

ZGP appliqué à la segmentation des images hyperspectrales

I. Grenet^{1,2}, Y. Bobichon¹, A. Girard¹

¹ Institut de Recherche Technologique Saint Exupéry - Sophia Antipolis, France

² MyDataModels - Sophia Antipolis, France

ingrid.grenet@mydatamodels.com
{yves.bobichon, adrien.girard}@irt-saintexupery.com

Résumé

La segmentation d'images satellites hyperspectrales est une tâche utile pour de nombreuses applications (cartographie, suivi de l'environnement, etc.) mais qui s'avère difficile, notamment à cause du manque de données annotées. Cet article présente les avantages de l'algorithme évolutionnaire ZGP pour le traitement de l'imagerie hyperspectrale dans les systèmes embarqués (frugalité, interprétabilité, simplicité) et montre comment il est capable de rivaliser avec des réseaux de neurones classiques en terme de performances.

Mots-clés

Hyperspectral, télédétection, ZGP, CNN, frugalité.

Abstract

Segmentation of hyperspectral satellite images is a useful task for many applications (mapping, environmental monitoring, etc.) but is difficult, especially due to the lack of annotated data. This paper presents the advantages of the ZGP evolutionary algorithm for processing hyperspectral images in embedded systems (frugality, interpretability, simplicity) and shows how it competes with classical neural networks in terms of performance.

Keywords

Hyperspectral, remote sensing, ZGP, CNN, frugality.

1 Introduction

L'imagerie satellite hyperspectrale utilise des mesures à haute résolution spectrale de la lumière permettant des applications dans de nombreux domaines comme l'agriculture, la gestion des ressources naturelles, la caractérisation des sols et des matériaux de surface, la prévention ou le suivi de catastrophes naturelles et le changement climatique. Les imageurs hyperspectraux génèrent d'énormes quantités de données qui nécessitent la mise en oeuvre de traitements embarqués au plus près du capteur afin d'assurer le stockage et la transmission de l'information utile avec la meilleure qualité possible [7] [12].

L'utilisation de réseaux de neurones profonds pour le traitement d'images à bord de satellites impose de lever de nombreux verrous liés à la disponibilité de bases de données annotées, la complexité des algorithmes et aux capacités de calculs et de stockage limités de ces systèmes embarqués

très contraints. Les techniques d'apprentissage frugal, semi ou non supervisé [3] offrent des solutions permettant de compenser le manque de données annotées au prix de performances souvent dégradées. De même, les algorithmes de vision par ordinateur développés pour les images RGB ne sont pas directement transposables aux données hyperspectrales qui portent par nature une information spectrale très riche nécessitant des méthodes d'analyse spécifiques. Parmi ces méthodes, la classification 1D, les réseaux de neurones récurrents (RNN), l'analyse en composantes principales (ACP) [6, 9, 8] s'avèrent souvent sous optimales. Les réseaux de neurones 3D permettant d'extraire simultanément des caractéristiques spectrales et spatiales semblent être les plus prometteurs pour l'analyse de données hyperspectrales [10]. Cependant, ces approches sont pénalisées par une complexité algorithmique souvent élevée qui a conduit certains auteurs à proposer des architectures spatio-spectrales séparables 2D+1D avec moins de paramètres [4]. Dans [11], Nalepa *et al.* proposent en particulier une architecture simplifiée de ce type de réseaux pour les images hyperspectrales. Récemment, une implémentation sur cible matérielle de cette approche a été menée afin d'être déployée prochainement sur un nano-satellite expérimental d'observation de la Terre [13].

Le travail présenté ici s'inscrit dans ce cadre de segmentation d'images hyperspectrales à bord des satellites d'observation de la Terre. Nous présentons des résultats obtenus avec un algorithme de programmation génétique unique, Zoetrope Genetic Programming (ZGP) [1], permettant de répondre aux nombreux points ouverts cités précédemment. En particulier, l'algorithme est frugal en données d'apprentissage et requiert peu de temps de calcul lors de l'entraînement. Aussi, il permet de générer des modèles simples et facilement interprétables comportant seulement quelques dizaines de paramètres et nécessitant peu de mémoire de stockage. Avec un temps d'inférence faible, les modèles ZGP sont aussi compatibles du traitement en temps réel des données issues des capteurs d'images satellites. Les modèles ZGP ont déjà été appliqués avec succès à la segmentation d'images satellites multisp spectrales [5] et ont démontré leur intérêt opérationnel lors d'une expérience menée sur le satellite OPS-SAT de l'Agence Spatiale Européenne en 2021 [2].

2 Méthodes

2.1 Algorithme évolutionnaire ZGP

ZGP est un algorithme de programmation génétique pour la régression symbolique dont l'originalité réside dans la manière de construire les individus de la population. En effet, ces individus sont des formules symboliques obtenues à partir d'un processus aléatoire intégrant diverses opérations de fusions. Les expressions mathématiques sont construites à partir des variables d'entrée qui caractérisent les données et de constantes sélectionnées aléatoirement. Des étapes de maturation consistant à appliquer plusieurs opérations de fusions donnent lieu aux "zoetropes" qui sont alors combinés et pondérés de façon linéaire pour générer un individu de la population. Les individus évoluent ensuite itérativement en fonction de leur adéquation aux données, grâce à des opérations génétiques (mutation et recombinaison). A chaque génération, l'ensemble des individus est évalué grâce à une fonction de fitness (qui dépend de la tâche effectuée, *i.e* régression ou classification) et le modèle final correspond au meilleur individu obtenu, toutes générations confondues. L'ensemble de ces processus est plus largement détaillé dans [1].

Dans ces travaux, nous appliquons ZGP à la segmentation des images hyperspectrales sur 3 jeux de données de référence. Plus précisément, ZGP est entraîné à faire de la classification multi-classes afin de prédire la classe de chaque pixel en fonction de ses informations spectrales.

2.2 Données

Les trois images hyperspectrales de référence, également utilisées dans [11] sont les suivantes :

1. Indian Pines : image acquise par le capteur aéroporté AVIRIS au Nord Ouest de l'Indiana (USA). La scène couvre une zone agricole de 145*145 pixels avec un pas au sol de 20m sur 224 bandes spectrales. La vérité terrain est composée de 16 classes de végétations dont certaines sont représentées par un très petit nombre de pixels, parfois moins de 100.
2. Pavia University : scène acquise par le capteur ROSIS au-dessus de l'Université de Pavie en Italie. L'image de 610*340 pixels à 1.3m de résolution comprend 103 bandes spectrales. La vérité terrain est labélisée en 9 classes constituées principalement de matériaux artificiels et de végétation.
3. Salinas : scène acquise par le capteur AVIRIS au-dessus de la vallée de Salinas en Californie (USA). Sa taille est de 512*217 pixels de 3.7m de résolution sur 224 bandes spectrales. La vérité terrain est composée de 16 classes de terrains agricoles.

Les classes fournies pour chacune de ces images étant différentes, les pixels d'apprentissage et de test doivent nécessairement être issus de la même image hyperspectrale. Ainsi, afin d'éviter la fuite d'information entre les données d'entraînement et de test qui conduirait à des résultats trop

optimistes, il est nécessaire de former des ensembles indépendants. Audebert *et al.* [10] et Nalepa *et al.* [11] ont respectivement proposé la méthode disjointe et la méthode par patches garantissant cette indépendance. Pour l'évaluation des modèles ZGP, nous utilisons une approche dans laquelle l'ensemble des pixels labélisés de chaque image est aléatoirement séparé en deux sous ensemble distincts : apprentissage (70%) et test (30%).

2.3 Entraînement de ZGP

ZGP est entraîné à la réalisation d'une tâche de classification multiple dont l'objectif est d'attribuer une classe à chacun des pixels de l'image hyperspectrale. Chaque pixel est traité indépendamment de ses voisins en considérant uniquement l'information spectrale associée localement au pixel courant. Les jeux de données utilisés pour l'entraînement sont constitués de pixels spectraux représentés par un vecteur dont le nombre de composantes est égal au nombre de bandes spectrales de l'image (103 ou 224). ZGP est entraîné sur 70% de tous les pixels labélisés des images de référence et testé sur les 30% des pixels restants. Bien que pris de manière aléatoire dans la même image, l'ensemble des pixels d'apprentissage est indépendant de celui des pixels de test, il n'y a donc pas de fuite possible entre les données d'apprentissage et de test lors de l'évaluation des modèles.

Stratégie d'enchaînement de modèles Dans le cas des images Indian Pines et Pavia University, une stratégie d'enchaînement de plusieurs modèles a été utilisée. Elle consiste à faire des prédictions avec un premier modèle et à les affiner en appliquant ensuite des modèles plus précis.

Pour Indian Pines, un premier modèle a été entraîné sur 10 classes de "haut niveau" regroupant certaines classes parmi les 16 originales, comme indiqué dans l'étape 1 de la Figure 1. Plus précisément, une classe "maïs" fusionne les classes originales "2 :maïs-notill", "3 :maïs-mintill" et "4 :maïs". La même chose a été faite pour l'herbe (classes originales 5, 6 et 7) et le soja (classes originales 10, 11 et 12). Ensuite, deux modèles intermédiaires permettent de distinguer des classes qui se sont révélées être difficiles à séparer par le modèle entraîné sur les 16 classes originales : un modèle binaire qui distingue les classes de haut niveau "maïs" et "soja" ainsi qu'un modèle qui distingue la classe de haut niveau "herbe" des classes originales 13, 14 et 15 (blé, bois, arbres) (resp. modèles 2 et 3 de l'étape 2 de la Figure 1). Enfin, trois modèles spécifiques des trois classes "maïs", "soja" et "herbe" permettent de séparer ces classes en leurs classes finales (modèles 4, 5 et 6 de l'étape 3 de la Figure 1). Finalement, 6 modèles au total sont utilisés afin de classifier les pixels dans l'une des 16 classes originales. Concernant le jeu de données de l'Université de Pavie, un premier modèle a été entraîné sur les neuf classes originales et un deuxième ensemble de trois modèles binaires différents a été utilisé pour déterminer les classes les plus difficiles à distinguer par le premier modèle : 1. classe 1 (asphalte) vs. classe 7 (bitume) ; 2. classe 3 (gravier) vs. classe 8 (briques) ; 3. classe 2 (pré) vs. classe 6 (sol nu).

Dans le cas de l'image Salinas, un modèle unique de multi-classification en 16 classes a directement été entraîné.

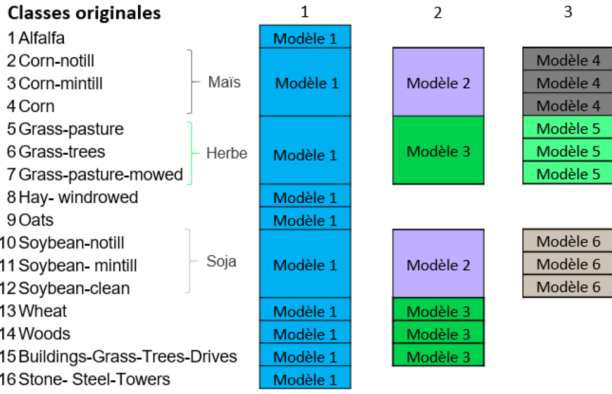


FIGURE 1 – Schéma de la stratégie d’enchaînement de modèles utilisée pour les données Indian Pines.

2.4 Métriques

Les performances sont évaluées de manière globale sur un ensemble de données par les métriques classiques suivantes : Précision, Rappel, F-score, Accuracy.

Plus précisément, nous calculons la moyenne (non pondérée) de la précision, du rappel et du F-score sur l’ensemble des classes que nous appelons « moyenne macro ». Pour l’accuracy, nous utilisons l’accuracy globale qui est celle obtenue sur l’ensemble des classes.

Nous comparons également la complexité des différents modèles en terme de nombre de paramètres. Pour les réseaux de neurones, il s’agit du nombre de paramètres entraîna- bles. Pour les modèles ZGP, cela correspond au nombre total de facteurs multiplicatifs dans les formules finales de chacun des modèles utilisés lors de la prédiction. La Figure 2 illustre le calcul du nombre de paramètres pour une formule ZGP obtenue dans l’un des modèles entraînés sur l’image Pavia University. Dans cet exemple, la complexité est de 43.

3 Résultats

Le Tableau 1 compare l’accuracy des modèles ZGP avec deux méthodes de deep learning hyperspectrales basées sur des réseaux de neurones convolutifs (1D et 3D CNN) [11]. Les résultats de deux approches d’évaluation sont présentés : "random" et "patch". L’approche "random" sélectionne les pixels de tests aléatoirement, tandis que l’approche "patch" garantit une indépendance des données d’entraî- nement et de test. Le réseau 3D CNN peut avoir une fuite d’information entre les données d’apprentissage et de test, car les couches de convolution prennent en considération les pixels voisins du pixel d’intérêt. Cette fuite d’informa- tion n’est pas valable pour le réseau 1D CNN, qui pos- sède une couche de convolution spectrale et non spatiale. ZGP utilise l’évaluation "random" et comme il ne prend pas en compte l’information spatiale, l’indépendance totale des pixels d’apprentissage et de test est garantie.

Les trois algorithmes sont moins performants sur Indian Pines en raison de la plus faible résolution de l’image, ce qui entraîne des pixels contenant des mélanges de plu-

$$\begin{aligned} \text{score} = & \exp(\mathbf{0.0069} \times B_{022} + \mathbf{0.0005} \times B_{042} - \mathbf{0.0031} \times \\ & B_{089} + \mathbf{0.0005} \times B_{103} \times \text{truncate}(B_{042}/(\mathbf{0.9271} \times \\ & (-B_{103} + B_{019}) + \mathbf{0.0729} \times \text{sqrt}(\text{abs}(B_{019})))) - \\ & \mathbf{0.0008} \times \cos(\mathbf{0.6596} \times \text{sqrt}(\text{abs}(B_{089})) + \mathbf{0.3404} \times \\ & \sin(B_{089})) \times B_{022} - \mathbf{0.0021} \times \text{sqrt}(\text{abs}(\mathbf{0.6596} \times \\ & \text{sqrt}(\text{abs}(B_{089})) + \mathbf{0.3403} \times \sin(B_{089}))) \times B_{022} - \\ & \mathbf{0.0004} \times \text{truncate}(B_{042}/(\mathbf{0.9270} \times (-B_{103} + B_{019}) + \\ & \mathbf{0.0729} \times \text{sqrt}(\text{abs}(B_{019})))) \times B_{019} - \mathbf{3.755e-05} \times \\ & \text{truncate}(B_{042}/(\mathbf{0.9270} \times (-B_{103} + B_{019}) + \mathbf{0.0729} \times \\ & \text{sqrt}(\text{abs}(B_{019})))) \times \text{sqrt}(\text{abs}(B_{019})) + \mathbf{3.150e-05} \times \\ & \text{sqrt}(\text{abs}(B_{089})) + \mathbf{1.060e-05} \times \text{sqrt}(\text{abs}(\mathbf{0.6596} \times \\ & \text{sqrt}(\text{abs}(B_{089})) + \mathbf{0.3403} \times \sin(B_{089}))) - \mathbf{0.0007} \times \\ & \text{sqrt}(\text{abs}(\mathbf{0.7159} \times \text{sqrt}(\text{abs}(\mathbf{0.6596} \times \text{sqrt}(\text{abs}(B_{089})) + \\ & \mathbf{0.3403} \times \sin(B_{089}))) + \mathbf{0.2840} \times \cos(\mathbf{0.6596} \times \\ & \text{sqrt}(\text{abs}(B_{089})) + \mathbf{0.3403} \times \sin(B_{089}))) + \mathbf{0.0064}(B_{042}) + \\ & \mathbf{0.0002}(B_{103}) + \mathbf{0.0048} \times \cos(B_{022}) + \mathbf{4.828e-05} \times \\ & \cos(B_{089}) + \mathbf{0.0003} \times \cos(\mathbf{0.9270} \times (-B_{103} + \\ & B_{019}) + \mathbf{0.0729} \times \text{sqrt}(\text{abs}(B_{019}))) + \mathbf{4.205e-06} \times \\ & \cos(\mathbf{0.6596} \times \text{sqrt}(\text{abs}(B_{089})) + \mathbf{0.3403} \times \sin(B_{089})) + \\ & \mathbf{0.0073} \times \text{abs}(\mathbf{0.9270} \times (-B_{103} + B_{019}) + \mathbf{0.0729} \times \\ & \text{sqrt}(\text{abs}(B_{019}))) + \mathbf{7.213e-05} \times \text{truncate}(B_{103})) \end{aligned}$$

FIGURE 2 – Exemple de formule ZGP avec en gras les constantes servant à calculer le nombre de paramètres. Les valeurs B_{xxx} correspondent aux bandes hyperspectrales. Ici on compte 43 constantes.

sieurs types de cultures et pouvant appartenir à plusieurs classes [10]. La forte ressemblance entre certaines classes et le petit nombre de pixels qui les représentent peuvent également rendre l’apprentissage plus difficile.

La Table 2 montre les performances de ZGP sur les trois images, avec une précision moyenne de 89% et 93% pour les images Pavia University et Salinas respectivement, mais seulement environ 60% pour l’image Indian Pines. Diffé- rentes hypothèses ont été évoquées ci-dessus pour expliquer ce dernier résultat.

Les Figures 3, 4 et 5 illustrent la segmentation faite par les modèles ZGP sur les trois images, montrant la qualité des prédictions réalisées par les modèles.

La Table 3 montre que les modèles ZGP sont beaucoup moins complexes que les réseaux de neurones 1D et 3D. En particulier, la complexité des modèles ZGP est indé- pendante de la dimension de l’image contrairement à celle des réseaux de neurones du fait des couches de sortie "fully connected".

Enfin, la Figure 6 montre l’intérêt des modèles ZGP qui atteignent des performances comparables voire meilleures que les réseaux de neurones pour une complexité significa- tivement inférieure, représentant un gain de stockage et de capacité de calcul pour l’implémentation dans des systèmes embarqués. La taille mémoire des modèles ZGP est de res- pectivement 393ko, 129ko et 135ko pour les images Indian Pines, Pavia University et Salinas.

TABLE 1 – Comparaison des performances (accuracy globale) obtenues par Nalepa *et al.* [11] avec les performances des modèles ZGP

Dataset	1D CNN		3D CNN		ZGP
	Random	Patch	Random	Patch	Random
Indian Pines	77.98	67.11	85.9	48.89	73.28
Pavia University	89.32	73.26	93.47	70.07	90.12
Salinas	83.57	64.2	94.27	69.72	90.21

TABLE 2 – Performances des modèles ZGP sur les 3 jeux de données de référence

	Indian Pines	Pavia U.	Salinas
Accuracy globale	73.28	90.12	90.21
Macro précision	62.37	89.19	93.86
Macro recall	60.28	85.98	93.24
Macro F-score	60.89	88.02	93.70

4 Conclusion

Ce texte décrit les performances de l'algorithme ZGP pour la segmentation hyperspectrale sur trois jeux de données de référence, comparant ses résultats à deux réseaux de neurones convolutifs évalués par deux méthodes différentes. Les résultats montrent que ZGP peut fournir des performances comparables, voire supérieures, aux CNN en utilisant uniquement les informations spectrales. ZGP sélectionne également des bandes spectrales spécifiques pour la prédiction, permettant un certain niveau de confiance et d'interprétabilité du résultat. Il possède plusieurs avantages adaptés aux contraintes de l'analyse d'images hyperspectrales et des systèmes embarqués : frugalité, facilité d'intégration et d'implémentation et méthode d'évaluation plus facile à mettre en oeuvre. Des travaux futurs visent à interpréter les modèles fournis en examinant les bandes spectrales utilisées pour la prédiction et à adapter l'algorithme pour une portabilité efficace et rapide sur FPGA. Nous participons également au challenge OrbitalAI organisé par l'Agence Spatiale Européenne.

Remerciements

Ces travaux ont été menés dans le cadre du projet CIAR ("Chaîne Image Autonome et Réactive") de l'Institut de Recherche Technologique Saint-Exupéry (www.irt-saintexupery.com). Les auteurs remercient les partenaires industriels et académiques du projet : Thales Alenia Space, ActiveEon, Avisto, Elsys Design, MyDataModels, GEO4i, Inria et LEAT/CNRS.

TABLE 3 – Nombre de paramètres des différents modèles

	Indian Pines	Pavia University	Salinas
ZGP	1784	511	496
1D CNN	10 105 040	5 086 537	11 333 840
3D CNN	333 216	250 009	352 416

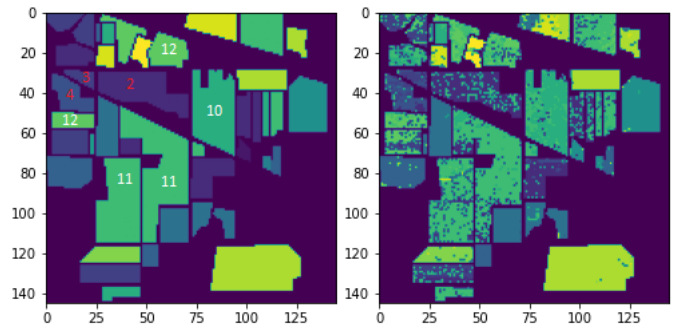


FIGURE 3 – Vérité terrain (gauche) et prédictions faites par ZGP (droite) de l'image Indian Pines. Les différentes classes de maïs (2,3,4) et de soja (10,11,12) sont repérées sur la vérité terrain.

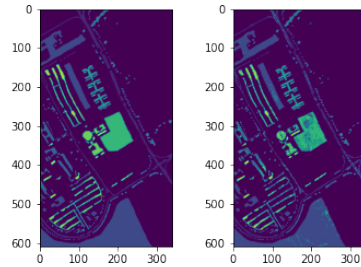


FIGURE 4 – Vérité terrain (gauche) et prédictions faites par ZGP (droite) de l'image Pavia University.

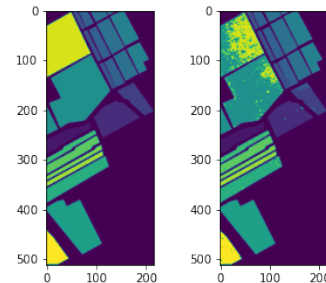


FIGURE 5 – Vérité terrain (gauche) et prédictions faites par ZGP (droite) de l'image Salinas.

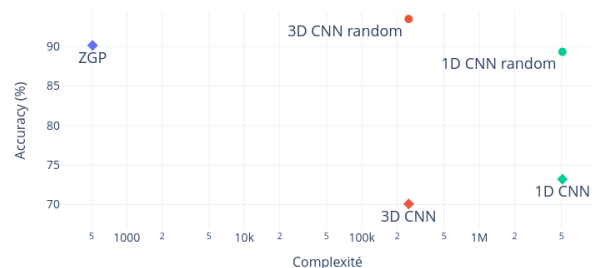


FIGURE 6 – Accuracy des modèles en fonction du nombre de paramètres (échelle logarithmique) des différents algorithmes pour l'image Pavia University.

Références

- [1] A. Boisbunon, C. Fanara, I. Grenet, J. Daeden, A. Vighi, and M. Schoenauer. Zoetrope genetic programming for regression. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pages 776–784, 2021.
- [2] F. Férésin, E. Kervennic, Y. Bobichon, E. Lemaire, N. Abderrahmane, G. Bahl, I. Grenet, M. Moretti, and M. Benguigui. In space image processing using AI embedded on system on module : example of OPS-SAT cloud segmentation. In *2nd European Workshop on On-Board Data Processing (OBDP)*, 2021.
- [3] Q. Fu, X. Yu, X. Wei, and Z. Xue. Semi-supervised classification of hyperspectral imagery based on stacked autoencoders. page 100332B, 08 2016.
- [4] Q. Gao, S. Lim, and X. Jia. Hyperspectral image classification using convolutional neural networks and multiple feature learning. *Remote Sensing*, 10(2) :299–7056, 2018.
- [5] I. Grenet, Y. Bobichon, A. Girard, and F. Férésin. Zgp : une alternative aux réseaux de neurones pour la segmentation sémantique de nuages dans les images satellites multi-spectrales. In *Conférence Nationale sur les Applications Pratiques de l'Intelligence Artificielle*, 2021.
- [6] W. Hu, Y. Huang, L. Wei, F. Zhang, and H. Li. Deep convolutional neural networks for hyperspectral image classification. *Journal of Sensors*, 2015 :1–12, 07 2015.
- [7] D. Lebedeff, M.F Foulon, R. Camarero, R. Vitulli, and Y. Bobichon. On-board cloud detection and selective spatial/spectral compression based on CCSDS 123.0-b-2 for hyperspectral missions. In *7th ESA International Workshop on On-Board Payload Data Compression Conference (OBPDC)*, 2020.
- [8] K. Makantasis, K. Karantzalos, A. Doulamis, and N. Doulamis. Deep supervised learning for hyperspectral data classification through convolutional neural networks. pages 4959–4962, 01 2015.
- [9] L. Mou, P. Ghamisi, and X. Zhu. Deep recurrent neural networks for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 55(7) :3639–3655, 2017.
- [10] Audebert N., B. Le Saux, and S. Lefèvre. Deep learning for classification of hyperspectral data : A comparative review. *IEEE Geoscience and Remote Sensing Magazine*, 7(2) :159–173, June 2019.
- [11] J. Nalepa, M. Myller, and M. Kawulok. Validating hyperspectral image segmentation. *IEEE Geoscience and Remote Sensing Letters*, 16(8) :1264–1268, Aug 2019.
- [12] R. Vitulli, M. Celesti, R. Camarero, G. Di Cosimo, F. Gascon, N. Longepe, J. Nieke, M. Rovatti, M.F Foulon, A. Grynagier, D. Lebedeff, Y. Bobichon, J. Nalepa, A. Wijata, M. Przeliorz, and T. Dullin. CHIME : The first AI-powered ESA operational mission. In *Small Satellite Systems and Services 4S symposium*, 2022.
- [13] M. Ziaja, P. Bosowski, M. Myller, G. Gajoch, M. Gumiela, J. Protich, K. Borda, D. Jayaraman, R. Dividino, and J. Nalepa. Benchmarking deep learning for on-board space applications. *Remote Sensing*, 13 :3981, Oct 2021.