

Proposition de thèse de doctorat

Analyse des contributions des algorithmes de squelettisation à l'apprentissage par réseaux de neurones à convolution

Contexte

Le domaine du Machine Learning a apporté d'importantes contributions à l'analyse d'images, et de nombreux algorithmes d'apprentissage artificiels, comme les réseaux de neurones à convolution, ont été développés spécifiquement dans le but d'analyser des données visuelles. Aujourd'hui, ces méthodes sont très souvent adoptées pour résoudre des problèmes de reconnaissance visuelle, pour peu qu'une base de données d'apprentissage suffisamment volumineuse ait pu être constituée. Les performances de ces algorithmes (aussi bien en taux de détection qu'en temps de calcul) ont fortement progressé ces dernières années, et continuent de progresser grâce, notamment, à de nouvelles architectures de réseaux de neurones plus performantes.

Après que les réseaux de neurones aient apportés une précieuse contribution au domaine de l'image, certains travaux de ces dernières années ont examiné comment les algorithmes « classiques » du domaine de traitement d'images pourraient contribuer aux réseaux de neurones, et améliorer leurs performances en reconnaissance visuelle [1][2]. Parmi les stratégies explorées pour améliorer les réseaux de neurones à convolution à l'aide d'approches plus classiques, la littérature a déjà proposé le prétraitement des données d'entrée [1], le post traitement des résultats produits par un réseau [5], ou la modification de l'architecture du réseau pour imiter le cadre sémantique de certains algorithmes reconnus [3][4]. De nombreuses applications adoptent ces approches hybrides, afin de ne pas se baser uniquement sur un réseau de neurones pour faire de la reconnaissance visuelle.

Les algorithmes de squelettisation sont des méthodes reposant sur la topologie discrète, dont l'objectif est la simplification de formes (2d ou 3d) en un sous ensemble de pixels appelé squelette, capturant les propriétés géométriques importantes de la forme initiale [6][7][8]. De nombreux algorithmes de squelettisation ont été proposés par le passé, et leur étude a permis d'établir des liens forts entre les squelettes et la ligne de partage des eaux [9][10], un algorithme parfois utilisé de pair avec des réseaux de neurones pour obtenir de meilleurs résultats dans des tâches de segmentation d'objets [5]. Les squelettes font aujourd'hui l'objet d'une attention particulière de la communauté de l'apprentissage, de récents travaux tentant de reproduire les résultats de ces algorithmes de squelettisation avec des réseaux de neurones [11][12].

Les squelettes sont des algorithmes d'analyse de formes, fonctionnant sur des images binaires où les objets d'intérêt sont déjà délimités. Ils sont souvent sensibles au bruit dans la forme, mais des travaux réalisés au LAGA dans ce domaine ont permis d'aboutir à des méthodes plus robustes, et avec des résultats visuels convaincants [13]. Il est possible de proposer des algorithmes de squelettisation sur les images en niveaux de gris, directement construits à partir des algorithmes fonctionnant sur les formes. Pour produire de bons résultats, ces algorithmes doivent être robustes au bruit (souvent présent dans les images en niveaux de gris) et produire de bons résultats visuels, comme le fait notre approche.

Objectifs du projet de thèse

Ce projet consistera en une analyse profonde des liens entre les réseaux de neurones et les algorithmes de squelettisation. Dans un premier temps, nous souhaiterions mesurer à quel point un prétraitement d'images à l'aide d'un algorithme de squelettisation en niveaux de gris permettrait d'améliorer les résultats d'un réseau de neurones s'entraînant sur ces images. Pour cela, plusieurs questions devront être abordées : quelle stratégie de squelettisation convient le mieux, quelle architecture de réseau de neurone répond favorablement à un tel prétraitement, et des processus de squelettisation ont-ils déjà lieu au sein des couches du réseau de neurone pendant son apprentissage ?

Dans un second temps, nous souhaiterions aussi comprendre comment un réseau de neurones peut aider à construire de nouveaux algorithmes de squelettisation : un réseau peut-il apprendre et produire une carte des informations géométriques importantes d'une forme, peut-il comprendre l'information topologique d'une forme et aider à la préserver lors d'un processus d'amincissement, peut-il aider à mettre en place de nouvelles méthodes d'amincissement ?

Les réseaux de neurones sont constitués de couches de neurones, et chacune de ces couches opèrent des fonctions précises. Parmi ces opérations, deux dominent le paysage des réseaux dédiés à l'information visuelle : le « max-pool » et la « convolution ». La couche « max-pool » réalise un maximum local des valeurs des pixels de l'image traitée, ce qui fait invariablement penser à l'opération de dilatation, à la base de la morphologie mathématique. Les liens forts existants entre morphologie et squelettisation nous permettent de penser que de nouveaux types de couches, réalisant des opérations morphologiques ou des opérations de squelettisation, pourraient être mis en place et évalués dans le cadre de ce travail.

Situé à la croisée de l'apprentissage et de l'analyse de formes (appliqué au traitement d'images), ce sujet s'inscrit dans le cadre des projets structurants de l'équipe MBI. Les résultats tirés de ce travail pourront avoir une application directe dans la collaboration actuellement en cours avec l'hôpital Beaujon.

Pré requis

Le candidat retenu devra avoir de bonnes compétences en programmation (C, Python ou C++), en apprentissage artificiel et devra savoir rapidement utiliser les outils de squelettisation mis en place dans l'équipe (codés en C). Une connaissance d'une librairie d'apprentissage, comme TensorFlow, serait souhaitable.

Encadrement de la thèse

La thèse sera encadrée par Françoise DIBOS (LAGA) et co-encadrée par John CHAUSSARD (LAGA).

Bibliographie

- [1] Shorten, C., Khoshgoftaar, T.M. A survey on Image Data Augmentation for Deep Learning. J Big Data 6, 60 (2019). <https://doi.org/10.1186/s40537-019-0197-0>
- [2] S. Dodge and L. Karam, "Understanding how image quality affects deep neural networks," 2016 Eighth International Conference on Quality of Multimedia Experience (QoMEX), Lisbon, 2016, pp. 1-6, doi: 10.1109/QoMEX.2016.7498955.
- [3] M. Bai and R. Urtasun, "Deep Watershed Transform for Instance Segmentation," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, 2017, pp. 2858-2866, doi: 10.1109/CVPR.2017.305.
- [4] Yim, Jonghwa & Sohn, Kyung-Ah. (2017). Enhancing the Performance of Convolutional Neural Networks on Quality Degraded Datasets.
- [5] Couprie, Camille & Farabet, Clement & Najman, Laurent & Lecun, Yann. (2014). Convolutional Nets and Watershed Cuts for Real-Time Semantic Labeling of RGBD Videos. Journal of Machine Learning Research. 15.
- [6] Bertrand, Gilles & Couprie, Michel. (2017). Parallel skeletonization algorithms in the cubic grid based on critical kernels. 10.1016/B978-0-08-101291-8.00008-0.
- [7] Chausard, John & Couprie, Michel & Talbot, Hugues. (2011). Robust skeletonization using the discrete λ -medial axis. Pattern Recognition Letters. 32. 1384-1394. 10.1016/j.patrec.2010.09.002.
- [8] Saha, Punam & Borgfors, G. & Sanniti di Baja, Gabriella. (2017). Skeletonization: Theory, Methods and Applications.
- [9] Fernand Meyer "Skeletons and watershed lines in digital spaces", Proc. SPIE 1350, Image Algebra and Morphological Image Processing, (1 November 1990); <https://doi.org/10.1117/12.23578>
- [10] Laurent Najman, Michel Schmitt. Watershed of a Continuous Function. Signal Processing, Elsevier, 1994, 38 (1), pp.99-112. (10.1016/0165-1684(94)90059-0).

- [11] Shen, Wei & Zhao, Kai & Jiang, Yuan & Wang, Yan & Bai, Xiang & Yuille, Alan. (2016). DeepSkeleton: Learning Multi-Task Scale-Associated Deep Side Outputs for Object Skeleton Extraction in Natural Images. IEEE Transactions on Image Processing. PP. 10.1109/TIP.2017.2735182.
- [12] O. Panichev and A. Voloshyna, "U-Net Based Convolutional Neural Network for Skeleton Extraction," 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Long Beach, CA, USA, 2019, pp. 1186-1189, doi: 10.1109/CVPRW.2019.00157.
- [13] Chaussard, John & Noël, Laurent & Biri, Venceslas & Couprie, Michel. (2013). A 3D Curvilinear Skeletonization Algorithm with Application to Path Tracing. 119-130. 10.1007/978-3-642-37067-0_11.

Informations de contact

Françoise DIBOS, Professeure émérite
Laboratoire LAGA, UMR 7539, Institut Galilée, Université Sorbonne Paris Nord (USPN)
email : dibos@math.univ-paris13.fr

John CHAUSSARD, Maître de Conférence
Laboratoire LAGA, UMR 7539, Institut Galilée, Université Sorbonne Paris Nord (USPN)
email : chaussard@math.univ-paris13.fr