

Techniques Avancées d'IA pour l'Extraction de Scénarios de Conduite en Milieu Urbain : IA Générative, meta-learning, Détection d'Objets et Augmentation de Données par Simulation

ENTREPRISES

L'**Institut pour la Transition Energétique (ITE) VEDECOM** est un institut français de recherche et de formation dédié aux mobilités innovantes et durables, à travers trois axes de R&D multidisciplinaire : l'électrification, le véhicule automatisé et connecté, les nouvelles solutions de mobilité et d'énergies partagées. Intégré dans le programme d'innovation "France 2030", l'Institut contribue également activement à de nombreux projets de recherche européens. Constitué de chercheurs, d'ingénieurs, de responsables d'études, de doctorants, ..., VEDECOM compte 80 collaborateurs qui ont à cœur de répondre aux problématiques de mobilité de demain.

L'**ESTACA** est une Ecole d'Ingénieurs spécialisée dans les transports (aéronautique, automobile, espace, transports guidés et naval). Ses équipes de recherche sont regroupées au sein d'ESTACA'LAB en deux Pôles : S2ET (« Systèmes et Energies Embarqués dans les Transports ») et MSCE (« Mécanique des Structures Composites et Environnement ») et sont réparties sur les 3 campus (Saint Quentin en Yvelines, Laval et Bordeaux). Le doctorant recruté sera intégré au pôle S2ET (« Systèmes et Energie Embarqués dans les Transports ») basé à Saint Quentin en Yvelines et à VEDECOM – Versailles.

CONTEXTE ET OBJECTIF

CONTEXTE DE LA THESE :

Cette thèse s'inscrit dans le cadre d'une collaboration entre l'Institut VEDECOM et l'ESTACA, dans le contexte du projet MOOVE Urban (**M**onitoring **O**utillé du **V**éhicule autonome dans son **E**nvironnement Urbain), soutenu par l'ANR ainsi que les industriels Renault et Stellantis.

La certification des systèmes avancés d'aide à la conduite (ADAS) repose principalement sur des essais basés sur des scénarios de conduite. Identifier de manière exhaustive l'ensemble des situations routières possibles et en apprendre la distribution statistique constitue un enjeu fondamental, mais extrêmement complexe [1]. Traditionnellement, l'extraction de scénarios à partir des données de roulage repose principalement sur des approches déterministes basées sur des règles expertes, qui présentent une forte transparence mais souffrent d'une rigidité limitant leur capacité à capturer la diversité, la complexité et l'imprévisibilité notamment l'environnements urbains. En particulier, ces méthodes montrent leurs limites face à la détection des scénarios rares, inconnus ou extrêmes, essentiels à la sécurité des véhicules autonomes [2].

Face à ces défis, l'intégration des techniques d'intelligence artificielle (IA) apparaît incontournable. Le projet antérieur MOOVE 4.0, sur la base d'une base de données de roulage de plus d'un million de kilomètres, a permis de développer des outils IA efficaces pour l'analyse de scénarios sur voies à chaussées séparées (VCS) [3,4,5]. Cependant, l'environnement urbain présente des défis considérablement accrus : grande variabilité du trafic, interactions multiples et complexes entre des usagers hétérogènes (véhicules, piétons, cyclistes, etc.), diversité accrue des infrastructures urbaines et faible quantité de données annotées disponibles pour l'apprentissage, particulièrement concernant les scénarios rares ou dangereux [6]

Dans ce contexte, le projet MOOVE Urban ambitionne non seulement d'étendre les outils existants, mais aussi d'explorer de nouvelles approches IA adaptées à ces spécificités urbaines. Afin de

compenser le manque de données annotées, la thèse explorera notamment des méthodes avancées d'augmentation des données à travers la simulation réaliste de scénarios critiques et rares. Par ailleurs, de nouvelles représentations des données, telles que la tokenisation des séries temporelles multivariées (combinant signaux continus et discrets), seront investiguées pour optimiser la gestion de la complexité structurelle et temporelle des scénarios urbains. Enfin, l'utilisation d'approches de meta-learning sera envisagée afin d'adapter rapidement les modèles aux scénarios rares ou inconnus, en les rendant également capables d'estimer et de gérer explicitement l'incertitude de leurs prédictions. Ces modèles seront ainsi conçus pour détecter lorsqu'un scénario est out-of-distribution (OOD), permettant au modèle d'être conscient de ses propres limites (savoir ce qu'il ne sait pas) [7]

PROBLEMATIQUES/VERROUS :

L'exploitation de données de conduite urbaines pour la détection automatique de scénarios se heurte d'abord à la variabilité intrinsèque du trafic : chaque occurrence peut impliquer un nombre très différent d'obstacles hétérogènes, produire des séries temporelles de longueurs inégales et exhiber une configuration spatiale propre au cas considéré. Cette hétérogénéité – nombre d'obstacles variable, durées inter- et intra-scénarios inconstantes, géométrie spécifique des interactions – empêche l'adoption d'un format d'entrée simple et complique l'entraînement de modèles d'apprentissage profond.

À cette complexité structurelle s'ajoutent les spécificités de l'environnement urbain. Les infrastructures (carrefours à feux, ronds-points, ...) multiplient considérablement les sources de variation contextuelle, tandis que la présence d'usagers vulnérables (piétons, cyclistes...) introduit des trajectoires moins prévisibles et des interactions plus fines à capturer. Transférer la représentation tensorielle développée pour les voies séparées (VCS) [3] vers ce milieu urbain riche sans perdre l'information de contexte ni exploser la dimensionnalité constitue donc un verrou majeur. À ce titre, l'exploration de nouvelles représentations (notamment par tokenisation) pour gérer efficacement cette complexité reste à investiguer [8].

En outre, la quantité limitée de données annotées disponibles constitue une contrainte significative, particulièrement critique pour les scénarios rares ou dangereux. Cette rareté accentue fortement le risque lié aux scénarios inconnus au moment de l'inférence : le modèle, n'ayant jamais rencontré certains scénarios, peut générer des faux positifs ou, pire, ignorer des situations critiques (faux négatifs). Ce problème est exacerbé par la co-occurrence fréquente de plusieurs manœuvres complexes simultanées – comme un freinage brutal du véhicule en amont (CIPV : *Closest-In-Path Vehicle*) couplé à un dépassement. Dès lors, l'utilisation de techniques d'apprentissage frugal (meta-learning, few-shot learning) [9] pour traiter efficacement ces cas rares ou inédits représente un véritable défi, en raison de la complexité et de la nature hétérogène des données issues des données de roulage.

Le pipeline actuel de détection repose sur une stratégie en deux étapes : balayage de fenêtres temporelles de tailles multiples suivi d'une suppression des redondances par la méthode du Non-Maximum Suppression (NMS). Cette méthode, très consommatrice de ressources, génère de nombreux doublons et provoque une latence élevée incompatible avec les contraintes industrielles. Réduire cette procédure lourde en un pipeline « one-stage » [10] et multilabel rapide, précis représente un verrou majeur pour automatiser l'extraction de scénarios.

Enfin, la faiblesse intrinsèque des modèles actuels à quantifier précisément leur incertitude représente un obstacle supplémentaire à la validation industrielle des solutions proposées. L'intégration d'approches capables de fournir des estimations fiables de confiance ou d'incertitude des prédictions [11,12] constitue ainsi une exigence fondamentale à satisfaire.

OBJECTIFS DE LA THESE :

- **Unification des données** : Concevoir une représentation tensorielle étendue capable d'englober simultanément les signaux continus et discrets, tout en préservant la configuration spatiale propre aux scénarios urbains. Cette structure doit notamment intégrer la complexité de l'infrastructure urbaine et la présence d'utilisateurs vulnérables afin d'offrir un format d'entrée stable et complet et exploitable par les modèles d'intelligence artificielle.
- **Apprentissage frugal par Meta-Learning** : Développer des méthodes d'apprentissage *few-shot* issues du méta-learning pour permettre aux modèles de s'adapter efficacement à de nouveaux scénarios urbains peu représentés ou totalement inédits, malgré la rareté des annotations disponibles. L'objectif est de renforcer la généralisation des modèles dans des contextes urbains variés et imprévus.
- **Classification multilabel des manœuvres co-occurentes** : Concevoir un modèle de classification multilabel capable d'identifier simultanément plusieurs manœuvres co-occurentes dans une même séquence temporelle. Le modèle s'appuiera sur des mécanismes d'attention et un partitionnement spatio-temporel adapté pour capturer la complexité dynamique des interactions urbaines.
- **Augmentation par simulation et analyse en espace latent**: Mettre en œuvre une boucle fermée « extraction → projection latente → analyse → simulation » : les scénarios détectés sont projetés dans l'espace latent d'un modèle génératif, où l'analyse des zones sous-représentées permet d'identifier les lacunes du jeu de données. À partir de ces "trous" dans l'espace latent, les variables d'entrée pertinentes sont déduites pour générer de nouveaux scénarios via simulation réaliste.
- **Détection « one-stage » à faible latence** : Remplacer le pipeline bi-étape actuel (fenêtrage + Non-Maximum Suppression) par une architecture légère, directe et efficace inspirée des paradigmes « one-stage » tels que YOLO, Faster R-CNN, CLIP, ALIGN ou Florence, adaptée aux séries temporelles longues. Cette nouvelle architecture devra permettre une détection rapide, précise et robuste des scénarios complexes en un seul passage.
- **Tokenisation des séries temporelles & backbones Transformer** : Explorer une représentation tokenisée de séries temporelles, découpées en patches temporels fusionnant les signaux continus et discrets, afin de tirer parti des architectures de type Transformer. Un pré-entraînement auto-supervisé sur un corpus massif de scénarios VCS non annotés sera envisagé, suivi d'un transfert vers le domaine urbain via un apprentissage *few-shot*.
- **Quantification et maîtrise de l'incertitude** : Intégrer des approches probabilistes, telles que les *deep ensembles* ou l'inférence variationnelle, pour estimer la confiance des prédictions. L'objectif est de fournir une évaluation fiable de l'incertitude, réduisant le taux de faux positifs et de faux négatifs, et apportant des garanties supplémentaires dans les processus de validation des systèmes AD/ADAS.

Mots Clés :

Scénarios de conduite • Données urbaines • Intelligence artificielle • Méta-learning • Classification multilabel • Simulation de données • Série temporelle • Tokenisation • Transformers • Incertitude • ADAS • Véhicule autonome • Apprentissage frugal • Représentation latente • One-stage detection

Références bibliographiques :

[1] Li, X., Song, R., Fan, J., Liu, M., & Wang, F. (2023). Development and Testing of Advanced Driver Assistance Systems Through Scenario-Based Systems Engineering. *IEEE Transactions on Intelligent Vehicles*, 8, 3968-3973. <https://doi.org/10.1109/TIV.2023.3297168>.

Campus Paris-Saclay / 12, avenue Paul Delouvrier, 78180 Montigny Le Bretonneux

- [2] Gao, Y. and Piccinini, M. and Zhang Y. and Wang D. and Moller K. and Roberto Brusnicki and Baha Zarrouki and Alessio Gambi and Jan Frederik Totz and Kai Storms and Steven Peters and Andrea Stocco and Bassam Alrifaae and Marco Pavone and Johannes Betz (2025) Foundation Models in Autonomous Driving: A Survey on Scenario Generation and Scenario Analysis
- [3] Kheriji, W., Hadj Selem, F., Ben Nejma, G., Bontemps, T., Durville, L., & Rahal, M.-C. (2024). Extracting scenarios for automated driving systems: A statistical-based occupation-grid approach with deep semi-supervised learning. International Conference on ITS (ITSC), September 24-27, 2024, Edmonton, Canada.
- [4] Hadj Selem, F., Ben Nejma, G., Kheriji, W., Durville, L., & Arnoux, E. (2024). Which scenarios must be tested for safety in automated driving: To cut testing budget. ITS World Congress Proceedings, Dubai, UAE, pages 15-28.
- [5] Hadj Selem, F., Ben Nejma, G., Kheriji, W., Durville, L., Rahal, M.-C., Geronimi, S., & Arnoux, E. (2024). Developing a methodology to assess data completeness of driving scenarios for testing autonomous vehicles: A focus on ODD-specific objectives. In: Proceedings of the Driving Simulation Conference 2024 Europe VR, Driving Simulation Association, Strasbourg, France, pp. 145-151.
- [6] Zhou, Z., Wang, Y., Xie, X., Chen, L., & Zhu, C. (2022). Foresee Urban Sparse Traffic Accidents: A Spatiotemporal Multi-Granularity Perspective. IEEE Transactions on Knowledge and Data Engineering, 34, 3786-3799. <https://doi.org/10.1109/tkde.2020.3034312>.
- [7] Eric Nalisnick and Akihiro Matsukawa and Yee Whye Teh and Dilan Gorur and Balaji Lakshminarayanan (2019) Do Deep Generative Models Know What They Don't Know?
- [8] Liu, C., Zhu, Z., Hao, W., & Sun, G. (2025). Heterogeneous multivariate time series imputation by transformer model with missing position encoding. *Expert Syst. Appl.*, 271, 126435. <https://doi.org/10.1016/j.eswa.2025.126435>.
- [9] Gharoun, H., Momenifar, F., Chen, F., & Gandomi, A. (2023). Meta-learning Approaches for Few-Shot Learning: A Survey of Recent Advances. *ACM Computing Surveys*, 56, 1 - 41. <https://doi.org/10.1145/3659943>.
- [10] Diwan, T., Anirudh, G., & Tembhurne, J. (2022). Object detection using YOLO: challenges, architectural successors, datasets and applications. *Multimedia Tools and Applications*, 82, 9243 - 9275. <https://doi.org/10.1007/s11042-022-13644-y>.
- [11] Lakshminarayanan, B. and Pritzel, A. and Blundell, C. (2017) Simple and Scalable Predictive Uncertainty Estimation using Deep Ensembles
- [12] J. Gawlikowski, C. R. N. Tassi, and M. Ali, "A survey of uncertainty in deep neural networks," Artificial Intelligence Review, July 2023

POSTE ET MISSIONS

Campus Paris-Saclay / 12, avenue Paul Delouvrier, 78180 Montigny Le Bretonneux

ESTACA, CREATEUR DE NOUVELLES MOBILITES

Le ou la doctorant(e) participera activement aux activités de recherche, en étroite collaboration avec les équipes de recherche académique et industrielle, et contribuera à la mise en œuvre de méthodes avancées d'IA pour faire face aux verrous identifiés

- **Début de la thèse** : septembre/octobre 2025 **Durée du contrat** : 36 mois

- **Missions principales** :
 - 1- **Etat de l'Art** : Mener une analyse complète des travaux récents sur la détection de scénarios de conduite, avec un accent particulier sur l'environnement urbain ; examiner les différentes représentations temporelles (séquences, patches, tokens) ainsi que les approches de classification multilabel ; dresser la synthèse des méthodes d'apprentissage frugal (few-shot, meta-learning) et des techniques d'augmentation de données par simulation; enfin, recenser les stratégies actuelles de quantification d'incertitude appliquées aux systèmes de détection.
 - 2- **Développement méthodologique** : Concevoir une représentation tensorielle unifiée capable de capturer la complexité des infrastructures urbaines et l'hétérogénéité des usagers ; explorer la tokenisation des séries temporelles et intégrer des backbones de type Transformer ; élaborer des modèles de classification multilabel spatio-temporels ; implanter des approches méta-apprenantes et bayésiennes afin de gérer l'incertitude et de détecter les données hors distribution.
 - 3- **Expérimentation et validation** : Évaluer les modèles sur des jeux de données réels ou simulés issus des campagnes VCS et MOOVE Urban ; développer un pipeline de détection « one-stage » à faible latence ; projeter les scénarios dans un espace latent pour identifier les régions sous-représentées et générer, via simulation, de nouveaux scénarios ciblés, assurant ainsi une meilleure couverture des cas critiques.
 - 4- **Valorisation scientifique** : Publier les résultats dans des revues et conférences internationales de référence ; contribuer aux livrables du projet MOOVE Urban et échanger étroitement avec les partenaires industriels pour transférer les avancées méthodologiques vers des applications concrètes.

PROFIL

- Ingénieure ou titulaire d'un Master 2 en informatique, intelligence artificielle ou mathématiques appliquées.

POLE DE RECHERCHE ET LIEU DU POSTE

S2ET (« SYSTEMES ET ENERGIES EMBARQUES DANS LES TRANSPORTS »)

Campus Paris-Saclay à Saint-Quentin-en-Yvelines / Institut VEDECOM à Versailles

CONTACTS

PERSONNES A CONTACTER :

TOUFIK AZIB (PR)
PROFESSEUR A L'ESTACA PARIS SACLAY
toufik.azib@estaca.fr

EQUIPE D'ENCADREMENT :

Campus Paris-Saclay / 12, avenue Paul Delouvrier, 78180 Montigny Le Bretonneux

ESTACA, CREATEUR DE NOUVELLES MOBILITES

CO-ENCADREMENT (ESTACA) : VINCENT JUDALET (DR)
ENSEIGNANT CHERCHEUR A L'ESTACA PARIS SACLAY
vincent.judalet@estaca.fr

CO-ENCADREMENT (VEDECOM) : WALID KHERIJI
walid.kheriji@vedecom.fr

CO-ENCADREMENT (VEDECOM) : FOUAD HADJ-SELEM
fouad.hadjselem@vedecom.fr