



HAL
open science

AI4GEO: A Path From 3D Model to Digital Twin

Pierre-Marie Brunet, Simon Baillarin, Pierre Lassalle, Flora Weissgerber, Bruno Vallet, Triquet Christophe, Gilles Foulon, Gaëlle Romeyer, Gwenaël Souille, Laurent Gabet, et al.

► **To cite this version:**

Pierre-Marie Brunet, Simon Baillarin, Pierre Lassalle, Flora Weissgerber, Bruno Vallet, et al.. AI4GEO: A Path From 3D Model to Digital Twin. IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium, Jul 2022, Kuala Lumpur, Malaysia. pp.4728-4731, 10.1109/IGARSS46834.2022.9883433 . hal-04108171

HAL Id: hal-04108171

<https://hal.science/hal-04108171v1>

Submitted on 26 May 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

AI4GEO: A PATH FROM 3D MODEL TO DIGITAL TWIN

Pierre-Marie Brunet¹, Simon Baillarin¹, Pierre Lassalle¹, Flora Weissgerber², Bruno Vallet³, Triquet Christophe⁴, Gilles Foulon², Gaëlle Romeyer³, Gwenaël Souille⁴, Laurent Gabet⁵, Cedrik Ferrero⁶, Thanh-Long Huynh⁷, Emeric Lavergne⁸

¹ Centre National d'Etudes Spatiales, Toulouse, France

² ONERA, Paris, France

³ LASTIG, Univ Gustave Eiffel, ENSG, IGN, F-94160 Saint-Mandé, France

⁴ CS Group, Toulouse, France

⁵ AIRBUS DS Intelligence, Toulouse, France

⁶ GEOSAT, Bordeaux, France

⁷ QuantCube, Paris, France

⁸ CLS, Toulouse, France

<https://www.ai4geo.eu>

ABSTRACT

3D Geospatial information plays a key role in many soaring sectors such as sustainable and smart cities, climate monitoring, ecological mobility, and economic intelligence. The availability of huge volumes of satellite, airborne and in-situ data now makes this production feasible at large scale. It needs nonetheless a certain level of manual intervention to secure the level of quality, which prevents mass production. This paper presents the AI4GEO program that aims at developing an end to end solution to produce automatically qualified 3D Digital model at scale together with multiple layers of information.

Index Terms— Optical Image processing, Photogrammetry, Semantic segmentation, CityGML, 3D, Digital Twin.

1. INTRODUCTION

The AI4GEO project aims at developing an automatic solution for producing 3D geospatial information and new added-value services.

Started at the end of 2019 for a 4-years duration the project timeframe is divided into 4 key stages (KS 1-4) and has already produced promising results for KS2.

This paper will first briefly present the AI4GEO project, context and overall objectives. It will then focus on the 3D pipeline that aims to produce Level Of Details (LoD) cityGML products as shown in Fig. 1 and related results achieved.



Fig. 1. LoDs illustration of CityGML 2.0 [1]

2. AI4GEO CONSORTIUM

AI4GEO is a French scientific and industrial program aimed at lifting the technological barriers to the automatic production of 2D and 3D Geographic Data. The consortium is composed of institutional partners (CNES, IGN, ONERA) and industrial groups (CS Group, AIRBUS Defense and Space, CLS, GEOSAT, QUANTCUBE) covering the whole value chain of Geospatial Information. To foster interaction and exchanges between academic and industrial partners, the project is structured around 2 axes which will progress simultaneously:

The first axis consists in developing a set of building blocks allowing the automated production of qualified 3D maps and their additional layers of information (3D objects and related semantics with temporal dynamic) as illustrated in Fig. 2. These blocks are provided with a new Data analytics platform [2] designed to access, manipulate, process and visualize at scale any kind of geospatial products (optical, radar, etc.).

This collaborative work will benefit from the latest research from all the partners (imagery, AI and Big Data technologies) as well as from an unprecedented database (satellite and airborne data (optics, radars, Lidars) combined with cartographic and in-situ data).



Fig. 2. 3D maps and semantics

The second axis consists in deriving from these technological bricks a variety of new services illustrated in Fig. 3 for fields targeted by industrial partners.

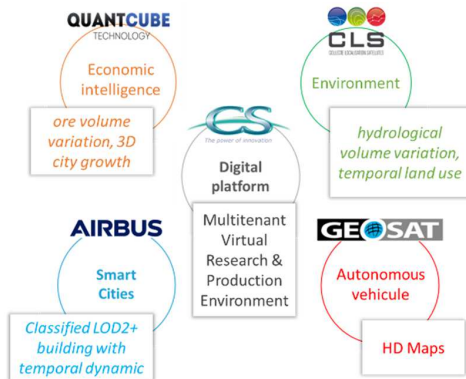


Fig. 3. Industrial services

This list could be extended in the future by early adopters (e. g. Precision Agriculture, or Robotics). It is also important to highlight that the project will benefit from the arrival of satellite constellations such as AIRBUS Pleiades-Neo program (in 2021) and above all CNES CO3D program [3] that will greatly increase the provision of accurate, Very High Resolution 3D data.

3. 3D PIPELINE

The 4-year project aspires to automatically produce 3D semantic maps at very high resolution and global scale. The 3D pipeline that has been developed and deployed onto the platform is presented below.

The objective is to extract building footprints, build the 3D models thanks to the Digital Elevation Models (DEM) and merge it with rich semantic layers (LULC, Change characterization, etc.).

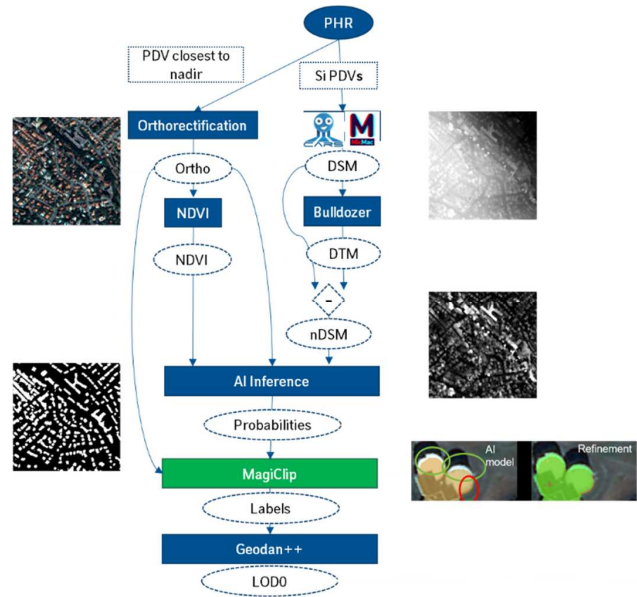


Fig. 4. 3D pipeline

3.1 Input Data

First, the input images of the pipeline consist of Pleiades-HR images of the same geographical area of interest with different viewing angles. All Pleiades-HR images are provided in perfect sensor geometry associated with a geometric sensor model (RPC). Native resolution of the panchromatic band is 0.7 m and native resolution of multi-spectral bands is 2.8 m.

Second, as our algorithms mostly rely on supervised deep learning, ground truths (GT) are also needed to train networks but also to assess the quality of the outputs. Analysis of the preliminary results pointed out a performance limitation due to multiple bias in the ground truths (GT), especially when considering alignment of object edges (highlighted by Intersection Over Union, IoU metric) even if the classification itself is good. Therefore, an activity on labelling tools has been initiated with two main axes. On the one hand a software, named *LabelCooker* is being developed and aims at merging existing GTs by combining them using different merging strategies such as completion or precision through geometrical approaches. On the other hand, a semi-automatic method is developed in the software *Autolabel*, which aims at producing new GT datasets using an active learning approach in order to increase labelling productivity.

3.2 Processing

The pipeline (Fig. 4) is designed to orchestrate modules via standardized API that allows to easily switch from one module to another depending on the objective or input data. The first step starts with DSM computation using CARS [4],

an open source stereo pipeline designed for scalability and robustness or MicMac, a photogrammetry software developed at the IGN. Each image of the stereoscopic pair of is converted in epipolar geometry and a disparity map is computed based on the Semi Global Matching algorithm [5]. In parallel a study is under way to use deep learning for the dense correlation computation. First result shows interesting performance improvement.

The second step of the pipeline aims at extracting the Digital Terrain Model (DTM) from the DSM at large scale using a specific module called Bulldozer. The algorithm is an improvement of the original method based on Lidar data [6] to better handle photogrammetric DSM and scalability. The idea of the algorithm is to simulate a rigid cloth falling on the inverted surface, as illustrated in Fig. 5. The high frequency distortions of DSM are therefore filtered whereas the low frequency variations of ground altitude are captured. At last, the Digital Height Model (DHM) is obtained. But the original algorithm adapted to lower resolution DSM coming from satellite imagery suffered from multiple problems. It hardly takes into account area with significant relief or topological discontinuity (cliff), further some artefacts were observed in the resulting DTM.

A preprocessing phase has been added to identify noisy pixels in the DSM and to filter them by bilinear interpolation. Then the number of iteration of the algorithm has been limited according to the maximum size of an over ground object in order to prevent unhooking of the DTM on relief.

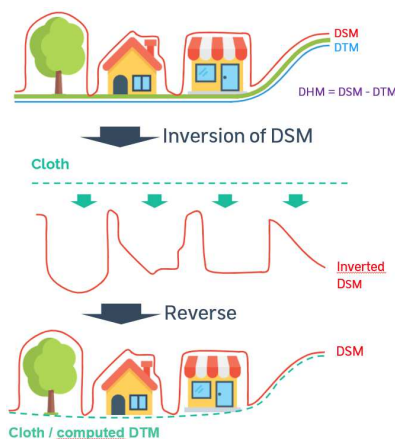


Fig. 5. DTM generation

Then the Pleiades-HR spectral bands (Red, Green, Blue, Near Infrared) the DHM, and two additional NDVI and C3 (shadow) channels are combined into a tensor. The building semantic segmentation process is performed using a U-Net architecture with an EfficientNet encoder and RefineNet decoder. The training database has been produced with LabelCooker by using a large scale but low precise OpenStreetMap Data (PHR acquisition and OSM vectorization can be temporally out of sync) together with an

in house high precision Label Database. Thus the model trained and fine-tuned provides a probability map that gives for each pixel the probability to belong to a building.

At this step, the building shapes are too smooth, the next step aims at generating a highly precise building vectors map.

The first optimization is performed on the raster data. It is based on a radiometric prior knowledge: homogeneity of the building's roof pixels whereas an important discontinuity is expected along edges. We developed a tool, called Magiclip, that implements a Grabcut algorithm in order to refine the shape of each building according to this rule.

A morphological post processing is then applied which consists of removing small buildings (when its area in pixels is lower than a user given threshold) and small inner courtyards. Then, a connected component segmentation is carried out in order to individualize each building.

The second optimization is performed on vector data with the benefit of the geometrical prior knowledge of building forms. A RANSAC regularization method is applied with geometric constraints considering the straight angles or parallel opposite segments to obtain a refined LOD0 (Level Of Detail 0) shape of each building. Fig.6 illustrates intermediate products and final optimized LOD0.

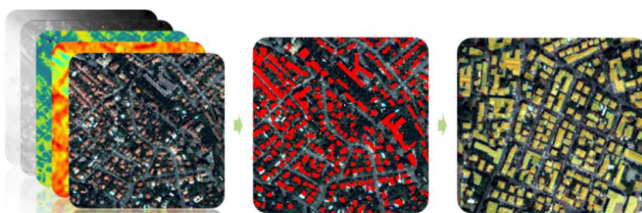


Fig. 6. LOD0 pipeline inputs (PHR multiview, C3, NDVI), intermediate extracted building and final LOD0

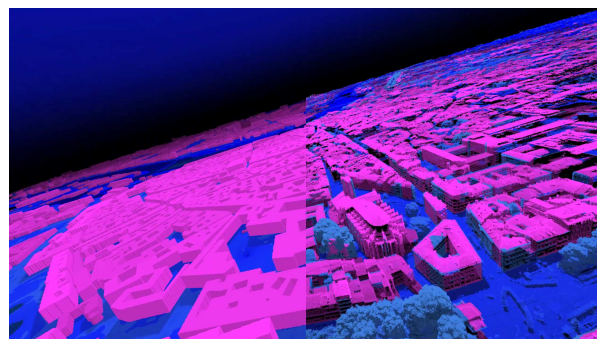


Fig. 7. LOD1 (left) compared to LIDAR (right) on Toulouse city

The next step consists of computing the LOD1 (flat roof) 3D reconstruction of buildings. Thanks to the DHM, statistical height measures such as the minimum, maximum, mean or median are computed for of each building and then 3D meshes are constructed for each one.

Current work is focusing on exploring algorithms to determine the roof shape of each building in order to compute a **LOD2** 3D reconstruction. To do so, roof faces need to be identified. Two methods have been applied: an unsupervised segmentation (Mean-Shift) and a supervised approach based on a MASK-RCNN to identify each roof face. The spatial arrangement of each roof facet for one building is then compared to a bank of standard roof types and the closest one is selected as a better approximation. Current results are promising but we need to optimize the performance and the generalization capability of our algorithms.

3.3 Results

A set of metrics has been defined to assess the quality of the LOD0 product.

First, the usual AI scores are computed on semantic segmentation prediction. The table 1 summarizes the performance obtained. The training phase consists of using 100k+ buildings over Toulouse, Montpellier and Biarritz cities and the prediction is performed on Paris and Strasbourg.

Table 1. Semantic segmentation performance

Accuracy	Mean IOU	F1Score	Precision	Recall
0.9333	0.818	0.835	0.895	0.782

Then, two sets of metrics are computed to evaluate the quality of building shape compared to the ground truth (BD IGN Topo): first topological metrics (pairing ratio, over and sub segmentation) gives information about building instances precision. Second, geometric metrics (contour and surface) tells us about the 2D shape precision:

- Intersection over Union (IoU) to evaluate the location and the global shape accuracy
- Ratio of the number of segments (between the resulted shape and the reference shape) to evaluate the simplicity of the result building shapes
- The difference of the orientation of the minimum oriented bounding box between the result shape and the reference one.

Table 2. Regularization performance

	Mean	Standard dev
IoU	0.692	0.11
Ratio segments	1.81	0.86
Orientation	0.15 rad	0.09 rad

Finally, 3D LOD1 performance evaluation implies to assess the altitude precision computed for each building that boils down to the assessment of DSM and DTM products. LOD1 ground truth of Strasbourg city (provided in opendata) was used in this case.

Table 3. LOD1 performance

Mean (bias)	3.14
Median	2.94
Standard deviation (error)	5.90

4. CONCLUSION

In this paper, the AI4GEO project has been presented and some of its current achievements reviewed. Many other research activities have been carried out within partner labs not reviewed in this paper (Urban&Global land cover, detection change, etc.). Ongoing work tackle new challenges towards the Global Smart Map (fully automatic LOD2, Generalization), together with performance improvements.

ACKNOWLEDGEMENTS

The project is funded for 4 years by the French future investment program led by the Secretariat General for Investment and operated by public investment bank BPI France.

REFERENCES

- [1] Biljecki, Filip & Ledoux, Hugo & Stoter, Jantien. (2016). An improved LOD specification for 3D building models. *Computers Environment and Urban Systems*. 59. 25-37. <https://doi.org/10.1016/j.compenvurbsys.2016.04.005>
- [2] M. Darques et al. "AI4GEO Engine: a hybrid HPC/Cloud AI oriented platform for Earth-wide EO data processing", *Proc. of BiDS'21*
- [3] Lebègue, L. & Cazala-Hourcade, E. & Languille, F. & Artigues, S. & Melet, Olivier. (2020). CO3D, A WORLDWIDE ONE ONE-METER ACCURACY DEM FOR 2025. *ISPRS 2020*. <https://doi.org/10.5194/isprs-archives-XLIII-B1-2020-299-2020>
- [4] Youssefi, David & Michel, Julien & Sarrazin, Emmanuelle & Buffe, Fabrice & Courmet, Myriam & Delvit, Jean-Marc & L'Helguen, Celine & Melet, Olivier & Emilien, Aurelie & Bosman, Julien. (2020). CARS: A Photogrammetry Pipeline Using Dask Graphs to Construct A Global 3D Model. 453-456. <https://doi.org/10.1109/IGARSS39084.2020.9324020>
- [6] Zhang, Wuming & Qi, Jianbo & Peng, Wan & Wang, Hongtao & Xie, Donghui & Wang, Xiaoyan & Yan, Guangjian. (2016). An Easy-to-Use Airborne LiDAR Data Filtering Method Based on Cloth Simulation. *Remote Sensing*. 8. 501. <https://doi.org/10.3390/rs8060501>