

Proposition de stage niveau Master 2

Segmentation de scène par deep learning pour une application de SLAM visuel robuste

Mots clés : Deep learning, segmentation de scène, détection d'objet, SLAM visuel, vision par ordinateur

1. Contexte

Le laboratoire IBISC se positionne à travers l'équipe IRA2 sur l'amélioration de l'interaction entre les personnes et les systèmes complexes, et s'organise en particulier autour de l'étude de la Perception, l'Interprétation et la Décision (PID) appliquée à des environnements réels ou virtuels.

Dans cette thématique, des travaux de recherche sont effectués sur la perception du mouvement d'un robot humanoïde et de son environnement grâce à des méthodes de vision par ordinateur et de robotique mobile (SLAM visuel entre autre). L'ensemble du système robotique prend place dans un système multiple comprenant une plateforme hybride de Gough-Stewart qui sert de simulateur de mouvement et un robot NAO placé sur la partie supérieure de la plateforme et qui se substitue à l'humain. Cette association est opérée en vue de concevoir des algorithmes et scénarios de simulation qui soient adaptés à la réhabilitation de personnes présentant un handicap moteur.

Ce stage de recherche s'articule autour de la conception d'une méthode de segmentation de scène afin de concevoir par la suite un système de SLAM visuel robuste capable de détecter les dynamiques de scènes en temps réel. Le candidat devra entreprendre une veille bibliographique sur les méthodes de segmentation de scène et de détection d'objet pour s'imprégner des concepts qui y sont liés. Le candidat devra ensuite proposer un algorithme d'apprentissage profond basé sur la méthode YOLOv4 qui réalise une segmentation sémantique de scène en plus de la détection d'objet. Une étude comparative sera à réaliser afin d'évaluer la famille de méthode la plus adaptée entre les méthodes DeepLab, Mask RCNN et l'algorithme YOLO proposé.

2. Objectifs

Les objectifs du stage sont les suivants :

- Réaliser un état de l'art des méthodes de segmentation de scène et de détection d'objet.
- Comprendre les notions et concept lié à la segmentation de scène et la détection d'objet.
- Comprendre et prendre en main le modèle de détection d'objet YOLOv4 basé sur les convolutions de réseaux de neurones.
- Proposer un algorithme d'apprentissage profond pour la segmentation de scène basé sur la méthode YOLOv4.
- Evaluer l'algorithme par rapport à l'état de l'art.

3. Profil recherché

Vous êtes curieux, créatif autonome et rigoureux. Votre intérêt pour l'innovation, la recherche en vision par ordinateur en traitement d'images vous caractérise. Vous présentez un réel intérêt pour le domaine de l'apprentissage profond appliqué à la vision par ordinateur.

4. Formation académique recherchée

- Titulaire d'un Bac+5 (Master ou ingénieur) en traitement d'image/vision par ordinateur ou informatique.
- Compétence en vision par ordinateur et/ou intelligence artificielle ainsi que développement logiciel (C/C++, Python).
- Maîtrise des outils appliqués à la vision par ordinateur: OpenCV, Matlab, Tensorflow, PyTorch...
- Anglais : bon niveau.

5. Modalités de candidature :

Les candidatures (CV + lettre de motivation) sont à envoyer directement par email au contact ci-dessous. La transmission de relevés de notes sera appréciée.

Durée du stage: 6 mois Date de début de stage : Février ou mars 2022

Encadrement et contact :

- Malik Mallem (Professeur, IBISC), malik.mallem@univ-evry.fr
- Ayman Beghdadi (Doctorant, IBISC), aymanaymar.beghdadi@univ-evry.fr

Références :

- [1] Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). Yolov4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*.
- [2] Mohamed, E., Shaker, A., Rashed, H., El-Sallab, A., & Hadhoud, M. (2021). INSTA-YOLO: Real-Time Instance Segmentation. *arXiv preprint arXiv:2102.06777*.
- [3] Sarmiento, J. A. (2021). Pavement Distress Detection and Segmentation using YOLOv4 and DeepLabv3 on Pavements in the Philippines. *arXiv preprint arXiv:2103.06467*.
- [4] Wang, Z. (2019). SEG-YOLO: Real-time instance segmentation using YOLOv3 and fully convolutional network.
- [5] Albahli, S., Nida, N., Irtaza, A., Yousaf, M. H., & Mahmood, M. T. (2020). Melanoma lesion detection and segmentation using YOLOv4-DarkNet and active contour. *IEEE Access*, 8, 198403-198414.
- [6] Cheng, B., Collins, M. D., Zhu, Y., Liu, T., Huang, T. S., Adam, H., & Chen, L. C. (2019). Panoptic-deeplab. *arXiv preprint arXiv:1910.04751*.

- [7] Wang, H., Zhu, Y., Adam, H., Yuille, A., & Chen, L. C. (2021). Max-deeplab: End-to-end panoptic segmentation with mask transformers. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 5463-5474).
- [8] Wang, H., Zhu, Y., Adam, H., Yuille, A., & Chen, L. C. (2021). Max-deeplab: End-to-end panoptic segmentation with mask transformers. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 5463-5474).
- [9] Liu, C., Chen, L. C., Schroff, F., Adam, H., Hua, W., Yuille, A. L., & Fei-Fei, L. (2019). Auto-deeplab: Hierarchical neural architecture search for semantic image segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 82-92).
- [10] He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision* (pp. 2961-2969).
- [11] Tsung-Yi Lin, Piotr Dollár, Ross Girshick, Kaiming He, Bharath Hariharan, and Serge Belongie. Feature pyramid networks for object detection, 2017.