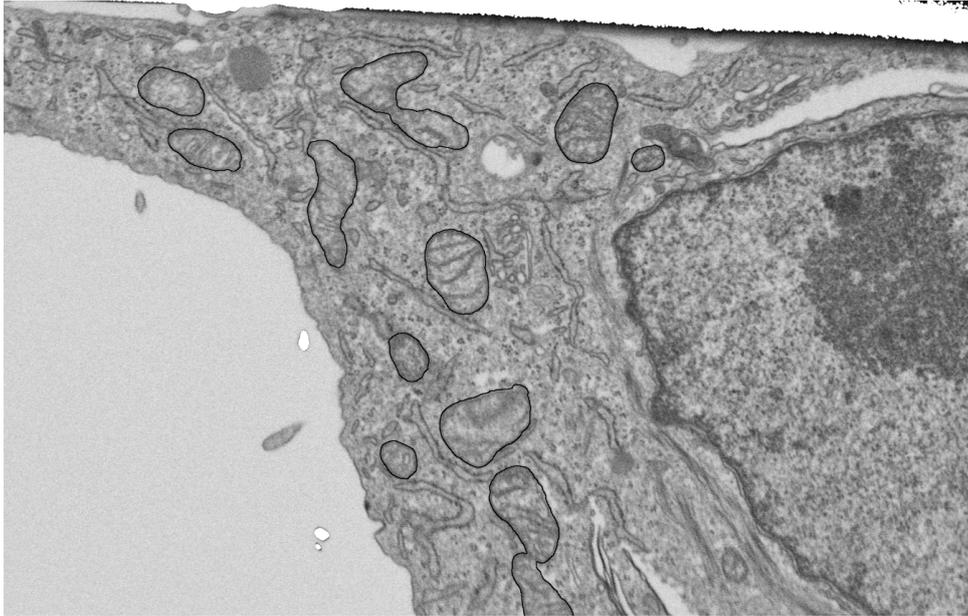


FIGURE 1 – Coupe d'une image FIB/SEM. Les mitochondries segmentées par un réseau convolutif sont délimitées en noir.

manière interactive, mais cette tâche est très fastidieuse et nécessite plusieurs jours pour chaque cellule. Dans un contexte de données massives (plusieurs dizaines de cellules et plusieurs dizaines de patients), il est difficile d'obtenir des informations quantitatives et statistiquement significatives de cette manière. L'automatisation de cette étape et le développement de méthodes de segmentation robustes représentent ainsi un enjeu majeur.

Dans ce domaine, un des verrous de la segmentation automatique est lié à la nature extrêmement bruitée des images de microscopie électronique (faible rapport signal/bruit, faible contraste) ainsi qu'à la nature complexe des images liée aux nombreux objets présents à l'intérieur de la cellule. Un second verrou est relatif à la double problématique de la segmentation sémantique, qui consiste d'une part à reconnaître différentes classes d'objets dont le nombre est inconnu, et d'autre part à la délimitation précise de leur contour. Des outils logiciels dédiés à la segmentation automatique d'image en microscopie électronique ont été développés : outils interactifs dans la plateforme Fiji [Schindelin et al. 2012], *active learning* à base de graphes nécessitant une base annotée [Nunez-Iglesias et al. 2014], techniques de seuillages après prétraitement de l'image [Hoang et al. 2017]. Pour la segmentation des régions, des stratégies fondées sur les superpixels sont utilisées afin de sur-segmenter l'image en régions pertinentes, puis une stratégie de scission/fusion est utilisée pour le regroupement des régions [Jones et al. 2015]. Enfin, des techniques de classification supervisée sont utilisées dans un cadre interactif afin d'identifier des régions d'intérêt [Sommer et al. 2011]. Ces techniques, qui requièrent pour la plupart un minimum d'interactions, ne sont pas adaptées et/ou pas suffisamment robustes pour notre problématique.

Les méthodes de segmentation à base de réseaux de neurones convolutifs ont récemment prouvé leur intérêt dans le cadre de l'imagerie biomédicale [Ronneberger et al. 2015]. Dans le contexte de la microscopie électronique, des travaux récents ont proposé des approches de segmentation utilisant ces techniques [Zeng et al. 2017, Haberl et al. 2018].

Travail à réaliser

Une première partie de cette thèse consistera à développer et améliorer des méthodes de segmentation dense à base de réseaux de neurones « tout convolutifs » adaptée au contexte de l'imagerie FIB/SEM. L'objectif sera de segmenter plusieurs types d'objets, mitochondries, réticulum endoplasmique, membrane cellulaire, noyau, en s'appuyant sur les travaux réalisés dans les équipes IMAGEs et SDC dans le domaine de l'imagerie IRM 3D [Sanches et al. 2019] et de l'imagerie histopathologique [Abreu et al. 2019].

Dans un second temps, l'objectif est d'évaluer l'utilisation des opérateurs connexes [Salembier and Wilkinson 2009] dans un cadre de simplification d'image afin de faciliter et/ou accélérer l'apprentissage des réseaux convolutifs. On peut noter en effet que les méthodes d'apprentissage profond, pour être

efficaces, s'appuient sur un certain nombre de prétraitements de l'image (normalisation, filtrage linéaire en particulier). Les opérateurs connexes sont des opérateurs non-linéaires qui ont la propriété de ne pas altérer les contours originaux de l'image (un contour est soit préservé soit supprimé, mais ne peut être déplacé), à l'inverse des opérateurs linéaires qui sont fondés sur l'opération de convolution. Les opérateurs connexes sont rapides à calculer (complexité linéaire) et permettent de préserver l'intégrité de l'information présente dans l'image (en particulier la localisation des contours) ce qui les rend particulièrement attractifs dans un contexte de quantification d'objets.

Le résultat de ces travaux permettra de modéliser sous forme d'atlas morphologique statistique les populations de cellules saines et pathologiques.

Programme envisagé et modalités d'encadrement

La première année de cette thèse sera focalisée d'une part sur l'étude de l'état de l'art dans le domaine de l'apprentissage profond et des opérateurs connexes en morphologie mathématique ainsi qu'en segmentation d'images en biologie structurale, et d'autre part sur la prise en main et l'expérimentation de premières méthodes (méthodes existantes dans la littérature ou méthodes nouvelles) sur la base d'images existantes. La seconde année sera dédiée à la poursuite des expérimentations et au développement de méthodes novatrices combinant opérateurs connexes et réseaux de neurones convolutionnels, en se concentrant sur l'analyse de méthodes et d'architectures pertinentes dans le cadre de l'imagerie FIB/SEM. La dernière année sera consacrée au développement de l'atlas morphologique.

Le doctorant mènera ses recherches au sein de l'équipe IMAGEs du laboratoire ICube, sur le site du campus d'Illkirch, et collaborera activement avec l'équipe de biologie structurale intégrative de l'IGBMC afin d'acquérir de l'expertise dans le domaine biologique et en particulier sur les processus d'acquisition d'image.

Compétences requises

- Formation en informatique ou mathématiques appliquées
- Expérience en traitement d'images et apprentissage machine
- Connaissances en apprentissage profond
- Expérience de programmation en C++ et Python
- Des connaissances en biologie seraient un plus.

Références

- [Abreu et al. 2019] Abreu, A., Franchet, C., Frenois, F.-X., Brousset, P., Girard, J.P., Deneffe, P., Naegel, B., Wemmert, C. Ensemble of neural networks for high endothelial venules detection in Meca-79 immunohistochemistry images International Symposium on Biomedical Imaging (ISBI), 2019
- [Birchfield 2018] Birchfield S. Image processing and analysis CENGAGE Publisher
- [Bosch et al. 2015] Bosch, C., Martínez, A., Masachs, N., Teixeira, C.M., Feraud, I., Ulloa, F., Pérez-Martínez, E., Lois, C., Comella, J.X., DeFelipe, J., Merchán-Pérez, A., Soriano, E. FIB/SEM technology and high-throughput 3D reconstruction of dendritic spines and synapses in GFP-labeled adult-generated neurons. *Front. Neuroanat.* 9, 60, 2015
- [De Winter et al. 2009] De Winter, D.A.M., Schneijdenberg, C.T.W.M., Lebbink, M.N., Lich, B., Verkleij, A.J., Drury, M.R., Humbel, B.M. Tomography of insulating biological and geological materials using focused ion beam (FIB) sectioning and low-kV BSE imaging. *J. Microsc.* 233, 372–383, 2009
- [Farabet et al. 2013] Farabet C., Couprie C., Najman, L. LeCun, Y. Learning Hierarchical Features for Scene Labeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence.*35(8), pp. 1915-1929, 2013
- [Haberl et al. 2018] Matthias G. Haberl, Christopher Churas, Lucas Tindall, Daniela Boassa, Sébastien Phan, Eric A. Bushong, Matthew Madany, Raffi Akay, Thomas J. Deerinck, Steven T. Peltier, Mark H. Ellisman CDeep3M—Plug-and-Play cloud-based deep learning for image segmentation *Nature Methods* 15, 677–680, 2018
- [Hoang et al. 2017] Hoang T.V., Kizilyaprak C., Spohner D., Humbel B.M. and Schultz P. Automatic Segmentation of High Pressure Frozen and Freeze-Substituted Mouse Retina Nuclei from FIB-SEM Tomograms *J. Struct. Biol.* 197(2) :123-134, 2017.
- [Jia et al. 2014] Jia Y., Shelhamer E., Donahue J. Karayev S. and Long J. Girshick, R., Guadarrama S. Darrell T. Caffe : Convolutional Architecture for Fast Feature Embedding arXiv :1408.5093, 2014

- [Jones et al. 2015] Jones, C., Liu, T., Cohan, N. W., Ellisman, M. Tasdizen, T. Efficient semi-automatic 3D segmentation for neuron tracing in electron microscopy images *Journal of neuroscience methods* 246, 13-21, 2015.
- [Kizilyaprak et al. 2014] Kizilyaprak, C., Bittermann, A. G., Daraspe, J., Humbel, B. M. FIB-SEM tomography in biology. *Methods in molecular biology* 1117, 541-558, 2014.
- [Martinez-Sanchez et al. 2014] Martinez-Sanchez, A., Garcia, I., Asano, S., Lucic, V., Fernandez, J. J. Robust membrane detection based on tensor voting for electron tomography *Journal of structural biology* 186, 49-61, 2014.
- [Merchán-Pérez et al. 2009] Merchán-Pérez, A., Rodriguez, J.-R., Alonso-Nanclares, L., Schertel, A., Defelipe, J. Counting Synapses Using FIB/SEM Microscopy : A True Revolution for Ultrastructural Volume Reconstruction. *Front. Neuroanat.* 3, 18., 2009
- [Najman and Talbot 2010] Najman L., Talbot H. *Mathematical Morphology : from theory to applications*. Najman Laurent and Talbot Hugues ed. ISTE-Wiley, ISBN : 978-1-84821-215-2, 2010
- [Nunez-Iglesias et al. 2014] Nunez-Iglesias, J., Kennedy, R., Plaza, S. M., Chakraborty, A., Katz, W. T. Graph-based active learning of agglomeration (GALA) : a Python library to segment 2D and 3D neuroimages *Frontiers in neuroinformatics* 8, 34, 2014
- [Romera-Paredes and Torr 2016] Romera-Paredes, B., and Torr, P.H.S. Recurrent Instance Segmentation *ECCV 2016 LNCS* : 312–329, 2016
- [Ronneberger et al. 2015] Ronneberger O., Fischer P., Brox T. U-Net : Convolutional Networks for Biomedical Image Segmentation Navab N., Hornegger J., Wells W., Frangi A. (eds) *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Lecture Notes in Computer Science, vol 9351. Springer, Cham, 2015
- [Salembier and Wilkinson 2009] Salembier P. and Wilkinson, M. H. F. Connected Operators : A review of region-based morphological image processing techniques *IEEE signal processing magazine* 26(6) :136–157
- [Sanches et al. 2019] Sanches, P. Meyer, C., Vigon, V., Naegel, B. Cerebrovascular network segmentation of MRA images with deep learning *International Symposium on Biomedical Imaging (ISBI)*, 2019
- [Schindelin et al. 2012] Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J. Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., Cardona, A. Fiji : an open-source platform for biological-image analysis *Nature methods* 9, 676-682, 2012.
- [Sommer et al. 2011] Sommer, C., Straehle, C., Kothe, U., Hamprecht, F. A. *ilastik : Interactive Learning and Segmentation Toolkit* *International Symposium on Biomedical Imaging*, 230-233, 2011.
- [Zeng et al. 2017] Tao Zeng, Bian Wu, Shuiwang Ji *DeepEM3D : approaching human-level performance on 3D anisotropic EM image segmentation* *Bioinformatics*, Volume 33, Issue 16, Pages 2555–2562, 2017