

Soutenance de thèse

Zhujun XU soutiendra sa thèse de doctorat préparée au sein de l'ISAE-ONERA SCANR et intitulée «*Détection et suivi d'objets 2D et 3D fondés sur l'apprentissage profond sur vidéos monoculaires dans le contexte des véhicules autonomes*»

Le 01 avril 2022 à 10h00, Amphi 4 - ISAE-SUPAERO

devant le jury composé de

M. Eric CHAUMETTE	Professeur ISAE-SUPAERO	Directeur de thèse
M. Damien VIVET	Maître de conférences ISAE-SUPAERO	Co-directeur de thèse
Mme Samia AINOUIZ	Professeure INSA Rouen Normandie	Rapporteuse
M. Thierry CHATEAU	Professeur Université Clermont Auvergne	Rapporteur

Résumé : L'objectif de ce travail de thèse est de développer des méthodes de détection et de suivi d'objets 2D et 3D fondés sur l'apprentissage profond sur vidéo monoculaire et de les appliquer au contexte du véhicule autonome. En effet, lorsque l'on utilise directement des détecteurs d'images fixes pour traiter un flux vidéo, la précision souffre d'un problème de qualité d'image du fait de l'échantillonnage. De plus, les annotations 3D des vidéos prennent du temps et sont coûteuses en raison du grand nombre d'images. Nous profitons donc des informations temporelles dans les vidéos, telle que la cohérence des objets, pour améliorer les performances. Les méthodes ne doivent pas introduire trop de charge de calcul supplémentaire, car le véhicule autonome exige une performance en temps réel. Plusieurs méthodes peuvent être utilisées dans différentes étapes, par exemple, la préparation des données, l'architecture du réseau et le post-traitement. Tout d'abord, nous proposons une méthode de post-traitement appelée heatmap propagation (propagation de carte de chaleur) fondée sur un one-stage détecteur CenterNet pour la détection d'objets dans les vidéos. Notre méthode propage la détection fiable effectuée sur les images précédentes sous la forme d'une heatmap pour la prochaine image. Ensuite, pour distinguer différents objets d'une même classe, nous proposons une architecture de réseau image par image pour la segmentation d'instances vidéo en utilisant les instance séquence queries (requête de séquence d'instances). Le suivi des instances est réalisé sans post-traitement supplémentaire pour l'association de données. Enfin, nous proposons une méthode d'apprentissage semi-supervisée pour générer des annotations 3D pour une base de données de suivi d'objets dans les vidéos 2D. Cela permet d'enrichir le processus d'apprentissage pour la détection d'objets 3D. Chacune des trois méthodes peut être appliquée individuellement pour étendre les détecteurs d'images dans le cadre d'applications sur vidéo. Nous proposons également deux structures de réseau complètes pour résoudre la détection et le suivi d'objets 2D et 3D sur vidéo monoculaire.

Mots-clés : Vision par ordinateur, Apprentissage profond, Analyse des vidéos, Véhicules autonomes

Summary: The objective of this project is to develop deep learning based 2d and 3d object detection and tracking methods on monocular video and apply them to the context of autonomous vehicles. Actually, when directly using still image detectors to process a video stream, the accuracy suffers from sampled image quality problem. Moreover, 3d video annotations are time-consuming and expensive due to large numbers of frames. We therefore want to take advantage of the temporal information in videos such as the object consistency, to improve the performance. Types of methods can be involved in different steps; for example, data preparation, network architecture and post-processing. The methods should not introduce too much extra computational burden, since the autonomous vehicle demands a real-time performance.

Keywords: Computer vision, Deep learning, Video analysis, Autonomous vehicles