

# Offre de post-doc au sein du programme de recherche AI4GEO 24 mois

(démarrage à partir de Septembre 2020)

## 1. Description

Dans le cadre du projet AI4GEO ( <http://ai4geo.io> ), les partenaires industriels et institutionnels ont convergé sur la création de produits à l'échelle globale valorisant les données spatiales 2D et 3D, à savoir :

- Séries temporelles de cartes d'occupation des sols à large échelle et à très haute résolution en utilisant conjointement les informations radiométriques et altimétriques.
- Création de séries temporelles de modèles LOD1 et LOD2 afin de surveiller l'aménagement urbain.
- Création de modèles 3D sémantisés calculés à partir d'un maillage de nuages de points 3D (stéréo et lidar) classifiés
- Détection et caractérisation des changements dans une série de nuages de points 3D (stéréo et lidar) pour des applications liées à la navigation autonome et au suivi des voies ferroviaires.

Dans ce contexte, plusieurs axes de recherche ont été identifiés pouvant aboutir à des travaux ayant un fort impact scientifique :

- Débruitage de Modèles Numériques de Surface : Les modèles numériques de surface et les nuages de points en sortie de chaîne de restitution altimétrique à partir d'images satellite multi-vues sont souvent très bruités et ce bruit a un impact sur la qualité de prédiction des modèles (machine learning et deep learning) ultérieurs. Cet axe de recherche consistera à étudier des méthodes de débruitage 2.5D ou 3D. Plusieurs solutions sont envisageables : IA et approches statistiques.
- Classification 3D multi-échelle : La classification OBIA 2.5D et 3D à plusieurs échelles sémantiques en utilisant l'information contextuelle (spatiale et temporelle). Aujourd'hui, produire une carte de classification d'objets est largement réalisable mais utiliser l'agencement spatial de ces objets entre eux pour en déduire des méta-objets (ensemble d'objets ayant une signification sémantique par exemple un groupe de maisons alignés avec des piscines et des jardins constitue un lotissement) constitue un vrai défi. De plus, l'utilisation de modèles « vrai 3D » (avec texture associée) permettra d'améliorer la performance de classification sémantique.
- Détection de changement 3D : La détection de changement dans des séries temporelles de nuages de points 3D. Cet axe de recherche fait déjà l'objet actuellement d'une thèse avec le laboratoire IRISA. Cependant, nous anticipons d'ores-et-déjà qu'il sera difficile de caractériser ce changement (car nous manquons notamment de vérités terrain associées) et de réaliser un passage à l'échelle de ces méthodes. L'étude aura pour objectif d'être en forte synergie avec la doctorante et d'augmenter la maturité des méthodologies proposées (passage à l'échelle et qualification des résultats).
- Fusion et recalage de nuages de points à différentes échelles (LIDAR terrestre / aéroporté / données satellites).
- Qualification des modèles IA : Aujourd'hui, pour mesurer la confiance associée à un modèle prédictif (machine learning et deep learning), la méthode classique reste de diviser un ensemble de données dont nous disposons des vérités terrain en 2 sous-ensembles : apprentissage et test. Des métriques telles que le score F1, la précision, le Recall, ROC, IOU sont évaluées sur le sous-ensemble test. Cependant, lorsque ces modèles sont déployés en phase opérationnelle, ils prédisent chaque nouvelle observation et donnent forcément une étiquette à celle-ci. L'idée est donc d'explorer des approches permettant d'identifier si la nouvelle observation est issue d'une réalisation d'une loi de distribution similaire à celle ayant généré les données d'apprentissage. Des méthodes nouvelles, telles que les approches de réseau de neurones profonds probabilistes, apparaissent notamment pour des modèles prédictifs embarqués dans des environnements critiques.

- Correction des biais d'apprentissage : Enfin un axe de recherche également très important et d'évaluer et de corriger le biais/bruit introduit par les erreurs soit humaines soit issues de mesures de capteur dans des données de référence (vérités terrain). En effet, la présence de ces données bruitées entraîne souvent une perte de généralisation des modèles IA car ceux-ci deviennent alors très sensibles à des perturbations mineures sur les caractéristiques de nouvelles observations ce qui peut donc induire une mauvaise classification.

Le post-doc sera mené au CNES à Toulouse dans le cadre du projet AI4GEO. Il sera encadré par le laboratoire de recherche d'accueil.

La personne recrutée sera intégrée à l'équipe AI4GEO et pourra interagir notamment avec différents services au sein du CNES, de l'ONERA et de l'IGN. Elle sera également en contact avec les équipes de recherche du 3IA Toulousain (ANITI).

## 2. Profil

Le candidat devra avoir une formation doctorale en informatique / traitement du signal / mathématiques appliquées et être intéressée par les domaines et techniques de l'apprentissage automatique (Machine Learning/Deep Learning), la vision par ordinateur et la télédétection (une expérience significative dans ces domaines serait évidemment un plus).

Pour candidater envoyer lettre de motivation, CV et liste de publications à [Pierre.lassalle@cnes.fr](mailto:Pierre.lassalle@cnes.fr), [Pierre-marie.brunet@cnes.fr](mailto:Pierre-marie.brunet@cnes.fr) & [Simon.baillarin@cnes.fr](mailto:Simon.baillarin@cnes.fr)

# Post-doc position within the AI4GEO research program 24 months

(starting from September 2020)

## 1. Description

As part of the AI4GEO project ( <http://ai4geo.io> ), industrial and institutional partners have converged on the creation of products on a global scale valuing 2D and 3D spatial data, namely:

- Time series of land cover maps on a large scale and at very high resolution using jointly radiometric and altimetric information.
- Creation of time series of LOD1 and LOD2 models in order to monitor urban development.
- Creation of semantic 3D models calculated from a mesh of classified 3D point clouds (stereo and lidar)
- Detection and characterization of changes in series of 3D point clouds (stereo and lidar) for applications related to autonomous navigation and monitoring of railway tracks.

In this context, several axes of research have been identified which can lead to work with a strong scientific impact:

- **Denoising of Digital Surface Models:** Digital surface models and point clouds at the output of the photogrammetric restitution chain from multi-view satellite images are often very noisy and this noise has an impact on the prediction quality of subsequent models (machine learning and deep learning). This line of research will consist in studying 2.5D or 3D denoising methods. Several solutions are possible: AI and statistical approaches.
- **Multi-scale 3D classification:** OBIA 2.5D and 3D classification on several semantic scales using contextual information (spatial and temporal). Today, producing an object classification map is largely feasible, but using the spatial arrangement of these objects among themselves to deduce meta-objects (set of objects having semantic meaning, for example a group of houses aligned with swimming pools and gardens is a housing development) is a real challenge. In addition, the use of "true 3D" models (with associated texture) will improve the performance of semantic classification.
- **3D change detection:** Change detection in 3D point cloud time series. This line of research is already the subject of a thesis with the IRISA laboratory. However, we already anticipate that it will be difficult to characterize this change (because we notably lack associated ground truths) and to carry out a scaling up of these methods. The objective of the study will be to be in strong synergy with the doctoral student and to increase the maturity of the methodologies proposed (scaling up and qualification of the results).
- **Fusion and registration of point clouds at different scales** (terrestrial / airborne LIDAR / satellite data).
- **Qualification of AI models:** Today, to measure the confidence associated with a predictive model (machine learning and deep learning), the classic method remains to divide a set of data of which we have field truths into 2 subsets: learning and test. Metrics such as F1 score, accuracy, Recall, ROC, IOU are evaluated on the test subset. However, when these models are deployed in the operational phase, they predict each new observation and necessarily give a label to it. The idea is therefore to explore approaches making it possible to identify if the new observation is the result of an implementation of a distribution law similar to that which generated the learning data. New methods, such as probabilistic deep neural network approaches, appear in particular for predictive models embedded in critical environments.
- **Correction of learning biases:** Finally, an axis of research that is also very important and of evaluating and correcting the bias / noise introduced by either human errors or from sensor measurements in reference data (field truths). Indeed, the presence of this noisy data often leads to a loss of generalization of IA models because these then become very sensitive to minor disturbances on the characteristics of new observations which can therefore induce a poor classification.

The post-doc will be conducted at CNES in Toulouse, France as part of the AI4GEO project. It will be supervised by the host research laboratory.

The recruited person will be integrated into the AI4GEO team and will be able to interact with various departments within CNES, ONERA and IGN. He/She will also be in contact with the ANITI research teams.

## 2. Profile

The candidate must have a PhD in computer science / signal processing / applied mathematics and be interested in the fields and techniques of Machine Learning / Deep Learning, Computer Vision or Remote Sensing.

To apply, send ASAP cover letter, CV & publications list: [Pierre.lassalle@cnes.fr](mailto:Pierre.lassalle@cnes.fr), [Pierre-marie.brunet@cnes.fr](mailto:Pierre-marie.brunet@cnes.fr) & [Simon.baillarin@cnes.fr](mailto:Simon.baillarin@cnes.fr)