

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

Monsieur Youssef FENJIRO

Soutiendra publiquement sa thèse de Doctorat en Informatique

Le Mercredi 27 Juillet 2022 à 9H30 au Grand amphi à l'ENSIAS

Intitulé de la thèse

**END-TO-END DEEP REINFORCEMENT LEARNING AND IMITATION LEARNING
MODELS FOR SELF-DRIVING CARS**

Devant le Jury composé de :

Président :

Pr. Karim BAINA, PES, ENSIAS, Université Mohammed V de Rabat

Directeur de thèse :

Pr. Houda BENBRAHIM, PES, ENSIAS, Université Mohammed V de Rabat

Rapporteurs :

Pr. Mounir GHOGHO, PES, Université Internationale de Rabat

Pr. Mostafa BELLAFKIH, PES, Institut National des Postes et Télécommunications, Rabat

Pr. Sanaa EL FKIHI, PES, ENSIAS, Université Mohammed V de Rabat

Examineur :

Pr. Younes TABII, PH, ENSIAS, Université Mohammed V de Rabat





جامعة محمد الخامس بالرباط
Université Mohammed V de Rabat

Résumé : L'humanité est à l'aube d'une transformation dans le monde des transports et de la mobilité. Les progrès technologiques des véhicules autonomes sont en train de révolutionner le secteur automobile. Il s'agit d'une rupture majeure et de l'avancée la plus bouleversante depuis l'invention de l'automobile, puisque le logiciel devient une partie importante de la voiture qui sera désormais, pilotée par l'intelligence artificielle. L'avènement de la voiture autonome changera fondamentalement la façon dont nous nous déplaçons et transportons les marchandises, ce qui induira des changements importants pour nos sociétés. De gros budgets sont investis dans ce domaine par les constructeurs de voitures, et aussi les sociétés high-tech qui se sont joint à cette course pour développer des modèles et prototypes capables de conduire de manière autonome, sans intervention humaine. Néanmoins, selon la SAE qui a établi cinq niveaux d'autonomie pour les systèmes de conduite autonomes (ADS), le degré de maturité n'est pas encore atteint. En effet, Tesla qui est précurseur dans le domaine est toujours au niveau deux. Il reste donc encore, plusieurs challenges à relever et un long chemin à parcourir entre la preuve de concept et la véritable extension de cette technologie pour le grand public, surtout que la sûreté et sécurité des humains sont la priorité numéro un, et la précision et la robustesse des algorithmes de conduite doivent être infaillibles et fiables à 100 %. Dans cette thèse, nous proposons d'ajouter des briques supplémentaires au socle global de la conduite autonome, à travers de nouveaux modèles d'apprentissage en profondeur de bout en bout, qui génèrent les commandes de conduite, orientation et vitesse, depuis l'image en entrée. On a utilisé pour cela les deux modes, supervisé et par renforcement, ainsi que des environnements de simulation, CARLA et OpenAI Gym, pour l'entraînement et les tests, vu les risques de sécurité et d'accidents dans le monde réel. Notre première contribution porte sur un nouvel algorithme DRLG basé sur le Deep Reinforcement Learning avec l'intégration en supplément à l'image de l'environnement, d'une entrée externe issue d'un système de planification d'itinéraire ou d'une intervention humaine, et ce pour indiquer la direction à suivre par la voiture autonome ou prendre une décision dans un carrefour ou rond-point. Notre modèle a réussi à conduire dans les routes tout en répondant aux commandes de guidage. La seconde contribution est un nouvel algorithme adoptant le modèle actor-critic, agissant dans un espace d'actions continu et combinant les deux modes d'intelligence artificielle. En première phase, on utilise l'apprentissage supervisé par imitation, qui permet au modèle d'acquérir rapidement des compétences de conduite, à l'aide de démonstrations d'experts et en deuxième phase, l'apprentissage par renforcement, qui prend le relais via une stratégie d'exploration, afin d'assurer la continuité et la progression de l'apprentissage. En comparaison avec les autres modèles de Deep Reinforcement et de Supervised Imitation Learning, les résultats de notre modèle étaient meilleurs. La troisième contribution est un nouvel algorithme OCRIL, qui est aussi une combinaison des deux modes d'apprentissage par imitation et par renforcement, avec l'optimisation de sa progression de façon à empêcher toute régression en terme de performance, via une contrainte appliquée lors de la phase d'exploration, sous forme d'un signal critique, qui conditionne la mise à jour du modèle durant son entraînement. Ce dernier algorithme a été ré-entraîné pour gérer également l'évitement d'obstacles statiques. Il a dépassé en terme de performance les autres algorithmes Deep Reinforcement et de Supervised Imitation Learning lors d'une étude comparative. Pour ces contributions, les résultats ont été très satisfaisants et ont permis d'apporter une participation à l'effort de recherche dans ce domaine.

Mots-clés : Algorithmes de vision et de contrôle par ordinateur, Apprentissage par imitation, Apprentissage en profondeur, Apprentissage en profondeur par renforcement, Voiture autonome.





جامعة محمد الخامس بالرباط
Université Mohammed V de Rabat

Abstract: Humanity is at the dawn of a transformation in the world of transport and mobility. Technological advancements in autonomous vehicles are revolutionizing the automotive industry. This is a major breakthrough and the most disruptive advancement since the invention of the automobile, as software becomes an important part of the car that will now be driven by artificial intelligence agents. The advent of the autonomous car will fundamentally change the way we move and transport goods, which will bring about significant changes in our societies. Large budgets are invested in this area by car manufacturers and high-tech companies that have joined this race to develop models and prototypes capable of driving autonomously, without human intervention. However, according to the SAE, which has established five levels of autonomy for autonomous driving systems (ADS), the degree of maturity has not yet been reached. Indeed, Tesla, which is a pioneer in the field, is still at level two. There are therefore still several challenges to be met and a long way to go between the proof of concept and the real extension of this technology for the general public, especially since the safety and security of humans are at stake, and the precision and robustness of the driving algorithms must be 100% reliable. In this thesis, we propose to bring additional bricks to the global foundation of autonomous driving, through new end-to-end deep learning models, which generate driving commands, orientation and speed, from the input image. Both modes, supervised and reinforcement, were used for this, as well as simulation environments, CARLA and OpenAI Gym, for training and testing, given the risks of safety and accidents in the real world. Our first contribution relates to a new DRLG algorithm based on Deep Reinforcement Learning with the integration of external input from a route planning system or human intervention, in addition to the image of the environment injected at the input of the model. These additional commands indicate to the autonomous car, the direction to follow when crossroads or roundabouts. Our model managed to drive on roads while responding to guidance commands. The second contribution is a new algorithm adopting the actor-critic model, acting in a continuous space of actions and combining the two modes of artificial intelligence, Deep reinforcement and Imitation Learning. In the first phase, supervised learning by imitation is used, which allows the model to quickly acquire driving skills, using expert demonstrations and in the second phase, reinforcement learning takes over via an exploration strategy, to ensure the continuity and progression of learning. Compared to the other Deep Reinforcement and Supervised Imitation Learning models, the results of our model were better in the simulation environment. The third contribution is a new OCRIL algorithm, which is also a combination of the two modes of Deep Reinforcement and Imitation Learning, with the optimization of its progression so as to prevent any regression in terms of performance, by applying a constraint during the exploration phase, in the form of a critical signal, which conditions the model's network weights updating during its training. In addition, this last algorithm has been re-trained to handle the avoidance of static obstacles. It outperformed the other Deep Reinforcement and Supervised Imitation Learning algorithms in a comparative study. For these contributions, the results were very satisfactory and made it possible to participate in the research effort in this field.

Keywords: Computer vision & control algorithms, Deep Learning, Deep Reinforcement Learning, Imitation Learning, Self-driving car

