

Stage : Prise en compte du contexte dans la segmentation sémantique d'un nuage de points

Contexte

La modification d'une installation industrielle implique souvent sa modélisation via des logiciels de Conception Assistée par Ordinateur (CAO). La création d'un tel modèle nécessite le déplacement de personnel sur le site de l'installation pour y relever des mesures. SEGULA Technologies développe un système d'acquisition et de modélisation rapide en se basant sur des techniques de photogrammétrie et lasergrammétrie. Une étape essentielle du projet est la segmentation des nuages de points acquis. Cette étape fait l'objet d'une thèse avec le LabSTICC.



Figure 1 Exemple d'un nuage de points de site industriel

La segmentation sémantique consiste à séparer un média (image, nuage de points, vidéo...) en ses différents composants (pixel, point), c'est la segmentation, tout en précisant la nature des différents composants (sémantique). C'est un problème difficile étudié le plus souvent en machine learning au même titre que la classification ou la détection.



Figure 2 Exemple de segmentation sémantique sur une image du jeu de données ADE20K [1], [2]

L'un des défis majeurs de la segmentation sémantique est de combiner des connaissances, robustes aux transformations, de caractéristiques et l'utilisation du contexte global d'une scène. C'est-à-dire allier le Quoi au Où (mettre les caractéristiques d'un objet dans son contexte pour mieux le reconnaître).

Stage proposé

L'application de méthode de deep learning pour segmenter sémantiquement un nuage de points est une idée plutôt récente (jeu de donnée S3DIS ([3]) en 2016). Contrairement aux travaux effectués sur des images, les méthodes actuelles ne permettent pas d'exploiter le contexte global d'une scène représentée sous forme de nuage de points.

Si les premières méthodes de segmentation basées directement sur les points ([4], [5]) pèchent sur l'utilisation des caractéristiques spatiales locales (le Quoi), de nouvelles méthodes, plus tournées vers la segmentation et que la classification maîtrisent mieux cet aspect ([6]–[8]). Cependant, l'utilisation du contexte (le Où), reste un aspect du problème moins traité.

Une méthode employée dans la littérature est l'utilisation de la récurrence. On peut par exemple découper la scène en plusieurs blocs de longueur et largeur prédéfinis et de hauteur identique à la scène ([9]). Après extraction des caractéristiques de chaque blocs, des unités LSTM ([10]) sont utilisées pour ajouter du contexte aux points composants ces blocs. D'autre part, la solution de [11] segmente tout d'abord la scène via une méthode « classique » (sans machine learning), puis assemble les différents ensembles de caractéristiques (Super points) en un graphe de connectivité. Le contexte est alors propagé le long des arrêtes de ce graphe via des GRUs ([12]).

Si ce genre de méthode apporte de bons résultats, elles souffrent encore de certains défauts. L'utilisation de blocs artificiels ([9]) pour découper la scène pourrait gêner la compréhension de grands objets, qui sont fort présents dans des scènes industrielles. D'autre part, l'ajout d'une découpe verticale est nécessaire pour comprendre ce genre de scènes, parfois très structurée verticalement.

D'autre part, la solution proposée par [11] dépend fortement des résultats de la méthode de segmentation utilisée et ne fait que raffiner les supers points obtenus. Il est par exemple montré dans l'article qu'un tableau blanc (du set S3DIS [3]) ne sera pas séparé du mur lors de la segmentation et donc ignoré. La méthode ainsi présentée est donc intrinsèquement limitée. Des efforts pour améliorer la sur-segmentation de départ existent cependant ([13]).

Le but de votre stage est donc d'étudier les différentes méthodes pouvant renforcer la notion de contexte global dans un réseau de segmentation sémantique 3D. Les efforts seront plus particulièrement concentrés sur l'utilisation de la récurrence et l'utilisation des superpoints.

Organisation et objectifs

1. Etat de l'art : recensement des différentes méthodes renforçant la notion de contexte dans un réseau.
2. Comparaison des méthodes de l'état de l'art.
3. Sélection et test de plusieurs méthodes sur un jeu de données courant dans la littérature
4. Application sur un jeu de données différent (industriel) et identification de problèmes liés au domaine d'application.
5. Propositions d'améliorations pour pallier les problèmes précédemment identifiés.
6. Implémentation et tests de ces propositions.

Profil et compétences

Master 2 recherche en informatique ou vision.

Programmation python avec connaissance de bibliothèques de machine learning telles que TensorFlow, Caffe ou Pytorch.

Des connaissances en C++ ou de CUDA peuvent être un plus.

Contacts

Doctorant	LabSTICC	SEGULA Technologies
Romain CAZORLA Doctorant LabSTICC - Ingénieur informatique Recherche et Innovation, SEGULA Technologies romain.cazorla@segula.fr T : +33 (0)2 40 00 74 83	Cédric BUCHE Professeur (HDR), ENIB buche@enib.fr T : +33 (0)2 98 05 89 66 Panagiotis PAPADAKIS Ass. Professor, IMT Atlantique panagiotis.papadakis@imt-atlantique.fr T : +33 (0)2 29 00 16 30	Line POINEL Responsable Recherche & Innovation, SEGULA Technologies line.poinel@segula.fr M: +33 (0)6 10 42 34 37

Références

- [1] B. Zhou, H. Zhao, X. Puig, S. Fidler, A. Barriuso, and A. Torralba, "Scene Parsing through ADE20K Dataset."
- [2] B. Zhou *et al.*, "Semantic Understanding of Scenes through the ADE20K Dataset."
- [3] I. Armeni *et al.*, "3D Semantic Parsing of Large-Scale Indoor Spaces," in *CVPR*, 2016, pp. 1534–1543.
- [4] C. R. Qi, H. Su, M. Kaichun, and L. J. Guibas, "PointNet : Deep Learning on Point Sets for 3D Classification and Segmentation," *IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017.
- [5] C. R. Qi, L. Yi, H. Su, and L. J. Guibas, "PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space," in *NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems*, 2017, pp. 5105–5114.
- [6] M. Jiang, Y. Wu, T. Zhao, Z. Zhao, and C. Lu, "PointSIFT: A SIFT-like Network Module for 3D Point Cloud Semantic Segmentation," *ArXiv Prepr.*, no. 2, Nov. 2018.
- [7] H. Zhao, L. Jiang, C.-W. Fu, and J. Jia, "PointWeb: Enhancing Local Neighborhood Features for Point Cloud Processing," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 5560–5568.
- [8] Y. U. E. Wang, Y. Sun, Z. Liu, S. E. Sarma, M. M. Bronstein, and J. M. Solomon, "Dynamic Graph CNN for Learning on Point Clouds," *ACM Trans. Graph.*, vol. 1, no. 2, Oct. 2019.
- [9] X. Ye, J. Li, H. Huang, L. Du, and X. Zhang, "3D Recurrent Neural Networks with Context Fusion for Point Cloud Semantic Segmentation," in *Proceedings of the European Conference on Computer Vision*, 2018, pp. 403–417.
- [10] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [11] L. Landrieu and M. Simonovsky, "Large-scale Point Cloud Semantic Segmentation with Superpoint Graphs," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 4558–4567.
- [12] H. Su, S. Maji, E. Kalogerakis, and E. Learned-Miller, "Multi-view Convolutional Neural Networks for 3D Shape Recognition," in *IEEE International Conference on Computer Vision*, 2015, pp. 945–953.
- [13] L. Landrieu and M. Boussaha, "Point Cloud Oversegmentation with Graph-Structured Deep Metric Learning," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 7440–7449.