



Classification automatique de nuages de points LiDAR par *Deep Learning*

Les défis d'une production industrielle

■ Baptiste BENARD - Jules ROGER - Jérôme LEROUX - Camille LHENRY - Maxime CHAUVIN

Ces dernières années, de nombreuses campagnes d'acquisition de nuages de points par LiDAR aérien ont été lancées avec pour objectif le relevé topographique précis et détaillé des territoires. Ces données sont de plus en plus riches en détails et couvrent des superficies importantes telles que la France dans le cadre du programme LiDAR HD.

*La classification automatique d'un nuage de points est un enjeu majeur dans la chaîne de production de données géospatiales. Leur traitement, complexe, nécessite une méthode industrielle afin de permettre la classification de données en grandes quantités tout en s'adaptant à des cahiers des charges variés. Cet article présente les travaux réalisés dans le cadre de la classification automatique de nuages de points LiDAR aériens, en particulier dans ce contexte de production industrialisée. Nous vous présentons notre retour d'expérience sur les défis d'industrialisation et d'exploitation de modèles *Deep Learning*, proposant des alternatives plus robustes et précises que les méthodes dites paramétriques.*

■ MOTS-CLÉS

Deep Learning, LiDAR, nuage de points, classification, segmentation sémantique, Intelligence Artificielle, industrialisation

Introduction

Dans le cadre de l'étude des territoires, la connaissance précise de la structure et la composition des espaces urbains et ruraux est un atout majeur dans de nombreux domaines (p. ex. aménagement du territoire, étude des littoraux, suivi de l'évolution forestière). De nombreuses institutions ont donc engagé des programmes nationaux d'acquisition LiDAR (*Light Detection and Ranging*) comme l'Institut national de l'information géographique et forestière (IGN) en France avec le programme LiDAR HD (>10 pts/m²) ou bien Swisstopo en Suisse avec swiss-SURFACE3D (~15-20 pts/m²). Lors de tels programmes, des entreprises comme GEOFIT et SINTEGRA réalisent l'acquisition et la classification de nuages de points aériens. La classification de nuages de points correspond dans ce cas à la tâche de segmentation sémantique : l'attribution à chaque point du nuage d'une classe (p. ex. bâtiment, végétation, sol). Cette classification est aujourd'hui un défi majeur de produc-

tion, car elle est extrêmement coûteuse en ressources humaines et matérielles du fait de la très grande quantité de données acquises. Aujourd'hui, des méthodes de préclassification automatique permettent de réduire le coût et le temps de classification en diminuant l'intervention manuelle d'annotation, de correction et de contrôle qualité du produit final, dans certains cas d'application, elle peut même atteindre un niveau de précision suffisant.

Bien que la classification de nuages ne soit pas un sujet récent, de forts progrès ont été faits dans le domaine de l'Intelligence Artificielle (IA), et plus précisément dans les domaines du *Machine Learning* (ML) et du *Deep Learning* (DL). L'évolution croissante de la puissance de calcul sur GPU (*Graphics Processing Unit*), les avancées de la recherche académique dans le domaine du *Deep Learning*, ainsi que l'accès à de fortes quantités de données d'apprentissage, permettent aujourd'hui d'entraîner des modèles robustes et de les utiliser pour effectuer des prédictions

automatiques. Cependant des difficultés peuvent apparaître lors de la mise en place de ces méthodes d'IA lorsqu'il s'agit, par exemple, de classer automatiquement un département, une région ou un pays entier. Pour répondre à cet enjeu, GEOFIT a constitué une équipe spécialisée pour développer, évaluer et intégrer ces nouvelles méthodologies au sein des autres services du groupe. Cet article est un premier retour d'expérience sur l'analyse de l'impact du *Deep Learning* sur les méthodes actuelles de classification des nuages de points, à travers : – la présentation de deux architectures *Deep Learning* (KP-FCNN et RandLA-Net) – l'explication des enjeux d'annotations de base de données – les contraintes de production sur des grands volumes de données – la valorisation de la classification avec la production de données vectorielles géospatiales.

Classifier un nuage de points

Du fait de la faible densité d'acquisition LiDAR (~1 pt/m²) et des ressources de traitement limitées, les premiers travaux de classification portaient sur la segmentation des classes sol et sursol afin de pouvoir réaliser des modèles numériques de terrain (MNT) (*figure 1*). Les méthodes paramétriques souvent utilisées dans des suites logicielles comme TerraScan® sont très efficaces pour réaliser cette classification. Ces méthodes reposent sur l'application de filtres successifs aux caractéristiques des nuages de points comme la planarité ou l'intensité, où chaque filtre est paramétrable. Ce principe repose sur une hypothèse de faible variabilité intraclasse et le paramétrage peut être sensible aux données d'entrée, mais très efficace dans la rapidité de mise en place. À partir d'une classification du sol



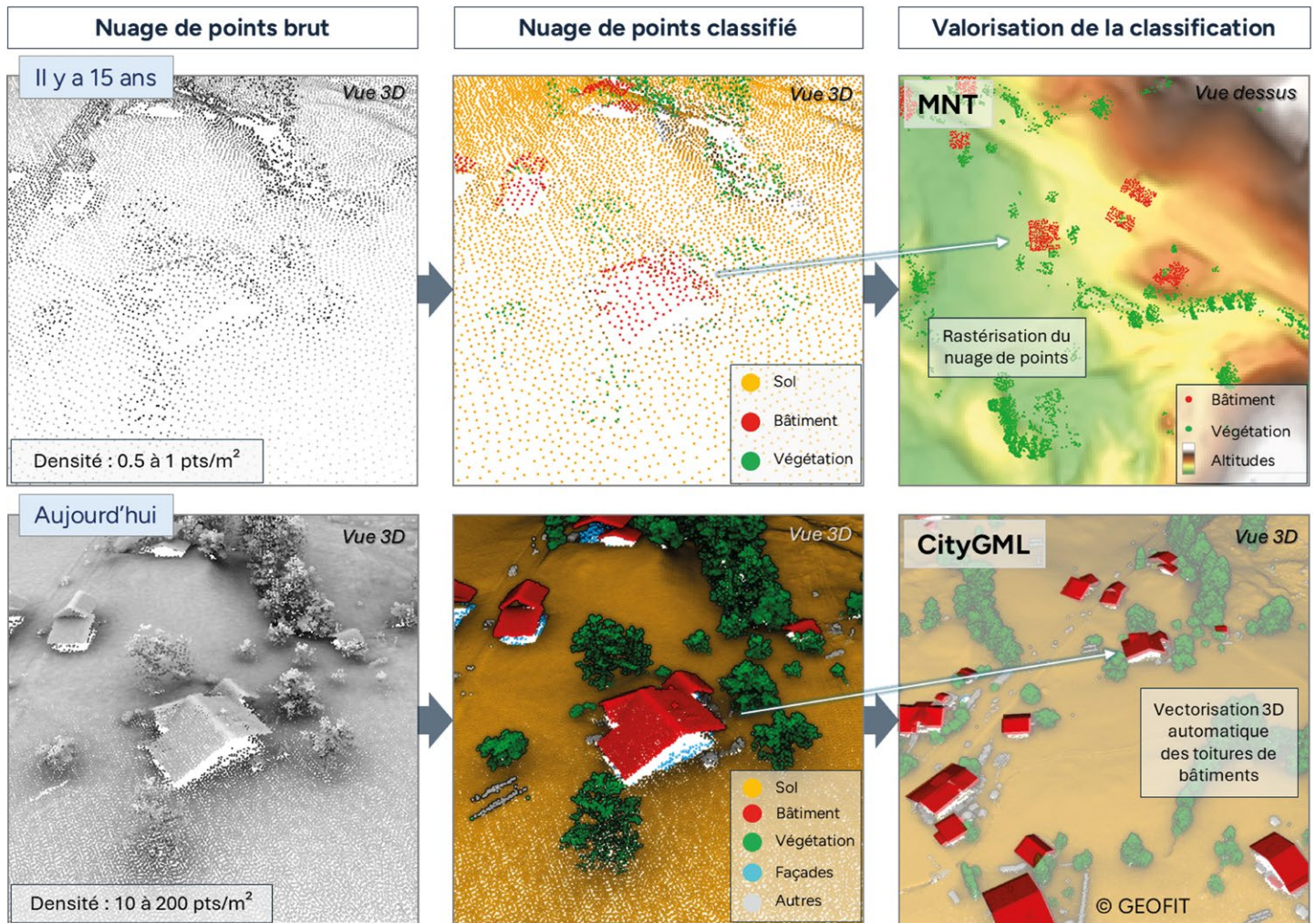


Figure 1. Comparaison des classifications et des valorisations sur un nuage de points acquis il y a 15 ans (production d'un MNT) et aujourd'hui (vectorisation 3D des toitures au format CityGML).

basée sur des critères géométriques, des classes comme les bâtiments et la végétation auront des signatures distinctes : les bâtiments sont constitués de plans en hauteur, tandis que les points de végétation sont disposés de manière plus dispersée (figure 1). Ces différentes signatures sont donc filtrables par paramétrage (Brovelli, M. and Lucca, S., 2011).

Aujourd'hui, les densités d'acquisition vont de 10 pts/m² pour le LiDAR HD à plus de 200 pts/m² pour certaines campagnes d'acquisition comme le suivi de lignes électriques. Cette richesse de données permet d'acquérir les objets dans toutes leurs complexités 3D : les façades et les cheminées sont maintenant parfaitement distinguables sur les toits des bâtiments (figure 1). Cette densité permet d'envisager de produire des données vectorielles 3D, au format CityGML par exemple, à

de hauts niveaux de détails (figure 1). Pour aborder la segmentation d'objets, des méthodes ont été développées pour distinguer les signatures géométriques de ces objets par une analyse multiéchelle (Brodu N. et Lague D. (2012)). Ces méthodes parviennent de manière semi-automatique à définir statistiquement les seuils pour distinguer, par exemple, un sol nu d'une zone de rochers. Le sol nu présente un aspect plan à toutes les échelles (de 50 cm à des dizaines de mètres), tandis qu'une zone de rochers apparaît plane à 50 cm, mais rugueuse à des échelles plus grandes. Ces méthodes sont cependant limitées dans le cadre de segmentations de scènes contenant des objets appartenant à des classes distinctes, mais avec des caractéristiques géométriques similaires.

Le développement des réseaux de neurones artificiels a permis l'essor de

nouveaux algorithmes plus robustes et plus précis. De nombreux travaux académiques se sont donc intéressés à des méthodes basées sur du *Deep Learning* pour la classification de nuages de points. Les méthodes basées sur le *Deep Learning* supervisé nécessitent des données segmentées qui vont servir de base d'apprentissage pour le modèle.

Différents types d'architectures existent qui, en plus d'utiliser des caractéristiques du signal (p. ex. intensité, nombre de retours), peuvent utiliser différentes représentations de la donnée. On a par exemple : les architectures basées sur les projections, celles qui vont voxeliser l'espace et celles qui vont traiter directement les nuages de points (figure 2).

En majorité, ces méthodes vont chercher à simplifier le problème afin de se rapprocher de traitements plus

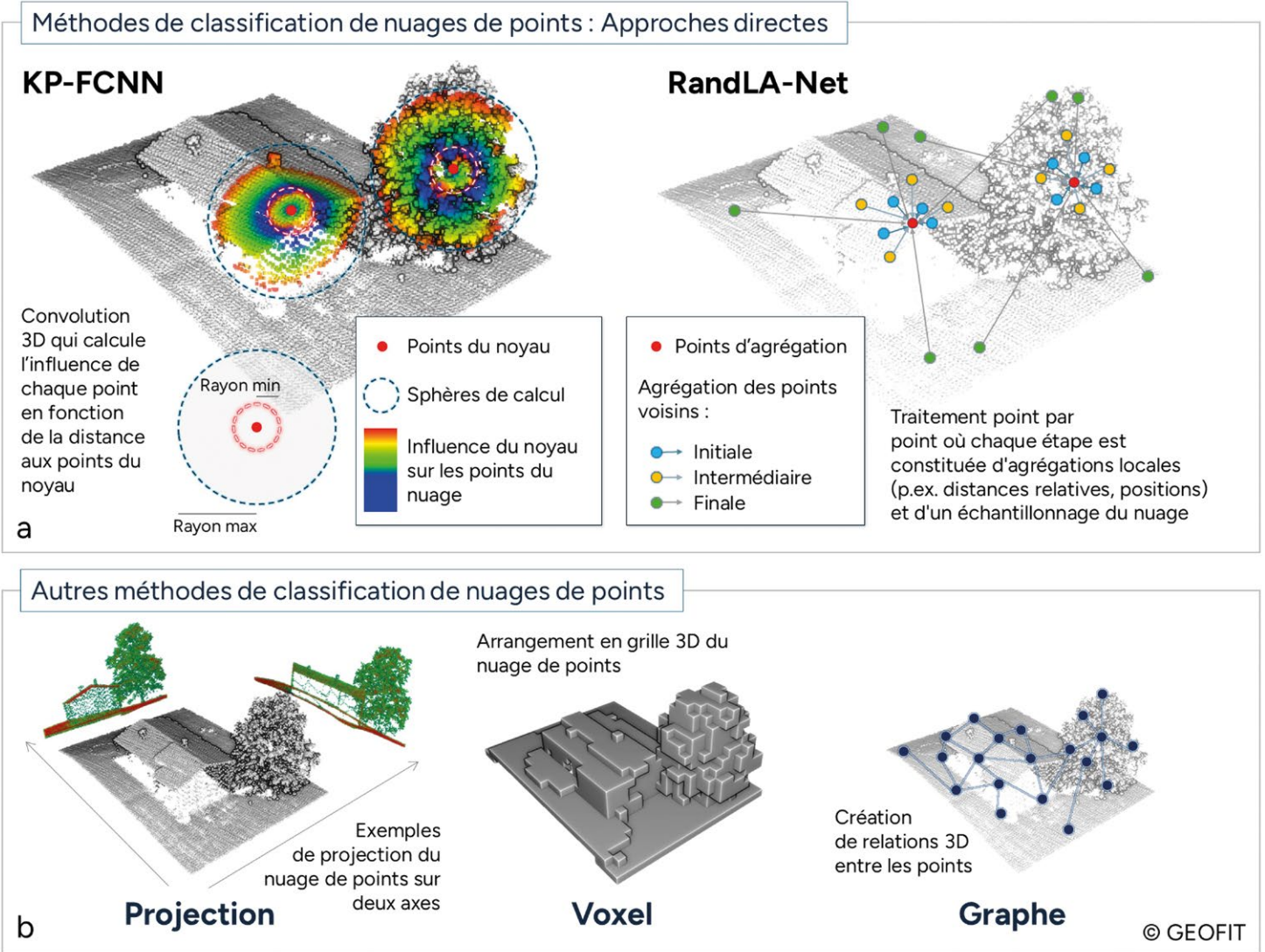


Figure 2. Méthodes de classification de nuages de points par *Deep Learning* : a. Méthodes directes : KP-FCNN et RandLA-Net. b. Autres méthodes de classification : Projection, Voxel et Graphe.

simples, similaires à ceux appliqués sur les images. Pour les projections, il s'agit de considérer plusieurs vues du nuage de points (Su, H. et al. (2015), Feng, Y. et al. (2018)), tandis que pour la voxelisation, l'objectif est de structurer la donnée selon une grille régulière 3D pour la traiter (Sanchez-Caballero, A. et al., 2022). Pour le traitement sous forme de graphe, l'information géométrique va être transformée en relation pour simplifier les traitements (Kipf, T. et Welling M. (2016)).

D'autres méthodes traitent directement les nuages de points. Les méthodes de cette famille peuvent définir des opérations comme les convolutions : calculs qui permettent de prendre en compte le voisinage d'une donnée. Elles peuvent également modifier les prétraitements

afin de pouvoir être utilisées sur des méthodes plus classiques tout en utilisant l'information de la disposition des points dans l'espace. Dans cette catégorie, on retrouve notamment les architectures avec des convolutions adaptées au calcul 3D comme les KPConv (Thomas, H. et al., 2019), mais aussi des méthodes qui vont agréger les relations géométriques locales à l'aide de MLP (*MultiLayer Perceptron*) pour des calculs moins coûteux tels que RandLA-Net (Hu, Q. et al., 2019) (figure 2a).

Classifier par Deep Learning

Les architectures de *Deep Learning* facilitent l'analyse de vastes ensembles de données, améliorant ainsi la généralisation dans la prédiction des objets

et répondant efficacement aux variations des méthodes d'acquisition, des paysages et des densités. Certaines architectures *Deep Learning*, bien que plus performantes, requièrent d'importantes ressources de calcul et, en pratique, une quantité excessive de données d'entraînement pour être mises en œuvre. Pour répondre à nos problématiques de production, nous avons choisi de baser nos travaux sur deux architectures adaptées : KP-FCNN et RandLA-Net.

■ Caractéristiques de KP-FCNN et de RandLA-Net

Les KP-FCNNs (*Kernel Point Fully Convolutional Neural Networks*) sont des réseaux de neurones convolutifs dont les convolutions sont des KPConv



► (*Kernel-Point Convolutions*). Ce sont des convolutions qui reposent sur des noyaux de points (kernel points) utilisant les distances entre les points du nuage et ceux du noyau pour faire des convolutions directement sur les points en 3D (figure 2a).

L'approche de RandLA-Net se base sur plusieurs éléments. Les informations géométriques sont agrégées localement en utilisant l'information des voisins les plus proches et traitées avec des MLP et des blocs d'attention. Ces derniers vont permettre de densifier l'information en ne conservant que les données pertinentes afin de simplifier les calculs (figure 2a).

Voici quelques caractéristiques importantes de ces méthodes :

Géométrie : Une KPConv est basée sur des noyaux de points pour capturer les caractéristiques locales dans l'espace 3D, tandis que RandLA-Net utilise un échantillonnage aléatoire suivi d'une agrégation locale pour simplifier et traiter les nuages de points.

Voisinage : Les KP-FCNNs fonctionnent avec des voisinages de rayon fixe, ce qui permet des représentations plus régulières des objets, mais limitera la distance de perception alors que RandLA-Net utilise un voisinage avec un nombre fixe de voisins qui permet des calculs plus simples et mieux opti-

misés, mais qui impactera l'échelle de perception en fonction de la densité du voisinage.

Bien que différentes, ces architectures capturent les informations locales et les agrègent, leur permettant d'obtenir de bons scores sur les tâches de segmentation sémantique. Ces architectures diffèrent dans leurs utilisations, voici notre retour d'expérience :

Efficacité : RandLA-Net demande moins de ressources de calculs et de mémoire, ce qui le rend plus adapté aux très grands nuages de points.

Quantité de données : Les KP-FCNNs nécessitent moins de données d'entraînement que RandLA-Net et apprennent plus rapidement ce qui les rend plus performants pour des jeux de données réduits.

Complexité vs Précision : Un KP-FCNN peut offrir une meilleure sensibilité pour capturer des détails géométriques subtils, mais avec une plus grande complexité de calcul. RandLA-Net équilibre la précision et l'efficacité, offrant une solution plus rapide et moins gourmande en ressources.

Bien que les caractéristiques des architectures impactent les performances, un jeu de données labellisées de qualité est indispensable pour effectuer un bon apprentissage supervisé.

■ Apprentissage supervisé

Lors d'un apprentissage supervisé, le modèle va apprendre à reproduire les classifications des nuages de points qui lui seront montrées successivement au cours de l'entraînement. À chaque exemple, le réseau de neurones prédit une classe pour chacun des points. S'ensuit une phase d'évaluation de l'erreur en comparant la prédiction avec la vérité terrain (figure 3). En utilisant cette évaluation de l'erreur, il est possible de modifier les paramètres internes du modèle afin de réduire cette erreur de prédiction sur l'échantillon. C'est un apprentissage itératif reposant sur ces nombreuses modifications permettant d'améliorer les performances.

■ Jeu de données

Un grand jeu de données améliore la robustesse du modèle en exposant celui-ci à une grande variété d'exemples, de paysages, de densités ou de types de capteurs. Cette diversité contribue à une meilleure généralisation des classes à détecter et renforce la fiabilité du modèle. Il faut néanmoins préserver une forme d'équilibre entre les classes, il est souvent préférable d'utiliser moins de nuages, mais contenant des environnements plus diversifiés : villes, villages, ports, gares, etc., plutôt que de grandes quantités

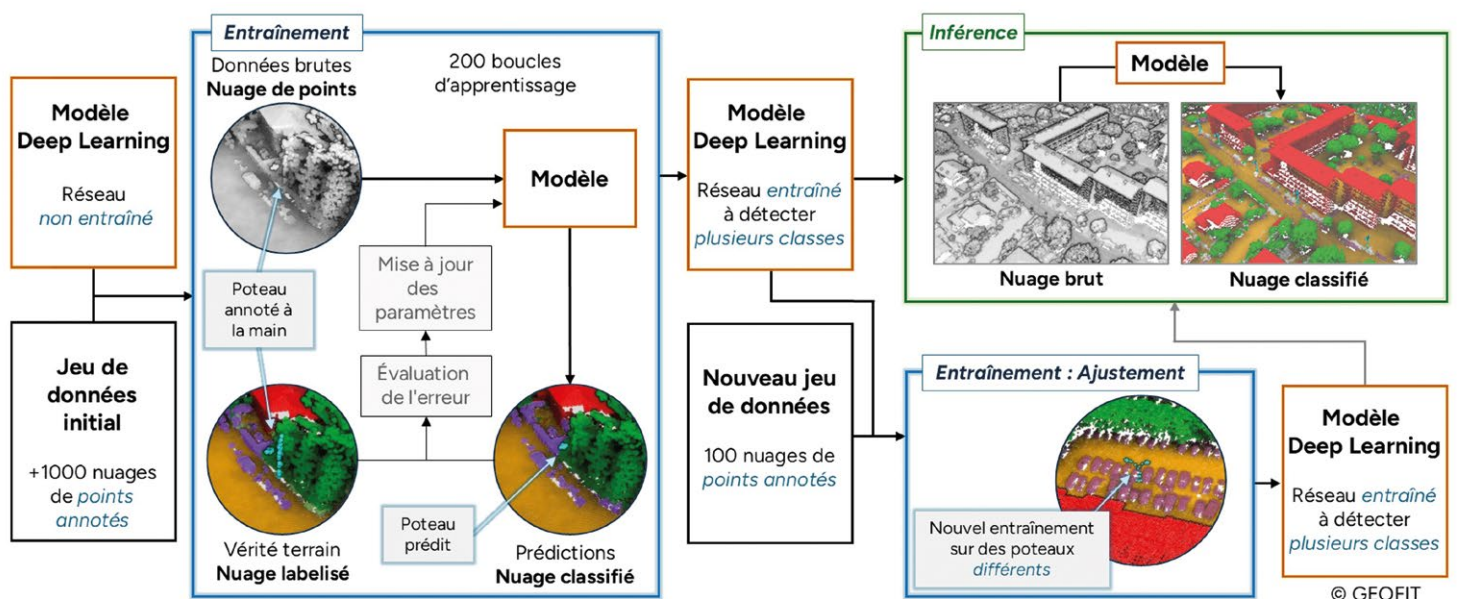


Figure 3. Chaîne de traitement simplifiée de l'apprentissage d'un modèle *Deep Learning*.

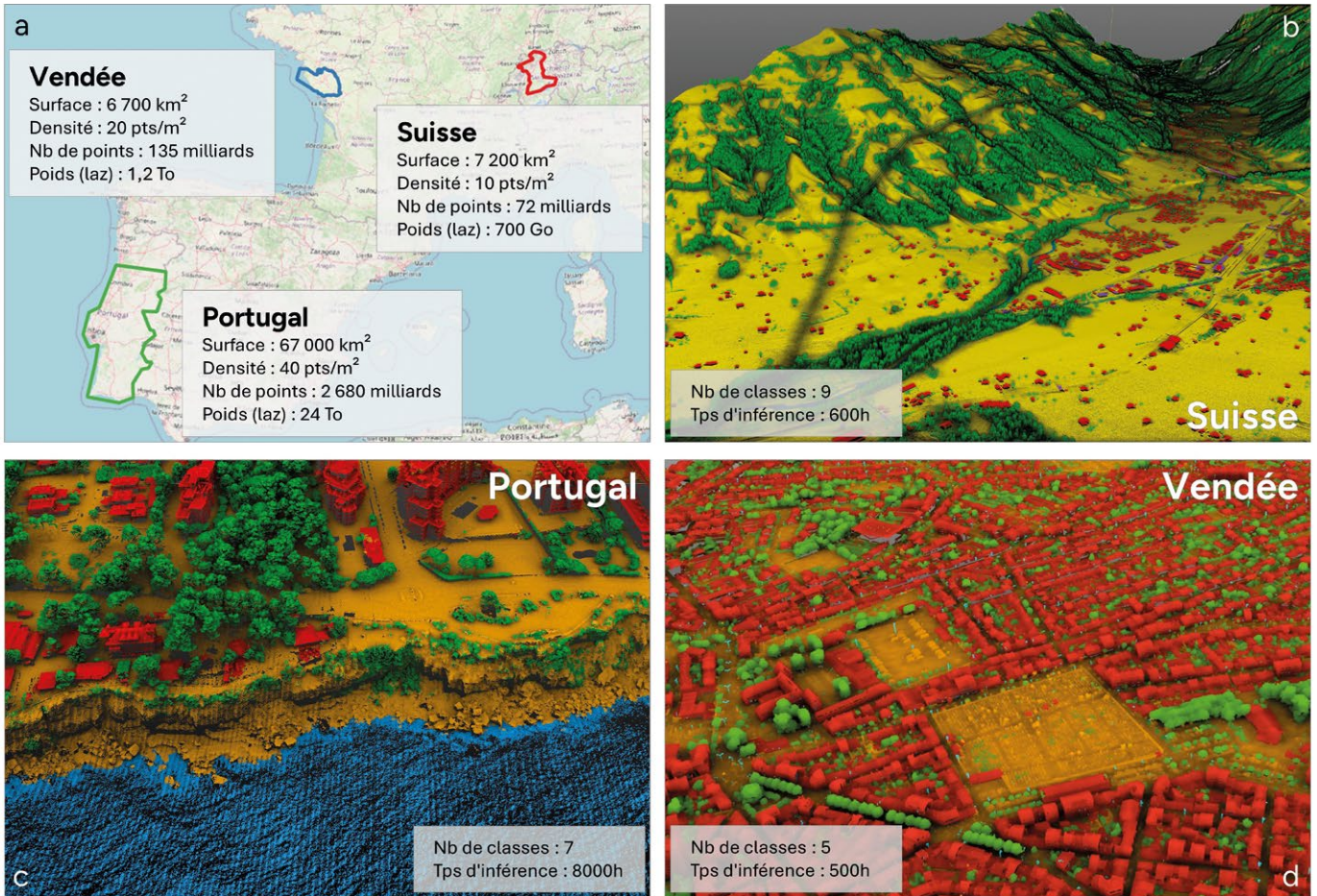


Figure 4. a. Surfaces de données acquises et classifiées sur les zones de la Suisse, du Portugal et de la Vendée. b. Illustration de classification en zone de montagne, c. en zone côtière et d. en zone urbaine.

de nuages contenant peu de classes : forêts, plaines, etc. La qualité de l'annotation des nuages de points est essentielle pour assurer de hautes performances aux modèles. Un jeu de données mal annoté induit des biais d'apprentissage et le modèle risque de répéter ces erreurs.

■ Méthodologies d'apprentissage

Entraînement

Lors d'un entraînement supervisé, le jeu de données est réparti en trois ensembles distincts qui interviennent dans des phases différentes de l'entraînement. Le jeu d'entraînement (~70 % des données) sert de base d'apprentissage pour le réseau de neurones. À chaque itération, le modèle apprend sur des exemples du jeu de données et modifie ses paramètres internes. Pour chaque itération, le modèle va calculer ses performances sur le jeu de données de validation (~15 % des données). L'erreur entre la prédiction et la vérité

terrain est calculée sans modifier les paramètres du réseau, permettant de suivre l'évolution des performances au cours de l'apprentissage.

Évaluation

Le modèle final est évalué sur le jeu de données de test (~15 % des données) à la fin de l'entraînement. Cette évaluation est cruciale pour comparer les performances lors de sessions d'entraînement successives. Lors de l'évaluation, le modèle prédit les classes sur l'ensemble de test. Pour chaque prédiction, l'évaluation des performances se base sur de nombreuses métriques comme les matrices de confusion, la précision, le rappel ou l'intersection sur l'union. Ces différentes métriques permettent un suivi et des comparatifs des différents entraînements ainsi qu'une analyse fine des erreurs.

Ajustement

La principale caractéristique des réseaux de neurones est l'adaptabilité

à de nouvelles données. Ainsi, une stratégie assez courante est l'entraînement par ajustement, ou *fine-tuning* en anglais. Cette stratégie repose sur le postulat qu'un modèle entraîné sur beaucoup de données va réussir à avoir une compréhension suffisante des objets pour servir de base pour un nouvel entraînement sur des données différentes. Ainsi, pour de nouvelles classes, de nouveaux paysages ou de nouveaux capteurs, un nouvel entraînement avec une quantité de données réduite est possible. Cette stratégie est utile car très économique : relativement peu de données annotées sont nécessaires et le temps d'apprentissage pour spécialiser un réseau à un nouveau jeu de données est réduit. Ainsi le réseau conserve les composantes élémentaires apprises lors du premier apprentissage et s'ajuste aux éventuelles nouvelles composantes des données, ce qui rend la méthode flexible et adaptable.



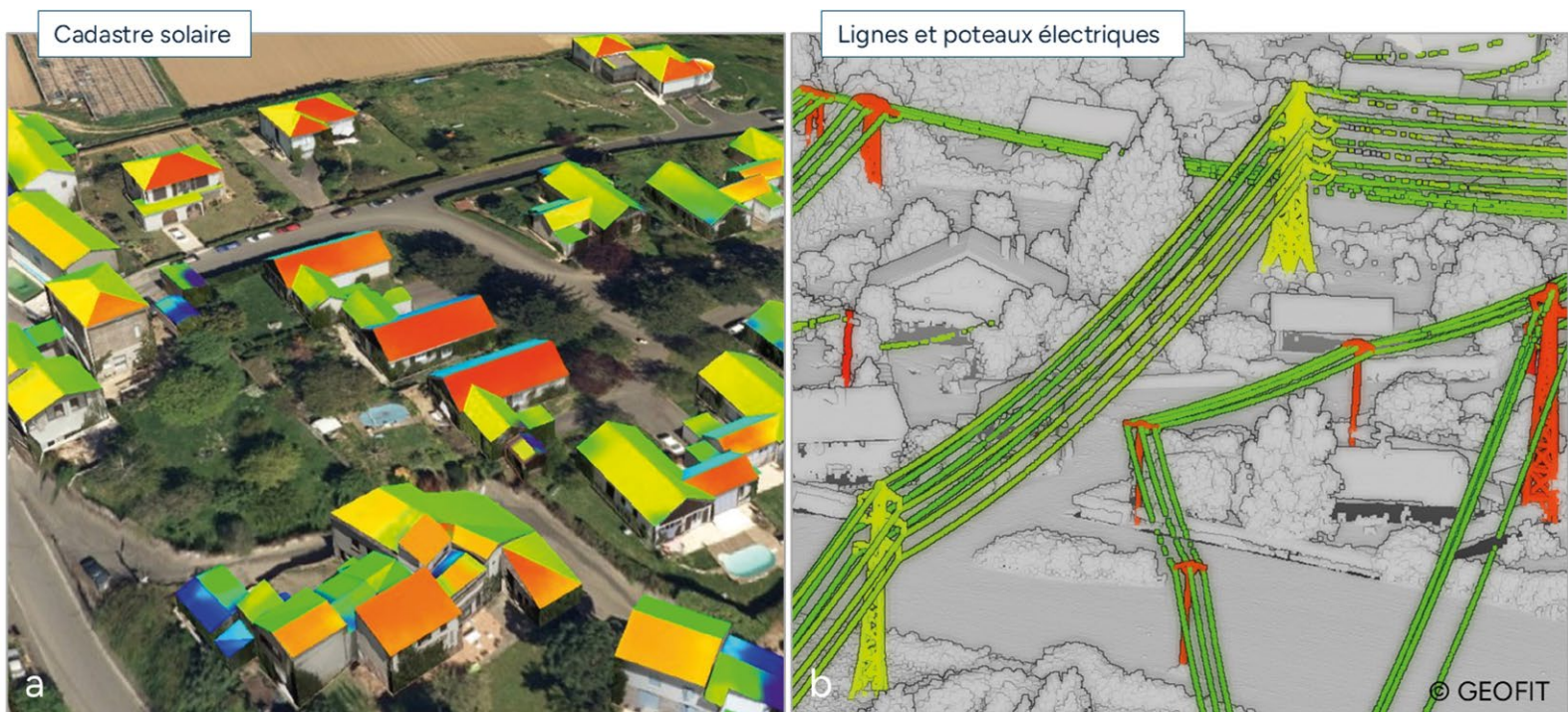


Figure 5. Exemples de valorisations de nuages de points classifiés. a. Cadastre solaire. b. Lignes et poteaux électriques.



■ Mise en production industrielle Inférence

La phase d'inférence est la dernière phase du processus, elle consiste à stopper la modification des paramètres du modèle et à effectuer des prédictions sur de nouvelles données. Cette phase peut être appliquée à de vastes zones, telles que des départements, régions ou pays, dans le cadre de la segmentation de nuages de points. Elle exige souvent des optimisations du code afin d'améliorer la qualité des résultats et l'efficacité des calculs.

Classification de large quantité de données

L'exploitation d'un modèle *Deep Learning* disponible en *open-source* est accessible à chacun, mais l'enjeu repose sur son exploitation et sa robustesse à l'échelle industrielle : s'adapter à des conditions d'acquisition, des objets ou des topographies différentes (figure 4) tout en respectant des temps calculs cohérents avec une chaîne de production de données.

L'utilisation d'un jeu de données important constitué de paysages variés (figure 4) avec une représentation de classes équilibrée va permettre au modèle de généraliser : c'est-à-dire être capable de segmenter de façon précise

un nuage de points présentant un paysage ou des objets nouveaux. Cette généralisation nécessite souvent un jeu de données de plusieurs centaines, voire milliers de nuages de points. La gestion de cette quantité de données peut poser de vraies contraintes de ressources mémoire et de stockage. De plus, la structure même des nuages de points en fait un type de données extrêmement volumineux, ce qui induit un traitement parfois difficile à optimiser par rapport à des données classiques comme les images.

En production, un modèle doit pouvoir effectuer des inférences (prédictions) sur des milliers de kilomètres carrés de données. Les architectures académiques ne sont pas toujours adaptées aux problématiques de traitement et d'optimisation sur de larges volumes de données, ce qui nécessite parfois une refonte profonde du code.

Valoriser la classification

L'étape de classification n'est que la première étape du traitement et de la valorisation d'un nuage de points. Lorsqu'il s'agit de cartographier un territoire, l'un des défis est d'exploiter chaque classe du nuage de points

pour en produire une donnée raster ou vectorielle précise et exploitable. Divers cas d'usages existent, on retrouve par exemple :

- la création de jumeaux numériques comme les bâtiments au format CityGML (figure 1) avec une vectorisation automatique des pans de toitures ;
- la génération de plan PCRS avec la vectorisation des objets aux abords des routes ;
- l'optimisation de réseaux de distribution ;
- l'étude de zones ombragées en ville avec la détection des zones piétonnes sous couvert végétal ;
- la création d'un cadastre solaire avec le calcul de l'exposition, de l'ombrage et potentiel solaire des toits des bâtiments (figure 5a) ;
- l'auscultation des lignes haute tension (figure 5b) et leurs distances avec les éléments proches.

Conclusions et perspectives

Les avancées de la recherche académique dans le domaine du *Deep Learning* ouvrent de nouvelles voies pour le traitement de nuages de points. L'augmentation de la qualité et de la quantité de nuages de points



LiDAR permet, grâce à ces nouvelles méthodes, une classification plus détaillée et un changement de paradigme : il ne s'agit plus de cartographier des surfaces, mais de modéliser des objets. Les volumes de données acquis ont augmenté drastiquement, entraînant une recherche d'automatisation pour permettre de tirer profit de ces nouvelles données plus riches et étendues. Malgré un coût technique de mise en place conséquent, les modèles *Deep Learning* offrent l'avantage de pouvoir généraliser à des environnements variés et permettent une analyse fine des objets. Les perspectives dans ce domaine sont prometteuses avec des développements continus de nouvelles méthodes comme les *Transformers* ou le *Self Supervised Learning*. ●

Contact

Baptiste BENARD¹ et Jules ROGER¹
Jérôme LEROUX², Camille LHENRY¹,
Maxime CHAUVIN²

Responsable Service Innovation
m.chauvin@geofit.fr

1. Ingénieur chercheur, Innovation GEOFIT
2. PhD, Innovation GEOFIT

Références

Brodu, N. and Lague, D., (2012) "3D Terrestrial LiDAR data classification of complex natural scenes using a multi-scale dimensionality criterion: applications in geomorphology", ISPRS Journal of Photogrammetry and Remote Sensing, vol. 68, pp 121-134.

Brovelli, M. and Lucca, S., (2011), "Filtering LiDAR with GRASS: overview of the method and comparisons with Terrascan", 2011 Italian Journal of Remote Sensing, vol. 42, no. 2, pp. 93-105.

Feng, Y., Zhang, Z., Zhao, X., Ji, R. and Gao, Y., (2018), "GVCNN: group-view convolutional neural networks for 3d shape recognition", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 264-272.

Hu, Q., Yang, B., Xie, L., Rosa, S., Guo, Y., Wang, Z., Trigoni, A. and Markham, A., (2019), "RandLA-Net: Efficient Semantic Segmentation of Large-Scale Point Clouds", 2020 IEEE/CVF Conference on Computer

Vision and Pattern Recognition (CVPR), 11105-11114.

Kipf, T. and Welling, M., (2016), "Semi-supervised Classification with Graph Convolutional Networks".

Sanchez-Caballero, A., de López-Diz, S., Fuentes-Jimenez, D., Losada-Gutiérrez, C., Marrón-Romera, M., Casillas-Perez, D. and Ibrahim Sarker, M., (2022), "3DFCNN: Real-Time Action Recognition using 3D Deep Neural Networks with Raw Depth Information", *Multimed Tools Appl* 81, 24119-24143.

Su, H., Maji, S., Kalogerakis, E. and Learned-Miller, E., (2015), "Multi-view convolutional neural networks for 3d shape recognition", Proceedings of the IEEE International Conference on Computer Vision, pp. 945-953 (2015).

Thomas, H., Ruizhongtai Qi, C., Deschaud, J-E., Marcotegui, B., Goulette, F. and Guibas, L. (2019), "KPConv: Flexible and Deformable Convolution for Point Clouds", Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2019, pp. 6411-6420.

ABSTRACT

In recent years, numerous campaigns to acquire point clouds using aerial LiDAR have been launched, with the aim of producing precise, detailed topographic surveys of territories. This data is increasingly rich in detail, covering large areas such as France with the LiDAR HD program. Automatic classification of point clouds is a major challenge in the geospatial data production chain. It is a complex process that requires an industrial method for classifying large quantities of data while adapting to a variety of specifications. This article presents the work carried out on the automatic classification of aerial LiDAR point clouds, particularly in the context of industrialised production. We present our feedback on the challenges of industrialising and exploiting Deep Learning models, offering more robust and accurate alternatives to so-called parametric methods.

COMITÉ DE LECTURE D'XYZ

BOSSER Pierre, professeur associé, ENSTA Bretagne, Brest

CLÉDAT Emmanuel, enseignant-chercheur, ENSG, Marne-la-Vallée

DURAND Stéphane, maître de conférences, ESGT, Le Mans

HULLO Jean-François, dr. ingénieur, EDF, Paris

KOEHL Mathieu, maître de conférences, INSA Strasbourg

LEQUEUX James, astronome émérite à l'Observatoire de Paris

MAINAUD DURAND Hélène, ingénieur topographe, CERN Genève

MISSIAEN Dominique, ingénieur topographe, CERN Genève

MOPIN Irène, ingénieur recherche, ENSTA Bretagne, Brest

MOREL Laurent, professeur des universités, ESGT, Le Mans

PANTAZIS N. Dimos, professeur, TEI Athènes

PARRA Camille, doctorante, Technische Universität München, Munich

POLIDORI Laurent, professeur, Universidade Federal do Pará, Brésil

REIS Olivier, ingénieur, traducteur, Sarreguemines

ROCHE Stéphane, professeur, Université Laval, Québec

TOUZÉ Thomas, dr. ingénieur géomètre, EDF, Grenoble

TROUILLET Michel, ingénieur topographe, Lyon

Font partie du comité de lecture les membres du comité de rédaction et la rédaction (la directrice des publications et son adjoint, le rédacteur en chef et son adjoint), voir la liste en page sommaire.