

Sujet de thèse CIFRE :

Segmentation biophysique par réseaux de neurones appliquée à l'imagerie satellitaire multi-spectrale

Mots clés : Intelligence artificielle, Réseaux de neurones, *Deep Learning*, Classification, Observation de la terre, *Remote Sensing*.

1 Encadrement

Directeur de thèse : Michel CRUCIANU, professeur Cnam, laboratoire CEDRIC
Co-encadrant : Mihai DATCU, professeur au Deutsches Zentrum für Luft- und Raumfahrt (Allemagne) et titulaire de la chaire Blaise Pascal au Cnam, laboratoire CEDRIC
Encadrant CS SI : Sébastien DORGAN (ingénieur de recherche)

2 Contexte

Au cours des dernières années, l'évolution des méthodes d'intelligence artificielle de l'apprentissage statistique classique (SVM, forêts aléatoires, etc.) à l'apprentissage profond (réseaux de neurones profonds) ont ouvert la voie à d'énormes progrès dans des domaines tels que la vision par ordinateur ou la reconnaissance vocale. Ces progrès ont permis l'émergence d'innovations qui sont maintenant au cœur de notre quotidien : assistants vocaux, voitures autonomes, etc.

En parallèle, les méthodes d'apprentissage statistique appliquées aux données d'imagerie satellitaire ont permis de créer des chaînes de segmentation biophysique semi-automatique relativement efficaces^{1, 2}. Il serait donc naturel de penser que l'utilisation des méthodes d'apprentissage profond sur ces données conduirait à des avancées comparables à celles connues dans le domaine de la vision par ordinateur. Cependant, appliquées à des cas concrets, les résultats ne sont pas aussi probants et il est encore aujourd'hui impossible de fournir une segmentation biophysique à l'échelle d'un continent ou d'un pays de façon automatique et (quasi) temps réel.

Pour comprendre ce phénomène contre-intuitif, nous avons mené une étude des méthodes de classification existantes utilisant des techniques d'apprentissage profond sur des données satellitaires multispectrales [2, 5, 3, 1]. Les conclusions que nous avons tirées ne remettent pas en question les capacités des méthodes d'apprentissage profond, mais davantage les approches utilisées pour construire des bases de données d'apprentissage et la grande sensibilité au bruit des réseaux neuronaux.

3 Analyse

Les études que nous avons examinées forment leur base de données d'apprentissage et de validation en étiquetant chaque pixel à l'aide d'une base de données de référence et

1. "LC map at 20m over Africa", <http://2016africallandcover20m.esrin.esa.int/>
2. "The Sentinel 2-Agri algorithms", <http://www.esa-sen2agri.org/operational-system/algorithms/>

en divisant le produit multispectral dans le domaine spatial en tenseurs de petites tailles. Les tenseurs ainsi produits sont alors utilisés pour alimenter des réseaux de neurones convolutifs de type Inception, VGG ou ResNet, performants pour la classification ou la segmentation d’images colorisées.

Cette procédure conduit à injecter dans le réseau un grand nombre de pixels mal étiquetés provenant de l’imperfection des bases de données de référence. Or les réseaux de neurones convolutifs sont très sensibles aux bruits de la base d’apprentissage et de validation qui les amènent à créer un grand nombre de faux positifs.

Certaines études choisissent de ne pas tenir compte de la dimension temporelle du problème et obtiennent donc des résultats satisfaisants uniquement sur les classes pseudo-invariantes (constructions, rivières, lacs...) mais sont inefficaces sur les classes à fortes variabilités temporelles (cultures, forêts...). D’autres abordent la composante temporelle en créant des tenseurs tridimensionnels sur l’espace, le temps et le spectre. Dans ce cas, ils obtiennent de meilleurs résultats sur les classes avec des variabilités temporelles plus élevées mais pour des coûts de traitement énormes et pour un gain de précision très faible, voire nul par rapport aux méthodes d’apprentissage classique. Enfin, aucune de ces études ne prend en compte des paramètres exogènes comme les données climatiques ou n’utilise d’indicateur consolidé pour réduire les biais de mesures ni de données fusionnées pour profiter des synergies entre capteurs satellitaires de nature différente (optique, radar, interféromètre...).

4 Objectifs et approche

Compte tenu de cette analyse il nous semble nécessaire d’aborder la problématique de la segmentation biophysique appliquée aux données satellitaires par l’élaboration d’une approche générique pour la création des bases de données d’entraînement et de validation permettant de :

- Réduire le nombre de pixels mal étiquetés injectés dans le réseau pendant l’entraînement.
- Prendre en compte la dimension temporelle du problème sans faire exploser les coûts de calcul.
- Prendre en compte la variation des conditions climatiques.
- Assurer la continuité avec les séries chronologiques à long terme existantes.
- Réduire l’impact des changements dans les conditions d’observation.
- Combiner des données provenant de bases de données de référence hétérogènes (région d’intérêt, échelle, précision...).

Compte tenu de la diversité des données géospatiales et des « objets » observés, il faudra examiner, adapter et combiner des méthodes d’intelligence artificielle et des approches statistiques pour créer une méthode hybride s’adaptant au mieux à la nature des données d’entraînement et à l’apprentissage des classes cibles.

Cette première partie de l’étude conduira à la création d’une chaîne de traitement produisant une segmentation biophysique sur une zone d’intérêt géographique où l’on dispose de séries temporelles d’images satellitaires (SITS) multi-spectrales complètes et de bases de données de référence denses. Un cas d’étude pourrait être de produire une segmentation en s’appuyant sur les produits Sentinel 2 pour les séries temporelles et sur

le Registre Parcellaire Graphique (RPG) pour la base de données de référence.

D'un point de vue plus général, les observations multi-spectrales sont affectées par les occlusions de nuages (en raison des conditions atmosphériques), sont à des intervalles de temps irréguliers, emploient des capteurs multiples et différents. Après avoir abordé le cas des SITS complètes, cette thèse étudiera et proposera des solutions pour l'apprentissage avec des SITS incomplètes, irrégulières et multi-capteurs. S'il est envisageable de transposer dans ce contexte des techniques d'imputation de données manquantes (par ex. basées sur des GAN, voir [4]), nous souhaitons plutôt explorer des solutions architecturales nouvelles avec des techniques de régularisation adaptées.

Enfin, la durée de vie d'une mission satellite d'observation de la Terre n'étant que de quelques années, il est important de pouvoir adapter une chaîne de traitement existante aux données d'une nouvelle mission. Pour cette raison, cette thèse s'intéressera également au transfert de l'apprentissage acquis d'un capteur satellite vers un autre afin de proposer une méthode permettant d'assurer la continuité de l'exploitation de la chaîne de traitement dans le temps. L'évaluation pourra être réalisée, par exemple, sur le transfert de Landsat 8 vers Sentinel 2 ou de SPOT vers Pléiade. La nature de ces transferts permet d'envisager des solutions qui s'éloignent des techniques actuellement employées pour le transfert d'apprentissage dans les réseaux profonds.

Références

- [1] P. Benedetti, Dino Ienco, Raffaele Gaetano, Kenji Ose, Ruggero G. Pensa, and Stéphane Dupuy. M3Fusion : A deep learning architecture for multi-{Scale/Modal/Temporal} satellite data fusion. *CoRR*, abs/1803.01945, 2018.
- [2] Amina Ben Hamida, Alexandre Benoit, Patrick Lambert, Louis Klein, Chokri Ben Amar, Nicolas Audebert, and Sébastien Lefèvre. Deep learning for semantic segmentation of remote sensing images with rich spectral content. In *2017 IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2017, Fort Worth, TX, USA, July 23-28, 2017*, pages 2569–2572, 2017.
- [3] Shunping Ji, Chi Zhang, Anjian Xu, Yun Shi, and Yulin Duan. 3d convolutional neural networks for crop classification with multi-temporal remote sensing images. *Remote Sensing*, 10(1) :75, 2018.
- [4] Jinsung Yoon, James Jordon, and Mihaela van der Schaar. GAIN : Missing data imputation using generative adversarial nets. In Jennifer Dy and Andreas Krause, editors, *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pages 5689–5698, Stockholmsmässan, Stockholm Sweden, 10–15 Jul 2018. PMLR.
- [5] Atsushi Yoshihara, Tetsuya Takiguchi, and Yasuo Ariki. Feature extraction and classification of multispectral imagery by using convolutional neural network. In *International Workshop on Frontiers of Computer Vision*, 2017.