



UNIVERSITÉ DU QUÉBEC À RIMOUSKI

Détection du niveau de fatigue au volant à l'aide de l'intelligence artificielle

MÉMOIRE DE MAÎTRISE

*Ce mémoire est présenté comme exigence pour l'obtention du diplôme
de Maîtrise en informatique – volet recherche (3017)*

Maître ès sciences (M.Sc.)

Département de mathématique, d'informatique et génie

PAR

© DORRA LAMOUCI

30 août 2024

Composition du jury :

- **Pr. Yacine Benahmed**, président du jury, Université du Québec à Rimouski
- **Pr. Yacine Yaddaden**, directeur de recherche, Université du Québec à Rimouski
- **Pr. Raef Cherif**, codirecteur de recherche, Université du Québec à Rimouski
- **Pr. Abdenour Bouzouane**, examinateur externe, Université du Québec à Chicoutimi

Moi, Dorra LAMOUCI, déclare que ce mémoire intitulé, « Détection du niveau de fatigue au volant à l'aide de l'intelligence artificielle » et le travail qui y est présenté sont de ma propre rédaction. Je confirme que :

- Ce travail a été réalisé entièrement et principalement pendant ma candidature à un diplôme de recherche au sein de cette université.
- Lorsque toute partie de cette thèse a été précédemment soumise pour l'obtention d'un diplôme dans cette université, cela a été clairement indiqué.
- Lorsque j'ai consulté les travaux publiés par d'autres, cela est toujours clairement attribué.
- Lorsque j'ai cité la source est toujours indiquée. À l'exception de telles citations, cette thèse est entièrement le fruit de mon propre travail.
- J'ai reconnu toutes les principales sources d'aide.
- Lorsque la thèse est basée sur un travail que j'ai réalisé en collaboration avec d'autres, j'ai clairement indiqué ce qui a été fait par les autres et ce à quoi j'ai contribué moi-même.

Signé : Dorra Lamouchi

UNIVERSITÉ DU QUÉBEC À RIMOUSKI

Service de la bibliothèque

Avertissement

La diffusion de ce mémoire ou de cette thèse se fait dans le respect des droits de son auteur, qui a signé le formulaire « Autorisation de reproduire et de diffuser un rapport, un mémoire ou une thèse ». En signant ce formulaire, l'auteur concède à l'Université du Québec à Rimouski une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de son travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, l'auteur autorise l'Université du Québec à Rimouski à reproduire, diffuser, prêter, distribuer ou vendre des copies de son travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de la part de l'auteur à ses droits moraux ni à ses droits de propriété intellectuelle. Sauf entente contraire, l'auteur conserve la liberté de diffuser et de commercialiser ou non ce travail dont il possède un exemplaire.

« La vie n'est facile pour aucun de nous. Mais quoi, il faut avoir de la persévérance, et surtout de la confiance en soi. Il faut croire que l'on est doué pour quelque chose, et que, cette chose, il faut l'atteindre coûte que coûte. »

Marie Curie

UNIVERSITÉ DU QUÉBEC À RIMOUSKI

Résumé

Département de mathématique, d'informatique et génie

Maîtrise en informatique – volet recherche (3017)

**Détection du niveau de fatigue au volant à l'aide de l'intelligence
artificielle**

par Dorra LAMOUCHI

Ces dernières années, le nombre d'accidents mortels sur les routes a augmenté de manière significative, en parallèle à la croissance du nombre de véhicules, en particulier sur les longs trajets entre des régions éloignées. Plusieurs facteurs peuvent être à l'origine de ces accidents, mais la plupart d'entre eux sont directement liés à l'état du conducteur. La fatigue, en particulier, revêt une importance cruciale, car elle entraîne une diminution des capacités cognitives et une réduction des réflexes, ce qui augmente considérablement les risques d'accident. De plus, les conducteurs ont souvent du mal à prendre conscience de leur propre niveau de fatigue.

Grâce aux récentes avancées en intelligence artificielle, il est désormais possible de concevoir des systèmes de détection automatique de la fatigue au volant. Bien que différents systèmes aient déjà été proposés, leurs performances ne sont pas encore optimales, ouvrant ainsi la voie au développement de nouvelles méthodes. En explorant d'autres modalités et sources de données, il devient possible de caractériser de manière plus précise la fatigue et d'alerter le conducteur au moment opportun, contribuant ainsi à la réduction des risques d'accidents dus à la fatigue sur la route.

On a commencé par un examen approfondi des méthodes existantes de détection de la fatigue. Les approches physiologiques, bien que précises, sont souvent coûteuses et intrusives, tandis que les méthodes comportementales peuvent être affectées par des facteurs externes, ce qui réduit leur fiabilité.

Pour surmonter ces défis, ce mémoire propose le développement de deux systèmes de détection distincts à partir de données physiologiques et comportementales. La première approche intègre des données physiologiques comme le rythme cardiaque pour évaluer les niveaux de somnolence. La seconde approche utilise la reconnaissance faciale avancée pour analyser les signes de fatigue tels que les clignements des yeux et les bâillements de la bouche. Ces deux systèmes sont élaborés et validés à l'aide de données publiques pour permettre des comparaisons avec des recherches antérieures.

En conclusion, ce mémoire vise à fournir une solution complète pour la détection de la fatigue des conducteurs combinant les avantages des approches comportementales et physiologiques tout en surmontant leurs limitations respectives.

UNIVERSITY OF QUEBEC AT RIMOUSKI

Abstract

Department of Mathematics, Computer Science and Engineering

Research Master's degree in Computer Science (3017)

Driver fatigue detection using Artificial Intelligence

by Dorra LAMOUCHI

In recent years, the number of fatal road accidents has increased significantly, in parallel with the growth in the number of vehicles, especially on long journeys between remote regions. Several factors can be the cause of these accidents, but most of them are directly related to the driver's condition. Fatigue, in particular, is of crucial importance, as it leads to a decrease in cognitive abilities and a reduction in reflexes, which significantly increases the risk of accidents. Additionally, drivers often have difficulty recognizing their own level of fatigue.

Thanks to recent advances in artificial intelligence, it is now possible to design automatic fatigue detection systems for drivers. Although different systems have already been proposed, their performance is not yet optimal, paving the way for the development of new methods. By exploring other modalities and data sources, it becomes possible to more precisely characterize fatigue and alert the driver at the right time, thus contributing to the reduction of fatigue-related road accidents.

This study began with an in-depth examination of existing methods for detecting fatigue. Physiological approaches, although precise, are often expensive and intrusive, while behavioral methods can be affected by external factors, reducing their reliability.

To overcome these challenges, this thesis proposes the development of two distinct detection systems based on physiological and behavioral data. The first approach integrates physiological data such as heart rate to assess levels of drowsiness. The second approach uses advanced facial recognition to analyze signs of fatigue such as eye blinks and yawns. These two systems are developed and validated using public data to allow comparisons with previous research.

In conclusion, this thesis aims to provide a comprehensive solution for detecting driver fatigue by combining the advantages of both behavioral and physiological approaches while overcoming their respective limitations.

Remerciements

J'aimerais remercier toute personne ayant contribué de près ou de loin à la réalisation de ce travail. Tout d'abord, je tiens à remercier particulièrement mon directeur de recherche **Pr. Yacine Yaddaden**, pour son suivi, sa patience et son encouragement. Son œil critique m'a été très précieux pour structurer le travail et pour améliorer la qualité des différentes sections.

Je remercie également mon co-directeur de recherche **Pr. Raef Cherif** pour son aide immense ainsi que pour tous les conseils et les informations qu'il m'a prodigués avec un degré de patience et de professionnalisme sans égal.

Que les membres de jury **Pr. Yacine Benahmed** et **Pr. Abdenour Bouzouane** trouvent, ici, l'expression de mes sincères remerciements pour l'honneur qu'ils me font en prenant le temps de lire et d'évaluer ce travail.

Je tiens également à exprimer ma profonde gratitude à tous mes professeurs qui ont donné d'énormes efforts à me guider et ont sculpté mes connaissances sans lesquels je ne serais pas en mesure d'établir ce travail.

Pour finir, je remercie mes chers parents qui m'ont guidée, soutenue et qui ont fait de moi la personne que je suis aujourd'hui et à qui je serai éternellement reconnaissante, mes grands-parents qui m'ont accablé par leur amour et leur intérêt infinis et que j'essaie de rendre fiers de moi, ma sœur et mon frère qui m'ont soutenu et encouragé à aller de l'avant et à dépasser mes frontières, mon futur mari qui est ma source de motivation infinie et mes amis à Rimouski, à Lévis, à Québec, à Drummondville et à Montréal qui ont su adoucir le mal du pays et la sensation d'être une étrangère dans un nouveau monde.

Table des matières

Résumé	xii
Abstract	xiv
Remerciements	xv
1 Introduction générale	1
1.1 Mise en contexte	1
1.2 Problématique	3
1.2.1 Caractère variable de la fatigue	4
1.2.2 Difficulté de reconnaissance des symptômes	4
1.2.3 Exigences des systèmes de détection de fatigue	4
1.2.4 Défis éthiques	5
1.3 Objectifs	5
1.4 Méthodologie suivie	6
1.5 Contributions	9
1.6 Structure du mémoire	10
2 Revue de littérature	11
2.1 Les fondamentaux	11
2.1.1 Théorie de la fatigue	11
2.1.2 La différence entre <i>fatigue</i> et <i>somnolence</i>	13
2.1.3 La fatigue au volant	13
2.1.4 Les systèmes de détection de la fatigue	14
Caractéristiques centrées sur le conducteur	15
Caractéristiques orientées véhicule	16
2.1.5 Les techniques de détection de la fatigue	16
2.2 Bases de données	18
2.2.1 National Tsing Hua University (NTHU-DDD) [58] . . .	19
2.2.2 Real-Life Drowsiness Dataset (UTA-RLDD) [23]	19
2.2.3 ULg multimodality drowsiness database (DROZY) [39]	20
2.2.4 Yawning detection dataset (YawDD) [1]	21

2.2.5	MRL Eye dataset [22]	22
2.2.6	Consented Activities of People (CAP) [57]	23
2.2.7	Closed Eyes in the Wild dataset (CEW) [51]	23
2.2.8	Bases de données utilisées	23
2.3	Méthodes et approches existantes	24
2.3.1	Détection centrée sur le véhicule	24
2.3.2	Détection centrée sur le conducteur	25
	Selon les caractéristiques physiologiques	25
	Selon les caractéristiques comportementales	26
2.4	Bilan et discussion	27

3	Comparative Study : Physiological-based Driver Drowsiness Detection Utilizing Traditional and Hybrid Methods	31
	Abstract	32
3.1	Introduction	32
3.2	Fundamentals	33
	3.2.1 Human-centered features	33
	3.2.2 Vehicle-centered measures	33
3.3	Related works	33
	3.3.1 Unimodal DDD systems	33
	3.3.2 Multimodal DDD systems	34
3.4	Feature extraction techniques	34
	3.4.1 Time domain features	34
	3.4.2 Frequency domain features	34
	3.4.3 Time-Frequency domain features	34
3.5	Classification techniques	35
	3.5.1 Decision Tree	35
	3.5.2 Support Vector Machine	35
	3.5.3 Random Forest	35
	3.5.4 Multilayer Perceptron	35
	3.5.5 Quadratic Discriminant Analysis	36
3.6	Proposed system	36
	3.6.1 Pre-processing	36
	3.6.2 Feature extraction	36
	3.6.3 Dimensionality reduction	36
	3.6.4 Classification	36
3.7	Evaluation protocol	36
	3.7.1 Data description	36

3.7.2	Evaluation metrics	37
3.7.3	Validation strategy	37
3.8	Results & discussion	37
3.9	Concluding remarks	38
	Acknowledgment	38
	References	38
4	Efficient Driver Drowsiness Detection Using Spatiotemporal Features with Support Vector Machine	41
	Abstract	43
4.1	Introduction	44
4.2	Background	45
4.2.1	Fundamentals	45
4.2.2	Related Works	46
4.3	Proposed System	49
4.3.1	Pre-processing	50
4.3.2	Spatiotemporal Feature Extraction	50
4.3.3	Dimensionality Reduction	52
4.3.4	Classification	53
4.4	Data Description	53
4.4.1	UTA-RLDD	53
4.4.2	DROZY	54
4.5	Evaluation Protocol	54
4.5.1	Performance Metrics	54
4.5.2	Validation Strategy	55
4.6	Results & Discussion	56
4.7	Conclusion	58
4.7.1	Limitations	58
4.7.2	Perspectives & Improvements	59
	Acknowledgment	59
	References	59
5	Conclusion Générale	63
5.1	Objectifs accomplis	63
5.2	Comparaison des approches développées	64
5.3	Limitations connues	65
5.4	Perspectives et travaux futurs	66
5.5	Retour d'expérience	67

Annexe A	69
Annexe B	79
Bibliographie	83

Table des figures

1.1	Proportion des causes d'accidents dues à des facteurs opérationnels [63].	3
1.2	Processus de la détection de l'état de fatigue.	8
2.1	La catégorisation des sources de données de la fatigue.	15
2.2	Les différents techniques de détection de la fatigue.	17
2.3	Aperçu de la collecte de données à partir de divers capteurs [18].	18
2.4	Exemples de données NTHU-DDD [58].	19
2.5	Exemples de données UTA-RLDD [23].	20
2.6	Exemples de données DROZY [39].	21
2.7	Exemples de données YawDD [1].	22
2.8	Exemples de données MRL Eye [22].	23
3.1	The time-frequency domain feature extraction pipeline.	35
3.2	Overview of signals collected from various sensors.	36
3.3	Overview of the proposed DDD system.	37
3.4	Comparison of the explored methods.	38
4.1	Overview of the proposed DDD system.	49
4.2	Region of interest extraction : (a) eyes (b) mouth.	50
4.3	Overview of the basic LBP.	51
4.4	Overview of the LBP-TOP.	51
4.5	Sample data from UTA-RLDD.	54
4.6	Sample data from DROZY.	55
4.7	The result of the experiment on various LBP-TOP parameters.	56
5.1	Comparaison des résultats des méthodes élaborées.	65

Liste des tableaux

2.1	Les bases de données utilisées.	24
2.2	Bilan des travaux physiologiques.	27
2.3	Bilan des travaux comportementaux.	28
3.1	List of the used feature extraction techniques.	36
3.2	The Karolinska Sleepiness Scale (KSS).	37
3.3	Overview of the classification results.	37
3.4	Confusion Matrix (%).	38
3.5	Comparison in terms of accuracy with state-of-the-art methods.	39
4.1	The Karolinska Sleepiness Scale (KSS).	46
4.2	Comparison with state-of-the-art methods using UTA-RLDD.	57
4.3	The confusion matrix using 2 classes on the UTA-RLDD dataset.	57
4.4	The confusion matrix using 3 classes on the UTA-RLDD dataset.	57
4.5	Comparison with state-of-the-art methods using DROZY.	58
4.6	The confusion matrix using the DROZY dataset.	58

Liste des abréviations

OMS	Organisation Mondiale de la Santé
ML	Machine Learning
DL	Deep Learning
EEG	ÉlectroEncéphaloGramme
ECG	ÉlectroCardioGramme
EOG	ÉlectroOculoGramme
EMG	ÉlectroMyoGramme
SVM	Support Vectors Machine
DT	Decision Tree
RF	Random Forest
RL	Regression Logistique
MLP	Perceptron Multicouche
KNN	K-Nearest Neighbors
CNN	Convolutional Neural Network
DNN	Deep Neural Network
ANN	Artificial Neural Network
LSTM	Long Short-Term Memory
GBM	Gradient Boosting Machine
GAN	Generative Adversarial Network
NTHU	National Tsing Hua University
UTA-RLDD	Real-Life Drowsiness Dataset
CAP	Consented Activities of People
YawDD	Yawning Detection Dataset
CEW	Closed Eyes in the Wild
YEC	YouTube Eye Classification
ABD	Autonomus Blink Dataset
SWM	Steering Wheel Movement
SDLP	Standard Deviation of Lateral Position
ApEn	Approximate Entropy
SC	Skin Conductance
SCL	Skin Conductance Level
SCR	Skin Conductance Response
IHM	Interaction Homme-Machine
ROI	Region Of Interest
KSS	Karolinska Sleepiness Scale

*Je dédie ce projet à toute ma famille et ceux qui me
sont chers, merci pour votre amour et votre
encouragement.*

Chapitre 1

Introduction générale

1.1 Mise en contexte

De nos jours, Les routes sont de plus en plus longues et complexes, souvent inadaptées à la densité croissante du trafic, alors que le nombre de véhicules continue à augmenter. Les trajets prolongés, notamment ceux reliant des régions éloignées, nécessitent des périodes de conduite étendues [11]. De plus, les voyages d'affaires sont devenus plus fréquents, obligeant les conducteurs à passer de nombreuses heures sur la route. Par ailleurs, l'augmentation du trafic de transport de marchandises par camion sur de longues distances exacerbe le problème, car les chauffeurs de poids lourds sont souvent soumis à des horaires exigeants et à des conditions de travail difficiles. Les réglementations sur les heures de conduite sont parfois insuffisantes pour prévenir la fatigue et les pauses obligatoires peuvent s'avérer inadéquates. La conduite nocturne présente également des défis importants tels que la réduction de la visibilité et la diminution de la vigilance [49].

Tous ces facteurs contribuent au risque des accidents, ce qui constitue une préoccupation majeure. Selon un nouveau rapport de l'organisation mondiale de la santé (OMS) publié en 2023, le nombre annuel de décès dus aux accidents de la route atteint 1,19 million par an [41], ainsi que des dizaines de millions d'autres étaient blessés. Les conséquences des accidents de la route sont considérables, aussi bien pour la santé et le bien-être des personnes affectées que pour leur famille et leur communauté sans oublier les énormes répercussions économiques qu'ils entraînent. La nature, souvent tragique des

accidents, souligne l'importance de développer des méthodes efficaces pour les prévenir.

Plusieurs facteurs [14] contribuent aux accidents de la route :

État du conducteur :

- La distraction du conducteur, souvent accentuée par l'utilisation d'appareils mobiles et des activités tel que manger ou par la déconcentration liée à la présence d'autres passagers.
- L'alcool, les drogues le stress et la somnolence perturbent les réflexes et le jugement du conducteur, ce qui rend la conduite extrêmement risquée.

Façon de conduite :

- La vitesse excessive est un facteur important, augmentant à la fois la probabilité d'accidents et la gravité des dommages subis.

Facteurs environnementaux :

- Des conditions météorologiques défavorables comme la pluie, le brouillard ou la neige accentuent les dangers vu la réduction de la visibilité et le manque d'adhérence à la route.
- Les animaux sauvages qui traversent les routes, les objets qui tombent des véhicules en mouvement ou les débris sur la route sont des exemples de facteurs imprévus qui surprennent les conducteurs.
- Les routes mal conçues et mal entretenues, le manque de visibilité aux intersections et le manque d'éclairage nocturne représentent des défaillances de l'infrastructure favorisant les collisions.

Plusieurs études démontrent que la fatigue au volant est un facteur important de mortalité chez les conducteurs sur la route. Selon l'administration nationale de la sécurité routière [40], plus de 40% des conducteurs ayant fait des accidents ont déclaré dormir moins de 6 heures par nuit. Conduire fatigué peut être aussi dangereux que conduire sous l'influence de l'alcool [21]. Sur le plan physiologique, la fatigue réduit les facultés cognitives essentielles pour une conduite sûre. Elle affecte la vigilance, la concentration et le temps de réaction. Le manque de sommeil peut entraîner des courtes périodes de

somnolence qui durent quelques secondes, pendant lesquelles un conducteur peut perdre le contrôle d'un véhicule. Cela est particulièrement dangereux sur les autoroutes à grande vitesse ou dans des conditions de circulation dense.

Plusieurs facteurs peuvent contribuer à la fatigue des conducteurs, notamment le manque de sommeil, les longues heures de conduite sans pauses suffisantes, la conduite nocturne, les conditions médicales non diagnostiquées telles que l'apnée du sommeil et le travail en horaires décalés. Une étude récente examine le poids des différents facteurs de fatigue dans une population donnée (figure 1.1). La fatigue est particulièrement prévalente chez les conducteurs expérimentés tels que les chauffeurs de camion, mais elle affecte également les conducteurs quotidiens.

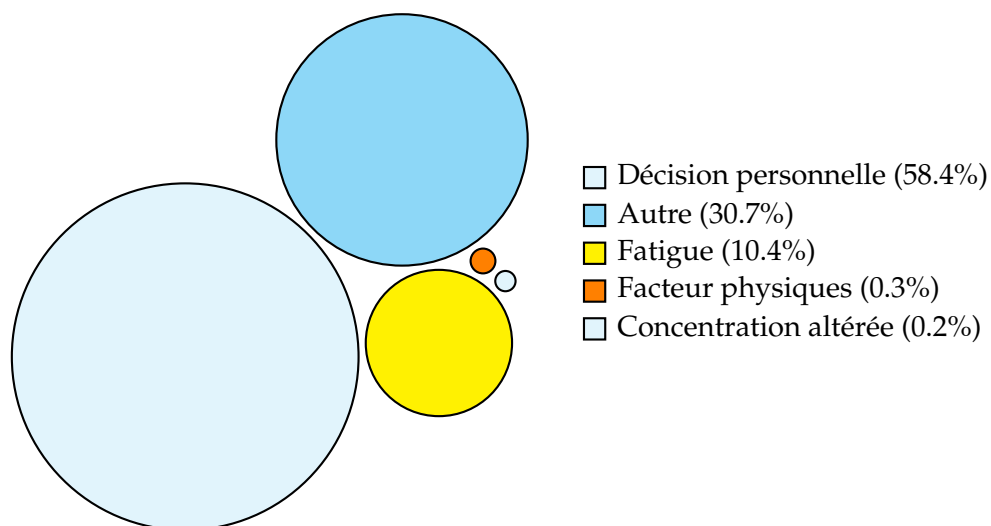


FIGURE 1.1 – Proportion des causes d'accidents dues à des facteurs opérationnels [63].

1.2 Problématique

Habituellement, la détection de la fatigue reposait sur l'auto-évaluation du conducteur, une méthode généralement peu fiable [53]. Avec le développement progressif des technologies, des méthodes plus objectives ont été élaborées. Bien qu'avancées, ces méthodes présentent encore des limites [10].

1.2.1 Caractère variable de la fatigue

Le phénomène de la fatigue se caractérise par son caractère subjectif et variable, ce qui lui confère un degré de complexité important. En effet, comme les manifestations de la fatigue varient considérablement d'une personne à l'autre [46], il est extrêmement difficile de normaliser les mécanismes de détection. Par exemple, certains conducteurs peuvent ressentir une baisse de vigilance, caractérisée par des moments de distraction ou de perte de concentration. D'autres peuvent avoir des réactions physiques telles que des bâillements fréquents, des yeux qui piquent ou des mouvements de tête répétitifs indiquant une lutte contre le sommeil.

1.2.2 Difficulté de reconnaissance des symptômes

Ces symptômes ne sont pas toujours évidents pour un conducteur. La fatigue peut s'installer graduellement, de sorte qu'il est difficile pour le conducteur de réaliser qu'il est trop fatigué pour conduire en toute sécurité. Cela est particulièrement plus grave pour les trajets longs ou monotones, où la fatigue peut augmenter sans que le conducteur n'en soit totalement conscient [37]. En outre, la détection de la fatigue ne se limite pas à observer les signes physiques ou comportementaux. Des facteurs tels que la durée du voyage, les conditions de luminosité et les conditions de santé personnelles jouent également un rôle important nécessitant ainsi des systèmes de détection qui peuvent intégrer et analyser plusieurs données pour bien évaluer le niveau de fatigue.

1.2.3 Exigences des systèmes de détection de fatigue

Les systèmes de détection de fatigue doivent être abordables et disponibles pour une large gamme de véhicules, pas seulement pour les modèles haut de gamme afin de garantir les bénéfices en matière de sécurité pour la majorité des conducteurs. Par contre, sur le plan technique, l'intégration de

ces systèmes est un défi majeur. Une fausse alerte pourrait être aussi dangereuse qu'une non-détection. Ces systèmes doivent fonctionner dans diverses conditions de conduite, être résistants aux interférences extérieures et surtout ils ne doivent pas nécessiter une interaction complexe de la part du conducteur, ce qui pourrait potentiellement détourner son attention de la route.

1.2.4 Défis éthiques

Parmi les défis éthiques, on trouve la question de la discrétion des systèmes de détection qui impliquent souvent une certaine forme de surveillance, soulevant des préoccupations en matière de vie privée. Par exemple, l'utilisation de caméras pour surveiller les signes de fatigue peut être perçue comme intrusive, engendrant des problèmes de confort pour les utilisateurs.

1.3 Objectifs

Le but principal de ce projet est de développer un système fiable pour la détection de la fatigue au volant afin de diminuer les accidents routiers ainsi que les dommages humains et matériels qui en découlent. Cette démarche permet d'aborder la problématique en se concentrant sur deux dimensions fondamentales de l'expérience humaine : les signaux physiologiques et les aspects comportementaux.

Détection de la fatigue par une analyse physiologique

Développer un système de détection de la fatigue basé sur les signaux physiologiques. Cet objectif se décompose en sous-objectifs :

- **Identification des signaux** : Réaliser une étude de corrélation des signaux physiologiques les plus liés à la fatigue.
- **Caractérisation des signaux** : Trouver les pipelines idéaux pour qualifier les signaux.

Détection de la fatigue par une analyse faciale

Développer un système de détection de la fatigue basé sur les aspects comportementaux. Cet objectif comprend les sous-objectifs suivants :

- **Uniformiser les données à utiliser** : Garantir la cohérence des données provenant de différentes bases de données.
- **Développement de modèles comportementaux** : Développer des modèles qui apprennent à détecter les schémas de fatigue à partir des signes comportementaux observés.

1.4 Méthodologie suivie

Dans ce projet, nous avons adopté une approche méthodologique progressive pour approfondir notre problématique et atteindre des objectifs pertinents, tout en tenant compte des recherches existantes. Les étapes que nous avons suivies sont présentées ci-dessous.

Revue de littérature

Cette étape consiste principalement à examiner plusieurs articles de conférences ou de revues scientifiques dans le domaine lié au projet de recherche, ce qui nous permettra d'obtenir un aperçu détaillé de l'état de l'art et des avancées déjà réalisées dans le domaine de l'amélioration de la vigilance au volant. De plus, cela nous permettra d'acquérir une compréhension générale du processus suivi lors du développement de ce type de méthode, d'évaluer les performances existantes, d'identifier les lacunes ainsi que les améliorations nécessaires pour obtenir un système optimal. Cette étape est essentielle pour définir les axes de recherche à explorer. En particulier, notre revue de littérature se concentrera sur les approches physiologiques et comportementales pour la détection de la fatigue au volant.

Définition des axes de travail

Pour la détection de la fatigue, diverses sources de données peuvent être exploitées. Dans le cadre de ce projet, on va s'intéresser à deux types de mesures :

- *Physiologiques* : en analysant les signaux liés à l'activité des ondes cérébrales, du rythme cardiaque, des mouvements oculaires et de la réaction musculaire.
- *Comportementales* : à travers l'analyse faciale, plus précisément en surveillant l'activité oculaire et les mouvements de la bouche.

Identification et récupération des données

En raison de contraintes de temps, il ne sera pas possible d'effectuer une collecte de données. Heureusement, lors de l'étape de revue de la littérature, nous avons trouvé des bases de données de référence [39, 23]. Nous devons néanmoins les préparer en effectuant plusieurs étapes de pré-traitement, telles que le nettoyage, la transformation, la normalisation et l'extraction des informations pertinentes. Ces bases de données publiques largement reconnues dans le domaine de la détection de fatigue au volant nous permet d'assurer la possibilité de se comparer à d'autres travaux déjà explorés sans avoir les contraintes logistiques et éthiques liées à la collecte de données.

Conception du système de la détection de la fatigue

Les caractéristiques extraites servent au développement de modèles qui apprennent à identifier les schémas associés à la fatigue en utilisant des techniques d'intelligence artificielle tels que les réseaux de neurones et les techniques d'apprentissage automatique. Ces modèles sont validés et évalués en utilisant des métriques particulières en fonction desquels le système peut être optimisé et ajusté.

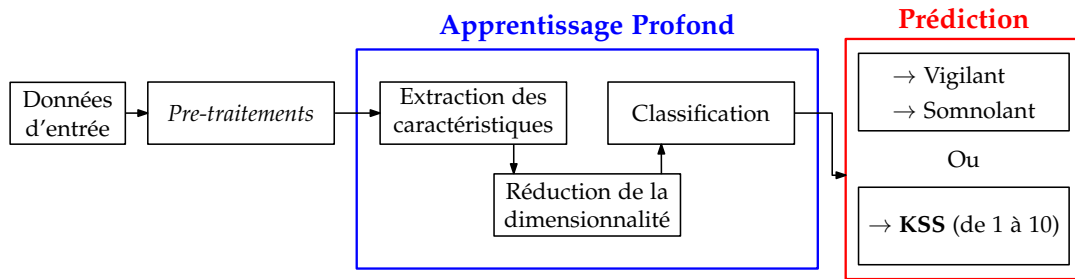


FIGURE 1.2 – Processus de la détection de l'état de fatigue.

La figure 1.2 illustre le déroulement des étapes qui seront expliquées en détail ci-dessous selon la nature des données :

1. Système de détection basé sur les données physiologiques :
 - Traitement des données : Pour les données physiologiques, un pré-traitement concernant le type des fichiers a été appliqué.
 - Extraction des caractéristiques : Nous avons utilisé des techniques de traitement du signal pour analyser les données choisies.
 - Sélection des attributs et réduction de dimensionnalité : c'est une étape optionnelle éliminant les caractéristiques redondantes et non pertinentes.
 - Analyse et interprétation des données : Des tests statistiques adaptés d'une part et un processus d'apprentissage et validation d'un modèle d'apprentissage automatique d'autre part ont permis de déterminer la relation entre les caractéristiques extraites et les niveaux de fatigue.
2. Système de détection basé sur les données comportementales :
 - Traitement des données : La première étape consiste à identifier la nature des données d'entrée en fonction de l'approche à adopter, qu'il s'agisse de vidéos ou d'images. Ensuite, on va procéder à une étape de normalisation et de résolution des problèmes de qualité tels que la résolution.
 - Extraction des caractéristiques : Nous avons employé des algorithmes avancés de reconnaissance faciale et de traitement d'images pour extraire des indicateurs comportementaux clés de la fatigue.

- Analyse et interprétation des données : Les caractéristiques extraites ont été analysées afin d'identifier des patterns récurrents associés à la fatigue ce qui a permis de définir des seuils spécifiques indiquant les niveaux de vigilance réduits chez le conducteur.

Ces étapes traditionnelles de traitement des données peuvent être remplacées par des réseaux neuronaux profonds ayant la capacité d'extraire automatiquement des caractéristiques pertinentes à partir de données brutes.

1.5 Contributions

Le présent travail de recherche a pour objectif principal de proposer un système de détection de la fatigue au volant afin de diminuer les accidents routiers en explorant les caractères physiologiques et les caractères comportementaux.

Conférences Internationales

Dans le cadre de ce mémoire, nous avons eu l'occasion de présenter notre progression ainsi que nos méthodes proposées lors de conférences internationales. Voici une liste de nos contributions acceptées et présentées lors d'événements de vulgarisation scientifique :

D. Lamouchi, Y. Yaddaden, R. Cherif, "Comparative Study : Physiological-based Driver Drowsiness Detection Utilizing Traditional and Hybrid Methods," in *International Conference on Image and Signal Processing and their Applications (ISPA'2024)*, IEEE, 2024, pp. 1-8, DOI :10.1109/ISPA59904.2024.10536845. (Statut : Présenté)

N. Cabot, **D. Lamouchi**, Y. Yaddaden, "Drowsiness Detection through Yawning and Eye Blinking Models Using Convolutional Neural Networks and Transfer Learning" in *International Conference On Intelligent Computing in Data Sciences (ICDS2024)*. (Statut : Accepté)

Revues Internationales

Nous avons également soumis nos travaux à des revues scientifiques internationales de renom. Voici le papier actuellement en cours de traitement :

D. Lamouchi, Y. Yaddaden, J. Parent, R. Cherif, "Efficient Driver Drowsiness Detection Using Spatiotemporal Features with Support Vector Machine" in *International Journal of Intelligent Transportation Systems Research*. (Statut : Soumis)

1.6 Structure du mémoire

Ce mémoire est découpé comme suit :

- Le chapitre Revue de littérature présente une revue de littérature relative au sujet. Il aborde également les limites identifiées.
- Le Chapitre 3 présente une étude comparative basée sur les signaux physiologiques pour la détection de fatigue au volant.
- Le Chapitre 4 illustre la détection de la somnolence du conducteur à l'aide des motifs binaires locaux de trois plans orthogonaux.
- La Conclusion Générale présente un bilan des méthodes explorées, en mettant en avant leurs forces et leurs limites, ainsi que des perspectives potentielles pour notre recherche.
- L'Annexe A met en valeur le développement de techniques automatisées pour la détection des signes de fatigue : les bâillements et les clignements des yeux.
- L'Annexe B expose les deux posters de mes participations au Forum Innovation, Ingénierie, Informatique et Entrepreneuriat (FI3E).

Chapitre 2

Revue de littérature

Ce chapitre explore les recherches et les efforts déployés par différentes équipes concernant la problématique de la détection de la fatigue chez les conducteurs.

La section 2.1 présente les approches fondamentales, suivie par la section 2.2 qui décrit les principales bases de données utilisées. Ensuite, la section 2.3 examine le travail effectué dans d'autres recherches, et enfin, la section 2.4 offre un bilan et une discussion sur les limitations des approches présentées dans ce chapitre.

2.1 Les fondamentaux

Dans cette partie, les principes de base qui déterminent la compréhension de la fatigue du conducteur sont abordés.

2.1.1 Théorie de la fatigue

Selon les perspectives théoriques classiques, la fatigue est le résultat de transformations biochimiques affectant l'organisme dans son ensemble, ou spécifiquement les organes actifs. D'un point de vue opérationnel, la fatigue peut être décrite comme un état psychophysique réversible induit par des périodes d'activité prolongées [46]. Cet état se caractérise par un ensemble de répercussions négatives qui se manifestent par des changements physiologiques, psychologiques et comportementaux. Les chercheurs ont souvent lié la fatigue à un apport insuffisant d'oxygène et de nutriments dans le sang

ou le plasma sanguin, ainsi qu'à une surcharge de substances résiduelles et de produits métaboliques. Les recherches sur la fatigue ont révélé des corrélations étroites avec le fonctionnement du système cérébral. Ces aspects comprennent l'épuisement des neurones au cours de longues périodes d'activité, la dynamique neuronale qui facilite le transport des nutriments vers les organes en activité, la rupture de l'équilibre au sein des activités neurophysiologiques et le rôle critique joué par les processus d'inhibition du cerveau [33].

La fatigue peut survenir de manière soudaine, et dans ce cas, elle est appelée **fatigue aiguë**, ou persister sur une longue période, alors elle est désignée comme **fatigue chronique** [16]. Ses causes sont multiples et interdépendantes :

Raisons physiques

- Manque de sommeil.
- Changements hormonaux : le déséquilibre dans la production de *cortisol* et *mélatonine* affectent le niveau d'énergie et la qualité du sommeil.
- Les maladies, telles que *l'hypothyroïdie*¹ ou *l'anémie*² : provoquent une fatigue chronique et influencent le métabolisme ainsi que le niveau d'oxygène dans le sang.
- Mauvaise alimentation.
- Déshydratation.

Raisons psychologiques

- Stress et anxiété : sollicitent les ressources mentales.
- Excès d'efforts mentaux : ces tâches demandent une concentration intense, épuisant ainsi les ressources cognitives.
- Difficultés à gérer les émotions.

Raisons environnementaux

- Bruit excessif.
- Éclairage inadéquat.

1. La glande thyroïde ne produit pas suffisamment d'hormones

2. Une insuffisance de globules rouges dans le sang

- Températures extrêmes.
- Manque de confort général selon la posture et l'environnement.

2.1.2 La différence entre *fatigue* et *somnolence*

La fatigue et la somnolence sont deux états distincts mais souvent liés. Malgré leur différence, les termes "fatigue" et "somnolence" sont souvent utilisés de manière interchangeable [13]. La fatigue est décrite comme l'incapacité d'exécuter une activité particulière en raison d'un effort mental ou physique ou de l'exécution d'une même tâche pendant une longue période. D'autre part, la somnolence est caractérisée par l'envie de dormir. Elle est causée par un besoin biologique impérieux de dormir. La somnolence peut avoir diverses causes notamment les problèmes de santé, les effets secondaires de médicaments, le travail pendant de longues heures, les difficultés de sommeil, etc. Il est donc évident que la fatigue contribue immédiatement à la somnolence.

Bien que la somnolence et la fatigue soient des concepts distincts, plusieurs travaux de recherche [19, 29] les ont assimilées en raison de leurs conséquences comparables. En intégrant les deux concepts, les chercheurs visent à développer des systèmes plus robustes capables d'identifier les signes précoces de danger, indépendamment de la source spécifique du problème, pour améliorer la sécurité des individus.

2.1.3 La fatigue au volant

La fatigue au volant est un terme utilisé pour désigner la fatigue ressentie par les conducteurs de divers types de véhicules tels que les voitures, les autobus, les camions, les trains, ou encore les opérateurs de grues de construction et d'autres engins lourds. Les causes de cette fatigue, les symptômes et les risques associés varient fortement en fonction du contexte et surtout la détection nécessite une approche adaptée à chaque situation spécifique [26].

Les conducteurs routiers professionnels subissent une pression constante afin de respecter des délais serrés, ce qui accroît leur niveau de stress et, par

conséquent, leur fatigue. En plus, elle peut être parfois accumulée au cours de longues périodes de travail sans repos adéquat. Cela est particulièrement valable pour les conducteurs de longue distance et les opérateurs d'engins lourds dans le cadre de projets de construction ou de travaux miniers [3].

Les domaines suivants illustrent des exemples typiques où maintenir une vigilance accrue au volant est essentiel :

Transport de voitures simples : Les conducteurs de véhicules personnels effectuant de longs trajets, souvent pour des raisons professionnelles ou personnelles, sans prendre suffisamment de pauses.

Transport de passagers : Les conducteurs d'autobus qui doivent respecter des horaires stricts.

Transport de marchandises : Les camionneurs effectuant de longs trajets pour livrer des marchandises dans des délais serrés.

Projets de construction : Les opérateurs d'engins lourds tels que les grues et les bulldozers, travaillant sur des chantiers de construction avec des échéances serrées.

Travaux miniers : Les conducteurs et les opérateurs d'équipements lourds dans les mines font face à des périodes de travail prolongées et à des conditions difficiles qui sont fréquentes.

2.1.4 Les systèmes de détection de la fatigue

Un système informatique automatisé basé sur l'intelligence artificielle nécessite une phase initiale d'apprentissage durant laquelle il acquiert la capacité, avec des niveaux de précision variables, de prédire les résultats attendus. Cette phase d'apprentissage est intimement liée à la qualité et à la diversité des données d'entrée utilisées.

Pour la détection de la fatigue au volant, plusieurs sources de données peuvent être exploitées tel que illustré par la figure 2.1. Ces modalités se divisent en deux catégories : des caractéristiques liées au véhicule et des caractéristiques liées au conducteur qui se divise à son tour selon la nature des données à deux sous-catégories : les mesures comportementales et les mesures physiologiques.

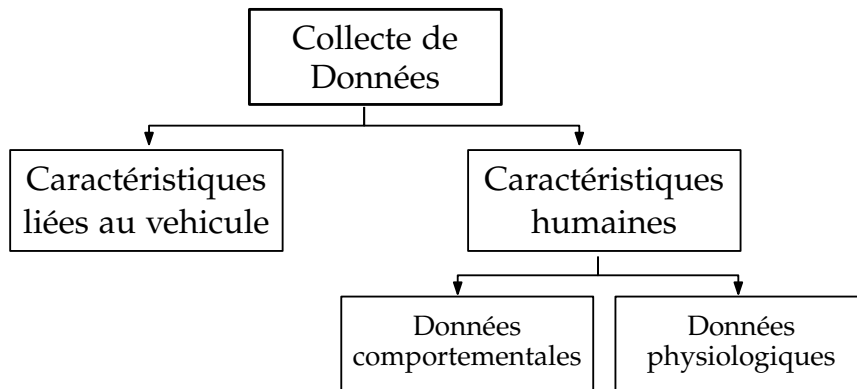


FIGURE 2.1 – La catégorisation des sources de données de la fatigue.

Caractéristiques centrées sur le conducteur

Mesures comportementales : L'analyse comportementale et plus spécifiquement l'analyse faciale se concentre sur la surveillance et l'interprétation des expressions faciales, de gestes corporels, de postures, ainsi que d'autres signaux non verbaux. L'activité oculaire, telle que la fréquence de clignement, la durée des clignotements et les mouvements des paupières présente un indicateur important de la fatigue [8]. De même pour les mouvements de la bouche, y compris le bâillement. Grâce aux avancées en matière de reconnaissance faciale et de traitement d'images, il est possible de capturer et d'analyser ces signes fournissant ainsi des prédictions sur l'état de vigilance du conducteur. Les outils utilisés incluent des caméras RGB (couleur), des caméras de profondeur pour mesurer la distance ainsi que des caméras infrarouges pour détecter la chaleur émise par le visage améliorant la précision de la détection des expressions faciales.

Mesures physiologiques : Les mesures physiologiques offrent une vision plus précise de l'état interne du conducteur [36]. L'analyse des signaux liés à l'activité des ondes cérébrales, du rythme cardiaque, de la réponse oculaire et de la réaction musculaire peut révéler des niveaux de fatigue qui ne sont pas immédiatement apparents à travers l'observation comportementale seule. Par exemple, les changements dans les ondes cérébrales peuvent indiquer des stades de somnolence ou de fatigue intense. De même, des variations dans le rythme cardiaque et la tension musculaire peuvent être des indicateurs de fatigue. Pour mesurer ces signaux physiologiques, des capteurs avancés comme l'électroencéphalographie pour les ondes cérébrales, l'électrocardiographie pour le rythme cardiaque, les caméras oculométriques pour les mouvements oculaires, l'électromyographie pour l'activité musculaire, et les capteurs de conductance de la peau (galvanomètres) pour la réponse sudorale sont utilisés.

Caractéristiques orientées véhicule

Les caractéristiques orientées véhicule analysent le mouvement du volant, la vitesse, le freinage et la position pour évaluer le comportement au volant de manière détaillée. En identifiant les mouvements inhabituels tels que les changements de voie brusques, les accélérations rapides et les irrégularités de la trajectoire, les systèmes basés sur ces données peuvent détecter les altérations potentielles dans le comportement du conducteur [56].

2.1.5 Les techniques de détection de la fatigue

Les systèmes de détection de la fatigue peuvent être réparties, selon les sources de données et les techniques utilisées, en trois catégories principales : les techniques basées sur le comportement du conducteur, les techniques basées sur les signaux physiologiques et les techniques basées sur les capteurs de le véhicule comme le montre la figure 2.2.

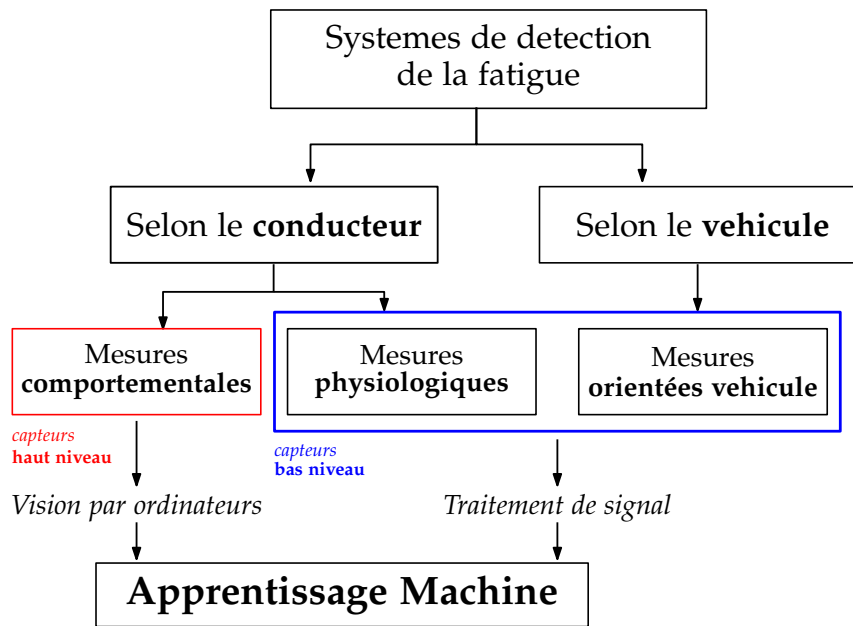


FIGURE 2.2 – Les différents techniques de détection de la fatigue.

La figure 2.3 illustre un exemple de capteurs utilisés dans un processus de collecte de données pour la détection de la fatigue.

Les techniques comportementales : utilisent le traitement d’image et la vision par ordinateur pour surveiller des indicateurs tels que le clignement et la fermeture des yeux, le bâillement, les expressions faciales et la position de la tête [30]. Les capteurs utilisés comprennent :

- Caméras RGB : capturer en couleur le visage et les yeux.
- Caméras infrarouges : suivre les mouvements des yeux et du visage dans des conditions de faible luminosité.
- Caméras de profondeur : mesurer les distances et les mouvements tridimensionnels du visage et de la tête.

Les techniques physiologiques : comprennent la surveillance des signaux tels que la température corporelle, la fréquence respiratoire, le rythme cardiaque et l’activité cérébrale électrique [17]. Les capteurs utilisés comprennent :

- Capteurs ECG : surveiller le rythme cardiaque.
- Capteurs EEG : mesurer l’activité cérébrale électrique.
- Capteurs EMG : détecter l’activité musculaire.

- Capteurs EOG : pour mesurer les mouvements oculaires.
- Capteurs de fréquence respiratoire : enregistrer les cycles de respiration.

Les techniques basées sur le véhicule : se servent de capteurs intégrés pour mesurer le mouvement du volant, la vitesse de rotation du volant, la position des mains, l'absence des mains et l'écart de la trajectoire [56]. Les capteurs utilisés comprennent :

- Capteurs de position du volant : mesurer l'angle et la vitesse de rotation du volant.
- Capteurs de pression : détecter la présence ou l'absence des mains sur le volant.
- Capteurs de trajectoire : surveiller l'écart par rapport à la voie de circulation.

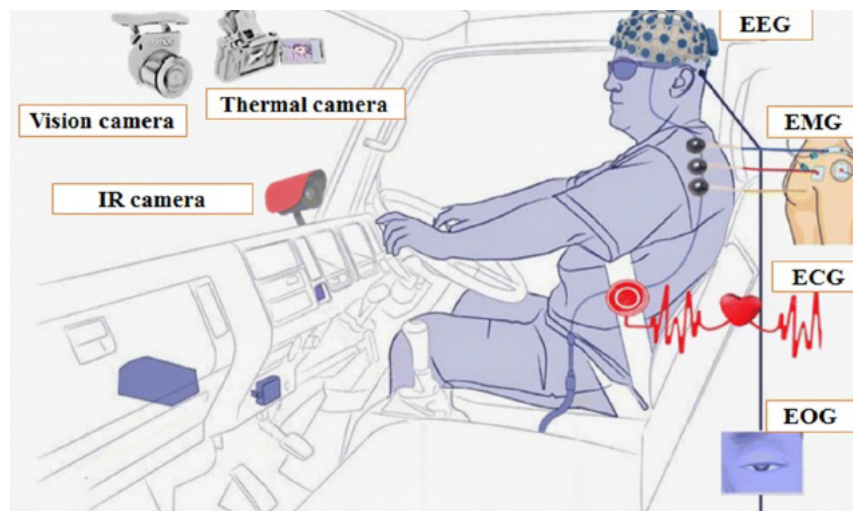


FIGURE 2.3 – Aperçu de la collecte de données à partir de divers capteurs [18].

2.2 Bases de données

Pour aborder efficacement le problème de la détection de la fatigue au volant, l'utilisation de bases de données bien structurées est essentielle. Ces bases de données, souvent issus de recherches antérieures et des études empiriques fournissent une base solide pour le développement et la validation des modèles de détection de fatigue.

2.2.1 National Tsing Hua University (NTHU-DDD) [58]

La base de données NTHU-DDD³ est reconnue pour ses riches données multimodales collectées dans des conditions de conduite simulées. Elle comprend des enregistrements vidéo du visage, des données physiologiques et des informations sur les performances de conduite. Les chercheurs l'ont utilisé pour entraîner des algorithmes à reconnaître les signes de fatigue grâce aux changements dans les expressions faciales et le comportement des yeux. Des exemples de données sont présentés dans la figure 2.4.

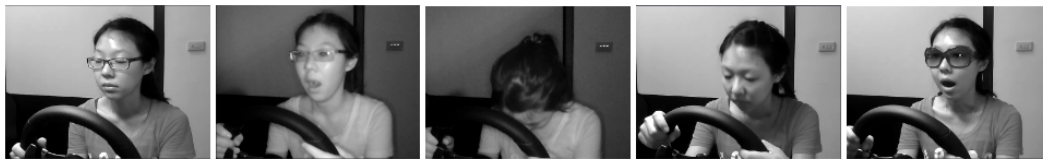


FIGURE 2.4 – Exemples de données NTHU-DDD [58].

2.2.2 Real-Life Drowsiness Dataset (UTA-RLDD) [23]

La base de données UTA-RLDD⁴ comprend des vidéos de participants exprimant divers degrés de somnolence dans des conditions de vie réelle. Elle a l'avantage de capturer des comportements spontanés, contrairement aux données recueillies dans un environnement contrôlé, ce qui aide les systèmes de détection de la fatigue à mieux généraliser dans des situations réelles. La figure 2.5 illustre quelques échantillons des données collectées.

3. <http://cv.cs.nthu.edu.tw/php/callforpaper/datasets/DDD/>

4. <https://sites.google.com/view/utarldd/home>



FIGURE 2.5 – Exemples de données UTA-RLDD [23].

2.2.3 ULg multimodality drowsiness database (DROZY) [39]

La base de données DROZY⁵ se distingue comme le seul, à notre connaissance, à contenir à la fois des données physiologiques et comportementales. Elle contient des enregistrements vidéo des visages des conducteurs comme le montre la figure 2.6, permettant l'analyse des expressions faciales, du clignement des yeux et du bâillement, ainsi que des données physiologiques telles que l'EEG, l'ECG, l'EMG et l'EOG.

5. <http://www.drozy.ulg.ac.be/>



FIGURE 2.6 – Exemples de données DROZY [39].

2.2.4 Yawning detection dataset (YawDD) [1]

La base de données YawDD⁶ est axée sur la détection du bâillement de la bouche, un indicateur comportementale clé de la fatigue, tel que montré dans la figure 2.7. Les séquences vidéo de YawDD permettent de développer des systèmes capables de détecter les signes précoces de somnolence chez les conducteurs.

6. <http://ieee-dataport.org/2561>

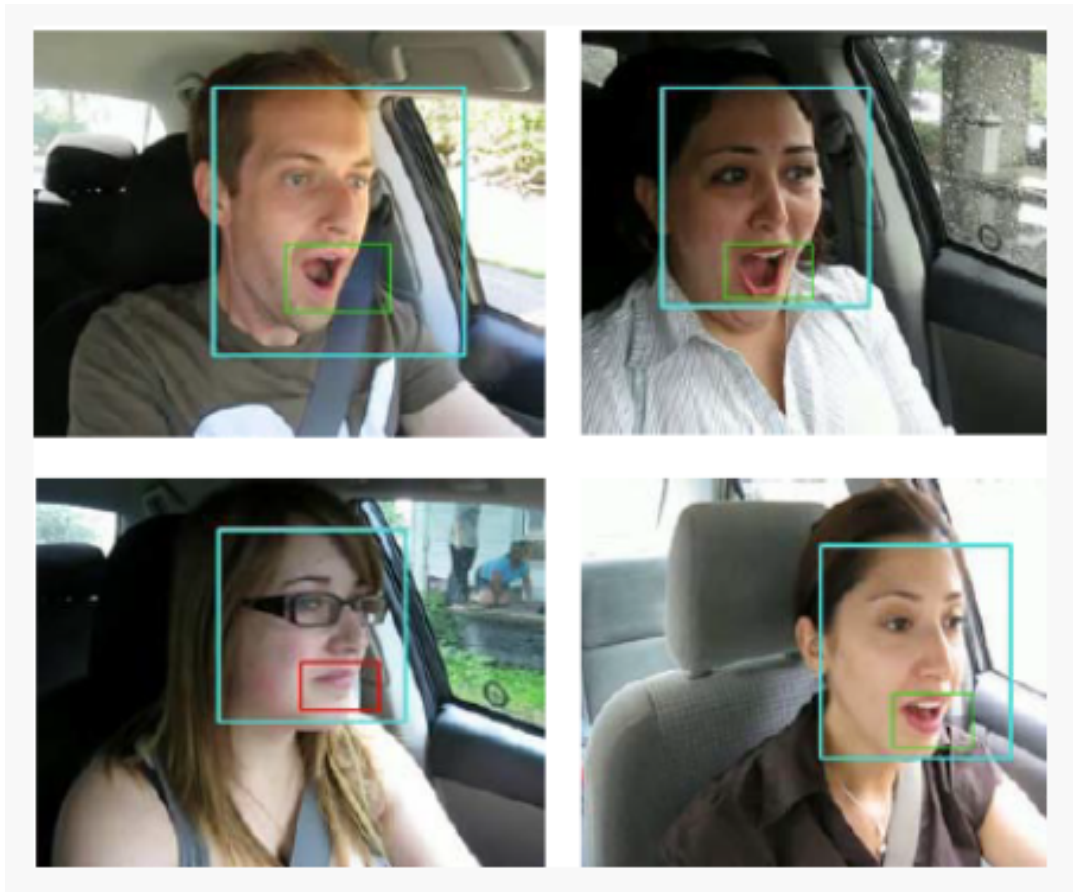


FIGURE 2.7 – Exemples de données YawDD [1].

2.2.5 MRL Eye dataset [22]

La base de données MRL Eye⁷, illustrée dans la figure 2.8, contient des enregistrements des mouvements oculaires, des clignements et des mesures de la pupille qui sont des indicateurs biométriques pouvant indiquer la fatigue du conducteur. Les signes de fatigue tels que la fréquence des clignements, la durée des clignements et les variations de la taille de la pupille peuvent être identifiés à partir de ces données.

7. <http://mrl.cs.vsb.cz/eyedataset>

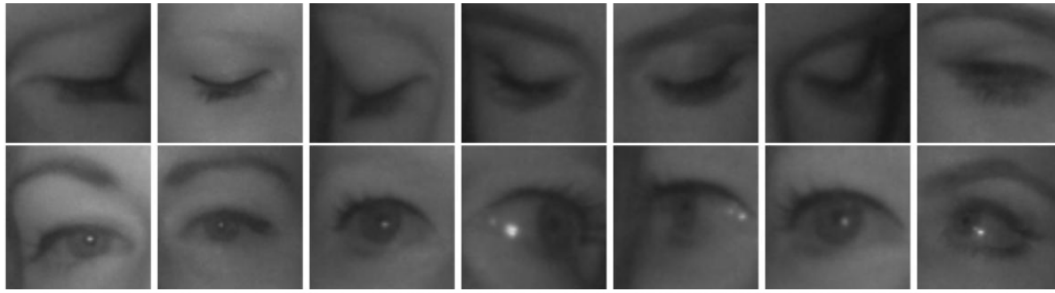


FIGURE 2.8 – Exemples de données MRL Eye [22].

2.2.6 Consented Activities of People (CAP) [57]

La base de données CAP⁸ est destinée pour l'analyse du sommeil et de ses troubles, utilisée indirectement dans la recherche sur la détection de la fatigue. Elle contient des données enregistrées pendant le sommeil, y compris des signaux EEG qui sont des indicateurs de la phase de sommeil et de la profondeur du sommeil à partir desquels on peut déduire les niveaux de fatigue.

2.2.7 Closed Eyes in the Wild dataset (CEW) [51]

La base de données CEW⁹ a été créée pour étudier la détection de la fermeture des yeux dans des conditions réelles non limitatives comprenant des variations dues à des différences individuelles ou bien à des changements environnementaux tels que l'éclairage, les clignotements, l'occlusion et le masquage. Ce sont des photos des patchs oculaires dérivées à partir de 2423 sujets.

2.2.8 Bases de données utilisées

Dans notre recherche, notre méthodologie se divise en deux systèmes de détection principaux exploitant respectivement les données comportementales et physiologiques comme le montre le Tableau 2.1.

8. <https://www.physionet.org/content/capslpdb/1.0.0/>

9. https://opendatalab.com/OpenDataLab/CEW_Closed_eyes_in_the_wild/

- Données comportementales : nous avons utilisé les bases de données UTA-RLDD et DROZY. Ils fournissent des vidéos enregistrées dans des conditions de conduite simulées mettant en évidence les comportements faciaux des conducteurs.
- Données physiologiques : nous avons utilisé la base de données DROZY. Elle comprend des enregistrements de signaux d'électrocardiogramme (ECG), d'électromyogramme (EMG), des deux canaux d'électrooculogramme (EOG) des cinq canaux d'électroencéphalogramme (EEG).

TABLE 2.1 – Les bases de données utilisées.

Catégorie	Nom	Participants	Scénarios	Échantillons	Type
Comportementale	UTA-RLDD	60	1	60	vidéos
	DROZY	14	3	36	images
Physiologique	DROZY	14	3	36	EEG, EOG, ECG et EMG

2.3 Méthodes et approches existantes

2.3.1 Détection centrée sur le véhicule

Dans la détection basée sur le véhicule, le mouvement du volant (SWM) et l'écart type de la position latérale (SDLP) sont souvent utilisés comme indicateurs de jugement en utilisant des capteurs placés sur différents composants du véhicule pour détecter la somnolence du conducteur [27].

Li et al. [34] ont calculé l'entropie approximative (ApEn) et l'écart-type de la SDLP et la complexité des séries temporelles à court terme de l'angle du volant en collectant des données provenant de capteurs montés sur la barre de direction du véhicule et ont établi un modèle de détermination de la fatigue à plusieurs niveaux.

CAI et al. [52] ont utilisé un réseau de contrôle de zone pour collecter des données sur l'état de fonctionnement d'un véhicule réel et ont extrait 18 caractéristiques liées au comportement de conduite, telles que la vitesse du véhicule, la vitesse de rotation, la pédale d'accélérateur, la pédale de frein et les informations sur les feux. Le coefficient de corrélation de Pearson a

été employé pour analyser la corrélation entre les caractéristiques extraites et la fatigue. Enfin, l'algorithme de la forêt aléatoire a été utilisé pour détecter l'état de fatigue du conducteur.

Toutefois, ces méthodes de détection sont limitées dans leur application pratique en raison des particularités du véhicule et de la route, ainsi que des différences entre les habitudes de conduite des différents conducteurs, de sorte que la fiabilité des approches de détection basées sur le véhicule est insuffisante.

2.3.2 Détection centrée sur le conducteur

Selon les caractéristiques physiologiques

Un article de revue récent par Albadawi et al. [4] met en avant plusieurs systèmes utilisant diverses mesures pour suivre et détecter la somnolence.

Dans l'article de Ye et al. [60], l'EMG, la pression de la main et les signaux de conductance cutanée (SC) sont mesurés de manière peu invasive à l'aide d'une plateforme de capteurs intégrée montée sur une manchette de volant. Ceci permet d'installer cette plateforme de capteurs dans n'importe quel véhicule existant, facilitant ainsi une surveillance étendue et pratique des conducteurs. Le signal SC peut être décomposé en deux composants principaux : *une composante tonique* qui varie lentement dans le temps, également connue sous le nom de niveau de conductance cutanée (SCL), et une *composante phasique* caractérisée par des changements rapides de l'amplitude du signal, également connue sous le nom de réponse de conductance cutanée (SCR), généralement associée à des événements liés à des stimuli.

La collecte de multiples signaux et biomarqueurs pour parvenir à une détection fiable de la somnolence a été exploitée dans la littérature. Par exemple, dans l'article de Awais et al. [7] ont proposé l'acquisition et le traitement combinés de signaux EEG et ECG pour discriminer entre les états alerte et somnolent, atteignant une acceptable précision avec le classificateur Machine à Vecteurs de Support (SVM).

Selon les caractéristiques comportementales

Dans l'article de De Lima Medeiros et al. [35], les auteurs ont introduit le clignement des yeux volontaire comme une Interaction Homme-Machine (IHM), en utilisant un outil de détection basé sur la vision par ordinateur avancé pour un traitement en temps réel avec une caméra générique. La détection du clignement des yeux a été réalisée en plus de l'identification de l'état des yeux, impliquant la reconnaissance faciale, la modélisation, l'extraction de la Région d'Intérêt (ROI) et la classification de l'état des yeux. Le système comprenait un correcteur de posture, un évaluateur de ROI et un filtre à moyenne mobile. Deux bases de données appelées YEC dataset [20] et ABD dataset¹⁰ ont été créés, atteignant une bonne précision pour les réseaux de neurones convolutifs (CNN) et les SVM sur divers ensembles de données.

Dans l'article de Safarov et al. [47], la détection de la somnolence basée sur le clignement des yeux a été mise en œuvre, analysant les motifs de clignement des yeux. Des données personnalisées ont été utilisées pour l'entraînement du modèle, et les fluctuations en temps réel des repères oculaires ont été mesurées à l'aide de méthodes d'apprentissage profond. L'analyse expérimentale a démontré une corrélation entre les bâillements et les yeux fermés, classés comme somnolents, avec une excellente performance globale pour la détection des yeux somnolents et la détection des bâillements.

Dans l'article de Hasan et al. [24], les auteurs ont initié leur travail sur la détection de la somnolence en détectant les bâillements. Ils ont proposé une méthode en trois étapes utilisant le détecteur de visage pour localiser le visage du conducteur et un modèle de Réseaux de Neurones Profonds (DNN) pour classifier le visage et détecter les bâillements. Le modèle DNN a été entraîné en utilisant un ensemble de données obtenu à partir du système développé Conduire en Sécurité.

Dans l'article de Rajkar et al. [44], une approche basée sur l'apprentissage profond a été présentée pour détecter la somnolence du conducteur à l'aide

10. <https://repositorio.ufrn.br/handle/123456789/53567>

de CNNs. Les yeux et les bâillements ont été utilisés pour détecter la somnolence en utilisant la base de données public YawDD.

2.4 Bilan et discussion

Afin de synthétiser les différentes études examinées dans cette revue de littérature, nous présentons ci-dessous un tableau récapitulatif pour chaque section, regroupant les travaux récents sélectionnés et mettant en évidence les auteurs, les méthodes utilisées, les bases de données exploitées et les taux de précision obtenus. Ce résumé nous permet de comparer les approches physiologiques et comportementales, et d'identifier les tendances actuelles et les lacunes de la recherche.

TABLE 2.2 – Bilan des travaux physiologiques.

Année	Auteur	Signaux	Méthode	Données	Taux d'exactitude
2020	Krishnan et al. [32]	EEG	KNN et SVM	DROZY	96.10%
2020	Yaacob et al. [59]	ECG et EEG	KNN et SVM		97.20%
2022	Zhao et al. [62]	ECG	CNN et BiLSTM		99.60%
2023	Perkins et al. [43]	EEG, ECG, EMG et EOG	ResNet-101 et LSTM		93.10%
2022	Alharbey et al. [5]	EEG	SVM		98.00%
2022	Ramos et al. [45]	EEG	Bagging et de vote		90.00%
2022	Hasan et al. [25]	EOG, EEG et ECG	KNN, SVM, RF et ANN	Collectées	83.50%
2021	Chaabene et al. [12]	EEG	CNN		90.42%
2021	Siddiqui et al. [50]	Taux de respiration	SVM, DT, RL, GBM, Extra Tree et MLP		87.00%
2023	Amidei et al. [6]	Conductance de la peau	RF, Bagging et Boosting		89.40%
2023	Peivandi et al. [42]	EEG	GAN et CNN		96.80%

Dans le tableau 2.2, nous présentons un bilan récapitulatif des travaux présentés pour la partie physiologique. Les principaux défis liés à l'usage des systèmes basés sur les données physiologiques concernent les caractéristiques des dispositifs employés et la configuration matérielle. Plusieurs dispositifs portables disponibles sur le marché ou sous forme de prototypes de recherche ont été utilisés, tels que des bracelets et des bagues intelligents. Cependant, la revue de littérature montre que les signaux collectés par des

dispositifs portables présentent des limitations en termes de qualité par rapport à ceux acquis par des équipements de laboratoire et des cartes d'acquisition de données. Les mouvements effectués lors des activités quotidiennes peuvent affecter le signal avec des artefacts de bruit. Par conséquent, des étapes de pré-traitement appropriées doivent être appliquées aux signaux pour les rendre adaptés à une classification fiable de la somnolence.

TABLE 2.3 – Bilan des travaux comportementaux.

Année	Auteur	Méthode	Données	Taux d'exactitude
2021	Suresh et al. [54]	Forward Deep-Learning CNN	CEW dataset et YawDD	86.05%
2021	Salman et al. [48]	Ensemble CNN	YawDD	97.00%
2022	Rajkar et al. [44]	CNN		96.00%
2022	De Lima Medeiros et al. [35]	CNN et SVM	YEC dataset	97.44%
2021	Adhinata et al. [2]	FaceNet et KNN	UTA-RLDD	94.68%
2021	Tamanani et al. [55]	LeNet		88.00%
2022	Krishna et al. [31]	YoloV5	UTA-RLDD + collectées	97.40%
2019	Yu et al. [61]	3D-DCNN	NTHU	76.20%
2021	Bakheet et al. [9]	HOG et Naïve Bayesian		85.62%
2017	Huynh et al. [28]	3D-CNN et GBM		87.46%
2020	Maior et al. [38]	SVM	DROZY	94.44%
2023	Safarov et al. [47]	CNN	Collectées	95.80%
2021	Hasan et al. [24]	DNN		95.20%
2019	Chirra et al. [15]	Deep-stacked CNN		96.42%

Le tableau 2.3 met en évidence les principaux travaux et résultats concernant la partie comportementale. Les défis liés à l'exploitation des données comportementales, en particulier celles issues de l'analyse vidéo du visage, des yeux et de la bouche, sont principalement dus aux variations et aux complexités de ces données tel que les conditions d'éclairage variables, les différents angles de caméra et les obstructions partielles du visage. De plus, les différences individuelles dans les expressions faciales et les comportements peuvent rendre difficile l'établissement de critères universels pour détecter la fatigue. La qualité et la résolution des vidéos affectent également l'exactitude

de l'analyse, nécessitant des algorithmes de traitement d'image et d'apprentissage machine très performants pour interpréter correctement ces indices subtils. Par conséquent, il est essentiel d'appliquer des techniques avancées de vision par ordinateur et d'apprentissage profond pour traiter et analyser efficacement ces types de données afin d'identifier de manière fiable les signes de fatigue du conducteur.

Chapitre 3

Titre :

Méthode basée sur les signaux physiologiques en utilisant des méthodes traditionnelles et hybrides

Résumé - La somnolence au volant est identifiée comme un facteur majeur contribuant aux accidents de la circulation et aux décès dans le monde entier. Pour répondre à cette préoccupation publique critique, les chercheurs développent divers systèmes de détection de la somnolence centrés sur le conducteur. Cependant, la plupart des études dans ce domaine s'appuient soit sur des techniques de surveillance visuelle, soit sur des indicateurs physiologiques. Cet article se concentre sur ce dernier point et présente une étude comparative visant à identifier la méthode la plus fiable et prometteuse pour détecter le niveau de somnolence à l'aide de signaux physiologiques. Cette étude explore diverses techniques utilisant des méthodes traditionnelles d'apprentissage automatique ou des réseaux neuronaux profonds pré-entraînés. Les premières se concentrent sur la génération de caractéristiques en exploitant l'aspect temporel des signaux à l'aide de séries temporelles et de représentations en fréquence par le biais de la Transformée de Fourier Rapide, de la Transformée en Cosinus Discrète et de la Transformée en Ondelettes Discrètes. Le second type utilise le spectrogramme, qui encapsule à la fois les informations temporelles et fréquentielles à travers une représentation visuelle, par la suite employée au sein d'un modèle de Réseau Neuronal Convolutif pré-entraîné pour l'extraction de caractéristiques. Les différentes approches explorées dans cette étude, utilisant le jeu de données ULg DROZY disponible publiquement, ont produit une multitude de résultats. Cependant, l'approche la plus prometteuse a atteint une précision maximale de 88 %.

Mots clés :

Détection de la somnolence, Signaux physiologiques, Spectrogramme, Transformée de Fourier Rapide, Transformée en Cosinus Discrète, Transformée en Ondelettes Discrètes, VGG-16, ResNet-50, PCA (Analyse en Composantes Principales).

Contributions associées :

D. Lamouchi, Y. Yaddaden, R. Cherif, "Comparative Study : Physiological-based Driver Drowsiness Detection Utilizing Traditional and Hybrid Methods," in *International Conference on Image and Signal Processing and their Applications (ISPA'2024)*, IEEE, 2024, pp. 1-8, DOI :10.1109/ISPA59904.2024.10536845. (Statut : Présenté)

Comparative Study: Physiological-based Driver Drowsiness Detection Utilizing Traditional and Hybrid Methods

Dorra Lamouchi
Département de mathématiques,
informatique et génie
Université du Québec à Rimouski
Rimouski, Québec, Canada.
dorra.lamouchi@uqar.ca

Yacine Yaddaden
Département de mathématiques,
informatique et génie
Université du Québec à Rimouski
Rimouski, Québec, Canada.
yacine_yaddaden@uqar.ca

Raef Cherif
Département de mathématiques,
informatique et génie
Université du Québec à Rimouski
Rimouski, Québec, Canada.
raef_cherif@uqar.ca

Abstract—Driver drowsiness is identified as a major factor leading to traffic accidents and fatalities worldwide. To address this critical public concern, researchers are developing various driver-centered drowsiness detection systems. However, most of the studies in this field rely on either visual-based monitoring techniques or physiological-based indicators. This paper focuses on the latter and presents a comparative study aimed at identifying the most reliable and promising method for detecting the level of drowsiness using physiological signals. This study explores various techniques employing traditional machine learning methods or pre-trained deep neural networks. The former focus on generating features leveraging the temporal aspect of the signals using time series and frequency representation through Fast Fourier Transform, Discrete Cosine Transform, and Discrete Wavelet Transform. The latter type utilizes the spectrogram, which encapsulates both temporal and frequency information through visual representation, subsequently employed within a pre-trained Convolutional Neural Network model for feature extraction. The different approaches explored in this study, using the publicly available ULg DROZY dataset, have produced a multitude of results. However, the most promising approach achieved a top accuracy of 88%.

Index Terms—Drowsiness detection, Physiological signals, Spectrogram, Fast Fourier Transform, Discrete Cosine Transform, Discrete Wavelet Transform, VGG-16, ResNet-50, PCA.

I. INTRODUCTION

Drowsiness [1] refers to the transitory status between alertness and sleep, marked by a decreased level of consciousness and generally associated with a diminution of natural body reflexes. According to the National Highway Traffic Safety Administration (NHTSA) [2], about 40% of traffic fatalities are due to falling asleep at the wheel. Every year around 100,000 accidents are caused by drowsy driving, leading to 2,000 deaths and 70,000 injuries. Hence, it is essential to develop a drowsiness monitoring system in order to enhance human safety.

Drowsiness can impair cognitive functions, creating difficulties for tasks demanding heightened focus and rapid reactions. This situation poses devastating consequences in several fields, including transport particularly on extended road trips or late-night travels. Furthermore, Driver Drowsiness Detection

(DDD) systems are mainly available in newer, often more expensive vehicle models, making them inaccessible to a wide consumer audience. Aiming to ensure the safety and well-being of all drivers, researches are still being carried out.

To address this issue, several types of drowsiness monitoring systems are used to detect and evaluate drowsiness levels within individuals [3]. A commonly used classification for drowsiness detection indicators involves categorizing them into two primary sources of information: *human-centered* and *vehicle-centered* [1].

Once signs of drowsiness are detected, various warning mechanisms to alert the driver and prevent potential accidents can immediately be initiated. These methods are specifically designed to engage the driver's focus and encourage him to react appropriately when they are in a drowsy state.

In this paper, we present a comparative study where various DDD systems using a variety of classical and hybrid approaches are rigorously tested. We aimed to conduct a comparison among *frequency* domain-based, *time* domain-based, and *time-frequency* domain-based representations to determine their respective strengths and weaknesses through the application of Machine Learning (ML) techniques, including *deep learning neural networks*. We opted for physiological-based signals such as the *electroencephalogram* (EEG), the *electrocardiogram* (ECG), the *electromyogram* (EMG) and the *electrooculogram* (EOG).

The remaining sections of this paper are organized as follows. First, Section II outlines the key concepts essential for our work. Previous methods and approaches are discussed in Section III. Sections IV and V detail the used feature extraction and classification techniques, respectively. Subsequently, the proposed process for drowsiness detection is described in Section VI. Section VII defines the experimental protocol and introduces the public dataset used for evaluation. Results and discussions are presented in Section VIII. Finally, Section IX offers concluding remarks and perspectives

II. FUNDAMENTALS

A **DDD** system aims to achieve similar objectives as Advanced Driver Assistance Systems (**ADAS**) by leveraging the latest technological advancements to enhance road safety. However, there are notable distinctions between them. The **DDD** system primarily focuses on monitoring the driver's behavior and state, whereas **ADAS** focuses on the vehicle itself, utilizing various sensors to gather data related to its surroundings and performance.

A driver's drowsiness detection system shares components with a typical pattern recognition system: 1) the *input* varies depending on the chosen modality for detection, 2) the *pre-processing* operations may enhance the input data, 3) the *feature extraction* component generates a relevant representation of the input, with the type of feature depending on the chosen modality, 3) the *dimensionality reduction* may then be employed to decrease the size of the feature vector and eliminate redundant information, 4) the *classification* typically involves supervised **ML** techniques, 5) the *output* may be *binary* (*drowsy* or *alert*) and may also indicate the level of drowsiness based on the Karolinska Sleepiness Scale (**KSS**).

Furthermore, two primary categories of data sources are typically utilized for detecting levels of drowsiness, as detailed below.

A. Human-centered features

They encompass all the features that could be collected and measured from the driver. There are two main sub-categories:

1) *Behavioral-based*: They are related to *body movements* and *facial expressions* [4], [5]. Some vehicles are already equipped with systems utilizing *facial analysis* and *eye-tracking* technology to detect signs of drowsiness or distraction. However, it is important to note that these methods suffer from limitations such as occlusions, where objects may obstruct the camera's view of the driver's face [6]. Additionally, the reliability of the gathered measurements can be affected by factors such as fluctuating illumination and the driver's seating posture.

2) *Physiological-based*: They involve the use of several *physiological* sensors to measure signals such as **ECG**, **EEG**, **EMG**, and **EOG** [7]. These signals are associated with *brain activity*, *heart rate patterns*, *muscle response*, and *eye movement*, respectively. They provide dynamic indications of the driver's state under natural driving conditions. Despite their high accuracy, these methods can be intrusive and impractical for driving use. The requirement for sensors and physical contact with the body makes implementation challenging in real-world scenarios where comfort and convenience are essential.

B. Vehicle-centered measures

They focus on various measures related to the vehicle, such as steering wheel movements, driving speed, braking, and trajectory deviations. By continuously monitoring vehicle movement patterns over time, it becomes possible, using **ML** techniques, to detect abnormal behaviors such as sudden

lane changes, rapid acceleration or deceleration, and other irregularities that may be caused by the driver's drowsiness or distraction [8]. While these methods can be highly effective, they are often restricted to high-end models due to their high costs, rendering them inaccessible to the average consumer.

III. RELATED WORKS

In recent years, numerous methods have been proposed to detect the level of driver fatigue or drowsiness utilizing a variety of sensors. Below, we present and discuss some of the *state-of-the-art* methods. Drawing upon existing literature, we will explore two common types of **DDD** systems:

A. Unimodal DDD systems

They focus on utilizing a single *modality* or type of data to detect and assess the driver's drowsiness level. For instance, a system may exclusively rely on physiological signals to determine the level of sleepiness.

Chaabene et al. [9] introduced a **DDD** system utilizing **EEG** signals, favoring single signal processing. Their methodology relies on a Convolutional Neural Network (**CNN**) [5] architecture designed to a *binary* classification (*drowsy* or *alert*). Krishnan et al. [10] employed spectral and band power properties extracted from **EEG** signals to identify drowsiness. In this study, the authors used two different classifiers, namely the *k*-Nearest Neighbors (*k*-**NN**) and Support Vector Machine (**SVM**). To evaluate the performance of the proposed **DDD** system, they relied on a public dataset named ULg Multimodality Drowsiness (**DROZY**) [11]. Ramos et al. [12] proposed an **EEG**-based system employing ensemble learning. They developed and validated both a *bagging* model with either five or three **EEG** channels and a *voting* model with a single channel across various scenarios. The performance of the bagging model was found to be comparable to that of classifiers utilizing a one-channel approach, suggesting that the utilization of multiple **EEG** channels may not necessarily enhance performance. Wang et al [13] utilized the Modified Multi-Scale Entropy (**MMSE**) approach on the driver's **EEG** signals. The signals were decomposed, and the optimal scaling factors were determined. This study was tested under both normal driving conditions and auditory stimulation, enabling effective classification of the driver's condition. Rundo et al. [14] developed a pipeline capable of detecting the level of drowsiness from a driver's **EEG** signal. The algorithm is based on the Discrete Cosine Transform (**DCT**), followed by a deep neural network for classification. To recognize both driver drowsiness and stress, Chui et al. [15] focused on the **ECG** signal for their study. They designed a generic model using the Multiple-Objective Genetic Algorithm (**MOGA**) and optimized Deep Multiple Kernel Learning **SVM** (**D-MKL-SVM**) to perform this task. The research conducted by Siddiqui et al. [16] primarily relied on respiratory rate data, favoring non-invasive and non-camera-based approaches to develop a system for **DDD**. The collected data were used to train and validate various traditional **ML** models, among which the **SVM** demonstrated higher accuracy compared to

the other models. Picot et al. [17] proposed a method based on the fuzzy fusion of multiple blinking characteristics measured each second over a 20-second sliding window. They were able to achieve high performance, according to features extracted from **EOG** signal. Their work shown that the use of fuzzy logic enables good temporal granularity of features, making drowsiness detection more reliable.

B. Multimodal DDD systems

They combine multiple modalities or data sources to enhance performance. For instance, a *multimodal* system may integrate *physiological* measurements with *behavioral* indicators.

Budak et al. [18] utilized both raw **EEG** signals and their corresponding *spectrograms* in their study. The system comprises three blocks. The first involves calculating the *energy distribution* and extracting *spectral entropy* from the *spectrogram* images. The second stage employs deep feature extraction directly from **EEG spectrogram** images using **AlexNet** and **VGGNet**. The third unit utilizes the Tunable Q-factor Wavelet Transform (**TQWT**) to decompose **EEG** signals into sub-bands and derive statistical features such as the mean and standard deviation of the sub-bands frequencies. These features are then fed to a Long-Short Term Memory (**LSTM**) network for classification. The results are combined via a majority voting layer. The study presented by Yaacob et al. [3] combined the **EEG** and **ECG** signals to enhance detection accuracy. Signal pre-processing techniques were applied to remove noise and artifacts from the raw data. Feature extraction methods were then used to obtain relevant information capturing key patterns and characteristics associated with drowsiness. The best result is obtained with the use of *k*-NN classifier. Zhao et al. [19] employed a *multimodal* approach utilizing both **ECG** and video features via hybrid characteristic selection. They explored both traditional **ML** and Deep Learning (**DL**) techniques, and the system was evaluated using the **DROZY** dataset as well as a collected one. Gwak et al. [20] relied on **ML** algorithms, including logistic regression, **SVM**, *k*-NN classifier, and Random Forest (**RF**), to classify alert and drowsy states of drivers using driving performance, behavioral, and physiological features. The results shown that the **RF** method is the most reliant.

IV. FEATURE EXTRACTION TECHNIQUES

A. Time domain features

Statistical properties play an essential role in obtaining essential information about signals. They include several metrics such as the *maximum*, *minimum*, *mean*, *kurtosis*, *skewness*, *variance* and *standard deviation*. The *variance* refers to the dispersion of data points around the average, *standard deviation* reflects the distributional extent of the data, showing the closeness of values within a range, *skewness* shows the distribution's asymmetry, identifying its inclination towards higher or lower levels and *kurtosis* enables the assessment of the identification of potential outliers.

B. Frequency domain features

Various techniques might be used to extract meaningful information related the frequency content of signals.

1) *Discrete Cosine Transform (DCT)*: is a mathematical technique commonly used in signal processing. It transforms a signal into a sum of cosine functions with varying frequencies and amplitudes (as defined in the Eq. 1).

$$X(k) = \sum_{n=0}^{N-1} x(n) \cdot \cos \left[\frac{\pi}{N} \left(n + \frac{1}{2} \right) k \right] \quad (1)$$

where $x(n)$ refers to the input discrete signal, $X(k)$ the frequency domain representation, N the number of samples in the signal, k the frequency index, and n the time index.

2) *Fast Fourier Transform (FFT)*: is a well-known method for converting a signal from its time-based representation into its frequency-based one (as defined in the Eq. 2). Next, Power Spectral Density (**PSD**) provides the distribution of signal power in different frequency ranges, revealing the dominant frequencies and corresponding energy levels. The *spectral centroid* recognizes the center of mass of the frequency distribution, indicating where most of the energy is concentrated. Additionally, *spectral entropy* determines both the complexity and uniformity of the signal's frequency distribution.

$$X(n) = \sum_{k=0}^{N-1} x(k) e^{-j \frac{2\pi}{N} kn} \quad (2)$$

where $x(n)$ refers to the input discrete signal, $X(n)$ the frequency domain representation, N the number of samples in the signal, k and n are indices corresponding to time and frequency domains respectively.

3) *Discrete Wavelet Transform (DWT)*: is a signal processing technique that enables detailed time and frequency analysis. It is particularly useful for analyzing signals with rapidly changing frequencies and transient properties. The **DWT** decomposes a signal into a set of *wavelet coefficients* representing different frequency components at various scales. This decomposition allows for the extraction of both *low-frequency* approximation (see Eq. 3) and *high-frequency* details (see Eq. 4) from the original signal.

$$c_j[n] = \sum_k h[k] \cdot x[2n - k] \quad (3)$$

$$d_j[n] = \sum_k g[k] \cdot x[2n - k] \quad (4)$$

where $c_j[n]$ refers to the approximation coefficient at level j , $d_j[n]$ the detail coefficient at level j , $h[k]$ the low-pass filter coefficients, $g[k]$ the high-pass filter coefficients and $x(n)$ the input discrete signal.

C. Time-Frequency domain features

This type of feature is extracted through different steps as shown in Fig. 1.

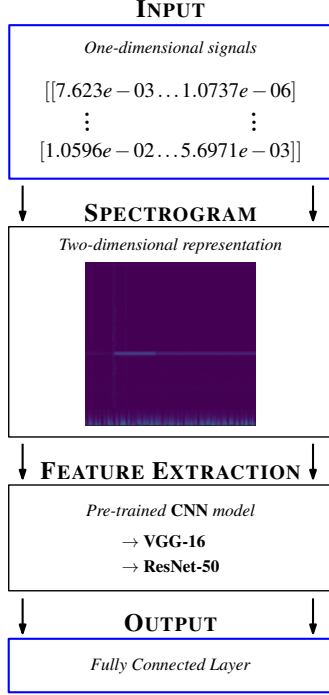


Fig. 1: The *time-frequency* domain feature extraction pipeline.

1) *Generation of the spectrogram*: Converting a one-dimensional signal into a *spectrogram* represents a way to visually display its frequency content over time. The process involves decomposing the signal into overlapping short segments, followed by applying a mathematical operation known as the Short-Time Fourier Transform (STFT) to each segment. As a result, the *spectrogram* represents frequency on the vertical axis, time on the horizontal axis, and color intensity reflecting the magnitude of the frequency components.

2) *DL-based feature extraction*: The *spectrogram* is then fed to a *pre-trained CNN* model in order to generate a feature vector that encapsulates most of the meaningful information. For that, we use two distinct *models*:

- The **VGG-16**, or the Visual Geometry Group-16 architecture [21], well-known for its image recognition capabilities, is used for feature extraction from the spectrogram by keeping the last fully connected layer as a feature vector.
- The **ResNet-50** is a variant of the Residual Network, publicly available and trained on the **ImageNet** dataset [22]. In this case, it is used for feature extraction from the spectrogram.

V. CLASSIFICATION TECHNIQUES

A. Decision Tree (DT)

is renowned for its partitioning recursive approach. It constructs a *tree structure* where each node represents a feature, and each branch represents a feature-based decision. This hierarchical structure is invaluable for discerning feature hierarchies and their significance in classification tasks. **DTs**

are widely utilized due to their interpretability, scalability, and ability to handle both numerical and categorical data. They serve as a foundational component in various **ML** algorithms and are particularly useful for tasks requiring transparent decision-making processes.

B. Support Vector Machine (SVM)

is a notably efficient algorithm for handling well-separated data points. It seeks to find the *optimal hyperplane* that maximizes the margin between different classes within the data (see Eq. 5). **SVM** is versatile, capable of handling both binary and *multi-class* classification tasks. It is particularly effective in scenarios where the data is linearly separable or can be transformed into separable feature space using *kernel tricks*. **SVMs** are known for their robustness, scalability, and ability to generalize well to unseen data, making them a popular choice across various domains in machine learning and data mining.

$$f(\mathbf{x}) = \mathbf{w} \cdot \mathbf{x} + b \quad (5)$$

where $\mathbf{x}$ represents the input feature vector, $\mathbf{w}$ the weight vector and b the bias term.

C. Random Forest (RF)

is an *ensemble* learning technique that combines multiple decision trees. It is particularly effective when dealing with high-dimensional data, as it provides insights into the interactions between different variables. By aggregating the predictions of individual decision trees, **RF** can achieve robust and accurate classification or regression results. Additionally, it offers built-in mechanisms for assessing feature importance, making it valuable for feature selection and interpretation tasks. **RF** is widely used in various fields, including bioinformatics, finance, and remote sensing, owing to its flexibility, scalability, and ability to handle complex datasets.

D. Multilayer Perceptron (MLP)

is a neural network architecture designed to analyze complex patterns in data. It consists of multiple layers of interconnected nodes, allowing it to capture and learn non-linear relationships within the data (see Eq. 6). **MLPs** are versatile models capable of handling a wide range of tasks, including classification, regression, and pattern recognition. By leveraging hidden layers and activation functions, **MLPs** can effectively model intricate relationships in the data, making them a popular choice in various fields such as image recognition, natural language processing, and financial forecasting.

$$f(\mathbf{x}) = \arg \max_k \sigma(\mathbf{W}^{(L)} \sigma(\mathbf{W}^{(L-1)} \dots \sigma(\mathbf{W}^{(1)} \mathbf{x} + \mathbf{b}^{(1)}) + \mathbf{b}^{(2)}) \dots + \mathbf{b}^{(L)}) \quad (6)$$

where $\mathbf{x}$ represents the input feature vector, σ the activation function applied element-wise, $\mathbf{W}^{(l)}$ the weights of l -th layer, $\mathbf{b}^{(l)}$ the biases of l -th layer and L the total number of layers. The decision is made by selecting the class k with the highest output of the final layer.

E. Quadratic Discriminant Analysis (QDA)

is a statistical technique used to model variations in class distributions. It assumes that each class follows a *Gaussian distribution* and computes quadratic decision boundaries to distinguish between classes (see Eq. 7). Unlike Linear Discriminant Analysis (**LDA**), which assumes equal co-variance matrices across classes, **QDA** allows for different co-variance matrices for each class, making it more flexible in capturing complex decision boundaries. **QDA** is particularly useful when the assumptions of **LDA** are not met, such as when the data has non-linear relationships or unequal variances between classes. It is commonly employed in classification tasks where the underlying distribution of the data is not well-understood or when the data violates the assumptions of simpler models.

$$f(\mathbf{x}) = \arg \max_k p(y = k | \mathbf{x}) \quad (7)$$

where $p(y = k | \mathbf{x})$ represents the probability of the input vector $\mathbf{x}$ belonging to class k , computed using Bayes' theorem.

VI. PROPOSED SYSTEM

The main objective of this work lies in comparing the performance of different descriptor combined with different classifiers as shown in Fig. 2. Each component of the pipeline is detailed in the following.

A. Pre-processing

In the context of this work, the preprocessing consists of extracting the signals collected using various sensors from the .edf format to the .csv format file since it is easier to manipulate. An overview of the extracted signals is displayed in Fig. 3, with four different sensors collecting data from the same subject through two different states (*drowsy* and *alert*).

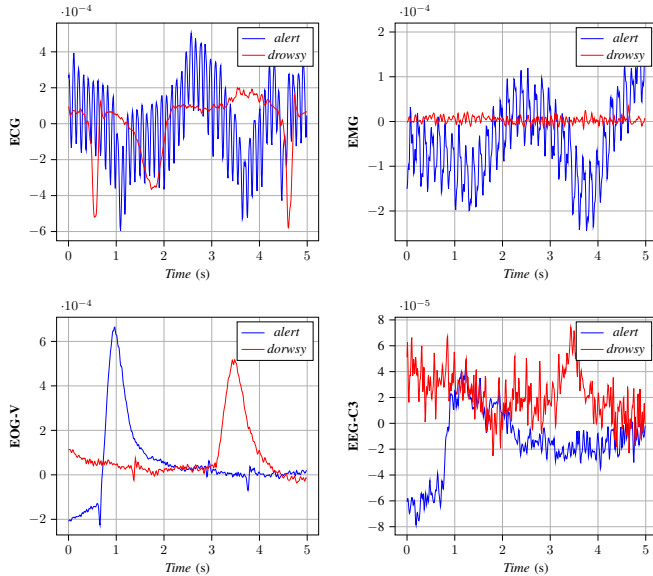


Fig. 3: Overview of signals collected from various *sensors*.

B. Feature extraction

Raw physiological signal data are not used directly with **ML** techniques since they usually have varying lengths and include noise. This is why feature extraction is a crucial step. In our case, we used different representations in different domains as shown in Table I.

TABLE I: LIST OF THE USED *feature extraction* TECHNIQUES.

CATEGORY	TECHNIQUE
Time domain	Statistical metrics
Frequency domain	FFT
	DCT
	DWT
Time-frequency domain	Spectrogram + VGG-16
	Spectrogram + ResNet-50

C. Dimensionality reduction

Principal Component Analysis (**PCA**) [23] is a data analysis technique used to reduce dimensionality. It calculates linear combinations of the input variables that are orthogonal to each other. The first principal component captures the largest variance in the data, the second captures the second-largest variance, and so on.

By reducing the dimensionality of the data using **PCA**, the complexity of the model would be decreased while retaining much of the information contained in the original data. It can improve model performance by reducing the risk of overfitting and speeding up training time.

D. Classification

In this study, we employed various *supervised ML* techniques. Classification serves as the cornerstone of our **DDD** system, as it enables the system to recognize the driver's state following a *training phase* where the model is trained on *labeled* data. Subsequently, the trained model is capable of classifying *unlabeled* data. To identify the most effective models, we *empirically* evaluated numerous *supervised ML* techniques, including **DT**, **SVM**, **MLP**, **QDA**, among others (refer to Section V for further details). Model performance evaluation is conducted during the *testing phase*.

VII. EVALUATION PROTOCOL

A. Data Description

The ULg Multimodality Drowsiness dataset (**DROZY**) [11] contains various *physiological* signals, including **EEG** signals recorded through the FZ, PZ, CZ, C3, and C4 channels (using the international 10-20 system), both *vertical* and *horizontal* **EOG** signals, as well as **EMG** and **ECG** signals. The dataset was collected from 14 healthy male and female subjects using a sampling frequency of 512 Hz. Three trials were conducted for each subject, each lasting 10 minutes.

The labeling was performed using the Karolinska Sleepiness Scale (**KSS**) [24], a subjective measure designed to assess an individual's level of drowsiness, as described in Table II.

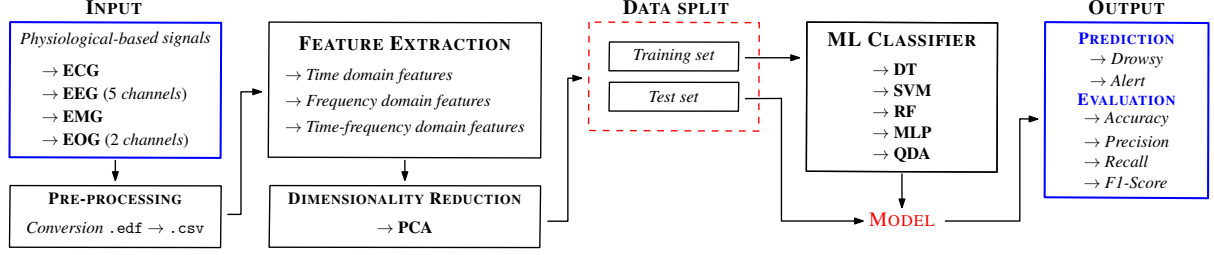


Fig. 2: Overview of the proposed DDD system.

TABLE II: THE KAROLINSKA SLEEPINESS SCALE (KSS).

STATE	SCORE	DESCRIPTION
Alert	1	Extremely alert
	2	Very alert
	3	Alert
	4	Rather alert
	5	Neither alert nor sleepy
Drowsy	6	Some signs of sleepiness
	7	Sleepy, but no effort to keep awake
	8	Sleepy, some effort to keep awake
	9	Very sleepy, great effort to keep awake
	10	Extremely sleepy, fighting sleep

B. Evaluation metrics

The performance of each approach was evaluated using three different metrics, namely the accuracy, precision, F1-score and recall. Moreover, the confusion matrix was also used to see the recognition rate of each of two driver's state.

The Eq. 8 defines the *accuracy*, where **TP**, **FN**, **FP**, and **TN** refer respectively to True Positives, False Negatives, False Positives, and True Negatives. The *precision* (see Eq. 9) measures the proportion of instances correctly identified as positive among all positive instances. The *recall* (see Eq. 10) quantifies the proportion of **TP** instances rated as positive by the model. The *F1-Score* (refer to Eq. 11) is particularly used when dealing with imbalanced datasets.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (8)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (9)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (10)$$

$$\text{F1-Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (11)$$

C. Validation Strategy

For this work, we employed both *stratified 5-fold cross-validation* and the *leave-one-out* strategy to assess the performance of our *models*. While the *leave-one-out* strategy is renowned for its precision, we encountered practical limitations when applying it with the hybrid approach (*spectrogram* + *pre-trained CNN models*), as it would have been considerably time-consuming and resource-intensive. Therefore, we opted for *stratified 5-fold cross-validation*.

VIII. RESULTS & DISCUSSION

Our experimentation involved testing various combinations of feature extraction and classification techniques, resulting in diverse outcomes that can help identify the most optimal combination. Table III presents the results obtained from different approaches, ranging from the Time Series (TS) analysis using statistical metrics to utilizing *pre-trained* deep neural networks for feature extraction. It is observed that the highest accuracy is achieved by combining the **FFT** and **DT** classifier reaching an accuracy of 88%. Following closely is the combination of **VGG-16** for feature extraction with a **DT** classifier.

TABLE III: OVERVIEW OF THE CLASSIFICATION RESULTS.

METHOD	CLASSIFIER	ACCURACY	PRECISION	RECALL	F1-SCORE
TS	DT	66.0%	65.6%	64.0%	64.0%
	SVM	55.0%	56.6%	54.0%	55.0%
	RF	61.0%	61.2%	61.0%	60.0%
	MLP	52.0%	51.8%	51.0%	51.0%
FFT	DT	88.0%	78.1%	88.0%	86.0%
	SVM	75.0%	72.3%	76.0%	75.0%
	RF	69.0%	68.8%	70.0%	70.0%
	MLP	66.0%	65.9%	67.0%	66.0%
DCT	DT	44.0%	47.2%	44.0%	45.0%
	SVM	56.0%	52.3%	56.0%	40.0%
	RF	64.0%	62.5%	64.0%	63.0%
	MLP	53.0%	48.1%	53.0%	42.0%
DWT	DT	28.0%	26.8%	28.0%	27.0%
	SVM	53.0%	48.2%	53.0%	46.0%
	RF	53.0%	52.7%	53.0%	52.0%
	MLP	61.0%	61.2%	61.0%	61.0%
VGG-16	DT	75.0%	76.9%	76.0%	75.0%
	SVM	63.0%	61.8%	63.0%	62.0%
	RF	58.0%	58.1%	59.0%	59.0%
	QDA	55.0%	56.4%	54.0%	54.0%
ResNet-50	DT	58.0%	59.1%	59.0%	58.0%
	SVM	58.0%	58.3%	59.0%	58.0%
	RF	63.0%	62.8%	64.0%	64.0%
	QDA	73.0%	72.5%	74.0%	74.0%

Considering only the most efficient classifiers along with the different method, the Fig. 4 presents the results comparing *traditional ML* approaches with *hybrid* ones based on **DL** techniques. It is observed that, except for the combination of **FFT** and **DT**, the *hybrid* approaches leveraging **DL** techniques generally outperform and achieve superior performance.

We also present the different *confusion matrices* allowing us to see the recognition rate for each driver's state (*drowsy* or *alert*) for each method, as shown in Table III. We notice that even though the combination of **FFT** and **DT** outperforms the other methods when detecting the drowsy state (with 100.0% accuracy), it is not the most efficient when detecting the alert state (with 56.3% accuracy) as shown in Table III (B). On the other hand, we observe that the combination of **VGG-16** and

DT improves the performance when recognizing the alert state by 12.5%, as shown in Table III (E). Therefore, in terms of balance, the combination of **VGG-16** and **DT** is more suited.

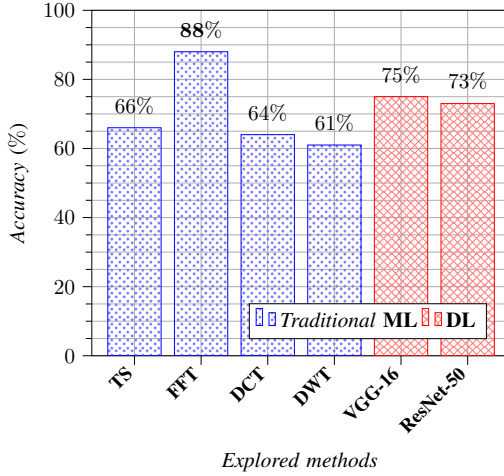


Fig. 4: Comparison of the explored methods.

TABLE IV: CONFUSION MATRIX (%) OF (A) **TS**-based, (B) **FFT**-based, (C) **DCT**-based, (D) **DWT**-based, (E) *Spectrogram* + **VGG-16**, AND (F) *Spectrogram* + **RESNET-50**.

(A)			(B)		
	alert	drowsy		alert	drowsy
alert	56.3	43.7	alert	56.3	43.7
drowsy	25.0	75.0	drowsy	0.00	100.0

(C)			(D)		
	alert	drowsy		alert	drowsy
alert	50.0	50.0	alert	62.5	37.5
drowsy	25.0	75.0	drowsy	40.0	60.0

(E)			(F)		
	alert	drowsy		alert	drowsy
alert	68.8	31.2	alert	50.0	50.0
drowsy	15.0	85.0	drowsy	5.0	95.0

Table V presents a comparison between our proposed system and *state-of-the-art* methods using *physiological*-based signal datasets. While many rely on datasets collected for their specific projects, our study, along with a few others, stands out for utilizing the publicly accessible **DROZY** dataset. Even though our proposed method does not outperform the existing ones, it still achieves *state-of-the-art* performance. Additionally, we were able to conduct a thorough comparison between the most commonly used descriptors for *physiologically*-based signals.

IX. CONCLUDING REMARKS

In the context of this work, we explored reliable techniques to perform **DDD** using various physiological markers, notably **ECG**, **EEG**, **EOG**, and **EMG** signals. A comparative analysis

based on three different domain feature representations was conducted: *time*, *frequency*, and *time-frequency*. For the latter, we leveraged the power of a *pre-trained CNN* for feature extraction. By comparing the different approaches, we observed that the best performances were achieved using the **FFT** along with a **DT** classifier.

As a future perspective, we are planning to extend this work to include an implementation for real-life deployments. By developing non-intrusive, portable devices suitable for use in actual situations, the benefits of driver sleepiness sensing technology may reach a broader audience.

ACKNOWLEDGMENT

We extend our sincere gratitude to the **ULg-INTELSIG** laboratory, responsible for the collection of the public **DROZY** dataset [11] used in our research.

REFERENCES

- [1] G. Zhang, K. K. Yau, X. Zhang, and Y. Li, "Traffic accidents involving fatigue driving and their extent of casualties," *Accident Analysis & Prevention*, vol. 87, pp. 34–42, 2016.
- [2] T. Stewart, "Overview of motor vehicle traffic crashes in 2021," Tech. Rep., 2023.
- [3] S. Yaacob, N. A. I. Affandi, P. Krishnan, A. Rasyadan, M. Yaakop, and F. Mohamed, "Drowsiness detection using eeg and ecg signals," in *2020 IEEE 2nd International Conference on Artificial Intelligence in Engineering and Technology (IICAET)*. IEEE, 2020, pp. 1–5.
- [4] Y. Yaddaden, "An efficient facial expression recognition system with appearance-based fused descriptors," *Intelligent Systems with Applications*, vol. 17, p. 200166, 2023.
- [5] Y. Yaddaden, M. Adda, A. Bouzouane, S. Gaboury, and B. Bouchard, "User action and facial expression recognition for error detection system in an ambient assisted environment," *Expert Systems with Applications*, vol. 112, pp. 173–189, 2018.
- [6] S. Rostaminia, A. Mayberry, D. Ganesan, B. Marlin, and J. Gummeson, "Iliid: low-power sensing of fatigue and drowsiness measures on a computational eyeglass," *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies*, vol. 1, no. 2, pp. 1–26, 2017.
- [7] M. Awais, N. Badruddin, and M. Drieberg, "A hybrid approach to detect driver drowsiness utilizing physiological signals to improve system performance and wearability," *Sensors*, vol. 17, no. 9, p. 1991, 2017.
- [8] D. Sałapatek, J. Dybała, P. Czapski, and P. Skalski, "Driver drowsiness detection systems," *Zeszyty Naukowe Instytutu Pojazdów/Politechnika Warszawska*, vol. 3, no. 112, pp. 41–48, 2017.
- [9] S. Chaabene, B. Bouaziz, A. Boudaya, A. Hökelmann, A. Ammar, and L. Chaari, "Convolutional neural network for drowsiness detection using eeg signals," *Sensors*, vol. 21, no. 5, p. 1734, 2021.
- [10] P. Krishnan, S. Yaacob, A. P. Krishnan, M. Rizon, and C. K. Ang, "Eeg based drowsiness detection using relative band power and short-time fourier transform," *J. Robotics Netw. Artif. Life*, vol. 7, no. 3, pp. 147–151, 2020.
- [11] Q. Massoz, T. Langohr, C. François, and J. G. Verly, "The ulg multimodality drowsiness database (called drozy) and examples of use," in *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE, 2016, pp. 1–7.
- [12] P. M. Ramos, C. B. Maior, M. C. Moura, and I. D. Lins, "Automatic drowsiness detection for safety-critical operations using ensemble models and eeg signals," *Process Safety and Environmental Protection*, vol. 164, pp. 566–581, 2022.
- [13] F. Wang, B. Lu, X. Kang, and R. Fu, "Research on driving fatigue alleviation using interesting auditory stimulation based on vmd-mmse," *Entropy*, vol. 23, no. 9, p. 1209, 2021.
- [14] F. Rundo, S. Rinella, S. Massimino, M. Coco, G. Fallica, R. Parenti, S. Conoci, and V. Perciavalle, "An innovative deep learning algorithm for drowsiness detection from eeg signal," *Computation*, vol. 7, no. 1, p. 13, 2019.
- [15] K. T. Chui, M. D. Lytras, and R. W. Liu, "A generic design of driver drowsiness and stress recognition using moga optimized deep mkl-svm," *Sensors*, vol. 20, no. 5, p. 1474, 2020.

TABLE V: COMPARISON IN TERMS OF ACCURACY WITH *state-of-the-art* METHODS.

METHOD	DESCRIPTION	DATASET	ACCURACY
Gwak et al. [20]	RF on ECG and EEG signals	Collected dataset	81.4%
Picot et al. [17]	Fuzzy Logic on EOG signals		81.7%
Siddiqui et al. [16]	SVM on respiratory rate data		87.0%
Chaabene et al. [9]	CNN on EEG signals		90.4%
Budak et al. [18]	AlexNet / VGG-16 + LSTM on EEG signals and spectrograms		94.3%
Chui et al. [15]	MOGA + D-MKL-SVM on ECG signals		98.3%
Krishnan et al. [10]	<i>k</i> - NN + <i>multi-class SVM</i> on EEG signals	DROZY	96.1%
Yaacob et al. [3]	<i>k</i> - NN on ECG and EEG signals		97.2%
Proposed	FFT + DT on ECG, EEG, EMG, and EOG signals	DROZY	88.0%

- [16] H. U. R. Siddiqui, A. A. Saleem, R. Brown, B. Bademci, E. Lee, F. Rustam, and S. Dudley, "Non-invasive driver drowsiness detection system," *Sensors*, vol. 21, no. 14, p. 4833, 2021.
- [17] A. Picot, S. Charbonnier, and A. Caplier, "Eog-based drowsiness detection: Comparison between a fuzzy system and two supervised learning classifiers," *IFAC Proceedings Volumes*, vol. 44, no. 1, pp. 14 283–14 288, 2011.
- [18] U. Budak, V. Bajaj, Y. Akbulut, O. Atila, and A. Sengur, "An effective hybrid model for eeg-based drowsiness detection," *IEEE sensors journal*, vol. 19, no. 17, pp. 7624–7631, 2019.
- [19] L. Zhao, M. Li, Z. He, S. Ye, H. Qin, X. Zhu, and Z. Dai, "Data-driven learning fatigue detection system: A multimodal fusion approach of ecg (electrocardiogram) and video signals," *Measurement*, vol. 201, p. 111648, 2022.
- [20] J. Gwak, M. Shino, and A. Hirao, "Early detection of driver drowsiness utilizing machine learning based on physiological signals, behavioral measures, and driving performance," in *2018 21st international conference on intelligent transportation systems (ITSC)*. IEEE, 2018, pp. 1794–1800.
- [21] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
- [22] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
- [23] K. Pearson, "Liii. on lines and planes of closest fit to systems of points in space," *The London, Edinburgh, and Dublin philosophical magazine and journal of science*, vol. 2, no. 11, pp. 559–572, 1901.
- [24] T. Åkerstedt and M. Gillberg, "Subjective and objective sleepiness in the active individual," *International journal of neuroscience*, vol. 52, no. 1-2, pp. 29–37, 1990.

Chapitre 4

Titre :

Méthode basée sur les caractéristiques spatio-temporelles et l'algorithme des machines à vecteurs de support

Résumé - Alors que la somnolence affecte considérablement la capacité humaine à se concentrer et à réagir, elle contribue également au risque d'accident. Les conducteurs ont souvent du mal à détecter leur somnolence jusqu'à ce qu'il soit trop tard, et beaucoup admettent qu'ils conduisent parfois en état de somnolence. Les systèmes antérieurs se fiaient au mouvement de la voiture et à l'angle du volant, ce qui est insuffisant pour éviter les dommages. Les systèmes actuels basés sur la vision sont généralement limités à l'utilisation des indices oculaires, dépendent du réglage chronophage des paramètres, ou ne peuvent pas fonctionner dans certaines circonstances. Il est nécessaire de travailler davantage pour garantir un système précis, accessible et non invasif. Dans cet article, nous présentons une approche pour reconnaître la présence d'un état de somnolence en nous appuyant sur les caractéristiques faciales en utilisant des motifs binaires locaux sur trois plans orthogonaux pour l'extraction de caractéristiques et des Machines à Vecteurs de Support pour la classification. Nos expériences ont été menées en utilisant deux ensembles de données de somnolence réalistes qui sont UTA-RLDD et DROZY, nous donnant une indication précise de 82% et 90% respectivement.

Mots clés :

Détection de la Fatigue, Apprentissage Automatique, Points de Repère Faciaux, Machine à Vecteurs de Support Multi-classes, PCA, UTA-RLDD, DROZY.

Contributions associées :

D. Lamouchi, Y. Yaddaden, J. Parent, R. Cherif, "Efficient Driver Drowsiness Detection Using Spatiotemporal Features with Support Vector Machine," in *International Journal of Intelligent Transportation Systems Research*. (Statut : Soumis)

N. Cabot, **D. Lamouchi**, Y. Yaddaden, "Drowsiness Detection through Yawning and Eye Blinking Models Using Convolutional Neural Networks and Transfer Learning" in *International Conference On Intelligent Computing in Data Sciences (ICDS2024)*. (Statut : Accepté)

International Journal of Intelligent Transportation Systems Research

Efficient Driver Drowsiness Detection Using Spatiotemporal Features with Support Vector Machine

--Manuscript Draft--

Manuscript Number:	IJIT-D-24-00279
Full Title:	Efficient Driver Drowsiness Detection Using Spatiotemporal Features with Support Vector Machine
Article Type:	Research Paper
Funding Information:	
Abstract:	Drowsiness significantly impairs human concentration and reflexes, leading to a heightened risk of accidents. Despite this, many drivers fail to recognize their drowsiness in time, often with serious consequences. Traditional detection systems based on vehicle movement and steering angles are inadequate in preventing such incidents. Existing vision-based systems, while promising, are typically limited to eye movement analysis, require extensive parameter tuning, and often struggle under varying conditions. To address these challenges, we propose a novel approach for Driver Drowsiness Detection that leverages facial features. Our method utilizes Local Binary Patterns on Three Orthogonal Planes for feature extraction and employs Support Vector Machines for classification. Experiments conducted on two benchmark drowsiness datasets, UTA-RLDD and DROZY, demonstrate our system's efficacy, achieving accuracy rates of 82% and 90%, respectively. These results indicate the potential for a more reliable and non-invasive drowsiness detection system.
Corresponding Author:	Yacine Yaddaden, Ph.D. Université du Québec à Rimouski: Universite du Quebec a Rimouski Lévis, Québec CANADA
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Université du Québec à Rimouski: Universite du Quebec a Rimouski
Corresponding Author's Secondary Institution:	
First Author:	Dorra Lamouchi, M.Sc. in computer science
First Author Secondary Information:	
Order of Authors:	Dorra Lamouchi, M.Sc. in computer science
	Yacine Yaddaden, Ph.D.
	Jérôme Parent, B.Sc. in computer science
	Raef Cherif, Ph.D.
Order of Authors Secondary Information:	
Author Comments:	

Noname manuscript No.
(will be inserted by the editor)

Efficient Driver Drowsiness Detection Using Spatiotemporal Features with Support Vector Machine

Dorra Lamouchi · Yacine Yaddaden ·
Jérôme Parent · Raef Cherif

Received: date / Accepted: date

Abstract Drowsiness significantly impairs human concentration and reflexes, leading to a heightened risk of accidents. Despite this, many drivers fail to recognize their drowsiness in time, often with serious consequences. Traditional detection systems based on vehicle movement and steering angles are inadequate in preventing such incidents. Existing vision-based systems, while promising, are typically limited to eye movement analysis, require extensive parameter tuning, and often struggle under varying conditions. To address these challenges, we propose a novel approach for Driver Drowsiness Detection that leverages facial features. Our method utilizes Local Binary Patterns on Three Orthogonal Planes for feature extraction and employs Support Vector Machines for classification. Experiments conducted on two benchmark drowsiness datasets, UTA-RLDD and DROZY, demonstrate our system’s efficacy, achieving accuracy rates of 82% and 90%, respectively. These results indicate the potential for a more reliable and non-invasive drowsiness detection system.

Keywords Drowsiness Detection · Fatigue Detection · Local Binary Patterns on Three Orthogonal Planes · Support Vector Machines · UTA-RLDD · DROZY

D. Lamouchi
Université du Québec à Rimouski, Lévis, QC, Canada.
E-mail: dorra.lamouchi@uqar.ca

Y. Yaddaden
Université du Québec à Rimouski, Lévis, QC, Canada.
E-mail: yacine_yaddaden@uqar.ca

J. Parent
Université du Québec à Rimouski, Rimouski, QC, Canada.
E-mail: jerome.parent@uqar.ca

R. Cherif
Université du Québec à Rimouski, Rimouski, QC, Canada.
E-mail: raef_cherif@uqar.ca

1 Introduction

Modern road networks are becoming increasingly extensive and complex, often struggling to accommodate the rising density of traffic as the number of vehicles continues to grow. Long-distance journeys, especially those connecting remote regions, require extended driving periods [7]. The frequency of business travel has also increased, compelling drivers to spend numerous hours on the road. Existing regulations on driving hours are sometimes inadequate to prevent drowsiness, and mandatory breaks may not be sufficient. Night driving introduces further challenges, including reduced visibility and decreased alertness [32]. According to a report by the World Health Organization (WHO) [27], the annual number of deaths due to road accidents stands at 1.19 million, with tens of millions more people injured. The impact of road accidents is substantial, threatening multiple human lives, including pedestrians, passengers, and the driver. Distraction is often linked to specific levels of fatigue, such as mental fatigue and drowsiness. Unlike tiredness, which fluctuates slowly, drowsiness can set in quickly. Some studies [37] have found that approximately 7,500 fatal crashes in the United States each year might involve drowsy drivers. This alarming statistic underscores the urgent need for an efficient and practical solution.

Fatigue manifests as a set of symptoms, including perceived drowsiness, and can be triggered by various factors that can be categorized into three main areas: vehicle-related factors, physiological characteristics, and behavioral features [31]. Vehicle-related factors include steering wheel angle, lane departure, sudden pressure changes on pedals, speed variations, and driver seating position. These features can be analyzed using appropriate sensors, which are typically installed only on high-end models due to cost. They can also be combined with other methods [33]. The second category encompasses methods based on biological measurements, such as heart or brain signals in the form of electrocardiograms (ECG), photoplethysmograms (PPG), electroencephalograms (EEG), or electromyograms (EMG). While data acquisition for these features can be distracting for the driver, the results are often highly accurate [19]. Finally, behavioral and physical features focus on body movements and gestures in general, and facial expressions in particular, to assess the driver's drowsiness. With the increasing popularity of computer vision, methods for detecting driver drowsiness through facial feature analysis are becoming more prevalent. The recognition of drowsy driving has emerged as a significant area of research in recent years. However, detecting human drowsiness remains a complex and challenging task. Drowsiness has a cumulative nature, meaning that a decision cannot be made based on a short observation period. The previous state of sleepiness significantly contributes to the current state, with the duration varying according to individual behavior [36].

In this paper, we propose a novel approach for Driver Drowsiness Detection (DDD) using Local Binary Patterns from Three Orthogonal Planes (LBP-TOP). This method captures the *spatiotemporal* features of facial expressions to differentiate between *alert* and *drowsy* states. The ultimate goal is to in-

tegrate these detection techniques into Advanced Driver Assistance Systems (ADAS), which can alert drivers when they are at risk of falling asleep at the wheel, thereby enhancing road safety and reducing accidents.

The paper is organized as follows. Section 2 provides an extensive review of existing research approaches. Section 3 establishes the foundational concepts required to understand our proposed system, which is detailed in Section 4. In Section 5, we outline our experimental setup and evaluation criteria. Section 6 presents the obtained results and their analysis. Finally, Section 7 concludes the paper.

2 Background

This section first explores the fundamental principles and technical aspects related to driver fatigue detection. It then evaluates various methodologies, assessing their effectiveness in identifying driver fatigue in real-world scenarios.

2.1 Fundamentals

2.2 Fatigue & Drowsiness

Fatigue and drowsiness are distinct yet interconnected states that are often confused in everyday language. Although these terms are sometimes used interchangeably, they describe different experiences [9]. Fatigue typically refers to a state of diminished ability to perform activities, whether due to physical exertion, mental strain, or repetitive tasks over an extended period. This condition may manifest as a lack of energy, reduced motivation, or impaired cognitive function.

Drowsiness, in contrast, is specifically associated with a strong urge to sleep. This state arises from a fundamental biological need for rest and can be exacerbated by various factors, including health issues, side effects of certain medications, long working hours, and inadequate or poor-quality sleep. Drowsiness is particularly problematic in activities requiring sustained attention, such as driving or operating machinery.

The relationship between fatigue and drowsiness is complex, as fatigue can contribute to the onset of drowsiness. For instance, prolonged mental or physical exertion without sufficient rest can lead to fatigue, which in turn increases the likelihood of drowsiness. Addressing these issues requires a multifaceted approach that includes managing workload, ensuring adequate rest, and addressing underlying health conditions to mitigate their impact on daily functioning and safety.

The Karolinska Sleepiness Scale (**KSS**) [3] is a subjective tool designed to assess an individual's level of drowsiness, as outlined in Table 1.

Table 1: The Karolinska Sleepiness Scale (KSS).

State	Score	Description
Alert	1	Extremely alert
	2	Very alert
	3	Alert
	4	Rather alert
	5	Neither alert nor sleepy
Drowsy	6	Some signs of sleepiness
	7	Sleepy, but no effort to keep awake
	8	Sleepy, some effort to keep awake
	9	Very sleepy, great effort to keep awake
	10	Extremely sleepy, fighting sleep

2.2.1 Advanced Driver Assistance System

Advanced Driver Assistance Systems (ADAS) encompass a range of on-board electronic systems designed to enhance safety and improve the driving experience. These include adaptive cruise control, lane departure warning systems, collision avoidance systems, and other technologies that help prevent accidents and improve vehicle control. ADAS are equipped with sensors and cameras that continuously capture and analyze data, providing real-time feedback to the driver and, in some cases, autonomously taking control of the vehicle to avert danger. Driver Drowsiness Detection (DDD) systems are a specialized subset of ADAS, specifically focused on monitoring and detecting driver fatigue and drowsiness.

2.3 Related Works

Extensive research has been conducted in the field of driver drowsiness detection, exploring a wide range of approaches and methodologies. For this study, the most relevant works have been selected, focusing on those that closely align with the objectives of our project.

2.3.1 Physiological-based

Physiological-based methods for detecting drowsiness leverage biological signals such as EEG, ECG, EMG, and EOG (electrooculogram). These approaches involve analyzing physiological data to assess alertness or drowsiness, often employing advanced signal processing and machine learning techniques.

Research by Krishnan et al. [19] demonstrated the effectiveness of EEG signals in capturing neural correlates of drowsiness. Using K-Nearest Neighbors (KNN) and Support Vector Machines (SVM), they achieved an accuracy of 96.10% on the DROZY dataset. This success underscores the utility of EEG in drowsiness detection. Complementarily, Yaacob et al. [38] combined ECG and EEG signals with KNN and SVM, reporting a slightly higher accuracy of

97.20%. This finding highlights the advantages of integrating different physiological signals.

Expanding on this, Zhao et al. [42] utilized ECG signals in conjunction with Convolutional Neural Networks (CNN) and Bidirectional Long Short-Term Memory (BiLSTM) networks, achieving an impressive accuracy of 99.60%. Their results suggest that deep learning techniques can significantly enhance the performance of physiological-based drowsiness detection systems.

Incorporating multiple physiological signals, Perkins et al. [29] employed EEG, ECG, EMG, and EOG with ResNet-101 and LSTM networks. Their approach achieved an accuracy of 93.10%, demonstrating the potential of multimodal methods in capturing a comprehensive range of physiological changes associated with drowsiness.

Further reinforcing the role of EEG, Alharbey et al. [4] achieved an accuracy of 98% using SVM for classification. Similarly, Ramos et al. [30] employed EEG with a bagging and voting ensemble method, achieving 90

Expanding the scope of physiological signals, Hasan et al. [15] analyzed EOG, EEG, and ECG signals using multiple classifiers such as KNN, SVM, Random Forest (RF), and Artificial Neural Networks (ANN). They reported an accuracy of 83.50%, suggesting that while integrating multiple signals can provide diverse insights, achieving higher accuracy may require more sophisticated modeling.

In addition to EEG-focused studies, Chaabene et al. [8] and Peivandi et al. [28] explored deep learning architectures for drowsiness detection. Chaabene et al. used a CNN, achieving 90.42% accuracy, whereas Peivandi et al. employed a combination of Generative Adversarial Networks (GAN) and CNN, reaching 96.80% accuracy. These studies highlight the evolving role of deep learning in improving drowsiness detection precision.

Finally, Siddiqui et al. [34] investigated respiratory rate data using various machine learning methods, including SVM, Decision Trees (DT), Reinforcement Learning (RL), Gradient Boosting Machines (GBM), Extra Trees, and Multilayer Perceptrons (MLP), achieving an accuracy of 87%. Similarly, Amidei et al. [5] examined skin conductance with RF, bagging, and boosting techniques, obtaining an accuracy of 89.4%. These studies underscore the potential of less commonly used physiological signals in enhancing drowsiness detection.

2.3.2 Behavioral-based

Numerous studies on computer vision-based fatigue detection primarily focus on eye and mouth characteristics. Common indicators of drowsiness include rapid and persistent eye blinking, head nodding, and repetitive yawning. However, despite advancements, these methods often require precise localization of the eyes and mouth, which can be challenging to achieve under real-life driving conditions [24].

Pandey et al. [26] utilized the University of Texas at Arlington (UTA) dataset, specifically the UTA-RLDD, to explore both spatial and temporal

features. Their approach incorporated several facial cues, including the eye-aspect ratio (EAR), mouth-aspect ratio (MAR), and coordinates of the nose and pupil centers. They compared two methods: the first focused on pupil motion with pre-processing treatments, achieving an accuracy of 79.9%, while the second employed a combination of LSTM and CNN with the pre-trained Inception-V3 model, reaching a higher accuracy of 97.5%. This comparison highlights the superior performance of deep learning methods, although they are more complex and time-consuming compared to the simpler approach.

Building on this, Ghoddoosian et al. [14] developed a time-based model using the UTA-RLDD dataset, focusing on features related to eye blinks, such as duration, amplitude, frequency, and speed. They applied a hierarchical multiscale long short-term memory network (HMLSTM) to identify three states of drowsiness: alert, low vigilant, and drowsy. Using a sliding window approach, they ranked video sequences based on regression scores, with the most frequently predicted class representing the final classification.

In a different approach, Tamanani et al. [35] created a system with two sequential components: an initial unit using the Haar cascade algorithm for face detection and preprocessing, followed by a CNN LeNet architecture for feature extraction and classification. Evaluated on the UTA-RLDD dataset, this model achieved an average accuracy of 91.8%, demonstrating the effectiveness of combining traditional and deep learning methods.

Meda et al. [23] explored various machine learning models on a subset of 50 extreme states videos (alert and drowsy) from 25 subjects in the UTA-RLDD dataset. They extracted facial features such as EAR, MAR, pupil circularity (PUR), and the mouth-aspect ratio over the eye-aspect ratio (MOE). Among the models tested, the Random Forest model delivered the best performance.

Adhinata et al. [2] utilized the FaceNet algorithm for feature extraction and tested two classifiers for drowsiness detection. The first classifier used K-Nearest Neighbors (K-NN), while the second employed a one-vs-all Support Vector Machine (SVM) with a polynomial kernel, achieving the highest accuracy of 94.68%. This study underscores the potential of combining advanced feature extraction with robust classification methods.

Expanding the approach, Magán et al. [20] proposed a technique using 60-second video segments for training drowsiness models. They compared two implementations: a deep learning-based model integrating convolutional and recurrent neural networks and a fuzzy logic approach for computing tiredness. This comparison emphasizes the evolving nature of drowsiness detection methodologies, from traditional models to more sophisticated, data-driven approaches.

Azmi et al. [6] explored Local Binary Patterns (LBP) and SVM for classifying eye states into open and closed categories. They used classical machine learning techniques on features such as EAR, MAR, and the percentage of eyelid closure over time (PERCLOS), achieving a 96% accuracy. Although PERCLOS calculation was based on a heuristic threshold, the method proved effective in detecting eye state.

In conjunction, Massoz et al. [22] introduced the ULg Multimodality Drowsiness Database (DROZY) and proposed two systems predicting post-stimulus reaction time (RT) using ocular parameters. The first system performed regression with Epsilon Support Vector Regression (epsilon-SVR), while the second conducted classification with SVM, determining RT thresholds. These systems highlight the utility of ocular parameters in assessing drowsiness.

Maier et al. [21] developed a real-time drowsiness detection method based on eye blink classification, testing models such as shallow multilayer perceptron (MLP), Random Forest, and SVM. SVM yielded the best results, validated on the DROZY database with input sequences of EAR, and categorized blinks into open eye, short blink, or long blink. The use of KSS levels for evaluation underscores the integration of subjective assessments with automated detection.

Dua et al. [11] utilized the National Tsing Hua University (NTHU) dataset to create an ensemble learning system combining FlowImageNet, AlexNet, VGG-FaceNet, and ResNet models. Data augmentation was performed to enhance dataset diversity, and binary classification was used to distinguish between drowsy and non-drowsy states.

Lastly, Jabbar et al. [16] proposed a lightweight CNN-based model suitable for embedded and Android devices, trained on the NTHU dataset. This approach, using a five-layer CNN, achieved an accuracy of 83.33% while maintaining a compact size of under 75KB. This development highlights the potential for deploying drowsiness detection systems in resource-constrained environments.

3 Proposed System

The Fig. 1 presents the different components of the proposed DDD system.

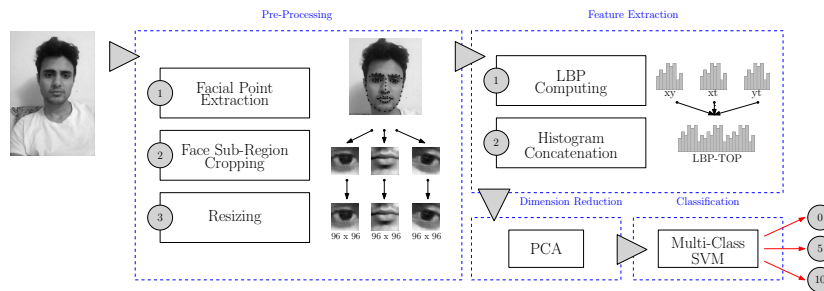


Fig. 1: Overview of the proposed DDD system.

3.1 Pre-processing

As shown in Fig. 2, the input data undergoes several pre-processing steps to obtain equal-dimension images for each region of interest [40]. To maintain a good compromise between data size control and the retention of valuable information from each recording, a sample was selected every fifteen frames, rather than processing each frame individually.

The first step in our pre-processing pipeline is dedicated to face detection. For this purpose, we used the facial landmark detection method proposed by Kazemi and Sullivan [17], which is known for its reliability and is implemented in **Dlib** [18]. This method is based on a combination of machine learning techniques, including both traditional shape predictors and deep learning models, designed to accurately identify 68 facial landmark points, encompassing facial features such as the eyes, nose, mouth, and various facial contours. Using these landmark points, which serve as essential spatial references, three distinct regions of interest are precisely isolated from the rest of the facial features: the right eye, the left eye, and the mouth. Finally, the cropped regions are resized to 96×96 pixels to ensure uniform standardization of dimensions.

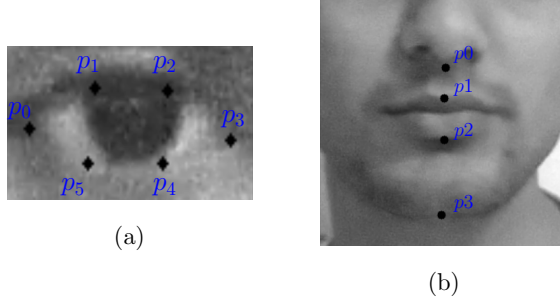


Fig. 2: Region of interest extraction: (a) eyes & (b) mouth.

3.2 Spatiotemporal Feature Extraction

3.2.1 Local binary patterns

Local Binary Patterns (LBP) [25] is a highly efficient texture operator and a key element in facial recognition tasks [6]. It characterizes the pixels of an image based on the grayscale values of their neighboring pixels. The resulting histogram can then be used as a texture descriptor. Fig. 3 provides an overview of the computation steps. The operator thresholds each pixel's neighborhood using the central value and interprets the result as a binary number, as shown in Eq. 1.

$$LBP_{p,r} = \sum_{i=0}^{p-1} s(g_i - g_0) \times 2^i, \quad s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (1)$$

where g_0 corresponds to the gray level of the center pixel of a local neighborhood and g_p to the gray levels of p equally spaced pixels on a circle of radius r .

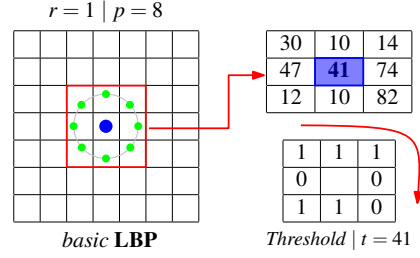


Fig. 3: Overview of the *basic LBP* [39].

3.2.2 Local Binary Patterns from three orthogonal planes

Due to the temporal aspect of the data, LBP-TOP [41] has been used to represent the extracted regions of interest over time. Instead of representing the input video as a stack of XY planes along the T axis, LBP-TOP is computed by taking stacks of XT planes along the Y axis and YT planes along the X axis, as shown in Fig. 4. The LBP histograms extracted from the XY , XT , and YT planes are concatenated into a single histogram, forming the LBP-TOP feature vector. The size of the histogram is 3×2^p , where P is the number of neighbors when applying the LBP.

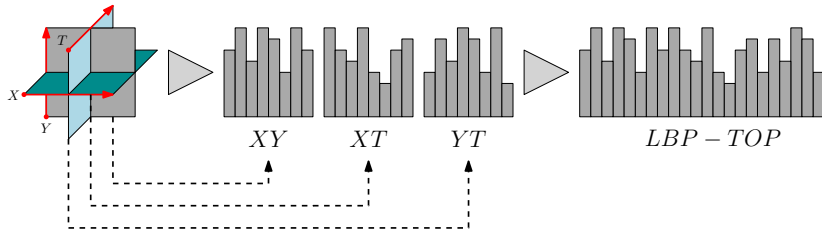


Fig. 4: Overview of the **LBP-TOP**.

3.2.3 Feature extraction using LBP-TOP

The main purpose of using LBP-TOP lies in its ability to provide the *spatiotemporal* information needed to accurately capture signs of drowsiness over windows of frames. For each frame, a circular neighborhood around a pixel is defined and divided into multiple sectors. Within each sector, a binary pattern is derived by comparing the central pixel's intensity with the intensities of the surrounding pixels. The value 1 is assigned if the surrounding pixel's intensity is higher; otherwise, it is 0. Binary patterns are independently calculated for each of the three orthogonal planes.

A series of histograms is produced, representing the spatial texture patterns in a specific plane. These histograms are then concatenated into a complete feature vector, characterizing both spatial and temporal texture variations within each facial region of interest.

3.3 Dimensionality Reduction

Due to the significant size of the feature vector resulting from the LBP-TOP feature extraction process, the need to reduce dimensionality was recognized. For this purpose, principal component analysis (PCA) [1] is applied, retaining the most significant information while maximizing the variance preserved in the data.

PCA operates by identifying principal components that represent linear combinations of the original features. By using a subset of these components, the dimensionality of the feature vector is effectively reduced, as shown in Eq. 2 where x is the original feature vector, containing n features, and z is the transformed feature vector, containing m principal components, where $m \leq n$. Each element z_j of the vector z is a linear combination of the original features x_i , with the coefficients $a_{i,j}$ representing the contribution of each original feature to the j -th principal component.

$$\begin{aligned} x &= x_1, x_2, \dots, x_n \\ z &= z_1, z_2, \dots, z_m \\ z_j &= a_j^T x = \sum_{i=1}^n a_{i,j} \cdot x_i \end{aligned} \tag{2}$$

The covariance matrix of the data points is computed, and its eigenvectors and associated eigenvalues are determined. Next, the eigenvectors are sorted in descending order. Selecting the first N (number of components) eigenvectors provides the new N dimensions.

3.4 Classification

The creation of systems that are able to carry out intricate tasks depends heavily on machine learning. Machine learning algorithms are employed in our context to evaluate data to ascertain whether a driver is drowsy. In these kinds of systems, recognition is the primary use of machine learning, where algorithms are trained to discriminate between various actions or states. Supervised learning, a subset of machine learning, is typically used for this. When learning under supervision, the system is trained on a set of labeled data with known intended results.

In this research, a supervised machine learning method is studied to perform the classification task. The use of linear SVM [10] is founded on its high performance and its ability to efficiently generalize even with relatively limited quantity of available data.

Our experiments, involving more than two sleepiness levels, employed the strategic approach called one-vs-all. It is particularly suitable for multiclass classification scenarios. The Machine Learning model is trained to generate a distinct binary classifier for each sleepiness level, by treating each specific level as the positive class while collapsing all the remaining levels into the negative class [12].

4 Data Description

Benchmark drowsiness datasets, such as the NTHU Driver Drowsiness Detection dataset [36], include videos in which subjects artificially simulate drowsiness. In these staged scenes, subjects often exaggerate behaviors, such as extended yawning or prolonged eye closure, which may not accurately capture the subtlety of real drowsiness. In contrast, both datasets used in our research are composed entirely of videos featuring genuinely drowsy subjects.

4.1 UTA-RLDD

The UTA-RLDD dataset [14] is the largest realistic driver drowsiness dataset to date. It comprises recordings from 60 volunteers who captured themselves using cellphones or web cameras in various real-life environments, with differing angles, lighting conditions, and backgrounds. Each of the 60 subjects recorded themselves three times in different states of drowsiness, which were self-assessed using the Karolinska Sleepiness Scale (KSS). The resulting dataset consists of 180 videos, each approximately 10 minutes long, recorded at 30 frames per second or less. The three states of drowsiness are classified as follows: **Alert**: KSS score of 1-3, no signs of sleepiness; **Low Vigilance**: KSS score of 6-7, some signs of sleepiness appear; **Drowsy**: KSS score of 8-9, actively trying not to fall asleep. Examples from the database are shown in Fig. 5.



Fig. 5: Samples from the **UTA-RLDD** dataset.

4.2 DROZY

This dataset [22] provides videos captured under conditions similar to those of the Real-Life Drowsiness Dataset. Unlike the UTA-RLDD dataset, all recordings in this dataset were conducted under controlled conditions, using the same camera, angle, lighting, and distance for each subject. Each of the 14 subjects was filmed three times for approximately 10 minutes per session, with recordings made at 30 frames per second. The dataset comprises 36 videos due to some lost data. Subjects were instructed to maintain a normal sleep schedule for one week prior to the tests and to have at least 7 to 8 hours of sleep the night before the first test. Before each test, subjects rated their level of drowsiness on the Karolinska Sleepiness Scale (KSS). By the time of the third test, subjects had been awake for 28 to 30 hours. Not all subjects were recorded in all states of drowsiness, and the imbalanced distribution of videos across states made it more appropriate to classify the videos into two states rather than three, as done in the UTA-RLDD dataset. We classified the two states as follows: **Alert**: KSS score of 1-5; **Drowsy**: KSS score of 6-9. Examples from the database are shown in Fig. 6.

5 Evaluation Protocol

5.1 Performance Metrics

To evaluate the performance of each model across both datasets, accuracy was used as the primary performance indicator, as defined by Eq. 3. Additionally,



Fig. 6: Samples from the **DROZY** dataset.

confusion matrices were employed to visually represent the effectiveness of the models.

$$accuracy(y, \hat{y}) = \frac{1}{n_{samples}} \sum_{i=0}^{n_{samples}-1} 1(\hat{y}_i = y_i) \quad (3)$$

where y represents the true labels, $\hat{y}$ denotes the predicted labels, $n_{samples}$ is the total number of samples, and $1(\hat{y}_i = y_i)$ is an indicator function that equals 1 when the predicted label matches the true label for the i -th sample and 0 otherwise.

5.2 Validation Strategy

To accurately assess the effectiveness of the proposed approach in real-world scenarios, classification is performed on entire videos. The experiments and evaluations for both datasets are conducted using stratified 5-fold cross-validation. In each iteration, four folds are used for training, and one fold is used for evaluation. The final metrics are obtained by averaging the scores from all iterations. Each video is composed of N windows of T length, and the model is trained to classify these windows into one of the drowsiness states. The predicted class for a given video is determined by the most frequently predicted state across all windows. The descriptor LBP-TOP with a number of neighbors of either 4 or 8 and a circle radius of 1, 2 or 3 has been tested as shown in Fig. 7.

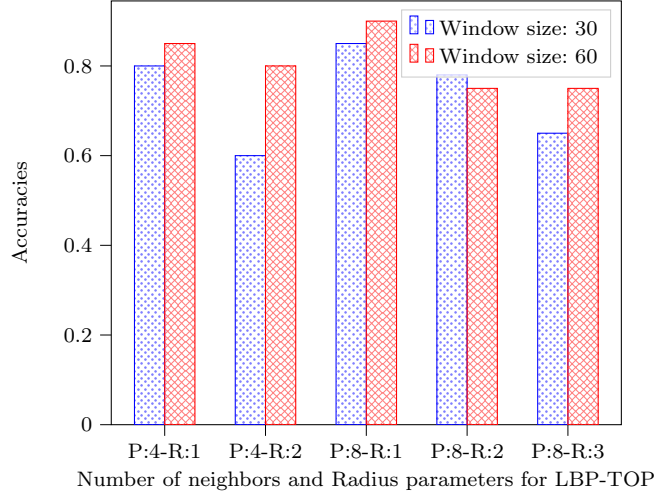


Fig. 7: The result of the experiment on various **LBP-TOP** parameters.

6 Results & Discussion

The model achieved a binary classification accuracy of 82.00% on the UTA-RLDD dataset. However, when introducing a third state of drowsiness (Low-Vigilance), the accuracy drops to 61.00%, as shown in Table 2. As expected, the middle class is the hardest to identify, which significantly impacts overall accuracy. These results align with those obtained by [14] using their baseline method. Deep learning models trained and evaluated on the UTA-RLDD dataset perform very well. Although [2] reported high testing accuracy, their model was trained and evaluated on a small sample of images rather than long video sequences. [23] compared the performance of seven different models based on facial features, with the highest-performing model achieving a binary classification accuracy of 78.00%. Compared with our results, this demonstrates the strength of SVM as a classifier for detecting states of drowsiness based on facial features.

Model accuracy is assessed using the confusion matrix, as shown in Table 3, that classifies two states: "Alert" and "Drowsy". The "Alert" cases represent an accuracy of 84% with a classification error rate of 16%, while recognition of the "Drowsy" state is achieved at 80%, but there is a possibility of misclassification of these patients as an alert state in 20% of cases.

Classification of three states can be achieved using the UTA-RLDD dataset, which gives us a better overview of the confusion matrix as presented in Table 4. "Alert" instances are accurately classified by the model at 71%. Furthermore, 47% of "Low-vigilance" cases were correctly classified. Finally, "Drowsy" states are identified by the model in 64% of cases.

Table 2: Comparison with *state-of-the-art* methods using **UTA-RLDD**.

METHOD	FEATURES	CLASSIFICATION	STATES	ACCURACY
[14]	Blink Sequences	HM-LSTM	3	65.20%
[26]	Facial Features	LSTM	2	79.90%
	CNN			97.00%
[23]	Facial Features	Random Forest	2	78.00%
[2]	CNN (FaceNet)	KNN	3	98.00%
		SVM		95.00%
Proposed	LBP-TOP	SVM	2	82.00%
			3	61.00%

Table 3: The confusion matrix using 2 classes on the **UTA-RLDD** dataset.

	Alert	Drowsy
Alert	0.84	0.16
Drowsy	0.20	0.80

Table 4: The confusion matrix using 3 classes on the **UTA-RLDD** dataset.

	Alert	Low-Vigilance	Drowsy
Alert	0.71	0.16	0.13
Low-Vigilance	0.24	0.47	0.29
Drowsy	0.15	0.22	0.64

On the DROZY dataset, the highest accuracy achieved by our model was 90.00%, as shown in Table 5. This dataset includes not only videos but also physiological signals and reaction times from Psychomotor Vigilance Tests (PVTs). Consequently, the literature features numerous methods using machine learning models trained on these additional features, making direct comparisons with our work challenging. Our approach yields higher accuracy than the one proposed by [22], although it is slightly lower than methods using biological features, which have proven to be highly accurate. The method proposed by [13] achieves an accuracy of 91.07% on video sequences by applying transfer learning to a MobileNet-224 model pre-trained on the Inception network. This slightly higher accuracy is expected from a well-established CNN architecture. The model developed by [21] does not directly predict drowsiness; instead, drowsiness is inferred from two rules based on blink patterns, as explained in Section 2. The system was validated on the DROZY dataset. The ground truth required no drowsiness warnings for alert subjects and at least one warning for drowsy subjects. Of the eight tested videos of alert subjects, no warnings were generated by the system. Of the ten tested videos of drowsy subjects, at least one warning was emitted in nine samples. Thus, the experiment correctly classified 17 out of 18 videos, resulting in an accuracy of 94.44%.

As shown in Table 6, the confusion matrix generated from the DROZY dataset pays attention to the model’s ability to classify ”Alert” and ”Drowsy”

Table 5: Comparison with *state-of-the-art* methods using **DROZY**.

METHOD	FEATURES	CLASSIFICATION	STATES	ACCURACY
[22]	Ocular Parameters	SVM	2	84.99%
[13]	Facial Features	MobileNet	2	91.07%
[21]	Blink Sequences	SVM	2	94.44%
[19]	EEG	KNN	2	95.80%
		SVM	2	94.30%
Proposed	LBP-TOP	SVM	2	90.00%

states. The values on the diagonal indicate that, for both "Alert" and "Drowsy" states, the model achieves 89% accuracy rates whereas there were 11% misclassifications for each case. Such a strong diagonal performance implies that this model is quite efficient at differentiating these two states in the DROZY dataset.

Table 6: The confusion matrix when using the **DROZY** dataset.

	Alert	Drowsy
Alert	0.89	0.11
Drowsy	0.11	0.89

7 Conclusion

In this paper, we propose a technique for detecting driver drowsiness that combines facial feature representation with classical machine learning methods. This study explores the application of facial representation techniques that have proven effective in other areas of facial analysis but have been scarcely used for fatigue recognition. Appearance-based features provide an alternative to the purely geometric methods commonly employed by researchers. Our model relies on traditional machine learning rather than the more recent deep learning approaches, and it has demonstrated reliability, efficiency, and interpretability. We successfully detected drowsiness signals using this model in two distinct real-world datasets.

7.1 Limitations

Despite the promising results demonstrated by the system, several limitations remain. One major limitation is its reliance on lighting conditions; variations in lighting can significantly affect image quality and, consequently, the accuracy of the system. Additionally, individual differences in how people exhibit drowsiness suggest that the system may need to be customized to account for

this variability. Another limitation is related to the real-time implementation of the module. While designed for efficiency, the module may struggle to operate in real time on devices with limited computational resources. Furthermore, the model's ability to generalize is contingent on the diversity of its training data, which may not encompass all possible scenarios of drowsiness.

7.2 Perspectives & Improvements

To address these limitations and enhance the system's capabilities, several improvements can be considered. One area for future work is to enhance the model's robustness to variations in lighting. This could involve developing advanced preprocessing techniques or adaptive modeling approaches to better handle diverse lighting conditions. Another avenue for improvement is increasing accuracy through personalization, which would allow the model to account for specific user characteristics, including individual physiological and behavioral traits. Additionally, optimizing support for embedded devices is crucial to enable real-time operation on hardware with limited resources. Expanding the training datasets to cover a wider range of conditions and subjects will further strengthen the model's generalization, making it applicable across diverse contexts.

Acknowledgements We extend our sincere gratitude to the **ULg-INTELSIG** laboratory, responsible for the collection of the **DROZY** dataset [22] and the **The University of Texas at Arlington** who collected the **UTA-RLDD** dataset [14] used in our research.

References

1. Abdi, H., Williams, L.J.: Principal component analysis. Wiley interdisciplinary reviews: computational statistics **2**(4), 433–459 (2010). DOI 10.1002/wics.101
2. Adhinata, F.D., Rakhmadani, D.P., Wijayanto, D.: Fatigue detection on face image using facenet algorithm and k-nearest neighbor classifier. Journal of Information Systems Engineering and Business Intelligence **7**(1), 22–30 (2021). DOI 10.20473/jisebi.7.1.22-30
3. Åkerstedt, T., Gillberg, M.: Subjective and objective sleepiness in the active individual. International journal of neuroscience **52**(1-2), 29–37 (1990). DOI 10.3109/00207459008994241
4. Alharbey, R., Dessouky, M.M., Sedik, A., Siam, A.I., Elaskily, M.A.: Fatigue state detection for tired persons in presence of driving periods. IEEE Access **10**, 79403–79418 (2022). DOI 10.1109/ACCESS.2022.3185251
5. Amidei, A., Spinsante, S., Iadarola, G., Benatti, S., Tramarin, F., Pavan, P., Rovati, L.: Driver drowsiness detection: a machine learning approach on skin conductance. Sensors **23**(8), 4004 (2023). DOI 10.3390/s23084004
6. Azmi, N., Rahman, A.M., Shirmohammadi, S., El Saddik, A.: Lbp-based driver fatigue monitoring system with the adoption of haptic warning scheme. In: 2011 IEEE International Conference on Virtual Environments, Human-Computer Interfaces and Measurement Systems Proceedings, pp. 1–4. IEEE (2011). DOI 10.1109/VECIMS.2011.6053852
7. Buchanan, C.: Traffic in Towns: A study of the long term problems of traffic in urban areas. Routledge (2015)
8. Chaabene, S., Bouaziz, B., Boudaya, A., Hökelmann, A., Ammar, A., Chaari, L.: Convolutional neural network for drowsiness detection using eeg signals. Sensors **21**(5), 1734 (2021). DOI 10.3390/s21051734

9. Chacon-Murguia, M.I., Prieto-Resendiz, C.: Detecting driver drowsiness: A survey of system designs and technology. *IEEE Consumer Electronics Magazine* **4**(4), 107–119 (2015). DOI 10.1109/MCE.2015.2463373
10. Cristianini, N., Shawe-Taylor, J.: An introduction to support vector machines and other kernel-based learning methods. Cambridge university press (2000). DOI 10.1017/CBO9780511801389
11. Dua, M., Singla, R., Raj, S., Jangra, A., et al.: Deep cnn models-based ensemble approach to driver drowsiness detection. *Neural Computing and Applications* **33**(8), 3155–3168 (2021). DOI 10.1007/s00521-020-05209-7
12. Duan, K.B., Rajapakse, J.C., Nguyen, M.N.: One-versus-one and one-versus-all multi-class svm-rfe for gene selection in cancer classification. In: *Evolutionary Computation, Machine Learning and Data Mining in Bioinformatics: 5th European Conference, EvoBIO 2007, Valencia, Spain, April 11–13, 2007. Proceedings 5*, pp. 47–56. Springer (2007). DOI 10.1007/978-3-540-71783-6_5
13. García-García, M., Caplier, A., Rombaut, M.: Sleep deprivation detection for real-time driver monitoring using deep learning. In: *Image Analysis and Recognition: 15th International Conference, ICIAR 2018, Póvoa de Varzim, Portugal, June 27–29, 2018, Proceedings 15*, pp. 435–442. Springer (2018). DOI 10.1007/978-3-319-93000-8_49
14. Ghoddoosian, R., Galib, M., Athitsos, V.: A realistic dataset and baseline temporal model for early drowsiness detection. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pp. 0–0 (2019). DOI 10.1109/CVPRW.2019.00027
15. Hasan, M.M., Watling, C.N., Larue, G.S.: Physiological signal-based drowsiness detection using machine learning: Singular and hybrid signal approaches. *Journal of safety research* **80**, 215–225 (2022). DOI 10.1016/j.jsr.2021.12.001
16. Jabbar, R., Shinoy, M., Kharbeche, M., Al-Khalifa, K., Krichen, M., Barkaoui, K.: Driver drowsiness detection model using convolutional neural networks techniques for android application. In: *2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT)*, pp. 237–242. IEEE (2020). DOI 10.1109/ICIOT48696.2020.9089484
17. Kazemi, V., Sullivan, J.: One millisecond face alignment with an ensemble of regression trees. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1867–1874 (2014). DOI 10.1109/CVPR.2014.241
18. King, D.E.: Dlib-ml: A machine learning toolkit. *The Journal of Machine Learning Research* **10**, 1755–1758 (2009)
19. Krishnan, P., Yaacob, S., Krishnan, A.P., Rizon, M., Ang, C.K.: Eeg based drowsiness detection using relative band power and short-time fourier transform. *J. Robotics Netw. Artif. Life* **7**(3), 147–151 (2020). DOI 10.2991/jrnal.k.200909.001
20. Magán, E., Sesmero, M.P., Alonso-Weber, J.M., Sanchis, A.: Driver drowsiness detection by applying deep learning techniques to sequences of images. *Applied Sciences* **12**(3), 1145 (2022). DOI 10.3390/app12031145
21. Maior, C.B.S., das Chagas Moura, M.J., Santana, J.M.M., Lins, I.D.: Real-time classification for autonomous drowsiness detection using eye aspect ratio. *Expert Systems with Applications* **158**, 113505 (2020). DOI 10.1016/j.eswa.2020.113505
22. Massoz, Q., Langohr, T., François, C., Verly, J.G.: The ulg multimodality drowsiness database (called drozy) and examples of use. In: *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 1–7. IEEE (2016). DOI 10.1109/WACV.2016.7477715
23. Meda, H., Ganesh, J.M.P., Sahani, A.: Machine learning models for drowsiness detection. In: *2021 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, pp. 1–5. IEEE (2021). DOI 10.1109/I2MTC50364.2021.9460017
24. Moujahid, A., Dornaika, F., Arganda-Carreras, I., Reta, J.: Efficient and compact face descriptor for driver drowsiness detection. *Expert Systems with Applications* **168**, 114334 (2021). DOI 10.1016/j.eswa.2020.114334
25. Ojala, T., Pietikäinen, M., Harwood, D.: A comparative study of texture measures with classification based on featured distributions. *Pattern recognition* **29**(1), 51–59 (1996). DOI 10.1016/0031-3203(95)00067-4

26. Pandey, N.N., Muppalaneni, N.B.: Temporal and spatial feature based approaches in drowsiness detection using deep learning technique. *Journal of Real-Time Image Processing* **18**(6), 2287–2299 (2021). DOI 10.1007/s11554-021-01114-x
27. Peden, M.M.: World report on road traffic injury prevention. World Health Organization (2004)
28. Peivandi, M., Ardabili, S.Z., Sheykhivand, S., Danishvar, S.: Deep learning for detecting multi-level driver fatigue using physiological signals: A comprehensive approach. *Sensors* **23**(19), 8171 (2023). DOI 10.3390/s23198171
29. Perkins, E., Sitaula, C., Burke, M., Manousakis, J., Anderson, C., Marzbanrad, F.: Investigating physiological and behavioural sensing modalities towards drowsiness detection. *IEEE Sensors Journal* (2023). DOI 10.1109/JSEN.2023.3326434
30. Ramos, P.M., Maior, C.B., Moura, M.C., Lins, I.D.: Automatic drowsiness detection for safety-critical operations using ensemble models and eeg signals. *Process Safety and Environmental Protection* **164**, 566–581 (2022). DOI 10.1016/j.psep.2022.06.039
31. Reddy, B., Kim, Y.H., Yun, S., Seo, C., Jang, J.: Real-time driver drowsiness detection for embedded system using model compression of deep neural networks. In: *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pp. 121–128 (2017). DOI 10.1109/CVPRW.2017.59
32. Sekkay, F., Imbeau, D., Dubé, P.A., Chinniah, Y., de Marcellis-Warin, N., Beauregard, N., Trépanier, M.: Assessment of physical work demands of long-distance industrial gas delivery truck drivers. *Applied Ergonomics* **90**, 103224 (2021). DOI 10.1016/j.apergo.2020.103224
33. Shi, S.Y., Tang, W.Z., Wang, Y.Y.: A review on fatigue driving detection. In: *ITM Web of Conferences*, vol. 12, p. 01019. EDP Sciences (2017). DOI 10.1051/itmconf/20171201019
34. Siddiqui, H.U.R., Saleem, A.A., Brown, R., Bademci, B., Lee, E., Rustam, F., Dudley, S.: Non-invasive driver drowsiness detection system. *Sensors* **21**(14), 4833 (2021). DOI 10.3390/s21144833
35. Tamanani, R., Muresan, R., Al-Dweik, A.: Estimation of driver vigilance status using real-time facial expression and deep learning. *IEEE Sensors Letters* **5**(5), 1–4 (2021). DOI 10.1109/LSENS.2021.3070419
36. Weng, C.H., Lai, Y.H., Lai, S.H.: Driver drowsiness detection via a hierarchical temporal deep belief network. In: *Computer Vision–ACCV 2016 Workshops: ACCV 2016 International Workshops, Taipei, Taiwan, November 20–24, 2016, Revised Selected Papers, Part III* 13, pp. 117–133. Springer (2017). DOI 10.1007/978-3-319-54526-4_9
37. Wheaton, A.G., Shults, R.A., Chapman, D.P., Ford, E.S., Croft, J.B.: Drowsy driving and risk behaviors—10 states and puerto rico, 2011–2012. *Morbidity and Mortality Weekly Report* **63**(26), 557 (2014)
38. Yaacob, S., Affandi, N.A.I., Krishnan, P., Rasyadan, A., Yaakop, M., Mohamed, F.: Drowsiness detection using eeg and ecg signals. In: *2020 IEEE 2nd International Conference on Artificial Intelligence in Engineering and Technology (IICAET)*, pp. 1–5. IEEE (2020). DOI 10.1109/IICAET49801.2020.9257867
39. Yaddaden, Y.: An efficient facial expression recognition system with appearance-based fused descriptors. *Intelligent Systems with Applications* **17**, 200166 (2023). DOI 10.1016/j.iswa.2022.200166
40. Yaddaden, Y., Adda, M., Bouzouane, A., Bouchard, B., et al.: Facial sub-regions for automatic emotion recognition using local binary patterns. In: *2018 International Conference on Signal, Image, Vision and their Applications (SIVA)*, pp. 1–6. IEEE (2018). DOI 10.1109/SIVA.2018.8661026
41. Zhao, G., Pietikainen, M.: Dynamic texture recognition using local binary patterns with an application to facial expressions. *IEEE transactions on pattern analysis and machine intelligence* **29**(6), 915–928 (2007). DOI 10.1109/TPAMI.2007.1110
42. Zhao, L., Li, M., He, Z., Ye, S., Qin, H., Zhu, X., Dai, Z.: Data-driven learning fatigue detection system: A multimodal fusion approach of ecg (electrocardiogram) and video signals. *Measurement* **201**, 111648 (2022). DOI 10.1016/j.measurement.2022.111648

Chapitre 5

Conclusion Générale

La sécurité routière reste une préoccupation majeure à l'échelle mondiale, notamment en ce qui concerne la détection de la fatigue du conducteur représentant un facteur clé de la réduction des accidents souvent mortels, ce qui démontre l'importance de développer des méthodes efficaces permettant de détecter la fatigue et d'alerter les conducteurs en temps adéquat.

La progression de l'IA et de l'apprentissage automatique a entraîné une évolution dans ce domaine favorisant une synthèse plus exhaustive avec des systèmes de détection plus adaptatifs. Les techniques étudiées dans ce projet varient en fonction de la nature des données analysées, notamment les signaux physiologiques et les aspects comportementaux du conducteur. Les méthodes physiologiques offrent une grande précision, mais leur mise en œuvre peut être coûteuse et intrusive, tandis que les méthodes comportementales sont potentiellement plus intégrables dans les véhicules actuels, bien qu'elles puissent être affectées par des variables externes.

5.1 Objectifs accomplis

Ce mémoire a traité la détection de la somnolence chez les conducteurs grâce à l'analyse de deux études différentes, chacune explorant un aspect distinct :

- ✓ Le premier article, intitulé "Étude comparative : Détection de la fatigue des conducteurs basée sur des signaux physiologiques via des méthodes traditionnelles et hybrides", souligne le rôle des signaux physiologiques dans la détection de la fatigue. En comparant les approches conventionnelles aux méthodes hybrides avancées, l'article révèle que l'utilisation conjointe d'algorithmes d'apprentissage profond et d'analyses spectrales permet d'obtenir des systèmes de détection plus précis.
- ✓ Dans le deuxième article intitulé "Détection de la somnolence du conducteur à l'aide de motifs binaires locaux de trois plans orthogonaux", une technique de vision est utilisée pour extraire des caractéristiques du visage. Cette approche démontre la validité des modèles binaires locaux pour détecter les signes de somnolence dans les vidéos des conducteurs.
- ✓ L'Annexe A de ce mémoire présente un article de conférence dont j'ai contribué à sa réalisation avec une étudiante de premier cycle dans le cadre de son stage d'initiation à la recherche. L'article, intitulé "Détection de la somnolence à travers les modèles de bâillements et de clignements des yeux utilisant les réseaux neuronaux convolutifs et l'apprentissage par transfert", se concentre sur l'élaboration des méthodes automatisées capables de détecter les deux principaux indicateurs de la fatigue : les bâillements et les clignements des yeux.

5.2 Comparaison des approches développées

Comme décrit dans la Revue de littérature, l'ensemble des méthodes de détection des signes de fatigue a été évalué à l'aide de deux bases de données publiques.

Dans ce qui suit, nous présenterons une comparaison entre les différentes méthodes proposées en termes de taux de reconnaissance. Dans la figure 5.1 sont représentées les meilleures performances pour chaque système, notamment celui basé sur les signaux physiologiques et celui utilisant les expressions faciales. Nous pouvons remarquer que le taux de reconnaissance le plus élevé a été atteint par la méthode comportementale et en utilisant la base de donnée **DROZY** [39].

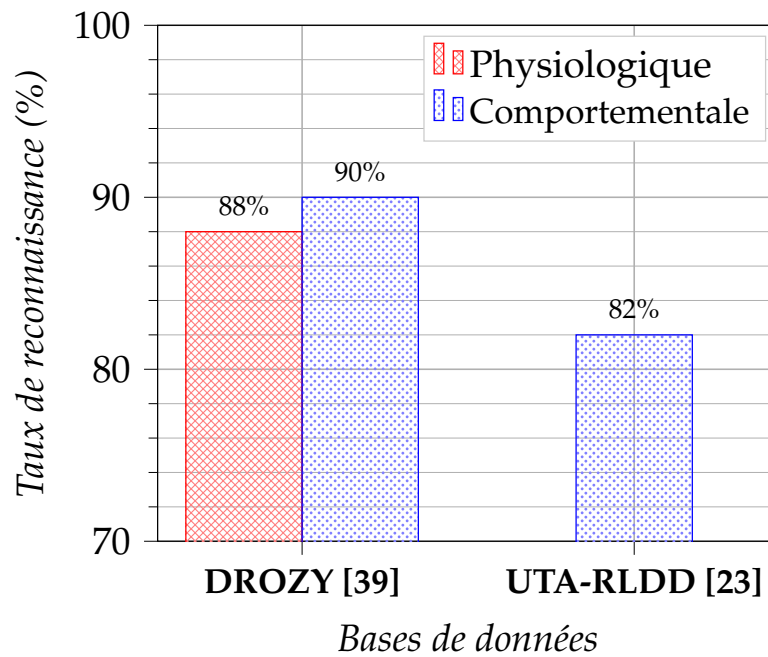


FIGURE 5.1 – Comparaison des résultats des méthodes élaborées.

5.3 Limitations connues

- × L'exclusion des données relatives au véhicule limite notre compréhension de l'interaction entre le conducteur et le véhicule, ce qui pourrait être utile pour notre recherche.

- × La séparation des systèmes selon leur nature, l'une axée sur les caractéristiques physiologiques et l'autre sur les caractéristiques comportementales, risque d'empêcher une analyse intégrée susceptible d'améliorer la qualité de la détection.
- × L'utilisation des bases de données publiques, bien qu'elle facilite la comparaison avec d'autres travaux, présente des défis en termes de taille et de diversité des données ce qui restreint la robustesse de nos conclusions.
- × La variabilité individuelle des sujets tels que par exemple l'âge, le sexe et l'état de santé, n'a pas été suffisamment prise en compte, ce qui devrait nécessiter des phases de pré-traitement plus adaptatifs.

5.4 Perspectives et travaux futurs

Bien que des progrès considérables aient été réalisés dans la détection de la somnolence au volant, il reste des défis à relever.

- ◇ Les systèmes doivent s'adapter aux différents environnements de conduite et aux profils variés des conducteurs.
- ◇ Les futures recherches se concentrent sur le développement de technologies plus polyvalentes et accessibles pour une application plus étendue et une efficacité accrue dans la prévention des accidents liés à la somnolence.
- ◇ D'autres approches et architectures d'apprentissage profond seront explorées pour améliorer la fiabilité des modèles.
- ◇ L'évolution de ce travail inclut le développement d'interventions préventives une fois la fatigue détectée chez le conducteur.
- ◇ Un système d'alerte efficace doit stimuler la vigilance sans induire d'anxiété. Cela peut se concrétiser par des alertes sonores ou tactiles

telles que des vibrations du siège ou du volant.

5.5 Retour d'expérience

Au cours de ma maîtrise, j'ai eu l'opportunité de me plonger profondément dans le domaine de l'intelligence artificielle. Ce projet n'est pas ma première incursion dans ce domaine puisque j'ai eu le privilège de contribuer à trois articles de conférence portant sur divers sujets. Toutefois, ce projet revêt un caractère singulier en raison de son impact humain et concret et de ses retombées bénéfiques pour la société. Cette perspective m'a motivé à choisir ce projet en particulier pour ma maîtrise.

L'encadrement offert pour ce projet est empreint d'un engagement solide, ce qui représente un facteur qui m'inspire et renforce mon efficacité. Les échanges constructifs et les conseils personnalisés que me procure régulièrement mon directeur de recherche représentent une assistance solide et pertinente, idéale pour l'avancement du projet. Tout cela a renforcé mon ambition de poursuivre mes études au niveau doctoral. En intégrant un programme de doctorat, je souhaite explorer des questions innovantes et contribuer significativement à l'avancement scientifique.

Je suis fière des réalisations accomplies dans le cadre de ce projet, ainsi que de l'obtention de la bourse d'excellence et de la bourse de persévérance scolaire. De plus, j'ai participé au Forum Innovation, Ingénierie, Informatique et Entrepreneuriat (FI3E) durant deux années consécutives, où mes présentations sous forme de posters (en Annexe B) m'ont valu le premier prix dans la catégorie informatique ainsi que le prix coup de cœur du public pour chacune de mes participations. Par ailleurs, j'ai occupé deux charges de cours

parallèlement à mes études, ce qui m'a permis de m'entraîner à devenir professeur universitaire. Grâce à cette expérience, j'ai su gérer mes responsabilités tout en obtenant le prix de meilleur chargé de cours en informatique.

Annexe A

Titre :

Détection de la somnolence par des modèles de bâillements et de clignements des yeux utilisant des réseaux neuronaux convolutifs et l'apprentissage par transfert

Résumé - La somnolence du conducteur est l'une des principales causes d'accidents de la route car elle diminue les capacités cognitives et les réflexes. Par conséquent, différentes méthodes basées sur l'intelligence artificielle ont été développées pour surveiller l'état du conducteur en temps réel. Cependant, ces méthodes nécessitent encore des améliorations car l'identification des signes de fatigue, tels que les bâillements et les clignements des yeux, reste un défi majeur. L'objectif principal de cet article est de créer deux modèles utilisant des réseaux neuronaux convolutifs (CNN) pré-entraînés pour détecter les bâillements et les clignements des yeux, ce qui devrait contribuer à prévenir les accidents de la route. Différents ensembles de données ont été utilisés, notamment YawDD et MRL Eye, pour entraîner divers modèles pré-entraînés basés sur des architectures bien connues telles que MobileNet, Xception, Inception-V3 et VGG-16, en utilisant le transfert d'apprentissage. Ces modèles générés ont démontré une efficacité élevée, MobileNet obtenant les meilleures performances globales avec une précision variant de 97,64% à 99,52%, selon l'ensemble de données utilisé pour l'entraînement et la validation. Les résultats indiquent que les modèles CNN pré-entraînés affinés par le transfert d'apprentissage sont très efficaces pour détecter les signes de somnolence, ce qui est prometteur pour les applications réelles dans les systèmes avancés d'aide à la conduite afin d'assurer la sécurité du conducteur.

Mots clés :

Détection de la somnolence, mesures comportementales, Réseaux neuronaux convolutifs, Apprentissage par transfert, MobileNet, Inception-V3, Xception, VGG-16, YawDD, MRL Eye

Drowsiness Detection through Yawning and Eye Blinking Models Using Convolutional Neural Networks and Transfer Learning

Noémie Cabot
Département de mathématiques,
informatique et génie
Université du Québec à Rimouski
Rimouski, Québec, Canada.
noemie.cabot@uqar.ca

Dorra Lamouchi
Département de mathématiques,
informatique et génie
Université du Québec à Rimouski
Rimouski, Québec, Canada.
dorra.lamouchi@uqar.ca

Yacine Yaddaden
Département de mathématiques,
informatique et génie
Université du Québec à Rimouski
Rimouski, Québec, Canada.
yacine_yaddaden@uqar.ca

Raef Cherif
Département de mathématiques,
informatique et génie
Université du Québec à Rimouski
Rimouski, Québec, Canada.
raef_cherif@uqar.ca

Abstract—Driver drowsiness is one of the main causes of road traffic accidents because it reduces cognitive capabilities and reflexes. Therefore, different artificial intelligence-based methods have been developed for monitoring the driver's state in real time. However, these methods still need improvement because identifying signs of fatigue, such as yawning and eye blinking, remains a significant challenge. The main objective of this paper is to create two models using pre-trained Convolutional Neural Networks (CNNs) to detect yawning and eye blinking, which should help prevent road traffic accidents. Different datasets were used, including YawDD and MRL Eye, to train various pre-trained models based on well-known architectures such as MobileNet, Xception, Inception-V3, and VGG-16, using transfer learning. These generated models demonstrated high efficiency, with MobileNet achieving the best overall performance with an accuracy ranging from 97.64% to 99.52%, depending on the dataset used for training and validation. The results indicate that pre-trained CNN models *fine-tuned* with transfer learning are highly effective in detecting signs of drowsiness, which is promising for real-world applications in Advanced Driver Assistance Systems (ADAS) to ensure the driver's safety.

Index Terms—Drowsiness Detection, Behavioral Measures, Convolutional Neural Networks, Transfer Learning, MobileNet, Inception-V3, Xception, VGG-16, YawDD Dataset, MRL Eye Dataset.

I. INTRODUCTION

In recent years, we have witnessed a significant increase in the number of vehicles on the road. Moreover, people are driving longer distances and in varying conditions. Naturally, this has led to an increase in road traffic accidents, resulting in many fatalities, with 1.19 million deaths in 2023 [1]. Various factors contribute to the increase in the accident rate, with the main ones related to the driver, the vehicle, and the environment. To prevent accidents and reduce casualties, several safety features have been deployed in road design,

such as rumble strips that alert the driver to slow down or to correct their lane [2]. Although these mechanisms contribute to accident prevention, they are still not very effective and require improvement.

With the increase in casualties due to road traffic accidents and advancements in information and communication technologies, smart vehicles and systems have been introduced to prevent these accidents. These systems are called Advanced Driver Assistance Systems (ADAS) and include several types of sensors that collect data from the driver, vehicle, and environment. Using the collected data, various approaches based on artificial intelligence techniques, especially deep neural networks such as Convolutional Neural Networks (CNNs) [3], have been explored.

Among the different factors that lead to road traffic accidents is the driver's mental or physical state, particularly the level of drowsiness. Indeed, drowsiness is directly related to the immediate risk of falling asleep at the wheel [4]. Thus, detecting the level of drowsiness might help prevent road traffic accidents by notifying drivers that they need to stop before falling asleep at the wheel.

To detect the level of drowsiness directly from the driver, two types of data are utilized: *physiological signals* [5] and *behavioral analysis*. This paper focuses on the latter by recognizing the main signs of drowsiness through facial analysis [6]. The two primary signs are *eye blinking* and *yawning*. Increased blinking rate and frequent yawning are indicators of drowsiness, as they reflect the driver's tendency to close their eyes and feel sleepy. These signs are detected in two facial regions, namely the eyes and the mouth. In the proposed Driver Drowsiness Detection (DDD) system, a *preprocessing* step is introduced to robustly and efficiently extract these regions.

Subsequently, several *pre-trained* and well-known deep neural network models are *fine-tuned* using *transfer learning* to create models for detecting each sign. Each model is trained and validated using various datasets.

The structure of this paper is organized as follows: Section II introduces the fundamental concepts that are essential to our study. Section III examines previous techniques and approaches. The **CNNs** are explained in Section IV, while Section V describes *transfer learning*. The proposed methodology for drowsiness detection is detailed in Section VI. Then, Section VII outlines the datasets used and the evaluation strategy. Results and discussion are presented in Section VIII. This paper concludes with Section IX, which provides concluding remarks and perspectives.

II. FUNDAMENTALS

The main aim of drowsiness detection systems lies in detecting early signs of drowsiness through close monitoring of the driver's behavior and physiological state. This information is used to inform the driver and prevent road traffic accidents, thereby ensuring road safety. In order to monitor the driver's state while driving, three categories of data can be collected using distinct sensors. They are described as follows:

DRIVER-CENTERED DATA (*direct*) that describe the state of the driver, including two subcategories:

- **Physiological signals** [5], which are more intrusive and include measurements of brain activity, heart rate, and muscle response signals, among others.
- **Behavioral analysis**, which typically employs cameras for facial analysis or to detect and recognize body movements

VEHICLE-CENTERED DATA (*indirect*) that provide information about the driving style through measures such as steering patterns, braking behavior, and lane positioning.

Designing a **DDD** system involves several steps similar to those of a common pattern recognition system. Depending on the type of data collected, which in our case are images or sequences of images, specific *preprocessing* steps are applied to prepare and enhance the quality of the input data. Subsequently, *feature extraction* techniques are employed to generate a more effective representation of the input data. A robust representation is crucial for achieving high accuracy. To further refine this representation, *feature selection* or *dimensionality reduction* techniques may be utilized. Finally, the *classification* phase employs *supervised* machine learning techniques to classify the driver's state into two classes: *drowsy* or *alert*.

In this work, we focus on detecting signs of drowsiness, specifically eye blinking and yawning. Instead of using traditional architectures with hand-engineered feature extraction, we employ **CNNs** that can autonomously extract relevant features from images. However, the regions of interest must be extracted, as has been done in previous work on facial expression recognition [7].

III. RELATED WORKS

Over the years, several **DDD** systems have been proposed using behavioral analysis, particularly facial analysis. Some works are based on traditional architectures, while others utilize deep neural networks. Below, we present some recent works in the field of **DDD**.

Sun et al. [8] utilized **Inception-V3** and **VGG-16** architectures to detect signs of drowsiness. The **MRL Eye** dataset was employed for training. After a series of preprocessing steps, including data augmentation and overfitting prevention, **VGG-16** demonstrated slightly higher accuracy during training, but **Inception-V3** outperformed it during validation.

Vijaypriya & Uma [9] proposed a **DDD** system based on the detection of a specific sign of drowsiness, namely eye blinking. The authors developed their own **CNN** architecture for feature extraction and classification. Before that, they applied preprocessing techniques such as normalization and data augmentation to improve the quality of the input images and ensure a higher accuracy. They used the **YawDD** dataset for training and validation.

In the paper by Gupta et al. [10], the authors proposed the development of a **DDD** system through three distinct phases. The first phase focuses on extracting *spatial* features using *pre-trained CNN* models to characterize eye blinking and yawning. Based on the obtained models, the authors used an **LSTM** (Long Short-Term Memory) network in the second phase to include *temporal* features. In both phases, the authors used two distinct benchmark datasets, namely the **MRL Eye** and **YawDD** datasets. In the final phase, the authors used another benchmark dataset, the **NTHU** dataset, on which they achieved great performance.

Another deep learning-based approach was presented by Rajkar et al. [11] for detecting driver drowsiness using **CNNs**. The authors focused on detecting eye blinks as the main sign of drowsiness. They began by using a Haar Cascade classifier to detect the regions of the eyes from the face image before feeding it to a **CNN** for classification.

Adhinata et al. [12] used the same technique, the Haar Cascade classifier, to detect the face and then employed a *pre-trained CNN*, namely **FaceNet**, to generate a relevant representation. They then used two types of classifiers with the **UTA-RLDD** dataset, namely the *multi-class* Support Vector Machine (**SVM**) and *k*-Nearest Neighbors (*k*-**NN**). They obtained the highest accuracy with the latter.

Other authors were more interested in the Eye Aspect Ratio (**EAR**) to describe drowsiness. Maior et al. [13] proposed extracting facial landmarks before computing the **EAR** and applying the **SVM** classifier. The authors performed the training and validation primarily on the **DROZY** dataset, on which they obtained relatively good performance.

Jabbar et al. [14] developed a mobile application that performs **DDD** through facial analysis. They used preprocessing techniques such as face detection, facial landmarks extraction, and image augmentation. They then trained a *lightweight CNN* model that can be run on a smartphone.

IV. CONVOLUTIONAL NEURAL NETWORKS

CNNs are the most common and popular deep neural networks, introduced by Yann LeCun [15] and successfully used for object detection and recognition by autonomously extracting relevant features without relying on *hand-engineered* features [3]. They are Inspired by the human visual cortex, where individual neurons correspond to specific areas of the visual field, CNNs are designed to process pixel data efficiently. They begin by identifying simple features such as edges in initial layers, progressing to more complex structures like shapes, and ultimately detecting even more intricate features such as faces and objects. CNN architectures typically comprise several types of layers:

CONVOLUTIONAL LAYERS apply various filters to the input data to generate *feature maps*,

ACTIVATION LAYERS introduce non-linearities to capture complex patterns,

POOLING LAYERS aim to reduce the dimensionality of the *feature maps* generated previously,

FULLY CONNECTED LAYERS aggregate information from all neurons in the preceding layer and process the extracted features to make final classifications.

CNNs have been widely used in various fields including pattern recognition, such as in healthcare [16], biometrics [17], and ambient assisted living [3]. They can be employed in different modes: for classification from scratch [3], fine-tuning existing models, or for feature extraction to generate relevant representations [16], [17]. In our case, we are interested in *fine-tuning pre-trained CNNs*. Below, we present the architectures that have been utilized:

A. Inception-V3

Inception-V3 [18] is a CNN architecture developed as part of Google's *Inception* series, designed for high-precision image processing tasks. It incorporates several advanced features, including 7×7 factorized convolutions, which reduce computational costs by decomposing large convolutional filters into smaller operations. Another notable feature is the grid reduction technique, which processes spatial dimensions without pooling to preserve crucial spatial information. These techniques allow the network to achieve increased depth without proportionally increasing computing resources. **Inception-V3** is well-suited for handling complex image processing tasks due to its optimal balance between computational power and efficiency.

B. MobileNet

MobileNet [19] is an optimized architecture designed for mobile and resource-constrained environments. It employs depth-separable convolutions to reduce the number of parameters and computational workload compared to networks of similar depth, without significantly sacrificing performance. The key innovation of **MobileNet** lies in its separation of convolution into two distinct layers: *depthwise* convolution and *pointwise* convolution. This approach enhances flexibility and

efficiency, making **MobileNet** suitable for various tasks such as image classification, object detection, and facial recognition.

C. VGG-16

Named after the Visual Geometry Group at Oxford that developed it, **VGG-16** [20] is a CNN renowned for its simplicity and depth. It was designed to enhance the accuracy of image recognition and classification tasks. In the **VGG-16** architecture, each layer employs 3×3 convolutional filters uniformly, and by stacking more layers than its predecessors, it achieves greater depth. Max pooling is also utilized to reduce spatial dimensions between convolutional layers. **VGG-16** consists of 16 layers in total: 13 convolutional layers followed by 3 fully connected layers. This design enables it to capture a broad spectrum of features, making it a popular choice for numerous deep learning applications in image processing. However, due to its substantial memory and computational requirements, **VGG-16** is less suitable for deployment in resource-constrained environments.

D. Xception

Xception [21] is a deep learning architecture developed by Google that improves upon the *Inception* model by incorporating *depthwise* separable convolutions. This approach divides the convolution process into two stages: *depthwise* convolutions and *pointwise* convolutions. It allows **Xception** to handle complex patterns with fewer parameters and minimal processing costs. The model consists of 36 convolutional layers organized into 14 modules, all of which feature linear residual connections around them, except for the first and last modules. This configuration enhances **Xception**'s effectiveness in tasks requiring extensive image processing capabilities.

V. TRANSFER LEARNING

Transfer learning [22] is an approach to machine learning where a pre-trained model on one task is used as the starting point for a new model on a new task. In the case of CNNs, the original architecture is trained on a large image dataset, such as **ImageNet**. During the training, the different layers learn to identify common features, such as edges and textures, found in a wide variety of images.

When applying transfer learning to image classification tasks, the convolutional core of the pre-trained model is used as a feature extractor for the new model, and its layers are initially frozen to maintain the learned weights. Custom layers are then added to the pre-trained base to adapt the model to a specific application.

Initially, only these newly added layers are trained, allowing the pre-extracted features to adjust to the current task. Once the new layers have been sufficiently trained, it is possible to unfreeze parts of the pre-trained model to allow for more fine-tuning. This can lead to performance improvements, especially if the dataset is large enough to avoid *overfitting*. Transfer learning reduces the required computational resources and training time while exploiting the feature-extracting abilities of the pre-trained network.

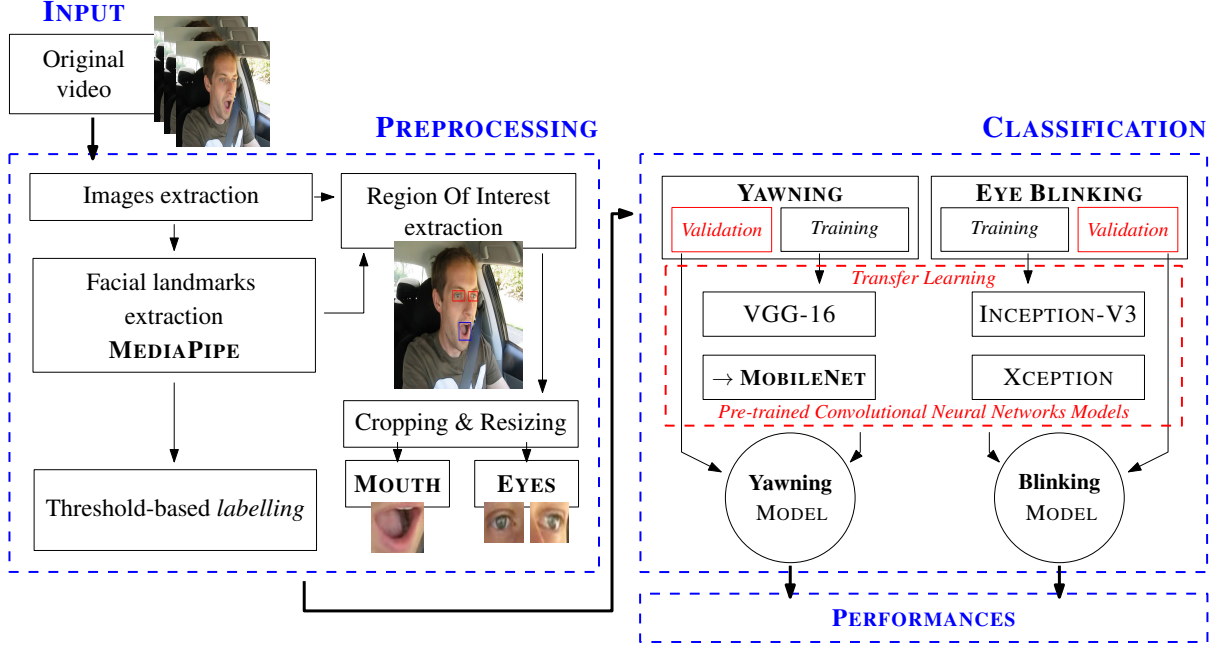


Fig. 1: Overview of the proposed DDD system.

VI. PROPOSED SYSTEM

The main objective of this work is to create efficient and accurate models based on *pre-trained CNNs* to detect signs of drowsiness, namely eye blinking and yawning. By creating both models, we aim to detect the driver's state in real-time. Fig. 1 represents the overall diagram of our proposed system. The proposed system consists of two main stages, which are detailed in the following subsections.

A. Preprocessing

In the context of this work, the images have been pre-processed in different ways depending on which benchmark dataset they belong to. For the **MRL Eye** dataset, since the images already included annotations indicating whether the eye is *open* or *closed*, the only step was to *resize* them to the same dimension of 128×128 pixels. For the **Drowsiness** and **Yawn** datasets, the images were resized, and annotations were associated to the images.

The **YawDD** dataset is more complex since it includes videos of the entire face. To extract the regions of interest, namely the eyes and the mouth, we had to follow the steps below:

- 1) The frames are extracted from each video at a frequency of 30 fps using OpenCV.
- 2) MediaPipe Face Mesh [23] is employed to identify all facial landmarks (a total of 478 points).
- 3) Using specific points, the three regions of interest, namely both eyes and the mouth, are extracted.
- 4) Each extracted image is resized to dimensions of 128×128 pixels.
- 5) Each image is annotated using the computed mouth and eye opening ratios.

The ratio r , which reflects the degree of openness of the mouth and eyes, is computed using the *width* and *height* of each eye and the mouth. Each *width* and *height* is determined using two different facial landmark points $P(x, y)$ around the regions of interest of the eyes and mouth. The height is computed using the point with the highest value of $y_{\max}$ and the point with the lowest value $y_{\min}$ with $height = y_{\max} - y_{\min}$. Similarly, the width is computed using the point with the highest value of $x_{\max}$ and the point with the lowest value $x_{\min}$ with $width = x_{\max} - x_{\min}$. Finally, the ratio r is computed using the following formula (Eq. 1):

$$r = \frac{height}{width} \quad (1)$$

If the $r_{\text{mouth}} > 0.75$, an annotation stating that the mouth is *open* is added to the image; otherwise, it indicates *closed*. The same method is used to determine if the eyes are open or closed. In this case, if the ratio $r_{\text{eye}} < 0.45$, the eye is considered *closed*; otherwise, it is considered *open*. Both thresholds were *empirically* determined based on tests to find the ratios that most accurately represented closed eyes or mouth.

B. Classification

After preprocessing all the images for each region of interest, namely the eyes and the mouth, the next step is to perform recognition. For that, we need to use a *supervised* machine learning technique, meaning that the chosen classifier has to be trained through a *learning phase* with *annotated* data before generating a *model* that performs the recognition.

Usually, authors either use traditional algorithms such as **SVM** or *k*-NN, while others leverage the power of deep neural

networks. In the context of this work, we are more interested in the latter, but with a specific training approach. Instead of performing the training from scratch, we use *pre-trained CNN* models where the weights have been already adjusted using large image datasets. For this work, we used **Inception-V3** [18], **MobileNet** [19], **VGG-16** [20], and **Xception** [21].

All the *pre-trained CNN* models used are *multi-class* classifiers since they include multiple outputs depending on the datasets they have been trained on. However, in the context of our work, we focus on *binary* classification, specifically the differentiation between the two states of the eyes (*blink* or *no blink*) or mouth (*yawn* or *no yawn*). Therefore, additional layers are added, including a Global Average Pooling 2D layer, a dense layer, a dropout layer, and a final dense layer with the *sigmoid* activation function needed for *binary* classification.

The final step is to use various benchmark datasets to *fine-tune* the weights of the pre-trained models through *transfer learning*. The resulting fine-tuned models are then evaluated and made ready for recognition.

VII. EXPERIMENTATION AND EVALUATION

In this section, we present the various datasets used for training and evaluating the models, along with the different metrics employed for evaluation and the validation strategy.

A. Data description

In this work, we used four distinct datasets for training and evaluating the models. Table I presents the different datasets along with their key information. Fig. 2 provides an overview of sample images from each dataset used.

The **YawDD** dataset [24] consists of a total of 322 videos (at 30 fps). It includes three distinct driving scenarios in which 57 male and 50 female volunteers subjects. The scenarios are: 1) normal driving (no talking), 2) talking or singing while driving, and 3) yawning while driving.

The **MRL Eye** dataset [25] is used for blink detection. It contains *infrared (IR)* eye images captured under different lighting conditions and using various acquisition devices. The dataset includes 84,898 images from 37 subjects in total, with 33 men and 4 women. It is annotated with information such as eye states and whether glasses are worn.

TABLE I: DETAILS ABOUT EACH BENCHMARK DATASET.

CATEGORY	NAME	SUBJECTS	SAMPLES	TYPE
Eye Blinking	MRL Eye [25]	37	84,898	IR images
	Drowsiness	—	1,452	images
Yawning	Yawn	—	5,119	images
Both	YawDD [24]	107	322	videos

The **Yawn** and **Drowsiness** datasets were downloaded from the well-known **Kaggle**¹ platform, which contains several datasets. The **Yawn** dataset contains 5,119 images in both color and black and white with different resolutions. It includes two states: *yawning* and *not yawning*. On the other hand, the **Drowsiness** dataset contains 1,452 images of open and closed eyes.

¹<https://www.kaggle.com/>

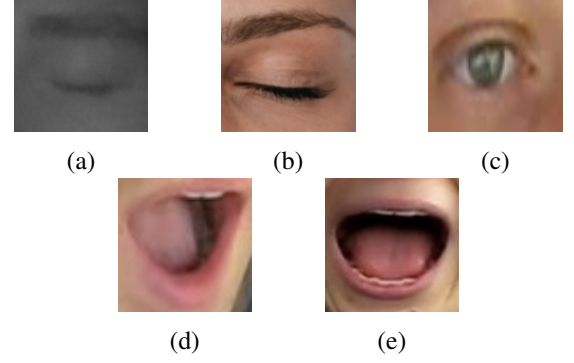


Fig. 2: Preprocessed samples from each dataset: (a) **MRL Eye** [25], (b) Drowsiness, (c) **YawDD** [24] (eyes), (d) **YawDD** [24] (mouth) & (e) Yawn.

B. Evaluation metrics

To assess the performance of each model on every dataset, four metrics were used: *accuracy*, *precision*, *recall*, and *F1-score*. In addition, confusion matrices were used to illustrate the effectiveness of the models in predicting the two classes.

Equation 2 represents the *accuracy* defined by the ratio of true results to the total number of results. The *precision* (see Eq. 3) is the ratio of true positive observations to the total predicted positives. The *recall* (see Eq. 4) measures the proportion of positive instances among actual positives. The *F1-Score* (see Eq. 5) is the harmonic mean of precision and recall.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (3)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (4)$$

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (5)$$

From the above equations, *TP*, *TN*, *FP*, and *FN* refer to the number of *true positive*, *true negative*, *false positive*, and *false negative*, respectively.

C. Validation Strategy

In order to evaluate the different models, we adopted the 10-folds *cross-validation* strategy. It divides the dataset into ten *stratified* subsets. The model is then trained on nine subsets and tested on the remaining one. This process is repeated ten times in order to use each subset exactly once as the test set.

VIII. RESULTS & DISCUSSION

Our experiment consisted in evaluating the performances of the four selected *pre-trained CNN* models on which has been applied *transfer learning* with different benchmark datasets in order to compare them and identify the models that offer the best performance for eye-blink and yawn detection.

Table II presents the results obtained by the blink detection models. It is observed that for the three datasets related to blinking, **MobileNet** provided the best results. For **MRL Eye** [25], **YawDD** [24] and **Drowsiness** datasets, the accuracies are 97.64%, 99.34% and 99.52%, respectively. On the other hand, Table III shows the results for the yawn detection models. Once again, the **MobileNet** architecture delivers the best performance, achieving accuracies of 98.75% and 98.44% on **YawDD** and **Yawn**, respectively.

TABLE II: PERFORMANCE OF MODELS FOR *eye blinking*.

DATASET	MODEL	PRECISON	RECALL	F1-SCORE	ACCURACY
MRL Eye	Inception-V3	96.20%	96.20%	96.20%	96.19%
	MobileNet	97.60%	97.60%	97.6%	97.64%
	VGG-16	92.30%	92.10%	92.10%	92.13%
	Xception	96.90%	96.8%	96.80%	96.85%
YawDD	Inception-V3	99.10%	99.10%	99.10%	99.09%
	MobileNet	99.30%	99.30%	99.30%	99.34%
	VGG-16	98.50%	98.50%	98.50%	98.51%
	Xception	99.10%	99.10%	99.10%	99.11%
Drowsiness	Inception-V3	97.80%	97.80%	97.80%	97.80%
	MobileNet	99.50%	99.50%	99.50%	99.52%
	VGG-16	97.10%	97.10%	97.10%	97.11%
	Xception	97.90%	97.90%	97.90%	97.87%

TABLE III: PERFORMANCE OF MODELS FOR *yawning*.

DATASET	MODEL	PRECISON	RECALL	F1-SCORE	ACCURACY
YawDD	Inception-V3	97.80%	97.80%	97.80%	97.81%
	MobileNet	98.80%	98.80%	98.70%	98.75%
	VGG-16	98.80%	98.80%	98.80%	97.96%
	Xception	97.70%	97.70%	97.70%	97.70%
Yawn	Inception-V3	96.70%	96.70%	96.70%	96.68%
	MobileNet	98.40%	98.40%	98.40%	98.44%
	VGG-16	96.8%	96.7%	96.7%	96.74%
	Xception	97.3%	97.3%	97.3%	97.34%

To assess the performance of **MobileNet** in detecting signs of drowsiness, Table IV shows the confusion matrices for this model across all datasets. Tables IV (A), (B), and (C) present the recognition rates for each state of the eye considered as signs of drowsiness (*blink* & *no blink*). We observe high recognition rates ranging from 97.3% to 99.6%.

Tables IV (D) and (E) present the recognition rates for each state of the mouth considered as signs of drowsiness (*yawn* & *no yawn*). Here as well, we observe high recognition rates ranging from 98.3% to 98.9%.

Fig. 3 shows the accuracy achieved when using **MobileNet** for each benchmark dataset and for both signs of drowsiness. In general, **MobileNet** performs well with high accuracy, but it performs slightly better with the **Drowsiness** dataset for yawning, achieving an accuracy of 99.52%, and with **YawDD** [24] for *blinking*, achieving an accuracy of 98.75%.

All models offer high accuracy, but **MobileNet** tends to perform better whereas **VGG-16** often performs slightly worse. This difference can be explained by the architectural advantages inherent in **MobileNet**. Its *lightweight* and fast design makes it suitable for applications such as drowsiness detection in driving scenarios. **MobileNet** uses depth-separable convolutions, which reduce the number of parameters compared to the convolutions used in **VGG-16**, thus lowering the risk of *overfitting* on smaller datasets.

Additionally, model performance was influenced by different dataset conditions, including lighting, image resolution, and subject diversity. Due to its deeper and more complex structure, which requires more data for effective generalization, **VGG-16** may have encountered more challenges than **MobileNet**. Consequently, **MobileNet** can maintain high accuracy under a variety of input conditions.

TABLE IV: CONFUSION MATRIX OF **MOBILENET** ON: (A) **MRL EYE** [25], (B) **DROWSINESS**, (C) **YAWDD** [24] (*eyes*), (D) **YAWDD** [24] (*mouth*) & (E) **YAWN**.

(A)			(B)		
	<i>no blink</i>	<i>blink</i>		<i>no blink</i>	<i>blink</i>
<i>no blink</i>	97.3%	2.7%	<i>no blink</i>	99.4%	0.6%
<i>blink</i>	2.0%	98.0%	<i>blink</i>	0.4%	99.6%
(C)			(D)		
	<i>no blink</i>	<i>blink</i>		<i>no yawn</i>	<i>yawn</i>
<i>no blink</i>	99.1%	0.9%	<i>no yawn</i>	98.6%	1.4%
<i>blink</i>	0.4%	99.6%	<i>yawn</i>	1.1%	98.9%
(E)				<i>no yawn</i>	<i>yawn</i>
	<i>no yawn</i>	<i>yawn</i>	<i>no yawn</i>	98.3%	1.7%
<i>no yawn</i>			<i>yawn</i>	1.4%	98.6%

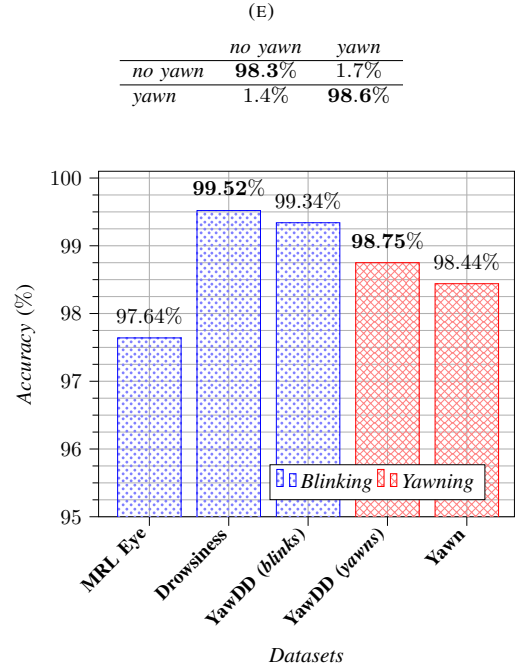


Fig. 3: Performance of the **MobileNet** on each dataset.

In both Table V and VI, we compare the performance of our method with existing *state-of-the-art* methods using

TABLE V: COMPARISON IN TERMS OF ACCURACY WITH STATE-OF-THE-ART METHODS (*eye blinking*).

METHOD	DESCRIPTION	DATASET	ACCURACY
Sun et al. [8]	Inception-V3	MRL Eye	92.30%
Rafid et al. [26]	Haar classifier + CNN		94.50%
Vijaypriya & Uma [9]	CNN	YawDD	97.00%
Mohanty et al. [27]	Eye Aspect Ratio + Threshold		93.21%
Proposed	MobileNet	MRL Eye	97.64%
		YawDD	99.34%

TABLE VI: COMPARISON IN TERMS OF ACCURACY WITH STATE-OF-THE-ART METHODS (*yawning*).

METHOD	DESCRIPTION	DATASET	ACCURACY
Mohanty et al. [27]	Mouth Aspect Ratio + Threshold	YawDD	96.71%
Rajkar et al. [11]	Haar cascade + CNN		95.76%
Proposed	MobileNet	YawDD	98.75%

the benchmark datasets **YawDD** [24] and **MRL Eye** [25]. We observe that even when certain authors used deep neural networks, our method still outperforms theirs in terms of accuracy. This proves the efficiency of our proposed method for detecting drowsiness.

IX. CONCLUDING REMARKS

In the context of this work, we proposed models that allow the detection of signs of drowsiness, namely *eye blinking* and *yawning*. We explored the performance of various *pre-trained CNN* models on which we applied *transfer learning* with benchmark datasets. While all models produced excellent results, **MobileNet** stood out by achieving more than 99% accuracy. In addition to being accurate, **MobileNet** is also considered a *lightweight* model, allowing it to run in resource-constrained environments. Therefore, it can be easily integrated into vehicles to warn drivers when signs of drowsiness are detected.

Integrating these **CNN** models into in-vehicle systems offers a promising way of improving driver safety by monitoring and reacting to signs of drowsiness. As a future perspective, we plan to improve and adapt the models to process sequences of images and take into account the temporal aspect that provides more information.

ACKNOWLEDGMENT

Our sincere appreciation goes to the Distributed and Collaborative Virtual Environment Research Laboratory (University of Ottawa) for providing the **YawDD** dataset [24] and to the Media Research Lab (Technical University of Ostrava) for providing the **MRL Eye** dataset [25].

REFERENCES

- [1] World Health Organization, *Global status report on road safety 2023*. World Health Organization, 2023.
- [2] M. M. Peden, *World report on road traffic injury prevention*. World Health Organization, 2004.
- [3] Y. Yaddaden, M. Adda, A. Bouzouane, S. Gaboury, and B. Bouchard, "User action and facial expression recognition for error detection system in an ambient assisted environment," *Expert Systems with Applications*, vol. 112, pp. 173–189, 2018.
- [4] A. Williamson, D. A. Lombardi, S. Folkard, J. Stutts, T. K. Courtney, and J. L. Connor, "The link between fatigue and safety," *Accident Analysis & Prevention*, vol. 43, no. 2, pp. 498–515, 2011.
- [5] D. Lamouchi, Y. Yaddaden, and R. Cherif, "Comparative study: Physiological-based driver drowsiness detection utilizing traditional and hybrid methods," in *2024 8th International Conference on Image and Signal Processing and their Applications (ISPA)*. IEEE, 2024, pp. 1–8.
- [6] Y. Yaddaden, "An efficient facial expression recognition system with appearance-based fused descriptors," *Intelligent Systems with Applications*, vol. 17, p. 200166, 2023.
- [7] Y. Yaddaden, M. Adda, A. Bouzouane, B. Bouchard *et al.*, "Facial sub-regions for automatic emotion recognition using local binary patterns," in *2018 International Conference on Signal, Image, Vision and their Applications (SIVA)*. IEEE, 2018, pp. 1–6.
- [8] Z. Sun, J. H. Chuah, and G. M. T. Chai, "Eye state recognition using inceptionv3 and vgg16," in *2023 International Conference on Artificial Intelligence, Blockchain, Cloud Computing, and Data Analytics (ICoABCD)*. IEEE, 2023, pp. 94–99.
- [9] V. Vijaypriya and M. Uma, "Comparative analysis for detecting drowsiness using deep learning approach," in *2023 Second International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT)*. IEEE, 2023, pp. 1–4.
- [10] A. Gupta, V. Maan, K. S. Sangwan *et al.*, "Optimized human drowsiness detection using spatial and temporal features of eye and mouth," in *2023 10th International Conference on Signal Processing and Integrated Networks (SPIN)*. IEEE, 2023, pp. 866–871.
- [11] A. Rajkar, N. Kulkarni, and A. Raut, "Driver drowsiness detection using deep learning," in *Applied Information Processing Systems: Proceedings of ICCET 2021*. Springer, 2022, pp. 73–82.
- [12] F. D. Adhinata, D. P. Rakhmadani, and D. Wijayanto, "Fatigue detection on face image using facenet algorithm and k-nearest neighbor classifier," *Journal of Information Systems Engineering and Business Intelligence*, vol. 7, no. 1, pp. 22–30, 2021.
- [13] C. B. S. Maior, M. J. das Chagas Moura, J. M. M. Santana, and I. D. Lins, "Real-time classification for autonomous drowsiness detection using eye aspect ratio," *Expert Systems with Applications*, vol. 158, p. 113505, 2020.
- [14] R. Jabbar, M. Shinoy, M. Kharbeche, K. Al-Khalifa, M. Krichen, and K. Barkaoui, "Driver drowsiness detection model using convolutional neural networks techniques for android application," in *2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT)*. IEEE, 2020, pp. 237–242.
- [15] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [16] H. Eltaif, Y. Yaddaden, R. Cherif, and Y. Benahmed, "Enhancing biometric authentication efficiency: A hybrid approach exploiting iris modality and leveraging one-class svm," in *2023 IEEE Third International Conference on Signal, Control and Communication (SCC)*. IEEE, 2023, pp. 01–06.
- [17] M. Baccar, Y. Yaddaden, M. Arbane, and R. Cherif, "A hybrid approach based on cnn and extremely randomized trees for the diagnosis of alzheimer's disease stages from mri images," in *2023 IEEE Third International Conference on Signal, Control and Communication (SCC)*. IEEE, 2023, pp. 01–06.
- [18] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the*

IEEE conference on computer vision and pattern recognition, 2016, pp. 2818–2826.

- [19] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, “Mobilenets: Efficient convolutional neural networks for mobile vision applications,” *arXiv preprint arXiv:1704.04861*, 2017.
- [20] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” *arXiv preprint arXiv:1409.1556*, 2014.
- [21] F. Chollet, “Xception: Deep learning with depthwise separable convolutions,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1251–1258.
- [22] M. Long, Y. Cao, J. Wang, and M. Jordan, “Learning transferable features with deep adaptation networks,” in *International conference on machine learning*. PMLR, 2015, pp. 97–105.
- [23] C. Lugaresi, J. Tang, H. Nash, C. McClanahan, E. Uboweja, M. Hays, F. Zhang, C.-L. Chang, M. Yong, J. Lee *et al.*, “Mediapipe: A framework for perceiving and processing reality,” in *Third workshop on computer vision for AR/VR at IEEE computer vision and pattern recognition (CVPR)*, vol. 2019, 2019.
- [24] S. Abtahi, M. Omidyeganeh, S. Shirmohammadi, and B. Hariri, “Yawdd: A yawning detection dataset,” in *Proceedings of the 5th ACM multimedia systems conference*, 2014, pp. 24–28.
- [25] R. Fusek, “Pupil localization using geodesic distance,” in *Advances in Visual Computing: 13th International Symposium, ISVC 2018, Las Vegas, NV, USA, November 19–21, 2018, Proceedings 13*. Springer, 2018, pp. 433–444.
- [26] A.-U.-I. Rafid, A. I. Chowdhury, A. R. Niloy, and N. Sharmin, “A deep learning based approach for real-time driver drowsiness detection,” in *2021 5th International Conference on Electrical Engineering and Information Communication Technology (ICEEICT)*. IEEE, 2021, pp. 1–5.
- [27] S. Mohanty, S. V. Hegde, S. Prasad, and J. Manikandan, “Design of real-time drowsiness detection system using dlib,” in *2019 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE)*. IEEE, 2019, pp. 1–4.

Annexe B

Titre :

Poster de participation au Forum Innovation, Ingénierie, Informatique et Entrepreneuriat (FI3E) en 2023 & 2024

Introduction - À chaque étape du projet où des résultats préliminaires sont obtenus, je fais des présentations afin d'obtenir des retours et des commentaires permettant d'apporter des améliorations. En 2023, mon poster de FI3E portait sur la partie des signaux physiologiques par contre en 2024 j'ai contribué avec l'étudiante de premier cycle pour un poster concernant la partie des aspects comportementaux. Pour les deux participations, on a obtenu le premier prix dans la catégorie informatique ainsi que le prix coup de cœur du public.

DÉTECTION DE LA FATIGUE AU VOLANT À L'AIDE DE L'INTELLIGENCE ARTIFICIELLE

Dorra Lamouchi (étudiante), Yacine Yaddaden & Raef Cherif

Département de Mathématiques d'Informatique et de Génie, Université du Québec à Rimouski



Problématique

- ✗ La somnolence au volant est une des principales causes des accidents mortels.
- ✗ Les solutions existantes sont relativement complexes, ce qui explique les installations effectuées directement par les constructeurs automobiles, notamment sur les véhicules haut de gamme.

Objectifs fixés

- ✓ Développer un système d'estimation de l'état du conducteur (*somnolant* ou *vigilant*) basé sur les signaux physiologiques en passant par des techniques d'apprentissage automatique [2].
- ✓ Le modèle généré sera exploité par une application mobile pour alerter adéquatement le conducteur afin qu'il puisse s'arrêter s'il atteint un certain seuil de fatigue.

Méthodologie suivie

1. Pré-traitement :

- Conversion du format de données de .edf à .csv afin de faciliter le traitement.

2. Extraction des caractéristiques :

- Utilisation des *métriques statistiques* (*max*, *min*, *mean*, *kurtosis*, *skewness*, ...),
- Représentation fréquentielle en utilisant la **transformée de Fourier**, puis extraire :
 - **Densité spectrale de puissance** mesure la quantité de puissance contenue dans la plage de fréquences.
 - **Entropie spectrale** quantifie l'incertitude d'une distribution de probabilité des amplitudes de chaque composante fréquentielle.
 - **Centroïde spectral** caractérise la fréquence moyenne ou le centre de gravité d'un spectre de fréquence.
- Extraction des coefficients de la **transformée en Ondelettes** (représentation temps/fréquence)
- Extraction des coefficients de la **transformée en Cosinus** (représentation fréquentielle).

3. Classification :

- Afin d'évaluer les performances de la méthode proposée, **Lazy Predict**, a été exploité afin de tester plusieurs algorithmes d'apprentissage automatique.
- Le logiciel de forage de données **Orange**, permettant de créer et analyser les données avec une interface graphique conviviale.
- Parmi les algorithmes d'apprentissage, uniquement les algorithmes suivants ont été comparés : arbre de décision, machine à vecteur de support, forêt aléatoire et réseau de neurones.

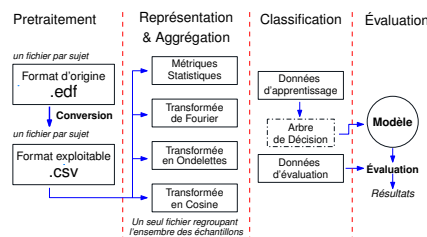


FIGURE 1 – Processus global de la solution proposée.

Bases de données

La **DROZY ULg de somnolence** [1] est *multimodale* et annotée suivant l'échelle du sommeil de Karolinska (**KSS**). Elle contient les signaux (voir Figure 1) :

- **Électroencéphalogramme (EEG)** → canaux C3, C4, Cz, Fz et Pz.
- **Électrooculogramme (EOG)** → électrodes verticale et horizontale.
- **Électrocardiogramme (ECG)**
- **Électromyogramme (EMG)**

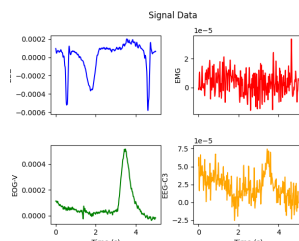


FIGURE 2 – Aperçu des signaux utilisés

Résultats obtenus

La Table 1 présente la *précision*, le *score F1* et le *rappel* obtenus suite à l'évaluation des performances des quatre modèles générés après la phase d'apprentissage.

TABLE 1 – Résultats obtenus pour quatre classifieurs.

Méthode	Précision	Score F1	Rappel
Machine à Vecteurs de Support	75%	75%	76%
Forêt aléatoire	69%	70%	70%
Réseau de neurones	66%	66%	67%
→ Arbre de décision	88%	86%	88%

L'**arbre de décision** a donné les meilleurs résultats. Les paramètres utilisés sont les suivants : *critère gini*, *profondeur maximale* = 100, *nombre minimal d'échantillons par feuille* = 5 et *nombre minimal d'observations* = 19.

		Predicted		
		0.0	1.0	Σ
Actual	0.0	9	7	16
	1.0	0	20	20
Σ		9	27	36

FIGURE 3 – Confusion matrix for the decision tree.

Conclusion

- ✓ Pour conclure, malgré le manque de données (seulement 36 fichiers), la méthode proposée a pu être évaluée et a permis d'obtenir des résultats prometteurs.
- ✓ Comme perspectives, il sera préférable d'appliquer d'autres pré-traitements et techniques d'extraction de caractéristiques afin d'atteindre de meilleures performances.

Références

- [1] Quentin Massoz, Thomas Langohr, Clémentine François, et Jacques G. Verly. The ulg multimodality drowsiness database (called drozy) and examples of use. In *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 1–7, 2016.
- [2] Sazali Yaacob, Nur Afrina Izzati Affandi, Pranesh Krishnan, Amir Rasyadan, Muhyi Yaakop, and Firdaus Mohamed. Drowsiness detection using eeg and eeg signals. In *2020 IEEE 2nd International Conference on Artificial Intelligence in Engineering and Technology (ICAET)*, pages 1–5. IEEE, 2020.

Problématique

- ✗ La somnolence au volant est l'un des facteurs majeurs qui causent des accidents de la route,
- ✗ Les méthodes de détection de la somnolence au volant basées sur l'intelligence artificielle restent à améliorer [3],
- ✗ L'identification des signes de fatigue tels que le bâillement et le clignement des yeux représente un défi important.

Objectifs

- ✓ Développer deux modèles de détection des clignements des yeux et des bâillements à l'aide des architectures de réseaux neuronaux convolutifs (CNN) afin de détecter la somnolence au volant et ainsi prévenir les accidents de la route,
- ✓ Évaluer l'efficacité des modèles développés pour détecter les signes de fatigue en utilisant des bases de données de référence.

Méthodologie suivie

La Figure 1 représente le système proposé.

1. Pré-traitements

- Conversion des vidéos en images,
- Extraction des régions d'intérêt (la bouche et les yeux),
- Redimensionnement des images à 128×128 pixels,
- Étiquetage des images en fonction du seuil d'ouverture de la bouche et des yeux.

2. Les réseaux de neurones convolutifs

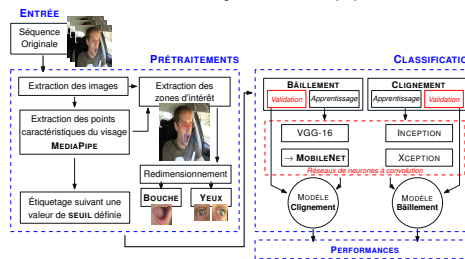
- Utilisation de modèles de réseaux neuronaux convolutifs pré-entraînés : VGG-16, MobileNet, Inception-V3 et Xception.

- Adapter et personnaliser les architectures afin d'améliorer les performances de reconnaissance.

3. Entraînement et validation des modèles

- Adoption de la validation croisée à 10 plis,
- Les métriques de performance : l'exactitude, la précision, le rappel et le score F1.

FIGURE 1 – Processus global de la solution proposée.



Les bases de données

- Dans le contexte de ce travail de recherche, plusieurs bases de données de référence ont été utilisées pour l'apprentissage et l'évaluation des modèles,
- Les détails de YawDD [1], MRL Eye [2], Yawn Dataset et Drowsiness Dataset sont présentés dans la Table 1.

TABLE 1 – Informations sur les bases de données utilisées.

Catégorie	Nom	Échantillons	Type
Clignements	MRL Eye	8000	images infrarouges
Drowsiness	Drowsiness	1452	images
Clignements et Bâillements	YawDD	16 000	vidéos
Bâillements	Yawn	5119	images

Résultats obtenus

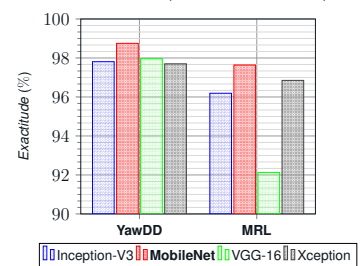
La Table 2 présente les métriques de performance du modèle MobileNet, développé et validé par la validation croisée sur les trois bases de données liées aux clignements et les deux bases de données liées aux bâillements.

TABLE 2 – Performances de MobileNet sur l'ensemble des bases de données.

Bases de données	Précision	Rappel	Score F1	Exactitude
MRL Eye	97,6%	97,6%	97,6%	97,6%
Drowsiness	99,5%	99,5%	99,5%	99,5%
YawDD (clignements)	99,3%	99,3%	99,3%	99,3%
YawDD (bâillements)	98,8%	98,8%	98,7%	98,8%
Yawn	98,4%	98,4%	98,4%	98,4%

MobileNet a les meilleures performances pour la détection des bâillements et des clignements des yeux comme montré à travers la Figure 2.

FIGURE 2 – Performances des quatre réseaux de neurones profonds.



La Table 3 représente les performances du réseau de neurones MobileNet avec la base de données de référence YawDD.

TABLE 3 – Matrices de confusion obtenues : (a) clignements & (b) bâillements.

Prédit	Prédit		Prédit	Prédit	
	Alerte	Somnolent		Alerte	Somnolent
Réel			Réel		
Alerte	99,1%	0,9%	Alerte	98,6%	1,4%
Somnolent	0,4%	99,6%	Somnolent	1,1%	98,9%
(a)			(b)		

Conclusion

- ✓ Les résultats obtenus soulignent l'efficacité des CNN pour les tâches de classification d'images, ce qui est encourageant pour la conception des systèmes de détection de fatigue au volant.,
- ✓ En perspective, il serait intéressant d'explorer d'autres architectures de CNN, telles que ResNet, qui a montré de bons résultats pour des tâches similaires.

Références

- [1] Shabnam Abtahi, Mona Omidyeganeh, Shervin Shirmohammadi, and Behnoosh Harii. Yawdd : A yawning detection dataset. In *Proceedings of the 5th ACM multimedia systems conference*, pages 24–28, 2014.
- [2] R. Fusek. Pupil localization using geodesic distance. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11241 LNCS :433–444, 2018.
- [3] Dorra Lamouchi, Yacine Yaddaden, and Raef Cherif. Comparative study : Physiological-based driver drowsiness detection utilizing traditional and hybrid methods. In *2024 8th international conference on image and signal processing and their applications (ISPA)*, pages 1–8. IEEE, 2024.

Bibliographie

- [1] Shabnam ABTAHI et al. « YawDD : A yawning detection dataset ». In : *Proceedings of the 5th ACM multimedia systems conference*. 2014, p. 24-28. DOI : 10.1145/2557642.2563678.
- [2] Faisal Dharma ADHINATA, Diovianto Putra RAKHMADANI et Danur WIJAYANTO. « Fatigue detection on face image using FaceNet algorithm and K-nearest neighbor classifier ». In : *Journal of Information Systems Engineering and Business Intelligence* 7.1 (2021), p. 22-30.
- [3] Torbjörn ÅKERSTEDT et al. « Mental fatigue, work and sleep ». In : *Journal of psychosomatic Research* 57.5 (2004), p. 427-433. DOI : 10.1016/j.jpsychores.2003.12.001.
- [4] Yaman ALBADAWI, Maen TAKRURI et Mohammed AWAD. « A review of recent developments in driver drowsiness detection systems ». In : *Sensors* 22.5 (2022), p. 2069. DOI : 10.3390/s22052069.
- [5] Riad ALHARBAY et al. « Fatigue state detection for tired persons in presence of driving periods ». In : *IEEE Access* 10 (2022), p. 79403-79418. DOI : 10.1109/ACCESS.2022.3185251.
- [6] Andrea AMIDEI et al. « Driver drowsiness detection : a machine learning approach on skin conductance ». In : *Sensors* 23.8 (2023), p. 4004. DOI : 10.3390/s23084004.

-
- [7] Muhammad AWAIS, Nasreen BADRUDDIN et Micheal DRIEBERG. « A hybrid approach to detect driver drowsiness utilizing physiological signals to improve system performance and wearability ». In : *Sensors* 17.9 (2017), p. 1991. DOI : 10.3390/s17091991.
- [8] Tanya BAFNA et John Paulin HANSEN. « Mental fatigue measurement using eye metrics : A systematic literature review ». In : *Psychophysiology* 58.6 (2021), e13828. DOI : 10.1111/psyp.13828.
- [9] Samy BAKHEET et Ayoub AL-HAMADI. « A framework for instantaneous driver drowsiness detection based on improved HOG features and naïve Bayesian classification ». In : *Brain Sciences* 11.2 (2021), p. 240. DOI : 10.3390/brainsci11020240.
- [10] Lawrence BARR et al. « A review and evaluation of emerging driver fatigue detection measures and technologies ». In : *National Transportation Systems Center, Cambridge. US Department of Transportation, Washington* 19.9 (2005), p. 51-59.
- [11] Colin BUCHANAN. *Traffic in Towns : A study of the long term problems of traffic in urban areas*. Routledge, 2015.
- [12] Siwar CHAABENE et al. « Convolutional neural network for drowsiness detection using EEG signals ». In : *Sensors* 21.5 (2021), p. 1734. DOI : 10.3390/s21051734.
- [13] Mario I CHACON-MURGUÍA et Claudia PRIETO-RESENDIZ. « Detecting driver drowsiness : A survey of system designs and technology ». In : *IEEE Consumer Electronics Magazine* 4.4 (2015), p. 107-119. DOI : 10.1109/MCE.2015.2463373.
- [14] Hong CHEN, Yang ZHAO et Xiaotong MA. « Critical factors analysis of severe traffic accidents based on Bayesian network in China ». In :

- Journal of advanced transportation* 2020.1 (2020), p. 8878265. DOI : 10 . 1155/2020/8878265.
- [15] Venkata Rami Reddy CHIRRA, Srinivasulu Reddy UYYALA et Venkata Krishna Kishore KOLLI. « Deep CNN : A Machine Learning Approach for Driver Drowsiness Detection Based on Eye State. » In : *Rev. d'Intelligence Artif.* 33.6 (2019), p. 461-466. DOI : 10.18280/ria.330609.
- [16] Neal F COOK et Jennifer RP BOORE. « Managing patients suffering from acute and chronic fatigue ». In : *British Journal of Nursing* 6.14 (1997), p. 811-815. DOI : 10.12968/bjon.1997.6.14.811.
- [17] Hugo D CRITCHLEY et Sarah N GARFINKEL. « The influence of physiological signals on cognition ». In : *Current Opinion in Behavioral Sciences* 19 (2018), p. 13-18. DOI : 10.1016/j.cobeha.2017.08.014.
- [18] Messaoud DOUDOU, Abdelmadjid BOUABDALLAH et Véronique BERGE-CHERFAOUI. « Driver drowsiness measurement technologies : Current research, market solutions, and challenges ». In : *International Journal of Intelligent Transportation Systems Research* 18 (2020), p. 297-319. DOI : 10.1007/s13177-019-00199-w.
- [19] Samy Abd EL-NABI et al. « Machine learning and deep learning techniques for driver fatigue and drowsiness detection : a review ». In : *Multimedia Tools and Applications* 83.3 (2024), p. 9441-9477. DOI : 10.1007/s11042-023-15054-0.
- [20] Ariel EPHRAT et al. « Looking to listen at the cocktail party : A speaker-independent audio-visual model for speech separation ». In : *arXiv preprint arXiv :1804.03619* (2018).
- [21] Yuki FUJITA et Akira SHIBATA. « Relationship between traffic fatalities and drunk driving in Japan ». In : *Traffic injury prevention* 7.4 (2006), p. 325-327.

- [22] R. FUSEK. « Pupil localization using geodesic distance ». In : *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 11241 LNCS (2018), p. 433-444. DOI : 10.1007/978-3-030-03801-4_38.
- [23] Reza GHODDOOSIAN, Marnim GALIB et Vassilis ATHITSOS. « A realistic dataset and baseline temporal model for early drowsiness detection ». In : *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. 2019, p. 0-0.
- [24] Fudail HASAN et Alexey KASHEVNIK. « State-of-the-art analysis of modern drowsiness detection algorithms based on computer vision ». In : *2021 29th Conference of Open Innovations Association (FRUCT)*. IEEE. 2021, p. 141-149. DOI : 10.23919/FRUCT52173.2021.9435480.
- [25] Md Mahmudul HASAN, Christopher N WATLING et Grégoire S LARUE. « Physiological signal-based drowsiness detection using machine learning : Singular and hybrid signal approaches ». In : *Journal of safety research* 80 (2022), p. 215-225. DOI : 10.1016/j.jsr.2021.12.001.
- [26] Xinyun HU et Gabriel LODEWIJKS. « Detecting fatigue in car drivers and aircraft pilots by using non-invasive measures : The value of differentiation of sleepiness and mental fatigue ». In : *Journal of safety research* 72 (2020), p. 173-187. DOI : 10.1016/j.jsr.2019.12.015.
- [27] Marwa K HUSSEIN et al. « Driver drowsiness detection techniques : A survey ». In : *2021 1st Babylon International Conference on Information Technology and Science (BICITS)*. IEEE. 2021, p. 45-51. DOI : 10.1109/BICITS51482.2021.9509912.
- [28] Xuan-Phung HUYNH, Sang-Min PARK et Yong-Guk KIM. « Detection of driver drowsiness using 3D deep neural network and semi-supervised

- gradient boosting machine ». In : *Computer Vision–ACCV 2016 Workshops : ACCV 2016 International Workshops, Taipei, Taiwan, November 20–24, 2016, Revised Selected Papers, Part III* 13. Springer. 2017, p. 134-145. DOI : 10.1007/978-3-319-54526-4_10.
- [29] M Ahmad KAMRAN, Malik M Naeem MANNAN et Myung Yung JEONG. « Drowsiness, fatigue and poor sleep's causes and detection : a comprehensive study ». In : *Ieee Access* 7 (2019), p. 167172-167186. DOI : 10.1109/ACCESS.2019.2951028.
- [30] Hang-Bong KANG. « Various approaches for driver and driving behavior monitoring : A review ». In : *Proceedings of the IEEE International Conference on Computer Vision Workshops*. 2013, p. 616-623.
- [31] Ghanta Sai KRISHNA, Kundrapu SUPRIYA, Jai VARDHAN et al. « Vision transformers and YoloV5 based driver drowsiness detection framework ». In : *arXiv preprint arXiv :2209.01401* (2022). DOI : 10.48550/arXiv.2209.01401.
- [32] Pranesh KRISHNAN et al. « EEG based Drowsiness Detection using Relative Band Power and Short-time Fourier Transform. » In : *J. Robotics Netw. Artif. Life* 7.3 (2020), p. 147-151. DOI : 10.2991/jrnal.k.200909.001.
- [33] Dragan KRIVOKAPIC. « Theories of tiredness in sport ». In : *Sport Mont* 15.1 (2017), p. 35-37.
- [34] Zuojin LI et al. « Online detection of driver fatigue using steering wheel angles for real driving conditions ». In : *Sensors* 17.3 (2017), p. 495. DOI : 10.3390/s17030495.
- [35] Paulo Augusto de LIMA MEDEIROS et al. « Efficient machine learning approach for volunteer eye-blink detection in real-time using webcam ».

- In : *Expert Systems with Applications* 188 (2022), p. 116073. DOI : 10.1016/j.eswa.2021.116073.
- [36] Monika LOHANI, Brennan R PAYNE et David L STRAYER. « A review of psychophysiological measures to assess cognitive states in real-world driving ». In : *Frontiers in human neuroscience* 13 (2019), p. 57. DOI : 10.3389/fnhum.2019.00057.
- [37] S LUBIN et E PRÉVOT. « Bilan de sommeil en médecine du travail ». In : *Médecine du sommeil* 3.8 (2006), p. 47-53.
- [38] Caio Bezerra Souto MAIOR et al. « Real-time classification for autonomous drowsiness detection using eye aspect ratio ». In : *Expert Systems with Applications* 158 (2020), p. 113505. DOI : 10.1016/j.eswa.2020.113505.
- [39] Quentin MASSOZ et al. « The ULg multimodality drowsiness database (called DROZY) and examples of use ». In : *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE. 2016, p. 1-7. DOI : 10.1109/WACV.2016.7477715.
- [40] NATIONAL HIGHWAY TRAFFIC SAFETY ADMINISTRATION. *Countermeasures that work drowsy driving*. <https://www.nhtsa.gov/book/countermeasures-that-work/drowsy-driving>. 2021.
- [41] ORGANISATION MONDIALE DE LA SANTÉ. *La situation mondiale de la sécurité routière 2023*. <https://news.un.org/fr/story/2023/12/1141622>. 2023.
- [42] Mohammad PEIVANDI et al. « Deep learning for detecting multi-level driver fatigue using physiological signals : A comprehensive approach ». In : *Sensors* 23.19 (2023), p. 8171. DOI : 10.3390/s23198171.

- [43] Emma PERKINS et al. « Investigating Physiological and Behavioural Sensing Modalities Towards Drowsiness Detection ». In : *IEEE Sensors Journal* (2023). DOI : 10.1109/JSEN.2023.3326434.
- [44] Ajinkya RAJKAR, Nilima KULKARNI et Aniket RAUT. « Driver drowsiness detection using deep learning ». In : *Applied Information Processing Systems : Proceedings of ICCET 2021*. Springer. 2022, p. 73-82. DOI : 10.1007/978-981-16-2008-9_7.
- [45] Plínio MS RAMOS et al. « Automatic drowsiness detection for safety-critical operations using ensemble models and EEG signals ». In : *Process Safety and Environmental Protection* 164 (2022), p. 566-581. DOI : 10.1016/j.psep.2022.06.039.
- [46] Emma REAM et Alison RICHARDSON. « Fatigue : a concept analysis ». In : *International journal of nursing studies* 33.5 (1996), p. 519-529. DOI : 10.1016/0020-7489(96)00004-1.
- [47] Furkat SAFAROV et al. « Real-time deep learning-based drowsiness detection : leveraging computer-vision and eye-blink analyses for enhanced road safety ». In : *Sensors* 23.14 (2023), p. 6459. DOI : 10.3390/s23146459.
- [48] Rais Mohammad SALMAN et al. « Driver drowsiness detection using ensemble convolutional neural networks on YawDD ». In : *arXiv preprint arXiv :2112.10298* (2021).
- [49] Firdaous SEKKAY et al. « Assessment of physical work demands of long-distance industrial gas delivery truck drivers ». In : *Applied Ergonomics* 90 (2021), p. 103224. DOI : 10.1016/j.apergo.2020.103224.
- [50] Hafeez Ur Rehman SIDDIQUI et al. « Non-invasive driver drowsiness detection system ». In : *Sensors* 21.14 (2021), p. 4833. DOI : 10.3390/s21144833.

- [51] Fengyi SONG et al. « Eyes closeness detection from still images with multi-scale histograms of principal oriented gradients ». In : *Pattern Recognition* 47.9 (2014), p. 2825-2838. DOI : 10.1016/j.patcog.2014.03.024.
- [52] CAI SU-XIAN et al. « Fatigue driving state detection based on vehicle running data ». In : *Journal of Transportation Systems Engineering and Information Technology* 20.4 (2020), p. 77.
- [53] Anna SUNDSTRÖM. « Self-assessment of driving skill—A review from a measurement perspective ». In : *Transportation Research Part F : Traffic Psychology and Behaviour* 11.1 (2008), p. 1-9.
- [54] Yeresime SURESH et al. « Driver drowsiness detection using deep learning ». In : *2021 2nd International Conference on Smart Electronics and Communication (ICOSEC)*. IEEE. 2021, p. 1526-1531. DOI : 10.1109/ICOSEC51865.2021.9591957.
- [55] Reza TAMANANI, Radu MURESAN et Arafat AL-DWEIK. « Estimation of driver vigilance status using real-time facial expression and deep learning ». In : *IEEE Sensors Letters* 5.5 (2021), p. 1-4. DOI : 10.1109/LSENS.2021.3070419.
- [56] Arash TAVAKOLI et Arsalan HEYDARIAN. « Multimodal driver state modeling through unsupervised learning ». In : *Accident Analysis & Prevention* 170 (2022), p. 106640. DOI : 10.1016/j.aap.2022.106640.
- [57] Mario Giovanni TERZANO et al. « Atlas, rules, and recording techniques for the scoring of cyclic alternating pattern (CAP) in human sleep ». In : *Sleep medicine* 2.6 (2001), p. 537-554.
- [58] Ching-Hua WENG, Ying-Hsiu LAI et Shang-Hong LAI. « Driver drowsiness detection via a hierarchical temporal deep belief network ». In :

- Computer Vision–ACCV 2016 Workshops : ACCV 2016 International Workshops, Taipei, Taiwan, November 20-24, 2016, Revised Selected Papers, Part III* 13. Springer. 2017, p. 117-133. DOI : 10.1007/978-3-319-54526-4_9.
- [59] Sazali YAACOB et al. « Drowsiness detection using EEG and ECG signals ». In : *2020 IEEE 2nd International Conference on Artificial Intelligence in Engineering and Technology (IICAET)*. IEEE. 2020, p. 1-5. DOI : 10.1109/IICAET49801.2020.9257867.
- [60] Chuwei YE et al. « Smart Steering Sleeve (S3) : A Non-Intrusive and Integrative Sensing Platform for Driver Physiological Monitoring ». In : *Sensors* 22.19 (2022), p. 7296. DOI : 10.3390/s22197296.
- [61] Jongmin YU et al. « Driver drowsiness detection using condition-adaptive representation learning framework ». In : *IEEE transactions on intelligent transportation systems* 20.11 (2018), p. 4206-4218. DOI : 10.1109/TITS.2018.2883823.
- [62] Liang ZHAO et al. « Data-driven learning fatigue detection system : A multimodal fusion approach of ECG (electrocardiogram) and video signals ». In : *Measurement* 201 (2022), p. 111648. DOI : 10.1016/j.measurement.2022.111648.
- [63] Zijie ZHOU et al. « Driver vigilance detection based on deep learning with fused thermal image information for public transportation ». In : *Engineering Applications of Artificial Intelligence* 124 (2023), p. 106604. DOI : 10.1016/j.engappai.2023.106604.