



Sujet de stage de fin d'étude/master 2

Etude des performances et quantification des incertitudes d'un réseau de neurones pour la classification de maladies des plantes lors d'un changement de contexte

Mots clés : deep learning, réseaux de neurones récurrents convolutifs (CNN), robustesse, méthodes d'ensemble, proxidétection, environnement.

Description du stage

Le secteur agricole fait actuellement face à de nombreux défis et à des changements structurels accentués par la démographie, le changement climatique, l'impact environnemental, les modes de consommation, la compétitivité, etc. Pour y faire face, les technologies du numérique (la proxidétection/téledétection, les capteurs, le traitement du signal et des images, l'intelligence artificielle, la robotique, ...), se présentent comme un des principaux leviers. Ces dernières années, la possibilité d'acquérir de grandes bases de données, les capacités de calcul accrues ainsi que des avancées théoriques clé ont permis l'essor de nouvelles méthodes d'apprentissage basées sur les réseaux de neurones profonds. En particulier, les réseaux de neurones convolutifs (CNN) sont devenus un standard pour de nombreuses tâches de classification. Néanmoins, plusieurs études ont montré que ces réseaux ne sont pas infaillibles notamment lorsque le contexte change entre la base d'entraînement et la base de test [1,2]. C'est notamment le cas lorsque l'on apprend un modèle sur des données acquises dans le sud de la France et qu'on le valide dans une région située au nord. De la même façon, la prédiction d'un CNN peut être erronée alors même qu'il indique accorder une forte confiance à cette prédiction. Dans ce contexte, l'objectif de ce stage est d'améliorer la robustesse des réseaux de neurones afin d'augmenter leur pouvoir prédictif, mais également de mieux quantifier la confiance sur leur prédiction.

Ce stage est réalisé en étroite collaboration avec une entreprise qui développe des solutions intelligentes pour les travaux agricoles et viticoles, afin d'améliorer la sécurité, la performance et l'impact environnemental. Dans le cadre de cette collaboration, le/la stagiaire développera des solutions algorithmiques et logicielles. Il/elle aura une vision d'ensemble sur les meilleures technologies, et une capacité d'action pour contribuer directement à l'adoption de ces technologies par les agriculteurs.

Le déroulement du stage s'articulera en trois parties :

- Dans un premier temps, le/la stagiaire réalisera une étude bibliographique sur les différentes approches de l'état de l'art afin d'améliorer la robustesse des réseaux de neurones. Il/elle étudiera en particulier les méthodes permettant de quantifier les performances ainsi que les incertitudes sur la prédiction faite par un CNN.
- Puis, dans un deuxième temps, à l'image des forêts aléatoires [3,4], nous proposerons une nouvelle architecture de CNN basée sur les méthodes d'ensemble afin d'améliorer la robustesse de ces réseaux de neurones.
- Et enfin, nous étudierons le potentiel de l'approche proposée en proxidétection sur des données réelles pour la classification de maladies de plantes.

Références

- [1] S. Dodge et L. Karam. *Understanding how image quality affects deep neural networks*. IEEE International conference on quality of multimedia experience (QoMEX), 2016
- [2] K. Ren, T. Zheng, Z. Qin et X. Liu. *Adversarial Attacks and Defenses in Deep Learning*. Engineering. Vol. 6, n° 3, pp. 346-360, 2020.
- [3] L. Breiman. *Random Forests*. Machine Learning. Vol. 45, n° 1, pp.5-32, 2001.
- [4] S. Akodad, L. Bombrun, J. Xia, Y. Berthoumieu et C. Germain. *Ensemble Learning Approaches Based on Covariance Pooling of CNN Features for High Resolution Remote Sensing Scene Classification*. Remote Sensing, Vol. 12, n° 3291, 2020.

Profil recherché

- Vous êtes en dernière année d'un **master ou école d'ingénieurs** avec une spécialisation en traitement du signal et des images ou mathématiques appliquées.
- Vous avez de bonnes connaissances en analyse d'images et classification de données. En particulier, vous avez des connaissances sur le **deep learning** et les **techniques d'apprentissage** supervisé.
- Vous avez de solides compétences dans les langages de développement tels que **Python, Matlab** et connaissez les principaux outils de **l'intelligence artificielle** (Pytorch, Tensorflow,...)
- Vous êtes **rigoureux et organisé**, pour aller au bout de vos missions dans des environnements complexes.
- Vous avez **l'esprit critique**, pour faire votre propre analyse dans la recherche de défauts de fonctionnement.
- Vous maîtrisez impérativement **l'anglais** à l'oral et l'écrit.

Ce stage peut donner suite à une poursuite en thèse de doctorat.

Organisation du stage et modalités de candidature

- Lieu du stage : Laboratoire IMS – Groupe Signal et Image, campus de Talence (33). Quelques missions dans les locaux de l'entreprise partenaire sont prévues.
- Durée : 6 mois à partir de février/mars 2021.
- Indemnité de stage : 1000€ bruts mensuels
- Envoyer par mail une lettre de candidature + un CV.

Contacts :

Lionel Bombrun, IMS, Tél : 05 4000 2473, lionel.bombrun@ims-bordeaux.fr

Jean-Pierre Da Costa, IMS, Tél : 05 4000 2634, jean-pierre.dacosta@ims-bordeaux.fr