



جامعة محمد الخامس بالرباط
Université Mohammed V de Rabat

École Nationale Supérieure d'Informatique et d'Analyse des Systèmes
Centre d'Études Doctorales en Sciences des Technologies de l'Information et de l'Ingénieur

AVIS DE SOUTENANCE DE THÈSE DE DOCTORAT

Madame Zainab Ouadirhi

Soutiendra publiquement sa thèse de Doctorat en Informatique

Spécialité : Génie Informatique

Le Lundi 19 Mai 2025 à 15h00 au Grand Amphi de l'ENSIAS de Rabat

Intitulé de la thèse

Occlusion-Aware Object Detection for Enhanced 2D/3D Vision Systems

Président :

Pr. El Mostafa DAOUDI, PES, Faculté des Sciences, Université Mohamed Premier, Oujda

Directeur de thèse :

Pr. Mostapha ZBAKH, PES, ENSIAS, Université Mohammed V de Rabat

Pr. Sidi Ahmed MAHMOUDI, Associate Professor, Faculté Polytechnique de Mons, Université de Mons, Belgique

Co-Directeur de thèse :

Pr. Mohammed BENJELLOUN, Professeur des Universités, Faculté Polytechnique de Mons, Université de Mons, Belgique

Rapporteurs :

Pr. Mahmoud NASSAR, PES, ENSIAS, Université Mohammed V de Rabat

Pr. Mostafa EL HABIB DAHO, Associate Professor, Université de Bretagne Occidentale (UBO), Brest, France

Pr. Matei MANCAS, Associate Professor, Faculté Polytechnique de Mons, Université de Mons, Belgique

Examineur :

Pr. Stéphane DUPONT, Associate Professor, Faculté des Sciences, Université de Mons, Belgique

Résumé: L'occlusion reste un défi majeur dans la détection d'objets, en particulier dans les environnements complexes où les objets sont partiellement ou totalement cachés. Cette thèse aborde ce problème à travers plusieurs contributions visant à améliorer la compréhension et l'atténuation des occlusions. Une nouvelle méthode d'évaluation du taux d'occlusion (OR) est introduite, combinant l'extraction de grille de voxels sensible à la densité et l'analyse de densité des voisins basée sur les diagrammes de Voronoï pour quantifier la sévérité de l'occlusion et guider la sélection du modèle.

Ces techniques de prétraitement optimisent l'utilisation des données de nuages de points 3D, assurant une meilleure performance dans les environnements fortement occlus.

La thèse propose FuDensityNet 2.0, un cadre de détection d'objets multimodal intégrant les données 2D et 3D. L'architecture fusionne ces deux modalités grâce à une stratégie de fusion efficace, garantissant une représentation robuste pour une détection précise sous occlusion. Une stratégie adaptative ajuste dynamiquement les performances du modèle en fonction du niveau d'occlusion, améliorant ainsi sa robustesse.

Des expériences menées sur KITTI, NuScenes et OccludedPascal3D valident ces contributions. FuDensityNet 2.0 atteint une précision moyenne (AP) de 76,6 % pour la détection de voitures sous « occlusion forte », surpassant les modèles de l'état de l'art de plus de 11 %. Les études comparatives et d'ablation démontrent l'efficacité de la voxelisation, de l'évaluation de OR et des techniques de fusion.

Enfin, cette recherche explore des approches basées sur le 2D pour générer des nuages de points via l'estimation de profondeur, réduisant ainsi la dépendance aux capteurs LiDAR. Ces contributions offrent des perspectives prometteuses pour des applications en véhicules autonomes, surveillance intelligente et robotique industrielle.

Mots-clés : Apprentissage profond, Détection d'objets 2D/3D, Intelligence artificielle, Occlusion, Vidéosurveillance.

Abstract: Occlusion remains a fundamental challenge in object detection, particularly in complex real-world environments where objects are partially or fully obscured. This thesis addresses the problem through multiple contributions that enhance the understanding and mitigation of occlusions. A novel Occlusion Rate (OR) evaluation method is introduced, combining density-aware voxel grid extraction and Voronoi-based neighbor density analysis to quantify occlusion severity and guide model selection. These preprocessing techniques optimize 3D point cloud data, ensuring improved performance in highly occluded settings.

Building on these foundations, the thesis proposes FuDensityNet 2.0, a multimodal object detection framework that integrates 2D image data and 3D voxelized point cloud data. The architecture integrates 2D and 3D features using a robust multimodal fusion strategy, which ensures efficient and complementary feature representation for accurate detection in occluded environments. An occlusion-aware detection strategy dynamically adapts to varying occlusion levels, enhancing robustness across diverse conditions.



جامعة محمد الخامس بالرباط
Université Mohammed V de Rabat

Extensive experiments conducted on benchmark datasets such as KITTI, NuScenes, and OccludedPascal3D validate the proposed contributions. FuDensityNet 2.0 achieves an average precision (AP) of 76.6 % for car detection under “hard” occlusion scenarios, surpassing state-of-the-art models by over 11 %. Comparative analyses and ablation studies further demonstrate the effectiveness of the voxelization strategy, occlusion rate assessment, and fusion techniques in addressing occlusions.

The contributions extend beyond model design. This work explores 2D-driven approaches for generating point clouds using depth estimation techniques, reducing reliance on LiDAR sensors and offering a foundation for cost-effective solutions. Additionally, the methodologies proposed in this research provide valuable insights for applications such as autonomous vehicles, smart surveillance, and industrial robotics, where robust object detection under occlusion is critical.

Keywords: 2D/3D object detection, artificial intelligence, deep learning, occlusion, video surveillance.