



AVIS DE SOUTENANCE DE THESE

Le Doyen de la Faculté des Sciences Dhar El Mahraz –Fès – annonce que

Mr BEN-FARESS Ilyass
Soutiendra : le Samedi 06/12/2025 à 14H30
Lieu : FSDM – Centre Visioconférence

Une thèse intitulée :

Explainable Machine Learning for Intelligent Transportation Systems

En vue d'obtenir le **Doctorat**

FD : **Sciences et Techniques**
Spécialité : **Intelligence Artificielle et Big Data**

Devant le jury composé comme suit :

| Nom et prénom | Etablissement | Grade | Qualité |
|----------------------|--|-------|-----------------------|
| LAMRINI Mohamed | Faculté des Sciences Dhar EL Mahraz, Fès | PES | Président |
| ZEMMOURI El Moukhtar | École Nationale Supérieure d'Arts et Métiers, Meknès | PES | Rapporteur |
| ADADI Amina | École Supérieure de Technologie, Meknès | MCH | Rapporteur |
| FARDOUSSE Khalid | Faculté Chariaa, Fès | MCH | Rapporteur |
| BOUAYAD Anas | Faculté des Sciences Dhar EL Mahraz, Fès | MCH | Examineur |
| BERRADA Ismail | Université Mohammed VI Polytechnique, Benguerir | MCH | Examineur |
| ZINEDINE Ahmed | Faculté des Sciences Dhar EL Mahraz, Fès | PES | Directeur de thèse |
| BOUHOUTE Afaf | Faculté des Sciences Dhar EL Mahraz, Fès | MCH | Co-Directeur de thèse |



Résumé :

Les avancées rapides de l'Intelligence Artificielle (IA) ont profondément transformé les Systèmes de Transport Intelligents (STI), permettant la prise de décision en temps réel, l'amélioration de la sécurité routière et l'automatisation avancée. Cependant, la dépendance croissante vis-à-vis des modèles d'apprentissage profond, souvent perçus comme des "boîtes noires", soulève des défis critiques en matière de transparence, de confiance et de responsabilité, en particulier dans les applications à fort enjeu de sécurité telles que la reconnaissance des panneaux de signalisation, la prédiction de la gravité des accidents et la détection des piétons.

Cette thèse aborde ces enjeux en explorant de manière systématique les techniques d'Intelligence Artificielle Explicable (XAI) afin de concilier performance prédictive et interprétabilité dans les STI. Elle débute par l'établissement d'un cadre théorique complet présentant les principales applications des STI, les modèles d'IA les plus utilisés, ainsi que les principes et taxonomies de la XAI. Un état de l'art des méthodes d'interprétabilité les plus récentes, telles que Grad-CAM, LIME, SHAP, les cartes de saillance et les explications contrefactuelles, y est également proposé.

Une revue systématique basée sur la méthodologie PRISMA est ensuite réalisée, analysant l'usage de la XAI dans quatre domaines des ITS : l'analyse et la prévention des accidents de la route, la reconnaissance des panneaux de signalisation (TSR), les véhicules autonomes et la prédiction ainsi que la gestion du trafic. Cette revue met en évidence les points forts, les limites et les perspectives de recherche.

La thèse démontre ensuite comment des techniques XAI représentatives peuvent être intégrées dans les pipelines de développement de modèles pour plusieurs problèmes ITS. Cela inclut des études de cas portant sur quatre tâches critiques : la reconnaissance des panneaux de signalisation, la prédiction de la gravité des accidents et la détection des piétons,

Les résultats confirment que l'explicabilité peut être atteinte sans compromettre la précision prédictive, renforçant ainsi la confiance, la responsabilité, la conformité réglementaire et l'analyse des erreurs lors des déploiements en conditions réelles.

Enfin, afin de pallier les principales limites des jeux de données existants, un nouveau jeu de données intitulé "Moroccan Traffic Sign Recognition (MTRSR)" a été introduit. Il comprend 2111 images de haute qualité réparties en 20 classes de panneaux de signalisation, collectées dans des conditions environnementales variées et enrichies de métadonnées spécifiques à la XAI. Les expériences de référence menées avec un modèle ResNet confirment des performances de classification compétitives. La comparaison des méthodes d'explicabilité montre que Score-CAM fournit les explications visuelles les plus fidèles et les plus stables, surpassant Grad-CAM, Grad-CAM++, XGrad-CAM et Integrated Gradients.

Dans l'ensemble, ce travail fournit un cadre solide pour le développement de modèles d'IA transparents et fiables dans les STI, contribuant à des infrastructures de transport plus sûres, plus éthiques et plus fiables.

Mots clés :

Intelligence Artificielle Explicable, Systèmes de Transport Intelligents, reconnaissance des panneaux de signalisation, prédiction de la gravité des accidents, détection des piétons.



Abstract :

The rapid advancement of Artificial Intelligence (AI) has significantly transformed Intelligent Transportation Systems (ITS), enabling real-time decision-making, improved road safety, and advanced automation. However, many deep learning models used in safety-critical tasks, such as traffic sign recognition, accident severity prediction, and pedestrian detection, operate as opaque black boxes, hindering trust, accountability, and regulatory acceptance.

This thesis addresses these challenges by systematically exploring Explainable Artificial Intelligence (XAI) techniques to reconcile predictive performance with interpretability in ITS.

First, a comprehensive theoretical framework is developed covering major ITS applications, commonly used AI models, and XAI taxonomies. State-of-the-art interpretability techniques, including Grad-CAM, LIME, SHAP, saliency maps, and counterfactual explanations, are also reviewed.

A PRISMA-based systematic review is then conducted, analyzing XAI usage across four ITS domains : Traffic Accident Analysis and Prevention, Traffic Sign Recognition (TSR), Autonomous Vehicles, and Traffic Prediction and Management. The review identifies strengths, limitations, and research gaps.

The thesis further demonstrates how representative XAI techniques can be incorporated into model development pipelines for multiple ITS problems, showing that transparent decision support can be achieved without degrading predictive performance. This includes case studies on traffic sign recognition, accident severity prediction, and pedestrian detection using CNN- and gradient-based explanation mechanisms.

To address dataset limitations, a novel Moroccan Traffic Sign Recognition (MTSR) dataset was introduced, comprising 2,111 high-quality images across 20 traffic sign classes collected under various environmental conditions. The dataset includes metadata supporting XAI evaluation. Baseline experiments using a ResNet-based model confirm competitive classification performance, and comparisons of attribution methods identify Score-CAM as offering the most stable and faithful explanations, outperforming Grad-CAM, GradCAM++, XGrad-CAM, and Integrated Gradients.

Overall, this work provides a robust framework for developing transparent and reliable AI models in ITS, supporting safer, more ethical, and more reliable transportation infrastructures.

Key Words:

Explainable AI, Intelligent Transportation Systems, traffic sign recognition, accident severity prediction, pedestrian detection.