

Sujet de la thèse	Segmentation sémantique des nuages de point 3D via l'apprentissage profond	
Spécialité de la thèse	Intelligence Artificielle / Vision par Ordinateur	
Mots-clefs	Nuages de point 3D, segmentation sémantique, apprentissage profond, interprétation des scènes	
Type de thèse	Cotutelle	
Pays	France	à définir
Laboratoire d'accueil	L@bISEN Yncréa Ouest, Equipe Vision-AD, Brest ou Nantes	à définir
Directeurs de thèse	Maher Jridi	à définir
Encadrants	Ayoub Karine, Thibault Napoléon	

Contexte

La compréhension et l'interprétation d'une scène 3D complexe est une tâche visuelle innée chez l'être humain qui peut être effectuée de manière instantanée et sans effort. Déléguer cette tâche à la machine afin de l'automatiser est un domaine qui a suscité la curiosité de plusieurs chercheurs de la communauté de la vision par ordinateur [1, 2, 3].

Dans la dernière décennie, les avancées technologiques ont permis l'acquisition d'un volume important et représentatif de scènes 3D réels sous forme de nuages de points. Parmi ces systèmes d'acquisition, on cite les scanners LIDAR (Light Detection And Ranging) et les caméras RGB-D (Red Green Blue – Depth). La segmentation de ces nuages de points en régions homogènes permet, ainsi, de générer une représentation sémantique de la scène observée. On s'intéressera plus précisément à l'interprétation des scènes d'intérieur (pièces, bureau...) et d'extérieur (bâtiments, villes, zones rurales...). Un exemple de segmentation sémantique d'une scène d'intérieur est donné dans la Figure 1.

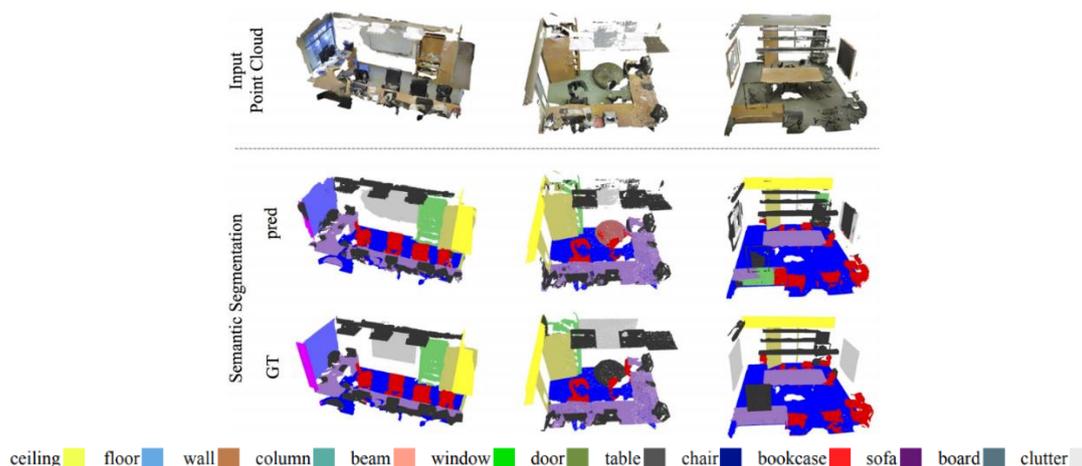


Figure 1 : Exemple d'une segmentation sémantique des nuages de points 3D [3]

Projet

L'objectif de la présente thèse est de proposer de nouvelles méthodes pour la segmentation sémantique des scènes d'intérieur et d'extérieur à partir des nuages de point 3D. L'objectif majeur serait de proposer de nouvelles méthodes et de les valoriser dans des conférences et journaux scientifiques. D'autre part, le sujet de thèse se veut aussi applicatif. Le candidat serait amené à

développer des prototypes opérationnels pour l'interprétation des scènes. Dans ce cadre, mise à part les bases de données publiques disponibles dans la littérature, des données privées provenant de nos partenaires industriels pourront aussi être exploitées dans le cadre d'applications bien précises.

Problématiques

Les nuages de points 3D modélise une représentation parcimonieuse d'une scène ou d'un objet. Ces nuages de points sont souvent non-structurés et non-ordonnés. Autrement dit, il n'existe pas un moyen pour les regrouper directement dans une grille afin de les considérer étant des pixels d'une image 2D.

La segmentation sémantique d'une scène, représentée par un nuage de points 3D, consiste à affecter des classes à chaque point 3D. Cette tâche est établie en utilisant une méthode de classification souvent supervisée. Pour ce faire, l'approche classique consiste à classifier les descripteurs (souvent liés à la forme) extraits à partir des nuages de points. Récemment, la disponibilité de bases de données composées d'un nombre important de nuages de point 3D annotés (NPM3D, S3DIS, Paris-Lille 3D...) a rendu difficile l'adoption de ces méthodes classiques. Pour remédier à cette difficulté, plusieurs chercheurs ont opté pour l'utilisation de l'apprentissage profond et plus précisément les réseaux de neurones convolutifs (Convolutional Neural Networks, CNN). Cette méthode a prouvé sa performance dans plusieurs applications liées à l'image 2D. Elle se base sur des convolutions discrètes exigeant une structure sous forme de grilles de données. De ce fait, l'utilisation directe de cette méthode pour le cas des nuages 3D reste impossible. Pour remédier à cela, deux familles de méthodes sont proposées dans littérature [4, 5]. La première famille consiste à projeter les points dans un autre espace à travers lequel une convolution discrète est possible [6, 7]. Quant à la seconde famille, elle tente de modifier la méthode CNN pour prendre en considération la nature complexe des nuages de points [3, 8]. Dans le cadre de la présente thèse, nous nous focaliserons sur le deuxième type d'approche qui est en début d'investigation.

Approches méthodologique et technique envisagées

- Etat de l'art sur la segmentation des nuages de points 3D
- Etude avancée et critique des méthodes de littérature traitant la segmentation des nuages de points 3D en utilisant l'apprentissage profond. L'implémentation de quelques-unes de ces méthodes en python sera un plus pour une bonne compréhension des approches existantes.
- Proposition de nouvelles contributions permettant d'améliorer l'existant en généralisant la convolution discrète utilisée dans le cadre des images 2D pour le cas des nuages de points 3D
- Développement d'un système opérationnel d'interprétation de scènes

Références succinctes

- [1] Boulch, A., Guerry, J., Le Saux, B., & Audebert, N. (2018). SnapNet: 3D point cloud semantic labeling with 2D deep segmentation networks. *Computers & Graphics*, 71, 189-198.
- [2] Pierre Biasutti, Vincent Lepetit, Mathieu Brédif, Jean-François Aujol, Aurélie Bugeau. LU-Net: An Efficient Network for 3D LiDAR Point Cloud Semantic Segmentation Based on End-to-End-Learned 3D Features and U-Net. *IEEE International Conference on Computer Vision Workshops (ICCV)*, Oct 2019, Séoul, South Korea
- [3] Qi, C. R., Su, H., Mo, K., & Guibas, L. J. (2017), "Pointnet: Deep learning on point sets for 3d classification and segmentation". In *Proceedings of the IEEE CVPR* (pp. 652-660).
- [4] Yulan Guo, Hanyun Wang, Qingyong Hu*, Hao Liu*, Li Liu, and Mohammed Bennamoun. (2019), "Deep Learning for 3D Point Clouds: A Survey". *arXiv preprint arXiv:1912.12033*
- [5] J. Zhang, X. Zhao, Z. Chen and Z. Lu, "A Review of Deep Learning-Based Semantic Segmentation for Point Cloud," in *IEEE Access*, vol. 7, pp. 179118-179133, 2019.

[6] Wu Z, Song S, Khosla A, Yu F, Zhang L, Tang X, et al. 3D shapenets: a deep representation for volumetric shapes. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2015. p. 1912–20. 702

[7] Maturana D, Scherer S. Voxnet: A 3D convolutional neural network for real-time object recognition. In: Proceedings of the IEEE/RSJ international conference on intelligent robots and systems (IROS). IEEE; 2015. p. 922–8.

[8] Boulch, A. (2020). ConvPoint: Continuous convolutions for point cloud processing. Computers & Graphics.

Profil recherché

Le ou la candidat(e) doit :

- Être titulaire d'un diplôme de Master et/ou Ingénieur, ou en cours de préparation, dans des domaines liés à l'informatique et la science de données
- Avoir un vif intérêt pour la recherche scientifique et être familier au moins avec l'un des outils/langages suivants (python, Keras, TensorFlow...)
- Avoir une aptitude au développement de prototypes
- Avoir de très bonnes connaissances théoriques et pratiques en Intelligence Artificielle, plus précisément en Deep Learning et en Vision par Ordinateur

Financement de la thèse en cotutelle

Le doctorant, au cours de sa période de présence en France (période de 4 à 6 mois par année), est financé avec une bourse de thèse à hauteur de 1000€/mois.

Le sujet de thèse est financé par les fonds propres du laboratoire L@bISEN de l'ISEN Yncréa Ouest.

Possibilité de faire des heures d'enseignement.

Modalités de candidature

Le dossier de candidature doit comprendre votre CV, lettre de motivation, Relevés de notes de L3, M1, M2 (ou années équivalentes) et une éventuelle lettre de recommandation. L'ensemble du dossier doit être adressé avant le 12/07/2020 à :

- maher.jridi@isen-ouest.yncrea.fr
- ayoub.karine@isen-ouest.yncrea.fr
- thibault.napoleon@isen-ouest.yncrea.fr

en précisant le moyen grâce auquel vous avez eu accès au sujet.

Les dossiers seront traités par ordre d'arrivée. Seuls les dossiers complets seront considérés et seuls les candidats retenus seront contactés et convoqués pour un entretien en visio-conférence.