

Utilisation des images satellites pour améliorer le recensement des logements à Mayotte

Maëlys Bernard
Insee
maelys.bernard@insee.fr

Raya Berova
Insee
raya.berova@insee.fr

Thomas Faria
Insee
thomas.faria@insee.fr

2025-07-10

Abstract Malgré leur disponibilité de plus en plus accrue, les images satellites très haute résolution sont jusqu'ici très peu utilisées au sein de la statistique publique. La manipulation de ces données non conventionnelles n'appartient pas encore aux pratiques usuelles des statisticiens publics, bien qu'elles offrent des opportunités significatives pour améliorer ou créer de nouveaux indicateurs statistiques. Cet article vise justement à démontrer qu'il est possible d'intégrer ces données satellitaires dans les processus statistiques afin d'assurer leur qualité. Mayotte constitue un cas emblématique où les difficultés d'accès au terrain et la forte dynamique de développement urbain complexifient l'organisation des enquêtes cartographiques qui permettent de repérer les logements en amont du recensement. Afin d'orienter les travaux de contrôle de la qualité de ces enquêtes, nous avons exploré l'utilisation des images satellitaires très haute résolution Pléiades (0,5 m de résolution spatiale), en combinaison avec des méthodes d'apprentissage profond, spécifiquement via des modèles de segmentation sémantique. La méthodologie développée consiste en l'entraînement d'un modèle SegFormer, fondé sur une architecture de type Transformer, largement reconnue dans la littérature puisqu'elle est à la base des modèles GPT. Le modèle a été entraîné grâce aux annotations issues du projet CoSIA de l'IGN, qui fournit une référence précise de la couverture du sol. À partir des prédictions du modèle, il est possible de produire des estimations précises des surfaces bâties sur un territoire donné, permettant ainsi d'analyser efficacement leur évolution dans le temps. Les résultats obtenus démontrent une performance élevée dans la détection automatisée des évolutions urbaines dans les DROM. Ces résultats sont disponibles sous la forme d'un tableau de bord interactif permettant aux équipes opérationnelles de cibler efficacement les zones prioritaires pour la réalisation d'enquêtes cartographiques de contrôle sur le terrain.

Après avoir décrit la méthode de prédiction de surface du bâti sur les images satellites, cet article présentera un cas pratique d'utilisation de ces prédictions à travers le contrôle

qualité des enquêtes cartographiques de la population à Mayotte. Depuis 2021, le recensement à Mayotte fonctionne sur un cycle quinquennal comme en métropole et dans les autres DOM. En amont de la collecte auprès des habitants, une enquête cartographique permet de repérer les logements. Le résultat de l'enquête cartographique détermine la base d'immeubles au sein de laquelle est sélectionné l'échantillon¹ de l'enquête annuelle de recensement. Ainsi, si l'enquête cartographique ne couvre pas bien l'intégralité du bâti du territoire à enquêter, les populations et nombre de logements déduits de l'enquête annuelle de recensement seront sous-estimés du fait de ce défaut de couverture. Les données satellites permettent de contrôler la qualité des enquêtes cartographiques en comparant l'évolution de la couverture de bâti avec celle du nombre de logements observé suite aux enquêtes cartographiques au sein de chaque îlot² à enquêter une année donnée. Les îlots sont répartis en cinq groupes de rotation correspondant à l'année à laquelle a lieu l'enquête annuelle de recensement. Afin d'assurer la qualité des enquêtes cartographiques, un deuxième passage sur le terrain a été planifié dans les grandes communes mahoraises pour certains îlots identifiés comme les plus prioritaires. Cet ordre de priorité a été défini par une méthode mobilisant les images satellites pour repérer les plus susceptibles d'être concernés par des manques potentiels dans les enquêtes cartographiques. L'enquête cartographique corrective ciblant ces îlots a permis d'assurer une meilleure couverture de ces enquêtes et d'éviter un risque de sous-estimation des logements.

Table des matières

Introduction	3
1 Contexte	3
2 Méthodologie	4
3 Résultats et inférence	5
4 Cas d'usage	6
5 Pistes et évolutions	7
Bibliographie	7

Introduction

blabla

1 Contexte

blabla

2 Méthodologie

1. Segmentation sémantique (def, méthode choisie (CV pour seg sem auto), historique IA (ancien to new model SAM))
2. Données **Raya** 2.1 Images satellites (def, produits, acquisition/livraison (partenariat IGN a promouvoir)) 2.2 Annotations (nerf de la guerre) : on a fait tout automatique, dire ce que ca coute pour belles annot. RIL (Insee), BDTPOPO (better), finally
3. Evaluation **Raya** IOU, Loss, zone de test manuel (calculer la taille des zones de test)
inférence (parler moyenne mobile) -> mise à dispo d'une API

3 Résultats et inférence

1. Métriques et images **Raya**

2. Diffusion 2.1 Webapp : pour enquêteurs 2.2 Fichiers parquet : pour statisticiens

IV - Cas d'usage (DMTR)

4 Cas d'usage

DMTR

5 Pistes et évolutions

blabla

Bibliographie