

Sujet de recherche - projet M1 ILC

Cytonet : Plateforme d'analyse d'images histopathologiques par réseau convolutionnel U-Net

Pr. Cédric Wemmert (ICube, Unistra)

Contexte scientifique

La pathologie consiste à essayer de comprendre des phénomènes microbiologiques en se basant sur l'examen de tissus vivants. Classiquement, cela implique une évaluation visuelle par des pathologistes d'un échantillon de cellules à l'aide d'un microscope, pour identifier les propriétés structurales des tissus. Actuellement, l'évaluation visuelle est en grande partie un processus très coûteux en temps humain et l'exactitude du diagnostic repose sur la formation et l'expérience personnelle du pathologiste.

Le domaine a connu plusieurs révolutions technologiques ces dernières années avec l'avènement de la microscopie virtuelle, qui consiste à convertir des biopsies de tissus en images digitales à très haute résolution (appelées lames virtuelles ou *whole-slide images* - WSI - voir un exemple sur la Figure 1) qui peuvent être consultées et analysées via un ordinateur. Ce nouveau domaine est souvent appelé pathologie numérique. Dans le cadre de la pathologie numérique, des efforts importants ont été déployés par la communauté informatique pour concevoir des outils d'analyse d'images, afin d'identifier par exemple les structures biologiques de base (stroma, cellules immunitaires, tumeurs, etc.) afin de faciliter la tâche des médecins / biologistes pour une interprétation (semi-)automatisée des lames virtuelles. La pathologie numérique est actuellement considérée comme l'une des techniques les plus prometteuses de la médecine diagnostique, afin d'obtenir plus rapidement un diagnostic plus fiable et un meilleur pronostic pour de nombreuses maladies importantes.

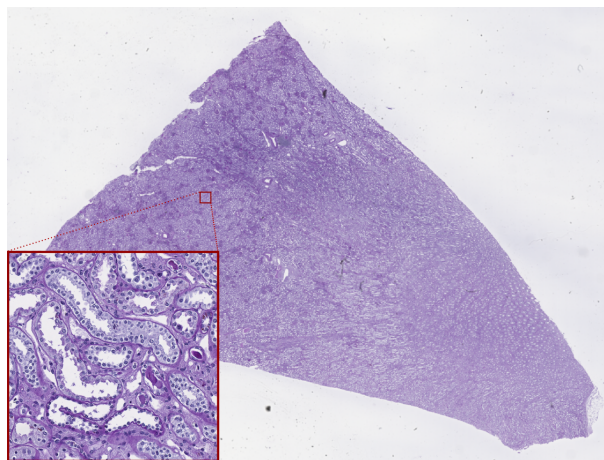


FIGURE 1 – Exemple de lame virtuelle complète (95615 × 72293 pixels).

Avec l'avènement récent de ces images, la pathologie numérique produit quotidiennement une quantité massive d'images associées à des métadonnées (informations sur le patient, diagnostic, traitement, etc.). Dans ce contexte de données massives, il existe de nombreux défis pour permettre l'utilisation de WSI dans les procédures de diagnostic en routine clinique. En effet, ces images contiennent des milliards de pixels et sont très hétérogènes, à la fois en terme de signal (dispositifs d'acquisition différents) et de sémantique (plusieurs types de cellules et de structures biologiques), ce qui provoque des difficultés techniques très importantes pour le traitement et la visualisation par des algorithmes classiques. En particulier, les approches traditionnelles d'analyse d'images ont maintenant atteint leurs limites, en raison de l'explosion de l'information induite par ces images. Notamment, avec l'augmentation de la résolution spatiale, les structures biologiques visibles dans les images deviennent de plus en plus complexes puisqu'elles apparaissent maintenant dans l'image comme une disposition spatiale de structures plus fondamentales. Pour développer la prochaine génération de systèmes informatiques, une direction de recherche clé consiste à transférer l'analyse d'images pathologiques d'un processus de bas niveau à un paradigme

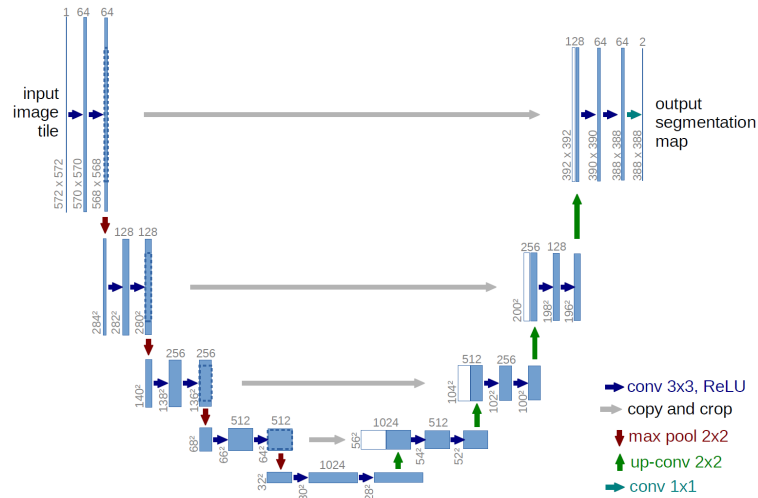


FIGURE 2 – Architecture du réseau convolutionnel U-Net

interactif de haut niveau, permettant d'intégrer judicieusement l'expertise utilisateur dans les outils d'analyse d'image.

Dans ce contexte médical, plusieurs défis informatiques majeurs sont soulevés. Le principal porte sur la conception de techniques d'analyse d'image efficaces, basées à la fois sur le traitement d'image et la fouille de données massives, pour faire face à la grande taille des images, leur variabilité importante et la complexité des objets biologiques à reconnaître. Pour faire face à ces défis, les techniques d'apprentissage profond par réseaux convolutionnels (*deep-learning*) ont montré leur pertinence sur de nombreuses applications [1, 3, 2].

Sujet

Il est admis que l'apprentissage par réseau de neurones profond nécessite un très grand nombre d'exemples annotés. U-net est une architecture particulière (voir Figure 2) de réseau de neurones convolutionnel dédié à la segmentation rapide et précise d'images. Elle a été proposée par Ronneberger *et al.* [4] lors du concours organisé durant la conférence ISBI 2015 (International Symposium on Biomedical Imaging - <http://biomedicalimaging.org/2015/>). U-Net propose une architecture et une stratégie d'apprentissage permettant de tirer partie plus efficacement des exemples disponibles qui s'appuie sur un principe d'augmentation des données. Il a été montré que cette technique permettait d'obtenir d'excellents résultats, plus rapidement et avec moins d'exemples que les méthodes existantes.

Une première version de plateforme logicielle (*Cytonet*) permettant d'entraîner un réseau U-Net sur des WSI a été mise en place dans notre équipe de recherche. L'objectif de ce projet est de poursuivre le développement de cette plateforme afin d'en simplifier l'utilisation pour le lancement de nouvelles expérimentations. Ce travail pourra être réalisé par 2 étudiants.

Contraintes logicielles - La plateforme *Cytonet* est développée en python et se base sur les librairies Keras et Tensorflow.

Références

- [1] G. Apou, N. Schaadt, B. Naegel, G. Forestier, R. Schönmeier, F. Feuerhake, Wemmert, C., and A. Grote. Detection of lobular structures in normal breast tissue. *Computers in Biology and Medicine*, 74 :91–102, Jul 2016.

- [2] Babak Ehteshami Bejnordi, Jimmy Linz, Ben Glass, Maeve Mullooly, Gretchen L Gierach, Mark E Sherman, Nico Karssemeijer, Jeroen van der Laak, and Andrew H Beck. Deep learning-based assessment of tumor-associated stroma for diagnosing breast cancer in histopathology images. *arXiv preprint arXiv:1702.05803*, 2017.
- [3] G. Litjens, T. Kooi, B. Ehteshami Bejnordi, A. Arindra Adiyoso Setio, F. Ciompi, M. Ghafoorian, J. A. W. M. van der Laak, B. van Ginneken, and C. I. Sánchez. A Survey on Deep Learning in Medical Image Analysis. *ArXiv e-prints*, February 2017.
- [4] O. Ronneberger, P. Fischer, and T. Brox. U-net : Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 9351 of *LNCS*, pages 234--241. Springer, 2015. (available on arXiv :1505.04597 [cs.CV]).