

RECONSTRUCTION AUTOMATIQUE DE CONTOURS DE TOITS EN 3D À L'AIDE DU FRAME FIELD LEARNING

Célestin Huet¹

1: Institut national de l'information géographique et forestière (IGN), France, celestin.huet@ign.fr

Résumé

Plusieurs modèles d'apprentissages profonds, comme le Frame Field Learning, ont été développés récemment pour détecter les contours de bâtiments sur images aériennes. Cependant, ils ne sont pas capables de fournir une estimation de la hauteur des bâtiments. De plus, de part le dévers, il peut exister un décalage en planimétrie entre la position détectée sur l'orthophoto et la position réelle. Par ailleurs, l'IGN dispose d'images orientées qui sont les images acquises par avion et dont leur vocation première est la production de la BD ORTHO®. L'intérêt de ces images est de voir chaque bâtiment en plusieurs fois depuis des points de vue différents, permettant ainsi des reconstructions 3D. Nous proposons ici un algorithme qui, en utilisant des inférences du Frame Field Learning sur ces images orientées, reconstitue les bords de bâtiments en trois dimensions. Le code est disponible sur https://github.com/IGNF/Samon_gouttieres

Mots clés : Détection de bâtiments, 3D, Frame Field learning

Abstract

Several deep learning models, such as Frame Field Learning, have been developed recently to detect buildings' outlines on aerial images. However, they are not able to predict a height estimation for these buildings. In addition, due to crossfall, a planimetric error between the detected position on the aerial image and the actual position may exist. Moreover, IGN owns oriented images which are images acquired from aircraft. Their primary purpose is to be used in the production of BD ORTHO®. Their interest is that each building is seen several times from different viewpoints. This multiplicity of viewpoints enables 3D reconstruction. We propose an algorithm which uses Frame Field Learning inferences outputs from oriented images to extract building outlines in 3D. The code is available here : https://github.com/IGNF/Samon_gouttieres

Keywords : Building detection, 3D, Frame Field Learning

1. Introduction

L'IGN est chargé de mettre à jour régulièrement la BD TOPO® (Institut national de l'information géographique et forestière, 2025b), une base de données vectorielles 3D décrivant le territoire français. Les bâtiments sont recensés dans une couche spécifique. Actuellement, lors d'une phase dite de restitution, des opérateurs visualisent les images aériennes en 3D et saisissent manuellement les bords de bâtiment. Cependant, cette phase est chronophage et il serait bénéfique de l'automatiser.

Récemment, plusieurs algorithmes d'apprentissages profonds ont été développés pour permettre la détection de bâtiments sous forme vectorielle comme le FFL (Frame Field Learning) (Girard et al., 2021), HiSup (Xu et al., 2023) ou Pix2Poly (Adimoolam et al., 2024). Toutefois, appliquer ces modèles directement sur les orthophotos a deux inconvénients. Le premier est l'impossibilité de déterminer la hauteur du bâtiment, ce qui est pourtant une information à renseigner dans la BD TOPO®. On peut envisager d'utiliser le Lidar HD (Institut national de l'information géographique et forestière, 2025c) mais aucune nouvelle campagne d'acquisition n'est prévue, donc c'est à éviter dans une perspective d'une mise à jour régulière de la donnée. On peut également utiliser le MNS (modèle numérique de sur-

face) de corrélation, mais il est très bruité. Le deuxième inconvénient est l'erreur planimétrique due au dévers. En effet, dans la BD ORTHO®, les images sont projetées sur un MNT (modèle numérique de terrain) qui ne modélise pas le sursol et donc en particulier les bâtiments. Ainsi, les bords de toit des bâtiments sur la BD ORTHO® peuvent être décalés de quelques mètres par rapport à la réalité. Cette erreur dépend de la hauteur du bâtiment et de l'angle de vue sur l'image orientée entre le bâtiment et le sommet de prise de vue, occasionnant un dévers pouvant atteindre jusqu'à 38% d'après les spécifications.

Il existe d'autres algorithmes qui reconstruisent les bâtiments en mesh à partir du Lidar : City3D (Huang et al., 2022), Point2Building (Liu et al., 2024). Cependant, pour la même raison qu'évoquée précédemment -pas d'autres campagnes d'acquisition prévues- ces méthodes sont à éviter pour une mise à jour régulière de la BD TOPO®.

La démarche que nous proposons est d'effectuer les inférences du FFL sur les images orientées. Vient ensuite un processus d'appariement de segments homologues, c'est-à-dire de segments issus de prédictions sur des images orientées différentes mais qui représentent le même bord de toit dans la réalité. Pour chaque groupe

de segments homologues, on effectue une intersection de plans dans l'espace pour obtenir la droite qui correspond au mieux à la position dans l'espace du bord de toit. Enfin, une dernière étape associe ces droites en polygones ou en mesh représentant chacun un bâtiment (figure 1).

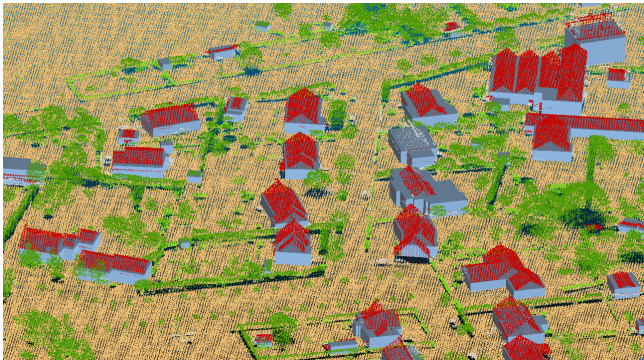


FIGURE 1: Résultat de la reconstruction des bâtiments (en bleu) superposés au Lidar HD.

2. État de l'art

2.1. Détection de contours de bâtiments

Une possibilité pour détecter les bâtiments sur les images aériennes est d'utiliser un modèle de segmentation sémantique. ICT-Net (Bodhiswatta et Charalambos, 2019) utilise une variante de U-Net. Cependant, le résultat est un raster qu'il faut ensuite vectoriser. Les bords de bâtiments obtenus sont très irréguliers à cause de cette structure en pixels. Le FFL (Girard et al., 2021) prédit une carte de prédiction de bâtiments, accompagnée d'une carte décrivant les orientations des angles et bords de bâtiments. Une deuxième étape permet d'en extraire les bâtiments sous forme de vecteurs. HiSup (Xu et al., 2023) a une démarche semblable : un modèle prédit une représentation qui contient une carte de segmentation des bâtiments, un champ d'attraction des bords de bâtiments, et les sommets concaves et convexes. Puis une deuxième étape en dérive les bâtiments.

Contrairement aux méthodes précédentes qui ont besoin de post-traitements importants après les inférences du modèle IA, Pix2Poly (Adimoolam et al., 2024) utilisant un modèle de transformers qui prédit directement les angles des bâtiments et les liaisons entre eux.

(Hattula et al., 2024) utilise un U-Net sur un MNS Lidar pour détecter le contour des bâtiments.

2.2. Reconstruction 3D de bâtiments

Une première méthode pour reconstruire les bâtiments en 3D est d'utiliser le lidar. La première étape est de récupérer les emplacements des bâtiments. (Amini et al., 2014) utilise un SVM pour classer des caractéristiques obtenues à partir du lidar et d'images aériennes. (Huang et al., 2022) nécessite en entrée les emprises des bâtiments. (Liu et al., 2024) a besoin d'un nuage de point limité à un seul bâtiment.

(Amini et al., 2014), (Chen et al., 2017), (Xu et al., 2020) et (Huang et al., 2022) recherchent ensuite des primitives (plans, cylindres) s'approchant au mieux des points en utilisant l'algorithme de Ransac ou ses variantes, puis une étape de post-traitement reconstruit les bâtiments à partir de ces plans. (Amini et al., 2014), (Xu et al., 2020) et (Barranquero et al., 2023), pour guider la reconstruction, utilisent des modèles de deep learning pour indiquer la forme du toit (plat, pointu, en croupe, en mansarde, en pignon). (Huang et al., 2023) et (Liu et al., 2024) utilisent des transformeurs pour déterminer la position des bords de pans de toits.

Il est également possible de partir d'images aériennes au lieu de nuages lidar. (Xu et al., 2020) et (Turki et al., 2025) calculent un nuage de point à partir d'une orthophoto et reviennent au cas où l'on a du lidar. (Gao et al., 2023) utilise un modèle qui détecte les contours des pans de toits sur une orthophoto vraie, puis utilise le MNS pour y projeter ces pans.

3. Données initiales

3.1. Les images orientées

Les campagnes d'acquisitions aériennes par l'IGN se font par département tous les trois ans. Chaque campagne respecte des spécifications de recouvrement entre les images de l'ordre de 70% en interbande et en intrabande. Ces spécifications assurent que tout point au sol est vu par plusieurs images. Elles varient en fonction de la nature du terrain. Le recouvrement dans une zone de montagne est plus élevé que dans une zone de plaine. Lors de ces acquisitions, la position de l'avion est enregistrée grâce à un GNSS (Global Navigation Satellite System) et son orientation au moyen d'une centrale inertielle. Pendant l'étape d'aérotriangulation, la position et l'orientation de chaque cliché sont affinées avec précision, ainsi que la focale de la caméra et les paramètres de distorsion. Les images sont ensuite corrigées de la distorsion. Elles ont une dimension de l'ordre de 26460x17004 pixels. Les images orientées servent alors à la production de la BD ORTHO® (Institut national de l'information géographique et forestière, 2025a). Elles ont donc des caractéristiques radiométriques et géométriques très proches.

3.2. Le modèle numérique de terrain

À plusieurs reprises dans l'algorithme, il est nécessaire de projeter des objets de la géométrie image en géométrie terrain. Pour cela, on utilise le RGE ALTI® (Institut national de l'information géographique et forestière, 2025d), un MNT de résolution 1 mètre. Contrairement à un MNS, il a l'avantage d'être beaucoup moins bruité.

3.3. Les prédictions de contours de bâtiments

Dans cette étude, le FFL (Girard et al., 2021) a été utilisé pour faire des prédictions de bâtiments. Il fournit une carte de segmentation ainsi qu'un champ de vecteurs

indiquant l'orientation locale des bâtiments. Un post-traitement utilise ensuite ces deux sorties pour convertir la carte de segmentation en format vectoriel. Le modèle utilisé a été ajusté sur un jeu de données constitué de 117 km² de BD ORTHO[®] répartis sur la France entière. On y trouve tout type de paysage et nature d'urbanisation (centres urbains, zones pavillonnaires, zones industrielles et commerciales). Les bâtiments ont été annotés à la main, au niveau du toit. Le jeu de données est donc de bonne qualité.

D'autres modèles ont été testés mais n'ont pas été retenus. Le modèle Pix2Poly (Adimoolam et al., 2024) utilise l'architecture d'un Transformer (Dosovitskiy et al., 2021) : le modèle prédit à chaque itération un nouveau point délimitant un bâtiment, en s'aidant des points qu'il a prédit précédemment. Il nécessite donc de faire une inférence pour obtenir un point, ce qui le rend plus lent que le FFL qui prédit en une seule inférence la carte de segmentation et le champ de vecteurs. Pour comparaison, il faut environ 8 minutes au FFL pour faire une inférence sur une image orientée et 6 heures à Pix2Poly.

Le modèle HiSup (Xu et al., 2023) fournit des contours de bâtiments trop irréguliers pour notre méthode.

Les images orientées et la BD ORTHO[®] étant très proches radiométriquement et géométriquement, le FFL n'a pas de difficultés pour effectuer des inférences de bonnes qualités sur les images orientées. Les inférences étant faites sur chaque image orientée, un même bâtiment est vu sur plusieurs images (Figure 2).



FIGURE 2: Un même bâtiment vu par plusieurs images orientées avec leur prédiction FFL associée.

4. Méthode

La démarche générale de l'algorithme est la suivante :

- lisser les prédictions IA
- apparier les bâtiments homologues
- estimer de la hauteur des bâtiments
- apparier les bords de toit homologues
- calculer la position en 3D des bords de toit
- réunir les bords de toit en bâtiments

4.1. Lissage des prédictions IA

Dans une prédiction IA (l'ensemble des polygones prédits par le FFL sur une image orientée), il est fréquent qu'un bord de toit soit divisé en plusieurs segments avec des angles très faibles entre les segments.

Pour la suite du traitement, il est inutile de conserver autant de segments. Deux segments consécutifs sont fusionnés si leur produit scalaire est supérieur à 0.99. On dispose ainsi de géométries où l'on a globalement sur chaque prédiction IA un bord de toit représenté par un seul segment (Figure 3).

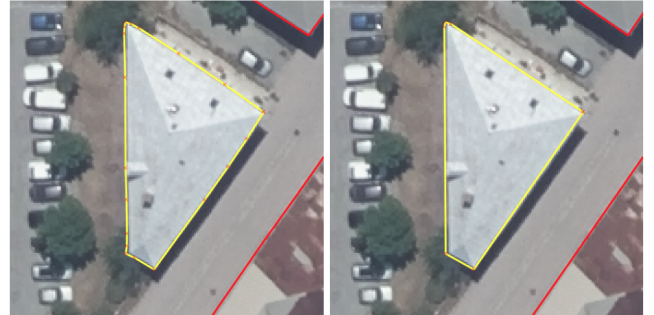


FIGURE 3: Les segments qui représentent un même bord de toit sont fusionnés.

4.2. Appariement des bâtiments homologues

L'objectif est d'associer les bâtiments homologues, c'est-à-dire les bâtiments issus de prédictions IA différentes, mais représentant dans la réalité la même construction.

On connaît la position, l'orientation et la focale de chaque image orientée. On peut donc projeter les bâtiments de chaque prédiction IA en géométrie terrain sur le MNT (figure 4).



FIGURE 4: Prédiction FFL projetées sur le MNT. Il y a une couleur par prédiction FFL.

Ensuite, dans chaque paire de bâtiments IA, chaque bâtiment est apparié avec le bâtiment de l'autre prédiction IA avec lequel il partage la plus grande surface. On obtient ainsi un graphe que l'on décompose en composantes connexes (figure 5).

4.3. Estimation de la hauteur des bâtiments

À cause du dévers, la distance entre les segments homologues projetés sur le MNT peut atteindre plusieurs mètres. Dans les zones avec une forte densité de bâti,



FIGURE 5: Appariement des bâtiments homologues. Il y a une couleur par groupe de bâtiments homologues.

cela perturbe l'appariement de segments (figure 6). L'objectif est donc d'estimer la hauteur des bâtiments pour projeter les prédictions IA sur MNT+hauteur estimée du bâtiment.



FIGURE 6: Projection sur le MNT des prédictions FFL. Il y a une couleur par prédiction FFL. Cette illustration laisse présager des difficultés pour appairer les segments.

On constate que les géométries de plusieurs bâtiments homologues sont semblables et que l'on peut utiliser cette propriété pour estimer la hauteur du bâtiment. On calcule un score de ressemblance entre deux prédictions IA qui tient compte de trois composantes :

- ratio de la surface
- différence du nombre de sommets
- erreur géométrique

En théorie, pour deux bâtiments homologues, les deux sommets de prise de vue des images et les deux barycentres des prédictions IA appartiennent au même plan (figure 7). L'erreur géométrique est la distance du barycentre d'une prédiction IA par rapport au plan

formé par les sommets de prise de vue des deux images et le barycentre de l'autre prédiction IA.

En appliquant le théorème de Thalès, on obtient une estimation de la hauteur du bâtiment.

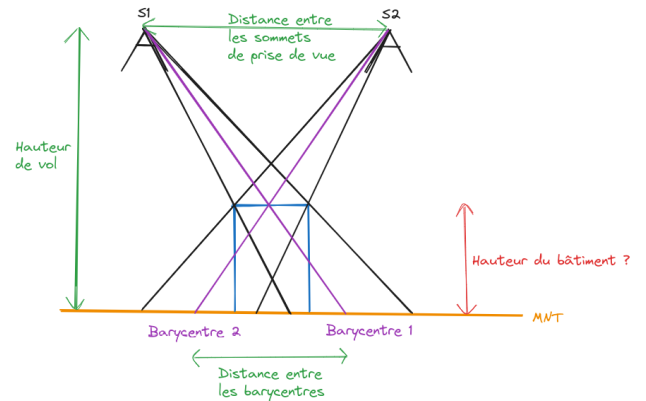


FIGURE 7: Schéma pour le calcul d'une estimation de la hauteur d'un bâtiment.

On projette ensuite les prédictions IA sur la somme du MNT et de la hauteur estimée (figure 8).



FIGURE 8: Projection des prédictions FFL sur MNT+hauteur estimée.

Cette méthode est rapide à mettre en œuvre. Cependant elle peut échouer dès lors que la géométrie des bâtiments devient complexe. C'est pourquoi nous proposons une deuxième méthode plus robuste mais également plus lente.

4.4. Deuxième méthode d'estimation de la hauteur des bâtiments

À l'intérieur de chaque groupe de bâtiments homologues, on sélectionne celui qui est le plus proche du nadir dans son image orientée. On sélectionne aléatoirement un point de référence qui est à deux mètres du

bord du bâtiment. On cherche par corrélation sur les autres images orientées les points homologues au point de référence. Pour chaque image orientée, on peut alors tracer une droite dans l'espace passant par le sommet de prise de vue et par le point de corrélation. En appliquant la méthode des moindres carrés, on peut alors trouver le point 3D qui intersecte au mieux ces droites. Si le calcul échoue, principalement à cause d'un échec dans la corrélation, alors on recommence le calcul avec un autre point de référence.

Dans cette méthode, on évite de prendre un point de référence qui est sur le bord de toit. En effet, la corrélation serait alors plus compliquée car, suivant le point de vue, on pourrait voir sur certaines vignettes de corrélation des façades, tandis que sur d'autres ces façades seraient masquées par le dévers. Si le toit est en pente, prendre comme point de référence un point à deux mètres du bord peut impliquer une erreur de quelques dizaines de centimètres, mais cela reste une approximation acceptable pour la suite du traitement.

On ne garde le résultat de l'estimation de la hauteur que si dans les moindres carrés, on utilise au moins trois images orientées. En effet, si seulement deux images orientées sont utilisées, les résultats sont rarement fiables (de l'ordre de la dizaine de mètres d'erreur).

4.5. Appariement des bords de toit homologues

On cherche à présent à appairer les segments homologues. Deux segments sont appariés si :

- ils appartiennent au même groupe de bâtiments homologues
- ils ont la même orientation : leur produit scalaire est supérieur à 0.98
- ils sont proches : la distance entre les deux centres de segment est inférieure à 1.5 mètre.

On obtient ainsi un graphe que l'on décompose en composantes connexes où chacune est constituée de segments homologues (figure 9).



FIGURE 9: Appariement des prédictions FFL par segment. Il y a une couleur par groupe de segments homologues.

4.6. Calcul du bord de toit en 3D

À ce stade de l'algorithme, on dispose de groupes de segments homologues, représentant donc le même bord de toit dans la réalité, mais sur des images orientées différentes. Pour chaque segment, on peut tracer un plan qui passe par le sommet de prise de vue de son

image orientée et par ce segment. Grâce à un calcul de moindres carrés, on détermine la droite qui passe au plus près des plans. Si le résidu normalisé d'un segment est supérieur à 2, alors on retire ce segment et on recommence le calcul d'intersection.

4.7. Fermeture des bâtiments

On dispose de bords de toits positionnés dans l'espace. La dernière étape est de les transformer en polygones avec un polygone par bâtiment. Pour cela, pour chaque groupe de bâtiments homologues, on choisit comme référence les bâtiments issus de l'image orientée la plus proche du nadir. Chaque segment des bâtiments référence est projeté sur l'altitude déterminée dans le calcul d'intersection de plans. Si le calcul n'a pas abouti pour un groupe de segments, on utilise l'altitude des segments voisins, et à défaut l'altitude moyenne des segments constituant les bâtiments.

Enfin, pour chaque couple de bords de toit voisins, on calcule la pseudo-intersection et on ajuste en conséquence leurs extrémités.

5. Zones d'études

L'algorithme a été testé sur trois zones d'étude.

5.1. Anjou

Cette zone représente un territoire rural de 661 km² dans le Maine-et-Loire entre la Loire au nord et Doué-la-Fontaine et Vihiers au sud. On y trouve des centres-villes de petits villages et petites villes, des lotissements et des zones d'activités. 346 images orientées ont été utilisées. Elles ont été acquises en juillet 2022.

5.2. Marseille

La ville de Marseille constitue la deuxième zone d'étude. Le cœur de la ville est composée d'immeubles hauts et de rues étroites. Les alentours sont composés de lotissements et de zones d'activités. 305 images orientées, acquises en juin 2023, ont été utilisées. La zone d'étude fait 119 km².

5.3. Mont-Dauphin

Cette troisième zone d'étude est beaucoup plus petite : 4 km². Elle est centrée sur la forteresse de Mont-Dauphin et la zone d'activité de Guillestre. Il s'agit d'une zone dans les Alpes avec un recouvrement entre les clichés bien supérieur à la zone sur l'Anjou. 31 images orientées de juillet 2022 ont été utilisées.

6. Résultats

Les résultats sur les trois zones d'études sont disponibles sur la page Github du projet.

6.1. Donnée de référence

La donnée utilisée pour contrôler les résultats est le Lidar HD produit par l'IGN (Institut national de l'information géographique et forestière, 2025c). Le lidar est classé automatiquement en différentes classes dont le bâti. Les données lidar n'ont pas été acquises au même moment que les images orientées. Pour le jeu de données Anjou, elles ont été acquises entre janvier et avril 2024 soit environ 18 mois après les images orientées. Pour le jeu de données Marseille, elles ont été acquises entre juin et juillet 2021, soit deux ans avant les images orientées. Pour Mont-Dauphin, elles l'ont été en juillet 2021, soit un an avant.

Pour chaque jeu de données, deux méthodes ont été utilisées : la méthode que nous proposons (Samon) et les inférences de FFL sur la BD Ortho. De plus, pour mesurer la précision géométrique, nous avons également comparé notre méthode aux bâtiments saisis manuellement par stéréo-restitution dans la BD TOPO®. Ceux-ci sont identifiés par l'attribut "photogrammétrie" dans le champs "ACQU_PLANI" et représentent un peu moins de 10 % des bâtiments de la base de données.

6.2. Résultats visuels

Dans les trois zones d'études, la très grande majorité des bâtiments est restituée (figures 10, 11 et 12). L'algorithme a un peu plus de difficultés dans la zone de Marseille sur certains immeubles. Du point de vue de la précision planimétrique, les limites de bâtiments sont cohérentes avec la classification du Lidar HD. Du point de vue altimétrique, les résultats sont un peu moins précis qu'en planimétrie mais restent convenables. Cependant, on ne parvient pas toujours à restituer les différences d'altitude entre deux immeubles adjacents partageant la même façade. Le bâti dans les cours d'immeubles laisse également à désirer.

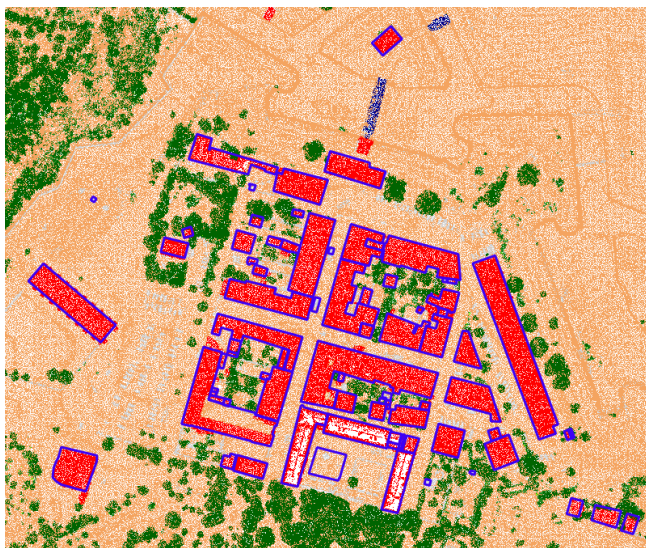


FIGURE 10: Résultat sur Mont-Dauphin. Les résultats de l'algorithme sont en bleu, superposés au Lidar HD

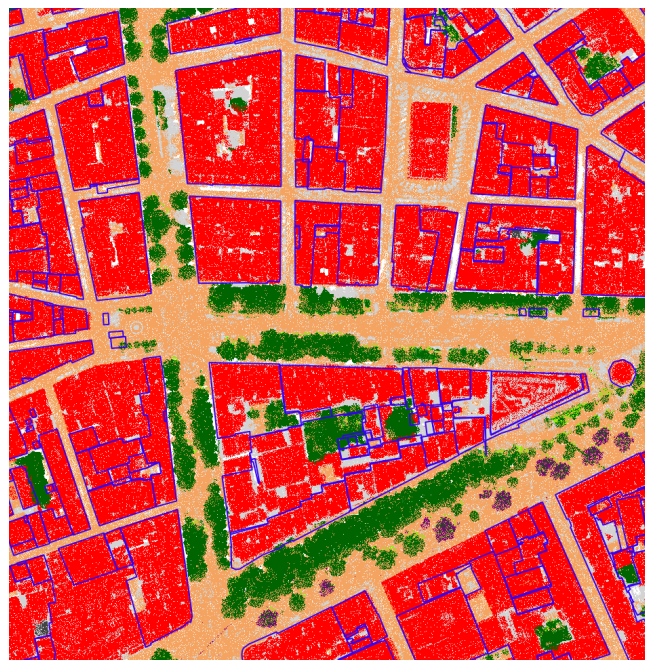


FIGURE 11: Résultat sur Marseille.

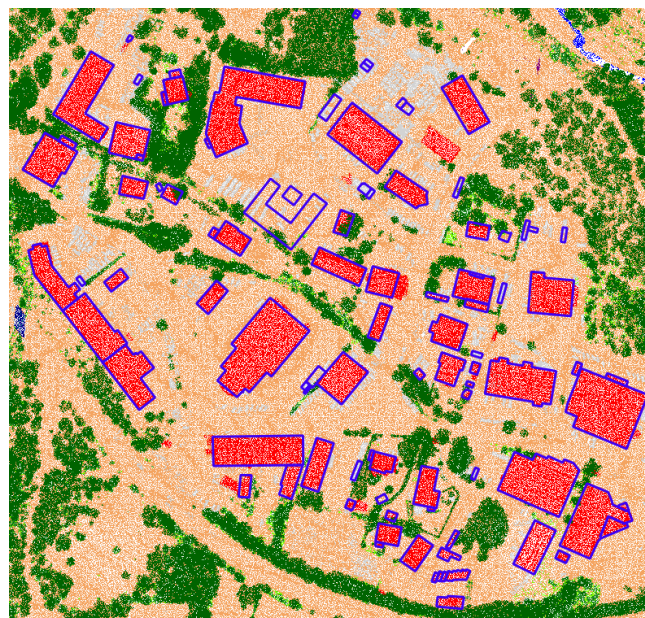


FIGURE 12: Résultat sur la zone d'activité de Mont-Dauphin. Au centre de l'image, le bâtiment en U a été construit entre les acquisitions Lidar et aériennes.

L'algorithme permet également de bien corriger les dévers. La BD ORTHO® est produite en projetant les images orientées sur un MNT qui ne tient donc pas compte du sursol. Les toits de bâtiments peuvent ainsi être décalés de plus d'un mètre par rapport à leur position réelle. Faire des inférences de FFL directement sur la BD ORTHO® implique une erreur en planimétrie à cause de ce dévers. La figure 13 montre que notre algorithme est insensible au dévers.

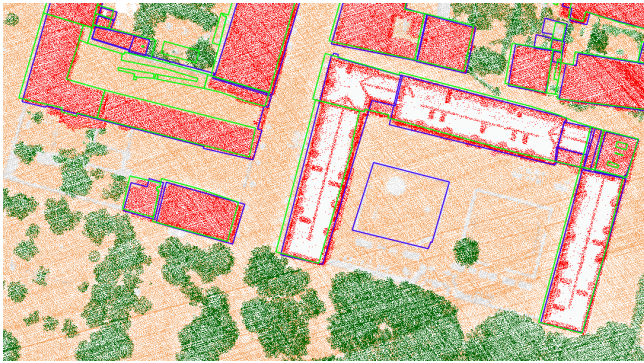


FIGURE 13: En bleu le résultat de Samon. En vert les prédictions du FFL sur la BD ORTHO®.

6.3. Exhaustivité des détections

Pour calculer la capacité de l'algorithme à détecter tous les bâtiments, nous avons calculé l'Intersection over Union (IoU) entre les résultats évalués et le Lidar HD. Il s'agit du rapport de l'intersection des prédictions et de la vérité terrain par l'union des prédictions et de la vérité terrain. Dans notre cas, on utilise comme unité le point lidar. La vérité terrain est constituée des points lidar classés comme bâti. Les points non classés du Lidar HD ont été retirés au préalable car ils peuvent représenter tout autant du bâti que du sol, des murets ou des ruines.

Jeu de données	Samon	BD Ortho
Anjou	0.7442	0.7309
Marseille	0.7583	0.7528
Mont-Dauphin	0.7760	0.7259

TABLE 1: Calcul de l'IoU

La table 1 montre que l'IoU reste du même ordre de grandeur pour les trois zones de tests et donc que notre algorithme reste robuste dans les différentes configurations de zones bâties : zones rurales, zones urbaines denses, zones d'activité. Parmi les erreurs, une partie est imputable aux données : l'écart temporel entre les acquisitions lidar et images, ainsi que de potentielles erreurs de classification du lidar. D'autres sont dues à notre algorithme : certaines extensions de bâtiments, voire des bâtiments entiers, surtout en centre-ville, ne sont pas détectés. Des conteneurs ou abris de jardin sont parfois identifiés comme bâtiments. Enfin, notre algorithme n'est pas toujours assez précis sur les bords de bâtiments.

Comparé aux inférences du FFL sur la BD ORTHO®, notre algorithme est systématiquement meilleur. Un premier élément qui l'explique est l'absence de correction du dévers dans le cadre des inférences sur la BD ORTHO®. De plus, lorsque le FFL omet de détecter un bâtiment, il n'a pas de seconde chance dans le cadre de la BD ORTHO®. Tandis que si un bâtiment n'a pas été détecté sur une image orientée, il est possible qu'il le soit sur une autre image. Le jeu sur Mont-Dauphin dispose d'une plus grande densité d'images orientées et c'est sur ce jeu que l'écart entre notre algorithme et les inférences sur la BD ORTHO® est le plus important.

6.4. Précision géométrique des contours des bâtiments

L'objectif est de calculer la distance des bords de bâtiments trouvés par notre algorithme à une vérité terrain, ici le Lidar HD. Pour chaque bord de bâtiment, on identifie les cinq points du Lidar HD les plus proches. À partir de ces points, on calcule la distance moyenne en altimétrie, en planimétrie et dans l'espace. Cette évaluation a cependant une limite. Dans le cas où la détection se trouve dans le plan du toit du bâtiment, les points lidar les plus proches ne sont pas ceux du bord de toit, mais ceux du toit, pouvant ainsi fausser les métriques.

Les inférences du FFL directement sur la BD Ortho ne donnent pas de valeur altimétrique. On ne peut donc les évaluer que sur la précision planimétrique. Il n'est pas possible de calculer la distance par rapport aux points lidar projetés au sol. En effet, dès que la prédiction sera à l'intérieur d'un bâtiment, la distance sera très faible car calculée par rapport aux points du toit du bâtiment. Pour éviter ce problème, on restreint les points lidar à ceux qui ont servi pour calculer les distances avec notre algorithme.

La table 2 montre la distance entre les reconstructions et le lidar. Notre algorithme réduit par deux les erreurs en planimétrie par rapport aux inférences sur la BD ORTHO®. Cela traduit une bien meilleure correction du dévers.

En revanche, la densité de recouvrement d'images orientées ne semble pas avoir une influence notable : les erreurs sont semblables entre le jeu de données sur l'Anjou et celui de Mont-Dauphin.

Enfin, comparé aux bâtiments saisis par stéréorestitution de la BD TOPO®, notre algorithme fournit des résultats équivalents sur l'Anjou et Mont-Dauphin, et légèrement inférieurs à Marseille, tout en étant entièrement automatisé.

7. Conclusion

Nous avons proposé un algorithme qui permet à la fois de détecter les bâtiments en 3D et de reconstruire en LOD 1 les bâtiments. Il permet également de corriger le dévers. Il reste robuste dans toutes les configurations de bâti.

	Anjou			Marseille			Mont-Dauphin		
	Samon	BD Ortho	BD Topo	Samon	BD Ortho	BD Topo	Samon	BD Ortho	BD Topo
Altimétrie	0.32	X	0.33	0.37	X	0.29	0.31	X	0.38
Planimétrie	0.28	0.51	0.27	0.36	0.67	0.25	0.28	0.65	0.30
Distance	0.48	X	0.48	0.59	X	0.43	0.46	X	0.56

TABLE 2: Distance absolue moyenne (en mètres) entre les reconstructions et les cinq points lidar classés bâti les plus proches

Pour aller plus loin, on pourrait envisager un modèle semblable qui soit entraîné à détecter le faîtage des toitures et autres lignes caractéristiques. D’une manière assez proche à l’algorithme proposé, on pourrait les reconstituer en 3D et ainsi atteindre le LoD 2 pour la reconstruction du bâtiment. On obtiendrait de la même manière une valeur pour l’altitude maximale du bâtiment, ce qui est également une valeur à renseigner dans la BD TOPO®.

Références

- Adimoolam, Y. K., Poullis, C., Averkiou, M., 2024. Pix2poly : A sequence prediction method for end-to-end polygonal building footprint extraction from remote sensing imagery. URL <https://arxiv.org/abs/2412.07899>
- Amini, H., Pahlavani, P., Karimi, R., 2014. 3d reconstruction of buildings with gabled and hipped structures using lidar data. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XL-2/W3, 47–52. URL <https://isprs-archives.copernicus.org/articles/XL-2-W3/47/2014/>
- Barranquero, M., Olmedo, A., Gómez, J., Tayebi, A., Hellín, C. J., Saez de Adana, F., 2023. Automatic 3d building reconstruction from openstreetmap and lidar using convolutional neural networks. Sensors 23 (5). URL <https://www.mdpi.com/1424-8220/23/5/2444>
- Bodhiswatta, C., Charalambos, P., 2019. On building classification from remote sensor imagery using deep neural networks and the relation between classification and reconstruction accuracy using border localization as proxy.
- Chen, D., Wang, R., Peethambaran, J., 2017. Topologically aware building rooftop reconstruction from airborne laser scanning point clouds. IEEE Transactions on Geoscience and Remote Sensing 55 (12), 7032–7052.
- Dosovitskiy, A., Beyler, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., Hounsby, N., 2021. An image is worth 16x16 words : Transformers for image recognition at scale. ICLR.
- Gao, W., Peters, R., Stoter, J., 2023. Unsupervised roofline extraction from true orthophotos for lod2 building model reconstruction. URL <https://arxiv.org/abs/2310.01067>
- Girard, N., Smirnov, D., Solomon, J., Tarabalka, Y., June 2021. Polygonal building extraction by frame field learning. Dans : Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 5891–5900.
- Hattula, E., Zhu, L., Raninen, J., 2024. Building extraction in urban and rural areas with aerial and lidar dsm. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences X-4/W4-2024, 73–79. URL <https://isprs-annals.copernicus.org/articles/X-4-W4-2024/73/2024/>
- Huang, J., Stoter, J., Peters, R., Nan, L., 2022. City3d : Large-scale building reconstruction from airborne lidar point clouds. Remote Sensing 14 (9).
- Huang, S., Wang, R., Guo, B., Yang, H., 2023. Pbwr : Parametric building wireframe reconstruction from aerial lidar point clouds. URL <https://arxiv.org/abs/2311.12062>
- Institut national de l’information géographique et forestière, 2025a. BD ORTHO®. Consulté le 28 juillet 2025. URL <https://geoservices.ign.fr/bdortho>
- Institut national de l’information géographique et forestière, 2025b. BD TOPO®. Consulté le 28 juillet 2025. URL <https://geoservices.ign.fr/bdtopo>
- Institut national de l’information géographique et forestière, 2025c. Lidar HD. Consulté le 28 juillet 2025. URL <https://geoservices.ign.fr/lidarhd>
- Institut national de l’information géographique et forestière, 2025d. RGE ALTI®. Consulté le 31 juillet 2025. URL <https://geoservices.ign.fr/rgealti>
- Liu, Y., Obukhov, A., Wegner, J. D., Schindler, K., 2024. Point2building : Reconstructing buildings from airborne lidar point clouds. URL <https://arxiv.org/abs/2403.02136>
- Turki, S., Panangian, D., Chaabouni-Chouayakh, H., Bittner, K., 2025. Aim2pc : Aerial image to 3d building point cloud reconstruction. URL <https://arxiv.org/abs/2503.18527>
- Xu, B., Xu, J., Xue, N., Xia, G.-S., 2023. Hisup : Accurate polygonal mapping of buildings in satellite imagery with hierarchical supervision. ISPRS Journal of Photogrammetry and Remote Sensing 198, 284–296.
- Xu, B., Zhang, X., Li, Z., Leotta, M. J., Chang, S.-F., Shan, J., 2020. Deep learning guided building reconstruction from satellite imagery-derived point clouds. CoRR abs/2005.09223. URL <https://arxiv.org/abs/2005.09223>