

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE
SCIENTIFIQUE



UNIVERSITE MOULOU MAMMERI DE TIZI-OUZOU
FACULTE DE GENIE ELECTRIQUE ET INFORMATIQUE
DEPARTEMENT : ELECTRONIQUE
MÉMOIRE DE FIN D'ÉTUDES DE MASTER ACADÉMIQUE

Domaine : **Sciences et Technologie.**

Filière : **Génie Electrique.**

Spécialité : **Electronique des Systèmes Embarqués.**

THEME

**Etude et réalisation d'un système
portatif pour la détection de
somnolence au volant**

Présenté par :

ABED Jugurtha.

BEN ABOU Kahina.

Encadré par :

Mr. ALOUACHE Djamel

Promotion 2023-2024



Remerciements

Nous remercions **Allah** le tout puissant, qui nous a donné la force et la patience pour l'accomplissement de ce travail.

Nos remerciements, les plus vifs, notre profonde gratitude et nos respects s'adressent à notre encadreur **Mr ALOUACHE**.

Nos vifs remerciements vont également au président et aux membres de jury d'avoir accepté d'évaluer notre travail.

Nous remercions tous ceux qui ont contribué de près ou de loin à la réalisation de ce travail.

Enfin nos profonds et sincères remerciements vont vers nos parents et non grands parents qui nous ont toujours soutenus tout au long de notre cursus.

KAHINA et DJUGURTHA



Dédicaces

Je dédie ce travail à ceux qui m'ont soutenu et encouragé tout au long de ce parcours :

À mes parents, pour leur amour inconditionnel, leur soutien infaillible, et leur confiance en moi. Merci pour avoir toujours cru en moi et pour m'avoir donné les moyens de poursuivre mes rêves.

A mes sœurs **Taous** et **Djahida** qui m'ont toujours épaulé, et surtout à leurs enfants Aris et Moustapha, qui ont apporté un grand bonheur dans notre vie. Sans oublier mon beau-frère **Belkacem**, qui a été un ami et plus qu'un frère pour moi.

A mes grands-parents Youcef et Zahra que la terre leur soit légère, qui ont toujours été fiers de moi, il aurait été très heureux de me voir réussir et arriver à cette étape de mon parcours.

À mes amis **YANIS, FOUAD, AMINE, AZIZE, NASSIM, OMAR, YANNI, ABDREHMAN** et **MASSI** pour leur amitié sincère, leur écoute et leurs encouragements, même dans les moments les plus difficiles. Merci pour les discussions et les moments partagés qui ont rendu cette aventure plus légère et agréable.

À mon encadreur, **Mr ALOUACHE** Djamal pour sa patience, ses conseils avisés, et son soutien tout au long de ce travail. Votre expertise et vos encouragements ont été essentiels à la réussite de ce projet.

A ma binôme Kahina et toute sa famille.

Et à tous ceux qui, de près ou de loin, ont contribué à la réussite de ce travail.

JUGURTHA



Dédicaces

Je dédie ce travail à :

À ma chère mère,

À mon cher père,

Merci pour votre amour inconditionnel, votre soutien indéfectible et vos encouragements tout au long de ce voyage. C'est grâce à vous que j'ai trouvé la force et la détermination pour atteindre cet objectif.

À mes grands-parents,

Vous avez toujours cru en moi quand je doutais de moi. Votre sagesse, vos conseils et vos prières m'ont accompagné tout au long de ce parcours. Je leur souhaite une bonne santé.

À mon frère,

Pour son soutien moral et son conseil tout au long de mes études.

À mon cher oncle,

Qui m'a aidé et supporté dans les moments difficiles.

À mon encadreur Mr ALOUACHE,

Pour son soutien inestimable tout au long de ce projet. Grâce à ces conseils éclairés, son expertise et sa disponibilité, j'ai pu surmonter les défis rencontrés et mener à bien ce travail.

À mes amies.

À mon binôme : **Jugurtha.**

KAHINA

Résumé

Ce projet de fin d'études se concentre sur la conception et la réalisation d'un système portable de détection de somnolence au volant, utilisant la méthode de Viola-Jones et la transformée de Hough circulaire. L'objectif est de prévenir les accidents de la route causés par la somnolence du conducteur en développant un dispositif embarqué capable de surveiller les signes de fatigue.

Le système est basé sur un Raspberry Pi 4, équipé d'une caméra infrarouge pour l'acquisition d'images en temps réel, et des baffles pour émettre des alertes sonores lorsque des signes de somnolence sont détectés. La méthode de Viola-Jones est utilisée pour la détection du visage et des yeux, tandis que la transformée de Hough circulaire permet de détecter l'état des yeux, en particulier pour identifier la fermeture ou l'ouverture des paupières, ce qui est un indicateur clé de la somnolence.

Les images capturées par la caméra infrarouge sont analysées en temps réel à l'aide de ces algorithmes, et lorsque le système détecte un état de somnolence chez le conducteur, une alerte sonore est déclenchée pour prévenir un accident potentiel.

Ce projet propose une solution portable et efficace pour améliorer la sécurité routière, avec des possibilités d'amélioration et d'intégration dans des systèmes intelligents de véhicules.

Mots clés : somnolence au volant, détection de somnolence, suivie du visage.

Abstract

This final year project focuses on the design and implementation of a portable system for detecting driver drowsiness, using the Viola-Jones method and the Hough Transform. The objective is to prevent road accidents caused by driver drowsiness by developing an embedded device capable of monitoring signs of fatigue.

The system is built around a Raspberry Pi 4, equipped with an infrared camera for real-time image acquisition, and speakers to emit audible alerts when signs of drowsiness are detected. The Viola-Jones method is used for face and eye detection, while the Hough Transform is employed to detect the state of the eyes, particularly to identify whether the eyelids are open or closed, which is a key indicator of drowsiness.

The images captured by the infrared camera are analyzed in real-time using these algorithms, and when the system detects drowsiness in the driver, an audible alert is triggered to prevent a potential accident.

This project offers a portable and efficient solution to improve road safety, with potential for further improvements and integration into intelligent vehicle systems.

Keywords: Drowsiness at the wheel, drowsiness detection, facial follow-up.

Liste_des tableaux

Liste des figures

Liste des abréviations

INTRODUCTION GENERALE 1

Chapitre 1 : Généralités sur la vision Artificielle

1.1	Préambule.....	2
1.2	Historique.....	2
1.3	Définition de l'image.....	2
1.4	Image numérique (numérisée).....	3
1.5	Les différents types d'image numérique	3
1.5.1	Image vectorielle	3
1.5.2	Image matricielle.....	4
1.6	Caractéristiques d'une image numérique	5
1.6.1	Pixel.....	5
1.6.2	La résolution.....	5
1.6.3	Dimension	6
1.6.4	La texture.....	6
1.6.5	Le bruit	7
1.6.6	La luminance	7
1.6.7	Contours	8
1.6.8	Histogramme	8
1.6.9	Egalisation d'histogramme.....	9
1.6.10	Espace de couleur.....	9
1.7	Segmentation.....	11
1.7.1	Binarisation	12
1.7.1.1	Binarisation globale.....	12
1.7.1.2	Binarisation locale.....	13
1.7.2	La détection des contours	13
1.8	Prétraitement	13
1.8.1	Filtrage.....	13
1.8.1.1	Filtres linaires	14
1.8.1.2	Filtres non linaires	15
1.9	Vidéo.....	15

1.9.1 Définitions de la vidéo.....	15
1.9.2 Vidéo numérique et analogique.....	16
La vidéo analogique	16
La vidéo numérique.....	16
1.9.3 Les vidéos entrelacées	16
1.10 Acquisition des données.....	17
1.11 Discussion	17

CHAPITRE 2 : Les méthodes de détection de la somnolence

2.1 Préambule.....	18
2.2 La somnolence.....	18
2.3 Les facteurs influençant la somnolence.....	19
2.3.1 L'insuffisance et la qualité du sommeil.....	19
2.3.2 Horaire de conduite	20
2.3.3 Durée de conduite.....	20
2.3.4 Environnement de conduite.....	20
2.3.5 Consommation de médicaments et autres substances	21
2.4 Les signes précurseurs de somnolence	21
2.5 Les méthodes de détection de somnolence.....	22
2.5.1 Méthode basée sur la physiologie.....	23
2.5.2 Méthode basée sur la performance du conducteur	24
2.5.3 Méthode basée sur le traitement d'image.....	25
2.5.3.1 PERCLOSE.....	26
2.5.3.2 Calcule l'EAR	26
2.5.3.3 Des réseaux LSTM.....	27
2.5.3.4 Analyse labiale	27
2.5.3.5 L'algorithme de Viola et Jones.....	27
2.5.3.6 Réseaux de neurone convolutifs.....	28
2.5.3.7 Calcule des distances oculaires.....	28
2.5.3.8 La transformée de Hough circulaire	30
2.6 Discussion	30

Chapitre 3 : Détection de la somnolence par vision artificielle

3.1 Préambule.....	31
Avertissements	31
3.2 Région d'intérêt (ROI).....	31
3.3 Reconnaissance d'objets	32

3.4 La détection du visage	33
3.4.1 Difficultés de la détection de visages	34
3.4.1.1 L'échelle.....	34
3.4.1.2 Pose	34
3.4.1.3 Occlusion.....	35
3.5 La Méthode de Viola et Jones	36
3.5.1 Définition.....	36
3.5.2 Fonctionnements	37
3.5.3 Principe.....	38
3.5.4 Eléments de la méthode.....	38
3.5.5 Caractéristiques de Haar.....	39
3.5.6 Image intégrale	39
3.5.7 Filtre rectangulaire.....	41
3.5.8 Cascade attentionnelle	42
3.6 La détection des yeux	42
3.7 Détection l'état des yeux	43
3.7.1 Détermination du rayon et du centre de la pupille et de l'iris	43
3.7.2 Détermination du Rayon de l'iris	44
3.8 Avertissements	45
3.9 Discussion	46
Chapitre 4 : Implémentation et résultats	
4.1 Préambule.....	46
4.2 Matériels et méthodes.....	46
4.2.1 Matériel	47
4.2.2 Méthodes	47
4.3 Détection du visage	49
4.4 Détection des yeux	50
4.5 Description de notre système.....	51
4.6 Résultats	53
4.7 Discussion	54
<i>CONCLUSION GENERALE</i>	55
<i>BIBLIOGRAPHIE</i>	56

Liste des tableaux

Tableau 2.1 : Bande de Fréquence EEG, (extrait de Larue, 2010)	24
---	----

Liste des figures

Figure 1.1 : Représentation d'une image numérique.	3
Figure 1.2 : Exemple d'une image vectorielle.	4
Figure 1.3 : Exemple d'une image matricielle.	4
Figure 1.4 : Groupe de pixel formant la lettre A.	5
Figure 1.5 : Une image sous 3 résolutions différentes.	6
Figure 1.6 : Les dimensions d'une image.	6
Figure 1.7 : Représenter la texture.	7
Figure 1.8 : Image sans bruit.	7
Figure 1.9 : Image avec bruit.	7
Figure 1.10 : Contours d'une image.	8
Figure 1.11 : Image avec son histogramme.	9
Figure 1.12 : Illustration d'un histogramme égalisé	9
Figure 1.13 : Palette de niveaux de gris (grayscale).	10
Figure 1.14 : Palette de RGB.	10
Figure 1.15 : Palette de HSV.	11
Figure 1.16 : Image segmentée.	12
Figure 1.17 : Image binarisée.	12
Figure 1.18 : Exemple de détection des contours.	13
Figure 1.19 : Exemple d'un filtre médian.	15
Figure 1.20 : Ensemble d'image formant une vidéo.	16
Figure 2.1 : Les statistiques des accidents de la route de 2015 à 2019.	19
Figure 2.2 : Les différents facteurs pouvant être liés à la somnolence.	19
Figure 2.3 : Les horaires à risque pour conduire.	20
Figure 2.4 : Symboles pouvant apparaître sur la boîte d'un médicament.	21

Figure 2.5 : Les signes de fatigue et de somnolence.....	22
Figure 2.6 : Diverses méthodes pour détecter la somnolence.....	22
Figure 2.7 : Illustration des capteurs utilisés dans les méthodes axées sur la performance du conducteur.....	23
Figure 2.8 : Détecteur de ligne sur véhicule.....	25
Figure 2.9 : Position des repères des yeux pour le calcul de l’EAR avec l’œil ouvert et l’œil fermé.....	26
Figure 2.10 : Architecteur utilisé dans l'approche de Bhargava Reddy et <i>al</i>	28
Figure 2.11 : Distance oculaire prévue (points 37 à 40 et 43 à 46) Pour recadrer le ROI.	29
Figure 2.12 : (a) image de l’œil, (b) différentes images de contours par la méthode Canny selon que l'on considère des contours diagonaux, (c) horizontaux ou (d) verticaux.....	30
Figure 3.1 : Schéma synoptique de la méthode proposée.	31
Figure 3.2 : Illustration qui montre les régions d’intérêt.....	32
Figure 3.3 : Un exemple de détection de la position d'objets.....	32
Figure 3.4 : Visages à différentes échelles.....	34
Figure 3.5 : Visages à différentes poses.....	35
Figure 3.6 : Visage occlus.....	35
Figure 3.7 : Schème synoptique de l’algorithme Viola et Jones.....	36
Figure 3.8_ : Un exemple de détection de visage par la méthode de Viola et Jones.....	36
Figure 3.9 : Exemple illustré de caractéristiques pseudo-Haar.....	37
Figure 3.10 : Illustration des caractéristiques de Haar : (A, b) deux rectangles, (c) trois rectangles, (d) quatre rectangles.....	39
Figure 3.11 : La valeur de l’image intégrale au point (x, y).....	40
Figure 3.12 : Exemple d’image intégrale.....	40
Figure 3.13 : Le filtre rectangulaire de Haar est utilisé par la cascade.	41
Figure 3.14 : Architecture de cascade.	42
Figure 3.15 : Schéma synoptique qui illustre les étapes de détection du centre de la pupille. ...	43

Figure 3.16 : Détermination du rayon et du centre de la pupille.....	43
Figure 3.17 : Détermination du Rayon de l'iris.	44
Figure 3.18 : (a) région de l'iris, (b) localisation de la pupille et de l'iris	44
Figure 4.1 : Architecture du système.	46
Figure 4.2 : Interface d'installation du système d'exploitation Raspberry Pi4 à l'aide de Raspberry Pi Imager.....	47
Figure 4.3 : Logiciel permettant la recherche des adresses IP.	48
Figure 4.4 : Logiciel permettant la Gestion à Distance.....	48
Figure 4.5 : Image illustrative de l'installation d'open cv et python.	49
Figure 4.6 : Image illustrative de la version récente d'open cv.	49
Figure 4.7 : Image illustrative de la détection du visage.....	50
Figure 4.8 : Image illustrative de la détection des yeux.....	50
Figure 4.9 : Organigramme de notre système.	52
Figure 4.10 : Image illustrative d'un conducteur en état de vigilance.	53
Figure 4.11 : Image illustrative d'un conducteur en état de somnolence.....	53
Figure 4.12 : Image illustrative d'un conducteur dont le regard est fixe à droite.	54
Figure 4.13 : Image illustrative d'un conducteur dont le regard est fixé à droite.	54

Liste des abréviations

Dpi	dots per inch.
Ppp	point par pouce.
RGB	rouge, vert, bleu.
HSV	Hue, Saturation, Value.
ROI	Région of Interest.
PAL	Phase Alternating Line.
NTSC	National Television System Committee.
SECAM	Séquentiel Couleur A Mémoire.
CCD	Charge Coupled Device.
EEG	Electroencephalographic.
EOG	Electro-oculographie.
ECG	Electrocardiography.
SVM	Support Vector Machine.
PNN	Prototype Nearest Neighbor.
MLP	Multi-Layer Perceptron.
KNN	K-Nearest Neighbors.
SDLP	Standard Deviation of Lane Position.
KSS	Karolinska Sleepiness Scale.
PERCLOSE	PERcentage of CLOSure.
EAR	Eye Aspect Ratio.
LSTM	Long short-term memory.
CNN	Convolutional Neural network.
RNN	Recurrent Neural Network.
MTCNN	Multi-Task Cascaded Convolutional Neural Networks.
IR	Infra-Rouge.
AdaBoost	Adaptive boosting.
THC	Transformée de Hough circulaire.

INTRODUCTION GENERALE

Selon l’OMS, environ 1,35 million de personnes meurent chaque année dans les accidents de la route à travers le monde et 20 à 50 millions de personnes blessées et gardant généralement des séquelles d’invalidité. La mortalité routière touche plusieurs pays et l’Algérie en fait partie. Selon les statistiques enregistrées par le centre national de la prévention routière (CNPR), la somnolence au volant est parmi les principales causes d’accidents, contribuant à 10% de ces derniers.

Afin de prévenir les accidents et de réduire le nombre de décès et de blessés, de nombreuses études ont été menées et diverses technologies ont été développées (comme les airbags, la ceinture de sécurité, ABS, etc...). Au cours des dernières années, une technique innovante a été présentée et mise en œuvre dans les voitures à la mode (telles que : Benz, Volvo, Audi, BMW, etc...). Cette technique est connue sous le nom de « système de détection de somnolence au volant du pilote.

L’objectif principal de notre mémoire est de réaliser un système portatif à base d’une carte Raspberry Pi4 et d’une caméra infrarouge qui permet de détecter la somnolence du conducteur afin d’aider ce dernier à éviter les accidents routiers causés par la collision, ce système peut reconnaître les visages et les yeux. Cette détection se fait en fonction de l’état des yeux en envoyant un signal sonore pour réveiller le conducteur.

Notre mémoire est organisé en quatre chapitres comme suit :

Le premier chapitre intitulé « Généralité sur la vision artificielle » explique les concepts de base sur le traitement d’image et vidéo.

Le deuxième chapitre intitulé « les méthodes pour la détection de la somnolence » porte sur la présentation de différentes méthodes pour détecter la somnolence et de divers résultats de recherche proposés par différents chercheurs.

Le troisième chapitre intitulé « détection de la somnolence par la vision artificielle » présente la méthode Viola et Jones pour la détection du visage et des yeux, ainsi que la transformée de Hough circulaire pour la détermination de l’état des yeux.

Le dernier chapitre « résultat et discussion » sera consacré à l’évaluation de la méthode proposée et les résultats obtenus.

CHAPITRE 1 :
Généralités sur la vision
Artificielle

1.1 Préambule

De nos jours, le domaine du traitement d'images est extrêmement vaste et a connu une évolution significative au cours des dernières dizaines d'années. Ce secteur a joué un rôle dans le développement de divers secteurs tels que la médecine, l'astronomie, l'informatique, la surveillance et bien d'autres encore. Le traitement d'image peut être défini comme l'ensemble des méthodes et techniques opérant sur l'image afin d'extraire les informations les plus pertinentes ou tout simplement pour fournir une image plus perceptible à l'œil humain.

Le traitement d'image repose sur trois étapes fondamentales : l'importation de l'image à l'aide d'outils d'acquisition d'image tels que la caméra, l'analyse et la manipulation de l'image, et enfin, l'extraction des informations.

Dans ce chapitre, nous abordons les généralités nécessaires à la compréhension des techniques de traitement d'images, prétraitement et Vidéos.

1.2 Historique

L'image constitue l'un des moyens les plus importants de la communication. Le traitement d'image a commencé à être étudié dans les années 1920 pour la transmission d'image par le câble sous-marin allant de New York à Londres. Harry G. Bartholomew et Maynard D. Macfarlane effectuèrent la première digitalisation d'image avec compression pour envoyer des fax de Londres à New York. Le temps de transfert passa ainsi de plus d'une semaine à moins de trois heures. Il n'y a pas vraiment eu dévolution pour l'image par la suite jusqu'à la période d'après-guerre.

1.3 Définition de l'image

L'image est une représentation d'une personne ou d'un objet par la peinture, le dessin, la photographie, le film...etc. C'est aussi un ensemble structure d'informations qui, après affichage sur l'écran, ont une signification pour l'œil humain.

Elle peut être décrite sous la forme d'une fonction $I(x, y)$ analogique continue, définie dans un domaine borne, tel que le x et y sont les coordonnées spatiales d'un point de l'image et I est une fonction d'intensité lumineuse et de couleur. Sous cet aspect, l'image est exploitable par la machine, ce qui nécessite sa numérisation.

1.4 Image numérique (numérisée)

L'image numérique est l'image dont la surface est divisée en éléments de taille fixe appelés cellules ou pixels, ayant chacun comme caractéristiques un niveau de gris ou de couleurs.

La numérisation d'une image est la conversion de celle-ci de son état analogique en une image numérique représentée par une matrice bidimensionnelle de valeurs numérique $F(x, y)$, comme montré dans la figure 1.2 où :

x, y : coordonnées caractéristiques d'un point d'image.

$F(x, y)$: niveau d'intensité.

La valeur en chaque point exprime la mesure d'intensité lumineuse perçue par le capteur. [1]

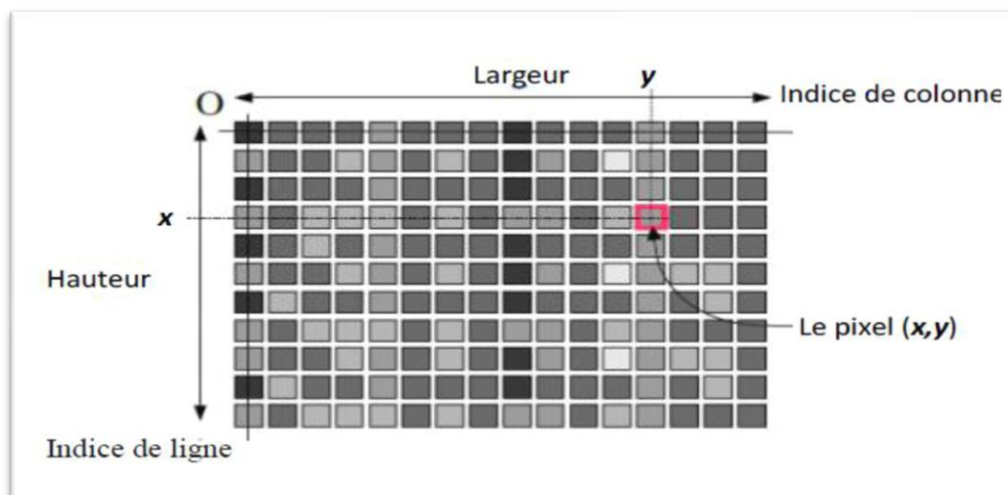


Figure 1.1: Représentation d'une image numérique.

1.5 Les différents types d'image numérique

Il existe deux catégories d'image numérique :

1.5.1 Image vectorielle

Une image vectorielle, également appelée image en mode ligne, est une image numérique composée d'expressions géométriques individuelles et de primitives géométriques (cercles, courbes de Bézier, lignes) définies pour diverses caractéristiques telles que la forme, la position et la couleur.

Ces modifications vous permettent de réaliser diverses transformations telles que l'écrasement, l'inclinaison et l'agrandissement des images sans perdre la qualité originale de l'image.



Figure 1.2 : Exemple d'une image vectorielle.

1.5.2 Image matricielle

Une image matricielle, également appelée bitmap, est une image constituée d'une matrice ou d'un tableau sous forme de grille, où chaque case a sa propre couleur (point de couleur).

Une série de points colorés qui forment ensemble d'une image ; ces points colorés sont appelés « pixels » dans le domaine de l'image numérique.



Figure 1.3 : Exemple d'une image matricielle.

1.6 Caractéristiques d'une image numérique

Comme nous l'avons vu, l'image est un ensemble structure d'informations parmi ses caractéristiques nous pouvons citer les paramètres suivants :

1.6.1 Pixel

Contraction de l'expression anglaise "Picture éléments" : éléments d'image, le pixel est le plus petit point de l'image, c'est une valeur numérique représentative des intensités lumineuses.

Si le bit est la plus petite unité d'information que peut traiter un ordinateur, le pixel est le plus petit élément que peuvent manipuler les matériels et logiciels sur l'image. La lettre A, par exemple, peut être affichée comme un groupe de pixels dans la figure ci-dessous. [2]

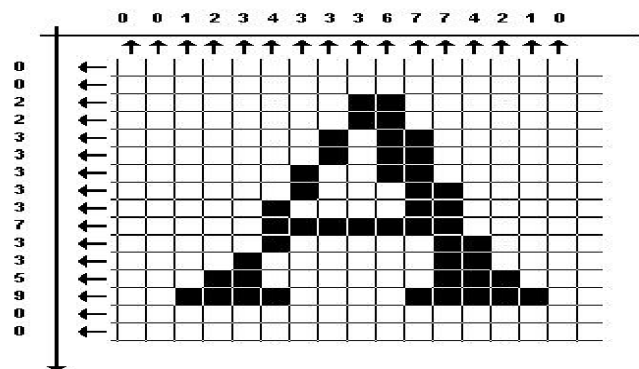


Figure 1.4 : Groupe de pixel formant la lettre A.

1.6.2 La résolution

La résolution d'une image correspond au niveau de détail qui va être représenté sur cette image. C'est le nombre de pixels par unité de longueur dans l'image à numériser. Elle est en dpi (dots per inch) ou en ppp (point par pouce).

Plus le nombre de pixels est élevé par unité de longueur de l'image à numériser, plus la quantité d'information qui décrit l'image est importante et plus la résolution est élevée. [3]



Figure 1.5 : Une image sous 3 résolutions différentes.

1.6.3 Dimension

C'est la taille de l'image cette dernière se présente sous forme de matrice dont les éléments sont des valeurs numériques représentatives des intensités lumineuses (pixels). [1]

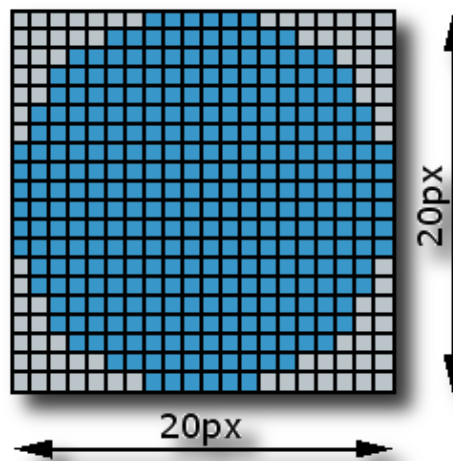


Figure 1.6: Les dimensions d'une image.

1.6.4 La texture

Une texture est une région dans une image numérique qui a des caractéristiques homogènes ces caractéristiques sont par exemple un motif basique qui se répète la texture est composée de Texel, l'équivalent des pixels. [4]

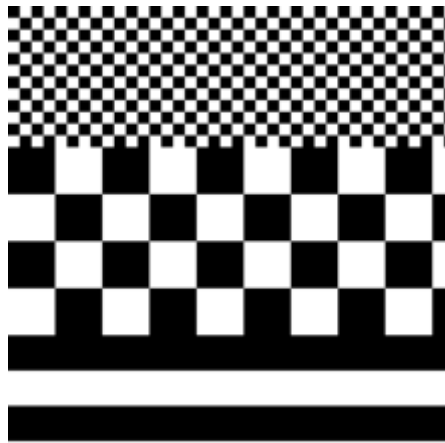


Figure 1.7 : Représenter la texture.

1.6.5 Le bruit

Un bruit (parasite) dans une image est considéré comme un phénomène de brusque variation de l'intensité d'un pixel par rapport à ses voisins, il provient de l'éclairage des dispositifs optiques et électroniques du capteur.[5]



Figure 1.8 : Image sans bruit. Figure 1.9 : Image avec bruit.

1.6.6 La luminance

C'est le degré de luminosité des points de l'image. Elle est définie aussi comme étant le quotient de l'intensité lumineuse d'une surface par l'aire apparente de cette surface, le mot luminances est substitué au mot brillance, qui correspond à l'éclat d'un objet.

Une bonne luminance se caractérise par :

- Des images lumineuses (brillantes) ;

- Un bon contraste : il faut éviter les images où la gamme de contraste tend vers le blanc ou le noir, ces images entraînent des pertes de détails dans les zones sombres ou lumineuses.
- L'absence de parasites.

1.6.7 Contours

Les contours représentent la frontière entre les objets de l'image, ou la limite entre deux pixels dont les niveaux de gris représentent une différence significative. Les textures décrivent la structure de ceux-ci. L'extraction de contour consiste à identifier dans l'image les points qui séparent deux textures. [4]



Figure 1.10 : Contours d'une image.

1.6.8 Histogramme

L'histogramme des niveaux de gris ou des couleurs d'une image est une fonction qui donne la fréquence d'apparition de chaque niveau de gris (couleurs) dans l'image. Pour diminuer l'erreur de quantification, pour comparer deux images obtenues sous des éclairages différents, ou encore pour mesurer certaines propriétés sur une image.

Il permet de donner un grand nombre d'informations sur la distribution des niveaux de gris (couleur) et de voir entre quelles bornes est répartie la majorité des niveaux de gris (couleur) dans le cas d'une image trop claire ou trop foncée. [6]

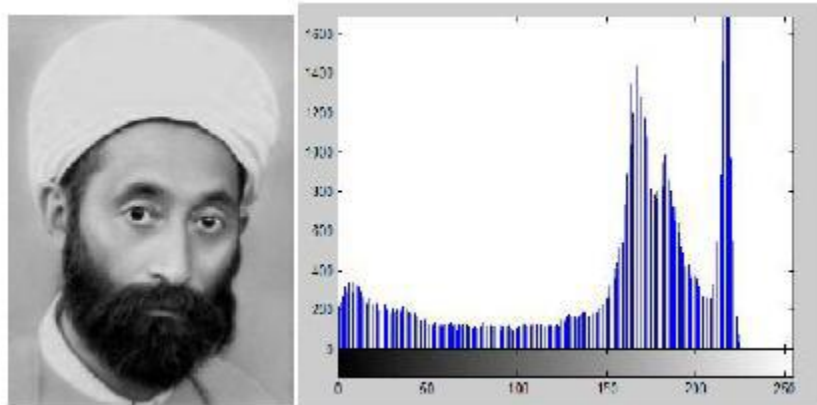


Figure 1.11 : Image avec son histogramme.

1.6.9 Egalisation d'histogramme

L'égalisation d'histogramme dans le traitement d'image est une méthode d'ajustement du contraste des images numériques à l'aide d'histogrammes. Elle consiste à appliquer une transformation à chaque pixel de l'image, obtenant une nouvelle image grâce à des opérations indépendantes sur chaque pixel.

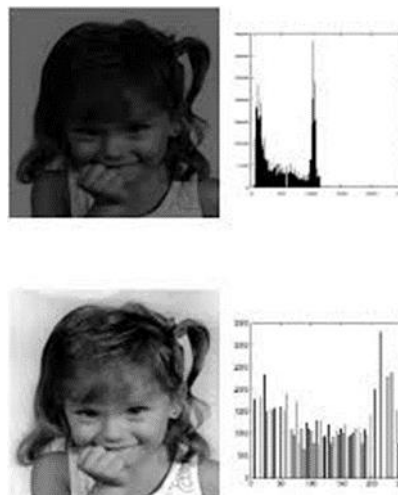


Figure 1.12 : Illustration d'un histogramme égalisé.

1.6.10 Espace de couleur

Chaque image numérique contient un ou plusieurs canaux de couleur, qui définit une intensité ou une certaine couleur dans une position (x, y) de l'image. Et la conversion d'une image de la représentation abstraite à la représentation réelle requiert la présence d'une palette de couleurs qui associe chaque valeur de la représentation abstraite à une couleur de cette palette.

Les palettes de couleurs les plus connues sont les niveaux de gris (grayscale), RGB et HSV :

- **Images à niveaux de gris (grayscale) :** C'est un système de codage informatique des couleurs avec un seul Canal où chaque valeur de ce canal varie de 0 (noir) à 255 (blanc). [1]

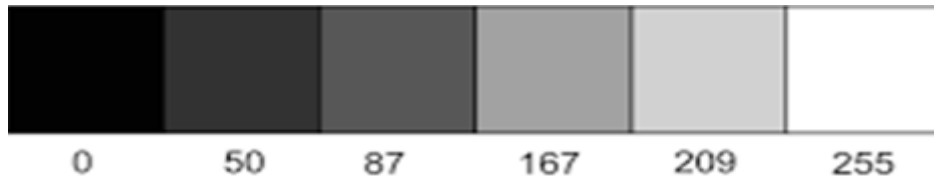


Figure 1.13 : Palette de niveaux de gris (grayscale).

- **RGB (rouge, vert, bleu) :** C'est un système de codage informatique des couleurs avec 3 canaux où le premier représente le rouge, le deuxième représente le vert et le dernier représente le bleu. Chaque valeur de ces canaux est variée de 0 à 255. [3]

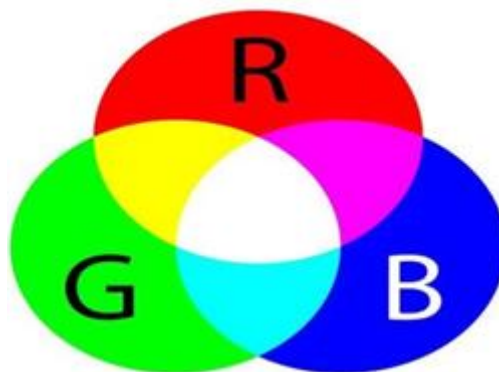


Figure 1.14 : Palette de RGB.

- **HSV (Teinte, saturation, luminosité) :** C'est un système de codage informatique des couleurs avec 3 canaux où le premier représente la teinte, le deuxième représente la saturation et le dernier représente la luminosité.

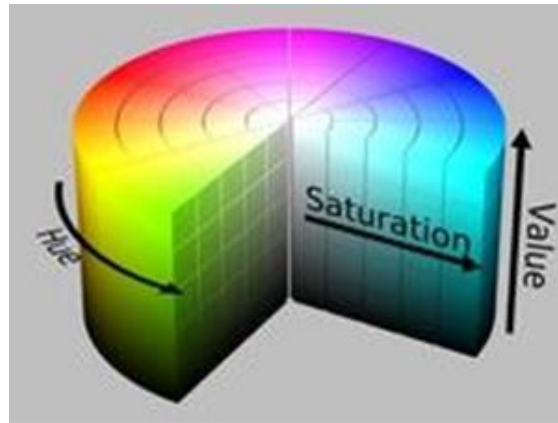


Figure 1.15 : Palette de HSV.

1.7 Segmentation

La segmentation d'image est une opération qui vise à regrouper les pixels selon des critères prédéfinis, il ne s'agit généralement que d'une première étape importante dans le processus d'interprétation d'une scène. De cette manière, les pixels sont regroupés en régions qui font partie d'une image c'est ce qu'on appelle également la binarisation.

La segmentation est une étape essentielle du traitement d'image. Il existe à ce jour de nombreuses méthodes de segmentation, qui peuvent être principalement classées en quatre classes :

- La segmentation fondée sur les régions (en anglais : region-based segmentation). On y trouve par exemple : la croissance de région (en anglais : region-growing), décomposition/fusion (en anglais : split and merge).
- La segmentation fondée sur les contours (en anglais : edge-based segmentation).
- La segmentation fondée sur la classification ou le seuillage des pixels en fonction de leur intensité (en anglais : classification ou thresholding).
- La segmentation fondée sur la coopération entre les trois premières segmentations.

Segmenter une image en objets revient à regrouper des régions qui ont une signification sémantique commune, même si leurs propriétés sont différentes. Par exemple, cela inclut de distinguer une personne du contexte dans lequel elle agit, même si cette personne porte des vêtements de couleurs différentes ou ne bouge que certaines parties du corps.

De nombreuses études ont été menées sur ce sujet dans divers domaines tels que le domaine médical et militaire. Cela reste un sujet d'actualité et une question ouverte. L'objectif est de rechercher des caractéristiques spéciales dans l'ensemble ou dans des parties de

l'image. Il s'agit du « ROI » « Région of Interest ». Le but est d'extraire des informations pertinentes sur une application particulière.

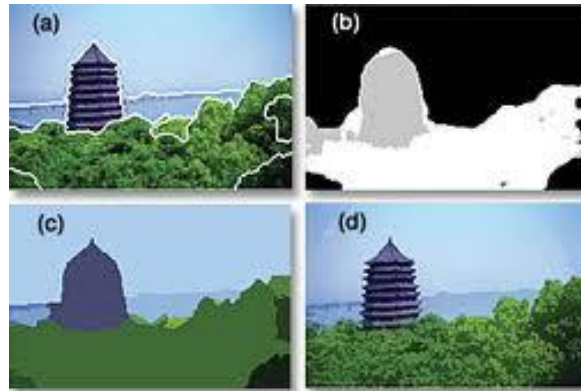


Figure 1.16 : Image segmentée.

1.7.1 Binarisation

La binarisation est de convertir les images originales en couleur ou en niveaux de gris en images binaires composées de petits pixels noirs et blancs elle peut être utilisée pour améliorer les caractéristiques d'une image en choisissant le seuil approprié. Si la valeur du pixel est en dessous du seuil nous lui associons la valeur (0). Si la valeur du pixel est égale ou supérieure au seuil nous lui donnons la valeur (1). La binarisation d'images est la base de l'analyse d'images et constitue l'étape clé du traitement d'images utilisé pour l'identification des particules.

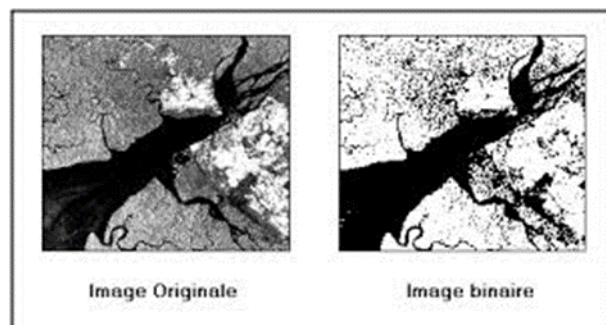


Figure 1.17 : Image binarisée.

1.7.1.1 Binarisation globale

La méthode de binarisation globale consiste à calculer un seuil unique en utilisant une mesure globale sur l'ensemble de l'image. Il nous donne la possibilité de déterminer si un pixel est noir ou blanc.

1.7.1.2 Binarisation locale

Dans cette forme de binarisation, le seuil n'est pas une valeur particulière pour l'ensemble de l'image, mais il est principalement déterminé par la région de taille définie par l'utilisateur. Ainsi, il convient de comparer chaque pixel avec le seuil calculé par sa région.

1.7.2 La détection des contours

Dans le domaine du traitement d'image et de la vision par ordinateur, la détection de contours désigne les méthodes qui permettent de détecter les points d'une image matricielle qui sont associés à un changement brusque de l'intensité lumineuse.

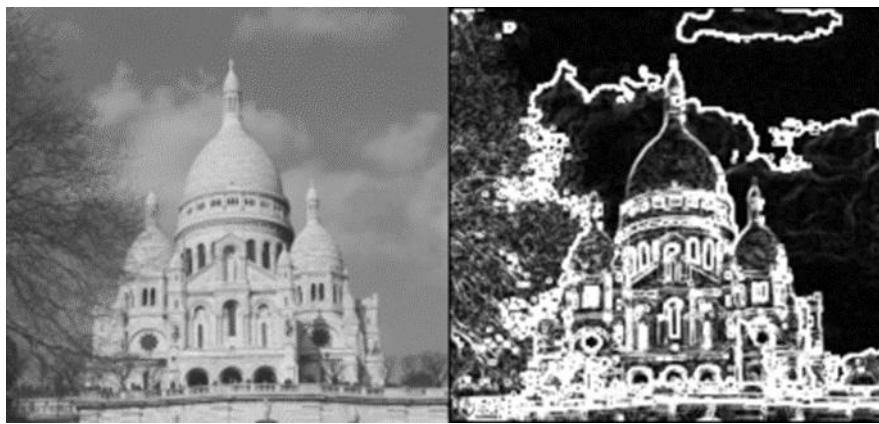


Figure 1.18 : Exemple de détection des contours.

1.8 Prétraitement

1.8.1 Filtrage

Le filtrage est l'application d'une transformation (appelée filtre) à n'importe quelle partie d'une image numérique. Le principe du filtrage est de modifier les valeurs des pixels d'une image pour améliorer son apparence. Cela crée une nouvelle image en utilisant les valeurs de pixels de l'image d'origine. Le filtrage fonctionne uniquement sur les données d'image numérisées. Les résultats obtenus dépendent fortement de la qualité du signal de l'image originale. Si le signal est fortement dégradé, le filtrage seul n'aura que peu d'effet. Dans ces situations, le filtrage n'est que la première étape d'un processus plus complexe.

Il existe plusieurs filtres utilisés en traitement d'image :

- Filtrage linaires
- Filtrage non linaires

1.8.1.1 Filtres linaires

Un filtre linéaire transforme un ensemble de données d'entrée en un ensemble de données de sortie selon une opération mathématique appelée convolution. Lorsqu'il s'agit des données numérisées, comme dans le traitement d'images, la relation entre les valeurs des pixels de sortie et d'entrée est décrite par un tableau de nombres (souvent carrés) appelés matrice de convolution. En général, le temps de calcul est réduit en divisant le filtre en deux filtres pouvant être récupérés par convolution l'un avec l'autre. En particulier, cette technique est utilisée pour créer un filtre bidimensionnel à partir de deux filtres unidimensionnels (vecteurs) dans les directions horizontale et verticale.

- **Filtre passe-bas :**

Un filtre passe-bas est un filtre qui laisse passer les basses fréquences et qui atténue les hautes fréquences, c'est-à-dire les fréquences supérieures à la fréquence de coupure. Il pourrait également être appelé filtre coupe-haut. Ce filtrage est souvent utilisé pour réduire le bruit de l'image, on parle dans ce cas de lissage.

Lissage : est une opération importante dans le traitement d'image, généralement utilisée pour réduire le bruit qui déforme les informations avant un traitement ultérieur. Cette opération consiste généralement à appliquer un filtre linéaire passe-bas numérique à l'image.

Filtre moyen : est un filtre passe-bas dont le principe de calculer la moyenne, des pixels situés dans le voisinage de chaque pixel. Cette famille de filtres permet de réduire le bruit dans l'image, ce qui rend les zones homogènes plus lisses.

Filtre gaussien : C'est un filtre utilisé pour lisser une image, mais en prenant en considération le fait que plus un pixel du voisinage est proche du pixel actuellement traité, plus il a de chances de faire partie du même objet et donc d'avoir la même couleur.

- **Filtre passe-haut**

Un filtre passe-haut est un filtre qui laisse passer les hautes fréquences et qui atténue les basses fréquences, c'est-à-dire les fréquences inférieures à la fréquence de coupure. Il pourrait également être appelé filtre coupe-bas.

1.8.1.2 Filtres non linéaires

Le filtrage non linéaire considère les pixels voisins selon des lois non linéaires, le filtre le plus classique et le plus utilisé est le filtre médian. Semblables aux filtres convolutifs, les filtres non linéaires fonctionnent sur des quartiers spécifiques.

Filtre médian : est un filtre numérique non linéaire couramment utilisé pour la réduction du bruit. La réduction du bruit est une étape de prétraitement classique visant à améliorer les résultats des traitements futurs (comme la détection des contours). Les techniques de filtre médian sont souvent utilisées dans le traitement d'images numériques car elles peuvent réduire le bruit tout en préservant les contours de l'image dans certaines conditions. L'idée principale d'un filtre médian est de remplacer chaque entrée par la valeur médiane de son voisinage.

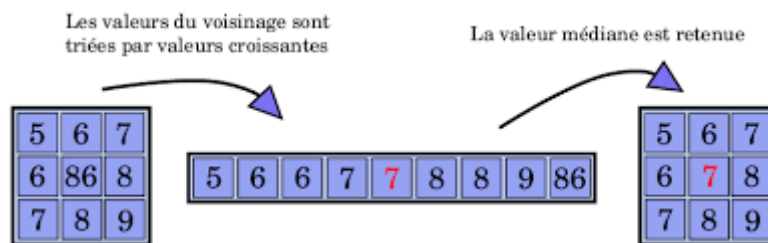


Figure 1.19 : Exemple d'un filtre médian.

1.9 Vidéo

1.9.1 Définitions de la vidéo

Le mot vidéo vient du mot latin « vidéo » qui signifie « je vois ». Il s'agit d'un terme qui inclut toutes les technologies permettant de visualiser ou d'enregistrer des images animées avec du son sur des supports électroniques plutôt que sous forme de film. Une vidéo est une séquence d'images à une fréquence d'images spécifique, l'œil humain peut distinguer environ 20 images par seconde.

La fluidité (vitesse) d'une vidéo est caractérisée par le nombre d'images par seconde (frame rate) exprimé en FPS (images par seconde).

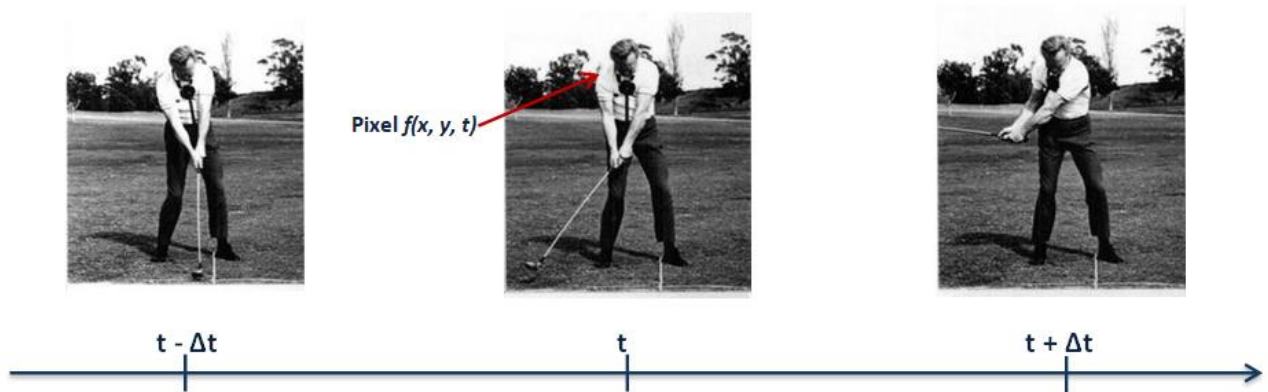


Figure 1.20 : Ensemble d'images formant une vidéo.

1.9.2 Vidéo numérique et analogique

On distingue généralement plusieurs grandes familles

La vidéo analogique

Représentant l'information comme un flux continu de données analogiques, destiné à être affichées sur un écran de télévision (basé sur le principe du balayage. Il existe plusieurs normes pour la vidéo analogique. Les trois principales sont : (PAL, NTSC, SECAM).

La vidéo numérique

La vidéo numérique consiste à afficher une succession d'images numériques. Puisqu'il s'agit d'images numériques affichées à une certaine cadence, il est possible de connaître le débit nécessaire pour l'affichage d'une vidéo, c'est-à-dire le nombre d'octets affichés (ou transférés) par unité de temps. Ainsi le débit nécessaire pour afficher une vidéo (en octets par seconde) est égal à la taille d'une image que multiplie le nombre d'images par seconde.

1.9.3 Les vidéos entrelacées

Entrelacement vidéo ou Le balayage entrelacé, en anglais « Interlaced video » est une technique d'affichage d'image de télévision, d'informatique ou d'appareils vidéo. Elle consiste à transmettre alternativement deux demi-images ou « trames », l'une comprenant les lignes numérotées impaires, l'autre celles numérotées paires.

L'entrelacement vidéo est conçu pour réduire l'effet de papillotement sur les écrans à tube cathodique, sans augmenter le nombre d'images à afficher par seconde et par conséquent devoir doubler la bande passante nécessaire lors de la transmission.

1.10 Acquisition des données

L'acquisition d'images joue un rôle crucial dans toute phase de conception et de production d'images. Afin de pouvoir manipuler une image sur un système informatique, il est essentiel de la transformer pour qu'elle soit lisible et manipulable par ce système. La numérisation permet de transformer l'image réelle en une image numérique capable d'être interprétée par le système. On peut distinguer deux catégories principales de ces systèmes de saisie : les scanners et les caméras numériques. [7]

Scanner : est un dispositif informatique qui permet de numériser des documents ou d'autres éléments tels que les empreintes digitales, par exemple.

Caméra numérique : Une caméra numérique est un appareil numérique qui utilise un capteur électronique (CCD) pour convertir les informations lumineuses en signaux électrique. Un CCD (Charge Coupled Device) transforme les photons lumineux qu'il reçoit sur l'électrode Semi-transparente en électron par effet photoélectrique. Ces électrons sont collectés dans la zone désertée. Le nombre d'électrons collectes est proportionnel à la quantité de lumière reçue.

1.11 Discussion

A travers de ce chapitre, nous avons exploré des concepts fondamentaux et établi des définitions élémentaires liées au traitement des images et des vidéos post-acquisition, destinées à être intégrées dans notre système. Nous avons procédé à la définition de l'image numérique et de la vidéo, en mettant en lumière leurs caractéristiques essentielles telles que le pixel, la résolution, la taille et l'espace de couleurs. De plus, nous avons examiné les instruments d'acquisition qui jouent un rôle crucial dans la capture de ces éléments visuels pour les inclure ultérieurement dans nos processus systémiques.

CHAPITRE 2 :

Les méthodes de détection de la sommolence

2.1 Préambule

La somnolence au volant représente un facteur de risque majeur dans les accidents de la route, et sa détection précoce est essentielle pour garantir la sécurité des usagers.

Dans ce chapitre nous explorerons les différentes approches technologiques et algorithmiques utilisées pour détecter la somnolence chez les conducteurs : les méthodes utilisant les signaux physiologiques, les méthodes qui basé sur le véhicule et les méthodes qui se concentre sur le traitement d'image.

2.2 La somnolence

La plupart des accidents de la route sont causés par la somnolence au volant, qui est un état qui est caractérisé par un fort désir de dormir, des difficultés à rester éveillé et une diminution de la vigilance et de l'attention. La somnolence peut présenter des risques lorsque l'on doit se concentrer à un certain niveau, comme lors de la conduite de véhicules motorisés. La somnolence ne se manifeste pas uniquement par la fatigue, mais peut également survenir lorsque la personne est sous l'influence de substances douteuses de médicaments ou d'alcool. Les collisions routières qui en découlent sont graves et fréquemment fatales. D'après les récents résultats de l'étude de référence menée par le laboratoire de transport et logistique de l'université de Batna, l'Algérie continue de se classer parmi les pays les plus touchés par les accidents de la route, avec une moyenne de 4 000 décès et 60 000 blessés constatés chaque année sur les routes, ce qui entraîne des pertes financières équivalentes à 100 milliards de dinars par an. En effet, ces dernières années, les accidents de la route ont pris la forme d'un véritable fléau. En moyenne, il y a 12 morts par jour.

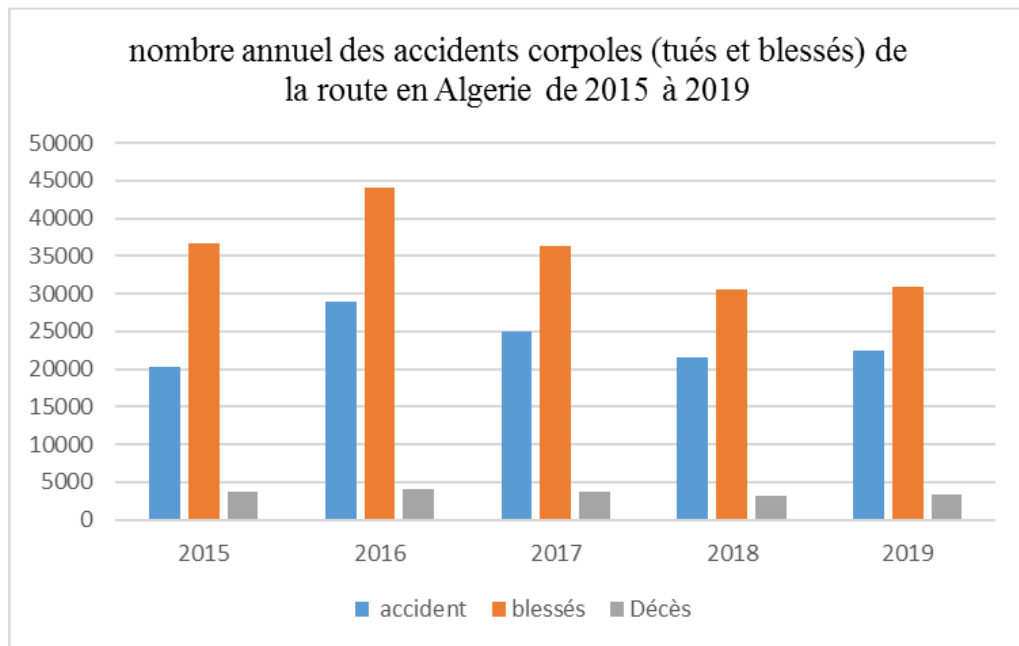


Figure 2.1 : Les statistiques des accidents de la route de 2015 à 2019.

2.3 Les facteurs influençant la somnolence

Des nombreux facteurs peuvent influencer l'état somnolence des conducteurs, ce qui peut augmenter le risque d'accident tel que l'insuffisance et la qualité du sommeil, l'heure et la durée de conduite, l'environnement de conduite, ainsi que des facteurs internes tels que la prise de médicaments ou d'autres substances...etc.

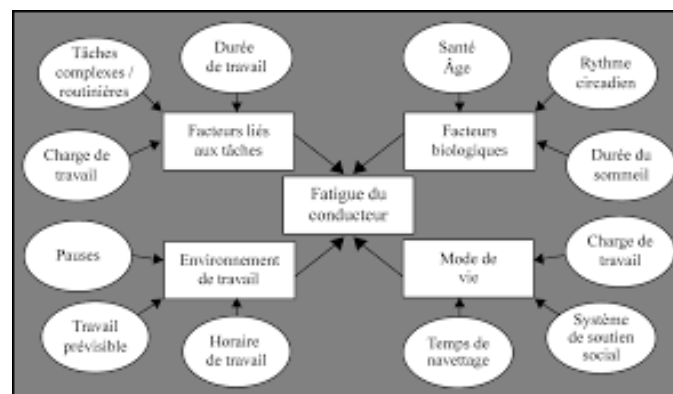


Figure 2.2 : Les différents facteurs pouvant être liés à la somnolence.

2.3.1 L'insuffisance et la qualité du sommeil

Il est évident que l'absence de sommeil adéquat et de mauvaise qualité pour les conducteurs peut provoquer une extrême somnolence. En outre, une faible quantité de sommeil ou une mauvaise qualité réduit la vigilance, ralentit les réactions et augmente ainsi le

risque d'accident de la route. Il est également possible que les conducteurs épuisés soient victimes de micro-sommeils involontaires, ce qui constitue un risque extrêmement dangereux. De plus, des éléments tels que des conditions de travail insatisfaisantes, comme des horaires de travail excessifs, des emplois du temps inadéquats et un climat organisationnel stressant peuvent être responsables de la fatigue chronique.

2.3.2 Horaire de conduite

Il est également souligné que l'horaire de conduite est un élément essentiel pouvant entraîner une somnolence. Effectivement, l'étude des accidents associés à une diminution de la vigilance révèle deux pics d'accidents. Le premier est ouvert entre 13h et 15h, tandis que le second, beaucoup plus important, est situé vers 2h du matin à 5h.



Figure 2.3 : Les horaires à risque pour conduire.

2.3.3 Durée de conduite

La durée de conduite un autre facteur joue également un rôle essentiel dans la fatigue au volant. En effet, la conduite prolongée sans pauses entraîne une augmentation de la fatigue et de la somnolence, qui sont évaluées de manière subjective et une diminution des performances de conduite. Il est crucial de faire régulièrement des pauses afin de rester vigilant.

2.3.4 Environnement de conduite

En plus de ces facteurs, différents éléments liés à l'environnement de conduite peuvent également avoir un impact sur l'état du conducteur, qui doit s'adapter à un environnement en évolution. Le risque de somnolence peut être amplifié par trois éléments contextuels : la monotonie, l'état de la route et le trafic. C'est essentiel pour assurer la sécurité et la performance du véhicule.

2.3.5 Consommation de médicaments et autres substances

Il est bien connu que l'alcool est une substance dangereuse pour la conduite, même en petites quantités. En plus des effets directs sur les réactions du conducteur, la consommation d'alcool a également une forte influence sur le niveau de somnolence.

Aussi certaines substances chimiques et certains médicaments peuvent également altérer l'état du conducteur, ce qui peut provoquer une somnolence diurne excessive causé par ces effets secondaires tels que les sédatifs hypnotiques, les anxiolytiques, les antidépresseurs, les neuroleptiques, les antihistaminiques, les amphétamines, les opiacés et les produits dérivés de la cocaïne. Il faut donc éviter toute substance susceptible de perturber les capacités lors de la conduite, pour éviter les accidents et respecter la loi.



Figure 2.4 : Symboles pouvant apparaître sur la boîte d'un médicament.

2.4 Les signes précurseurs de somnolence

Le conducteur peut détecter rapidement la fatigue et la somnolence en observant les signes précurseurs suivants :

- Difficulté à maintenir la tête droite.
- Bâillements à répétition.
- Lourdeur des paupières.
- Picotement des yeux.
- Difficultés à maintenir une vitesse constante.
- Troubles de la concentration.
- Engourdissement des jambes.

Au bout de 30 minutes après l'apparition des premiers signes de somnolence, le risque d'accident est quatre fois plus élevé.



Figure 2.5 : Les signes de fatigue et de somnolence.

2.5 Les méthodes de détection de somnolence

Plusieurs études ont été menées pour limiter la mortalité liée à la fatigue au volant, et diverses méthodes existent pour détecter la somnolence. Différents types des nouvelles technologies ont été étudiés. La première étude a permis de définir le comportement des conducteurs et d'analyser leur niveau de fatigue. La seconde étude est l'étude physiologique, vient ensuite l'étude de véhicule sur la route, enfin l'étude basée sur le traitement d'image.

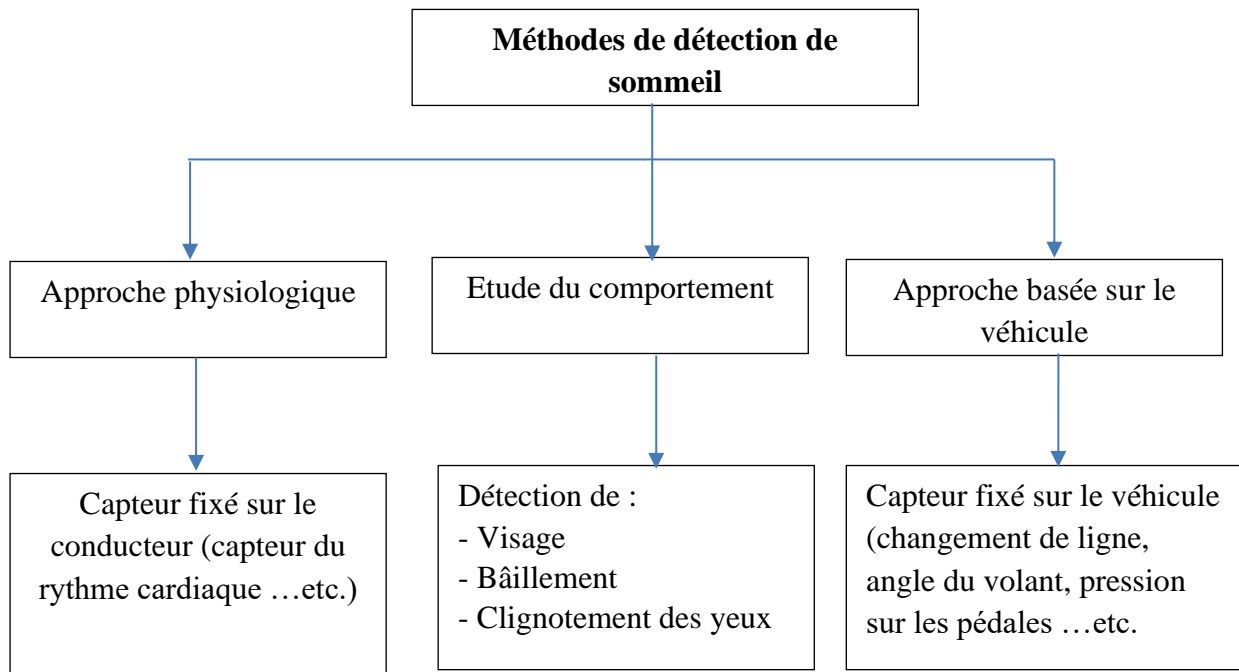


Figure 2.6 : Diverses méthodes pour détecter la somnolence.

2.5.1 Méthode basée sur la physiologie

Lorsque le conducteur se sent somnolent ou fatigué, son corps effectue de nombreux changements différents. Les changements se traduisent par une modification de la vitesse de rythme cardiaque, l'état de cerveau, l'état des muscles et l'activité oculaire. Sur la base de ces changements, plusieurs études et méthodes ont été proposées. Dans ces méthodes, les capteurs portés par le conducteur sont souvent utilisés. Ces capteurs permettent l'acquisition d'une variété de signaux EEG (état du cerveau), EOG (mouvements oculaires) et ECG (fréquence cardiaque). Dans certaines études, ils ont utilisé les signaux EOG pour surveiller l'état du conducteur pendant la conduite, En utilisant des algorithmes de l'apprentissage SVM [8], PNN (Prototype Nearest Neighbors) et MLP (Multi-Layer perceptron) [9]. Cela se fait en utilisant les signaux EOG verticaux et horizontaux obtenus du capteur pour classer les mouvements oculaires.

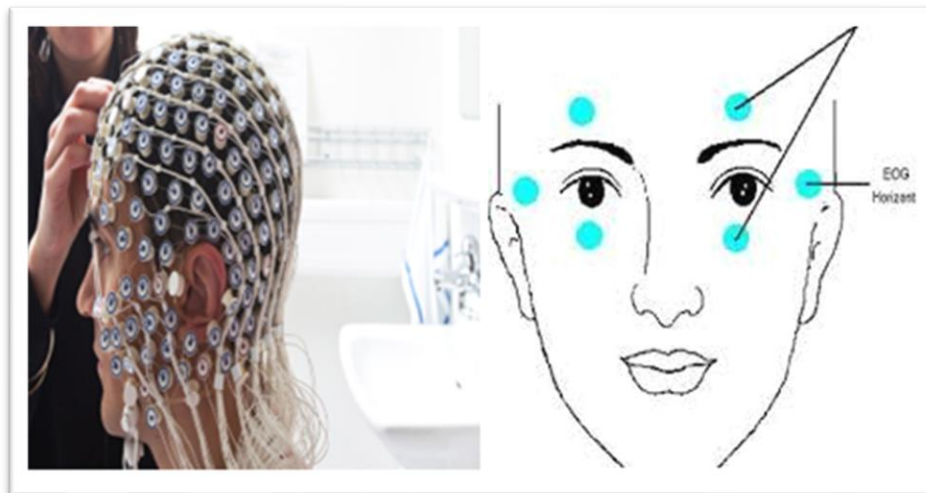


Figure 2.7 : Illustration des capteurs utilisés dans les méthodes axées sur la performance du conducteur [8].

On utilise également les signaux ECG pour identifier l'état du conducteur. On peut réaliser cela en utilisant la variabilité du rythme cardiaque, où une augmentation du rythme cardiaque indique une diminution du niveau de conscience, ce qui peut se produire pendant la somnolence du conducteur [10]. L'EEG est aussi l'un des signaux les plus efficaces pour étudier l'état du conducteur. Le spectre EEG comprend diverses bandes de fréquences, avec la bande delta de 0.5 à 4 Hz correspondant à l'activité du sommeil, la bande thêta de 4 à 8 Hz correspondant à la somnolence, la bande alpha de 8 à 13 Hz correspondant à la relaxation et la bande beta de 13 à 25 Hz correspondant à la vigilance [11].

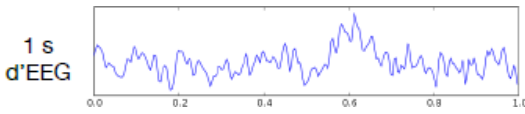
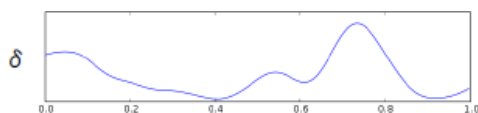
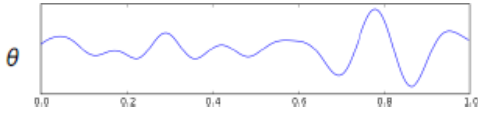
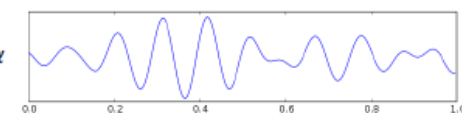
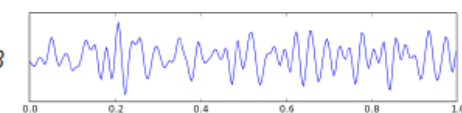
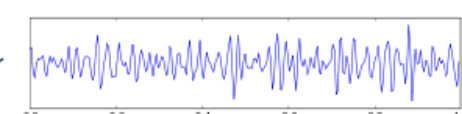
Bandes de fréquences	Noms Onde	Rôle	1 Second d'EEG
			
De0 à4 Hz	Onde delta	Sommeil long et profond	
De4 à8 Hz	Onde thêta	Diminution importante de la vigilance	
De8 à12 Hz	Onde alpha	Diminution modérée de la vigilance	
De12 à30 Hz	Onde beta	État éveillé actif	
Supérieur à 30 Hz	Onde gamma	Tache congestive ou fonction motrice	

Tableau 2.1 : Bande de Fréquence EEG, (extrait de La rue, 2010).

2.5.2 Méthode basée sur la performance du conducteur

Plusieurs facteurs influencent généralement le mode de conduite, tels que les tâches de conduite (changement de vitesses et accélération, courbure et largeur de la voie, etc.) et les caractéristiques du conducteur (expérience de conduite). En utilisant ces facteurs, il est possible de repérer l'état du conducteur. Dans la majorité des situations, des capteurs sont utilisées pour Prendre en compte les éléments précédents en positionnant ces capteurs dans les divers éléments de la voiture tels que le volant et la pédale d'accélération. Dans ces approches,

diverses études étaient suggérées, mais les mesures les plus fréquemment utilisées sont la détection de mouvement de volant et la détection de l'écart type de position.

Un système développé par Krajewski et ses collègues [12] repose sur l'extraction des propriétés du mouvement de volant, qui est lié au temps, à la fréquence et à l'environnement. Les algorithmes d'apprentissage SVM (Support Vector Machine) [8] et KNN (K-Nearest Neighbors) [9] ont ensuite été employés pour classer ces caractéristiques. La précision de ce système est de 86,1%. Une autre mesure a été suggérée, connue sous le nom de "Standard Deviation of Lane Position" (SDLP), qui permet d'évaluer les niveaux de somnolence du conducteur. En général, la voie est surveillée par une caméra. M.Inger et ses collègues [13] ont comparé le SDLP avec le KSS [14] afin d'évaluer le SDLP et ont constaté qu'il existe une corrélation directe entre le KSS et SDLP. Quand le SDLP augmente le KSS augmente aussi.

Plusieurs facteurs influencent ces approches, tels que le modèle de voiture, l'état de la route et l'expérience du conducteur. Afin de gérer ces mesures, il est nécessaire de consacrer beaucoup de temps et de données afin de créer un système avec une grande précision.



Figure 2.8 : Détecteur de ligne sur véhicule.

2.5.3 Méthode basée sur le traitement d'image

Les techniques mentionnées précédemment sont considérées comme invasives car elles ont un impact sur l'attention du conducteur. De nos jours, grâce au progrès technologique, la recherche s'est orientée vers des techniques de traitement d'images où une caméra est employée pour extraire les signes de somnolence du conducteur, comme les mouvements des yeux, les bâillements fréquents et les mouvements de la tête.

2.5.3.1 PERCLOSE

L'utilisation du traitement d'images a principalement été étudiée en se concentrant sur le mouvement oculaire, en particulier la mesure PERCLOSE [15], car certaines études ont démontré que cette mesure est une mesure fiable pour prédire la somnolence.

Selon Garcia et ses collègues [16], un système est proposé qui comprend trois étapes principales. Dans la première étape, ils effectuent le prétraitement en utilisant les algorithmes de Viola et Jones pour repérer les visages et les yeux. Ensuite, ils localisent les pupilles et évaluant l'état des yeux en utilisant la projection intégrale. Enfin, ils prédisent l'état du conducteur en utilisant la mesure PERCLOSE.

$$Perclose = \frac{\text{nombre des yeux fermés}}{\text{nombre des yeux fermés} + \text{nombre des yeux ouverts}} \dots\dots\dots(1)$$

W. Zhang et ses collègues [17] ont suggéré un autre système où ils ont également utilisé la mesure PERCLOSE ainsi que plus de cinq autres mesures telles que la durée maximale de la fermeture, la fréquence de clignement, le niveau moyen d'ouverture des yeux, la vitesse de fermeture des yeux et la vitesse d'ouverture des yeux. Ces mesures ont été classifiées pour différentes catégories. Ainsi, W. Zhang et ses collègues ont employé le Discriminant linéaire de Fisher.

2.5.3.2 Calcule l'EAR

Une autre méthode proposée par Adrian [18] consiste à calculer l'EAR (Eyes aspect ratio) après avoir détecté le visage et ses repères, afin de déterminer l'état des yeux (ouvert ou fermé). Enfin, l'alarme est déclenchée si les yeux sont fermés dans 48 trames consécutives.

L'équation ci-dessous illustre la relation appelée Eye aspect ratio (EAR).

$$EAR = \frac{\|p2-p6\| + \|p3-p5\|}{2\|p1-p4\|} \dots\dots\dots(2)$$

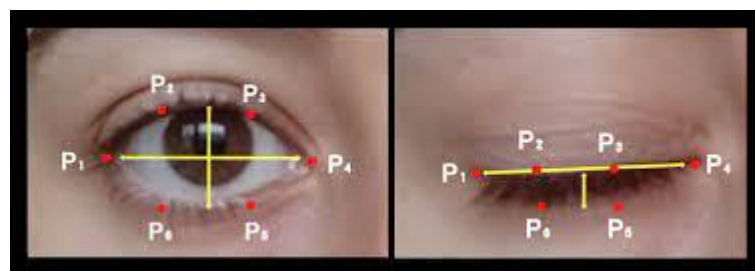


Figure 2.9 : Position des repères des yeux pour le calcul de l'EAR avec l'œil ouvert et l'œil fermé.

2.5.3.3 Des réseaux LSTM

Dans leur étude, Reza Ghoddoosian et ses collègues [19] ont suggéré un autre système qui ne tient compte que de trois caractéristiques : la fréquence de clignement, le niveau moyen d'ouverture des yeux et la vitesse d'ouverture des yeux. Pour classifier ces mesures, ils ont utilisé des réseaux de neurones récurrents (l'architecteur LSTM).

2.5.3.4 Analyse labiale

Même si des recherches ont démontré que les mouvements oculaires sont un indicateur fiable de la somnolence, d'autres symptômes ne sont pas plus significatifs que les mouvements oculaires. Dariusz Sawicki [20] a suggéré un mécanisme qui permet de différencier les fréquents bâillements de la somnolence. Quand vous parlez, chantez ou parlez, ouvrez la bouche et souriez. Les deux étapes principales du système sont la distinction des lèvres, qui est complétée par l'un des algorithmes de détection des marques faciales, et le calcul du rapport entre la hauteur et la largeur de la bouche.

2.5.3.5 L'algorithme de Viola et Jones

À la suite des symptômes similaires et d'autres méthodes, M. Omidyeganeh et ses collègues [21] ont suggéré un autre système pour anticiper l'état du conducteur. L'algorithme de Viola et Jones a été employé pour repérer le visage et la bouche, puis la théorie de back projection a été utilisée pour évaluer à la fois le taux et l'ampleur des modifications dans la bouche afin de prédire son état ainsi que celui du conducteur.

Les mouvements de tête sont également des signes qui peuvent être utilisés pour surveiller l'état du conducteur. Selon les chercheurs Lam Thanh Hien et Do Nang Toan [22], une méthode consiste à détecter le visage humain et à extraire certaines caractéristiques du visage, telles que les deux coins éloignés des yeux, les deux points de la bouche et le bout du nez. Ensuite, ces données sont utilisées pour calculer les paramètres prédéterminés tels que la distance point à point, la distance point à bord et la zone triangulaire contrainte par deux des coins de la bouche et des points du nez. En conclusion, on évalue la somnolence en se basant sur les seuils préétablis pour ces paramètres.

Tous les systèmes cités ci-dessus se basent sur un seul symptôme pour identifier la somnolence. Cependant, il existe d'autres méthodes suggérées qui combinent plusieurs caractéristiques. Selon M. Omidyeganeh et ses collègues [23], deux fonctionnalités ont été intégrées : le bâillement fréquent et le PERCLOSE, afin de diminuer les fausses alarmes et

améliorer la précision du système. Selon EsraVural et ses collègues [24], un autre système combine différentes caractéristiques telles que le bâillement fréquent, les mouvements de tête et les clignements des yeux, entre autres. La majorité des méthodes développées extraient les caractéristiques en utilisant les méthodes mentionnées dans les méthodes précédentes.

2.5.3.6 Réseaux de neurone convolutifs

Dernièrement, certaines méthodes qui font appel à l'apprentissage en profondeur, comme l'utilisation de CNN et RNN, ont fait leur apparition. L'utilisation des CNNs et des RNNs pour l'apprentissage en profondeur représente une avancée majeure, notamment dans les domaines de la vision par ordinateur tels que la classification d'images, la détection d'objets, la reconnaissance des émotions, la segmentation de scènes, etc. Selon ce concept, différentes méthodes sont suggérées pour repérer la somnolence.

Un système reposant sur les réseaux de neurones est proposé par Bhargava Reddy et ses collègues [25]. Ce système est constitué de deux étapes, où lors de la première étape, ils font appel à MTCNN (Multi-Task Cascaded Convolutional Neural Networks) afin de repérer le visage, les yeux, le nez, les lèvres gauche et droite, etc. Ce résultat correspond à l'entrée de la CNN de deuxième étape, qui va extraire les caractéristiques déjà extraites lors de l'étape précédente et prédire l'état du conducteur.

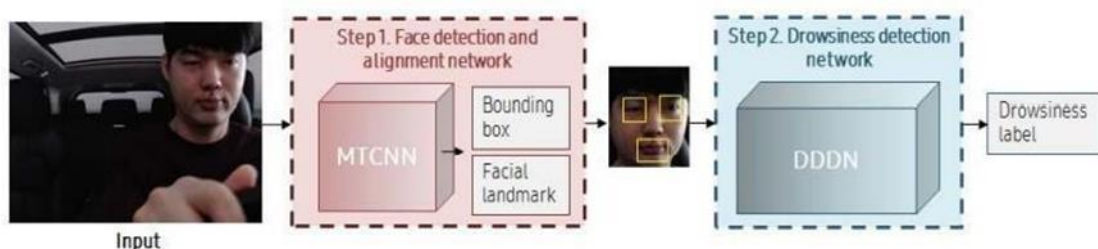


Figure 2.10 : Architecteur utilisé dans l'approche de Bhargava Reddy et al [25].

Un système similaire créé par Nisha Gandhi et ses collègues [26] utilise deux types d'apprentissage en profondeur : ils utilisent le CNN (réseau neuronal convolutif) pour identifier l'état des yeux (fermés ou ouverts) et le RNN (réseau de neurones récurrent) pour identifier l'état du conducteur.

2.5.3.7 Calcul des distances oculaires

Il est nécessaire de commencer par évaluer la boîte de tête afin de repérer les yeux. L'algorithme de Viola et Jones [27] est utilisé pour détecter la tête, tandis que l'approche de repère facial est mise en place par Kazemi et Sullivan [28]. Ce procédé d'apprentissage

automatique de l'arbre de régression a été développé en utilisant les données de repères de face iBUG 300-W [29]. Pour les autres, l'avantage de cette méthode réside dans la grande précision de la détection dans diverses positions de la tête.

Une fois que le point de repère de l'œil est atteint, le rendement sur investissement est réduit. Le visage présente une forme asymétrique, donc pour détecter la somnolence, il suffit d'avoir un seul œil pour être pris en compte. Cette méthode réduit considérablement le temps de calcul nécessaire pour la détection. Il est recommandé de choisir le front de l'œil de la caméra car il renferme davantage et réduit le taux d'erreur.

Les distances calculées à partir du point le plus à droite et du point le plus à gauche des yeux droit et gauche sont comparées par l'algorithme afin de choisir la distance la plus grande pour le recadrage. La figure 2.11 présente la distance prévue.

L'algorithme identifie les repères oculaires, évalue la distance absolue entre les points 37 et 40, 43 et 46, et choisit une distance plus importante comme front de l'œil de la caméra. Afin de résoudre le problème de l'éclairage, les écrivains font appel à un égaliseur d'histogramme afin de rendre l'œil contraste égal. Cette méthode améliore la précision de la méthodologie pour détecter la proximité des yeux. Une fois que les yeux ont été égalisés et détectés, l'œil égalisé est transmis au réseau.

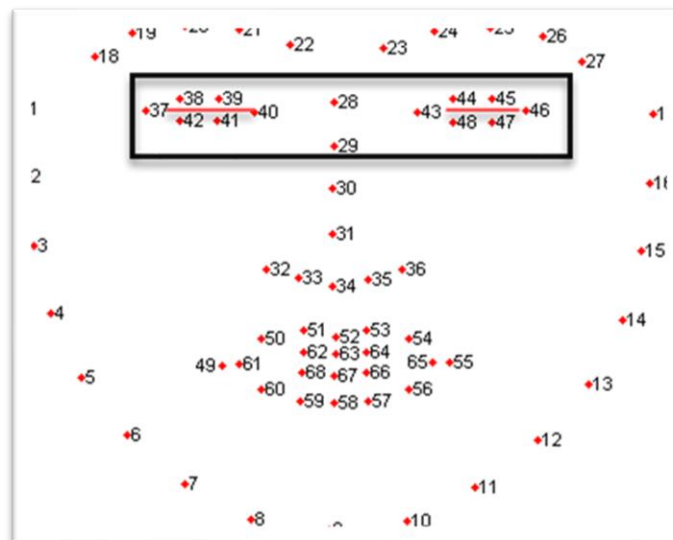


Figure 2.11 : Distance oculaire prévue (points 37 à 40 et 43 à 46) pour recadrer le ROI.

2.3.5.8 La transformée de Hough circulaire

On peut utiliser la transformée de Hough circulaire (CHT) pour détecter l'état des yeux chez les conducteurs. Cette méthode permet d'isoler des objets de formes géométriques simples dans l'image. En général, on se limite aux lignes, cercles ou ellipses présents dans l'image. Une des principales qualités de la transformation de Hough circulaire réside dans sa capacité à supporter les obstacles dans les objets recherchés et à rester relativement indépendante des bruits. Cette méthode nous offre la possibilité de distinguer les lignes (droites), les cercles ou toute autre forme présente dans une image. Les éléments à repérer dans l'image de l'œil (iris, pupille, paupières) présentent des formes circulaires ou ellipsoïdaux, ce qui les rend parfaitement accessibles à la recherche par la transformée de Hough circulaire [30].

Les différentes étapes de la transformation de Hough circulaire sont les suivantes :

- Une image de contours est produite à l'aide d'une méthode de génération de contours. La figure 2.12 présente plusieurs illustrations de contours obtenues à l'aide de la méthode Canny.
- Un vote est organisé sur l'image de contours obtenue. Chaque point de contour vote pour les cercles auxquels il appartient et le cercle qui obtient le plus de vote est le cercle recherché. Dans cette situation, nous effectuons une comptabilisation pour chaque cercle dans la zone recherchée ; le cercle qui présente le plus de points de contours est celui qui est recherché.

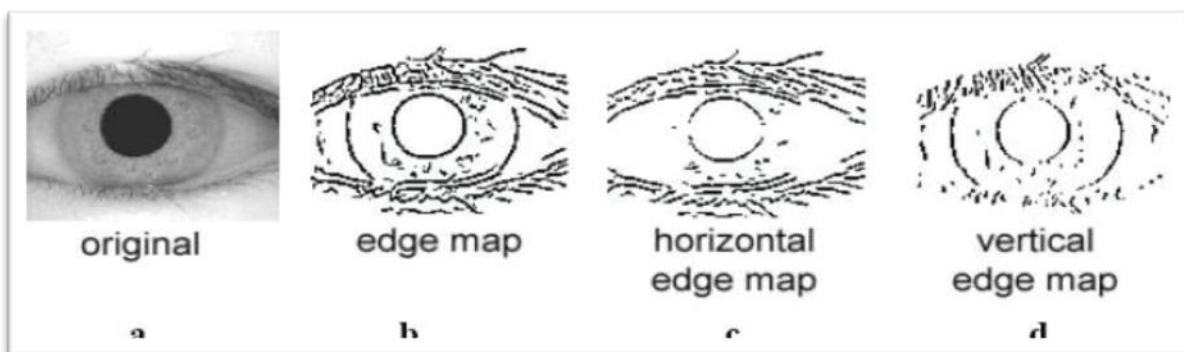


Figure 2.12 : (a) image de l'œil, (b) différentes images de contours par la méthode Canny selon que l'on considère des contours diagonaux, (c) horizontaux ou (d) verticaux.

2.6 Discussion

Dans ce chapitre, nous avons exposé les éléments qui influencent la somnolence, les symptômes préliminaires de somnolence et enfin les divers systèmes de détection automatique de la somnolence du conducteur, qui sont classés en trois catégories principales :

- Les méthodes qui se basent sur les performances du conducteur : elles sont non intrusives, mais elles sont influencées par divers facteurs tels que le type de voiture, l'état de route et l'expérience du conducteur.
- Les méthodes basées sur les signaux physiologiques : elles sont efficaces, mais elles sont intrusives et perturbent le conducteur tout en étant équipées de capteurs.
- Les systèmes qui reposent sur le traitement des images : ces techniques sont complexes, mais elles offrent des résultats satisfaisants et sont également non intrusives.

CHAPITRE 3 :

Détection de la somnolence par vision artificielle

3.1 Préambule

La détection de somnolence chez les conducteurs est importante, car la fatigue au volant est une cause fréquente d'accident grave. Le traitement d'image offre une solution efficace pour identifier en temps réel les signes de somnolence en analysant des caractéristiques visuelles telles que la détection de visage, les mouvements oculaires et clignements d'yeux.

Dans notre approche, nous avons utilisé l'algorithme de Viola et Jones pour la détection du visage, permettant de localiser rapidement le visage et les yeux du conducteur. Ensuite, pour évaluer l'état des yeux nous avons opté pour la transformée de Hough circulaire une méthode performante pour détecter les formes des yeux et les contours.

Le schéma synoptique ci-dessous illustre les différentes étapes suivies dans notre méthode pour la détection de la somnolence au volant, allant de la détection de visage à l'analyse de l'état des yeux.

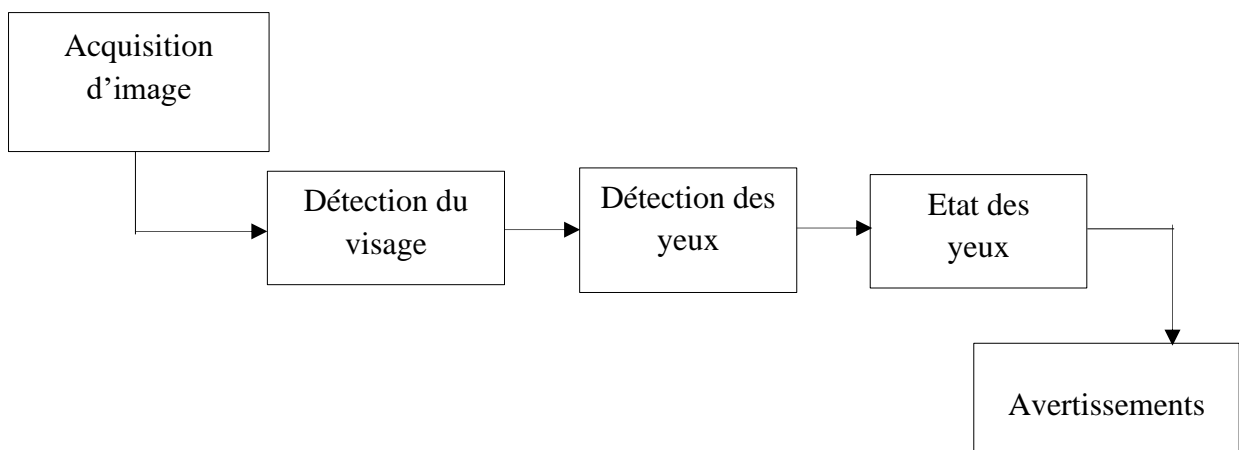


Figure 3.1 : Schéma synoptique de la méthode proposée.

Les algorithmes de vision artificielle, tels que la méthode Viola-Jones pour la détection de visage et des yeux, ou encore la transformation de Hough circulaire pour identifier l'iris, sont utilisés dans notre système pour traiter les images en temps réel.

3.2 Région d'intérêt (ROI)

Dans notre mémoire de fin d'études, nous nous focalisons sur le problème de détection de la somnolence du conducteur en se basant sur l'état des yeux (ouverts ou fermés). Pour identifier cet état, nous extrayons les régions du visage et des yeux afin de les analyser

séparément, comme illustré dans la figure 3.2 ces zones extraites sont désignées sous le nom de "régions d'intérêt" (ROI, pour Région of Interest).

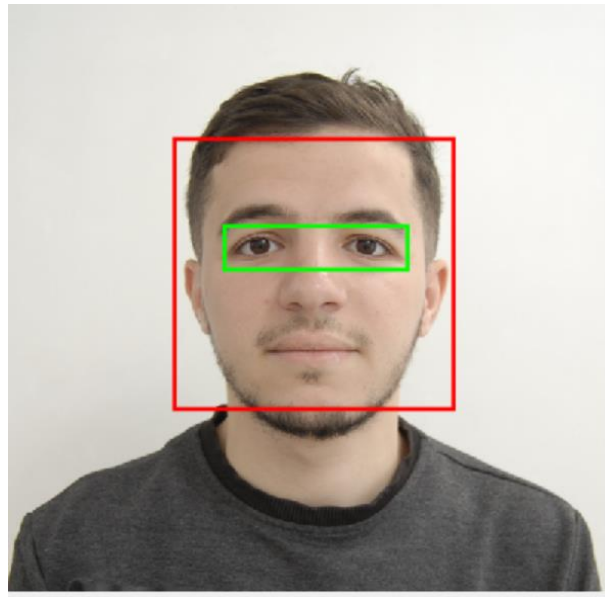


Figure 3.2 : Illustration qui montre les régions d'intérêt.

3.3 Reconnaissance d'objets

Reconnaître des objets en général, et des visages en particulier, est une tâche relativement facile pour les humains, mais extrêmement difficile pour les ordinateurs.

Cette tâche implique généralement deux concepts : Détection de position qui Utilise le point de départ de l'objet pour identifier la longueur et la largeur de l'objet et la classification identifie ensuite le type d'objet localisé à l'étape précédente.

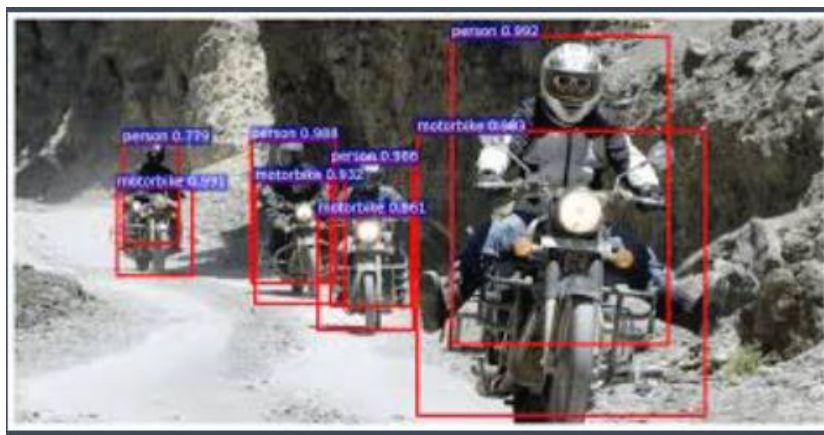


Figure 3.3 : Un exemple de détection de la position d'objets.

3.4 La détection du visage

De nos jours, la détection de visage est l'une des technologies les plus utilisées dans le traitement d'image pour identifier et localiser les visages dans des images ou des vidéos, avec de nombreuses méthodes développées au fil des ans telle que :

- **L'histogramme de gradient orienté (HOG)**

L'histogramme de gradient orienté est une méthode utilisée en vision par ordinateur pour décrire l'apparence locale d'un objet dans une image. Il se base sur la distribution de l'intensité des gradients ou de la direction des contours. Cette méthode est similaire à d'autres techniques comme les histogrammes d'orientation de bord, mais elle se distingue par le calcul sur une grille dense de cellules. Le HOG présente l'avantage d'être invariant à certaines transformations géométriques et photométriques, ce qui le rend robuste aux changements d'éclairage ou de pose.

- **Les réseaux de neurones convolutionnels (CNN)**

Les réseaux de neurones convolutionnels (sont des outils d'apprentissage profond particulièrement adaptés à la reconnaissance d'images, notamment pour la détection de visages. Ces modèles apprennent à identifier les caractéristiques distinctives d'un visage directement à partir des pixels d'une image, sans nécessiter de prétraitement complexe. Grâce à leurs multiples couches de convolution, les CNN peuvent extraire des caractéristiques de plus en plus abstraites, allant de simples contours à des représentations complexes du visage. Cette capacité d'apprentissage automatique les rend très robustes aux variations d'éclairage, de pose et d'expression, tout en étant relativement simples à mettre en œuvre.

- **Viola et Jones**

La méthode de Viola et Jones est une des techniques les plus importantes pour la détection de visage, notamment en raison de son efficacité en temps réel. Son principal avantage réside dans son utilisation d'un classificateur en cascade, qui permet de filtrer rapidement des régions non pertinentes de l'image et de se concentrer sur les zones susceptibles de contenir un visage. De plus, l'utilisation de caractéristiques Haar permet de simplifier les calculs tout en conservant une bonne précision. Cette méthode est particulièrement adaptée aux dispositifs à faible puissance de calcul, tels que les caméras embarquées ou les systèmes en temps réel, où les ressources sont limitées.

En raison de ces avantages offerts, nous avons choisi d'utiliser cette technique pour la détection du visage dans notre projet.

3.4.1 Difficultés de la détection de visages

La détection automatique du visage rencontre de nombreuses difficultés en raison de la grande variabilité des formes à identifier, ainsi que le changement dans les expressions faciales, les occlusions partielles et les variations d'éclairage, qui modifient l'aspect général des visages. En effet, de nombreux défis en détection faciale proviennent des conditions strictes liées à la classification des formes et aux techniques d'apprentissage. Les facteurs influençant les performances d'un système de détection faciale comprennent l'échelle, la pose et l'occlusion [32].

3.4.1.1 L'échelle

Un groupe de visages peut se présenter dans une image donnée à différentes échelles. On peut traiter l'échelle (ou la taille) d'un visage en utilisant un processus simple de mesure. [32]



Figure 3.4 : Visages à différentes échelles.

3.4.1.2 Pose

Lorsqu'il y a des variations de pose, les performances d'un système de détection de visages diminuent substantiellement. On observe des poses variables lorsque le point de vue change ou lorsque la tête est tournée dans les trois-dimensions. Ainsi, il est possible que certaines parties du visage comme l'œil ou le nez soient partiellement ou totalement occultées. [32]



Figure 3.5 : Visages à différentes poses.

3.4.1.3 Occlusion

Le visage peut être partiellement masqué par des objets dans la scène, ou par le port d'accessoire tels que lunettes, écharpe...etc.

Les systèmes de détection basés sur des modèles classiques peuvent échouer à reconnaître un visage incomplet ou déformer les résultats, car ils s'appuient sur des caractéristiques visibles. Même les algorithmes d'apprentissage profond, bien que plus performants, ont du mal à gérer les cas où une grande partie du visage est masquée. [32]



Figure 3.6 : Visage occlus.

3.5 La Méthode de Viola et Jones

L'algorithme de Viola et Jones, reconnu pour sa capacité à effectuer des détections rapides et précise de visage dans des images, se distingue pour son efficacité en temps réel, une caractéristique cruciale pour système de sécurité embarqué.

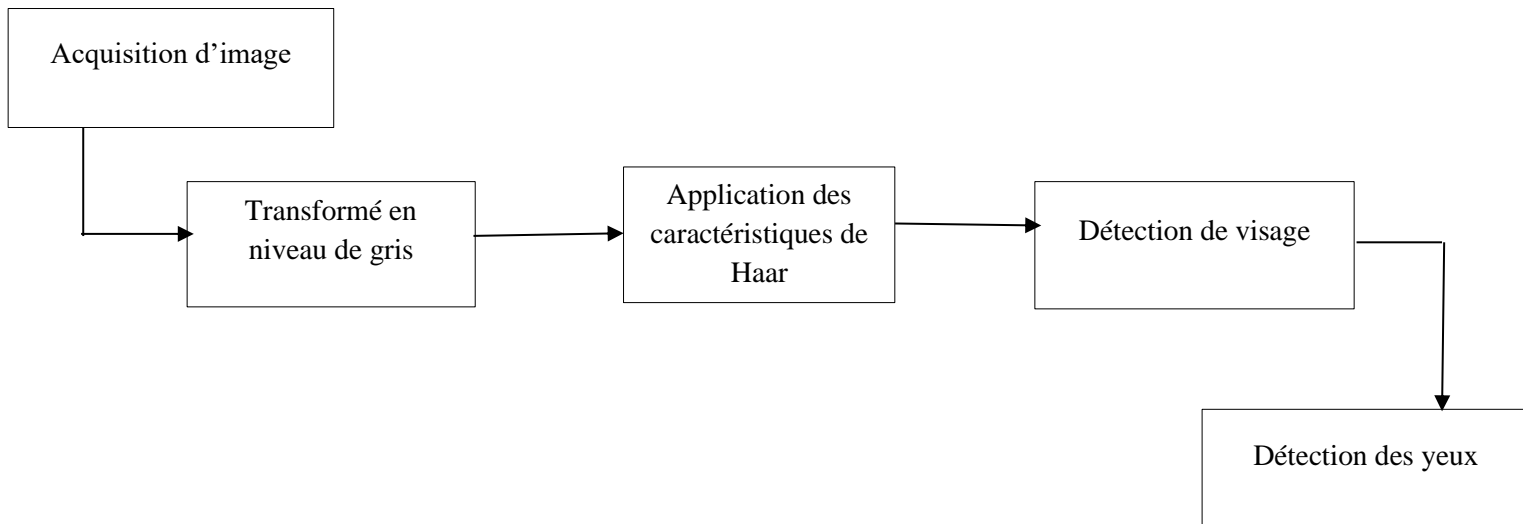


Figure 3.7 : Schème synoptique de l'algorithme Viola et Jones.

3.5.1 Définition

Les chercheurs Paul Viola et Michael Jones ont développé la méthode de Viola et Jones en 2001 pour détecter un objet dans une image numérique. C'est l'une des premières méthodes qui est considérée comme la méthode la plus recommandée et qui permet de repérer efficacement et en temps réel des objets dans l'image.

Originellement conçue pour repérer des visages, elle peut aussi servir à repérer d'autres objets tels que des voitures ou des avions. Viola et Jones ont été l'une des méthodes les plus célèbres et les plus couramment employées, notamment pour la détection de visages et la détection de personnes. [33]

La méthode de Viola et Jones, en tant que méthode d'apprentissage supervisé, exige de quelques centaines à plusieurs milliers d'exemples de l'objet que l'on veut détecter pour entraîner un classifieur. Après avoir été appris, ce classifieur sert à repérer la présence possible de l'objet dans une image en la parcourant entièrement, à toutes les positions et dans toutes les tailles envisageables. [34]

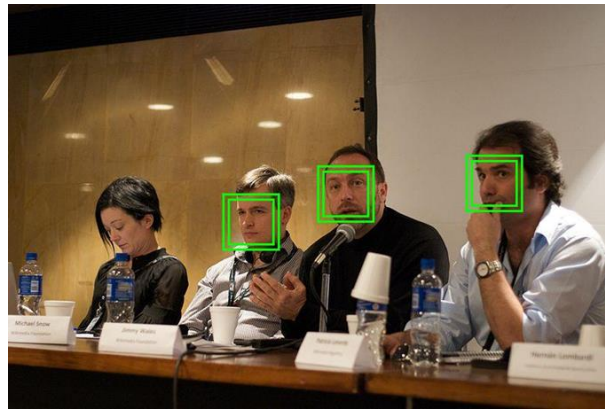


Figure 3.8 : Un exemple de détection de visage par la méthode de Viola et Jones.

3.5.2 Fonctionnements

La méthode de Viola et Jones n'a pas besoin d'étudier chaque pixel d'une image pour l'analyser et ce grâce à l'image intégral qui permet de définir plusieurs zones rectangulaires au sein d'une image. L'intérêt de cette notion est le fait d'offrir la possibilité d'accéder aux valeurs des autres zones à gauche et au-dessus de la zone à laquelle nous sommes. Ces zones permettent de créer des caractéristiques pseudo-Haar [35], qui sont en fait des masques permettant de déterminer plusieurs patterns.



Figure 3.9 : Exemple illustré de caractéristiques pseudo-Haar.

On arrive à reconnaître des visages en détectant des motifs est cela est rendu possible par la variation de l'intensité de la lumière entre les yeux et les pommettes (caractéristique n°3), et l'intensité de la lumière entre les yeux et le nez (caractéristique n°2).

La méthode de Viola et Jones repose sur l'utilisation des images intégrales et les caractéristiques pseudo-Haar, plus deux mécanismes : le classifieur et le boosting.

Le classifieur permet de déterminer l'ensemble des zones rentrant sous la coupe d'une caractéristique pseudo-Haar, en déterminant les seuils pouvant déterminer les exemples positifs des négatifs. Ceci demande donc une phase d'apprentissage, qui permet de définir les seuils

utilisés. Au final, un classifieur est une association entre une caractéristique pseudo-Haar et un seuil. C'est ce qu'on appelle un classifieur faible. Les caractéristiques sont calculées en soustrayant la somme des pixels noirs à la somme des pixels blancs. [12]

3.5.3 Principe

Il est possible de calculer les caractéristiques à toutes les positions et à toutes les échelles dans une fenêtre de détection de petite taille, généralement de 24×24 pixels² ou de 20×15 pixels¹³. Viola et Jones fournissent l'exemple d'une fenêtre de taille 24×24 qui génère environ 160 000 caractéristiques, ce qui représente un nombre considérable de caractéristiques par fenêtre.

Au moment de la détection, on parcourt toute l'image en déplaçant la fenêtre de détection d'un certain pas dans les directions horizontale et verticale. Les échelles sont modifiées en augmentant successivement la taille de la fenêtre.

Selon Viola et Jones, un facteur multiplicatif de 1,25 est utilisé jusqu'à ce que la fenêtre couvre l'ensemble de l'image.

Enfin, et pour être plus résistantes aux fluctuations de l'éclairage, les fenêtres sont standardisées à l'aide de la variance².

Ces décisions techniques, en particulier l'utilisation d'images intégrales, entraînent une amélioration significative de l'efficacité, car les caractéristiques peuvent être évaluées rapidement, peu importe la taille de la fenêtre.

3.5.4 Eléments de la méthode

Viola et Jones utilisent la méthode de l'apparence [18], qui consiste à explorer l'image en calculant un certain nombre de caractéristiques dans des zones rectangulaires qui se superposent. Elle se distingue par l'utilisation de caractéristiques très basiques mais très variées. Les images intégrales sont une première innovation de la méthode, car elles permettent de calculer rapidement ces caractéristiques. La sélection de ces caractéristiques par boosting est une autre innovation majeure, en utilisant les caractéristiques comme des classifieurs. Enfin, la proposition de la méthode consiste à intégrer les classifieurs boostés dans un processus en cascade, ce qui permet d'obtenir un gain considérable en temps de détection.

La méthode Viola et Jones, en tant que méthode d'apprentissage supervisé, est structurée en deux phases : une phase d'apprentissage du classifieur à partir d'un grand

nombre d'exemples positifs (objets d'intérêt, visages) et d'exemples négatifs, et une phase de détection par application de ce classifieur à des images inconnues.

3.5.5 Caractéristiques de Haar

Les classificateurs de Viola et Jones utilisent les caractéristiques de type Haar qui sont des caractéristiques locales calculées en soustrayant la somme d'une sous-région de la caractéristique de la somme de la région restante de la caractéristique. Viola et Jones utilisent trois types des caractéristiques : deux rectangles, trois rectangles et quatre rectangles. La valeur d'une caractéristique à deux rectangles est la différence entre la somme des pixels dans deux régions rectangulaires. Les régions ont la même taille et la même forme et sont adjacentes horizontalement ou verticalement. Une caractéristique à trois rectangles calcule la somme dans deux rectangles extérieurs soustraits de la somme dans un rectangle central. Enfin, une caractéristique à quatre rectangles calcule la différence entre les paires diagonales de rectangles. [26]

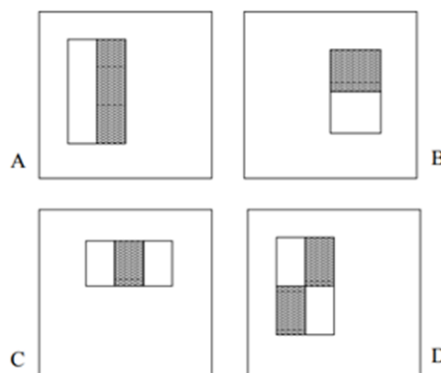


Figure 3.10 : Illustration des caractéristiques de Haar : (A, b) deux rectangles, (c) trois rectangles, (d) quatre rectangles. [26]

3.5.6 Image intégrale

L'image intégrale est une matrice de la même dimension que l'image a traitée, ou l'amplitude de chaque élément de coordonnée (x,y) dans l'image intégrale représente la somme des amplitudes de tous les pixels au-dessus et à gauche de l'image y compris le pixel de coordonnée (x,y) d'où le nom image intégrale. Cette image sera donc une représentation intermédiaire dans le traitement pour diminuer le temps de calcul et cela pour calculer d'une manière efficace et rapide les caractéristique pseudo-Haar. Soit (ii) l'image intégrale de l'image initiale(i) et $i(x, y)$ la valeur de l'image intégrale au point (x, y) .

On peut également définir l'image intégrale I_i par :

$$I_i(x, y) = \sum_{x' \leq x; y' \leq y} i(x', y') \dots\dots\dots(3)$$

Où $i(x, y)$ est l'image intégrale et $i(x', y')$ l'image principal.

Étant donné que nous employons cette nouvelle représentation afin de réduire le temps de calcul, on peut expliquer ses bénéfices. Premièrement, elle peut être déterminée de manière efficace en utilisant la paire des récurrences suivante :

$$S(x, y) = S(x, y - 1) + i(x, y) \dots\dots\dots(4)$$

$$I_i(x, y) = I_i(x - 1, y) + S(x, y) \dots\dots\dots(5)$$

C'est pourquoi $S(x, y)$ est la somme totale. [36]

$$\forall x, s(x - 1) = 0 ; \text{ et } \forall y, I_i(-1, y) = 0 \dots\dots\dots(6)$$

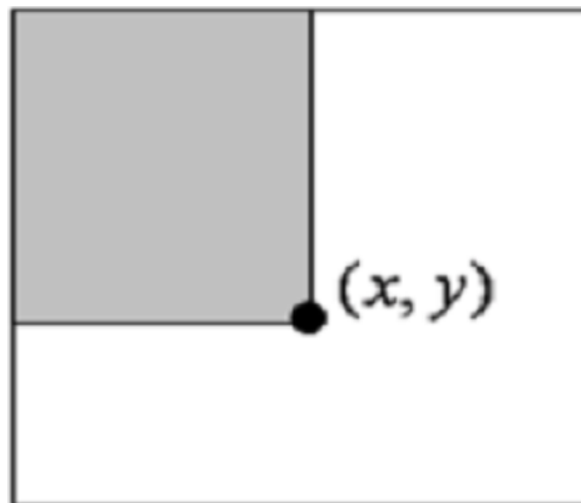


Figure 3.11 : La valeur de l'image intégrale au point (x, y) .

Avec l'image intégrale, toute somme rectangulaire peut être calculée en quatre opérations. Il est clair que la différence entre deux sommes rectangulaires peut être calculée en huit références. Comme les entités à deux rectangles définis ci-dessus impliquent des sommes rectangulaires adjacentes, elles peuvent être calculées en six références de matrice, huit dans le cas des trois entités rectangulaires et neuf pour des entités à quatre rectangles. [36]

1	1	1
1	1	1
1	1	1

Image d'entrée

1	2	3
2	4	6
3	6	9

Image intégrale

Figure 3.12 : Exemple d'image intégrale.

3.5.7 Filtre rectangulaire

Un détecteur de visages a été créé par Viola & Jones, connu sous le nom d'AdaBoost, qui repose sur des filtres rectangulaires de Haar, comme illustré dans la figure 3.12. [37]

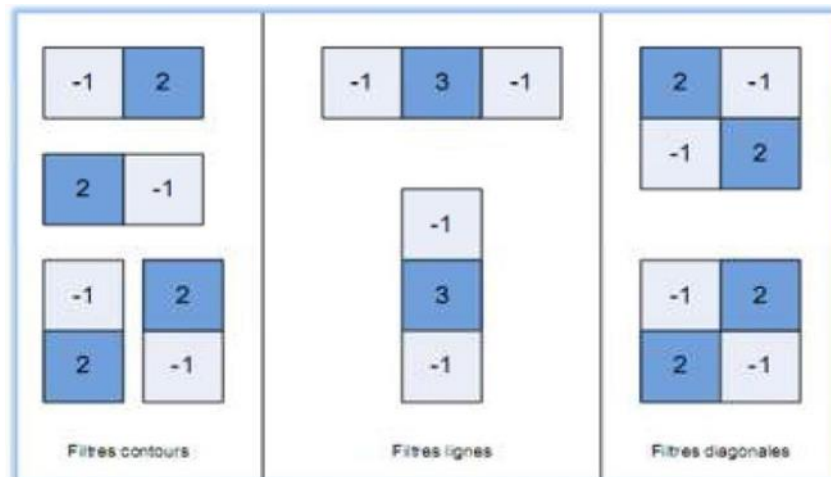


Figure 3.13 : Le filtre rectangulaire de Haar est utilisé par la cascade.

Les filtres rectangulaires permettent d'obtenir des données sur la répartition des niveaux de gris entre deux zones adjacentes dans l'illustration. L'ensemble des filtres de Haar utilisés est illustré dans la figure 3.12.

Les filtres sélectionnés sont les rectangles à deux et à trois. Dans le cas d'un filtre à trois rectangles, la valeur (sortie) d'un filtre appliqué à une région de l'image est calculée en soustrayant la somme des pixels dans le rectangle bleu clair de la somme des pixels dans le rectangle bleu foncé (multipliée par un coefficient) [38].

3.5.8 Cascade attentionnelle

Pendant la phase d'entraînement, Viola et Jones ont utilisé l'algorithme d'AdaBoost (Adaptive boosting). Où L'idée générale de cet algorithme consiste pour chaque classificateur faible ultérieur à corriger les erreurs du classificateur précédent. Pour ce faire, ils attribuent une pondération à chaque exemple d'apprentissage, entraînent les classificateurs, choisissent le meilleur classificateur et mettent à jour les pondérations en fonction de l'erreur du classificateur. Les exemples mal étiquetés auront des poids plus importants et seront donc correctement classés par le prochain classificateur choisi. Tous les classificateurs obtenus à l'étape précédente sont des classificateurs faibles et pour créer un classificateur puissant, Viola et Jones ont utilisé la cascade attentionnelle. Cette technique permet la création d'une série des classificateurs qui sont organisés de complexe a plus complexe.

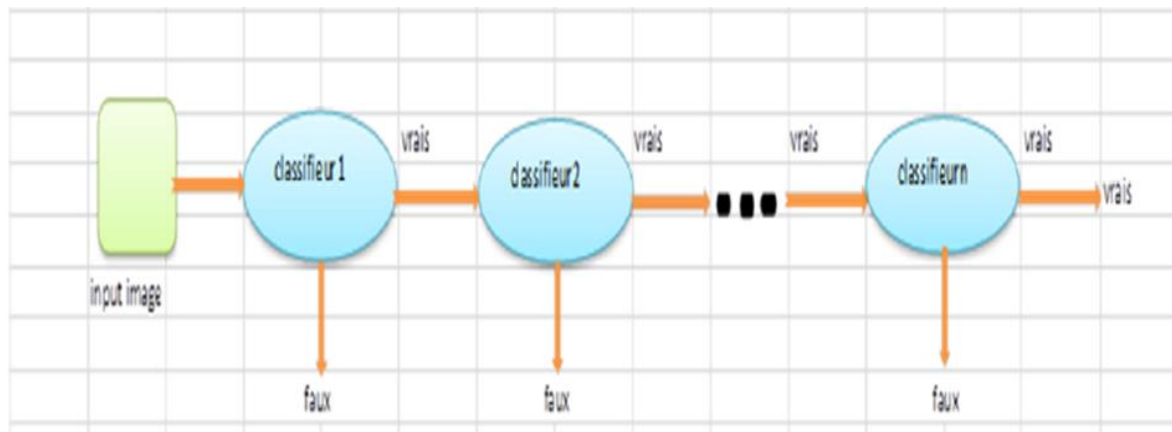


Figure 3.14 : Architecture de cascade.

3.6 La détection des yeux

La détection des yeux à l'aide de la méthode de Viola et Jones est une technique couramment utilisée dans la vision artificielle pour repérer des caractéristiques spécifiques dans des images, comme des visages ou des yeux. Cette méthode repose sur l'apprentissage automatique et utilise des caractéristiques de Haar, une approche efficace pour la détection d'objets en temps réel.

Étapes de la détection des yeux avec la méthode de Viola et Jones :

- **Préparation de l'image :** L'image est d'abord convertie en niveaux de gris pour simplifier le traitement, car la méthode de Viola et Jones fonctionne plus rapidement avec des images en noir et blanc.

- **Utilisation de caractéristiques de Haar** : La méthode applique des "caractéristiques de Haar", qui sont des motifs rectangulaires simples (par exemple, des zones claires et sombres) utilisés pour détecter les contours et les structures dans une image. Ces caractéristiques permettent d'identifier des motifs similaires à ceux des yeux.
- **Fenêtre de détection glissante** : La méthode scanne l'image avec une "fenêtre glissante", testant chaque section de l'image pour voir si elle contient les caractéristiques d'un œil. Cette fenêtre change de taille à chaque étape pour détecter des yeux à différentes échelles.
- **Algorithme de type cascade** : La méthode Viola-Jones utilise une cascade de classifieurs entraînés pour filtrer rapidement les régions qui ne contiennent probablement pas d'yeux, tout en se concentrant sur celles qui sont susceptibles de les contenir. Cela accélère le processus de détection en éliminant rapidement les zones inutiles.
- **Post-traitement** : Une fois que les yeux sont détectés, d'autres techniques peuvent être appliquées, telles que l'analyse de l'état des yeux (ouverts ou fermés) en fonction de la forme et de la position de l'iris ou des paupières.

3.7 Détection l'état des yeux

Dans notre projet nous avons utilisé la transformée de Hough circulaire pour évaluer l'état des yeux (ouverts/ fermés). Cette méthode a été choisie en raison de sa capacité à détecter des formes spécifiques, comme le cercle dans l'image (voir le chapitre 2). En appliquant la transformée de Hough circulaire nous avons pu identifier avec précision les contours et les caractéristiques des yeux tels que : le centre de la pupille et de l'iris, ce qui est essentiel pour notre analyse.

3.7.1 Détermination du rayon et du centre de la pupille et de l'iris

La détermination du centre de la pupille se fait de la manière suivante :

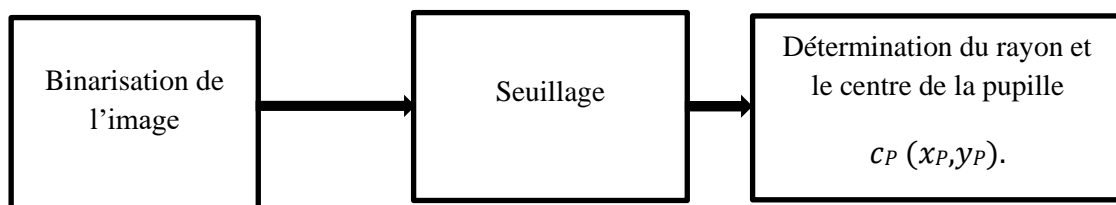


Figure 3.15 : Schéma synoptique qui illustre les étapes de détection du centre de la pupille.

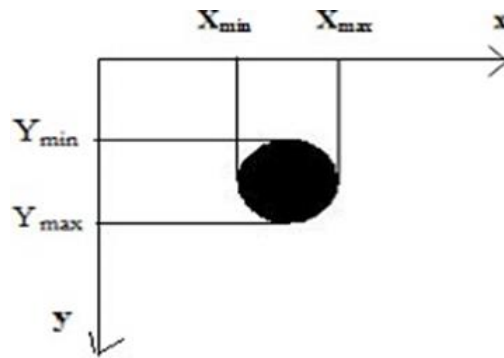


Figure 3.16 : Détermination du rayon et du centre de la pupille.

Le rayon et le centre de la pupille sont donnés par les formules suivantes :[30]

$$RP = (xmax - xmin)/2$$

$$xp = Rp + xmin$$

$$yp = Rp + ymin$$

3.7.2 Détermination du Rayon de l'iris

Une fois que CP(xP, yP) a été déterminé, il est possible d'extraire le rayon de l'iris Ri à partir du centre de la pupille. En partant de ce point (CP) que nous avons identifié, nous progressons jusqu'à obtenir un second contour. Le point (Xi) où nous sommes arrivés est donc identifié : il se trouve au bord de l'iris. [30]

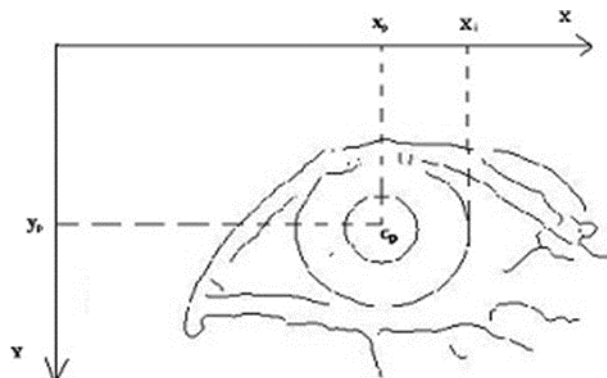


Figure 3.17 : Détermination du Rayon de l'iris.

Ainsi, le rayon de l'iris est calculé selon l'équation suivante :[30]

$$Ri = CP - xi$$

En utilisant ces équations, il est facile d'isoler la région de l'iris.

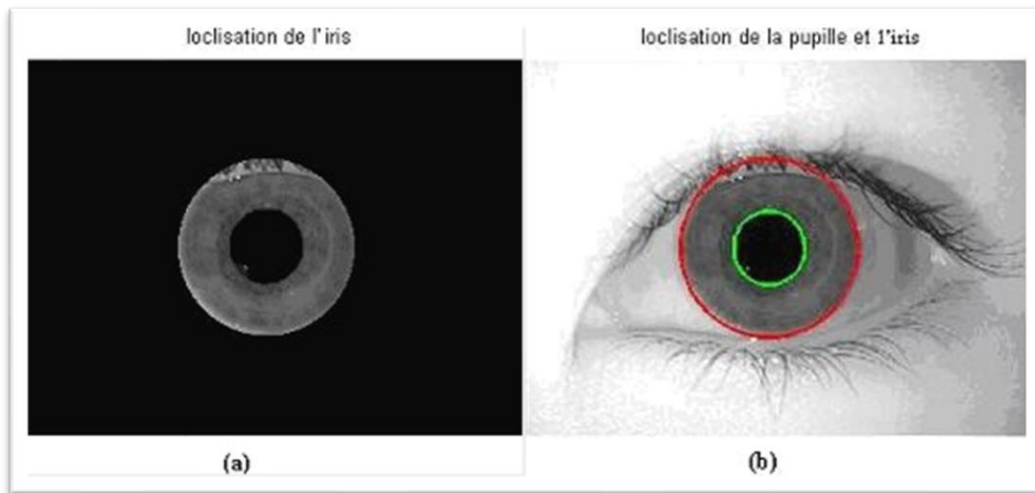


Figure 3.18 : (a) région de l'iris, (b) localisation de la pupille et de l'iris.

3.8 Avertissements

Dans les systèmes de détection de somnolence au volant, la mise en place d'un système d'avertissement efficace est cruciale pour assurer la sécurité du conducteur et des passagers. Lorsqu'un risque de somnolence est identifié, des alertes sont émises pour prévenir le conducteur et parmi ces derniers on trouve :

- ***Alerte Vibratoire***

Le volant ou le siège du conducteur peut vibrer pour signaler une alerte, permettant ainsi de capter l'attention de manière subtile mais efficace.

- ***Alerte Visuelle***

Des voyants lumineux ou des messages affichés sur le tableau de bord informent le conducteur d'un danger potentiel lié à la somnolence.

- ***Système de Notification***

Une notification peut être envoyée à un appareil mobile connecté pour inciter le conducteur à faire une pause.

- ***Alerte sonore***

Une alerte sonore émise par des haut-parleurs (baffles) avertit le conducteur en cas de somnolence, afin de prévenir les accidents et inciter à corriger son comportement.

Dans notre projet, nous avons opté pour l'alerte sonore via des baffles. Lorsqu'un signe de somnolence est détecté, un signal sonore est diffusé à travers les haut-parleurs.

3.9 Discussion

Dans ce chapitre, nous avons présenté notre approche de détection de la somnolence basée sur la vision artificielle, nous avons utilisé la méthode viola et Jones pour localiser rapidement et précisément le visage et les yeux, où nous avons approfondi la méthode viola et Jones de manière générale. Ainsi que la transformée de Hough circulaire pour analyser l'état des yeux et qui permettent de déterminer s'ils sont ouverts ou fermés.

CHAPITRE 4 :

Implémentation et résultats

4.1 Préambule

Ce chapitre pratique explore les étapes nécessaires pour réaliser un système portatif de détection de somnolence au volant, en intégrant des techniques avancées de traitement d'images et de détection en temps réel.

En plus dans ce chapitre, nous détaillerons les différentes phases de développement du système, y compris la configuration des outils nécessaires, la création des algorithmes de détection, et l'intégration des composants matériels et logiciels pour aboutir à une solution fonctionnelle.

Et enfin, les résultats de l'application des algorithmes viola / Jones et transformée de Hough circulaire seront présentés et discuté.

4.2 Matériels et méthodes

Dans ce projet, nous avons proposé une méthode de détection de somnolence basée sur l'algorithme de viola et Jones et la transformée de Haugh circulaire. En premier lieu, le système détecte le visage du conducteur à l'aide de la bibliothèque spécialisée en traitements d'images et d'apprentissage automatique (Open CV), Permet de capturer et d'analyser les images en temps réel. Et d'extraire les images à partir de séquences vidéo.

Une fois les images extraites, le système détecte les visages et les yeux avec la méthode de Viola et Jones, l'algorithme localise la pupille de l'œil et détecte l'état des yeux (ouverts /fermés) et cela se fait avec la transformée de Hough circulaire.

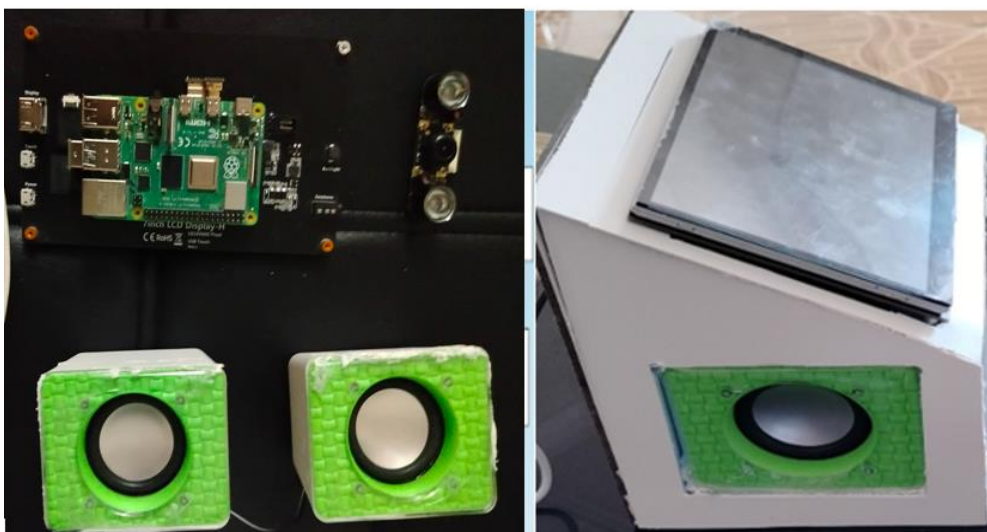


Figure 4.1 : Architecture du système.

4.2.1 Matériel

Pour la réalisation de ce projet nous avons utilisé :

- Un Raspberry pi 4 modèle b. (2GB de ram).
- Camera infrarouge.
- Baffle.
- Ordinateur portable (dell i5 7generation).
- Câblemicro-HDMI.
- Téléviseur.
- Clavier et souris. (Sans fils).

4.2.2 Méthodes

Pour la réalisation de ce projet nous avons en premier lieu installé le logiciel RASPBERRY PI IMAGER V 1.8.5 qui est un outil développé par la Fondation Raspberry Pi4 pour faciliter l'installation des systèmes d'exploitation sur les cartes SD destinées aux appareils Raspberry.

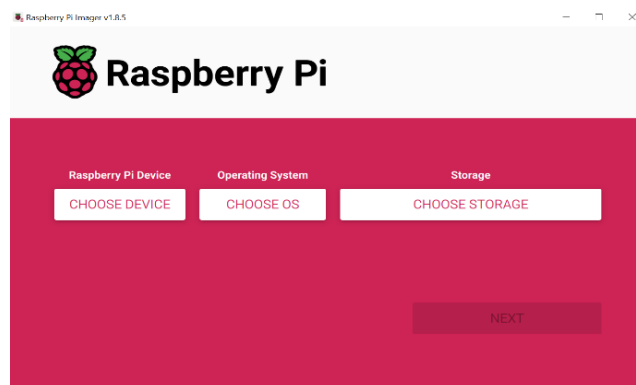


Figure 4.2 : Interface d'installation du système d'exploitation Raspberry Pi4 à l'aide de Raspberry Pi Imager.

Puis nous avons téléchargé Raspberry Pi4 OS (64-bit) à partir du site officiel de la Fondation Raspberry Pi4. Qui est une version du système d'exploitation officiel pour les appareils Raspberry Pi4, optimisée pour les architectures 64 bits.

Ensuite nous avons téléchargé le logiciel ADVANCED IP SCANNER afin de pouvoir récupérer l'adresse IP de notre raspberry pi pour y accéder via SSh depuis le cmd avec la commande suivante : `ssh pi@ 192.000.0.0` (dont pi est le nom d'utilisateur).

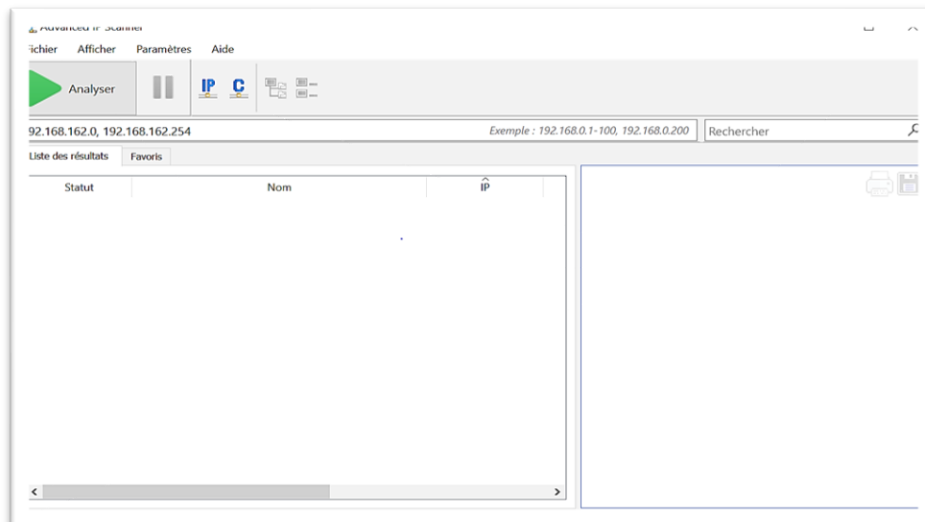


Figure 4.3 : Logiciel permettant la recherche des adresses IP.

Une fois connectés au Raspberry Pi4, nous avons installé Raspberry Pi4 Connect avec la commande suivante : `sudo apt install raspberry-connect`. Ensuite, nous avons créé un compte sur le site Raspberry Pi4 Connect depuis notre PC, ce qui nous a permis d'accéder au Raspberry Pi4 avec nos appareils (téléphone et PC).

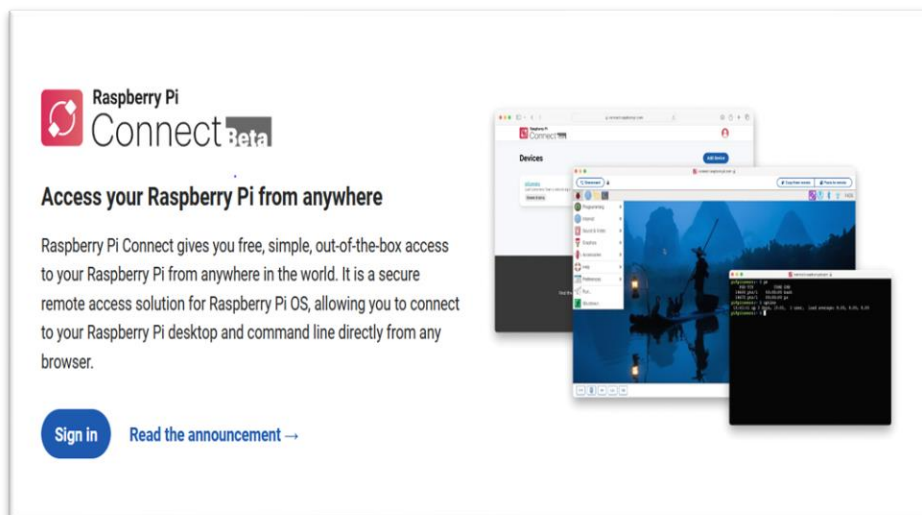


Figure 4.4 : Logiciel permettant la Gestion à Distance.

Une fois accédé, nous avons vérifié que Python3 était bien installé en tapant `python` dans le terminal de notre carte. Ensuite, nous avons testé la présence de la bibliothèque Open CV en lançant la commande '`import cv2`' pour confirmer son installation.

detectMultiScale, il identifie les visages dans chaque image et dessine des rectangles autour de ces visages détectés. L'image a noté est affichée dans une fenêtre.

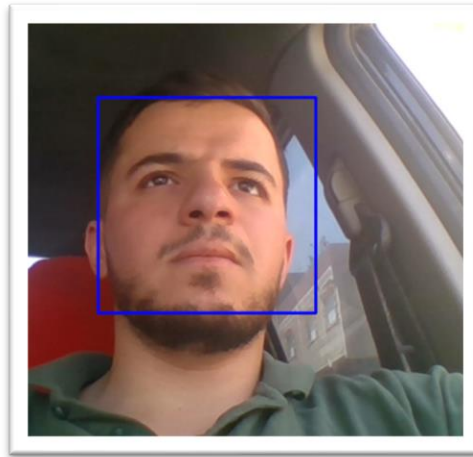


Figure 4.7 : Image illustrative de la détection du visage.

4.4 Détection des yeux

Une fois qu'un visage est détecté, le programme définit une région d'intérêt (ROI) autour de la région du visage. Cette REINE est ensuite utilisée pour localiser les yeux à l'intérieur du visage. Le classificateur en cascade ('haarcascade_eye.xml') pour les yeux est appliqué à cette région spécifique, permettant de détecter les yeux plus précisément en se concentrant uniquement sur la partie du visage. Après la détection des yeux, des rectangles sont dessinés autour de chaque œil détecté, permettant de visualiser leur emplacement exact dans l'image.

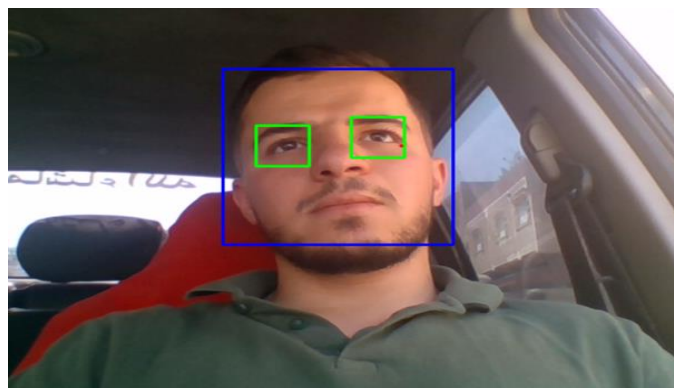


Figure 4.8 : Image illustrative de la détection des yeux.

4.5 Description de notre système

Le programme commence par charger les classificateurs en cascade pour la détection des visages et des yeux, et initialise la caméra. Il vérifie si la caméra s'ouvre correctement, et si ce n'est pas le cas, il affiche un message d'erreur et termine le programme. Ensuite, il initialise le module de mixage de Pygame et charge un fichier audio pour l'alerte sonore.

Le programme définit des variables pour suivre le nombre de frames où les yeux sont fermés et un seuil au-delà duquel une alerte est déclenchée. Une boucle infinie est ensuite utilisée pour capturer les frames de la caméra. Chaque frame est converti en niveaux de gris pour simplifier le processus de détection.

Les visages sont détectés dans chaque frame en utilisant le classificateur en cascade pour les visages. Pour chaque visage détecté, une région d'intérêt (ROI) est définie, et le classificateur en cascade pour les yeux est appliqué à cette région pour détecter les yeux. Les détections des yeux sont ensuite filtrées pour garder uniquement celles situées dans la partie supérieure du visage, car les yeux doivent se trouver dans cette région.

Le programme vérifie ensuite si les yeux sont fermés. Si aucun œil n'est détecté dans la région définie, il incrémente un compteur de frames pour les yeux fermés. Si des yeux sont détectés, le compteur est réinitialisé à zéro et l'alerte sonore est arrêtée si elle était en cours de lecture. Les yeux détectés sont entourés de rectangles verts, et la transformation de Hough circulaire est utilisée pour détecter les cercles dans les yeux, représentant les iris. Si le nombre de frames consécutives avec les yeux fermés dépasse le seuil défini, une alerte visuelle est affichée sur l'écran, et l'alerte sonore est déclenchée si elle ne l'était pas déjà. Le programme continue d'afficher les frames capturés avec les rectangles et les cercles détectés.

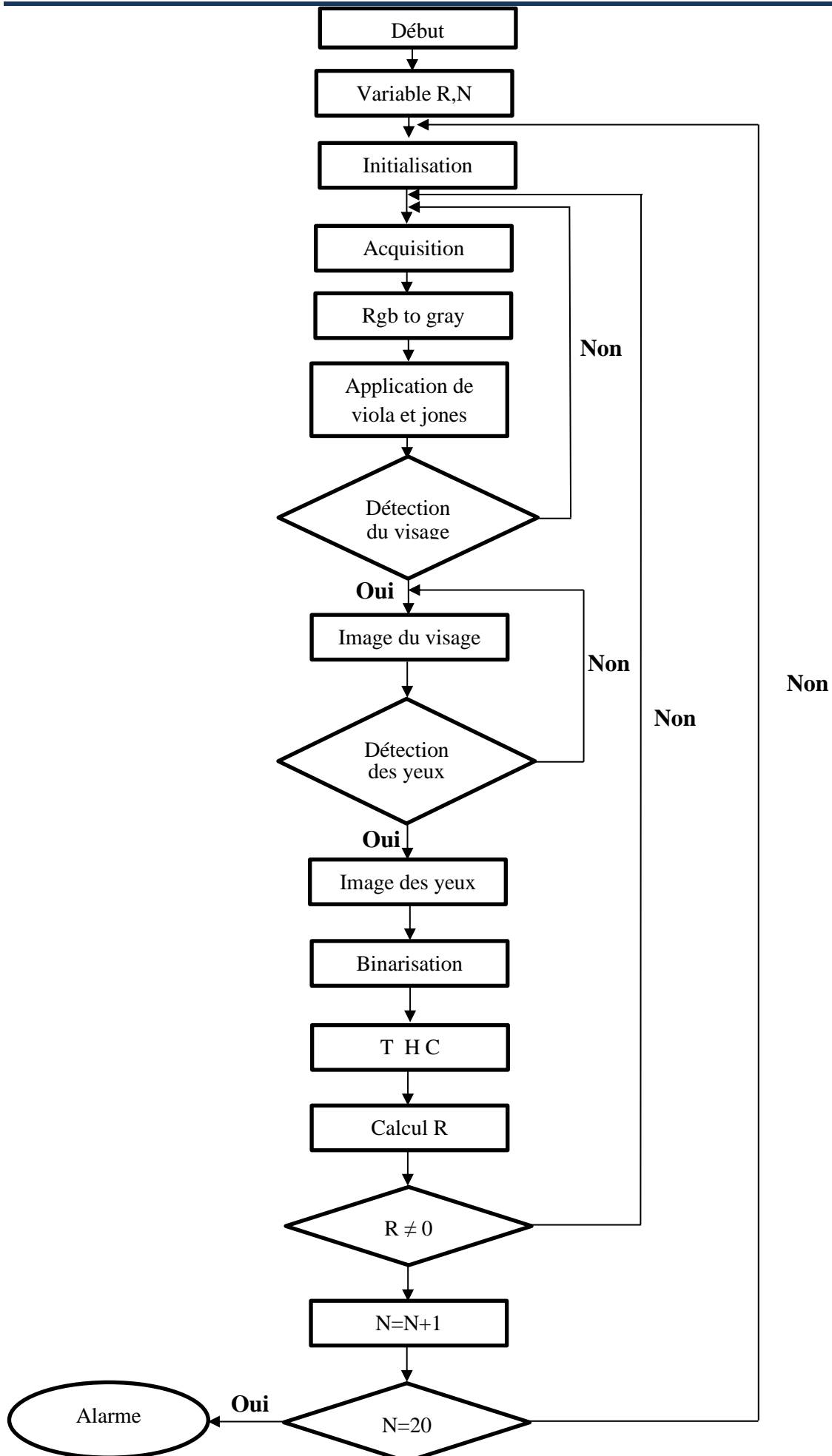


Figure 4.9 : Organigramme de notre système.

4.6 Résultats

L'objectif de cette section est de présenter les résultats obtenus à partir des tests de notre méthode proposée, l'algorithme de Viola et Jones. Les résultats obtenus sont les suivants :

- Lorsque les yeux restent immobiles, c'est-à-dire lorsque le regard est fixé droit devant, cela indique que le conducteur est vigilant.

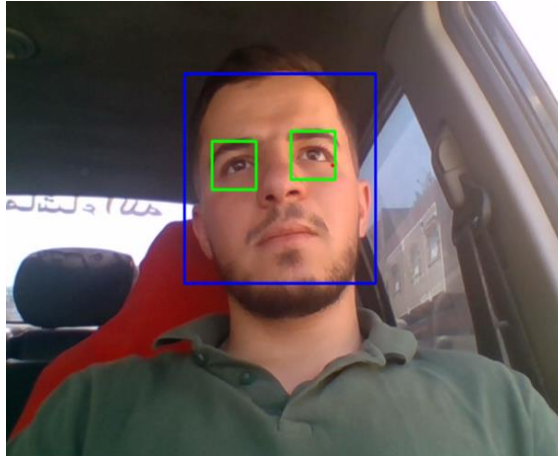


Figure 4.10 : Image illustrative d'un conducteur en état de vigilance.

- Lorsqu'il y a un obstacle entre l'œil et la webcam, un mouvement aléatoire, ou si les yeux ne sont pas détectés, cela indique une somnolence, et l'alarme se déclenche.



Figure 4.11 : Image illustrative d'un conducteur en état de somnolence.

- Cette image illustre la direction à droite des yeux capturée par la caméra suivant la direction à droite du globe oculaire.

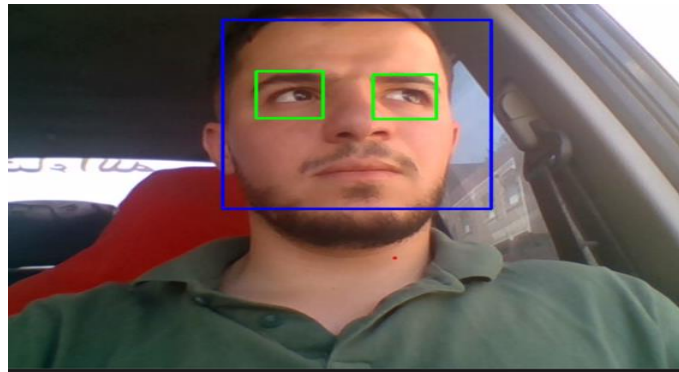


Figure 4.12 : Image illustrative d'un conducteur dont le regard est fixe à droite.

- Cette image illustre la direction à gauche des yeux capturée par la caméra suivant la direction à gauche du globe oculaire.

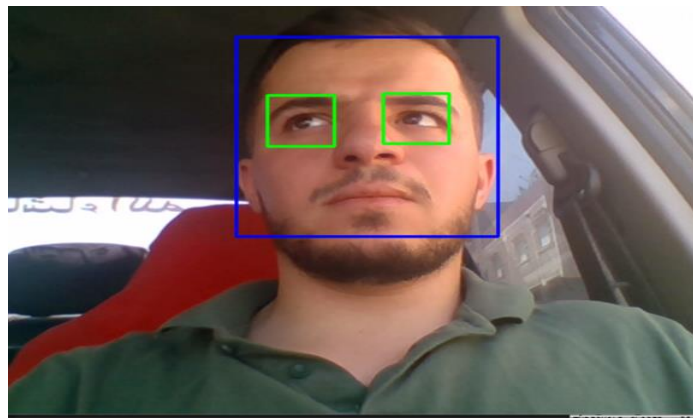


Figure 4.13: image illustrative d'un conducteur dont le regard est fixé à droite.

4.7 Discussion

Dans ce chapitre, nous avons présenté les différentes étapes de La détection de somnolence au volant à l'aide du Raspberry pi4 et la caméra infrarouge placée en face du conducteur, à savoir :

- Détection du visage.
- Détection des Yeux.
- Détection de somnolence.

Dans cette section, nous avons également défini les étapes et le fonctionnement du programme proposé sous Python, et nous avons expliqué les résultats obtenus à l'aide de captures illustrant chaque étape du travail.

CONCLUSION GENERALE

Dans ce projet nous avons abordé plusieurs méthodes de détection de somnolence au volant nous avons opté pour la méthode de Viola et Jones combinée avec la transformée de Hough circulaire parce que la rapidité de l'algorithme permet de détecter des objets en temps réel, ce qui est essentiel pour les applications interactives telles que la surveillance et les interfaces utilisateur. L'utilisation de l'image intégrale contribue à l'efficacité de calcul en permettant de calculer rapidement les caractéristiques nécessaires, réduisant ainsi le temps de traitement et augmentant l'efficacité globale.

En outre, les caractéristiques de Haar utilisées par l'algorithme sont robustes aux variations d'éclairage et de contraste, permettant une détection fiable dans diverses conditions d'éclairage. Le classificateur en cascade permet une détection multi-échelle, essentielle pour identifier des objets de tailles variées dans une même image.

Enfin, la précision de l'algorithme est renforcée par l'apprentissage supervisé à partir de grands ensembles de données d'exemples positifs et négatifs, rendant le classificateur très précis pour la détection des objets spécifiques qu'il a été entraîné à reconnaître.

Nous avons conçu un système portatif et non intrusif qui surveille les yeux du conducteur pour détecter la somnolence. Ce système a base d'une carte Raspberry Pi4, utilise l'algorithme Viola-Jones sous python (bibliothèque open cv) pour effectuer cette détection.

Les résultats obtenus montrent l'efficacité de l'algorithme de Viola et Jones pour l'identification des visages ainsi la transformée de Hough circulaire pour évaluer l'état des yeux du conducteur. Ce travail représente une avancée significative vers la création de systèmes de sécurité routière plus accessibles et adaptables. Bien que le prototype développé nécessite encore des améliorations avant une adoption à grande échelle, les bases posées par ce mémoire offrent un cadre solide pour le développement futur de systèmes de détection de somnolence fiables et économiques.

Après avoir testé notre système avec succès, nous avons identifié certains aspects à améliorer. Notamment, si le conducteur est atteint de surdit  ou bien malentendants, il faut envisager une autre approche pour le syst me d'alerte tels que, l'am lioration du syst me d'alerte (nous pourrions int grer des vibrations dans le volant ou le si ge du conducteur.) et l'int gration avec d'autres syst mes de Surveillance   distance (connexion au cloud pour surveiller les conducteurs en cas de somnolence pour pouvoir intervenir dans les d lais).

BIBLIOGRAPHIE

- [1] C. Solomon and T. Breckon, Fundamentals of digital Image Processing, Hoboken: wiley Blackwell, 2011.
- [2] L. Diane, cours de traitement d'images, nice: Laboratoire I3S, 2004, p. 10.
- [3] G. Peyré., "le traitement numérique des images," Archive ouverte HAL, 2011.
- [4] N. MERABE, M MAHLIA, recherche d'images par le contenu, université Abou Bakr belkaid- tlemcen.2011
- [5] C. Houqssine, segmentation d'images par une approche biomimetique hybride, université m'hamed bougara- Boumerdes. 2012.
- [6] LOURCI Mohamed Amine, "Segmentation de Blocs d'Images de documents Imprimés", Mémoire de Magister, Université de SAAD DAHLEB de BLIDA, Juin 2015.
- [7] H. KHAMELI and H. KALLAL, "Segmentation d'images satellitaire par Cuckoo," UNIVERSITE ABDELHAMID IBN BADIS, MOSTAGANEM, 2012.
- [8] S. Hu and G. Zheng, "Driver Drowsiness Detection with Eyelid related Parameters by Support Vector Machine", Expert Systems with Application, vol.36, pp. 7651-7658, 2009.
- [9] z. Zhang, "Introduction to machine learning: K-nearest neighbors," Annals of translational Medicine, vol. 4, p. 218, 2016.
- [10] M. Patel, S. K. Lal, D. Kavanagh and P. Rossiter, "Applying neural network analysis on heart rate variability data to assess driver fatigue," Expert Systems with Applications, vol. 38, p. 7235–7242, 2011.
- [11] S. Arun, S. Kenneth and M. Murugappan, "Detecting Driver Drowsiness Based on Sensors: A Review," {Sensors (Basel, Switzerland, vol. 12, pp. 16937-16953, 2012.
- [12] j. krajewski, D. Sommer, U. Trutschul, D. Edwards and M. Golz, "Steering Wheel Behavior Based Estimation of Fatigue", Driven assessment conference, 2009.
- [13] M. Inger, T. Åkerstedt, B. Peters, A. Anund and G. Kecklund, "Subjective sleepiness, simulated driving performance and blink duration: Examining individual differences," Journal of sleep research, vol. 15, pp. {47-53, 2006.

- [14] M. Gillberg, G. Kecklund and T. Åkerstedt, "Relations between performance and subjective ratings of sleepiness during a night awake," *Sleep: Journal of Sleep Research & Sleep Medicine*, vol. 17, pp. 236-241, 1994.
- [15] J. May and C. Baldwin, "Driver fatigue: The importance of identifying causal factors of fatigue when considering detection and countermeasure technologies," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 12, pp. 218-224, 2009.
- [16] I. Gracia, S. Bronte, L. M. Bergasa, J. Almazan and J. Yebes, "vision-based drowsiness detector for real driving conditions" *IEEE Intilligent Vehicles Symposium*, pp. 618-623, 2012.
- [17] W. Zhang, B. Cheng and Y. Lin, "Driver drowsiness recognition based on computer vision technology," *Tsinghua Science and Technology*, vol. 17, pp. 354-362, 2012.
- [18] A. Rosebrock, "Drowsiness detection with OpenCV," 08 May 2018. [Online]. Available: <https://www.pyimagesearch.com/2017/05/08/drowsiness-detection-opencv/>. [Accessed 29May 2019].
- [19] G. Reza, G. Marnim and A. Vassilis, "A Realistic Dataset and Baseline Temporal Model for Early Drowsiness," *CoRR*, vol. abs/1904.07312, 2019.
- [20] M. Ochocki and D. Sawicki, "Yawning Recognition based on Dynamic Analysis and Simple Measure," in *Proceedings of the International Conference on Computer- Human Interaction Research and Applications*, in Funchal, Madeira, Portugal, pp. 111- 117, 2017.
- [21] M. Omidyeganeh, S. Shirmohammadi, S. Abtahi, A. Khurshid, M. Farhan, J. Scharcanski, B. Hariri, D. Laroche and L. Martel, "Yawning Detection Using Embedded Smart Cameras," *IEEETransactions on Instrumentation and Measurement*, vol. 65, pp. 1- 13, 2016.
- [22] T. H. Lam and N. T. Do, "Algorithm to detect driver's drowsiness based on nodding behavior," *International Journal of Soft Computing, Mathematics and Control (IJSCMC)*, vol.5, 2016.
- [23] M. Omidyeganeh, A. Javadtalab and S. Shirmohammadi, "Intelligent driver drowsiness detection through fusion of yawning and eye closure," *IEEE International Conference on Virtual Environments, Human-Computer Interfaces and Measurement Systems Proceedings*, pp. 1-6, 2011.

- [24] V. Esra, C. Mujdat, E. Aytul, L. Gwen, B. Marian and M. Javier, "Automated Drowsiness Detection for Improved Driving Safety," International Conference on Automotive Technologies, 2008.
- [25] B. Reddy, Y. Kim, S. Yun, C. Seo and J. Jang, "Real-Time Driver Drowsiness Detection for embedded System Using Model Compression of Deep Neural Networks," 2017.
- [26] G. Nisha, N. Tejas and Y. Aditya, "Drowsy Driver Detection in Video Sequences using LSTM with CNN Features," 31 Des 2017. [Online]. Available: <https://github.com/nishagandhi/DrowsyDriverDetection>. [Accessed 21 July 2019].
- [27] Viola, P., & Jones, M.: Rapid object detection using a boosted cascade of simple features. In Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. Vol. 1, pp. I-I (2001)
- [28] 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)
- [29] Intelligent Behaviour Understanding Group (iBUG).
<https://ibug.doc.ic.ac.uk/resources/facial-point-annotations/>
- [30] Li Ma, Yunhong Wang, Tieniu Tan, « Personal identification based on iris texture analysis », IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 25, N° 12, pp. 1519-1533, December 2003.
- [31] Détection de Visages par Méthode Hybride AdaBoost et Couleur de Peau : Mémoire de Magister Présenté à l'Université de Guelma Faculté des Sciences et de l'Ingénierie.
- [32] G. Pierre, <<la méthode de Viola et Jones>>, 2013.
- [33] Ligne [Archive] en Paul Viola et Michael Jones<<Robust Real-time Face detection>>, IJCV, 2004, p.137_154.
- [34] P. Viola and M. Jones, "Robust Real-Time Face Detection," International Journal of Computer Vision, vol. 57, pp. 137--154, 01 May 2001.

[35] H. Vietman, reconnaissance des visages sur le téléphone portable, mémoire de Master. Institut de la francophonie pour l'informatique (IFI), octobre 2010.

[36] A. Vegri Pablo, P.lionel, C, Xavier, « cascade de classifieur génératifs et discriminants pour la détection des véhicules », Article, institut des systèmes Intelligents et robotiques, université Pierre et Marie Paris 6.

[37] L. Alexis, « Bagging et Bosting », conférence, université de Caen, février 2014.