



Avis de Soutenance

Monsieur Fahad LATEEF

Présentera ses travaux en soutenance

Soutenance prévue le **mercredi 15 décembre 2021** à 14h00

Lieu : UTBM, Montbéliard Campus, 1 Cr Louis Leprince-Ringuet, 25200 MONTBELIARD

Salle : M101

Titre des travaux : Analyse sémantique de l'environnement de conduite dans les scénarios urbains

Ecole doctorale : SPIM - Sciences Physiques pour l'Ingénieur et Microtechniques

Section CNU : 61

Unité de recherche : CIAD - Connaissance et Intelligence Artificielle Distribuées

Directeur de thèse : Yassine RUICHEK

Codirecteur de thèse : HDR NON HDR

Soutenance : Publique A huis clos

Membres du jury :

<u>Nom</u>	<u>Qualité</u>	<u>Etablissement</u>	<u>Rôle</u>
M. Yassine RUICHEK	Professeur des universités	Université Bourgogne - Franche-Comté	Directeur de thèse
Mme Samia AINOUS	Professeure des universités	INSA Rouen	Rapporteuse
M. Hedi TABIA	Professeur des universités	Université d'Evry	Rapporteur
M. Franck DAVOINE	Chargé de recherche	Institute for Information Sciences and Technologies (INS2I)	Examineur
M. You LI	Ingénieur de recherche	Renault	Examineur
M. Jean-Christophe BURIE	Professeur des universités	La Rochelle Université	Examineur

Mots-clés : Apprentissage Profond, Conduite Autonome, Segmentation Sémantique, Attention Visuelle, Détection d'objets en Mouvement, Compensation de Mouvement

Résumé de la thèse (en français) :

La tâche de compréhension des scènes urbaines nécessite la reconnaissance des constituants sémantiques de la scène et les interactions complexes entre eux. Par le biais de cette thèse, nous explorons et fournissons des représentations efficaces pour comprendre les scènes urbaines basées sur la perception, qui peuvent être utiles pour la planification et la prise de décision dans divers environnements urbains complexes et conditions environnementales variées. Nous présentons d'abord une taxonomie des méthodes d'apprentissage profond dans le domaine de la segmentation sémantique, en vue de l'intérêt que porte la communauté scientifique à ce sujet pour la compréhension des scènes de conduite urbaine. Ainsi, nous avons d'abord classifié ces méthodes en fonction de leur structure architecturale afin d'élaborer ensuite une discussion sur leurs avantages, limites possibles et orientations futures. En suite, nous avons proposé une nouvelle approche de l'attention visuelle pour la conduite basée sur un réseau génératif conditionnel (GAN). Présentation des algorithmes de saillance bien connus, à la fois des approches classiques et d'apprentissage profond utilisées pour l'attention visuelle. Dans ce contexte, nous avons mis en place une large base de données d'attention visuelle basée sur une nouvelle stratégie d'extraction de cartes de saillance à partir d'un ensemble de données de conduite existant. Nous avons ensuite proposé un nouveau cadre d'identification d'objets qui combine des indices de mouvement et de géométrie pour comprendre l'environnement de conduite urbain. Par ailleurs, un nouveau modèle de détection d'objets en mouvement a été développé en intégrant un réseau codeur-décodeur couplé avec la segmentation sémantique et un réseau d'estimation de disparité. Un algorithme d'enregistrement d'image est proposé avec le flux optique pour compenser l'ego-mouvement. De nombreuses évaluations approfondies sur divers ensembles de données de conduite montrent que toutes les méthodes proposées atteignent des performances remarquables en termes de précision et démontrent l'efficacité des techniques essentielles pour la compréhension de la scène en conduite autonome.

Abstract (in English):

Understanding urban scenes require recognizing the semantic constituents of a scene and the complex interactions between them. In this work, we explore and provide effective representations for understanding urban scenes based on in situ perception, which can be helpful for planning and decision-making in various complex urban environments and under a variety of environmental conditions. We first present a taxonomy of deep learning methods in the area of semantic segmentation, the most studied topic in the literature for understanding urban driving scenes. The methods are categorized based on their architectural structure and further elaborated with a discussion of their advantages, possible limitations, and future directions. Then, we proposed a new approach to visual attention for driving based on a conditional generative adversarial network. We have presented the well-known saliency algorithms, both classical and Deep Learning approaches, used for visual attention. We built a large visual attention database on a new strategy for mining saliency heatmaps from existing driving datasets. We then proposed a novel object identification framework that combines motion and geometry cues to understand the urban driving environment. A new moving object detection model is developed by integrating an encoder-decoder network with semantic segmentation and a disparity estimator. An image registration algorithm is proposed along with the optical flow to compensate for ego-motion. Extensive empirical evaluations on various driving datasets show that all the proposed methods achieve remarkable performance in terms of accuracy and demonstrate the effectiveness of the essential techniques for scene understanding in autonomous driving.