



**HAL**  
open science

# Segmentation sémantique de nuages de points multisources de parcelles et d'ouvrages existants à l'aide de techniques d'apprentissage machine

Jérémy Montlahuc

## ► To cite this version:

Jérémy Montlahuc. Segmentation sémantique de nuages de points multisources de parcelles et d'ouvrages existants à l'aide de techniques d'apprentissage machine. Génie des procédés. HESAM Université; École de technologie supérieure (Montréal, Canada), 2023. Français. NNT : 2023HESAE044 . tel-04502254

**HAL Id: tel-04502254**

**<https://pastel.hal.science/tel-04502254v1>**

Submitted on 13 Mar 2024

**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.

**ÉCOLE DOCTORALE SCIENCES ET MÉTIERS DE L'INGÉNIEUR**  
**LISPEN - Campus d'Aix-en-Provence**  
**THÈSE**

*Présentée par :* **Jérémy MONTLAHUC**

*Soutenue le :* **27 juin 2023**

*pour obtenir le grade de :* **Docteur d'HESAM Université et Doctorat en génie Ph.D.**

*préparée à :* **l'École Nationale Supérieure d'Arts et Métiers**  
**et à**

**l'École de Technologie Supérieure**

*Spécialité :* **Génie de la conception**

**Segmentation sémantique de nuages de points  
multisources de parcelles et d'ouvrages existants à l'aide  
de techniques d'apprentissage machine**

**THÈSE dirigée par : M. Louis RIVEST**

**et co-encadrée par : M. Jean-Philippe PERNOT, M. Antoine TAHAN et M. Arnaud POLETTE**

**Jury**

**M. Jean-Luc MARI**, Professeur des Universités, LIS, Aix-Marseille Université

**M. Yann QUINSAT**, Professeur agrégé, LURPA, Université Paris-Saclay

**M. Marc-Antoine DROUIN**, Ph. D., CNRC

**M. Michel GUÉVREMONT**, Ph. D., Hydro-Québec

**M. Louis RIVEST**, Professeur, LIPPS, ÉTS

**M. Antoine TAHAN**, Professeur, LIPPS, ÉTS

**M. Jean-Philippe PERNOT**, Professeur des Universités, LISPEN, Arts et Métiers

**M. Arnaud POLETTE**, Maître de Conférences, LISPEN, Arts et Métiers

Président

Rapporteur

Examineur

Examineur

Examineur

Examineur

Examineur

Examineur

**T  
H  
È  
S  
E**

Segmentation sémantique de nuages de points multisources de parcelles et d'ouvrages existants à l'aide de techniques d'apprentissage machine

par

Jérémy MONTLAHUC

THÈSE PRÉSENTÉE À L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE ET À  
L'ÉCOLE NATIONALE SUPÉRIEURE DES ARTS ET MÉTIERS COMME  
EXIGENCE PARTIELLE À L'OBTENTION DU DOCTORAT EN GÉNIE  
Ph. D.  
ET DU GRADE DE DOCTEUR DE L'ÉCOLE NATIONALE SUPÉRIEURE  
DES ARTS ET MÉTIERS

MONTRÉAL, LE 27 JUIN 2023

ÉCOLE DE TECHNOLOGIE SUPÉRIEURE  
UNIVERSITÉ DU QUÉBEC



Jérémy Montlahuc, 2023



Cette licence [Creative Commons](https://creativecommons.org/licenses/by-nc-nd/4.0/) signifie qu'il est permis de diffuser, d'imprimer ou de sauvegarder sur un autre support une partie ou la totalité de cette œuvre à condition de mentionner l'auteur, que ces utilisations soient faites à des fins non commerciales et que le contenu de l'œuvre n'ait pas été modifié.

## **PRÉSENTATION DU JURY**

CETTE THÈSE A ÉTÉ ÉVALUÉE

PAR UN JURY COMPOSÉ DE :

M. Louis RIVEST, directeur de thèse  
Professeur, Département de génie des systèmes à l'École de technologie supérieure

M. Antoine TAHAN, codirecteur de thèse  
Professeur, Département de génie mécanique à l'École de technologie supérieure

M. Jean-Philippe PERNOT, co encadrant de thèse  
Professeur des Universités, LISPEN, Arts et Métiers HESAM Université

M. Arnaud POLETTE, co encadrant de thèse  
Maître de Conférences, LISPEN, Arts et Métiers HESAM Université

M. Jean-Luc MARI, président du jury  
Professeur des Universités, LIS, à la Faculté des Sciences, Aix-Marseille Université

M. Yann QUINSAT, membre du jury  
Professeur agrégé, LURPA, à l'Université Paris-Saclay

M. Marc-Antoine DROUIN, examinateur externe  
CNRC

M. Michel GUÉVREMONT, examinateur externe  
Hydro-Québec

ELLE A FAIT L'OBJET D'UNE SOUTENANCE DEVANT JURY ET PUBLIC

LE 27 JUIN 2023

À L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE



# **Segmentation sémantique de nuages de points multisources de parcelles et d'ouvrages existants à l'aide de techniques d'apprentissage machine**

Jérémy MONTLAHUC

## **RÉSUMÉ**

Les technologies d'acquisition à distance sont de plus en plus disponibles et performantes. Leur déploiement comme technologie de mesure sans contact est plus fréquent et les possibilités d'utilisation ne font qu'augmenter. Ces technologies, couplées à la puissance de calcul des ordinateurs, rendent possibles l'acquisition et le traitement d'un grand nombre de données. La combinaison des données issues de diverses technologies devient donc envisageable, et même souhaitable pour pallier les limites de chaque technologie.

Ce projet de recherche s'inscrit dans cette lignée. Il a pour domaine d'application la numérisation de larges zones géographiques, allant de  $10^2$  m<sup>2</sup> au 10 km<sup>2</sup> avec l'exploitation simultanée de plusieurs technologies d'acquisition. Dans notre cas spécifique, nous avons accès à des relevés de quatre technologies d'acquisition distinctes que sont le lidar aéroporté à bord d'avions, lidar aéroporté à bord d'hélicoptères, la photogrammétrie par drones et le lidar terrestre. Le principal requis est la création automatisée, à partir de plusieurs types de relevés et via un algorithme de segmentation sémantique, d'un nuage de points multisource complet et consolidé. Ce sujet soulève la question fondamentale qui est de comment exploiter au mieux les différents relevés, qui sont divers par leurs densités et caractéristiques, dans le but de créer et de segmenter précisément le nuage de points multisource.

Les différentes sources de données fournissant des données hétérogènes, leur fusion et leur consolidation sont des sujets peu traités dans la littérature. La principale raison est le manque de projets industriels qui impliquent plusieurs technologies d'acquisition. Ces dernières, lorsqu'elles sont mises en place simultanément pour un projet, doivent répondre au mieux aux critères qui se résume à un nuage cohérent et une segmentation précise. L'utilisation d'une unique technologie par projet est de plus en plus remise en cause. En effet, l'utilisation croissante des drones et de la photogrammétrie aérienne pour soutenir les lidars terrestres procure un accès grandissant à des captations dites multisources. Cependant, la littérature scientifique est, pour le moment, maigre en articles prenant plusieurs sources, et est encore plus pauvre quand il s'agit de sources autre qu'un couple images/lidar.

Nos travaux s'inscrivent dans l'optique d'une ouverture des projets de captation numérique pour l'usage de diverses sources complémentaires. Notre principale contribution propose différents modules développés pour traiter et segmenter sémantiquement les nuages de points multisources. La propagation des attributs acquis par les différentes sources est basée sur la notion de proximité. Les points avoisinants sont interjumelés et les attributs sont ainsi consolidés, ce qui a permis de segmenter sémantiquement tout en bénéficiant des avantages de chaque technologie. Une étude d'ablation est réalisée dans le but d'évaluer l'intérêt des différents modules de notre proposition. Notre proposition est validée sur des nuages de points multisources mis à disposition par notre partenaire industriel. Également, les performances de notre proposition sont comparées, avantageusement, avec d'autres algorithmes récemment proposés dans la littérature scientifique sur ce sujet. Notre proposition a permis d'obtenir de

meilleures performances de segmentation sémantique. Toutefois, le temps des calculs est légèrement plus grand.

**Mots-clés** : Segmentation sémantique, Réseaux de neurones artificiels, Lidar aéroporté, Lidar terrestre, Fusion de données, Photogrammétrie



# **Semantic segmentation of multisources point clouds of existing parcels and structures using machine learning techniques**

Jérémy MONTLAHUC

## **ABSTRACT**

Remote acquisition technologies are becoming more and more accessible and are no longer reserved for professionals. They are therefore being used more frequently and opportunities to use them are only increasing. These technologies, coupled with the growing computing power of computers, make it possible to acquire and process a large amount of data. Combining data from different technologies is therefore becoming possible.

This research project aligns with this topic and involves the digitization of large geographic areas ranging in size from several hundred square meters to one square kilometer with several acquisition technologies. In our case, we had access to surveys from four distinct acquisition sources: airborne lidar onboard aircraft and helicopters, photogrammetry from UAVs and terrestrial lidar. The main requirements to fulfill the project are to be able to create a complete point cloud from several types of sources and be able to semantically segment multisources point clouds. These requirements raise the question of how best to use the various sources available to create and segment a multisource point cloud.

The use of multiple data sources that provide heterogeneous data for the same project is not well covered in the literature for a variety of reasons. One of the main reasons is that there are very few projects involving multiple technologies in the industry. The acquisition technologies that are selected for a project are those that best meets the project's needs. The idea of using only one technology per project is increasingly being challenged with the use of UAVs and photogrammetry to support ground-based lidar. However, the body of scientific literature currently includes few articles that take into account multiple sources and even fewer that use other combinations of sources than photography + lidar. This project therefore provides multiple contributions. First, it includes multiple modules that have been developed to process and semantically segment multisources point clouds. Second, the propagation of the attributes acquired by the different sources on the neighboring points and, finally, a simple system to take into account the neighborhood. Our proposed method enabled us to semantically segment multisources point clouds by taking into account the attributes of each source using different modules. An ablation study was then performed to evaluate the interest of the different modules in our proposed method. Our method provided very good semantic segmentation results in comparison with algorithms used in existing works on this subject. Our work aligns with the openness to complementary sources that is observed in digital capture projects.

**Keywords** : Semantic Segmentation, Artificial Neural Network, Airborne Lidar, Terrestrial Lidar, Data Fusion, Photogrammetry



## TABLE DES MATIÈRES

	Page
INTRODUCTION .....	1
CHAPITRE 1 REVUE DE LA LITTÉRATURE.....	7
1.1 Introduction.....	8
1.2 Numérisation de terrains de grande échelle .....	9
1.2.1 Principes physiques de captation .....	11
1.2.2 Techniques et technologies de captation associées.....	12
1.2.3 Synthèse des technologies de captation .....	18
1.2.4 Jeux de données existants .....	19
1.3 Segmentation sémantique de nuages de points .....	24
1.3.1 Segmentation s'appuyant sur les bordures – <i>Edge-Based</i> .....	27
1.3.2 Segmentation par croissance de régions – <i>Region Growing</i> .....	27
1.3.3 Segmentation par ajustement de modèles – <i>Model Fitting</i> .....	28
1.3.4 Segmentation reposant sur la théorie des graphes – <i>Graph-Based</i> .....	29
1.3.5 Segmentation par apprentissage machine .....	29
1.3.6 Méthodes hybrides de segmentation.....	33
1.3.7 Caractéristiques supplémentaires.....	34
1.3.8 Évaluation et validation de la segmentation sémantique .....	39
1.4 Fusion de données géospatiales .....	43
1.4.1 Fusion au niveau des données d'entrée.....	44
1.4.2 Fusion au niveau des attributs.....	45
1.4.3 Fusion au niveau des résultats.....	45
1.4.4 Fusions multiniveaux .....	47
1.5 Synthèse de l'état de l'art.....	47
1.5.1 Positionnement de notre sujet par rapport à l'état de l'art .....	48
1.5.2 Critiques de l'état de l'art .....	49
CHAPITRE 2 OBJECTIFS ET MÉTHODOLOGIE DE LA RECHERCHE .....	51
2.1 Introduction.....	51
2.2 Problématique et question de recherche .....	51
2.3 Objectifs de la recherche.....	53
2.3.1 Objectif principal .....	53
2.3.2 Sous-objectifs.....	54
2.4 Méthodologie de la recherche.....	54
2.4.1 Frontières du projet et hypothèses de recherche .....	54
2.4.2 Outils utilisés .....	55
2.4.3 Synthèse de la méthodologie .....	56
CHAPITRE 3 PROPOSITION DE SEGMENTATION SÉMANTIQUE .....	59
3.1 Introduction.....	60
3.2 Étude de segmentation sémantique monosource .....	61

3.2.1	Sélection de voisinages monosource en vue du calcul d'attributs .....	61
3.2.2	Calcul d'attributs monosource pour la segmentation sémantique .....	66
3.2.3	Sélection d'attributs et réduction dimensionnelle .....	70
3.2.4	Apprentissage machine - Réseau de neurones .....	71
3.3	Étude de segmentation sémantique multisource .....	73
3.3.1	Architecture de la proposition de segmentation sémantique multisource .....	73
3.3.2	Calculs d'attributs pour l'approche multisource .....	77
3.3.3	Propagation des attributs acquis .....	78
3.3.4	Réseau de neurones multisource (NN 1) .....	80
3.3.5	Calcul des attributs de classes avoisinantes .....	82
3.3.6	Seconde segmentation sémantique (NN 2) .....	84
3.4	Optimisation des nuages de points .....	84
3.4.1	Segmentation en entités distinctes .....	85
3.4.2	Sélection de la source .....	86
3.5	Synthèse et contributions .....	88
CHAPITRE 4 RÉSULTATS .....		91
4.1	Introduction .....	91
4.2	Présentation des données utilisées .....	92
4.2.1	Classes représentées dans les nuages de points utilisés .....	93
4.2.2	Données monosources .....	98
4.2.3	Données multisources .....	102
4.3	Résultats de l'étude sur l'apprentissage monosource .....	106
4.3.1	Segmentation sémantique monosource .....	106
4.3.2	Synthèse sur l'apprentissage monosource .....	114
4.4	Résultats de l'étude d'ablation sur la proposition multisource .....	117
4.4.1	Segmentation sémantique monosource (chemin A1+A3, B) .....	118
4.4.2	Segmentation sémantique multisource (chemin A1+A2+A3, B) .....	121
4.4.3	Segmentation sémantique en cascade (chemin A1+A2+A3, B, D) .....	124
4.4.4	Segmentation sémantique avec prises en compte du voisinage (chemin A1+A2+A3, B, C, D) .....	126
4.4.5	Synthèse de l'étude d'ablation et comparaison des différents résultats ..	129
4.5	Synthèse globale .....	135
CHAPITRE 5 ÉTUDE COMPARATIVE DES PERFORMANCES DE LA SOLUTION PROPOSÉE .....		137
5.1	Introduction .....	137
5.2	Sélection d'algorithmes pour l'étude comparative .....	137
5.2.1	Critères de sélection .....	138
5.2.2	Algorithmes sélectionnés .....	138
5.3	Résultats de l'étude comparative .....	144
5.3.1	Résultats monosources .....	144
5.3.2	Résultats multisources .....	145
5.4	Conclusion .....	149
CONCLUSION, RECOMMANDATIONS ET PERSPECTIVES .....		151

ANNEXE I	INFORMATIONS SUPPLÉMENTAIRES CONCERNANT LA CAPTATION DE LA LOCALISATION #1 .....	157
	LISTE DE RÉFÉRENCES BIBLIOGRAPHIQUES .....	159



## LISTE DES TABLEAUX

	Page
Tableau 1.1	Schématisation des principes physiques permettant la numérisation à distance Tiré de Beraldin et al. (2015).....11
Tableau 1.2	Représentation de deux autres technologies utilisant le principe de triangulation : la projection de franges et le théodolite Tiré de « Skinexigence - Projection de franges » (2022) et « Théodolite Spectra Precision NE-101 pour l’arpentage   Cansel » (2022).....17
Tableau 1.3	Caractéristiques et certains ordres de grandeurs des technologies et méthodes de numérisation qui ont été utilisées pour l’obtention de nuage de points .....18
Tableau 1.4	Tableau des principaux jeux de données répertoriés en ligne.....22
Tableau 1.5	Tableau des principaux jeux de données répertoriés en ligne (suite) .....23
Tableau 1.6	Représentation d’une classification et d’une segmentation sémantique sur un objet et sur un terrain Tirée de Yi et al. (2016).....25
Tableau 1.7	Matrice de confusion pour deux classes 0 et 1 .....40
Tableau 1.8	Matrice de confusion lors d’une segmentation sémantique. Chaque ligne de la matrice représente une classe et le pourcentage des points qui la compose, les colonnes représentant les classes estimées .....42
Tableau 3.1	Représentation de l’extraction de lignes électriques à l’aide de sphères de voisinage et des vecteurs des composantes principales .....65
Tableau 3.2	Ensemble des attributs calculés pour la segmentation sémantique monosource .....67
Tableau 4.1	Représentation et composition des nuages de points monosource .....100
Tableau 4.2	Représentation schématique et composition des nuages de points multisources .....103
Tableau 4.3	Matrice de confusion de la Localisation #4 avec l’apprentissage réalisé sur la Localisation #5 avec un jeu de données ayant des classes avec un nombre de points uniforme.....107
Tableau 4.4	Ensemble des résultats pour la segmentation sémantique des Localisation #4 et #5 avec jeu d’apprentissage uniformisé et réel .....108

Tableau 4.5	Matrice de confusion de la Localisation #4 avec l'apprentissage réalisé sur la Localisation #5 avec un jeu de données ayant des classes avec un nombre de points non uniforme.....	111
Tableau 4.6	Matrice de confusion de la Localisation #4 avec les classes de « Végétation » regroupées et la classe de « Route » regroupée avec la classe « Sol ».....	112
Tableau 4.7	Matrice de confusion de la Localisation #4 après l'utilisation d'un post-traitement de lissage des classes .....	113
Tableau 4.8	Matrice de confusion de la Localisation #1 après la segmentation sémantique monosource.....	119
Tableau 4.9	Matrice de confusion de la Localisation #2 après la segmentation sémantique monosource.....	120
Tableau 4.10	Matrice de confusion de la Localisation #1 après la segmentation sémantique multisource .....	122
Tableau 4.11	Matrice de confusion de la Localisation #2 après la segmentation sémantique multisource .....	123
Tableau 4.12	Matrice de confusion de la Localisation #1 après la segmentation sémantique en cascade .....	125
Tableau 4.13	Matrice de confusion de la Localisation #2 après la segmentation sémantique en cascade .....	126
Tableau 4.14	Matrice de confusion de la Localisation #1 après segmentation sémantique avec prise en compte du voisinage .....	128
Tableau 4.15	Matrice de confusion de la Localisation #2 après segmentation sémantique avec prise en compte du voisinage .....	129
Tableau 4.16	Valeurs de la <i>MCC</i> pour la Localisation #1 au travers des différentes parties de l'étude d'ablation.....	130
Tableau 4.17	Valeurs de la <i>MCC</i> pour la Localisation #2 au travers des différentes parties de l'étude d'ablation.....	131
Tableau 4.18	Valeurs de la <i>MCC</i> pour la Localisation #3.1 au travers des différentes parties de l'étude d'ablation.....	132
Tableau 4.19	Valeurs de la <i>MCC</i> pour la Localisation #3.2 au travers des différentes parties de l'étude d'ablation.....	133



Tableau 4.20	Valeurs de la <i>MCC</i> pour la Localisation #3.3 au travers des différentes parties de l'étude d'ablation.....	134
Tableau 5.1	Résultats comparatifs de la segmentation sémantique monosource de la Localisation #4 avec les algorithmes RandLA-Net et PointNet++.....	145
Tableau 5.2	Résultats comparatifs de la segmentation sémantique de la Localisation #1 avec l'étude d'ablation .....	146
Tableau 5.3	Résultats comparatifs de la segmentation sémantique de la Localisation #2 avec l'étude d'ablation .....	147
Tableau 5.4	Résultats comparatifs de la segmentation sémantique de la Localisation #3.1 avec l'étude d'ablation .....	147
Tableau 5.5	Résultats comparatifs de la segmentation sémantique de la Localisation #3.2 avec l'étude d'ablation .....	148
Tableau 5.6	Résultats comparatifs de la segmentation sémantique de la Localisation #3.3 avec l'étude d'ablation .....	148



## LISTE DES FIGURES

		Page
Figure 0.1	Exemple de nuage de points représentant la Localisation #5 avec la structure de ses données (coordonnées et attributs).....	2
Figure 0.2	Schématisme d'une segmentation de classe en entités distinctes.....	4
Figure 1.1	Représentation des zones d'utilisation des différents principes physiques pour la numérisation à distance en fonction de la distance de captation z et de la résolution relative à cette distance Tirée de Barbero & Ureta (2011).....	10
Figure 1.2	Représentation des principales technologies de captation à distance en fonction de la taille de l'objet à numériser et de la précision Tirée de Franceschini et al. (2014) .....	10
Figure 1.3	a) Schématisation de la numérisation à l'aide d'un lidar aéroporté, b) Schématisation du nombre de retours et numéro de retour.....	14
Figure 1.4	a) Schéma du principe physique de la triangulation, b) Représentation d'un nuage de points avec les différentes positions du capteur photographique et, c) Artefact d'acquisition nécessaire pour augmenter la précision de la photogrammétrie Tirée de Beraldin et al. (2015).....	15
Figure 1.5	Arbre représentant les méthodes permettant la segmentation de nuages de points .....	26
Figure 1.6	Flux représentant les différentes étapes retrouvées dans les projets de segmentation sémantique utilisant l'apprentissage machine non-profond.....	30
Figure 1.7	Représentation d'un nuage de points sous forme de voxels Tirée de Pedram Ghamisi et al. (2018) .....	32
Figure 1.8	Représentation d'un nuage de points d'un avion avec un maillage et de multiples photos pour l'utilisation de CNN 2D avec l'approche <i>MultiView</i> Tirée de H. Huang et al. (2017).....	33
Figure 1.9	Sphères englobantes sur le même nuage de points à différentes échelles afin de calculer des attributs à différentes échelles et prendre en compte différents contextes Tirée de Brodu & Lague (2012) .....	35

Figure 1.10	Schématisation d'une segmentation sémantique avec le modèle <i>One-vs-all</i> nécessitant plusieurs <i>classifier</i> Tirée de Brodu & Lague (2012).....	36
Figure 1.11	Visualisation de l'attribut « verticalité » sur le nuage de points « Localisation #1 » du PI. L'attribut « verticalité » a été calculé en prenant une sphère de taille fixe de diamètre 1m .....	38
Figure 1.12	Schématisation d'une fusion de données au niveau des données d'entrée .....	45
Figure 1.13	Schématisation d'une fusion de données au niveau des attributs .....	45
Figure 1.14	Schématisation d'une fusion de données au niveau du résultat.....	46
Figure 1.15	Schématisation d'une fusion de données au niveau du résultat.....	46
Figure 1.16	Schématisation de fusions de données multiniveaux.....	47
Figure 2.1	Diagramme de la méthodologie de recherche.....	57
Figure 3.1	Schématisation des étapes pour la segmentation sémantique de nuages de points en utilisant un algorithme d'apprentissage machine et un vecteur d'attributs .....	62
Figure 3.2	Représentation d'une sphère et d'un cylindre englobant pour délimiter le voisinage d'un point $P_i = (x_i, y_i, z_i)$ .....	63
Figure 3.3	Schématisation de sphères de voisinage englobant une portion de lignes électriques .....	64
Figure 3.4	Schématisation d'une captation par ALS avec l'intensité du laser et les différents retours (sur la gauche). Nuage de points et histogramme représentant les différents retours présents dans un cylindre $C_{i, k}$ (sur la droite).....	69
Figure 3.5	Représentation du MNT pour la Localisation #4 à l'aide d'un maillage (en haut). Le nuage de points est coloré selon la différence de hauteur entre le point et le maillage du MNT (en bas) .....	70
Figure 3.6	Schématisation du réseau de neurones pour la segmentation sémantique monosource .....	72
Figure 3.7	<i>Workflow</i> de notre proposition de segmentation multisource.....	76
Figure 3.8	Schématisation des différentes possibilités C1, C2 et C3 pour la propagation des attributs acquis.....	79

Figure 3.9	Schématisation de la création des attributs de classes avoisinantes sur la Localisation #1 .....	83
Figure 3.10	Schématisation d'une segmentation de classe en entité distinctes.....	85
Figure 3.11	Arbre de priorisation des technologies en fonction de la classe du point et des sources disponibles .....	87
Figure 4.1	a) Exemple de la classe Sol de la Localisation #1, b) Exemple de la classe Végétation basse de la Localisation #1 .....	93
Figure 4.2	a) Exemple de la classe Végétation haute de la Localisation #1, b) Exemple de la classe Bâtiments de la Localisation #6 .....	95
Figure 4.3	a) Exemple de la classe Surfaces hydrologiques de la Localisation #4, b) Exemple de la classe Conducteurs de la Localisation #3 .....	96
Figure 4.4	Exemple de la classe Routes de la Localisation #3.....	97
Figure 4.5	Exemples de la classe Pylônes de la Localisation #3 .....	98
Figure 4.6	Représentation et composition des nuages de points des Localisation #4.....	100
Figure 4.7	Représentation et composition des nuages de points des Localisation #5.....	101
Figure 4.8	Représentation et composition du nuage de points de la Localisation #6.....	102
Figure 4.9	Représentation et composition du nuage de points de la Localisation #1 .....	104
Figure 4.10	Représentation et composition du nuage de points de la Localisation #2.....	104
Figure 4.11	Représentation et composition des nuages de points de la Localisation #3.1.....	105
Figure 4.12	Représentation et composition des nuages de points des Localisation #3.2 et Localisation #3.3 .....	106
Figure 4.13	Localisation #6 après segmentation sémantique.....	109
Figure 4.14	Localisation #4 après segmentation sémantique et comparaison entre le modèle de référence et la segmentation sémantique .....	111

Figure 4.15	Zoom sur une mauvaise classification de la Localisation #5 à cause du nombre de retour .....	115
Figure 4.16	Zoom sur l'effet de bordure. La colonne de gauche correspond à la classe « Sol », la colonne de droite à la classe « Surfaces hydrologiques ». L'attribut de planéité calculé avec une sphère de rayon 3 m a été pris en exemple pour les classes de « Sol » et « Surfaces hydrologiques » de la Localisation #6.....	116
Figure 4.17	<i>Workflow</i> pour la première étape de l'étude d'ablation sur la segmentation sémantique monosource .....	118
Figure 4.18	Localisation #1 après segmentation sémantique monosource .....	119
Figure 4.19	<i>Workflow</i> pour la deuxième étape de l'étude d'ablation sur la segmentation sémantique multisource .....	121
Figure 4.20	Localisation #1 après segmentation sémantique multisource.....	122
Figure 4.21	<i>Workflow</i> pour la troisième étape de l'étude d'ablation sur la segmentation sémantique en cascade.....	124
Figure 4.22	Localisation #1 après segmentation sémantique en cascade.....	125
Figure 4.23	Ensemble du <i>Workflow</i> avec prise en compte du voisinage dans le scénario de segmentation en cascade .....	127
Figure 4.24	Localisation #1 après segmentation sémantique avec prise en compte du voisinage .....	128
Figure 4.25	Localisation #3.1 après segmentation sémantique avec prise en compte du voisinage .....	132
Figure 4.26	Localisation #3.2 après segmentation sémantique avec prise en compte du voisinage .....	133
Figure 4.27	Localisation #3.3 après segmentation sémantique avec prise en compte du voisinage .....	134
Figure 4.28	Problématique des classes pour l'apprentissage .....	136
Figure 5.1	Schéma de la proposition RandLA-Net Tirée de Hu et al. (2019).....	139
Figure 5.2	Schéma de la proposition PointNet++ Tirée de Qi, Yi, Su, & Guibas (2017).....	141

Figure 5.3 Schéma de la proposition PVCNN Tirée de Liu, Tang, Lin, & Han (2019).....143

Figure 6.1 a) Nuage de points obtenu avec la photogrammétrie par drone avec les attributs RGB obtenu lors de l'acquisition. b) Attributs R, G et B obtenu à l'aide d'un réseau de neurones prenant en entrée le vecteur d'attributs calculé.....155





## LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES

2D	2 Dimensions
3D	3 Dimensions
ACP	Analyse de Composantes Principales
ALS	<i>Airborne Lidar Scanner</i>
ASPRS	<i>American Society of Photogrammetry and Remote Sensing</i>
CAO	Conception Assistée par Ordinateur
CNN	<i>Convolutional Neural Network</i>
CRF	<i>Conditionnal Random Field</i>
CRP	<i>Close-Range Photogrammetry</i> .....
FN	<i>False Negative</i>
FP	<i>False Positive</i>
FPS	<i>Farthest Point Sampling</i>
GTA V	<i>Grand Theft Auto V</i>
HLS	<i>Hue Light Saturation</i>
IoU	<i>Intersection over Union</i>
LIDAR	<i>LIght Detection And Ranging</i>
MCC	<i>Matthew's Correlation Coefficient</i>
MIoU	<i>Mean Intersection over Union</i>
MLS	<i>Mobile Lidar Scanner</i>
MNS	Modèle Numérique de Surface
MNT	Modèle Numérique de Terrain
MRF	<i>Markov Random Field</i>

NN	<i>Neural Network</i>
OA	<i>Overall Accuracy</i>
PI	Partenaire Industriel
Pts	Points
ReLU	<i>Rectified Linear Unit</i>
RGB	<i>Red Green Blue</i>
RGB-D	<i>Red Green Blue - Depth</i>
SVM	<i>Support Vector Machine</i>
TLS	<i>Terrestrial Lidar Scanner</i>
TN	<i>True Negative</i>
TP	<i>True Positive</i>
ULS	<i>Unmanned Lidar Scanner</i>

## **LISTE DES SYMBOLES ET UNITÉS DE MESURE**

### **UNITÉS DE BASE**

m	mètre (unité de longueur)
cm	centimètre (unité de longueur)
km	kilomètre (unité de longueur)
Go	Giga octet (unité de mesure de stockage en informatique)

### **SYMBOLES DE BASE**

x	axe spatial des abscisses
y	axe spatial des ordonnées
z	axe spatial des cotes



## INTRODUCTION

L'accessibilité croissante des technologies de télédétection ou de mesure à distance, telles que les lidars et la photogrammétrie, engage leur déploiement dans de nombreux domaines allant de l'archéologie à la navigation autonome pour voiture ou robot. Ces technologies permettent en général de numériser à distance des parcelles de terrain, des chantiers, des intérieurs de bâtiments ou tous autres objets afin de pouvoir travailler ou archiver le modèle numérisé dans le cadre d'applications très variées. Le modèle obtenu peut être de différentes natures. Nous parlerons, par la suite, principalement de nuage de points, dont un exemple est présenté Figure 0.1 avec sa structure de données (coordonnées et attributs). La Figure 0.1 représente l'un des six nuages de points qui ont été utilisés dans nos études. Ces nuages de points proviennent de notre Partenaire Industriel (PI). Afin d'anonymiser l'emplacement, un chiffre a été associé à chaque localisation. On retrouve par exemple ici la Localisation #5. Un nuage de points correspond à l'agrégation de différents points mesurés sous forme de nuages en prenant les coordonnées  $(x, y, z)$  dans l'espace tridimensionnel. Chaque point peut avoir d'autres attributs que ces coordonnées (exemple : couleur, affiliation, intensité).

Ces différences d'échelles, allant de la centaine de km<sup>2</sup> pour des forêts ou littoraux à quelques cm<sup>2</sup> pour des artefacts de musée ou les sites archéologiques, dépendent de la nature du projet et de la technologie employée. Il est plutôt rare de voir des projets qui nécessitent ou utilisent plusieurs technologies de mesure à distance en même temps. Toutefois, dans certains cas spécifiques, et c'est précisément le cadre de notre projet, l'obtention du nuage de points final n'est possible qu'en croisant et consolidant les informations obtenues par différentes technologies. Ce travail de collaboration métrologique est nécessaire dans ces cas pour obtenir des nuages de points plus complets et plus riches (attributs) que la numérisation de terrains accidentés impose. On peut citer comme exemple l'utilisant un lidar aéroporté en collaboration avec un lidar terrestre.

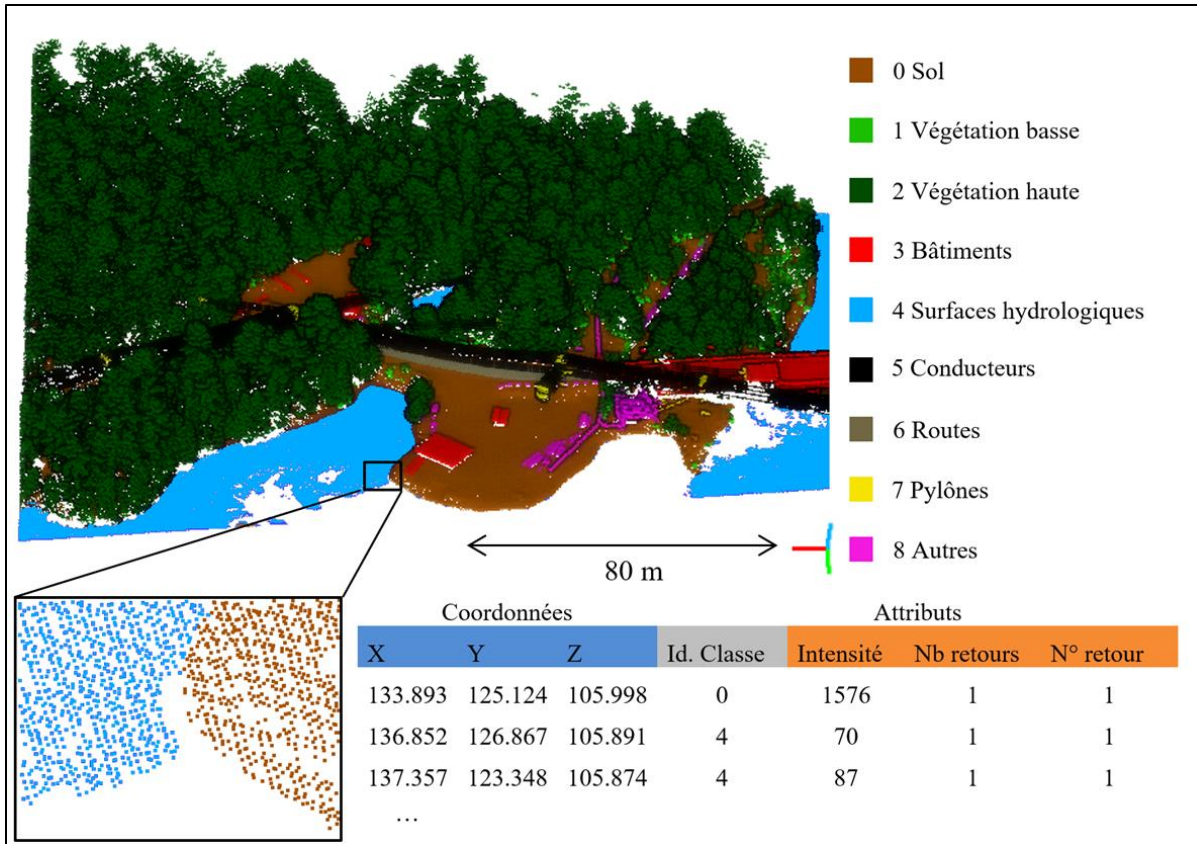


Figure 0.1 Exemple de nuage de points représentant la Localisation #5 avec la structure de ses données (coordonnées et attributs)

Cette collaboration de différentes méthodes de captation pose la question de l'exploitation en synergie des différentes données captées par chaque technologie.

Un projet de numérisation peut être découpé en trois parties. La première phase se situe avant l'acquisition, elle consiste en la préparation du terrain ainsi qu'aux vérifications des lois et normes en fonction du lieu que l'on veut numériser. La deuxième phase est l'acquisition en elle-même. Enfin, la dernière phase est celle du traitement des données (nettoyage, consolidation, géoréférencement, fusion, étiquetage).

C'est par exemple le cas pour l'utilisation des données pour la navigation des voitures autonomes. Pour que le véhicule puisse se constituer une carte de son environnement, il est nécessaire de segmenter sémantiquement le nuage de points capté. La segmentation sémantique réalise le découpage du nuage de points selon la sémantique des points présents.

En prenant l'exemple de la voiture autonome, un point peut appartenir à une entité comme « Routes », « Bâtiments » comme illustré sur la Figure 0.1. La segmentation sémantique revient alors à donner à chaque point la classe à laquelle il appartient. La segmentation sémantique est ici représentée par les différentes couleurs des points (un ombrage artificiel a été utilisé afin de reproduire l'aspect 3D du nuage de points).

Dans le cas de notre partenaire, l'usage *in fine* du nuage de points peut être diversifié :

- Réaliser une modélisation des installations et/ou des bâtiments afin d'avoir un modèle tel que construit (*As is*). Ces données permettent d'identifier des fissures, de faire un suivi des bâtiments et des structures tout au long de leur cycle de vie (glissement, tassement différentiel).
- Calculer un volume de biomasse afin de prévoir le matériel nécessaire pour enfouir un câble ou élaguer la végétation autour de câbles.
- Identifier des débris, suivre l'évolution de l'érosion des sols, ou prévoir des zones inondables proches des surfaces hydrologiques numérisées.
- Identifier les zones praticables selon les moyens de transports et de maintenance nécessaires.

Le présent projet de thèse s'intéresse à la combinaison des différentes données et informations fournies par différentes technologies de mesure pour permettre les usages ci-dessus. Le postulat de base est que l'utilisation appariée de plusieurs technologies sur un même site de numérisation permettrait d'acquérir plus d'information qu'au travers d'une seule technologie. Dans ce projet, les données sont directement fournies par le PI qui a été anonymisé. Ces données proviennent de deux à quatre technologies en fonction du site de numérisation (terrain urbain, terrain escarpé, forêt). Les technologies à disposition du PI sont : (i) le lidar aéroporté à bord d'avion, (ii) le lidar aéroporté à bord d'hélicoptère, (iii) le lidar terrestre et (iv) un nuage de points produit à l'aide de photogrammétrie par drone.

Le projet est constitué de plusieurs étapes qui peuvent être découpées comme suit :

- La propagation des attributs entre les différentes sources est nécessaire dans le cas d'un nuage de point comportant plusieurs sources. Elle permet d'avoir des nuages de points avec les mêmes attributs.
- La segmentation sémantique permet de segmenter le nuage de points selon les objets qu'il représente. Un résultat de segmentation sémantique est visible dans la Figure 0.1.
- La segmentation en entités distinctes permet de segmenter chaque classe en ses différentes entités qui la compose. Dans le cas de bâtiments par exemple, cette division peut être schématisée par la Figure 0.2.



Figure 0.2 Schématisation d'une segmentation de classe en entités distinctes

- Le choix d'une source permet de ne conserver qu'une seule des sources. Cette sélection peut mettre en évidence certaines parties du nuage de points tout en allégeant le nuage de point en ne présentant qu'une partie des données.

La thèse présentée dans ce manuscrit traite l'ensemble de ces points, cependant la majeure partie du travail porte sur les deux premiers points (segmentation sémantique et propagation d'attributs) et est découpée de la manière suivante :

Le **CHAPITRE 1** est une revue de la littérature comportant trois parties. La première partie porte sur les technologies de captation à distance qui sont utilisées dans ce projet. Cette première partie se conclut avec les différents jeux de données de nuages de points disponibles dans la littérature. La deuxième partie porte sur la segmentation sémantique des nuages de points et les métriques utilisées afin de valider la segmentation sémantique. Enfin, la troisième partie porte sur la fusion de données géospatiales. Cette partie permet d'illustrer certaines des



nouveautés apparaissant au travers de notre sujet et de se positionner par rapport à l'état de l'art dans le domaine.

Le **CHAPITRE 2** décrit la problématique et la question de recherche ainsi que l'objectif et les sous-objectifs qui y sont liés. La méthodologie de recherche est ensuite présentée. Ce chapitre permet aussi de présenter nos postulats, les frontières du projet, ainsi que les outils utilisés au cours de nos travaux. Ce chapitre permet enfin d'expliquer les hypothèses retenues dans le projet.

Le **CHAPITRE 3** présente la proposition qui permettra de répondre à la question de recherche. Ce chapitre est décomposé en trois parties. La première partie porte sur l'étude de la segmentation sémantique en utilisant une seule source, la deuxième partie porte sur l'utilisation de plusieurs sources de données. Enfin, la dernière partie est une synthèse de la proposition. Cette dernière partie présente aussi les différentes contributions réalisées.

Le **CHAPITRE 4** présente les résultats et les études de cas réalisés à l'aide de notre proposition. Ce chapitre est introduit avec une présentation des différents jeux de données ainsi que les localisations anonymisées où les jeux de données ont été relevés. Ces différents jeux de données sont par la suite utilisés pour évaluer les performances de notre proposition. La suite du chapitre est découpée selon les différentes parties de la proposition et leur évaluation. Enfin, une synthèse revenant sur les différents résultats conclut ce chapitre.

Le **CHAPITRE 5** expose une comparaison des performances de notre approche avec d'autres approches présentées dans la revue de la littérature. Ce chapitre est structuré en trois parties. La première porte sur les algorithmes sélectionnés pour la comparaison. La deuxième sur les résultats comme tels, et la troisième partie est une discussion permettant de confronter et comparer les résultats obtenus grâce à notre approche avec ceux obtenus grâce aux algorithmes retenus de la littérature.

La **CONCLUSION** achève ce manuscrit en revenant sur l'originalité de nos travaux en rappelant les contributions, les principaux résultats et en suggérant des pistes et des recommandations pour de futures recherches sur le sujet.

# CHAPITRE 1

## REVUE DE LA LITTÉRATURE

### Sommaire

1.1	Introduction.....	8
1.2	Numérisation de terrains de grande échelle.....	9
1.2.1	Principes physiques de captation.....	11
1.2.2	Techniques et technologies de captation associées.....	12
1.2.2.1	Lidar terrestre.....	13
1.2.2.2	Lidar aéroporté.....	13
1.2.2.3	Photogrammétrie.....	15
1.2.3	Synthèse des technologies de captation.....	18
1.2.4	Jeux de données existants.....	19
1.3	Segmentation sémantique de nuages de points.....	24
1.3.1	Segmentation s'appuyant sur les bordures – <i>Edge-Based</i> .....	27
1.3.2	Segmentation par croissance de régions – <i>Region Growing</i> .....	27
1.3.3	Segmentation par ajustement de modèles – <i>Model Fitting</i> .....	28
1.3.4	Segmentation reposant sur la théorie des graphes – <i>Graph-Based</i> .....	29
1.3.5	Segmentation par apprentissage machine.....	29
1.3.5.1	Apprentissage machine – <i>Machine Learning</i> .....	30
1.3.5.2	Apprentissage profond – <i>Deep Learning</i> .....	30
1.3.6	Méthode hybride de segmentation.....	33
1.3.7	Caractéristiques supplémentaires.....	34
1.3.8	Évaluation et validation de la segmentation sémantique.....	39
1.4	Fusion de données géospatiales.....	43
1.4.1	Fusion au niveau des données d'entrée.....	44
1.4.2	Fusion au niveau des attributs.....	45
1.4.3	Fusion au niveau des résultats.....	45
1.4.4	Fusions multiniveaux.....	47
1.5	Synthèse de l'état de l'art.....	47
1.5.1	Positionnement de notre sujet par rapport à l'état de l'art.....	48
1.5.2	Critiques de l'état de l'art.....	49

## 1.1 Introduction

Ce premier chapitre porte sur la revue de la littérature concernant notre projet de thèse. Après avoir évoqué les étapes qui constituent un projet de captation numérique de terrain dans l'introduction de cette thèse, la revue de la littérature porte dans un premier temps sur les technologies de captation à distance. Le fonctionnement physique de ces technologies ainsi que leurs avantages et inconvénients sont rappelés. Cette partie sur les technologies de captation se termine par la présentation de jeux de données qui existent dans la littérature et qui peuvent être utilisés afin de comparer différentes approches.

La deuxième partie de cette revue de littérature porte sur la segmentation sémantique automatique des nuages de points. Cette partie permet de présenter les différentes familles d'algorithmes de segmentation sémantique automatique qui existent, ainsi que les métriques utilisées afin de mesurer un résultat et de le valider ou non.

La troisième partie porte sur la fusion de données géospatiales. Elle présente d'abord le principe de la fusion de données puis ensuite les différentes approches qui existent dans la littérature.

Enfin, une synthèse nous permet de positionner notre sujet vis-à-vis de la revue de littérature et permet une critique de notre état de l'art.

Cette revue de la littérature a été principalement réalisée à l'aide de bases de données de publications scientifiques telles que *Scopus*, *Web of Science* ou encore *Google Scholar*. Les journaux scientifiques ainsi que les conférences d'où sont extraits ces articles, ainsi que les principales catégories dans lesquelles ces revues scientifiques apparaissent au sein d'outils tels que *Journal Citation Reports*, ont aussi été investigués. Ce type d'outil permet aussi de partiellement retirer certaines revues prédatrices. Le but est d'avoir une revue de littérature la plus exhaustive possible vis-à-vis de la question de recherche. Au vu des choix effectués, comme le fait d'utiliser certaines bases de données, il est possible que certains articles n'aient pas été identifiés. C'est pourquoi d'autres outils tels que *ResearchRabbit* ont pu être utilisés afin de compléter notre revue de littérature par d'autres articles. L'outil *ResearchRabbit* identifie des articles en lien avec ceux de notre littérature en cherchant les références et

citations des articles ajoutés dans notre revue. Cet outil permet aussi de visualiser les articles de notre revue sous forme de graphes, de voir les liens entre les différents articles sélectionnés et de découvrir de nouveaux auteurs.

## **1.2 Numérisation de terrains de grande échelle**

Les technologies de captation disponibles utilisent différents principes physiques, décrits dans le prochain paragraphe. La taille du terrain à numériser est l'un des principaux paramètres de leur catégorisation, visible dans la Figure 1.1 (Barbero & Ureta, 2011). Cette figure représente les différents principes physiques de captation en fonction de la distance  $z$  de captation, ainsi que de la résolution relative à cette distance  $z$ . On retrouve les mêmes ordres de grandeur dans la Figure 1.2 proposés par (Franceschini, Galetto, Maisano, & Mastrogiacomo, 2014) qui, quant à elle, présente directement les solutions techniques.

Au vu des distances de captation du projet considéré dans cette thèse, qui sont de l'ordre du mètre à la centaine de mètres, les techniques de triangulation et de temps de vol seront présentées plus en détail. Elles correspondent aux principes physiques pouvant au mieux répondre aux problématiques de distance que demande la numérisation de terrains et d'ouvrages existants.

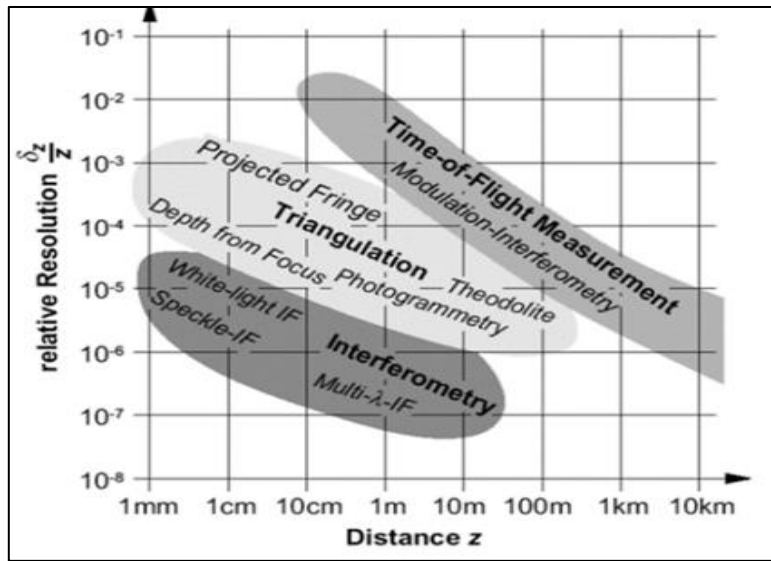


Figure 1.1 Représentation des zones d'utilisation des différents principes physiques pour la numérisation à distance en fonction de la distance de captation  $z$  et de la résolution relative à cette distance  
Tirée de Barbero & Ureta (2011)

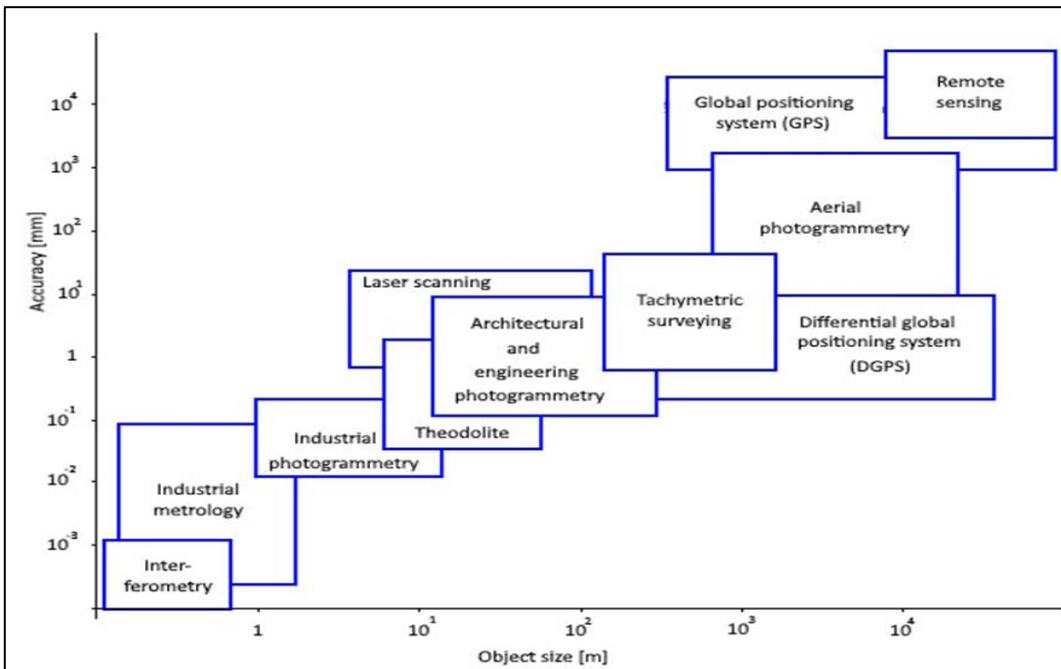


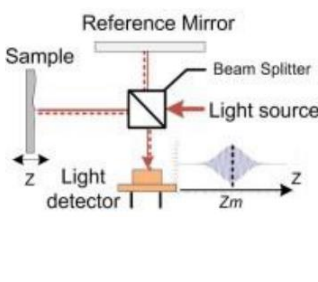
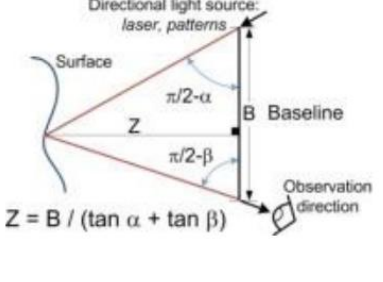
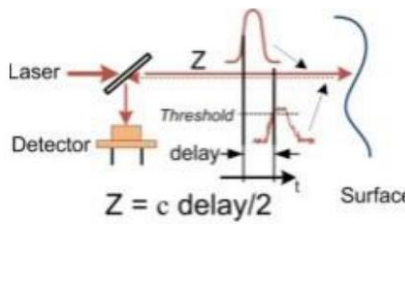
Figure 1.2 Représentation des principales technologies de captation à distance en fonction de la taille de l'objet à numériser et de la précision  
Tirée de Franceschini et al. (2014)

### 1.2.1 Principes physiques de captation

Les technologies d'acquisitions à distance fonctionnent, à l'heure actuelle, sur les trois principes physiques suivants : l'interférométrie, la triangulation et le temps de vol.

L'**interférométrie** utilise 2 sources cohérentes, afin de créer une figure d'interférence due à la différence de distance que les rayons parcourent. Ces deux sources sont souvent une même source qui est séparée en deux à l'aide d'un miroir semi-réfléchissant. Les rayons séparés prennent chacun un chemin différent avant de venir créer sur le détecteur une figure d'interférence. Il est ensuite possible de calculer la différence de distance entre les deux chemins et donc de savoir la distance à laquelle se trouve l'objet numérisé. Ce principe physique, représenté dans le Tableau 1.1 a) (Beraldin, Mackinnon, & Cournoyer, 2015), est, à l'heure actuelle, surtout utilisé pour mesurer des petites dimensions, mais permet d'avoir des mesures allant jusqu'au mètre ou à la dizaine de mètres comme le montre la Figure 1.1. Il est aussi l'un des principes les plus précis pour la captation à distance.

Tableau 1.1 Schématisation des principes physiques permettant la numérisation à distance  
Tiré de Beraldin et al. (2015)

a) Interférométrie	b) Triangulation	c) Temps de vol
		

La **triangulation** correspond à la mesure de distance par calcul angulaire. Dans le cas d'un laser, une source émet un signal qui est observé (avec un certain angle) par un capteur. Connaissant les angles d'émission et de réception, ainsi que la distance entre la source et le capteur, la distance de l'objet est calculable. Le cas de la triangulation dans la photogrammétrie

sera traité de manière plus précise ultérieurement dans ce rapport. Le principe est représenté dans le Tableau 1.1 b).

Le **temps de vol** repose sur le calcul du temps entre l'émission du signal et le retour du signal une fois qu'il a été réfléchi par la surface à mesurer. Le signal peut être émis dans différentes longueurs d'onde selon le résultat voulu. Le domaine visible permet par exemple d'atteindre des objets situés à de plus longues distances. L'utilisation d'infrarouge permet d'avoir moins de perturbation et est plus facilement différenciable des sources de lumière artificielle pour le capteur. Le principe est représenté Tableau 1.1 c).

Chacun de ces trois principes est ensuite développé en différentes technologies ou techniques. La triangulation est par exemple utilisée dans la photogrammétrie, la projection de franges, l'utilisation de théodolite et le *Depth From Focus*. Ces familles et technologies se retrouvent dans la Figure 1.1. Cette représentation est utile pour une première sélection du principe physique d'acquisition à choisir selon l'utilisation nécessaire (distance de captation ou résolution).

Dans notre cas, les principes qui seront utilisés sont la triangulation avec de la photogrammétrie par drone et la mesure de temps de vol avec les différents types de lidars. Ces choix sont dus à la nature des terrains à numériser qui sont de grandes dimensions.

### **1.2.2 Techniques et technologies de captation associées**

Les principes physiques de captation à distance peuvent être utilisés à travers différentes technologies. Cette section porte sur les principales technologies de captation à distance et plus particulièrement sur les technologies utilisées lors de notre projet.

Le LiDAR-*Light Detection And Ranging* est à l'heure actuelle l'une des technologies les plus utilisées pour l'acquisition à distance. On le retrouve sous différentes formes. Le lidar peut être utilisé sur des bases mobiles ou sur des trépieds fixes. Son fonctionnement peut varier selon la technologie (laser fixe, en rotation sur une plateforme, etc.), mais le principe reste le même. Dans le cas de ce sujet, les LiDAR terrestre (*TLS-Terrestrial Laser Scanner*) et aéroporté



(ALS-*Airborne Laser Scanner*) à bord d'avions ou d'hélicoptères seront mis en avant comme il s'agit des technologies dont les nuages de points à notre disposition sont extraits. Les MLS (*Mobile Laser Scanner*) et ULS (*Unmanned Laser Scanner*) ont des fonctionnements similaires et sont pour le MLS souvent à bord ou sur des véhicules comme des voitures quant aux ULS ils sont montés sur des drones.

### **1.2.2.1 Lidar terrestre**

Les lidars terrestres sont des capteurs actifs utilisant le principe de mesure par temps de vol. Leur utilisation est complexe et nécessite la plupart du temps des personnes qualifiées (minimum deux) afin de le manipuler sans danger (le lidar pouvant projeter un laser dans le domaine visible, infrarouge ou ultraviolet).

Malgré leur grande zone d'action, une dizaine de numérisations sont généralement nécessaires pour compléter la numérisation de la majeure partie d'un barrage hydraulique (Buffi, Manciola, Grassi, Barberini, & Gambi, 2017). Chaque numérisation peut durer entre 20 à 30 minutes.

La précision des lidars est parmi les meilleures avec un ordre de grandeur du millimètre pour une distance d'acquisition de plusieurs dizaines à centaines de mètres. La technologie de lidar permet d'obtenir un attribut qui correspond à l'intensité de retour du laser. Certaines technologies associent une couleur à chaque point, en plus de l'intensité de retour, donc trois attributs correspondant à la décomposition en Rouge, Vert et Bleu (*Red, Green, Blue RGB*) de la couleur.

### **1.2.2.2 Lidar aéroporté**

Le fonctionnement d'un lidar aéroporté est similaire à celui du lidar terrestre. Il est cependant transporté et utilisé à bord d'un avion (ou hélicoptère) en vol. L'utilisation de cette technologie en vol crée une distribution particulière des points qui se retrouvent en forme de ligne (visible Figure 1.3 a)). Ces nuages de points sont de plus faible densité que le TLS (de l'ordre de 1 à 4 pts par m<sup>2</sup> pour un avion selon la distance du sol et la vitesse de vol). Comme pour le lidar terrestre, le lidar aéroporté est manipulé par un opérateur. La disposition des points est la même

que celle présentée dans la Figure 1.3. Cette technologie permet d'acquérir les informations suivantes :

- L'intensité de retour du laser est l'intensité lumineuse reçue par le capteur après l'aller-retour.
- Le nombre de retours visibles Figure 1.3 b)
- Le numéro du retour visible Figure 1.3 b) avec le 1<sup>er</sup>, 2<sup>e</sup> et 3<sup>e</sup> retour
- L'angle du laser lors de la numérisation est visible à la Figure 1.3 a) (le lidar effectuant des allers-retours perpendiculaires au sens de la marche).

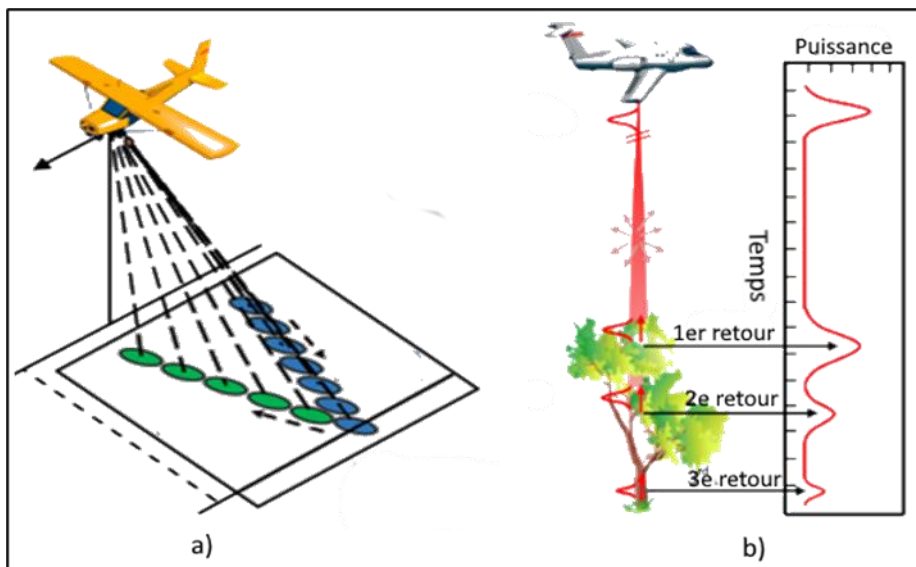


Figure 1.3 a) Schématisation de la numérisation à l'aide d'un lidar aéroporté, b) Schématisation du nombre de retours et numéro de retour

L'altitude de vol et la vitesse d'avance font que le lidar aéroporté est moins précis que le lidar terrestre (Daneshmand et al., 2018). Le lidar aéroporté permet de couvrir de grandes zones géographiques rapidement et à moindre coût. Cependant, sa densité et sa précision peuvent être des inconvénients dans certains projets.

### 1.2.2.3 Photogrammétrie

La photogrammétrie est une technique permettant la création d'un nuage de points provenant d'images. Elle est définie comme « la technique permettant, à partir d'un couple stéréoscopique de photographies, d'étudier et de définir avec précision les formes, les dimensions et la position dans l'espace d'un phénomène quelconque. » (“Le grand dictionnaire terminologique,” 2022). Même si l'on peut utiliser la photogrammétrie de manière terrestre, la plupart des cas d'usage sont réalisés à l'aide de photographies aériennes. Contrairement au lidar, il s'agit donc d'un capteur passif (il recueille l'énergie rayonnante de la zone). Il existe différents types de photogrammétrie, la moins coûteuse étant la *Structure from Motion* capable de reconnaître la position du capteur en fonction des photographies prises à différents points de vue. Néanmoins la *Close-Range Photogrammetry* (CRP) est la plus utilisée grâce à sa précision et l'*Aerial Photogrammetry*, la plus versatile dans le cas de terrains accidentés (Daneshmand et al., 2018). Ces deux principes ne sont pas opposés. La CRP signifie que l'objet numérisé se situe à moins de 1000 pieds, soit environ 300m (Luhmann, 2010). La photogrammétrie aérienne signifie que les photos sont aériennes.

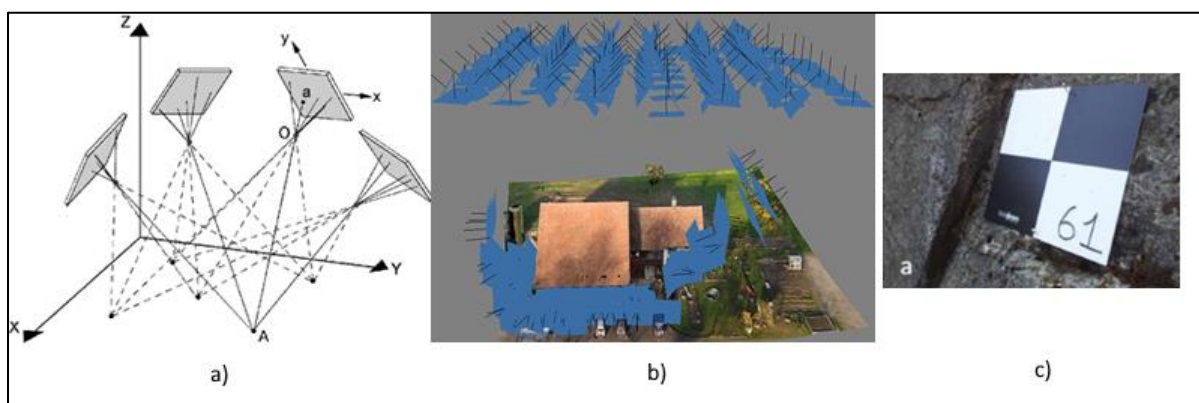


Figure 1.4 a) Schéma du principe physique de la triangulation, b) Représentation d'un nuage de points avec les différentes positions du capteur photographique et, c) Artefact d'acquisition nécessaire pour augmenter la précision de la photogrammétrie  
Tirée de Beraldin et al. (2015)

Le principe de fonctionnement de la photogrammétrie est de collecter différentes photographies, sous différents angles d'un même objet ou d'une même scène afin de

reconstruire la scène. Cette reconstitution est réalisée à partir de points de repère et des angles de vues comme le montre le schéma de principe de la Figure 1.4 (Beraldin et al., 2015). La reconstruction est souvent réalisée à l'aide d'un logiciel de calcul spécialisé capable de sélectionner automatiquement les points de référence entre les différentes images et capable de retrouver les positions successives prises par le capteur.

Dans le cas de notre projet, c'est un drone qui prendra les photographies. Le drone permet de capter l'entièreté de la zone à numériser (les zones sont plus ou moins faciles d'accès pour les opérateurs), tout en permettant d'être proche de la zone à numériser si on le souhaite. Le drone permet d'automatiser le vol tout en démocratisant la technique de photogrammétrie aérienne. Afin d'avoir une meilleure précision au niveau des points, des points de contrôle peuvent être placés au sol. Les photographies doivent aussi être prises avec différents angles de vue afin d'augmenter la précision des coordonnées sur les trois axes (x, y et z). Des photographies prises uniquement avec un angle dirigé vers le nadir (angle correspondant à la verticale de la scène observée) réduisent la précision des coordonnées sur l'axe z comme le rappellent (Rossi, Mancini, Dubbini, Mazzone, & Capra, 2017).

L'un des principaux avantages de cette technologie est la vitesse d'acquisition. En effet, selon la taille de la scène et la précision voulue, le temps de vol du drone peut être court. De plus l'acquisition peut être automatisée en programmant le vol du drone auparavant.

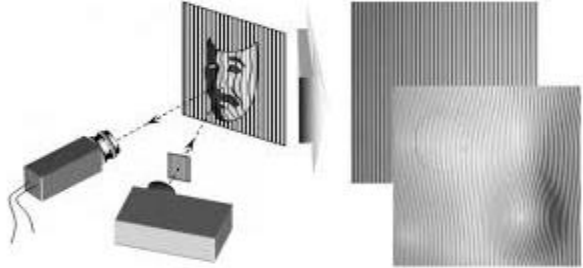

Bien que le temps d'acquisition soit très court, le temps global est rallongé par le traitement des photographies et la reconstruction de la scène globale qui peut être de plusieurs heures. Les logiciels de traitement reconstruisent automatiquement la scène en retrouvant la position d'où a été prise la photographie et en calant les images les unes par rapport aux autres grâce à des points de références.

Un inconvénient majeur vient du fait que les nuages de points sont très liés à la luminosité et aux textures des photographies. Les ombres, la végétation, la neige et l'eau peuvent créer des artefacts ou des aberrations rendant le traitement automatique plus difficile. Enfin, selon la taille et la résolution des images, les données peuvent prendre un espace de stockage non négligeable (de l'ordre de la dizaine de Go) en plus de l'espace de stockage nécessaire aux nuages de points.

Les autres technologies souvent rencontrées pour la télédétection en triangulation sont les suivantes :

- La projection de franges fonctionne sur le principe de triangulation. Un système projette un *patern* de lumière connue et une caméra enregistre la déformation du *patern* sur l'objet à numériser. Ce principe offre des résultats probants dans des salles peu éclairées et pour des objets à numériser de dimension de la taille allant du centimètre au mètre. L'utilisation de lumière bleue permet de numériser des objets avec une lumière ambiante plus importante. Cependant cette technologie n'est pas employée pour les relevés de grandes dimensions et en extérieur (tel que des bâtiments entiers). (Tableau 1.2 a))

Tableau 1.2 Représentation de deux autres technologies utilisant le principe de triangulation : la projection de franges et le théodolite  
Tiré de « Skinexigence - Projection de franges » (2022) et  
« Théodolite Spectra Precision NE-101 pour l'arpentage | Cansel » (2022)

a) Projection de franges	b) Théodolite
	

- Le théodolite : est une technologie utilisant la triangulation et est utilisée par les arpenteurs-géomètres lors de relevé de terrain. Elle est très précise, mais lente du fait qu'elle ne relève qu'un point à la fois. Cette technologie n'est donc pas adaptée pour la numérisation d'un terrain ou d'un ouvrage complet. Cependant, cette technologie peut être utilisée afin de relever les points de contrôle avec une plus grande précision. (Tableau 1.2 b))

### 1.2.3 Synthèse des technologies de captation

Bien qu'il existe d'autres technologies de captation à distance, cette synthèse reviendra principalement sur les technologies à disposition pour ce projet. Les valeurs et caractéristiques ont été prises de différents ordres de grandeur disponibles dans les travaux (Barbero & Ureta, 2011 ; Franceschini et al., 2014 ; Huber et al., 2010 ; Kyle et al., 2001, 2001 ; Matese et al., 2015 ; Moon, Chung, Kwon, Seo, & Shin, 2019 ; Schmitt et al., 2016 ; Siebke et al., 2018 ; Summan et al., 2015 ; Xie, Tian, & Zhu, 2020 ; Yan & Shaker, 2018) ainsi qu'avec les données fournies par le PI et sont visible dans le Tableau 1.3.

Tableau 1.3 Caractéristiques et certains ordres de grandeurs des technologies et méthodes de numérisation qui ont été utilisées pour l'obtention de nuage de points

Critères	Photogrammétrie	Lidar terrestre (TLS)	Lidar aéroporté (ALS)
<b>Vitesse d'acquisition</b>	Rapide (mètres carrés par minute)	Rapide	Très rapide (selon la vitesse de vol)
<b>Vitesse de traitement</b>	Lent	Rapide	Rapide
<b>Densité du nuage</b>	De >10 à <400pts/m <sup>2</sup> dépend du capteur de l'appareil utilisé et du nombre de photographies réalisées	> 100 pts/m <sup>2</sup> selon la distance entre l'objet numérisé et le capteur	< 20 pts/m <sup>2</sup> selon la vitesse de vol et l'altitude
<b>Précision du nuage</b>	Moyenne à faible (2 à 30 cm)	Très bonne (ordre du millimètre)	Bonne (< 15 cm)
<b>Bruitage</b>	Très bruité	Peu bruité	Bruité
<b>Portée</b>	Ordre de la dizaine ou centaine de mètres	Jusqu'au kilomètre	Ordre du kilomètre
<b>Coût de la technologie</b>	Faible (1k à 15k\$)	Coûteux (50k\$ à 100k\$)	Faible (≈35\$ le km <sup>2</sup> ) avec un avion

#### 1.2.4 Jeux de données existants

Afin de comparer les performances (rapidité d'exécution et exactitude de la segmentation ou de la classification) de différentes méthodes de segmentation sémantique, des bases de données sont utilisées. Les bases de données peuvent aussi servir pour l'entraînement des algorithmes ayant besoin de phases d'entraînement avant la classification. La majeure partie de ces bases de données sont des données réelles, captées et classées par des experts. Ces tâches étant fastidieuses, peu de jeux de données existent et certaines publications, telles que (H. Huang et al., 2017), ou (B. Wu, Wan, Yue, & Keutzer, 2017) essaient donc de créer leurs propres jeux de données. L'utilisation des jeux de données existants, ou des techniques plus inventives, permet d'augmenter le nombre de jeux de données avec l'utilisation de scènes tirées de jeux vidéo permet de créer de nouveaux jeux de nuage de points. Huang *et al.* ont par exemple créé de nouveaux nuages de points en translatant aléatoirement dans un petit espace les points existants. Wu *et al.* ont quant à eux utilisé le jeu GTA V afin de créer leur jeu de données. Le jeu de données provenant de GTA V est donc un jeu de données synthétique comme il existe des méthodes permettant de générer de tels jeux de données avec l'utilisation de modèle CAO dans (Montlahuc, Shah, Polette, & Pernot, 2019) ou encore dans (M. Chen et al., 2022 ; M. Chen, Feng, Prasad, McAlinden, & Soibelman, 2020b, 2020a). Ces jeux de données synthétiques doivent être le plus proche possible du réel afin d'apporter des données utilisables pour l'apprentissage des algorithmes dans le cadre de l'apprentissage machine.

Le problème majeur des jeux de données recensés dans la suite de ce paragraphe est qu'ils ne proviennent que d'une technologie. La plupart des nuages de points proviennent de lidars terrestres ou de lidars aéroportés. Des nuages de points artificiels pourraient alors être créés afin de reproduire certaines caractéristiques des nuages de points d'autres technologies.

Comme nous l'avons vu plus haut, certaines méthodes utilisent des images des nuages de points et non les données brutes. Cette technique permet l'utilisation de jeux de données d'images qui sont beaucoup plus répandues et d'utiliser directement des modèles préentraînés.

Les principaux jeux de données de nuages de points utilisés dans la littérature sont les suivants :

- Semantic-8 ou Semantic3D utilisés dans plusieurs publications dont (Landrieu & Simonovsky, 2017) est un jeu de données comportant 15 nuages de points allant de 20 à 280 millions de points pour le jeu de données d'entraînement. Il y a aussi 15 nuages de points allant de 20 à 400 millions de points pour le jeu de données de test. Ce jeu de données correspond à l'acquisition par un lidar terrestre de terrains ruraux ou urbains. Ce jeu de données est en libre accès et a été créé pour pouvoir réaliser des comparaisons de solutions. Il est utilisé par une dizaine de publications. Ces publications utilisent principalement des réseaux de neurones ou des algorithmes de segmentation sémantique automatique.
- S3DIS (Landrieu & Simonovsky, 2017) est un jeu de données contenant des nuages de points et des images (2,5 D) correspondant à l'intérieur de différents bâtiments de l'université de Stanford. Un total de six zones sont scannées avec un total de 13 classes.
- SemanticKITTI et KITTI (Behley et al., 2019 ; Boulch, Saux, & Audebert, 2017) sont des jeux de données qui contiennent des images et des nuages de points. Ces données étant prises à l'aide d'une station positionnée sur une voiture, les données représentent des espaces ruraux et urbains ainsi que des parties de campus. Le jeu compte plus d'une centaine d'images et de nuages de points. Ces données sont surtout utiles pour la recherche concernant les voitures autonomes, car les données classées correspondent aux éléments que l'on peut retrouver dans la circulation.
- MNIST (Boulch, 2019) correspond à un jeu d'images pour entraîner des réseaux de neurones de convolutions. Le jeu de données représente des chiffres écrits à la main.
- ModelNet10 et ModelNet40 (Atzmon, Maron, & Lipman, 2018) sont les jeux de données de l'université de Princeton. Ces modèles CAO représentent des objets de la vie courante. Le ModelNet10 compte 10 classes représentant 10 objets de la vie courante et le ModelNet40, 40 objets différents.
- ShapeNet (Boulch, 2019) représente lui aussi des objets du quotidien avec, cette fois-ci, 55 objets/classes différentes provenant de 51 300 modèles 3D.
- BHCP (H. Huang et al., 2017) *Benchmark* représente lui aussi des objets du quotidien. Ces nuages de points proviennent de modèles 3D.



- Nyuv2 (Maturana & Scherer, 2015) est un jeu de données contenant des images (2,5 D) représentant l'intérieur de bâtiments administratifs et de maisons. Ce jeu comporte 464 scènes classées.
- SUN RGB-D (Rahman, Tan, Xue, Shao, & Lu, 2019) est un jeu de données comportant 10 355 images RGB-D représentant l'intérieur de bâtiments.
- *GTA V Artificial Data Set* (B. Wu et al., 2017) est un jeu de données artificiel créé à l'aide du jeu vidéo GTA V et d'*Addon*. Ce jeu permet de créer des scènes et des images 2,5 D.
- ScanNet (Charles, Su, Kaichun, & Guibas, 2017) est un jeu de données contenant des vidéos RGB-D comportant plus de 1500 numérisations. Ce jeu représente l'intérieur de bâtiments.
- Oakland 3D CVPR 2009 (Weinmann, Jutzi, Hinz, & Mallet, 2015) comporte des nuages de points de la ville d'Oakland. Comportant 1,6 million de points qui sont répartis en trois jeux de données : entraînement, validation et test.
- Paris-Lille-3D (Roynard, Deschaud, & Goulette, 2018) est un jeu de données représentant des rues des villes de Paris et de Lille. Il comporte 50 classes et 140 millions de points.
- Toronto City (S. Wang et al., 2016) est un jeu de données de la ville de Toronto, comportant plusieurs sources, photographies aériennes, panoramas de rue, ALS, MLS ainsi que des cartes intégrant des informations sur les routes et bâtiments. Le jeu de données correspondant au nuage de points provenant de MLS est nommé Toronto-3D.
- RELLIS-3D (Jiang, Osteen, Wigness, & Saripalli, 2022) est un jeu de données de lidar et photographie, prise à l'aide d'un robot terrestre et représente un espace rural.
- D'autres nuages de points peuvent aussi être utilisés tels que des rues de Hong Kong, de Paris, des zones de l'Allier ou de la Garonne ou encore des zones géographiques en Chine (B. Guo, Huang, Zhang, & Sohn, 2015 ; Yuan Li, Wu, & Ge, 2019).

Le travail de (Xie et al., 2020) classe les différents jeux de données selon la source et le contenu du nuage de points (entre numérisations urbaines, de forêts ou à l'intérieur de

bâtiments). On y retrouve 5 jeux de données pour les nuages dérivants de photographies et tous étant des numérisations d’environnement urbain. Une dizaine provenant d’ALS et sont répartis entre des terrains urbains (20 jeux de données), des forêts (2 jeux de données). Les jeux de données provenant de MLS et TLS sont aussi largement dominés par les jeux de données urbains (qui représentent 28 jeux de données). Il existe en revanche un unique jeu de données forestier et trois jeux de données concernant l’intérieur de bâtiments. Cette classification selon le contenu des nuages de points permet de se rendre compte des projets les plus en vue. Ces projets sont portés pour une part par de la navigation de voiture autonome en ville et une autre part par l’agroforesterie ainsi que le dénombrement d’arbres. Les nuages relevés en intérieur sont quant à eux plus utilisés pour de la navigation autonome ou de la préhension dans le cas de jeux de données d’objets.

La majeure partie des jeux de données en libre accès repris dans le Tableau 1.4 répondent à des problématiques de navigations, de voitures autonomes, de drones, ou servent pour la reconnaissance d’objets et de végétations. Nos travaux s’inscrivent dans un cadre plus large allant de la surveillance de terrain, pour la détection de débris, l’identification de zones praticables, comme décrit dans l’introduction. Les jeux de données spécifiques au PI sont donc utilisés dans la suite de la thèse puisque ceux présents dans la littérature n’offrent pas la diversité et les problématiques des jeux de données fournies par le PI. Ces jeux de données ont été captés avec les technologies du PI et traités, classés selon leur besoin.

Tableau 1.4 Tableau des principaux jeux de données répertoriés en ligne

<b>Nom du jeu de données</b>	<b>Technologies</b>	<b>Natures des données</b>	<b>Nombre de classes</b>	<b>Données représentées</b>	<b>Nombre de données</b>
Semantic-8	TLS	Nuage de points	8	Urbain	15 nuages
S3DIS	Lidar	Nuages de points	13	Intérieur de bâtiments	6 nuages

Tableau 1.5 Tableau des principaux jeux de données répertoriés en ligne (suite)




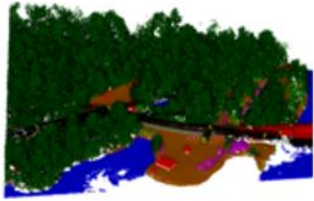
Nom du jeu de données	Technologies	Natures des données	Nombre de classes	Données représentées	Nombre de données
<i>GTA V Artificial Data Set</i>	Non spécifié	Images	19	Urbain	24966 images
ScanNet	Non spécifié	Vidéos RGB D	20	Intérieur de bâtiments	1513 scènes
Oakland 3D	MLS	Nuages de points	5	Urbain	1 nuage
Paris-Lille-3D	MLS	Nuages de points	50	Urbain	3 nuages
Toronto-3D	MLS	Nuages de points	8	Urbain	4 nuages
RELLIS-3D	MLS et caméra	Nuages de points et images	20	Rural	13556 nuages 6235 images
SemanticKITTI	MLS	Nuages de points	28	Urbain route	11 nuages
MNIST	Non spécifié	Images	10	Chiffres	
ModelNet10 et ModelNet40	Non spécifié	CAO	10 et 40	Objets	12311 modèles (ModelNet40)
ShapeNet	Non spécifié	CAO	55	Objets	51300 modèles
BHCP	Non spécifié	CAO	4	Objets	404 modèles
Nyuv2	Microsoft Kinect	Vidéos RGB D	894	Intérieur de bâtiments	1449 images
SUN RGB-D	Intel Realsense Asus Xtion Kinect v1 Kinect v2	Images RGB D	19	Intérieur de bâtiments	10335 images

### 1.3 Segmentation sémantique de nuages de points

La segmentation sémantique automatique de nuages de points est un domaine devenu plus populaire avec l'augmentation de la puissance de calcul et l'avènement de technologies et techniques telles que les réseaux de neurones artificiels. La segmentation sémantique est encore largement réalisée manuellement ou en partie de manière manuelle, mais cette tâche répétitive et chronophage fait l'objet de projet d'automatisation, comme c'est le cas dans notre thèse.

La segmentation sémantique peut être définie comme « une technique d'analyse de données dans laquelle celles-ci sont réparties dans des groupes distincts sur la base de leur sémantique » (« Le grand dictionnaire terminologique », 2022). La différence entre segmentation sémantique et classification de nuages de points est visible dans le Tableau 1.6. Cette différence porte sur le fait de vouloir classer l'ensemble du nuage ou de vouloir classer les points qui composent le nuage. En effet, dans le cas de la classification, nous allons donner à l'ensemble du nuage une classe, dans le cas de la segmentation sémantique, le nuage sera découpé et chaque point est classé. Il est possible de remarquer cette différence dans le Tableau 1.6 dans lequel nous avons utilisé les travaux de (Yi et al., 2016) pour la segmentation sémantique de nuages de points représentant des objets. Dans ce cas, la moto est découpée selon les zones d'intérêt avec le siège, les roues, et les poignées par exemple. Cette décomposition peut être utile par exemple dans la problématique de préhension automatique d'objet par un bras robotisé. Dans le cas de classification, le nuage a été classé en « Motocyclette ». Dans le cas de données semblables à nos données, la classification indique si le nuage correspond à une zone de prairie, une zone urbaine, ou encore une forêt, un glacier ou un désert. La segmentation sémantique permet de découper le nuage de points selon que le point appartient au sol, à de la végétation ou à une autre classe.

Tableau 1.6 Représentation d'une classification et d'une segmentation sémantique sur un objet et sur un terrain  
Tirée de Yi et al. (2016)

	Classification	Segmentation sémantique
Objets / pièces	 <p>Motocyclette</p>	 <p> <span style="color: brown;">■</span> Roues  <span style="color: green;">■</span> Phare <span style="color: orange;">■</span> Réservoir  <span style="color: pink;">■</span> Siège <span style="color: blue;">■</span> Poignées         </p>
Terrains / scènes	 <p>Forêt</p>	 <p> <span style="color: brown;">■</span> Sol <span style="color: green;">■</span> Végétation basse  <span style="color: darkgreen;">■</span> Végétation haute <span style="color: pink;">■</span> Autres  <span style="color: red;">■</span> Bâtiments <span style="color: yellow;">■</span> Pylônes  <span style="color: black;">■</span> Conducteurs <span style="color: grey;">■</span> Routes  <span style="color: blue;">■</span> Surfaces hydrologiques         </p>

Comme pour les technologies d'acquisition, de nombreux travaux ont été effectués avec des données 2D, images et photographies, dans le domaine de la segmentation sémantique. Certaines des techniques et méthodes sont donc dérivées de ces approches utilisant des données 2D.

Différentes approches pour segmenter les nuages de points sont possibles. Certaines de ces méthodes sont dérivées de la segmentation d'images. Dans la suite de cet état de l'art, nous

prendrons la structure que l'on retrouve dans les travaux (Grilli, Menna, & Remondino, 2017 ; Nguyen & Le, 2013 ; Xie et al., 2020) qui classent les différentes méthodes de segmentation selon les techniques et types d'algorithmes utilisés (Figure 1.5). Cette figure présente les différentes méthodes disponibles pour segmenter un nuage de points. Dans certains cas, ces méthodes sont des **méthodes de segmentation** et non pas des **méthodes de segmentation sémantique**. C'est par exemple le cas pour les méthodes qui se basent sur : les bordures, la croissance de région, un modèle, la théorie des graphes.

Nous les présenterons dans la suite car de nombreux algorithmes de segmentation sémantique hybride utilisent certaines de ces méthodes bien souvent couplées avec de l'apprentissage machine.

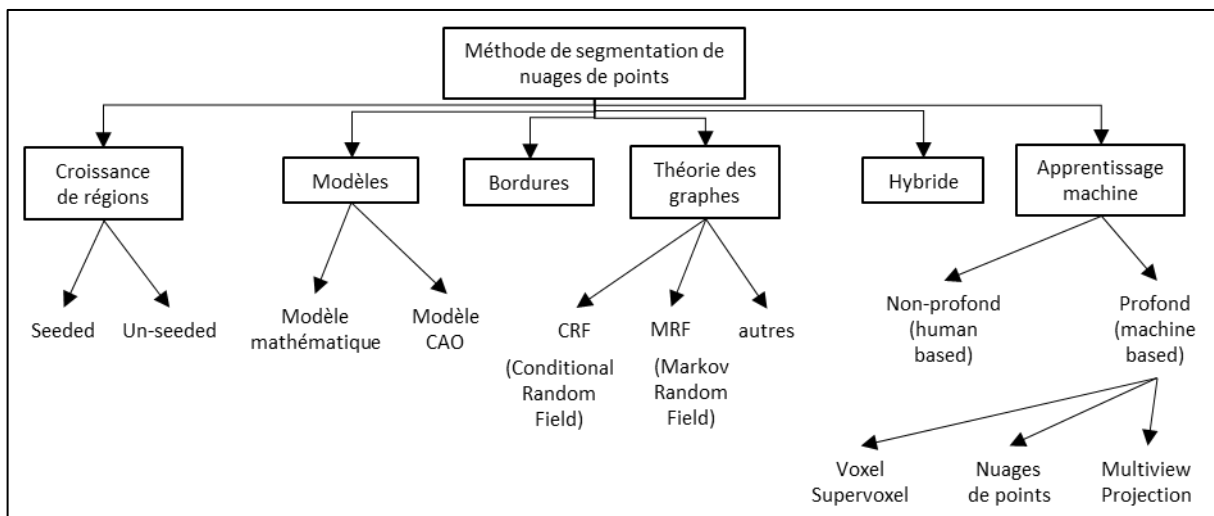


Figure 1.5 Arbre représentant les méthodes permettant la segmentation de nuages de points

Cette deuxième partie de l'état de l'art est découpée en deux avec, dans un premier temps, les différentes approches de segmentation qui existent. La seconde partie porte sur la validation de la segmentation et est étudiée avec une description des métriques que l'on retrouve dans la littérature pour la segmentation sémantique de nuages de points.

### 1.3.1 Segmentation s'appuyant sur les bordures – *Edge-Based*

L'utilisation de bordures pour la segmentation vient à l'origine de la segmentation d'images. Comme décrit dans les études (Grilli et al., 2017 ; Nguyen & Le, 2013), les algorithmes qui utilisent les bordures pour la segmentation de nuages de points sont découpés en deux étapes. La première étape consiste à découper le nuage de points selon les bordures. La seconde étape est de regrouper les points à l'intérieur des bordures comme faisant partie d'un même segment ou d'une même classe. Afin de repérer les bordures, les algorithmes de segmentation utilisent un critère ainsi qu'un seuil. Le critère peut être le changement de courbure ou les changements de normale au nuage de points. Ces méthodes sont beaucoup moins utilisées depuis l'avènement des méthodes qui utilisent de l'apprentissage machine qui sont plus performantes. Elles peuvent cependant être couplées à des méthodes d'apprentissage machine.

### 1.3.2 Segmentation par croissance de régions – *Region Growing*

Dans le cas d'algorithmes de segmentation par *Region Growing*, deux types d'algorithmes existent comme il est rappelé dans les travaux (Grilli et al., 2017, 2017 ; Nguyen & Le, 2013). L'approche *Top-Down*, aussi appelée *Un-Seeded* et les algorithmes *Bottom-Up* qui, quant à eux, utilisent des graines ou *Seed*. Dans les deux cas, un critère ainsi qu'un seuil doivent être utilisés afin de savoir si une nouvelle région doit être créée ou si la région doit être étendue.

Dans le cas des algorithmes *Top-Down*, l'ensemble du nuage sera alors considéré comme une région. La région sera subdivisée en sous-régions par le critère et le seuil donnés. Cette approche est peu utilisée, car il peut être difficile de créer un critère pertinent afin de subdiviser la région de départ. En effet, afin de bien subdiviser la première région en sous-régions, il faut sélectionner un ou des attributs pertinents. Des attributs peuvent être des caractéristiques géométriques. Par exemple, la planéité permettrait de subdiviser les différentes faces d'un cube au niveau des arêtes. Cependant, dans le cas de données plus complexes et inconnues, il est plus difficile de sélectionner les attributs appropriés. Afin d'être efficace, il est donc souvent nécessaire de bien connaître les données d'entrées.

Dans le cas des algorithmes *Bottom-Up*, un nombre de points est choisi comme graine initiale puis chaque région va croître. Cette propagation se réalise de la graine aux points voisins s'ils respectent un certain critère et un seuil définis en amont. On retrouve par exemple les attributs tels que la proximité et la planéité dans l'étude (J. Zhang, Lin, & Ning, 2013). Il est aussi possible de retrouver la méthode par croissance de région dans la création de *Supervoxels* comme dans l'étude (Papon, Abramov, Schoeler, & Worgotter, 2013) où des graines sont utilisées afin de segmenter le nuage de point avant de réaliser une classification. Cette méthode peut néanmoins prendre du temps et est sensible au bruit. Un autre inconvénient principal de cette méthode porte sur le choix des graines, le reste de la méthode dépendant grandement de ce premier choix.

### **1.3.3 Segmentation par ajustement de modèles – *Model Fitting***

La classification utilisant des primitives géométriques est plus spécifique. Elle est peu utilisée dans le cas de segmentation sémantique d'un nuage de points au complet. Dans ce cas, des primitives géométriques sont utilisées afin de se rapprocher au mieux des nuages de points existants. Les nuages de points seront donc segmentés en fonction des formes géométriques que l'on peut retrouver à l'intérieur des nuages. On retrouve principalement cette méthode lorsqu'un modèle précis est connu ou lorsque l'on recherche des formes simples. La méthode RANSAC ainsi que ses variantes sont alors utilisées afin de regrouper les points appartenant à la forme recherchée. Ces techniques sont plus souvent utilisées pour des bâtiments, composés de plans (Ye, Acikgoz, Pendrigh, Riley, & DeJong, 2018). Les méthodes de segmentation utilisant des modèles seront donc utilisées dans des projets avec une forte connaissance des objets à segmenter. L'inconvénient principal de cette technique est qu'il faut avoir un modèle numérique préexistant de l'objet à segmenter, ou encore que l'objet soit simple et qu'il puisse être découpé avec des formes géométriques simples telles que des plans, cylindres, sphères ou cônes.



### 1.3.4 Segmentation reposant sur la théorie des graphes – *Graph-Based*

Les techniques utilisant des graphes proviennent à l'origine de la segmentation d'images. Le principe est tout d'abord de créer un graphe représentant le nuage de points, puis de découper ce graphe selon certaines techniques comme la coupe minimum (*Minimum-Cut* ou min-cut), qui fonctionne en découpant le graphe avec le moins de coupe possible.

Un des moyens de construire le graphe pour un nuage de point est de prendre chaque point comme un nœud et de chercher les  $k$  plus proches voisins (*K Nearest Neighbor*) pour les relier au point comme dans les études (Kelbe, Aardt, Romanczyk, Leeuwen, & Cawse-Nicholson, 2017 ; Yuan Li et al., 2019 ; W. Shi, Rangunathan, & Rajkumar, 2020 ; Tao, Duan, Wei, Lu, & Zhou, 2022 ; Te, Hu, Guo, & Zheng, 2018 ; Y. Wang et al., 2019 ; Xia & Wang, 2018).

Les modèles probabilistes *Conditional Random Field* (CRF) et *Markov Random Field* (MRF) font partie des méthodes de segmentation reposant sur la théorie des graphes. Ils sont souvent utilisés comme post-traitement, comme dans les travaux (Charles et al., 2017 ; B. Guo et al., 2015 ; Yuan Li et al., 2019 ; Pham, Nguyen, Hua, Roig, & Yeung, 2019 ; Tchapmi et al., 2017 ; B. Wu et al., 2017), une fois que l'algorithme de segmentation sémantique a traité le nuage de points. Les CRF permettent une harmonisation de la segmentation sémantique. L'un des avantages de l'approche avec les graphes concerne sa robustesse face aux nuages de points ayant du bruit ou aux nuages ayant une densité de points variable. Parmi les inconvénients on retrouve le fait que l'algorithme peut être plus lent que d'autres algorithmes et est donc peu utilisable pour les besoins en temps réel.

### 1.3.5 Segmentation par apprentissage machine

Les algorithmes d'apprentissage machine peuvent être quant à eux rangés dans deux catégories, l'apprentissage machine (*Machine Learning*) d'un côté et l'apprentissage profond (*Deep Learning*) de l'autre. L'apprentissage machine fait appel à l'intelligence humaine afin de créer des attributs ou de calculer des valeurs afin de pouvoir utiliser ces valeurs pour différencier les cas. L'apprentissage profond quant à lui ne fera pas appel à l'homme et créera de lui-même les attributs, filtres ou valeurs dont il a besoin afin de maximiser la métrique sélectionnée. C'est pourquoi l'apprentissage profond peut aussi être appelé apprentissage

machine de bout en bout (*End-to-End*) pour signifier la non-utilisation d'attributs créés par l'homme. Ces deux approches ont des avantages et des inconvénients et, selon le projet, une approche peut être plébiscitée par rapport à l'autre. Nous les verrons plus en détail dans les deux paragraphes suivants.

### 1.3.5.1 Apprentissage machine – *Machine Learning*

Les algorithmes d'apprentissage machine fonctionnent en utilisant des attributs afin de retrouver des similarités parmi les points à classer. Les algorithmes d'apprentissage machine peuvent être de différentes natures comme le *Support Vector Machine* (SVM) dans le travail (Brodu & Lague, 2012) ou des réseaux de neurones artificiels en passant par les algorithmes de forêt aléatoire (*Random Forest*) de (Weinmann et al., 2015). Ces algorithmes fonctionnent selon les étapes représentées dans la Figure 1.6 (Weinmann et al., 2015 ; Xie et al., 2020). Une première étape consiste à regrouper des points avec l'étape de sélection du voisinage. Des attributs sont alors calculés puis sélectionnés. Enfin, la classification va permettre de classer chaque point parmi les différentes classes qui ont été proposées à l'algorithme. Ici on parle bien de classification, car c'est le point en lui-même qui appartient à une classe et non le nuage de point. Une fois chaque point classé, le nuage de points est alors segmenté sémantiquement.

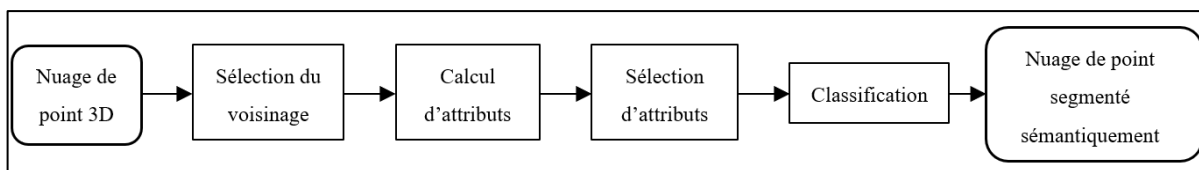


Figure 1.6 Flux représentant les différentes étapes retrouvées dans les projets de segmentation sémantique utilisant l'apprentissage machine non-profond

### 1.3.5.2 Apprentissage profond – *Deep Learning*

Les algorithmes d'apprentissage profond peuvent eux aussi être de différentes natures, mais une majeure partie d'entre eux se compose de réseaux de neurones artificiels. L'apprentissage profond peut être utilisé dans différents cas. Pour la segmentation sémantique de nuages de

points, ces algorithmes peuvent être regroupés dans différentes catégories proposées dans (Mirzaei et al., 2022 ; Y. Wu, Wang, Zhang, & Ogai, 2021 ; Xie et al., 2020). Il est en plus possible de découper ces algorithmes en trois parties (Mirzaei et al., 2022). La première partie de l’algorithme, non obligatoire, correspond à l’échantillonnage du nuage de points. La deuxième partie correspond au regroupement des points. La troisième partie correspond à la cartographie (*Mapping*) du nuage de points.

Une approche étant plus souvent utilisée et provenant à l’origine de la segmentation sémantique d’images est celle des réseaux de neurones de convolution *Convolutional Neural Network* (CNN). Les nuages de points n’étant pas ordonnés et structurés comme l’est une image, les nuages de points doivent être transformés afin d’être utilisables.

Les travaux (Mirzaei et al., 2022 ; Y. Wu et al., 2021 ; Xie et al., 2020) proposent de faire un découpage entre les trois catégories suivantes :

- *Voxel Based* : les voxels (mot-valise né du télescope des mots volume et pixel) sont des pixels 3D comme visibles dans la Figure 1.7 avec la représentation de deux arbres (Pedram Ghamisi et al., 2018). En choisissant une résolution (une taille de voxel), les points du nuage vont être regroupés à l’intérieur de voxels. Cette transformation appelée voxelisation permet de répondre au manque de structure et d’ordonnement du nuage de points original. Il est possible, par la suite, de concaténer les attributs au niveau du centroïde de chaque voxel afin d’utiliser ces nouveaux points comme dans (Tchapmi et al., 2017 ; X. Yang et al., 2021). Cette structure permet l’utilisation de CNN (Z. Hu et al., 2022 ; Liu, Tang, Lin, & Han, 2019 ; Maturana & Scherer, 2015 ; Tang et al., 2020). Selon la taille des voxels, la résolution des données après le traitement peut créer une segmentation sémantique de moins bonne qualité. Si l’on prend une grande résolution pour la voxelisation, le temps de traitement peut être plus long et ne pas permettre certains traitements en temps réel.

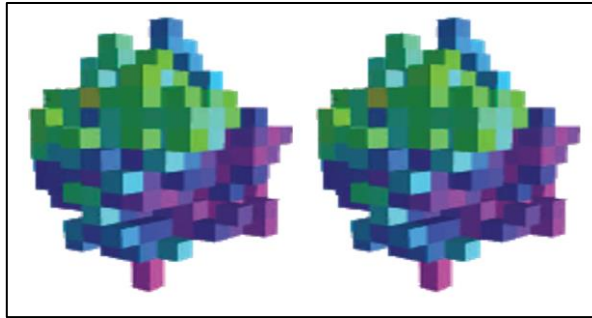


Figure 1.7 Représentation d'un nuage de points sous forme de voxels  
Tirée de Pedram Ghamisi et al. (2018)

- *Multiview Based* : les algorithmes multivues sont des méthodes prenant les nuages de points et les transformant en un ensemble de photographies. Ces photographies ou Snap, comme dans la Figure 1.8, sont réalisés selon différents angles de vues, à différentes échelles ou avec des zooms (H. Huang et al., 2017). Une fois les images réalisées, un CNN est alors utilisé pour segmenter sémantiquement les images. Les labels peuvent alors être transférés des images aux nuages de points. L'un des avantages de ces méthodes est le fait de pouvoir directement ou presque directement utiliser les algorithmes déjà existants pour la segmentation sémantique d'images. L'un des inconvénients est le fait que les images 2D ne reflètent pas totalement l'aspect 3D des nuages de points. Il est aussi parfois nécessaire de créer une colorisation, un ombrage fictif ou de passer par un maillage du nuage de points comme dans (Boulch et al., 2017 ; Xu & Yang, 2018 ; Z. Yang et al., 2017 ; Z. Yang, Tan, Pei, & Jiang, 2018) afin de pouvoir utiliser les algorithmes CNN 2D existants. Il est aussi possible de parler de projections du nuage de points. Par exemple, pour créer une image en 2.5D, il s'agit d'une image 2D prenant en compte la profondeur. Les projections les plus utilisées sont la projection sphérique, comme dans (Y. Zhang et al., 2020), et « *Bird's-Eye View* » comme dans (Zou & Li, 2021).

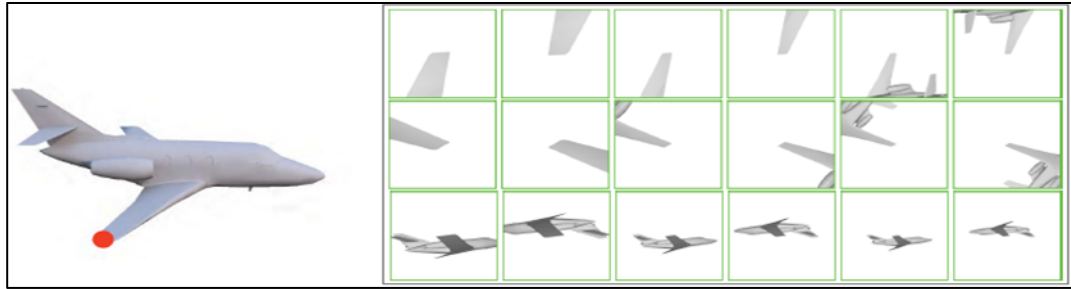


Figure 1.8 Représentation d'un nuage de points d'un avion avec un maillage et de multiples photos pour l'utilisation de CNN 2D avec l'approche *MultiView*  
Tirée de H. Huang et al. (2017)

- *Point Based* : les méthodes se basant directement sur les points sans utiliser des images ou des voxels pour représenter le nuage. On retrouve dans ces méthodes les algorithmes tels que PointNet et PointNet++ (Charles et al., 2017 ; Qi, Yi, Su, & Guibas, 2017). Les travaux utilisant directement les nuages de points comme source peuvent travailler avec des convolutions continues (Y. Guo et al., 2021). Le fait de ne pas passer par de la 2D avec des images ou une perte de résolution liée aux voxels est un avantage selon (Xie et al., 2020).

L'un des inconvénients principaux de l'apprentissage profond est le fait qu'il peut être difficile d'interpréter le fonctionnement ou les résultats obtenus. Un autre inconvénient porte sur les nuages de points à traiter. L'apprentissage profond nécessite un grand nombre de données de préférences représentant le plus fidèlement possibles l'ensemble des classes sous différentes conditions. Les données nécessaires doivent réussir à représenter l'ensemble des classes recherchées, en nombre suffisant, ainsi que l'ensemble de l'hétérogénéité des classes. Ces caractéristiques de représentativité des données font de l'apprentissage machine profond une méthode qui peut être difficile à utiliser dans certains projets comme dans le nôtre.

### 1.3.6 Méthodes hybrides de segmentation

Les méthodes hybrides correspondent aux méthodes utilisant plusieurs des méthodes vues dans la partie segmentation. L'essor des réseaux de neurones profonds ainsi que les résultats obtenus

dans la segmentation sémantique permettent d'imaginer de les coupler avec des méthodes utilisées par le passé.

On retrouve donc des méthodes hybrides d'apprentissage machine couplée avec :

- la croissance de région et les graphes dans les travaux de (Landrieu & Simonovsky, 2017) ;
- la théorie des graphes dans les travaux de (G. Li, Muller, Thabet, & Ghanem, 2019).

En plus de ces hybridations, une partie des méthodes actuelles font appel à un CRF afin d'affiner leur segmentation sémantique. Cette utilisation de deux méthodes différentes, théories des graphes et une autre, peut aussi en partie être vue comme une méthode hybride. Les méthodes hybrides ont de meilleurs résultats en particulier grâce à l'ajout de CRF, mais cet ajout d'étape rend la méthode légèrement plus lente. Le gain sur la précision de la segmentation sémantique est bien souvent privilégié.

### 1.3.7 Caractéristiques supplémentaires

D'autres possibilités de classer les différents algorithmes ont été créées. C'est par exemple le cas de (Nguyen & Le, 2013) qui utilise le terme de segmentation sémantique basé sur les attributs pour parler des modèles utilisant de l'apprentissage machine. Cependant, des attributs peuvent être utilisés dans d'autres méthodes. C'est le cas de la planéité ou du changement de courbure qui sont utilisés dans la segmentation par croissance de régions. C'est pourquoi nous allons regrouper dans cette sous-partie les différentes caractéristiques qui peuvent appartenir aux différentes méthodes de segmentation que nous avons vues tout au long de cette section.

Comme nous l'avons vu dans l'introduction, notre sujet est **multisource**. Différentes sources de données sont prises en compte en entrée afin de réaliser la segmentation sémantique des nuages de points. Dans le cas de plusieurs sources, le terme de fusion de données est souvent utilisé et est plus exploré dans la section 1.4 sur la fusion de données géospatiales. Les sources peuvent fournir des données de même format comme plusieurs lidars donnant chacun un nuage de points, ou des sources fournissant des données de formats divers. Les données multisources

proviennent souvent d'une technologie fournissant un nuage de points, ainsi qu'une technologie fournissant des images.

Une autre caractéristique visible dans plusieurs travaux est le fait d'utiliser une approche à différentes échelles aussi appelées approche **multiéchelles**. Ces approches, visibles dans les travaux de (Bai et al., 2018 ; Engelmann, Kontogianni, Hermans, & Leibe, 2017 ; S.-M. Hu, Cai, & Lai, 2018 ; H. Huang et al., 2017 ; S. Li et al., 2022 ; Yong Li et al., 2019 ; Teruggi, Grilli, Russo, Fassi, & Remondino, 2020 ; Z. Yang et al., 2018) partent du principe qu'une classe peut avoir différentes caractéristiques et différents contextes en changeant d'échelles.

Pour la végétation par exemple :

- une feuille peut donc apparaître comme un plan à l'échelle du millimètre,
- une branche comme une ligne à l'échelle du décimètre et
- un arbre comme un objet 3D à l'échelle du mètre.

La majeure partie des travaux proposent d'utiliser trois échelles comme visibles dans la Figure 1.9 (Brodu & Lague, 2012). Les valeurs des échelles vont dépendre des données à traiter ainsi que des classes à rechercher. L'un des inconvénients du fait d'utiliser trois échelles aux lieux d'une est le temps de calcul qui peut être multiplié par trois si l'on compare à une seule échelle. L'avantage est l'augmentation de la précision de la segmentation sémantique.

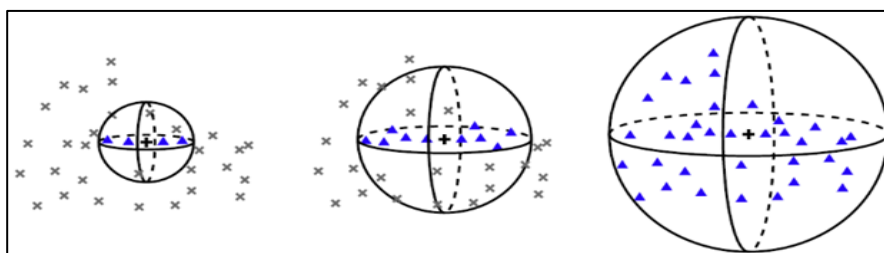


Figure 1.9 Sphères englobantes sur le même nuage de points à différentes échelles afin de calculer des attributs à différentes échelles et prendre en compte différents contextes  
Tirée de Brodu & Lague (2012)

La segmentation sémantique **hiérarchique** ou en **cascade** peut être vue de deux manières. Dans un cas, il s'agit d'une segmentation *All vs One*, qui comprend alors une cascade de *Classifier* permettant de sélectionner parmi les différentes classes (Brodu & Lague, 2012). On retrouve en partie le fait d'utiliser un second algorithme pour affiner la segmentation sémantique dans les travaux utilisant un CRF en deuxième partie comme décrit dans la section 1.3.4. L'autre possibilité de hiérarchisation est d'utiliser des classes plus générales dans un premier temps et par la suite prendre les classes plus précises. On le retrouve par exemple dans les travaux de (Teruggi et al., 2020) où les façades d'une basilique sont ensuite subdivisées en murs et vitraux. Cette segmentation sémantique sur plusieurs niveaux est intéressante dans de tels cas. Cependant, comme avec la segmentation *one vs all* visible dans la Figure 1.10, si les sous-classes sont trop diverses et représentent un ensemble trop grand, la classification aura de moins bons résultats. L'un des avantages de cette hiérarchisation apparaît lorsque deux classes ou plus sont très similaires. Ainsi, il est possible avec la hiérarchisation de la solution de créer un *Classifier* comparant seulement les deux classes similaires.

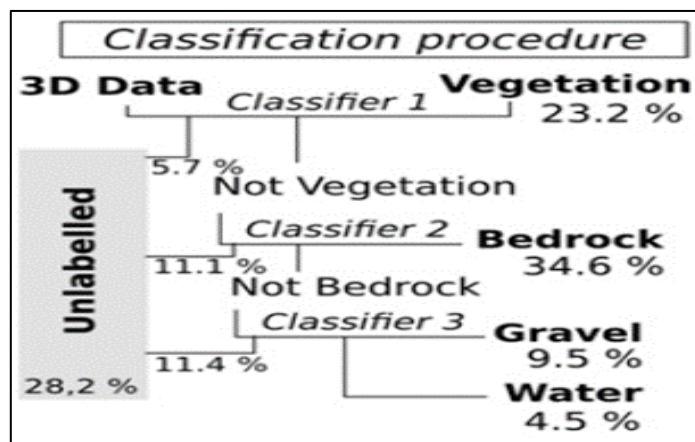


Figure 1.10 Schématisation d'une segmentation sémantique avec le modèle *One-vs-all* nécessitant plusieurs *classifier*  
Tirée de Brodu & Lague (2012)

Selon le projet, la segmentation sémantique peut être **spécifique** et cibler une unique classe du nuage de points. Il s'agit alors d'une segmentation ayant pour but de segmenter une classe spécifique. C'est le cas dans le projet (R. Huang et al., 2018) où la segmentation sémantique



porte uniquement sur le fait de reconnaître les bâtiments. Cette problématique est aussi partagée par les études (Ostrowski, Pilarska, Charyton, & Bakula, 2018 ; Selvaraju et al., 2021 ; Sun & Salvaggio, 2013 ; Xia & Wang, 2018). Il est aussi possible de retrouver un objectif spécifique en ne cherchant à segmenter sémantiquement que les routes (Belton & Bae, 2010 ; Vo, Truong-Hong, & Laefer, 2015 ; J. Wu et al., 2019) ou encore pouvoir retrouver les conducteurs électriques afin de vérifier et d'éviter les dangers liés à leurs balancements en cas de vent (Z. Shi, Lin, & Li, 2020). Le modèle mathématique du conducteur est plus spécifique. A contrario, une segmentation sémantique **générale** va segmenter sémantiquement l'ensemble du nuage de points.

Les segmentations sémantiques peuvent utiliser des **attributs**. Dans certains cas, les algorithmes de segmentation sémantique font appel à des attributs afin de classer les points. Ces attributs peuvent être utilisés avec des seuils dans le cas d'approches par *Region Growing* ou dans le cas d'algorithmes de *Random Forest*. Ils peuvent aussi faire appel à un vecteur d'attributs dans le cas d'algorithme d'apprentissage machine non-profond. Dans ce dernier cas, il est possible de les nommer *Human Based Approach*. Cette dénomination fait écho aux algorithmes d'apprentissage profond qui ne font pas appel à des attributs créés par des êtres humains aussi appelés *Handcrafted Features* par (Chew, Ji, & Zhang, 2022). Les algorithmes d'apprentissage profond peuvent aussi être appelés *End-to-End Machine Learning* ou encore *Machine Based Approach* comme les attributs utilisés ont été créés par la machine, donc sans, ou avec moins, de biais ou d'a priori humain. Les attributs peuvent être de différentes natures parmi lesquelles des attributs géométriques qui correspondent à la géométrie du point et de ses voisins comme dans les études (Antonarakis, Richards, & Brasington, 2008 ; Belton & Bae, 2010 ; Brodu & Lague, 2012 ; B. Guo et al., 2015 ; He, Chen, Wang, & Liu, 2017 ; Landrieu & Simonovsky, 2017 ; Yong Li et al., 2019 ; Yuan Li et al., 2019 ; Qin, Hu, & Dai, 2018 ; Sareen, Knopf, & Canas, s.d. ; Z. Shi et al., 2020 ; Teruggi et al., 2020 ; Weinmann et al., 2015 ; L. Wu, Zhu, Lawes, Dunkerley, & Zhang, 2019 ; Xu & Yang, 2018 ; Zhao, Pang, & Zhang, 2018), des attributs probabilistes. Afin de calculer les attributs, différentes méthodes sont possibles, la plus simple étant de prendre les plus proches voisins du point à classer, puis calculer l'attribut avec ce sous-nuage de points voisins. La Figure 1.11 est le nuage de points

qui représente la Localisation #1. La colorisation correspond à l'attribut de verticalité. Plus le nuage de points avoisinant du point à classer est vertical, plus la valeur de l'attribut est haute. Les murs et autres plans verticaux vont donc apparaître en rouge (avec une valeur pour l'attribut verticalité qui est grande). Les routes et les surfaces hydrologiques apparaîtront bleues (avec une verticalité presque nulle) avec l'échelle de valeurs allant de 0 à 1 et de couleur allant du bleu au rouge de la Figure 1.11.

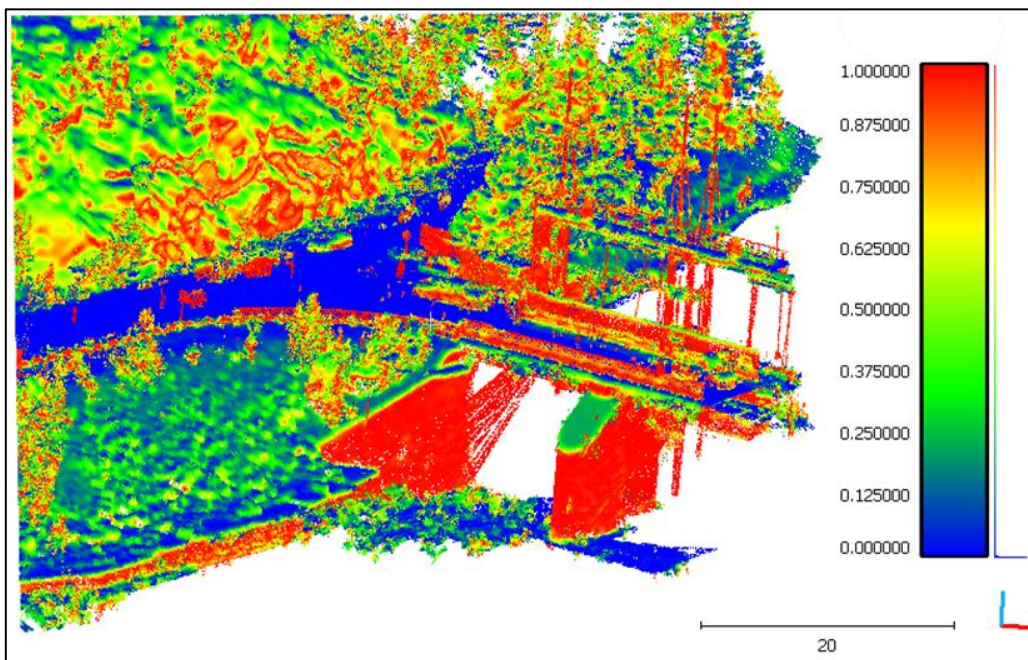


Figure 1.11 Visualisation de l'attribut « verticalité » sur le nuage de points « Localisation #1 » du PI. L'attribut « verticalité » a été calculé en prenant une sphère de taille fixe de diamètre 1m

Afin de pouvoir calculer des attributs, il est nécessaire de réaliser un **regroupement de points** que l'on peut aussi appeler **sélection de points** ou **sélection de voisinage** comme les points sélectionnés sont au voisinage du point à classer. Des sphères, cylindres, projections, grilles ou encore le nombre de voisins peuvent être utilisés afin de sélectionner les points voisins. Ces différentes méthodes peuvent aussi être liées à des attributs particuliers comme l'entropie des valeurs propres. Cet attribut représente l'ordonnement d'un nuage de points. Il est pris

comme attribut permettant de sélectionner la taille des sphères pour calculer les attributs dans les études (Weinmann et al., 2015 ; Xu & Yang, 2018).

### **1.3.8 Évaluation et validation de la segmentation sémantique**

Afin de vérifier les performances des algorithmes que nous allons mettre en place et de les comparer avec les algorithmes existant dans la littérature, un certain nombre d'éléments sont nécessaires.

Dans un premier temps, afin de juger de la pertinence des résultats concernant la segmentation sémantique des points des nuages et la suite des traitements, il faut avoir à disposition des nuages de points multisources déjà classés. Comme vu auparavant, la majeure partie des bases de données existantes ne comprennent qu'une technologie (le plus souvent des nuages de points de lidar aéroporté ou terrestre). Une des possibilités est donc de créer soit même le jeu de données en classifiant manuellement un jeu provenant des trois sources de données. Une autre possibilité est de créer un jeu de données artificielles comme celui de GTA V proposé par exemple dans (B. Wu et al., 2017).

Ces deux techniques peuvent prendre du temps et ne sont pas forcément réalistes. La première technique de segmentation sémantique manuelle prendra trop de temps pour un grand jeu de données. La seconde technique, à partir de données artificielles, ne rendrait pas compte d'une segmentation sémantique réelle, les données étant générées artificiellement. Une autre possibilité est de tester les parties des algorithmes indépendamment. Pour ce faire, la partie critique sera de comparer les performances de la segmentation sémantique des nuages de points sur différents jeux de données connus. Une fois les jeux de données sélectionnés, il sera possible de faire tourner différents algorithmes dessus.

Les principales caractéristiques recherchées pour la validation, lors de la segmentation sémantique de nuage de points sont liées :

- au temps de calcul de l'algorithme, les métriques liées au temps de calcul sont principalement utilisées pour les projets nécessitant une segmentation sémantique pour la conduite autonome de voitures ou de robots.
- à l'exactitude de la segmentation sémantique.

- à d'autres performances mesurables telles que :
  - la robustesse de la solution face aux variations de densités, aux bruits, aux trous ou aux aberrations.
  - l'invariance de la solution face à l'ordonnement des points, aux translations, rotations et changements d'échelles du nuage de points.

Le temps de calcul étant en partie lié à la puissance de calcul de l'ordinateur utilisé, il sera intéressant de tester différents algorithmes sur le même ordinateur. La plupart des publications de la littérature ne comparent que les résultats concernant la validité des classes et non les temps de calcul nécessaires. Les ordinateurs et cartes graphiques utilisés pouvant différer entre les publications, les temps de calcul ne sont pas comparables d'un article à l'autre.

Concernant l'exactitude des classes supposées, différentes métriques peuvent être calculées. La valeur principalement calculée pour la classification est la *MIoU* pour *Mean Intersection over Union* ou juste *IoU* (Eq. 1.1). Cette valeur correspond à la moyenne des valeurs de *IoU* qui sont :

$$IoU \text{ ou } Q = \frac{TP}{TP + FN + FP} \quad (1.1)$$

Dans le cas de segmentation sémantique de nuages de points, cette valeur permet de mesurer la précision avec laquelle les points d'une classe ont bien été classés ou non.

Ces variables, visibles dans le Tableau 1.7, sont : *TP* (*True Positive*) les points classés avec la bonne classe (*True Positive*), *FN* (*False Negative*) les points de la classe recherchée, mais étant mal classé (*False Negative*) et *FP* (*False Positive*) les points classés avec la classe recherchée, mais n'étant pas de cette classe (*False Positive*).

Tableau 1.7 Matrice de confusion pour deux classes 0 et 1

	Prédiction = 1	Prédiction = 0
Classe = 1	<i>True Positive (TP)</i>	<i>False Negative (FN)</i>
Classe = 0	<i>False Positive (FP)</i>	<i>True Negative (TN)</i>

D'autres valeurs peuvent être aussi calculées telles que la *Precision* (Eq. 1.2), le *Recall* (Eq. 1.3) ou encore le *F1-score* (Eq. 1.4). Ces valeurs sont moins utilisées, mais apparaissent par exemple dans la publication (Yuan Li et al., 2019) ou encore dans la présentation orale (He et al., 2017):

$$Precision \text{ ou } Cr = \frac{TP}{TP + FP} \quad (1.2)$$

$$Recall \text{ ou } Cp = \frac{TP}{TP + FN} \quad (1.3)$$

$$F1 - score = \frac{2(Recall * Precision)}{(Recall + Precision)} \quad (1.4)$$

$$Overall Accuracy \text{ ou } OA = \frac{TP}{TP + TN + FP + FN} \quad (1.5)$$

La *Precision* ( $Cr$ ) mesure le nombre de points pertinents (qui peuvent faire partie de la classe voulue) parmi l'ensemble des points recherchés. Le *Recall* ( $Cp$ ), ou rappel, mesure le nombre de points pertinents parmi l'ensemble des points pertinents du nuage de points.

Le *F1 - score* permet d'avoir une moyenne harmonique de la précision et du rappel. Le *F1 - score* est plus pertinent que la précision globale ou *Overall Accuracy* (Eq. 1.5) pour les nuages de points ayant des classes avec des nombres de points très différents.

Les jeux de données utilisés pourront vraisemblablement comporter des erreurs puisque la segmentation sémantique manuelle par des experts peut être fastidieuse. Cependant, il n'est pas possible de savoir quel pourcentage de points seront mal classés par les experts. Aucun jeu de données trouvé jusqu'ici ne parle de segmentation sémantique par différents experts afin de réaliser un test de répétabilité. Ceci est potentiellement dû au temps requis pour classer le nuage de points une première fois.

Les autres performances quant à elles sont intéressantes pour la segmentation sémantique des nuages de points provenant d'autres sources avec des densités, des niveaux de bruits ou niveaux d'aberrations différents.


Concernant la validation des performances de la segmentation sémantique, elle devrait pouvoir s'effectuer sur des nuages de points test qui ne seront pas forcément réalistes ou provenant de la réalité. Cependant, des tests devront aussi être effectués sur des nuages de points acquis de manière habituelle.

D'autres tests devront aussi être effectués pour le découpage des classes en entités distinctes, ainsi que la propagation des attributs (tel que la couleur ou encore l'intensité de retour).

La matrice de confusion est souvent utilisée afin de vérifier avec quelles classes est confondue chacune des classes recherchées. Dans le Tableau 1.8 représentant une matrice de confusion, il est possible de voir que les points correspondant à la classe « Surfaces hydrologiques » sont perçus comme des points appartenant à la classe Sol dans 15% des cas. Ces deux classes Sol et Surface hydrologiques sont proches d'un point de vue géométrique, il peut être possible de les différencier selon d'autres critères.

Cette matrice est plus utile dans le cas d'une segmentation sémantique où des attributs sont utilisés. Dans ce cas, il est possible d'imaginer et de créer des attributs spécifiques afin de remédier aux problèmes de confusion soulevés dans cette matrice.

Tableau 1.8 Matrice de confusion lors d'une segmentation sémantique.  
Chaque ligne de la matrice représente une classe et le pourcentage des points qui la compose, les colonnes représentant les classes estimées

Est estimé comme	Sol	Végétation	Bâtiments	Surfaces hydrologiques	Conducteurs	Pylônes
						
Sol	0.92	0.00	0.05	0.03	0.00	0.00
Végétation	0.00	0.99	0.00	0.00	0.00	0.00
Bâtiments	0.01	0.01	0.97	0.00	0.00	0.01
Surfaces hydrologiques	0.15	0.00	0.03	0.82	0.00	0.00
Conducteurs	0.00	0.00	0.01	0.00	0.97	0.01
Pylônes	0.00	0.37	0.10	0.00	0.01	0.52

Cependant, avec des nuages de points ayant des classes avec peu de points pour les représenter, la matrice de confusion semblera avoir de bons résultats alors que certaines classes sont mal

segmentées. C'est souvent le cas pour certaines classes spécifiques telles que des surfaces hydrologiques. Les surfaces hydrologiques sont de grandes étendues qui sont moins bien captées à cause de la réflexion de l'eau ainsi que des mouvements. Il est donc utile de la coupler avec une métrique permettant de rendre compte de cette différence. Le coefficient de corrélation de Matthew (Eq. 1.6) (Ying Li et al., 2021) permet cela et prend en compte le nombre de points bien classés ( $TP \times TN$ ) ainsi que le nombre de points mal classés ( $FP \times FN$ ). Ainsi la valeur MCC est grande si l'ensemble des valeurs de la matrice de confusion sont bonnes (TP et TN grande et FP et FN petite) :

$$MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (1.6)$$

Enfin, il est possible d'utiliser la visualisation des nuages de points afin d'effectuer certaines vérifications. Dans certains cas, la vérification visuelle permet de vérifier l'appartenance à une certaine classe pour certains points. Dans le cas où certains algorithmes se trompent systématiquement sur la segmentation de certaines classes, la visualisation permet aussi de vérifier la classe souhaitée.

#### 1.4 Fusion de données géospatiales

Dans la majeure partie des projets, une unique technologie de captation est utilisée. Cette unique technologie est choisie selon les besoins du projet. L'utilisation de multiples technologies pour la captation n'est pas encore totalement démocratisée comme le décrit le travail (P. Ghamisi et al., 2019). Le sujet est émergent dans le cas de fusion caméra/LIDAR, les études présentent l'utilisation de diverses technologies ayant des natures différentes telles que des fusions entre nuages de points et images. L'utilisation de différentes sources peut avoir plusieurs avantages. C'est en particulier le cas avec l'utilisation de lidar couplé à une caméra, les deux technologies étant complémentaires. Cependant, comme il faut avoir un résultat unique à la fin malgré les diverses sources, l'utilisation de fusion de données est nécessaire. Cette fusion peut apparaître à différentes étapes dans le processus de segmentation ou

d'utilisation des données. Pour cette partie, nous reprendrons la différenciation présentée dans (Cui et al., 2022 ; P. Ghamisi et al., 2019). Leur différenciation prend en compte la segmentation sémantique, mais aussi d'autres tâches telles que le fait de compléter un nuage de points avec des données ou zones manquantes par d'autres sources par exemple. La fusion peut prendre place au niveau des données d'entrée dès le début du *Workflow*. Le *Workflow* est l'ensemble des étapes nécessaires qui permettent d'arriver au produit fini, dans notre cas, un nuage de points segmenté sémantiquement. Il est aussi possible d'effectuer la fusion plus tardivement. C'est par exemple le cas lorsqu'elle est effectuée lors de la voxelisation ou encore à l'étape de calcul d'attributs. Il est aussi possible de fusionner les données à l'étape des résultats. Enfin, nous parlerons de fusions multiniveaux qui utilise des fusions à différentes étapes du processus de traitement des données.

#### **1.4.1 Fusion au niveau des données d'entrée**

Il est possible de trouver les termes *Data Fusion* dont une représentation est présente à la Figure 1.12. Ce type de fusion de données est plus simple dans le cas où les données ont le même format (uniquement des nuages de points ou uniquement des images par exemple). Cependant, la fusion de données de format différent est aussi possible. Pour ce faire, il est nécessaire de projeter les attributs manquants du nuage de points vers l'image ou vice versa.

Il est aussi possible de trouver les termes fusion au niveau des sources. Un nuage de points avec une partie manquante pourra être complété à l'aide d'images provenant d'autres sources et en créant le nuage de point associé aux images. Dans le cas de la fusion de données au niveau des sources (Z. Wang, Wu, & Niu, 2020) utilise l'expression « fusion d'attributs complémentaires » qui correspond au fait de fusionner des attributs propres à chaque source. Dans ce cas, seulement un jeu de données sur les deux a l'attribut à partager. Il peut s'agir de la couleur dans le cas d'image, ou bien d'intensité dans le cas de lidar. Cette fusion permet d'améliorer les résultats de la segmentation sémantique. Si les sources utilisées sont de natures différentes (images et nuages de points), le temps de traitement peut être allongé par rapport à l'utilisation d'une seule source.



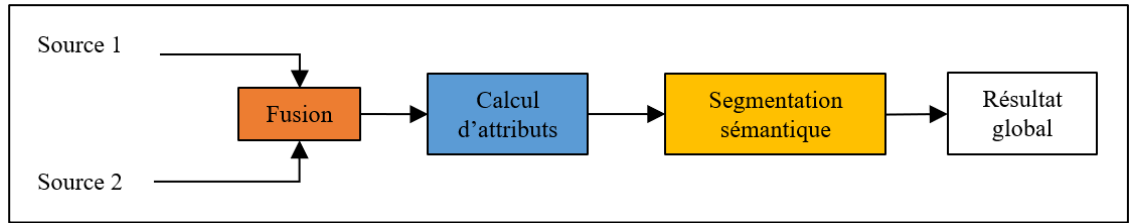


Figure 1.12 Schématisation d'une fusion de données au niveau des données d'entrée

### 1.4.2 Fusion au niveau des attributs

Dans le cas de la fusion au niveau des attributs visibles, Figure 1.13, la fusion revient à calculer des attributs propres aux points, tels que les attributs géométriques développés plus haut (C. Li, Hang, & Rasti, 2021 ; Qin et al., 2018 ; Siddiqui, Madhok, & O'Toole, 2020).

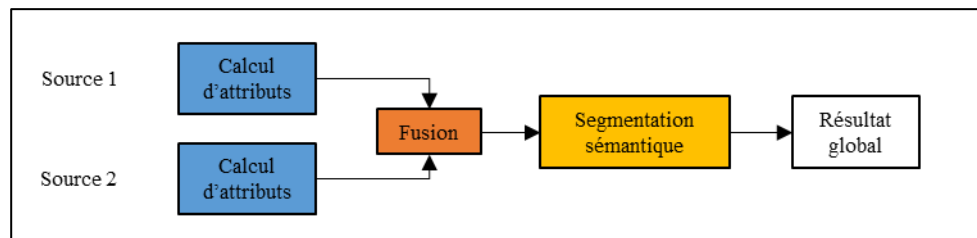


Figure 1.13 Schématisation d'une fusion de données au niveau des attributs

Dans le cas de EMFNet (C. Li et al., 2021) par exemple, un CNN est utilisé pour calculer des attributs sur les images (source 1) et un autre CNN est utilisé pour calculer des attributs sur des données lidar (source 2).

### 1.4.3 Fusion au niveau des résultats

La fusion au niveau du résultat est aussi appelée « stratégie de fusion basée sur la décision multisources » ou encore « fusion au niveau de la décision ». Elle est décrite par (Cui et al., 2022) comme le fait de prendre le résultat (dans notre cas le résultat de la segmentation sémantique) d'une source afin d'orienter le travail sur la seconde source.

Dans le cas de segmentation sémantique, il est possible de prendre le résultat de la branche travaillant sur les images qui peut être représenté comme une boîte 2D englobant l'objet à segmenter. Cette boîte 2D peut alors être projetée dans l'espace 3D afin de créer une boîte englobante 3D dans le nuage de points et d'y faire la segmentation sémantique dans cet espace plus restreint. L'un des inconvénients principaux est le fait que la boîte englobante créée à partir de l'image peut contenir d'autres entités et donc créer un biais visible au travers de la Figure 1.14. L'un des avantages est le fait de pouvoir utiliser les nombreux travaux déjà existants sur la segmentation sémantique des images.

Cette approche est donc principalement utilisée dans le cas de fusion de données : nuages de points et images, comme dans (Rahman et al., 2019).

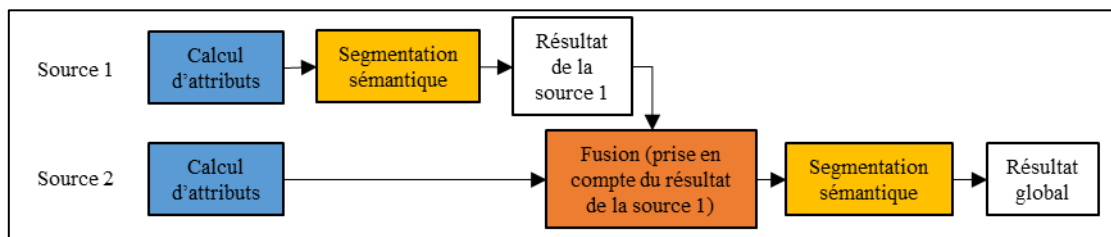


Figure 1.14 Schématisation d'une fusion de données au niveau des résultats en projetant la segmentation sémantique de la source 1 sur les données de la source 2

Il est aussi possible de trouver la fusion au niveau des résultats comme présentés dans la Figure 1.15. Dans ce cas, la fusion va prendre en entrée les résultats proposés par différentes sources comme proposé dans (Qin et al., 2018) et décrit dans (Z. Wang et al., 2020).

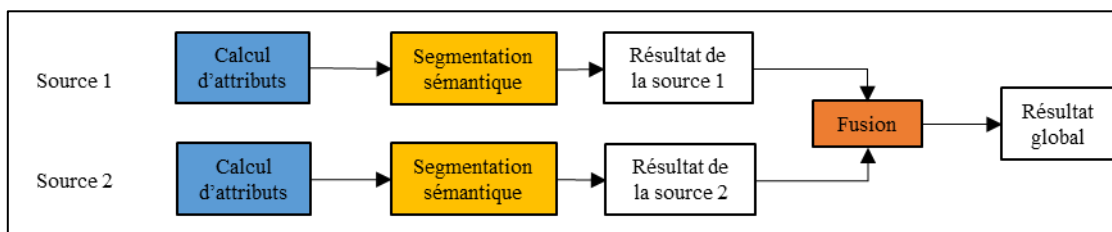


Figure 1.15 Schématisation d'une fusion de données au niveau du résultat

#### 1.4.4 Fusions multiniveaux

Les fusions dites multiniveaux apparaissent lorsque l'on utilise différentes méthodes de fusion à différentes étapes et dans le même travail (même finalité). Cette fusion utilise donc différentes méthodes décrites tout au long de cette partie comme dans (X. Chen, Ma, Wan, Li, & Xia, 2017 ; Qin et al., 2018) où les fusions multiniveaux sont réalisées en utilisant une fusion au niveau des attributs **et** une autre fusion au niveau des résultats.

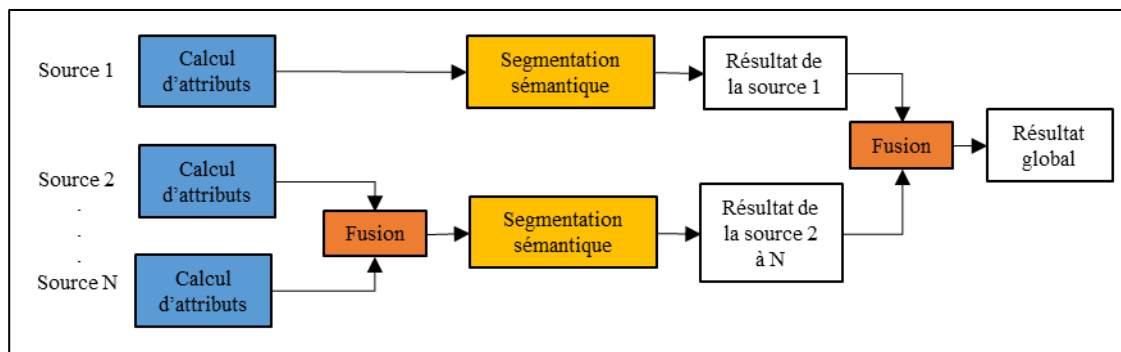


Figure 1.16 Schématisation de fusions de données multiniveaux

Les fusions multiniveaux présentées Figure 1.16 permettent donc d'allier les bénéfices de différentes fusions en propageant les différents attributs, en sélectionnant plus précisément la zone à segmenter. Un inconvénient de ce type de méthode peut être la complexité apportée par l'ajout de fusion supplémentaire ou encore le temps de calcul qui peut être allongé au vu du nombre de sources utilisées.

#### 1.5 Synthèse de l'état de l'art

Cette section va présenter notre positionnement par rapport à l'état de l'art ainsi qu'une critique de ce dernier principalement sur la méthodologie employée.

### 1.5.1 Positionnement de notre sujet par rapport à l'état de l'art

L'utilisation de diverses sources pour la segmentation sémantique est une problématique reconnue comme un sujet ayant besoin d'être approfondi en recherche comme le préconise les travaux de (Pedram Ghamisi et al., 2018). Il n'existe pas encore de solution robuste, automatisable, applicable à divers scénarios, déployable industriellement. C'est en partie ce que nous nous sommes fixés comme objectif. La disponibilité des données a été un atout majeur qui a contribué à nos travaux de développement et de validation. Ces données, fournies par le PI, sont des nuages de points provenant de divers sites et diverses technologies de captation.

De plus, les données traitées dans ce sujet diffèrent légèrement des numérisations de villes et de routes utiles pour l'apprentissage de voiture autonome, ou des relevés de végétation nécessaires pour les études de conservation des forêts. En effet, les données présentes dans notre étude comportent différentes sources technologiques, plus précisément, entre 2 et 3. Pour chaque localisation, différentes technologies sont utilisées et les localisations représentent des numérisations de terrains urbains, mais aussi de terrains ruraux proches de surfaces hydrologiques ou de forêts. Les jeux de données seront décrits plus en détail dans le CHAPITRE 4 Résultats. De ce fait, il est possible que les algorithmes proposés dans la littérature aient de moins bons résultats sur ces nouvelles données. Un algorithme utilisant le potentiel des données multisources est donc une piste de réflexion à avoir.

Le nombre restreint de nuages de points ayant des sources similaires nous a orientés vers l'utilisation des techniques d'apprentissage machine. Les nuages de points ayant été mis à notre disposition par le PI sont de sources diverses et déjà labélisées. Ces nuages de points ont été numérisés à des périodes différentes et représentent des localisations très diverses. Les algorithmes utilisant les réseaux de neurones profonds sont sensibles aux changements de sources dans les données d'apprentissage comme décrit dans (Xie et al., 2020).

Dans le cas des données du PI, la fusion de données correspond à une fusion avec des données de même format : deux nuages de points sous la forme de coordonnées spatiales géoréférencées

(dans certains cas, chaque point aura un ou plusieurs attributs). La littérature comporte plus d'articles utilisant une fusion de données de format : nuages de points et images comme le montrent les études de (P. Ghamisi et al., 2019 ; Ying Li et al., 2021). Les articles qui proposent des fusions du type nuages de points et nuages de points portent principalement sur le fait de combler les nuages de points existants et ils portent moins sur la segmentation sémantique (Kwon, Park, Moon, Jung, & Park, 2017 ; Ying Li et al., 2021 ; J. Zhang & Lin, 2017).

### **1.5.2 Critiques de l'état de l'art**

Dans un premier temps, le recensement de l'état de l'art du domaine a été principalement réalisé à l'aide de requêtes et mots clés dans des moteurs de recherche *Scopus*, *Web of Science* et *Google Scholar*. Ces sites sont en partie biaisés, toutes les revues scientifiques n'y sont pas forcément répertoriées. Le classement des articles qui y apparaissent lors de recherches peut amener à ne pas identifier certains articles comme les articles les plus souvent cités sont mis en avant. De plus, les mots clés utilisés peuvent être connotés ou n'être utilisés que dans certains domaines, fermant ainsi la porte à certaines recherches utilisant d'autres termes et mots clés. Pour remédier à ses problèmes, d'autres sites, techniques et algorithmes de recherche ont été utilisés comme *ResearchRabbit* ou encore le classement des revues scientifiques selon les disciplines représentées. Le fait de chercher les disciplines des revues nous a permis de trouver plus précisément où il était opportun pour nous de publier nos contributions. Cependant, pour la partie Revue de la littérature, il est plausible que cette recherche bibliographique nous ait borné sur certaines revues proches de notre sujet spécifique. Par conséquent, des questionnements et propositions qui ne sont pas encore exploités dans notre domaine ont pu nous échapper. Enfin, les publications dans les conférences nous ont permis de rencontrer des chercheurs proches du notre sujet de thèse permettant alors des échanges fructueux et des pistes de réflexion sur notre méthodologie.

Bien que nous ayons essayé de rendre l'état de l'art le plus exhaustif possible, il est donc, par définition, imparfait et inachevé. Néanmoins, nous sommes convaincus qu'il est *représentatif*.

Aussi, ce premier chapitre nous a permis de présenter un recensement scientifique et technologique des systèmes d'acquisition, de la segmentation sémantique ainsi que de la fusion de données géospatiales. Il nous a aussi permis d'alimenter nos réflexions et de réaliser une première critique sur la méthodologie retenue dans cette thèse. Le deuxième chapitre nous permet de présenter en détail les questions de recherche ainsi que les objectifs et la méthodologie employée.

## CHAPITRE 2

### OBJECTIFS ET MÉTHODOLOGIE DE LA RECHERCHE

#### Sommaire

2.1	Introduction.....	51
2.2	Problématique et question de recherche .....	51
2.3	Objectifs de la recherche.....	53
2.3.1	Objectif principal .....	53
2.3.2	Sous-objectifs.....	54
2.4	Méthodologie de la recherche.....	54
2.4.1	Frontières du projet et hypothèses de recherche .....	54
2.4.2	Outils utilisés pour le projet de doctorat .....	55
2.4.3	Synthèse de la méthodologie .....	56

#### **2.1 Introduction**

Comme nous l'avons vu au travers de la revue de la littérature, la segmentation sémantique de nuages de points est un domaine relativement récent où la recherche est encore nécessaire. C'est en particulier le cas dans l'utilisation de sources de données multiples.

Ce chapitre permet de situer la problématique de recherche, l'objectif principal, les sous-objectifs, les frontières du projet ainsi que les outils utilisés au cours de nos travaux. Ce chapitre permet aussi de présenter, expliquer et argumenter certains choix effectués au cours du doctorat. Enfin, un diagramme et une synthèse de la méthodologie viennent conclure ce chapitre.

#### **2.2 Problématique et question de recherche**

Les technologies d'acquisition à distance se démocratisent et ne sont plus réservées aux professionnels de certains domaines. Leurs usages sont donc plus fréquents et les possibilités

d'utilisation ne font qu'augmenter. Ces technologies, couplées à la puissance de calcul des ordinateurs, rendent possibles l'acquisition et le traitement d'un grand nombre de données. La combinaison des données issues de plusieurs technologies devient donc envisageable, et même souhaitable pour une métrologie collaborative.

Ce projet de recherche s'inscrit en partie dans cette lignée avec la numérisation de larges zones géographiques ( $\sim 10^2$  km<sup>2</sup>) avec quatre (4) différentes technologies d'acquisitions. Ces zones d'acquisitions comprennent différentes régions d'intérêt incluant de la végétation, des étendues d'eau, ainsi que d'autres structures d'origines humaines, telles que des bâtiments, des routes ou des assemblages supportant des câbles électriques. Dans notre cas, le projet est proposé par l'équipe de géomatique du PI. La mission de l'équipe géomatique est de traiter les données et informations géographiques de l'entreprise. L'entreprise partenaire de ce projet possède des relevés de quatre (4) technologies distinctes que sont le lidar aéroporté (à bord d'avions et d'hélicoptères), la photogrammétrie par drone et le lidar terrestre. Chaque localisation comprend cependant entre 2 et 3 technologies différentes.

Le principal objectif du PI dans ce projet est la création, à partir de plusieurs technologies de relevés, d'un nuage de points consolidé pour qu'il soit traité par un algorithme de segmentation de nuages de points multisources afin d'automatiser et de fiabiliser la reconnaissance des zones d'intérêt (exemple : pylônes électriques, structure, etc.).

L'utilisation de données multisources issues de technologies différentes, donc hétérogènes, est un sujet peu traité dans la littérature pour plusieurs raisons. L'une des raisons principales est le manque dans l'industrie de projets impliquant plusieurs technologies de mesure. En effet, les technologies d'acquisition mises en place pour un projet spécifique sont sélectionnées pour répondre, au mieux, aux critères du dit projet (en termes d'étendues de mesure, de résolution, de fréquence d'échantillonnage, etc.). Par conséquent, dans de nombreux cas, une seule technologie est considérée comme la plus adaptée pour le cas étudié. Cette affirmation est de plus en plus remise en cause avec l'utilisation accrue de drones et de la photogrammétrie pour réaliser des relevés rapides et moins coûteux. Ce projet pose donc la question suivante :



*Comment combiner et faire travailler en synergie collaborative les données fournies par différentes technologies de numérisation afin de recréer et segmenter un nuage multisource consolidé comprenant les informations pertinentes comprises dans les relevés des technologies employées ?*

Cette synergie peut permettre à un usager de répondre à divers problèmes comme décrits dans l'introduction. Les problèmes tels que l'estimation de biomasse à retirer, la détermination des chemins pour faire passer des équipements (comme des camions) ou encore la vérification de la dangerosité de la végétation entourant des conducteurs sont envisagés.

Cette problématique permet d'énoncer un objectif et des sous-objectifs liés à ce dernier qui sont présentés dans la section suivante.

## **2.3 Objectifs de la recherche**

### **2.3.1 Objectif principal**

La démarche devra générer, traiter et enrichir un nuage permettant d'identifier automatiquement les zones d'intérêt et sélectionner les sources au sein des parcelles faisant l'objet des relevés. Cette démarche devra prendre en compte les avantages spécifiques de chaque technologie d'acquisition et de la présenter dans une couche technologique, par zone d'intérêt, dans le nuage final.

Ce projet de recherche a donc pour sujet le traitement de nuages de points multisources et multicouches sur des relevés dimensionnels de grandes dimensions. Multisources signifie que les données sont acquises à partir de plusieurs technologies complémentaires. Multicouches signifie que certaines zones sont acquises et superposées les unes aux autres.

Le nuage de points combiné (résultant de la combinaison de nuages primaires) devra comporter plus d'informations qu'un nuage de points obtenu à l'aide d'une seule technologie. Les zones d'intérêt correspondent aux différentes instances des différentes classes présentes dans le

nuage. Le volume de données du nouveau nuage peut aussi être plus faible en optimisant le choix des technologies utilisées selon la zone d'intérêt.

Cet objectif peut être découpé en plusieurs sous-objectifs qui sont décrits dans la partie suivante.

### **2.3.2 Sous-objectifs**

L'objectif global de la recherche soulève des problèmes de différentes natures menant aux sous-objectifs suivants :

- 1) Proposer une méthode pour la fusion des différentes sources de données et la propagation des attributs.
- 2) Proposer une méthode permettant de segmenter et classer les points provenant des différents nuages de points.
- 3) Proposer une méthode afin d'optimiser et de sélectionner les sources du nuage combiné pour chaque zone segmentée. L'optimisation sera effectuée en créant un nuage de points à densité variable selon la classe et le niveau d'intérêt de la zone.

## **2.4 Méthodologie de la recherche**

### **2.4.1 Frontières du projet et hypothèses de recherche**

Le projet étant de provenance industrielle, certaines problématiques liées à l'acquisition de relevés à distance n'apparaissent pas pour différentes raisons. Parmi ces problématiques, soulignons celles qui touchent aux technologies et aux prétraitements des données.

- Les données étant déjà acquises, le choix des technologies n'a pas été de notre ressort.
- Les données fournies par le partenaire étant déjà repositionnées les unes par rapport aux autres, ces problèmes ne seront pas traités au cours de ce projet. Par conséquent, nous considérons les différentes sources de données dans un seul repère géolocalisé avec précision.

- D'autres problématiques, comme celles qui portent sur les incertitudes de mesures lors de l'acquisition par exemple, ont moins d'intérêt pour le PI (les données répondant aux critères d'incertitudes de l'entreprise). Ces problématiques de validité des mesures ne seront donc pas prises en compte pour le projet.
- Les problématiques concernant la création de nuages de points à partir d'images ou de photographies ne seront pas non plus traitées. Notre partenaire nous a fourni directement les nuages de points calculés à partir de photogrammétrie. Les images auraient pu être demandées afin d'ajouter un nouveau type de donnée. Cette proposition (d'utiliser des images en plus des nuages de points) a déjà été étudiée dans la littérature, par conséquent, elle n'a pas été traitée dans ce projet.
- Enfin, les nuages de points fournis par le PI étant déjà en partie classés par des experts du domaine, l'apprentissage supervisé se fera en considérant cette classification comme suffisamment juste pour être prise comme référence. Nous verrons cependant, lors de la partie discussion des résultats, qu'il est possible de classer certains points dans une ou plusieurs classes selon notre interprétation des lieux. Nous calculerons donc un pourcentage d'incertitude pour les segmentations sémantiques réalisées par les experts afin de valider ou non nos résultats.

#### 2.4.2 Outils utilisés

Cette section couvre les décisions prises concernant les différents outils à utiliser au cours des travaux ainsi que la justification de ces choix. La revue de la littérature du chapitre précédent a déjà fourni une description de ces outils et méthodes.

**Programmation :** Afin d'évaluer nos résultats, une partie du projet consiste à programmer et implémenter nos solutions pour la segmentation sémantique de nuages de points. Nous avons fait le choix de programmer sous Python. Ce choix est dû aux avantages qu'offrent les nombreuses bibliothèques disponibles dans ce langage. D'autres langages de programmation, tels que le C ou le C++, auraient pu être utilisés (plus efficace dans l'utilisation et la gestion de données volumineuses).

Les principales bibliothèques utilisées au cours de la programmation sont *Tensorflow*, *Pandas*, *Numpy*, *Matplotlib* et *Scikit-learn*.

*Tensorflow* est une bibliothèque principalement utilisée pour la création et l'utilisation de réseaux de neurones. En comparaison avec *Pytorch*, qui est aussi une bibliothèque permettant la création et l'utilisation d'algorithmes d'apprentissage machine, *Tensorflow* offre un plus grand assortiment de *tutoriels* facilitant la prise en main et la compréhension de la bibliothèque.

*Pandas* est une bibliothèque permettant une utilisation et gestion de données efficace.

*Numpy* est une bibliothèque permettant l'utilisation de matrice et de fonctions mathématiques.

*Matplotlib* est une bibliothèque de visualisation de données.

*Scikit-learn* a permis d'avoir un premier aperçu de l'utilisation d'algorithmes d'apprentissage machine. Cette bibliothèque est simple d'utilisation, mais nous arrivons rapidement à ses limites dans le cas de réseaux de neurones.

### **2.4.3 Synthèse de la méthodologie**

La méthodologie utilisée au cours de ce doctorat est synthétisée dans la Figure 2.1 Elle se reflète dans la structure de cette thèse.

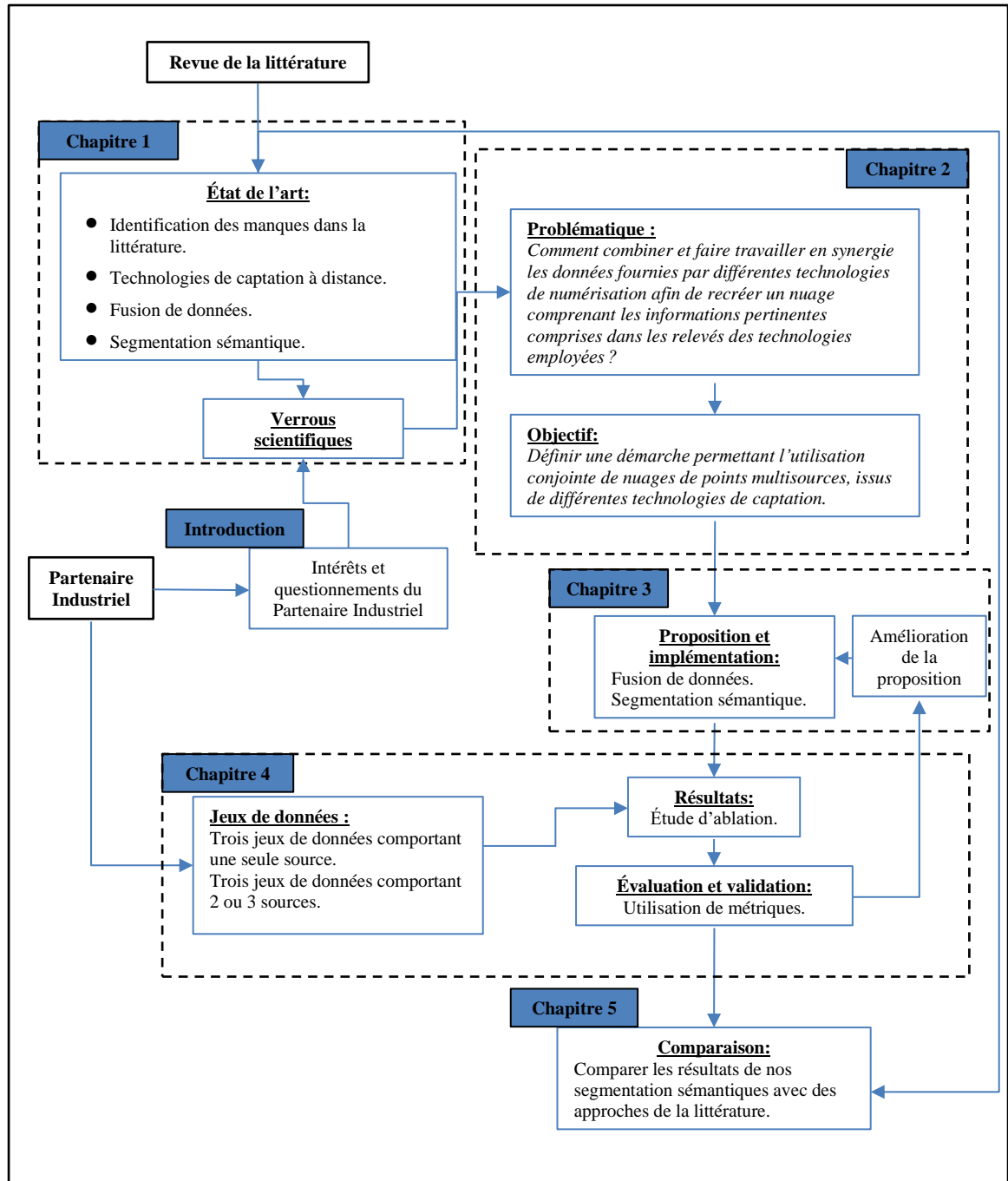


Figure 2.1 Diagramme de la méthodologie de recherche

Dans un premier temps, nous avons réalisé une Revue de la littérature ainsi que des réunions avec le PI. Ces réunions nous ont permis de définir les intérêts et les questionnements du PI. La revue nous a permis de réaliser un état de l'art du domaine, d'identifier les dernières

avancées et les algorithmes tels que proposés par la communauté des chercheurs, mais aussi d'identifier des manquements, et donc de potentiels sujets de développement. Ces étapes nous ont permis d'identifier des sujets qui peuvent être utile au PI et impliquer des verrous scientifiques.

Le sujet développé est celui pour lequel le PI porte un grand intérêt. Une problématique de recherche ainsi que des objectifs ont alors pu être développés pour ce sujet, présenté plus haut. L'étape suivante, présenté au CHAPITRE 3, concerne la création d'une proposition et son implémentation. Cette proposition fait suite aux manquements relevés dans la littérature et doit essayer de répondre aux questionnements soulevés. Cette proposition a été améliorée en fonction des résultats obtenus lors de l'implémentation.

Afin de valider notre approche, une partie *Résultats et évaluations* est nécessaire. Cette partie comprend les données et les métriques utilisées. Concernant les données utilisées, il s'agit de divers nuages de points présentés au CHAPITRE 4. Ces données nous ont été directement fournies par le PI. Les lieux représentés ainsi que les technologies de captations utilisées sont divers et recouvrent une grande partie des usages du PI. Les métriques d'évaluations choisies ont été sélectionnées afin de prendre en compte certaines particularités des nuages de points tels que le déséquilibre des classes présentes dans les nuages. Le déséquilibre des classes correspond à la représentation de chacune des classes dans le nuage. Certaines classes se trouvent en abondances, d'autres ayant peu de points pour les représenter.

D'autres algorithmes mentionnés dans la littérature seront utilisés dans la section de comparaison, qui se trouve au CHAPITRE 4. Le nombre d'algorithmes disponibles pour la comparaison a été influencé par le temps nécessaire pour maîtriser leur fonctionnement et la manière de les mettre en œuvre.

## CHAPITRE 3

### PROPOSITION DE SEGMENTATION SÉMANTIQUE

#### Sommaire

3.1	Introduction.....	60
3.2	Étude de segmentation sémantique monosource .....	61
3.2.1	Sélection de voisinages monosource en vue du calcul d'attributs.....	61
3.2.2	Calcul d'attributs monosource pour la segmentation sémantique .....	66
3.2.3	Sélection d'attributs et réduction dimensionnelle.....	70
3.2.4	Apprentissage machine - Réseau de neurones .....	71
3.2.4.1	Architecture du réseau de neurones .....	71
3.2.4.2	Choix des hyperparamètres pour la segmentation sémantique ..	72
3.2.4.3	Découpage du jeu de données.....	73
3.3	Étude de segmentation sémantique multisource .....	73
3.3.1	Architecture de la proposition de segmentation sémantique multisource	73
3.3.2	Calculs d'attributs pour l'approche multisource.....	77
3.3.2.1	Attributs .....	77
3.3.2.2	Sélection d'attributs multisource .....	77
3.3.3	Propagation des attributs acquis.....	78
3.3.4	Réseau de neurones multisource (NN 1) .....	80
3.3.4.1	Architecture du réseau de neurones .....	80
3.3.4.2	Hyperparamètres dans le cas de la segmentation sémantique multisource.....	81
3.3.4.3	Découpage des données en jeux d'apprentissage, de test et de validation.....	81
3.3.5	Calcul des attributs de classes avoisinantes .....	82
3.3.6	Seconde segmentation sémantique NN 2.....	84
3.4	Optimisation des nuages de points.....	84
3.4.1	Segmentation en entités distinctes .....	85
3.4.2	Sélection de la source.....	86
3.5	Synthèse et contributions .....	88

### 3.1 Introduction

Ce chapitre porte sur les étapes que nous avons mises en place pour segmenter sémantiquement les nuages de points à notre disposition. La revue de la littérature nous a informé que la segmentation sémantique peut être réalisée avec différentes approches. Dans un premier temps, nous avons décidé d'utiliser une approche basée sur l'apprentissage machine, et plus particulièrement, sur les réseaux de neurones non profonds couplés avec la création et le calcul d'un ensemble d'attributs. Ce choix a été motivé par les résultats obtenus lors de premiers essais et la disponibilité de nuages de points avec des classes et technologies de captations diverses. Aussi, l'apprentissage machine permet de faire évoluer l'algorithme en fonction des nouvelles données disponibles chez le PI.

L'apprentissage profond est aussi une solution intéressante, elle devrait être étudiée dans la suite du projet, nous avons formulé des recommandations dans ce sens. Comme décrit dans la revue de la littérature, l'apprentissage profond possède quelques inconvénients parmi lesquels la contrainte d'employer des données similaires. En effet, pour obtenir de meilleurs résultats sur des nuages de points, il est préférable d'utiliser la même source technologique, avec une densité de points quasi constante et des terrains similaires. L'apprentissage profond nécessite aussi une grande quantité de données. C'est aussi le cas pour l'apprentissage machine. Cependant, l'utilisation d'attributs géométriques, calculés en amont, permet de dépasser cette limitation, surtout s'ils sont calculés au niveau de chacun des points plutôt qu'au niveau du nuage complet.

Le chapitre est découpé comme suit : dans un premier temps, notre approche de segmentation sémantique est testée et évaluée sur des nuages de points qui proviennent d'une seule source. Par la suite, l'approche est perfectionnée pour répondre au cas d'une segmentation sémantique de nuages de points multisources. Ensuite, une section est dédiée pour présenter l'optimisation du nuage de points résultant avec, à la fois la sélection de la source et le découpage des classes en entités distinctes. Enfin, le chapitre se conclut avec une synthèse ainsi qu'un rappel des différentes contributions.



## 3.2 Étude de segmentation sémantique monosource

Le choix de commencer par de la segmentation sémantique monosource a été motivé par la nature des données fournies par le PI.

- Les premiers nuages de points fournis pour ce projet ont été les Localisation #1, #4, #5 et #6. Ces nuages ont pour la majorité (#4, #5 et #6) une seule source de donnée (un ALS).
- Par la suite, pour les nuages des Localisation #2 et #3, seuls les nuages qui proviennent d'ALS étaient segmentés sémantiquement par des experts. Ce qui nous a incités à étudier dans un premier temps la segmentation sémantique monosource.

Comme mentionné dans la littérature, pour rendre la segmentation sémantique la plus autonome possible, il est impératif d'avoir des données d'entraînements. Nos données d'entraînement ont donc été les données ALS déjà segmentées sémantiquement par des experts.

### 3.2.1 Sélection de voisinages monosource en vue du calcul d'attributs

Pour classer automatiquement chaque point, plusieurs options sont envisageables. Dans un premier temps, nous avons retenu une approche qui utilise des attributs *humains* tels que la **linéarité** et la **verticalité** par exemple. Les attributs humains sont des attributs imaginés et calculés par l'*humain*. Ils sont en partie biaisés car proviennent de concepts *humains* tel que la **linéarité** ou la **verticalité**. *A contrario*, les algorithmes d'apprentissage machine profond, qui n'en utilisent pas, créent des attributs qui sont potentiellement moins biaisés car ne correspondent pas forcément à des concepts *humains*. Pour calculer ces attributs, il est nécessaire de regrouper un petit nombre de points (étape « Sélection du voisinage » de la Figure 3.1) autour du point à classer dans le but de calculer ses attributs en appréhendant la configuration locale.

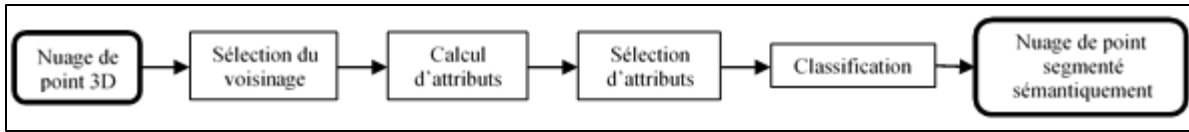


Figure 3.1 Schématisation des étapes pour la segmentation sémantique de nuages de points en utilisant un algorithme d'apprentissage machine et un vecteur d'attributs

Dans notre proposition, les voisinages peuvent correspondre soit à des points inclus dans une sphère de rayon déterminé, soit à des points contenus dans un cylindre vertical de rayon déterminé et de hauteur infinie comme visible dans la Figure 3.2. L'utilisation d'un cylindre orienté est particulièrement intéressante lorsqu'on considère des nuages de points acquis par lidar aéroporté, car elle permet d'exploiter *a priori* fort sur ce type de données lié aux données altimétriques en  $z$ . Il est aussi possible de trouver d'autres moyens de calculer des attributs tels que l'utilisation de grilles, de voxels ou encore de projeter des points sur des plans prédéterminés. L'ensemble de ces techniques n'ont pas été évaluées principalement par manque de temps.

Par la suite, nous avons utilisé le formalisme suivant (Figure 3.2) :

- $(x_i, y_i, z_i)$  sont les coordonnées spatiales du point  $P_i$  pour lequel des attributs doivent être extraits ;
- $N_S$  et  $N_C$  sont respectivement le nombre de sphères et de cylindres utilisés pour définir les voisinages pour chaque point  $P_i$ ;
- $S_{i,j}$  est le voisinage obtenu en utilisant une sphère centrée sur  $P_i$  avec un rayon  $R_{S_j}$  avec  $j \in [1..N_S]$ . L'indice  $i$  est omis par souci de clarté ;
- $C_{i,k}$  est le voisinage obtenu en utilisant un cylindre aligné avec  $P_i$  et orienté selon l'axe  $z$  avec un rayon  $R_{C_k}$  avec  $k \in [1..N_C]$ . Ici encore, l'indice  $i$  a été omis par souci de clarté ;
- $N_{p_{S_{i,j}}}$  et  $N_{p_{C_{i,k}}}$  sont respectivement le nombre total de points dans  $S_{i,j}$  et  $C_{i,k}$  ;
- $\lambda_l$  est la  $l^{\text{ème}}$  valeur propre du nuage de points liée à un voisinage obtenu avec une sphère, c'est-à-dire liée à un  $S_{i,j}$ . Par souci de clarté, les indices inférieurs  $i$  et  $j$  ne sont pas répétés lors de l'utilisation des valeurs propres dans la suite du document.

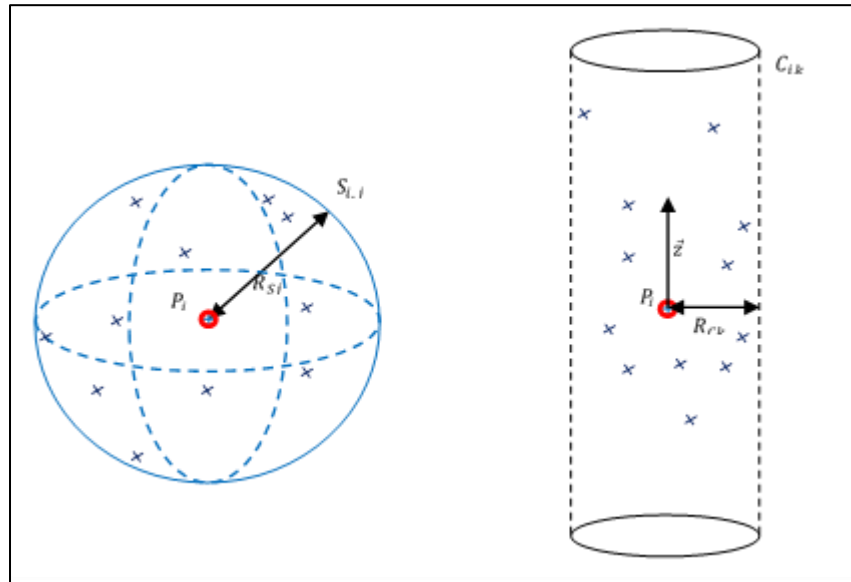


Figure 3.2 Représentation d'une sphère et d'un cylindre englobant pour délimiter le voisinage d'un point  $P_i = (x_i, y_i, z_i)$

Dans la littérature, deux types d'approches sont utilisées pour sélectionner les tailles optimales des sphères. L'une est de considérer leurs rayons comme fixes. Dans ce cas, pour un rayon donné, le nombre de voisins à l'intérieur de la sphère centrée sur un point  $P_i$  varie en fonction de la densité locale du nuage de points. L'autre possibilité est de choisir un nombre fixe de points voisins à englober par la sphère. Dans ce cas, le rayon de la sphère change en fonction de la densité locale du point considéré  $P_i$ . Dans notre cas, en gardant à l'esprit le fait d'avoir différentes technologies avec chacune une densité différente, l'utilisation de rayons fixes reste plus appropriée pour une analyse systématique à des échelles fixes. C'est cette approche qui a été retenue.

En effet, les classes ont approximativement les mêmes formes (*Shape*) à une échelle donnée, malgré les variations possibles de la densité locale. En d'autres termes, les vecteurs d'attributs extraits des différents voisinages sont plus sensibles à la forme locale de l'objet couvert qu'au nombre de points dans le voisinage.

Plutôt que de prendre une seule sphère, et comme proposé par Brodu (Brodu & Lague, 2012) et Yang (Yang et al., 2018), trois tailles de sphères différentes sont utilisées pour définir les

voisinages. Ce choix permet d'accumuler plus d'attributs et ainsi de mieux décrire les classes à plusieurs échelles.

Par exemple, un câble électrique d'une ligne à haute tension (de la classe des Conducteurs) peut être considéré comme un fil unidimensionnel lorsqu'on utilise des sphères de rayon  $R_{S_2} = 1\text{m}$ , mais il peut aussi être vu comme un nuage de points 2D ou 3D si plusieurs câbles parallèles sont inclus par une sphère de rayon  $R_{S_3} = 3\text{m}$ , comme nous l'avons schématisé dans la Figure 3.3 et représenté via des données réelles dans le Tableau 3.1. Le Tableau 3.1 représente un nuage de points qui correspond à des Conducteurs de la Localisation #4. Les sphères bleues représentent les sphères de captation de voisinage de rayon  $R_{S_2} = 1\text{ m}$  et  $R_{S_3} = 3\text{ m}$ . Dans la seconde partie du tableau à droite, le nuage à l'intérieur des sphères est extrait, et les vecteurs des principaux axes du nuage de points sont représentés en rouge.

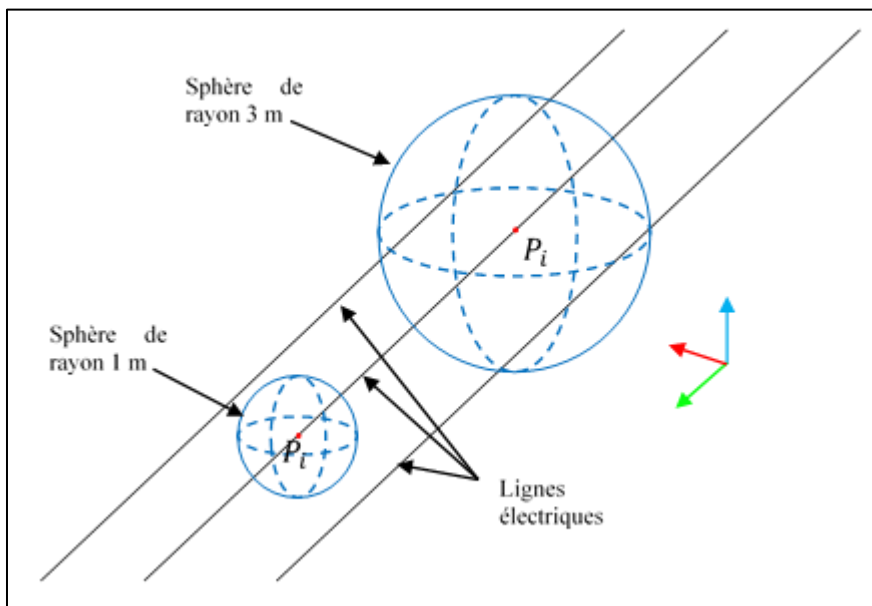
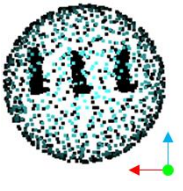
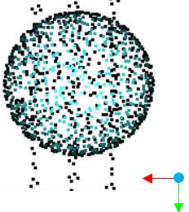

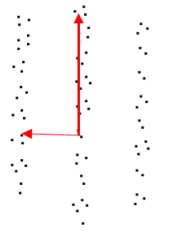
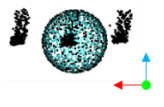
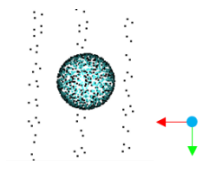




Figure 3.3 Schématisation de sphères de voisinage englobant une portion de lignes électriques

Naturellement, plus le rayon de la sphère est grand, plus le nombre de points dans la sphère sera important (à une densité donnée considérée comme fixe), et plus le calcul du voisinage ainsi que des attributs sera chronophage. Dans le cas d'acquisitions lidars aéroportées et après plusieurs essais empiriques, un rayon maximal de 3 m semble être un bon compromis entre le

nombre de points pris en compte et la rapidité de calcul des attributs. Le nombre de trois sphères relève lui aussi d'un choix, soutenu dans ce cas par la littérature, qui permet d'obtenir assez d'attributs tout en limitant le temps de calcul. L'équipe de Yang propose d'utiliser des sphères de rayons  $R_S = 0.5, 1$  et  $2.5$  m. Dans notre cas, nous avons retenu le choix de partir sur des rayons  $R_S = 0.5, 1$  et  $3$  m de diamètre.

Tableau 3.1 Représentation de l'extraction de lignes électriques à l'aide de sphères de voisinage et des vecteurs des composantes principales

	Extraction du nuage de points avec les sphères		Calcul des PCA une fois les nuages extraits	
	Vue de face	Vue de dessus	Vue de face	Vue de dessus
Sphère de 3m de rayon				
Sphère de 1m de rayon				

Comme décrit au début de cette section, pour percevoir certains phénomènes qui se produisent sur de grandes différences de hauteur, la sphère n'est pas l'outil de sélection le plus adéquat. Lors de l'acquisition avec un ALS, deux attributs qui lui sont propres sont acquis, le **nombre de retours** ainsi que le **rang des retours**. Ces attributs peuvent être exploités pour discriminer aisément certaines classes. Aujourd'hui, elles sont principalement utilisées pour soustraire le Sol d'autres classes comme la Végétation. Cependant, leur utilisation dépend fortement de la technologie utilisée, et l'approche proposée ne fonctionnerait pas sur des nuages de points

obtenus avec d'autres technologies pour lesquelles ces informations ne sont pas disponibles (par exemple, photogrammétrie ou encore le lidar terrestre).

L'utilisation de cylindres à axe vertical de hauteur infinie centrés sur chaque point permet de créer un substitut aux attributs acquis du lidar aéroporté (c'est-à-dire le nombre de retours et leur rang). Ainsi, ces attributs deviennent disponibles sur tous les nuages de points, quelle que soit la technologie adoptée. Comme pour les voisinages obtenus à partir de sphères, il est aussi possible de considérer plusieurs rayons afin d'extraire les attributs à plusieurs échelles. Après plusieurs essais, un seul cylindre ( $N_C = 1$ ) de rayon  $R_{C_1} = 0,1$  m s'est avéré suffisant. Bien qu'il soit possible de prendre en compte des cylindres de différents diamètres, en pratique, ce calcul apporte moins (du fait du nombre restreint d'attributs et du fait que les attributs varient peu entre les différents diamètres). Les principales différences entre l'utilisation du cylindre et les attributs acquis des lidars aéroportés sont **l'angle d'émission** du laser qui n'est pas nécessairement colinéaire au vecteur  $z$  jumelé au fait que le laser peut être vu plus comme **un cône** plutôt qu'un cylindre parfait. Il est donc possible d'imaginer l'utilisation de ces attributs dans d'autres cas avec non pas la direction  $z$  en vecteur, mais en prenant l'angle de la source en compte. Cette possibilité sera plus facilement réalisable avec les ALS qui disposent de l'angle de la source dans ses attributs, mais elle sera plus difficile à implémenter avec les TLS et la photogrammétrie par drone où les angles devront être calculés.

Une fois les voisinages de chaque point sélectionnés, il est nécessaire de passer au calcul d'attributs afin de créer un vecteur d'attributs pour chaque point du nuage. L'ensemble des vecteurs sera ensuite utilisé par notre algorithme d'apprentissage machine.

### 3.2.2 Calcul d'attributs monosource pour la segmentation sémantique

La revue de la littérature est riche en attributs proposés et utilisés pour caractériser la distribution des points dans un nuage. Ces attributs sont ensuite utilisés comme entrées d'un algorithme d'apprentissage. Les attributs adoptés ainsi que ceux développés sont énumérés dans le Tableau 3.2. La manière dont ils peuvent être calculés est également expliquée et, le cas échéant, la référence à un article où cet attribut est utilisé est également fournie. Les

attributs sont regroupés selon le type de voisinages utilisé pour les calculer (sphère ou cylindre) ou selon le type d'attribut (géométrique, statistique) auquel ils se réfèrent.

Tableau 3.2 Ensemble des attributs calculés pour la segmentation sémantique monosource

ID	Nom de l'attribut	Formule	Littérature dans laquelle on peut retrouver l'attribut
$f_{S1_{i,j}}$	$p_1$	$p_1 = \frac{\lambda_1}{\lambda_1 + \lambda_2 + \lambda_3}$	(Brodu & Lague, 2012)
$f_{S2_{i,j}}$	$p_2$	$p_2 = \frac{\lambda_2}{\lambda_1 + \lambda_2 + \lambda_3}$	(Brodu & Lague, 2012)
$f_{S3_{i,j}}$	$p_3$	$p_3 = \frac{\lambda_3}{\lambda_1 + \lambda_2 + \lambda_3}$	(Brodu & Lague, 2012)
$f_{S4_{i,j}}$	Linéarité ( $L_\lambda$ )	$L_\lambda = \frac{\lambda_1 - \lambda_2}{\lambda_1}$	(Landrieu & Simonovsky, 2017)
$f_{S5_{i,j}}$	Planéité ( $P_\lambda$ )	$P_\lambda = \frac{\lambda_2 - \lambda_3}{\lambda_1}$	(Xu & Yang, 2018)
$f_{S6_{i,j}}$	Sphéricité ( $S_\lambda$ )	$S_\lambda = \frac{\lambda_3}{\lambda_1}$	(He et al., 2017)
$f_{S7_{i,j}}$	Omnivariance ( $O_\lambda$ )	$O_\lambda = \sqrt[3]{\lambda_1 * \lambda_2 * \lambda_3}$	(Weinmann et al., 2015)
$f_{S8_{i,j}}$	Entropie des valeurs propres	$E_\lambda = -(\lambda_1 \ln(\lambda_1) + \lambda_2 \ln(\lambda_2) + \lambda_3 \ln(\lambda_3))$	(Xu & Yang, 2018)
$f_{S9_{i,j}}$	Somme	$\Sigma_\lambda = \lambda_1 + \lambda_2 + \lambda_3$	(Weinmann et al., 2015)
$f_{S10_{i,j}}$	Anisotropie	$A_\lambda = \frac{\lambda_1 - \lambda_3}{\lambda_1}$	(He et al., 2017)
$f_{S11_{i,j}}$	Verticalité	$V_s = 1 -  z \cdot N_{i,j} $ avec $z = (0,0,1)$ et $N_{i,j}$ la normale au point $i$ calculé en prenant en compte les points dans $S_{i,j}$	(Landrieu & Simonovsky, 2017)
$f_{C12_{i,k}}$	Nombre de retours calculés	$N_C = \log(Npc_{i,k})$	-
$f_{C13_{i,k}}$	Rang du retour	$Rg_C = \log(Rgc_{i,k})$	-
$f_{g14_i}$	Élévation	$\Delta DTM = Z_{i,MNS} - Z_{i,MNT}$	(Z. Yang et al., 2018)

Les attributs  $f_{S_{1,i,j}}$  à  $f_{S_{10,i,j}}$  sont calculés à partir du voisinage associé à la  $j^{\text{ème}}$  sphère  $S_{i,j}$  centrée sur le point  $P_i$ . Tous ces attributs proviennent des valeurs propres  $\lambda_1$ ,  $\lambda_2$  et  $\lambda_3$  calculées à partir des composantes principales du nuage de points  $S_{i,j}$ . Les attributs  $p_1$ ,  $p_2$  et  $p_3$  sont utilisés pour définir si le nuage de points dans  $S_{i,j}$  est un nuage représenté par une, deux ou trois dimensions. En effet, si le nuage de points représente plutôt un ensemble d'une dimension, nous aurons  $p_1 \gg p_2 > p_3$ . Si le nuage est plutôt 2D, nous aurons  $p_1 > p_2 \gg p_3$ .  $p_3$  est également l'attribut utilisé pour connaître le changement de courbure locale. Les attributs de linéarité, planéité et sphéricité (ou dispersion) sont utilisés pour caractériser les formes locales du nuage environnant. L'entropie des valeurs propres représente l'ordonnancement du nuage. L'attribut  $f_{S_{11,i,j}}$  correspond à la verticalité et permet d'ajouter des nuances aux attributs dimensionnels de linéarité et de planéité. Ces nuances permettent de différencier les surfaces planes entre les « Murs » et le « Sol » (qui sont des ensembles de surfaces planes), ou entre les « Conducteurs » et les « Pylônes » (qui peuvent être deux classes représentées par des nuages de points unidimensionnels).

Les attributs  $f_{C_{12,i,k}}$  et  $f_{C_{13,i,k}}$  sont calculés à partir d'un voisinage associé au  $k^{\text{ème}}$  cylindre  $C_{i,k}$  centré sur le point considéré  $i$ .  $N_{p_{C_{i,k}}}$  est le nombre de voisins présents dans le cylindre. Cette valeur peut être vue comme le nombre de retours capturés lors de l'acquisition avec les lidars aéroportés (Figure 3.4, gauche). Elle est représentée à l'aide d'un histogramme (Figure 3.4, droite).

$Rg_{C_{i,k}}$  est le rang du point  $i$  par rapport aux autres points de  $C_{i,k}$  et tel que le rang du point le plus bas est égal à 1 et le rang du point le plus haut est égal à  $N_{p_{C_{i,k}}}$ . Ces attributs offrent une nouvelle approche qui rend accessible, aux nuages de toutes sources, certains attributs traditionnellement exclusifs au lidar aéroporté. Ils permettent de discriminer rapidement certaines classes telles que le « Sol » ou la « Végétation ». Nous avons utilisé une échelle logarithmique pour mieux rendre compte de la grande variété des valeurs possibles.



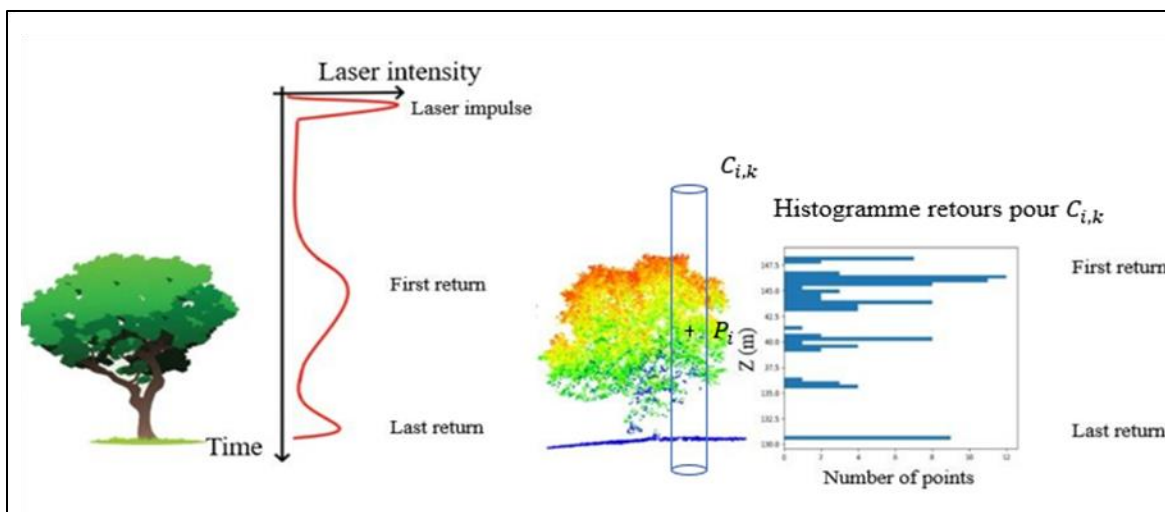


Figure 3.4 Schématisation d’une captation par ALS avec l’intensité du laser et les différents retours (sur la gauche). Nuage de points et histogramme représentant les différents retours présents dans un cylindre  $C_{i,k}$  (sur la droite)

L’attribut  $f_{g_{14}i}$  n’est pas associé à un voisinage. Il correspond à la différence d’altitude entre le modèle numérique de surface (MNS) et le modèle numérique de terrain (MNT). Plus précisément, il s’agit de la différence d’altitude entre le point considéré  $i$  (visible dans la couche MNS de la Figure 3.5) et le maillage représente la terre nue (c’est-à-dire le MNT de la figure 3-5). Il existe différentes techniques pour obtenir le MNT comme l’utilisation d’une grille (Deibe, Amor, & Doallo, 2020). Dans notre cas, nous avons retenus les classes Sol, Surface hydrologique ainsi que Route pour obtenir le maillage MNT. Dans le cas d’un nouveau nuage, non encore segmenté sémantiquement, il est possible de réaliser ce maillage en utilisant les points correspondant au dernier retour. Il est cependant nécessaire de traiter le nuage de points à la main afin d’enlever les points aberrants. Les points qui sont les derniers retours dans le cas de bâtiments par exemple doivent être enlevé. Nous avons calculé par la suite les distances des points avec ce maillage pour identifier les valeurs d’élévations  $\Delta DTM$ .

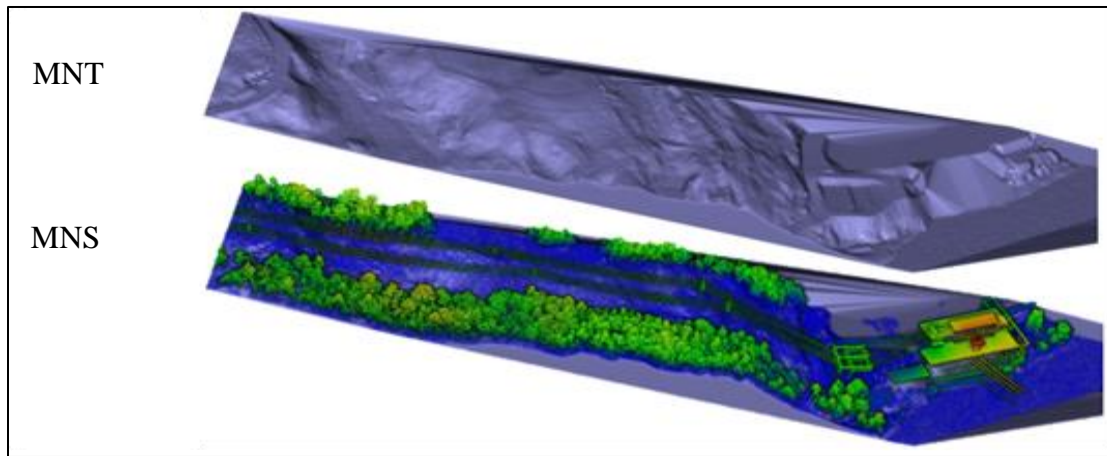


Figure 3.5 Représentation du MNT pour la Localisation #4 à l'aide d'un maillage (en haut). Le nuage de points est coloré selon la différence de hauteur entre le point et le maillage du MNT (en bas)

Une fois l'ensemble des attributs obtenus, une opération de centrage et de normalisation a été réalisée avant de lancer l'apprentissage. Néanmoins, dans certains cas, il s'avère nécessaire de sélectionner seulement une partie des attributs. Cette étape permet de soustraire des informations redondantes afin d'éviter en partie le surapprentissage et est expliquée dans la section suivante.

### 3.2.3 Sélection d'attributs et réduction dimensionnelle

La liste initiale des attributs était composée de 36 attributs (Tableau 3.2) calculés avec trois tailles de sphères (rayons 0,5 m, 1 m et 3 m) et une taille de cylindre (rayon 0,1 m). Bien entendu, ces quatre rayons sont des hyperparamètres qui peuvent être modifiés en fonction du jeu de données utilisées afin d'améliorer les scores de classification. Le nombre de sphères et de cylindres peut lui aussi être amené à changer selon les données utilisées. Ces paramètres sont donc propres à un modèle d'apprentissage et devront rester identiques lors de l'exploitation du modèle.

Dans notre cas, le nombre de sphères (3) est déjà utilisé dans la littérature ; c'est principalement le choix de leur taille qui a été fait de manière empirique (guidé par les tailles proposées dans la littérature). Pour le cylindre, les attributs extraits du cylindre (nombre de points et rang du

point) apportent moins d'informations à grande échelle et sur différentes échelles. Par conséquent, un unique rayon a été utilisé.

Pour sélectionner les attributs les plus pertinents, différentes approches peuvent être effectuées. On peut, par exemple, vérifier les corrélations entre les différents attributs, vérifier si certains attributs sont des combinaisons d'autres attributs, ou bien encore effectuer une analyse des composantes principales des attributs afin de réduire leur nombre (Analyse de Composantes Principales, ACP). L'utilisation de l'ensemble des attributs apporte souvent de meilleurs résultats, mais peut amener du surapprentissage. Cependant, pour réduire le temps de calcul, il est possible de réduire le nombre d'attributs en soustrayant ceux qui possèdent une petite contribution aux résultats finaux de la segmentation sémantique.

En effectuant une analyse de corrélation linéaire entre les attributs, les attributs  $A_\lambda$ ,  $p_2$  et  $S_\lambda$  sont apparus comme redondants, et ce pour les trois tailles de sphère. Ainsi, il est donc envisageable de les retirer tout en conservant de bonnes performances pour la segmentation sémantique, mais nous les avons gardés.

### 3.2.4 Apprentissage machine - Réseau de neurones

Les algorithmes de classification supervisée basés sur l'apprentissage machine sont nombreux. Comme le montre (Weinmann et al., 2015), il est possible avec le même vecteur d'attributs d'utiliser différentes méthodes. Dans un premier temps, pour nous familiariser avec les différents algorithmes et sélectionner celui à utiliser par la suite, nous avons utilisé la librairie *Scikit-Learn* qui permet de prendre en main des modèles simples de différentes méthodes tels que *Random Forest*, *Support Vector Classifier*, etc. Après les premiers essais, notre choix s'est porté sur l'utilisation de réseau de neurones au vu des meilleurs résultats et des possibilités qu'offrent les réseaux de neurones.

#### 3.2.4.1 Architecture du réseau de neurones

La structure du réseau de neurones ainsi développé consiste en 4 couches cachées avec respectivement 128, 64, 64 et 32 neurones (Figure 3.6). Les entrées sont les 36 attributs extraits pour chacun des  $N$  points composant l'ensemble d'apprentissage. La couche de sortie comporte

9 neurones, chacun d'entre eux représentant une classe. Les classes peuvent être la « Végétation Basse », les « Routes », les « Bâtiments », le « Sol », les « Conducteurs », les « Surfaces hydrologiques », la « Végétation haute », les « Pylônes », et les « Autres ». La fonction d'activation est une tangente hyperbolique pour toutes les couches sauf la dernière qui a une fonction *softmax*. Cette dernière permet d'avoir à la sortie, et pour chaque point considéré, les probabilités d'appartenance à chacune des classes. L'optimisation est réalisée à l'aide de l'algorithme d'*Adam* et la fonction de perte choisie est celle de perte d'entropie croisée, qui est adaptée dans le cas d'une classification à classes multiples. Cette architecture finale résulte de multiples essais empiriques. Le nombre de neurones pour la dernière couche correspond au nombre de classes, le nombre pour l'entrée correspond au nombre d'attributs du vecteur (dans notre cas 36). Le nombre de neurones par couche dans la suite a été choisi en essayant de garder des nombres de la forme  $2^n$  ainsi qu'en suivant des règles générales apprises lors de formation sur les réseaux de neurones.

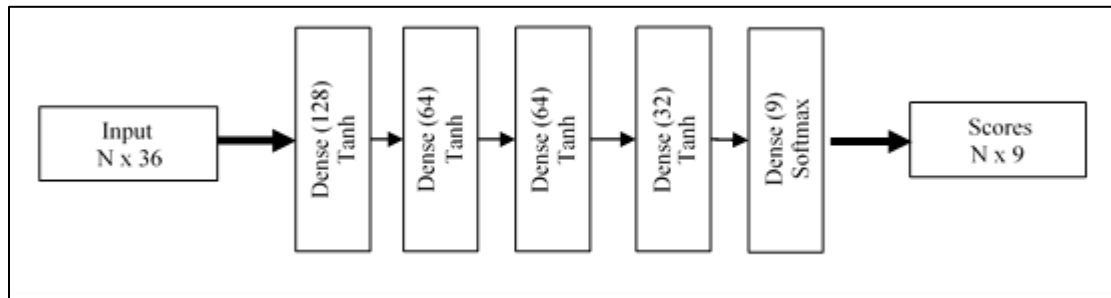


Figure 3.6 Schématisation du réseau de neurones pour la segmentation sémantique monosource

### 3.2.4.2 Choix des hyperparamètres pour la segmentation sémantique

Les hyperparamètres correspondent aux paramètres dont l'utilisateur fixe la valeur qui influe sur l'apprentissage de l'algorithme. Ces paramètres sont nombreux et peuvent être difficiles à régler. Leur réglage se fait principalement de manière empirique. Parmi les principaux hyperparamètres, on retrouve le taux d'apprentissage Alpha, le *Dropout* ou encore des poids d'initialisations. Les hyperparamètres correspondent aussi aux nombres d'époques (*Epoch*), aux nombres de lots (*Batch*), le nombre de points par lot. Les hyperparamètres sont aussi le

nombre de couches dans le réseau de neurones et le nombre de neurones par couche qui ont été discutés dans la section précédente. Dans les premiers essais sur les données monosource, nous n'avons pas fait varier l'ensemble de ces hyperparamètres. Les principaux paramètres que nous avons fait varier, en plus de l'architecture du réseau, sont le nombre d'époques que nous avons fait varier entre 20 pour les premiers essais à 500 époques (avec un arrêt de l'apprentissage dans le cas d'une détérioration de la métrique) pour les essais avec les architectures les plus prometteuses. Le nombre de points par lot a aussi évolué avec des essais de dizaines de points ( $10^1$ ) aux dizaines de milliers ( $10^4$ ) dans un premier temps pour se rendre compte de l'effet de ce paramètre sur nos résultats. Le nombre de points par lot a été finalement de l'ordre de grandeur de 5000 points par lot.

### 3.2.4.3 Découpage du jeu de données

Pour réaliser l'apprentissage du réseau de neurones, il est nécessaire de découper les jeux de données en trois parties. La partie **apprentissage** du jeu de données va permettre au réseau de neurones de modifier les poids des neurones afin d'améliorer le score des métriques sélectionnées. La partie de **validation** du jeu de données est utilisée pour vérifier et valider à la fin de chaque époque si l'apprentissage pour cette époque est allé dans la direction désirée. Enfin, la partie de **test** est une partie du jeu de données sur laquelle va être testée la prédictivité du modèle. L'apprentissage machine nécessite un nombre de données important. Ces données doivent être autant que possible diversifiées pour pouvoir généraliser la solution.

## 3.3 Étude de segmentation sémantique multisource

L'architecture globale de notre proposition est tout d'abord présentée. Les différentes parties qui composent notre proposition sont par la suite présentées plus en détail.

### 3.3.1 Architecture de la proposition de segmentation sémantique multisource

Cette section développe principalement la proposition développée pour traiter le cas de la segmentation sémantique multisource. Notre *Workflow* (Figure 3.7) décrit les différentes

étapes utilisées pour segmenter des nuages de points multisources. Le flux de travail comprend quatre (4) parties. La première (A) comprend le calcul de certains attributs et la propagation d'autres. La partie A comporte trois branches (A1, A2 et A3), qui sont utilisées pour calculer les attributs monosources et multisources, et propager les attributs dans certains nuages de points au fur et à mesure de leur acquisition. La deuxième partie (B) englobe un premier réseau de neurones (NN 1), qui classe les points en utilisant leur vecteur d'attributs. La troisième partie (C) permet de calculer les attributs concernant les classes voisines. La quatrième partie (D) est un autre réseau de neurone (NN 2).

L'entrée principale est donc un ensemble de nuages de points visibles à gauche de la Figure 3.7 et qui représente la Localisation #1 (premier des trois cas décrits dans les résultats). Ces nuages sont issus de diverses technologies qui seront notées  $k \in [1,2,3]$  pour les trois sources. Dans ce cas précis de la Localisation #1, les données initiales proviennent de trois sources : TLS, ALS et photogrammétrie à partir d'un drone. Comme différentes technologies d'acquisition cohabitent, certains des attributs acquis avec une technologie (Rouge Vert Bleu (RVB), intensité, etc.) pourraient être absents d'un autre nuage de points acquis avec une technologie différente. Par conséquent, les attributs acquis sont représentés en **orange** tandis que les attributs manquants sont représentés en **gris foncé** et les attributs propagés sont représentés en **rose** sur la Figure 3.7.

Pour chaque point,  $P_i$  ( $i \in [1 \dots N]$ ), le vecteur  $T_{1,i}$  est composé des éléments suivants : les technologies en **gris**, les trente-cinq (35) attributs monosources calculés en **vert**  $\{f_{mono}\} \in \mathbb{R}^{35}$ , les trente-huit (38) attributs multisources calculées en **jaune**  $\{f_{multi}\} \in \mathbb{R}^{38}$  et les six attributs propagés en **rose**  $\{f_{propagés}\} \in \mathbb{R}^6$ .

Dans la partie A, les coordonnées cartésiennes des points sont utilisées pour calculer les attributs géométriques (branches A1 et A2). Ceci est décrit dans la section 3.3.2. Ensuite, les coordonnées sont également utilisées pour propager les attributs acquis avec certaines technologies à d'autres nuages (branche A3), comme nous le verrons plus loin dans la section 3.3.3.

Le premier réseau de neurones, nommé Neural Network 1 (NN 1), est ensuite utilisé pour calculer la probabilité d'appartenance de chaque point à chaque classe de la partie B de la

Figure 3.7 (ces valeurs sont représentées en rouge sur la Figure 3.7). Le NN 1 et ses utilisations sont décrits dans la section 3.3.4.

Dans la partie C, de nouveaux attributs, appelés attributs de voisinage (en **mauve**) sont calculés pour être utilisés dans le second réseau de neurone, NN 2. Pour calculer ces nouveaux attributs, la partie C a besoin des coordonnées et de la prédiction de NN 1. Le processus et les attributs sont décrits dans la section 3.3.5.

Dans la partie D, le NN 2 est utilisé pour segmenter le nuage de points. NN 2 a le vecteur  $T_{2,i}$  comme entrée. Ce vecteur est une concaténation de  $T_{1,i}$  ainsi que des attributs des voisins et de la prédiction de NN 1. À la fin, NN 2 fournit une segmentation sémantique affinée du nuage de points.

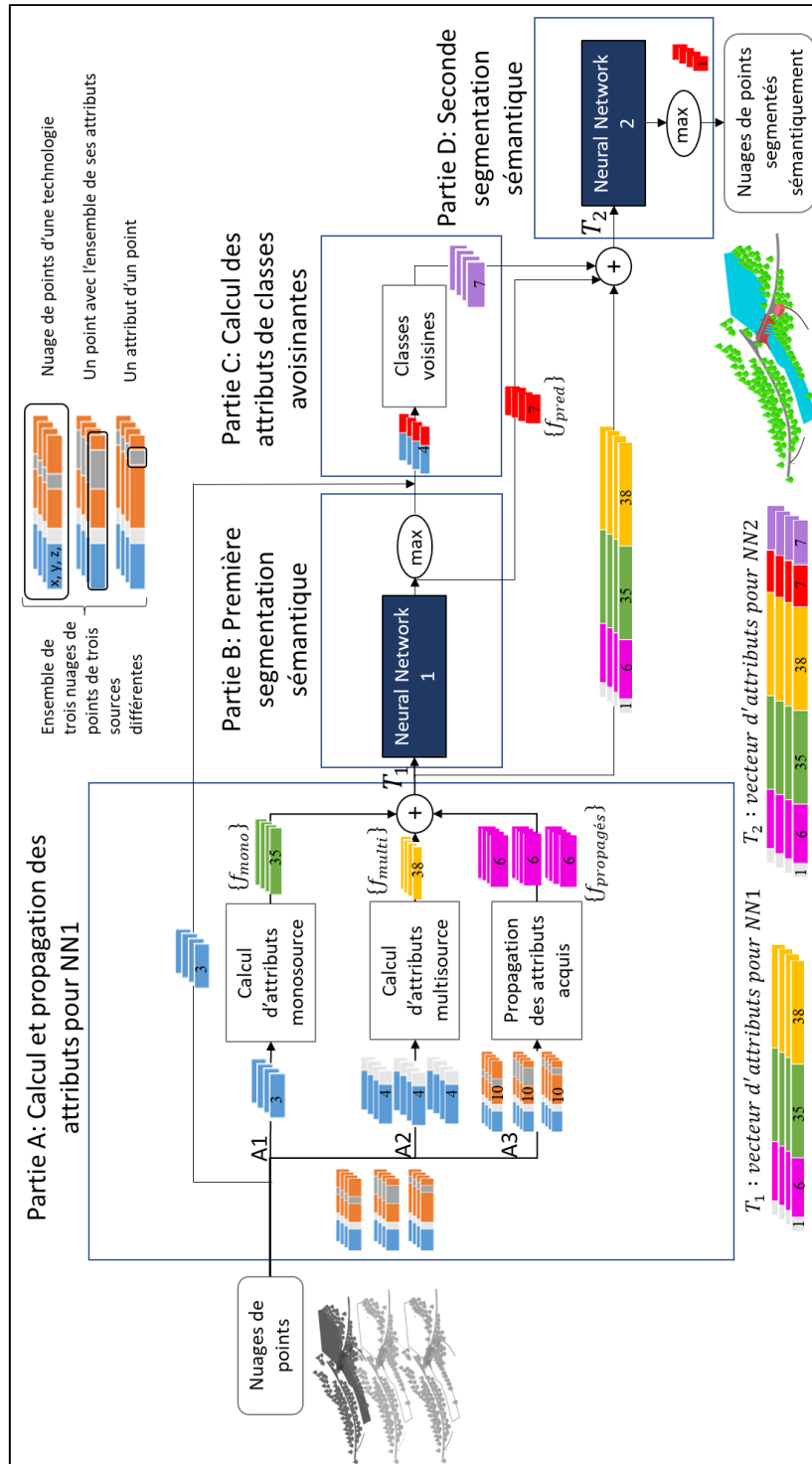


Figure 3.7 Workflow de notre proposition de segmentation multisource



### 3.3.2 Calculs d'attributs pour l'approche multisource

Comme nous l'avons vu au cours de la section 3.2, il existe différentes façons de calculer des attributs. Dans notre approche multisource, nous avons réutilisé les connaissances acquises lors de notre approche monosource dans le but de gagner du temps et de pouvoir comparer les deux approches. Le calcul des attributs ainsi que les attributs ne diffèrent pas entre les deux approches. Cependant, pour inclure l'aspect multisource, les attributs multisources prennent en compte l'ensemble des points présents dans la sphère ou le cylindre, peu importe leur source d'origine. Les attributs monosources quant à eux ne prennent en compte que les points provenant de la même source que le point dont les attributs sont calculés.

#### 3.3.2.1 Attributs

En plus des attributs déjà présentés dans le Tableau 3.2, de nouveaux attributs liés à l'aspect multisource ont été calculés et pris en compte pour l'approche multisource. Les attributs statistiques correspondent à la représentation de chaque source dans le nuage voisin. Le nombre de nouveaux attributs dépend donc du nombre de sources dans le nuage comprenant toutes les sources. Ces attributs ont été évalués pour mettre en évidence la différence de densité parmi les différentes sources et n'ont été calculés que pour une sphère de rayon  $R_{S_2} = 1$  m, l'ajout des sphères de rayon 0.5 et 3 m n'apportant que peu d'information supplémentaire.

#### 3.3.2.2 Sélection d'attributs multisource

La création du MNT de manière automatique avec des algorithmes fait encore l'objet d'études et n'est pas toujours viable selon les nuages de points dont on dispose. Après avoir essayé différentes techniques, les résultats pour les nuages de points des localisations multisources ont été moins probants qu'avec les localisations monosource (de multiples artefacts et erreurs apparaissaient). Le nombre d'attributs monosources passe donc à 35 comme visible (Figure 3.7). Comme pour la partie monosource, il est possible de retirer certains attributs qui apportent moins aux résultats finaux.

### 3.3.3 Propagation des attributs acquis

En plus des différents points de vue, les différentes technologies offrent une gamme d'attributs qui sont acquis pendant la numérisation de la zone. Ces différents attributs, qui sont inhérents à une ou plusieurs technologies, ne sont disponibles initialement que sur les points capturés par la ou les technologies respectives. Ces attributs sont représentés en orange sur la Figure 3.7. Les différentes technologies d'acquisition utilisées dans cette étude ont chacune des attributs différents qui sont acquis pendant la phase d'acquisition d'un projet. Les attributs à propager sont les couleurs (RVB), l'intensité du TLS, de l'ALS et du HLS, ainsi que le nombre de retours et les numéros de retour de l'ALS et du HLS.

La propagation de ces attributs aux autres points de la zone correspond à une fusion de données. En utilisant le voisinage, qui est le sous-ensemble du nuage de points contenu dans les sphères susmentionnées, combiné au calcul de la moyenne de l'attribut manquant dans le point, il est possible de diviser l'utilisation en trois (3) cas distincts visibles dans la Figure 3.8. Le premier cas (C1) est un point avec des attributs manquants qui a, dans son voisinage proche, un point avec les attributs qui lui manquent. Le deuxième cas (C2) est un point avec des attributs manquants qui n'a pas de valeur pour les attributs manquants dans son voisinage proche. Le troisième cas (C3) correspond à un point, avec des attributs manquants, qui n'a pas de voisinage proche.

Par exemple, avec C1, dans le scénario de la couleur manquante, seule une des trois technologies de la Localisation #1 possède des données RVB. Dans ce scénario, pour les points sans données RVB qui sont suffisamment proches d'un point avec des données RVB, le point le plus proche avec l'attribut manquant peut être considéré comme une approximation. Bien entendu, certaines conditions doivent être satisfaites pour que les attributs soient propagés. Une distance seuil ( $t_1$ ) de 5 cm est considérée comme une approximation raisonnable en raison des technologies utilisées et de la densité des nuages. La valeur des attributs manquante sera donc la même que celle du point le plus proche si la distance entre eux est inférieure à  $t_1$ . Cette approximation est efficace, mais dans le cas d'objets en mouvement, elle sera fautive. Si le nuage de points source et le nuage de points destination ne sont pas assez bien positionnés l'un par rapport à l'autre, ou si les deux nuages ont été numérisés à des périodes tellement éloignées

que des changements ont été apportés à la localisation, cette approximation sera beaucoup moins précise.

Dans certains cas (C2), lorsqu'un objet tel qu'un arbre est supprimé entre deux acquisitions, le plus proche voisin peut être trouvé à une distance supérieure à la valeur seuil. Une sphère de 1 m de rayon a été utilisée, et la médiane de l'élément manquant a été prise.

Dans le cas (C3), où ni le plus proche voisin ni la sphère de rayon 1 m ne sont suffisants pour inclure un point avec la caractéristique manquante, le point ne sera pas utilisé pour le processus de segmentation sémantique, et ne sera donc pas classé.

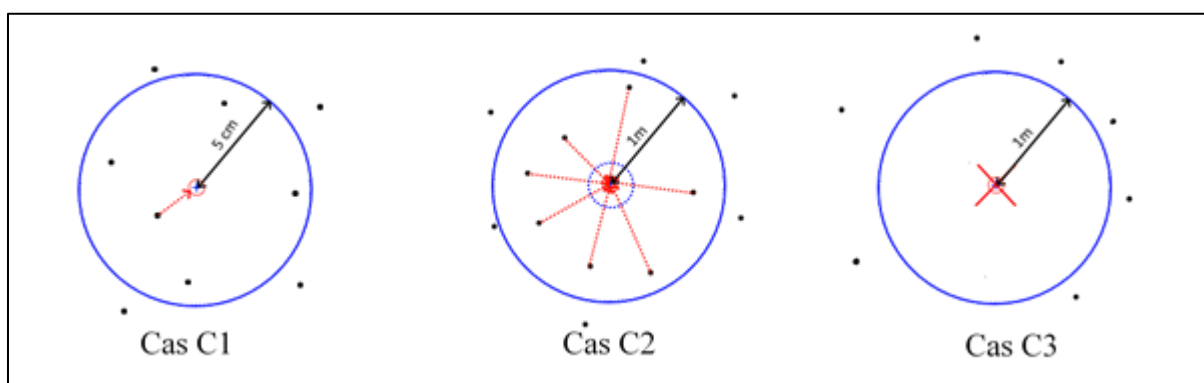


Figure 3.8 Schématisation des différentes possibilités C1, C2 et C3 pour la propagation des attributs acquis

D'autres méthodes ont été envisagées et/ou testées pour pallier le manque d'attributs acquis.

- Nous avons utilisé la valeur moyenne (ou médiane) comme valeur par défaut quand l'attribut était manquant.
- Nous avons essayé de ne pas utiliser les attributs acquis et de ne faire la segmentation sémantique qu'avec les attributs calculables.
- Nous avons envisagé l'utilisation de *masques* pour le réseau de neurones.
- Nous avons utilisé un réseau de neurones afin de déterminer la valeur de l'attribut (cette méthode est utilisée pour coloriser des images noir et blanc en couleurs).

Cependant, cette dernière méthode engendre un biais qui confirme la segmentation sémantique. Cela est dû à la boucle créée en utilisant les mêmes attributs calculés pour la création de la valeur de l'attribut manquant et pour la segmentation sémantique par la suite.

### 3.3.4 Réseau de neurones multisource (NN 1)

Comme pour la segmentation monosource, nous avons utilisé un premier réseau de neurones NN 1 afin de pouvoir classer nos points. Cette section élabore plus en détail les différents aspects de ce réseau de neurones.

#### 3.3.4.1 Architecture du réseau de neurones

Divers tests empiriques ont dû être réalisés pour obtenir l'architecture du réseau neuronal (NN 1) et ajuster les différents hyperparamètres. Tout d'abord, des architectures ont été testées avec deux, trois ou quatre couches cachées de neurones, quitte à avoir de *l'Overfitting* (surapprentissage) sur les données d'apprentissage pour réduire le biais. Ces premiers tests ont permis d'éviter le surapprentissage sur les données d'apprentissage tout en choisissant le taux d'apprentissage Alpha ( $10^{-5} \leq \alpha \leq 10^{-2}$ ) et le taux d'apprentissage le plus adapté aux données. Les tests ont été réalisés avec entre 16 et 1 024 neurones par couche cachée.

Dans un deuxième temps, le *Dropout* a été ajouté ( $0,2 \leq Dropout \leq 0,5$ ) et le poids régularisé ( $10^{-1} \leq l_2 \leq 10^{-5}$ ) pour réduire la variance. Le NN résultant est donc composé d'une première couche cachée de 512 neurones avec un *Dropout* de 0,5 et un poids régularisé de  $10^{-4}$  activé avec une unité linéaire rectifiée (*ReLU*). Les deuxième et troisième couches cachées étaient composées de 256 et 128 neurones, respectivement, avec un dropout de 0,5 et activées par une *ReLU*. La dernière couche cachée est composée de 72 neurones, avec un dropout de 0,5 et activé par une *ReLU*. La dernière couche est composée de 7 neurones - un pour chaque classe - et activée avec un *Softmax*. La *classe prédite* est celle qui a le score le plus élevé après la couche *Softmax*. Étant donné que les différents lieux géographiques, qui représentent chacun un ensemble de données différent, ne disposent pas nécessairement chacune de toutes les sources disponibles, l'architecture globale du NN reste la même, mais les entrées et les couches d'entrée varient légèrement. Comme mentionné précédemment, ces attributs ne sont pas présents dans tous les nuages de points ou dans tous les lieux géographiques ; c'est en partie pour cette raison que les algorithmes ont été entraînés et testés uniquement sur les mêmes types de nuages de points et dans les mêmes lieux.

Des poids ont été ajoutés aux classes pour améliorer l'apprentissage en fonction de l'emplacement et des classes présentes dans les nuages. On applique ces poids dans le réseau de neurone. Les poids permettent de prendre en compte la différence de nombre de points dans les différentes classes. La formule pour les poids des classes (Eq. 3.1) permet de prendre en compte le nombre de points par classe ( $n_j$ ) sans arriver à des poids trop élevés pour les classes dominantes ou trop faibles pour les classes qui sont représentées par moins de points.

$$w_j = \frac{1}{\sqrt{n_j}} \quad (3.1)$$

où  $w_j$  est le poids de la classe  $j$  et  $n_j$  est le nombre de points de la classe  $j$ .

Les entrées de NN1 sont : les technologies en gris, les attributs monosources calculés en **vert**  $\{f_{mono}\}$ , les attributs multi-sources calculés en **jaune**  $\{f_{multi}\}$  et les attributs propagés en **rose**  $\{f_{propagés}\}$ . Les sorties de NN1 sont les probabilités, en **rouge**  $\{f_{pred}\} \in \mathbb{R}^7$ , pour chaque point d'appartenir aux classes proposées.

### 3.3.4.2 Hyperparamètres dans le cas de la segmentation sémantique multisource

Les hyperparamètres, déjà énoncés dans la partie monosource, ont pu être davantage travaillés pour cette partie. Le nombre de couches, le nombre de neurones par couches, ainsi que certains paramètres tel qu'Alpha ont déjà été présentés dans la partie architecture. Les hyperparamètres qui restent sont : le nombre d'époques, de lots et le nombre de points par lot. Ils ont eux aussi été déterminés de manière empirique au cours de la création du réseau de neurones. L'ensemble de nos données ont été utilisées afin de trouver un compromis pour les paramètres.

### 3.3.4.3 Découpage des données en jeux d'apprentissage, de test et de validation

Comme pour la partie monosource, il est nécessaire pour entrainer le réseau de découper nos jeux de données en plusieurs parties. Encore une fois les parties apprentissage, test et validation ont été créées. Les différentes localisations qui représentent nos jeux de données sont hétérogènes en sources de données, classes, représentation de ces classes et attributs acquis.

De ce fait, nous avons préféré faire l'apprentissage, la partie test et validation sur la même localisation. Cette contrainte est forte, elle restreint l'apprentissage à chaque localisation et ne permet donc pas une généralisation de l'apprentissage sur l'ensemble des données disponibles. Afin de réaliser le découpage du jeu de données en différentes parties, nous avons fait le choix de découper pour les Localisation #1 et Localisation #2 en séparant le jeu de données de manière manuelle. Un échantillonnage aléatoire est parfois utilisé pour réaliser cette action. Cependant, cette méthode créera, dans notre cas, un résultat biaisé. Des points de l'apprentissage apparaissent dans les sphères des points pour le test ou la validation. Dans le cas où 70% du jeu de données est utilisé pour l'apprentissage, un point du jeu de test a des chances d'être entouré de points d'apprentissage. Les rayons des sphères étant de l'ordre du mètre, les valeurs des attributs de deux points adjacents sont très proches. Ce biais, a donc été en partie enlevé avec la découpe manuelle des jeux de données. Ce découpage a aussi permis de s'assurer une représentation suffisante des classes.

Dans le cas de la Localisation #3, le jeu de données est composé de trois emplacements. Deux d'entre eux ont servi pour l'apprentissage. Le dernier a ensuite été découpé pour la partie test et la partie validation.

### **3.3.5 Calcul des attributs de classes avoisinantes**

À l'issue du premier réseau de neurone NN 1, et afin de tirer le meilleur parti de la première segmentation sémantique, il a été décidé de créer des attributs prenant en compte les classes du voisinage. Ces attributs correspondent à un dénombrement du nombre de points appartenant à chaque classe dans le nuage de points voisins après la première classification. Pour une classe donnée, les points voisins ne peuvent appartenir qu'à un nombre de classes limité. C'est pourquoi les « Conducteurs » ont comme classe voisine d'autres « Conducteurs », des « Pylônes », des « Bâtiments », ou dans le cas de voisinage plus grands, de la « Végétation ». Chaque point aura donc un nombre d'attributs correspondant au nombre de classes (7), dont leurs valeurs sont le nombre de points voisins. Pour un point appartenant au « Sol », par exemple, on trouve que les voisins sont issus des classes « Sol » et « Routes » (histogramme 1 de la Figure 3.9). Pour des points appartenant à la classe « Pylône », on retrouve dans les

classes voisines, des points qui sont du « Pylône », des « Conducteurs » une partie de la « Végétation » et des points du « Sol » (histogramme 2 de la Figure 3.9). Pour ces représentations Figure 3.9, un point du sol histogramme 1 et un point d'un pylône histogramme 2 ont été pris en exemple. Les deux histogrammes supplémentaires correspondent à un point de la surface hydrologique pour l'histogramme 3 et un point du sol pour l'histogramme 4. La visualisation créée pour la Figure 3.9 est réalisée avec les voisinages d'un point à chaque fois. Les valeurs pour chaque classe sont exprimées en pourcentage (%) de la représentation du voisinage puis présentées dans une échelle logarithmique pour obtenir une meilleure visualisation.

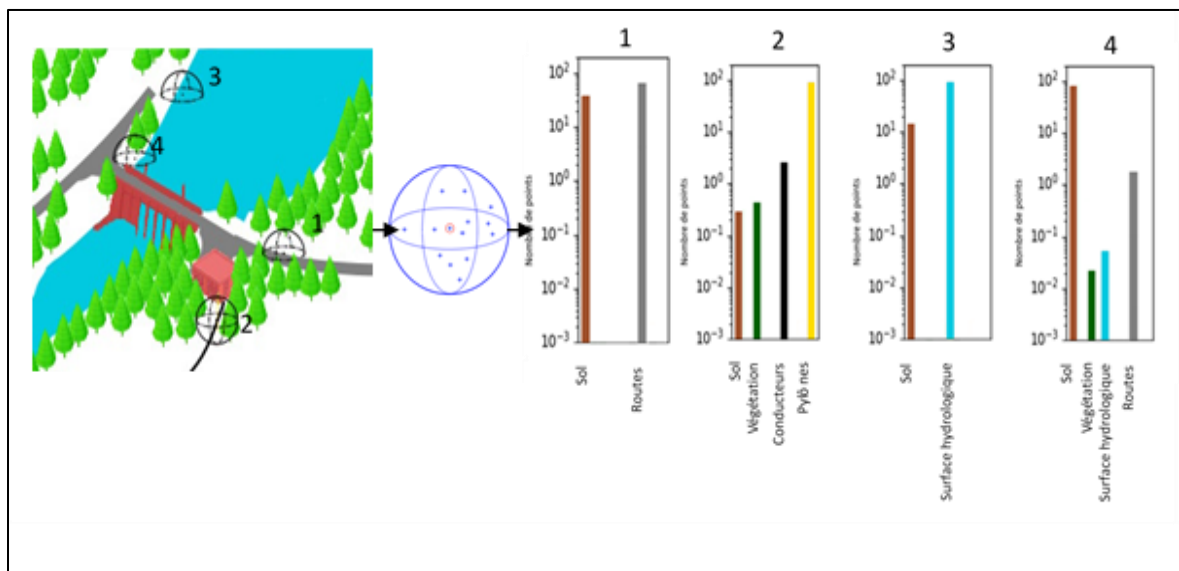


Figure 3.9 Schématisation de la création des attributs de classes avoisinantes sur la Localisation #1

Aussi, l'image ne représente pas l'ensemble des points de la classe. Certains points peuvent avoir des voisinages plus complexes. Un point du « Sol » peut avoir entre autres des points voisins appartenant aux classes « Surfaces Hydrologiques », « Pylônes », ou encore faisant partie de la « Végétation » par exemple.

Bien qu'il soit possible d'utiliser plusieurs sphères de tailles différentes, nous avons opté pour une seule sphère de rayon  $R_{S_2} = 1$  m afin de réduire le temps de calcul après avoir réalisé

différents essais. Ces attributs de classes voisines peuvent être retrouvés dans l'utilisation de *Conditional Random Field* (CRF) ou des réseaux de neurones de convolution.

### 3.3.6 Seconde segmentation sémantique (NN 2)

Le deuxième réseau de neurones, NN 2, utilisé dans la partie D a le même nombre de couches et la même composition que le NN 1. Le seul changement réside dans la taille de la couche d'entrée, car le nombre d'attributs a augmenté entre la première et la deuxième segmentation sémantique. Les entrées de NN 2 sont : les technologies (en **gris**), les attributs monosources calculés (en **vert**  $\{f_{mono}\}$ ), les attributs multi-sources calculés (en **jaune**  $\{f_{multi}\}$ ) et les attributs propagés (en **rose**  $\{f_{propagés}\}$ ), les attributs voisins (en **mauve**) et la prédiction du premier NN (en **rouge**). Les sorties de ce second NN sont les probabilités (en **rouge**  $\{f_{pred}\}$ ) de chaque point appartenant aux (7) classes proposées. La classe qui récolte le plus haut pourcentage est alors conservée et le point est classé dans cette classe.

L'apprentissage des réseaux de neurones a été découplé. Dans un premier temps, l'apprentissage pour le réseau NN 1 a été réalisé puis une fois les attributs de voisinage calculés, l'apprentissage de NN 2 a été réalisé.

## 3.4 Optimisation des nuages de points

Les nuages de points sont donc segmentés sémantiquement et chaque point encapsule les informations provenant de sa source ainsi que des sources proches. Cependant, les nuages de points provenant de différentes sources peuvent se recouper, se chevaucher et se recouvrir. Dans l'optique de ne garder que les informations les plus pertinentes, nous avons envisagé d'optimiser le nuage de point résultant. Cette optimisation revient à ne garder par zone segmentée qu'une source de données. Pour y arriver, il est nécessaire de segmenter chaque classe en ses entités qui la compose. Puis de réaliser une sélection des points ou des ensembles de points selon les sources disponibles, ainsi que d'autres caractéristiques comme la classe représentée par le point ou la date de captation ou encore le niveau d'importance de la zone.



### 3.4.1 Segmentation en entités distinctes

Afin de segmenter le plus de classes possibles en ses différentes entités qui la composent, nous avons décidé de prendre en considération la classe de chaque point ainsi que la distance qui le sépare des points qui appartiennent à la même classe (comme schématisé Figure 3.10). Le paramètre distance est à l'heure actuelle choisie par l'opérateur. L'opérateur doit sélectionner ce paramètre en fonction de la densité du nuage de points sur lequel il travaille. Une fois le paramètre de distance sélectionné, un algorithme de *Connected-Components Labeling* est utilisé. Cet algorithme fonctionne avec la théorie des graphes et va créer un graphe à partir des points ayant reçu la même classification, puis le graphe sera découpé en fonction du paramètre de distance sélectionné.



Figure 3.10 Schématisation d'une segmentation de classe en entité distinctes

Des arbres proches les uns des autres ne seront pas segmentés en entités individuelles, mais seront regroupés comme une seule entité. Le fait que les arbres ne soient pas segmentés de cette manière peut porter préjudice dans les projets d'élagage par exemple. Des méthodes qui permettent de segmenter les arbres individuellement existent et sont en usage pour différents projets. Ces méthodes restent de l'ordre de la segmentation particulière, elles ne visent que la végétation et sont souvent liées à un type de végétation. Elle repose sur des méthodes qui utilisent les paramètres de la végétation comme la couronne des arbres et leur tronc pour les découper en entités distinctes.

On retrouve des découpes particulières pour les bâtiments pour des projets de reconstruction de ville par exemple. Dans ce cas, la segmentation en entités distinctes peut reposer sur la reconnaissance des toits puis, une fois les toitures reconnues, les bâtiments sont découper les

uns des autres. Avec notre méthode, les bâtiments collés les uns aux autres formeront un regroupement, les bâtiments ne seront pas distincts.

### 3.4.2 Sélection de la source

Une fois les différents nuages de points segmentés, il sera possible de sélectionner pour chaque zone d'intérêt la source que l'on veut utiliser. Cette opération permettra d'obtenir un nuage unique ne comprenant que les informations et la densité des points utiles. On peut parler d'une représentation optimale pour le stockage de l'information.

Le terme d'optimisation découle du fait que le nuage final sera réduit en taille et ne comportera pas autant de points que les ensembles originaux des nuages. Pour l'optimisation du nuage final, il faudra utiliser les nuages correspondant à la bonne technologie pour chaque couple zone géographique et classe. Pour arriver à ce résultat, différents critères seront pris en compte tels que la classe a priori du point selon la technologie d'acquisition disponible. Cette partie est réalisée en collaboration avec le PI, c'est lui qui fixe les informations dont il a besoin ainsi que les zones d'intérêt. Dans un premier temps, les caractéristiques du nuage telles que sa densité et sa précision ont été envisagées pour sélectionner la technologie parmi les technologies disponibles. De nouveaux attributs ont été envisagés. Ils sont à l'heure actuelle au nombre de deux et correspondent à la **date d'acquisition** et au niveau **d'intérêts**.

- La **date d'acquisition** peut jouer sur certains éléments contenus dans le nuage de points. En effet, des bâtiments ou de la végétation pourront disparaître ou apparaître selon la date du relevé. Une comparaison des nuages de points devra donc être effectuée afin de s'assurer qu'il y a peu de différence. S'il y en a, il faudra alors caractériser les manques ou ajouts de points et les faire apparaître sur le nuage final de points. Pour juger des différences entre les différents nuages de points issus d'une même zone géographique (avec des temporalités différentes), des algorithmes de comparaison présents dans la littérature sont à explorer dans de futurs travaux (ex. algorithmes M3C2 *Multiscale Model to Model Cloud Comparison* ou encore C2C *Cloud to Cloud* par exemple).

- Les **niveaux d'intérêts** proviennent quant à eux à la fois de la classe des points et de la zone géographique. Le PI souhaite que le nuage de points qui correspondent à ses bâtiments ou à ses postes de travail soit le plus précis possible, mais porte moins d'intérêt aux bâtiments ne faisant pas partie des siens. Ainsi, certaines zones géographiques devront comporter un nuage de points provenant d'une source autre, afin d'augmenter la précision, malgré la nature similaire de la classe du point. Les niveaux d'intérêts influent aussi selon le projet à effectuer. Par exemple, dans le cas d'un projet concernant le volume de biomasse à enlever, la végétation aura un plus grand niveau d'intérêt que pour un projet concernant des bâtiments par exemple.

Un arbre de décision a ainsi été créé (Figure 3.11). Cet arbre doit permettre la sélection de la source à retenir en fonction de la classe. Ce premier jet a été créé après un entretien avec l'équipe de géomatique du PI. La classe de Surface hydrologique est principalement numérisée par les ALS comme ils peuvent survoler les étendues d'eau. Les classes de Pylônes et Conducteurs sont critiques et nécessitent les données les plus précises possibles. Pour les autres classes, le premier choix dépend de l'intérêt de la zone pour le PI ou si les Bâtiments ou leur accès appartiennent au PI.

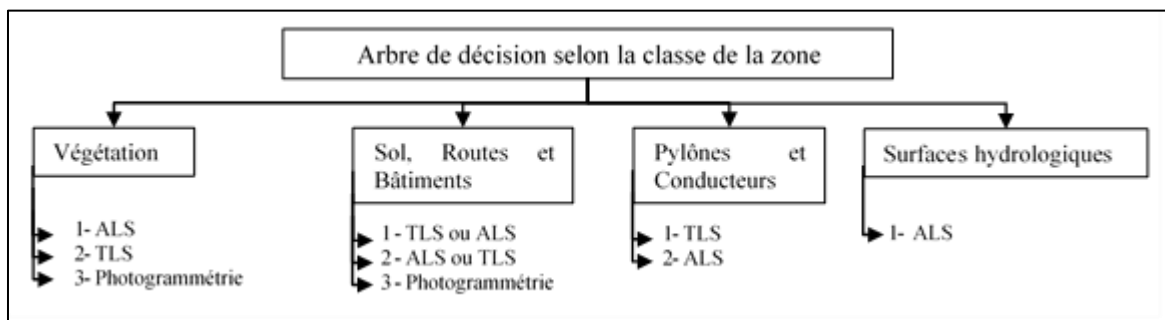


Figure 3.11 Arbre de priorisation des technologies en fonction de la classe du point et des sources disponibles

À noter qu'un système de poids, qui permet de pondérer en fonction de la date d'acquisition, de la source (et donc de la précision) ainsi que du niveau d'intérêt, permettrait de créer un

système plus complet qui prendrait en compte l'ensemble des attributs à notre disposition. Cette proposition pourra, elle aussi, être étudiée dans de futurs travaux.

### **3.5 Synthèse et contributions**

Notre proposition propose une nouvelle méthode de segmentation sémantique en lien avec l'utilisation de multiple source de données sur la même zone géographique. Les premiers tests sur les données monosources nous ont permis de prendre en main les données, comprendre les avantages et inconvénients de différents algorithmes sur des jeux de données complexes. Cette première partie nous a aussi permis de mettre en œuvre une stratégie de *Transfert Learning*.

La suite de notre proposition porte plus en détail sur les données multisources, à l'usage des différents attributs acquis et calculés. Cette seconde approche, développée et publiée dans un article de conférence, a permis la proposition de nouveaux attributs. Ceci nous a permis de développer une méthode de fusion de données par propagation, en plus de la fusion de données sur le calcul d'attributs.

Enfin, une seconde segmentation en cascade qui prend en compte les résultats de la première segmentation ainsi que les classes avoisinantes a été mise en place, validée et testée. Cette contribution a été soumise dans notre second article pour journal avec comité de lecture.

À noter que l'accroissement de la taille du nuage de points et son optimisation via la sélection de sources sont, pour le moment, surtout une approche théorique qui permet au PI de mettre en place l'arbre de décision de sélection de sources vu dans ce chapitre. L'intérêt de cette optimisation peut mener à débattre sur la sélection d'une technologie spécifique de captation ou de supprimer les données provenant des sources qui ne sont pas sélectionnées. Enfin, les aspects de datation et de niveaux d'intérêt, bien que décrit dans ce chapitre n'ont pas encore été mis en place en réalité.

Le prochain chapitre présente et discute l'ensemble des résultats obtenus grâce au *workflow* développé.



## CHAPITRE 4

### RÉSULTATS

#### Sommaire

4.1	Introduction.....	91
4.2	Présentation des données utilisées .....	92
4.2.1	Classes représentées dans les nuages de points utilisés .....	93
4.2.2	Données monosource .....	98
4.2.2.1	Localisation #4 et #5 .....	99
4.2.2.2	Localisation #6.....	101
4.2.3	Données multisources .....	102
4.2.3.1	Localisation #1.....	102
4.2.3.2	Localisation #2.....	105
4.2.3.3	Localisation #3.....	105
4.3	Résultats de l'étude sur l'apprentissage monosource .....	106
4.3.1	Segmentation sémantique monosource.....	106
4.3.1.1	Réduction du nombre de classes .....	112
4.3.1.2	Post-traitement .....	113
4.3.2	Synthèse sur l'apprentissage monosource .....	114
4.4	Résultats de l'étude d'ablation sur la proposition multisource .....	117
4.4.1	Segmentation sémantique monosource (chemin A1+A3, B).....	118
4.4.2	Segmentation sémantique multisource (chemin A1+A2+A3, B) .....	121
4.4.3	Segmentation sémantique en cascade (chemin A1+A2+A3, B, D).....	124
4.4.4	Segmentation sémantique avec prises en compte du voisinage (chemin A1+A2+A3, B, C, D).....	126
4.4.5	Synthèse de l'étude d'ablation et comparaison des différents résultats..	129
4.5	Synthèse globale .....	135

#### 4.1 Introduction

Pour évaluer notre proposition et nos contributions, des expérimentations sur nos différents jeux de données ont été mises en œuvre. Ces expérimentations permettent de mesurer l'adéquation et la précision de notre proposition. Dans notre cas, les données pour nos

expérimentations ont été fournies par le PI. Elles sont présentées dans la première partie du présent chapitre. Dans un deuxième temps, les résultats de la première étude sur les données qui n'ont qu'une source en entrée sont présentés. Il s'agit de notre approche monosource. Dans un troisième temps, une étude d'ablation est proposée comme une analyse de sensibilité dans le but d'identifier l'apport des différentes parties de notre proposition. Une synthèse présentée à la fin du présent chapitre permet de mettre en avant les avantages de notre proposition. La comparaison à des solutions existantes est proposée dans le chapitre suivant.

## 4.2 Présentation des données utilisées

Cette partie présente les jeux de données utilisés au cours de cette thèse de doctorat dans le but d'évaluer les performances de notre proposition. Les données employées ont été fournies par le PI et sont ici présentées selon leurs localisations ainsi que selon les technologies de captation utilisées. Dans un premier temps, les différentes classes rencontrées sont décrites selon les éléments qui les composent. Ces classes sont aussi catégorisées selon la diversité et la représentativité au travers des nuages disponibles. En effet, pour caractériser les classes de points, deux notions sont nécessaires : la diversité et la représentativité d'une classe. La **diversité** d'une classe correspond à la différence qui existe entre les différents éléments qui la composent. Cette valeur est d'autant plus grande que la classe peut être subdivisée en sous-classes représentant plus fidèlement les éléments qui la composent. Il est possible ainsi de prendre comme exemple une classe englobant l'ensemble de la végétation. Cette classe a une grande diversité, il est possible de la subdiviser en classe plus spécifique (ex. une classe qui n'est composée que de sapins). Quant à la **représentativité** d'une classe, il s'agit d'un critère qui mesure la présence dans les nuages de points de l'ensemble de la diversité possible de la classe, en nombre suffisant pour être pris en compte par un algorithme d'apprentissage machine. En reprenant l'exemple de la végétation, la classe de végétation possède une haute représentativité si et seulement si elle est composée de l'ensemble des éléments que l'on peut trouver dans cette classe. Par conséquent, et pour garantir les meilleures performances possibles, il est impératif de manipuler des nuages de points représentatifs des classes que l'on recherche à segmenter.



#### 4.2.1 Classes représentées dans les nuages de points utilisés

Lors d'une segmentation sémantique d'un nuage de points, chaque point se voit assigner une classe qui représente l'entité dont fait partie le point. Dans notre cas, les classes possibles pour la segmentation sont dans un premier temps au nombre de neuf (9), mais nous verrons que certaines d'entre elles ont été fusionnées dans le reste de l'étude.

- Le « **Sol** » est la classe qui est la plus représentée dans les nuages de points. Un exemple de cette classe est donné dans la Figure 4.1. Selon la localisation et la nature du terrain (ex. forêts, vallée, ville ou montagne) l'aspect géométrique du sol est amené à changer mais, dans l'ensemble, le sol reste une classe avec peu de diversité. Selon la saison, le sol peut être recouvert de neige, d'eau, de végétation ou de verglas. Dans ces cas, il est possible que certains attributs tels que la couleur ou encore l'intensité de retour du laser changent à travers les saisons.
  - o Diversité : basse. Représentativité : haute.

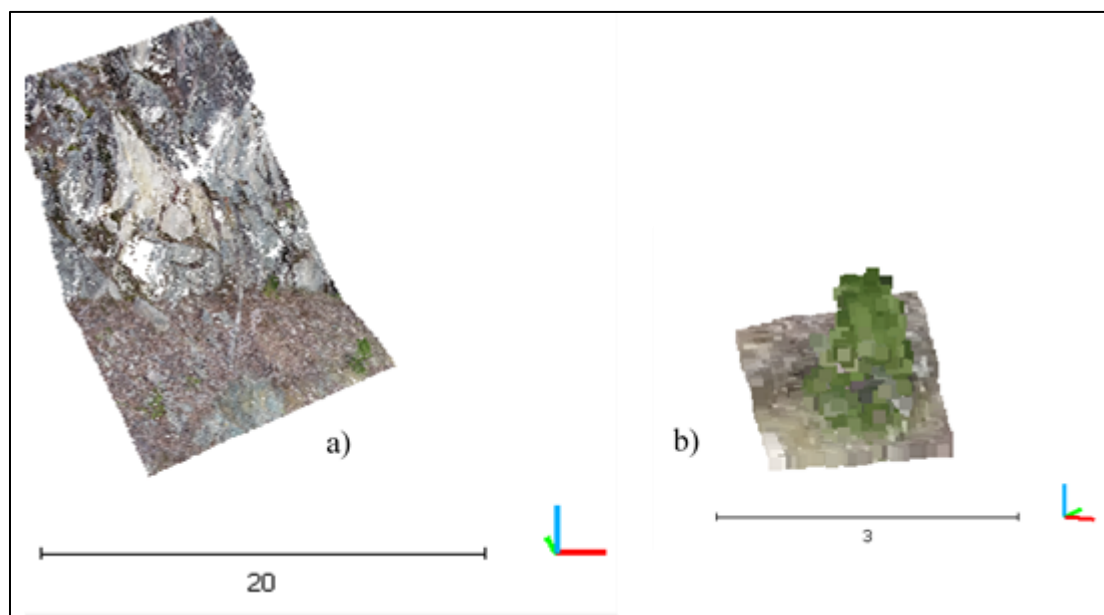


Figure 4.1 a) Exemple de la classe Sol de la Localisation #1, b) Exemple de la classe Végétation basse de la Localisation #1

- La « **Végétation basse** » est constituée de végétation inférieure à 2 m de hauteur selon l'ASPRS. On peut y retrouver des arbustes et autres types de végétation basse. Selon le nuage de points retenu, il est possible que la classe végétation basse corresponde à tous points de végétation inférieure à 2 m. Il est aussi possible de retrouver la classe de végétation moyenne qui est quant à elle comprise entre 2 m et 5 m de hauteur. Cette classe est peu représentée en termes de nombre de points, malgré la diversité potentielle de la végétation. Un exemple de cette classe est donné dans la Figure 4.1.
  - Diversité : moyenne. Représentativité : basse.
  
- La classe « **Végétation haute** » est, quant à elle, constituée des éléments de végétation supérieure à une hauteur de 5 m. À part les nuages de points des zones urbaines, il s'agit de la classe la plus représentée dans les nuages fournis par le PI. La diversité des éléments composants la classe peut la rendre plus difficile à segmenter sémantiquement. Dans la suite de nos expérimentations, nous avons retenu le maintien d'une seule classe réunissant les végétations basses et hautes que nous désignons pour la suite comme « Végétation ». Un exemple de cette classe est donné dans la Figure 4.2.
  - Diversité : moyenne. Représentativité : haute.

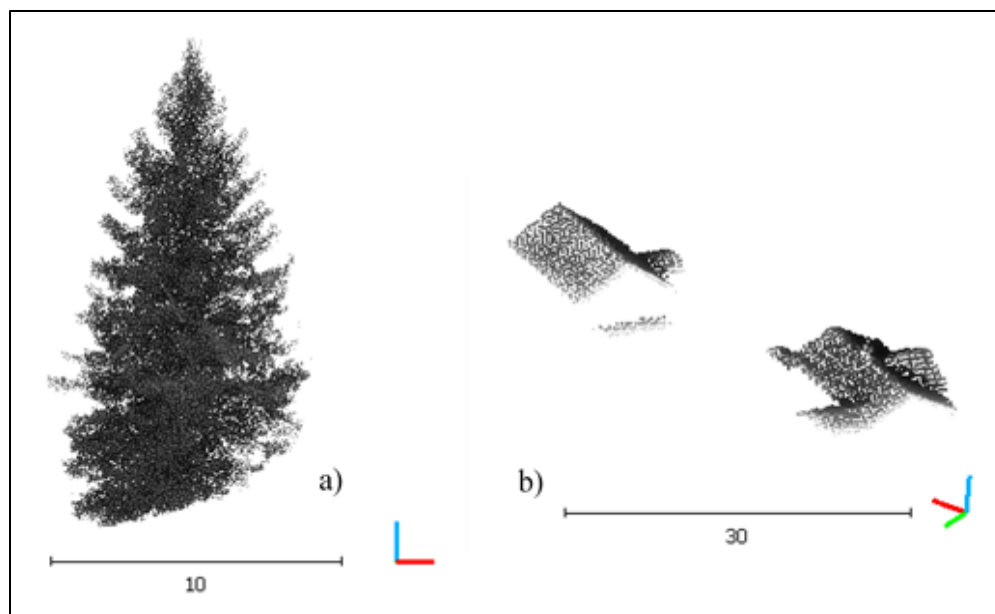


Figure 4.2 a) Exemple de la classe Végétation haute de la Localisation #1,  
b) Exemple de la classe Bâtiments de la Localisation #6

- La classe des « **Bâtiments** » est constituée principalement de bâtiments ainsi que d'autres infrastructures de fabrication humaine (ex. ponts et barrages de retenue d'eau). Selon la localisation, les bâtiments peuvent être de diverses natures (ex. maisons, immeubles, etc.). Certains bâtiments tels que les édifices religieux peuvent aussi augmenter la diversité de cette classe. Les infrastructures peuvent augmenter de manière considérable la diversité de cette classe. Un exemple de cette classe est donné dans la Figure 4.2.
  - Diversité : haute. Représentativité : basse.
  
- La classe des « **Surfaces hydrologiques** » est constituée de surfaces appartenant à des étendues d'eau. On y retrouve principalement des lacs, des rivières et des fleuves. Il est aussi possible d'y retrouver des piscines ou encore des bassins. Ces surfaces peuvent réfléchir et interagir avec les lidars en créant des artefacts qui brulent les captations de données. Un exemple de cette classe est donné dans la Figure 4.3.
  - Diversité : moyenne. Représentativité : basse.

- La classe des « **Conducteurs** » est constituée de conducteurs d'électricité. Il s'agit de câbles de transmission d'énergie électrique (pour différentes tensions, hauteurs et configurations). Ils ont été regroupés dans la même classe malgré des différences notables dans leurs morphologies. Leurs caractéristiques géométriques uniques les rendent plus simples à segmenter. Elle n'en reste pas moins que c'est l'une des classes ne comprenant qu'un petit nombre de points. Un exemple de cette classe est donné dans la Figure 4.3.
  - Diversité : basse. Représentativité : haute.

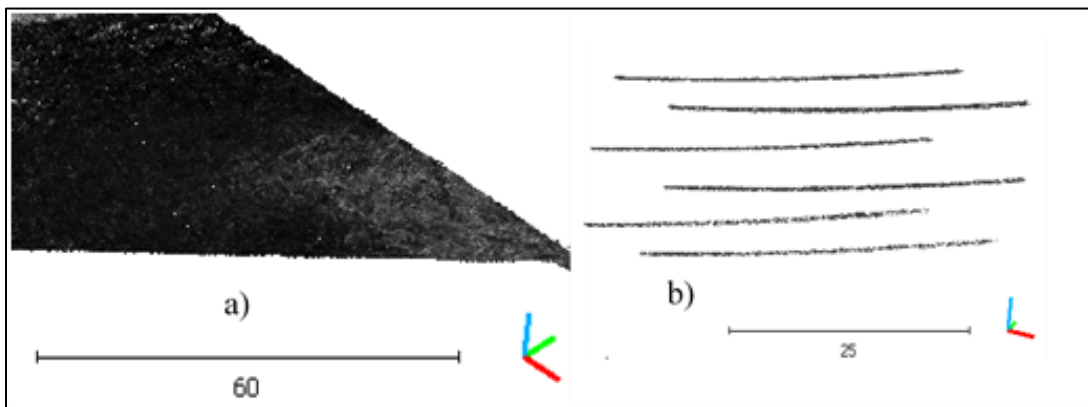


Figure 4.3 a) Exemple de la classe Surfaces hydrologiques de la Localisation #4,  
 b) Exemple de la classe Conducteurs de la Localisation #3

- La classe de « **Routes** » est dissociée du « Sol » en permettant la circulation de véhicules. Cette classe est donc constituée principalement de routes, mais elle comprend aussi les chemins ainsi que des voies ferrées. Un exemple de cette classe est donné dans la Figure 4.4.
  - Diversité : basse. Représentativité : basse.

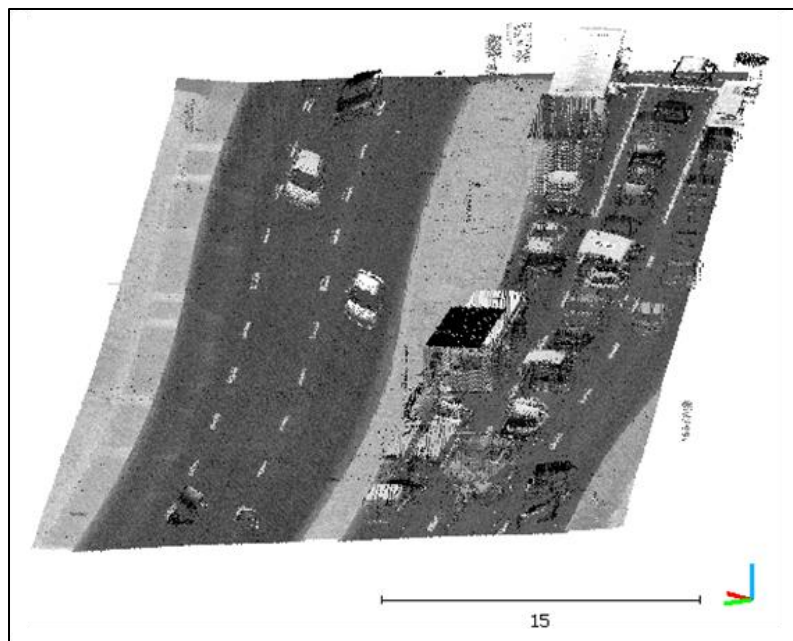


Figure 4.4 Exemple de la classe Routes de la Localisation #3

- La classe des « **Pylônes** » est constituée de structures de fabrication humaine servant au maintien des conducteurs. Ces structures peuvent varier de forme, de taille et des matériaux employés (bois, acier, etc.). Plusieurs sous-classes auraient donc pu être créées afin de représenter cette diversité. Là aussi, il s'agit d'une classe qui ne comporte que peu de points. Il est possible de trouver sur les sites de fournisseurs d'énergies une représentation exhaustive des différents types de pylônes qui existent (Figure 4.5). Cette classe prend aussi en compte les postes de transformateurs de puissance.
  - Diversité : moyenne. Représentativité : haute.

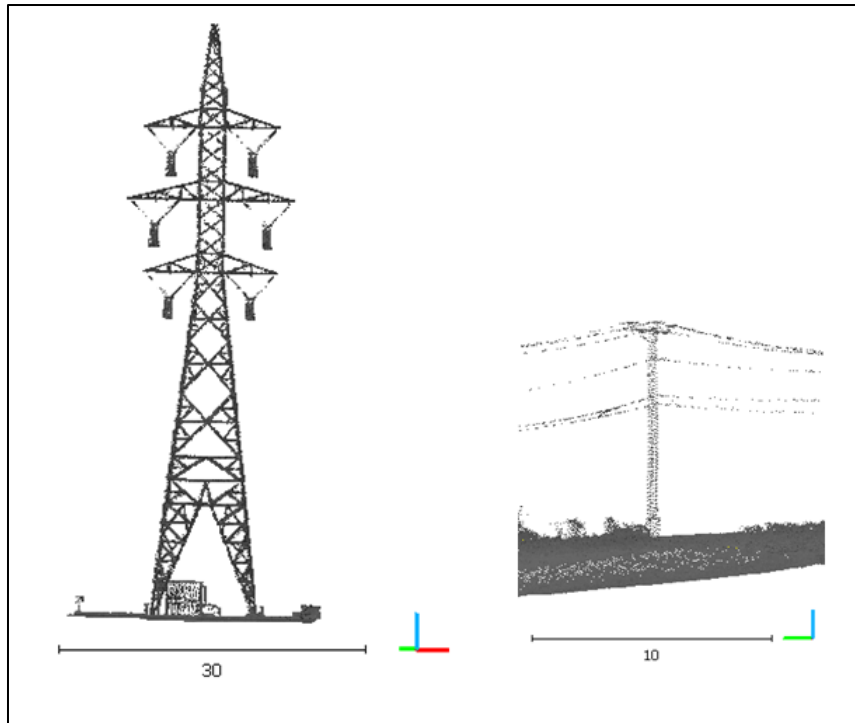


Figure 4.5 Exemples de la classe Pylônes de la Localisation #3

- Enfin, la classe « **Autres** » est constituée des éléments restants qui n'appartiennent à aucune classe mentionnée précédemment. Il peut s'agir de panneaux de signalisation, de rambardes de sécurité, glissières, lampadaires ou encore des personnes et véhicules. Les objets en mouvement, tels que les véhicules et personnes, sont retirés des nuages de TLS suite à un prétraitement du PI qui élimine les points qui n'apparaissent pas sur plusieurs de leurs numérisations TLS. À cause de la diversité de cette classe ainsi que sa sous-représentation dans les nuages, cette classe est ignorée pour la suite.
  - Diversité : haute. Représentativité : basse.

#### 4.2.2 Données monosources

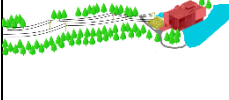


Comme exposé au long de la méthodologie, notre étude porte sur l'utilisation de données provenant de différentes sources captées sur une même localisation. Cependant, il est possible que, pour certaines localisations, une seule source de données soit disponible.

Les nuages de points monosources ont permis de démarrer nos travaux, ils ont contribué à interpréter la segmentation sémantique et à tester les premiers algorithmes avant de passer au modèle plus complexe qui intègre différentes sources de données.

#### **4.2.2.1 Localisation #4 et #5**

Ces deux premières localisations, visibles dans la Figure 4.6, sont numérisées à l'aide d'ALS. Elles représentent des lieux proches géographiquement et similaires dans les classes présentes et dans leurs représentations. Elles comptent une grande partie de végétation, quelques bâtiments et elles sont proches d'une retenue d'eau. On y retrouve les mêmes classes avec un nombre de points similaire qui sont visibles dans le Tableau 4.1 ainsi que dans la Figure 4.6 et Figure 4.7. La segmentation sémantique, utilisée pour l'apprentissage, est réalisée manuellement par des experts du domaine et est vérifiée par le PI. Comme la numérisation est réalisée à l'aide d'ALS, les attributs acquis sont les mêmes, ils sont aux nombres de six (6). Ces attributs sont l'identifiant de la source, l'angle du laser lorsque le point est numérisé, le nombre de retours et le numéro du retour, le temps GPS de la captation et, enfin, l'intensité de retour du point. Les deux nuages étant proches géographiquement et ayant des similitudes, ils sont utilisés concurremment pour l'apprentissage machine.

Tableau 4.1 Représentation et composition des nuages de points monosource

	Localisation #4	Localisation #5	Localisation #6
			
<b>Technologies</b>	ALS	ALS	ALS
<b>Sol</b>	561,970 (38.7%)	539,105 (30.7%)	2,466,636 (49.5%)
<b>Végétation basse</b>	17,508 (1.2%)	23,276 (1.3%)	-
<b>Végétation haute</b>	673,326 (46.4%)	894,527 (51.0%)	2,215,111 (44.5%)
<b>Bâtiments</b>	65,191 (4.5%)	95,756 (5.5%)	49,520 (1.0%)
<b>Surfaces hydrologiques</b>	12,543 (0.9%)	149,355 (8.5%)	21,802 (0.4%)
<b>Conducteurs</b>	9,297 (0.6%)	11,167 (0.6%)	50,520 (1.0%)
<b>Routes</b>	28,195 (1.9%)	10,019 (0.6%)	120,664 (2.4%)
<b>Autres</b>	75,596 (5.2%)	14,785 (0.8%)	-
<b>Pylônes</b>	8,179 (0.6%)	17,666 (1.0%)	55,045 (1.1%)
<b>Total</b>	<b>1,451,805 (100%)</b>	<b>1,755,656 (100%)</b>	<b>4,929,298 (100%)</b>

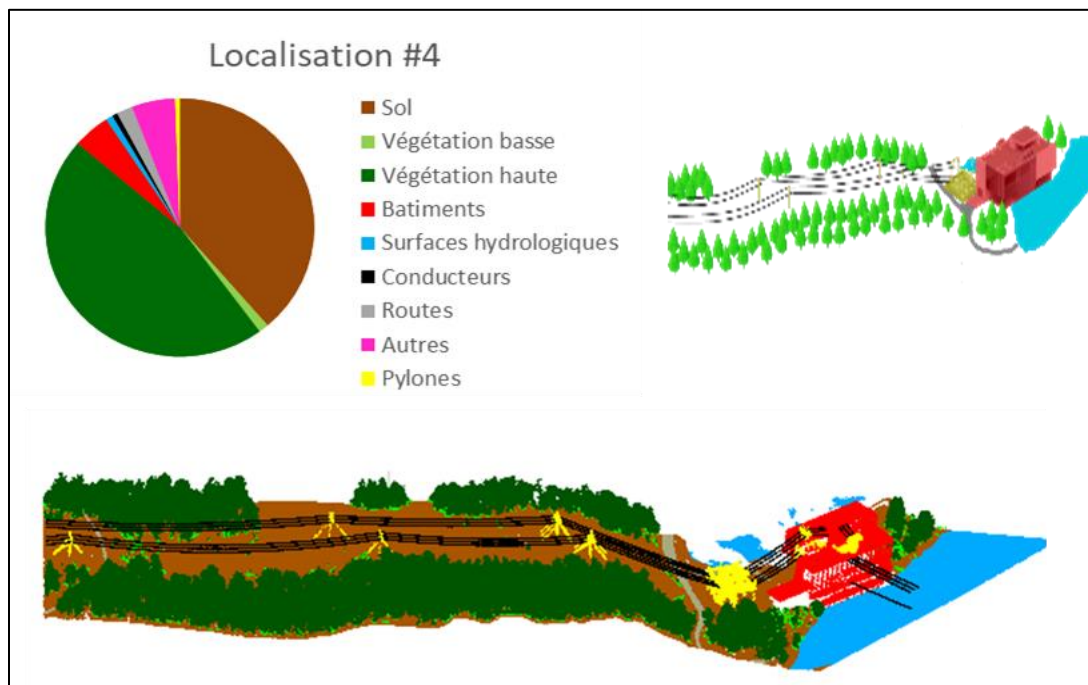


Figure 4.6 Représentation et composition des nuages de points des Localisation #4



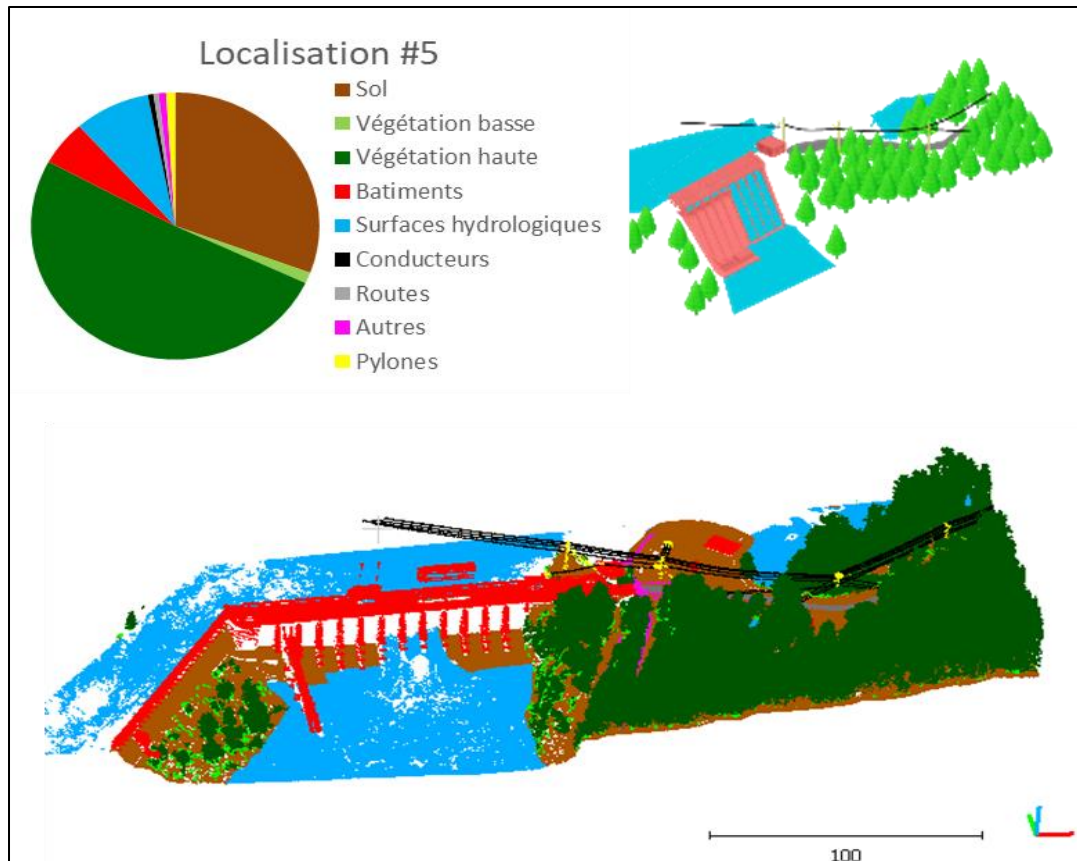


Figure 4.7 Représentation et composition des nuages de points des Localisation #5

#### 4.2.2.2 Localisation #6

Le nuage de points situé à la Localisation #6 est visible à la Figure 4.8. Il provient d'un ALS. Les classes représentées sont semblables aux Localisations #4 et #5. Leurs répartitions dans le nuage sont indiquées au Tableau 4.1. Ce nuage représente une zone rurale. Toutefois, cette dernière est composée de plusieurs petits bâtiments qui sont des habitations. Pour ce nuage, afin d'avoir une bonne représentativité de chaque classe, nous utilisons une partie pour l'apprentissage machine, une partie pour le test et une troisième partie pour la validation. Chacune de ces parties englobe une diversité suffisante des classes et des éléments.

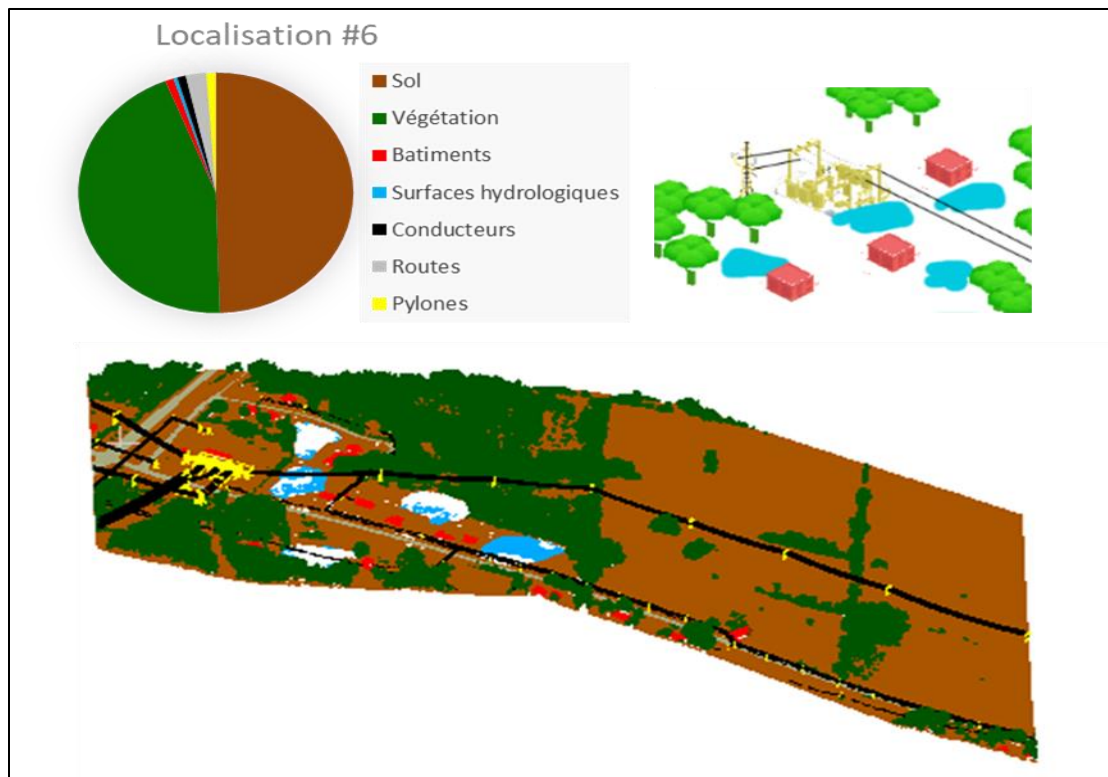


Figure 4.8 Représentation et composition du nuage de points de la Localisation #6

### 4.2.3 Données multisources

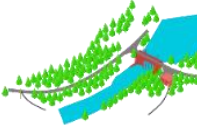
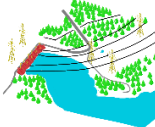
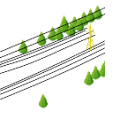
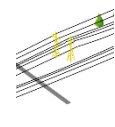
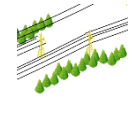
Le cœur de notre problématique concerne les données multisources. Dans les sections suivantes, les différents jeux de données multisources sont présentés.

#### 4.2.3.1 Localisation #1

La Localisation #1 est représentée dans la proposition visible Figure 4.9 et reprise sous forme stylisée dans le Tableau 4.2. Les technologies de captation utilisées pour la Localisation #1 sont au nombre de trois (3). On y retrouve un nuage provenant d'un ALS, un deuxième provenant d'un TLS et enfin un troisième nuage produit à partir de photographies prises par drone. Le relevé ALS provient, pour cette localisation, du Ministère des Forêts, de la Faune et de Parcs (Gouvernement du Québec). Il n'a donc pas les mêmes densités ou segmentation sémantique habituellement demandées par le PI. Le nuage représente une zone rurale qui

comporte un déversoir, une route, des forêts ainsi qu'un petit bâtiment. Les différentes classes ainsi que le nombre de points par classe sont indiqués dans le Tableau 4.2. Ce dernier nous permet aussi de présenter les attributs acquis qui sont utilisés lors de la segmentation sémantique. Le nuage étant presque symétrique au niveau de la représentation des classes, nous avons retenu de réaliser le découpage apprentissage/test en suivant simplement le cours de l'eau.

Tableau 4.2 Représentation schématique et composition des nuages de points multisources

	Localisation #1	Localisation #2	Localisation #3.1	Localisation #3.2	Localisation #3.3	
						
<b>Technologies</b>	ALS, TLS et UAV (photogrammétrie)	ALS et TLS	ALS, HLS et TLS			
Attributs acquis	RGB, Intensité, Nb de retour, No du retour	RGB, Intensité, Nb de retour, No du retour	Intensité (ALS, HLS, TLS), Nb de retour (ALS, HLS), No du retour (ALS, HLS)			
<b>Classes</b>	Sol	9,219,176 (36%)	16,512,741 (31%)	60,653,208 (82%)	78,727,491 (85%)	55,254,561 (81%)
	Végétation	13,102,686 (52%)	16,088,473 (30%)	3,370,985 (5%)	4,945,552 (5%)	3,359,675 (5%)
	Bâtiments	764,485 (3%)	11,386,493 (21%)	1,232,888 (2%)	2,535,929 (3%)	4,054,964 (6%)
	Surfaces hydrologiques	1,050,067 (4%)	4,341,752 (8%)	68,113 (0.1%)	0 (0%)	128,802 (0.2%)
	Conducteurs	1,561 (0%)	335,759 (0.6%)	615,008 (0.8%)	956,192 (1%)	750,014 (1%)
	Routes	1,244,555 (5%)	1,706,033 (3%)	6,856,725 (9%)	4,723,342 (5%)	4,695,847 (7%)
	Pylônes	25,564 (0.1%)	2,588,986 (5%)	981,236 (1%)	843,458 (1%)	339,631 (0.5%)
	Total	25,408,094 (100%)	52,960,237 (100%)	73,778,163 (100%)	92,861,022 (100%)	68,583,494 (100%)

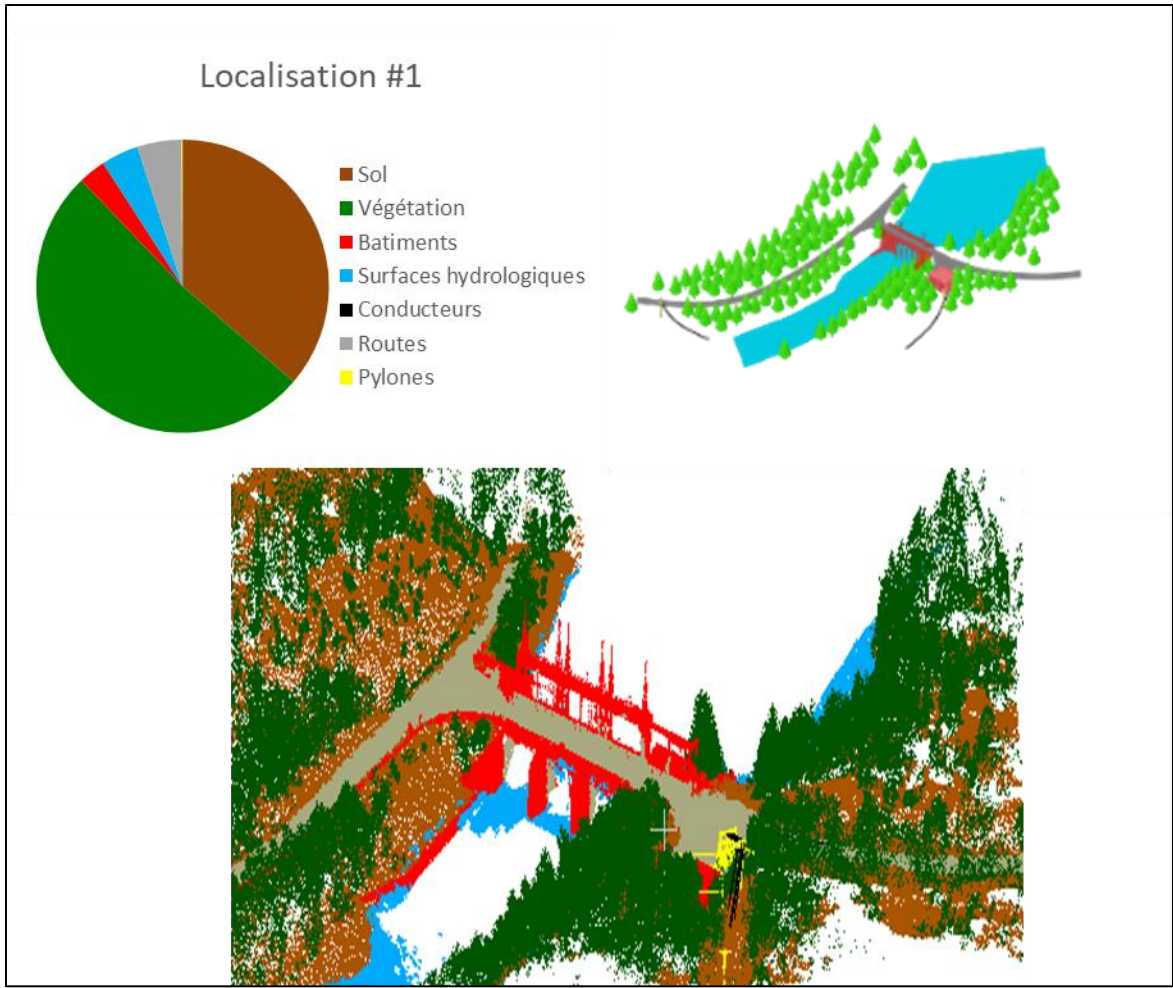


Figure 4.9 Représentation et composition du nuage de points de la Localisation #1

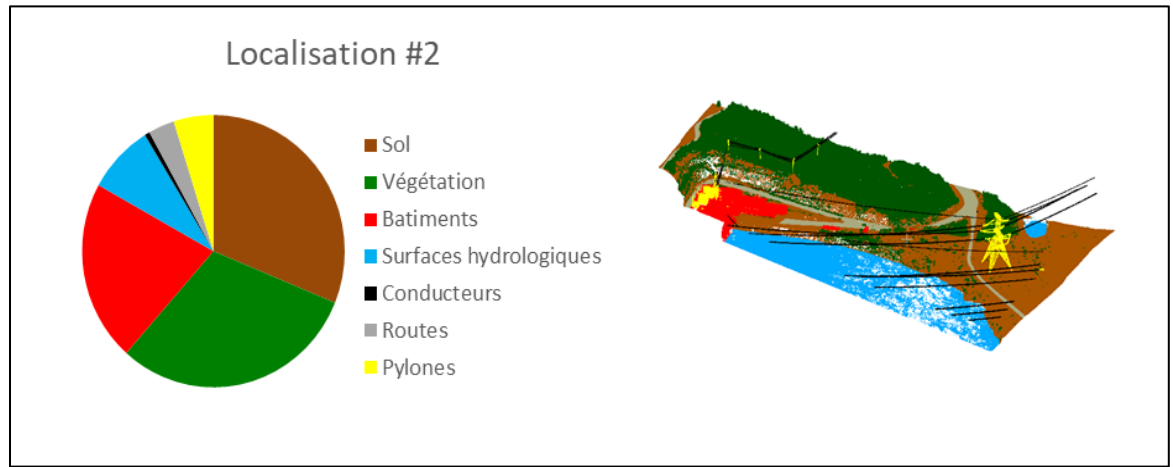


Figure 4.10 Représentation et composition du nuage de points de la Localisation #2

#### 4.2.3.2 Localisation #2

Le nuage de points de la Localisation #2 est visible à la Figure 4.10. Les technologies de captation utilisées pour cette localisation sont un ALS et un TLS. Le nuage du TLS contient à l'origine l'intérieur des bâtiments. Ces parties ne sont pas le point d'intérêt, par conséquent, elles sont retirées de notre étude. Ce nuage correspond à un bâtiment et son environnement qui est constitué de routes, chemins, végétation et d'autres bâtiments plus petits.

#### 4.2.3.3 Localisation #3

La Localisation #3 est un ensemble de trois nuages de points notés Localisation #3.1, #3.2 et #3.3 qui se trouvent à des zones géographiques proches (mêmes sources de captation). Ces nuages sont visibles à la Figure 4.11 et Figure 4.12 ou encore dans le Tableau 4.2. Les technologies de captations sont au nombre de trois (3) : un ALS, un TLS ainsi qu'un HLS. Les zones sont urbaines, elles comportent principalement des routes, de la végétation et des bâtiments.

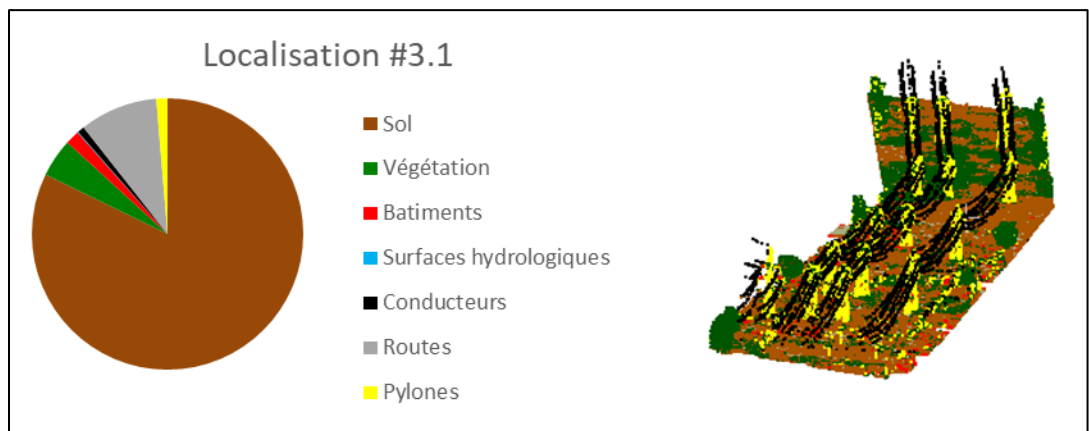


Figure 4.11 Représentation et composition des nuages de points de la Localisation #3.1

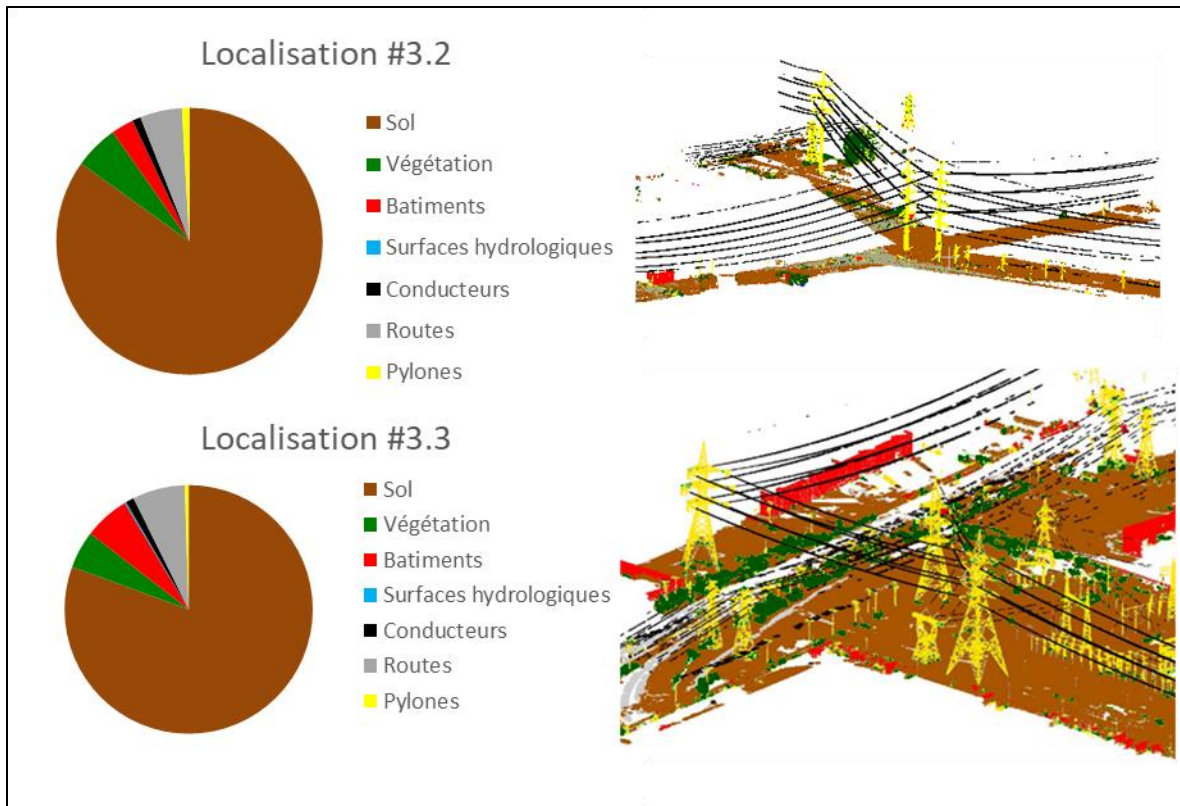


Figure 4.12 Représentation et composition des nuages de points des Localisation #3.2 et Localisation #3.3

### 4.3 Résultats de l'étude sur l'apprentissage monosource

Cette partie concerne la segmentation sémantique sur un nuage monosource. Dans notre cas, nous utilisons les nuages de points qui proviennent des ALS sur les Localisations #4, #5 et #6. Cette première partie porte sur l'utilisation de réseaux de neurones dans la segmentation sémantique en utilisant des attributs principalement calculés à l'aide de sphères et de cylindres. Une image de chaque localisation (après segmentation sémantique) est présentée.

#### 4.3.1 Segmentation sémantique monosource

Cette partie des résultats est consacrée principalement au nuage de la Localisation #4. La matrice de confusion obtenue est présentée dans le Tableau 4.3. Pour toutes les matrices de confusion, la carte de couleurs est la suivante : la diagonale principale est colorée en **rouge**

pour le minimum (qui correspond à 0%) et en **vert** pour le maximum (qui correspond à 100%, valeur cible). Le reste de la carte de couleurs des matrices va du **blanc** pour le minimum (0%, valeur cible) au **rouge** pour le maximum (100%). Cette carte de couleur permet de mettre l'accent sur les bons points (en **vert**) et les échecs (en **rouge**). L'utilisation du gradient de couleur du blanc au rouge pour le reste de la matrice permet de souligner les performances.

Cette matrice de confusion est le résultat de la classification de la Localisation #4 avec un apprentissage sur la Localisation #5 (avec la méthode des classes uniformément distribuées) et la Localisation #4 comme ensemble de tests. On y constate que pour certaines classes (par exemple, « Végétation haute » et « Végétation basse »), la méthode affiche une performance moindre. Une grande partie de la végétation basse est confondue avec la végétation haute. Le même phénomène se produit avec la classe « Autre ».

Tableau 4.3 Matrice de confusion de la Localisation #4 avec l'apprentissage réalisé sur la Localisation #5 avec un jeu de données ayant des classes avec un nombre de points uniforme

Localisation #4	Sol	Végétation basse	Végétation haute	Bâtiments	Surfaces hydrologiques	Conducteurs	Routes	Autres	Pylônes
Sol	0.81	0.00	0.00	0.02	0.01	0.00	0.09	0.08	0.00
Végétation basse	0.00	0.69	0.20	0.00	0.00	0.00	0.00	0.06	0.05
Végétation haute	0.00	0.00	0.92	0.01	0.00	0.00	0.00	0.02	0.02
Bâtiments	0.00	0.00	0.00	0.85	0.00	0.00	0.00	0.11	0.03
Surfaces hydrologiques	0.03	0.00	0.00	0.02	0.89	0.00	0.03	0.03	0.00
Conducteurs	0.00	0.00	0.00	0.00	0.00	0.99	0.00	0.00	0.00
Routes	0.06	0.00	0.00	0.04	0.01	0.00	0.88	0.00	0.00
Autres	0.09	0.02	0.04	0.16	0.00	0.00	0.00	0.68	0.03
Pylônes	0.00	0.00	0.02	0.05	0.00	0.01	0.00	0.02	0.89

Le réseau de neurones est entraîné pour tenir en compte de la représentation de chaque classe en utilisant une uniformisation des classes. Pour ce faire, un nouvel ensemble d'entraînement est construit, incluant 70% des points de chaque classe. Les résultats sont présentés dans le Tableau 4.4. La première colonne est le résultat de l'apprentissage sur la Localisation #5 (avec des classes uniformément distribuées) et qui est testé sur la Localisation #4. La deuxième colonne est le résultat de l'apprentissage sur la Localisation #5 (avec des classes non uniformément distribuées) et qui est testé sur la Localisation #4. La troisième colonne correspond à la même méthodologie (en utilisant des classes non uniformément distribuées), mais en effectuant l'apprentissage sur la Localisation #4 et en testant sur la Localisation #5. Enfin, la quatrième colonne est identique à la deuxième colonne, mais en ne prenant en compte qu'une seule classe pour la végétation. Les valeurs présentent en **gras**, dans le Tableau 4.4, correspondent aux meilleurs résultats pour la classe en question.

Tableau 4.4 Ensemble des résultats pour la segmentation sémantique des Localisation #4 et #5 avec jeu d'apprentissage uniformisé et réel

Appris sur	Localisation #5 uniforme	Localisation #5 non- uniforme	Localisation #4 non- uniforme	Localisation #5 avec Végétation groupée
Testé sur	Localisation #4	Localisation #4	Localisation #5	Localisation #4
Sol	0.79	<b>0.91</b>	0.84	0.91
Végétation basse	<b>0.38</b>	0.31	0.40	-
Végétation haute	0.92	<b>0.97</b>	0.96	<b>0.98</b>
Bâtiments	0.58	<b>0.60</b>	0.53	0.57
Surfaces hydrologiques	<b>0.61</b>	0.56	0.45	0.60
Conducteurs	<b>0.99</b>	0.99	0.92	0.95
Routes	0.31	<b>0.36</b>	0.17	0.34
Autres	<b>0.36</b>	0.28	0.16	0.28
Pylônes	0.28	<b>0.45</b>	0.23	0.35
<b>Moyenne IoU</b>	0.58	0.60	0.52	<b>0.62</b>
<b>Précision globale</b>	0.86	<b>0.92</b>	0.90	0.92

On peut remarquer que les classes qui sont géométriquement similaires donnent de moins bons résultats. C'est le cas pour les classes telles que les « Routes » qui sont similaires au « Sol ». Cela peut s'expliquer par la similarité des attributs utilisés pour décrire les deux classes (« Routes » et « Sol »). Le même problème se retrouve, dans une moindre mesure, pour les



« Surfaces hydrologiques » qui, comme le « Sol » et les « Routes », ressemblent à une surface plane à différentes échelles et sont visibles Figure 4.13. Les points étiquetés « Autres » représentent un ensemble de classes et donnent également de moins bons résultats à cause de sa grande diversité. Le réseau de neurones classe mal certains objets de cette classe, tels que les lampadaires dans la catégorie « Végétation haute » ou des garde-corps de sécurité routière dans la catégorie « Bâtiments » (les « Bâtiments » ayant eux aussi des barrières et rambardes). La confusion entre rambardes des « Bâtiments » et rambardes pour la sécurité routière est cependant intéressante. Elle met en avant l'utilité d'avoir des classes spécifiques plutôt que des classes qui englobent des objets trop divers.



Figure 4.13 Localisation #6 après segmentation sémantique

La technique de classification proposée fonctionne particulièrement bien pour les classes ayant une forme typique, comme les « Conducteurs ».

On remarque également que l'uniformisation des classes permet une meilleure classification des classes sous-représentées. Cependant, cette classification demeure imparfaite en raison du manque de diversité dans chaque nuage. Le Tableau 4.4 montre qu'avec ou sans uniformisation des classes, des classes telles que « Végétation haute » et « Conducteurs » sont bien classées.

En effet, les classes géométriquement simples telles que les « Conducteurs » présentent peu de diversités, dans la plupart des cas, il s'agit d'un câble (1D) et dans de rares cas de plusieurs câbles (une forme 2D). Grâce à cette faible diversité, il est possible de les caractériser simplement avec peu de données d'entraînement. D'autres classes ne sont pas aussi simples à caractériser en raison de la diversité de représentation qui les caractérise. Dans ce cas, la diversité souhaitée ne peut pas être obtenue avec une augmentation artificielle des données comme ce qui a été fait ici. Il est néanmoins possible d'élargir cette diversité en ajoutant des exemples de points provenant d'autres nuages.

Comme mentionné, la « Végétation basse » est souvent confondue avec la « Végétation haute » (Tableau 4.3 et Tableau 4.5). Ce manque de performance peut avoir plusieurs origines. Une piste d'explication se trouve dans le fait que la « Végétation basse » est souvent spatialement (géographiquement) proche de la « Végétation haute ». Cette proximité, couplée à la faible densité de points, signifie que lors de la sélection dans le voisinage, une grande partie des points présents dans les sphères sont des points de haute végétation. La « Végétation basse » est par conséquent décrite avec des points de « Végétation haute ». L'ajout de caractéristiques liées aux classes voisines peut s'avérer intéressant pour contrer ce problème. Cette idée sera approfondie dans la section 4.4.4. L'ajout de telles caractéristiques peut également être intéressant pour des classes telles que les « Pylônes » qui sont géométriquement similaires à la « Végétation haute ». Cependant, les « Pylônes » sont typiquement voisins des classes « Bâtiments » et « Conducteurs » dans ces localisations.

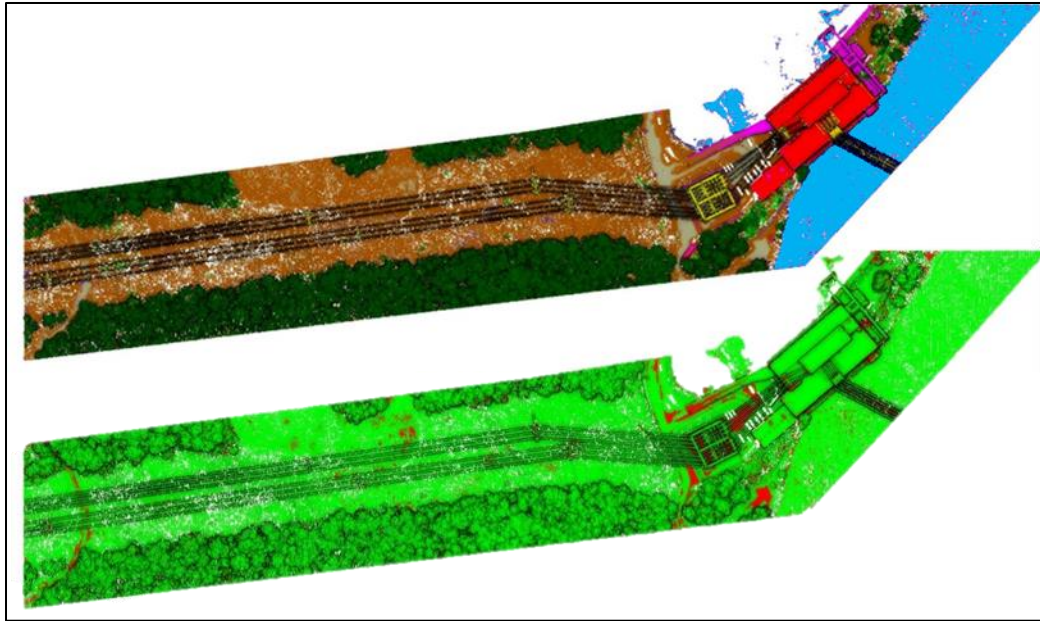


Figure 4.14 Localisation #4 après segmentation sémantique et comparaison entre le modèle de référence et la segmentation sémantique

Tableau 4.5 Matrice de confusion de la Localisation #4 avec l'apprentissage réalisé sur la Localisation #5 avec un jeu de données ayant des classes avec un nombre de points non uniforme

Localisation #4	Sol	Végétation basse	Végétation haute	Bâtiments	Surfaces hydrologiques	Conducteurs	Routes	Autres	Pylônes
Sol	0.95	0.00	0.00	0.01	0.01	0.00	0.01	0.01	0.00
Végétation basse	0.02	0.32	0.64	0.00	0.00	0.00	0.00	0.02	0.00
Végétation haute	0.00	0.00	0.99	0.00	0.00	0.00	0.00	0.00	0.00
Bâtiments	0.02	0.00	0.02	0.90	0.00	0.00	0.00	0.04	0.02
Surfaces hydrologiques	0.07	0.00	0.00	0.02	0.90	0.00	0.00	0.01	0.00
Conducteurs	0.00	0.00	0.00	0.01	0.00	0.97	0.00	0.00	0.01
Routes	0.50	0.00	0.00	0.00	0.07	0.00	0.43	0.00	0.00
Autres	0.15	0.01	0.13	0.37	0.00	0.00	0.00	0.31	0.03
Pylônes	0.00	0.00	0.25	0.08	0.00	0.01	0.00	0.02	0.65

#### 4.3.1.1 Réduction du nombre de classes

En fonction de la résolution nécessaire, une réduction du nombre de classes peut être accomplie. En effet, lorsque la distinction entre végétations basse et haute n'a pas d'importance ou d'intérêt, le résultat est sensiblement amélioré en ne considérant qu'une seule classe de végétation (Tableau 4.4, colonne 4). Le même principe s'applique avec les classes « Autres » (qui représentent la plupart du temps des constructions humaines) qui peuvent être rattachées à la classe « Bâtiment ». Enfin, dans certains projets, la distinction entre la route et le sol n'est pas essentielle. C'est principalement le cas dans les zones plus recluses (ou isolées) où les « Routes » sont moins praticables et où les équipements sont transportés en dehors des routes. Dans ces rares cas, il est possible de réduire les classes présentes pour obtenir de meilleurs résultats. Le Tableau 4.6 montre la matrice de confusion pour la Localisation #4 ayant appris sur la Localisation #5.

Tableau 4.6 Matrice de confusion de la Localisation #4 avec les classes de « Végétation » regroupées et la classe de « Route » regroupée avec la classe « Sol »

Localisation #4	Sol/Routes	Végétation	Bâtiments	Surfaces hydrologiques	Conducteurs	Pylônes
Sol/Routes	0.92	0.00	0.05	0.03	0.00	0.00
Végétation	0.00	0.99	0.00	0.00	0.00	0.00
Bâtiments	0.01	0.01	0.97	0.00	0.00	0.01
Surfaces hydrologiques	0.15	0.00	0.03	0.82	0.00	0.00
Conducteurs	0.00	0.00	0.01	0.00	0.97	0.01
Pylônes	0.00	0.37	0.10	0.00	0.01	0.52

#### 4.3.1.2 Post-traitement

Une fois la segmentation sémantique accomplie, il est possible d'améliorer encore les résultats obtenus en appliquant un post-traitement comme un lissage des classes par exemple. Dans notre cas, un point entouré par des points d'une même classe possède une grande probabilité d'appartenir à cette classe. La classe majoritairement représentée dans le voisinage peut donc être considérée comme la classe du point central de ce voisinage. Cette première approximation permet de soustraire le bruit présent dans certaines régions. Dans la matrice présentée dans le Tableau 4.7, un rayon de 1 m est utilisé. En utilisant un rayon plus grand, par exemple 3 m, les classes peu représentées disparaissent au profit des classes dominantes ; les résultats n'ont donc pas été retenus. C'est le cas de la classe « Végétation basse » qui passe de 32% dans le Tableau 4.5 à 28% dans le Tableau 4.7. Les points de la classe « Végétation basse » qui sont classés en « Végétation haute » passe de 64% à 68%.

Tableau 4.7 Matrice de confusion de la Localisation #4 après l'utilisation d'un post-traitement de lissage des classes

Localisation #4	Sol	Végétation basse	Végétation haute	Bâtiments	Surfaces hydrologiques	Conducteurs	Routes	Autres	Pylônes
Sol	0.96	0.00	0.00	0.02	0.00	0.00	0.00	0.01	0.00
Végétation basse	0.02	0.28	0.68	0.00	0.00	0.00	0.00	0.02	0.00
Végétation haute	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00
Bâtiments	0.02	0.00	0.02	0.90	0.00	0.00	0.00	0.04	0.02
Surfaces hydrologiques	0.06	0.00	0.00	0.02	0.92	0.00	0.00	0.01	0.00
Conducteurs	0.00	0.00	0.00	0.01	0.00	0.98	0.00	0.00	0.01
Routes	0.50	0.00	0.00	0.00	0.08	0.00	0.40	0.00	0.00
Autres	0.15	0.02	0.14	0.37	0.00	0.00	0.00	0.25	0.05
Pylônes	0.00	0.00	0.40	0.08	0.00	0.01	0.00	0.02	0.49

### 4.3.2 Synthèse sur l'apprentissage monosource

Cette première approche, monosource, a permis de saisir le fonctionnement de certains algorithmes d'apprentissage machine pour la segmentation sémantique. Aussi, elle a permis de produire les premiers résultats jugés comme satisfaisant.

Cette première proposition permet de souligner les premières complications issues de l'exploitation des données du PI avec notre méthode. En utilisant des sphères pour calculer des attributs, les classes qui sont proches les unes des autres et qui se retrouvent *de facto* dans les mêmes sphères peuvent être confondues (ex. confusion entre la « Végétation basse » et la « Végétation haute »).

Toujours par rapport aux choix *a priori* des classes, la classe « Autres » n'est pas une bonne solution. Cette classe a tendance à englober des objets trop divers (ex. des lampadaires, des voitures, des garde-corps, etc.). Cette diversité, couplée à notre proposition qui utilise des attributs en partie géométriques, contribue à brouter les différences entre les classes ce qui réduit par la suite la performance de la segmentation sémantique.

Nous constatons aussi que certains attributs ont un poids plus important dans la décision. Dans la Figure 4.15, des points de « Surfaces hydrologiques » sont classés en « Sol » et « Routes » à cause de l'attribut *Numéro de retour*. Les points de « Surfaces hydrologiques » ayant dans une grande partie des cas un *Numéro de retour* égal à 1, le réseau de neurones va privilégier de classer les points dans des classes proches géométriquement, mais qui ont plus de points avec des *Numéro de retour* >1 tels que le « Sol » et la « Routes ». On remarque donc des points marron « Sol » et gris « Routes » au milieu de la « Surfaces hydrologiques » bleu.

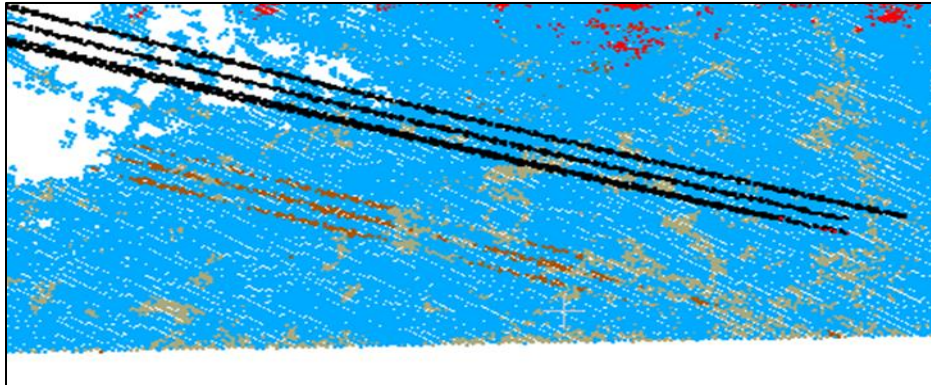


Figure 4.15 Zoom sur une mauvaise classification de la Localisation #5 à cause du nombre de retour

Un autre problème soulevé porte sur les bordures des nuages de points. En effet, le voisinage des points lors du calcul d'attribut est tronqué par rapport à la réalité. Dans la réalité, il n'existe pas de bordure, l'espace est continu. La sphère pour un point de la bordure ne sera donc pas entièrement remplie comme elle le serait dans un espace continu. Ce problème est cependant localisé à la fois dans l'espace, mais aussi dans les classes qu'il touche, ainsi que dans leurs mauvaises interprétations. Ce problème touche par exemple les points de la classe de « Végétation haute » qui sont principalement des grandes structures verticales en 3D. La troncature peut alors transformer ces grandes structures verticales en trois dimensions, en grandes structures verticales à deux dimensions qui ressemblent plus à certains « Pylônes ».

On retrouve aussi dans la Figure 4.16 les points du « Sol » qui sont confondus principalement avec des points de « Surfaces hydrologiques ». Cette confusion découle du fait que les zones de « Surfaces hydrologiques » sont semblables géométriquement aux points du « Sol ». En plus, l'attribut de planéité, calculé à l'aide d'une sphère de rayon 3 m visible à la Figure 4.16, est plus faible pour les points appartenant aux bordures. On observe que des points des bordures ont des valeurs de planéité entre 0.3 et 0.6 alors que les points n'étant pas sur les bordures ont des valeurs de planéité qui avoisinent 0.9. Considérant que les points de bordures correspondent à moins de 10% (visible au travers des histogrammes de la Figure 4.16) pour les points de la classe « Sol » alors qu'ils correspondent à environ 20% des points des « Surfaces hydrologiques », le réseau de neurones a donc plus de chances de prédire que les points des bordures du « Sol » sont plutôt des points de « Surfaces hydrologiques ».

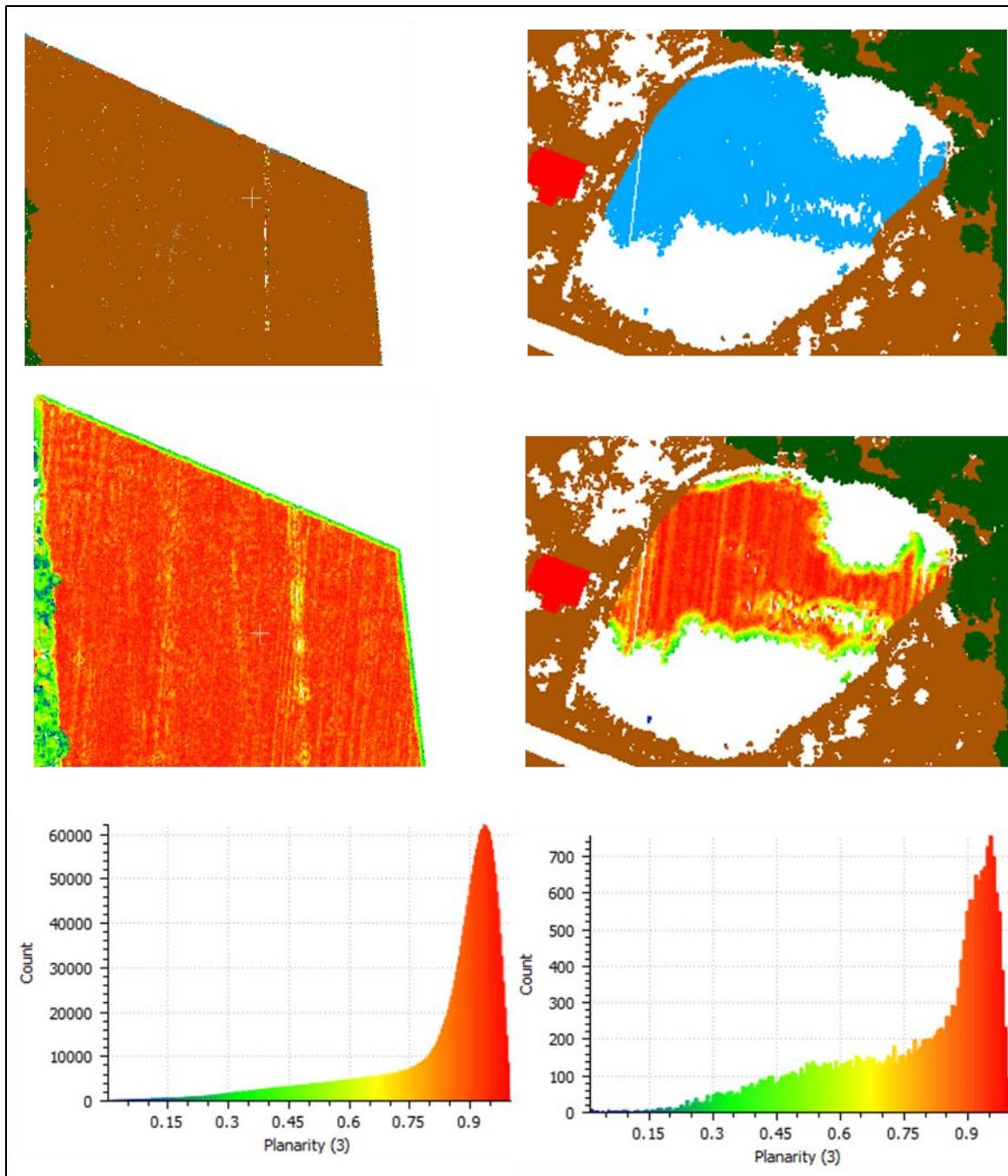


Figure 4.16 Zoom sur l'effet de bordure. La colonne de gauche correspond à la classe « Sol », la colonne de droite à la classe « Surfaces hydrologiques ». L'attribut de planéité calculé avec une sphère de rayon 3 m a été pris en exemple pour les classes de « Sol » et « Surfaces hydrologiques » de la Localisation #6



Finalement, la représentation hétérogène des classes pour l'apprentissage conduit vers l'utilisation automatique de poids dans le réseau de neurones dans le but d'améliorer la représentation des classes qui comportent peu de points.

#### **4.4 Résultats de l'étude d'ablation sur la proposition multisource**

Pour évaluer l'importance (et la valeur ajoutée) des différentes parties de notre proposition, une étude d'ablation est réalisée. Une étude d'ablation permet d'étudier une solution en enlevant ou en ajoutant les parties qui composent la solution afin de comprendre au mieux les contributions de chaque partie. Les performances de segmentation sémantique sont d'abord testées en ne considérant que les attributs monosources et les attributs propagés qui correspondent à l'utilisation des branches A1 et A3 des parties A et B de la Figure 4.17 (section 4.4.1). Ensuite, les caractéristiques monosources et multisources ainsi que les caractéristiques propagées sont considérées (partie A et partie B, section 4.4.2). Le résultat obtenu en utilisant le deuxième réseau de neurones avec les classes prédites (correspondant à la partie A, la partie B et la partie D) est décrit dans la section 4.4.3. La section 4.4.4 correspond à l'utilisation des parties A, B, C et D.

#### 4.4.1 Segmentation sémantique monosource (chemin A1+A3, B)

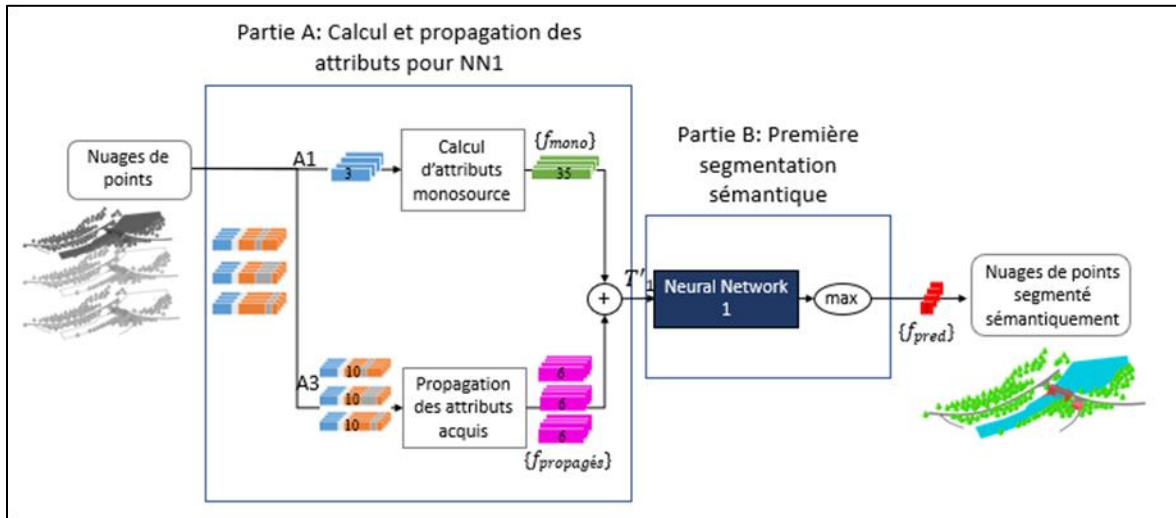


Figure 4.17 *Workflow* pour la première étape de l'étude d'ablation sur la segmentation sémantique monosource

Cette section présente les résultats et l'interprétation de la segmentation sémantique en utilisant uniquement les caractéristiques monosources et les caractéristiques acquises par chaque nuage (correspondant aux branches A1 et A3 de la Figure 4.17). Les Tableaux 4.8 et 4.9 présentent les matrices de confusion des deux premières localisations.

La matrice de confusion du Tableau 4.8 montre que les classes avec peu d'exemples (comme les « Conducteurs » et les « Pylônes » dans la Localisation #1) sont bien classées. Malgré le peu d'exemples, leurs caractéristiques géométriques étant très particulières, il est possible pour le réseau de neurones de bien classer les points appartenant à ces classes. Les « Conducteurs » sont souvent vus comme des nuages de points 1D dans les sphères de voisinage. Cependant, la proximité de ces classes dans le nuage de points est en partie responsable du résultat de 11% des points appartenant à la classe « Pylônes » perçus comme des points de « Conducteurs ». La confusion existe aussi dans l'autre sens avec 6% des points de la classe « Conducteurs » qui sont perçus comme des points de la classe « Pylônes ».

La classe « Surfaces hydrologiques » quant à elle souffre principalement de ses similarités géométriques avec le « Sol ». On retrouve donc des points « Surfaces hydrologiques » à la place de certains points de « Sol » dans la Figure 4.18.

Tableau 4.8 Matrice de confusion de la Localisation #1 après la segmentation sémantique monosource

Localisation #1	Sol	Végétation	Bâtiments	Surfaces hydrologiques	Conducteurs	Routes	Pylônes
Sol	0.94	0.02	0.01	0.00	0.00	0.03	0.01
Végétation	0.06	0.88	0.01	0.02	0.00	0.00	0.04
Bâtiments	0.00	0.05	0.88	0.02	0.00	0.04	0.01
Surfaces hydrologiques	0.07	0.05	0.03	0.80	0.00	0.01	0.03
Conducteurs	0.00	0.00	0.00	0.00	0.94	0.00	0.06
Routes	0.01	0.00	0.00	0.01	0.00	0.97	0.01
Pylônes	0.01	0.03	0.00	0.00	0.11	0.00	0.85

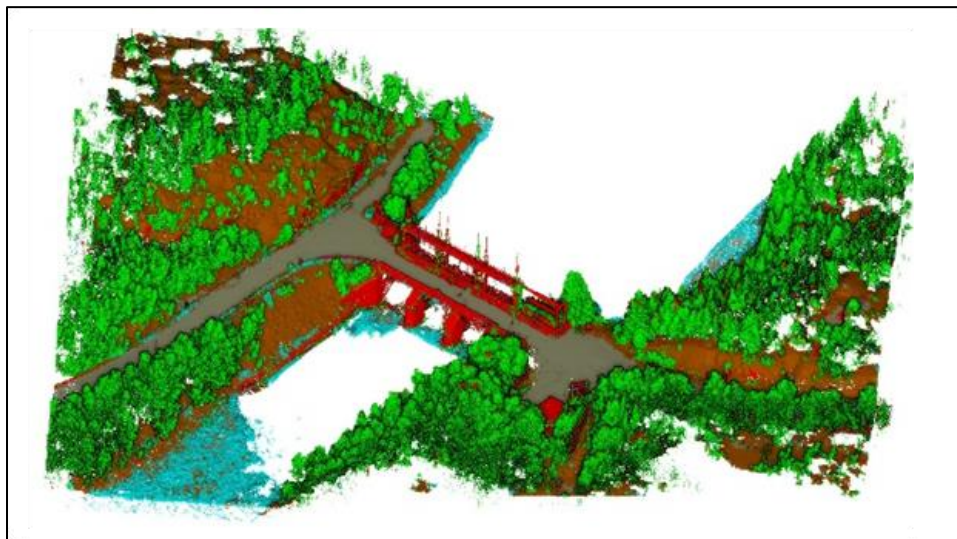


Figure 4.18 Localisation #1 après segmentation sémantique monosource

Tableau 4.9 Matrice de confusion de la Localisation #2  
après la segmentation sémantique monosource

Localisation #2	Sol	Végétation	Bâtiments	Surfaces hydrologiques	Conducteurs	Routes	Pylônes
Sol	0.89	0.00	0.02	0.03	0.00	0.05	0.01
Végétation	0.02	0.90	0.00	0.01	0.01	0.00	0.05
Bâtiments	0.06	0.00	0.88	0.02	0.00	0.03	0.01
Surfaces hydrologiques	0.33	0.04	0.14	0.47	0.00	0.01	0.01
Conducteurs	0.00	0.02	0.00	0.00	0.93	0.00	0.05
Routes	0.45	0.01	0.10	0.02	0.00	0.40	0.00
Pylônes	0.03	0.03	0.00	0.00	0.07	0.00	0.87

Le Tableau 4.9 montre aussi la confusion entre les classes « Surfaces hydrologiques » « Routes » et « Sol », qui possèdent des caractéristiques géométriques similaires. Toutes les trois sont des classes relativement planes, ayant des nombres et un numéro de retour similaire. Ainsi 33% des points des « Surfaces hydrologiques » sont confondus avec du « Sol » et 14% avec des points de « Bâtiments ».

Afin d'essayer d'améliorer les résultats nous avons donc décidé d'ajouter l'aspect multisource dont les résultats sont présentés section suivante.

#### 4.4.2 Segmentation sémantique multisource (chemin A1+A2+A3, B)

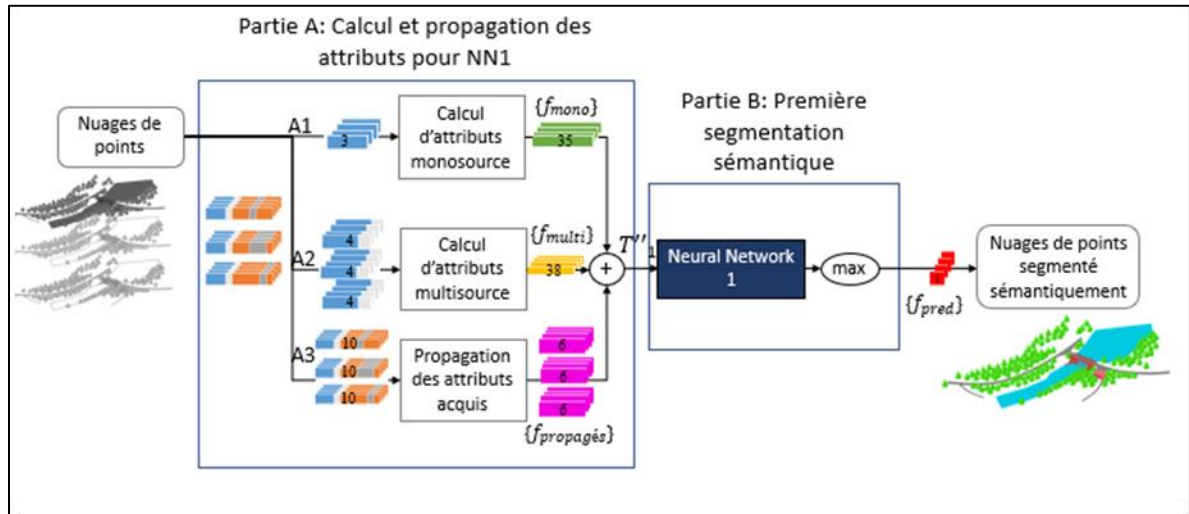


Figure 4.19 *Workflow* pour la deuxième étape de l'étude d'ablation sur la segmentation sémantique multisource

Dans cette partie, les branches A1, A2 et A3 de la Figure 4.19 sont utilisées. Avec l'ajout d'éléments multisources, certains problèmes liés à la représentation de certaines classes apparaissent encore dans les Tableaux 4.10 et 4.11.

Certaines classes, dont les « Routes » ou les « Surfaces hydrologiques » dans le Tableau 4.11 et les « Pylônes » ou les « Surfaces hydrologiques » dans le Tableau 4.10, ont des valeurs plus faibles que dans les Tableaux 4.8 et 4.9.

C'est le cas des classes qui ne sont pas suffisamment représentées dans les données, comme les « Pylônes » de la Localisation #1 ou les « Surfaces hydrologiques » de la Localisation #3 (voir Tableaux 4.16 à 4.20).

Tableau 4.10 Matrice de confusion de la Localisation #1  
après la segmentation sémantique multisource

Localisation #1	Sol	Végétation	Bâtiments	Surfaces hydrologiques	Conducteurs	Routes	Pylônes
Sol	0.91	0.03	0.02	0.01	0.00	0.03	0.00
Végétation	0.04	0.94	0.01	0.00	0.00	0.00	0.01
Bâtiments	0.04	0.03	0.91	0.00	0.00	0.00	0.01
Surfaces hydrologiques	0.04	0.06	0.04	0.84	0.00	0.02	0.00
Conducteurs	0.00	0.00	0.00	0.00	0.99	0.00	0.00
Routes	0.02	0.00	0.02	0.01	0.00	0.95	0.00
Pylônes	0.02	0.09	0.00	0.00	0.11	0.00	0.79

En comparaison avec l'approche monosource, certaines classes telles que la « Végétation », les « Bâtiments », les « Surfaces hydrologiques » ou les « Routes » vont avoir de meilleurs résultats avec leurs scores dans la matrice de confusion qui augmentent entre 3 et 6 points. Le résultat de cette segmentation sémantique est visible dans la Figure 4.20. Les différences entre les résultats sont difficilement discernables, cependant, il est possible de remarquer que certains points de « Sol » se sont retrouvés classés en « Bâtiments ».



Figure 4.20 Localisation #1 après segmentation sémantique multisource

Tableau 4.11 Matrice de confusion de la Localisation #2 après la segmentation sémantique multisource

Localisation #2	Sol	Végétation	Bâtiments	Surfaces hydrologiques	Conducteurs	Routes	Pylônes
Sol	0.95	0.01	0.02	0.03	0.00	0.00	0.00
Végétation	0.02	0.93	0.00	0.02	0.00	0.01	0.03
Bâtiments	0.04	0.01	0.92	0.00	0.00	0.01	0.03
Surfaces hydrologiques	0.40	0.05	0.04	0.50	0.00	0.00	0.01
Conducteurs	0.00	0.00	0.00	0.00	0.98	0.00	0.02
Routes	0.43	0.02	0.15	0.07	0.00	0.30	0.00
Pylônes	0.03	0.01	0.04	0.00	0.02	0.00	0.91

Pour la Localisation #2, visible dans le Tableau 4.11, une amélioration des résultats est constatée par rapport au Tableau 4.8 et 4.9. Cependant, quelques classes qui ont déjà de mauvaises performances avec l'approche monosource (la classe « Routes ») possèdent de moins bons résultats avec l'approche multisource. Ce défaut (ou carence) peut être dû à la nature de cette classe dans la Localisation #2. En effet, on remarque qu'une plus grande proportion de la classe « Routes » est cette fois confondue avec la classe « Bâtiments ». Une partie de la classe « Bâtiments » possède sur son toit des « Routes » et des chemins de passage pour véhicules.

Pour résumer, à quelques exceptions, l'utilisation de toutes les caractéristiques (acquises, monosources et multisources) aide indéniablement à la segmentation sémantique des différents nuages de points. Mais, nous avons remarqué que si la première proposition de classe (avec le plus haut pourcentage) n'était pas la bonne, la seconde était souvent la bonne classe pour le point. Nous avons donc décidé de donner cette information à notre algorithme à l'aide d'une seconde segmentation sémantique avec le même réseau de neurones.

#### 4.4.3 Segmentation sémantique en cascade (chemin A1+A2+A3, B, D)

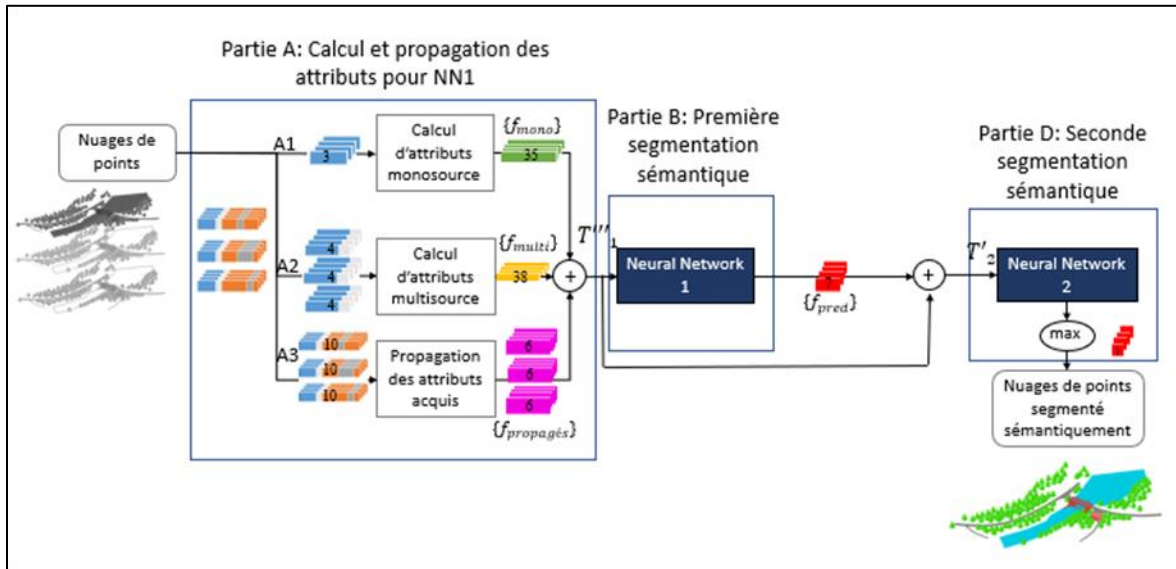


Figure 4.21 *Workflow* pour la troisième étape de l'étude d'ablation sur la segmentation sémantique en cascade

Cette section présente les matrices de confusion de la Localisation #1 et de la Localisation #2 après la deuxième classification (en utilisant les parties A, B et D visibles, Figure 4.21). Cette seconde segmentation qui utilise la prédiction de la première segmentation comme entrée peut donner de moins bonnes performances que la segmentation multisource, surtout avec les classes qui ne sont représentées que par quelques points.

En effet, on peut constater dans le Tableau 4.12 que les points des « Pylônes » se transforment en points de « Végétation » car cette dernière classe est proche des "Pylônes" et est plus représentée dans le nuage de points. Les points de la classe « Routes » sont classés en points de « Sol ». Une représentation de ses résultats est visible dans la Figure 4.22.



Tableau 4.12 Matrice de confusion de la Localisation #1 après la segmentation sémantique en cascade

Localisation #1	Sol	Végétation	Bâtiments	Surfaces hydrologiques	Conducteurs	Routes	Pylônes
Sol	0.90	0.02	0.00	0.07	0.00	0.01	0.00
Végétation	0.02	0.95	0.00	0.02	0.00	0.01	0.01
Bâtiments	0.06	0.00	0.93	0.00	0.00	0.01	0.00
Surfaces hydrologiques	0.02	0.01	0.00	0.93	0.00	0.04	0.00
Conducteurs	0.00	0.01	0.00	0.00	0.96	0.00	0.04
Routes	0.11	0.00	0.02	0.02	0.00	0.85	0.00
Pylônes	0.01	0.16	0.00	0.00	0.14	0.00	0.70



Figure 4.22 Localisation #1 après segmentation sémantique en cascade

Tableau 4.13 Matrice de confusion de la Localisation #2 après la segmentation sémantique en cascade

Localisation #2	Sol	Végétation	Bâtiments	Surfaces hydrologiques	Conducteurs	Routes	Pylônes
Sol	0.89	0.01	0.04	0.01	0.00	0.05	0.00
Végétation	0.02	0.91	0.02	0.00	0.00	0.00	0.05
Bâtiments	0.09	0.00	0.90	0.01	0.00	0.00	0.00
Surfaces hydrologiques	0.16	0.01	0.05	0.73	0.00	0.04	0.00
Conducteurs	0.00	0.00	0.01	0.00	0.95	0.00	0.04
Routes	0.04	0.00	0.17	0.08	0.00	0.71	0.00
Pylônes	0.01	0.05	0.02	0.00	0.09	0.00	0.83

On observe presque le même changement pour la Localisation #2 dans le Tableau 4.13, les « Routes » deviennent des points de « Bâtiment », car cette dernière classe est enfouie dans le « Sol » et une partie de la classe « Routes » se trouve sur la toiture des bâtiments.

Les observations concernant les classes avoisinantes vues dans la littérature et dans nos résultats nous ont donné l'idée d'ajouter des attributs liés aux points avoisinants avant notre deuxième segmentation sémantique. C'est pourquoi la prochaine section prend en compte ces nouveaux attributs et, ce faisant, l'ensemble de notre proposition.

#### 4.4.4 Segmentation sémantique avec prises en compte du voisinage (chemin A1+A2+A3, B, C, D)

Dans cette partie, l'ensemble de la proposition présentée dans la Figure 4.23 est utilisée pour exécuter une segmentation sémantique.

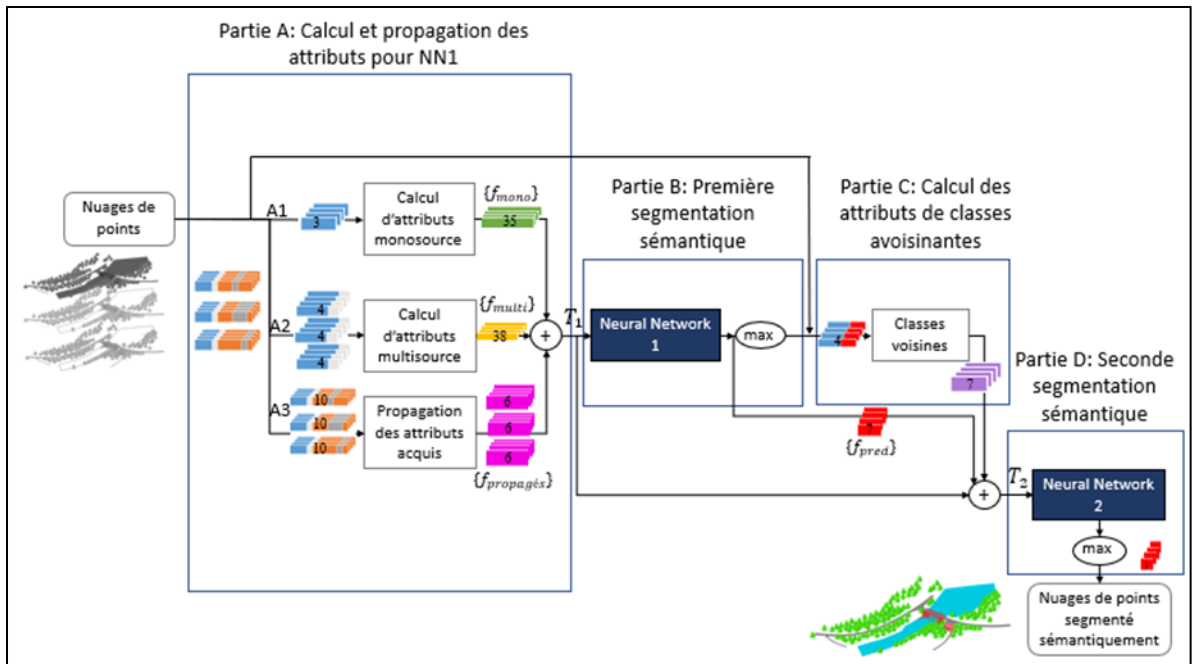


Figure 4.23 Ensemble du *Workflow* avec prise en compte du voisinage dans le scénario de segmentation en cascade

Comme on le constate avec la deuxième classification (Tableaux 4.12 et 4.13), la plupart des valeurs des Tableaux 4.14 et 4.15 ne changent pas. Les classes qui sont peu représentées (comme les « Pylônes » pour la Localisation #1) augmentent leur score passant de 70% (Tableau 4.12) à 89% (Tableau 4.14). Le plus grand changement apparaît dans les Tableaux 4.16 à 4.20, qui présentent les valeurs de la *MCC* pour chaque classe dans chaque scénario présenté dans cette étude d'ablation. La Localisation #1 une fois segmentée sémantiquement avec l'ensemble de notre proposition est visible Figure 4.24.



Figure 4.24 Localisation #1 après segmentation sémantique avec prise en compte du voisinage

Tableau 4.14 Matrice de confusion de la Localisation #1 après segmentation sémantique avec prise en compte du voisinage

Localisation #1	Sol	Végétation	Bâtiments	Surfaces hydrologiques	Conducteurs	Routes	Pylônes
Sol	0.92	0.01	0.00	0.05	0.00	0.01	0.00
Végétation	0.02	0.95	0.00	0.02	0.00	0.01	0.01
Bâtiments	0.07	0.00	0.93	0.00	0.00	0.00	0.00
Surfaces hydrologiques	0.02	0.01	0.00	0.91	0.00	0.06	0.00
Conducteurs	0.00	0.01	0.00	0.00	0.93	0.00	0.06
Routes	0.09	0.00	0.00	0.02	0.00	0.89	0.00
Pylônes	0.01	0.00	0.00	0.00	0.10	0.00	0.89

Tableau 4.15 Matrice de confusion de la Localisation #2 après segmentation sémantique avec prise en compte du voisinage

Localisation #2	Sol	Végétation	Bâtiments	Surfaces hydrologiques	Conducteurs	Routes	Pylônes
Sol	0.90	0.01	0.04	0.01	0.00	0.05	0.00
Végétation	0.01	0.94	0.01	0.00	0.00	0.00	0.04
Bâtiments	0.11	0.00	0.89	0.00	0.00	0.00	0.00
Surfaces hydrologiques	0.13	0.01	0.03	0.78	0.00	0.06	0.00
Conducteurs	0.00	0.00	0.00	0.00	0.95	0.00	0.04
Routes	0.04	0.00	0.08	0.06	0.00	0.81	0.00
Pylônes	0.01	0.05	0.02	0.00	0.13	0.00	0.79

#### 4.4.5 Synthèse de l'étude d'ablation et comparaison des différents résultats

Dans cette section, nous comparons les différents résultats de segmentation sémantique obtenus par nos méthodes. Chaque tableau (Tableau 4.16 au Tableau 4.20) représente une Localisation, et par conséquent, un nuage de points multisource. Chaque colonne représente une méthode. Dans ces tableaux, nos méthodes sont appelées **Mono**, pour la segmentation sémantique monosource (section 4.4.1), **Multi** pour la méthode proposée pour la segmentation multisource (section 4.4.2), **Cascade** pour la méthode incluant la seconde segmentation sémantique (section 4.4.3) et enfin, **Voisin** pour la méthode complète.

La Localisation #1 est un nuage de points qui comporte sept classes, avec deux parmi elles « Pylônes » et « Conducteurs » qui ne sont représentées que par quelques points. Ce manque de points entraîne une mauvaise représentation et se traduit par une faible valeur de *MCC* dans la plupart de nos méthodes, allant de 0,16 pour la méthode monosource jusqu'à 0,46 avec les méthodes des voisins. Dans la plupart des cas, les meilleurs résultats sont obtenus avec le *Workflow* complet. La classe des « Routes » décroît légèrement de 0,88 à 0,85 lorsque tous les

éléments de notre proposition sont utilisés. Il y a une différence de 2 ou 4 points pour l'*Overall Accuracy (OA)* entre chaque méthode. Quant à la valeur de l'indice *MCC*, elle peut augmenter jusqu'à 8 points.

Tableau 4.16 Valeurs de la *MCC* pour la Localisation #1 au travers des différentes parties de l'étude d'ablation

Localisation #1				
	Mono	Multi	Cascade	Voisin
Sol	0.88	0.88	0.89	<b>0.90</b>
Végétation	0.86	0.90	0.93	<b>0.94</b>
Bâtiments	0.83	0.77	0.95	<b>0.96</b>
Surfaces hydrologiques	0.78	<b>0.85</b>	0.68	0.69
Conducteurs	0.47	0.55	0.53	<b>0.58</b>
Routes	<b>0.88</b>	0.86	0.83	0.85
Pylônes	0.16	0.31	0.35	<b>0.46</b>
Moyenne <i>MCC</i>	0.69	0.73	0.74	<b>0.77</b>
<i>Overall Accuracy (OA)</i>	0.90	0.92	0.92	<b>0.94</b>

Dans le cas de la Localisation #2, le Tableau 4.17 indique que la segmentation sémantique multisource peut donner de meilleurs résultats que les méthodes en cascade ou de voisinage. Pour les classes telles que les « Routes », le « Sol » et la « Végétation », l'utilisation de nouvelles caractéristiques avec des sources multiples, puis une deuxième classification et des caractéristiques voisines a pour effet d'augmenter l'indice *MCC*.

Bien que les points des « Conducteurs » soient correctement identifiés comme des « Conducteurs », les « Conducteurs » et les « Pylônes » ont tendance à se mélanger au fur et à mesure que le nombre d'attributs augmente, ce qui entraîne une valeur plus faible de l'indice *MCC* avec seulement 0.58 pour le Tableau 4.16. La faible valeur de l'indice *MCC* pour les « Routes » et les « Surfaces hydrologiques » avec la méthode Mono est due à la confusion entre trois classes, à savoir les « Routes », les « Surfaces hydrologiques » et le « Sol ». La valeur moyenne de l'indice *MCC* augmente jusqu'à 10 points, tandis qu'une augmentation plus conséquente est observée après l'ajout des attributs multisources dans les données

d'apprentissage. En effet, l'indice *OA* augmente jusqu'à 5 points. L'augmentation la plus importante se produit après l'utilisation d'une segmentation multisource.

Tableau 4.17 Valeurs de la *MCC* pour la Localisation #2 au travers des différentes parties de l'étude d'ablation

Localisation #2				
	Mono	Multi	Cascade	Voisin
Sol	0.78	0.82	0.82	<b>0.83</b>
Végétation	0.92	0.92	0.91	<b>0.94</b>
Bâtiments	0.86	<b>0.89</b>	0.85	0.86
Surfaces hydrologiques	0.54	0.56	0.79	<b>0.83</b>
Conducteurs	0.66	<b>0.92</b>	0.74	0.68
Routes	0.36	0.41	0.61	<b>0.66</b>
Pylônes	0.75	<b>0.80</b>	0.76	0.76
Moyenne <i>MCC</i>	0.70	0.76	0.78	<b>0.80</b>
<i>OA</i>	0.84	0.88	0.88	<b>0.89</b>

Pour la Localisation #3, le principal problème concerne la classe « Surfaces hydrologiques », qui représente de petites rivières et de petits lacs, souvent entourés de végétation ; en conséquence, la *MCC* indiquée dans les Tableaux 4.18 à 4.20 présente une valeur extrêmement faible pour la classe « Surfaces hydrologiques ». Les valeurs moyennes de l'indice *MCC* et de l'indice *OA* s'améliorent lorsque les caractéristiques multisources sont prises en compte, comme le montrent les Tableaux 4.18 à 4.20 et les Figures 4.25 à 4.27. À chaque emplacement, l'*OA* augmente d'au moins 2 points et la *MCC* moyen augmente de 4 à 7 points pour la Localisation #3.2 passant pour *OA* de 0.86 à 0.90 dans le Tableau 4.19 quand la *MCC* augmente de 0.67 à 0.78. Le meilleur résultat est observé à la Localisation #3.2 par rapport aux Localisation #3.1 et #3.3, il est dû à l'absence de surfaces d'eau qui peuvent perturber les performances sur les Localisation #3.1 et #3.3. L'utilisation de la deuxième segmentation, Cascade, diminue les résultats moyens des indices *MCC* qui passe de 0.68 à 0.67 entre le Multi et la Cascade dans le Tableau 4.18 et de 0.71 à 0.70 dans le Tableau 4.20.

Le principal inconvénient de notre méthode est le temps consommé par l'utilisation de trois technologies au lieu d'une. Le temps de calcul nécessaire pour les attributs de chaque technologie est ajouté au temps de calcul général pour tous les nuages de points.

Tableau 4.18 Valeurs de la *MCC* pour la Localisation #3.1 au travers des différentes parties de l'étude d'ablation

Localisation #3.1				
	Mono	Multi	Cascade	Voisin
Sol	0.62	0.70	0.74	<b>0.86</b>
Végétation	0.69	0.69	0.65	<b>0.70</b>
Bâtiments	0.58	<b>0.65</b>	0.64	0.63
Surfaces hydrologiques	0.14	0.17	0.11	<b>0.18</b>
Conducteurs	0.97	0.98	0.98	<b>0.99</b>
Routes	0.52	0.67	0.70	<b>0.76</b>
Pylônes	0.84	0.87	0.88	<b>0.90</b>
Moyenne MCC	0.62	0.68	0.67	<b>0.72</b>
OA	0.89	0.91	<b>0.92</b>	<b>0.92</b>

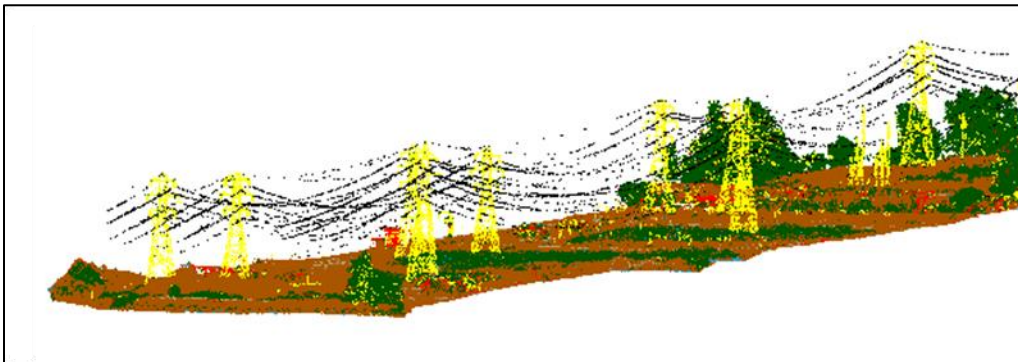


Figure 4.25 Localisation #3.1 après segmentation sémantique avec prise en compte du voisinage



Tableau 4.19 Valeurs de la *MCC* pour la Localisation #3.2 au travers des différentes parties de l'étude d'ablation

Localisation #3.2				
	Mono	Multi	Cascade	Voisin
Sol	0.56	0.65	0.66	<b>0.72</b>
Végétation	0.78	0.78	0.75	<b>0.82</b>
Bâtiments	0.62	<b>0.85</b>	0.80	0.80
Surfaces hydrologiques	-	-	-	-
Conducteurs	0.97	<b>0.97</b>	0.94	0.96
Routes	0.29	0.40	0.44	<b>0.54</b>
Pylônes	0.78	0.81	0.79	<b>0.82</b>
Moyenne MCC	0.67	0.74	0.74	<b>0.78</b>
OA	0.86	<b>0.90</b>	0.88	<b>0.90</b>

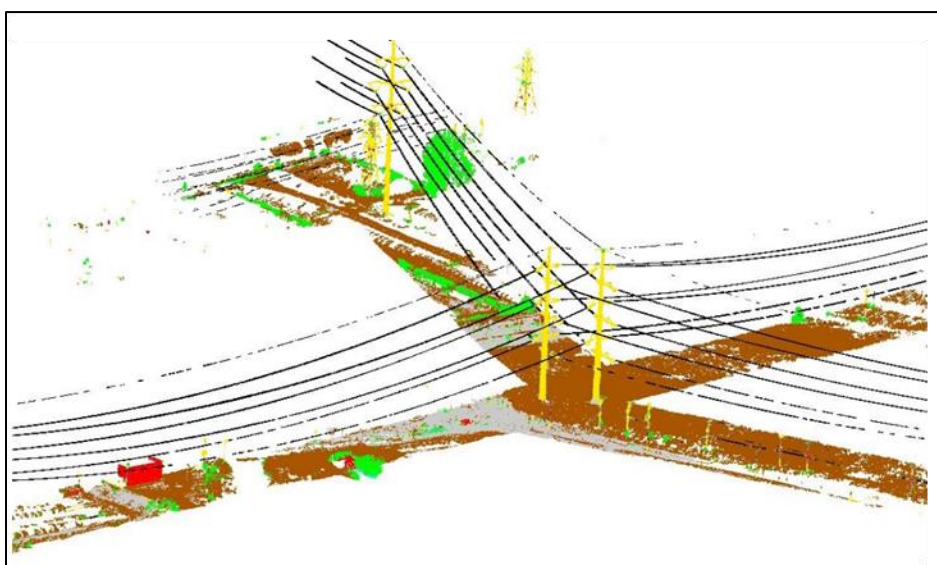


Figure 4.26 Localisation #3.2 après segmentation sémantique avec prise en compte du voisinage

Tableau 4.20 Valeurs de la *MCC* pour la Localisation #3.3 au travers des différentes parties de l'étude d'ablation

Localisation #3.3				
	Mono	Multi	Cascade	Voisin
Sol	0.68	0.76	0.76	<b>0.84</b>
Végétation	0.71	0.65	0.66	<b>0.81</b>
Bâtiments	0.73	<b>0.82</b>	0.80	0.78
Surfaces hydrologiques	0.07	<b>0.18</b>	0.13	0.09
Conducteurs	0.99	0.98	0.97	<b>0.99</b>
Routes	0.69	<b>0.79</b>	0.76	0.77
Pylônes	0.79	0.76	<b>0.80</b>	<b>0.80</b>
<b>Moyenne MCC</b>	0.67	0.71	0.70	<b>0.73</b>
<b>OA</b>	0.90	0.92	0.91	<b>0.93</b>

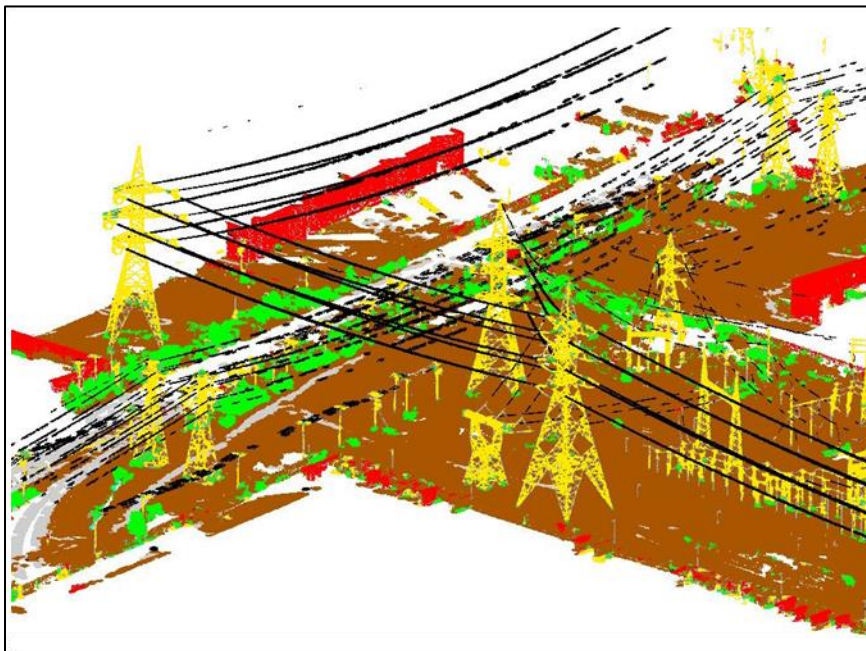


Figure 4.27 Localisation #3.3 après segmentation sémantique avec prise en compte du voisinage

## 4.5 Synthèse globale

Notre approche monosource a permis de prendre un premier contact avec les données disponibles et à les manipuler tout en appréhendant le sujet multisource. Cette première approche a aussi permis d'obtenir les premiers résultats convenables pour la segmentation sémantique monosource.

Notre proposition permet de discriminer de manière efficace les différentes classes recherchées par le PI du projet. Les attributs calculés ainsi que ceux propagés permettent d'apporter des informations supplémentaires à l'ensemble des nuages de points.

Le fait de ne pas utiliser un apprentissage profond peut être vu comme un point positif dans le cas où une nouvelle classe doit être discriminée selon ses caractéristiques géométriques. L'utilisation de la matrice de confusion permet d'améliorer la compréhension des résultats et des performances de l'algorithme. Dans certains cas, elle permet d'imaginer de nouveaux attributs afin de discriminer des classes trop semblables.

Au cours de notre étude, nous avons pu constater certaines problématiques lors de l'examen des résultats. Comme nous l'avons mentionné, ces problèmes sont intrinsèques à notre proposition. Par exemple, c'est le cas pour les points qui sont situés sur les bords du nuage de points. Le voisinage de ces points n'est pas complet, les valeurs d'attributs calculés à l'aide de sphères ou de cylindre correspondent alors à des attributs semblables à d'autres classes.

Une autre problématique est le traitement des classes en amont. Dans notre cas, hormis certains nuages de points, nous avons considéré la classification comme parfaite. Cependant, selon la compréhension de l'expert qui a classé manuellement les points à la main, certains points peuvent appartenir à plus d'une classe. Il est possible de voir, par exemple dans la Figure 4.28, que des points appartenant à la classe « Sol » peuvent être aussi classés dans la classe « Routes » ; des chemins de passages sont aussi visible en affichant l'attribut *Intensité de*

*retour* du lidar aéroporté (droite). Les points ici représentés en vert sur la Figure 4-28 (gauche) correspondent à des véhicules.

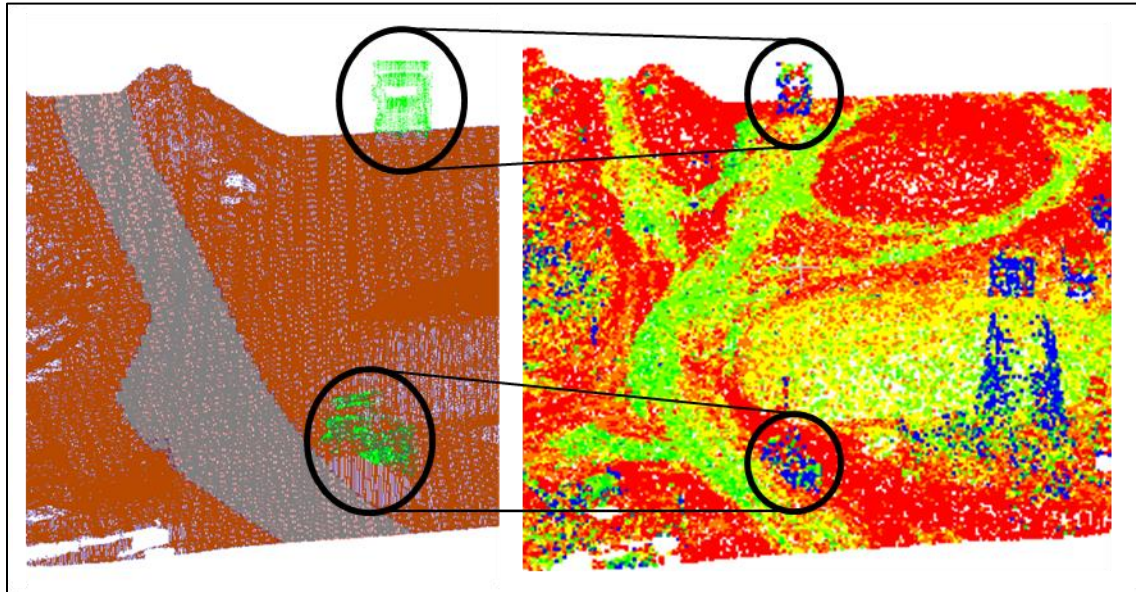


Figure 4.28 Problématique des classes pour l'apprentissage. À gauche segmentation sémantique proposée par le PI (en marron le « Sol », en gris les « Routes » et en verre des « Camions »), à droite l'intensité de retour du lidar aéroporté révèle des chemins de passages pour les camions

Suite à nos résultats, il est nécessaire de se comparer aux algorithmes qui peuvent être trouvés dans la littérature. Le prochain chapitre nous permettra cette comparaison.

## CHAPITRE 5

### ÉTUDE COMPARATIVE DES PERFORMANCES DE LA SOLUTION PROPOSÉE

#### Sommaire

5.1	Introduction.....	137
5.2	Sélection d’algorithmes pour l’étude comparative .....	137
5.2.1	Critères de sélection.....	138
5.2.2	Algorithmes sélectionnés .....	138
5.2.2.1	Algorithme RandLA-Net .....	139
5.2.2.2	Algorithme PointNet++ .....	141
5.2.2.3	Algorithme PVCNN.....	142
5.3	Présentation des résultats de l’étude comparative .....	144
5.3.1	Résultats monosources.....	144
5.3.2	Résultats multisources .....	145
5.4	Conclusion .....	149

#### 5.1 Introduction

Comme corollaire aux premiers résultats de notre proposition, nous cherchons à réaliser une étude comparative (*Benchmarking*) pour sonder les performances (avantages et limites) de la solution proposée par rapport aux performances de différents algorithmes recensés dans la littérature. À partir des jeux de données du PI, nous tentons d’opposer les résultats de notre segmentation sémantique aux résultats obtenus par d’autres algorithmes publiés dans le domaine. Ce chapitre revient donc dans un premier temps sur le processus de sélection des algorithmes retenus. Par la suite, leur fonctionnement est brièvement décrit. Enfin, une comparaison est réalisée et commentée.

#### 5.2 Sélection d’algorithmes pour l’étude comparative

Comme nous l’avons déjà mentionné dans la revue de la littérature, il existe une multitude d’algorithmes qui permettent la segmentation sémantique de nuages de points. Cette section

nous permet de présenter, dans un premier temps, les critères pour la sélection des algorithmes, puis, dans un second temps, les algorithmes qui ont été sélectionnés pour notre comparaison.

### 5.2.1 Critères de sélection

Étant donnée la multitude des algorithmes, des critères sont donc retenus pour sélectionner, au mieux, une partie de ces algorithmes pour l'étude comparative :

1. la disponibilité du code source,
2. la similitude des données d'entrée, bien qu'aucun des algorithmes identifiés ne traitent de données multisources comme les données du PI,
3. les résultats des algorithmes sur différents jeux de données de références disponibles en ligne.

Les algorithmes sélectionnés ont été utilisés tels que publiés sans tenter de les modifier pour prendre en compte l'aspect multisource de nos données. Cela a pour avantage de limiter le temps consacré à chaque algorithme, cependant, les résultats obtenus auraient pu être meilleurs en transformant les algorithmes sélectionnés afin de mieux prendre en compte l'aspect multisource. Nos jeux de données multisources ont donc été traité comme si l'ensemble des points ne provenait que d'une seule source.

### 5.2.2 Algorithmes sélectionnés

Les algorithmes qui répondent à nos critères sont nombreux (de l'ordre de la centaine). Par conséquent, le choix effectué n'est que partiel. En effet, le **RandLA-Net** (Randomized Local Aggregation Network) (Q. Hu et al., 2019) et le **PVCNN** (Point-Voxel CNN) (Liu et al., 2019) affichent de bonnes performances sur les jeux de données disponibles en ligne. Ils sont retenus. Néanmoins, l'algorithme **PointNet++** (Qi et al., 2017) est aussi retenu, car il est considéré, encore aujourd'hui, comme une valeur de référence (les résultats de PointNet++ ont, aujourd'hui, des performances légèrement inférieures à celles du RandLA-Net et PVCNNs). Les sections suivantes permettent de décrire les différents algorithmes et leur fonctionnement.

Comme recensé dans la Revue de la littérature, il est possible de décomposer les propositions d'apprentissage machine en différentes parties : échantillonnage (*sampling*), le regroupement des points (*grouping*) et enfin le fait de cartographier (*mapping*) les points.

### 5.2.2.1 Algorithme RandLA-Net

Le RandLA-Net présente une architecture de réseaux de neurones pour le traitement de nuages de points. Il est proposé par Hu et al. en 2019. Pour la segmentation sémantique, il affiche des performances de pointe dans plusieurs jeux de données. L'algorithme RandLA-Net est capable d'apprendre des représentations de ces nuages de points en les traitant à **plusieurs échelles**. Dans le cas de RandLA-Net, les différentes échelles sont représentées par le nombre de voisins  $K$  pris en compte pour regrouper les points en prenant un  $K$  petit, l'algorithme va décrire plusieurs échelles mésoscopique du centimètre à macroscopique pour le mètre. L'algorithme RandLA-Net se compose de plusieurs étapes visibles dans la Figure 5.1.

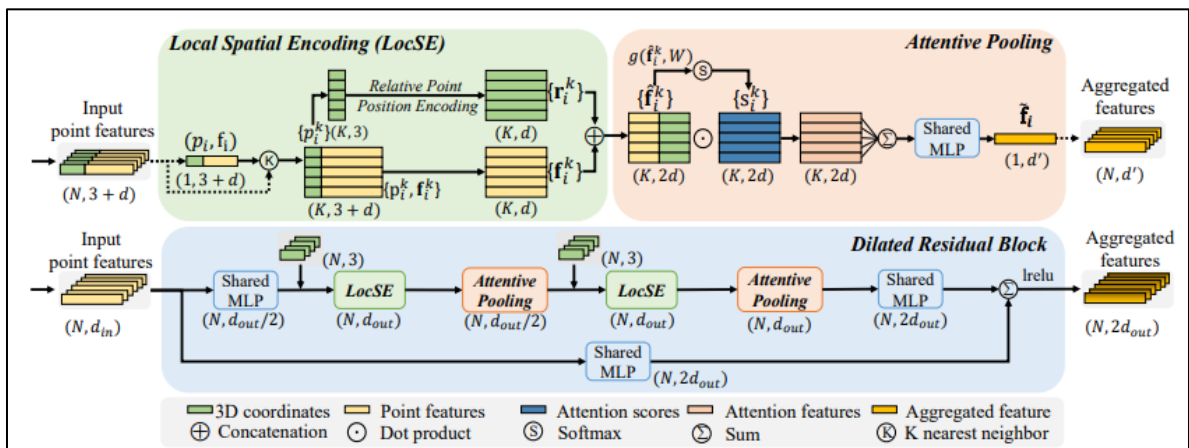


Figure 5.1 Schéma de la proposition RandLA-Net  
Tirée de Hu et al. (2019)

- (i) L'échantillonnage : le nuage de points d'entrée est sous-échantillonné de manière **aléatoire** pour réduire sa taille. Un sous-ensemble de points est sélectionné de manière aléatoire dans le nuage de points original et un graphe de voisinage local est ainsi construit autour de chaque point.

- (ii) L'extraction des attributs : le **graphe de voisinage local** de chaque point est traité par une série de **couches convolutives** afin d'extraire les attributs. Ces couches convolutives opèrent sur les coordonnées des points et sur leurs attributs (par exemple RGB). Elles sont conçues pour capturer les informations géométriques et sémantiques.
- (iii) Les attributs de chaque point sont ensuite traités par une série de couches convolutives pour capturer les caractéristiques à **différentes échelles**. Ces couches opèrent sur les caractéristiques locales et sur leurs graphes de voisinage correspondants. Elles sont conçues pour capturer les caractéristiques locales et globales du nuage de points.
- (iv) Les attributs extraits à différentes échelles sont fusionnés pour produire un vecteur d'attributs qui capture la forme générale du nuage de points. On déploie une série de couches entièrement connectées qui opèrent sur les attributs concaténés à différentes échelles.
- (v) L'agrégation des attributs : les attributs de chaque point sont ensuite agrégés pour les points voisins afin d'obtenir des informations contextuelles. On déploie une opération de mise en commun des attributs pour combiner les attributs des points voisins avec différents poids, en fonction de leur distance spatiale par rapport au point central.
- (vi) Le *mapping* : le vecteur final des caractéristiques passe par plusieurs couches entièrement connectées (MLP) pour produire la sortie du réseau qui peut être utilisée pour des tâches telles que la classification, la segmentation sémantique ou encore la détection d'objets.

En reprenant les mots clés de l'état de l'art, la proposition RandLA-Net peut donc être classée dans les catégories de proposition utilisant une **méthode hybride** qui utilise la **théorie des graphes**. Cette proposition est aussi **monosource**, **multiéchelle** et **générale**. Enfin, elle utilise des **attributs** produits à l'aide d'un **CNN**.



### 5.2.2.2 Algorithme PointNet++

L'algorithme PointNet++ présente une architecture de réseaux de neurones pour le traitement des nuages de points. Il est proposé par Qi et al. en 2017 comme une extension de l'architecture de l'algorithme originale PointNet. L'algorithme PointNet++ est capable d'apprendre des représentations à **plusieurs échelles**. Les différentes étapes qui composent cet algorithme sont visible dans la Figure 5.2 et sont les suivantes :

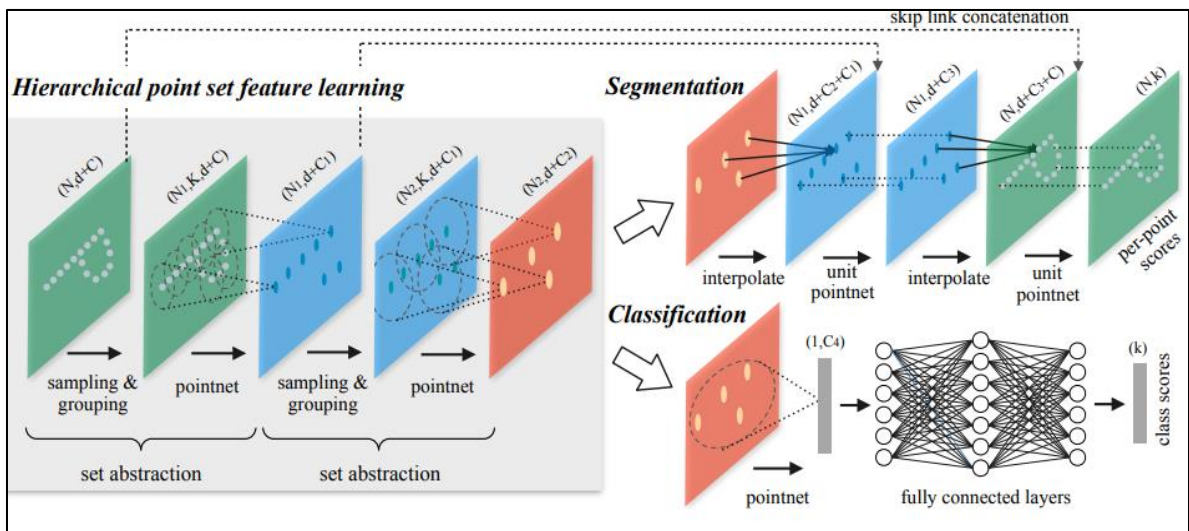


Figure 5.2 Schéma de la proposition PointNet++  
Tirée de Qi, Yi, Su, & Guibas (2017)

- (i) Transformation de l'entrée : le nuage de points d'entrée est transformé pour le rendre invariant à la translation et à la rotation.
- (ii) L'échantillonnage (*sampling*) : le nuage de points transformé est alors échantillonné en utilisant l'algorithme *Farthest Point Sampling* (FPS) qui a pour avantage de maintenir une meilleure représentation du nuage de points.
- (iii) L'extraction des attributs : le nuage de points est traité par un module PointNet pour extraire les attributs de chaque point. Ce module se compose de plusieurs **couches entièrement connectées** qui opèrent sur chaque point indépendamment, suivies d'une opération de mise en commun (*MaxPooling*) pour agréger les attributs de tous les points. Le module PointNet est ensuite appliqué de manière réursive à

**plusieurs échelles.** À chaque échelle, un sous-ensemble de points est sélectionné à l'aide de FPS. Ces points sont traités par le module PointNet pour extraire les attributs à cette échelle. Les attributs des points échantillonnés sont ensuite utilisés pour mettre à jour les attributs de leurs voisins dans le nuage original. Ce processus est répété à plusieurs échelles pour capturer les attributs à différents niveaux d'abstraction.

- (iv) L'agrégation des attributs : les attributs extraits à différentes échelles sont agrégés à l'aide d'un module de fusion des attributs. Ce module se compose de plusieurs couches entièrement connectées qui opèrent sur les attributs concaténés à différentes échelles, suivies d'une opération de mise en commun (*MaxPooling*) pour agréger les caractéristiques de toutes les échelles.
- (v) Le *mapping* : le vecteur d'attributs passe par plusieurs couches entièrement connectées pour produire la sortie, qui peut être utilisée pour des tâches telles que la classification, la segmentation sémantique ou encore la détection d'objets.

En reprenant les mots clés de l'état de l'art, la proposition PointNet++ peut donc être classée dans les catégories de proposition utilisant une méthode ***Point Based***, **monosource**, **multiéchelle**, **générale**, utilisant des **attributs** produits à l'aide d'un **réseau de neurones**.

### 5.2.2.3 Algorithme PVCNN

Le PVCNN présente une architecture de réseaux de neurones pour le traitement des nuages de points. Il est proposé par Liu et al. en 2019. Il reprend en partie le travail de PointNet comme structure. La structure est visible Figure 5.3.

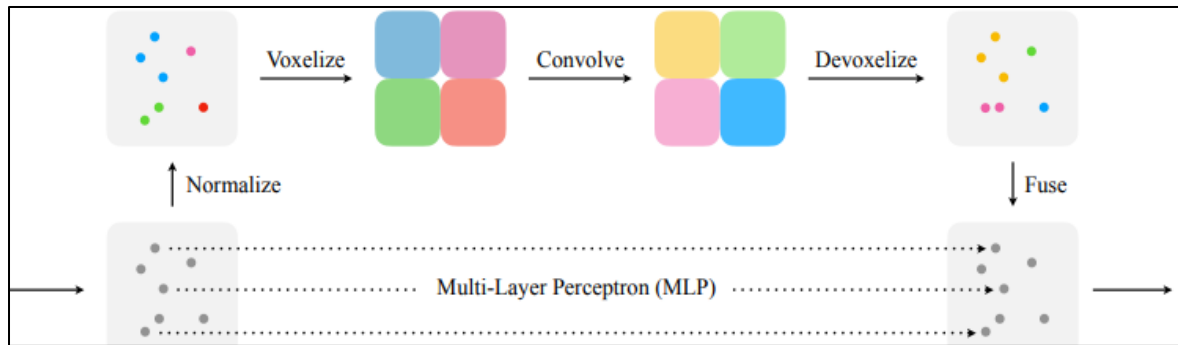


Figure 5.3 Schéma de la proposition PVCNN  
Tirée de Liu, Tang, Lin, & Han (2019)

- (i) Transformation de l'entrée : le nuage de points est transformé en une représentation de taille fixe, en voxélisant le nuage de points.
- (ii) Extraction des attributs des points : les points du nuage sont traités par un module de type PointNet pour extraire les attributs de chaque point. Ce module se compose de plusieurs couches entièrement connectées qui opèrent sur chaque point indépendamment.
- (iii) Extraction des attributs des voxels : en parallèle du travail sur les points, les voxels sont traités par une série de couches convolutives pour extraire les attributs des voxels. Ces couches convolutives opèrent sur les voxels et leurs attributs.
- (iv) Fusion point-voxel : les attributs extraits du nuage de points et de la grille de voxels sont combinés à l'aide d'un module de fusion des attributs. Ce module se compose de plusieurs couches entièrement connectées.
- (v) Le mapping : le vecteur de caractéristiques final passe par plusieurs couches entièrement connectées pour produire le résultat, qui peut être utilisé pour des tâches telles que la classification, la segmentation et la détection d'objets.

En reprenant les mots clés de l'état de l'art, la proposition PVCNN peut donc être classée dans les catégories de proposition utilisant une méthode **point-voxel**, **monosource**, **multiéchelle**, **générale**, utilisant des **attributs** produits à l'aide d'un **CNN** et d'un **réseau de neurones**.

### 5.3 Résultats de l'étude comparative

Nous présentons et commentons les résultats obtenus avec ces différentes propositions, dans un premier temps avec les résultats monosource puis, par la suite, avec les résultats multisources.

#### 5.3.1 Résultats monosources

Les résultats décrits dans cette section correspondent aux résultats monosource et plus précisément la Localisation #4. Le Tableau 5.1 présente les résultats obtenus. Pour chaque classe, la colonne avec le meilleur résultat est indiquée en **gras**. Un fond **vert** est employé dans le cas de performances jugées comme similaires, et un fond de couleur **orange** illustre des divergences notables entre les trois approches (notre proposition, RandLANet et PointNet++). L'approche PVCNN n'avait pas encore été étudié à ce moment de notre étude. L'étalon de l'étude comparative est toujours le classement réalisé par les experts (réf. Chapitre 4). Les métriques M IoU et OA sont employées comme mesures globales des performances de chaque algorithme pour cette section.

On y constate que les classes ayant une géométrie simple (telle que la classe « Conducteurs ») conservent de bons résultats, peu importe l'algorithme utilisé. Ce n'est cependant pas le cas de toutes les classes. On retrouve par exemple plus de difficultés à classer les points appartenant à la classe « Pylônes » (phénomène accentué pour le cas de l'algorithme PointNet++). Certains des défauts de la segmentation sémantique sont expliqués par le peu de points qui représentent la classe ou encore par sa proximité avec d'autres classes dans le nuage de points.

En résumé, notre proposition est compétitive. Elle affiche des performances honorables (sinon, meilleures) par rapport à l'État de l'art du domaine. Toutefois, nous tenons à souligner que les algorithmes de la littérature ne sont pas optimisés pour les données du PI. Par conséquent, toute généralisation de cette conclusion doit impérativement être appuyée par d'autres études de cas.

Tableau 5.1 Résultats comparatifs de la segmentation sémantique monosource de la Localisation #4 avec les algorithmes RandLA-Net et PointNet++

Méthode utilisée	Proposition	RandLA-Net	PointNet++
<b>Appris sur</b>	<b>Localisation #5</b>	<b>Localisation #5</b>	<b>Localisation #5</b>
<b>Testé sur</b>	<b>Localisation #4</b>	<b>Localisation #4</b>	<b>Localisation #4</b>
Sol	0.91	<b>0.93</b>	0.87
Végétation basse	0.31	0.36	<b>0.41</b>
Végétation haute	<b>0.97</b>	0.95	0.90
Bâtiments	0.60	<b>0.61</b>	0.46
Surfaces hydrologiques	0.56	0.37	<b>0.63</b>
Conducteurs	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>
Routes	<b>0.36</b>	0.33	0.28
Autre	0.28	0.37	<b>0.44</b>
Pylônes	<b>0.45</b>	0.40	0.12
<b>M IoU</b>	<b>0.60</b>	0.59	0.57
<b>OA</b>	<b>0.92</b>	0.91	0.87
<b>Temps de calcul</b>	3h	45 min	1h10

### 5.3.2 Résultats multisources

Cette section revient sur les résultats obtenus dans notre étude d'ablation en les comparant avec les résultats obtenus avec les propositions de RandLA-Net et PVCNN. PointNet++ n'a pas été utilisé cette fois, le PVCNN peut être considéré comme une version légèrement améliorée de PointNet++.

À travers les résultats visibles dans les Tableaux 5.2 à 5.6, nous pouvons constater que les performances de notre proposition se retrouvent en tête (ou proche) de tous les algorithmes. On y retrouve certains défauts (lacunes) des propositions qui utilisent l'apprentissage machine avec peu de points pour la phase de l'apprentissage. Ce défaut est particulièrement visible avec la classe « Surface hydrologique » pour l'ensemble de la Localisation #3 où la MCC maximum est de 0.2.

Grâce à la prépondérance de certaines classes comme le « Sol » et la « Végétation » les valeurs d’OA arrivent à être proche ou légèrement au-dessus des 90% de bonne classification sur l’ensemble des nuages des points. En ne comparant que notre proposition au complet, les plus gros écarts pour l’OA sont de 2 points (visible pour la Localisation #2 Tableau 5.3 et Localisation #3.2 Tableau 5.5).

Tableau 5.2 Résultats comparatifs de la segmentation sémantique de la Localisation #1 avec l’étude d’ablation

MCC pour les classes	Localisation #1					
	Mono	Multi	Cascade	Voisin	RandLA-Net	PVCNN
Sol	0.88	0.88	0.89	0.90	<b>0.94</b>	0.91
Végétation	0.86	0.90	0.93	<b>0.94</b>	0.92	0.89
Bâtiments	0.83	0.77	0.95	<b>0.96</b>	0.90	0.91
Surfaces hydrologiques	0.78	<b>0.85</b>	0.68	0.69	0.77	0.66
Conducteurs	0.47	0.55	0.53	<b>0.58</b>	0.50	<b>0.58</b>
Routes	<b>0.88</b>	0.86	0.83	0.85	0.82	0.87
Pylônes	0.16	0.31	0.35	0.46	<b>0.51</b>	0.38
<b>Moyenne MCC</b>	0.69	0.73	0.74	<b>0.77</b>	<b>0.77</b>	0.74
<b>OA</b>	0.90	0.92	0.92	<b>0.94</b>	0.93	0.93
<b>Temps de calcul</b>	5h	6h	8h	9h	2h	3h30

Tableau 5.3 Résultats comparatifs de la segmentation sémantique de la Localisation #2 avec l'étude d'ablation

	Localisation #2					
MCC pour les classes	Mono	Multi	Cascade	Voisin	RandLA-Net	PVCNN
Sol	0.78	0.82	0.82	0.83	<b>0.86</b>	0.82
Végétation	0.92	0.92	0.91	<b>0.94</b>	0.92	0.90
Bâtiments	0.86	0.89	0.85	0.86	<b>0.90</b>	0.88
Surfaces hydrologiques	0.54	0.56	0.79	<b>0.83</b>	0.63	0.51
Conducteurs	0.66	<b>0.92</b>	0.74	0.68	0.88	0.90
Routes	0.36	0.41	0.61	0.66	0.64	<b>0.75</b>
Pylônes	0.75	<b>0.80</b>	0.76	0.76	0.55	0.70
<b>Moyenne MCC</b>	0.70	0.76	0.78	<b>0.80</b>	0.77	0.78
<b>OA</b>	0.84	0.88	0.88	0.89	<b>0.91</b>	0.90

Tableau 5.4 Résultats comparatifs de la segmentation sémantique de la Localisation #3.1 avec l'étude d'ablation

	Localisation #3.1					
MCC pour les classes	Mono	Multi	Cascade	Voisin	RandLA-Net	PVCNN
Sol	0.62	0.70	0.74	<b>0.86</b>	0.83	0.84
Végétation	0.69	0.69	0.65	0.70	<b>0.75</b>	0.67
Bâtiments	0.58	0.65	0.64	0.63	0.60	<b>0.66</b>
Surfaces hydrologiques	0.14	0.17	0.11	0.18	0.09	<b>0.20</b>
Conducteurs	0.97	0.98	0.98	<b>0.99</b>	0.98	0.98
Routes	0.52	0.67	0.70	<b>0.76</b>	0.73	0.65
Pylônes	0.84	0.87	0.88	<b>0.90</b>	<b>0.90</b>	0.87
<b>Moyenne MCC</b>	0.62	0.68	0.67	<b>0.72</b>	0.70	0.70
<b>OA</b>	0.89	0.91	0.92	0.92	<b>0.93</b>	0.91

Tableau 5.5 Résultats comparatifs de la segmentation sémantique de la Localisation #3.2 avec l'étude d'ablation

	Localisation #3.2					
MCC pour les classes	Mono	Multi	Cascade	Voisin	RandLA-Net	PVCNN
Sol	0.56	0.65	0.66	0.72	0.80	<b>0.82</b>
Végétation	0.78	0.78	0.75	<b>0.82</b>	0.77	0.73
Bâtiments	0.62	<b>0.85</b>	0.80	0.80	0.72	0.77
Surfaces hydrologiques	-	-	-	-	-	-
Conducteurs	0.97	0.97	0.94	0.96	<b>0.98</b>	0.96
Routes	0.29	0.40	0.44	<b>0.54</b>	0.49	0.33
Pylônes	0.78	0.81	0.79	<b>0.82</b>	0.81	0.78
<b>Moyenne MCC</b>	0.67	0.74	0.74	<b>0.78</b>	0.76	0.73
<b>OA</b>	0.86	0.90	0.88	0.90	0.91	<b>0.92</b>

Tableau 5.6 Résultats comparatifs de la segmentation sémantique de la Localisation #3.3 avec l'étude d'ablation

	Localisation #3.3					
MCC pour les classes	Mono	Multi	Cascade	Voisin	RandLA-Net	PVCNN
Sol	0.68	0.76	0.76	0.84	0.78	<b>0.88</b>
Végétation	0.71	0.65	0.66	<b>0.81</b>	0.70	0.69
Bâtiments	0.73	<b>0.82</b>	0.80	0.78	0.65	0.73
Surfaces hydrologiques	0.07	<b>0.18</b>	0.13	0.09	0.12	0.15
Conducteurs	<b>0.99</b>	0.98	0.97	<b>0.99</b>	<b>0.99</b>	0.98
Routes	0.69	0.79	0.76	0.77	<b>0.80</b>	0.77
Pylônes	0.79	0.76	0.80	0.80	<b>0.85</b>	0.82
<b>Moyenne MCC</b>	0.67	0.71	0.70	<b>0.73</b>	0.71	0.72
<b>OA</b>	0.90	0.92	0.91	<b>0.93</b>	0.92	0.91



## 5.4 Conclusion

L'étude comparative a permis de mesurer les performances de notre proposition en la comparant avec des propositions récentes issues de la littérature spécialisée du domaine.

Notre approche multisource offre, dans une majeure partie des cas, de meilleures performances que les approches proposées récemment dans le domaine. Certes, l'étude comparative est limitée aux études de cas fournies par le PI du projet. Une retenue dans la généralisation de cette conclusion est donc de mise.

Néanmoins, les performances affichées démontrent le potentiel de déploiement de notre proposition et sa précision dans le cas spécifique des données du PI du projet.



## CONCLUSION, RECOMMANDATIONS ET PERSPECTIVES

La numérisation et la captation à distance sont des domaines voués à croître dans les années à venir. Cette croissance est alimentée par une plus grande accessibilité aux technologies de captation et par l'augmentation de la puissance de calcul pour le traitement de données. Les projets de captation à distance couvrent un large éventail de domaines et de besoins. À titre d'exemples, on peut citer la numérisation inverse pour la création de jumeaux numériques, la conduite autonome des véhicules ou encore la captation des sites archéologiques. De nos jours, diverses technologies sont matures et sont disponibles. On retrouve, par exemple, des lidars embarqués sur des véhicules motorisés pour les rendre autonomes, des TLS pour numériser avec précision des terrains et bâtiments, ou encore l'utilisation de lidars embarqués sur drones ou des appareils photographiques pour numériser des terrains accidentés ou escarpés difficiles d'accès.

Cependant, la majeure partie des projets recensés n'utilisent qu'une seule technologie à la fois. Typiquement, la technologie la plus adaptée est sélectionnée en fonction des requis et des contraintes du projet. Depuis quelques années, certains projets et initiatives entreprennent l'utilisation conjointe d'un ensemble de technologies. Il s'agit principalement de lidar couplé à des images. La motivation de telles approches est de conjuguer les avantages des deux technologies : l'accessibilité de l'image à la précision volumétrique des lidars.

Il apparaît donc pertinent de développer des méthodes qui permettent d'utiliser et de fusionner les informations issues de plusieurs technologies pour un même projet. Notre première contribution est de proposer une approche qui s'appuie sur des données de multiples sources et qui permet de segmenter sémantiquement les différentes sources de nuages de points en propageant des informations contenues dans les divers points des nuages. Cette nouvelle approche permet de traiter des nuages de points multisources tout en affichant des résultats et des performances compétitives en comparant à ceux récemment publiés dans la littérature scientifique du domaine.

Notre deuxième contribution est une proposition pour réaliser la fusion de données multiniveaux pour différents nuages de points. Bien que quelques techniques de fusion de données géospatiales et multiniveaux soient proposées dans la littérature, les données exploitées dans le cadre du présent projet permettent de se questionner sur le type de fusion nécessaire (et possible) dans le cas de données non complètes (c'est-à-dire, les données qui ne recouvrent pas la même zone géographique).

Notre troisième contribution concerne l'implantation de l'utilisation des classes avoisinantes pour améliorer une deuxième segmentation sémantique. Cette approche permet d'améliorer les performances de l'opération de segmentation sémantique sur la majeure partie des classes des nuages de points mis à disposition pour le présent projet.

Cet ensemble de contribution nous permet *in fine* de réaliser une segmentation sémantique sur les nuages de points fournis par le PI avec des résultats qui se comparent avantageusement par rapport à ceux issus d'autres algorithmes de la littérature. En reprenant les sous-objectifs énoncés dans le Chapitre 2, notre proposition nous permet de faire collaborer différentes sources en fusionnant leurs attributs. Cette proposition nous permet de réaliser la segmentation sémantique de nos données avec une précision satisfaisante.

Bien sûr, des contraintes logistiques et temporelles nous poussent à '*mettre un terme*' à nos travaux et à produire un rapport. Néanmoins, nous proposons ici quelques pistes pour de futurs travaux d'investigation et quelques pistes à explorer que notre expérience nous dicte.

(i) Nos travaux (c.-à-d. apprentissage, exploration, investigation et validation) exploitent un ensemble restreint de nuages de points mis à disposition par le PI. Des études à de plus grandes échelles, dans le sens du nombre de nuages, prenant en compte plus de données, sont pertinentes et devraient être menées. Nous émettons une conjecture qu'un ensemble plus important de nuages de points permet une amélioration de la segmentation sémantique en proposant simplement un éventail plus large de cas d'étude, et par conséquent, plus

représentatif des mêmes classes de points. Comme les couleurs jouent un rôle important comme attributs dans la segmentation sémantique automatique, ceci peut atténuer les problèmes liés à la représentativité des couleurs tout au long d'une journée et de l'année.

(ii) Nos travaux portent sur l'usage d'un algorithme d'apprentissage machine non-profond. Une seule comparaison avec des algorithmes d'apprentissage profond a été réalisée. Nous proposons d'analyser l'usage des réseaux de neurones profonds et de comparer leurs performances avec celles obtenues dans le cadre de la présente thèse.

(iii) Comme présenté dans le Chapitre 4, les classes des points et surtout leurs découpages *a priori*, jouent un grand rôle sur les performances finales. Subséquemment, si certaines classes peuvent être considérées comme une agglomération de plusieurs classes, ceci pourrait avoir une incidence sur la qualité de la segmentation. Nous proposons d'examiner l'effet (la sensibilité) de les dissocier. Par exemple, les *Routes*, *Chemins de terre*, *Voies ferroviaires* pourraient se détacher de la classe primaire *Routes*. De manière similaire, une différenciation sur la classe *Sol* pourra aider à obtenir une meilleure segmentation sémantique.

(iv) Une lacune de notre méthode, qui emploi essentiellement des attributs géométriques, porte sur les bordures (frontières) des nuages de points. En effet, les attributs sont calculés à l'aide de sphères. Par conséquent, pour un point situé sur une bordure, la moitié (ou le trois-quarts) de la sphère est vide (en s'imaginant un nuage de points coupé de manière quadrillée).

Lors du calcul d'attributs, les possibilités envisagées, mais que partiellement essayées, sont :

- de sélectionner (ou identifier) les points la bordure et de créer un *classificateur* différent, uniquement pour ces zones;
- de dupliquer le nuage de point pour avoir une sorte de miroir qui permet d'avoir une sphère pleine d'attributs au moment du calcul d'attributs;
- de découper le nuage de points en sous nuage pour avoir plus de points avec des caractéristiques de bordures.

Chacune de ses approches engendre des biais, et n'est donc pas parfaite. Cependant, elles n'ont pas pu être toutes investiguées lors de ce doctorat. Nous recommandons d'étudier ces trois avenues pour identifier une approche optimale et adaptée au cas d'étude.

(v) Lors de l'apprentissage, nous avons conservé l'ensemble des attributs acquis et calculés. Cependant, certains de ses attributs sont redondants, il est donc envisageable dans le futur de réaliser un choix parmi les attributs à conserver. Ainsi, en plus du gain de temps réalisé, notre approche aura moins de chance de réaliser un surapprentissage.

(vi) Une étude concernant la robustesse de notre algorithme pourrait aussi être réalisée en ajoutant du bruit, de manière artificielle, aux nuages de points. Cet ajout permettrait de quantifier la sensibilité de notre algorithme.

(vii) L'ajout de données hétérogènes, telles que des images, afin de réaliser une nouvelle fusion de données avec une première segmentation sémantique sur les images est une piste envisagée pour de futurs travaux. Comme nous l'avons dit dans la revue de la littérature, ce type de fusion de données permet une amélioration des métriques de la segmentation sémantique. Elle peut aussi permettre d'améliorer notre modèle en lui permettant de prendre en compte des données de différentes natures.

Comme mentionné dans la méthodologie, avec les jeux de données utilisés, certaines localisations ne possèdent pas de nuages de points avec les attributs de couleurs RGB. Dans le cas d'images en noir et blanc, les CNN sont couramment utilisés pour coloriser les images. Coloriser signifie de mettre en couleur de manière automatique (le plus souvent des images par interprétation des nuances de gris). Dans le cas d'images, cette colorisation permet de les rendre plus vivantes et de mieux pouvoir les interpréter.

Nous avons observé au cours de notre projet que la colorisation du nuage de points pouvait aussi être réalisée. Dans notre cas spécifique, nous avons utilisé les attributs géométriques pré-sélectionnés et un réseau de neurones pour recréer des couleurs. L'utilisation de CNN a probablement le potentiel d'obtenir de meilleures performances. Cette colorisation (dont les

premiers résultats sont visible Figure 6.1) permet dans un premier temps d'améliorer l'interprétation du nuage de points par les opérateurs qui le visualisent.

Une étude pourrait être réalisée pour tenter d'identifier les biais qu'introduit l'utilisation de cette colorisation pour la segmentation sémantique des nuages de point avec un algorithme.

Il serait alors possible de comparer :

- la segmentation sémantique obtenue avec un algorithme **sans** l'utilisation des attributs **RGB**;
- la segmentation sémantique obtenue avec un algorithme en prenant en compte l'utilisation des attributs **RGB réels**;
- la segmentation sémantique obtenue avec un algorithme en prenant en compte l'utilisation des attributs **RGB recréés**.



Figure 6.1 a) Nuage de points obtenu avec la photogrammétrie par drone avec les attributs RGB obtenu lors de l'acquisition. b) Attributs R, G et B obtenu à l'aide d'un réseau de neurones prenant en entrée le vecteur d'attributs calculé

Cette étude pourrait aussi porter sur du *Transfert Learning* en utilisant les sources TLS et Photogrammétrie comme sources primaires qui possèdent l'information RGB et les sources ALS et HLS comme sources secondaires n'ayant pas l'information RGB.

À l'heure actuelle, le manque de jeux de données (points et de localisation) est un obstacle pour le projet. La saisonnalité, ainsi que la météorologie, la luminosité, doivent être prises en

compte comme ces changements qui affectent les couleurs présentent dans les nuages de points.

Au niveau des classes représentées dans les études de cas du PI, la diversité notable des différents nuages de points rend la segmentation sémantique complexe et ardue. Comme nous l'avons constaté dans la revue de la littérature, la classification de terrains permet de donner à un nuage de points une classe qui décrive l'ensemble du nuage. Il est possible d'avoir des classes telles que *Forêt*, *Prairie*, *Zone Urbaine*, *Zone Industrielle* par exemple. L'utilisation des classes voisines lors de la seconde segmentation a permis d'entrevoir les différences qui existent entre ces classes et de nouvelles classes (*Végétation*, *Sol*, *Surface Hydrologique*, etc.). Il est donc intéressant de classer adéquatement les nuages de points avant de les segmenter sémantiquement. Cette classification *a priori* des nuages de points peut être alors utilisée pour sélectionner un *classifieur* spécifique ou de changer les paramètres du *classifieur* existant pour prendre en compte cet *a priori* et obtenir ainsi une segmentation sémantique différente, et qui peut être plus performante. Cependant, cette nouvelle étape rallonge la durée de traitement des nuages de points en plus de créer potentiellement plusieurs *classifieurs*.



## ANNEXE I

### INFORMATIONS SUPPLÉMENTAIRES CONCERNANT LA CAPTATION DE LA LOCALISATION #1

Cette annexe concerne les relevés de la Localisation #1.

La technologie de captation utilisée pour le relevé du lidar aéroporté est un lidar de l'entreprise Leica de modèle ALS70HP.

Les paramètres de vol de l'avion ont été retranscrits dans le Tableau-A I-1 suivant :

Tableau-A0 I-1 Caractéristique du vol lors de la captation de la Localisation #1 avec le lidar aéroporté

Fréquence de balayage	37 Hz
Angle de balayage	20 degrés (de part et d'autre du nadir)
Hauteur de vol	1200 m au-dessus du sol
Recouvrement latéral	30%
Vitesse de vol	71 m/s
Largeur du balayage	780 m
Espacement entre lignes de vol	550 m
Empreinte au laser au sol	29 cm
Densité théorique	1 point au mètre carré

Pour les données provenant de la photogrammétrie par drone, la captation a été réalisée à l'aide d'un drone de la marque DJI avec une caméra FC6520 et une lentille MFT 15mm. Certaines informations de la captation pour la Localisation #1 ont été retranscrites dans le Tableau-A I-2 et le Tableau-A I-3.

Tableau-A0 I-2 Caractéristiques de la captation de la Localisation #1 avec la photogrammétrie par drone

Superficie couverte	0.0825 $km^2$
Nombre d'images	516 photographies géoréférencées, dont 488, calibrées
Points de contrôle	5 <i>Ground Control Points</i> (GCP)

Tableau-A0 I-3 Incertitudes de positionnement et de rotations de la caméra lors de la captation pour la Localisation #1 avec la photogrammétrie par drone

	X (m)	Y (m)	Z (m)	$\Omega$ (°)	$\Phi$ (°)	K (°)
Moyenne	0.093	0.108	0.073	0.086	0.071	0.038
Écart type	0.051	0.064	0.048	0.061	0.042	0.025

## LISTE DE RÉFÉRENCES BIBLIOGRAPHIQUES

- Antonarakis, A. S., Richards, K. S., & Brasington, J. (2008). Object-based land cover classification using airborne LiDAR. *Remote Sensing of Environment*, 112(6), 2988-2998. <https://doi.org/10.1016/j.rse.2008.02.004>
- Atzmon, M., Maron, H., & Lipman, Y. (2018). Point Convolutional Neural Networks by Extension Operators. Repéré à <https://arxiv.org/abs/1803.10091v1>
- Bai, T., Sun, K., Deng, S., Li, D., Li, W., & Chen, Y. (2018). Multi-scale hierarchical sampling change detection using Random Forest for high-resolution satellite imagery. *International Journal of Remote Sensing*, 39(21), 7523-7546. <https://doi.org/10.1080/01431161.2018.1471542>
- Barbero, B. R., & Ureta, E. S. (2011). Comparative study of different digitization techniques and their accuracy. *Computer-Aided Design*, 43(2), 188-206. <https://doi.org/10.1016/j.cad.2010.11.005>
- Behley, J., Garbade, M., Milioto, A., Quenzel, J., Behnke, S., Stachniss, C., & Gall, J. (2019, 16 août). SemanticKITTI: A Dataset for Semantic Scene Understanding of LiDAR Sequences. arXiv. Repéré à <http://arxiv.org/abs/1904.01416>
- Belton, D., & Bae, K.-H. (2010). Automating post-processing of terrestrial laser scanning point clouds for road feature surveys (p. 6).
- Beraldin, J., Mackinnon, D., & Cournoyer, L. (2015). Metrological characterization of 3D imaging systems: progress report on standards developments. Dans *17th International Congress of Metrology* (p. 21). <https://doi.org/10.1051/metrology/20151303>
- Boulch, A. (2019). Generalizing discrete convolutions for unstructured point clouds. *arXiv:1904.02375 [cs]*. Repéré à <http://arxiv.org/abs/1904.02375>
- Boulch, A., Saux, B. L., & Audebert, N. (2017). Unstructured Point Cloud Semantic Labeling Using Deep Segmentation Networks. The Eurographics Association. <https://doi.org/10.2312/3dor.20171047>
- Brodu, N., & Lague, D. (2012). 3D terrestrial lidar data classification of complex natural scenes using a multi-scale dimensionality criterion: Applications in geomorphology.

*ISPRS Journal of Photogrammetry and Remote Sensing*, 68, 121-134.  
<https://doi.org/10.1016/j.isprsjprs.2012.01.006>

- Buffi, G., Manciola, P., Grassi, S., Barberini, M., & Gambi, A. (2017). Survey of the Ridracoli Dam: UAV-based photogrammetry and traditional topographic techniques in the inspection of vertical structures. *Geomatics, Natural Hazards and Risk*, 8(2), 1562-1579. <https://doi.org/10.1080/19475705.2017.1362039>
- Charles, R. Q., Su, H., Kaichun, M., & Guibas, L. J. (2017). PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. Dans *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 77-85). Honolulu, HI : IEEE. <https://doi.org/10.1109/CVPR.2017.16>
- Chen, M., Feng, A., Prasad, P. B., McAlinden, R., & Soibelman, L. (2020a). Generating Synthetic Photogrammetric Data for Training Deep Learning based 3D Point Cloud Segmentation Models. Dans *Interservice/Industry Training, Simulation, and Education Conference (IITSEC) 2020* (p. 12). <https://doi.org/10.48550/arXiv.2008.09647>
- Chen, M., Feng, A., Prasad, P. B., McAlinden, R., & Soibelman, L. (2020b). Semantic Segmentation and Data Fusion of Microsoft Bing 3D Cities and Small UAV-based Photogrammetric Data. Dans *Interservice/Industry Training, Simulation, and Education Conference (IITSEC)* (p. 12). <https://doi.org/10.48550/arXiv.2008.09648>
- Chen, M., Hu, Q., Yu, Z., Thomas, H., Feng, A., Hou, Y., ... Soibelman, L. (2022, 13 octobre). STPLS3D: A Large-Scale Synthetic and Real Aerial Photogrammetry 3D Point Cloud Dataset. arXiv. Repéré à <http://arxiv.org/abs/2203.09065>
- Chen, X., Ma, H., Wan, J., Li, B., & Xia, T. (2017). Multi-view 3D Object Detection Network for Autonomous Driving, 6526-6534. <https://doi.org/10.1109/cvpr.2017.691>
- Chew, A. W. Z., Ji, A., & Zhang, L. (2022). Large-scale 3D point-cloud semantic segmentation of urban and rural scenes using data volume decomposition coupled with pipeline parallelism. *Automation in Construction*, 133, 103995. <https://doi.org/10.1016/j.autcon.2021.103995>
- Cui, Y., Chen, R., Chu, W., Chen, L., Tian, D., Li, Y., & Cao, D. (2022). Deep Learning for Image and Point Cloud Fusion in Autonomous Driving: A Review. *IEEE Transactions on Intelligent Transportation Systems*, 23(2), 722-739. <https://doi.org/10.1109/TITS.2020.3023541>

- Daneshmand, M., Helmi, A., Avots, E., Noroozi, F., Alisinanoglu, F., Arslan, H. S., ... Anbarjafari, G. (2018). 3D Scanning: A Comprehensive Survey. *arXiv:1801.08863 [cs]*. Repéré à <http://arxiv.org/abs/1801.08863>
- Engelmann, F., Kontogianni, T., Hermans, A., & Leibe, B. (2017). Exploring Spatial Context for 3D Semantic Segmentation of Point Clouds. Dans *2017 IEEE International Conference on Computer Vision Workshops (ICCVW)* (pp. 716-724). Venice : IEEE. <https://doi.org/10.1109/ICCVW.2017.90>
- Franceschini, F., Galetto, M., Maisano, D., & Mastrogiacomo, L. (2014). Large-scale dimensional metrology (LSDM): from tapes and theodolites to multi-sensor systems. *International Journal of Precision Engineering and Manufacturing*, 15(8), 1739-1758. <https://doi.org/10.1007/s12541-014-0527-2>
- Ghamisi, P., Rasti, B., Yokoya, N., Wang, Q., Hofle, B., Bruzzone, L., ... Benediktsson, J. A. (2019). Multisource and Multitemporal Data Fusion in Remote Sensing: A Comprehensive Review of the State of the Art. *IEEE Geoscience and Remote Sensing Magazine*, 7(1), 6-39. <https://doi.org/10.1109/MGRS.2018.2890023>
- Ghamisi, Pedram, Rasti, B., Yokoya, N., Wang, Q., Hofle, B., Bruzzone, L., ... Benediktsson, J. A. (2018). Multisource and Multitemporal Data Fusion in Remote Sensing. *arXiv:1812.08287 [cs, eess, stat]*. Repéré à <http://arxiv.org/abs/1812.08287>
- Grilli, E., Menna, F., & Remondino, F. (2017). A review of point clouds segmentation and classification algorithms. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-2/W3, 339-344. <https://doi.org/10.5194/isprs-archives-XLII-2-W3-339-2017>
- Guo, B., Huang, X., Zhang, F., & Sohn, G. (2015). Classification of airborne laser scanning data using JointBoost. *ISPRS Journal of Photogrammetry and Remote Sensing*, 100, 71-83. <https://doi.org/10.1016/j.isprsjprs.2014.04.015>
- Guo, Y., Wang, H., Hu, Q., Liu, H., Liu, L., & Bennamoun, M. (2021). Deep Learning for 3D Point Clouds: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(12), 4338-4364. <https://doi.org/10.1109/TPAMI.2020.3005434>
- He, E., Chen, Q., Wang, H., & Liu, X. (2017). A curvature based adaptive neighborhood for individual point cloud classification. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-2/W7, 219-225. <https://doi.org/10.5194/isprs-archives-XLII-2-W7-219-2017>

- Hu, Q., Qingyong Hu, Yang, B., Xie, L., Rosa, S., Guo, Y., ... Markham, A. (2019). RandLA-Net: Efficient Semantic Segmentation of Large-Scale Point Clouds. *arXiv: Computer Vision and Pattern Recognition*. <https://doi.org/10.1109/cvpr42600.2020.01112>
- Hu, S.-M., Cai, J.-X., & Lai, Y.-K. (2018). Semantic Labeling and Instance Segmentation of 3D Point Clouds using Patch Context Analysis and Multiscale Processing. *IEEE transactions on visualization and computer graphics*. <https://doi.org/10.1109/TVCG.2018.2889944>
- Hu, Z., Bai, X., Shang, J., Zhang, R., Dong, J., Wang, X., ... Tai, C.-L. (2022, 25 juillet). VMNet: Voxel-Mesh Network for Geodesic-Aware 3D Semantic Segmentation. *arXiv*. Repéré à <http://arxiv.org/abs/2107.13824>
- Huang, H., Kalogerakis, E., Chaudhuri, S., Ceylan, D., Kim, V. G., & Yumer, E. (2017). Learning Local Shape Descriptors from Part Correspondences With Multi-view Convolutional Networks. *arXiv:1706.04496 [cs]*. Repéré à <http://arxiv.org/abs/1706.04496>
- Huang, R., Yang, B., Liang, F., Dai, W., Li, J., Tian, M., & Xu, W. (2018). A top-down strategy for buildings extraction from complex urban scenes using airborne LiDAR point clouds. *Infrared Physics & Technology*, 92, 203-218. <https://doi.org/10.1016/j.infrared.2018.05.021>
- Huber, D., Akinci, B., Tang, P., Adan, A., Okorn, B., & Xiong, X. (2010). Using laser scanners for modeling and analysis in architecture, engineering, and construction. Dans *2010 44th Annual Conference on Information Sciences and Systems (CISS)* (pp. 1-6). <https://doi.org/10.1109/CISS.2010.5464818>
- Jiang, P., Osteen, P., Wigness, M., & Saripalli, S. (2022, 25 mai). RELLIS-3D Dataset: Data, Benchmarks and Analysis. *arXiv*. Repéré à <http://arxiv.org/abs/2011.12954>
- Kelbe, D., Aardt, J. van, Romanczyk, P., Leeuwen, M. van, & Cawse-Nicholson, K. (2017). Multiview Marker-Free Registration of Forest Terrestrial Laser Scanner Data With Embedded Confidence Metrics. *IEEE Transactions on Geoscience and Remote Sensing*, 55(2), 729-741. <https://doi.org/10.1109/TGRS.2016.2614251>
- Kwon, S., Park, J.-W., Moon, D., Jung, S., & Park, H. (2017). Smart Merging Method for Hybrid Point Cloud Data using UAV and LIDAR in Earthwork Construction. Dans *Procedia Engineering* (Vol. 196, pp. 21-28). <https://doi.org/10.1016/j.proeng.2017.07.168>

- Kyle, S., Robson, S., Chapman, D., Cross, P., Oldfield, S., & Iliffe, J. (2001). Understanding large scale metrology - Maths. Harrogate.
- Landrieu, L., & Simonovsky, M. (2017). Large-scale Point Cloud Semantic Segmentation with Superpoint Graphs. *arXiv:1711.09869 [cs]*. Repéré à <http://arxiv.org/abs/1711.09869>
- Le grand dictionnaire terminologique. (2022). Repéré à <http://gdt.oqlf.gouv.qc.ca/index.aspx>
- Li, C., Hang, R., & Rasti, B. (2021). EMFNet: Enhanced Multisource Fusion Network for Land Cover Classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14, 4381-4389. <https://doi.org/10.1109/JSTARS.2021.3073719>
- Li, G., Muller, M., Thabet, A., & Ghanem, B. (2019). DeepGCNs: Can GCNs Go As Deep As CNNs? Dans *2019 IEEE/CVF International Conference on Computer Vision (ICCV)* (pp. 9266-9275). Seoul, Korea (South) : IEEE. <https://doi.org/10.1109/ICCV.2019.00936>
- Li, S., Chen, X., Liu, Y., Dai, D., Stachniss, C., & Gall, J. (2022). Multi-Scale Interaction for Real-Time LiDAR Data Segmentation on an Embedded Platform. *IEEE Robotics and Automation Letters*, 7(2), 738-745. <https://doi.org/10.1109/LRA.2021.3132059>
- Li, Ying, Ma, L., Zhong, Z., Liu, F., Chapman, M. A., Cao, D., & Li, J. (2021). Deep Learning for LiDAR Point Clouds in Autonomous Driving: A Review. *IEEE Transactions on Neural Networks and Learning Systems*, 32(8), 3412-3432. <https://doi.org/10.1109/TNNLS.2020.3015992>
- Li, Yong, Tong, G., Du, X., Yang, X., Zhang, J., & Yang, L. (2019). A Single Point-Based Multilevel Features Fusion and Pyramid Neighborhood Optimization Method for ALS Point Cloud Classification. *Applied Sciences*, 9(5), 951. <https://doi.org/10.3390/app9050951>
- Li, Yuan, Wu, B., & Ge, X. (2019). Structural segmentation and classification of mobile laser scanning point clouds with large variations in point density. *ISPRS Journal of Photogrammetry and Remote Sensing*, 153, 151-165. <https://doi.org/10.1016/j.isprsjprs.2019.05.007>
- Liu, Z., Tang, H., Lin, Y., & Han, S. (2019). Point-Voxel CNN for Efficient 3D Deep Learning. Dans *Advances in Neural Information Processing Systems* (Vol. 32). Curran Associates, Inc. Repéré à <https://proceedings.neurips.cc/paper/2019/hash/5737034557ef5b8c02c0e46513b98f90-Abstract.html>

- Luhmann, T. (2010). Close range photogrammetry for industrial applications. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(6), 558-569. <https://doi.org/10.1016/j.isprsjprs.2010.06.003>
- Matese, A., Toscano, P., Di Gennaro, S. F., Genesio, L., Vaccari, F. P., Primicerio, J., ... Gioli, B. (2015). Intercomparison of UAV, Aircraft and Satellite Remote Sensing Platforms for Precision Viticulture. *Remote Sensing*, 7(3), 2971-2990. <https://doi.org/10.3390/rs70302971>
- Maturana, D., & Scherer, S. (2015). VoxNet: A 3D Convolutional Neural Network for real-time object recognition. Dans *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (pp. 922-928). Hamburg, Germany : IEEE. <https://doi.org/10.1109/IROS.2015.7353481>
- Mirzaei, K., Arashpour, M., Asadi, E., Masoumi, H., Bai, Y., & Behnood, A. (2022). 3D point cloud data processing with machine learning for construction and infrastructure applications: A comprehensive review. *Advanced Engineering Informatics*, 51, 101501. <https://doi.org/10.1016/j.aei.2021.101501>
- MONTLAHUC, J., SHAH, G. A., Polette, A., & Pernot, J.-P. (2019). As-scanned point clouds generation for virtual Reverse Engineering of CAD assembly models. *Computer-Aided Design and Applications*, 16(6), 1171-1182. <https://doi.org/10.14733/cadaps.2019.1171-1182>
- Moon, D., Chung, S., Kwon, S., Seo, J., & Shin, J. (2019). Comparison and utilization of point cloud generated from photogrammetry and laser scanning: 3D world model for smart heavy equipment planning. *Automation in Construction*, Volume 98, 322-331. <https://doi.org/10.1016/j.autcon.2018.07.020>
- Nguyen, A., & Le, B. (2013). 3D point cloud segmentation: A survey. Dans *2013 6th IEEE Conference on Robotics, Automation and Mechatronics (RAM)* (pp. 225-230). Manila, Philippines : IEEE. <https://doi.org/10.1109/RAM.2013.6758588>
- Ostrowski, W., Pilarska, M., Charyton, J., & Bakula, K. (2018). Analysis of 3d building models accuracy based on the airborne laser scanning point clouds. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-2, 797-804. <https://doi.org/10.5194/isprs-archives-XLII-2-797-2018>
- Papon, J., Abramov, A., Schoeler, M., & Worgotter, F. (2013). Voxel Cloud Connectivity Segmentation - Supervoxels for Point Clouds. Dans *2013 IEEE Conference on*



*Computer Vision and Pattern Recognition* (pp. 2027-2034). Portland, OR, USA : IEEE. <https://doi.org/10.1109/CVPR.2013.264>

- Pham, Q.-H., Nguyen, D. T., Hua, B.-S., Roig, G., & Yeung, S.-K. (2019, 5 avril). JSIS3D: Joint Semantic-Instance Segmentation of 3D Point Clouds with Multi-Task Pointwise Networks and Multi-Value Conditional Random Fields. arXiv. Repéré à <http://arxiv.org/abs/1904.00699>
- Qi, C. R., Yi, L., Su, H., & Guibas, L. J. (2017). PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space. *arXiv:1706.02413 [cs]*. Repéré à <http://arxiv.org/abs/1706.02413>
- Qin, N., Hu, X., & Dai, H. (2018). Deep fusion of multi-view and multimodal representation of ALS point cloud for 3D terrain scene recognition. *ISPRS Journal of Photogrammetry and Remote Sensing*, 143, 205-212. <https://doi.org/10.1016/j.isprsjprs.2018.03.011>
- Rahman, M. M., Tan, Y., Xue, J., Shao, L., & Lu, K. (2019). 3D object detection: Learning 3D bounding boxes from scaled down 2D bounding boxes in RGB-D images. *Information Sciences*, 476, 147-158. <https://doi.org/10.1016/j.ins.2018.09.040>
- Rossi, P., Mancini, F., Dubbini, M., Mazzone, F., & Capra, A. (2017). Combining nadir and oblique UAV imagery to reconstruct quarry topography: methodology and feasibility analysis. *European Journal of Remote Sensing*, 50(1), 211-221. <https://doi.org/10.1080/22797254.2017.1313097>
- Roynard, X., Deschaud, J.-E., & Goulette, F. (2018). Paris-Lille-3D: A Point Cloud Dataset for Urban Scene Segmentation and Classification. Dans *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (pp. 2108-21083). Salt Lake City, UT : IEEE. <https://doi.org/10.1109/CVPRW.2018.00272>
- Sareen, K. K., Knopf, G. K., & Canas, R. (2011). Hierarchical data clustering approach for segmenting colored three-dimensional point clouds of building interiors | Request PDF. *ResearchGate*. <http://dx.doi.org/10.1117/1.3599868>
- Schmitt, R. H., Peterek, M., Morse, E., Knapp, W., Galetto, M., Härtig, F., ... Estler, W. T. (2016). Advances in Large-Scale Metrology – Review and future trends. *CIRP Annals*, 65(2), 643-665. <https://doi.org/10.1016/j.cirp.2016.05.002>
- Selvaraju, P., Nabail, M., Loizou, M., Maslioukova, M., Averkiou, M., Andreou, A., ... Kalogerakis, E. (2021). BuildingNet: Learning to Label 3D Buildings. Dans *2021*

*IEEE/CVF International Conference on Computer Vision (ICCV)* (pp. 10377-10387). Montreal, QC, Canada : IEEE. <https://doi.org/10.1109/ICCV48922.2021.01023>

Shi, W., Rangunathan, & Rajkumar. (2020). Point-GNN: Graph Neural Network for 3D Object Detection in a Point Cloud. *arXiv:2003.01251 [cs]*. Repéré à <http://arxiv.org/abs/2003.01251>

Shi, Z., Lin, Y., & Li, H. (2020). Extraction of urban power lines and potential hazard analysis from mobile laser scanning point clouds. *International Journal of Remote Sensing*, 41(9), 3411-3428. <https://doi.org/10.1080/01431161.2019.1701726>

Siddiqui, T. A., Madhok, R., & O'Toole, M. (2020). An Extensible Multi-Sensor Fusion Framework for 3D Imaging. Dans *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (pp. 4344-4353). <https://doi.org/10.1109/CVPRW50498.2020.00512>

Siebke, I., Campana, L., Ramstein, M., Furtwängler, A., Hafner, A., & Lösch, S. (2018). The application of different 3D-scan-systems and photogrammetry at an excavation — A Neolithic dolmen from Switzerland. *Digital Applications in Archaeology and Cultural Heritage*, 10, e00078. <https://doi.org/10.1016/j.daach.2018.e00078>

Skinexigence - Projection de franges. (2023). Repéré à <https://www.skinexigence.com/fr/projection-de-franges>

Summan, R., Pierce, S. G., Macleod, C. N., Dobie, G., Gears, T., Lester, W., ... Smyth, P. (2015). Spatial calibration of large volume photogrammetry based metrology systems. *Measurement*, 68, 189-200. <https://doi.org/10.1016/j.measurement.2015.02.054>

Sun, S., & Salvaggio, C. (2013). Aerial 3D Building Detection and Modeling From Airborne LiDAR Point Clouds. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6(3), 1440-1449. <https://doi.org/10.1109/JSTARS.2013.2251457>

Tang, H., Liu, Z., Zhao, S., Lin, Y., Lin, J., Wang, H., & Han, S. (2020, 13 août). Searching Efficient 3D Architectures with Sparse Point-Voxel Convolution. arXiv. Repéré à <http://arxiv.org/abs/2007.16100>

Tao, A., Duan, Y., Wei, Y., Lu, J., & Zhou, J. (2022). SegGroup: Seg-Level Supervision for 3D Instance and Semantic Segmentation. *IEEE Transactions on Image Processing*, 31, 4952-4965. <https://doi.org/10.1109/TIP.2022.3190709>

- Tchapmi, L. P., Choy, C. B., Armeni, I., Gwak, J., Savarese, S., & Silvio Savarese. (2017). SEGCloud: Semantic Segmentation of 3D Point Clouds. *2017 International Conference on 3D Vision (3DV)*, 537-547. <https://doi.org/10.1109/3dv.2017.00067>
- Te, G., Hu, W., Guo, Z., & Zheng, A. (2018). RGCNN: Regularized Graph CNN for Point Cloud Segmentation. *Proceedings of the 26th ACM international conference on Multimedia*. Repéré à <http://arxiv.org/abs/1806.02952>
- Teruggi, S., Grilli, E., Russo, M., Fassi, F., & Remondino, F. (2020). A Hierarchical Machine Learning Approach for Multi-Level and Multi-Resolution 3D Point Cloud Classification. *Remote Sensing*, *12*(16), 2598. <https://doi.org/10.3390/rs12162598>
- Théodolite Spectra Precision NE-101 pour l'arpentage | Cansel. (2023). Repéré à <https://www.cansel.ca/store/cansel/fr/Industries/G%C3%A9ospatiale-et-ing%C3%A9nierie/Outils-de-mesure/Th%C3%A9odolites-et-Transits/Th%C3%A9odolite-Spectra-Precision-NE-101-pour-l%27arpentage/p/508841>
- Vo, A., Truong-Hong, L., & Laefer, D. F. (2015). Aerial laser scanning and imagery data fusion for road detection in city scale. Dans *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (pp. 4177-4180). <https://doi.org/10.1109/IGARSS.2015.7326746>
- Wang, S., Bai, M., Mattyus, G., Chu, H., Luo, W., Yang, B., ... Urtasun, R. (2016, 1 décembre). TorontoCity: Seeing the World with a Million Eyes. arXiv. Repéré à <http://arxiv.org/abs/1612.00423>
- Wang, Y., Sun, Y., Liu, Z., Sarma, S. E., Bronstein, M. M., & Solomon, J. M. (2019). Dynamic Graph CNN for Learning on Point Clouds. *ACM Transactions on Graphics*, *38*(5), 146:1-146:12. <https://doi.org/10.1145/3326362>
- Wang, Z., Wu, Y., & Niu, Q. (2020). Multi-Sensor Fusion in Automated Driving: A Survey. *IEEE Access*, *8*, 2847-2868. <https://doi.org/10.1109/ACCESS.2019.2962554>
- Weinmann, M., Jutzi, B., Hinz, S., & Mallet, C. (2015). Semantic point cloud interpretation based on optimal neighborhoods, relevant features and efficient classifiers. *ISPRS Journal of Photogrammetry and Remote Sensing*, *105*, 286-304. <https://doi.org/10.1016/j.isprsjprs.2015.01.016>

- Wu, B., Wan, A., Yue, X., & Keutzer, K. (2017). SqueezeSeg: Convolutional Neural Nets with Recurrent CRF for Real-Time Road-Object Segmentation from 3D LiDAR Point Cloud. *arXiv:1710.07368 [cs]*. Repéré à <http://arxiv.org/abs/1710.07368>
- Wu, J., Xu, H., Zheng, Y., Zhang, Y., Lv, B., & Tian, Z. (2019). Automatic Vehicle Classification using Roadside LiDAR Data: *Transportation Research Record*. <https://doi.org/10.1177/0361198119843857>
- Wu, L., Zhu, X., Lawes, R., Dunkerley, D., & Zhang, H. (2019). Comparison of machine learning algorithms for classification of LiDAR points for characterization of canola canopy structure. *International Journal of Remote Sensing*, 40(15), 5973-5991. <https://doi.org/10.1080/01431161.2019.1584929>
- Wu, Y., Wang, Y., Zhang, S., & Ogai, H. (2021). Deep 3D Object Detection Networks Using LiDAR Data: A Review. *IEEE Sensors Journal*, 21(2), 1152-1171. <https://doi.org/10.1109/JSEN.2020.3020626>
- Xia, S., & Wang, R. (2018). Extraction of residential building instances in suburban areas from mobile LiDAR data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 144, 453-468. <https://doi.org/10.1016/j.isprsjprs.2018.08.009>
- Xie, Y., Tian, J., & Zhu, X. X. (2020). Linking Points With Labels in 3D: A Review of Point Cloud Semantic Segmentation. *IEEE Geoscience and Remote Sensing Magazine*, 8(4), 38-59. <https://doi.org/10.1109/MGRS.2019.2937630>
- Xu, Z., & Yang, Z. (2018). Eigenentropy based convolutional neural network based als point clouds classification method. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-3, 2017-2022. <https://doi.org/10.5194/isprs-archives-XLII-3-2017-2018>
- Yan, W. Y., & Shaker, A. (2018). Airborne LiDAR intensity banding: Cause and solution. *ISPRS Journal of Photogrammetry and Remote Sensing*, 142, 301-310. <https://doi.org/10.1016/j.isprsjprs.2018.06.013>
- Yang, X., Zou, H., Kong, X., Huang, T., Liu, Y., Li, W., ... Zhang, H. (2021, 23 septembre). Semantic Segmentation-assisted Scene Completion for LiDAR Point Clouds. arXiv. Repéré à <http://arxiv.org/abs/2109.11453>
- Yang, Z., Jiang, W., Xu, B., Zhu, Q., Jiang, S., & Huang, W. (2017). A Convolutional Neural Network-Based 3D Semantic Labeling Method for ALS Point Clouds. *Remote Sensing*, 9(9), 936. <https://doi.org/10.3390/rs9090936>

- Yang, Z., Tan, B., Pei, H., & Jiang, W. (2018). Segmentation and Multi-Scale Convolutional Neural Network-Based Classification of Airborne Laser Scanner Data. *Sensors (Basel, Switzerland)*, 18(10). <https://doi.org/10.3390/s18103347>
- Ye, C., Acikgoz, S., Pendrigh, S., Riley, E., & DeJong, M. J. (2018). Mapping deformations and inferring movements of masonry arch bridges using point cloud data. *Engineering Structures*, 173, 530-545. <https://doi.org/10.1016/j.engstruct.2018.06.094>
- Yi, L., Kim, V. G., Ceylan, D., Shen, I.-C., Yan, M., Su, H., ... Guibas, L. (2016). A scalable active framework for region annotation in 3D shape collections. *ACM Transactions on Graphics*, 35(6), 1-12. <https://doi.org/10.1145/2980179.2980238>
- Zhang, J., & Lin, X. (2017). Advances in fusion of optical imagery and LiDAR point cloud applied to photogrammetry and remote sensing. *International Journal of Image and Data Fusion*, 8(1), 1-31. <https://doi.org/10.1080/19479832.2016.1160960>
- Zhang, J., Lin, X., & Ning, X. (2013). SVM-Based Classification of Segmented Airborne LiDAR Point Clouds in Urban Areas. *Remote Sensing*, 5(8), 3749-3775. <https://doi.org/10.3390/rs5083749>
- Zhang, Y., Zhou, Z., David, P., Yue, X., Xi, Z., Gong, B., & Foroosh, H. (2020). PolarNet: An Improved Grid Representation for Online LiDAR Point Clouds Semantic Segmentation. Dans *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 9598-9607). Seattle, WA, USA: IEEE. <https://doi.org/10.1109/CVPR42600.2020.00962>
- Zhao, R., Pang, M., & Zhang, Y. (2018). Robust shape extraction for automatically segmenting raw LiDAR data of outdoor scenes. *International Journal of Remote Sensing*, 39(23), 9181-9205. <https://doi.org/10.1080/01431161.2018.1508914>
- Zou, Z., & Li, Y. (2021, 19 septembre). Efficient Urban-scale Point Clouds Segmentation with BEV Projection. arXiv. Repéré à <http://arxiv.org/abs/2109.09074>

## Segmentation sémantique de nuages de points multisources de parcelles et d'ouvrages existants à l'aide de techniques d'apprentissage machine

**Résumé :** Les technologies d'acquisition à distance sont de plus en plus disponibles et performantes. Leur déploiement comme technologie de mesure sans contact est plus fréquent et les possibilités d'utilisation ne font qu'augmenter. Ces technologies, couplées à la puissance de calcul des ordinateurs, rendent possibles l'acquisition et le traitement d'un grand nombre de données. La combinaison des données issues de diverses technologies devient donc envisageable, et même souhaitable pour pallier les limites de chaque technologie. Ce projet de recherche s'inscrit dans cette lignée. Il a pour domaine d'application la numérisation de larges zones géographiques, allant de  $10^2$  m<sup>2</sup> au 10 km<sup>2</sup> avec l'exploitation simultanée de plusieurs technologies d'acquisition. Nos travaux s'inscrivent dans l'optique d'une ouverture des projets de captation numérique pour l'usage de diverses sources complémentaires. Notre principale contribution propose différents modules développés pour traiter et segmenter sémantiquement les nuages de points multisources. Les points avoisinants sont interjumelés et les attributs sont ainsi consolidés, ce qui a permis de segmenter sémantiquement tout en bénéficiant des avantages de chaque technologie. Une étude d'ablation est réalisée dans le but d'évaluer l'intérêt des différents modules de notre proposition. Notre proposition est validée sur des nuages de points multisources mis à disposition par notre partenaire industriel.

**Mots clés :** Segmentation sémantique, Réseaux de neurones artificiels, Lidar aéroporté, Lidar terrestre, Fusion de données, Photogrammétrie

**Abstract :** Remote acquisition technologies are becoming more and more accessible and are no longer reserved for professionals. They are therefore being used more frequently and opportunities to use them are only increasing. These technologies, coupled with the growing computing power of computers, make it possible to acquire and process a large amount of data. Combining data from different technologies is therefore becoming possible. The use of multiple data sources that provide heterogeneous data for the same project is not well covered in the literature for a variety of reasons. This project therefore provides multiple contributions. First, it includes multiple modules that have been developed to process and semantically segment multisources point clouds. Second, the propagation of the attributes acquired by the different sources on the neighboring points and, finally, a simple system to take into account the neighborhood. Our proposed method enabled us to semantically segment multisources point clouds by taking into account the attributes of each source using different modules.

**Keywords :** Semantic Segmentation, Artificial Neural Network, Airborne Lidar, Terrestrial Lidar, Data Fusion, Photogrammetry