

Rapport INDG - Relevé Automatique Réseaux

Antoine Carreaud, Adrien Gressin

22 décembre 2022

1 Administratif

Suite à la décision de l'INDG d'un premier financement du projet RAR, avec une focalisation du champ de recherche sur la comparaison des méthodes d'acquisition sur différentes thématiques relevantes de l'INDG, de la constitution d'une base de données d'apprentissage et d'une étude préliminaire sur les méthodes d'extraction automatique des objets d'intérêts, nous avons proposé d'adapter le programme de notre projet tel que décrit ici.

Le périmètre de cette version réduite du projet RAR correspond à la 1ère année du projet initiale et sera composé des trois phases suivantes :

- WP1 : Évaluation des méthodes d'acquisition et choix d'une méthode de relevé adéquate ;
- WP2 : Constitution des données d'apprentissage en vue de l'entraînement d'une méthode automatique ;
- WP3 : Étude préliminaire sur les méthodes d'extraction automatique des objets d'intérêts à partir des données du WP2.

Le détail des objectifs et travaux effectués sur ces trois phases est largement détaillé en Section 2.1, 2.2 et 2.3. Le planning de ces différents WPs avait été envisagé comme indiqué sur le Tableau 1 en 12 mois.

Mois	01	02	03	04	05	06	07	08	09	10	11	12
WP1												
WP2												
WP3												

TABEAU 1 – Planning prévisionnel des différents WP du projet RAR.

Malgrès une version réduite du projet, nous nous sommes efforcé de garder des contacts privilégiés avec les différents partenaires du projet. En effet une collaboration avec Dr. Mathieu Salzmann et un de ses doctorants (Lis Krzysztof Maciej) de l'EPFL a été menée afin de nous aiguiller sur le WP3. Concernant le rapprochement avec le STDL, nous avons pris contact avec la FHNW et l'UNIGE, et nous les avons conviés à notre premier COPROJ. Mais compte tenu des changements de stratégie du STDL par rapport aux écoles partenaires, nous n'avons pas continuer ces collaborations dans le cadre de ce projet. Cependant, comme on le décrira dans la partie Perspectives (3.3), nous collaborons étroitement avec l'UNIGE sur la suite de ce projet. Par ailleurs nous remercions Fabio Mariani des SIG qui a suivi de très prêt ce projet avec beaucoup d'intérêt.

2 Différents WP

Le projet RAR (Relevé automatique des réseaux) a pour objectif la mise à jour rapide et à moindre coût des réseaux d'eau potable, de gaz, etc., (en raison du financement amoindri du projet, seul le réseau d'eau sera traité lors de ce petit projet) par relevé lors des fouilles permettant d'avoir accès aux éléments des réseaux. Il s'intéresse à l'automatisation de la reconnaissance d'objets, grâce à des techniques d'apprentissage automatique (apprentissage profond). Ce point sera abordé rapidement dans le WP3 (faute de budget). En amont de la reconnaissance automatique, nous avons évalué différentes méthodes d'acquisition lors du WP1 et constituer un jeu de données d'entraînement lors du WP2.

2.1 WP1 - Évaluation des méthodes d'acquisition et choix d'une méthode de relevé adéquate

2.1.1 Méthode d'acquisition actuelle

La méthode actuelle de mise à jour du cadastre sous-terrain de la ville de Genève se base sur un relevé GNSS avec deux opérateurs. L'un relève avec une canne GNSS les conduites, l'autre reporte dans la base de données sur une tablette le type de conduites, diamètres et les caractéristiques nécessaires à la mise à jour. Cette méthode est déjà performante, mais il est possible d'améliorer la partie reconnaissance des objets à cartographier qui pourrait-être automatisée. De plus, les relevés GNSS peuvent avoir des lacunes dans les zones urbaines denses, et ne donnent pas d'informations sur la géométrie des pièces par rapport à d'autres solutions envisagées pour ce projet.

2.1.2 Définition des besoins

Comme décrit dans le cadre de ce projet, la méthode d'acquisition retenue devra permettre : d'une part l'identification automatique des objets du cadastre souterrain et d'autre part de localiser ces objets avec une précision compatible avec la base de données du cadastre souterrain, le tout de manière simple, rapide et peu onéreuse. L'étape de reconnaissance automatique des objets imposera donc une résolution d'image et/ou une densité de points 3D suffisante. De manière générale, en photo interprétation, on considère qu'un objet est reconnaissable sur une image s'il a une taille d'une vingtaine de pixels sur l'image. Dans le cas de cette étude, les objets les plus petits peuvent avoir une taille de quelques centimètres, il faudra garantir une résolution d'image sub-centimétrique. De manière analogue, on peut appliquer le même raisonnement sur l'espacement moyen entre les points d'un nuage de points 3D, qui devra être en dessous du centimètre. Par ailleurs, le cadastre sous-terrain doit permettre de retrouver les éléments du sous-sol lors d'une fouille future, ainsi que de calculer des écoulements sur le réseau hydraulique. Ce dernier devra être mis à jour avec des données d'une classe de précision spécifique de 10 cm. L'étape d'acquisition devrait être au moins aussi simple et rapide que la méthode actuelle, décrite dans la partie précédente et ne devrait pas générer de surplus de travail important au bureau après la journée d'acquisition. De plus, la quantité de données générées, ainsi que les temps et puissance de calcul nécessaires, ne devrait pas impacter le coût global de la méthode. Pour récapituler, les critères pour le choix de la solution retenue sont :

- Une précision absolue (géoréférencement dans un système de coordonnées global) de 10 cm ;
- Une précision relative du système centimétrique (niveau de bruit faible pour distinguer correctement la géométrie des objets à détecter) ;
- Un niveau de détail permettant de distinguer les différents types d'objets composant le réseau sous-terrain ;
- Acquisition simple et rapide ;
- Un système robuste (travail sur des chantiers) ;
- Un prix abordable ;
- Simplicité et rapidité des traitements ;
- Quantité de données générées (données brutes et données post-traitées).

En utilisant ces critères définis par rapport aux besoins du projet, nous pourrions comparer différentes solutions (matériels + logiciels) afin d'aider à la décision.

2.1.3 Chantiers tests

Afin de tester différents matériels et pouvoir les comparer, nous avons fait quelques acquisitions sur des chantiers réels. Les différentes acquisitions de données se sont déroulées du 20 au 22 janvier et du 3 au 5 novembre 2020 dans la région de Genève sur 19 fouilles différentes. Pour chacune de ces fouilles, plusieurs méthodes de mesure ont été utilisées en vue de la numérisation des objets souhaités. Les systèmes d'acquisition ont été employés sur les chantiers listés ci-dessous, une carte des différents chantiers visités est visible sur la Figure 1.

- Jussy (3 fouilles),
- Athenaz,
- Lignon,
- Le tram à Bernex,
- Plan les Ouates,
- Pictet,
- Avenue de Champel,
- Carl Vogt (3 fouilles),
- Mont-Blanc (2 fouilles),
- Jeunes,
- Rambossons (2 fouilles),
- Tronchin (2 fouilles).

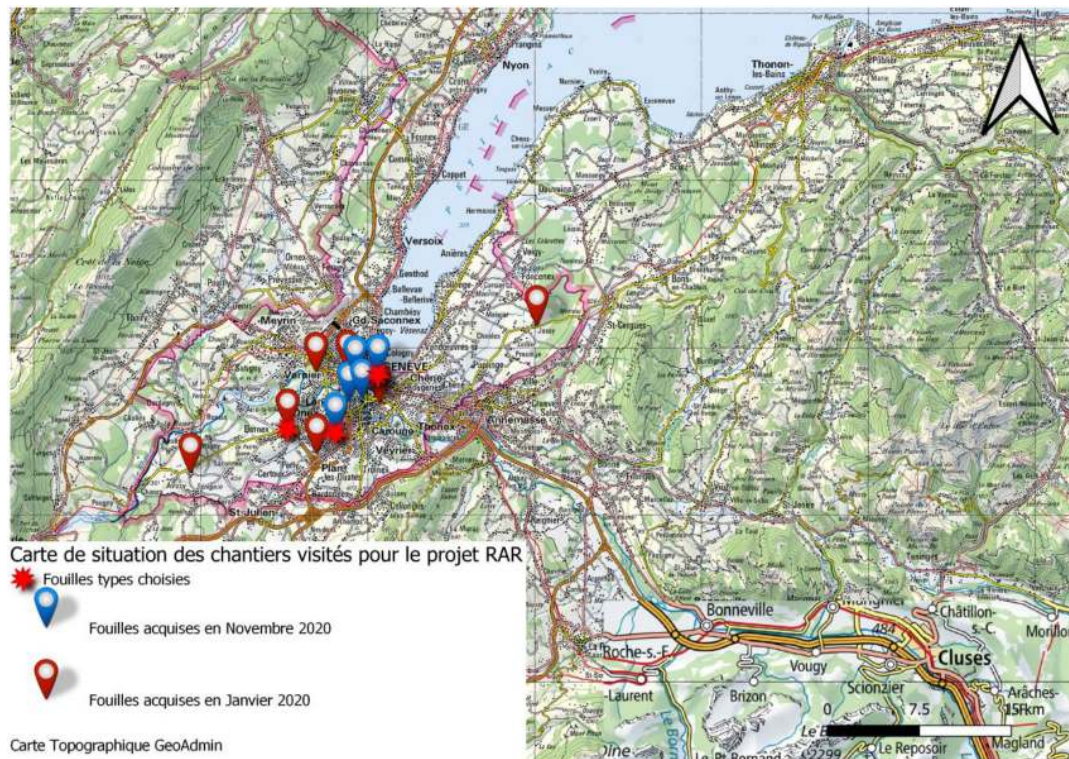


FIGURE 1 – Carte des chantiers visités lors des journées d'acquisition, 3 fouilles types ont été choisies pour évaluer les différentes méthodes d'acquisition.

2.1.4 Systèmes d'acquisition testés

Afin de pouvoir comparer les différentes méthodes d'acquisition existantes, différents systèmes ont été testés sur le terrain, ils seront présentés dans cette partie. Sans pouvoir être exhaustif au vu du nombre de fabricants, de familles de systèmes d'acquisition, ainsi que de solutions de traitement existantes, nous nous sommes efforcés de choisir un panel représentatif de l'état du marché en 2020, en prenant tout particulièrement le soin d'intégrer les systèmes les plus novateurs. Les 9 systèmes retenus peuvent se classer en deux familles :

- Les méthodes lasergrammétriques (basées sur des données LiDAR) :
 - Le Leica RTC360,
 - Le GeoSLAM ZebRevo,
 - Le Blk2Go,
 - Le LiDAR de l'iPad 12 Pro.
- Les méthodes photogrammétriques (basées sur des données images) :
 - Un appareil photo reflex,
 - Un iPhone 11,
 - Un iPhone 11 avec l'application KickTheMap,
 - Le GS18i de Leica,
 - Les Hololens 2 de Microsoft.



FIGURE 2 – Les systèmes d'acquisition testés durant cette phase.

2.1.5 Définition des Uses cases

Vu les systèmes d'acquisition existants et présentés dans la partie précédente, nous proposerons ici différents scénarios d'utilisation, incluant les étapes d'acquisition et de traitement qui sont par essence même indissociables. En effet, le traitement des données est très différent suivant la méthode et le système choisi. Finalement deux chaînes de traitements peuvent être envisageables suivant le choix d'une méthode lasergrammétrique ou photogrammétrique. Les différents "uses cases" testés sont présentés sur la Figure 3.

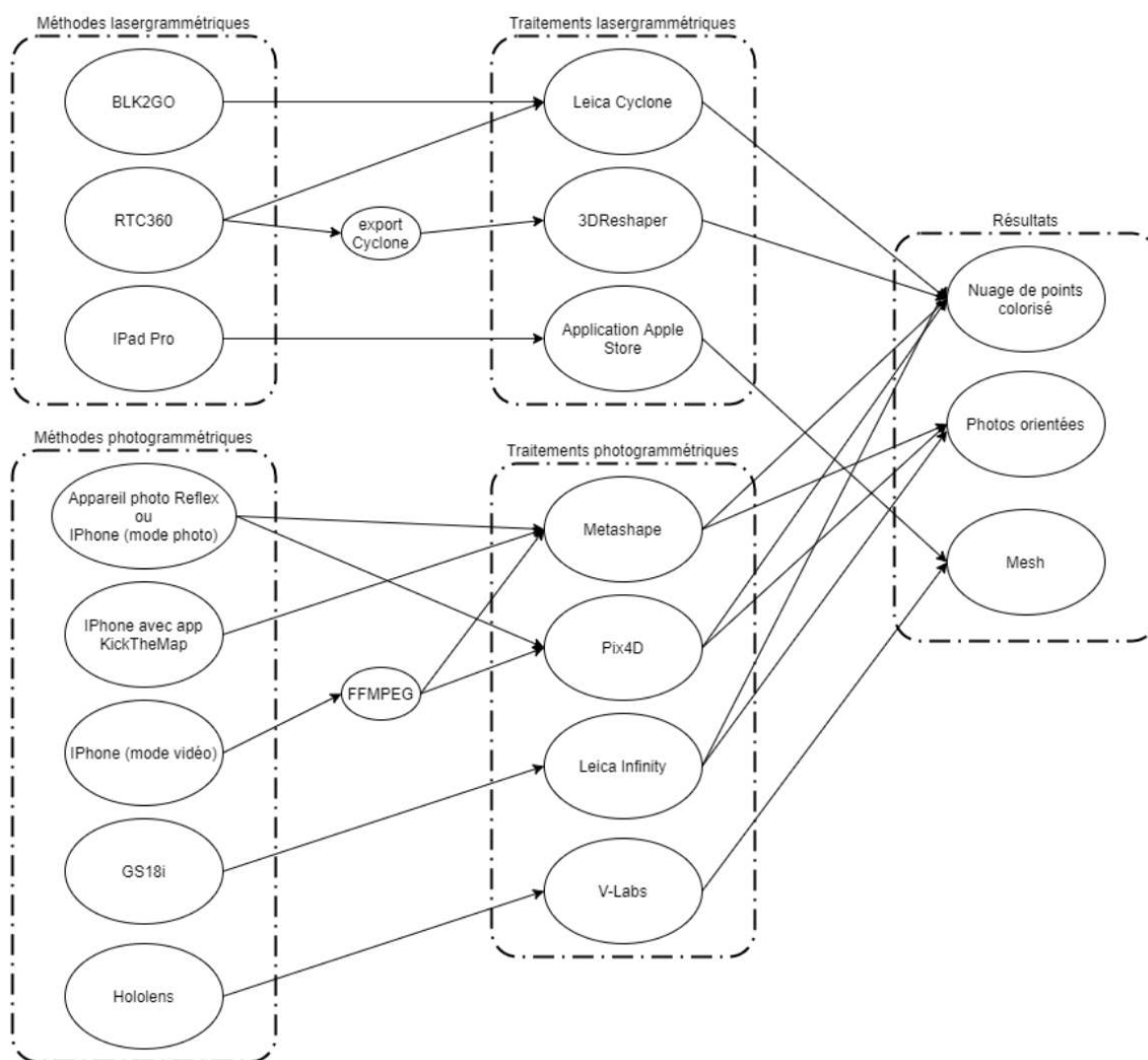


FIGURE 3 – Schéma récapitulatif des différents Uses Cases testés.

2.1.6 Différentes analyses sur les trois fouilles types

Différentes précisions ont été calculées sur chaque fouille type par rapport à des données de référence (celles du RTC360). Des analyses qualitatives (visuelles) ont aussi été faites et seront présentées par fouille type.

1ère fouille type "Tram de Bernex" Nous avons effectué des comparaisons quantitatives (précisions calculées par comparaison avec une référence) ainsi que qualitatives (analyse visuelle) sur les résultats obtenus avec les méthodes suivantes :

- Appareil photo Reflex,
- L'application KickTheMap sur iPhone,
- Les photos de l'iPhone11,
- Le ZebRevo de GeoSLAM.

. Ces comparaisons sont présentées dans les paragraphes suivants. La zone d'étude est visible sur la Figure 4.



FIGURE 4 – Zone d'étude pour les comparaisons sur la fouille "Tram de Bernex".

Analyse quantitative Afin de récapituler, les histogrammes de comparaison des nuages de points des différents systèmes avec leur précision calculée sont visibles sur la Figure 5. Cette analyse quantitative a été faite sur une zone de 10 mètres de long.

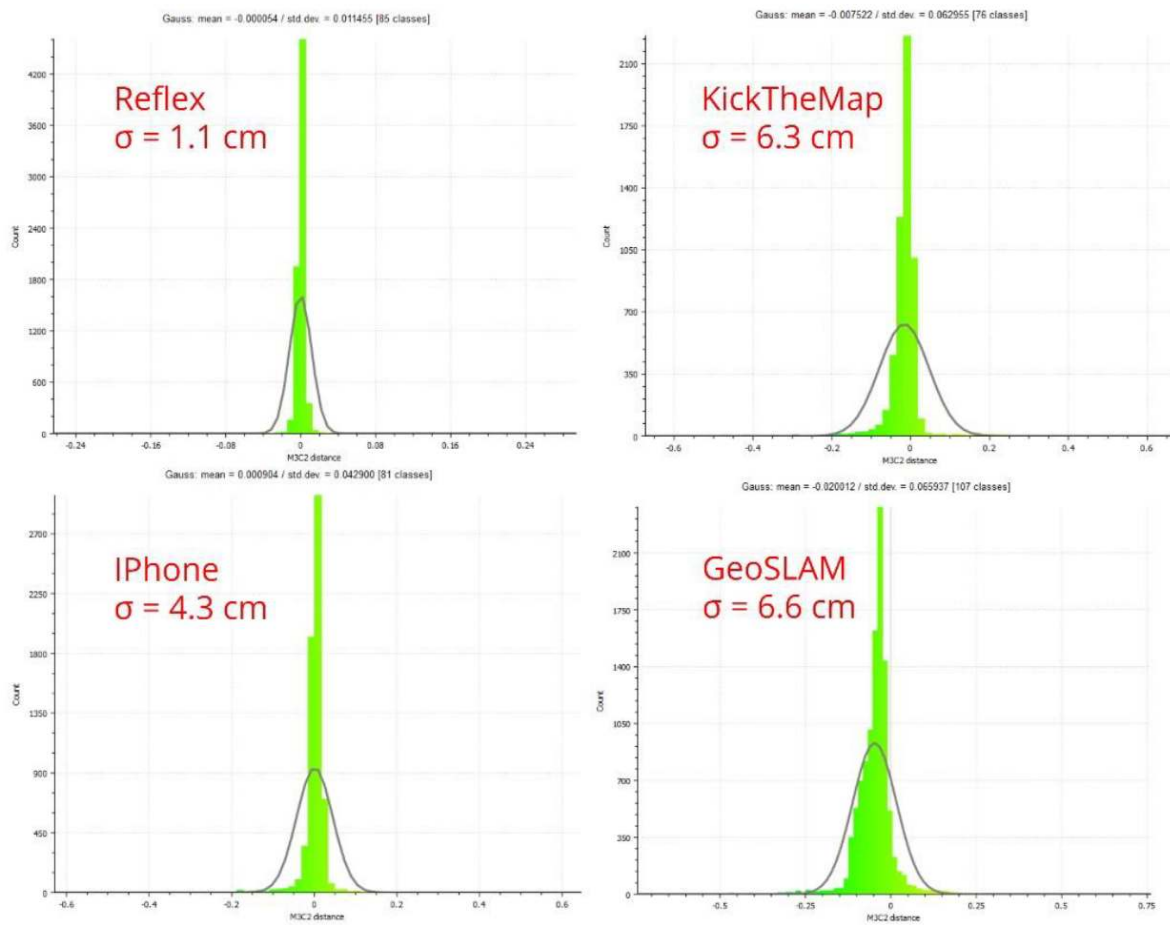


FIGURE 5 – Comparaison des nuages de points acquis avec les différents systèmes d'acquisition, la référence pour la comparaison est le nuage de points du RTC360.

On remarque sur la Figure 5 que la meilleure précision est obtenue avec le Reflex suivi par les photos de l'iPhone puis l'iPhone avec KickTheMap et le GeoSLAM.

Analyse qualitative Les Figures 6 et 7 nous montrent à chaque fois une comparaison des 4 nuages de points obtenus (en noir sur les figures).

La comparaison a été faite avec le nuage de référence acquis au RTC360 (en rouge sur les figures). Les zones d'études sont très petites afin de pouvoir faire une analyse visuelle. En plus de la comparaison visuelle des résultats, une analyse quantitative est à nouveau faite mais cette fois seulement sur l'échantillon visible (c'est pour cela que les précisions changent par rapport à celles annoncées avant). On remarque que les meilleurs résultats sont obtenus avec l'appareil photo Reflex et l'iPhone 11. Ces coupes montrent aussi que KickTheMap a peut-être un problème de géoréférencement (ces traitements sont faits directement sur un serveur). La méthode GeoSLAM semble aussi avoir un problème de géoréférencement mais les coupes permettent en plus de remarquer le bruit de mesure important.

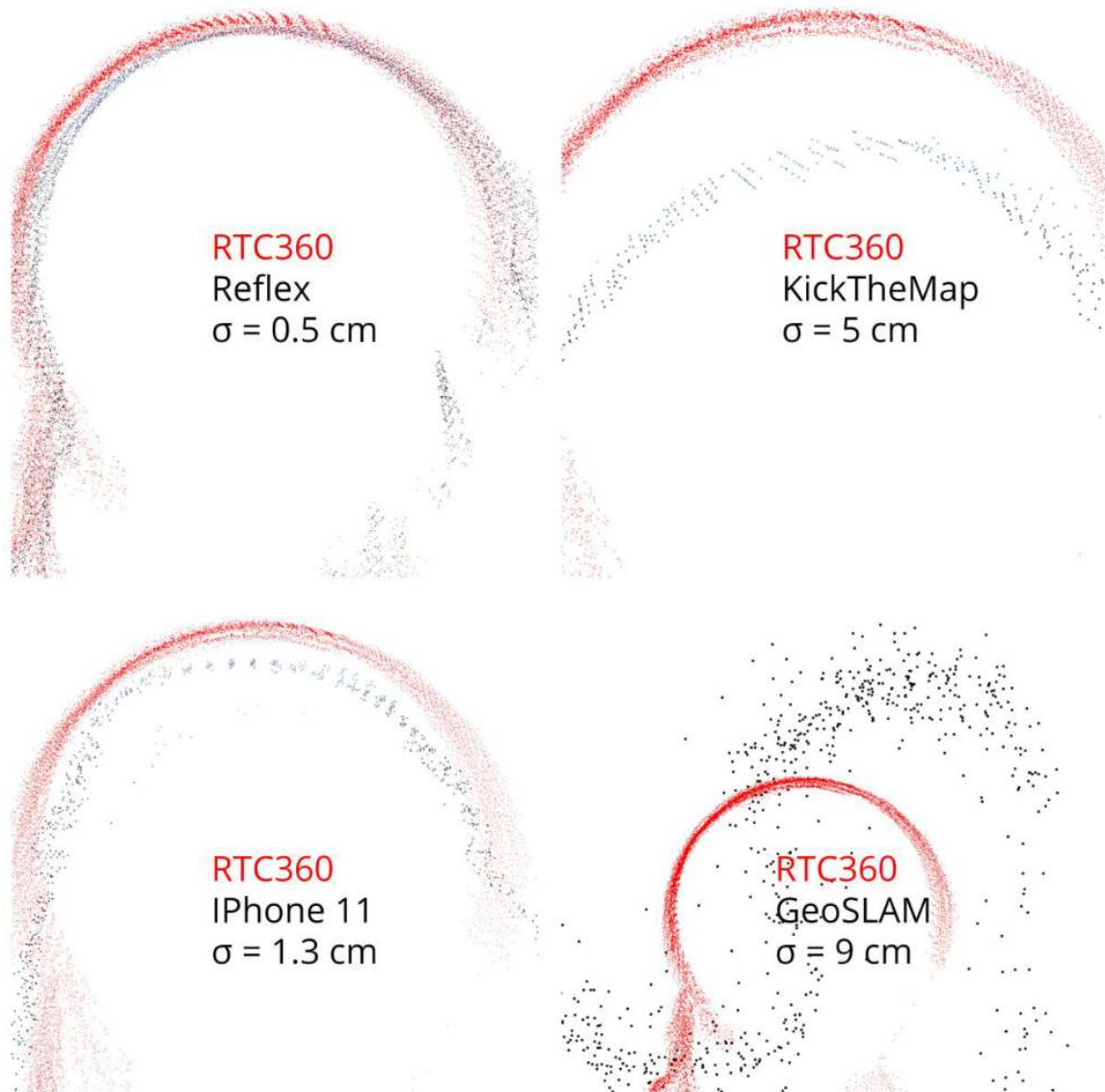


FIGURE 6 – Comparaison d'une coupe d'un tuyau acquis avec les 4 méthodes testées (appareil photo Reflex, l'application KickTheMap, les photos de l'iPhone 11, le ZebRevo de GeoSLAM). En rouge le nuage de référence, en noir le nuage de points pour chaque méthode.

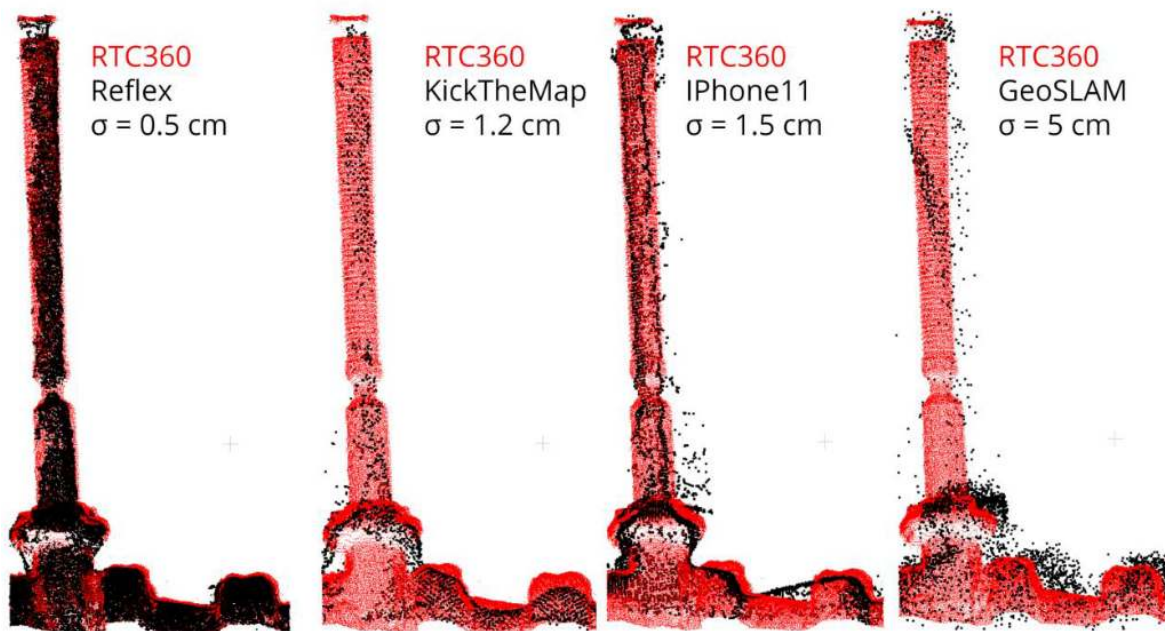


FIGURE 7 – Comparaison d'une coupe d'un tuyau vertical acquis avec les 4 méthodes testées (appareil photo Reflex, l'application KickTheMap, les photos de l'iPhone 11, le ZebRevo de GeoSLAM). En rouge le nuage de référence, en noir le nuage de points pour chaque méthode.

Analyse de la complétude des nuages de points Afin de comparer la capacité à distinguer tous les objets sur les nuages de points, nous avons défini un critère de comparaison se présentant sous la forme d'un coefficient adimensionnel appelé "C2". Il représente le taux (en %) de points du nuage de référence présent à moins de 2 cm du nuage de points testé. Ainsi plus ce taux est haut plus nous serons capables de détecter tous les objets car plus le nuage de points sera complet. Pour s'affranchir des problèmes de géoréférencement et ainsi se concentrer uniquement sur la complétude des nuages de points, nous avons effectué un recalage par une ICP sur la référence. Ces résultats sont donnés sur la première ligne du Tableau 2.

Densité des nuages de points Sur la zone de comparaison déjà présentée, nous avons comparé les densités de points obtenues avec les différents systèmes. Les 2ème et 3ème ligne du Tableau 2 présente ces résultats (pour une zone de 20 m2 de fouille).

Les densités annoncées ici ne sont pas très pertinentes car nous pouvons augmenter ou diminuer la densité soit lors de l'acquisition pour les méthodes lasergrammétriques soit lors du post-traitement pour les méthodes photogrammétriques mais elles donnent un ordre de grandeur des densités atteignables.

	RTC360	GeoSLAM	Reflex	IPhone	IPhone+KTM
Coefficient C2 (%)	Référence	100	82	98.5	95
Nombre de points	11 M	0.6 M	4.8 M	1 M	0.3 M
Densité (pts/m2)	550'000	30'000	240'000	50'000	17'500

TABLEAU 2 – Récapitulatif des coefficients C2, du nombre de points 3D et de la densité de points pour chaque méthode d'acquisition.

2ème fouille type "Mont Blanc 1" Les analyses qualitatives et quantitatives pour cette fouille sont présentées ci-dessous. La zone d'étude est visible sur la Figure 8.



FIGURE 8 – Zone d'étude pour les comparaisons sur la fouille "Mont Blanc 1".

Analyse quantitative Afin de récapituler, les histogrammes de comparaison des nuages de points des différents systèmes avec leur précision calculée est visible sur la Figure 9.

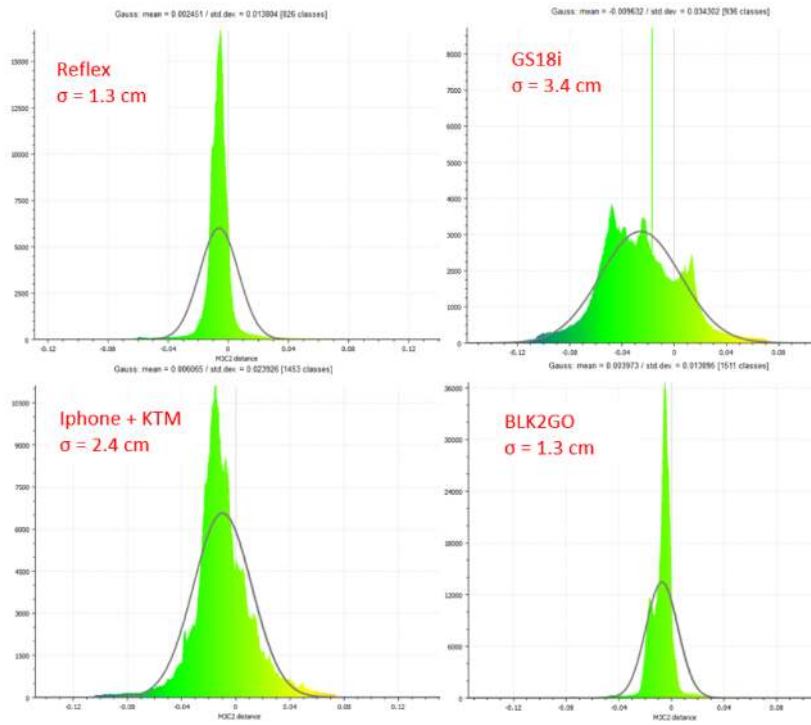


FIGURE 9 – Analyse quantitative des différents systèmes testés sur la fouille "Mont Blanc 1".

Sur la Figure 9 on remarque que toutes les méthodes possèdent une précision suffisante pour ce projet. L'étape suivante est de voir si les objets qui doivent être identifiés par les SIG sont identifiables sur les résultats en sortie.

Analyse qualitative et complétude des nuages de points De même que pour la première fouille type, nous avons calculé l'indice C2, sur la Figure 10, une comparaison visuelle est possible avec les informations de précision et du coefficient C2.

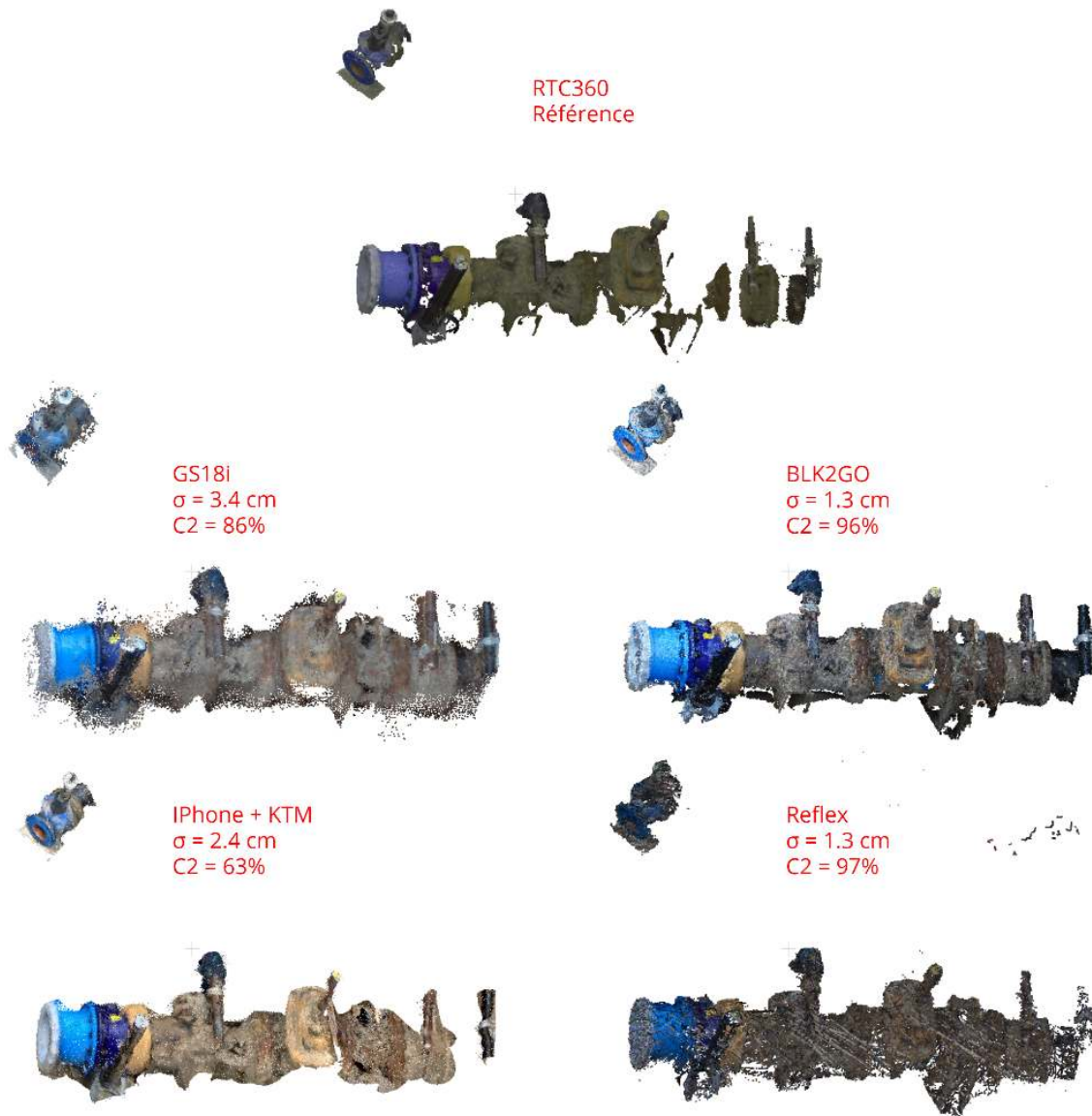


FIGURE 10 – Analyse qualitative des différents systèmes testés sur la fouille "Mont Blanc 1".

On remarque sur la Figure 10 que le système BLK2GO possède un excellent score, cependant il faut faire attention car les tuyaux noirs ou brillants ne sont pas acquis par ce système (dû à la longueur d'onde de 830 nm utilisée par le scanner laser). Sur d'autres acquisitions nous avons pu remarquer cela mais ici les tuyaux ne sont ni noirs ni brillants. Le nuage de points du reflex est très bon, les autres systèmes possèdent un score qu'on juge suffisant car les éléments non détectés sur le nuage de points le seront sur les photos.

3ème fouille type "Rambossons 1" Les mêmes comparaisons que pour les deux fouilles précédentes ont été faites et sont présentées ci-dessous. La zone d'étude est visible sur la Figure 11.



FIGURE 11 – Zone d'étude pour les comparaisons sur la fouille "Rambossons 1".

Analyse quantitative Afin de récapituler, les histogrammes de comparaison des nuages de points des différents systèmes avec leur précision calculée est visible sur la Figure 12.

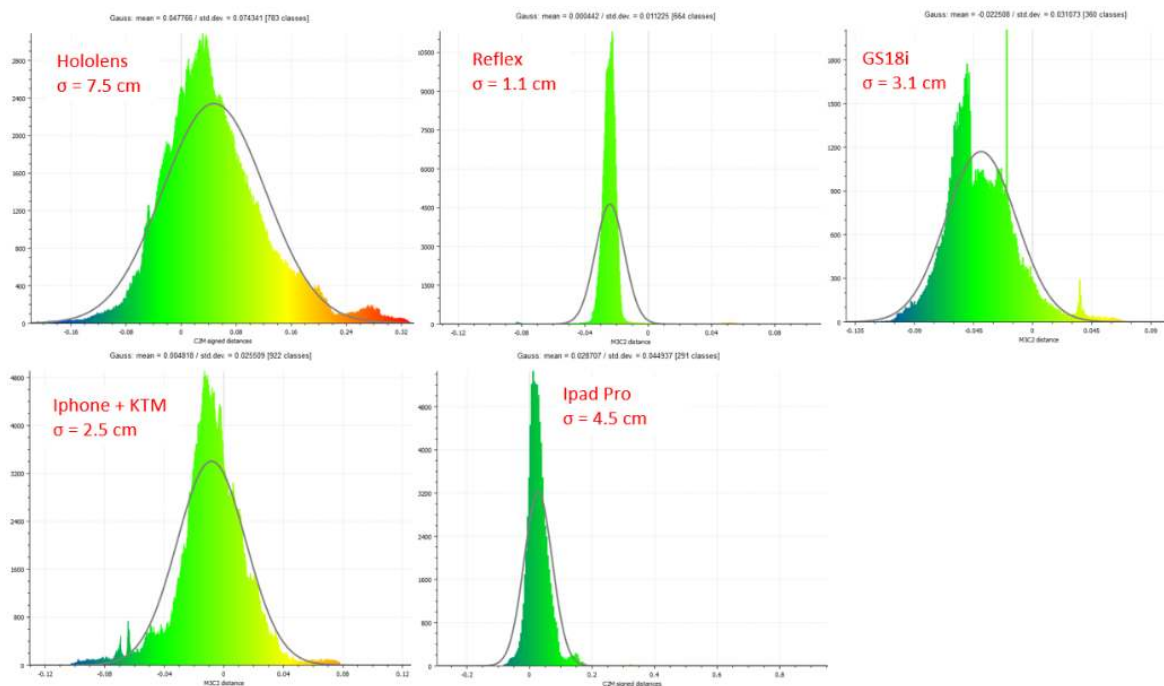


FIGURE 12 – Analyse quantitative des différents systèmes testés sur la fouille "Rambossons 1".

Sur la Figure 12 on remarque que toutes les méthodes possèdent une précision suffisante pour ce projet mise à part le système Hololens. L'étape suivante est de voir si les objets qui doivent être identifiés par les SIG sont identifiables sur les résultats en sortie.

Analyse qualitative et complétude des nuages de points De même que pour les autres fouilles, nous avons calculé l'indice C2, sur la Figure 13, une comparaison visuelle est possible avec les informations de précision et du coefficient C2.

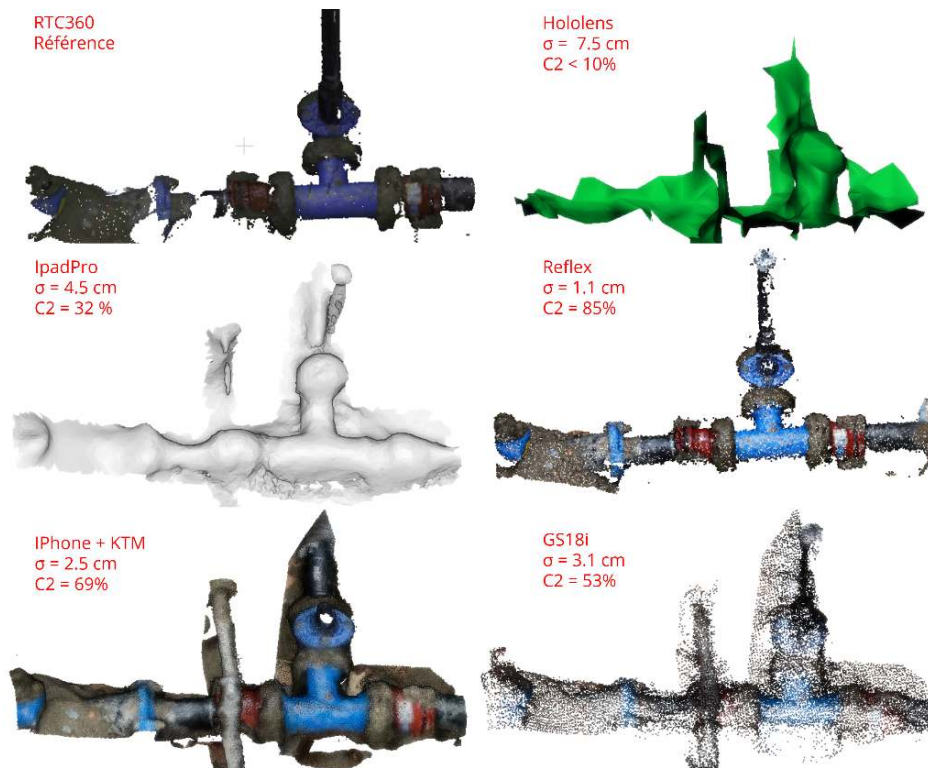


FIGURE 13 – Analyse qualitative des différents systèmes testés sur la fouille "Rambossons 1".

On remarque sur la Figure 13 que les systèmes iPadPro et Hololens possèdent un score très bas ne permettant pas de détecter tous les éléments indispensables. Le nuage de points du reflex est très bon, suivi par les autres systèmes (iPhone + KTM) qui possèdent un score qu'on juge suffisant car les éléments non détectés sur le nuage de points le seront sur les photos.

Finalement une synthèse de toutes ces analyses a été faite dans un article de la Gazette de l'IGSO : [Article \(pages 19-24\)](#).

2.1.7 Choix définitif du système d'acquisition

Au vu des différentes analyses de précision, complétude, etc. nous avons décidé de s'orienter vers un appareil photo reflex équipé d'un module GNSS en mode RTK (visible sur la Figure 14 qui permet de taguer chaque photo avec une position GNSS RTK. Ce système nous permet d'avoir accès aux images brutes de très bonnes qualité ainsi qu'après traitements photogrammétriques aux nuages de points 3D.



FIGURE 14 – Système d'acquisition choisi, appareil photo reflex et module GNSS RTK.

2.2 WP2 - Constitution des données d'apprentissage en vue de l'entraînement d'une méthode automatique

Pour mener à bien ce Work Package nous avons dû développer toute une infrastructure permettant de simplifier et automatiser au maximum, le dépôt, le traitement 3D et la labellisation des données.

2.2.1 Dépôt des données

Pour simplifier les échanges de données entre les équipes de terrain des SIG et la HEIG-VD, nous avons développé un serveur web avec une page HTML qui leur permet de déposer une acquisition (images et fichier GPS) en donnant un nom et une date. Un lien de partage est automatiquement créé sur une instance de nextcloud hébergée sur un serveur à la HEIG-VD. Enfin, une fois les données déposées, l'opérateur peut vérifier la qualité des données avec des indicateurs simples pour vérifier que le fichier GPS est présent et correspond aux images déposées ainsi que le nombre de positions en mode RTK fixe. La page web développée est visible sur la Figure 15.

I

Nom de la fouille :

testAntoine2

Date :

20.05.2021

Créer lien de partage

II

Lien de partage pour déposer les données :

<http://rar-indg.heig-vg.ch/nextcloud/s/n8oNiAyNLazRMf>

Check des données

III

Type de fichiers	Nombre ou ok
Photos (.jpg)	20
Fichier GPS (.000)	OK
Nombre de tag GPS	20
Positions RTK	19
Check Global	OK

Finir et nouveau dépôt

FIGURE 15 – Serveur et page web développés pour permettre aux utilisateurs de déposer facilement l'acquisition d'une fouille.

2.2.2 Traitements 3D automatiques

Vu le nombre d'acquisitions nous ne pouvions pas traiter les données manuellement, nous avons donc développé des scripts de téléchargement automatique sur le cloud. Ces scripts sont exécutés automatiquement chaque jour, si une nouvelle acquisition est détectée, un script permettant de réaliser les traitements 3D de manière automatique est exécuté. Grâce à cela nous obtenons chaque jour le nuage de points géoréférencé, l'orthophoto, un rapport de qualité et le projet Agisoft de manière automatique pour chaque fouille acquise. Nous regardons ensuite les rapports pour voir si le traitement s'est bien passé ou si nous devons corriger manuellement quelque chose.

2.2.3 Labellisation des données

La labellisation des données images est une étape cruciale d'un projet de reconnaissance automatique afin d'entraîner un réseau de convolutions. Au vu du nombre d'images à labelliser il était

impensable de labelliser chaque image une par une manuellement. C'est pourquoi nous avons développé une méthode de labellisation rapide qui utilise l'apport des traitements photogrammétriques. Un article scientifique¹ a été écrit à ce sujet pour présenter cette nouvelle méthode à la conférence Swiss Geoscience Meeting de 2021. Cette nouvelle méthode est inspirée d'une publication scientifique². Comme on peut le voir sur la Figure 16 (qui est un site utilisé pour entraîner le réseau de neurones), de nombreuses images voient les mêmes objets, le but est donc d'éviter de faire le travail plusieurs fois. Dans cette méthode, seul le nuage de points doit être labellisé à la main, tandis que la reprojection des objets dans chaque image est entièrement automatisée. Cela nous permet de gagner 60 % du temps par rapport à une méthode entièrement manuelle. Notre méthode est composée de trois étapes principales. Tout d'abord, une reconstruction 3D par traitement photogrammétrique fournit des images localisées et orientées avec un nuage de points dense (données 3D). Ce processus a été entièrement automatisé en utilisant l'API Python d'Agisoft Metashape. Ensuite, le nuage de points est labellisé manuellement par un opérateur dans le logiciel CloudCompare. Afin de gagner du temps, aucune séparation d'instance n'est effectuée manuellement sur les données 3D. Les instances seront calculées automatiquement lors de la dernière étape. Enfin, les données 3D labellisées sont astucieusement projetées sur chaque image pour obtenir des images labellisées (au format COCO).

Cette dernière étape est le cœur de notre méthode (voir Figure 17) : (1) le nuage de points est trié par classes ; (2) l'algorithme DBSCAN est utilisé sur chaque classe pour créer des instances d'objets ; (3) les cartes de profondeur (issues de la reconstruction 3D) sont utilisées pour projeter uniquement les points visibles sur l'image ; (4) l'algorithme alpha-shape permet de créer l'enveloppe autour des objets ; (5) enfin les images labellisées et les fichiers d'annotation au format COCO sont calculés et exportés. Toutes les étapes du traitement ont été automatisées dans un script Python.

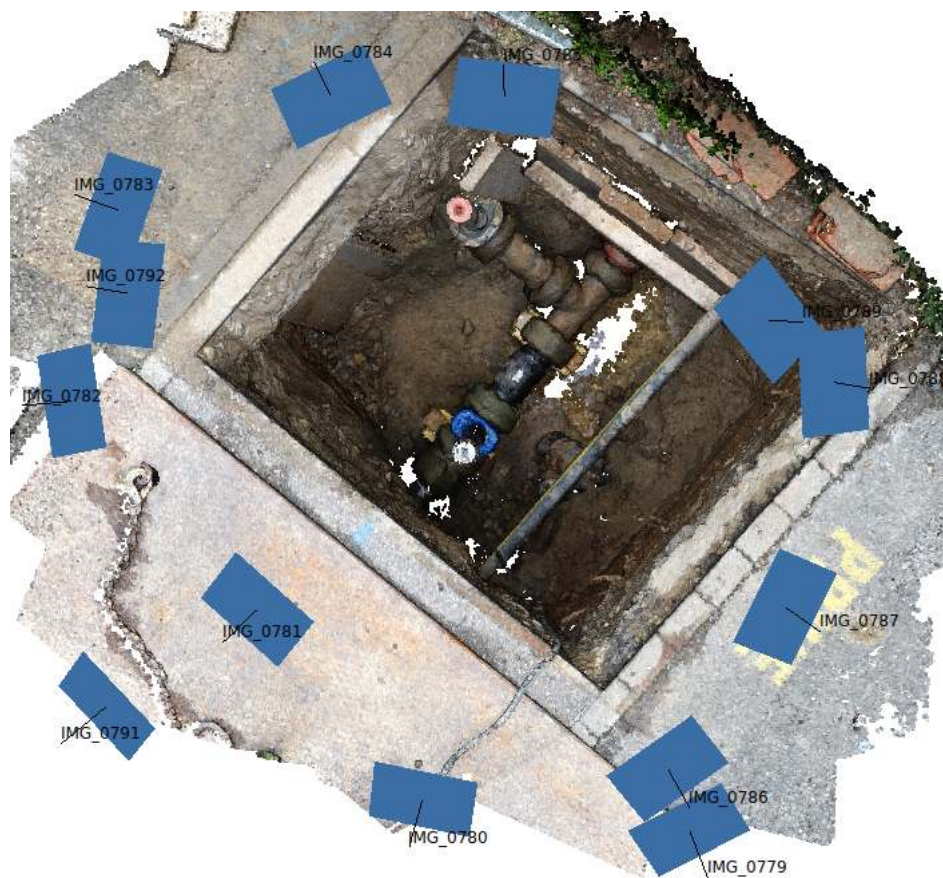


FIGURE 16 – Reconstruction 3D d'une fouille sur le réseau d'eau basée sur des images depuis plusieurs points de vues.

1. Carraud *et al.*, Fast image labelling using 3d reconstruction, SGM 2021 ([pdf](#))

2. Braun *et al.*, Automated Progress Monitoring Based on Photogrammetric Point Clouds and Precedence Relationship Graphs, ISARC2015 ([pdf](#))

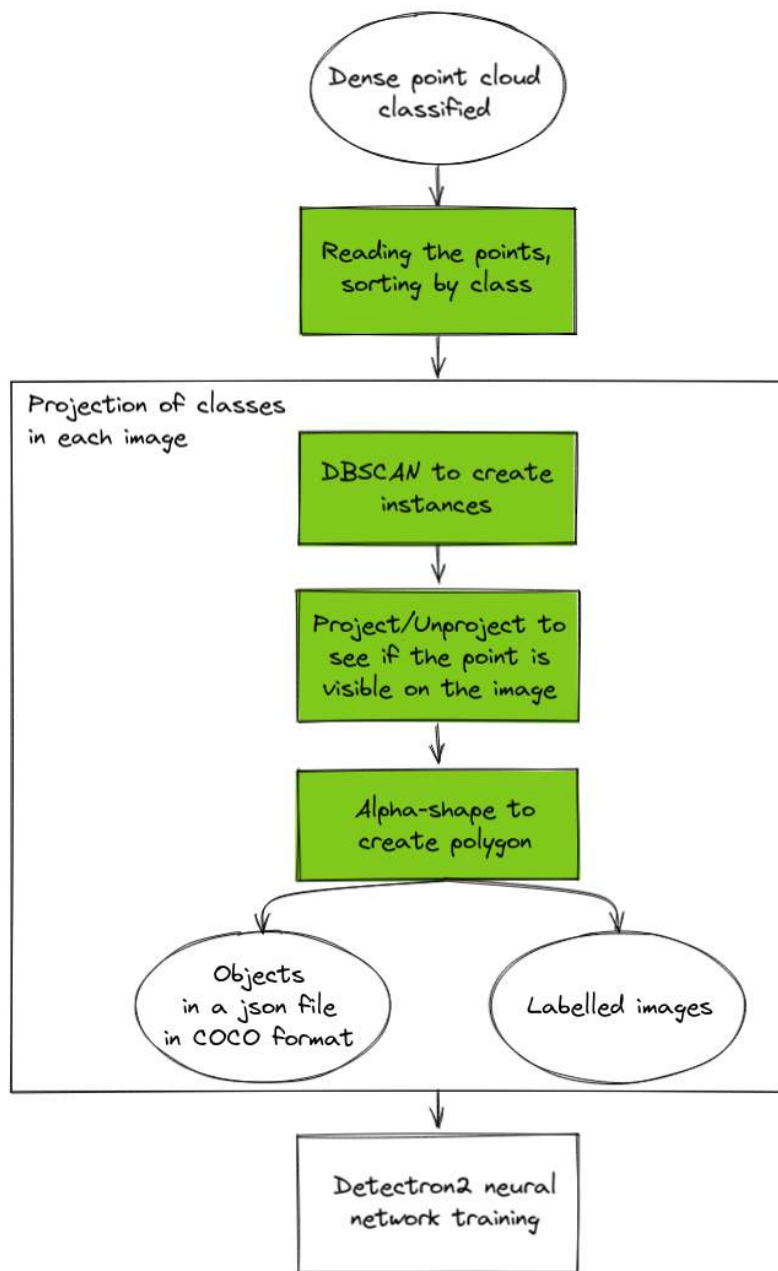


FIGURE 17 – Chaîne de traitements de la méthode de labellisation rapide développée. En vert plein les étapes automatiques, les rectangles indiquent les traitements et les ellipses les données.

2.3 WP3 - Étude préliminaire sur les méthodes d'extraction automatique des objets d'intérêts à partir des données du WP2

Dans ce Work Package nous avons entraîné un réseau de convolutions avec des données d'entraînement créées lors du WP2. Compte tenu de la durée réduite du projet, il s'agit ici que de résultats préliminaires : la partie entraînement pourrait être optimisée et la quantité de données augmentée. Nous présenterons ici la partie entraînement (2.3.1) et la partie production (2.3.2).

2.3.1 Entraînement d'un modèle

Pour la mise à jour du cadastre souterrain, nous avons besoin de détecter les classes des objets, ainsi que de les différencier, c'est-à-dire de savoir qu'il y a deux tuyaux et pas seulement qu'il y a une classe de tuyaux sur l'image. Dans notre recherche, nous nous sommes donc concentrés sur la segmentation d'instances, qui consiste à trouver la position et les masques de pixels d'objets connus dans une image et à les différencier. La segmentation d'instances peut être définie comme la technique permettant de résoudre simultanément le problème de la détection d'objets (en fournissant les classes et les boîtes limites), ainsi que celui de la segmentation sémantique (en fournissant une inférence fine par la prédiction d'étiquettes pour chaque pixel). L'entraînement d'un réseau de segmentation d'instances nécessite des dizaines de centaines d'images représentant l'objet, idéalement avec des points de vue et des arrière-plans différents. Une fois notre base de données constituée (elle contient 1 700 images étiquetées), nous avons pu tester les premiers entraînements avec la librairie Detectron2.

Pour entraîner un réseau de convolutions avec des dizaines ou des centaines de millions de poids, il est nécessaire de disposer d'une très grande puissance de calcul et d'un ensemble de données d'entraînement extrêmement important. Dans de nombreuses applications, il n'est pas possible de disposer d'autant de données. Ainsi, nous avons utilisé le principe de l'apprentissage par transfert pour utiliser un réseau pré-entraîné sur des classes différentes des nôtres. Le principe est de figer les poids des opérations de convolutions et d'ajuster les poids de la classification (les couches de sortie) lors d'un entraînement rapide sur des données étiquetées avec les classes que nous voulons détecter. Il existe encore de nombreux paramètres ajustables, appelés hyperparamètres, qui ont fait l'objet de nombreux tests. Les hyperparamètres qui ont eu le plus fort impact sur nos résultats sont tout d'abord le taux d'apprentissage, la pyramide des tailles des boîtes d'ancrage avec leur ratio afin de mieux représenter la taille des objets à trouver dans nos images et la pyramide des tailles d'images à utiliser afin de résoudre partiellement le problème des différentes échelles. Enfin, le nombre d'images par lot a été augmenté à 4 ; les tailles recommandées sont 32, 64 ou 128 pour éviter une convergence lente et chaotique. Cependant, la RAM de la carte graphique (une Tesla V100 24Gb RAM) est très vite saturée et ne nous permet pas d'augmenter la taille des lots, ce qui aboutit à un cas de descente de gradient stochastique. L'utilisation d'un serveur de calcul est une des perspectives de cette recherche. Le réseau que nous avons sélectionné est un réseau d'une profondeur de 50, le R-50 FPN, pré-entraîné sur le jeu de données COCO 2017 (118000 images étiquetées) sur 37 époques. Afin d'entraîner au mieux le réseau, c'est-à-dire de pouvoir généraliser au mieux et/ou de ne pas surajuster sur les données d'entraînement, nous avons utilisé la structure habituelle train/dev/test. Cette structure permet également de ne pas surajuster sur l'ensemble de test mais uniquement sur l'ensemble de dev au fur et à mesure des différentes expériences. En d'autres termes, nous utilisons cette structure pour optimiser à la fois le biais (avoir un biais faible, c'est-à-dire avoir la plus grande précision possible sur les données d'entraînement, en se rapprochant de l'erreur de Bayes) et la variance (avoir une variance faible, c'est-à-dire avoir une précision sur les jeux de données de développement et de test aussi proche que possible de celle de l'entraînement).

2.3.2 Etapes de production

Dans cette étape, nous supposons que la méthode de détection des objets a déjà été entraînée lors de l'étape d'entraînement. Nous utilisons le modèle entraîné sur nos données personnalisées pour faire des prédictions sur d'autres images. Pour obtenir de très bons résultats, nous pouvons utiliser un seuil élevé pour la phase d'inférence (supérieur à 90% de confiance), ce qui a pour effet de diminuer le taux de rappel et d'augmenter la précision. Mais comme dans notre cas chaque objet est vu dans au moins trois images, cela nous permet d'optimiser la précision au détriment du rappel, car un objet non détecté dans une image sera certainement détecté dans la suivante.

Segmentation d’objets 3D Nous détectons les objets dans les images puis nous créons des masques par classe afin de construire des nuages de points par classe à l’aide d’un logiciel de photogrammétrie. Cela nous permet d’obtenir diverses informations utiles pour la mise à jour (altitude d’un objet, diamètre d’un tuyau, etc.). La méthode est extrêmement simple, nous l’avons déjà utilisée dans le cadre d’un mémoire de Master pour extraire des raisins d’un nuage de points pour l’agriculture de précision et d’autres l’utilisent également pour classifier des nuages de points de bâtiments historiques.

Mise à jour de la base de données Dans cette étape, nous devons introduire une étape de vectorisation des objets 3D détectés et une comparaison avec la base de données existante pour détecter les changements. En outre, plusieurs attributs liés aux objets doivent être connus pour la mise à jour de la base de données. Ces processus spécifiques feront partie de la suite du projet.

2.3.3 Analyses de précision de reconnaissances

Pour les résultats, nous avons utilisé une fouille de l’ensemble de test. Ce site ne comprend pas tous les types d’objets que nous cherchons à détecter, mais il nous permet de comprendre dans une étude de cas assez simple les résultats qui peuvent être obtenus avec la chaîne de traitements présentée. D’autres fouilles comportent d’autres types d’objets, dont certaines classes sont extrêmement sous-représentées dans nos données d’entraînement et dégradent donc fortement les résultats. La détection de ces objets fera l’objet de développements et de recherches lors de la suite du projet. Pour les analyses de précision, la librairie que nous avons utilisée (Detectron2) permet d’accéder facilement aux informations de précision et d’Intersection sur Union (IoU) sur chaque classe. Cependant nous avons développé un script permettant de le faire par nous même et ainsi calculer des matrices de confusion. Dans cette section, nous présentons les résultats au niveau de la segmentation des images ainsi qu’au niveau de la segmentation des nuages de points 3D.

Résultats au niveau image La Figure 18 montre un résultat brut après inférence sur une image d’une fouille test. À partir de ce résultat, nous extrayons les masques par classe afin d’appliquer la méthode de segmentation 3D (présentée dans la Section 2.3.2) et obtenons le nuage de points labellisé visible sur la Figure 19.



FIGURE 18 – Image brute après inférence.

Le tableau 3 présente les différents taux de précision et rappel obtenus pour chaque classe, ainsi que la précision moyenne. Dans l’ensemble, la précision et le taux de rappel sont très bons, l’erreur de Bayes a été estimée à 99%, il serait donc théoriquement possible d’améliorer un peu ces résultats.

	Précision [%]	Rappel [%]
Bride	95.9	96.9
T	99.5	97.8
Tuyau	97.4	96.9
Vanne papillon	95.2	89.8
Vanne à tuile	96.6	99.2
Précision moyenne = 96.8 %		

TABLEAU 3 – Précision, rappel et précision moyenne obtenus sur chaque classe au niveau image.

Résultats au niveau nuage de points (3D) La Figure 19 montre (a) le nuage de points brut après traitement photogrammétrique, ainsi que (b) le nuage de points segmenté par classe avec la méthode présentée (voir Section 2.3.2) superposé au nuage de points brut.

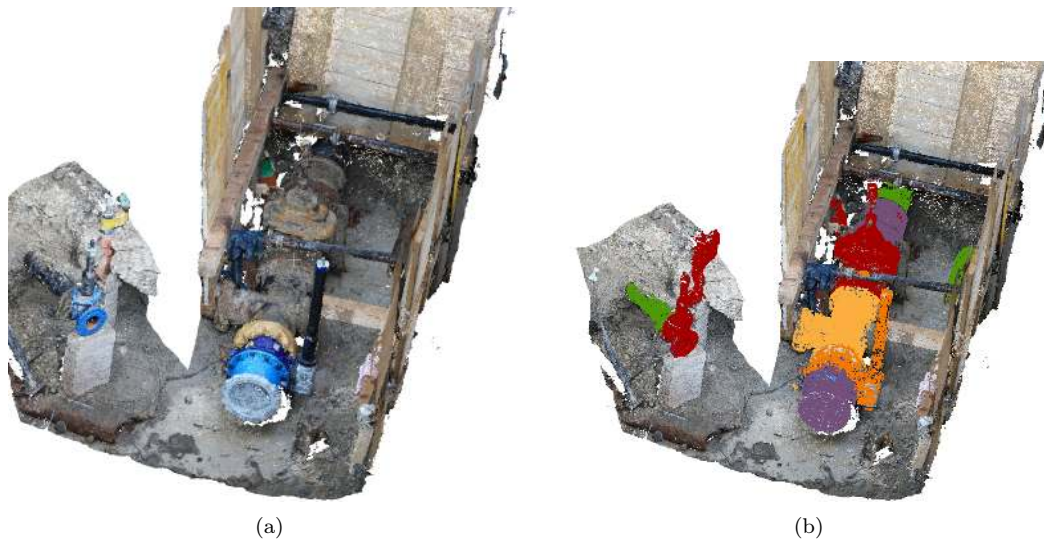


FIGURE 19 – (a) Nuage de points brut et (b) nuage de points segmenté par classe superposé sur le nuage de points brut.

Le tableau 4 montre les différents taux de précision et de rappel obtenus sur chaque classe ainsi que la précision moyenne. Nous remarquons que nous perdons un peu de précision sur toutes les classes lorsque nous passons au nuage de points 3D, passant de 96,8 % au niveau image à 96,0 % de précision moyenne au niveau 3D. Le résultat est tout de même très bon.

	Précision [%]	Rappel [%]
Bride	92.1	95.7
T	97.9	97.9
Tuyau	97.5	95.3
Vanne papillon	93.2	90.0
Vanne à tuile	97.9	98.9
Précision moyenne = 96.0 %		

TABLEAU 4 – Précision, rappel et précision moyenne obtenus sur chaque classe au niveau 3D (nuage de points).

Des développements plus poussés sur cette partie du projet seront à venir lors de la suite du projet.

3 Conclusion et perspectives

3.1 Conclusion

Malgré un budget et une durée très réduits par rapport à la demande initiale, les objectifs de ce projet ont été tenus, même dépassés. Les trois WP ont été traités en lien avec les SIG qui sont très enthousiastes et motivés avec le travail que nous leur avons présenté.

Le WP1 nous a permis de faire un choix parmi toutes les méthodes d'acquisition. En effet le choix d'un système a impliqué le choix d'un logiciel donc d'une chaîne de traitements entière. Le choix dépendait de beaucoup de critères et finalement nous nous sommes tourné vers un appareil photo reflex équipé d'une antenne GNSS RTK. Ce n'était pas un système testé pendant cette phase mais plutôt l'assemblage des résultats de l'appareil photo reflex (testé pendant le WP1) et de la qualité de géoréférencement de la GS18i (testée également durant le WP1). Une synthèse des comparaisons des différents systèmes testés a été publiée dans la Gazette de l'IGSO³.

Lors du WP2 nous avons développé toute une infrastructure permettant aux équipes de terrain d'acquérir et de déposer facilement leur acquisition sur notre serveur. L'infrastructure nous permet aussi de traiter en grande quantité des fouilles en peu de temps grâce à beaucoup d'automatisation de processus. Enfin une nouvelle méthode de labélisation rapide a été développée. Tous ces développements ont rendu possible l'élaboration rapide d'un premier jeu de données d'entraînement raisonnable (1700 images labélisées) sur le réseau d'eau. La nouvelle méthode de labélisation rapide a été présentée à deux conférences différentes, au Swiss Geoscience Meeting 2021 avec un résumé étendu⁴ et une présentation orale ainsi qu'aux Swiss Remote Sensing Days 2022 avec un poster.

Enfin le WP3 nous a permis de choisir une méthode de segmentation automatique, nous avons pu entraîner un réseau de convolutions et atteindre de bons résultats assez rapidement. Nous avons atteint une précision de 96 % de reconnaissance automatique sur les objets communs du réseau d'eau (les objets très représentés dans notre base de données d'entraînement). Des développements ont été fait pour pouvoir utiliser la segmentation des images afin d'obtenir des nuages de points 3D segmentés automatiquement.

Finalement la chaîne de traitements entière a été présentée lors du congrès international 2022 de l'ISPRS à Nice avec un poster et un article scientifique⁵.

3.2 Résultats supplémentaires

Plusieurs résultats qui n'étaient pas prévus dans le projet financé par l'INDG ont été obtenus en plus, notamment grâce à un étudiant (Chahine Nicolas Zede) en projet de fin d'études de master dans notre laboratoire. Nous allons présenter ici rapidement la méthode développée qui pourrait être un point de départ utile pour le géoréférencement sans GNSS pour la suite de ce projet. La motivation initiale de cette recherche était de trouver une méthode de géoréférencement des images en se passant du GNSS qui peut-être inutilisable en milieu urbain dense. L'idée qui a été développée durant ce stage est celle d'utiliser la position approchée GNSS d'une caméra panoramique pour ensuite l'ajuster. La méthode permettant d'ajuster cette position utilise des points caractéristiques de l'image et essaye de les faire correspondre aux même points dans le LiDAR SwissSURFACE3D. Ensuite une transformation peut-être calculée pour corriger la position approximative initiale de la caméra. Plus de détails sur la méthode sont disponibles dans la thèse de Master de Chahine Nicolas Zede⁶.

3.3 Perspectives

Suite à ce projet, nous avons pu continuer de collaborer avec les SIG en montant un projet In-nosuisse. Ce projet est également en partenariat avec l'UNIGE. Le but de ce projet est de permettre d'industrialiser la méthode en partie développée lors du projet INDG chez les SIG. Pour ce faire, plusieurs objectifs sont à atteindre, nous devons élargir au maximum la base de données d'entraînement, reconnaître non seulement les objets du réseau d'eau (limité à ce réseau pour le projet INDG) mais aussi les objets de tous les autres (gaz, électricité, fibre, thermique chaud et froid). Nous devons aussi

3. Gazette de l'IGSO 26, article (pages 19-24) ([pdf](#))

4. Carreaud *et al.*, FAST IMAGE LABELLING USING 3D RECONSTRUCTION, SGM 2021 ([pdf](#))

5. Carreaud *et al.*, AUTOMATING THE UNDERGROUND CADASTRAL SURVEY : A PROCESSING CHAIN PROPOSAL, ISPRS 2022 ([pdf](#))

6. Zede 2022, Géolocalisation d'une caméra panoramique par appariement de références ([pdf](#))

développer une méthode de squelettisation permettant de mettre à jour la base de données. L'UNIGE travaille sur l'ajout d'un raisonneur sémantique basé sur des ontologies et des règles. La Figure 20 décrit le processus qui sera développé lors du projet.

Lors de ce projet, plusieurs pistes d'amélioration ont été identifiées, nous présenterons ici celles qui nous paraissent les plus pertinentes :

- La détection d'objets rares reste un problème avec la chaîne de traitements développée. Une perspective serait d'utiliser des réseaux adverses génératifs (GAN) pour créer des données artificiellement réelles. Les méthodes d'augmentation permettent une certaine amélioration sans coûts supplémentaires en temps mais ne sont pas suffisantes pour correspondre aux objets hautement représentés dans la base de données d'entraînement.
- Afin d'optimiser l'entraînement du réseau de convolutions, il est prévu d'avoir accès à un véritable serveur de calcul avec plusieurs cartes graphiques, ce qui permettra de ne pas être limité dans le choix de certains paramètres et de tester plus rapidement plusieurs réseaux.
- Une amélioration possible de la chaîne de traitements serait l'utilisation d'un réseau de convolutions directement en 3D ou sur des données RGB-D qui tirerait théoriquement avantage de la géométrie 3D ainsi que de la couleur. À l'heure actuelle, ces réseaux sont beaucoup moins développés que les réseaux de convolutions 2D.

Enfin, les travaux développés dans le cadre de ce projet ouvrent de belles perspectives dans des domaines variés. En effet, ces travaux offrent la possibilité d'apporter une sémantique à des modèles 3D qui en sont généralement dépourvus. Ainsi, on pourrait imaginer utiliser cette méthode pour automatiser la classification de nuage de points LiDAR à partir des images obliques, avec l'introduction de classes très spécifiques (gaine technique, antenne, Velux sur les toits, balcon, fenêtre en façade...), ou encore enrichir la sémantique d'un modèle BIM par exemple en rajoutant des objets (luminaire, prise électrique...) ou en caractérisant les types de matériaux des murs (béton, brique apparente, peinture...).

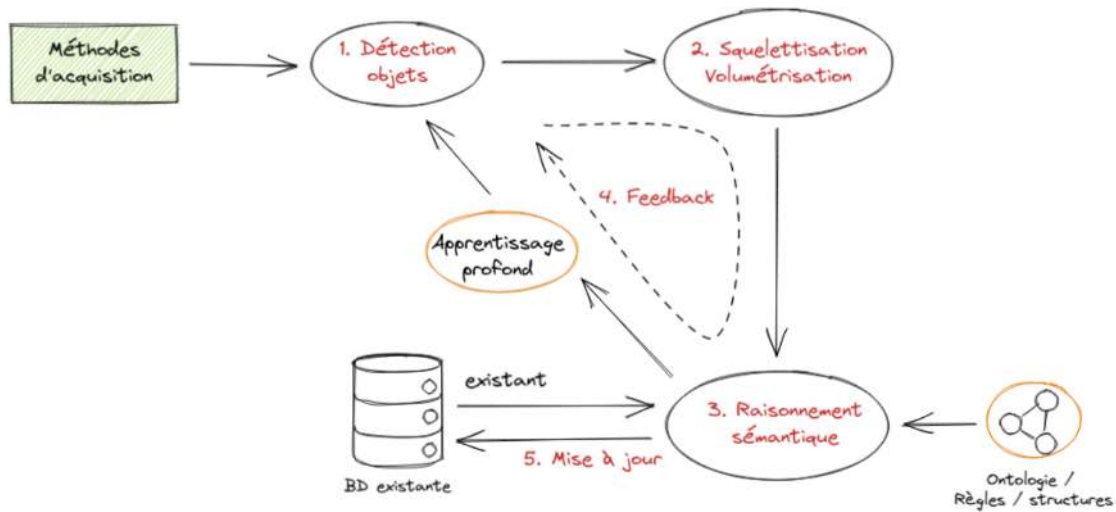


FIGURE 20 – Schéma général du projet Innosuisse.