



## AVIS DE SOUTENANCE DE THESE

Le Doyen de la Faculté des Sciences Dhar El Mahraz –Fès – annonce que

Mme **BENALLAL Hafsa**  
Soutiendra : **le Samedi 06/06/2026 à 10H00**  
Lieu : **FSDM – Centre Visioconférence**

Une thèse intitulée :

**Semantic Segmentation of 3D Point Clouds Using Deep Learning**

En vue d'obtenir le **Doctorat**

**FD : Sciences et Technologies de l'Information et de la Communication**  
**Spécialité : Informatique et intelligence artificielle**

Devant le jury composé comme suit :

Nom et prénom	Etablissement	Grade	Qualité
Pr. Ali YAHYAOUY	Faculté des Sciences Dhar EL Mahraz, Fès	PES	Président
Pr. Benaïssa BELLACH	Ecole Nationale des Sciences Appliquées, Oujda	PES	Rapporteur
Pr. Abderazzak TAIME	École Supérieure de Technologie, Khénifra	MCH	Rapporteur
Pr. Mohamed Adnane MAHRAZ	Faculté des Sciences Dhar EL Mahraz, Fès	PES	Rapporteur
Pr. Khalid EL FAZAZY	Faculté des Sciences Dhar EL Mahraz, Fès	PES	Examineur
Pr. Noura AHERRAHOUCHE	Faculté des Sciences Dhar EL Mahraz, Fès	MCH	Examineur
Pr. Jamal RIFFI	Faculté des Sciences Dhar EL Mahraz, Fès	MCH	Co-encadrant
Pr. Nadine Abdallah SAAB	École d'ingénieurs, ISEN Ouest, Brest, France	MC	Expert
Pr. Ayman AL FALOU	École d'ingénieurs, ISEN Ouest, Brest, France	PES	Expert
Pr. Hamid TAIRI	Faculté des Sciences Dhar EL Mahraz, Fès	PES	Directeur de thèse



## Résumé :

Cette thèse porte sur la segmentation sémantique de nuages de points 3D acquis en environnement extérieur par des capteurs LiDAR, avec pour objectif l'attribution d'une étiquette sémantique à chaque point de la scène. Ce problème, central pour la conduite autonome, la cartographie urbaine et la perception robotique, se heurte à la parcimonie et à l'irrégularité des données tridimensionnelles, au déséquilibre sévère entre classes sémantiques, et à l'hétérogénéité des modalités capteurs. Les travaux s'articulent autour de trois contributions architecturales complémentaires, chacune répondant aux limitations identifiées dans la précédente. La première contribution introduit la distillation de connaissances pondérée par capsules (CWKD), un cadre exploitant les magnitudes d'activation des réseaux de capsules comme signaux de confiance spatiale pour moduler le transfert de connaissances d'un encodeur 2D vers un encodeur 3D. Le modèle atteint 66,8 % de mIoU sur SemanticKITTI avec 21,4 M de paramètres, soit le meilleur rapport performance/paramètres parmi les méthodes comparées. La deuxième contribution propose CapsX3D : Capsule Networks with Cross-Modal Attention for MultiSensor 3D Semantic Segmentation, une architecture intégrant un module d'attention croisée bidirectionnelle (CMAF) permettant un échange réciproque d'informations entre les caractéristiques capsules 2D et les caractéristiques 3D éparses. Ce mécanisme permet d'atteindre un mIoU de 67,2%, avec des performances de pointe sur trois classes (camion : 79,8%, cycliste : 87,3%, végétation : 87,9%). La troisième contribution propose CyRA-Net, une architecture légère pour la segmentation sémantique 3D multimodale, combinant un encodeur compact, des modules de contexte multi-échelle optimisés et un mécanisme d'attention cross-modale efficace pour la fusion robuste des informations. CyRA-Net atteint 63,19 % de mIoU en validation sur SemanticKITTI avec seulement 6,3 M de paramètres, offrant le ratio d'efficacité paramétrique le plus élevé parmi les méthodes comparées (9,96 points/million). L'ensemble des contributions est évalué sur le benchmark SemanticKITTI et partage un fil conducteur méthodologique : l'exploitation de représentations en coordonnées cylindriques et de mécanismes de pondération adaptatifs de sophistication croissante.

## Mots clés :

3D point cloud segmentation, semantic segmentation, LiDAR-camera fusion, knowledge distillation, capsule networks, multi-representation fusion, cylindrical voxels, gated crossrepresentation fusion, SemanticKITTI, autonomous driving.



## Semantic Segmentation of 3D Point Clouds Using Deep Learning

### Abstract :

This thesis focuses on the semantic segmentation of 3D point clouds acquired in outdoor environments by LiDAR sensors, with the objective of assigning a semantic label to each point in the scene. This problem, which is central to autonomous driving, urban mapping, and robotic perception, faces challenges such as the sparsity and irregularity of 3D data, severe class imbalance, and the heterogeneity of sensor modalities. The work is structured around three complementary architectural contributions, each addressing the limitations identified in the previous one. The first contribution introduces Capsule-Weighted Knowledge Distillation (CWKD), a framework that leverages capsule activation magnitudes as spatial confidence signals to modulate the knowledge transfer from a 2D encoder to a 3D encoder. The model achieves 66.8% mIoU on SemanticKITTI with 21.4M parameters, representing the best performance-to-parameter ratio among the compared methods. The second contribution proposes CapsX3D : Capsule Networks with Cross-Modal Attention for MultiSensor 3D Semantic Segmentation, an architecture integrating a bidirectional cross-modal attention module (CMAF) that enables reciprocal information exchange between 2D capsule features and sparse 3D features. This mechanism raises the mIoU to 67.2%, with state-of-the-art performance on three classes (truck : 79.8%, cyclist : 87.3%, vegetation : 87.9%). The third contribution presents CyRA-Net, a lightweight architecture that combines a compact encoder, optimized multi-scale context modules, and an efficient cross-modal attention mechanism for robust fusion. CyRA-Net achieves 63.19% mIoU on the SemanticKITTI validation set with only 6.3M parameters, delivering the highest parameter-efficiency ratio (9.96 mIoU-points per million parameters) among all compared methods, and making it suitable for embedded 3D semantic segmentation. All contributions are evaluated on the SemanticKITTI benchmark and share a methodological thread : the use of cylindrical coordinate representations and increasingly sophisticated adaptive weighting mechanisms.

### Key Words :

3D point cloud segmentation, semantic segmentation, LiDAR-camera fusion, knowledge distillation, capsule networks, multi-representation fusion, cylindrical voxels, gated crossrepresentation fusion, SemanticKITTI, autonomous driving.