République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université Mohamed El Bachir El Ibrahimi – Bordj Bou Arréridj
Faculté des Mathématiques et d'Informatique
Département d'Informatique



MÉMOIRE
Présenté en vue de l'obtention du diplôme
Master en Informatique
Spécialité : Réseaux et Multimédia.

THÈME :
Détection de la somnolence du conducteur à l'aide de techniques d'intelligence artificielle

Présenté par :
SALI Oussama
Benameur Zahra

Devant le jury composé de :
Présidente : Dr, Benabid Sonia
Examinateur 1 : Dr, Saha Adel
Encadrante *:* Dr, REGOUID Meryem

2025/2026

Remerciements

En premier lieu, nous rendons grâce à notre Seigneur, Dieu Tout-Puissant, pour la grâce de succès qui nous a accompagnés tout au long de notre parcours jusqu'à ce travail que nous honorons ; Louange à Dieu pour ce beau travail.

Nous exprimons notre sincère gratitude à notre directrice de mémoire, Mme REGOUID Meryem, pour son encadrement, ses conseils, son dévouement à nous accompagner tout au long de ce travail. Nous en sommes très reconnaissants. Nous la remercions d'avoir partagé son expérience et ses conseils qui ont permis de bien travailler.

Nous remercions enfin les membres du jury pour le temps qu'ils ont bien voulu accorder à évaluer ce mémoire et à leurs commentaires pertinents.

Nous remercions aussi les enseignants du département d'informatique de l'université Mohamed El Bachir El Ibrahim de Bordj Bou Arréridj pour leur bonne qualité de formation et suivi dont ils ont su nous faire bénéficier durant notre formation théorique et pratique.

Dédicace

Je dédie ce modeste travail:

*À **mes chers parents**,*

Pour votre amour infini, vos prières, votre patience et vos encouragements constants. Vous avez toujours cru en moi et m'avez donné la force d'avancer. Ce travail est avant tout le fruit de vos sacrifices et de votre soutien inestimable.

*À **ma camarade et amie, Benameur Zahra**,*

Pour ton sérieux, ta persévérance et ta précieuse collaboration tout au long de ce projet. Merci pour ton esprit d'équipe, ton aide et ta bonne humeur.

*À **mes amis**,*

Pour leur présence, leur soutien et les moments de partage qui ont rendu ce parcours plus agréable et motivant.

*Enfin, à **tous mes professeurs**,*

Pour leurs efforts, leur disponibilité et la qualité de l'enseignement qu'ils nous ont transmis. Grâce à eux, j'ai acquis des connaissances et une passion qui m'accompagneront tout au long de ma carrière.

SALI Oussama

Dédicace

À l'occasion d'arriver au terme d'une étape décisive de mon parcours académique ,je dédie ce projet de fin d'études :

*À **ma famille** et **mes parents**,*
qui m'ont soutenue tout au long de ma vie,
qui m'ont conseillée et guidée vers le bon chemin,
et qui ont souffert pour que j'arrive là où je suis aujourd'hui.
Salutations à vous depuis cette plateforme et merci du fond du cœur.

*À **mon camarade, Sali Oussama**,*
merci pour la coopération, le respect et les efforts partagés
tout au long de la réalisation de ce projet.

*À **tous mes enseignants et professeurs,***
qui ont contribué à mon éducation et à mon orientation,
merci beaucoup pour vos efforts.
*Enfin, à **toutes les personnes ayant contribué à la réussite de ce travail**.*

BENAMEUR Zahra

RÉSUMÉ

Ce mémoire porte sur la conception et la réalisation d'un système de détection de la somnolence du conducteur basé sur l'analyse de la fermeture des yeux à partir d'images vidéo. L'objectif est d'identifier automatiquement les signes de somnolence en temps réel, en surveillant l'état des paupières du conducteur : lorsque le système détecte que les deux yeux restent fermés au-delà d'une durée critique, une alarme sonore est immédiatement déclenchée afin d'alerter le conducteur et prévenir tout accident potentiel.

Le système s'appuie sur des techniques de traitement d'images et d'intelligence artificielle. La vidéo est d'abord analysée afin de détecter le visage du conducteur à l'aide d'un détecteur de visage basé sur l'architecture Single Shot Detector (SSD), implémentée à l'aide du module Deep Neural Network (DNN) d'OpenCV. Les yeux sont ensuite localisés à l'aide des classifieurs de Haar. Les régions oculaires extraites sont analysées par un réseau de neurones convolutionnel (Convolutional Neural Network - CNN) spécialement entraîné pour distinguer les yeux ouverts des yeux fermés. L'état de vigilance du conducteur est suivi dynamiquement à l'aide d'un système de score permettant de différencier un simple clignement des yeux d'un véritable épisode de somnolence. Lorsque le score dépasse un seuil prédéfini, une alarme sonore est déclenchée afin d'alerter le conducteur.

Mots-clés : détection de la somnolence, détection de visage, classifieurs Haar, apprentissage profond, vision par ordinateur.

ABSTRACT

This thesis focuses on the design and implementation of a driver drowsiness detection system based on the analysis of eye closure from video images. The objective is to automatically identify signs of drowsiness in real time by monitoring the driver's eyelid state. When the system detects that both eyes remain closed for a critical period, an audible alarm is immediately triggered to alert the driver and help prevent potential accidents.

The proposed system relies on image processing and artificial intelligence techniques. The video is first analyzed to detect the driver's face using a face detector based on the Single Shot Detector (SSD) architecture, implemented through OpenCV's Deep Neural Network (DNN) module. The eyes are then localized using Haar Cascade classifiers. The extracted eye regions are analyzed by a Convolutional Neural Network (CNN) specifically trained to distinguish between open and closed eyes. The driver's level of alertness is continuously monitored using a scoring system that differentiates normal blinking from actual drowsiness episodes. When the score exceeds a predefined threshold, an audible alarm is activated to warn the driver.

Keywords: driver drowsiness detection, face detection, Haar classifiers, deep learning, computer vision.

ملخص

تتناول هذه المذكرة تصميم وإنجاز نظام للكشف عن نعاس السائق اعتمادًا على تحليل إغلاق العينين انطلاقًا من مقاطع الفيديو. يهدف النظام إلى التعرف تلقائيًا على علامات النعاس في الزمن الحقيقي من خلال مراقبة حالة جفني السائق. وعندما يكتشف النظام بقاء العينين مغلقتين لمدة حرجة، يتم تشغيل إنذار صوتي بشكل فوري لتنبيه السائق والمساهمة في الوقاية من الحوادث المحتملة.

يعتمد النظام المقترح على تقنيات معالجة الصور والذكاء الاصطناعي. يتم أولاً تحليل الفيديو للكشف عن وجه السائق باستخدام كاشف وجوه يعتمد على بنية SSD (Single Shot Detector)، والمُنفَّذة بواسطة وحدة DNN (Deep Neural Network) في مكتبة OpenCV. بعد ذلك يتم تحديد موضع العينين باستعمال مصنفات Haar Cascade. ثم تُحلَّل مناطق العين المستخرجة بواسطة شبكة عصبية التفافية CNN (Convolutional Neural Network) مدربة خصيصًا للتمييز بين العين المفتوحة والعين المغلقة. كما تتم متابعة مستوى يقظة السائق بشكل ديناميكي باستخدام نظام يعتمد على مقياس (Score) يسمح بالتمييز بين الرمش الطبيعي وحالات النعاس الحقيقية. وعند تجاوز هذا المقياس لعتبة محددة مسبقًا، يتم تشغيل إنذار صوتي لتنبيه السائق.

الكلمات المفتاحية : كشف نعاس السائق، كشف الوجه، مصنفات هار، التعلم العميق، الرؤية الحاسوبية.

Table des matières

LISTE DES FIGURES XI

LISTE DES TABLEAUX XIII

LISTE DES ABRÉVIATIONS XIV

Introduction générale 1

Chapitre 1 : Contexte et enjeux de la somnolence au volant 4

1.1 Introduction 4

1.2 Définition de la somnolence 4

1.2.1 Indicateurs de la somnolence 5

1.2.2 Facteurs favorisant la somnolence 5

1.3 Conséquences de la somnolence sur la conduite et la sécurité routière 6

1.4 Statistiques sur la somnolence au volant 7

1.4.1 À l'échelle mondiale 7

1.4.2 En Europe 7

1.4.3. Aux États- Unis 8

1.4.4 En Algérie 8

1.5 Problématique de la détection automatisée de la somnolence 9

1.6 Conclusion 10

Chapitre 2 : Avancées récentes dans la détection de la somnolence du conducteur 12

2.1 Introduction 12

2.2 Méthodes de détection de somnolence 12

2.2.1 Approche physiologique 13

2.2.2 Approche basée sur le comportement 14

2.2.3 Approche basée sur le véhicule 16

2.3 Travaux existants 17
2.3.1 Systèmes basés sur l'analyse des signaux physiologiques 17
2.3.2 Systèmes basés sur le comportement du conducteur 18
2.3.3 Systèmes basée sur le comportement du véhicule 19
2.4 Apports de la vision par ordinateur dans la détection 20
2.4.1 Détection du visage 20
2.4.2 Détection et localisation des yeux 21
2.4.3 Métriques d'évaluation des modèles 22
2.5 Conclusion 24
Chapitre 3 : Système de détection de la somnolence en conduite proposé 26
3.1 Introduction 26
3.2 Objectif du système proposé 26
3.3 Apprentissage profond (Deep Learning) 27
3.4 Architecture du système proposé 27
3.5 Modèle CNN proposé 28
3.5.1 Le jeu de données utilisé pour l'entraînement 28
3.5.2 Collecte et préparation des données 31
3.5.3 Préparation des données pour l'entraînement du modèle 32
3.5.4 Conception du modèle CNN 32
3.5.5 Entraînement du modèle 34
3.5.6 Suivi de l'apprentissage 34
3.5.7 Évaluation des performances 35
3.5.8 Sauvegarde du modèle 35
3.6 Intégration du modèle CNN dans le système proposé 35
3.6.1 Détection du visage 36
3.6.2 Détection des yeux 37
3.6.3 Prétraitement et préparation des images 38

3.6.4 Prédiction de l'état des yeux 38
3.6.5 Prédiction de l'état des yeux et calcul du score de somnolence 38
Figure 3.5 : Architecture de calcule du score de somnolence. 41
3.6.6 Jeu de données utilisé pour évaluation 44
3.7 Conclusion 46
Chapitre 4 : Évaluation expérimentale des performances du système de détection de la somnolence 48
4.1 Introduction 48
4.2 Environnement de travail 48
4.2.1 Équipements utilisés 48
4.2.2 Langages, bibliothèques, outils et plates-formes 49
4.3 Évaluation du modèle CNN 50
4.3.1 Analyse descriptive du jeu de données MRL Eye 50
4.3.2 Analyse des courbes d'apprentissage 52
4.3.3 Matrice de confusion 53
4.3.4 Discussion scientifique des résultats 56
4.3.5 Analyse critique des performances 58
4.3.6 Interprétation dans le contexte applicatif 59
4.4 Évaluation du système proposé 59
4.4.1 Analyse descriptive du jeu de données DDVD 60
4.4.2 Répartition des vidéos de test du jeu de données DDVD 63
4.4.3 Méthodologie d'évaluation 63
4.4.4 Résultats expérimentaux 64
4.4.4 Analyse des performances du système proposé 68
4.4.5 Discussion 69
4.5 Interface graphique utilisateur 70
4.5 Conclusion 73

Conclusion générale **74**

REFERENCES BIBLIOGRAPHIQUES **76**

LISTE DES FIGURES

Figure 1.1 : Statistiques de la somnolence au volant à l'échelle mondiale. 7
Figure 1.2 : Conséquences de la somnolence au volant en Europe 8
Figure 1.3 : Statistique sur les accidents causés par la somnolence en Algérie 9
Figure 2.1 : Les Méthodes de détection de somnolence [11]. 13
Figure 2.2 : Système en cours d'exécution. 14
Figure 2.3 : Prototype de l'œuvre proposée. 14
Figure 2.4 : Détection du visage par la caméra du Smartphone. 15
Figure 2.5 : Des exemples sur des illustration du fonctionnement du système. 15
Figure 2.6 : Détecteur de ligne sur véhicule. 17
Figure 2.7 : Capteurs de pression intégrés dans un siège pour la détection des signes 18
Figure 2.8 : Système de surveillance du conducteur DMS intégré 19
Figure 2.9 : Exemple du système de surveillance du conducteur DriverFocus 19
Figure 2.10 : L'affichage des alertes en cas de quitter le trajectoire ou gardez 20
Figure 3.1 : Diversité des images du jeu de données MRL Eye. 29
Figure 3.2 : Exemple d'annotation des images du jeu de données. 31
Figure 3.3 : Prétraitement d'une image du jeu de données MRL Eye. 32
Figure 3.4 : Architecture du modèle CNN proposé. 34
Figure 3.5 : Architecture de calcule du score de somnolence. 41
Figure 3.6 : L'architecture globale du système proposé. 43
Figure 4.1 : Analyse statistique du jeu de données MRL Eye. 51
Figure 4.2 : Distribution des classes du jeu de données MRL Eye. 51
Figure 4.3 : Évolution de l'exactitude durant l'entraînement et la validation. 52
Figure 4.4 : Évolution de la perte durant l'entraînement et la validation. 53
Figure 4.5 : Évaluation des performances par matrice de confusion. 54
Figure 4.6 : Distribution et caractéristiques du jeu de données DDVD. 62
Figure 4.7 : Premier exemple de simulation sur une vidéo générée 64
Figure 4.8 : Deuxième exemple de simulation sur une vidéo générée 65

Figure 4.9 : Exemple de simulation sur une vidéo réelle enregistrée. 65
Figure 4.10 : Exemple de simulation sur une vidéo téléchargée. 65
Figure 4.11 : Interface utilisateur du système de détection de la somnolence 71
Figure 4.12 : Interface de vidéo séléctionnée. 71
Figure 4.13 : : Interface détection en cours. 72
Figure 4.14 : Interface détection terminée. 72

LISTE DES TABLEAUX

Tableau 2.1 : Les différentes techniques de détection du visage. 21
Tableau 2.2 : Les différentes techniques de détection et localisation des yeux. 22
Tableau 3.1 : Description des annotations utilisées dans le jeu de données MRL Eye 30
Tableau 3.2 : Architecture détaillée du modèle CNN proposé. 33
Tableau 3.3 : Principe de mise à jour du score de somnolence 42
Tableau 3.4 : Structure d'annotation et convention de nommage des fichiers vidéo de jeu ... 45
Tableau 4.1 : Comparaison des performances du modèle proposé avec celles de travaux utilisant le jeu de données MRL Eye. 55
Tableau 4.2 : Comparaison des performances du modèle proposé avec celles de travaux utilisant d'autres jeux de données. 56
Tableau 4.3 : Vue globale du jeu de données DDVD. 61
Tableau 4.4 : Résultats de test sur le jeu de données DDVD. 67
Tableau 4.5 : Résultats globaux du système proposé. 68

LISTE DES ABRÉVIATIONS

SAAC :Systèmes Avancés d'Aide à la Conduite
BCG :Ballistocardiographie
CEW :Closed Eyes in the Wild
CNN :Convolutional Neural Network
CNPR :Centre National de Prévention Routière
DAC :Driver Alert Control
DDVD :Driver Drowsiness Video Dataset
DMS :Driver Monitoring System
DNN :Deep Neural Network
EAR :Eye Aspect Ratio
ECG :Électrocardiogramme
EEG :Électroencéphalogramme
OESR :Observatoire européen de la sécurité routière
FIFO :First In, First Out
FN :False Negative
FP :False Positive
GPU :Graphics Processing Unit
GSM :Global System for Mobile Communications
HOG :Histogram of Oriented Gradients
IA :Intelligence Artificielle
MRL :Media Research Lab
MTCNN :Multi-task Cascaded Convolutional Neural Network
NHTSA :National Highway Traffic Safety Administration
NumPy :Numerical Python
ONISR :Observatoire National Interministériel de la Sécurité Routière

OMS :Organisation Mondiale de la Santé

PCS :Pre-Collision System

RAM :Random Access Memory

SCAC :Système de Contrôle de l'Attention du Conducteur

SCG :Sismocardiographie

SSD :Single Shot MultiBox Detector

SSD :Solid-State Drive

TN :True Negative

TP :True Positive

YOLO :You Only Look Once

YawDD :Yawning Detection Dataset

Introduction générale

Avec l'augmentation rapide du nombre de véhicules à travers le monde, les accidents de la route représentent aujourd'hui un enjeu majeur de sécurité publique, entraînant chaque année des pertes humaines et matérielles considérables. Parmi les différentes causes identifiées, la somnolence du conducteur constitue un facteur particulièrement critique, car elle altère directement la concentration, le temps de réaction ainsi que la capacité de prise de décision.

En effet, un conducteur en état de somnolence peut perdre progressivement le contrôle de son véhicule, augmentant ainsi le risque d'accidents graves. Plusieurs études ont démontré que la somnolence, notamment caractérisée par une fermeture prolongée des yeux, est responsable d'un nombre important d'accidents de la circulation.

Afin de faire face à cette problématique, le développement de systèmes intelligents capables de détecter automatiquement l'état de somnolence du conducteur s'impose comme une solution prometteuse. Bien que certaines technologies de détection de la somnolence soient déjà intégrées dans des véhicules modernes, leur utilisation reste limitée à une catégorie restreinte de voitures.

Par ailleurs, quelques dispositifs existent sous forme de produits commercialisés, mais leur accessibilité demeure limitée et leur adoption reste faible, en particulier dans les pays en développement comme l'Algérie, où la majorité des véhicules ne disposent pas de telles technologies. Ce constat met en évidence le besoin de solutions accessibles, efficaces et adaptées au contexte local.

Grâce aux avancées réalisées dans le domaine de la vision par ordinateur et de l'intelligence artificielle, il est désormais possible d'analyser des images en temps réel et de détecter les signes de somnolence. Parmi les techniques utilisées, le détecteur de visage SSD permet de localiser efficacement le visage du conducteur, tandis que les classifieurs Haar Cascade assurent la détection des yeux. Les régions oculaires extraites sont ensuite analysées par un réseau de neurones convolutionnel (CNN) capable de distinguer avec précision les yeux ouverts des yeux fermés, permettant ainsi d'identifier les épisodes de somnolence.

Dans ce contexte, notre travail consiste à concevoir et à développer un système intelligent de détection de la somnolence basé sur l'analyse de l'état des yeux, en utilisant les techniques de vision par ordinateur et l'apprentissage profond. Le système analyse une vidéo du conducteur, détecte le visage à l'aide d'un détecteur SSD implémenté via le module DNN d'OpenCV, localise les yeux grâce aux classifieurs Haar Cascade, puis utilise un modèle CNN pour déterminer si les yeux sont ouverts ou fermés.

Enfin, un système de score permet d'évaluer l'état de vigilance du conducteur et de déclencher une alarme sonore lorsque des signes de somnolence sont détectés.

L'objectif principal de ce projet est de contribuer à l'amélioration de la sécurité routière en proposant une solution efficace, automatique et fonctionnant en temps réel, tout en répondant au manque de systèmes similaires accessibles sur le marché, notamment dans le contexte algérien.

Le présent mémoire est structuré en quatre chapitres :

- Le premier chapitre traite de la somnolence au volant, de ses causes et de ses conséquences sur la sécurité routière. Il met également en évidence l'importance de cette problématique à travers diverses statistiques et souligne la nécessité de disposer d'un système de détection en temps réel.
- Le deuxième chapitre présente les principales approches de détection de la somnolence, notamment les méthodes physiologiques et comportementales. Il comprend également une étude des travaux existants ainsi qu'une présentation des techniques de détection basées sur la vision par ordinateur.
- Le troisième chapitre décrit la conception du système proposé, fondé sur un modèle d'apprentissage profond (CNN) destiné à la classification de l'état des yeux. Il détaille également les étapes de préparation des données, d'entraînement du modèle et de son intégration dans un système de détection de la somnolence.
- Enfin, le quatrième chapitre est consacré à l'évaluation des performances du modèle CNN ainsi que du système de détection de la somnolence. Les résultats obtenus confirment la robustesse de l'approche proposée et son efficacité pour une utilisation en temps réel dans un contexte de sécurité routière.

Chapitre 1:

Contexte et enjeux de la somnolence au volant.

Chapitre 1
Contexte et enjeux de la somnolence au volant

1.1 Introduction

La sécurité routière représente un enjeu majeur à l'échelle mondiale. Les accidents liés à la somnolence du conducteur causent chaque année de nombreuses pertes humaines et matérielles. La somnolence affecte la vigilance, la concentration et le temps de réaction, augmentant ainsi le risque d'accidents graves.

Dans ce chapitre, nous abordons la problématique de la somnolence du conducteur en définissant ce phénomène et en identifiant ses principales causes ainsi que ses conséquences sur la sécurité routière. Nous présentons également des statistiques mondiales et nationales relatives à ce problème, avant d'exposer la problématique de la détection automatisée de la somnolence et la nécessité de développer un système intelligent capable d'alerter le conducteur en temps réel.

1.2 Définition de la somnolence

La somnolence est un état physiologique caractérisé par une forte envie de dormir, accompagnée d'une diminution du niveau de vigilance, d'une baisse de l'attention et d'un ralentissement des capacités de réaction. Elle correspond à un état intermédiaire entre l'éveil et le sommeil et peut affecter significativement les performances cognitives et comportementales d'un individu. Dans le contexte de la conduite automobile, la somnolence constitue un facteur de risque majeur, car elle réduit la capacité du conducteur à percevoir son environnement, à prendre des décisions rapides et à réagir efficacement aux situations imprévues [1].

1.2.1 Indicateurs de la somnolence

La somnolence se manifeste par plusieurs signes observables permettant d'évaluer le niveau de vigilance d'un individu. Parmi ces signes, la fermeture prolongée des yeux est considérée comme l'un des indicateurs les plus fiables. Lorsqu'un conducteur maintient les yeux fermés pendant une durée anormalement longue, sa perception visuelle diminue fortement, ce qui augmente considérablement le risque d'accident.

D'autres manifestations peuvent également révéler un état de somnolence, notamment le clignement fréquent ou ralenti des yeux, les bâillements répétés, les difficultés à maintenir le regard fixé sur la route, les mouvements irréguliers des yeux. ainsi que les mouvements involontaires de la tête.

Ces indicateurs sont largement utilisés dans les systèmes modernes de détection de la somnolence basés sur la vision par ordinateur. Toutefois, la fermeture prolongée des yeux demeure l'un des indicateurs les plus utilisés pour évaluer la perte de vigilance d'un conducteur [2].

1.2.2 Facteurs favorisant la somnolence

La somnolence au volant peut être provoquée par plusieurs facteurs, souvent combinés :

- **Manque de sommeil** : Une durée de sommeil insuffisante réduit les capacités cognitives, diminue l'attention et augmente le temps de réaction du conducteur.
- **Fatigue physique et mentale** : Les efforts prolongés, le stress, la charge de travail importante ou les activités nécessitant une concentration soutenue favorisent la somnolence.
- **Conduite prolongée** : Les longs trajets, particulièrement sur autoroute, peuvent entraîner une diminution progressive de la vigilance et favoriser l'apparition de micro-sommeils.
- **Conduite nocturne** : Le rythme biologique humain favorise naturellement le sommeil pendant la nuit, ce qui augmente le risque de somnolence chez les conducteurs.
- **Conditions physiologiques et médicales** : Certaines pathologies, telles que l'apnée du sommeil, ainsi que la prise de médicaments sédatifs ou de substances pouvant réduire la vigilance, augmentent le risque de somnolence au volant.

Selon Abulkhair et al. [3], la combinaison d'un manque de sommeil, d'une conduite prolongée et d'un déplacement effectué durant les heures nocturnes constitue l'une des situations les plus critiques en matière de sécurité routière.

1.3 Conséquences de la somnolence sur la conduite et la sécurité routière

La somnolence au volant représente un risque important pour la sécurité routière en raison de ses effets sur les capacités de conduite. Ses principales conséquences sont présentées dans les points suivants :

- **Diminution de la vigilance et de la concentration** : La somnolence réduit la capacité du conducteur à maintenir une attention constante sur la route et à réagir efficacement aux changements de circulation. Selon l'Observatoire européen de la sécurité routière (OESR), la somnolence entraîne une baisse significative de la vigilance ainsi qu'une dégradation des performances de conduite, augmentant ainsi le risque d'accident [4].
- **Allongement du temps de réaction** : Un conducteur somnolent présente un temps de réaction plus long face aux situations dangereuses, augmentant ainsi le risque de collision. l'Administration nationale de la sécurité routière (National Highway Traffic Safety Administration - NHTSA) rapporte que la somnolence altère la prise de décision rapide et ralentit les réponses du conducteur [5].
- **Perte de contrôle du véhicule** : Les épisodes de micro-sommeil peuvent entraîner une perte totale de contrôle du véhicule, provoquant des sorties de route ou des collisions graves. Selon l'Observatoire National Interministériel de la Sécurité Routière (ONISR), la somnolence constitue un facteur important des accidents mortels sur autoroute, en raison de la diminution de la vigilance et de la perte de contrôle du conducteur [5].
- **Danger pour les autres usagers** : La somnolence peut également provoquer des comportements imprévisibles tels que des déviations involontaires, des freinages brusques ou des accélérations inadaptées, mettant en danger les autres usagers de la route. Des études en sécurité routière montrent que le manque de sommeil peut être aussi dangereux que la conduite sous l'influence de l'alcool ou de drogues.
- **Autres conséquences indirectes** : La somnolence entraîne également plusieurs conséquences indirectes susceptibles d'affecter la sécurité routière. Elle provoque une diminution des performances cognitives du conducteur, réduisant ainsi sa capacité de concentration, d'analyse et de prise de décision. Elle favorise également l'apparition de quasi-accidents, qui constituent souvent des indicateurs d'un danger imminent. En outre, un conducteur somnolent représente un risque important pour les autres usagers de la route, car il peut provoquer des collisions impliquant plusieurs véhicules, entraînant des blessures corporelles et des dommages matériels parfois considérables.

1.4 Statistiques sur la somnolence au volant

1.4.1 À l'échelle mondiale

- Selon une revue systématique de la littérature scientifique, la somnolence au volant constitue un problème majeur de sécurité routière à l'échelle mondiale. Les études analysées montrent qu'elle contribuerait à environ 3 % à plus de 30 % des accidents de la route, selon les pays et les méthodes d'évaluation utilisées. Cette variabilité reflète les difficultés méthodologiques liées à l'identification précise de la somnolence comme facteur contributif des accidents. Néanmoins, les résultats de cette revue confirment que la somnolence représente un facteur de risque important pouvant altérer la vigilance, le temps de réaction et la capacité de prise de décision du conducteur [6]. La figure 1.1 présente une synthèse de l'impact de la somnolence au volant à l'échelle mondiale.



Figure 1.1 : Statistiques de la somnolence au volant à l'échelle mondiale [6].

1.4.2 En Europe

- Selon Gonçalves et al. (2015), une enquête menée auprès de 12 434 conducteurs dans 19 pays européens montre qu'environ 17 % des conducteurs ont déclaré avoir somnolé ou s'être endormis au volant au cours des deux dernières années. Parmi eux, 7 % ont indiqué qu'un épisode de somnolence avait entraîné un accident, dont 13,2 % ayant nécessité une hospitalisation et 3,6 % ayant conduit à un décès [7]. La figure 1.2 résume les résultats de cette enquête menée auprès des conducteurs européens.

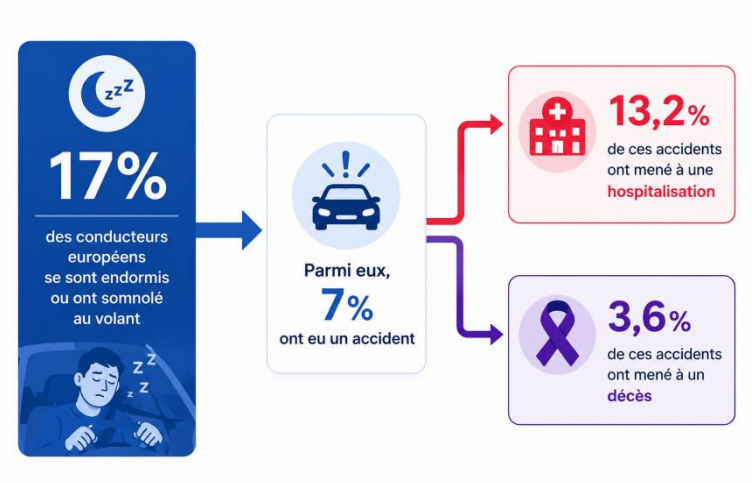


Figure 1.2 : Conséquences de la somnolence au volant en Europe
Selon Gonçalves et al (2015) [7].

- Selon une enquête réalisée par Ipsos pour la Fondation VINCI Autoroutes (2017), 38 % des conducteurs européens considèrent la somnolence comme l'une des principales causes des accidents mortels sur autoroute [8].
- Par ailleurs, environ 26 % des conducteurs déclarent avoir déjà ressenti un assoupissement au volant, et 16 % indiquent avoir déjà dévié de leur trajectoire ou empiété sur le bas-côté en raison de la somnolence [8].

1.4.3. Aux États- Unis

- Selon la National Highway Traffic Safety Administration (NHTSA), environ 91000 accidents impliquant des conducteurs somnolents ont été enregistrés en 2017, entraînant près de 50 000 blessés et environ 800 décès liés à la somnolence au volant [9].

1.4.4 En Algérie

- En Algérie, les données relatives à la somnolence au volant restent limitées. Selon M. Naït Hocine, directeur général du Centre national de prévention routière (CNPR), environ 10 % des accidents de la route enregistrés en 2016, soit près de 2 500 accidents, seraient attribués à la somnolence au volant. Ces déclarations ont été rapportées par l'Algérie Presse Service (APS, 2016) [10], à l'occasion d'un congrès consacré aux troubles du sommeil et à la sécurité routière. La figure 1.3 illustre la part des accidents de la route liés à la somnolence au volant en Algérie, estimée à 10 % des accidents enregistrés en 2016.

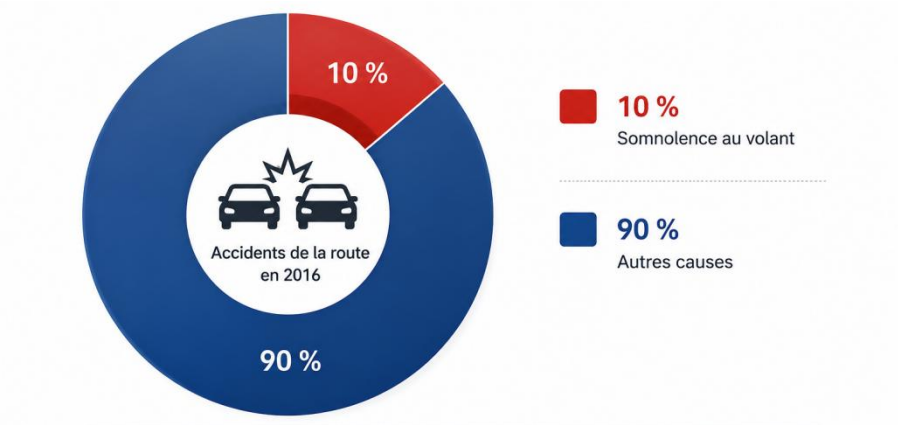


Figure 1.3 : Statistique sur les accidents causés par la somnolence en Algérie en 2016 selon CNPR [10].

Certaines sources mentionnent que, sur certaines grandes routes ou autoroutes, la somnolence pourrait être une cause significative derrière les accidents les plus graves, mais ces chiffres ne sont pas systématiquement mesurés par des organismes officiels récents.

1.5 Problématique de la détection automatisée de la somnolence

Malgré les progrès technologiques dans le domaine automobile, la majorité des véhicules ne disposent pas encore de systèmes capables de détecter automatiquement la somnolence du conducteur. Les solutions existantes sont souvent limitées par :

- **Le coût élevé** : rendant ces systèmes peu accessibles au grand public.
- **La complexité technique** : nécessitant des équipements sophistiqués ou des capteurs multiples.
- **L'accessibilité réduite** : avec des dispositifs généralement intégrés uniquement dans des véhicules haut de gamme.

Ces limites soulignent la nécessité de développer une solution intelligente, abordable et efficace permettant de détecter en temps réel les signes de somnolence du conducteur à travers l'analyse de la fermeture prolongée des yeux et de déclencher une alerte préventive afin de réduire le risque d'accidents.

La problématique centrale de ce projet consiste à concevoir une solution fiable, accessible et adaptable à différents types de véhicules, capable d'identifier automatiquement les états de somnolence du conducteur et de déclencher une alerte en cas de détection de ce phénomène. Cette solution doit également présenter un fort potentiel de commercialisation en tant

qu'innovation technologique contribuant au renforcement de la sécurité routière, aussi bien en Algérie qu'à l'international.

Cette approche repose sur la combinaison des techniques de vision par ordinateur et d'intelligence artificielle afin de proposer une solution automatisée, performante et applicable dans des conditions réelles de conduite, contribuant ainsi à l'amélioration de la sécurité routière.

1.6 Conclusion

Ce chapitre a mis en évidence l'importance de la somnolence comme facteur de risque majeur sur la route, en montrant ses causes et ses conséquences sur la sécurité des conducteurs et des usagers. Les statistiques consultées confirment l'ampleur du problème et soulignent la nécessité de développer des solutions technologiques fiables.

Le chapitre suivant présentera un état de l'art détaillé des méthodes de détection actuelles, ainsi que les techniques de vision par ordinateur qui constituent une partie importante de notre système.

Chapitre 2 :

Avancées récentes dans la détection de la somnolence du conducteur.

Chapitre 2

Avancées récentes dans la détection de la somnolence du conducteur

2.1 Introduction

L'augmentation des accidents de la route dans le monde a incité les chercheurs à développer des systèmes capables de réduire ce problème. Plusieurs solutions ont été proposées pour détecter la somnolence du conducteur en temps réel.

Ce chapitre présente les principales méthodes utilisées pour la détection de la somnolence. Il met également en évidence les travaux de recherche existants dans ce domaine et souligne l'apport des techniques de vision par ordinateur dans l'analyse du visage et des yeux pour la détection de la somnolence.

2.2 Méthodes de détection de somnolence

Plusieurs méthodes ont été proposées dans la littérature pour détecter la somnolence du conducteur en temps réel. Selon Ramzan et al. [11], ces méthodes sont classées en trois catégories. La figure 2.1 présente cette classification.

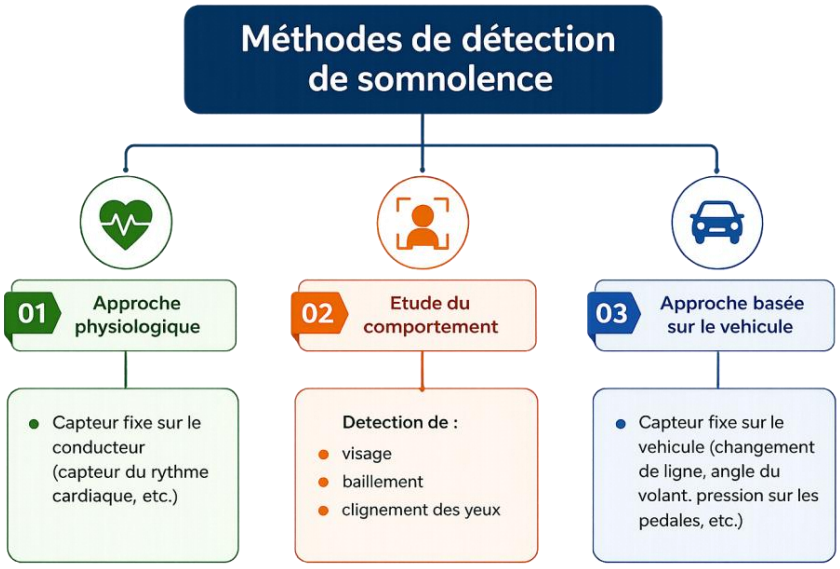


Figure 2.1 : Les Méthodes de détection de somnolence [11].

Comme le montre la Figure 2.1, les méthodes de détection de la somnolence sont classées en trois catégories principales : les approches physiologiques, les approches basées sur le comportement du conducteur et les approches basées sur le véhicule. Chacune de ces approches utilise des indicateurs différents pour évaluer l'état de vigilance du conducteur.

2.2.1 Approche physiologique

L'approche physiologique correspond à une méthode de détection de la somnolence basée sur l'analyse des données biologiques collectées à l'aide de capteurs placés sur le corps du conducteur, tels que les capteurs de rythme cardiaque ou de clignement des yeux.

E. Rogado, J.L. García, R. Barea et L.M. Bergasa 2009 [12], les auteurs ont proposé une méthode permettant de détecter les premiers signes de somnolence en se basant sur l'analyse de plusieurs paramètres physiologiques et environnementaux, notamment la fréquence cardiaque, la pression exercée sur le volant et la variation de température entre l'intérieur et l'extérieur du véhicule. L'acquisition et le traitement de ces données permettent d'estimer le niveau de somnolence du conducteur de manière indirecte et d'améliorer la sécurité de la conduite.

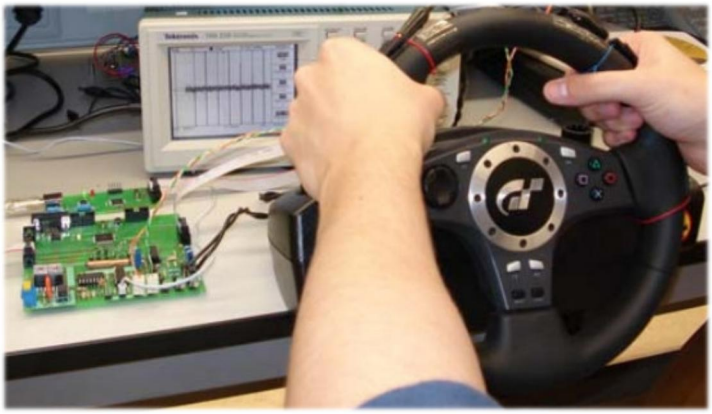


Figure 2.2: Système en cours d'exécution.

Selon M. Arunasalam, N. Yaakob, A. Amir, M. Elshaikh et N. F. Azahar, 2019 [13], la détection de la somnolence peut être réalisée en exploitant les données issues d'un capteur de fréquence cardiaque et d'un capteur de clignement des yeux. Les données collectées sont ensuite traitées afin de déterminer l'état du conducteur. En cas de détection de somnolence, le système agit automatiquement en réduisant progressivement la vitesse du véhicule jusqu'à son arrêt complet. Par ailleurs, l'état du conducteur ainsi que sa position géographique sont transmis vers le Cloud à l'aide d'un module Global System for Mobile Communications (GSM) pour assurer le suivi en temps réel.

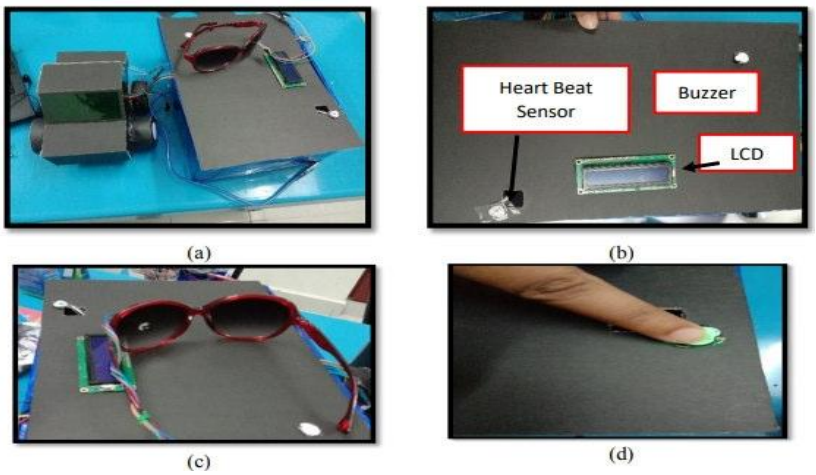


Figure 2.3 : Prototype de l'œuvre proposée.

2.2.2 Approche basée sur le comportement

L'approche basée sur le comportement du conducteur consiste à analyser les changements visibles sur le visage afin d'identifier les signes de somnolence. Cette méthode repose principalement sur la surveillance de certains indicateurs, notamment l'état des yeux

et les mouvements de la bouche, qui représentent des paramètres importants pour évaluer le niveau de vigilance.

Maysoon Abulkhair, Arwa H. Alsahli, Kawther M. Taleb, Atheer M. Bahran, Fatimah M. Alzahrani, Hend A. Alzahrani et Lamiaa Fattouh Ibrahim, 2015 [14], les auteurs ont développé un système de détection de la somnolence adapté à un Smartphone. Cette solution a été proposée afin de remplacer les ordinateurs, qui sont difficiles à installer dans un véhicule. Le Smartphone est positionné devant le conducteur à l'aide d'un support, et sa caméra est utilisée pour capturer les images du visage. Ces images sont ensuite traitées à l'aide d'un algorithme de traitement d'image afin de déterminer l'état des yeux. Lorsqu'un état de somnolence est détecté, une alarme sonore est déclenchée pour avertir le conducteur.



Figure 2.4 : Détection du visage par la caméra du Smartphone.

Jie Yi Wong et Phooi Yee Lau, 2019 [15], les auteurs ont proposé un système embarqué basé sur la carte Raspberry Pi 3 Model B pour surveiller l'état du conducteur en temps réel. Une caméra connectée au Raspberry Pi permet de capturer les images, qui sont ensuite analysées pour détecter la fermeture des yeux. La méthode utilisée repose sur le calcul du rapport entre la hauteur et la largeur des yeux. Lorsque les yeux restent fermés pendant une durée supérieure à deux secondes, le système considère que le conducteur est en état de somnolence et active une alarme sonore afin de l'alerter.



Figure 2.5 : Des exemples sur des illustration du fonctionnement du système.

2.2.3 Approche basée sur le véhicule

L'approche basée sur le comportement du véhicule consiste à détecter la somnolence du conducteur en analysant les données fournies par des capteurs installés sur différents éléments du véhicule, tels que le volant, la pédale d'accélération ou encore les systèmes de suivi de trajectoire. Ces capteurs permettent de mesurer certains paramètres liés à la conduite, qui peuvent refléter le niveau de vigilance du conducteur.

Selon Arun Sahayadhas, Kenneth Sundaraj et Murugappan Murugappan, 2012 [16], la fatigue et la privation de sommeil ont un impact direct sur les performances de conduite, notamment sur la stabilité du véhicule et le contrôle du volant. Parmi les indicateurs les plus utilisés, on distingue le mouvement du volant et l'écart type de la position sur la voie.

Le mouvement du volant est mesuré à l'aide d'un capteur d'angle installé sur la colonne de direction. Ce paramètre permet d'évaluer le nombre de corrections effectuées par le conducteur pour maintenir la trajectoire du véhicule. En condition normale, le conducteur réalise des micro-corrections fréquentes. En revanche, en cas de somnolence, ces corrections deviennent moins nombreuses, ce qui traduit une diminution de l'attention. Il est important de noter que le mouvement du volant ne mesure pas directement la position des mains sur le volant, mais analyse les variations de l'angle de braquage et les mouvements du volant. Ainsi, des mouvements moins fréquents, retardés ou irréguliers peuvent constituer des indicateurs de fatigue ou de somnolence chez le conducteur.

L'écart type de la position sur la voie correspond à la variation de la position du véhicule par rapport à sa trajectoire. Cette mesure est généralement obtenue à l'aide d'une caméra installée à l'extérieur du véhicule. Une augmentation de cette variation peut indiquer un état de somnolence du conducteur. Toutefois, ce paramètre peut être influencé par des facteurs externes tels que l'état de la route, les conditions climatiques ou la qualité de l'éclairage.

En résumé, bien que les méthodes basées sur le comportement du véhicule permettent de surveiller l'état du conducteur de manière indirecte, leur précision reste limitée, car les paramètres mesurés peuvent être affectés par d'autres facteurs non liés à la somnolence, tels que l'alcool, les médicaments ou les conditions de conduite.



Figure 2.6 : Détecteur de ligne sur véhicule.

2.3 Travaux existants

De nombreuses recherches et développements industriels ont été menés pour détecter la somnolence du conducteur et améliorer la sécurité routière. Ces travaux existants se concentrent principalement sur trois types de systèmes :

2.3.1 Systèmes basés sur l'analyse des signaux physiologiques

Les systèmes de détection de la somnolence basés sur les signaux physiologiques exploitent des paramètres biologiques du conducteur, notamment l'électrocardiographie (ECG), la respiration et les micro-mouvements corporels. L'analyse de la variabilité de la fréquence cardiaque ainsi que des techniques comme la ballistocardiographie (BCG) et la sismocardiographie (SCG) permettent d'estimer l'état de vigilance de manière non intrusive.

Selon des travaux scientifiques [17][18], ces capteurs peuvent être intégrés dans les sièges ou le volant des véhicules afin de surveiller le conducteur en temps réel.

Toutefois, malgré leur potentiel élevé, ces systèmes restent principalement au stade de recherche et de prototypes industriels en raison de contraintes techniques et d'intégration.

Ainsi, ils sont considérés comme des technologies émergentes au sein des Systèmes Avancés d'Aide à la Conduite (SAAC). La figure 2.7 présente un siège équipé de capteurs de pression permettant de surveiller la posture du conducteur et de détecter les signes de fatigue ou de somnolence à partir des variations de pression exercées sur le siège.



Figure 2.7 : Capteurs de pression intégrés dans un siège pour la détection des signes de fatigue ou de somnolence

2.3.2 Systèmes basés sur le comportement du conducteur

Les systèmes basés sur le comportement du conducteur visent à détecter la somnolence, l'inattention ou la distraction en analysant directement les signaux comportementaux du conducteur, notamment les mouvements des yeux, les expressions du visage et la direction du regard. Ces systèmes utilisent généralement des caméras embarquées, souvent infrarouges, ainsi que des techniques de vision par ordinateur pour évaluer en temps réel l'état de vigilance du conducteur.

Par exemple, le système de surveillance du conducteur (DMS, Driver Monitoring System), également appelé Système de Contrôle de l'Attention du Conducteur(SCAC), est un dispositif de sécurité embarqué permettant d'évaluer le niveau d'attention du conducteur afin de détecter les signes de distraction ou de somnolence et de déclencher des alertes appropriées. Introduit par Toyota en 2006, notamment sur le véhicule Lexus GS 450h, ce système fonctionne en complément du système de pré-collision afin d'améliorer la sécurité du véhicule. Il repose sur une caméra et des capteurs infrarouges installés sur la colonne de direction, permettant de surveiller en temps réel le visage du conducteur ainsi que les mouvements de ses yeux. En cas de distraction détectée dans une situation à risque, le système émet d'abord des alertes visuelles et sonores. Si le conducteur ne réagit pas, il peut activer automatiquement le freinage afin de réduire le risque d'accident [19]. La figure 2.8 illustre un système de surveillance du conducteur intégré au tableau de bord d'un véhicule.



Figure 2.8 : Système de surveillance du conducteur DMS intégré au tableau de bord.

De même, Subaru a intégré dans certains de ses véhicules le système de surveillance du conducteur DriverFocus, qui utilise une caméra embarquée pour analyser le visage et le regard du conducteur. Ce système permet de détecter l'inattention ou la somnolence, de prévenir le conducteur en cas de distraction, et peut également reconnaître différents conducteurs afin d'ajuster automatiquement certains paramètres du véhicule comme la position du siège et des rétroviseurs [20]. La figure 2.9 présente le système DriverFocus développé par Subaru, qui surveille le visage et le regard du conducteur afin de détecter l'inattention et la somnolence en temps réel.



Figure 2.9 : Exemple du système de surveillance du conducteur DriverFocus développé par Subaru.

2.3.3 Systèmes basée sur le comportement du véhicule

Certains systèmes de détection de la somnolence reposent sur des capteurs de comportement du véhicule, installés pour surveiller les actions du conducteur. Toyota et Ford, par exemple, utilisent des capteurs qui suivent les mouvements du volant et la position du

véhicule sur la voie. Ces technologies permettent de détecter des comportements anormaux, comme une réduction des corrections de trajectoire, un indicateur fréquent de somnolence. Lorsqu'une telle anomalie est détectée, une alarme est déclenchée pour prévenir le conducteur. Ces systèmes utilisent les données collectées par des capteurs dans le véhicule, offrant ainsi une solution efficace pour détecter la somnolence sans intrusivité excessive (Sahayadhas et al., 2012 ; Toyota Safety Report, 2017) [16][21]. La figure 2.10 illustre un exemple d'affichage des alertes générées en cas de sortie de voie ou de retrait des mains du volant.



Figure 2.10 : L'affichage des alertes en cas de quitter le trajectoire ou gardez les mains hors du volant.

2.4 Apports de la vision par ordinateur dans la détection

La vision par ordinateur est une branche de l'intelligence artificielle qui permet à une machine d'analyser et d'interpréter des images ou des vidéos de manière automatique. Elle est largement utilisée pour la détection de la somnolence du conducteur, car elle permet de suivre en temps réel des indices visuels tels que la position du visage et la localisation des yeux.

Le processus de détection visuelle repose principalement sur deux étapes successives. La première consiste à détecter le visage du conducteur, tandis que la seconde vise à localiser précisément les yeux à l'intérieur du visage détecté.

2.4.1 Détection du visage

La détection du visage consiste à localiser la région du visage dans une image ou une séquence vidéo, ce qui constitue une étape préalable essentielle avant toute analyse détaillée des yeux ou des expressions faciales. Le tableau 2.1 présente les principales méthodes de détection du visage ainsi que leurs avantages et leurs limites.

Tableau 2.1 : Les différentes techniques de détection du visage.

Méthode	Description	Avantages	Limites
Haar Cascades (Viola & Jones, 2001) [22]	Méthode basée sur les caractéristiques de Haar et un classificateur en cascade permettant la détection rapide du visage.	Détection très rapide en temps réel, adaptée aux systèmes embarqués.	Sensible aux variations d'éclairage, aux angles de tête et aux occlusions partielles, ce qui la rend moins robuste dans des conditions réelles.
HOG (Dalal & Triggs, 2005) [23]	Analyse la distribution des gradients d'intensité pour extraire les contours du visage.	Plus robuste que Haar, notamment pour les visages inclinés ou partiellement masqués.	Présente des performances inférieures à celles des méthodes d'apprentissage profond dans des environnements complexes.
Deep Neural Networks (DNN) [24]	Utilisation de réseaux de neurones profonds (SSD, ResNet, YOLO (You Only Look Once), etc.) pour la détection du visage, souvent via le module de réseaux de neurones profonds d'OpenCV (DNN).	Très haute précision et bonne robustesse face aux variations d'éclairage et de pose.	Coût de calcul plus élevé et besoin potentiel de ressources matérielles importantes.
MTCNN (Zhang et al., 2016) [25]	Réseau convolutionnel en cascade permettant la détection du visage et des points clés faciaux (yeux, nez, bouche).	Très précis et adapté aux applications nécessitant les landmarks faciaux.	Complexité et consommation de ressources supérieures aux méthodes classiques.

2.4.2 Détection et localisation des yeux

Une fois le visage détecté, la localisation des yeux permet de repérer précisément leur position pour les étapes ultérieures d'analyse comportementale.Plusieurs techniques ont été proposées pour la détection et la localisation des yeux.

Le tableau 2.2 présente les principales méthodes utilisées ainsi que leurs avantages et leurs limites.

Tableau 2.2 : Les différentes techniques de détection et localisation des yeux.

Méthode	Description	Avantages	Limites
Cascade de Haar pour les yeux (Viola & Jones, 2001) [22]	Extension de la méthode Cascade de Haar appliquée spécifiquement à la détection des yeux.	Rapide, simple à mettre en œuvre et adaptée aux systèmes fonctionnant en temps réel.	Sensible aux variations d'éclairage, aux rotations de la tête et aux conditions d'acquisition non contrôlées.
Détection des points caractéristiques du visage (Kazemi & Sullivan, 2014) [26]	Détection des points clés du visage (yeux, nez, bouche, etc.) à l'aide de modèles tels que dlib (68 points) ou MediaPipe Face Mesh [27].	Localisation précise des yeux même lorsque le visage n'est pas parfaitement orienté vers la caméra.	Peut être moins précise lorsque la qualité de l'image est faible ou en présence d'occlusions importantes.
Réseaux de neurones convolutifs (RNC ou CNN) [24]	Modèles d'apprentissage profond entraînés pour détecter directement les régions oculaires à partir des images.	Très robustes face aux variations d'éclairage, de pose et d'environnement.	Nécessitent davantage de ressources de calcul ainsi qu'un volume important de données d'entraînement.
Réseau convolutif multitâche en cascade (MTCNN) (Zhang et al., 2016) [25]	Architecture en cascade permettant la détection du visage et l'extraction automatique des coordonnées des yeux et d'autres points caractéristiques.	Très précis et efficace en temps réel grâce à la détection simultanée du visage et des points faciaux.	Complexité computationnelle plus élevée que celle des méthodes classiques.

2.4.3 Métriques d'évaluation des modèles

Pour évaluer les performances des modèles de vision par ordinateur et d'intelligence artificielle, plusieurs métriques d'évaluation sont couramment utilisées. Ces métriques permettent de mesurer la capacité d'un modèle à effectuer des prédictions correctes et à distinguer efficacement les différentes classes. Elles jouent un rôle essentiel dans l'analyse

des performances des modèles de classification, notamment dans les problèmes de classification binaire tels que la distinction entre les yeux ouverts et les yeux fermés [29]. Le tableau 2.3 présente les métriques d'évaluation les plus couramment utilisées pour analyser les performances des modèles de classification.

Tableau 2.3 : Métriques d'évaluation des performances des modèles.

Métrique	Formule	Rôle
Exactitude	$\frac{\text{Prédictions correctes}}{\text{Nombre total de prédictions}}$	Mesure la proportion de prédictions correctement classées parmi l'ensemble des prédictions réalisées par le modèle.
Précision	$\frac{TP}{TP + FP}$	Évalue la capacité du modèle à limiter les fausses détections en mesurant la proportion de prédictions positives correctes.
Rappel	$\frac{TP}{TP + FN}$	Mesure la capacité du modèle à détecter correctement les instances positives présentes dans les données.
F1-Score	$2 \times \frac{\text{Précision} \times \text{Rappel}}{\text{Précision} + \text{Rappel}}$	Fournit une mesure équilibrée entre la précision et le rappel, particulièrement utile lorsque les classes sont déséquilibrées.

où :

- **TP (True Positive - Vrais positifs)** : nombre de cas positifs correctement détectés.
- **TN (True Negative - Vrais négatifs)** : nombre de cas négatifs correctement détectés.
- **FP (False Positive - Faux positifs)** : nombre de cas négatifs incorrectement classés comme positifs.
- **FN (False Negative - Faux négatifs)** : nombre de cas positifs incorrectement classés comme négatifs.

Ces métriques permettent de comparer différents modèles et de sélectionner celui offrant les meilleures performances pour la détection de la somnolence du conducteur[29], [30].

2.5 Conclusion

Dans ce chapitre, nous avons présenté un état de l'art des principales approches utilisées pour la détection de la somnolence du conducteur, notamment les méthodes basées sur les signaux physiologiques, la véhicule et le comportement du conducteur. Notre travail s'oriente vers une méthode basée sur le comportement, jugée la plus efficace, en utilisant un modèle CNN.

Le chapitre suivant sera consacré à la description détaillée de la méthodologie adoptée ainsi qu'à l'architecture du système proposé.

Chapitre 3 :

Système de détection de la somnolence en conduite proposé.

Chapitre 3
Système de détection de la somnolence en conduite proposé

3.1 Introduction

Ce chapitre présente la méthodologie adoptée pour la conception et le développement du système de détection de la somnolence du conducteur. L'approche proposée repose sur un modèle d'apprentissage profond basé sur les réseaux de neurones convolutifs (CNN), permettant la classification de l'état des yeux (ouverts ou fermés) à partir d'images. L'objectif est d'intégrer ce modèle au sein d'un système intelligent capable d'assurer une détection fiable et en temps réel de la somnolence, afin de renforcer la sécurité routière.

3.2 Objectif du système proposé

L'objectif principal de ce projet est de concevoir et développer un système intelligent capable de surveiller la vigilance du conducteur afin de prévenir les accidents liés à la somnolence. Ce système s'appuie sur deux technologies majeures :

- **Détection du visage** : localiser le visage du conducteur à l'aide du détecteur SSD (Single Shot Detector) basé sur un réseau de neurones convolutionnel (CNN).
- **Correction de l'orientation de la tête** : réalisée à l'aide de MediaPipe Face Mesh afin de compenser les inclinaisons du visage et d'améliorer la précision de la détection.
- **Détection des yeux** : localiser les yeux du conducteur à l'aide des classifieurs de Haar Cascade.
- **Classification de l'état des yeux** : déterminer si les yeux sont ouverts ou fermés à l'aide du modèle CNN proposé.
- **Analyse temporelle** : utiliser une logique de score afin d'évaluer l'état de vigilance du conducteur et de détecter les signes de somnolence.

- **Déclenchement d'une alerte** : avertir le conducteur par une alarme sonore lorsque le score de somnolence dépasse un seuil prédéfini.

3.3 Apprentissage profond (Deep Learning)

L'apprentissage profond (Deep Learning) est une branche de l'apprentissage automatique reposant sur l'utilisation de réseaux de neurones artificiels comportant plusieurs couches de traitement. Grâce à son architecture hiérarchique, il est capable d'apprendre automatiquement des représentations de plus en plus complexes à partir des données brutes, sans nécessiter une étape manuelle d'extraction des caractéristiques.

Cette approche a démontré d'excellentes performances dans de nombreux domaines, notamment la vision par ordinateur, la reconnaissance vocale et le traitement du langage naturel. Parmi ses principaux avantages, on peut citer :

- La capacité à exploiter efficacement de grandes quantités de données ;
- L'extraction automatique des caractéristiques pertinentes à partir des données d'entrée ;
- Une grande robustesse face aux variations des données ;
- Des performances élevées dans les tâches de classification, de détection et de reconnaissance d'images.

Dans ce projet, nous utilisons un modèle d'apprentissage profond appelé réseau de neurones convolutionnel (CNN) pour classifier l'état des yeux.

3.4 Architecture du système proposé

Le système de détection de somnolence est composé de deux parties principales :

- La première consiste à créer et entraîner un modèle CNN capable de reconnaître l'état des yeux (ouverts, fermés) à partir d'images prétraitées.
- La deuxième partie consiste à intégrer le modèle CNN dans un système complet de détection de somnolence à partir d'une vidéo. L'utilisateur fournit une vidéo au système, qui extrait les images, détecte le visage et les yeux du conducteur, puis analyse l'état des yeux à l'aide du modèle entraîné. En cas de fermeture prolongée des yeux, un système de score est utilisé pour déclencher une alarme sonore. Cette architecture permet de tester efficacement le modèle dans des conditions proches de la conduite réelle.

3.5 Modèle CNN proposé

3.5.1 Le jeu de données utilisé pour l'entraînement

- **Présentation générale**

Dans le cadre du développement du système de détection de la somnolence du conducteur, nous avons utilisé le jeu de données MRL Eye Dataset, une base de données publique largement utilisée dans les travaux de recherche consacrés à l'analyse de l'état des yeux. Ce jeu de données constitue une référence dans le domaine de la vision par ordinateur et de la détection de la vigilance du conducteur.

Le jeu de données a été développé par le Media Research Lab (MRL), laboratoire de recherche rattaché au département d'informatique de VSB – Technical University of Ostrava en République tchèque. Il a été conçu pour soutenir les recherches liées à l'analyse visuelle du comportement humain, notamment la détection de l'état des yeux, le suivi du regard et l'étude des clignements.

Le jeu de données MRL Eye contient un grand nombre d'images d'yeux capturées dans des conditions variées, incluant différentes configurations d'éclairage, plusieurs dispositifs d'acquisition, diverses résolutions d'image ainsi que des sujets portant ou non des lunettes. Cette diversité permet de représenter des situations proches des conditions réelles d'utilisation.

Grâce à la richesse de ses annotations et à la variété des données qu'elle contient, ce jeu de données est particulièrement adaptée à l'entraînement, à la validation et à l'évaluation des modèles d'intelligence artificielle destinés à la classification de l'état des yeux (ouverts ou fermés). Dans ce travail, elle a été utilisée pour entraîner le modèle CNN proposé, chargé de reconnaître automatiquement l'état des yeux du conducteur. La figure 3.1 présente quelques exemples d'images issues du jeu de données MRL Eye.

.

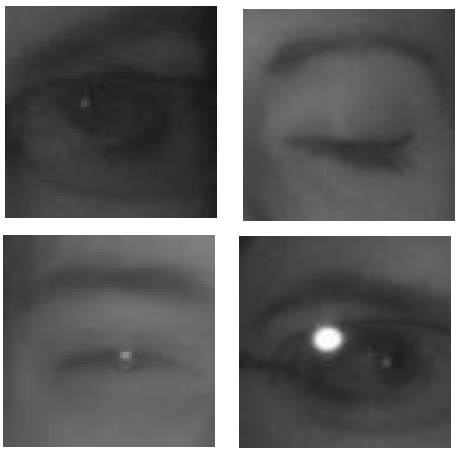


Figure 3.1 : Diversité des images du jeu de données MRL Eye.

- **Structure des annotations du jeu de données MRL Eye**

Dans le jeu de données MRL Eye, les informations descriptives associées à chaque image sont directement intégrées dans le nom du fichier. Cette méthode permet de stocker les annotations sans recourir à un fichier d'étiquetage externe, ce qui facilite l'organisation, la gestion et l'exploitation du jeu de données lors des phases de prétraitement et d'apprentissage.

Le format général des noms de fichiers est défini comme suit :

sXXXX_XXXXX_X_X_X_X_X_XX.png

Chaque segment du nom de fichier correspond à une annotation spécifique décrivant soit les caractéristiques du sujet, soit les conditions d'acquisition de l'image. Ces informations comprennent notamment l'identifiant du sujet, le numéro de l'image, le genre, la présence de lunettes, l'état de l'œil, les reflets éventuels, la qualité de l'image ainsi que le type de capteur utilisé pour la capture. Le tableau 3.1 présente la signification des différentes annotations utilisées dans le jeu de données MRL Eye. Ces annotations permettent d'analyser l'influence de plusieurs facteurs sur les performances du modèle et contribuent à la construction d'un système de classification plus robuste face aux variations des conditions réelles d'utilisation.

Tableau 3.1 : Description des annotations utilisées dans le jeu de données MRL Eye

Annotation	Description	Valeurs possibles	Remarques
Identifiant du sujet	Identifiant de la personne à laquelle appartient l'image	sXXXX	Permet d'éviter que le modèle apprenne uniquement les caractéristiques d'une seule personne.
Numéro de l'image	Numéro de l'image dans la séquence capturée	XXXXX	Permet d'identifier chaque image de manière unique.
Genre	Sexe du sujet	0 : Homme 1 : Femme	Utile pour analyser les performances du modèle selon différents profils.
Lunettes	Indique si la personne porte des lunettes	0 : Sans lunettes 1 : Avec lunettes	Les lunettes peuvent générer des reflets ou masquer partiellement les yeux.
État de l'œil	État de l'œil du sujet	0 : Œil fermé 1 : Œil ouvert	Variable principale utilisée pour l'entraînement du modèle.
Reflets	Présence de reflets sur l'image	0 : Aucun reflet 1 : Reflet faible 2 : Reflet fort	Les reflets peuvent perturber la détection et la classification des yeux.
Qualité de l'image	Qualité visuelle de l'image selon les conditions d'acquisition	0 : Mauvaise qualité 1 : Bonne qualité	Influence directement les performances du modèle de classification.
Type de capteur	Caméra utilisée pour capturer l'image	01 : RealSense SR300 02 : IDS Imaging 03 : Aptina Imaging	Permet d'identifier le dispositif d'acquisition utilisé.

La figure 3.2 présente une capture d'écran d'un exemple d'annotation des images du jeu de données MRL Eye

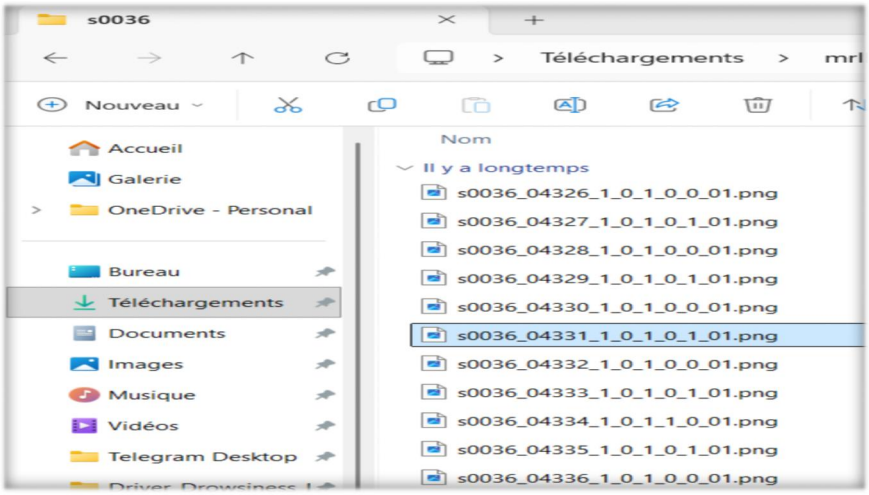


Figure 3.2 : Exemple d'annotation des images du jeu de données.

3.5.2 Collecte et préparation des données

- **Téléchargement du jeu de données**

  Le téléchargement du jeu de données a été effectué à partir de la ressource présentée précédemment, à savoir le jeu de données MRL Eye, (VSB – Technical University of Ostrava, 2018).

- **Extraction et organisation des fichiers**

  Après téléchargement et décompression de l'archive principale, des sous-archives correspondant à chaque sujet (ex : `s001`, `s002`, etc.) sont regroupées puis extraites afin d'obtenir toutes les images dans un même dossier de travail.

- **Prétraitement et préparation des images**

  Chaque image du jeu de données subit les étapes suivantes :

  - Conversion en niveaux de gris afin de simplifier la représentation de l'image et de réduire la complexité computationnelle.
  - Redimensionnement à une taille fixe de 52 × 52 pixels afin d'uniformiser les dimensions des données d'entrée.
  - Conversion de l'image en tableau NumPy pour permettre son traitement par les bibliothèques d'apprentissage automatique.
  - Normalisation des valeurs des pixels dans l'intervalle [0,1] en divisant chaque pixel par 255, ce qui facilite et accélère l'apprentissage du modèle.

- Attribution d'une étiquette (0 : œil fermé, 1 : œil ouvert) à partir du nom du fichier image.

La figure 3.3 présente un exemple du prétraitement appliqué à une image du jeu de données MRL Eye. Elle illustre la transformation de l'image originale en une image prétraitée.

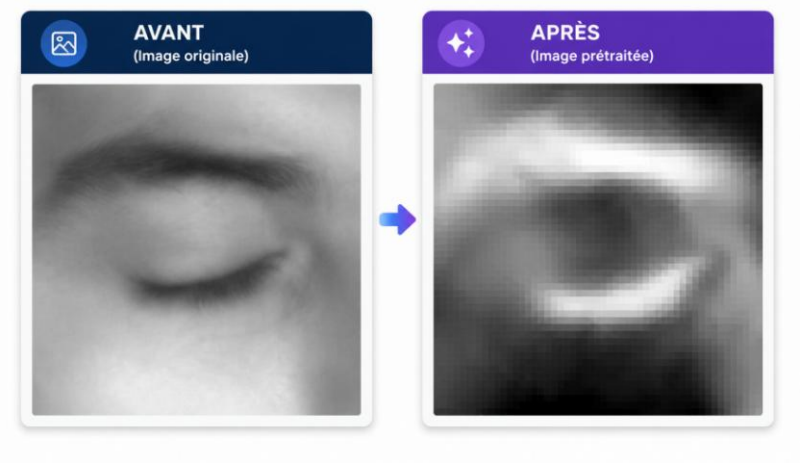


Figure 3.3 : Prétraitement d'une image du jeu de données MRL Eye.

3.5.3 Préparation des données pour l'entraînement du modèle

- **Encodage des étiquettes** : Les étiquettes associées aux images sont converties en vecteurs One-Hot Encoding à l'aide de la fonction to_categorical(). Ainsi, la classe « œil fermé » est représentée par [1,0] et la classe « œil ouvert » par [0,1]. Cette représentation est adaptée à la classification binaire réalisée par le réseau de neurones.
- **Division du jeu de données :** Le jeu de données est divisé en deux sous-ensembles à l'aide de la fonction train_test_split(). Ainsi, 67 % des images sont utilisées pour l'entraînement du modèle, tandis que les 33 % restants sont réservés à l'évaluation de ses performances. Cette séparation est réalisée avec le paramètre random_state = 42, garantissant la reproductibilité des résultats.

3.5.4 Conception du modèle CNN

- **Architecture du réseau**

La structure du modèle, conçue avec Keras (TensorFlow backend), comprend les éléments suivants :

- **Couches de convolution (**Conv2D**)** :Extraction des caractéristiques spatiales à l'aide de trois couches convolutionnelles successives contenant respectivement 32, 64 et 128 filtres de taille *(3×3)* avec une fonction d'activation ReLU.

- **Couches de pooling (**MaxPooling2D**)** : Deux opérations de sous-échantillonnage *(2×2) so*nt appliquées après les deux premières couches convolutionnelles afin de réduire progressivement la dimension spatiale des cartes de caractéristiques.
- **Couche Dropout** : Une première couche de régularisation est introduite après la troisième convolution avec un taux de 25 %, afin de limiter le surapprentissage.
- **Flatten et couches denses** : Les cartes de caractéristiques sont transformées en un vecteur unidimensionnel via une couche Flatten, puis traitées par deux couches entièrement connectées de 128 neurones chacune, utilisant la fonction d'activation ReLU.
- **Deuxième Dropout** : Une seconde couche Dropout avec un taux de 50 % est appliquée entre les deux couches denses pour renforcer la généralisation du modèle.
- **Couche de sortie (**Softmax**)** : Une couche finale composée de 2 neurones avec une activation Softmax, permettant la classification binaire (œil ouvert / œil fermé).

Le tableau 3.2 présente l'architecture détaillée du réseau de neurones convolutionnel (CNN) proposé. Il décrit les différentes couches du modèle, leurs paramètres ainsi que les dimensions des données à chaque étape du processus de classification.

Tableau 3.2 : Architecture détaillée du modèle CNN proposé.

Couche	**Type**	**Nb. filtres / unités**	**Taille filtre / Pool**	**Fonction d'activation**	**Dimension de sortie**
Entrée	Input	—	—	—	52 × 52 × 1
Conv1	Conv2D	32	3 × 3	ReLU	50 × 50 × 32
Pool1	MaxPooling2D	—	2 × 2	—	25 × 25 × 32
Conv2	Conv2D	64	3 × 3	ReLU	23 × 23 × 64
Pool2	MaxPooling2D	—	2 × 2	—	11 × 11 × 64
Conv3	Conv2D	128	3 × 3	ReLU	9 × 9 × 128
Dropout	Dropout (50 %)	—	—	—	9 × 9 × 128
Flatten	Flatten	—	—	—	10368
Dense1	Dense	128	—	ReLU	128
Dropout	Dropout (50 %)	—	—	—	128
Dense2	Dense	128	—	ReLU	128

La figure 3.4 illustre les principales couches du modèle CNN proposé pour la classification des yeux ouverts et fermés.

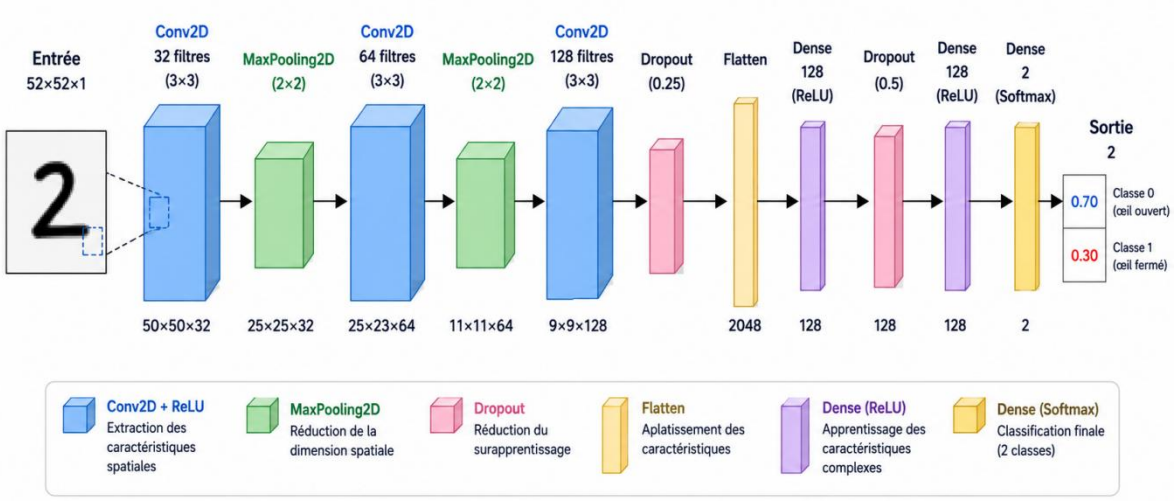


Figure 3.4 : Architecture du modèle CNN proposé.

- **Compilation**
  - Fonction de perte : `binary_crossentropy`.
  - Optimiseur : `adam`, reconnu pour sa rapidité et robustesse.
  - Métrique d'évaluation : Exactitude(Accuracy).

3.5.5 Entraînement du modèle

Une fois le modèle construit, l'entraînement est réalisé à l'aide de la fonction : model.fit(). Les paramètres principaux utilisés sont les suivants :

- Nombre d'époques : 15
- Batch size : 128
- Fonction de perte : binary_crossentropy
- Optimiseur : Adam

Pendant l'entraînement, le modèle ajuste progressivement ses poids afin de minimiser l'erreur entre les prédictions et les labels réels.

3.5.6 Suivi de l'apprentissage

Pendant l'entraînement, plusieurs métriques sont enregistrées :

- Exactitude sur les données d'entraînement ;
- Fonction de perte sur les données d'entraînement ;
- Exactitude sur les données de validation ;
- Fonction de perte sur les données de validation.

Les métriques sont représentées graphiquement pour analyser le comportement du modèle. Ces courbes permettent de détecter un surapprentissage, lorsque le modèle s'adapte trop aux données d'entraînement, ou un sous-apprentissage, lorsque le modèle n'a pas suffisamment appris. Cette visualisation aide à évaluer la qualité de l'entraînement et à identifier les améliorations possibles.

3.5.7 Évaluation des performances

Après l'entraînement, le modèle est évalué afin de mesurer sa capacité à généraliser sur le jeu de test. Cette étape est réalisée à l'aide de la fonction model.evaluate(), qui permet de calculer l'exactitude (Accuracy) du modèle sur de nouvelles données ainsi que la valeur de la fonction de perte.

Pour une analyse plus détaillée des performances du modèle, une matrice de confusion est également construite. Cette matrice compare les prédictions du modèle aux classes réelles et permet d'obtenir les valeurs de TP, TN, FP et FN, qui sont déjà présentées dans la section 2.4.3.

La matrice de confusion est représentée sous forme d'une carte de chaleur (heatmap) à l'aide de la bibliothèque Seaborn, facilitant ainsi l'interprétation des résultats.

À partir de cette matrice, plusieurs métriques complémentaires peuvent également être calculées afin d'évaluer plus finement les performances du modèle, notamment la précision , le rappel, le F1-score et le taux d'erreur de classification. Ces métriques offrent une vision plus complète de l'efficacité du modèle et permettent d'identifier ses points forts ainsi que ses limites dans la classification de l'état des yeux.

3.5.8 Sauvegarde du modèle

Le modèle entraîné est sauvegardé au format `.h5` via `model.save`, ce qui permet sa réutilisation ou intégration dans une chaîne applicative sans nécessité de ré-entraînement.

3.6 Intégration du modèle CNN dans le système proposé

Le modèle CNN entraîné pour la détection de l'état des yeux est intégré dans un système complet capable d'analyser le conducteur en temps réel et de déclencher des alertes en cas de somnolence. Le système est composé de plusieurs modules interconnectés, assurant précision et réactivité.

Dans le cadre du système proposé, la détection du visage et la localisation des yeux constituent une étape fondamentale permettant d'extraire les régions d'intérêt nécessaires à l'analyse de l'état de vigilance du conducteur. Cette étape repose sur une approche hybride combinant un modèle de détection basé sur les réseaux de neurones profonds (DNN) et des classificateurs Haar Cascade.

3.6.1 Détection du visage

La détection du visage repose sur un modèle DNN basé sur l'architecture SSD ResNet-10, chargé via le module cv2.dnn d'OpenCV à partir des fichiers deploy.prototxt et res10_300x300_ssd_iter_140000_fp16.caffemodel. Ce modèle offre une meilleure robustesse que les méthodes classiques face aux variations d'éclairage, aux rotations de la tête ainsi qu'aux occultations partielles du visage.

Avant l'étape de détection DNN, le système applique une correction automatique de l'orientation du visage grâce à MediaPipe Face Mesh. Cette étape constitue une amélioration importante du système car elle permet d'augmenter la précision de détection lorsque le conducteur incline sa tête.

MediaPipe Face Mesh détecte en temps réel jusqu'à 478 landmarks faciaux permettant de localiser précisément les principales caractéristiques du visage.

À partir de ces coordonnées, l'angle d'inclinaison de la tête est calculé à l'aide de la fonction trigonométrique np.arctan2(dy, dx). Lorsque l'angle absolu dépasse 3°, le frame est automatiquement redressé par une transformation affine (cv2.warpAffine) avant l'exécution de la détection faciale. Cette correction améliore la stabilité de localisation du visage et réduit les erreurs causées par les inclinaisons latérales de la tête.

Après correction géométrique, le frame est redimensionné (SCALE_FACTOR = 0.75) afin de diminuer la charge computationnelle et d'accélérer l'inférence du modèle DNN. Une image blob de taille 300×300 est ensuite générée avec cv2.dnn.blobFromImage avant d'être transmise au réseau neuronal.

Lorsque plusieurs visages sont détectés, le système sélectionne automatiquement le visage principal selon plusieurs critères :

- la plus grande surface du rectangle de détection .
- une position centrée horizontalement dans l'image (entre 20 % et 80 % de la largeur) .
- une localisation dans la partie supérieure de l'image.

Cette stratégie permet de privilégier le conducteur et d'ignorer les éventuels passagers ou faux positifs situés en arrière-plan.

Le rectangle englobant le visage est affiché dynamiquement avec une couleur dépendant du niveau d'alerte du système :

- vert : état normal.
- orange : niveau d'attention.
- rouge : danger de somnolence.

Afin d'optimiser les performances et de réduire la charge de calcul du processeur, la détection DNN du visage n'est pas exécutée sur chaque frame vidéo. Le système effectue la détection une frame sur deux, tandis que les informations du dernier visage détecté sont temporairement conservées et réutilisées entre les frames intermédiaires.

Cette approche agit comme un mécanisme de cache permettant d'éviter des calculs redondants, d'améliorer la fluidité du traitement en temps réel et de maintenir une précision satisfaisante.

3.6.2 Localisation des yeux

Une fois le visage détecté, la région correspondante est extraite, puis limitée à sa partie supérieure afin de se concentrer sur la zone des yeux. Cette région est convertie en niveaux de gris pour faciliter le traitement.

La détection des yeux est ensuite effectuée à l'aide de deux classificateurs Haar Cascade distincts, dédiés respectivement à l'œil gauche et à l'œil droit. Ces classificateurs reposent sur des caractéristiques visuelles simples et permettent une détection rapide en temps réel.

Pour chaque œil détecté, une sélection est effectuée afin de conserver la région la plus pertinente (basée sur la taille). Cette étape permet de réduire les faux positifs et d'améliorer la précision globale du système.

La détection des yeux ne s'exécute pas sur chaque frame, mais une frame sur deux, afin d'optimiser les performances et de réduire la charge de calcul du processeur.

Le choix des classificateurs Haar Cascade pour la détection des yeux repose sur leur rapidité d'exécution et leur faible coût computationnel. En effet, cette méthode est particulièrement adaptée aux applications en temps réel, où la vitesse de traitement est un facteur critique. De plus, étant appliquée sur une région déjà restreinte (le visage), elle permet d'obtenir des résultats satisfaisants avec une complexité réduite.

3.6.3 Prétraitement et préparation des images

Chaque œil détecté subit les étapes de prétraitement suivantes afin de répondre aux exigences du modèle CNN :

- Conversion en niveaux de gris afin de simplifier la représentation de l'image et de réduire la complexité computationnelle.
- Redimensionnement à une taille fixe de 52 × 52 pixels afin d'uniformiser les dimensions des données d'entrée.
- Conversion de l'image en tableau NumPy pour permettre son traitement par les bibliothèques d'apprentissage automatique.
- Normalisation des valeurs des pixels dans l'intervalle [0,1] en divisant chaque pixel par 255, ce qui facilite et accélère l'apprentissage du modèle.

Ce prétraitement reproduit exactement le traitement appliqué aux images lors de l'entraînement du modèle, garantissant la cohérence entre les données d'entraînement et les données réelles.

3.6.4 Prédiction de l'état des yeux

Les images prétraitées sont passées dans le modèle CNN qui prédit si l'œil est ouvert ou fermé. Ces prédictions sont ensuite utilisées pour calculer le score de somnolence du conducteur.

3.6.5 Prédiction de l'état des yeux et calcul du score de somnolence

Après le prétraitement des images oculaires, les régions correspondant aux yeux sont transmises au modèle CNN afin de déterminer l'état de vigilance du conducteur. Le modèle réalise une classification binaire permettant d'identifier si les yeux sont ouverts ou fermés.

Pour chaque œil détecté, le modèle retourne des probabilités associées aux deux classes (« Open » et « Closed »). Les résultats des deux yeux sont ensuite combinés afin d'obtenir une décision globale plus fiable. Le conducteur est considéré comme ayant les yeux fermés lorsque les deux yeux sont prédits comme fermés.

Afin d'améliorer la robustesse du système et de limiter les fausses alertes causées par les clignements naturels, les erreurs ponctuelles de détection ou les variations rapides de luminosité, une logique temporelle est intégrée au système. Cette logique repose sur un buffer (eye_buffer) stockant les résultats des dernières frames analysées.

Chaque frame traitée reçoit une valeur binaire :

- 1 : yeux fermés.
- 0 : yeux ouverts.

Le système considère l'état « yeux fermés » comme valide uniquement lorsque plusieurs prédictions consécutives indiquent une fermeture des yeux. Dans l'implémentation actuelle, les trois dernières prédictions doivent être identiques pour confirmer l'état de fermeture. Lorsque le buffer est plein, chaque nouvelle prédiction remplace la plus ancienne selon le principe FIFO (First In, First Out), ce qui permet de conserver en permanence les trois observations les plus récentes. Cette approche améliore considérablement la stabilité et la fiabilité des résultats.

Le système calcule ensuite un score de somnolence dynamique (score) évoluant progressivement dans le temps :

- lorsque les yeux restent fermés pendant plusieurs frames successives, le score est incrémenté.
- lorsque les yeux sont ouverts, le score diminue progressivement jusqu'à revenir à zéro.

Cette méthode permet d'évaluer non seulement l'état instantané des yeux, mais également la durée de fermeture oculaire, qui représente un indicateur important de somnolence chez le conducteur.

Le niveau de vigilance du conducteur est déterminé selon la valeur du score :

- Niveau normal : score faible ;
- Niveau attention : score intermédiaire ;
- Niveau danger : score supérieur au seuil critique (ALARM_THRESHOLD = 6).

Lorsque le score dépasse ce seuil critique, le système déclenche automatiquement une alarme sonore afin d'alerter le conducteur. En parallèle :

- une bordure rouge animée apparaît autour de la vidéo .
- le message « DROWSINESS DETECTED » est affiché .
- le rectangle entourant le visage devient rouge afin de signaler un état critique.

Pour améliorer les performances globales du système et réduire la charge du processeur, l'analyse complète n'est pas exécutée sur chaque frame vidéo. Le système applique une stratégie de saut de frames (FRAME_SKIP = 2) consistant à traiter une frame sur deux. Les frames intermédiaires sont simplement affichées sans exécuter les opérations de détection et de classification.

Durant les frames ignorées, le système conserve les dernières informations calculées, notamment :

- le visage détecté .
- l'état des yeux .
- le score de somnolence .
- le niveau d'alerte.

Ainsi, une frame ignorée n'est pas considérée comme une nouvelle prédiction d'yeux fermés ou ouverts. Le score de somnolence est mis à jour uniquement sur les frames effectivement analysées par le modèle CNN. Cette stratégie permet d'obtenir une exécution plus fluide tout en maintenant une précision satisfaisante pour la détection de la somnolence en temps réel.

Enfin, le système génère un graphe dynamique affichant l'évolution du score de somnolence (score_history) en temps réel. Ce graphe permet de visualiser l'état du conducteur sur plusieurs frames successives grâce à un code couleur progressif :

- vert : état normal .
- orange : vigilance .
- rouge : danger élevé.

Cette combinaison entre prédiction CNN, logique temporelle et optimisation des performances permet d'obtenir un système de détection de somnolence fiable, réactif et adapté aux contraintes du temps réel. La figure 3.5 présente une architecture simplifiée du calcul du score de somnolence. Elle montre comment les prédictions du modèle CNN sont combinées avec un buffer temporel afin de confirmer la fermeture des yeux, mettre à jour le score et déclencher une alerte lorsque le seuil critique est atteint.

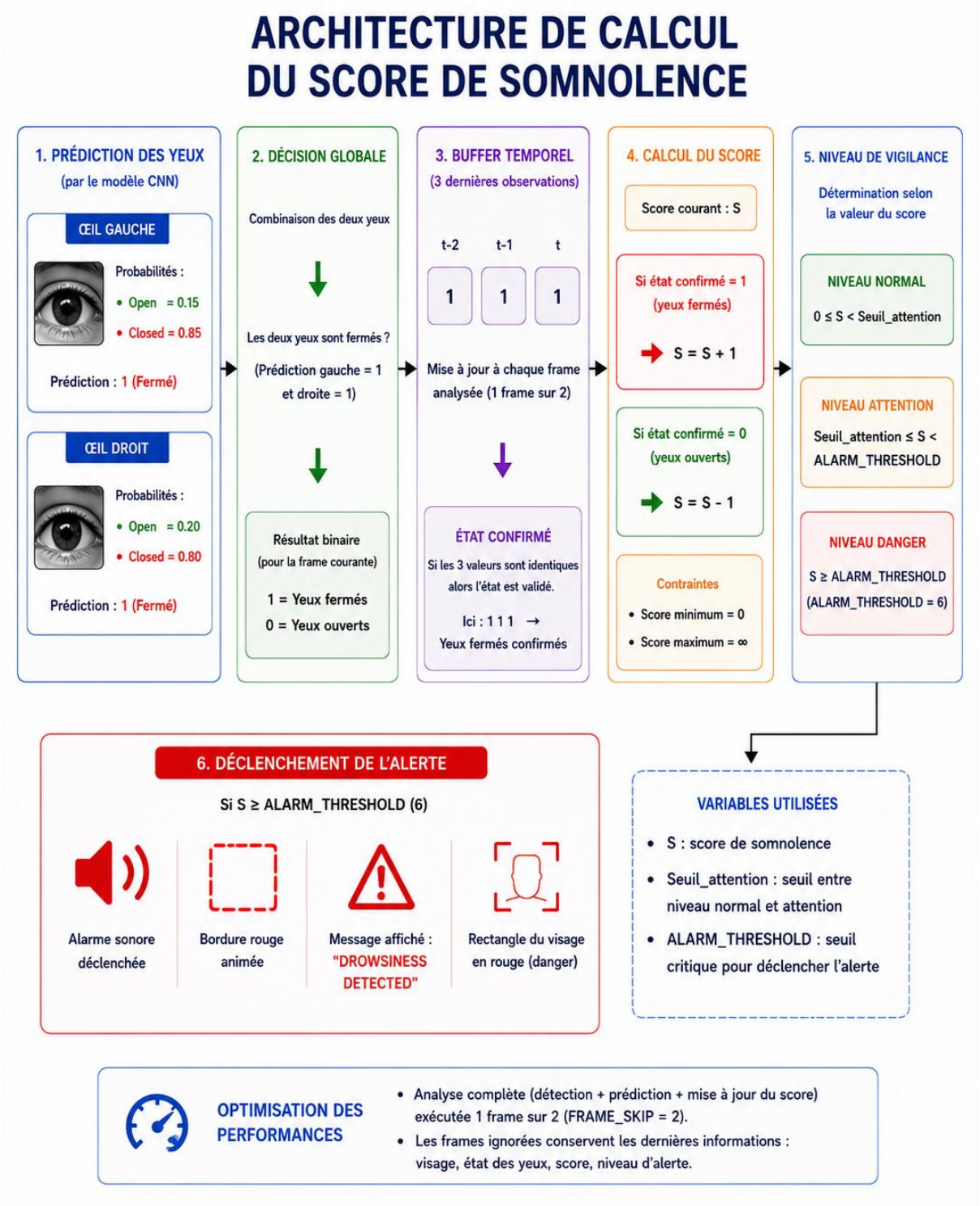


Figure 3.5 : Architecture de calcule du score de somnolence.

Le tableau 3.3 présente le mécanisme de mise à jour du score de somnolence à partir des trois dernières prédictions stockées dans un buffer temporel fonctionnant selon le principe FIFO (First In, First Out). Lorsque le buffer est plein, chaque nouvelle prédiction remplace automatiquement la plus ancienne afin de conserver en permanence les trois observations les plus récentes.

Tableau 3.3 : Principe de mise à jour du score de somnolence à l'aide d'un buffer temporel FIFO.

Frame	**Observations (3 dernières)**	**Fermeture confirmée ?**	**Score**
1	[1, 0, 0]	Non	0
2	[1, 1, 0]	Non	0
3	[1, 1, 1]	Oui	1
4	[1, 1, 1]	Oui	2
5	[0, 1, 1]	Non	1
6	[0, 0, 0]	Non	0

La figure 3.6 illustre l'architecture globale du système proposé. Elle présente les principales étapes du processus, depuis l'acquisition de la vidéo jusqu'à la détection de la somnolence et le déclenchement de l'alerte.

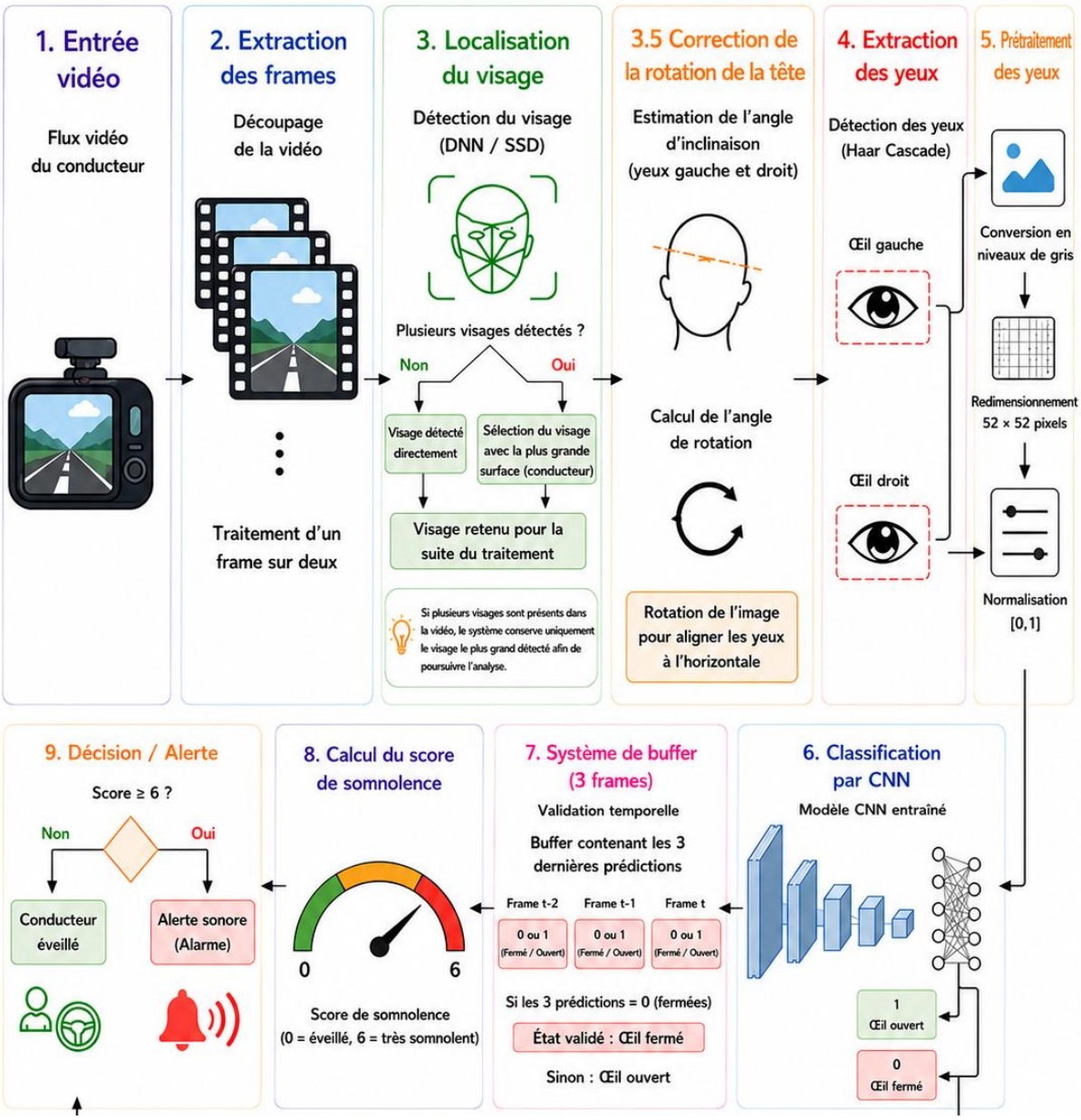


Figure 3.6 : L'architecture globale du système proposé.

3.6.6 Jeu de données utilisé pour évaluation

- **Présentation générale**

En l'absence d'un jeu de données public adapté aux exigences spécifiques de notre système, nous avons constitué notre propre jeu de données d'évaluation, nommé Driver Drowsiness Video Dataset (DDVD). Ce choix s'est imposé en raison de la difficulté à trouver des ressources existantes couvrant simultanément l'ensemble des situations que nous souhaitons évaluer. Pour pallier ce manque, nous avons combiné trois sources de données complémentaires : des séquences vidéo réelles, des vidéos générées par intelligence artificielle, ainsi que des vidéos collectées sur Internet. Cette diversité permet de tester le système dans une large variété de situations et d'assurer sa robustesse.

- **Vidéos réelles de conducteurs** : Ces enregistrements ont été capturés à partir de plusieurs conducteurs dans des conditions de conduite variées (jour, nuit, luminosité variable, différents types de routes), avec ou sans présence de passagers. Cette diversité permet de simuler des situations proches de la réalité et d'évaluer la capacité du système à détecter la somnolence dans des environnements réels.
- **Vidéos générées par intelligence artificielle** : Ces vidéos ont été créées afin de simuler différents états de somnolence chez les conducteurs. Elles incluent des variations telles que les mouvements de tête, les clignements des yeux, les expressions faciales ainsi que la présence de passagers. L'avantage principal de cette approche est de permettre la génération de scénarios contrôlés, y compris des cas rares ou difficiles à capturer dans des conditions réelles.
- **Vidéos publiées sur Internet** : Des vidéos ont également été collectées à partir de sources en ligne (plateformes publiques, bases de données ouvertes, etc.). Ces données présentent une grande diversité en termes de profils de conducteurs, de qualité d'image, d'angles de vue et de conditions d'éclairage. Elles permettent d'enrichir le jeu de données et d'améliorer la capacité de généralisation du système.

En combinant ces différentes sources de données, le système est évalué sur un ensemble riche et varié de scénarios, impliquant plusieurs conducteurs, différentes conditions environnementales et diverses configurations. Cette approche contribue à garantir une évaluation plus réaliste et à renforcer la robustesse et la fiabilité du système proposé.

- **Structure proposée des annotations du jeu de données DDVD**

Le jeu de données DDVD est organisé en deux dossiers principaux correspondant respectivement à l'absence de somnolence et à la présence de somnolence. Chaque dossier

contient plusieurs vidéos enregistrées dans différentes conditions afin d'assurer une diversité suffisante des données et d'améliorer la robustesse du système lors des phases d'entraînement, de validation et d'évaluation.

Afin de faciliter l'identification et l'organisation des vidéos, une convention de nommage spécifique a été adoptée. Cette convention permet d'associer à chaque fichier des informations descriptives relatives au sujet et aux conditions d'enregistrement. Le format général des noms de fichiers est défini comme suit : Subject_xxx_xxx_x_x_x_x.mp4.

Le tableau 3.4 présente la structure d'annotation utilisée pour le nommage des vidéos du jeu de données DDVD. Chaque élément du nom de fichier correspond à une caractéristique précise de l'enregistrement, notamment l'identifiant du sujet, le numéro de la vidéo, les conditions d'éclairage, la présence ou non de lunettes, le genre du participant ainsi que la source de la vidéo. Cette convention facilite la gestion du jeu de données, la recherche des séquences vidéo et l'analyse des performances du système selon différents scénarios de test.

Tableau 3.4 : Structure d'annotation et convention de nommage des fichiers vidéo de jeu du données DDVD.

Élément	Description	Valeurs possibles
Identifiant du sujet	Code unique représentant le participant	Subject001, Subject002, ...
Numéro de la vidéo	Indique l'ordre de l'enregistrement pour un même sujet	001, 002, 003, ...
Condition d'éclairage	Indique les conditions lumineuses lors de la capture	0 = Jour 1 = Nuit
Présence de lunettes	Indique si le sujet porte des lunettes	0 = Sans lunettes 1 = Avec lunettes
Genre	Indique le sexe du sujet	M = Masculin F = Féminin
Source de la vidéo	Indique l'origine de la vidéo	R = Réelle I = Internet T = IA (générée)

Exemple complet : Subject001_002_0_1_M_R.mp4

Ce nom indique qu'il s'agit de la 2e vidéo du sujet 001, enregistrée de jour, avec lunettes, d'un sujet masculin, et provenant d'un enregistrement réel.

3.7 Conclusion

Dans ce chapitre, nous avons présenté la méthodologie adoptée pour la conception et l'implémentation du système de détection de la somnolence du conducteur, fondé sur la vision par ordinateur et les techniques d'apprentissage profond. Les différentes phases du développement, depuis la préparation des données jusqu'à l'intégration du modèle proposé dans un environnement de traitement de vidéos simulées, ont été détaillées.

Chapitre 4 :

Évaluation expérimentale des performances du système de détection de la somnolence

Chapitre 4
Évaluation expérimentale des performances du système de détection de la somnolence

4.1 Introduction

Dans ce chapitre, nous présentons l'évaluation des performances du système de détection de la somnolence du conducteur.

A cet effet, nous allons présenter la fiabilité de notre approche en deux étapes principales, dans un premier temps, nous évaluons les performances du modèle d'apprentissage profond (Deep Learning) pour la classification de l'état des yeux, puis nous analysons le comportement du système complet par son application sur des expériences réelles sous forme de vidéos.

4.2 Environnement de travail

4.2.1 Équipements utilisés

- Un ordinateur portable présentant les caractéristiques suivantes :
  - Processeur : Intel® Core™ i5 8^{e} génération ;
  - Mémoire vive (RAM – Random Access Memory) : 8 Go ;
  - Stockage : SSD (Solid-State Drive) de 256 Go ;
  - Carte graphique (GPU – Graphics Processing Unit) : Intel® UHD Graphics.
- Unité de calcul utilisée pour l'entraînement et l'évaluation du modèle CNN : GPU (via Google Colab)
- Caméra : Appareil Redmi Note 14 Pro utilisé pour l'enregistrement de vidéos dans des conditions réelles afin d'évaluer le système.

4.2.2 Langages, bibliothèques, outils et plates-formes

- **Langages de programmation**

Le développement de notre système de détection de somnolence repose principalement sur le langage Python. Ce langage a été choisi en raison de sa simplicité, de sa lisibilité et de la richesse de son écosystème, notamment dans les domaines de l'intelligence artificielle et du traitement d'images.

- **Bibliothèques logicielles et cadres de développement (frameworks)**

Afin de réaliser les différentes fonctionnalités du système, plusieurs bibliothèques ont été utilisées :

  - **TensorFlow** : Utilisé comme framework principal de deep learning pour supporter l'exécution du modèle de réseau de neurones convolutif (CNN). Il fournit l'infrastructure nécessaire pour le calcul des opérations matricielles et l'inférence du modèle en temps réel.
  - **Keras** : Utilisé comme interface de haut niveau intégrée à TensorFlow, permettant de charger facilement le modèle CNN pré-entraîné (load_model) et d'effectuer la classification de l'état des yeux (ouverts ou fermés).
  - **OpenCV** : pour la détection du visage et des yeux ainsi que le traitement des images et des vidéos.
  - **NumPy** : Utilisée pour la manipulation des matrices et des tableaux numériques, notamment pour le traitement des images et la gestion des coordonnées des boîtes de détection.
  - **Pandas** : utilisé lors de la phase de préparation et d'analyse du jeu de données. Il permet de structurer les étiquettes, d'analyser la distribution des classes (yeux ouverts / fermés) et de générer des statistiques descriptives utiles pour vérifier l'équilibre des données.
  - **Matplotlib et Seaborn** : pour la visualisation des performances du modèle (courbes et matrice de confusion).
  - **Tkinter** : pour la création de l'interface graphique utilisateur.
  - **Pygame** : pour la gestion du système d'alerte sonore.

- **Outils et applications**

Le développement et l'expérimentation ont été réalisés à l'aide des outils suivants :

- **Google Colab** : utilisé pour l'entraînement du modèle CNN en exploitant des ressources de calcul accélérées (GPU), ce qui a permis de réduire le temps d'apprentissage et d'optimiser les performances du modèle.
- **Visual Studio Code** : utilisé comme environnement de développement pour l'écriture, l'organisation et la gestion du code source du projet, offrant une meilleure structuration et lisibilité du programme.

4.3 Évaluation du modèle CNN

L'évaluation du modèle CNN a été réalisée en utilisant le jeu de données MRL Eye, qui contient des images d'yeux ouverts et fermés dans différentes conditions. Cette diversité permet de tester la robustesse du modèle.

Le modèle, déjà entraîné sur ce jeu de données, est basé sur une architecture CNN composée de plusieurs couches (convolution, pooling et couches entièrement connectées), permettant d'extraire efficacement les caractéristiques visuelles des yeux, comme expliqué dans le chapitre précédent.

Enfin, le modèle est évalué sur des données non utilisées lors de l'entraînement afin de vérifier sa capacité de généralisation et son efficacité dans la classification de l'état des yeux.

4.3.1 Analyse descriptive du jeu de données MRL Eye

- **Analyse statistique**

Une analyse statistique du jeu de données MRL Eye a été réalisée afin d'étudier la répartition des classes utilisées pour l'entraînement du modèle. Cette étape constitue un point essentiel pour évaluer la qualité des données et leur adéquation avec le problème de classification considéré.

Les résultats obtenus sont les suivants :

- Nombre total d'yeux ouverts : **42952** ;
- Nombre total d'yeux fermés : **41946** ;
- Total des images : **84898**.

La figure 4.1 présente un diagramme en barres illustrant la répartition des images du jeu de données MRL Eye selon l'état des yeux (ouverts ou fermés)

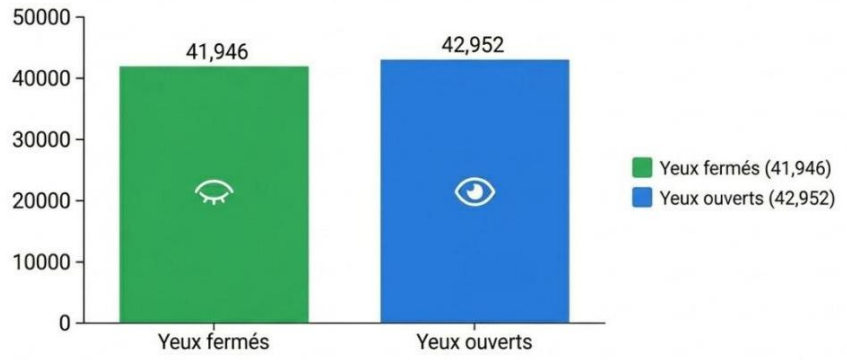


Figure 4.1 : Analyse statistique du jeu de données MRL Eye.

- **Distribution des classes**

Les pourcentages des classes sont :

- Yeux ouverts : **50.59 %** ;
- Yeux fermés : **49.40 %.**

La figure 4.2 illustre la distribution des classes du jeu de données MRL Eye. Les images sont réparties de manière quasi équilibrée entre les yeux ouverts (50,59 %) et les yeux fermés (49,40 %).

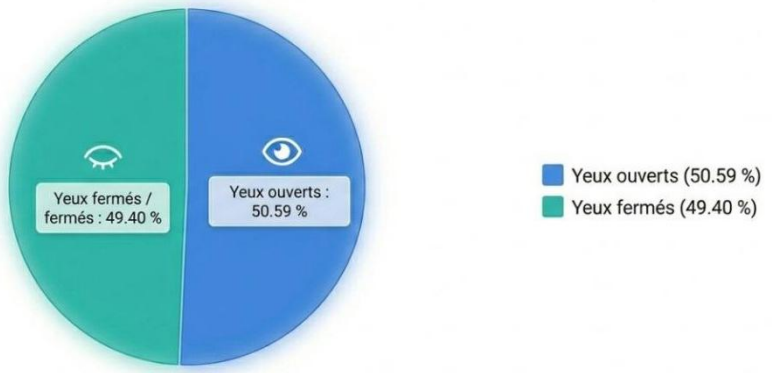


Figure 4.2 : Distribution des classes du jeu de données MRL Eye.

Cette répartition montre que le jeu de données MRL Eye est quasiment équilibré entre les deux classes. Cet équilibre est un avantage important dans le cadre de l'apprentissage supervisé, car il permet de limiter les biais de classification. En effet, dans le cas où une classe serait dominante, le modèle risquerait de privilégier cette dernière, ce qui entraînerait une dégradation des performances globales et une mauvaise généralisation.

Ainsi, la quasi-équivalence des deux classes garantit un apprentissage plus stable et plus fiable, contribuant à améliorer la robustesse du modèle CNN utilisé pour la détection de la somnolence.

4.3.2 Analyse des courbes d'apprentissage

L'analyse des courbes d'apprentissage constitue une étape essentielle pour évaluer le comportement du modèle durant la phase d'entraînement. Ces courbes représentent généralement l'évolution de la fonction de perte (loss) et de l'exactitude (Accuracy) en fonction du nombre d'itérations (epochs), aussi bien pour les données d'entraînement que pour les données de validation. La figure 4.3 illustre l'évolution de l'exactitude du modèle au cours des phases d'entraînement et de validation.

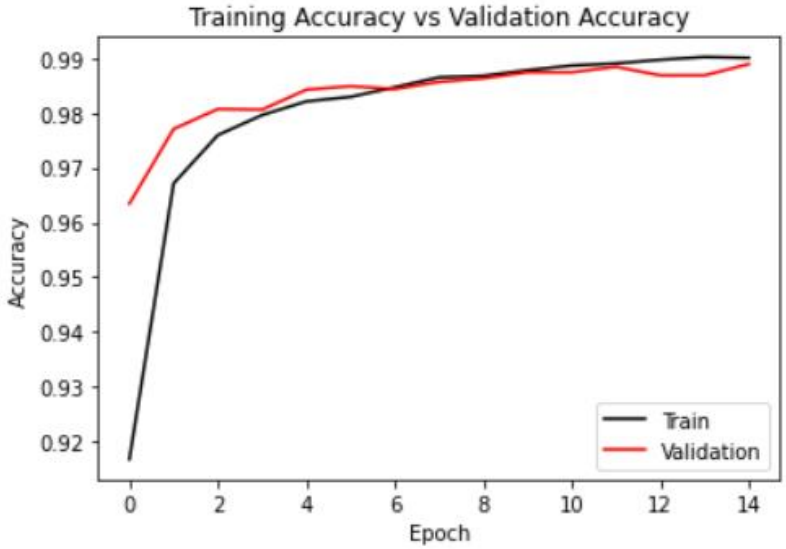


Figure 4.3 : Évolution de l'exactitude durant l'entraînement et la validation.

La courbe d'exactitude met en évidence une amélioration progressive des performances du modèle au fil des époques, jusqu'à atteindre une valeur proche de 99 % en fin d'apprentissage.

Par ailleurs, l'exactitude obtenue sur l'ensemble de validation évolue de manière similaire à celle observée sur l'ensemble d'entraînement, ce qui témoigne d'une bonne capacité de généralisation et de l'absence de surapprentissage notable.

La figure 4.4 illustre l'évolution de la fonction de perte du modèle au cours des phases d'entraînement et de validation

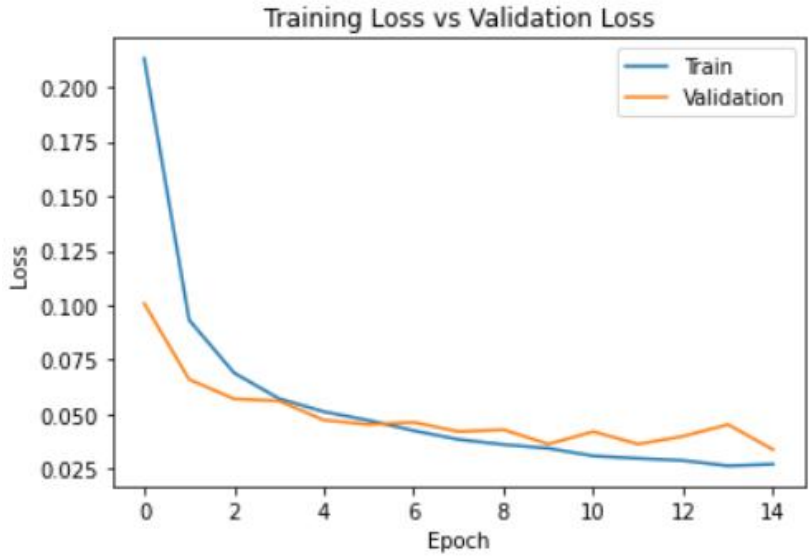


Figure 4.4 : Évolution de la perte durant l'entraînement et la validation.

En parallèle, la courbe de perte diminue de manière progressive et régulière au fil des époques, aussi bien pour l'ensemble d'entraînement que pour l'ensemble de validation. Cette évolution indique que le modèle apprend efficacement les caractéristiques discriminantes des données et converge vers une solution stable.

On observe également que l'écart entre la perte d'entraînement et la perte de validation reste faible, ce qui suggère une bonne capacité de généralisation et l'absence d'un surapprentissage significatif. Ainsi, la diminution simultanée des deux courbes confirme la robustesse du modèle ainsi que la qualité de son apprentissage.

4.3.3 Matrice de confusion

Afin d'évaluer de manière détaillée et approfondie les performances du modèle, une matrice de confusion a été utilisée. Cet outil permet d'analyser la répartition des prédictions correctes et incorrectes en distinguant les vrais positifs, vrais négatifs, faux positifs et faux négatifs, offrant ainsi une vision plus précise de la capacité du modèle à bien classifier les différentes classes. La figure 4.5 présente la matrice de confusion obtenue lors de l'évaluation du modèle sur l'ensemble de test.

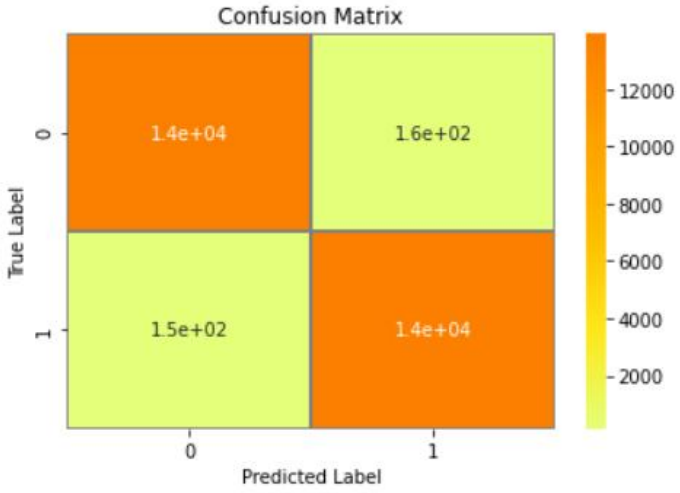


Figure 4.5 : Évaluation des performances par matrice de confusion.

Les valeurs obtenues sont :

- Vrais négatifs (TN) ≈ 14 000
- Faux positifs (FP) ≈ 160
- Faux négatifs (FN) ≈ 150
- Vrais positifs (TP) ≈ 14 000

À partir de ces valeurs, les métriques suivantes ont été calculées :

- Exactitude (Accuracy)

$$\text{Exactitude} = \frac{TP + TN}{TP + TN + FP + FN} \quad \textbf{(4.1)}$$
$$\text{Exactitude} \approx 98.89\%$$

- Précision (Precision)

$$\text{Précision} = \frac{TP}{TP + FP} \quad \textbf{(4.2)}$$
$$\text{Précision} \approx 98.87\%$$

- Rappel (Recall)

$$\text{Rappel} = \frac{TP}{TP + FN} \quad \textbf{(4.3)}$$
$$\text{Rappel} \approx 98.94\%$$

- F1-score

$$\mathrm{F1} = 2 \times \frac{\text{Précision} \times \text{Rappel}}{\text{Précision} + \text{Rappel}} \qquad (\mathbf{4.4})$$
$$\mathrm{F1} \approx 98.90\%$$

- Taux d'erreur de classification (Classification Error Rate)

$$\text{Taux d'erreur classification} = 1 - \text{Exactitude} \qquad (\mathbf{4.5})$$
$$\text{Taux d'erreur classification} \approx 1.11\%$$

Le tableau 4.1 présente une comparaison des performances du modèle proposé avec celles d'autres modèles évalués sur le même jeu de données MRL Eye.

Tableau 4.1 : Comparaison des performances du modèle proposé avec celles de travaux utilisant le jeu de données MRL Eye.

Méthode / Auteur	**Année**	**Jeu de données**	**Exactitude (%)**	**Précision (%)**	**Rappel (%)**	**F1-score (%)**
Kayadibi et al. [30]	2022	MRL Eye	97.00	—	—	—
Hassan et al. [31]	2025	MRL Eye	99.15	99.00	99.00	99.00
Notre modèle proposé	2026	MRL Eye	98.89	98.87	98.94	98.90

Le tableau 4.2 présente une comparaison des performances du modèle proposé avec celles de différents travaux de la littérature évalués sur d'autres jeux de données.

Tableau 4.2 : Comparaison des performances du modèle proposé avec celles de travaux utilisant d'autres jeux de données.

Méthode / Auteur	Année	Jeu de données	Exactitude (%)	Précision (%)	Rappel (%)	F1-score (%)
Salman et al. [33]	2021	Yawning Detection Dataset (YawDD)	—	97.00	90.30	93.50
Kayadibi et al. [32]	2022	Closed Eyes in the Wild (CEW)	97.93	97.11	98.74	97.92
Florez-Zela et al. [34]	2023	NITYMED	99.71	—	99.47	—
Notre modèle proposé	2026	MRL Eye	98.89	98.87	98.94	98.90

4.3.4 Discussion scientifique des résultats

La comparaison des performances du modèle CNN proposé avec plusieurs travaux récents montre que celui-ci atteint une exactitude de 98,89 %, une précision de 98,87 %, un rappel de 98,94 % et un F1-score de 98,90 % sur le jeu de données MRL Eye. Ces résultats démontrent la capacité du modèle à distinguer efficacement les états « œil ouvert » et « œil fermé » tout en maintenant un équilibre entre précision et capacité de généralisation.

Le modèle proposé par Salman et al. (2021), évalué sur le jeu de données YawDD, obtient une précision de 97,00 %, un rappel de 90,30 % et un F1-score de 93,50 %. La différence observée au niveau du rappel est particulièrement importante. Un rappel plus faible signifie qu'un nombre plus élevé d'exemples positifs n'est pas correctement détecté. Dans le contexte de la détection de la somnolence, cette situation est critique puisqu'elle peut conduire à ne pas identifier certains cas réels de fatigue. Notre modèle atteint un rappel de 98,94 %, ce qui traduit une meilleure capacité à détecter les états critiques et à réduire les faux négatifs.

En 2022, Kayadibi et al. obtiennent une exactitude de 97,93 %, une précision de 97,11 %, un rappel de 98,74 % et un F1-score de 97,92 % sur le jeu de données CEW. Les

performances légèrement supérieures obtenues par notre modèle peuvent être attribuées à l'efficacité de l'architecture CNN proposée ainsi qu'à la qualité du prétraitement appliqué aux images. De plus, le jeu de données MRL Eye offre une grande diversité d'images permettant au modèle d'apprendre des représentations robustes des états oculaires.

La même année, Kayadibi et al. rapportent également une exactitude de 97,00 % sur le jeu de données MRL Eye. Notre modèle présente ainsi une amélioration de 1,89 %. Cette différence peut être expliquée par la capacité de notre architecture à extraire efficacement les caractéristiques discriminantes des yeux ouverts et fermés, tout en limitant le surapprentissage grâce aux couches de régularisation utilisées.

Le modèle ResNet50V2 proposé par Florez-Zela et al. (2023), évalué sur le jeu de données NITYMED, atteint une exactitude de 99,71 % ainsi qu'un rappel de 99,47 %, soit des performances supérieures à celles de notre modèle. Cette amélioration peut être expliquée par la profondeur importante de l'architecture ResNet50V2 et par l'utilisation des connexions résiduelles qui facilitent l'apprentissage de caractéristiques complexes. Cependant, cette architecture nécessite davantage de ressources de calcul, de mémoire et de temps d'entraînement. Malgré un léger écart de performance, notre modèle reste plus léger et mieux adapté à une intégration dans un système embarqué fonctionnant en temps réel.

Le modèle basé sur les transformeurs proposé par Hassan et al. (2025) représente l'une des meilleures performances du tableau avec une exactitude de 99,15 %, une précision de 99,00 %, un rappel de 99,00 % et un F1-score de 99,00 %. Cette légère supériorité peut être attribuée aux mécanismes d'attention utilisés par les architectures Transformer, capables de capturer des relations spatiales complexes dans les images. Toutefois, l'écart avec notre modèle demeure très faible, seulement 0,26 % en exactitude. En contrepartie, ces architectures présentent une complexité computationnelle élevée et nécessitent des ressources matérielles importantes, ce qui peut limiter leur utilisation dans des systèmes embarqués à faible capacité.

Enfin, notre modèle CNN proposé atteint une exactitude de 98,89 %, une précision de 98,87 %, un rappel de 98,94 % et un F1-score de 98,90 %. Les résultats obtenus montrent qu'il offre un excellent compromis entre performance de classification et efficacité computationnelle. Le rappel élevé constitue un avantage particulièrement important dans le domaine de la sécurité routière, car il permet de réduire le nombre de faux négatifs et d'améliorer la détection des situations réelles de somnolence. De plus, sa structure

relativement simple facilite son déploiement dans des applications temps réel tout en maintenant un niveau de performance proche de celui des architectures plus complexes.

Il convient également de souligner que les travaux comparés utilisent différents jeux de données tels que YawDD, CEW, NITYMED et MRL Eye. Les différences liées aux conditions d'éclairage, aux angles de capture, à la qualité des images et à la diversité des sujets peuvent influencer directement les performances obtenues. Ainsi, cette comparaison permet principalement de situer les performances globales du modèle proposé par rapport aux approches récentes de la littérature.

4.3.5 Analyse critique des performances

L'analyse des résultats expérimentaux montre que le modèle CNN proposé atteint des performances élevées et stables sur l'ensemble des métriques d'évaluation. Les valeurs obtenues en termes d'exactitude, de précision, de rappel et de F1-score confirment l'efficacité du modèle dans la classification des états des yeux ouverts et fermés, ainsi que sa capacité à fournir des prédictions fiables. La proximité entre ces différentes métriques indique également un bon équilibre entre la capacité de détection des cas positifs et la limitation des erreurs de classification.

Cependant, malgré ces résultats satisfaisants, une analyse plus approfondie permet de mettre en évidence plusieurs points importants. Tout d'abord, la comparaison avec les travaux de l'état de l'art montre que certains modèles plus complexes, notamment les architectures de type ResNet50V2 et Transformer, obtiennent des performances légèrement supérieures. Néanmoins, ces améliorations restent marginales et ne dépassent pas un faible pourcentage.

En revanche, ces modèles plus avancés présentent une complexité beaucoup plus élevée, ce qui constitue une limitation importante dans le contexte des applications embarquées. En effet, leur utilisation implique :

- une forte demande en puissance de calcul ;
- une consommation mémoire plus élevée ;
- un temps d'exécution plus long ;
- une dépendance à des ressources matérielles avancées (GPU).

À l'inverse, le modèle proposé se distingue par sa légèreté et sa simplicité architecturale, ce qui lui permet d'être facilement déployé dans des systèmes temps réel de

surveillance du conducteur. Ce point est essentiel, car dans ce type d'application, la rapidité de détection est aussi importante que la performance globale.

Par ailleurs, il est important de souligner que les performances obtenues dépendent fortement des conditions expérimentales et du jeu de données utilisé. Les différences entre les jeux de données (qualité des images, conditions d'éclairage, diversité des sujets et angles de capture) peuvent influencer directement les résultats. Ainsi, les performances rapportées doivent être interprétées dans un cadre comparatif et non absolu.

Enfin, le principal avantage du modèle proposé réside dans son compromis entre performance et efficacité computationnelle. Contrairement aux modèles très profonds, il offre une solution équilibrée, capable d'assurer une détection précise de l'état des yeux tout en restant adapté à une implémentation en temps réel sur des dispositifs à ressources limitées.

4.3.6 Interprétation dans le contexte applicatif

Dans un système embarqué de détection de somnolence, les résultats obtenus montrent que le modèle CNN proposé est adapté à une utilisation en temps réel grâce à sa rapidité d'exécution et sa simplicité d'architecture.

Le modèle est suffisamment fiable pour être intégré dans un dispositif de surveillance du conducteur, car il permet de détecter efficacement les états des yeux et de générer des alertes en cas de risque de somnolence. Cela contribue directement à la prévention des accidents.

De plus, sa robustesse face aux variations des conditions réelles ainsi que sa faible complexité facilitent son intégration dans des systèmes embarqués sans nécessiter de ressources matérielles très élevées.

Ainsi, le modèle proposé représente une solution pratique et efficace pour les applications de sécurité routière en temps réel.

4.4 Évaluation du système proposé

Après l'évaluation du modèle CNN sur le jeu de donnée MRL Eye, une évaluation du système complet est réalisée en conditions réelles afin de valider son efficacité dans un environnement proche de l'utilisation réelle.

Dans le cadre de cette évaluation, nous avons également constitué un jeu de données personnalisée DDVD composée de vidéos enregistrées en conditions réelles ,de vidéos

générées et de vidéos téléchargées depuis Internet. Ces vidéos simulent différents scénarios de conduite, incluant l'état normal du conducteur, le clignement naturel des yeux ainsi que des situations de somnolence caractérisées par une fermeture prolongée des yeux. Ce jeu de données a été utilisée afin de tester la robustesse et la performance du système proposé dans des conditions variées et réalistes.

Le système est testé sur ces vidéos en temps réel. Le pipeline de traitement comprend plusieurs étapes successives : la détection du visage, la localisation des yeux, la classification de leur état à l'aide du modèle CNN entraîné, puis l'application d'une logique temporelle permettant le calcul du score de somnolence et la prise de décision concernant le déclenchement de l'alerte.

4.4.1 Analyse descriptive du jeu de données DDVD

Le jeu de données DDVD, constitué pour le test du système de détection de somnolence proposé, comprend 46 vidéos réparties en trois catégories selon leur mode d'acquisition : des vidéos générées par intelligence artificielle, des vidéos enregistrées par des participants à l'aide de smartphones, ainsi que des vidéos collectées à partir de sources ouvertes disponibles sur Internet. Ces vidéos couvrent une variété de scénarios, de conditions d'éclairage, de genres et de profils de conducteurs afin d'évaluer la robustesse du système proposé dans des conditions proches de la réalité. Le tableau 4.3 présente les caractéristiques générales du jeu de données DDVD, incluant la répartition des sujets, les sources des vidéos, le nombre d'enregistrements et leur durée moyenne.

Tableau 4.3 : Vue globale du jeu de données DDVD.

Type de vidéo	N° de sujet	Genre		Lunettes		N° de vidéo	Durée moy(s)	Source
		Homme	Femme	Avec	Sans			
Vidéos générées par IA	7	4	3	1	6	7	4 - 8 s	IA générative
Vidéos enregistrées	6	4	2	mix		28	4 - 24 s	Smartphone
Vidéos téléchargées	5	2	3	1	4	11	9 - 21 s	internet
Total général	18	10	8	—	—	46		

La figure 4.6 illustre les principales caractéristiques du jeu de données DDVD utilisée dans cette étude. Elle présente la répartition des vidéos selon leur source, la distribution des sujets par type de vidéo et par genre, la répartition du port de lunettes, ainsi que la durée moyenne des vidéos. Elle fournit également un aperçu global du jeu de données à travers le nombre total de vidéos et de sujets.

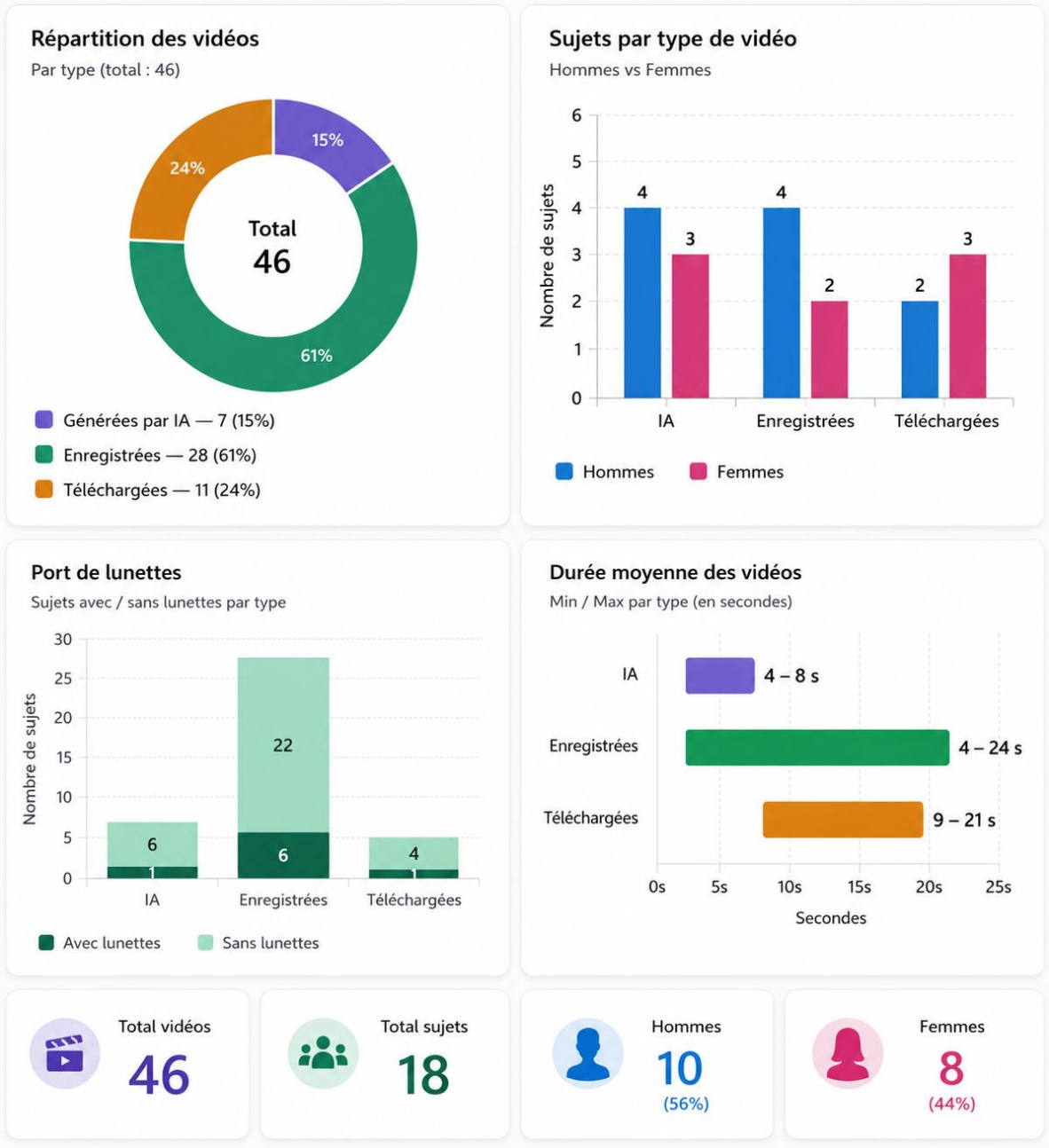


Figure 4.6 : Distribution et caractéristiques du jeu de données DDVD.

Les vidéos enregistrées représentent 61 % du jeu de données DDVD, ce qui garantit une forte représentation des conditions réelles. Les vidéos générées par intelligence artificielle apportent une diversité contrôlée des scénarios, tandis que les vidéos téléchargées introduisent des conditions d'acquisition variées et plus hétérogènes.

Cette diversité de sources permet d'assurer un ensemble de données riche et équilibré, contribuant ainsi à une meilleure robustesse du modèle. Elle garantit également que le

système est testé dans différents contextes, ce qui améliore sa capacité de généralisation et sa fiabilité lors de son utilisation en conditions réelles.

Le jeu de données DDVD présente une répartition entre les vidéos des hommes et des femmes. Cette diversité de genre permet de mieux représenter les variations naturelles des comportements et des expressions faciales selon les individus.

Il est important de souligner que l'évaluation du système proposé sur les deux genres est essentielle afin de garantir des performances stables et fiables, indépendamment du sexe du sujet. En effet, tester le modèle sur des profils masculins et féminins permet de vérifier sa capacité à fonctionner efficacement sur différents utilisateurs et de réduire les biais éventuels liés à un seul groupe.

4.4.2 Répartition des vidéos de test du jeu de données DDVD

De manière globale, les vidéos de notre jeu de données DDVD sont réparties en deux grandes catégories : présence de somnolence et absence de somnolence. Cette structuration permet d'évaluer la capacité du système à distinguer efficacement entre un état de vigilance normale et un état de somnolence.

Chaque catégorie contient plusieurs scénarios regroupant différents sujets. Chaque vidéo constitue un scénario de test indépendant. Ainsi, chaque scénario correspond à une vidéo unique associée à un sujet spécifique. Un même sujet peut donc apparaître dans plusieurs scénarios.

Ces scénarios prennent en compte divers facteurs tels que le moment de la journée (jour ou nuit), le port de lunettes ou non, ainsi que le genre du sujet (homme ou femme). Cette diversité permet de tester le modèle dans des environnements variés, contribuant ainsi à une évaluation plus complète, robuste et représentative des conditions réelles de conduite.

4.4.3 Méthodologie d'évaluation

L'évaluation du système proposé est réalisée dans des conditions variées à partir des vidéos constituant le jeu de données DDVD. Chaque vidéo est traitée en temps réel par notre système afin d'évaluer ses performances dans la détection de la somnolence du conducteur.

Le processus commence par la détection du visage, suivie de l'extraction de la région des yeux. Ces régions sont ensuite classifiées par le modèle CNN entraîné.

Une logique temporelle, décrite en détail dans la section 3.6.5, est appliquée pour détecter un état de somnolence en fonction de la durée de fermeture des yeux. Une alerte est générée lorsque cette durée dépasse un seuil prédéfini.

4.4.4 Résultats expérimentaux

Les résultats expérimentaux obtenus permettent d'évaluer les performances du système de détection de somnolence dans différentes conditions de test. L'évaluation est réalisée à partir du jeu de données DDVD, composée de vidéos variées couvrant plusieurs scénarios. Les Figures 4.7 à 4.10 présentent différents exemples de simulation réalisés sur des vidéos de natures diverses. Les figures 4.7 et 4.8 illustrent des vidéos générées par intelligence artificielle, la figure 4.9 présente une vidéo réelle enregistrée, tandis que la figure 4.10 correspond à une vidéo téléchargée à partir d'une source ouverte. Ces exemples permettent d'évaluer le comportement du système dans différentes conditions d'utilisation.

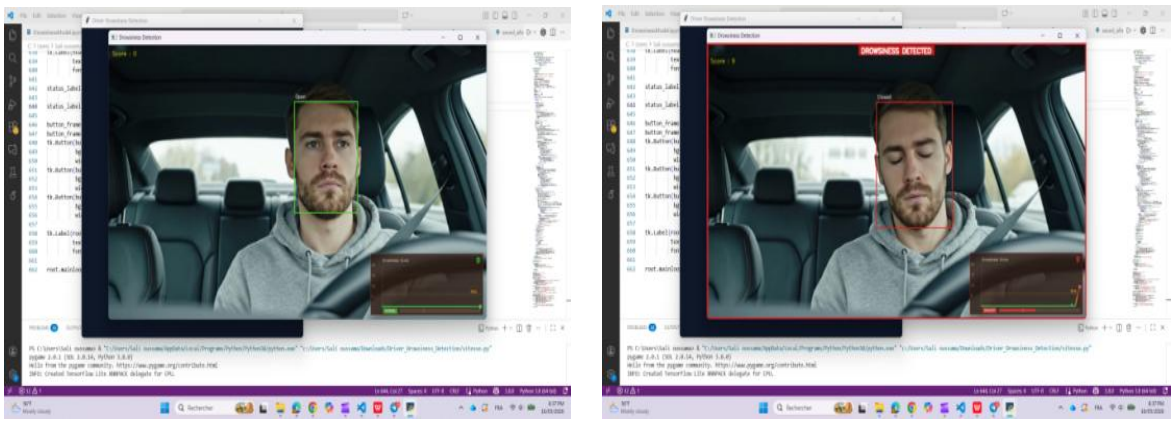


Figure 4.7 : Premier exemple de simulation sur une vidéo générée par intelligence artificielle.

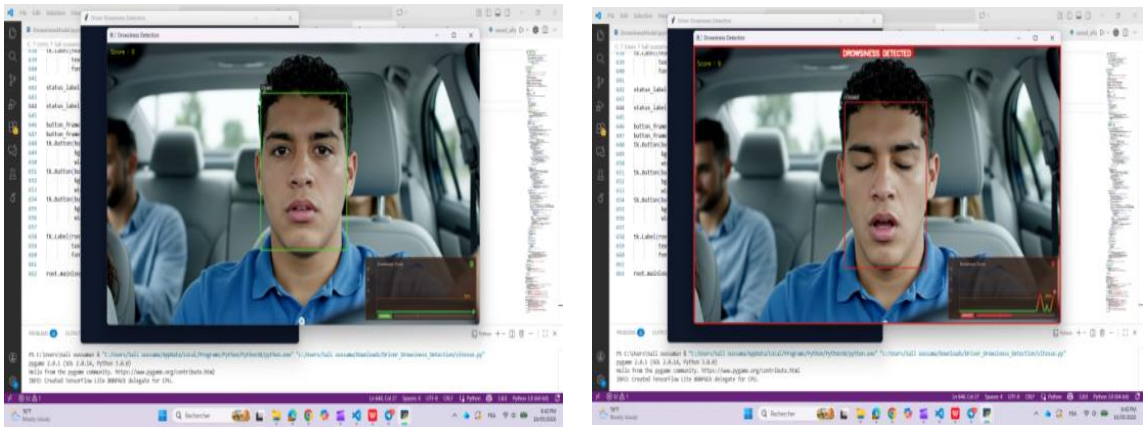


Figure 4.8 : Deuxième exemple de simulation sur une vidéo générée par intelligence artificielle.

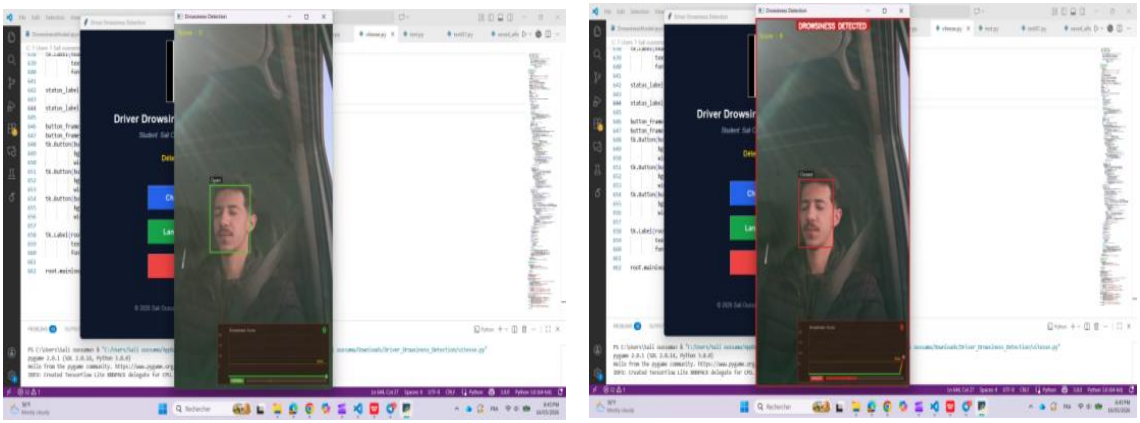


Figure 4.9 : Exemple de simulation sur une vidéo réelle enregistrée.

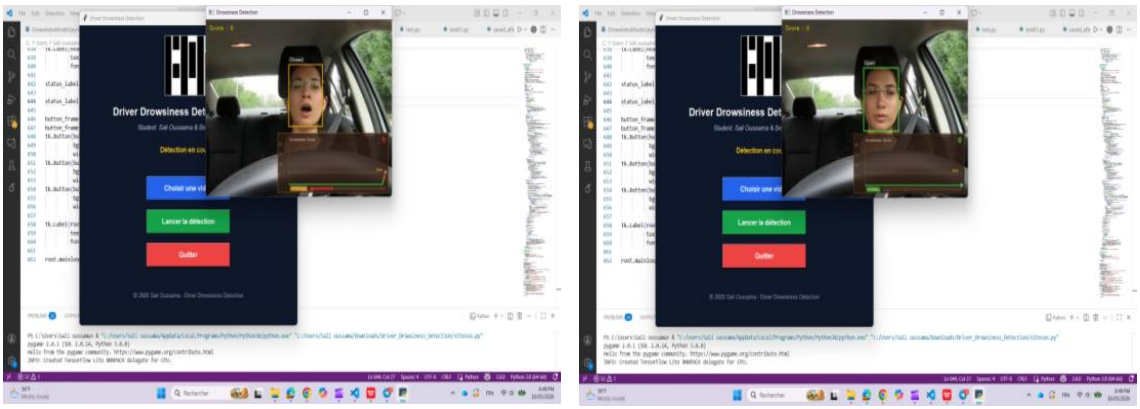


Figure 4.10 : Exemple de simulation sur une vidéo téléchargée.

Le tableau 4.4 présente les résultats obtenus lors de l'évaluation du système proposé sur les différents scénarios vidéo du jeu de données DDVD. Afin de faciliter la lecture, les vidéos consécutives appartenant à un même sujet et ayant produit le même résultat de classification ont été regroupées sous forme d'intervalles. Ainsi, le résultat indiqué s'applique à l'ensemble des vidéos comprises dans l'intervalle correspondant. Par exemple, l'entrée « Subject 1 [1-6] » signifie que les vidéos 1 à 6 du sujet 1 ont toutes produit le même résultat de classification. La colonne « Scénario (Vidéo) » identifie le sujet testé ainsi que les vidéos concernées, la colonne « Catégorie de vidéo » indique la présence ou l'absence de somnolence dans les vidéos de référence, tandis que la colonne « Résultats » présente le résultat de classification obtenu par le système.

Tableau 4.4 : Résultats de test sur le jeu de données DDVD.

Scénario (Vidéo)	Catégorie de vedio	Résultats
Subject 1 [1-6]	Présence de somnolence	TP
Subject 1 [7-9]	Absence de somnolence	TN
Subject 2 [1-4]	Présence de somnolence	TP
Subject 2 [5-6]	Absence de somnolence	TN
Subject 3 [1]	Présence de somnolence	TP
Subject 4 [1]	Absence de somnolence	FN
		TP
Subject 5 [1-3]	Présence de somnolence	TP
Subject 5 [2-5]	Absence de somnolence	TN
Subject 6 [1]	Présence de somnolence	TP
Subject 7 [1]	Présence de somnolence	TP
Subject 8 [1]	Présence de somnolence	TP
Subject 9 [1]	Présence de somnolence	TP
Subject 10 [1]	Présence de somnolence	TP
Subject 10 [2]	Absence de somnolence	TN
Subject 11 [1-3]	Absence de somnolence	TN
Subject 12 [1-2]	Présence de somnolence	TP
Subject 12 [3-4]	Absence de somnolence	TN
Subject 13 [1]	Présence de somnolence	FN
Subject 13 [2]	Absence de somnolence	TN
Subject 14 [1-2]	Présence de somnolence	TP
Subject 15 [1-3]	Présence de somnolence	TP
Subject 16 [1]	Présence de somnolence	TP
Subject 17 [1]	Présence de somnolence	TP
Subject 18 [1-2]	Présence de somnolence	TP

Le tableau 4.4 montre que le système proposé a correctement identifié la majorité des scénarios de test. Les vidéos contenant des signes de somnolence ont principalement été

classées comme des vrais positifs (TP), tandis que les vidéos ne présentant pas de somnolence ont été majoritairement reconnues comme des vrais négatifs (TN). Toutefois, quelques erreurs de classification ont été observées sous la forme de faux négatifs (FN), correspondant à des situations où la somnolence était présente mais n'a pas été détectée par le système. À partir de ces résultats, les valeurs de TP, TN, FP et FN ont été déterminées afin de calculer les différentes métriques d'évaluation du système. Le Tableau 4.5 présente les résultats globaux obtenus ainsi que les performances associées.

Tableau 4.5 : Résultats globaux du système proposé.

Métrique	**Valeur**
TP (Vrais positifs)	36
TN (Vrais négatifs)	15
FP (Faux positifs)	0
FN (Faux négatifs)	3
Exactitude	94,44 %
Précision	100 %
Rappel	92,31 %
F1-score	96,00 %

4.4.4 Analyse des performances du système proposé

Les résultats expérimentaux obtenus montrent que le système proposé présente des performances élevées pour la détection automatique de la somnolence du conducteur en temps réel. Les performances globales observées sont cohérentes avec celles obtenues lors de l'évaluation du modèle CNN.

Le système affiche un taux de détection de la somnolence particulièrement satisfaisant, ce qui confirme sa capacité à reconnaître efficacement les signes de somnolence dans différentes conditions d'utilisation. Les métriques obtenues, notamment une exactitude de 94,44 %, une précision de 100 % et un F1-score de 96 %, démontrent la robustesse et la fiabilité de l'approche proposée.

L'intégration du modèle d'apprentissage profond dans un système de vision par ordinateur fonctionnant en temps réel s'est révélée stable et efficace. Le système est capable d'analyser automatiquement les vidéos et de détecter les épisodes de somnolence, ce qui le rend adapté aux applications intelligentes d'assistance à la conduite.

Le nombre de fausses alertes (FP) reste très faible durant les expérimentations réalisées. Cette caractéristique est particulièrement importante dans les systèmes embarqués, car elle permet d'éviter les déclenchements inutiles des alarmes et d'améliorer le confort ainsi que l'expérience utilisateur.

Les résultats obtenus dans différentes conditions de test indiquent que le système présente de bonnes performances pour une application à d'autres vidéos. Ainsi, il pourrait être capable de généraliser ses performances à de nouvelles données, sous réserve d'une validation complémentaire sur un ensemble plus large et plus diversifié de vidéos.

L'approche basée sur la vision par ordinateur combinée à l'apprentissage profond apparaît ainsi comme une solution pertinente et efficace pour les applications de détection de la somnolence du conducteur en temps réel. Elle offre un bon compromis entre précision, rapidité d'exécution et simplicité d'intégration dans des systèmes intelligents de sécurité routière.

4.4.5 Discussion

Les résultats expérimentaux obtenus confirment que le système proposé constitue une solution fiable et efficace pour la détection de la somnolence du conducteur en temps réel. Grâce à l'utilisation des techniques de vision par ordinateur et de l'apprentissage profond, le système proposé est capable d'identifier automatiquement les signes de somnolence à partir de l'analyse de l'état des yeux, qu'ils soient ouverts ou fermés.

L'approche développée présente plusieurs avantages par rapport aux systèmes traditionnels basés sur des capteurs physiologiques. Tout d'abord, elle est non intrusive puisqu'elle ne nécessite aucun capteur porté par le conducteur. Le système utilise uniquement une caméra pour analyser les images vidéo, ce qui améliore le confort d'utilisation et facilite son intégration dans un véhicule.

De plus, la solution proposée reste relativement simple à mettre en œuvre et présente un coût matériel réduit comparativement aux systèmes nécessitant des équipements spécialisés. Le modèle CNN permet d'extraire automatiquement les caractéristiques visuelles pertinentes des régions oculaires et de classifier avec précision l'état des yeux en distinguant les yeux ouverts des yeux fermés. Cette classification constitue une étape essentielle dans l'évaluation de l'état de vigilance du conducteur. Par ailleurs, l'approche proposée offre de meilleures performances que les méthodes classiques reposant sur des règles fixes.

Les tests réalisés ont montré que le système proposé possède une bonne capacité de généralisation face aux variations d'éclairage, aux mouvements de la tête et aux différences morphologiques entre les sujets. Toutefois, l'évaluation expérimentale n'a pas inclus des scénarios contenant des conducteurs portant des lunettes noires ou fortement opaques. Par conséquent, les performances du système dans ce type de situation ne peuvent pas être garanties.

En effet, la présence de lunettes noires empêche la détection précise des yeux et limite fortement l'extraction des caractéristiques visuelles nécessaires à l'identification des signes de somnolence. Cette limitation constitue l'un des principaux défis des systèmes basés uniquement sur l'analyse visuelle du visage.

Cependant, pour améliorer davantage le système, plusieurs pistes peuvent être envisagées :

- Intégration de techniques de suivi du visage (Face Tracking) afin d'assurer une détection plus stable en présence de mouvements de la tête ;
- Amélioration des performances du système dans des conditions d'éclairage difficiles, notamment en faible luminosité ou durant la conduite nocturne ;
- Intégration de paramètres comportementaux complémentaires, tels que les bâillements ou les mouvements de la tête, pour une détection plus fiable de la somnolence.

4.5 Interface graphique utilisateur

L'interface graphique développée avec Tkinter permet à l'utilisateur de charger facilement une vidéo et de lancer l'analyse. Un message d'état informe en temps réel du déroulement de la détection. Les résultats sont visualisés dans une fenêtre OpenCV, affichant le visage détecté, les yeux et le score de somnolence, ce qui facilite l'interprétation des résultats. La figure 4.11 présente l'interface utilisateur graphique du système proposé. Elle permet à l'utilisateur de sélectionner une vidéo, de lancer le processus de détection de la somnolence et de suivre l'état d'exécution du système. L'interface affiche également l'identité du projet ainsi que les informations relatives aux développeurs, offrant ainsi une utilisation simple et intuitive.

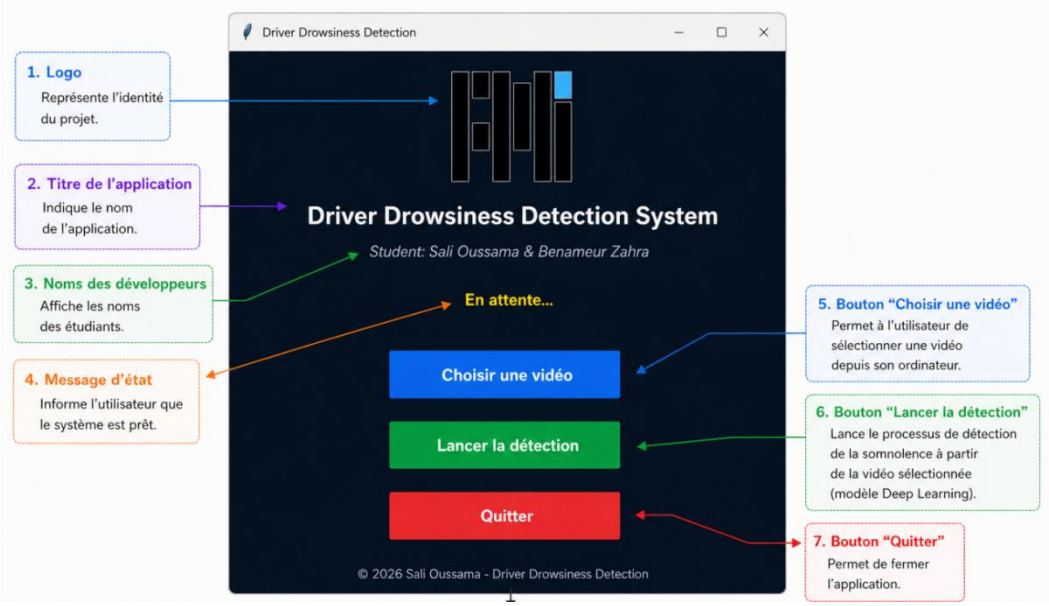


Figure 4.11 : Interface utilisateur du système de détection de la somnolence du conducteur.

La figure 4.12 présente l'interface du système après la sélection d'une vidéo par l'utilisateur

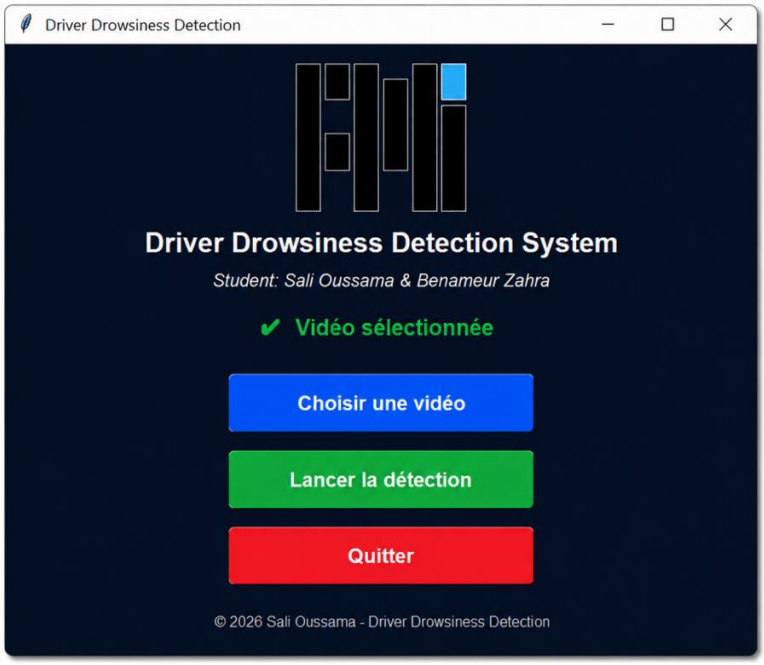


Figure 4.12 : Interface de vidéo séléctionnée.

La figure 4.13 présente l'interface du système durant le processus de détection. Elle indique que l'analyse de la vidéo est en cours d'exécution.

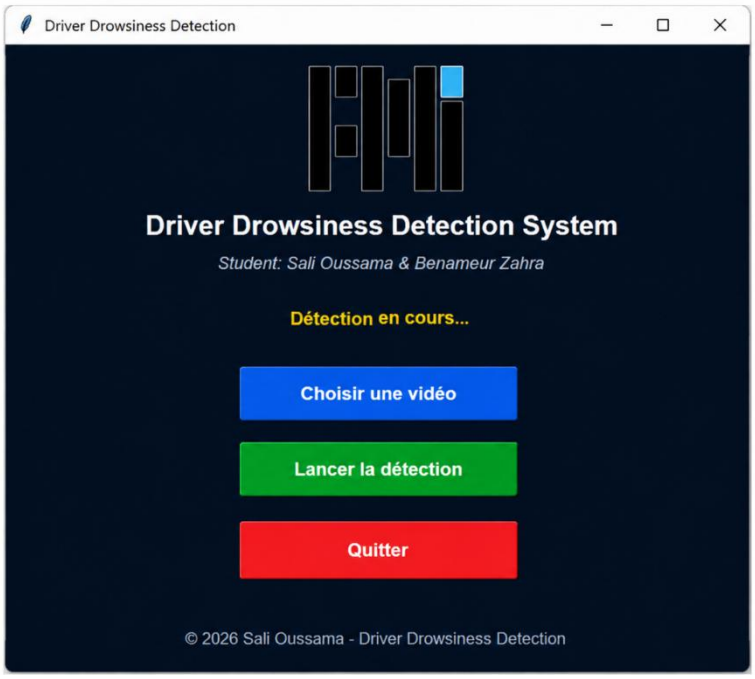


Figure 4.13 : : Interface détection en cours.

La figure 4.14 présente l'interface du système après la fin du processus de détection. Elle indique que l'analyse de la vidéo a été réalisée avec succès et informe l'utilisateur de l'achèvement du traitement.

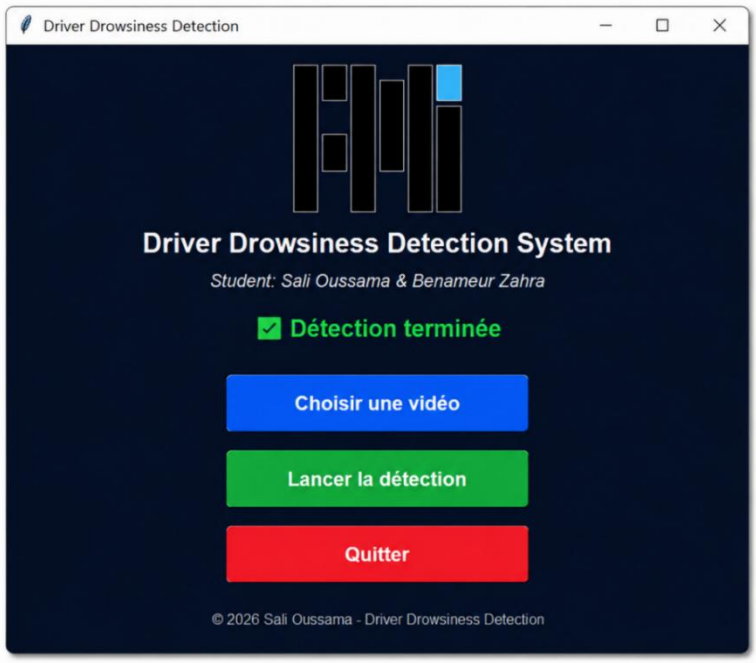


Figure 4.14 : Interface détection terminée.

4.5 Conclusion

Dans ce chapitre, nous avons présenté l'évaluation du modèle CNN ainsi que celle du système complet de détection de la somnolence proposé. Les résultats obtenus montrent une très bonne performance globale, avec une précision élevée et une bonne stabilité en conditions réelles.

Le système proposé constitue donc une solution efficace et prometteuse pour la détection de la somnolence du conducteur, tout en restant léger et adapté à une utilisation en temps réel.

Cependant, certaines limitations subsistent, ouvrant la voie à des améliorations futures.

Conclusion générale

Au terme de ce travail, nous avons abordé la problématique de la somnolence du conducteur, qui constitue aujourd'hui l'une des principales causes d'accidents de la route à travers le monde. Cette étude nous a permis de mettre en évidence les effets significatifs de la somnolence sur la vigilance, le temps de réaction et la capacité de prise de décision du conducteur. Ces conséquences peuvent compromettre la sécurité routière et soulignent la nécessité de développer des solutions de détection fiables et efficaces en temps réel.

Dans un premier temps, nous avons présenté un état de l'art des différentes approches utilisées pour la détection de la somnolence, notamment les méthodes basées sur les signaux physiologiques, le comportement du conducteur et le comportement du véhicule. Cette analyse a montré que, malgré leur efficacité, certaines de ces approches demeurent invasives, coûteuses ou difficiles à mettre en œuvre dans des conditions réelles. En revanche, les méthodes basées sur le comportement du conducteur, en particulier celles exploitant la vision par ordinateur et l'intelligence artificielle, constituent des solutions prometteuses grâce à leur caractère non intrusif, leur efficacité et leur capacité à fonctionner en temps réel.

Dans ce contexte, nous avons proposé un système intelligent de détection de la somnolence basé sur l'analyse de l'état des yeux à l'aide de techniques de vision par ordinateur et d'un modèle d'apprentissage profond (CNN). Le système développé repose sur plusieurs étapes, allant de la détection du visage, à la localisation des yeux, jusqu'à la classification de leur état (ouverts ou fermés), avec l'intégration d'un mécanisme de décision basé sur un score de somnolence permettant de déclencher une alarme en cas de fermeture prolongée des yeux.

Les résultats obtenus lors de l'évaluation du modèle CNN ont mis en évidence des performances élevées, avec une précision proche de 99 %, confirmant son aptitude à reconnaître efficacement les différents états des yeux. En outre, l'évaluation du système complet sur un ensemble de vidéos réelles, générées artificiellement et téléchargées depuis des sources publiques a permis de confirmer son efficacité dans des conditions d'utilisation

proches de la réalité, tout en maintenant une bonne performance globale et une réactivité satisfaisante.

Cependant, certaines limitations ont été observées lors de l'utilisation du système, notamment dans des conditions difficiles telles que la faible luminosité, les reflets sur les lunettes ou les mouvements importants de la tête. De plus, le système présente une limite importante en présence de lunettes de soleil, car celles-ci masquent les yeux et empêchent leur détection correcte, ce qui peut entraîner une baisse de performance ou l'absence de détection de la somnolence.

Plusieurs perspectives d'amélioration peuvent être envisagées afin d'optimiser davantage le système proposé. Parmi celles-ci, la conception et le développement d'un modèle personnalisé de détection du visage et des yeux constituent une piste prometteuse pour améliorer la précision, la robustesse et l'adaptabilité du système dans différentes conditions d'utilisation. L'intégration d'un tel modèle au sein du système proposé pourrait ainsi contribuer à renforcer ses performances globales.

Par ailleurs, l'intégration de techniques de suivi du visage (Face Tracking) permettrait d'assurer une détection plus stable en présence de mouvements de la tête. De même, l'amélioration des performances du système dans des conditions d'éclairage difficiles, notamment en faible luminosité ou lors de la conduite nocturne, représente un axe de développement important.

En outre, l'intégration d'indicateurs comportementaux complémentaires, tels que les bâillements ou les mouvements de la tête, pourrait contribuer à une détection plus fiable et plus précise de la somnolence du conducteur. Cette approche permettrait également d'améliorer la robustesse du système face à certaines situations où l'analyse des yeux devient difficile, notamment en présence de lunettes, d'occlusions partielles du visage ou de conditions d'éclairage défavorables.

En conclusion, ce travail démontre que l'utilisation conjointe de la vision par ordinateur et de l'apprentissage profond permet de concevoir des systèmes fiables et efficaces pour la détection de la somnolence du conducteur, contribuant ainsi à la réduction des risques d'accidents et à l'amélioration de la sécurité routière.

REFERENCES BIBLIOGRAPHIQUES

[1] M. H. Kryger, T. Roth, and W. C. Dement, Principles and Practice of Sleep Medicine, 6th ed. Philadelphia, PA, USA: Elsevier, 2017.

[2] T. Soukupová and J. Čech, “Real-Time Eye Blink Detection Using Facial Landmarks,” in Proceedings of the 21st Computer Vision Winter Workshop (CVWW), Rimske Toplice, Slovenia, 2016, pp. 1–8.

[3] M. F. Abulkhair, H. A. Salman, and L. F. Ibrahim, “Using Mobile Platform to Detect and Alert Driver Fatigue,” International Journal of Computer Applications, vol. 123, no. 8, pp. 27–35, Aug. 2015, doi: 10.5120/IJCA2015905428.

[4] European Road Safety Observatory (ERSO), Fatigue and Drowsiness in Road Traffic. Brussels, Belgium: European Commission, 2024.

[5] National Highway Traffic Safety Administration (NHTSA), Drowsy Driving and Automobile Crashes, Traffic Safety Facts, Report No. DOT HS 811 449. Washington, DC, USA: U.S. Department of Transportation, 2011.

[6] S. Saleem, “Risk Assessment for Road Traffic Accidents Related to Sleepiness During Driving: A Systematic Review,” Eastern Mediterranean Health Journal, vol. 28, no. 9, pp. 695–700, 2022, doi: 10.26719/emhj.22.055.

[7] M. Gonçalves, A. Amici, G. Lucas, et al., “Sleepiness at the Wheel Across Europe: A Survey of 19 Countries,” Journal of Sleep Research, vol. 24, no. 3, pp. 242–253, 2015, doi: 10.1111/jsr.12267.

[8] Ipsos and Fondation VINCI Autoroutes, Baromètre Européen de la Conduite Responsable 2017. Paris, France: Fondation VINCI Autoroutes, 2017.

[9] National Highway Traffic Safety Administration (NHTSA), “Drowsy Driving,” U.S. Department of Transportation, 2024. [Online]. Available: https://www.nhtsa.gov/risky-driving/drowsy-driving. [Accessed: Jun. 07, 2026].

[10] Algérie Presse Service (APS), “La somnolence au volant responsable d’environ 10 % des accidents de la route en Algérie,” APS, Alger, Algérie, 2016.

[11] M. Ramzan, H. U. Khan, S. M. Awan, A. Ismail, M. Ilyas, and A. Mahmood, “A Survey on State-of-the-Art Drowsiness Detection Techniques,” IEEE Access, vol. 7, pp. 61904–61919, 2019, doi: 10.1109/ACCESS.2019.2914373.

[12] E. Rogado, J. L. García, R. Barea, L. M. Bergasa, and E. López, “Driver Fatigue Detection System,” in Proc. IEEE International Conference on Robotics and Biomimetics (ROBIO), Bangkok, Thailand, 2009, pp. 186–191, doi: 10.1109/ROBIO.2009.4913155.

[13] M. Arunasalam, N. Yaakob, A. Amir, M. Elshaikh, and N. F. Azahar, “Real-Time Drowsiness Detection System for Driver Monitoring,” IOP Conference Series: Materials Science and Engineering, vol. 767, no. 1, Art. no. 012066, 2020, doi: 10.1088/1757-899X/767/1/012066.

[14] M. Abulkhair, A. H. Alsahli, K. M. Taleb, A. M. Bahran, F. M. Alzahrani, H. A. Alzahrani, and L. F. Ibrahim, “Mobile Platform Detect and Alerts System for Driver Fatigue,” Procedia Computer Science, vol. 65, pp. 813–822, 2015, doi: 10.1016/j.procs.2015.09.032.

[15] J. Y. Wong and P. Y. Lau, “Real-Time Driver Alert System Using Raspberry Pi,” ECTI Transactions on Electrical Engineering, Electronics and Communications, vol. 17, no. 2, pp. 149–158, 2019.

[16] A. Sahayadhas, K. Sundaraj, and M. Murugappan, “Detecting Driver Drowsiness Based on Sensors: A Review,” Sensors, vol. 12, no. 12, pp. 16937–16953, 2012, doi: 10.3390/s121216937.

[17] S. Leonhardt, L. Leicht, and D. Teichmann, “Unobtrusive Vital Sign Monitoring in Automotive Environments: A Review,” Sensors, vol. 18, no. 9, Art. no. 3080, 2018, doi: 10.3390/s18093080.

[18] M. Walter, M. Eilebrecht, T. Wartzek, and S. Leonhardt, “Vital Sign Monitoring in Car Seats Based on ECG, BCG and SCG,” Sensors, vol. 20, no. 19, Art. no. 5699, 2020, doi: 10.3390/s20195699.

[19] Toyota Motor Corporation, “Driver Monitoring System (DMS) for Lexus GS 450h,” Toyota Technical Report, Toyota City, Japan, 2006.

[20] Subaru Canada, “DriverFocus Driver Monitoring System,” [Online]. Available: https://www.subaru.ca/WebPage.aspx?WebSiteID=284&WebPageID=22757. [Accessed: Jun. 07, 2026].

[21] Toyota Motor Corporation, Toyota Safety Report 2017, Toyota City, Japan, 2017.

[22] P. Viola and M. Jones, “Rapid Object Detection Using a Boosted Cascade of Simple Features,” in Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Kauai, HI, USA, 2001, pp. I-511–I-518, doi: 10.1109/CVPR.2001.990517.

[23] N. Dalal and B. Triggs, “Histograms of Oriented Gradients for Human Detection,” in Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), San Diego, CA, USA, 2005, pp. 886–893, doi: 10.1109/CVPR.2005.177.

[24] Y. LeCun, Y. Bengio, and G. Hinton, “Deep Learning,” Nature, vol. 521, no. 7553, pp. 436–444, 2015, doi: 10.1038/nature14539.

[25] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, “Joint Face Detection and Alignment Using Multi-task Cascaded Convolutional Networks,” IEEE Signal Processing Letters, vol. 23, no. 10, pp. 1499–1503, 2016, doi: 10.1109/LSP.2016.2603342.

[26] V. Kazemi and J. Sullivan, “One Millisecond Face Alignment with an Ensemble of Regression Trees,” in Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Columbus, OH, USA, 2014, pp. 1867–1874, doi: 10.1109/CVPR.2014.241.

[27] C. Lugaresi et al., “MediaPipe: A Framework for Building Perception Pipelines,” arXiv preprint arXiv:1906.08172, 2019.

[28] T. Fawcett, “An Introduction to ROC Analysis,” Pattern Recognition Letters, vol. 27, no. 8, pp. 861–874, 2006, doi: 10.1016/j.patrec.2005.10.010.

[29] D. M. W. Powers, “Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness and Correlation,” Journal of Machine Learning Technologies, vol. 2, no. 1, pp. 37–63, 2011.

[30] I. Kayadibi, G. E. Güraksın, U. Ergün and N. Özmen Süzme, "An Eye State Recognition System Using Transfer Learning: AlexNet-Based Deep Convolutional Neural Network", International Journal of Computational Intelligence Systems, vol. 15, no. 1, 2022. DOI: 10.1007/s44196-022-00108-2.

[31] O. F. Hassan, A. F. Ibrahim, A. Gomaa, M. A. Makhlouf and B. Hafiz, "Real-time Driver Drowsiness Detection Using Transformer Architectures: A Novel Deep Learning Approach", Scientific Reports, vol. 15, 2025. DOI: 10.1038/s41598-025-02111-x.

[32] R. M. Salman, M. Rashid, R. Roy, M. M. Ahsan and Z. Siddique, "Driver Drowsiness Detection Using Ensemble Convolutional Neural Networks on YawDD", arXiv:2112.10298, 2021.

[33] R. Florez-Zela, F. Palomino-Quispe, R. J. Coaquira-Castillo, J. C. Herrera-Levano, T. Paixão and A. B. Alvarez, "A CNN-Based Approach for Driver Drowsiness Detection by Real-Time Eye State Identification", Applied Sciences, vol. 13, no. 13, 7849, 2023. DOI: 10.3390/app13137849.