



Des images de synthèse pour entraîner les réseaux de neurones

Génération d'images aériennes simulées

Marc Jaeger¹, Marc Chaumont², Frédéric Borne¹

1. UMR Amap (botAnique et Modélisation de l'Architecture des Plantes et des végétations), Montpellier.
2. UMR LIRMM (laboratoire de Recherche en Informatique, Robotique et Microelectronique de Montpellier), Montpellier.
borne@cirad.fr, marc.chaumont@lirmm.fr

Objet d'étude :

- a) **Synthèse de scènes 3D** contenant des arbres (l'architecture est réaliste) ressemblant à une photo aérienne prise par un drone (i.e. génération d'un DOE - Design of Experiences).
- b) **Apprentissage d'un réseau de neurones de type Faster-RCNN** [Ren et al. 2021 - Faster-RCNN] **ou Yolo** [YOLOV5], [Ge et al. 2021 - YOLOX] dédié à la détection d'arbres à partir des images de synthèse.
- c) **Transfert de l'apprentissage pour que le réseau de neurones** soit capable de réaliser des détections dans des images réelles. Il y a de nombreuses solutions pour cette « adaptation de domaine ».

Problématique :

Identifier et dénombrer des arbres de manière automatique à partir d'images RGB satellitaires/drones, et plus encore en zones de forêts naturelles, est un vrai challenge. Les variétés, l'âge de la plantation, l'état phytosanitaire, la densité impliquent en effet une grande variabilité dans le contenu des images ce qui rend la détection automatique difficile, même par deep learning.

Objectif du stage

Les approches de détection dans des images RGB par deep learning sont prometteuses [Arce et al. 2021]. Cependant, de manière générale, afin d'obtenir des performances intéressantes, il est nécessaire de disposer d'un grand nombre d'exemples d'apprentissage annotés. Dans notre cas, il est nécessaire de disposer d'images « réelles » vues depuis un drone ou un satellite et de connaître la position et la variété de chaque arbre dans l'image ; c'est ce que l'on appelle la « vérité terrain ». Or, actuellement, il n'y a pas de base de données suffisamment grande pour les forêts naturelles, d'où l'idée d'utiliser des images de synthèse.

Le laboratoire AMAP développe des modèles structure fonction de plantes (FSPM) qui permettent de simuler la structure 3D [Jaeger & de Reffye, 1992] et le fonctionnement de la plante. Des outils permettent ensuite de créer des scènes virtuelles et de les visualiser off line ou en temps réel [Zhang et al., 2017] en mode « réaliste » ou label. On peut ainsi générer un grand nombre d'images « exemples annotés » qui peuvent être utilisées pour l'apprentissage d'un réseau de neurones. Ainsi, dans le cadre de ce stage, nous souhaitons **vérifier que**



l'utilisation d'image de synthèse peut, in fine, améliorer les performances d'un algorithme de détection dans des images réelles.

Descriptif du stage

Dans un premier temps, nous souhaitons utiliser l'algorithme de synthèse afin de simuler des *images aériennes virtuelles, censées être comparables à des images réelles*. Il faudra ainsi prendre en main l'algorithme de synthèse qui est entre autre paramétré par la prise de vue, l'âge de la plantation, la densité des individus et le mélange de variétés et d'espèces. Afin d'avoir des images de synthèse architecturalement similaires aux images réelles (dont nous disposons de la vérité terrain ; ce sont des parcelles témoins), nous utiliserons des paramètres de synthèse semblables aux paramètres des images réelles.

Dans un second temps, nous réaliserons l'apprentissage d'un réseau de neurones de type Faster-RCNN ou Yolo à partir des images de synthèse et nous évaluerons ses performances.

Enfin, différents mécanismes de transfert ou d'adaptation de domaine seront envisagés afin de démontrer l'intérêt d'utiliser des images de synthèse dans un apprentissage [Ghifary et al. 2016 - DRCN], [Zhu et al. 2017 – CycleGAN], etc.

Ainsi, les étapes principales sont :

1. Création de fichiers de configurations d'entraînement (DOEs) à partir de l'analyse de parcelles témoins « réelles ».
2. Classification et évaluation par deep learning des images de synthèse.
3. Transfert ou adaptation de domaines pour obtenir un réseau de détection plus performant.

L'étudiant travaillera avec un spécialiste de la modélisation des plantes et de la génération d'images de synthèse. Il interagira avec d'autres étudiants et chercheurs plus spécifiquement en charge de traitement d'images ou de deep-learning.

Références :

[Arce et al. 2021] Arce, L.S.D., Osco, L.P., Arruda, M.d.S.d. et al. *Mauritia flexuosa palm trees airborne mapping with deep convolutional neural network*. *Sci Rep* 11, 19619 (2021). <https://doi.org/10.1038/s41598-021-98522-7> <https://www.nature.com/articles/s41598-021-98522-7.pdf>

[Ren et al. 2015 - Faster-RCNN] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," in *Proceedings of the Advances in Neural Information Processing Systems, NIPS'2015*, Montreal, Canada, Dec. 2015, vol. 28, pp. 91–99, Curran Associates, Inc

Jaeger M., De Reffye P., 1992. "Basic concepts of computer simulation of plant growth". In: *Journal of Biosciences*, 1992. vol. 17, no. 3, p. 275-291. https://www.imwa.info/docs/imwa_1999/IMWA1999_Jaeger_125.pdf

[YOLOv5] <https://github.com/ultralytics/yolov5/3><https://github.com/ultralytics/yolov5/>

[Ge et al. 2021 - YOLOX] Zheng Ge, Songtao Liu, Feng Wang, Zeming Li, and Jian Sun, "YOLOX: exceeding YOLO series in 2021," CoRR, vol. abs/2107.08430, 2021

[Ghifary et al. 2016 - DRCN] Deep Reconstruction-Classification Networks for Unsupervised Domain Adaptation Ghifary M., Kleijn W.B., Zhang M., Balduzzi D., Li W. (2016) Deep Reconstruction-Classification Networks for Unsupervised Domain Adaptation. In: Leibe B., Matas J., Sebe N., Welling M. (eds) Computer Vision – ECCV 2016. ECCV 2016. Lecture Notes in Computer Science, vol 9908. Springer, Cham. https://doi.org/10.1007/978-3-319-46493-0_36 <https://arxiv.org/pdf/1607.03516.pdf> :

[Zhu et al. 2017 – CycleGAN] Jun-Yan Zhu*, Taesung Park*, Phillip Isola, and Alexei A. Efros. "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks", in IEEE International Conference on Computer Vision (ICCV), 2017. <https://arxiv.org/pdf/1703.10593.pdf>

[Zhang et al., 2017] Zhang, Y. X., Bao, G., Meng, W., Jaeger, M., Li, H., Deussen, O., Chen, B., 2017. Tree branch level of detail models for forest navigation. Computer Graphics Forum, 36 (8): 402-417. <https://hal-lirmm.ccsd.cnrs.fr/lirmm-01925641/document>

Exemples d'images virtuelles pseudo réalistes et label.



Profil :

Elève Master 2 ou Master 1 ou Ingénieur : traitement d'images et signal/téledétection/bioinformatique

- Maîtrise d'un langage de programmation de bas niveau (ex: C++),
- Maîtrise d'un langage de programmation de script (ex. Python),



- Expérience en apprentissage profond (TensorFlow ou Pytorch),
- Connaissances en traitement d'images,
- Aptitude au travail en interdisciplinarité
- Prise d'initiative et bon relationnel,
- Anglais scientifique courant

Laboratoire d'accueil :

UMR Amap <http://amap.cirad.fr/fr/index.php> - CIRAD Montpellier

Conditions de stage :

Durée : 4 à 6 mois pendant le premier semestre 2022.

Indemnités : ~600 € / mois environ.

Contacts :

Marc Jaeger,

Frédéric Borne,

Marc Chaumont,

borne@cirad.fr,

marc.chaumont@lirmm.fr

Tél. 04 67 61 49 20

Contacts pour candidater (avec CV et notes du M1 et si possible du M2) :

Frédéric Borne,

Marc Chaumont,

borne@cirad.fr,

marc.chaumont@lirmm.fr