

الجمهورية الشعبية الديمقراطية الجزائرية
People's Democratic Republic of Algeria
وزارة التعليم العالي و البحث العلمي
Ministry of Higher Education and Scientific Research
المدرسة العليا للإعلام الآلي 8 ماي 1945 - سيدي بلعباس
Higher School of Computer Science
8 Mai 1945 - Sidi Bel Abbès



Thesis

To obtain the diploma of **Engineering**
Field of Study: **Computer Science**
Specialization: **Artificial Intelligence and Data Science (AIDS)**

Theme

Design and Implementation of a Modular System for Real-Time Driver Inattention Detection

Presented by
YAHIA Abdelmadjid
Defended on: **July, 2025**
In front of the jury composed of

Dr. BOUSMAHA Rabab
Dr. NAOUM Hanae
Pr. BENSLIMANE Sidi Mohammed
Dr. KECHAR Mohamed
Dr. ELHANNANI Souad
Ms. GUENDIL Kawther Yasmine

President of the Jury
Supervisor
Co-Supervisor
Examiner
Examiner
Representative of the Economic Partner

Academic Year: 2024/2025

Abstract

This project addresses the critical issue of driver inattention by developing a non-intrusive, real-time monitoring system. A key contribution is its modular architecture, which leverages AI for feature extraction and rule-based logic for decision-making to enhance robustness. The system utilizes MediaPipe Face Mesh for drowsiness detection via adaptive Eye Aspect Ratio (EAR) and PERCLOS metrics, and a state-of-the-art D-FINE object detector to identify distractions by analyzing hand-object interactions.

A significant finding from our research is that a high-quality, curated dataset with efficient on-the-fly augmentation yields superior performance over larger, heavily augmented datasets. The resulting proof-of-concept was validated on custom real-world video scenarios, demonstrating functional viability at 15-20 FPS on a GPU. This work delivers a functional prototype that confirms the architectural approach and provides a strong foundation for a practical, low-cost driver safety solution.

Keywords— Driver Drowsiness Detection, Distraction Detection, Computer Vision, Modular Architecture, Deep Learning, Real-Time Systems, D-FINE, MediaPipe

ملخص

يعالج هذا المشروع المشكلة الحرجة المتمثلة في عدم انتباه السائق من خلال تطوير نظام مراقبة غير تدخلي وفي الوقت الفعلي. تتمثل إحدى المساهمات الرئيسية في هندسته المعمارية المعيارية (modular)، التي تستخدم الذكاء الاصطناعي لاستخراج الميزات ومنطقاً قائماً على القواعد لاتخاذ القرار، مما يعزز المتانة. يستخدم النظام MediaPipe Face Mesh لكشف التعابير عبر مقاييس تكيفية مثل نسبة أبعاد العين (EAR) و PERCLOS، وكاشف أجسام متطور، D-FINE، لتحديد مصادر التشتت عبر تحليل التفاعلات بين اليد والأجسام. من النتائج الهامة لبحثنا أن مجموعة البيانات عالية الجودة والمنسقة بعناية، مع زيادة فعالة للبيانات بشكل فوري، تحقق أداءً متفوقاً على مجموعات البيانات الأكبر حجماً والمعززة بشكل مكثف. تم التحقق من صحة النموذج الأولي الناتج على سيناريوهات فيديو واقعية مخصصة، مما أظهر جدواه الوظيفية بسرعة معالجة تتراوح بين 15-20 إطاراً في الثانية على وحدة معالجة الرسومات. يقدم هذا العمل نموذجاً أولياً وظيفياً يؤكد النهج المعماري المتبع ويوفر أساساً متيناً لحل عملي ومنخفض التكلفة لسلامة السائق.

الكلمات المفتاحية--- كشف تعابير السائق، كشف التشتت، الرؤية الحاسوبية، الهندسة المعيارية، التعلم العميق، أنظمة الوقت الفعلي، MediaPipe،D-FINE.

Résumé

Ce projet aborde le problème critique de l'inattention du conducteur en développant un système de surveillance non-intrusif et en temps réel. Une contribution clé est son architecture modulaire, qui utilise l'IA pour l'extraction de caractéristiques et une logique basée sur des règles pour la prise de décision, améliorant ainsi la robustesse. Le système utilise MediaPipe Face Mesh pour la détection de la somnolence via des métriques adaptatives comme le ratio d'aspect des yeux (EAR) et le PER-CLOS, et un détecteur d'objets de pointe, D-FINE, pour identifier les distractions en analysant les interactions main-objet.

Un résultat significatif de notre recherche est qu'un jeu de données de haute qualité, soigneusement sélectionné, avec une augmentation de données efficace à la volée, offre des performances supérieures à celles des jeux de données plus volumineux et fortement augmentés. Le prototype qui en résulte a été validé sur des scénarios vidéo réels personnalisés, démontrant sa viabilité fonctionnelle à 15-20 FPS sur un GPU. Ce travail présente un prototype fonctionnel qui confirme l'approche architecturale et fournit une base solide pour une solution de sécurité routière pratique et à faible coût.

Keywords— Détection de la somnolence du conducteur, Détection de la distraction, Vision par ordinateur, Architecture modulaire, Apprentissage profond, Systèmes temps réel, D-FINE, MediaPipe.