

L'insertion au service de l'intelligence artificielle : Modèle d'apprentissage pour l'annotation d'images satellite pour le consortium AI4GEO

Fabien Amarger¹, Nathalie Noblecourt¹, Jehanne Portefaix¹, Pierre-Marie Brunet²

¹ Digitanie

Cour Guillaut 09700 Saverdun

prénom.nom@digitanie.org

² CNES

Centre spatial de Toulouse

18 avenue Edouard Belin 31401 Toulouse

prénom.nom@cnes.fr

Résumé

Le consortium AI4GEO avait pour volonté de regrouper différents acteurs, publics ou privés, du milieu du traitement de données géographiques pour mutualiser les solutions et les moyens. Un des besoins communs à plusieurs membres du consortium a été de pouvoir générer une représentation en 3D à partir d'une image satellite. Pour cela, il est nécessaire d'avoir deux types de données : l'altitude des éléments et l'emprise des éléments. La hauteur est calculée par photogrammétrie à partir de plusieurs prises de vues du même point. L'emprise est plus complexe à déterminer puisqu'il faut un modèle permettant d'analyser l'image satellite, par exemple les contours d'un immeuble ou le tracé d'une route. Un modèle de segmentation sémantique d'images satellites nécessite une grande quantité de données d'apprentissage pour avoir un résultat suffisamment précis, de même qu'un enjeu fort de généralisation (conserver la performance de prédiction des zones géographiques variées). Pour obtenir ces données, la seule méthode viable est d'effectuer cette annotation manuellement pour optimiser la fiabilité. Ce travail est long et redondant, et nécessite une organisation spécifique de la production. L'entreprise d'insertion Digitanie s'est spécialisée dans la production industrielle de ce type de données. La collaboration entre le CNES (membre du consortium AI4GEO) et Digitanie a permis de produire une grande quantité de données précises et fiables. Nous proposons, dans cet article, de présenter le consortium AI4GEO, l'implication du CNES dans ce consortium et les objectifs visés. Ensuite, nous présenterons la collaboration avec l'entreprise Digitanie et l'organisation de la production qui a permis la création d'autant de données fiables. Enfin, nous concluons sur les résultats obtenus, aussi bien en terme technique et scientifique que l'apport pour l'insertion.

Mots-clés

Apprentissage, Segmentation sémantique, Analyse d'image,

Image satellite, AI4GEO, Insertion par l'activité économique

Abstract

The aim of the AI4GEO consortium was to bring together different actors, both public and private, in the field of geographic data processing, in order to share solutions and resources. One of the needs shared by several consortium members was to be able to generate a 3D representation from a satellite image. To achieve this, two types of data are required : the entity height and the entity ground surface. The height is calculated photogrammetrically from several shots taken from the same point. The ground surface, on the other hand, is more complex to determine, since it requires a model that can be used to analyze the satellite image, such as the contours of a building or the layout of a road. A semantic segmentation model for satellite imagery requires a large amount of training data to obtain a sufficiently accurate result, as well as a high level of generalizability (maintaining prediction performance for varied geographical areas). To obtain this data, the only viable method is to carry out this annotation manually, to optimize reliability. This work is time-consuming and redundant, and requires a specific production organization. The insertion company Digitanie specializes in the industrial production of this type of data. Collaboration between CNES (a member of the AI4GEO consortium) and Digitanie has produced a large quantity of accurate, reliable data. In this article, we present the AI4GEO consortium, CNES's involvement and its objectives. We will then present the collaboration with the Digitanie company and the production organization that enabled the creation of so much reliable data. Finally, we'll conclude with a look at the results obtained, both in technical and scientific terms, and in terms of their contribution to the integration process.

Keywords

Machine Learning, Semantic segmentation, Image analysis, Satellite image, AI4GEO, Integration company

1 Introduction

AI4GEO [1] est un programme français scientifique et industriel composé de plusieurs instituts publics (CNES, IGN, ONERA) et de groupes industriels (CS Group, AIRBUS Defence and Space, CLS, GEOSAT, QUANTCUBE) couvrant toute la chaîne de valeur de l'information géospatiale. Il vise à produire automatiquement des cartes sémantiques en 3D à très haute résolution et à l'échelle mondiale. Pour ce faire, une partie essentielle consiste à produire des cartes d'occupation du sol de haute qualité.

Ces cartes sémantiques ont plusieurs objectifs applicatifs et peuvent apporter un intérêt majeur pour différents domaines comme les véhicules autonomes, l'urbanisme, l'intelligence économique ou encore l'agriculture.

Pour cette tâche, les méthodes d'apprentissage profond (Deep Learning) constituent aujourd'hui l'état de l'art. En particulier, les réseaux entièrement convolutifs (Fully Convolutional Networks - FCN) [4] sont désormais courants pour effectuer une telle tâche de segmentation sémantique. Ces modèles ont réussi à surpasser les méthodes traditionnelles grâce à leur capacité à extraire et à apprendre automatiquement les caractéristiques pertinentes pour la tâche considérée et en prenant en compte le contexte spatial [3, 5]. Ce dernier point est nécessaire pour améliorer les performances du traitement des images satellites de très haute définition en vue de produire des cartes précises de l'occupation des sols.

Dans cet article, nous allons présenter la méthode mise en place pour l'entraînement d'un modèle pour l'analyse d'images satellites en vue d'obtenir ces cartes précises de l'occupation des sols. Nous présenterons ensuite comment nous avons pu obtenir un grand nombre de données de labellisation, des images de grande qualité pour permettre cet entraînement grâce à l'aide de l'activité par l'insertion économique. Enfin, nous concluons sur l'intérêt, pour l'avenir de l'intelligence artificielle, d'obtenir des données d'apprentissage de qualité, produites en France et permettant à des personnes éloignées de l'emploi de contribuer à ces sujets.

2 Vectorisation sémantique multi-classe d'images satellites très haute définition [6]

La base de données AI4GEO est constituée d'images satellites Pléiades¹ stockées sous forme de géotiffs optimisés pour le cloud (Cloud Optimized Geotiffs - COG) de 16 bits et d'une vérité de terrain annotée manuellement qui a été produite tout au long du projet. Les images utilisées pour l'entraînement et l'inférence sont des images ortho rectifiées d'une résolution de 50 cm avec une correction du

1. <https://dinamis.data-terra.org/en/pleiades-products/>

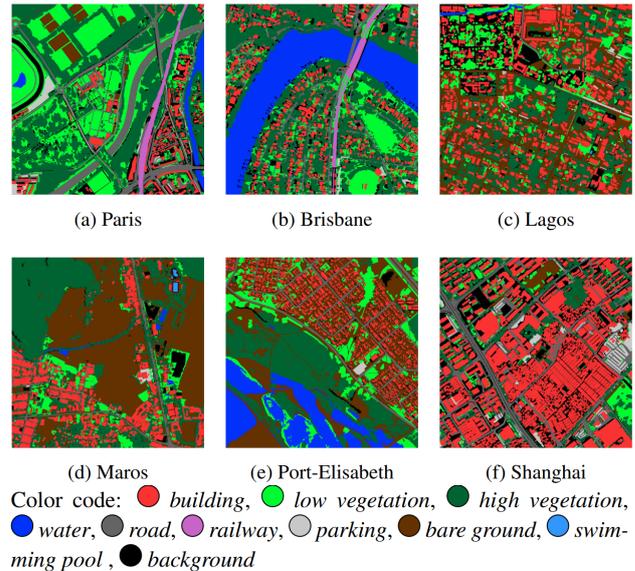


FIGURE 1 – Exemples de résultats de la vectorisation sémantique

sommet de l'atmosphère (Top Of Atmosphere - TOA). Il s'agit d'une opération de prétraitement standard qui génère des images de réflectance (corrigées radiométriquement à partir de l'étalonnage du capteur et des effets atmosphériques systématiques de l'atmosphère). Chaque image est composée de quatre bandes spectrales : Rouge, Vert, Bleu et Proche-Infrarouge. Quatorze villes présentant un intérêt industriel ont été choisies : Arcachon, Biarritz, Brisbane, Lagos, Maros, Montpellier, Munich, Nantes, Paris, Port-Elisabeth, Shanghai, Strasbourg, Tianjin, Toulouse. Pour toutes ces villes, les responsables du projet au CNES ont sélectionné 10 tuiles de 2048x2048 pixels représentant la diversité du paysage sémantique. Ces tuiles ont été segmentées manuellement, puis chaque polygone a été classé par des opérateurs en photo interprétation. Nous avons mis en place un processus incrémental tout au long de la production de la vérité terrain afin d'assurer un dialogue efficace entre les opérateurs et les experts en Machine Learning. Nous reviendrons plus en détails sur ce processus dans une section dédiée dans cet article. Nous obtenons une vérité terrain de haute qualité jusqu'à 137 classes différentes. Pour la phase d'apprentissage du modèle, uniquement 10 classes ont été utilisées. Pour ce faire, nous avons forcé le classement de chaque pixel parmi les classes suivantes : bâtiment, route, chemin de fer, végétation haute, végétation basse, eau et arrière-plan (autres). Ce choix a été fait après une discussion entre les différents partenaires pour décider quelles classes étaient les plus utiles pour les différents cas d'utilisation. Les classes sont déséquilibrées dans cet ensemble de données, mais cela représente une difficulté structurelle de la tâche. La figure 1 donne des exemples de vérités terrain obtenues par labellisation manuelle. OpenStreetMap (OSM) [2] est une base de données géographique ouverte et collaborative qui fournit des caractéristiques géographiques dans le monde entier. L'utilisation de ce jeu de

données permet d'exploiter des images Pléiades complètes (environ 40000x40000 pixels) et de ne pas se limiter aux tuiles présélectionnées. Cependant, la qualité de la vérité terrain extraite de ce jeu de données peut varier en fonction de la zone considérée. Plus précisément, la vérité terrain de certains éléments peut être légèrement mal placée ou même complètement manquante. En outre, la date d'acquisition de l'image Pléiades peut impliquer des différences de mise à jour avec OSM. Les données OSM ne peuvent donc pas suffire comme base d'entraînement pour l'entraînement du modèle de vectorisation sémantique, mais ces données peuvent être une aide, précieuse, pour la vectorisation manuelle. C'est donc dans ce contexte que l'entreprise Digitanie a été sollicitée pour annoter manuellement des images Pléiades par photo-interprétation avec l'aide des données OSM.

3 Processus de production

Dans le cadre d'un partenariat entre le CNES (membre du consortium AI4GEO) et Digitanie, un processus de production a été mis en place pour répondre au besoin de données géographiques de très haute qualité pour l'apprentissage d'un modèle de segmentation sémantique. À l'instar d'une usine de production mécanique de précision, un processus de production a été mis en place pour labelliser des images satellites par photo-interprétation.

L'objectif est d'annoter un sous-ensemble d'images, représentatives, des villes à labelliser. Ces images sont fournies par le CNES, Digitanie réalise la vectorisation manuelle. Ce travail nécessite un gros volume horaire pour réussir à analyser et vectoriser l'intégralité des images, c'est pour cela que nous parlons d'usine du numérique pour de la production de données géographiques.

3.1 Organisation de production

Dans un objectif de proposer une labellisation de qualité et à moindre coût, il a été nécessaire de penser l'organisation de la production dans son ensemble pour optimiser l'intégralité du processus de production.

Pour effectuer la labellisation, nous utilisons l'outil QGIS² qui est un outil libre de traitement de données géographiques.

Échantillon, méthodologie et métriques

La première étape consiste à établir une méthodologie à partir d'un échantillon suffisamment représentatif de données ; À partir de cet échantillon, les référents techniques définissent des indicateurs ou métriques, qui vont déterminer un temps de production unitaire pour une surface d'image traitée (découpée en tuiles). La taille de l'image doit être suffisamment importante pour permettre une production sur un temps donné par un opérateur, mais ne doit pas être trop grande pour permettre une livraison et un contrôle de façon continue. Les référents techniques définissent une méthodologie précise, partagée et validée par le CNES, qui permet de déterminer les classes à labelliser et les interprétations à effectuer lors d'ambiguïtés. De plus, ils

se rendent disponibles régulièrement pour aider les opérateurs. Cette proximité et la clarté de la méthodologie sont d'autant plus importants, et nous le verrons plus en détail plus tard, qu'elles permettent à des personnes n'ayant pas de connaissance en traitement de données géographiques de pouvoir contribuer aisément au projet.

Formation

Une fois cette étape de méthodologie établie à partir de l'échantillon, Les salariés doivent être formés. Effectivement, cette phase de formation permet aux opérateurs de s'approprier la méthodologie, de l'éprouver, voir de l'améliorer dans une démarche d'amélioration continue. Cette phase de montée en compétences, est essentielle pour permettre un rendement maximum dès la mise en production des opérateurs et ne doit pas être négligée.

Production

Une fois la méthodologie définie, les métriques établies, les opérateurs formés, il est maintenant possible de passer à la phase de production. Elle est organisée par vagues pour permettre d'observer les métriques à la fin de chaque vague et d'ajuster la méthodologie itérativement. Chaque vague correspond à un ensemble de tuiles à labelliser provenant d'une ville en particulier. Un travail préparatoire, effectué par les référents techniques, permet aux opérateurs de se concentrer sur la production. Une répartition des tâches de production est faite en début de semaine, pour que chaque opérateur n'ait ni trop, ni pas assez de travail pour la semaine et pour paralléliser la production au maximum. Les opérateurs remplissent, quotidiennement, les indicateurs pour pouvoir suivre l'état de la production et pouvoir planifier le travail pour permettre la continuité de la production. Digitanie et le CNES ont mis un point d'honneur, tout au long du projet, à s'assurer que cette répartition du travail était en adéquation avec des conditions respectueuses des opérateurs. Cette répartition prend en compte l'ancienneté, mais aussi la diversité des personnes. Ce respect des personnes permet de s'assurer que le travail n'est pas trop fastidieux et de ne pas démotiver les opérateurs. De plus, un espace de communication dédié au projet a été mis en place pour faciliter l'entraide et la collaboration entre les opérateurs.

Contrôle

Un contrôle croisé est effectué par les opérateurs, ce qui permet de sensibiliser à la qualité du travail attendu. Une fois les images produites et contrôlées, elles sont à nouveau contrôlées (de façon plus générale) par un référent technique avant d'être livrées pour les clients.

Retours d'expériences

Afin de garantir un niveau de qualité optimum de la production des données, il est nécessaire d'organiser des retours d'expériences régulièrement. Nous avons mis en place trois catégories de retours d'expériences. La première catégorie est un retour d'expériences hebdomadaire entre les référents techniques et les clients afin de s'assurer de la satisfaction de ces derniers quant à la qualité de la production, mais aussi de l'efficacité de celle-ci. Ensuite, nous organisons des retours d'expériences entre les référents techniques et

2. <https://qgis.org/>

les opérateurs pour clarifier certains points de la méthodologie qui auraient pu évoluer, mais aussi pour s'assurer qu'elle est soutenable par les opérateurs. Enfin, le dernier type de retours d'expériences est entre les opérateurs et les clients. Ces retours sont moins fréquents, mais sont très importants et ont plusieurs objectifs : permettre aux clients de se rendre compte du travail réalisé par les opérateurs et permettre aux opérateurs de comprendre réellement les besoins et les objectifs du projet pour donner du sens à leur travail quotidien. Tout au long du projet, un suivi des métriques est effectué par les référents techniques, en toute transparence avec les clients. Cette transparence est un pilier de tout le processus de production, autant d'un point de vue suivi économique, de qualité du résultat attendu mais aussi d'un point de vue scientifique. Effectivement, la transparence permet un ajustement permanent de la méthodologie et du processus de production. Par exemple, certaines difficultés relevées par les opérateurs ont influencé les choix des futures tuiles à labelliser.

3.2 L'intérêt de la production numérique pour l'insertion

Comme évoqué précédemment, nous recrutons des opérateurs à tous niveaux techniques concernant le traitement de données géographiques, mais aussi concernant l'utilisation de l'outil informatique de manière générale. Digitanie est une entreprise d'insertion et de ce fait, elle accompagne des personnes éloignées de l'emploi à s'insérer par l'activité économique (IAE)³ et qui plus est dans un milieu rural et donc assez éloigné des grandes zones technologiques que peuvent connaître les grandes villes. Nous avons mis en place tout un parcours de suivi de compétence de ces personnes pour les aider à acquérir le niveau attendu pour pouvoir contribuer aux projets de traitement de données géographiques et notamment à la labellisation d'images satellites pour le CNES dans le cadre du consortium AI4GEO. Nous avons observé une moyenne d'un mois de formation nécessaire pour qu'une personne, quel que soit son niveau initial, puisse commencer à contribuer à labelliser ces images. Il est particulièrement gratifiant pour des personnes qui n'avaient jamais travaillé sur un ordinateur auparavant, de pouvoir contribuer rapidement à un sujet de pointe autour de l'intelligence artificielle et l'observation de la Terre. Cette activité crée aussi des vocations puisqu'un certain nombre de personnes se réorientent vers la géomatique après leur passage à Digitanie.

4 Conclusion

Nous avons tout d'abord décrit les objectifs du consortium AI4GEO en termes de création de cartes géospatiales 3D de très hautes qualités. Nous avons détaillé les outils techniques et technologiques mis en œuvre pour atteindre ces objectifs avec un besoin en termes de labellisation manuelle d'images satellites. Nous avons ensuite précisé le processus de production manuelle mis en place pour permettre cette

labellisation avec un très haut niveau de qualité. Cette production permet d'améliorer la qualité des modèles obtenus et donc d'améliorer la labellisation automatique des images satellites d'observation de la Terre. Enfin, nous avons expliqué en quoi l'organisation d'une entreprise d'insertion pour de la production numérique permet de répondre à ces besoins. Les modèles d'intelligence artificielle sont toujours plus performants en absorbant toujours plus de données. Mais la qualité des données utilisées pour l'apprentissage de ces modèles impact directement la qualité des modèles eux-mêmes. Pouvoir disposer des données de très haute qualité pour l'apprentissage permet d'optimiser cette phase d'apprentissage. De plus, ce travail qui est long et fastidieux, peut être une chance pour des personnes éloignées de l'emploi de retrouver une stabilité économique grâce à l'IAE. C'est cette complémentarité entre l'IA et l'IAE que nous avons voulu mettre en avant dans cet article et que nous souhaitons continuer à faire évoluer sur d'autres sujets de traitement de données (géographiques ou non) et plus généralement sur cette notion de production industrielle du numérique tout en continuant à accompagner et à aider des personnes éloignées de l'emploi, notamment en valorisant leurs compétences sur des sujets de pointes.

Références

- [1] Pierre-Marie Brunet, Pierre Lassalle, Simon Baillarin, Bruno Vallet, Arnaud Le Bris, Gaëlle Romeyer, Guy Le Besnerais, Flora Weissgerber, Gilles Foulon, Vincent Gaudissart, et al. Ai4geo : A data intelligence platform for 3d geospatial mapping. In *24th ISPRS Congress Commission II : Imaging Today, Foreseeing Tomorrow*, volume 43, pages 817–823, 2021.
- [2] OpenStreetMap Contributors. Planet dump retrieved from <https://planet.osm.org>. 2022.
- [3] Alberto Garcia-Garcia, Sergio Orts-Escolano, Sergiu Oprea, Victor Villena-Martinez, and Jose Garcia-Rodriguez. A review on deep learning techniques applied to semantic segmentation. *arXiv preprint arXiv :1704.06857*, 2017.
- [4] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3431–3440, 2015.
- [5] Bipul Neupane, Teerayut Horanont, and Jagannath Aryal. Deep learning-based semantic segmentation of urban features in satellite images : A review and meta-analysis. *Remote Sensing*, 13(4) :808, 2021.
- [6] Axel Rochel, Clément Deschesne, Adrien Chan-Hon-Tong, and Pierre-Marie Brunet. Multiclass semantic segmentation with very high-resolution satellite images. In *Proceedings of the 2023 conference on Big Data from Space*, pages 21–24. Publications Office of the European Union, 2023.

3. <https://travail-emploi.gouv.fr/insertion-par-lactivite-economique-retour-sur-le-congres-des-5-et-6-decembre-2024>