UNIVERSITÉ PARIS 8 DES CRÉATIONS

Inspection structurale multi-source et détection d'anomalies basées sur la modélisation 3D et l'analyse optimisée des nuages de points

Thèse de doctorat présentée et soutenue par:

Mohamed Abdelmadjid TALBI

Le 2 Décembre 2025

Directeur de thèse: Prof. Mehdi AMMI – Université Paris 8
Unité de recherche: Laboratoire d'Intelligence Artificielle et Sémantique des Données (LIASD)
École doctorale n° 224: Cognition, langage, interaction

Membres du jury :

Prof. Fakhreddine ABABSA – ENSAM (Rapporteur)
Prof. Rani EL MEOUCHE – ESTP (Rapporteur)
Prof. Florimond GUENIAT – Université de Birmingham (Examinateur)
Dr. Hamdi AMROUN – LG Labs (Examinateur)
Prof. Boubaker DAACHI – Université Paris 8 (Examinateur)

Abstract

The inspection of reinforcing bars (rebars) in reinforced concrete (RC) structures is crucial for ensuring structural safety, durability, and compliance with construction standards such as ACI 117-10. The mechanical integrity of RC depends on the correct placement, spacing, alignment, and overlap of rebars. Traditional inspection methods, based on visual checks and manual measurements, remain limited: they are labor-intensive, time-consuming, hazardous in construction environments, and prone to variability. Non-destructive testing (NDT) technologies such as terrestrial LiDAR, photogrammetry, and Ground Penetrating Radar (GPR) have improved inspection accuracy, yet they are often costly, complex to operate, and difficult to deploy routinely in the field.

In addition to the challenges of data acquisition, the massive point clouds generated by these technologies introduce significant computational bottlenecks. Processing these large-scale datasets for segmentation and analysis requires algorithms that are not only accurate but also computationally efficient to ensure timely feedback on construction sites. Current deep learning models often struggle to balance precision with scalability, limiting their practical deployment in real-time or resource-constrained inspection workflows.

This thesis addresses these limitations by proposing a comprehensive approach for automated rebar inspection, structured around three major contributions. These contributions respectively address the need for a framework capable of reliably exploiting point clouds from diverse sources, the improvement of 3D reconstruction quality, and the optimization of the computational efficiency of point cloud analysis.

The first contribution focuses on developing a complete framework capable of exploiting data from consumer-grade sensors, such as the LiDAR integrated into smartphones, to perform reliable and standardized rebar inspections. This framework integrates two key modules: the *Component Recognition Module* (CRM), based on an adapted RandLA-Net for semantic segmentation of rebar components, and the

Dimensional Assessment Module (DAM), which uses a RANSAC-based geometric fitting procedure to automatically measure rebar spacing. Experiments demonstrate that this solution achieves 82.1% segmentation accuracy and a mean absolute error of 0.9 cm for spacing estimation, showing that low-cost hardware can be effectively leveraged for precise and field-ready inspections that comply with construction standards.

The second contribution concerns the improvement of 3D reconstruction. The use of Neural Radiance Fields (NeRF) as an alternative to photogrammetry is experimented and integrated into the previously introduced framework. The reconstructions produced by NeRF are shown to be denser, more visually homogeneous, and slightly more stable in dimensional evaluation, highlighting the potential of these models as a flexible solution for inspection applications in civil engineering.

The third contribution targets the computational efficiency of point cloud analysis. We propose HENS-Net, a lightweight model designed to optimize semantic segmentation while reducing computational demand. HENS-Net relies on Hilbert curve serialization to replace costly k-NN searches with efficient sequential queries, and incorporates an optimized SEPooling aggregation mechanism as an alternative to computationally expensive attention modules. Experimental results on public datasets (S3DIS, Toronto3D) show that HENS-Net reduces inference time by 11% and model size by 13% compared to RandLA-Net, while achieving comparable or even superior accuracy.

Taken together, these three contributions bridge the gap between advances in 3D vision and the operational requirements of civil engineering, paving the way for the integration of automated inspection frameworks into quality control and construction monitoring workflows.

Keywords: Reinforced Concrete, Rebar Inspection, Civil Engineering, Structural Monitoring, Quality Control, LiDAR, Neural Radiance Fields, 3D Reconstruction, Semantic Segmentation, Hilbert curve, Deep Learning.

Résumé

L'inspection des armatures métalliques dans les structures en béton armé (BA) est cruciale pour garantir la sécurité structurelle, la durabilité et la conformité des ouvrages aux normes en vigueur, telles que l'ACI 117-10. L'intégrité mécanique du béton armé dépend directement de la disposition correcte des barres, de leur espacement, de leur alignement et de leur recouvrement. Les méthodes traditionnelles d'inspection, basées sur des vérifications visuelles et des mesures manuelles, restent limitées : elles sont chronophages, exigeantes en main-d'œuvre, dangereuses dans les environnements de chantier et sujettes à une forte variabilité. Les technologies de contrôle non destructif (CND), telles que le LiDAR terrestre, la photogrammétrie ou le radar à pénétration de sol (GPR), ont permis d'améliorer la précision des inspections. Toutefois, elles demeurent coûteuses, complexes à déployer et difficiles à intégrer dans les pratiques courantes sur le terrain.

Au-delà des défis liés à l'acquisition des données, les nuages de points massifs générés par ces technologies soulèvent d'importants limitations computationnels. Le traitement et l'analyse de ces ensembles de données à grande échelle, notamment pour la segmentation et la reconnaissance automatique, nécessitent des algorithmes capables de concilier précision et efficacité. Les modèles actuels d'apprentissage profond peinent à trouver cet équilibre, ce qui limite leur déploiement pour des inspections rapides et continues sur le terrain. Une meilleure efficacité computationnelle est donc indispensable pour fournir un retour d'information fiable et en temps utile dans les processus de contrôle qualité et de suivi de chantier.

Cette thèse propose une approche exhaustive pour l'inspection automatique des armatures, articulée autour de trois contributions majeures. Celles-ci répondent respectivement à la nécessité de disposer d'un framework capable d'exploiter de manière fiable des nuages de points issus de sources variées, à l'amélioration de la qualité des reconstructions 3D, et à l'optimisation de l'efficience de l'analyse des nuages de points.

La première contribution ne se limite pas à l'utilisation de capteurs grand public,

mais vise à concevoir un framework complet capable de tirer parti de ces capteurs pour réaliser une inspection fiable et standardisée. Le pipeline proposé s'appuie sur deux modules principaux : le Component Recognition Module (CRM), basé sur une version adaptée de RandLA-Net pour la segmentation sémantique des composants d'armature, et le Dimensional Assessment Module (DAM), reposant sur un ajustement géométrique basé sur RANSAC pour mesurer automatiquement l'espacement entre les barres. Les expérimentations montrent que ce système atteint une précision de segmentation de 82,1% et une erreur absolue moyenne de 0,9 cm, prouvant qu'un matériel peu coûteux peut être exploité efficacement pour des inspections précises conformes aux normes de chantier.

La deuxième contribution porte sur l'amélioration de la reconstruction 3D. L'usage des Neural Radiance Fields (NeRF), en alternative à la photogrammétrie, est exploré et intégré au framework précédent. Les reconstructions produites par NeRF se révèlent plus denses, visuellement plus homogènes et légèrement plus stables en évaluation dimensionnelle, ce qui souligne le potentiel de ces modèles comme solution flexible pour des applications d'inspection en génie civil.

La troisième contribution vise l'efficience computationnelle de l'analyse des nuages de points. Nous proposons HENS-Net, un modèle léger conçu pour optimiser la segmentation sémantique tout en réduisant les besoins en calcul. HENS-Net s'appuie sur une sérialisation par courbes de Hilbert pour remplacer les recherches k-NN coûteuses par des requêtes séquentielles efficiente, et intègre un mécanisme d'agrégation SEPooling optimisé, en alternative aux modules attentionnels coûteux. Les résultats expérimentaux sur des jeux de données publics (S3DIS, Toronto3D) montrent que HENS-Net réduit le temps d'inférence de 11% et la taille du modèle de 13% par rapport à RandLA-Net, tout en atteignant des performances comparables, voire supérieures en terme de précisions.

En réunissant ces trois contributions, cette thèse rapproche les avancées scientifiques en vision 3D des besoins opérationnels du génie civil, ouvrant la voie à l'intégration de frameworks d'inspection automatique dans les processus de contrôle qualité et de suivi de chantier.

Mots-clés : Béton Armé, Inspection des Armatures, Génie Civil, Suivi Structurel, Contrôle de la Qualité, LiDAR, Champs de Radiance Neuronaux, Reconstruction 3D, Segmentation Sémantique, Courbe de Hilbert, Apprentissage Profond.

Contents

Abstract **i**

Résumé **iii**

Contents **v**

List of Figures **ix**

List of Tables **xi**

Nomenclature **xiii**

1 Introduction **1**

1.1 Contexte général 1
1.2 Objectifs de la thèse 3
1.3 Questions de recherche 4
1.4 Contributions principales 5
1.5 Organisation du manuscrit 6

2 Etat de l'art **9**

2.1 Introduction 9
2.2 Inspection des structures en béton armé 9
2.2.1 Enjeux et paramètres critiques 10
2.2.2 Méthodes traditionnelles et limites 10
2.2.3 Vers les approches numériques et automatisées 10
2.3 Technologies de collecte de données pour l'inspection du béton armé 11
2.3.1 Vision par ordinateur et capteurs RGB/RGB-D 11
2.3.2 Technologies LiDAR 12
2.3.3 Radar à pénétration de sol (GPR) 14

2.4 Méthodes d'analyse et de traitement pour l'inspection des armatures . . . 15
2.4.1 Méthodes basées sur des règles et la géométrie computationnelle . . . 15
2.4.2 Méthodes exploitant les approches d'apprentissage profond . . . 16
2.5 Paradigmes de la reconstruction 3D . . . 17
2.6 Paradigmes de segmentation des nuages de points 3D . . . 19
2.6.1 Méthodes basées sur la projection . . . 19
2.6.2 Méthodes basées sur les voxels . . . 19
2.6.3 Méthodes basées sur les points . . . 20
2.6.4 Méthodes de sérialisation . . . 21
2.7 Discussion . . . 22

3 Framework d'inspection automatique des armatures à partir de nuages de points 25
3.1 Introduction . . . 25
3.2 Framework proposé . . . 27
3.2.1 Vue d'ensemble du framework . . . 27
3.2.2 Prétraitement des données . . . 28
3.2.3 Module de Reconnaissance de Composants (CRM) . . . 30
3.2.4 Module d'Évaluation Dimensionnelle (DAM) . . . 31
3.3 Expérimentations . . . 32
3.3.1 Jeu de données . . . 35
3.3.2 Configuration expérimentale . . . 35
3.3.3 Résultats . . . 36
3.4 Conclusion . . . 38

4 Expérimentation des champs de radiance neuronaux pour l'inspection automatique des armatures 41
4.1 Introduction . . . 41
4.2 Méthodologie . . . 43
4.2.1 Vue d'ensemble du framework . . . 44
4.2.2 Acquisition des données . . . 44
4.2.3 Prétraitement des données . . . 47
4.2.4 Module de Reconnaissance de Composants (CRM) . . . 50

4.2.5 Module d'Évaluation Dimensionnelle (DAM) 52
4.3 Expérimentations 55
4.3.1 Jeu de données 55
4.3.2 Configuration expérimentale 55
4.3.3 Résultats et discussion 57
4.4 Conclusion 63

5 Architecture HENS-Net pour la segmentation sémantique efficiente des nuages de points 65
5.1 Introduction 65
5.2 HENS-Net 68
5.2.1 Sérialisation Hilbert 68
5.2.2 Sélection de voisinage 70
5.2.3 Agrégation de caractéristiques 71
5.2.4 Fonction de perte 75
5.3 Expérimentations 76
5.3.1 Jeux de données 76
5.3.2 Configuration expérimentale 77
5.3.3 Segmentation sémantique sur S3DIS 77
5.3.4 Segmentation sémantique sur Toronto3D 78
5.3.5 Efficience de HENS-Net 78
5.4 Conclusion 79

6 Discussion 83

7 Conclusion 87

Bibliographie 91

List of Figures

1.1 Synthèse des contributions. 6

2.1 Approche proposée par [15] pour l'évaluation probabiliste de conduites en béton armé. 13
2.2 Histogrammes de comparaison nuage-à-nuage (C2C) entre les nuages de points smartphone et le nuage de points de référence [12]. 13
2.3 Illustration du mécanisme d'identification de barres proposé par [14]. 14
2.4 Vue d'ensemble de la méthode de calcul de l'espacement entre barres adjacentes proposée par [64]. 17
2.5 Nuages de points photogrammétriques vs NeRF [13]. 18
2.6 Illustration de l'architecture VoxelNet [76]. 20
2.7 Architecture du modèle PointTransformerV3 proposée par [63]. 21
2.8 Illustration de la sérialisation d'un nuage de points [63]. 21

3.1 Procédure globale du framework proposé pour l'évaluation automatique de l'espacement des armatures à partir de données LiDAR iPhone. 28
3.2 Dispositif expérimental utilisé pour la collecte des données. 36
3.3 Comparaison visuelle entre la vérité terrain (a) et la segmentation prédite (b), illustrant la performance du module CRM. 38

4.1 Illustration des paramètres critiques dans une structure en BA. 42
(a) espacement des armatures 42
(b) enrobage 42
(c) chevauchement des nappes 42
4.2 Principe général de l'entraînement NeRF, adapté de la version originale proposée par Mildenhall et al. [33]. 43
4.3 Architecture générale du framework proposé pour l'évaluation automatique de l'espacement des armatures. 45
4.4 Pipeline de reconstruction 3D basé sur NeRF. 48
4.5 Pipeline de reconstruction 3D basé sur la photogrammétrie. 49

4.6 Architecture du modèle RandLA-Net utilisé dans le module CRM. N représente le nombre de points. FC : couche entièrement connectée, LFA : agrégation locale de caractéristiques, RS : échantillonnage aléatoire, MLP : perceptron multicouche partagé, US : sur-échantillonnage. 51
4.7 Pipeline du module DAM pour l'estimation automatique de l'espacement des armatures. 53
4.8 Configurations expérimentales. 56
(a) grande structure 56
(b) petite structure 56
4.9 Comparaison des nappes d'armatures : nuages NeRF vs. Photogrammétrie. 58
4.10 Comparaison de la surface de maquette : nuages NeRF vs. Photogrammétrie. 58
4.11 Temps vSLAM en fonction de la surface inspectée. 59

5.1 Sérialisation de nuages de points avec la courbe de Hilbert à différentes résolutions. 67
5.2 Vue d'ensemble de l'architecture de HENS-Net. Le réseau est organisé en plusieurs étapes (N), chacune reposant sur des couches d'agrégation intégrant un encodage spatial et un module de SEPooling. MLP désigne un perceptron multicouche partagé, et FC une couche entièrement connectée. 69
5.3 Visualisation de l'ordonnancement des points sur trois objets exemples : chaise, table et sofa. La rangée supérieure montre les nuages de points bruts (non ordonnés), où des points consécutifs peuvent être spatialement éloignés, formant des voisinages fragmentés et désorganisés. La rangée inférieure illustre les nuages de points après sérialisation Hilbert, où les points consécutifs restent spatialement cohérents, constituant des voisinages compacts et significatifs. 72
5.4 Structure de la couche d'agrégation de caractéristiques. MLP désigne un Multi-Layer Perceptron partagé, SE correspond au module de *Spatial Encoding*, et $\sum$ représente l'opérateur de somme. 73

List of Tables

3.1 Répartition des points par classe dans le jeu de données 35
3.2 Résultats globaux de l'évaluation du framework proposé 37
3.3 Résultats détaillés par classe pour la segmentation sémantique 38

4.1 Analyse de la distribution des classes : NeRF vs. Photogrammétrie . 55
4.2 Répartition des temps de traitement en fonction de la surface pour la photogrammétrie (Scaniverse) et la reconstruction NeRF. 59
4.3 Comparaison de la précision et de l'IoU : Photogrammétrie vs. NeRF 60
4.4 Comparaison de l'évaluation de l'espacement des armatures selon différentes approches 61

5.1 Résultats de segmentation sémantique sur S3DIS (Area 5). 78
5.2 Résultats de segmentation sémantique sur S3DIS (validation croisée 6 plis). 79
5.3 Résultats de segmentation sémantique sur Toronto3D. 79
5.4 Comparaison du temps de traitement et du nombre de paramètres entre HENS-Net et RandLA-Net [22] sur un sous-ensemble de 5 scènes de S3DIS [22], zone 5. 80

Nomenclature

Symbole	**Description**	**Unité SI**
x, y, z	coordonnées spatiales d'un point	m
r, g, b	intensités de couleur rouge, vert, bleu	–
h	indice de Hilbert associé à un point	–
N	nombre total de points dans un nuage	–
k	nombre de voisins utilisés (k-NN)	–
$d_{i,k}$	distance euclidienne entre p_i et son $k^{\text{ème}}$ voisin	m
$\mathbf{\Delta p}_{i,k}$	vecteur de décalage entre p_i et $p_{i,k}$	m
σ	densité volumique dans NeRF	–
(θ, ϕ)	angles de direction de vue	rad
$\mathbf{f}_i$	vecteur de caractéristiques associé au point p_i	–
$\mathbf{s}$	poids d'attention (SEPooling)	–
α_i	poids associé à la classe i (WCE)	–

Abréviation	**Description**	**Unité SI**
CRM	*Component Recognition Module*	–
DAM	*Dimensional Assessment Module*	–
NeRF	Neural Radiance Fields	–
SfM	Structure-from-Motion	–
MVS	Multi-View Stereo	–
LiDAR	Light Detection and Ranging	–
mIoU	*mean Intersection over Union*	%
OA	*Overall Accuracy*	%
MAE	*Mean Absolute Error*	m ou cm
MSE	*Mean Squared Error*	m^2 ou cm^2
SD	*Standard Deviation*	m ou cm

1 | Introduction

L'objectif de ce chapitre est de poser les bases générales de la recherche présentée dans cette thèse. Nous y introduisons le contexte et les enjeux liés à l'inspection des structures en béton armé, en mettant en évidence les limites des approches traditionnelles et les opportunités offertes par l'intelligence artificielle et les technologies 3D. Nous formulons ensuite la problématique centrale et les motivations qui guident nos travaux, avant de préciser les objectifs, les questions de recherche et les contributions principales de la thèse. Enfin, nous présentons l'organisation du manuscrit, afin de donner une vue d'ensemble du cheminement adopté.

1.1. Contexte général

L'inspection des structures en béton armé constitue un enjeu majeur pour la sécurité et la durabilité des infrastructures. Qu'il s'agisse de bâtiments résidentiels, d'ouvrages d'art tels que les ponts, ou encore d'infrastructures stratégiques comme les tunnels et les barrages, ces ouvrages sont essentiels à notre environnement bâti. Leur qualité dépend en grande partie de la mise en place correcte des armatures, chargées d'assurer la résistance mécanique et la stabilité globale de la structure. Des défauts tels qu'un espacement irrégulier, un recouvrement insuffisant ou une superposition mal contrôlée des barres peuvent compromettre dès l'origine la performance de l'ouvrage et conduire, à terme, à des défaillances graves, aux conséquences humaines et économiques considérables [27, 71].

Traditionnellement, l'évaluation de la qualité des armatures repose sur l'inspection visuelle manuelle. Des experts se déplacent sur site pour examiner l'état des surfaces et effectuer des mesures ponctuelles. Bien que cette pratique reste largement utilisée, elle présente de nombreuses limites : dépendance au jugement subjectif de l'opérateur, variabilité inter-experts [31], coûts et durée élevés, ainsi que difficulté d'accès à certaines zones. La couverture spatiale demeure souvent incomplète, laissant subsister des défauts non détectés susceptibles d'évoluer silencieusement [14].

Ces limites ont stimulé l'émergence de solutions reposant sur l'intelligence artificielle (IA), et plus particulièrement sur la vision par ordinateur et l'apprentissage profond. L'essor de capteurs tridimensionnels, tels que les scanners LiDAR, les caméras RGB-D, permet aujourd'hui de générer des nuages de points denses et précis. Ces données tridimensionnelles offrent une représentation géométrique riche des structures, ouvrant la voie à une inspection plus rapide, reproductible et moins dépendante de l'expertise humaine [56, 64]. Représentées sous forme de nuages de points (ensembles de points définis par leurs coordonnées (x, y, z) éventuellement enrichis d'attributs comme la couleur, l'intensité ou les normales), offrent une description géométrique explicite de la scène. Contrairement aux images 2D, qui ne fournissent qu'une projection plane, le nuage de points capture directement la métrique, la profondeur et la topologie des surfaces, permettant ainsi de mesurer des distances et identifier des éléments structuraux avec une précision spatiale fine. Cette richesse informationnelle en fait un support privilégié pour les tâches d'inspection structurelle, où les déviations géométriques (espacement, alignement, recouvrement) doivent être quantifiées de manière absolue et non simplement interprétées qualitativement.

Couplés à des algorithmes de segmentation sémantique et de reconnaissance automatique, ces dispositifs constituent une base prometteuse pour le contrôle qualité en génie civil.

Plus récemment, l'intégration de capteurs 3D dans des dispositifs grand public, tels que les smartphones équipés de LiDAR, a profondément transformé l'écosystème technologique. Ces capteurs, accessibles et peu coûteux, offrent de nouvelles perspectives pour démocratiser l'acquisition de données 3D. Des études ont déjà confirmé leur potentiel dans des contextes variés, allant de l'archéologie au suivi industriel [12, 40]. Cependant, leur précision reste inférieure à celle des capteurs industriels, en raison du bruit et de la présence des imprécisions géométriques qu'ils introduisent [41]. Ces contraintes imposent le développement de méthodes robustes capables de compenser ces limites, tout en exploitant l'opportunité unique de rendre l'inspection automatisée plus largement accessible.

Au-delà de l'acquisition, l'exploitation de nuages de points massifs soulève un défi computationnel. Les algorithmes de segmentation et d'analyse actuels, bien qu'efficaces, sont souvent coûteux en mémoire et en temps de calcul, ce qui limite leur déploiement sur le terrain. Il devient donc nécessaire de concevoir des architectures plus légères et scalables, capables de traiter efficacement de larges volumes de données sans compromettre la précision géométrique.

Actuellement, l'inspection des structures par IA se situe à la croisée de deux dynamiques complémentaires : d'une part, la nécessité croissante d'automatiser et d'objectiver le diagnostic des structures ; d'autre part, la volonté de rendre ces technologies accessibles grâce à des solutions accessibles, rapides et généralisables. Ce double enjeu, scientifique et économiques, constitue le socle sur lequel s'appuie la présente thèse.

1.2. Objectifs de la thèse

L'objectif principal de cette thèse est de contribuer à l'automatisation et à la démocratisation de l'inspection des structures en béton armé en s'appuyant sur les avancées récentes en vision par ordinateur et en apprentissage profond appliquées aux nuages de points tridimensionnels. L'ambition est de développer des approches capables de rendre l'inspection plus accessible, plus efficace et plus rapide, tout en répondant aux contraintes pratiques et économiques du génie civil, ainsi qu'aux exigences de conformité définies par des normes telles que l'ACI 117-10 [2].

Un premier objectif consiste à définir un **framework** reproductible permettant d'exploiter des nuages de points issus de capteurs à faible coût, en particulier les LiDAR intégrés aux smartphones, dans un contexte d'inspection structurelle tout en garantissant une précision compatible avec les exigences de conformité. L'enjeu n'est pas uniquement de comparer ces capteurs aux scanners industriels tels que les TLS (Terrestrial Laser Scanners), dont le coût et les contraintes d'exploitation limitent l'usage sur chantier, mais de déterminer dans quelles conditions ils peuvent être intégrés dans un processus d'inspection fiable et opérationnel.

Un deuxième objectif est d'examiner le potentiel des modèles de reconstruction de type **Neural Radiance Fields (NeRF)** afin de dépasser certaines limites observées dans les approches photogrammétriques, notamment en termes de complétude et de stabilité visuelle. L'objectif est d'évaluer la pertinence de ces méthodes récentes pour produire des reconstructions 3D exploitables dans des tâches de contrôle qualité.

Enfin, un troisième objectif concerne l'efficience de l'analyse des nuages de points. Les approches actuelles de segmentation 3D, qu'elles reposent sur la projection, la voxelisation ou le traitement direct des nuages de points, souffrent de limitations bien identifiées telles que l'explosion mémoire ou la complexité des recherches de voisinage. Cette thèse vise à concevoir une **architecture** de segmentation plus légère, exploitant des représentations compactes et des mécanismes d'agrégation

optimisés, de manière à rendre l'analyse 3D plus efficiente sans sacrifier la précision.

1.3. Questions de recherche

Les objectifs identifiés soulèvent une série de questions de recherche structurantes, ces questions portent à la fois sur la performance algorithmique, l'accessibilité des outils, la qualité des reconstructions 3D et l'efficacité/efficience de l'analyse de nuages de points.

La première question consiste à déterminer dans quelles conditions un framework d'inspection automatique peut être considéré comme fiable et reproductible dans un contexte de contrôle qualité des armatures, y compris lorsqu'il s'appuie sur des nuages de points de moindre qualité issus de capteurs à faible coût tels que le LiDAR embarqué sur smartphone. Il s'agit également d'évaluer dans quelle mesure ce framework demeure conforme aux tolérances imposées par des normes comme l'ACI 117-10 [2] et présente un niveau de robustesse suffisant pour constituer une alternative aux dispositifs plus coûteux.

Les approches traditionnelles telles que la photogrammétrie et le LiDAR professionnel présentent des limites distinctes : la photogrammétrie souffre d'un manque de réalisme visuel et d'une densité irrégulière, tandis que le LiDAR industriel implique un coût opérationnel élevé malgré sa précision métrique. La deuxième question consiste à examiner dans quelles conditions les modèles NeRF peuvent dépasser ces limitations et produire des reconstructions 3D exploitables pour l'évaluation des armatures dans un processus de contrôle qualité automatisé, au point de constituer une alternative crédible aux méthodes classiques.

Les méthodes actuelles de segmentation sémantique imposent un compromis marqué entre précision et coût computationnel : les approches par projection entraînent une perte de fidélité géométrique, les méthodes voxel-based requièrent une mémoire importante, et la plupart des modèles basés point présentent une complexité élevée dans les recherches de voisinage. La troisième question consiste à déterminer dans quelles conditions il est possible de concevoir une architecture légère et scalable, fondée sur des représentations compactes, capable de maintenir un niveau de précision élevé tout en réduisant de manière significative les besoins en calcul et les temps de traitement.

Ces trois questions structurent directement les contributions de cette thèse, qui visent respectivement : la définition d'un framework d'inspection 3D applicable à

des données nuages de points multi-source, l'évaluation du potentiel des approches de reconstruction NeRF pour le contrôle qualité des armatures, et la conception d'une architecture de segmentation efficace et efficiente.

1.4. Contributions principales

Afin de répondre aux questions de recherche identifiées, cette thèse propose trois contributions majeurs, chacun visant à lever une limitation spécifique mise en évidence dans l'état de l'art.

La première contribution introduit un framework de bout en bout pour l'inspection automatique des armatures, structuré autour de deux modules : le Component Recognition Module (CRM), dédié à la segmentation sémantique et à l'identification des composants, et le Dimensional Assessment Module (DAM), dédié à l'évaluation des espacements. Ce framework est appliqué à des nuages de points issus de différentes modalités d'acquisition et évalué en particulier sur des données acquises avec le LiDAR embarqué sur iPhone, afin de démontrer sa capacité à rester exploitable sur des capteurs grand public, rendant ainsi l'inspection plus accessible tout en maintenant une conformité aux tolérances prescrites par des normes telles que l'ACI 117-10 [2].

La deuxième contribution consiste à expérimenter l'usage des Neural Radiance Fields (NeRF) comme alternative de reconstruction 3D dans le cadre de l'inspection des armatures. Les reconstructions produites par NeRF sont systématiquement comparées à celles obtenues par des approches classiques, en particulier la photogrammétrie, reconnue pour ses limites en termes de complétude et de stabilité visuelle. Cette comparaison est réalisée au sein du même framework d'inspection défini dans la première contribution, ce qui permet d'évaluer dans quelles conditions NeRF produit des modèles plus cohérents et plus exploitables pour des tâches de contrôle qualité.

La troisième contribution introduit un modèle de segmentation sémantique 3D baptisé HENS-Net, reposant sur une stratégie de sérialisation Hilbert pour la sélection des voisinages et sur un mécanisme d'agrégation optimisé. Cette architecture vise à concilier précision et efficience computationnelle en réduisant le coût des recherches de voisinage, qui constitue l'un des principaux limitations des approches traditionnelles. Validée sur plusieurs jeux de données publics, tels que S3DIS et Toronto3D, HENS-Net démontre des performances compétitives par rapport aux méthodes existantes tout en réduisant significativement les besoins en calcul et le temps de

traitement.

En résumé, ces trois contributions répondent de manière cohérente aux défis identifiés : la mise en place d'un framework d'inspection applicable à des données issue de capteur grand public, l'intégration de modèles de reconstruction NeRF pour améliorer la qualité des reconstruction 3D, et la conception d'une architecture de segmentation 3D alliant précision et efficience computationnelle. Ces contributions constituent l'apport scientifique principal de cette thèse et sont détaillées dans les chapitres suivants.

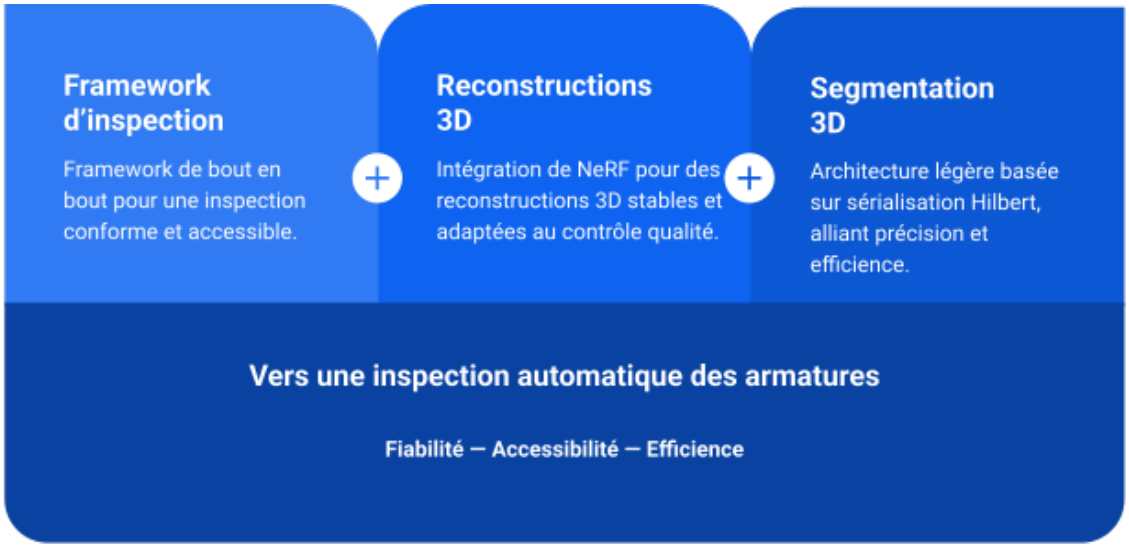


Figure 1.1: Synthèse des contributions.

1.5. Organisation du manuscrit

Le manuscrit est structuré en sept chapitres, organisés de manière à conduire progressivement le lecteur du contexte général vers les contributions spécifiques de cette recherche, puis vers leur discussion et leur conclusion.

Le **chapitre 1** introduit le contexte général de l'inspection des structures en béton armé et les enjeux liés à l'utilisation de la vision par ordinateur et de l'apprentissage profond dans ce domaine. Il expose la problématique, les objectifs, les questions de recherche et les contributions principales de la thèse.

Le **chapitre 2** est consacré à l'état de l'art. Il présente d'abord les enjeux et paramètres critiques de l'inspection des structures en béton armé, puis les méthodes traditionnelles et leurs limites. Il décrit ensuite les technologies de collecte de don-

nées, ainsi que les méthodes d'analyse et de traitement, des approches géométriques classiques aux techniques d'apprentissage profond. Une section spécifique est dédiée aux NeRFs, avant d'aborder les paradigmes de segmentation des nuages de points 3D et une conclusion ouvrant sur les contributions.

Le **chapitre 3** présente la première contribution de la thèse, à savoir un framework d'inspection automatique des armatures à partir de nuages de points. Il expose le contexte, la méthodologie, les expérimentations réalisées, ainsi que les résultats obtenus et leur analyse.

Le **chapitre 4** introduit la deuxième contribution, consacrée à l'expérimentation du NeRF pour la reconstruction et l'inspection des armatures. Après un rappel des fondements de NeRF, le chapitre présente la méthodologie adoptée, les expérimentations comparatives avec la photogrammétrie, ainsi que les résultats obtenus.

Le **chapitre 5** est dédié à la troisième contribution, à savoir la présentation de HENS-Net, un modèle de segmentation sémantique 3D reposant sur la sérialisation Hilbert et un schéma d'agrégation optimisé. Ce chapitre décrit l'architecture proposée, la méthodologie d'évaluation, les expérimentations menées sur des jeux de données publics, ainsi que les analyses comparatives associées.

Le **chapitre 6** propose une discussion générale croisant les trois contributions. Il met en évidence leurs complémentarités, analyse les limites globales de la recherche et ouvre sur des perspectives de travaux futurs.

Enfin, le **chapitre 7** conclut la thèse en synthétisant les principaux apports, en rappelant leur positionnement par rapport à l'état de l'art, et en soulignant les retombées scientifiques et applicatives de cette thèse.

2 | Etat de l'art

2.1. Introduction

Ce chapitre présente un panorama synthétique et critique des approches existantes pour l'inspection des structures en béton armé. Il s'articule comme suit : (1) un rappel des enjeux métier et des paramètres critiques à contrôler, ainsi que les limites des pratiques traditionnelles ; (2) un tour d'horizon des technologies de collecte illustrées par des exemples issus de la littérature ; (3) les méthodes d'analyse et de traitement pour l'inspection, des approches géométriques classiques aux méthodes exploitant l'apprentissage profond ; (4) la reconstruction 3D à partir d'images, en situant la photogrammétrie face aux approches neuronales récentes, avec un focus sur NeRF pour discuter apports et limites ; (5) les paradigmes de segmentation des nuages de points (projection, voxelisation, traitement direct des points, sérialisation). Le chapitre se conclut par une discussion transversale positionnant ces solutions selon l'accessibilité/coût, la qualité géométrique et visuelle, et l'efficacité computationnelle, afin de dégager les limites scientifiques et besoins opérationnels auxquels répondront les contributions de la thèse.

2.2. Inspection des structures en béton armé

L'inspection des structures en béton armé (RC) constitue un enjeu fondamental dans le domaine du génie civil, car elle conditionne la durabilité, la sécurité et la conformité des ouvrages. En effet, les armatures (barres de renfort en acier) jouent un rôle critique dans la résistance mécanique et la ductilité du béton, et leur mauvaise disposition peut avoir des conséquences graves sur la performance structurelle. Cette section présente d'abord les enjeux et paramètres clés à contrôler lors de l'inspection des structures RC, puis les méthodes traditionnelles historiquement utilisées et leurs limites, afin de mettre en évidence la nécessité de solutions plus automatisées et numériques.

2.2.1. Enjeux et paramètres critiques

La qualité d'une structure RC repose sur le respect de paramètres géométriques précis liés aux armatures. Parmi eux, l'espacement entre barres, le recouvrement des nappes, l'enrobage (distance entre la surface externe du béton et la barre la plus proche) et l'alignement constituent des critères déterminants. Un mauvais espacement peut provoquer des concentrations de contraintes, tandis qu'un enrobage insuffisant expose les barres à la corrosion, réduisant ainsi la durabilité de la structure [53, 57]. De plus, des tolérances dimensionnelles trop larges ou non respectées entraînent des écarts entre la conception et la réalisation, ce qui compromet la conformité aux normes de qualité et de sécurité, telles que celles fixées par l'ACI 117-10 [2]. Ces constats expliquent pourquoi les agences de construction et de contrôle qualité accordent une importance cruciale à l'inspection régulière et fiable des armatures avant le bétonnage.

2.2.2. Méthodes traditionnelles et limites

Historiquement, l'inspection des armatures s'est appuyée sur des méthodes manuelles simples, telles que la mesure au ruban ou la vérification visuelle par les ingénieurs de chantier. Ces approches, encore très répandues aujourd'hui, souffrent de plusieurs limitations : elles sont chronophages, dépendent fortement de l'expérience de l'inspecteur, et ne garantissent pas une précision suffisante, en particulier dans des configurations complexes comme les cages d'armatures multi-couches [67]. De plus, elles ne permettent pas d'obtenir un enregistrement numérique exploitable a posteriori, ce qui complique la traçabilité des inspections et la détection d'éventuelles anomalies ultérieures. Dans un contexte où les délais de construction sont de plus en plus contraints, ces méthodes s'avèrent peu adaptées et risquent d'introduire des erreurs de conformité non détectées.

2.2.3. Vers les approches numériques et automatisées

Pour pallier ces limites, de nombreux travaux se sont orientés vers des approches instrumentées ou automatisées. Par exemple, l'intégration de la laser-scanning [10] et du LiDAR portatif [69, 71] a montré qu'il était possible de capturer la disposition réelle des barres et d'évaluer leur conformité aux plans de conception. De même, des approches hybrides combinant vision par ordinateur et réalité augmentée permettent d'assister les inspecteurs en visualisant directement les écarts et en générant des

instructions de correction [10].

Cependant, malgré leur potentiel, ces solutions demeurent coûteuses et nécessitent une expertise technique avancée pour être déployées efficacement sur les chantiers. Elles ne répondent donc pas pleinement aux besoins opérationnels en termes d'accessibilité, de rapidité et de facilité d'utilisation.

Dans les sections suivantes, nous détaillerons d'abord les technologies de collecte de données pour l'inspection du béton armé, en présentant leurs principes, avantages et limites. Nous aborderons ensuite les méthodes d'analyse et de traitement pour l'inspection des armatures, en explorant les algorithmes et modèles permettant d'exploiter efficacement les données acquises pour le contrôle qualité et la détection des non-conformités.

2.3. Technologies de collecte de données pour l'inspection du béton armé

La réussite de toute inspection des structures en béton armé dépend en grande partie de la qualité et de la fiabilité des données acquises. Différentes technologies de collecte permettent de générer des représentations 2D ou 3D exploitables pour analyser la disposition des armatures, détecter des défauts, et vérifier la conformité aux plans. Dans cette section, nous présentons quatre grandes familles de technologies : la vision par ordinateur, le LiDAR, le radar à pénétration de sol (GPR), et la photogrammétrie. Chacune de ces approches offre des avantages distincts mais présente également des limites, qui motivent le recours à des méthodes hybrides et à de nouvelles solutions comme les champs de radiance neuronaux (NeRF).

2.3.1. Vision par ordinateur et capteurs RGB/RGB-D

Les méthodes basées sur la vision exploitent des images RGB ou RGB-D pour extraire automatiquement des informations relatives aux armatures. Grâce aux progrès récents en apprentissage profond, la vision par ordinateur est devenue une approche incontournable pour la détection et la classification des éléments de construction.

Par exemple, Xi et al. [64] ont proposé une méthode de contrôle automatisé de l'espacement des distanciers d'armature, combinant deep learning (Faster R-CNN) et géométrie computationnelle. Leur système intègre également un module basé sur l'estimation de pose pour détecter les points clés nécessaires aux calculs de distances,

démontrant l'efficacité de la vision pour des inspections précises et automatisées. D'autres études s'appuient sur la détection d'objets (YOLO, Faster-RCNN) pour identifier des défauts de surface ou réaliser un comptage automatique des barres via drones [32, 55]. Ces approches offrent rapidité et flexibilité, mais elles souffrent de limitations liées aux conditions d'éclairage, aux occlusions, et à l'absence d'information géométrique complète, ce qui rend parfois nécessaire la combinaison avec des données 3D.

2.3.2. Technologies LiDAR

Le LiDAR est largement utilisé pour capturer des nuages de points denses et précis, permettant une modélisation 3D directe des structures RC. Des travaux ont montré son intérêt tant dans des contextes professionnels (TLS, scanners mobiles) que dans des contextes plus abordables grâce aux capteurs intégrés dans les smartphones.

Yuan et al. [69–72] ont proposé plusieurs frameworks exploitant des scanners LiDAR portatifs ou mobiles pour évaluer la qualité des cages d'armature en chantier. Leur approche consiste à extraire automatiquement la position des barres et à comparer le résultat avec les plans pour générer un indice de qualité (RLQI), réduisant ainsi la subjectivité des inspections manuelles. À l'échelle industrielle, Calvi et al. [7] ont appliqué la technologie TLS pour analyser les dommages d'un bâtiment RC soumis à des essais sur table vibrante, en caractérisant les déplacements résiduels et les fissures du béton. De même, Ebrahimi et al. [15] ont utilisé le LiDAR pour l'évaluation probabiliste de conduites en béton armé, illustrant la polyvalence de cette technologie dans différents contextes structurels.

Enfin, l'émergence de solutions grand public a ouvert la voie à des approches accessibles : Rutkowski et al. [40] ainsi que Costantino et al. [12] ont démontré l'utilité du LiDAR de l'iPhone dans des tâches comme l'inspection de puits de mines, l'archéologie ou l'urbanisme. Bien que la précision métrique reste inférieure à celle des scanners professionnels (erreurs de l'ordre de 1–3 cm), ces dispositifs présentent un excellent compromis coût–accessibilité, qui motive leur intégration dans des cadres de QC/QA abordables, même si leur utilisation n'a pas encore été explorée dans le contexte spécifique de l'inspection des structures en béton armé.

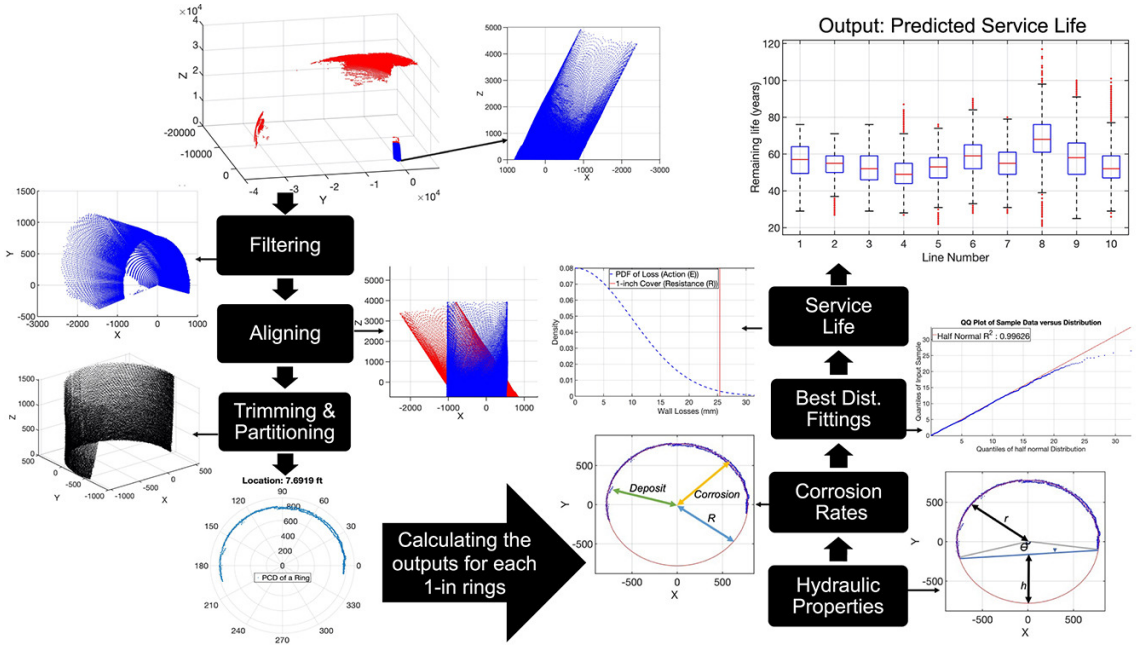


Figure 2.1: Approche proposée par [15] pour l'évaluation probabiliste de conduites en béton armé.

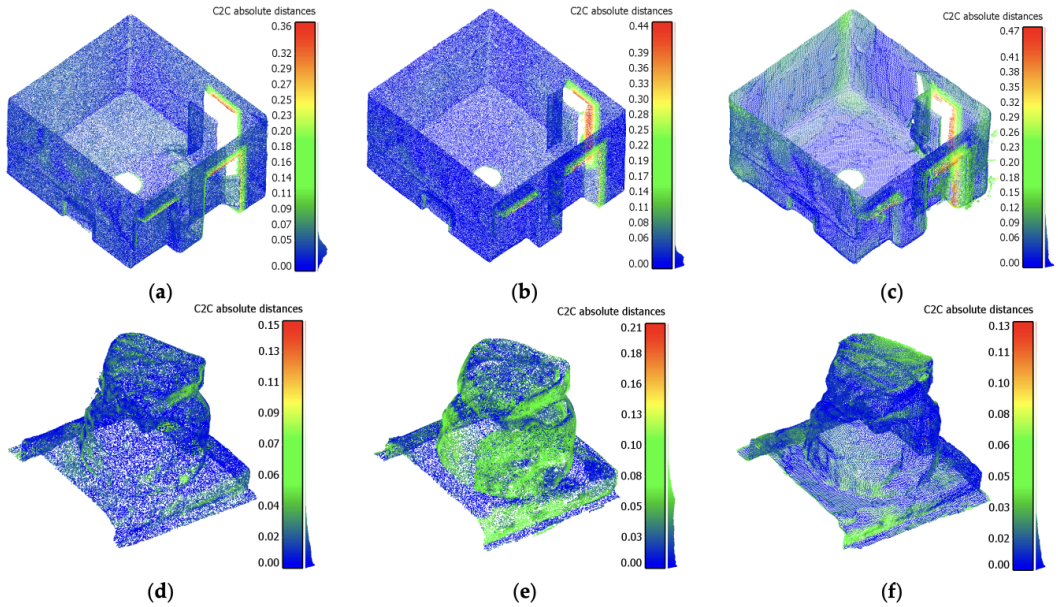


Figure 2.2: Histogrammes de comparaison nuage-à-nuage (C2C) entre les nuages de points smartphone et le nuage de points de référence [12].

2.3.3. Radar à pénétration de sol (GPR)

Le GPR est particulièrement adapté à la détection d'éléments enfouis et non visibles, ce qui en fait un outil précieux pour l'évaluation de l'enrobage ou l'identification de défauts internes (cavités, corrosion). Ma et al. [31] ont proposé un algorithme de reconnaissance de motifs hyperboliques dans les données GPR, permettant la détection automatique des barres avec une précision de 88%. De leur côté, Dinh et al. [14] ont développé une technique de migration hyperbolique simplifiée pour l'identification de barres, atteignant une précision supérieure à 98%.

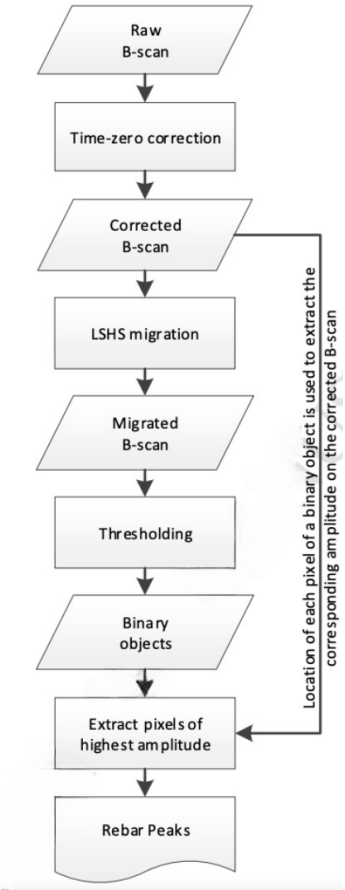


Figure 2.3: Illustration du mécanisme d'identification de barres proposé par [14].

Plus récemment, Kim et al. [25] ont montré que l'analyse des signaux A-scan peut être exploitée pour la détection de cavités dans les parois RC, tandis que d'autres travaux [1, 16] ont évalué l'impact de la corrosion sur les signaux GPR et proposé des modèles d'analyse automatisés. Bien que puissant pour révéler des caractéristiques invisibles aux approches optiques, le GPR reste coûteux et nécessite des opérateurs spécialisés. De plus, la qualité des résultats dépend fortement des propriétés du

béton (taux d'humidité, présence de chlorures), ce qui en limite l'applicabilité universelle.

2.4. Méthodes d'analyse et de traitement pour l'inspection des armatures

Si les technologies de collecte constituent une première étape essentielle pour obtenir des données exploitables, leur valeur réside principalement dans la capacité à analyser et interpréter ces données de manière automatique. Dans le domaine de l'inspection des structures en béton armé, un large éventail de méthodes ont été développées pour détecter, segmenter et mesurer les armatures. Ces approches peuvent être classées en deux grandes familles : les méthodes fondées sur des algorithmes "handcrafted" (géométrie, règles explicites) et celles basées sur l'apprentissage profond.

2.4.1. Méthodes basées sur des règles et la géométrie computationnelle

Les premières tentatives d'automatisation reposent sur des algorithmes "handcrafted", conçus pour extraire des caractéristiques spécifiques des données acquises. Yuan et al. [69] ont proposé une approche basée sur un algorithme de reconnaissance géométrique pour identifier la position des barres à partir de nuages de points LiDAR. Leur méthode calcule automatiquement un indice de qualité d'armature (RLQI), en comparant les mesures obtenues aux plans d'exécution. Une extension de ce travail a été présentée par les mêmes auteurs dans le contexte des ponts [71, 72], en développant deux algorithmes complémentaires (projection et découpage) pour extraire la disposition des barres à différents niveaux d'échelle. D'autres travaux ont misé sur la détection de motifs géométriques : Kim et al. [27] ont exploité RANSAC pour modéliser des cercles correspondant aux sections de barres dans des panneaux préfabriqués. Ce type d'approche se distingue par sa simplicité et sa faible dépendance aux données d'entraînement, mais il reste sensible au bruit et aux configurations complexes. Dans le domaine du GPR, Ma et al. [31] ont utilisé une stratégie de reconnaissance de motifs hyperboliques à partir de cartes SSD pour détecter les barres dans les signaux radar, tandis que Dinh et al. [14] ont proposé une technique de migration hyperbolique simplifiée pour améliorer la localisation automatique des armatures. Ces approches illustrent la pertinence de règles physiques ou géométriques explicites, mais elles manquent souvent de généralisation face à des

environnements hétérogènes.

2.4.2. Méthodes exploitant les approches d'apprentissage profond

Avec l'essor des réseaux de neurones, les approches basées sur l'apprentissage profond ont progressivement remplacé les méthodes "handcrafted, en raison de leur capacité à apprendre automatiquement des caractéristiques discriminantes. Shu et al. [46] ont introduit PCCR-Net, un réseau de segmentation dédié aux composants préfabriqués, qui intègre une fonction de perte focalisée et un protocole de mesure dimensionnelle. Cette approche a démontré qu'il était possible d'automatiser l'évaluation des dimensions des éléments en combinant segmentation et métriques quantitatives. Xi et al. [64] ont appliqué Faster-RCNN pour la détection des distanciers d'armature, associé à une étape d'estimation de pose par CPN pour calculer automatiquement les espacements. Cette combinaison illustre l'intérêt de l'intégration de modules complémentaires (détection + keypoints + calcul géométrique). D'autres travaux ont élargi le spectre des applications : Bolourian et al. [6] ont adapté PointNet++ pour la segmentation de défauts de surface (fissures, éclats), en intégrant des normales de surface afin de mieux capturer les caractéristiques locales des défauts. Bahreini et Hammad [3] ont exploré l'usage de DGCNN pour la segmentation de défauts de surface et l'intégration des résultats dans un modèle BIM enrichi, montrant la valeur de lier vision, nuages de points et gestion de l'information dans un cadre pratique.

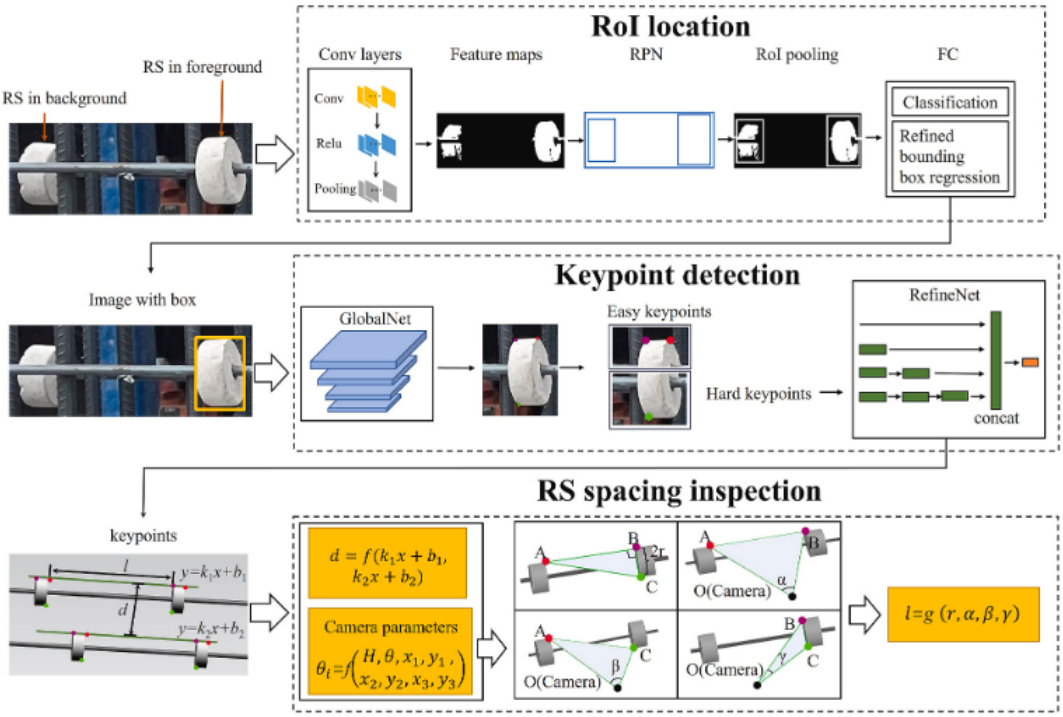


Figure 2.4: Vue d'ensemble de la méthode de calcul de l'espacement entre barres adjacentes proposée par [64].

Enfin, des travaux récents appliquent l'apprentissage profond à des tâches encore plus spécialisées : classification automatique des diamètres d'armatures [27], inspection probabiliste des canalisations RC via LiDAR [15], ou encore détection de cavités internes par GPR combiné à l'IA [16, 25].

2.5. Paradigmes de la reconstruction 3D

La reconstruction tridimensionnelle est un élément clé pour l'inspection des structures en béton armé, permettant de modéliser avec précision la disposition des armatures et d'évaluer leur conformité. Parmi les approches les plus utilisées, la photogrammétrie s'impose comme une solution largement adoptée grâce à son coût matériel réduit et à la disponibilité d'outils logiciels performants. Basée sur des algorithmes tels que Structure-from-Motion (SfM) [42] et Multi-View Stereo (MVS) [45], elle génère des modèles géométriquement détaillés à partir de simples images RGB.

Face à ces limites, de nouvelles approches ont émergé pour améliorer la qualité visuelle et la stabilité des reconstructions. Les Neural Radiance Fields (NeRF) [33] constituent une avancée récente en modélisation 3D. Ils utilisent un réseau de neu-

rones pour apprendre une fonction volumétrique continue à partir d'images calibrées, permettant de restituer des scènes complètes et photoréalistes.

NeRF a rapidement été adopté dans des domaines variés. Dans le domaine culturel et patrimonial, Croce et al. [13] ont montré que cette approche pouvait rivaliser avec la photogrammétrie pour la numérisation d'artefacts et de monuments, en particulier lorsque les conditions de capture sont limitées. Dans le secteur de l'architecture et de la construction, Inspection-NeRF [19] a été utilisé pour l'inspection de surfaces par robots grimpeurs, tandis que NeRF-to-BIM [18] a été proposé pour améliorer les processus de Scan-to-BIM, en offrant une alternative plus flexible et moins coûteuse que les relevés laser traditionnels.

À ce jour, aucun travail n'a exploré l'application de NeRF à l'inspection des armatures dans le béton armé. Croce et al. [13] ont toutefois réalisé une comparaison directe entre la photogrammétrie et NeRF dans le contexte patrimonial, montrant que si la photogrammétrie génère des modèles globalement complets, elle reste sensible aux conditions d'acquisition et présente un réalisme visuel limité. À l'inverse, NeRF se distingue par la stabilité et la richesse visuelle de ses reconstructions, suggérant un fort potentiel pour des tâches d'inspection exigeant des reconstructions fiables et détaillées.

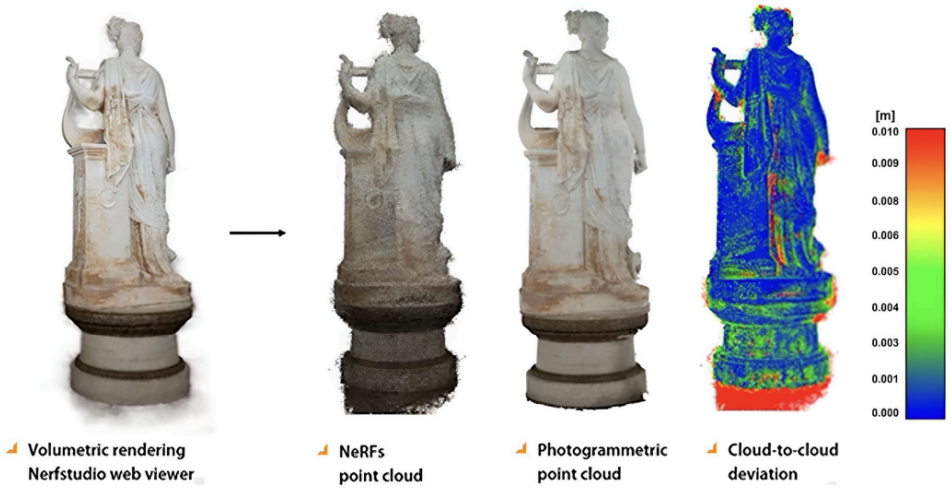


Figure 2.5: Nuages de points photogrammétriques vs NeRF [13].

Ces résultats mettent en évidence l'intérêt d'explorer le potentiel de NeRF pour l'inspection des armatures.

2.6. Paradigmes de segmentation des nuages de points 3D

L'analyse des nuages de points 3D constitue un défi majeur en vision par ordinateur en raison de leur nature non structurée, de la variabilité de densité et de la présence fréquente de bruit. Pour répondre à ces contraintes, plusieurs paradigmes ont été développés au fil du temps, chacun offrant un compromis particulier entre efficacité, précision et scalabilité. Ces paradigmes se répartissent principalement en quatre catégories : méthodes basées sur la projection, méthodes voxelisées, méthodes point-based et approches récentes de sérialisation.

2.6.1. Méthodes basées sur la projection

Les approches de projection cherchent à exploiter la maturité des réseaux de neurones convolutifs (CNN) 2D en transformant les données 3D en représentations 2D régulières. SqueezeSeg [60] a été l'un des premiers à projeter des nuages LiDAR sur des images sphériques, permettant l'application directe de CNN 2D. Cette approche a inspiré RangeNet++ [34], qui introduit une représentation apprise et un raffinement par CRF pour corriger les erreurs de segmentation aux frontières. D'autres variantes comme PolarNet [73] exploitent des projections cylindriques afin de mieux préserver les relations radiales, tandis que SqueezeSegV3 [65] a ajouté des modules d'agrégation contextuelle et une perte sensible à l'incertitude. Ces méthodes se distinguent par leur rapidité et leur scalabilité, en particulier dans des contextes tels que la conduite autonome. Toutefois, la transformation 3D–2D engendre des artefacts de projection et des distorsions géométriques, entraînant une perte d'information spatiale fine.

2.6.2. Méthodes basées sur les voxels

Les méthodes voxel-based consistent à discrétiser l'espace 3D en une grille régulière, permettant l'application de convolutions 3D. VoxelNet [76] a initié ce paradigme en intégrant des représentations de points au sein de voxels pour appliquer des convolutions 3D classiques. Pour pallier la complexité cubique, Submanifold Sparse Convolutional Networks [17] ont introduit des opérations clairsemées, améliorées par MinkowskiNet [11], qui offre une scalabilité accrue. Cylinder3D [77] a, quant à lui, proposé une discrétisation cylindrique adaptée aux scènes LiDAR outdoor, tandis

que SPVNAS [50] exploite la recherche d'architectures neuronales pour optimiser le compromis entre précision et efficacité. Bien que puissantes, ces méthodes souffrent de la croissance cubique du coût mémoire et computationnel en haute résolution, et introduisent des erreurs de quantification liées au maillage spatial.

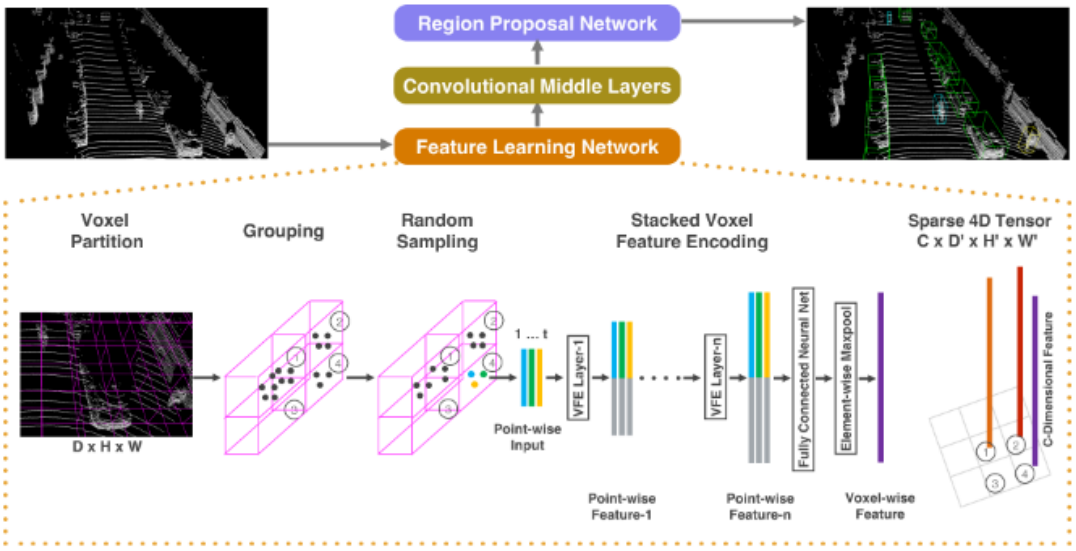


Figure 2.6: Illustration de l'architecture VoxelNet [76].

2.6.3. Méthodes basées sur les points

Les approches point-based traitent directement les nuages sous forme de points non ordonnés, conservant ainsi toute la richesse géométrique. PointNet [36] a ouvert la voie en utilisant des MLP partagés et un pooling symétrique pour assurer l'invariance par permutation. PointNet++ [37] a étendu cette idée par un apprentissage hiérarchique des caractéristiques locales. Par la suite, de nombreux modèles ont cherché à enrichir le traitement du voisinage : DGCNN [59] introduit les graphes dynamiques pour capter les relations topologiques, KPConv [52] propose des noyaux déformables pour mieux épouser la géométrie, tandis que PointCNN [29] apprend des transformations locales pour ordonner les points avant convolution. RandLA-Net [22] a particulièrement marqué le domaine en proposant un échantillonnage aléatoire efficace et une agrégation attentive, ce qui a permis d'atteindre un bon compromis entre précision et rapidité. Les modèles plus récents comme Point Transformer [74] et ses évolutions [62, 63] intègrent des mécanismes d'attention vectorielle pour capturer à la fois la localité et les dépendances globales.

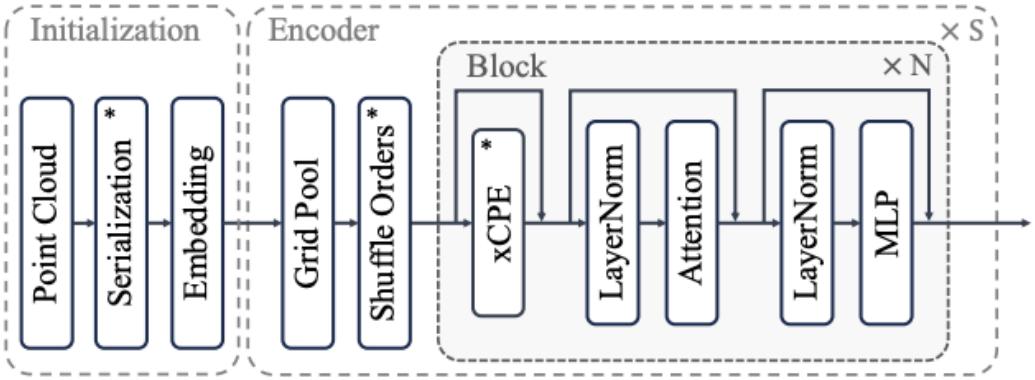


Figure 2.7: Architecture du modèle PointTransformerV3 proposée par [63].

Ces méthodes offrent une précision élevée mais souffrent d'un coût computationnel important lié aux recherches k-NN, ce qui limite leur application sur des nuages massifs.

2.6.4. Méthodes de sérialisation

La sérialisation des nuages de points est une approche émergente qui vise à ordonner les points 3D le long de courbes de remplissage d'espace afin de préserver la localité spatiale dans une séquence 1D. Les courbes de Hilbert et Morton (Z-order) sont les plus exploitées, car elles garantissent une cohérence spatiale locale [9].

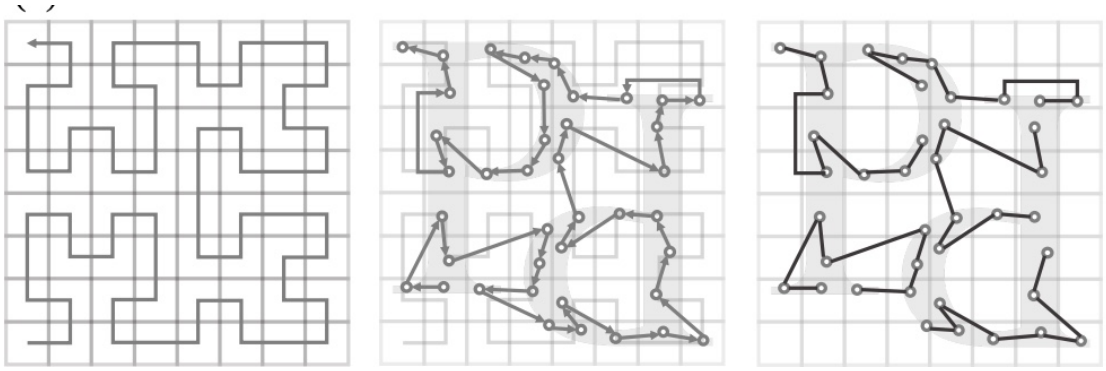


Figure 2.8: Illustration de la sérialisation d'un nuage de points [63].

Chen et al. ont proposé HilbertNet, qui applique cette idée dans un cadre récurrent. Des méthodes récentes comme PointMamba [30] ont montré qu'une sérialisation efficace combinée à des modèles séquentiels pouvait rivaliser avec des architectures

beaucoup plus lourdes. De plus, Point Transformer V3 [63] a introduit une hybridation des courbes d'ordre et des mécanismes d'attention pour traiter de larges scènes à moindre coût. Ces approches présentent un compromis prometteur : elles permettent de remplacer les requêtes k-NN coûteuses par des fenêtres séquentielles, réduisant ainsi la complexité tout en préservant la fidélité géométrique.

2.7. Discussion

L'analyse de la littérature met en lumière plusieurs défis qui freinent encore l'adoption à grande échelle de l'inspection automatisée des structures en béton armé. Ces défis concernent l'ensemble de la chaîne de traitement, depuis la collecte des données jusqu'à leur reconstruction et leur exploitation pour des tâches de mesure et de contrôle qualité.

En premier lieu, les technologies de capture restent un facteur limitant majeur. Les dispositifs professionnels tels que le LiDAR terrestre (TLS) [7, 15, 69–72] ou le radar à pénétration de sol (GPR) [14] offrent une précision métrique et une robustesse éprouvées, mais ils demeurent coûteux, encombrants et nécessitent une expertise technique pour être déployés sur le terrain. Leur usage est donc réservé à des contextes spécifiques, comme la recherche ou les grands projets d'infrastructure. À l'inverse, des solutions plus accessibles comme les capteurs intégrés aux smartphones, notamment le LiDAR de l'iPhone, ouvrent des perspectives intéressantes pour démocratiser l'acquisition 3D. Plusieurs travaux ont montré que ces capteurs pouvaient produire des nuages exploitables dans des tâches d'inspection [12, 40], mais leur précision reste inférieure à celle des systèmes professionnels et les nuages générés sont souvent bruités et incomplets. Ce manque de qualité impacte directement la fiabilité des étapes suivantes, en particulier la segmentation et la mesure dimensionnelle. Ainsi, un compromis doit être trouvé entre accessibilité et précision, en tenant compte des contraintes opérationnelles de chantier.

La reconstruction 3D constitue une étape critique du processus, car elle conditionne la qualité des modèles utilisés pour l'inspection et la mesure des armatures. Les méthodes traditionnelles de reconstruction comme la photogrammétrie permettent de produire des modèles géométriquement relativement précis avec des temps de traitement relativement rapides. Elles sont donc bien adaptées aux besoins actuels de la construction, mais présentent certaines limites, notamment une sensibilité aux conditions d'acquisition telles que l'éclairage, les occlusions ou la présence de

textures répétitives, ce qui peut affecter la complétude et l'homogénéité des nuages de points générés [13].

Les avancées récentes autour du NeRF offrent une alternative novatrice, centrée sur la génération de représentations implicites capables de capturer simultanément la géométrie et l'apparence visuelle d'une scène. Contrairement aux approches traditionnelles, NeRF met l'accent sur le réalisme visuel et la continuité des reconstructions, produisant des modèles plus riches et homogènes à partir de simples images calibrées et d'un matériel d'acquisition léger. Des études dans des domaines tels que le patrimoine culturel ou le Scan-to-BIM ont montré que NeRF pouvait surpasser la photogrammétrie en termes de stabilité visuelle et de gestion des détails complexes [13]. Toutefois, cette approche n'a pas été explorée pour l'inspection des armatures dans le béton armé, un domaine qui exige à la fois une grande précision métrique et une robustesse face aux conditions de terrain. Cette absence de travaux souligne un potentiel de recherche important, tout en mettant en évidence la nécessité d'adapter et d'évaluer rigoureusement ces modèles.

Enfin, l'exploitation des nuages de points pour l'inspection repose sur des techniques de segmentation et d'analyse. La segmentation sémantique est une étape incontournable pour identifier automatiquement les composants structurels, comme les nappes d'armatures, les distanciers ou les éléments contextuels. Les paradigmes existants présentent toutefois des limites importantes. Les méthodes basées sur la projection [34, 60, 73] sont rapides et bien adaptées à des environnements comme la conduite autonome, mais elles introduisent une perte de fidélité géométrique en raison de la transformation 3D–2D. Les approches voxelisées [11, 76, 77], quant à elles, conservent la structure spatiale mais souffrent d'une explosion du coût mémoire et computationnel lorsque la résolution augmente, ce qui les rend peu pratiques pour des scènes complexes et de grande taille. Les méthodes basés point [22, 29, 52, 59, 62, 74] préservent la richesse géométrique des nuages mais dépendent de recherches k-NN coûteuses, qui ralentissent considérablement le traitement. Ces limites expliquent l'émergence récente des approches de sérialisation, comme l'utilisation de courbes de Hilbert ou de Morton [30, 63]. En ordonnant les points 3D dans une séquence 1D tout en préservant la localité spatiale, ces méthodes permettent de remplacer les requêtes k-NN par des fenêtres glissantes plus légères, offrant ainsi un compromis intéressant entre précision et efficacité.

En synthèse, la littérature met en évidence plusieurs tensions qui structurent la problématique de l'inspection automatisée des armatures. D'une part, les capteurs

professionnels offrent une qualité métrique élevée mais au prix d'un coût et d'une logistique qui en limitent l'usage en pratique, tandis que les capteurs grand public rendent l'acquisition plus accessible au prix d'une moindre fidélité géométrique.

D'autre part, les approches de modélisation 3D traditionnelles présentent chacune des limites structurelles : le LiDAR professionnel demeure coûteux et difficile à mobiliser pour des usages courants, tandis que la photogrammétrie souffre de limites de complétude et d'homogénéité visuelle dans les scènes complexes. Dans ce contexte, des méthodes telles que NeRF offrent la perspective de reconstructions plus homogènes et flexibles, bien que leur exploration reste encore marginale dans le domaine de l'inspection des armatures. Enfin, les modèles actuels d'analyse de nuages de points peinent à concilier précision et efficience computationnelle, ce qui justifie l'intérêt porté aux architecture de segmentation allégées.

Ces constats soulignent la nécessité de concevoir des pipelines complets capables d'intégrer ces dimensions de manière cohérente : exploiter des données potentiellement de moindre qualité issues de capteurs accessibles au sein d'un framework fiable, améliorer la qualité des reconstructions dans un cadre de contrôle dimensionnel, et optimiser l'efficience des modèles de segmentation. Ces trois axes structurent les contributions présentées dans les chapitres suivants.

3 | Framework d'inspection automatique des armatures à partir de nuages de points

3.1. Introduction

Le béton armé constitue l'un des matériaux les plus utilisés dans la construction moderne grâce à la complémentarité mécanique entre le béton, résistant à la compression, et les armatures en acier, qui assurent la résistance en traction [57, 69]. La qualité de la mise en place des armatures, notamment en termes d'espacement, d'alignement et de recouvrement, détermine directement la sécurité structurelle, la durabilité et la conformité normative des ouvrages [10, 53].

L'inspection des armatures est réalisée en phase de ferraillage afin de détecter d'éventuelles non-conformités avant le coulage du béton. Historiquement, elle repose sur des méthodes manuelles, telles que l'inspection visuelle ou les mesures au ruban, encore largement répandues sur les chantiers [27]. Bien qu'elles soient simples à mettre en œuvre, ces méthodes présentent plusieurs limites, car elles sont chronophages, dépendent fortement de l'expérience des opérateurs et sont difficiles à appliquer dans des environnements contraints tels que la présence de nappes superposées, des zones à accès limité ou des conditions dangereuses [67].

Pour pallier ces limites, des technologies de contrôle non destructif ont été introduites, offrant des niveaux de précision et d'automatisation supérieurs. Parmi ces technologies figurent le LiDAR terrestre [10, 71], la photogrammétrie [6, 12], le radar à pénétration de sol (GPR) [14, 31] et les capteurs RGB-D [67]. Ces approches ont permis des avancées notables dans l'inspection automatique des structures en béton armé, avec des travaux portant sur la segmentation sémantique de défauts comme les fissures et éclats [6], la reconnaissance automatique des composants préfabriqués et l'évaluation dimensionnelle [46], ou encore l'estimation précise de l'espacement

des barres par des algorithmes combinant vision par ordinateur et géométrie computationnelle [64]. D'autres études se sont appuyées sur le LiDAR mobile pour extraire automatiquement la position des armatures à grande échelle [71] ou sur des méthodes basées sur le GPR pour détecter la position des barres dans le béton [31].

Cependant, ces approches présentent des contraintes importantes. Les technologies avancées, telles que les scanners LiDAR professionnels ou les systèmes GPR, offrent une grande précision mais restent coûteuses et nécessitent une expertise spécialisée, limitant leur adoption sur les chantiers courants. Les systèmes de vision purement basés sur l'image, bien que plus accessibles, impliquent souvent un compromis sur la précision métrique, ce qui peut poser problème pour le respect des tolérances exigées dans les normes de construction.

Parallèlement, l'émergence de capteurs de profondeur intégrés dans des appareils grand public, comme le LiDAR embarqué sur les iPhones, a ouvert de nouvelles perspectives pour la capture 3D à faible coût [12, 40]. Ce capteur repose sur un principe de direct Time-of-Flight (dToF), qui mesure le temps de vol des impulsions lumineuses pour estimer la distance et reconstruire la géométrie en trois dimensions [51] .

Plusieurs études ont exploré l'utilisation de ces capteurs dans des contextes variés, tels que l'inspection minière [40] ou l'analyse géométrique d'objets et d'environnements urbains [12]. Ces travaux démontrent que, bien que la précision absolue soit inférieure à celle des équipements professionnels, ces capteurs peuvent fournir des nuages de points exploitables dans des scénarios où la rapidité et la simplicité priment sur la précision millimétrique.

Les travaux exploitant le LiDAR embarqué sur iPhone pour l'inspection des armatures demeurent extrêmement limités, et l'on ne recense aucune proposition de framework complet et validé dédié à cette tâche. Les travaux existants se concentrent soit sur des dispositifs coûteux et complexes, soit sur des approches fondées sur l'image, sans réelle intégration d'un capteur grand public dans un cadre d'inspection automatisée. Cette lacune conduit à une question centrale : est-il possible de concevoir une solution à la fois fiable, reproductible et accessible, adaptée aux contraintes de chantier, en s'appuyant exclusivement sur un capteur de ce type ?

Dans ce contexte, l'objectif est de proposer un framework de bout en bout adapté à des nuages de points issus de modalités d'acquisition variées, avec pour cas d'usage le LiDAR embarqué sur iPhone, afin d'investiguer la faisabilité d'une inspection

automatisée reposant sur une solution accessible. Ce framework repose sur deux modules complémentaires. Le premier, appelé Component Recognition Module (CRM), utilise une adaptation de RandLA-Net [20] pour réaliser la segmentation sémantique des nuages de points et identifier les composantes structurelles. Le second, nommé Dimensional Assessment Module (DAM), applique une variante de l'algorithme RANSAC pour modéliser la géométrie des barres et calculer automatiquement les espacements inter-barres, étape essentielle pour vérifier la conformité au plan de ferraillage.

Sur le jeu de données collecté, le framework atteint une précision globale de 82,1% et une mIoU de 71,8% pour la segmentation, ainsi qu'une erreur absolue moyenne de 0,9 cm pour l'évaluation des espacements. Ces performances se situent dans la tolérance de ±1,3 cm spécifiée par la norme ACI 117-10 [2] pour l'espacement des armatures, démontrant ainsi que la solution proposée permet un contrôle conforme aux standards. Ces résultats montrent qu'il est possible de réaliser une inspection automatisée de faible coût avec un capteur grand public, sans compromettre la précision métrique requise pour la validation réglementaire. La suite du chapitre détaille la méthodologie proposée, le protocole expérimental et les résultats obtenus, ainsi qu'une discussion sur les limites et perspectives.

3.2. Framework proposé

Dans cette section, nous présentons notre framework exploitant les nuages de points acquis à l'aide du capteur LiDAR intégré aux iPhones pour l'inspection des armatures. Nous donnons d'abord une vue d'ensemble du pipeline proposé. Ensuite, nous détaillons les étapes de prétraitement appliquées aux données pour préparer leur analyse. Enfin, nous décrivons la structure et le fonctionnement des deux modules principaux : le **Component Recognition Module (CRM)** et le **Dimensional Assessment Module (DAM)**.

3.2.1. Vue d'ensemble du framework

Le framework proposé vise à résoudre le problème de l'évaluation de l'espacement des armatures dans les projets de construction, un aspect critique pour assurer l'intégrité structurelle et la conformité des ouvrages en béton armé. Il s'articule autour de deux modules principaux : le **CRM** (Component Recognition Module) et le **DAM** (Dimensional Assessment Module).

Le module CRM repose sur le modèle d'apprentissage profond **RandLA-Net** [20], optimisé et entraîné pour reconnaître deux composants : la nappe d'armatures (*Rebar Mat*) et la cage d'armatures (*Rebar Cage*). Ce choix permet de reformuler la reconnaissance des armatures comme une tâche de *segmentation sémantique*, assurant une identification précise des différents composants. Cette étape est essentielle pour l'évaluation de l'espacement. La robustesse de RandLA-Net et son efficacité sur des nuages de grande taille en font un choix adapté.

Ensuite, le module DAM analyse les nuages de points classés comme *Rebar Mat*. Il applique une approche en deux étapes fondée sur **RANSAC** : (i) ajustement de plans afin de séparer les nappes supérieure et inférieure, puis (ii) ajustement de droites pour détecter les barres individuelles, extraire les barres parallèles et adjacentes, et calculer les distances entre elles. L'évaluation de ces distances permet non seulement d'identifier les composants mais aussi de fournir une analyse dimensionnelle complète.

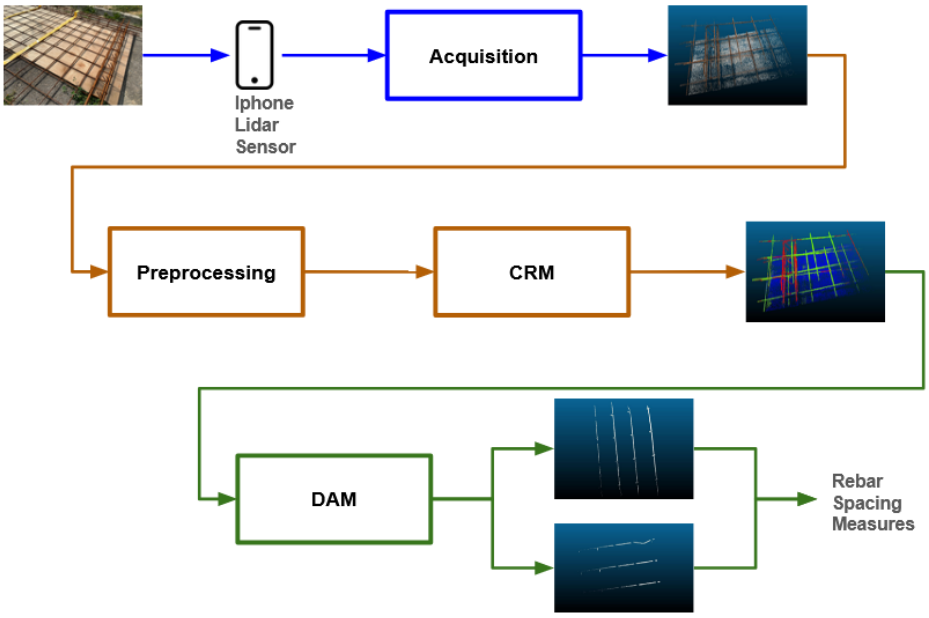


Figure 3.1: Procédure globale du framework proposé pour l'évaluation automatique de l'espacement des armatures à partir de données LiDAR iPhone.

3.2.2. Prétraitement des données

Les nuages de points capturés à l'aide de capteurs grand public, tels que le LiDAR intégré aux iPhones, présentent plusieurs défis avant d'être exploités pour une in-

spection automatisée. Ces données brutes contiennent généralement du bruit, des zones manquantes et une densité irrégulière, et peuvent être affectées par des imprécisions liées à des erreurs de calibration ou de mouvement pendant l'acquisition. Le prétraitement constitue donc une étape essentielle, car il conditionne directement la qualité des résultats des étapes ultérieures, notamment la segmentation sémantique et l'évaluation dimensionnelle. Il vise à stabiliser les données, réduire les artefacts et préparer les nuages de points à l'entraînement des modèles d'apprentissage profond. Dans notre framework, il repose sur une série d'opérations successives soigneusement définies pour répondre à ces problématiques spécifiques.

Enregistrement des nuages de points. La première étape consiste à aligner les différents nuages capturés lors de l'acquisition. Les structures d'armatures étant souvent complexes et partiellement occluses, un seul scan ne suffit pas à obtenir une couverture complète. Plusieurs acquisitions sont donc réalisées sous des angles variés, puis fusionnées dans un même repère spatial grâce à un recalage multi-vues. Cette opération, appelée *registration*, repose sur l'algorithme ICP (*Iterative Closest Point*) qui minimise l'erreur de distance entre les nuages partiels. Elle permet ainsi de reconstituer la géométrie complète de la structure, de réduire les zones aveugles liées aux occlusions et d'améliorer la densité et l'homogénéité du nuage final. Dans le cas d'un capteur mobile tel qu'un iPhone, cette étape compense également les variations de trajectoire et d'orientation de l'appareil, garantissant un modèle cohérent avant toute analyse.

Suppression statistique des outliers. Une fois les nuages alignés, les points aberrants ou bruités, fréquemment générés par les capteurs grand public en raison de réflexions parasites ou d'erreurs de mesure, sont supprimés. Ces outliers, s'ils ne sont pas filtrés, peuvent perturber la segmentation et introduire des biais dans les résultats. Pour y remédier, un filtre statistique basé sur le voisinage est appliqué. Celui-ci repose sur l'hypothèse que les points valides présentent des distances homogènes par rapport à leurs voisins, tandis que les points aberrants se distinguent par des distances anormalement élevées. Pour chaque point, la distance moyenne à ses k plus proches voisins est calculée, et le point est supprimé si cette distance dépasse un seuil défini à partir de la distribution statistique globale, généralement fixé à la moyenne augmentée de deux écarts-types. Cette étape réduit considérablement les artefacts tout en préservant la géométrie utile et améliore la robustesse de la reconstruction.

Sous-échantillonnage par grille (Grid Sampling). Le nuage est ensuite sous-échantillonné afin de réduire sa taille et de rendre les traitements ultérieurs plus efficaces. Les scans initiaux peuvent contenir plusieurs centaines de milliers de points, ce qui représente un coût computationnel élevé. Pour limiter cette complexité tout en conservant la structure géométrique essentielle, un sous-échantillonnage par grille régulière est réalisé. L'espace 3D est discrétisé en cellules cubiques de taille fixe, et un seul point représentatif est conservé par cellule. Cette méthode, contrairement à un échantillonnage aléatoire, garantit une répartition homogène des points et permet de contrôler explicitement la densité finale du nuage, tout en préservant les informations nécessaires à la segmentation. Elle constitue un compromis optimal entre la qualité des données et l'efficacité computationnelle, ce qui est essentiel pour un déploiement pratique sur le terrain.

Mise à l'échelle et normalisation. Enfin, les coordonnées (x, y, z) des points sont normalisées et mises à l'échelle dans un repère unifié. Cette étape est indispensable pour les réseaux de segmentation, car elle favorise la convergence de l'apprentissage, limite les biais liés aux différences d'échelle entre plusieurs jeux de données et améliore la stabilité numérique des calculs. La normalisation prend une importance particulière lorsque les nuages proviennent de contextes variés, tels que des structures de tailles différentes ou des résolutions de capteur hétérogènes. Elle constitue la dernière opération du prétraitement avant l'entrée des données dans le module CRM pour la segmentation.

En résumé, ce processus de prétraitement garantit que les nuages de points utilisés dans le pipeline soient propres, homogènes et compacts, tout en préservant les informations géométriques nécessaires à une inspection précise et fiable.

3.2.3. Module de Reconnaissance de Composants (CRM)

L'identification précise des composants constitue une étape essentielle du pipeline proposé, car elle conditionne directement la fiabilité de l'évaluation dimensionnelle des armatures. Toute erreur dans la segmentation entraîne inévitablement des imprécisions dans la mesure des espacements, ce qui compromet l'efficacité globale du cadre d'inspection.

Nous formulons la reconnaissance des composants du béton armé comme une tâche de **segmentation sémantique** de nuages de points 3D. Pour ce faire, nous adoptons le modèle RandLA-Net [20], reconnu pour sa capacité à traiter efficacement de larges

nuages sans recourir à la voxelisation. Son mécanisme d'agrégation hiérarchique permet de capturer simultanément l'information locale et le contexte global, ce qui le rend particulièrement adapté à la complexité géométrique des armatures. Par ailleurs, RandLA-Net a été largement validé sur plusieurs jeux de données publics de référence, où il s'est imposé comme un modèle à la fois puissant et efficient, ce qui en fait un choix pertinent.

Dans notre implémentation, le CRM est entraîné pour distinguer les principales composantes des armatures. L'objectif est d'attribuer à chaque point du nuage une étiquette sémantique reflétant son appartenance à l'un de ces composants.

Compte tenu du fort déséquilibre entre les classes, certaines étant surreprésentées tandis que d'autres apparaissent rarement, nous adoptons la fonction de perte **Weighted Cross-Entropy** (WCE) [8]. Elle permet d'ajuster l'importance de chaque classe en introduisant un poids α_i, ce qui évite que les classes minoritaires ne soient négligées par le modèle. La fonction de perte est définie comme suit :

$$WCE(y, \hat{y}) = -\sum_i \alpha_i \, p(y_i) \, \log\left(p(\hat{y}_i)\right)$$

où :

- y_i et $\hat{y}_i$ représentent respectivement les labels réels et prédits de la classe i,
- α_i désigne le poids associé à la classe i,
- $p(y_i)$ correspond à la probabilité réelle de la classe i,
- $p(\hat{y}_i)$ est la probabilité prédite de la classe i.

Ce choix méthodologique garantit une meilleure robustesse du CRM face aux déséquilibres de données et permet d'obtenir une segmentation suffisamment fiable pour alimenter le module d'évaluation dimensionnelle (DAM).

3.2.4. Module d'Évaluation Dimensionnelle (DAM)

Une fois le nuage de points classifié par le CRM, nous extrayons les sous-ensembles correspondant spécifiquement aux nappes d'armatures. L'objectif du DAM est d'analyser ces points afin de séparer les niveaux d'armatures et de détecter les barres individuelles, pour ensuite évaluer automatiquement les espacements.

La première étape consiste à séparer les nappes d'armatures dans le nuage de points.

Pour cela, nous utilisons une variante de l'algorithme RANSAC, adapté à la détection de multiples plans. Chaque plan est ajusté en minimisant la distance entre les points et le modèle géométrique défini par l'équation suivante :

$$ax + by + cz + d = 0$$

où :

- (a, b, c) désigne le vecteur normal au plan,
- (x, y, z) sont les coordonnées des points du nuage,
- d représente la distance du plan par rapport à l'origine.

L'algorithme RANSAC est paramétré par le nombre maximal d'itérations et par un seuil de distance permettant de décider si un point est un inlier. Ce seuil est choisi expérimentalement de manière à bien séparer les deux nappes, garantissant une correspondance fidèle avec les niveaux réels sans recouvrement significatif.

L'algorithme 3.1 décrit la procédure de détection des nappes.

Après cette étape, chaque nappe est isolée et peut être analysée pour identifier les barres individuelles. Pour cela, nous appliquons une seconde version de l'algorithme RANSAC, cette fois adaptée à l'ajustement de lignes. Une droite passant par deux points $\mathbf{p}_1$ et $\mathbf{p}_2$ peut être définie par l'équation paramétrique suivante :

$$\mathbf{r}(t) = \mathbf{p}_1 + t \cdot \mathbf{d}$$

où $\mathbf{d} = \frac{\mathbf{p}_2 - \mathbf{p}_1}{\|\mathbf{p}_2 - \mathbf{p}_1\|}$ est le vecteur directeur normalisé.

L'algorithme 4.1 décrit la procédure de détection des barres individuelles.

Enfin, les droites détectées sont regroupées en paires adjacentes et parallèles. La distance entre chaque paire est ensuite calculée afin de vérifier si l'espacement respecte les normes de conception et de construction.

3.3. Expérimentations

Dans cette section, nous évaluons les performances du framework proposé, et plus particulièrement celles des modules CRM et DAM, à l'aide de notre jeu de données

Algorithm 3.1 Algorithme de détection des nappes d'armatures

Initialisation:
remaining_points ← points
detected_planes ← liste vide
detected_inliers ← liste vide
epoch ← 0
repeat
  Incrémenter *epoch* de 1
  Réinitialiser *best_plane* ← None, *best_inliers_indices* ← liste vide
  for $i = 1$ à *max_iterations* **do**
    Sélectionner trois points aléatoires dans *remaining_points*
    Calculer le vecteur normal au plan défini par ces points
    Normaliser le vecteur normal
    Déterminer les coefficients du plan
    Calculer la distance de chaque point à ce plan
    Déterminer les inliers selon le seuil de distance
    if le nombre d'inliers est supérieur au précédent meilleur **then**
      Mettre à jour *best_inliers_indices*
      Mettre à jour *best_plane*
    end if
  end for
  Ajouter *best_inliers_indices* à *detected_inliers*
  Retirer les inliers de *remaining_points*
  Ajouter *best_plane* à *detected_planes*
until *epoch* = 2
return *detected_planes*, *detected_inliers* =0

Algorithm 3.2 Algorithme de détection des barres d'armatures

Initialisation:
remaining_points ← points
lines ← liste vide
while taille(*remaining_points*) ≥ *min_points_for_line* **do**
  best_inliers_count ← 0
  best_line_model ← None
  for $i = 1$ à *num_iterations* **do**
    Sélectionner deux points p_1 et p_2 dans *remaining_points*
    Définir une ligne par p_1 et p_2
    Calculer les distances de tous les points à cette ligne
    Déterminer les inliers selon le seuil de distance
    if nombre d'inliers > *best_inliers_count* **then**
      Mettre à jour *best_line_model*
      Mettre à jour *best_inliers_count*
    end if
  end for
  if *best_inliers_count* ≥ *min_points_for_line* **then**
    Ajouter *best_line_model* à *lines*
    Retirer les inliers de *remaining_points*
  else
    Quitter la boucle
  end if
end while
return *lines* =0

construit spécifiquement pour cette étude. Nous décrivons d'abord le jeu de données, puis les paramètres expérimentaux retenus, avant de présenter et d'analyser les résultats obtenus.

3.3.1. Jeu de données

Le jeu de données utilisé pour l'évaluation est constitué de 46 nuages de points pour l'entraînement, 16 pour la validation et 16 pour le test. En moyenne, chaque nuage contient environ 377 020 points et couvre une surface de taille moyenne équivalente à une grille de 5×5.

L'annotation a été réalisée en trois classes : *Nappe d'armatures*, *Chaînage* et *Autres*. Pour l'entraînement et l'évaluation, seules les coordonnées 3D (x, y, z) et les informations de couleur (r, g, b) ont été utilisées.

La répartition détaillée des points par classe est présentée dans le Tableau 3.1. La distribution est fortement déséquilibrée : la classe *Autres* contient le plus grand nombre de points, suivie de la classe *Nappe d'armatures*, tandis que la classe *Chaînage* reste très minoritaire. Ce déséquilibre marqué justifie l'utilisation de la fonction de perte *Weighted Cross Entropy* (WCE) [8], qui permet de mieux représenter les classes minoritaires et d'éviter que le modèle ne se concentre excessivement sur les classes dominantes.

Table 3.1: Répartition des points par classe dans le jeu de données

Classe	Nombre de points
Autres	16 226 270
Nappe d'armatures	10 978 460
Chaînage	2 202 847

3.3.2. Configuration expérimentale

Les expérimentations ont été menées à l'aide du dispositif illustré à la Figure 3.2. Le processus de collecte des données repose sur l'utilisation du capteur LiDAR de l'iPhone 14 Pro, en combinaison avec l'application *Scaniverse* (version 2.1.8), afin de scanner des structures d'armatures et générer des nuages de points contenant des informations de position (x, y, z) et de couleur (r, g, b). Chaque nuage de points est ensuite soumis aux étapes de prétraitement détaillées dans la Section 3.2.2.

Pour la tâche de segmentation sémantique, toutes les expériences ont été réalisées



Figure 3.2: Dispositif expérimental utilisé pour la collecte des données.

avec PyTorch sur un GPU Nvidia RTX A2000. L'optimiseur Adam a été utilisé avec un taux d'apprentissage initial de 1×10^{-2} et une taille de lot de 2. Le modèle RandLA-Net a été entraîné pendant 100 époques, avec une décroissance exponentielle du taux d'apprentissage (facteur 0.95 à chaque époque). Les nuages de points en entrée étaient composés de 40 960 points, et le voisinage considéré était fixé à $k = 16$.

Concernant l'évaluation des espacements entre barres, les deux modèles RANSAC (extraction des nappes d'armatures et détection des barres individuelles) ont été configurés avec un nombre maximal d'itérations fixé à 1000.

3.3.3. Résultats

L'évaluation du framework proposé repose sur deux ensembles de métriques complémentaires :

- **La précision globale** (Overall Accuracy, OA) et l'**Intersection sur Union**

moyenne (mIoU) pour évaluer la performance du module CRM de segmentation sémantique des nuages de points ;

- **L'erreur absolue moyenne** (Mean Absolute Error, MAE) pour mesurer la précision du module DAM dans l'estimation automatique des espacements entre barres.

Les résultats globaux sont présentés dans le Tableau 3.2. Notre approche atteint une précision globale de **82,1 %** et une mIoU de **71,8 %** pour la segmentation sémantique. Concernant l'évaluation dimensionnelle, le framework obtient une MAE de seulement **0,9 cm**, ce qui est inférieur à la tolérance maximale de $\pm 1,3$ cm (0,5 inch) spécifiée par la norme **ACI 117-10** [2]. Cela démontre que les mesures générées par notre système sont conformes aux exigences industrielles pour l'inspection des armatures en béton armé.

Table 3.2: Résultats globaux de l'évaluation du framework proposé

Métrique	**OA (%)**	**mIoU (%)**	**MAE (cm)**
Valeur obtenue	82.1	71.8	0.9

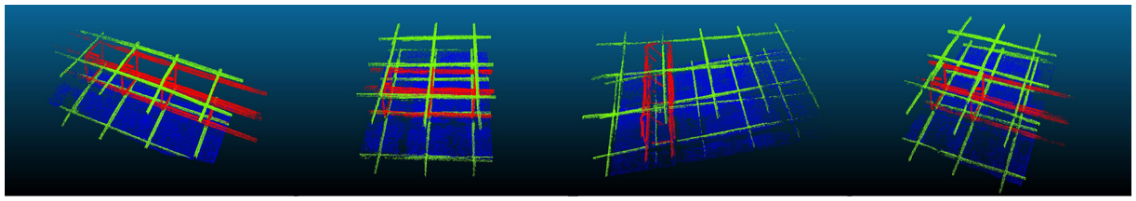
Afin d'illustrer la performance visuelle du module CRM, la Figure 3.3 présente une comparaison entre les annotations de vérité terrain et la segmentation prédite.

- La sous-figure (a) montre la vérité terrain avec les trois classes annotées : *Nappe d'armatures*, *Chaînage*, et *Autres éléments*.
- La sous-figure (b) présente la segmentation produite automatiquement par le modèle RandLA-Net intégré au CRM.

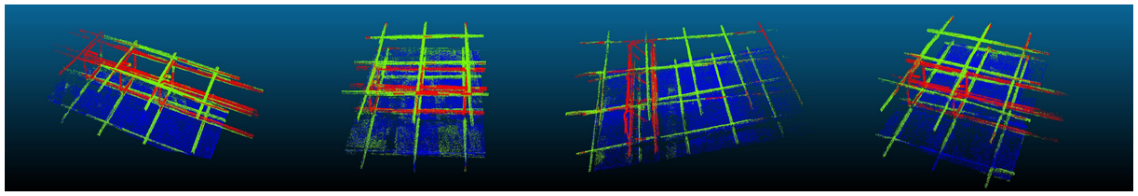
Cette visualisation met en évidence la capacité du modèle à détecter correctement la majorité des barres et des chaises, tout en révélant certaines zones de confusion localisées, notamment entre les nappes et le chaînage.

Une analyse détaillée par classe est présentée dans le Tableau 3.3.

- La classe *Autres* est la mieux reconnue, avec une précision moyenne élevée et un IoU de **89,7 %**, reflétant une bonne différenciation des éléments contextuels.
- La classe *Nappe d'armatures* obtient une précision satisfaisante mais un IoU plus modeste (**67,1 %**), indiquant que certains points sont encore confondus avec la classe *Chaînage*.
- La classe *Chaînage* reste la plus difficile à identifier, avec des confusions



(a) Annotations de vérité terrain avec les classes : nappe d'armatures, chaînage et autres éléments.



(b) Segmentation prédite par le framework proposé.

Figure 3.3: Comparaison visuelle entre la vérité terrain (a) et la segmentation prédite (b), illustrant la performance du module CRM.

fréquentes dues à la forte similarité géométrique et visuelle avec les nappes d'armatures.

Table 3.3: Résultats détaillés par classe pour la segmentation sémantique

Classe	**Précision (%)**	**IoU (%)**
Autres	93.7	89.7
Nappe d'armatures	82.5	67.1
Chaînage	70.7	57.1

Ces résultats démontrent la capacité du framework à réaliser une inspection automatisée répondant aux contraintes du terrain. La précision atteinte pour la mesure des espacements confirme que notre système est adapté aux besoins opérationnels de contrôle qualité, conformément à la norme ACI 117-10 [2].

3.4. Conclusion

Dans ce chapitre, nous avons proposé un framework de bout en bout pour l'inspection automatique des armatures, avec comme source de données, les nuages de points issues du capteur LiDAR intégré aux iPhones. Cette approche s'inscrit dans la continuité des travaux présentés dans l'état de l'art, en cherchant à répondre à l'une des principales limitations identifiées : l'accessibilité des approches et outils de capture et la nécessité de disposer de solutions à la fois fiables et conformes aux exigences du génie civil.

Le pipeline proposé, articulé autour du Component Recognition Module (CRM) et du Dimensional Assessment Module (DAM), permet non seulement d'identifier automatiquement les composantes des armatures dans un nuage de points, mais aussi d'évaluer avec précision l'espacement entre les barres. Les résultats obtenus confirment la pertinence de cette démarche : une precision globale de 82,1 % et une mIoU de 71,8 % pour la segmentation sémantique, ainsi qu'une erreur absolue moyenne de seulement 0,9 cm pour la mesure des espacements.

Ces performances démontrent qu'il est possible de démocratiser l'inspection des armatures en recourant à des capteurs grand public, tout en maintenant un niveau de précision compatible avec les exigences du génie civil. Toutefois, plusieurs limitations demeurent, notamment la robustesse face à des environnements de chantier variés, la généralisation à d'autres types de structures, et la qualité des reconstructions 3D.

Cette première contribution établit ainsi les bases d'une inspection automatisée, accessible et reproductible. Elle ouvre la voie à l'exploration de méthodes de reconstruction plus riches, notamment à travers les Neural Radiance Fields (NeRF), qui seront étudiés dans le chapitre suivant.

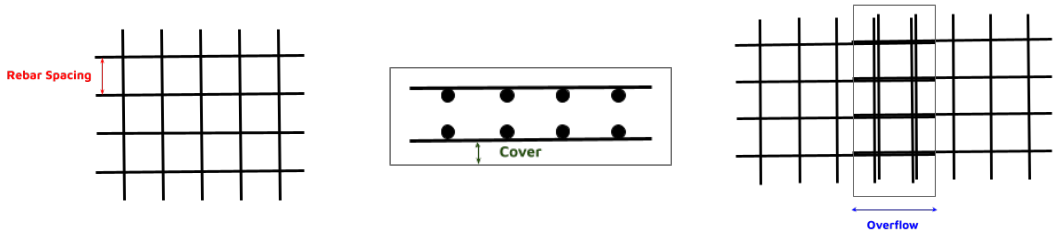
4 | Expérimentation des champs de radiance neuronaux pour l'inspection automatique des armatures

4.1. Introduction

Garantir l'intégrité structurelle et la sécurité des ouvrages en béton armé (BA) constitue un enjeu central dans le secteur de la construction [47]. Le BA associe deux composants principaux : le béton, qui fournit la résistance en compression, et les armatures métalliques (rebars), qui assurent la résistance en traction [70]. Cette combinaison confère au BA des propriétés mécaniques supérieures, expliquant son utilisation massive dans les bâtiments et les infrastructures modernes [58]. La durabilité et la fiabilité de ces structures reposent directement sur le positionnement, l'espacement et l'alignement corrects des armatures.

Plusieurs paramètres doivent être inspectés avec précision afin de garantir la conformité aux normes de conception et aux performances structurelles attendues : l'*espacement des barres*, qui assure une bonne répartition des charges et un coulage homogène du béton ; l'*alignement des armatures*, garant de la continuité des chemins de charge ; le *enrobage* (cover), épaisseur de béton (généralement de 2,5 à 4 cm en bâtiment) qui protège les aciers contre la corrosion et le feu ; ainsi que le *chevauchement des nappes* (overlap), nécessaire pour assurer la continuité mécanique dans les configurations multi-couches [54]. Toute déviation par rapport à ces paramètres peut entraîner des faiblesses structurelles ou une perte de durabilité.

Historiquement, l'inspection des armatures reposait sur des méthodes manuelles et intrusives, telles que le carottage ou l'examen visuel, chronophages, coûteuses en main-d'œuvre et parfois dangereuses en environnement de chantier [26]. Pour pallier



(a) espacement des armatures (b) enrobage (c) chevauchement des nappes

Figure 4.1: Illustration des paramètres critiques dans une structure en BA.

ces limitations, des techniques non destructives ont été introduites, parmi lesquelles le radar à pénétration de sol (GPR), les systèmes de vision et le LiDAR. Si ce dernier est reconnu pour sa capacité à produire des nuages de points denses et précis, son coût élevé et la complexité de son utilisation limitent sa généralisation.

Parallèlement, les technologies de reconstruction 3D ont connu des avancées rapides, ouvrant la voie à des représentations numériques détaillées et exploitables de scènes complexes [75]. Ces progrès ont déjà trouvé des applications en patrimoine, en robotique, en santé ou encore dans le domaine AEC (Architecture, Engineering and Construction) [24].

Dans ce contexte, les **Neural Radiance Fields (NeRF)** [33] se distinguent comme une approche récente, capable de générer des reconstructions implicites et photoréalistes à partir d'un simple jeu d'images RGB. NeRF modélise une scène comme une fonction continue reliant les coordonnées spatiales (x, y, z) et la direction de vue (θ, ϕ) à une couleur (R, G, B) et une densité volumique σ. Ce formalisme permet de synthétiser de nouvelles vues réalistes depuis des points de vue inédits, en intégrant à la fois la géométrie et l'apparence dans un réseau neuronal. Contrairement aux méthodes traditionnelles basées sur des modèles géométriques explicites, NeRF encode implicitement la structure et les textures, offrant ainsi un réalisme inédit.

Depuis son introduction, NeRF a suscité un vif intérêt et a été décliné en de multiples variantes pour traiter les scènes dynamiques, les environnements de grande échelle ou encore les applications temps réel. S'il a déjà démontré son potentiel dans des domaines variés comme la conduite autonome, la préservation patrimoniale ou l'imagerie médicale, son utilisation pour l'inspection des structures en béton armé reste encore inexplorée.

Dans notre contribution précédente, nous avons proposé un **framework** exploitant

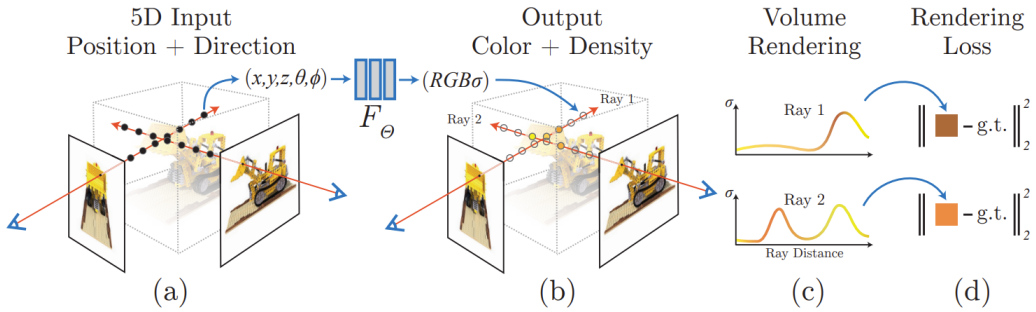


Figure 4.2: Principe général de l'entraînement NeRF, adapté de la version originale proposée par Mildenhall et al. [33].

le LiDAR des iPhones pour l'évaluation de l'espacement des armatures. Cette contribution a démontré la faisabilité d'une inspection automatisée accessible, tout en révélant certaines limitations liées à l'incomplétude des nuages de points et à la difficulté de passage à l'échelle.

La présente contribution cherche à dépasser ces limites en intégrant NeRF comme alternative de reconstruction 3D à partir d'images RGB. L'objectif est double : d'une part, évaluer la pertinence de NeRF pour générer des nuages de points de meilleure qualité pour l'inspection des armatures, et d'autre part, comparer ses performances à celles d'une méthode établie comme la photogrammétrie. En procédant ainsi, nous visons à enrichir le **framework** existant et à examiner dans quelle mesure NeRF peut améliorer la précision, la robustesse et l'accessibilité des inspections de structures en béton armé.

4.2. Méthodologie

Cette section présente l'approche méthodologique adoptée pour explorer l'application des Neural Radiance Fields (NeRF) à l'inspection des armatures dans les structures en béton armé. La méthodologie proposée s'articule en plusieurs étapes : une vue d'ensemble du framework général, la description des processus d'acquisition et de prétraitement des données, puis l'intégration des modules CRM (Component Recognition Module) et DAM (Dimensional Assessment Module) dans la chaîne de traitement.

4.2.1. Vue d'ensemble du framework

Le framework proposé vise à répondre au défi de l'évaluation de l'espacement des armatures dans les structures en béton armé, paramètre essentiel pour garantir l'intégrité structurelle et la conformité réglementaire. Dans notre contribution précedente, l'application *Scaniverse* a été exploitée pour générer des nuages de points en combinant deux modes distincts d'acquisition :

- **Mode LiDAR :** basé sur le capteur LiDAR intégré aux iPhones et sur ARKit, qui fusionne le suivi visuel avec les données de l'accéléromètre et du gyroscope afin de produire des maillages 3D.
- **Mode photogrammétrique :** basé uniquement sur des images RGB, dont les poses sont estimées à partir de caractéristiques visuelles et reconstruites par stéréovision multi-vues, sans recours au LiDAR.

Si ces deux approches permettent de générer des données relativement précises, elles présentent néanmoins des limites en termes de *scalabilité* et de complétude des nuages de points.

Afin de dépasser ces limites, nous investiguons dans ce travail le potentiel des NeRF pour la génération de nuages de points. Les reconstructions NeRF sont réalisées à l'aide de *Nerfstudio* [49], qui transforme des séquences d'images RGB en modèles 3D détaillés. Pour évaluer leur pertinence, nous comparons les performances des NeRF avec celles de la photogrammétrie intégrée à Scaniverse. Les deux approches reposant exclusivement sur l'iPhone comme dispositif d'acquisition, elles garantissent une solution accessible et peu coûteuse pour l'inspection des armatures.

Une fois les nuages de points générés, le module CRM exploite le modèle RandLA-Net [21], optimisé pour la segmentation sémantique de grandes scènes, afin d'identifier automatiquement les composantes des armatures. Ensuite, le module DAM applique un algorithme RANSAC adapté pour modéliser les barres d'armature et calculer l'espacement entre les barres parallèles. L'ensemble du framework méthodologique est illustré à la Figure 4.3.

4.2.2. Acquisition des données

Cette partie s'intéresse à l'application des NeRF pour la génération de nuages de points destinés à l'inspection des armatures. Les NeRF exploitent des réseaux de neurones pour reconstruire des modèles 3D détaillés à partir d'images 2D, ouvrant

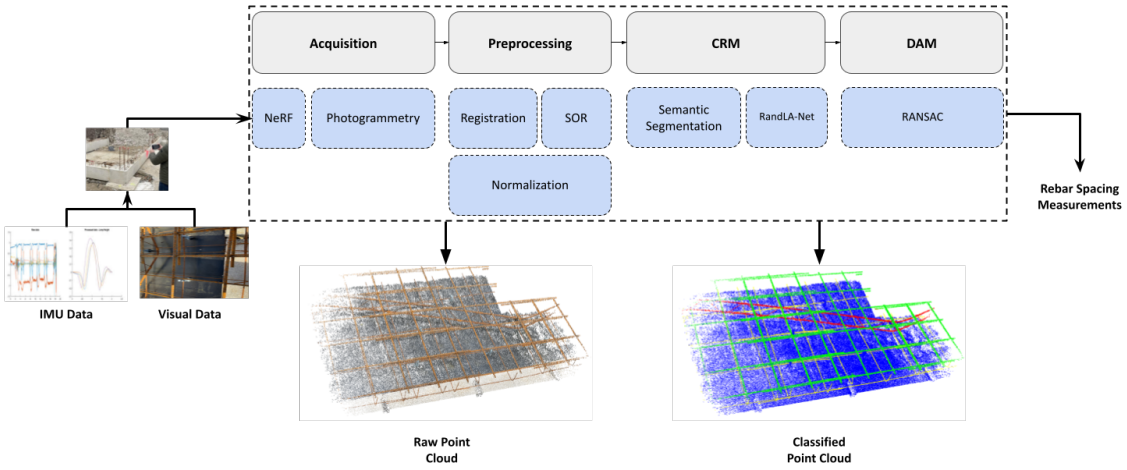


Figure 4.3: Architecture générale du framework proposé pour l'évaluation automatique de l'espacement des armatures.

ainsi la voie à des reconstructions riches et précises à partir de données d'entrée accessibles.

Pour évaluer l'efficacité des NeRF, nous utilisons également la photogrammétrie intégrée à l'application *Scaniverse*, qui combine images RGB et données issues des capteurs inertiels (gyroscope et accéléromètre) afin de produire des reconstructions 3D. Ainsi, deux méthodes sont comparées :

- **NeRF :** utilisé comme méthode principale de reconstruction, générant des nuages de points à partir d'images RGB capturées avec un iPhone.
- **Photogrammétrie :** employée à titre de référence (benchmark), afin de mesurer les apports spécifiques des NeRF par rapport à une technique déjà éprouvée.

Les deux approches partagent l'avantage d'un processus d'acquisition peu coûteux, basé uniquement sur un iPhone, ce qui les rend particulièrement adaptées pour envisager une démocratisation de l'inspection automatisée des armatures.

NeRF

Dans cette étude, nous avons déployé les NeRF à l'aide du framework *Nerfstudio* [49], en utilisant spécifiquement la méthode *depth-enhanced Nerfacto*. Ce choix repose sur son équilibre entre vitesse d'entraînement et qualité de reconstruction, intégrant des techniques issues de modèles récents tels que MipNeRF-360 [4] et Instant-NGP [35].

Lors de nos expérimentations, Instant-NGP a également été testé comme alternative. Bien qu'il permette un entraînement nettement plus rapide, nous avons constaté l'apparition d'artefacts visuels, en particulier dans les zones de recouvrement complexe des armatures, ce qui compromet sa fiabilité pour l'évaluation de l'espacement des barres.

En comparaison, *Nerfacto* a produit des reconstructions plus précises et plus stables, offrant une fidélité géométrique essentielle pour une évaluation dimensionnelle fiable. Bien que légèrement plus lent qu'Instant-NGP, *Nerfacto* reste bien plus efficace que les NeRF classiques, ce qui en fait un compromis pertinent pour notre cas d'étude.

L'intégration des NeRF dans ce framework dépasse l'objectif classique de reconstruction géométrique et de texturage. Notre ambition est d'obtenir des mesures dimensionnelles précises dans un environnement réel. Pour atteindre cet objectif, nous avons recours à la localisation et cartographie simultanées visuo-inertielles (*vSLAM*), qui fusionne les données issues des capteurs inertiels (IMU) et des images afin de calculer des poses caméras fiables.

Le cœur de notre système repose sur l'algorithme HybVIO [44], développé par l'équipe SpectacularAI [43]. Ce système associe un front-end de filtrage basé sur un filtre de Kalman étendu (EKF) et un back-end d'optimisation SLAM, garantissant des estimations robustes des trajectoires caméras grâce à :

- un suivi visuel des caractéristiques locales entre images et la construction d'une carte clairsemée de points d'intérêt,
- la fusion des mesures inertielles (accéléromètre, gyroscope, magnétomètre) avec les indices visuels via l'EKF,
- une optimisation globale des poses pour corriger la dérive accumulée et assurer la cohérence à long terme.

Afin de collecter les données, nous avons développé une application iOS appelée *LiAcquire*, qui capture de manière synchronisée :

- des images et cartes de profondeur à 30 fps,
- des données IMU à environ 99 Hz (accéléromètre, gyroscope et magnétomètre).

Les données collectées sont ensuite traitées via l'API SpectacularAI, qui applique l'algorithme HybVIO pour générer des poses caméras calibrées et corrigées de la dérive. Ces poses, associées aux images correspondantes, servent à l'entraînement

du modèle *depth-nerfacto* dans Nerfstudio [49].

La calibration (intrinsèque et extrinsèque) est effectuée automatiquement lors de la configuration initiale de HybVIO, assurant ainsi la synchronisation précise entre capteurs visuels et inertiels.

La Figure 4.4 illustre le processus de reconstruction 3D par NeRF.

Photogrammétrie

Pour établir une comparaison, nous avons utilisé la photogrammétrie intégrée à l'application *Scaniverse*. Cette méthode exploite des images RGB acquises autour de la structure, en les combinant avec les données issues de l'accéléromètre et du gyroscope pour améliorer l'estimation des poses caméra.

La reconstruction repose sur deux algorithmes fondamentaux :

- **Structure-from-Motion (SfM)** [42] : estimation simultanée des paramètres de la caméra et de la structure 3D en détectant et associant des points caractéristiques entre plusieurs images. Une triangulation épipolaire permet ensuite de générer un nuage de points clairsemé.
- **Multi-View Stereo (MVS)** [45] : densification du nuage de points en estimant la profondeur de chaque pixel à partir des images multi-vues, afin de produire une reconstruction dense et détaillée.

Cette chaîne photogrammétrique est résumée dans la Figure 4.5.

4.2.3. Prétraitement des données

Afin de préparer les nuages de points générés par NeRF et la photogrammétrie pour les étapes d'analyse, plusieurs opérations de prétraitement ont été appliquées. Ces étapes visent à améliorer la qualité des données et à garantir leur compatibilité avec les modules de segmentation et d'évaluation dimensionnelle.

Enregistrement des nuages de points. Lorsque plusieurs séquences sont acquises depuis des angles différents, elles doivent être alignées dans un même système de coordonnées. L'enregistrement permet de fusionner ces scans partiels en un nuage de points cohérent, assurant une couverture complète de la scène.

Suppression statistique des valeurs aberrantes. Des points parasites peuvent apparaître lors des reconstructions 3D, en particulier dans les zones de faible texture

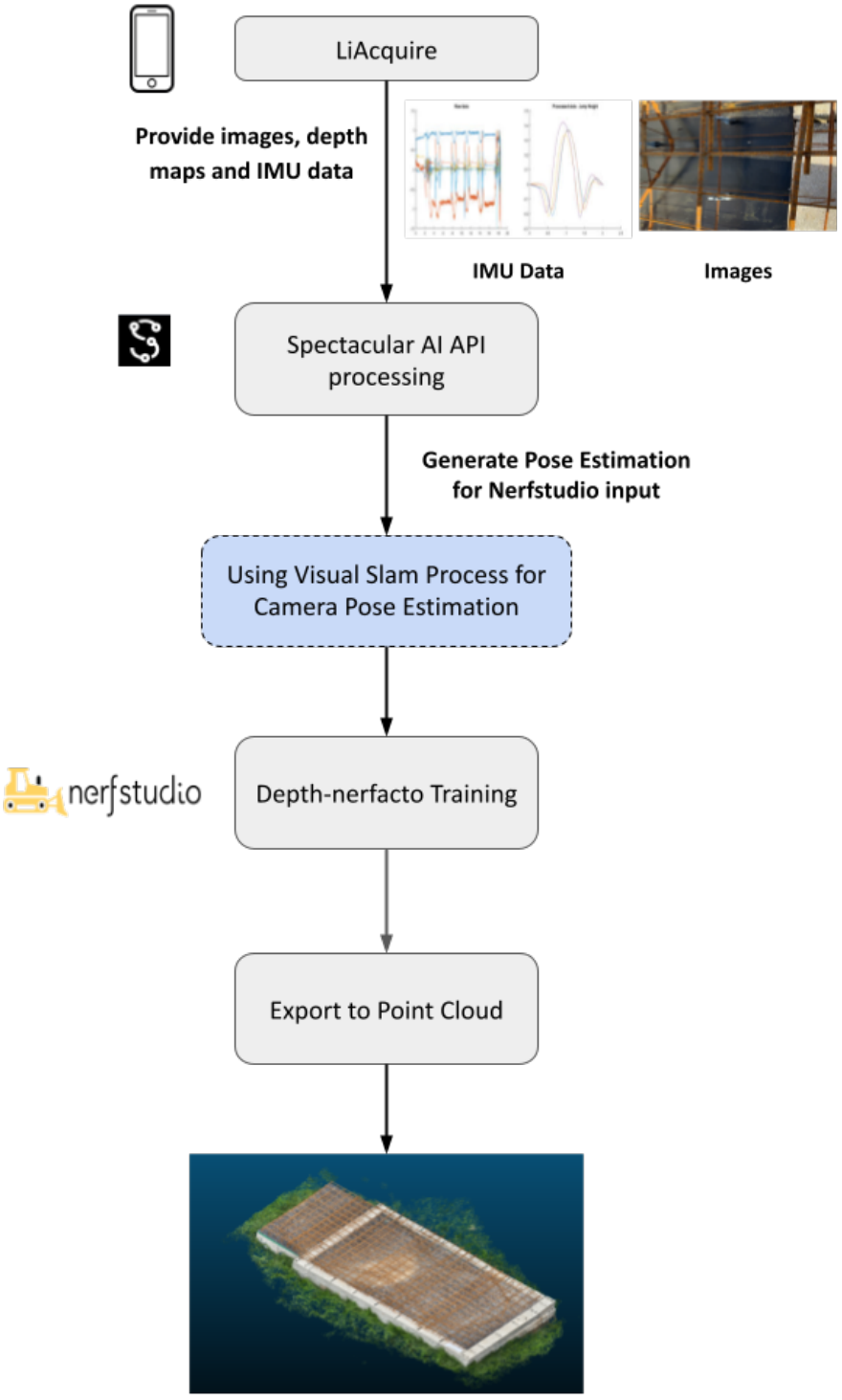


Figure 4.4: Pipeline de reconstruction 3D basé sur NeRF.

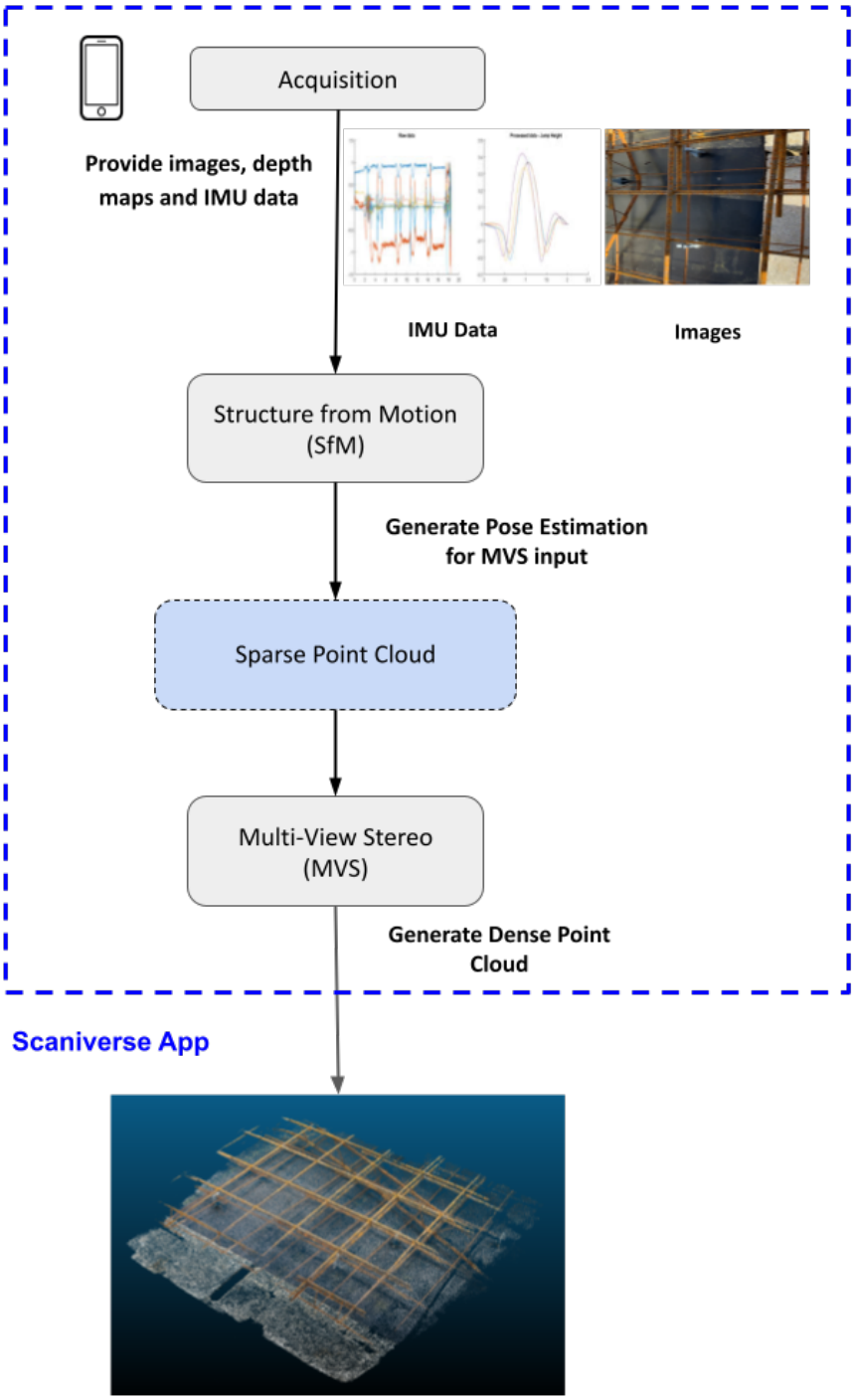


Figure 4.5: Pipeline de reconstruction 3D basé sur la photogrammétrie.

ou de recouvrement. Pour y remédier, une méthode de suppression statistique des outliers a été utilisée, éliminant les points dont la distance aux voisins dépasse un seuil défini par une analyse statistique locale.

Mise à l'échelle et normalisation. Afin de stabiliser l'apprentissage et de faciliter la convergence du modèle de segmentation, les nuages de points sont normalisés dans un espace de coordonnées standardisé. Cette opération réduit la variabilité due à la taille des scènes et assure une représentation plus homogène des données.

Ces opérations constituent une étape indispensable pour garantir que les nuages de points issus des deux approches (NeRF et photogrammétrie) puissent être exploités de manière fiable et efficace par le framework d'inspection.

4.2.4. Module de Reconnaissance de Composants (CRM)

L'identification précise des composants d'armatures constitue une étape clé pour l'évaluation dimensionnelle. En effet, sans une segmentation fiable des différents éléments, la mesure de l'espacement des barres devient imprécise et sujette à des erreurs significatives. Le rôle du CRM est donc de transformer la tâche de reconnaissance des composants en un problème de segmentation sémantique de nuages de points.

Le module CRM s'appuie sur le modèle RandLA-Net [21], choisi pour sa capacité à traiter efficacement des nuages de points de grande taille sans recourir à la voxelisation. Sa structure encodeur-décodeur est conçue pour capturer à la fois l'information locale et le contexte global, tout en maintenant un coût computationnel réduit. Cette efficacité le rend adapté aux données issues de NeRF et de la photogrammétrie, souvent denses et hétérogènes.

Comme illustré dans la Figure 4.6, l'encodeur applique une série d'agrégations locales et d'opérations de sous-échantillonnage, tandis que le décodeur reconstruit progressivement la résolution ponctuelle grâce à des opérations d'interpolation et à des connexions de type *skip connections*. Ces connexions permettent de préserver des détails fins en combinant les caractéristiques extraites à différents niveaux de profondeur.

Dans le cadre de cette contribution, le modèle a été entraîné pour détecter quatre classes :

- **Nappe supérieure** : points correspondant à la nappe de barres située en

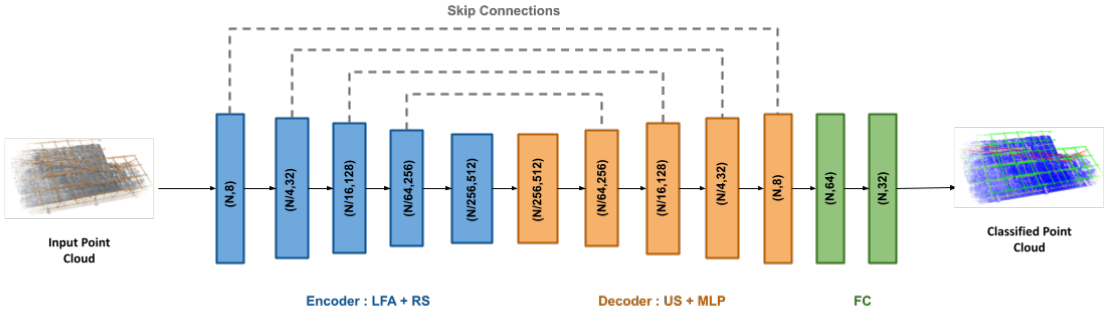


Figure 4.6: Architecture du modèle RandLA-Net utilisé dans le module CRM. N représente le nombre de points. FC : couche entièrement connectée, LFA : agrégation locale de caractéristiques, RS : échantillonnage aléatoire, MLP : perceptron multicouche partagé, US : sur-échantillonnage.

partie haute,

- **Nappe inférieure** : points correspondant à la nappe située en partie basse,
- **Distanciers** : points associés aux éléments maintenant l'écartement entre les nappes,
- **Classe "Autre"** : regroupe l'ensemble des points qui ne correspondent pas aux trois classes précédentes.

Cette catégorisation point par point est indispensable pour isoler les éléments structuraux pertinents, qui serviront ensuite à l'évaluation des espacements dans le module suivant.

Compte tenu du déséquilibre marqué entre les classes, certaines étant surreprésentées tandis que d'autres sont rares, une fonction de perte *Weighted Cross Entropy* (WCE) [8] a été utilisée. Cette approche pénalise davantage les erreurs sur les classes minoritaires, afin d'éviter que le modèle ne privilégie les classes dominantes. La fonction est définie par :

$$WCE(y, \hat{y}) = -\sum_{i} \alpha_i \, p(y_i) \log(p(\hat{y}_i)) \tag{4.1}$$

où :

- y_i et $\hat{y}_i$ représentent respectivement les étiquettes réelles et prédites,

- α_i est le poids associé à la classe i,
- $p(y_i)$ est la probabilité réelle de la classe i,
- $p(\hat{y}_i)$ est la probabilité prédite de la classe i.

Cette optimisation améliore la robustesse du CRM, garantissant une reconnaissance plus équilibrée des composants dans les nuages de points issus de NeRF et de la photogrammétrie.

4.2.5. Module d'Évaluation Dimensionnelle (DAM)

L'évaluation de l'espacement des barres constitue un paramètre critique pour vérifier la conformité et la sécurité des structures en béton armé. Après la segmentation réalisée par le CRM, le DAM exploite le sous-ensemble de points correspondant aux nappes d'armatures pour détecter individuellement les barres et mesurer automatiquement les distances entre elles.

La procédure repose sur un algorithme RANSAC adapté à la **modélisation d'un cylindre**, permettant de modéliser chaque barre comme un cylindre. Cette approche est particulièrement robuste face au bruit et aux données incomplètes, deux caractéristiques fréquentes dans les nuages de points issus de reconstructions 3D. Son fonctionnement itératif consiste à rechercher les modèles géométriques qui maximisent le nombre d'inliers, tout en rejetant les points aberrants. Une fois les cylindres identifiés, le module regroupe les barres en paires parallèles et adjacentes, puis calcule les distances entre elles pour en extraire l'espacement effectif.

La Figure 4.7 illustre la chaîne de traitement mise en place dans le DAM.

L'algorithme 4.1 détaille le processus utilisé pour l'extraction individuelle des barres.

Une fois les cylindres détectés, les paires de barres sont automatiquement identifiées sur la base de leur parallélisme et de leur proximité spatiale. Les distances entre chaque paire sont ensuite calculées et comparées aux tolérances prévues par les normes de construction, permettant de vérifier la conformité de l'espacement.

Cette approche garantit une mesure robuste et automatisée, en s'appuyant sur une modélisation géométrique explicite des barres d'armature.

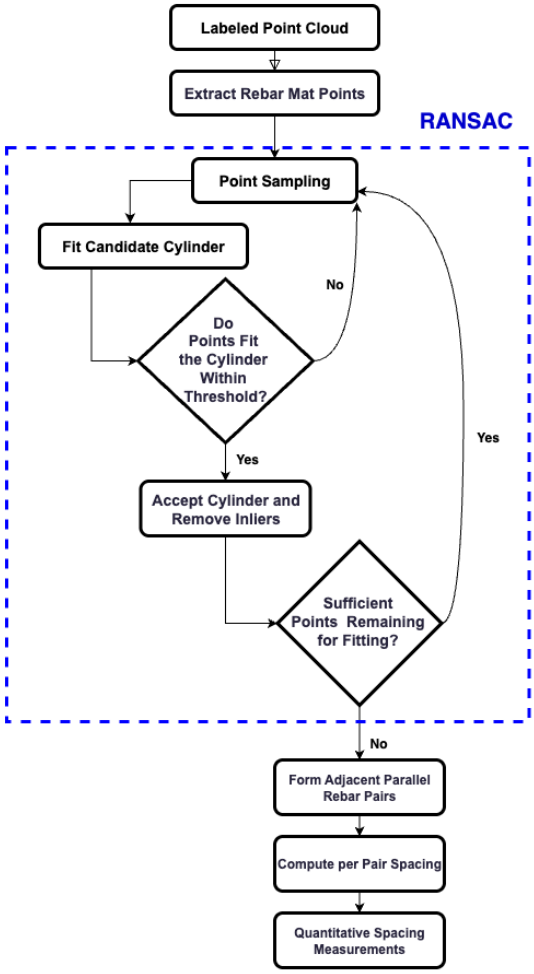


Figure 4.7: Pipeline du module DAM pour l'estimation automatique de l'espacement des armatures.

Algorithm 4.1 Algorithme d'extraction des barres d'armature

Initialisation :
Définir *remaining_points* comme le nuage de points d'entrée.
Créer une liste vide *cylinders* pour stocker les résultats.
while taille de *remaining_points* $\geq$ *min_points_for_cylinder* **do**
  Définir *best_inliers_count* = 0
  Définir *best_cylinder_model* = None
  for chaque itération de 1 à *num_iterations* **do** {ex. 1000 itérations}
    Sélectionner aléatoirement trois points p_1, p_2, p_3 dans *remaining_points*
    Calculer *cylinder_radius*, *cylinder_point* et *cylinder_direction*
    Calculer les distances de tous les points au cylindre défini
    Identifier les inliers dont la distance est inférieure au seuil fixé
    if nombre d'inliers > *best_inliers_count* **then**
      Mettre à jour *best_inliers_count*
      Mettre à jour *best_cylinder_model*
    end if
  end for
  if *best_inliers_count* > *min_points_for_cylinder* **then**
    Ajouter *best_cylinder_model* à *cylinders*
    Supprimer les inliers de *remaining_points*
  else
    break
  end if
end while
return *cylinders* =0

4.3. Expérimentations

Cette section présente une comparaison entre les reconstructions obtenues par NeRF et par photogrammétrie, en évaluant leurs performances respectives en termes de qualité des nuages de points et de temps de reconstruction. Dans un second temps, l'efficacité du framework proposé est analysée à partir des nuages de points issus des deux méthodes, afin de déterminer leur pertinence pour l'évaluation de l'espacement des armatures.

4.3.1. Jeu de données

Un jeu de données équilibré a été constitué afin de garantir des conditions expérimentales identiques : il comprend 100 scans pour la photogrammétrie et 100 scans pour NeRF. En moyenne, les nuages de points générés par la photogrammétrie comptent 666 175 points par scan, tandis que ceux issus de NeRF en contiennent environ 537 731.

Les deux jeux présentent un déséquilibre marqué dans la distribution des classes, comme indiqué dans le Tableau 4.1. Chaque nuage de points est annoté en quatre classes : *Autre*, *Nappe supérieure*, *Nappe inférieure* et *Distancier*. Les attributs utilisés pour l'entraînement et l'évaluation incluent à la fois les coordonnées 3D (x, y, z) et les couleurs (r, g, b).

Table 4.1: Analyse de la distribution des classes : NeRF vs. Photogrammétrie

Classe	**NeRF (%)**	**Photogrammétrie (%)**
Autre	72.59	65.54
Nappe supérieure	18.81	20.86
Nappe inférieure	6.42	10.02
Distancier	2.19	3.58

4.3.2. Configuration expérimentale

L'environnement expérimental comprend une structure en béton armé composée de deux niveaux de nappes de barres, maintenues par des distanciers afin de garantir une séparation correcte entre les couches (voir Figure 4.8). Différentes configurations de diamètres de barres ont été utilisées afin de créer un environnement de test généralisé. Cette variabilité permet d'évaluer l'adaptabilité et la précision du framework proposé pour la mesure des espacements d'armatures dans des conditions



(a) grande structure (b) petite structure

Figure 4.8: Configurations expérimentales.

diverses.

Pour cela, plusieurs maquettes ont été construites avec des dimensions variables :

- **Grandes structures :** environ 4 m × 2.5 m,
- **Petites structures :** environ 1.5 m × 1.5 m.

Toutes les structures ont été montées sur des plateformes surélevées afin de faciliter le processus de numérisation et garantir une visibilité optimale pour une reconstruction 3D précise. L'acquisition des données a été réalisée à l'aide d'un iPhone 14 Pro.

Les nuages de points photogrammétriques ont été générés à l'aide de l'application *Scaniverse* (V4.0.1), tandis que les nuages NeRF ont été produits avec *Nerfstudio* (V1.1.4).

Pour la tâche de segmentation sémantique des nuages de points, l'ensemble des expériences a été conduit sur un GPU Nvidia RTX 4060 avec PyTorch. L'optimisation a été effectuée à l'aide de l'algorithme Adam, avec un taux d'apprentissage initial fixé à $1e^{-2}$ et une taille de lot de 2. Le modèle RandLA-Net a été entraîné pendant 200 époques, avec un facteur de décroissance exponentielle de 0.95 appliqué à chaque époque, en utilisant une taille de nuage d'entrée de 40,960 points et une taille de voisinage $k = 16$.

Pour la tâche d'évaluation de l'espacement des armatures, le modèle RANSAC a été utilisé avec un nombre maximal d'itérations fixé à 1000. Les espacements estimés par le module DAM ont ensuite été comparés à des mesures de terrain collectées à l'aide d'un mètre ruban calibré.

4.3.3. Résultats et discussion

Cette section évalue les méthodes proposées selon plusieurs axes : qualité des nuages de points, temps de reconstruction 3D, précision de la segmentation sémantique et mesure des espacements des armatures. La comparaison entre NeRF et la photogrammétrie permet de mettre en évidence les forces et limites de chaque approche. Les métriques clés retenues incluent la précision, le temps de traitement et l'erreur de mesure.

Qualité des nuages de points

Une comparaison entre les nuages générés par NeRF et ceux issus de la photogrammétrie a été réalisée à l'aide de l'outil *cloud-to-cloud distance* de CloudCompare. Le nuage NeRF a été produit avec *Nerfstudio*, tandis que le nuage photogrammétrique provient de l'application *Scaniverse*. Les deux nuages ont été alignés dans CloudCompare afin d'analyser les écarts spatiaux. L'échelle de couleur (allant de 0.032 mm en bleu à 10 mm en rouge) permet de visualiser précisément les variations de distance entre les deux nuages.

La Figure 4.9 illustre la comparaison des nappes d'armatures. Le nuage photogrammétrique présente une représentation plus dense et précise, capturant efficacement les détails fins des barres. À l'inverse, le nuage NeRF, bien que moins dense et légèrement plus bruité, parvient à restituer certains détails situés aux extrémités des nappes, là où la photogrammétrie montre des lacunes.

La Figure 4.10 montre la comparaison sur la surface en bois supportant la structure. Ici, NeRF propose une représentation plus homogène et continue de la planche, tandis que la photogrammétrie présente des zones clairsemées et incomplètes.

Coût computationnel

La performance computationnelle de la photogrammétrie (Scaniverse) et de la reconstruction NeRF a été évaluée sur différentes surfaces afin d'analyser la scalabilité temporelle. Le pipeline NeRF se compose de trois étapes : (i) acquisition par iPhone, (ii) SLAM visuo-inertiel via l'algorithme SpectacularAI pour reconstruire les poses caméra et l'échelle métrique, et (iii) entraînement NeRF dans Nerfstudio avec un budget par défaut de 30k itérations. Un arrêt anticipé a été testé à 20–40 % du budget d'entraînement, niveau auquel une géométrie exploitable apparaît déjà (les formes globales sont correctes bien que les couleurs restent bruitées), tandis que

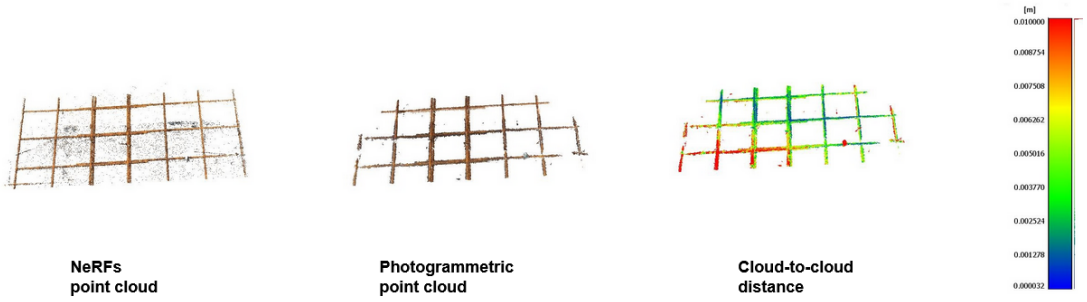


Figure 4.9: Comparaison des nappes d'armatures : nuages NeRF vs. Photogrammétrie.

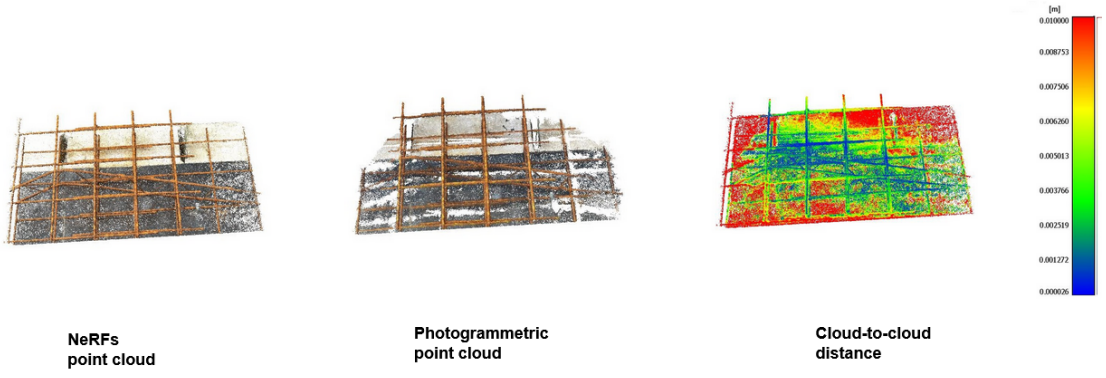


Figure 4.10: Comparaison de la surface de maquette : nuages NeRF vs. Photogrammétrie.

Table 4.2: Répartition des temps de traitement en fonction de la surface pour la photogrammétrie (Scaniverse) et la reconstruction NeRF.

Surface (m^2)	Photogrammétrie (Scaniverse)	NeRF					
	Total (min)	Acq.	vSLAM	Train 100%	Total 20%	Total 40%	Total (100%)
0.6	3	0:50	2:24	29:35	8:19	14:14	32:49
1.0	3–5	1:22	4:03	30:35	10:10	16:17	36:00
2.0	5–10	2:01	6:49	30:07	12:51	18:52	38:56
3.0	–	2:36	10:20	30:43	16:29	22:37	43:03
4.0	–	3:25	16:54	30:48	23:04	29:13	51:07
6.0	–	4:28	29:56	31:15	36:11	42:26	65:39

Note — Les colonnes Total 20% et Total 40% illustrent le coût d'arrêt anticipé sur un entraînement de 30k itérations.

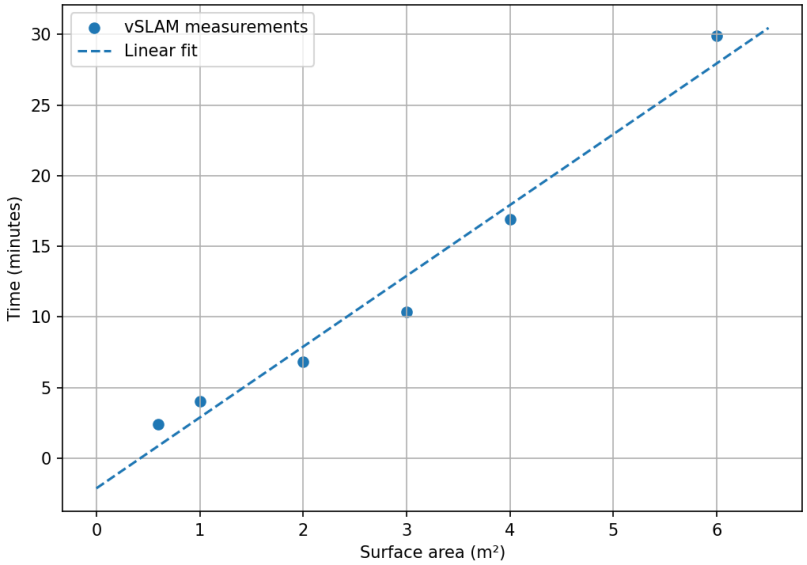


Figure 4.11: Temps vSLAM en fonction de la surface inspectée.

les 100 % permettent principalement de raffiner les détails fins et la cohérence des couleurs.

Sur le plan computationnel, NeRF requiert des ressources équipées de GPU et consomme typiquement 5–6 GB de VRAM durant l'entraînement. Comme indiqué dans le Tableau 4.2, sur l'intervalle testé (0.6–6 m^2) le temps total NeRF est passé d'environ 33 min à 65 min, principalement en raison de la phase vSLAM, dont la durée croît presque linéairement avec la surface inspectée, comme illustré à la Fig. 4.11. La phase d'entraînement reste quant à elle quasi constante (∼30 min pour 30k itérations). Lorsqu'un arrêt anticipé est appliqué (20–40% du budget), le temps total de reconstruction est réduit à ∼8–42 minutes.

À l'inverse, la photogrammétrie via l'application Scaniverse offre un traitement plus rapide (3–10 min) pour de petites surfaces, mais échoue au-delà d'environ 2 m^2 en

raison des limites de calcul et de mémoire embarquées. NeRF, bien que plus coûteux en calcul, demeure stable pour des surfaces de plus grande taille ou plus complexes.

Afin d'optimiser le temps de traitement sur des projets de grande envergure, la zone d'inspection peut être partitionnée en tuiles spatiales de 2–5 m^2 reconstruites en parallèle puis fusionnées.

Tâche de segmentation sémantique

Pour la tâche de segmentation sémantique, nous avons évalué les deux méthodes selon deux métriques : la précision (Accuracy) et l'Intersection/Union (IoU). Les résultats, présentés dans le Tableau 4.3, montrent que la photogrammétrie obtient de meilleures performances globales que NeRF, bien que l'écart reste limité pour certaines classes.

Table 4.3: Comparaison de la précision et de l'IoU : Photogrammétrie vs. NeRF

Classe	**Précision (%)**		**IoU (%)**	
	Photogrammétrie	**NeRF**	**Photogrammétrie**	**NeRF**
Autre	99.63	94.27	98.16	93.39
Nappe supérieure	87.98	86.97	87.57	72.80
Nappe inférieure	85.93	82.35	81.88	65.94
Distancier	86.96	71.14	39.66	30.73
Moyenne/Globale	88.90	83.69	73.65	65.72

Ces résultats indiquent que la photogrammétrie reste globalement plus performante, atteignant une précision globale de 88.9 % et une mIoU de 73.65 %. Toutefois, NeRF obtient des performances proches sur des classes essentielles comme la nappe supérieure (86.97 % de précision contre 87.98 % pour la photogrammétrie), confirmant sa pertinence pour l'identification des éléments structuraux clés.

L'écart le plus marqué est observé pour la classe *distancier*, avec une IoU de seulement 30.73 % pour NeRF contre 39.66 % pour la photogrammétrie. Cette différence s'explique par la complexité géométrique et la petite taille des distanciers, ainsi que par la similarité visuelle avec les nappes, qui accentuent la difficulté de segmentation.

En résumé, si la photogrammétrie conserve une avance en termes de précision globale, NeRF démontre une capacité compétitive, notamment pour la détection des nappes, ce qui confirme son potentiel pour une utilisation pratique dans l'inspection des armatures.

Évaluation de l'espacement des armatures

L'évaluation de l'espacement des armatures a été réalisée en comparant les mesures automatiques issues du module DAM avec des mesures de référence collectées sur chantier à l'aide d'un mètre ruban étalonné. Trois métriques ont été retenues : l'erreur absolue moyenne (MAE), l'erreur quadratique moyenne (MSE) et l'écart-type (SD). Ces métriques permettent de juger non seulement de la précision des mesures mais également de leur stabilité. Les résultats sont présentés dans le Tableau 4.4.

Table 4.4: Comparaison de l'évaluation de l'espacement des armatures selon différentes approches

Métrique	NeRF	Photogrammétrie	LiDAR (ZEB Horizon) [70]	Capteur RGB-D (Azure Kinect) [68]
MAE (cm)	0.92	0.997	1.14	—
MAE (in)	0.36	0.39	0.45	—
MSE (cm^2)	1.11	1.31	—	—
MSE (in^2)	0.44	0.52	—	—
SD (cm)	0.88	1.14	—	0.86
SD (in)	0.35	0.45	—	0.34

Les résultats montrent que NeRF obtient des performances légèrement supérieures à la photogrammétrie. L'erreur absolue moyenne est de 0.92 cm pour NeRF contre 0.997 cm pour la photogrammétrie, avec une erreur quadratique moyenne plus faible (1.11 contre 1.31 cm^2). De plus, NeRF présente une variabilité plus réduite (SD = 0.88 cm contre 1.14 cm), traduisant une meilleure stabilité des estimations.

La différence de MAE entre les deux approches reste faible (moins d'un millimètre). Cependant, la plus faible dispersion observée avec NeRF suggère une plus grande robustesse, ce qui est essentiel dans un contexte opérationnel.

Comparée aux approches de l'état de l'art, la méthode proposée se montre compétitive. Par exemple, le ZEB Horizon associé à l'algorithme SARI [70] atteint un MAE de 1.14 cm, supérieur à celui de NeRF. De même, l'Azure Kinect [68] obtient un SD de 0.86 cm, comparable à celui de NeRF (0.88 cm). Ces résultats confirment que NeRF, en s'appuyant uniquement sur des images RGB acquises par smartphone, peut offrir des performances proches, voire supérieures, à celles de capteurs spécialisés beaucoup plus coûteux.

Discussion générale

L'ensemble des résultats expérimentaux confirme le potentiel de NeRF pour l'inspection automatisée des armatures. Bien que la photogrammétrie conserve un léger avan-

tage en termes de densité et de précision locale des nuages de points, NeRF se distingue par une meilleure homogénéité des reconstructions et une stabilité accrue dans les mesures d'espacement. Cette robustesse est particulièrement notable pour les configurations complexes, où la photogrammétrie tend à produire des zones incomplètes.

Un autre atout majeur de la solution proposée réside dans son accessibilité. Là où les approches traditionnelles reposent sur des capteurs spécialisés tels que les LiDAR portatifs ou les capteurs RGB-D, NeRF exploite uniquement des images RGB issues de smartphones, ce qui réduit considérablement les coûts d'équipement et simplifie le déploiement sur chantier.

Sur le plan computationnel, NeRF impose un coût de traitement significatif mais maîtrisable. Le temps de calcul croît de manière quasi linéaire avec la surface inspectée, principalement en raison de l'étape de vSLAM, tandis que la phase d'entraînement conserve un comportement stable en mémoire et en durée par scène. L'arrêt anticipé à 20–40 % du budget d'itérations peut fournir une géométrie exploitable pour l'inspection, ce qui permet de réduire le temps de traitement.

Le contrôle qualité et l'assurance qualité (QC/QA) des armatures s'appuient généralement sur des standards internationaux établis, tels que l'ACI 117-10 [2], qui définissent des tolérances admissibles pour l'espacement des barres, l'enrobage ou encore l'alignement. Par exemple, l'ACI 117-10 [2] prescrit que l'espacement des armatures doit, en règle générale, respecter une tolérance de ±13 mm (±0,5 in) par rapport à l'espacement de conception. Les évaluations réalisées dans notre cadre montrent que les mesures issues de NeRF se situent systématiquement à l'intérieur de cette tolérance de ±13 mm, validant ainsi la pertinence de cette approche dans des scénarios QC/QA conformes aux pratiques actuelles.

Au-delà de la précision métrique, NeRF, en combinaison avec notre framework, s'intègre naturellement dans les workflows BIM et les processus QC/QA, en apportant un support direct à la prise de décision et à la traçabilité documentaire lors des inspections de terrain. Il permet d'alimenter les chaînes existantes de contrôle qualité et de suivi de chantier avec des informations structurées et exploitables, facilitant notamment la génération de rapports de non-conformité (NCR). Ce lien fonctionnel entre inspection numérique, validation technique et documentation formelle assure une continuité fiable entre le diagnostic sur site et la gouvernance documentaire du projet.

4.4. Conclusion

Cette contribution a exploré l'utilisation des Neural Radiance Fields (NeRF) pour l'inspection des armatures dans les structures en béton armé, en la comparant à la photogrammétrie et en l'évaluant face à plusieurs approches de l'état de l'art. L'objectif principal était de maintenir des outils d'acquisition abordables afin de favoriser l'applicabilité sur le terrain.

Les résultats montrent que NeRF présente des avantages significatifs pour les surfaces de grande taille, en produisant des nuages de points plus complets et plus homogènes que la photogrammétrie. Sur le plan de la segmentation sémantique, les performances obtenues sont proches de celles de la photogrammétrie, confirmant la pertinence de NeRF pour l'analyse de structures de grande dimension. Concernant l'évaluation des espacements, NeRF surpasse légèrement la photogrammétrie en termes de précision et de régularité, avec une erreur absolue moyenne (MAE) et une erreur quadratique moyenne (MSE) plus faibles, ainsi qu'une variabilité réduite dans les mesures.

Comparée à des approches de référence utilisant des capteurs spécialisés (LiDAR portatif ou capteurs RGB-D), NeRF affiche des performances compétitives. L'erreur moyenne observée reste systématiquement en deçà de la tolérance de ± 13 mm (0.5 in) spécifiée par la norme ACI 117-10 [2], attestant de sa fiabilité pour des applications pratiques. Tout en étant une technologie récente encore en évolution, NeRF offre une alternative crédible et économique puisqu'elle repose uniquement sur des images RGB standard, sans recourir à des équipements coûteux.

Ces résultats confirment le potentiel de NeRF comme outil d'inspection des armatures dans le béton armé, particulièrement adapté aux scénarios nécessitant la couverture de larges surfaces. Néanmoins, des validations supplémentaires en conditions réelles, sur des configurations variées de structures et d'environnements, restent nécessaires pour confirmer son applicabilité opérationnelle. Des perspectives de recherche incluent également l'évaluation de variantes plus rapides de NeRF, afin de concilier précision et temps de calcul, et d'ouvrir la voie à une intégration plus fluide dans les pipelines d'inspection automatisée.

5 | Architecture HENS-Net pour la segmentation sémantique efficiente des nuages de points

5.1. Introduction

L'analyse de nuages de points 3D à grande échelle joue un rôle critique dans la compréhension des scènes pour des applications variées, allant de la robotique et de la conduite autonome à la réalité virtuelle [74], mais également dans des contextes industriels tels que l'inspection des structures en béton armé (RC). Dans ce dernier cas, les nuages de points constituent une base de données essentielle pour identifier et caractériser les composants des armatures, permettant par exemple l'évaluation de l'espacement des barres, du recouvrement des nappes ou encore du positionnement des distanciers. Ces tâches exigent des méthodes robustes de segmentation sémantique, capables de traiter efficacement des scènes complexes et de grande taille.

Contrairement aux images 2D structurées, les nuages de points sont constitués de millions de points échantillonnés de manière irrégulière et désordonnée [38], ce qui rend difficile l'application directe des techniques classiques d'apprentissage profond. Ces difficultés sont accentuées par la taille limitée et la diversité des jeux de données 3D existants [63], ce qui impose souvent des compromis entre précision et efficacité computationnelle. De plus, les nuages de points issus de capteurs LiDAR ou d'appareils similaires présentent des densités variables et un caractère aléatoire marqué [22], compliquant davantage leur segmentation sémantique à grande échelle.

Parmi les différentes tâches de traitement 3D, la segmentation sémantique est essentielle puisqu'elle consiste à attribuer une étiquette de classe à chaque point. Dans le cadre du béton armé, cette étape permet de discriminer automatiquement les

nappes supérieures et inférieures, les distanciers ou encore les éléments extérieurs à l'armature, rendant les nuages de points exploitables pour des modules d'évaluation dimensionnelle.

Les approches existantes pour la segmentation sémantique des nuages de points peuvent être regroupées en trois grandes familles : **basées projection**, **basées voxels**, et **basées points**.

Les méthodes basées projection convertissent les données 3D en représentations 2D afin de tirer parti de l'efficacité des réseaux convolutifs 2D, mais elles induisent généralement une perte de détails géométriques [34, 65, 73]. Les approches basées voxels conservent la structure spatiale via des grilles volumiques, mais au prix d'un coût mémoire et computationnel élevé, surtout à haute résolution [11, 50, 76, 77]. Enfin, les méthodes basées points traitent directement les ensembles bruts, permettant une extraction locale précise des caractéristiques. Toutefois, elles s'appuient sur des requêtes coûteuses de voisinage de type k plus proches voisins (k-NN), qui deviennent un goulet d'étranglement lors du passage à grande échelle [22, 36, 37, 62, 74]. Les recherches répétées de type k-NN sur des millions de points exigent une consommation mémoire importante et rallongent fortement les temps de traitement, limitant l'applicabilité temps réel et la scalabilité pour des inspections de grande envergure.

Une piste prometteuse pour atténuer ces contraintes computationnelles est la *sérialisation de nuages de points* à l'aide de courbes de remplissage d'espace, telles que la courbe de Hilbert [9, 30, 63]. Ces courbes imposent un ordre unidimensionnel aux nuages de points tout en préservant la localité spatiale, ce qui facilite une sélection de voisinage plus efficace [9, 30, 63]. En ordonnant les points selon leur position sérialisée, les voisinages locaux peuvent être approximés par de simples fenêtres glissantes, réduisant drastiquement le recours aux calculs géométriques coûteux [9, 30, 63]. Comme l'illustre la Figure 5.1, l'augmentation de la résolution de la courbe de Hilbert permet une couverture plus fine de l'espace 3D, où les points voisins dans le nuage original tendent à rester adjacents dans la séquence sérialisée. Cette propriété permet un traitement plus structuré et scalable des nuages de points, offrant une alternative computationnellement plus légère aux approches classiques basées k-NN.

Dans ce contexte, nous introduisons *HENS-Net* (Hilbert-Efficient Neighborhood Selection Network), une architecture légère conçue pour la segmentation sémantique de nuages de points 3D à grande échelle. HENS-Net exploite la sérialisation Hilbert

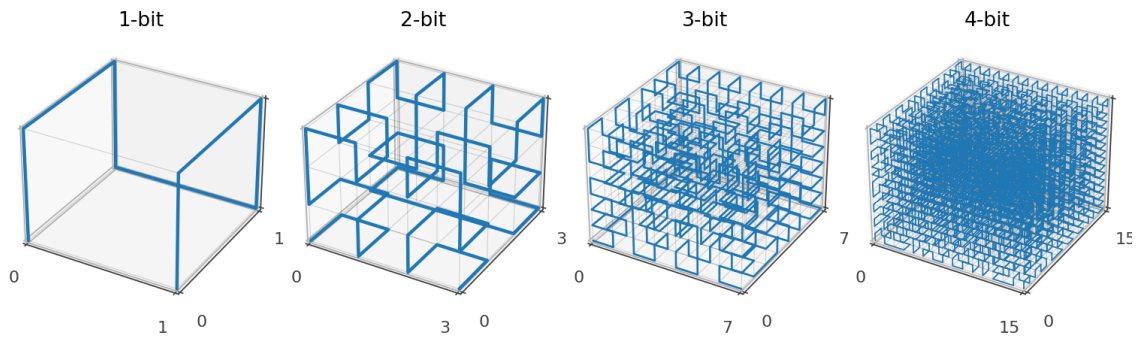


Figure 5.1: Sérialisation de nuages de points avec la courbe de Hilbert à différentes résolutions.

pour imposer un ordre spatialement cohérent, guidant ainsi la sélection de voisinages de manière efficace. Conçue pour équilibrer précision et rapidité, HENS-Net obtient de solides performances sur plusieurs jeux de données tout en maintenant une faible latence de traitement, ce qui la rend adaptée aux scénarios pratiques tels que l'inspection des structures en béton armé.

Les apports spécifiques de ce chapitre peuvent être résumés comme suit :

- Proposition d'une **stratégie de sérialisation Hilbert** préservant la localité spatiale et permettant une sélection de voisinages rapide et scalable, réduisant fortement la dépendance aux recherches globales k-NN.
- Conception d'un **mécanisme hybride de sélection de voisinages** combinant fenêtres glissantes et raffinement local par k-NN, afin d'assurer à la fois efficacité et précision spatiale.
- Introduction d'une architecture légère de type **Encodeur–Décodeur**, intégrant des couches d'agrégation avec encodage spatial et un bloc SEPooling optimisé, conciliant précision de segmentation et rapidité de traitement.
- Validation expérimentale approfondie sur des jeux de données intérieurs et extérieurs, montrant que HENS-Net atteint ou dépasse les méthodes existantes tout en restant léger et efficient. Les évaluations menées sur S3DIS [22] et Toronto3D [48] confirment la robustesse et la généralisabilité de l'approche.

Dans les sections suivantes, nous passons en revue les principales approches de segmentation sémantique 3D. Nous présentons ensuite l'architecture de HENS-Net, en

détaillant la stratégie de sélection des voisinages basée sur Hilbert et le pipeline d'agrégation de caractéristiques. Nous discutons enfin des résultats expérimentaux obtenus sur S3DIS [22] et Toronto3D [48], avant de conclure par des perspectives de recherche futures.

5.2. HENS-Net

Hilbert-Efficient Neighborhood Selection Network (HENS-Net) est une architecture de segmentation sémantique légère, conçue pour le traitement de nuages de points de grande taille. Elle combine les avantages de la sérialisation par courbe de Hilbert, qui permet de préserver la localité spatiale lors de l'ordonnancement des points, avec une structure compacte de type encodeur–décodeur inspirée de RandLA-Net [22].

Le principe consiste à transformer un nuage de points 3D non structuré en une séquence unidimensionnelle suivant l'ordre de la courbe de Hilbert. Cette organisation séquentielle facilite des requêtes de voisinage rapides et cohérentes spatialement, réduisant ainsi la dépendance aux recherches coûteuses de type k-NN et améliorant l'efficacité de l'extraction de caractéristiques.

L'architecture, illustrée à la Figure 5.2, est composée d'un encodeur et d'un décodeur. L'encodeur repose sur une succession de couches d'agrégation, chacune intégrant des modules d'Encodage Spatial et de *SEPooling*, destinés à capturer à la fois les relations géométriques et les informations contextuelles locales. Au fil des étapes, le nuage de points est sous-échantillonné et encodé, ce qui élargit progressivement le champ réceptif. Le décodeur reproduit une structure symétrique, en employant des opérations de sur-échantillonnage et des connexions de saut (*skip connections*) afin de restaurer les détails spatiaux fins nécessaires à une segmentation précise.

Les sous-sections suivantes détaillent les éléments constitutifs de l'architecture proposée, en particulier le mécanisme de sérialisation Hilbert, les couches d'agrégation et le module *SEPooling*.

5.2.1. Sérialisation Hilbert

Au cœur de HENS-Net se trouve une stratégie de sérialisation efficace des nuages de points, qui permet d'accélérer considérablement la sélection des voisinages. Dès l'initialisation, chaque nuage de points brut est converti en une séquence unidi-

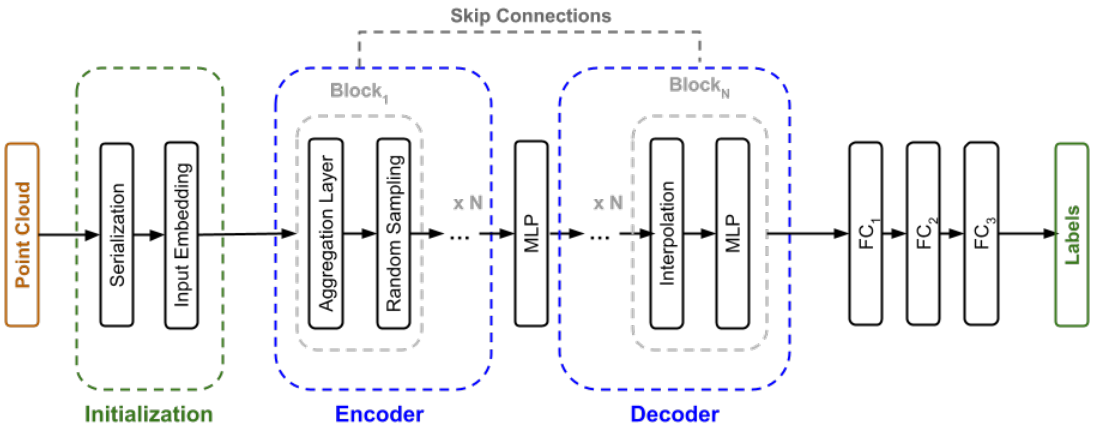


Figure 5.2: Vue d'ensemble de l'architecture de HENS-Net. Le réseau est organisé en plusieurs étapes (N), chacune reposant sur des couches d'agrégation intégrant un encodage spatial et un module de SEPooling. MLP désigne un perceptron multicouche partagé, et FC une couche entièrement connectée.

mensionnelle suivant une courbe de Hilbert tridimensionnelle. L'un des principaux avantages de cette courbe est la préservation de la localité spatiale : deux points proches dans l'espace 3D restent voisins dans l'ordre sérialisé en 1D. En triant les points selon leurs indices Hilbert, on garantit que les régions adjacentes dans l'espace original sont projetées sur des segments contigus. Cette organisation rend possible l'utilisation de fenêtres glissantes de taille fixe pour sélectionner les voisinages locaux, éliminant ainsi le recours systématique à des recherches coûteuses de type k-NN. Ce prétraitement réduit la charge computationnelle et accélère à la fois l'entraînement et l'inférence.

Pour le calcul des indices Hilbert, nous adoptons l'implémentation proposée dans [61].

Quantification. Chaque point $\mathbf{p}_i \in \mathbb{R}^3$ est d'abord normalisé dans un cube borné puis discrétisé en utilisant B bits par axe :

$$\mathbf{q}_i = \left\lfloor \frac{\mathbf{p}_i - \mathbf{p}_{\min}}{\mathbf{p}_{\max} - \mathbf{p}_{\min}} \cdot (2^B - 1) \right\rfloor, \quad \mathbf{q}_i \in \{0, \ldots, 2^B - 1\}^3, \tag{5.1}$$

où $\mathbf{p}_{\min}$ et $\mathbf{p}_{\max}$ correspondent respectivement aux minima et maxima élément-par-élément du nuage de points.

Indexation Hilbert. Chaque point quantifié $\mathbf{q}_i$ est ensuite associé à un indice Hilbert scalaire :

$$h_i = H(\mathbf{q}_i), \quad H : \{0, \ldots, 2^B - 1\}^3 \to \{0, \ldots, 2^{3B} - 1\}, \tag{5.2}$$

où $H(\cdot)$ désigne la fonction d'indexation Hilbert, qui applique une conversion en code Gray, des permutations binaires, et génère une représentation entière compacte (64 bits) de la position du point le long de la courbe de Hilbert. L'indice obtenu h_i sert ensuite de clé pour le tri.

Tri. Enfin, le nuage de points est trié selon les indices Hilbert h_i, produisant un ordre linéaire qui facilite les requêtes de voisinage tout en préservant la cohérence spatiale.

5.2.2. Sélection de voisinage

Un élément central de HENS-Net est sa stratégie efficace de sélection des voisinages, qui combine la sérialisation basée sur la courbe de Hilbert avec une étape de raffinement local. Traditionnellement, la sélection de voisinage dans les nuages de points repose sur des recherches répétées de type k plus proches voisins (k-NN) en distance euclidienne, ce qui est coûteux en calcul, en particulier pour les jeux de données de grande taille. Pour surmonter cette limitation, HENS-Net introduit une approche hybride qui réduit significativement le coût computationnel tout en maintenant la précision spatiale.

Le processus débute par la sérialisation du nuage de points à l'aide des indices Hilbert décrits dans la section précédente. Les points sont d'abord triés en fonction de leurs indices Hilbert, produisant une séquence unidimensionnelle qui préserve la localité spatiale. Cet ordonnancement permet de récupérer rapidement un ensemble de voisins candidats pour chaque point de requête en utilisant une simple fenêtre glissante.

Pour chaque point de requête, une fenêtre de taille fixe est extraite autour de sa position dans la liste triée par indices Hilbert. Cette fenêtre constitue un ensemble initial de voisins candidats. Grâce à la propriété de localité spatiale de la courbe de Hilbert, la majorité de ces candidats sont proches du point de requête dans l'espace 3D. Toutefois, en raison de légères distorsions spatiales induites par la projection 1D, certains candidats peuvent ne pas être géométriquement proches.

Algorithm 5.1 k-NN basé Hilbert (forme essentielle)

Entrée: points $\mathcal{P}$, indices Hilbert $\mathbf{h}$, nombre de voisins k, demi-largeur de fenêtre w
Sortie: liste des voisins $\mathbf{N}$

$\mathbf{s} \leftarrow \text{argsort}(\mathbf{h})$ *(tri par indice Hilbert)*
for $i = 1$ **à** N **do**
  $\mathcal{C} \leftarrow$ indices dans l'intervalle $[i - w, i + w] \setminus \{i\}$ sur $\mathbf{s}$
  $\mathbf{N}_i \leftarrow k$ points les plus proches de $\mathcal{P}_h[i]$ dans $\mathcal{C}$ *(distance euclidienne)*
end for
retourner $\mathbf{N}$ =0

Pour pallier ce problème, une étape de raffinement est appliquée en effectuant une recherche locale k-NN euclidienne uniquement à l'intérieur de la fenêtre de candidats. Cette étape supplémentaire garantit que l'ensemble final de voisins est non seulement obtenu de manière efficace, mais aussi géométriquement pertinent.

La procédure de sélection de voisinage est formalisée dans l'algorithme 5.1.

Ce mécanisme hybride offre un excellent compromis entre efficacité computationnelle et précision spatiale. En réduisant drastiquement l'espace de recherche, il supprime le besoin de requêtes k-NN globales répétées tout en capturant la structure géométrique locale nécessaire à un apprentissage efficace des caractéristiques.

La comparaison visuelle de la figure 5.3 met en évidence l'avantage pratique de la sérialisation Hilbert. Les nuages de points bruts (rangée supérieure) présentent des connexions longues et aléatoires qui masquent les relations spatiales, compliquant ainsi les requêtes de voisinage. À l'inverse, les nuages de points ordonnés par Hilbert (rangée inférieure) conservent une forte cohérence spatiale, avec des points consécutifs proches géométriquement. Cette organisation réduit drastiquement la complexité de recherche lors des requêtes de voisinage, tout en préservant la qualité de l'information spatiale pour une extraction locale efficace des caractéristiques.

5.2.3. Agrégation de caractéristiques

Les couches d'agrégation de caractéristiques constituent les blocs fondamentaux de l'encodeur de HENS-Net. Elles sont responsables de la capture des structures géométriques locales et de leur transformation en représentations riches au niveau des points.

La conception de la couche d'agrégation repose sur deux modules complémentaires :

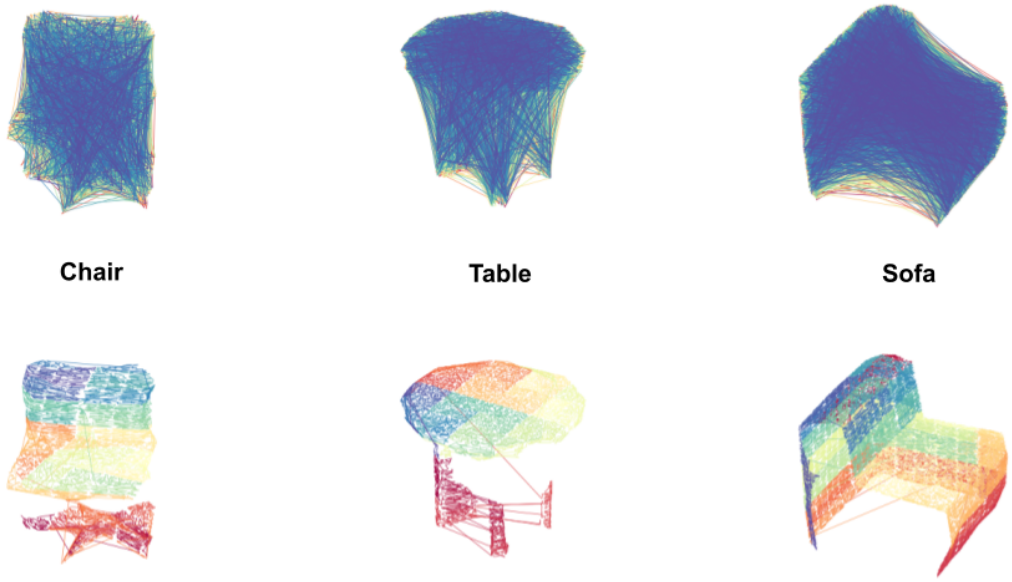


Figure 5.3: Visualisation de l'ordonnancement des points sur trois objets exemples : chaise, table et sofa. La rangée supérieure montre les nuages de points bruts (non ordonnés), où des points consécutifs peuvent être spatialement éloignés, formant des voisinages fragmentés et désorganisés. La rangée inférieure illustre les nuages de points après sérialisation Hilbert, où les points consécutifs restent spatialement cohérents, constituant des voisinages compacts et significatifs.

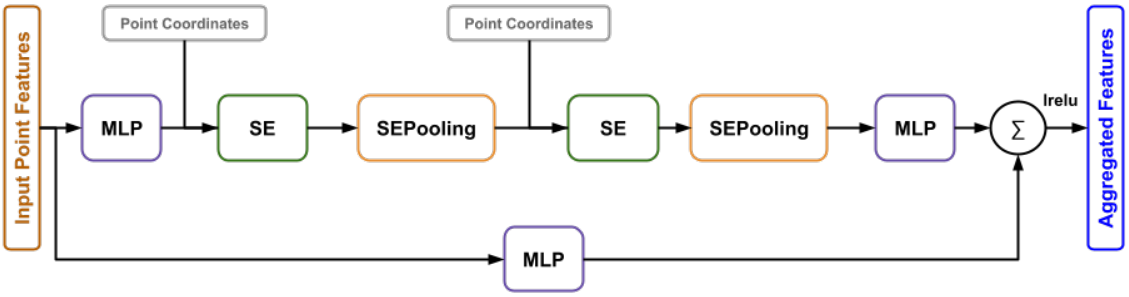


Figure 5.4: Structure de la couche d'agrégation de caractéristiques. MLP désigne un Multi-Layer Perceptron partagé, SE correspond au module de *Spatial Encoding*, et $\sum$ représente l'opérateur de somme.

le module *Spatial Encoding* et le module *SEPooling*. Le module *Spatial Encoding* a pour objectif d'apprendre les relations géométriques entre chaque point et ses voisins, en intégrant à la fois les positions absolues et les déplacements relatifs. Le bloc *SEPooling*, quant à lui, réajuste de manière adaptative et agrège les caractéristiques de voisinage grâce à une combinaison de max-pooling et d'un mécanisme léger de pondération par canal inspiré du principe de *squeeze-and-excitation*.

Enfin, la sortie agrégée est fusionnée avec l'entrée initiale via une connexion résiduelle de type MLP partagé, ce qui permet d'augmenter la dimension des caractéristiques tout en préservant les informations de bas niveau. Cette voie résiduelle contribue à stabiliser l'apprentissage et à favoriser une propagation efficace du gradient dans l'ensemble du réseau. La structure de la couche d'agrégation est illustrée dans la figure 5.4.

Encodage spatial

Le bloc d'encodage spatial vise à capturer la structure géométrique locale autour de chaque point en encodant ses relations avec un nombre fixe de K voisins. Contrairement aux méthodes classiques qui nécessitent des recherches globales coûteuses de type k-NN, HENS-Net exploite l'ordre imposé par la courbe de Hilbert (voir section 5.2.1) pour récupérer efficacement le contexte de voisinage.

Pour chaque point requête $\mathbf{p}_i$ et son k-ième voisin $\mathbf{p}_{i,k}$, on calcule le décalage relatif et la distance euclidienne :

$$\Delta\mathbf{p}_{i,k} = \mathbf{p}_i - \mathbf{p}_{i,k}, \quad d_{i,k} = \|\Delta\mathbf{p}_{i,k}\|_2. \tag{5.3}$$

Ces descripteurs géométriques, combinés aux positions absolues du point requête et de son voisin, sont concaténés dans un vecteur brut :

$$\mathbf{g}_{i,k} = [\,\mathbf{p}_i,\ \mathbf{p}_{i,k},\ \Delta\mathbf{p}_{i,k},\ d_{i,k}\,].$$

Ce vecteur est ensuite passé dans un MLP afin de produire une représentation géométrique compacte :

$$\mathbf{r}_{i,k} = \mathrm{MLP}(\mathbf{g}_{i,k}). \tag{5.4}$$

Pour enrichir la représentation locale, cette représentation géométrique est concaténée avec la caractéristique du voisin $\mathbf{f}_{i,k}$:

$$\mathbf{y}_{i,k} = [\,\mathbf{f}_{i,k},\ \mathbf{r}_{i,k}\,]. \tag{5.5}$$

La caractéristique $\mathbf{y}_{i,k}$ intègre ainsi à la fois le contexte géométrique et les attributs du voisin, et est transmise au bloc SEPooling pour un traitement ultérieur.

SEPooling

Le bloc SEPooling est une stratégie d'agrégation locale à la fois légère et efficace, combinant un max-pooling classique avec un mécanisme de *Squeeze-and-Excitation* afin d'améliorer la pondération des canaux.

Étant donné le tenseur des caractéristiques de voisinage :

$$\mathbf{X} \in \mathbb{R}^{B\times C\times N\times K},$$

où B est la taille du batch, C le nombre de canaux, N le nombre de points et K le nombre de voisins, SEPooling suit quatre étapes principales.

Max-pooling local. On commence par condenser l'information en sélectionnant l'activation maximale par canal :

$$\mathbf{U} = \max_k \mathbf{X}_{:,:,:,k}, \quad \mathbf{U} \in \mathbb{R}^{B\times C\times N}. \tag{5.6}$$

Squeeze global. Pour résumer les statistiques globales des canaux, on calcule la moyenne sur l'ensemble des points :

$$\mathbf{z} = \frac{1}{N} \sum_{n=1}^{N} \mathbf{U}_{:,:,n}. \tag{5.7}$$

Excitation adaptative. Une MLP à deux couches est utilisée pour réajuster les poids des canaux :

$$\mathbf{s} = \sigma\left(W_2 \, \mathrm{ReLU}(W_1 \, \mathbf{z})\right), \tag{5.8}$$

où W_1 réduit la dimension par un facteur r, W_2 restaure la taille initiale des canaux, et σ est la fonction sigmoïde.

Recalibrage. Enfin, les caractéristiques sont recalibrées et projetées dans l'espace de sortie :

$$\mathbf{Y} = \mathrm{MLP}(\mathbf{s} \odot \mathbf{U}), \quad \mathbf{Y} \in \mathbb{R}^{B \times C' \times N \times 1}, \tag{5.9}$$

où $\odot$ désigne une multiplication canal-par-canal et la MLP est implémentée sous forme de convolution 1×1.

Bien que SEPooling soit moins expressif que des mécanismes d'attention plus complexes, il présente l'avantage d'un coût de calcul réduit. L'augmentation progressive du rayon de voisinage dans les couches d'agrégation permet de compenser ce compromis, en élargissant le champ réceptif et en capturant un contexte plus global. Cette conception, à la fois simple et efficace, assure un équilibre favorable entre précision de segmentation et rapidité de traitement.

5.2.4. Fonction de perte

L'entraînement de HENS-Net repose sur une formulation de perte duale combinant l'entropie croisée pondérée et le surrogate Lovász-Softmax [5]. Cette combinaison vise à équilibrer la classification fine au niveau des points individuels avec une optimisation explicite de l'IoU au niveau régional.

La perte totale est définie comme une combinaison convexe :

$$\mathcal{L} = \lambda_{\mathrm{CE}} \, \mathcal{L}_{\mathrm{CE}} + \lambda_{\mathrm{Lov}} \, \mathcal{L}_{\mathrm{Lov}}, \quad \text{avec} \quad \lambda_{\mathrm{CE}} + \lambda_{\mathrm{Lov}} = 1, \tag{5.10}$$

où λ_{Lov} contrôle la contribution du terme d'optimisation de la cohérence régionale,

et λ_{CE} représente la pondération résiduelle assignée au terme de classification.

L'entropie croisée est pondérée par les fréquences inverses des classes afin de limiter l'impact du déséquilibre de labels. Le terme Lovász-Softmax agit sur les sorties normalisées par softmax pour améliorer la cohérence spatiale des prédictions. Dans l'ensemble de nos expériences, nous fixons $\lambda_{\text{Lov}} = 0.5$, ce qui fournit un signal d'apprentissage stable et efficace, particulièrement adapté aux classes minoritaires ou présentant une complexité géométrique élevée.

5.3. Expérimentations

Dans cette section, nous évaluons les performances de HENS-Net sur deux jeux de données de référence en segmentation sémantique 3D : S3DIS pour les environnements intérieurs et Toronto3D pour les scènes urbaines extérieures. Cette combinaison permet d'analyser à la fois la robustesse et la capacité de généralisation de notre approche dans des contextes variés. Nous comparons HENS-Net à plusieurs modèles de l'état de l'art, évalués sur les mêmes jeux de données. Les performances sont rapportées en termes de précision de segmentation et d'efficacité computationnelle.

5.3.1. Jeux de données

S3DIS

Le jeu de données S3DIS [22] contient des scans 3D de six vastes zones intérieures, comprenant 272 pièces annotées avec 13 classes sémantiques telles que le sol, les murs, le plafond ou encore les chaises. Nous adoptons deux protocoles d'évaluation couramment utilisés : (i) la séparation *Area 5*, où le modèle est entraîné sur les zones 1–4 et 6 puis testé sur la zone 5, et (ii) la validation croisée à 6 plis, où chaque zone est tour à tour utilisée comme ensemble de test et la performance finale est obtenue en moyennant les résultats.

Toronto3D

Toronto3D [48] est un jeu de données urbain extérieur acquis à l'aide d'un système LiDAR mobile. Il fournit des annotations denses au niveau des points pour 8 classes sémantiques, incluant notamment les bâtiments, les voitures et les poteaux. Le jeu couvre environ 1 km de route et contient environ 78,3 millions de points. Il se

compose de quatre scènes annotées (L001–L004). Nous suivons le protocole standard : L001, L003 et L004 sont utilisées pour l'entraînement, tandis que L002 sert à l'évaluation.

5.3.2. Configuration expérimentale

L'ensemble des expériences a été réalisé sur un seul GPU NVIDIA RTX A3000. Le modèle HENS-Net a été implémenté avec PyTorch et entraîné à l'aide de l'optimiseur Adam, avec un taux d'apprentissage initial fixé à 0.001. Le nombre de voisins K a été défini à 32, tandis que la taille de la fenêtre candidate est fixée à 256. À chaque itération d'entraînement, 6×10^4 points sont échantillonnés dans chaque nuage de points en entrée. Le taux d'apprentissage décroît d'un facteur de 0.995 après chaque époque, et l'entraînement est mené sur 200 époques.

Afin d'améliorer la capacité de généralisation, des techniques d'augmentation classiques ont été appliquées, incluant des rotations aléatoires et du *jittering*. Chaque nuage de points en entrée est sérialisé au moyen d'une courbe de Hilbert 3D codée sur 16 bits lors de la phase d'initialisation. Pour l'entraînement comme pour le test, seules les coordonnées 3D et les informations de couleur sont utilisées comme descripteurs d'entrée.

Pour l'évaluation des performances, nous adoptons deux métriques standards : l'intersection/union moyenne (mIoU) et la précision globale (OA).

5.3.3. Segmentation sémantique sur S3DIS

Le Tableau 5.1 présente les performances de segmentation sémantique obtenues par HENS-Net sur le benchmark *S3DIS Area 5*. Notre méthode atteint une précision globale (OA) de 90.7% et une mIoU de **75.3%**, dépassant le précédent état de l'art, PTv3 [63], de 0.6%.

Sous le protocole d'évaluation en validation croisée à 6 plis (Tableau 5.2), HENS-Net obtient une mIoU de 77.54%, surpassant la majorité des méthodes récentes, notamment RandLA-Net [22] de 7.5 points. Ces résultats confirment l'efficacité de la stratégie de sélection de voisinage proposée ainsi que du design d'agrégation de caractéristiques adopté.

Table 5.1: Résultats de segmentation sémantique sur S3DIS (Area 5).

Méthode	OA (%)	mIoU (%)
PointNet [36]	78.6	47.6
PointNet++ [37]	81.0	54.5
DGCNN [59]	84.1	56.1
SPG [28]	85.5	62.1
PointCNN [29]	88.1	65.4
KPConv [52]	–	70.6
RandLA-Net [22]	86.0	67.9
SPT [39]	–	68.9
PTv1 [74]	–	70.4
PTv2 [62]	91.24	71.6
PTv3 [63]	**91.96**	74.7
HENS-Net (Proposé)	90.7	**75.3**

5.3.4. Segmentation sémantique sur Toronto3D

Le jeu de données Toronto3D constitue un environnement extérieur particulièrement exigeant, notamment en raison de la présence de classes sous-représentées comme les clôtures, qui compliquent la tâche de segmentation sémantique. Malgré ces difficultés, HENS-Net montre des performances solides. Comme le rapporte le Tableau 5.3, HENS-Net atteint une mIoU de 81.4%, dépassant largement PointNet++ [37], DGCNN [59], KPFCNN [48] et MS-TGNet [48], obtenant des résultats comparables à RandLA-Net [23], et se rapprochant du modèle de tête MappingConvSeg [66]. Ces résultats démontrent que HENS-Net peut fournir des performances compétitives en segmentation dans des environnements urbains extérieurs tout en restant efficace en termes de calcul.

5.3.5. Efficience de HENS-Net

Dans cette expérience, nous avons évalué l'efficience de HENS-Net en termes de temps de traitement total et de nombre de paramètres. La comparaison est effectuée avec RandLA-Net, considéré comme l'un des modèles de référence pour le traitement efficace de nuages de points. Pour cette analyse, cinq nuages de points issus de la zone 5 du jeu de données S3DIS ont été sélectionnés. Le temps de traitement a été mesuré sur les scènes complètes, sans appliquer de sous-échantillonnage. Comme le montre le Tableau 5.4, HENS-Net traite les données en 225 secondes contre 254 secondes pour RandLA-Net, soit une réduction relative d'environ 11%. De plus, HENS-Net contient 1.08 million de paramètres, soit environ 13% de moins que

Table 5.2: Résultats de segmentation sémantique sur S3DIS (validation croisée 6 plis).

Méthode	**OA (%)**	**mIoU (%)**
PointNet [36]	–	41.1
PointNet++ [37]	–	53.5
DGCNN [59]	–	47.9
PointCNN [29]	–	57.3
KPConv [52]	–	70.6
RandLA-Net [22]	88.0	70.0
SPT [39]	–	76.0
PTv1 [74]	–	65.4
PTv2 [62]	90.76	75.17
PTv3 [63]	**92.59**	**80.81**
HENS-Net (Proposé)	90.57	77.54

Table 5.3: Résultats de segmentation sémantique sur Toronto3D.

Méthode	**OA (%)**	**mIoU (%)**
PointNet++ [37]	84.88	41.81
DGCNN [59]	94.24	61.79
KPFCNN [48]	95.39	69.11
MS-TGNet [48]	**95.71**	70.50
RandLA-Net [23]	94.37	81.77
MappingConvSeg [66]	94.72	**82.89**
HENS-Net (Proposé)	90.20	81.40

les 1.24 millions de RandLA-Net.

Cette amélioration de l'efficience repose principalement sur deux choix architecturaux. D'une part, l'adoption d'une stratégie rapide de sélection des voisinages basée sur la courbe de Hilbert, qui accélère la recherche locale et réduit la charge de calcul par rapport aux approches classiques k-NN. D'autre part, l'intégration d'un mécanisme d'agrégation SEPooling simple mais efficace. Contrairement aux stratégies d'attention plus coûteuses, qui nécessitent des opérations softmax complexes, SEPooling reste léger tout en permettant l'exploitation de voisinages plus larges sans compromettre la vitesse de traitement.

5.4. Conclusion

Ce travail propose une nouvelle approche pour la segmentation sémantique efficace de nuages de points 3D à grande échelle. Contrairement à la tendance actuelle con-

Table 5.4: Comparaison du temps de traitement et du nombre de paramètres entre HENS-Net et RandLA-Net [22] sur un sous-ensemble de 5 scènes de S3DIS [22], zone 5.

Méthode	**Temps (s)**	**Paramètres (M)**
RandLA-Net [22]	254	1.24
HENS-Net (Proposé)	**225**	**1.08**

sistant à recourir à des modèles complexes basés sur des mécanismes d'attention, HENS-Net montre qu'il est possible d'obtenir des performances compétitives sans une telle sophistication. L'originalité de l'approche repose sur l'utilisation d'une sérialisation par courbe de Hilbert, qui impose une structure unidimensionnelle préservant la localité sur des nuages de points non ordonnés. Cette organisation permet une sélection rapide et cohérente des voisinages au moyen de fenêtres glissantes, réduisant considérablement les coûts de calcul.

L'architecture s'appuie sur un schéma encodeur–décodeur intégrant des couches d'agrégation combinant encodage spatial et module SEPooling, assurant une agrégation locale à la fois légère et efficace. L'objectif poursuivi est de concevoir un modèle conciliant efficience computationnelle et précision de segmentation pour l'analyse de scènes 3D à grande échelle.

Les résultats expérimentaux obtenus sur S3DIS et Toronto3D confirment la pertinence de l'approche : HENS-Net atteint une précision de segmentation comparable, voire supérieure, à celle des méthodes de référence, tout en réduisant la taille du modèle et le temps global de traitement. Ces résultats mettent en évidence le potentiel de la sérialisation des nuages de points comme levier vers une compréhension 3D en temps réel.

Au-delà des validations expérimentales, cette contribution s'inscrit directement dans le contexte général de cette thèse consacré à l'inspection automatisée des structures en béton armé. En proposant un modèle léger et scalable, HENS-Net constitue un pas vers des solutions capables de traiter efficacement de larges nuages de points acquis sur le terrain, tout en restant compatibles avec des contraintes opérationnelles (coût, rapidité, robustesse, conformité). Ce positionnement rapproche les avancées académiques en vision 3D des besoins pratiques de l'ingénierie civile, ouvrant la voie à des applications concrètes en inspection et suivi de chantier.

En perspective, il sera intéressant d'évaluer HENS-Net sur d'autres jeux de don-

nées afin de mieux cerner ses capacités de généralisation. Par ailleurs, des travaux récents explorent la combinaison de stratégies de sérialisation avec des réseaux de neurones graphiques. L'intégration de ces approches pourrait ouvrir la voie à des représentations de voisinage plus adaptatives et expressives dans le cadre de notre méthodologie.

6 | Discussion

L'inspection automatisée des structures en béton armé (BA) à partir de données 3D constitue un enjeu scientifique et technologique majeur, répondant à des besoins croissants en matière de contrôle qualité (QC/QA), de sécurité et de durabilité dans le secteur de la construction. Les travaux présentés dans cette thèse ont proposé trois axes complémentaires qui contribuent à combler des lacunes identifiées dans l'état de l'art. Ces avancées s'inscrivent dans la continuité des efforts entrepris pour démocratiser l'inspection des armatures, améliorer la précision de la modélisation 3D et développer des architectures de segmentation plus efficiente. Nous revenons dans ce chapitre sur les apports de ces trois contributions, en soulignant leurs améliorations par rapport aux travaux existants et les perspectives qu'elles ouvrent.

Une première avancée majeure réside dans la conception d'un framework d'inspection à faible coût basé sur le LiDAR intégré aux smartphones. Jusqu'ici, la majorité des approches reposaient soit sur des méthodes manuelles, chronophages et risquées, soit sur des dispositifs spécialisés comme le LiDAR terrestre [7, 15, 69–72] et le GPR [14], dont le coût et la complexité d'utilisation limitent la diffusion à grande échelle. Dans ce contexte, notre contribution démontre que le LiDAR embarqué sur un iPhone, bien que moins précis qu'un capteur industriel, peut fournir des données exploitables lorsqu'il est intégré dans un pipeline de traitement robuste. En particulier, la combinaison d'un module de reconnaissance des composants (CRM), basé sur un modèle de segmentation sémantique, et d'un module d'évaluation dimensionnelle (DAM), reposant sur un algorithme RANSAC, permet d'obtenir des mesures fiables de l'espacement des barres d'armature. Les résultats obtenus montrent que cette approche peut satisfaire les tolérances spécifiées dans des standards comme l'ACI 117-10 [2], tout en réduisant considérablement les coûts associés. Ainsi, cette contribution illustre une voie vers la démocratisation de l'inspection automatisée, en ouvrant la possibilité d'un déploiement à grande échelle sur chantier.

La deuxième contribution s'attache à dépasser certaines limites observées dans notre première approche, notamment en termes de complétude et de fidélité des nuages de

points générés par le LiDAR iPhone. Dans ce but, nous avons exploré le potentiel des Neural Radiance Fields (NeRF), une méthode récente de reconstruction 3D basée sur l'apprentissage profond, encore largement inexploitée dans le domaine de l'ingénierie civile. En comparant NeRF à une méthode établie comme la photogrammétrie, nous avons montré que NeRF est capable de produire des reconstructions plus uniformes et plus cohérentes sur de larges surfaces, avec une capacité accrue à restituer des détails géométriques fins. Si la photogrammétrie conserve un léger avantage en termes de densité locale et de rapidité sur de petites zones, NeRF se distingue par sa stabilité et sa robustesse lorsque la surface à inspecter devient plus étendue. De plus, l'intégration de NeRF dans notre framework a démontré sa capacité à fournir des mesures d'espacement précises, avec un écart absolu moyen inférieur à la tolérance de ±13 mm fixée par les standards. Ces résultats positionnent NeRF comme une alternative crédible aux méthodes traditionnelles. Cette contribution élargit donc le champ des possibilités pour l'inspection des structures en BA, en ouvrant la voie à des solutions hybrides combinant photogrammétrie et NeRF en fonction des contraintes de terrain.

La troisième contribution s'attaque à une problématique transversale, au cœur de toutes les approches de traitement de données nuages de points : la segmentation sémantique de nuages de points à grande échelle. La majorité des méthodes existantes, qu'elles soient basées sur des projections [34, 60, 73], des voxels [11, 76, 77] ou un traitement direct des points [22, 29, 52, 59, 62, 74], doivent composer avec un compromis difficile entre précision et efficience. Les approches voxelisées souffrent d'une explosion des coûts mémoire, les méthodes par projection entraînent une perte d'information géométrique, tandis que les modèles point-based, bien que précis, sont souvent freinés par la lourdeur des requêtes k-NN répétées. Pour répondre à ce défi, nous avons proposé HENS-Net (Hilbert-Efficient Neighborhood Selection Network), une architecture légère qui repose sur la sérialisation des nuages de points par courbe de Hilbert. Ce procédé de transformation impose un ordre unidimensionnel préservant la localité spatiale, permettant de sélectionner efficacement les voisins via des fenêtres glissantes, complétées par un raffinement local en k-NN. Cette stratégie permet à HENS-Net de rivaliser avec des méthodes de pointe comme PTv3 [63] ou RandLA-Net [22], tout en réduisant le temps de traitement et le nombre de paramètres. Les résultats obtenus sur S3DIS et Toronto3D confirment la capacité de HENS-Net à généraliser sur des scènes variées, en atteignant des performances comparables aux modèles les plus récents mais avec une empreinte computationnelle significativement allégée. Cette contribution démontre que la sérialisation constitue

une alternative prometteuse aux requêtes géométriques traditionnelles, ouvrant des perspectives pour des architectures à la fois plus légères et plus adaptées à des environnements contraints.

Pris dans leur ensemble, ces trois axes constituent une réponse cohérente aux principaux défis identifiés dans l'état de l'art. L'intégration du LiDAR d'iPhone au sein d'un framework dédié permet d'explorer la faisabilité d'une inspection reposant sur des capteurs accessibles tout en restant conforme aux exigences du génie civil. L'exploration des NeRFs apporte une solution innovante pour la modélisation 3D, ouvrant la voie à des inspections plus riches et visuellement détaillées. Enfin, HENS-Net apporte une nouvelle architecture proposant une nouvelle manière de traiter efficacement des nuages de points à grande échelle.

Ces apports ne doivent pas occulter certaines limites qui appellent de futurs développements. Concernant l'inspection à partir du LiDAR iPhone, si la précision obtenue est satisfaisante pour l'évaluation des espacements, d'autres paramètres comme le recouvrement des nappes ou l'épaisseur de l'enrobage nécessiteront des investigations supplémentaires. Pour NeRF, malgré ses performances prometteuses, le temps de reconstruction reste un frein pour une adoption généralisée en contexte chantier ; l'intégration de variantes plus rapides comme Instant-NGP constitue une piste intéressante. Quant à HENS-Net, bien qu'efficace sur deux benchmarks de référence, sa robustesse face à d'autres contextes reste à valider.

En conclusion, cette thèse propose une approche intégrée qui, de la capture au traitement des données, vise à rendre l'inspection des structures en béton armé plus accessible, plus fiable et plus efficiente. Les contributions développées apportent à la fois des solutions pratiques et des avancées méthodologiques, et ouvrent de nouvelles perspectives de recherche à l'intersection de la vision 3D, de l'apprentissage profond et du génie civil.

7 | Conclusion

L'inspection des structures en béton armé est un enjeu central pour la sécurité, la durabilité et la conformité des infrastructures. Historiquement réalisée à travers des méthodes manuelles ou des technologies spécialisées coûteuses, elle reste aujourd'hui limitée par des contraintes opérationnelles, économiques et humaines. Dans ce contexte, cette thèse a proposé un framework complet d'inspection automatique s'appuyant sur les avancées de la vision 3D et de l'apprentissage profond, avec l'ambition de rendre ce processus plus fiable, plus accessible et plus efficient. L'objectif était de répondre à trois défis majeurs identifiés dans l'état de l'art : (i) définir une solution capable d'exploiter de manière fiable des données issues de capteurs accessibles, en particulier le LiDAR embarqué sur smartphone, (ii) explorer des méthodes de reconstruction 3D récentes, telles que NeRF, afin d'améliorer la cohérence et la qualité des modèles 3D utilisés pour le contrôle dimensionnel, et (iii) concevoir une architecture de segmentation 3D efficiente, capable de traiter efficacement de grands volumes de données sans compromis sur la précision.

La première contribution a porté sur la conception d'un framework complet d'inspection automatique des armatures, évalué en particulier sur des données acquises via le LiDAR embarqué sur iPhone. À l'inverse des approches existantes reposant sur des dispositifs professionnels lourds et coûteux, ce framework montre qu'un capteur grand public peut être intégré dans un processus d'inspection automatique des armatures. Il s'articule autour de deux modules complémentaires : le Component Recognition Module (CRM), dédié à la segmentation sémantique des composants des armatures, et le Dimensional Assessment Module (DAM), chargé d'évaluer automatiquement l'espacement entre les barres au moyen d'algorithmes géométriques robustes. Les expérimentations ont confirmé la pertinence de ce pipeline, avec une précision globale de segmentation de 82,1 %, une mIoU de 71,8 % et une erreur absolue moyenne de 0,9 cm pour l'évaluation des espacements, respectant la tolérance de $\pm$13 mm spécifiée par l'ACI 117-10.

La deuxième contribution s'est intéressée à la qualité des reconstructions 3D, en

explorant l'usage des NeRF comme alternative ou complément aux approches classiques, notamment la photogrammétrie. Comparativement à cette dernière, NeRF produit des nuages de points plus uniformes et stables, particulièrement sur de larges surfaces. Pour l'évaluation des espacements, NeRF a obtenu une erreur absolue moyenne de 0,92 cm, légèrement inférieure à celle de la photogrammétrie (0,997 cm), avec une variabilité réduite (SD = 0,88 cm contre 1,14 cm). Ces performances sont systématiquement restées en conformité avec la tolérance définie par l'ACI 117-10, démontrant que NeRF peut être une alternative crédible tout en utilisant uniquement des images RGB capturées avec un smartphone. Bien que son temps de traitement soit actuellement plus long, l'intégration de variantes plus rapides, telles qu'Instant-NGP, ouvre la perspective d'une adoption pratique sur le terrain.

La troisième contribution a visé à relever le défi de l'analyse efficiente des nuages de points. Les approches existantes présentent des limites en termes de coût mémoire et de complexité des recherches de voisinage. Pour répondre à ce problème, nous avons proposé HENS-Net, une architecture légère et efficiente exploitant la sérialisation des nuages de points via des courbes de Hilbert pour structurer les données et réduire le coût des recherches locales. Sur les benchmarks S3DIS et Toronto3D, HENS-Net a démontré une performance de segmentation compétitive, comparable voire supérieure aux modèles de référence, confirmant son efficacité et sa capacité de généralisation sur des environnements intérieurs et extérieurs variés. En termes d'efficience, HENS-Net réduit le temps de traitement de 11% et le nombre de paramètres de 13% par rapport à RandLA-Net.

Malgré ces avancées, certaines limites subsistent et définissent des perspectives claires pour de futurs travaux. La généralisation à des configurations d'armatures plus variées, dans des environnements fortement encombrés ou soumis à des conditions de capture hétérogènes, devra être validée par des campagnes d'évaluation plus larges et diversifiées. L'accélération des reconstructions NeRF, grâce à des variantes plus rapides, est essentielle pour un usage pratique. L'extension de l'analyse au-delà de l'espacement, vers des paramètres tels que le recouvrement des nappes, l'enrobage ou l'alignement nécessitera de nouveaux modules géométriques et une prise en compte explicite de la précision et des tolérances de mesure.

En conclusion, cette thèse apporte une contribution significative à l'inspection automatisée des armatures en démontrant la possibilité de concilier accessibilité des données, fiabilité des traitements et efficience de l'analyse au sein d'un même pipeline. En combinant les trois contributions, ce travail établit les fondations d'un framework

opérationnel pour le contrôle qualité en génie civil.

Bibliographie

[1] M. A. Almalki and K. F. Almutairi. Inspection of reinforced concrete structures using ground penetrating radar : Experimental approach. *Journal of King Saud University – Science*, 36 :103140, 2024. doi : 10.1016/j.jksus.2024.103140.

[2] American Concrete Institute. *Specification for tolerances for concrete construction and materials.* American Concrete Institute, Farmington Hills, MI, USA, aci 117-10 edition, 2015.

[3] F. Bahreini and A. Hammad. Dynamic graph CNN-based semantic segmentation of concrete defects and as-inspected modeling. *Automation in Construction*, 159 :105282, 2024. doi : 10.1016/j.autcon.2024.105282.

[4] J. T. Barron, B. Mildenhall, M. Tancik, P. Hedman, R. Martin-Brualla, and P. P. Srinivasan. Mip-NeRF : A multiscale representation for anti-aliasing neural radiance fields. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 5855–5864, 2021. doi : 10.1109/ICCV48922.2021.00580.

[5] M. Berman, A. R. Triki, and M. B. Blaschko. The lovasz-softmax loss : A tractable surrogate for the optimization of the intersection-over-union measure in neural networks. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4413–4421, 2018. doi : 10.1109/CVPR.2018.00464.

[6] N. Bolourian, M. Nasrollahi, F. Bahreini, and A. Hammad. Point cloud–based concrete surface defect semantic segmentation. *Journal of Computing in Civil Engineering*, 37(2) :04022056, 2023. doi : 10.1061/JCCEE5.CPENG-5009. URL `https://ascelibrary.org/doi/abs/10.1061/JCCEE5.CPENG-5009`.

[7] P. M. Calvi, E. Che, T. Sweet, L. N. Lowes, and J. W. Berman. Data collection using terrestrial laser scanners from the shake-table test of a full-scale reinforced concrete building. *Journal of Structural Engineering*, 150(2) :04023219, 2024. doi : 10.1061/JSENDH.STENG-12627.

[8] J. Chen, Y. Zhao, C. Meng, and Y. Liu. Multi-feature aggregation for semantic segmentation of an urban scene point cloud. *Remote Sensing*, 14(20) :5134, 2022. doi : 10.3390/rs14205134.

[9] W. Chen, X. Zhu, G. Chen, and B. Yu. Efficient point cloud analysis using hilbert curve. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 730–747, 2022.

[10] H.-L. Chi, M.-K. Kim, K.-Z. Liu, J. Thedja, J. Seo, and D.-E. Lee. Rebar inspection integrating augmented reality and laser scanning. *Automation in Construction*, 136 :104183, 2022. ISSN 0926-5805. doi : https://doi.org/10.1016/j.autcon.2022.104183. URL `https://www.sciencedirect.com/science/article/pii/S0926580522000565`.

[11] C. B. Choy, J. Gwak, and S. Savarese. 4d spatio-temporal convnets : Minkowski convolutional neural networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3075–3084, 2019.

[12] D. Costantino, G. Vozza, M. Pepe, and V. S. Alfio. Smartphone lidar technologies for surveying and reality modelling in urban scenarios : Evaluation methods, performance and challenges. *Applied System Innovation*, 5(4), 2022. ISSN 2571-5577. doi : 10.3390/asi5040063. URL `https://www.mdpi.com/2571-5577/5/4/63`.

[13] V. Croce, D. Billi, G. Caroti, A. Piemonte, L. De Luca, and P. Véron. Comparative assessment of neural radiance fields and photogrammetry in digital heritage : Impact of varying image conditions on 3D reconstruction. *Remote Sensing*, 16(2) :301, 2024. doi : 10.3390/rs16020301.

[14] K. Dinh, N. Gucunski, and T. H. Duong. Migration-based automated rebar picking for condition assessment of concrete bridge decks with ground penetrating radar. *NDT E International*, 98 :45–54, 2018. ISSN 0963-8695. doi : https://doi.org/10.1016/j.ndteint.2018.04.009. URL `https://www.sciencedirect.com/science/article/pii/S0963869517306308`.

[15] M. Ebrahimi, H. Hojat Jalali, and S. Sabatino. Probabilistic condition assessment of reinforced concrete sanitary sewer pipelines using LiDAR inspection data. *Automation in Construction*, 150 :104857, 2023. doi : 10.1016/j.autcon.2023.104857.

[16] N. Faris, T. Zayed, E. M. Abdelkader, and A. Fares. Corrosion assessment using

ground penetrating radar in reinforced concrete structures : Influential factors and analysis methods. *Automation in Construction*, 156 :105130, 2023. doi : 10.1016/j.autcon.2023.105130.

[17] B. Graham, M. Engelcke, and L. Van Der Maaten. 3d semantic segmentation with submanifold sparse convolutional networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 9224–9232, 2018.

[18] S. Hachisuka, A. Tono, and M. Fisher. Harbingers of NeRF-to-BIM : A case study of semantic segmentation on building structure with neural radiance fields. In *Proceedings of the 2023 European Conference on Computing in Construction and the 40th International CIB W78 Conference*, 2023. doi : 10.35490/EC3.2023.284.

[19] K. Hong, H. Wang, and B. Yuan. Inspection-NeRF : Rendering multi-type local images for dam surface inspection task using climbing robot and neural radiance field. *Buildings*, 13(1) :213, 2023. doi : 10.3390/buildings13010213.

[20] Q. Hu, B. Yang, L. Xie, S. Rosa, Y. Guo, Z. Wang, N. Trigoni, and A. Markham. Randla-net : Efficient semantic segmentation of large-scale point clouds. *CoRR*, abs/1911.11236, 2019. URL `http://arxiv.org/abs/1911.11236`.

[21] Q. Hu, B. Yang, L. Xie, S. Rosa, Y. Guo, Z. Wang, N. Trigoni, and A. Markham. RandLA-Net : Efficient semantic segmentation of large-scale point clouds. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 11108–11117, 2020.

[22] Q. Hu, B. Yang, L. Xie, S. Rosa, Y. Guo, Z. Wang, N. Trigoni, and A. Markham. RandLA-Net : Efficient semantic segmentation of large-scale point clouds. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 11108–11117, 2020.

[23] Q. Hu, B. Yang, L. Xie, S. Rosa, Y. Guo, Z. Wang, N. Trigoni, and A. Markham. Learning semantic segmentation of large-scale point clouds with random sampling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(11) :8338–8354, 2022. doi : 10.1109/TPAMI.2021.3083288.

[24] M. Jia and M. Zhang. An overview of methods and applications of 3D reconstruction. *International Journal of Computer Science and Information Technology*, 3(1) :16–23, 2024. doi : 10.62051/ijcsit.v3n1.03.

[25] J. Kim, D. Kim, and H. Youn. An alternative approach for detecting cavities in reinforced concrete structures using GPR A-scan data. *Developments in the Built Environment*, 19 :100479, 2024. doi : 10.1016/j.dibe.2024.100479.

[26] M. Kim, J. P. P. Thedja, H. Chi, and D. Lee. Automated rebar diameter classification using point cloud data based machine learning. *Automation in Construction*, 122 :103476, 2021. doi : 10.1016/j.autcon.2020.103476.

[27] M.-K. Kim, J. P. P. Thedja, H.-L. Chi, and D.-E. Lee. Automated rebar diameter classification using point cloud data based machine learning. *Automation in Construction*, 122 :103476, 2021. ISSN 0926-5805. doi : https://doi.org/10.1016/j.autcon.2020.103476. URL `https://www.sciencedirect.com/science/article/pii/S0926580520310566`.

[28] L. Landrieu and M. Simonovsky. Large-scale point cloud semantic segmentation with superpoint graphs. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4558–4567, 2018. doi : 10.1109/CVPR.2018.00479.

[29] Y. Li, R. Bu, M. Sun, W. Wu, X. Di, and B. Chen. Pointcnn : convolution on $\mathcal{X}$-transformed points. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, NIPS'18, page 828–838, Red Hook, NY, USA, 2018. Curran Associates Inc.

[30] D. Liang, X. Zhou, W. Xu, X. Zhu, Z. Zou, X. Ye, X. Tan, and X. Bai. Pointmamba : A simple state space model for point cloud analysis. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2024.

[31] X. Ma, H. Liu, M. L. Wang, and R. Birken. Automatic detection of steel rebar in bridge decks from ground penetrating radar data. *Journal of Applied Geophysics*, 158 :93–102, 2018. ISSN 0926-9851. doi : https://doi.org/10.1016/j.jappgeo.2018.07.010. URL `https://www.sciencedirect.com/science/article/pii/S0926985117310510`.

[32] X. Ma, Z. He, Z. Xu, W. Wang, and J. Xiao. Machine vision-based surface inspection system for rebar. In *2022 IEEE 10th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)*, pages 2634–2637, 2022. doi : 10.1109/ITAIC54216.2022.9836480.

[33] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and

R. Ng. NeRF : Representing scenes as neural radiance fields for view synthesis, 2020. URL `https://arxiv.org/abs/2003.08934`.

[34] A. Milioto, I. Vizzo, J. Behley, and C. Stachniss. RangeNet++ : Fast and accurate lidar semantic segmentation. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4213–4220, 2019.

[35] T. Müller, A. Evans, C. Schied, and A. Keller. Instant neural graphics primitives with a multiresolution hash encoding. *ACM Transactions on Graphics*, 41(4) : 102 :1–102 :15, 2022. doi : 10.1145/3528223.3530127.

[36] C. R. Qi, H. Su, K. Mo, and L. J. Guibas. Pointnet : Deep learning on point sets for 3d classification and segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 652–660, 2017.

[37] C. R. Qi, L. Yi, H. Su, and L. J. Guibas. PointNet++ : Deep hierarchical feature learning on point sets in a metric space. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 5105–5114, 2017.

[38] F. Remondino et al. Semantic segmentation of large-scale outdoor point clouds by encoder–decoder shared mlps with multiple losses. *Remote Sensing*, 13(16) : 3121, 2021.

[39] D. Robert, H. Raguet, and L. Landrieu. Efficient 3d semantic segmentation with superpoint transformer. In *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 17149–17158, 2023. doi : 10.1109/ICCV51070.2023.01577.

[40] W. Rutkowski and T. Lipecki. Use of the iphone 13 pro lidar scanner for inspection and measurement in the mineshaft sinking process. *Remote Sensing*, 15(21), 2023. ISSN 2072-4292. doi : 10.3390/rs15215089. URL `https://www.mdpi.com/2072-4292/15/21/5089`.

[41] W. Rutkowski and T. Lipecki. Use of the iphone 13 pro LiDAR scanner for inspection and measurement in the mineshaft sinking process. *Remote Sensing*, 15(21) :5089, 2023. doi : 10.3390/rs15215089.

[42] J. L. Schönberger and J. Frahm. Structure-from-motion revisited. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4104–4113, 2016. doi : 10.1109/CVPR.2016.445.

[43] O. Seiskari, J. Ylilammi, V. Kaatrasalo, P. Rantalankila, M. Turkulainen,

J. Kannala, and A. Solin. Spectacular AI mapping tools. URL `https://spectacularai.github.io/docs/sdk/tools/nerf.html`.

[44] O. Seiskari, P. Rantalankila, J. Kannala, J. Ylilammi, E. Rahtu, and A. Solin. HybVIO : Pushing the limits of real-time visual-inertial odometry. In *2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pages 287–296, 2022. doi : 10.1109/WACV51458.2022.00036.

[45] S. M. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski. A comparison and evaluation of multi-view stereo reconstruction algorithms. In *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, volume 1, pages 519–528, 2006. doi : 10.1109/CVPR.2006.19.

[46] J. Shu, W. Li, C. Zhang, Y. Gao, Y. Xiang, and L. Ma. Point cloud-based dimensional quality assessment of precast concrete components using deep learning. *Journal of Building Engineering*, 70 :106391, 2023. ISSN 2352-7102. doi : https://doi.org/10.1016/j.jobe.2023.106391. URL `https://www.sciencedirect.com/science/article/pii/S2352710223005703`.

[47] M. A. Talbi, H. Amroun, and M. Ammi. Cost-effective rebar inspection using point cloud data. In *2024 16th International Conference on Human System Interaction (HSI)*, pages 1–6, 2024. doi : 10.1109/HSI61632.2024.10613561.

[48] W. Tan, N. Qin, L. Ma, Y. Li, J. Du, G. Cai, K. Yang, and J. Li. Toronto-3D : A large-scale mobile lidar dataset for semantic segmentation of urban roadways. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pages 202–203, 2020.

[49] M. Tancik, E. Weber, E. Ng, R. Li, B. Yi, T. Wang, A. Kristoffersen, J. Austin, K. Salahi, A. Ahuja, D. McAllister, J. Kerr, and A. Kanazawa. Nerfstudio : A modular framework for neural radiance field development. In *ACM SIGGRAPH 2023 Conference Proceedings*, 2023. doi : 10.1145/3588432.3591516.

[50] H. Tang, Z. Liu, S. Zhao, Y. Lin, J. Lin, H. Wang, and S. Han. Searching efficient 3d architectures with sparse point-voxel convolution. In *European Conference on Computer Vision (ECCV)*, 2020.

[51] L. Teppati Losè, A. Spreafico, F. Chiabrando, and F. Giulio Tonolo. Apple lidar sensor for 3d surveying : Tests and results in the cultural heritage domain. *Remote Sensing*, 14(17), 2022. ISSN 2072-4292. doi : 10.3390/rs14174157. URL `https://www.mdpi.com/2072-4292/14/17/4157`.

[52] H. Thomas, C. R. Qi, J.-E. Deschaud, B. Marcotegui, F. Goulette, and L. J. Guibas. Kpconv : Flexible and deformable convolution for point clouds. *Proceedings of the IEEE International Conference on Computer Vision*, 2019.

[53] L. Tierney and M. Safiuddin. Insights into concrete forming, reinforcing, and pouring in building construction. *Buildings*, 12(9), 2022. ISSN 2075-5309. doi : 10.3390/buildings12091303. URL `https://www.mdpi.com/2075-5309/12/9/1303`.

[54] L. Tierney and M. Safiuddin. Insights into concrete forming, reinforcing, and pouring in building construction. *Buildings*, 12(9) :1303, 2022. doi : 10.3390/buildings12091303.

[55] S. Wang, M. Kim, H. Hae, M. Cao, and J. Kim. The development of a rebar-counting model for reinforced concrete columns : Using an unmanned aerial vehicle and deep-learning approach. *Journal of Construction Engineering and Management*, 149(11) :04023111, 2023. doi : 10.1061/JCEMD4.COENG-13686.

[56] S. Wang, L. Deng, J. Guo, M. Liu, and R. Cao. Automatic quality inspection of rebar spacing using vision-based deep learning with RGB-D camera. In *Proceedings of the 41st International Symposium on Automation and Robotics in Construction*, pages 57–64, Lille, France, 2024. doi : 10.22260/ISARC2024/0009.

[57] W. Wang. Mechanical research on reinforced concrete materials. *Materials*, 16 (21), 2023. ISSN 1996-1944. doi : 10.3390/ma16216892. URL `https://www.mdpi.com/1996-1944/16/21/6892`.

[58] W. Wang. Mechanical research on reinforced concrete materials. *Materials*, 16 (21) :6892, 2023. doi : 10.3390/ma16216892.

[59] Y. Wang, Y. Sun, Z. Liu, S. E. Sarma, M. M. Bronstein, and J. M. Solomon. Dynamic graph CNN for learning on point clouds. *ACM Transactions on Graphics*, 38(5) :146 :1–146 :12, 2019.

[60] B. Wu, A. Wan, X. Yue, and K. Keutzer. Squeezeseg : Convolutional neural nets with recurrent crf for real-time road-object segmentation from 3d lidar point cloud. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1887–1893, 2018. doi : 10.1109/ICRA.2018.8462926.

[61] X. Wu and K. Xu. numpy-hilbert-curve : Hilbert Order Implementation in

NumPy. https://github.com/PrincetonLIPS/numpy-hilbert-curve, 2020. Accessed : 2025-04-17.

[62] X. Wu, Y. Lao, L. Jiang, X. Liu, and H. Zhao. Point transformer v2 : Grouped vector attention and partition-based pooling. In *NeurIPS*, 2022.

[63] X. Wu, L. Jiang, P.-S. Wang, Z. Liu, X. Liu, Y. Qiao, W. Ouyang, T. He, and H. Zhao. Point transformer v3 : Simpler, faster, stronger. In *CVPR*, 2024.

[64] J. Xi, L. Gao, J. Zheng, D. Wang, C. Tu, J. Jiang, Y. Miao, and J. Zhong. Automatic spacing inspection of rebar spacers on reinforcement skeletons using vision-based deep learning and computational geometry. *Journal of Building Engineering*, 79 :107775, 2023. ISSN 2352-7102. doi : https://doi.org/10.1016/j.jobe.2023.107775. URL https://www.sciencedirect.com/science/article/pii/S2352710223019551.

[65] C. Xu, B. Wu, Z. Wang, W. Zhan, P. Vajda, K. Keutzer, and M. Tomizuka. Squeezesegv3 : Spatially-adaptive convolution for efficient point-cloud segmentation. In *European Conference on Computer Vision (ECCV)*, 2020.

[66] K. Yan, Q. Hu, H. Wang, X. Huang, L. Li, and S. Ji. Continuous mapping convolution for large-scale point clouds semantic segmentation. *IEEE Geoscience and Remote Sensing Letters*, 19 :1–5, 2022. doi : 10.1109/LGRS.2021.3107006.

[67] X. Yuan, F. Moreu, and M. Hojati. Cost-effective inspection of rebar spacing and clearance using rgb-d sensors. *Sustainability*, 13(22), 2021. ISSN 2071-1050. doi : 10.3390/su132212509. URL https://www.mdpi.com/2071-1050/13/22/12509.

[68] X. Yuan, F. Moreu, and M. Hojati. Cost-effective inspection of rebar spacing and clearance using RGB-D sensors. *Sustainability*, 13(22) :12509, 2021. doi : 10.3390/su132212509.

[69] X. Yuan, A. Smith, R. Sarlo, C. D. Lippitt, and F. Moreu. Automatic evaluation of rebar spacing using lidar data. *Automation in Construction*, 131 :103890, 2021. ISSN 0926-5805. doi : https://doi.org/10.1016/j.autcon.2021.103890. URL https://www.sciencedirect.com/science/article/pii/S0926580521003411.

[70] X. Yuan, A. Smith, R. Sarlo, C. D. Lippitt, and F. Moreu. Automatic evaluation of rebar spacing using LiDAR data. *Automation in Construction*, 131 :103890, 2021. doi : 10.1016/j.autcon.2021.103890.

[71] X. Yuan, A. Smith, F. Moreu, R. Sarlo, C. D. Lippitt, M. Hojati, S. Alampalli, and S. Zhang. Automatic evaluation of rebar spacing and quality using lidar data : Field application for bridge structural assessment. *Automation in Construction*, 146 :104708, 2023. ISSN 0926-5805. doi : https://doi.org/10.1016/j.autcon.2022.104708. URL `https://www.sciencedirect.com/science/article/pii/S0926580522005787`.

[72] X. Yuan, A. Smith, F. Moreu, R. Sarlo, C. D. Lippitt, M. Hojati, S. Alampalli, and S. Zhang. Automatic evaluation of rebar spacing and quality using LiDAR data : Field application for bridge structural assessment. *Automation in Construction*, 146 :104708, 2023. doi : 10.1016/j.autcon.2022.104708.

[73] Y. Zhang, Z. Zhou, P. David, X. Yue, Z. Xi, B. Gong, and H. Foroosh. Polarnet : An improved grid representation for online lidar point clouds semantic segmentation. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.

[74] H. Zhao, L. Jiang, J. Jia, P. Torr, and V. Koltun. Point transformer. In *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 16239–16248, 2021.

[75] L. Zhou, G. Wu, Y. Zuo, X. Chen, and H. Hu. A comprehensive review of vision-based 3D reconstruction methods. *Sensors*, 24(7) :2314, 2024. doi : 10.3390/s24072314.

[76] Y. Zhou and O. Tuzel. Voxelnet : End-to-end learning for point cloud based 3d object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4490–4499, 2018.

[77] X. Zhu, H. Zhou, T. Wang, F. Hong, Y. Ma, W. Li, H. Li, and D. Lin. Cylindrical and asymmetrical 3d convolution networks for lidar segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 9939–9948, 2021.