



REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université Mohamed Khider – BISKRA

Faculté des Sciences Exactes, des Sciences de la Nature et de la Vie

Département d'informatique

N° d'ordre : IVA03/IVA/M2/2023

Mémoire

Présenté pour obtenir le diplôme de master académique en

Informatique

Parcours : Image et Vie Artificielle (IVA)

Titre : Conception et mise en œuvre d'un système d'assistance et d'aide à la conduite

Par :

LEBBAZ ZOULIKHA

Soutenu le. 20/06/2023 devant le jury composé de :

Babahenini Djihane

MCB

Président

FEKRAOUI Farah

MCB

Rapporteur

Belaiche Hamza

MCB

Examineur

Année universitaire 2022-2023

شكر و عرفان

رب اوزعني ان اشكر نعمتك التي انعمت علي وعلى والدي وان اعمل صالحا
ترضاه وأدخلني برحمتك في عباده الصالحين
الآية 19 من سورة النمل

أول الشكر وأخره لله عز وجل الذي رزقنا العقل وحسن التوكل عليه سبحانه وتعالى وألهمنا
الصبر ووقفنا إلى إتمام هذه المذكرة المتواضعة
بعد رحلة بحث طويلة وجهد واجتهاد

لا يسعنا إلا أن نخص بالذكر باسمي عبارات الشكر والتقدير إلى من شرفني بإشرافها
الأستاذة فكاوي فرح لما قدمته من توجيهات ونصائح طيلة فترة انجاز هذه المذكرة جزاك الله كل خير
كما لا ننسى أن أتقدم بأرقى وأتمن عبارات الامتنان لأعضاء لجنة التقييم على تكريمهم لقبول مناقشة هذه
المذكرة اسأل الله أن يبارك فيهم وفي عملهم
وأخيرا أجدد الشكر لكل من ساندني في انجاز هذه المذكرة

أهداء

أهدي ثمرة هذا العمل إلى من أفتخر بأبوته وأحمل اسمه أبي العزيز بن زكري، وإلى من علمتني الصبر والكفاح
إلى حبيبتي أمي الغالية مسعودة حفصهما الله وأطال في عمرهما
إلى من بهم أكبر وعليهم أعتمد إخوتي عبدو، سعيد، ضياء وأخواتي زينب، امينة، فطومة، مريم
حفظهم الله ورعاهم وجعلهم خير السند
إلى صديقات المشوار اللاتي قاسمني لحضاته رعاهم الله ووفقهم
إلى من رفعوا رايات العلم والتعليم أساتذتي الأفاضل
وإلى كل من احبهم قلبي

Résumé

De nos jours, les systèmes de vision sont de plus en plus utilisés dans le contexte routier. Ils le sont par exemple pour la vidéo surveillance et les ADAS (systèmes avancés d’assistance et d’aide à la conduite). Ils permettent ainsi d’assurer la sécurité et faciliter la mobilité.

Le nombre d’accidents enregistrés sur le territoire national est considérable, ils font partie des événements les plus malheureux et les plus dangereux auxquels les gens peuvent être confrontés en conduisant. Des millions de personnes souffrent de blessures graves et d’invalidité. Cependant, la gravité des accidents de la route n’est pas moins importante que celle en milieu urbain. Il s’avère donc nécessaire de concevoir et de mettre en œuvre des systèmes d’aide à la conduite fiables et en temps réel basé sur l’analyse d’images. La conception de ce système est en effet une tâche essentielle pour améliorer la sécurité routière et la mobilité.

L’objectif de ce sujet est le développement d’un système capable de percevoir l’environnement, de détecter et de reconnaître les panneaux de signalisation et les dangers potentiels pour un conducteur en mouvement dans son environnement, et ce, afin d’éviter les collisions et d’assurer la sécurité des conducteurs et des piétons.

Nous avons donc développé un système de détection et de reconnaissance des panneaux de signalisation et des dangers potentiels (personnes, chats et chiens traversant la route et voitures) en utilisant de puissantes techniques de reconnaissance d’objets, à savoir les réseaux neuronaux profonds.

A cette fin, nous avons créé une nouvelle base de données à partir de deux bases d’images grand public, Cette base ne contient que les classes présentes en milieu urbain, où pour chaque classe, nous avons tâché de choisir les images les plus représentantes de l’objet en question.

Dans le but d’augmenter la performance du système proposé, nous avons utilisé

la notion de réseau spécialisé, c'est-à-dire, un réseau profond capable de détecter un seul type d'objet. Pour cela nous avons combiné entre deux réseaux profonds modèle yolov5s et qui travaillent conjointement, le premier pour les panneaux de signalisation et le deuxième pour les dangers présents sur une route en milieu urbain.

Mots-clés : Systèmes de vision, systèmes d'aide à la conduite fiables, temps réel, détection des dangers potentiels, Reconnaissance des panneaux de signalisation, réseaux neuronaux profonds.

Abstract

Nowadays, vision systems are increasingly used in road environments. They are used, for example, in video surveillance and ADAS (advanced driver assistance systems). They help to ensure safety and facilitate mobility.

The number of accidents recorded nationwide is considerable, and they are among the most unfortunate and dangerous events people can face while driving. Millions of people suffer serious injury and disability. However, road accidents are no less serious than those in urban areas. It is therefore necessary to design and implement reliable, real-time driver assistance systems based on image analysis. The design of such systems is essential to improving road safety and mobility.

The subject's objective is to develop a system capable of perceiving the environment, detecting and recognizing road signs and potential hazards for a moving driver in his environment, in order to avoid collisions and ensure the safety of drivers and pedestrians.

We have therefore developed a system for detecting and recognizing road signs and potential hazards (people, cats and dogs crossing the road and cars) using powerful object recognition techniques, namely deep neural networks.

To this end, we created a new database from two databases of images for the general public. This database contains only the classes present in the urban environment, where for each class we tried to select the images most representative of the object in question.

In order to increase the performance of the proposed system, we used the notion of a specialized network, i.e. a deep network capable of detecting a single type of object. To this end, we combined two deep networks, the yolov5s model, which work together : the first for traffic signs and the second for road hazards in an urban environment.

Keywords : vision systems, reliable driver assistance systems, real time, potential hazard detection, Traffic Sign Recognition, Deep Neural Networks.

المخلص

في الوقت الحاضر، يتم استخدام أنظمة الرؤية بشكل متزايد في بيئات الطرق. يتم استخدامها، على سبيل المثال، في المراقبة بالفيديو و ADAS (أنظمة مساعدة السائق المتقدمة). أنها تساعد على ضمان السلامة وتسهيل التنقل.

عدد الحوادث المسجلة على الصعيد الوطني كبير، وهي من بين أكثر الأحداث المؤسفة والخطيرة التي يمكن أن يواجهها الناس أثناء القيادة. يعاني ملايين الأشخاص من إصابات خطيرة وإعاقة. ومع ذلك، فإن حوادث الطرق لا تقل خطورة عن تلك التي تقع في المناطق الحضرية. لذلك من الضروري تصميم وتنفيذ أنظمة مساعدة السائق الموثوقة في الوقت الفعلي بناءً على تحليل الصور. يعد تصميم مثل هذه الأنظمة ضروريًا لتحسين السلامة على الطرق والتنقل.

الهدف من هذا الموضوع هو تطوير نظام قادر على إدراك البيئة، الكشف والتعرف على الإشارات المرورية والأخطار المحتملة للسائق المتحرك في بيئته، وذلك لتجنب الاصطدامات وضمان سلامة السائقين والمشاة.

لذلك قمنا بتطوير نظام لاكتشاف والتعرف على إشارات الطرق والمخاطر المحتملة (الأشخاص والقطط والكلاب الذين يعبرون الطريق والسيارات) باستخدام تقنيات التعرف على الأشياء القوية، وهي الشبكات العصبية العميقة.

تحقيقاً لهذه الغاية، أنشأنا قاعدة بيانات جديدة من قاعدتي بيانات للصور لعامة الناس. تحتوي قاعدة البيانات هذه فقط على الفئات الموجودة في البيئة الحضرية ، حيث حاولنا لكل فئة تحديد الصور الأكثر تمثيلاً للكائن المعني.

من أجل زيادة أداء النظام المقترح ، استخدمنا فكرة الشبكة المتخصصة ، أي شبكة عميقة قادرة على اكتشاف نوع واحد من الكائنات. لتحقيق هذه الغاية ، قمنا بدمج شبكتين عميقتين ، نموذج yolov5s ، اللتان تعملان معاً: الأولى لإشارات المرور والثانية لمخاطر الطرق في البيئة الحضرية.

الكلمات المفتاحية : أنظمة الرؤية، أنظمة مساعدة السائق الموثوقة، الوقت الفعلي، اكتشاف المخاطر المحتملة، التعرف على علامات المرور، الشبكات العصبية العميقة.

Liste d'abreviation

ACC	Adaptative Cruise Control
ADAS	Advenced Driver Assisstant System
AEB	autonomous emergency braking
ANN	Artificial Neural Networks
CNN	Convolutional Neural Network
DL	Deep Learning
Faster R-CNN	Faster Region Convolutional Neural Network
FPN	Pyramid of Feature Networks
GPS	Global Positioning Systems
HOG	Histogram of Oriented Gradients
HSI	Hue, Saturation, Intensity
IA	Intelligence aArtificielle
IHM	Interface Homme-Machine
IMU	Inertial Measurement Units
IOU	Intersection over Union généralisé
ITS	Intelligent Transport Systems
LCH	Lightness, Chroma et Hue
LC	Lane Centering
LDW	Lane Departure Warning
LIDAR	Laser Detection And Ranging
LKA	Lane Keeping Assist
Mask R-CNN	Mask Region Convolutional Neural Network
ML	Machine Learning
NMS	Non Maximum Suppression
PAN	Perceptual Adversary Networks

PMD	Photonic Mixer Device
RADAR	Radio Detection And Ranging
SSD	Single Shot Detection
SVM	Support Vector Machine
TSR	Traffic Signs Recognition
YOLO	You Only Look Once

Table des matières

Liste d'abreviation	i
Table des matières	i
Liste des Figures	iv
Liste des tableaux	vii
Introduction générale	1
Chapitre 1: Revue de littérature	4
1.1 Introduction	4
1.2 Traitement d'image et vision par ordinateur	4
1.2.1 L'image numérique	4
1.2.2 Traitement d'image	6
1.2.2.1 Étapes de base du traitement d'image	7
1.2.3 vision par ordinateur	8
1.2.3.1 Technologies de vision par ordinateur	8
1.2.3.2 Fonctionnement de la vision par ordinateur	10
1.2.3.3 Domain d'application	10
1.3 Les propriétés des panneaux de signalisation routière	11
1.4 Techniques de reconnaissance des panneaux de signalisation routière . .	14
1.4.1 Sélection colorimétrique	14
1.4.2 Caractéristiques géométriques	14
1.4.3 Combinaison des critères colorimétrique et géométriques	15
1.4.4 Méthodes à base apprentissage	16
1.4.4.1 Apprentissage automatique	16
1.4.4.2 Apprentissage en profondeur	17
1.4.4.3 les réseaux de neurones	18
1.5 Conclusion	18

Chapitre 2: Systèmes avancés d'aide à la conduite	19
2.1 Introduction	19
2.2 Définition	19
2.3 Types de capteurs utilisés	20
2.3.1 Le capteur de vision	20
2.3.2 Le capteur de RADAR	20
2.3.3 Le capteur de LIDAR	21
2.3.4 Le capteur ultrasonique	21
2.3.5 Autres capteurs	21
2.4 Les ADAS basés sur Le capteur de vision	22
2.4.1 Les types des caméras utilisées par l'ADAS basé sur la vision	22
2.4.1.1 Les caméras monoculaires	22
2.4.1.2 Les caméras stéréos	23
2.4.1.3 Les caméras thermiques (infrarouges)	23
2.4.2 Fonctionnement de l'ADAS basé sur la vision	24
2.4.3 Types d'ADAS basés sur la vision	26
2.5 Les avantages des ADAS	29
2.6 Limites des systèmes ADAS	29
2.7 Etude de l'existant	30
2.8 Conclusion	32
Chapitre 3: Réseau de neurones convolutionnel YOLO	33
3.1 Introduction	33
3.2 Définition du réseau de neurones YOLO	33
3.3 Architecture du réseau profond Yolov5	35
3.4 Fonctionnement de l'algorithme YOLOv5	36
3.5 La fonction Loss (Loss function)	42
3.6 Conclusion	43
Chapitre 4: Conception, Implementation et Discussion des resultats	45
4.1 Introduction	45
4.2 Environnements et outils de développement	45
4.2.1 Matériel utilisé	45
4.2.2 Langages de programmation	45

4.2.2.1	python	45
4.2.3	Environnements de développement	46
4.2.3.1	Visual Studio Code (VS Code)	46
4.2.3.2	Roboflow	47
4.2.3.3	Google colab	47
4.2.4	bibliothèque utilisé	47
4.3	Etude de conception	50
4.3.1	Présentation de l'architecture générale	50
4.3.2	Architecture détaillée	50
4.3.2.1	Conception de la base des images format yolov5s	50
4.3.2.2	Apprentissage du réseau yolov5	51
4.3.2.3	Test et validation	51
4.3.2.4	Sorties du système	53
4.4	Implémentation et discussion des résultats	54
4.4.1	Choix de la base d'images	54
4.4.2	Discussion des résultats obtenus	55
4.4.2.1	Apprentissage du réseau profond yolov5s	56
4.4.3	Détection des objets par le réseau yolov5s	57
4.4.4	Tests et résultats obtenus	58
4.5	conclusion	65
Chapitre 5: conclusion générale		66
5.1	Conclusion	66
5.2	Bilan des travaux réalisés et des résultats obtenus	66
5.3	Perspectives	68
Bibliographie		69

Table des figures

1.1	Représentation d'image numérique [49]	5
1.2	Exemples de types d'images numériques	6
1.3	Filtrage d'une image par un noyau de convolution [23]	7
1.4	La segmentation d'une image [23]	7
1.5	Différence entre Classification et Détection d'objet	9
1.6	Un exemple de suivi d'objets	10
1.7	Panneaux des obligations [30]	12
1.8	Panneaux d'avertissement de danger [30]	13
1.9	Panneaux d'interdiction ou de restriction [30]	13
1.10	Panneaux de priorité [30]	13
1.11	Emplacement de certains panneaux de signalisation	14
1.12	Stratégie générale des trois techniques [41]	15
1.13	Illustration montre l'architecture de Deep Learning [26]	17
2.1	Les types de capteurs utilisés d'un ADAS [32]	20
2.2	Exemples d'applications ADAS [1]	22
2.3	Standard camera on left ; integrated Active-IR on right [27]	23
2.4	Capture d'une image à partir d'une vidéo de caméra thermique passive [26]	24
2.5	exemple de détection de fatigue [26]	27
2.6	Exemples Régulateur de vitesse adaptatif [2]	27
2.7	exemple Le système de freinage d'urgence [3]	28
2.8	Détecteurs d'objets à une et deux étapes [4].	31
3.1	Architecture et composition de YOLO [26]	34

3.2	L'architecture. Notre réseau de détection comporte 24 couches convolutionnelles suivies de 2 couches entièrement connectées. Alternance 1×1 les couches convolutionnelles réduisent l'espace des caractéristiques des couches précédentes. Nous pré-entraînons les couches convolutives sur la classification ImageNet tâche à la moitié de la résolution (image d'entrée 224×224), puis doublez la résolution pour la détection [37].	34
3.3	Frise chronologique de l'algorithmes de détection d'objets yolo	35
3.4	Architecture réseau Yolov5	35
3.5	Le modèle YOLO avec une cellule de grille 7x7 a été appliqué sur l'image d'entrée [44]	36
3.6	Paramètres pour bounding box dans une cellule de grille 3x3 [44]	38
3.7	La spécification du vecteur d'étiquette Y dans un modèle YOLO a des cellules de grille 3×3 et prédit l'objet pour 3 Des classes [44]	39
3.8	Un exemple d'une voiture et d'un piéton au centre de la cellule	40
3.9	Caption	41
3.10	(a) L'intersection des boîtes englobantes. (b) L'union des boîtes englobantes [35]	42
4.1	Logo python	46
4.2	Logo Visual Studio Code	46
4.3	Logo Roboflow	47
4.4	Logo Google colab	47
4.5	Logo OpenCV	48
4.6	Logo Kivy	48
4.7	Logo Pillow	49
4.8	Logo PyTorch	49
4.9	Logo tensorflow	49
4.10	Représentation générale du système.	50
4.11	l'apprentissage d'un model.	51
4.12	ouvrir la camera	52
4.13	Processus de détection par le modèle yolov5s	52
4.14	Différentes classes de panneaux de signalisation en Urbain	55

4.15 Exemples des images de la base utilisée pour l'apprentissage pour la reconnaissance des obstacles	55
4.16 20 itérations	56
4.17 50 itérations	56
4.18 150 itérations	56
4.19 200 itérations	56
4.20 Matrice de confusion relatives aux panneaux de signalisation	57
4.21 Chargement du modèle	57
4.22 Détection des objets	57
4.23 Affectation du nom de l'objet	58
4.24 Taux.Recon=93%	58
4.25 Taux.Recon=94%	58
4.26 Taux.Recon=92%	59
4.27 Taux.Recon=50%	59
4.28 Taux.Recon=79%	59
4.29 Taux.Recon=96%	59
4.30 Taux de reconnaissance des plaques de signalisation	60
4.31 Taux de reconnaissance des plaques de signalisation	60
4.32 Taux.Recon=96%	61
4.33 Taux.Recon=50%	61
4.34 Taux.Recon=81%	61
4.35 Taux.Recon=83%	61
4.36 Taux.Recon=76%	62
4.37 Taux.Recon=83%	62
4.38 Taux.Recon=79%	62
4.39 Taux.Recon=76%	62
4.40 Taux de reconnaissance des plaques de signalisation	63
4.41 Détection de panneaux de signalisation et une voiture	63
4.42 Détection des obstacles	64
4.43 Détection des obstacles	64

Liste des tableaux

4.1	Description des deux bases d'images	54
4.2	Résultats de l'apprentissage du réseau yolov5s	56

Introduction générale

Le développement de la technologie au cours des vingt dernières années a simplifié l'accès aux systèmes numériques dans notre vie quotidienne. Parmi les éléments de ces systèmes, une grande importance est accordée à l'image. La présentation et la manipulation d'images numériques est actuellement un sujet de recherche actif. Le domaine du traitement d'images est un domaine très vaste, et il a connu un développement remarquable au cours des dernières décennies. La vision par ordinateur est l'une des branches les plus puissantes et les plus convaincantes de l'intelligence artificielle. Il se concentre sur la reproduction d'aspects de la complexité du système de vision humaine, permettant aux ordinateurs d'identifier et de manipuler des objets dans des images et des vidéos de la même manière que les humains. Dans le passé, la vision par ordinateur fonctionnait de manière limitée, mais grâce aux progrès de l'intelligence artificielle et aux innovations en matière d'apprentissage en profondeur, le domaine a fait de grands progrès ces dernières années et a pu dépasser les capacités humaines dans certaines tâches telles que la détection et la reconnaissance d'objets.

Motivation

Le transport routier est le mode de transport le plus utilisé générant ainsi un trafic important qui ne cesse d'augmenter. Malheureusement, cette augmentation du trafic routier a des conséquences sur la santé publique. Les accidents de la circulation font partie des événements les plus malheureux et les plus dangereux auxquels les gens peuvent être confrontés en conduisant. Des millions de personnes souffrent de blessures graves et d'invalidité. Cependant, d'après les experts en accidentologies, le facteur humain représente la première cause dans 90% des accidents. C'est pourquoi les industriels automobiles intègrent de plus en plus de systèmes d'aide à la conduite lors de la conception de nouveaux véhicules. Les systèmes avancés d'aide à la conduite (Advanced Driver Assistance System - ADAS) ont un grand intérêt dans la sécurité

routière, ils assistent le conducteur en contrôlant son mode de conduite, et en le tenant informé des situations potentiellement dangereuses, ou en changeant son comportement afin d'éviter une collision ou limiter les conséquences d'un accident. Parmi ces systèmes il y a le système de reconnaissance de panneaux routiers (TSR). Le système de reconnaissance de panneaux routiers (TSR) est un des systèmes avancés d'aide à la conduite. Ce système assiste le conducteur en le tenant informé de la signalisation présente sur la route parcourue. Un système TSR est usuellement basé sur une caméra embarquée sur l'avant du véhicule.

Buts poursuivis

Le travail présenté dans ce mémoire vise à développer un système basé sur des techniques de vision par ordinateur et de traitement d'images et qui permettra la détection, la localisation et la reconnaissance de signaux routiers par acquisition vidéo. Ainsi, il sera également capable de détecter les obstacles susceptibles de présenter un danger pour les conducteurs et les piétons.

Le système sera conçu de façon à améliorer la sécurité de la conduite et éviter les accidents de la circulation, en milieu urbain, en alertant constamment et en temps réel les conducteurs de la présence de panneaux de signalisation et de dangers tels que les personnes, les véhicules, les animaux, etc.

Les buts poursuivis peuvent être résumés par les points suivants :

- Analyse du flux vidéo fourni par la caméra, la détection et la reconnaissance des panneaux de signalisation ainsi que les dangers présents sur la route en milieu urbain.
- Exécution en temps réel : C'est-à-dire alerter le conducteur constamment de la présence des panneaux de signalisation et des dangers ou obstacles (être humain, animaux, véhicules).

À cette fin, nous avons fait appels aux réseaux de neurones convolutifs profond CNN car cette technique d'apprentissage profond largement utilisée dans les domaines de la classification d'images et de la détection d'objets. Nous nous sommes intéressés plus particulièrement au réseau dit yolov5s.

Afin de rendre notre système plus rapide et fiable en temps réel, nous avons introduit la notion de réseau spécialisé, c'est-à-dire, un réseau qui soit capable de reconnaître un type d'objets spécifique. Pour cela, nous avons combiné deux réseaux

yolov5s qui travaillent conjointement et en même temps, l'un reconnaît les panneaux de signalisation et le deuxième les dangers qui peuvent se présenter en milieu urbain.

Organisation du manuscrit

Ce mémoire est composé de quatre chapitres qui sont organisés comme suit : **Le premier chapitre** est une revue littérature sur le traitement d'images et la vision par ordinateur. Nous avons donné également une description des panneaux de signalisation ainsi que les techniques de reconnaissance répondant à notre problématique. Il s'agit tout d'abord de présenter certains des algorithmes qui y sont associés dans la littérature et qui répondent à tout ou partie des besoins allant des méthodes de détection par descripteur jusqu'à l'apprentissage profond. **Le second chapitre** est dédié à couvrir principalement deux parties. Dans la première, nous introduisons le concept d'un système d'aide à la conduite et les capteurs utilisés dans ce système, et nous discuterons en détail des systèmes d'aide à la conduite basés sur la vision, et identifierons les principaux avantages et défis qui affectent ces systèmes. Dans la seconde, nous introduisons une méthode capable d'identifier automatiquement les feux de circulation et les objets dangereux (tels que les piétons, les véhicules, les animaux, etc.), basée sur des techniques de vision par ordinateur et des architectures d'apprentissage en profondeur. Dans **Le troisième chapitre** sera consacré au réseau de neurone convolutionnel profond modèle yolov5s utilisé dans notre système pour la détection et le suivi des panneaux de signalisations et des dangers présents sur un milieu urbain. Nous décrirons ensuite dans **le quatrième chapitre**, la conception de notre système et nous discuterons les résultats obtenus par l'implémentation de ce système. **Enfin**, nous concluons ce mémoire par une synthèse générale et proposons des axes d'évolutions pour l'amélioration de ce travail.

Chapitre 1

Revue de littérature

1.1 Introduction

Au fil des décennies, le traitement d'images est devenu un domaine très vaste qui a connu un développement considérable. Il englobe toutes les techniques utilisées pour améliorer ou extraire des informations d'une image numérique. La disponibilité d'images et de vidéos sur des plateformes en ligne telles que Facebook, Instagram et Google a permis l'émergence de la vision par ordinateur, une technologie qui permet aux machines de comprendre et d'analyser le monde réel. Des systèmes capables de reconnaître automatiquement des formes et des objets, y compris les panneaux de signalisation, ont été développés. Les chercheurs ont commencé à utiliser des technologies de pointe telles que les réseaux de neurones artificiels et les techniques d'apprentissage en profondeur. Pour améliorer la reconnaissance automatisée des panneaux de signalisation.

Dans ce chapitre, nous présenterons un état d'art sur le traitement d'image et la vision par ordinateur. Nous survolons par la suite les approches utilisées pour rechercher et détecter les panneaux de signalisation dans des images de scènes routières.

1.2 Traitement d'image et vision par ordinateur

1.2.1 L'image numérique

L'image numérique est une image dont la surface est divisée en éléments de taille fixe appelés cellules ou pixel, ayant chacun comme caractéristique de couleurs ou un

niveau de gris. 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 numériques $f(x, y)$.

où : x, y : coordonnées cartésiennés d'un point de l'image, et $f(x, y)$: niveau d'intensité. La valeur à chaque point est une mesure de l'intensité lumineuse perçue par le capteur [39].

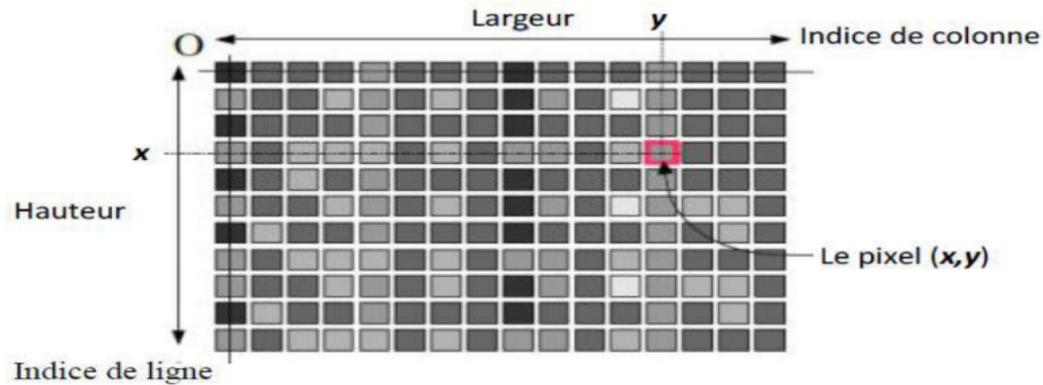


FIGURE 1.1 – Représentation d'image numérique [49]

Il existe trois types d'images numérique, à savoir :

Images binaires :

C'est une image composée de couleur noir et blanc seulement, et chaque pixel contient une valeur numérique 0 qui représente la couleur noire, ou la valeur numérique 1 qui représente la couleur blanche [26].

Image en niveaux de gris :

C'est une image composite de noir et blanc avec des nuances de gris. Chaque pixel contient des nombres de 0 à 255, qui représentent l'intensité de la couleur. La couleur est un blanc pur à la valeur 255 [26].

Images en couleur :

C'est une image qui contient des couleurs multiples et cohérentes dans sa composition, et exprime la réalité d'une manière plus précise et réaliste. Chaque pixel contient trois couleurs de base (rouge, vert et bleu) [26].



FIGURE 1.2 – Exemples de types d’images numériques

1.2.2 Traitement d’image

Le traitement d’images est l’ensemble techniques et méthodes opérant sur l’image, dont le but est d’améliorer son aspect visuel ou d’en extraire des informations jugées pertinentes que peut contenir les pixels. Il se définit comme un ensemble de tâches destinées à extraire de l’image des informations qualitatives et quantitatives [23]. C’est un type de traitement du signal dans lequel l’entrée est une image et la sortie une autre. Ces opérations comprennent :

- ★ **Numérisation** : Pour traiter une image sur un système informatique, il faut avant tout la soumettre à un processus de numérisation qui la rende lisible et manipulable par ce système. Ce processus est réalisé en convertissant les images traditionnelles (l’image originale) en formats numériques (dans l’unité de traitement) [23].
- ★ **Filtrage** : Ce procédé est utilisé pour supprimer le bruit afin d’améliorer la qualité de l’image ou de masquer certains détails. le filtrage est obtenu par convolution de l’image avec un noyau spécifique. Ce noyau peut être interprété comme une micro-image ou une vignette qui contient un modèle de transformation (linéaire ou non linéaire) et est appliqué à chaque pixel de l’image à filtrer pour

créer une nouvelle image [23].

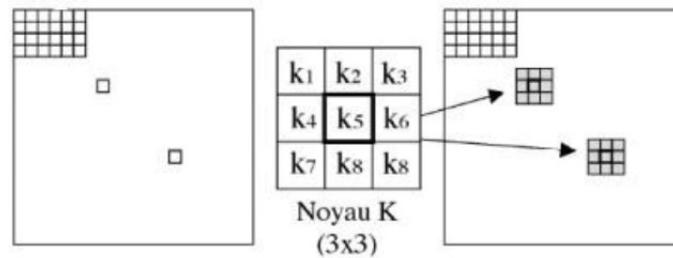


FIGURE 1.3 – Filtrage d’une image par un noyau de convolution [23]

- ★ **La segmentation** : Ce procédé consiste à diviser l’image étudiée en sous-groupes appelés régions. Une région est un ensemble connexe de pixels ayant des propriétés communes (intensité, texture) qui les différencient des pixels des régions voisines [23].

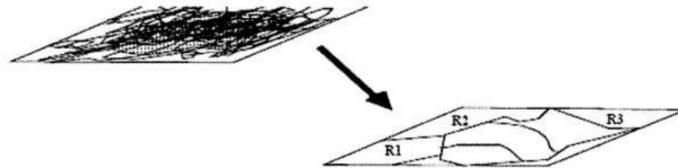


FIGURE 1.4 – La segmentation d’une image [23]

- ★ **Extraction des caractéristiques** : C’est un processus qui utilise des techniques de vision par ordinateur pour extraire des informations importantes d’une image, telles que la couleur, la forme et la texture.

Les fonctionnalités utilisées sont classées en propriétés générales et propriétés d’application. Ainsi, les propriétés générales incluent la correction des couleurs, les réglages de luminosité, de saturation et de contraste, la correction des couleurs, la suppression des contours et la reconnaissance des objets, tandis que les propriétés appliquées incluent Des applications telles que les visages humains, les empreintes digitales et Caractéristiques de conception [38].

1.2.2.1 Étapes de base du traitement d’image

Il se compose de six étapes consécutives [36] :

- ★ Acquisition de l'image par un capteur de lumière (comme un appareil photo, un capteur laser, etc.)
- ★ Traitement préliminaire de l'image, comme l'élimination du bruit (filtrage) ou la conversion en une image binaire.
- ★ Découpage de l'image pour séparer les informations importantes (par exemple, un objet dans l'image) du fond
- ★ Extraction de caractéristiques ou de propriétés
- ★ Classification des caractéristiques et leur association avec le modèle auquel elles appartiennent, ainsi que la reconnaissance des modèles
- ★ Compréhension de l'image

1.2.3 vision par ordinateur

La vision par ordinateur est une branche de l'intelligence artificielle dont l'objectif principal est de concevoir des applications et des programmes intelligents capables d'analyser, de comprendre et d'interpréter des images et des séquences vidéo afin de reconnaître et de distinguer des objets et divers détails de manière efficace et précise de la même manière que les humains [5].

1.2.3.1 Technologies de vision par ordinateur

Il existe trois fonctions principales de traitement d'image avec des techniques de vision par ordinateur, de sorte que les informations extraites des images peuvent être utilisées pour produire des décisions telles que la reconnaissance d'objet, la détection sur l'objet ou la détermination de certaines propriétés des objets dans l'image [6].

1. Classification d'objets :

C'est le processus de reconnaissance du contenu d'une image et de classification des objets qu'elle contient selon une catégorie spécifique. Par exemple, avec la classification d'objets, l'ordinateur peut distinguer les personnes des objets dans une image et déterminer combien de personnes apparaissent dans l'image [6] [34].

2. Détection d'objets :

C'est une technologie qui nous permet d'identifier le type et l'emplacement des objets sur une image ou une vidéo. Chaque objet est séparé par une boîte englo-

bante approximative. Cette technologie peut déterminer l'emplacement exacte des objets dans une scène donnée [34].



FIGURE 1.5 – Différence entre Classification et Détection d'objet

La détection d'objets et la classification d'objets sont parfois confondues, l'image ci-dessus montre la différence entre eux :

dans la classification d'objets, une image donnée est classée dans des catégories pré-programmées où l'image représente l'entrée du processus et la sortie représente le nom de l'élément dans l'image. Par exemple, si l'entrée est une image contenant un chat, la sortie sera le mot "chat". Si l'entrée est une image contenant deux chats, la sortie sera également "chat". De ce point de vue, le processus de classification des objets, malgré son importance, se limite aux résultats qu'il nous donne. Contrairement du processus de détection d'objet, chacun des éléments détectés est identifié dans une trame, en plus de la localisation précise de cet élément dans la scène. La détection d'objets est donc un processus plus complexe que la classification d'objets et donne plus d'informations sur l'image [28].

3. suivi d'objets :

C'est le processus de suivi d'un objet spécifique ou de plusieurs objets dans une scène spécifique. Ces applications traditionnelles sont soit le suivi de l'objet dans un clip vidéo, soit le suivi de l'objet en temps réel. Ceci est très important pour les systèmes de conduite autonome tels que les entreprises de véhicules autonomes [34].

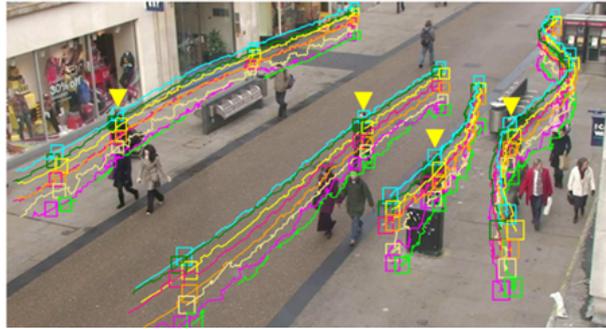


FIGURE 1.6 – Un exemple de suivi d'objets

1.2.3.2 Fonctionnement de la vision par ordinateur

Les applications de vision par ordinateur utilisent les entrées des appareils de détection, l'intelligence artificielle (IA), L'apprentissage automatique (ML) et L'apprentissage profond (DL) pour répliquer la manière dont fonctionne le système de vision humaine. Les applications de vision par ordinateur s'exécutent sur des algorithmes qui sont entraînés sur de grandes quantités de données visuelles ou d'images dans le cloud. Elles reconnaissent les modèles dans ces données visuelles et utilisent ces modèles pour déterminer le contenu d'autres images [6].

1.2.3.3 Domain d'application

L'utilisation des techniques de vision par ordinateur dans le domaine du traitement d'images comprend un certain nombre d'applications telles que : [6]

- **Réalité augmentée** : Les objets physiques sont détectés et suivis en temps réel avec la vision par ordinateur. Ces informations sont ensuite utilisées pour placer des objets virtuels de manière réaliste dans un environnement physique.
- **Reconnaissance faciale** : La vision par ordinateur peut être appliquée pour identifier les individus.
- **Extraction de texte** : La reconnaissance optique de caractères peut être utilisée pour booster la détectabilité du contenu pour les informations contenues dans de grandes quantités de texte et pour activer le traitement des documents pour les scénarios d'automatisation des processus robotisés.
- **Organisation du contenu** : La vision par ordinateur peut être utilisée pour identifier des personnes ou des objets dans des photos et les organiser en fonction de cette identification. Les applications de reconnaissance de photos comme

celles-ci sont couramment utilisées dans le stockage de photos et les réseaux sociaux.

- **Santé** : Les photos ou images capturées par d'autres appareils médicaux peuvent être analysées pour permettre aux médecins d'identifier les problèmes et d'effectuer des diagnostics plus rapidement et avec plus de précision.
- **Véhicules autonomes** : Les voitures autonomes utilisent l'identification et le suivi des objets en temps réel pour recueillir des informations sur ce qui se passe autour d'elles et se diriger en conséquence.
- **Agriculture** : Des images de cultures provenant de satellites, de drones ou d'avions peuvent être analysées pour surveiller les récoltes, détecter la levée des mauvaises herbes ou identifier la carence nutritive des cultures.
- **Analyse spatiale** : Le système identifie les personnes ou les objets, tels que les voitures, dans un espace et suit leur mouvement au sein de cet espace.
- **Fabrication** : La vision par ordinateur peut surveiller les machines de fabrication à des fins de maintenance. Elle peut également être utilisée pour surveiller la qualité et l'emballage des produits sur une ligne de production.
- **Sports** : La détection et le suivi des objets sont utilisés pour l'analyse des stratégies et du jeu.

1.3 Les propriétés des panneaux de signalisation routière

La signalisation routière est la signalisation traditionnelle installée sur la route et vise à assurer la sécurité des usagers de la route, en les informant des dangers et prescriptions liés à la circulation ainsi que des éléments utiles à la prise de décision. L'importance du rôle de la signalisation routière augmente avec une bonne conception et construction et le développement du trafic, où il facilite la circulation et réduit les causes d'accidents. En effet la signalisation routière a pour objectif de rendre plus sûre la circulation routière [24].

Les principales propriétés de la signalisation routière sont :

1. **Couleur** : Les panneaux de signalisation sont généralement rouges ou bleus et les idéogrammes sont souvent dessinés en noir sur fond blanc. La couleur est utilisée pour indiquer différents types de messages tels que des interdictions, des avertissements ou des obligations [41].
2. **forme** : Les panneaux de signalisation sont circulaires, rectangulaires, triangulaires ou octogonal. La forme est utilisée pour aider à identifier rapidement le type de message que le panneau porte [41].
3. **symbole** : Des illustrations ou des textes sont utilisés dans les signalisations routières pour transmettre des informations et attirer l'attention sur des mots-clés. Les symboles peuvent représenter des objets, des situations ou des actions spécifiques tels que des feux de circulation pour les véhicules ou les piétons, ou des flèches [30].

Grâce à ces caractéristiques, Les panneaux de signalisation sont classés en différents types :

- **Les panneaux obligatoires** : ont pour but d'obliger les usagers à respecter certaines règles et prescriptions au niveau du panneau ou devant celui-ci. Ces panneaux indiquent la direction que l'utilisateur doit suivre, également une vitesse minimale à respecter, voire un itinéraire obligatoire pour certains types d'usagers. Cette panneau sont de forme est ronde et de couleur est bleue [30]. Comme le montre la figure suivante :



FIGURE 1.7 – Panneaux des obligations [30]

- **Panneaux d'avertissement de danger** : ont pour but d'indiquer la nature du danger, leur but est de respecter les règles de comportement et d'éviter d'éventuelles collisions pendant la conduite. Ce panneau a la forme d'un triangle avec une bordure rouge [30]. Comme le montre la figure suivante :



FIGURE 1.8 – Panneaux d’avertissement de danger [30]

- **Panneaux d’interdiction ou de restriction** : Ces types des panneaux interdisent ou limitent certaines actions à ceux qui les trouvent devant dans la direction de leur marche et de l’endroit où ils se trouvent. Cette panneau sont de forme circulaires et ont un bord rouge [30]. Comme le montre la figure suivante :



FIGURE 1.9 – Panneaux d’interdiction ou de restriction [30]

- **Panneaux de priorité** : Ils sont destinées à informer les usagers de la route aux règles de priorité spéciales aux intersections ou aux passages étroits, en général ils n’ont pas de formulaire ou une couleur spécifique [30]. Comme montre la figure suivante :



FIGURE 1.10 – Panneaux de priorité [30]

4. **position** : Des panneaux sont stratégiquement placés sur les routes pour diriger les conducteurs. Certains sont érigés près des zones dangereuses, d’autres sont placés aux intersections, et d’autres sont utilisés pour indiquer les limites de vitesse ou les zones de stationnement.



FIGURE 1.11 – Emplacement de certains panneaux de signalisation

1.4 Techniques de reconnaissance des panneaux de signalisation routière

Il existe de nombreuses techniques utilisées pour identifier les panneaux de signalisation :

1.4.1 Sélection colorimétrique

Cette technique est utilisée pour détecter les panneaux de signalisation routière en détectant les régions rouges et bleues de l'image. Le principal problème est la lumière et l'âge différents des balises, ce qui rend la détection très difficile. Différents espaces colorimétriques sont utilisés. Par exemple, l'espace RVB standard, qui utilise les rapports G et B sur R pour classer les pixels. L'espace colorimétrique HSI (Hue, Saturation, Intensity) est sans doute le plus utilisé dans ce domaine. Des espaces plus complexes tels que LCH (Lightness, Chroma et Hue) sont également utilisés. Le pouvoir discriminant d'un critère de couleur dans la détection de peinture dépend du type de scène. Meilleur efficace dans les zones rurales et les autoroutes. Au contraire, en milieu urbain, il existe un grand nombre de choses qui ont les mêmes couleurs que les panneaux comme les publicités, les enseignes de magasins, etc. diminue l'utilité de ce critère [41].

1.4.2 Caractéristiques géométriques

Cette technique est la technique de reconnaissance de la forme et de la taille des panneaux de signalisation routière. Où il est basé sur la découverte des formes géométriques des panneaux (cercles, triangles, carrés) à l'intérieur des régions d'intérêt. Cela renforce la fiabilité de la détection et permet de trouver la catégorie du panneau,

ce qui facilite la validation et l'identification de son type [41].

Les algorithmes de détection basés sur la forme sont largement étudiés. Par exemple la transformation de Hough qui est appliquée pour détecter les signes triangulaires, rectangulaires et fléchés. Une autre méthode comme le vote à symétrie radiale utilise la caractéristique de l'apparence des panneaux de signalisation [30].

1.4.3 Combinaison des critères colorimétrique et géométriques

Il s'agit de la technologie de reconnaissance des panneaux de signalisation améliorés par des normes de couleur et de géométrie. Cette technologie cherche à appairer par corrélation d'images, chaque candidat avec le bon panneau parmi l'ensemble des panneaux de références de même catégorie [41]. **Par exemple**, les algorithmes de traitement d'images peuvent d'abord utiliser des normes de couleur pour déterminer la couleur dominante dans une zone de l'image, puis des normes de géométrie pour vérifier si la forme correspond au type de signalisation attendu ou non.

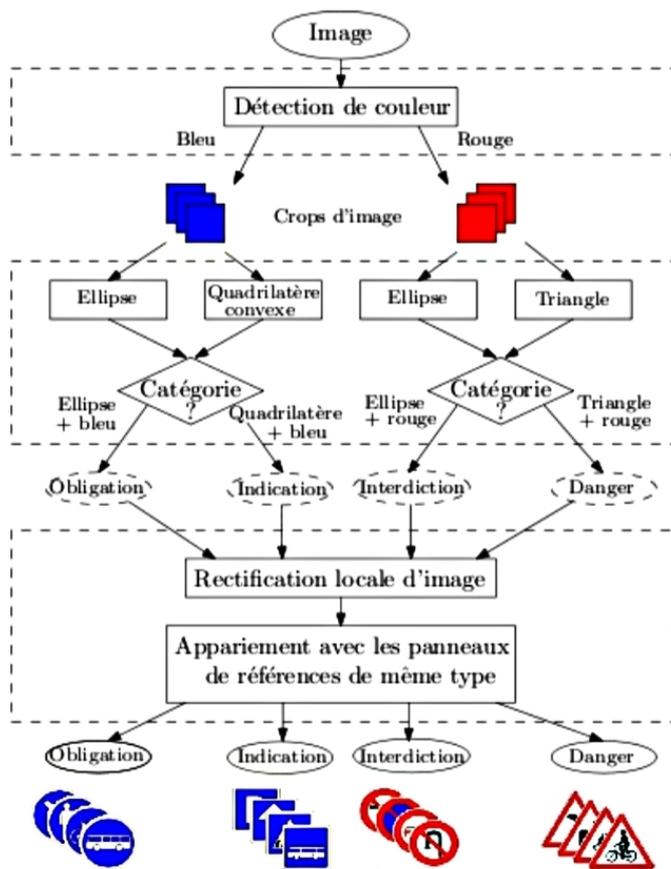


FIGURE 1.12 – Stratégie générale des trois techniques [41]

La figure ci-dessus montre la stratégie générale des trois techniques étudiées dans une séquence successive : La première étape consiste à détecter des pixels d'une couleur spécifique pour obtenir des régions d'intérêt dans l'image. La seconde étape consiste à détecter des formes géométriques connues de panneaux (cercles, triangles, carrés) dans les régions d'intérêts précédemment obtenues. Cette étape propose des candidats potentiels de panneaux. La catégorie de chaque candidat est déterminée en fonction de sa couleur et de sa forme. La dernière étape peut être vue comme une validation permettant de rejeter ou d'accepter ces candidats [41].

1.4.4 Méthodes à base apprentissage

Les méthodes calorimétriques et géométriques présentent une faiblesse commune à plusieurs facteurs tels que les changements d'éclairage, les occlusions, le changement d'échelle, la rotation et les translations. Cependant, ces problèmes pourraient également être traités en utilisant l'apprentissage automatique, mais cela nécessite une grande base de données annotées [48].

1.4.4.1 Apprentissage automatique

L'apprentissage automatique (en anglais Machine Learning), est un domaine d'étude au sein de l'intelligence artificielle (IA) qui s'appuie sur des approches statistiques pour permettre aux ordinateurs d'apprendre à partir de données. Cela signifie que les ordinateurs peuvent améliorer leurs performances dans la résolution de tâches sans être explicitement programmés pour chacune d'entre elles. Généralement, l'apprentissage automatique comporte deux phases. La première consiste à calculer un modèle à partir de données, appelé apprentissage. La deuxième phase est la phase de production, où le modèle est déterminé et de nouvelles données peuvent être soumises pour obtenir le résultat correspondant à la tâche souhaitée, telle que la prédiction. Dans le domaine de la classification d'images, les descripteurs caractéristiques servent de base de données au modèle pour apprendre à classer le contenu des images, comme reconnaître la présence d'un chat sur une photographie [26].

Une branche de l'apprentissage automatique :

Apprentissage supervisé : L'apprentissage supervisé est une méthode d'apprentissage automatique qui consiste à apprendre une fonction spécifique pour prédire des

valeurs futures sur la base d'un ensemble d'exemples pré-classés. En général, un modèle d'apprentissage supervisé est formé en lui fournissant un ensemble de données pré-étiquetées qui ont des entrées et des sorties correspondantes. Avec cette combinaison, le modèle apprend la relation entre les entrées et les sorties, puis ce modèle peut être utilisé pour prédire de nouvelles valeurs de sortie en fonction des nouvelles entrées [40].

Apprentissage non supervisé : L'apprentissage non supervisé a l'avantage de ne pas avoir de données pré-étiquetées ou d'informations d'entrée cibles pour l'algorithme. Au lieu de cela, le système s'appuie sur l'analyse de données non étiquetées et sur l'extraction du modèle ou de la structure sans indication explicite. Bien que les données étiquetées ne soient pas disponibles, l'apprentissage non supervisé peut être utile pour extraire des modèles et des informations précieuses à partir de données non étiquetées, telles que la classification d'images, la compilation de documents ou l'analyse de discours [40].

1.4.4.2 Apprentissage en profondeur

L'apprentissage en profondeur (En anglais Deep learning) est un ensemble de réseaux de neurones artificiels (ANN) qui effectuent des calculs complexes sur de grandes quantités de données.

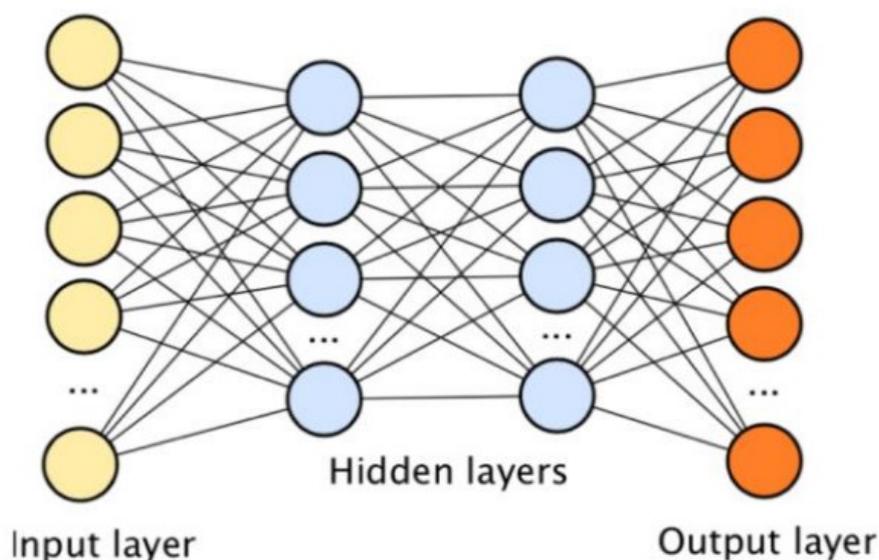


FIGURE 1.13 – Illustration montre l'architecture de Deep Learning [26]

De nombreux chercheurs le décrivent comme un sous-ensemble de l'apprentissage automatique qui, à son tour, est un sous-ensemble de l'intelligence artificielle (IA). Les modèles communs d'apprentissage en profondeur peuvent être utilisés dans n'importe quelle tâche de vision par ordinateur, ce qui inclut divers domaines tels que les prévisions météorologiques, la reconnaissance vocale, la détection d'objets et bien d'autres [25].

1.4.4.3 les réseaux de neurones

Les réseaux de neurones sont l'unité fonctionnelle de Deep Learning et sont connus pour imiter le comportement du cerveau humain pour résoudre des problèmes complexes axés sur les données. Les données d'entrée sont traitées à travers différentes couches de neurones artificiels empilés pour produire la sortie souhaitée [7].

Il existe différents type de réseaux de neurones : [26]

- Les réseaux de neurones profonds : Ces réseaux sont similaires aux réseaux RNA mais avec plus de couches cachées.
- Les réseaux de neurones récurrents.
- Les réseaux de neurones convolutifs.

1.5 Conclusion

Nous avons introduit dans ce chapitre les notions de base de la vision par ordinateur qui servent de fondement à la compréhension de différentes techniques de traitement d'images. Nous avons également présenté les méthodes basées sur l'apprentissage et reposant sur l'intelligence artificielle utilisées pour la détection des objets et qui sont largement utilisées dans la création de systèmes d'aide à la conduite. Ces derniers seront le sujet du chapitre suivant.

Chapitre 2

Systemes avancés d'aide à la conduite

2.1 Introduction

Chaque année, plusieurs centaines de personnes sont tuées en raison d'accidents de la route (Selon l'Organisation mondiale de la santé, plus de 1,2 million de personnes meurent et plus de 50 millions de personnes sont blessées chaque année dans des accidents de la route). Le facteur humain (sommolence, conduite sous influence, usage du téléphone...) est responsable de plus de %90 des accidents sur la route. C'est pourquoi, les constructeurs automobiles cherchent à développer des technologies permettant de rendre les trajets routiers plus surs, souples et confortables. C'est justement tout l'intérêt des ADAS «Advanced Driver Assistance Systems».

2.2 Définition

Un systèmes avancés d'aide à la conduite (ADAS) est un système embarqué intelligent qui reconnaît diverses conditions, telles que l'état du conducteur, l'état du véhicule et les conditions environnementales pour augmenter la capacité du conducteur à s'adapter aux dangers de la route pendant la conduite [42].

Il existe deux types de systèmes ADAS [8] :

- **Systeme ADAS passif** : l'ordinateur avertit le conducteur d'une condition dangereuse (des sirènes, des feux clignotants,etc) afin qu'il porter les meilleurs jugements sur la route. Dans ce cas, l'automobiliste doit prendre les mesures nécessaires pour éviter un accident. Le système ADAS passif fournit des données

en temps réel sur l'environnement de conduite et des alertes sur les dangers potentiels via l'interface utilisateur de l'appareil (IHM),

- **Système ADAS actif** : le véhicule fonctionne seul pour éviter les pires scénarios, par exemple : le système applique les freins sans l'assistance du conducteur lorsqu'il détecte un accident imminent. Les caractéristiques fonctionnelles comprennent le régulateur de vitesse adaptatif (ACC), le centrage de voie (LC), l'assistance au maintien de voie (LKA) et l'assistance aux embouteillages [45].

2.3 Types de capteurs utilisés

Il existe différentes classes de capteurs utilisées par les ADAS à savoir : Capteurs de vision, LIDAR, RADAR et ultrasons [26].

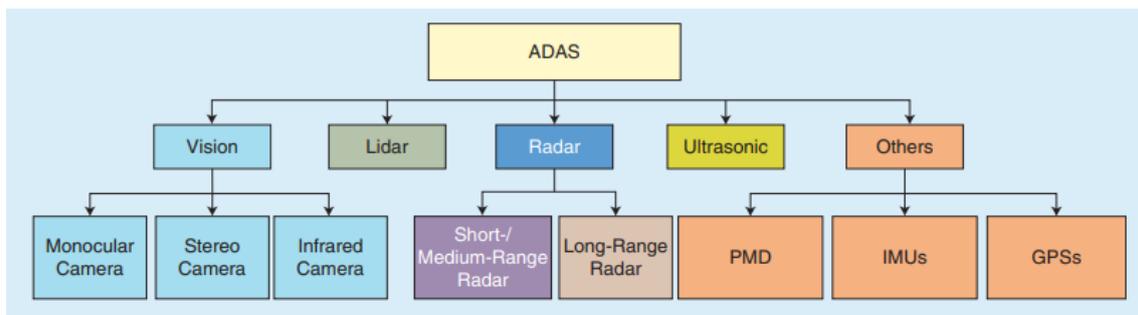


FIGURE 2.1 – Les types de capteurs utilisés d'un ADAS [32]

2.3.1 Le capteur de vision

Les caméras sont les capteurs de vision les plus largement utilisés dans les véhicules. ADAS basé sur la vision utilise une ou plusieurs caméras pour surveiller l'intérieur et l'extérieur d'un véhicule et détecter, analyser et suivre divers objets. Les avantages de l'utilisation d'une caméra sont son faible coût, sa facilité d'installation de plus qu'elle fournit au système de nombreuses informations précises sur l'environnement [32].

2.3.2 Le capteur de RADAR

Les capteurs de détection et de télémétrie radio (en anglais : Radio Detection and Ranging) sont utilisés pour détecter des objets et pour mesurer la vitesse et la distance en mesurant le changement de fréquence de l'onde réfléchi. De plus, il n'est pas

affecté par les conditions météorologiques (brouillard ou pluie) et est relativement peu coûteux. Parmi les applications des radars la détection des angles morts et les alertes de trafic transversal [32].

2.3.3 Le capteur de LIDAR

La détection et la télémétrie de la lumière (En anglais Light Detection And Ranging) deviennent très populaires dans les véhicules autonomes. Il fonctionne en tirant un faisceau laser sur un objet, puis en mesurant le temps nécessaire à la lumière pour rebondir sur le capteur et en calculant la distance à l'objet. LIDAR offre une réalisation d'images 3D haute résolution à 360 degrés avec des informations de profondeur très précises. Malgré les avantages mentionnés, ces types de capteurs sont coûteux, lourds et encombrants, et les conditions météorologiques telles que la pluie ou le brouillard affectent la couverture et la précision de ces systèmes [32].

2.3.4 Le capteur ultrasonique

Pour mesurer la distance à un objet, les capteurs à ultrasons utilisent des ondes sonores. Fondamentalement, ces types de capteurs sont utilisés pour détecter des objets très proches du véhicule. Le stationnement automatique est l'une de ses applications [32].

2.3.5 Autres capteurs

Il existe d'autres capteurs utilisés pour compléter et améliorer la fonctionnalité du système [32], comme GPS (Global Positioning Systems) qui est utilisé pour connaître l'emplacement du véhicule dans l'environnement, PMD (Photonic Mixer Device) est utilisé pour faire une détection optique rapide et une démodulation simultanée des signaux lumineux incohérents, et IMU (Inertial Measurement Units) qui est utilisé pour fournir une position fiable et un discernement de mouvement pour les applications d'ADAS de stabilisation et de navigation [26].

2.4 Les ADAS basés sur Le capteur de vision

S'appuyant sur la vision artificielle (caméra), les constructeurs automobiles cherchent à créer un ADAS plus sécurisés. Tesla a créé le Système Autopilot, qui est une application d'ADAS basé sur la vision offrant des fonctionnalités telles que la détection de voie sur la route, régulateur de vitesse adaptatif et le stationnement automatique. Toute fois, le conducteur est responsable de toutes ces fonctionnalités et le véhicule nécessite une surveillance constante [26].

2.4.1 Les types des caméras utilisées par l'ADAS basé sur la vision

Certains systèmes ADAS n'utilisent qu'une seule caméra dans le processus d'analyse et d'évaluation, tandis que d'autres utilisent une caméra couplée à un RADAR, un LIDAR et d'autres capteurs. La caméra peut être installée à l'intérieur ou à l'extérieur du véhicule en fonction de la conception et des exigences du système.

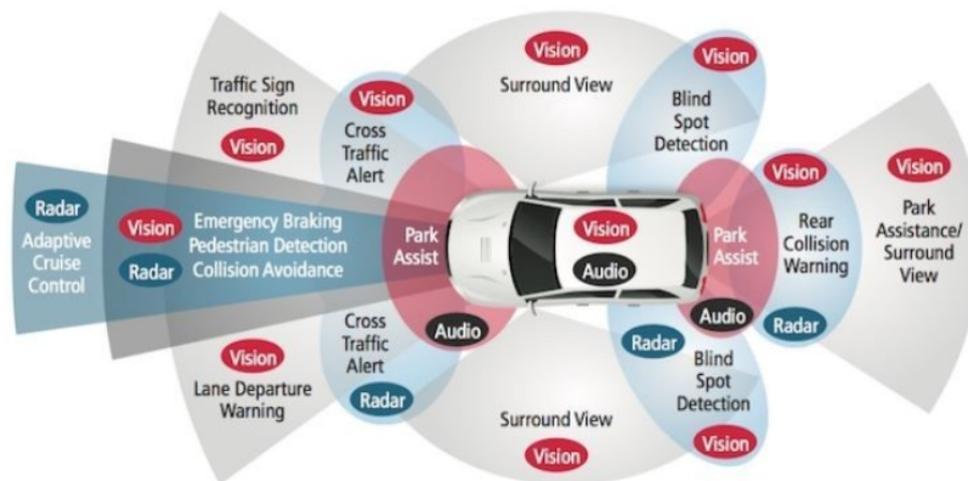


FIGURE 2.2 – Exemples d'applications ADAS [1]

Parmi les types de caméras utilisées par les ADAS basé sur la vision nous pouvons citer :

2.4.1.1 Les caméras monoculaires

Ce type de caméra se compose d'une seule lentille et il ne donne qu'une seule image à chaque instant. Ces types de caméras peuvent être utilisés dans diverses applications

telles que la détection des panneaux de signalisation, des piétons, des obstacles et surveiller même le conducteur à l'intérieur du véhicule (comme le système de détection de fatigue) [26].

2.4.1.2 Les caméras stéréos

Ce type de caméras se compose de deux ou plusieurs lentilles. Les caméras stéréos sont plus utiles pour extraire les informations en trois dimensions provenant de deux ou plusieurs images bidimensionnelles en appariant des paires stéréos (images des capteurs gauche et droit) et en utilisant une carte de disparité pour estimer la profondeur relative d'une scène. Ces caméras peuvent être utilisées pour une variété d'applications comme la reconnaissance des panneaux de signalisation, les piétons, les voies et l'estimation de la distance, avec beaucoup plus de précision par rapport aux caméras monoculaires [26].

2.4.1.3 Les caméras thermiques (infrarouges)

Il existe deux types de caméra thermique : une caméra thermique active et une caméra thermique passive.

Une caméra thermique active utilise une source lumineuse proche infrarouge qui peut être intégré directement dans les véhicules pour illuminer la scène (impossible de voir avec l'œil humain), et une caméra numérique standard pour capturer la lumière réfléchie. La figure 2.3 montre la différence entre les images dans la caméra normale et la caméra thermique active dans la nuit [26].



FIGURE 2.3 – Standard camera on left ; integrated Active-IR on right [27]

Une caméra thermique passive utilise des capteurs infrarouges où chaque pixel sur le capteur infrarouge peut être considéré comme un capteur de température qui peut capturer le rayonnement thermique émis par toute matière et il ne nécessite aucune

illumination spéciale de la scène. Ce type de caméra est le plus utilisé dans les applications d'ADAS par rapport aux caméras infrarouges active. [26].



FIGURE 2.4 – Capture d'une image à partir d'une vidéo de caméra thermique passive [26]

2.4.2 Fonctionnement de l'ADAS basé sur la vision

Grâce à l'utilisation de caméras qui prennent des photos ou des vidéos, le système prend des décisions sur l'environnement du véhicule. La qualité de l'image est d'abord améliorée, puis des algorithmes sont utilisés pour identifier l'environnement autour de la voiture (Détection des voies et des véhicules par exemple). Par la suite, les décisions appropriées et les actions nécessaires (telles que l'avertissement, le freinage d'urgence, les signaux sonores, etc.) sont prises, sur la base de la situation concernée.

Un système ADAS basé sur la vision suit les étapes suivantes :[31]

- **Acquisition d'image :**

Ce processus implique d'abord de capturer l'image et de la soumettre à un algorithme de traitement d'image qui peut gérer des images en niveaux de gris purs ou en couleur selon les exigences de mise en œuvre de l'algorithme.

- **Contrôleur :**

Avant le traitement, il est d'usage d'effectuer quelques vérifications sur l'image, vérification de l'état de l'image et de la fonction elle-même, et s'assurer qu'ils peuvent fournir des résultats utiles. Ces vérifications peuvent inclure (si des informations pertinentes sont disponibles sur le système) : La précision de l'étalonnage, des la contamination et la qualité de l'image elle même tels que le niveau de bruit et la luminance.

- **Pré-traitement :**

Il s'agit d'améliorer l'image obtenue, afin d'augmenter l'efficacité du post-traitement. Le prétraitement implique de nombreuses opérations différentes, notamment : l'extraction d'une région d'intérêt, le sous-échantillonnage d'une image, la génération d'une pyramide d'image (pour évaluer le niveau de traitement), le lissage d'image (comme le gaussien pour réduire le bruit de l'image) ou la détection des contours (comme Sobel pour accentuer les dégradés).

- **Segmentation :**

C'est le processus de division d'une image en régions similaires en localisant les limites d'objets, tels que le bord d'une route ou le ciel, pour simplifier la représentation de l'image, améliorer les performances de post-traitement et réduire le temps de traitement.

- **L'estimation de mouvement :**

Est un processus informatique crucial utilisé dans les applications ADAS et qui implique l'analyse du mouvement dans les images ou les vidéos afin de détecter les changements dans le mouvement des objets ou des personnes. Pour ce faire, on fait correspondre les caractéristiques d'images successives à l'aide d'algorithmes de suivi des mouvements. Il existe deux méthodes principales pour le suivi des mouvements : le suivi clairsemé et le suivi dense. Le suivi clairsemé est rapide et efficace, tandis que le suivi dense offre une plus grande couverture d'image et davantage d'informations statistiques. Ces deux méthodes sont plus robustes au bruit de l'image et aux transformations projectives, mais elles sont coûteuses en termes de calcul.

- **Estimation de la profondeur :**

Les algorithmes d'estimation de la profondeur visent à représenter la structure spatiale d'un environnement. Il existe deux types de techniques de perception de la profondeur pour les caméras : la stéréo et la mono. Les caméras stéréo utilisent une carte de contraste pour déterminer la distance entre les pixels et peuvent résoudre la position 3D de chaque pixel dans le monde en triangulant les rayons projetés entre les deux caméras. Les systèmes monoculaires détectent également la profondeur, mais nécessitent le mouvement de la caméra pour établir une base de référence pour la triangulation. Les méthodes stéréoscopiques ont une

meilleure résolution et nécessitent un traitement moins intensif, mais elles sont limitées à la zone de chevauchement des deux caméras et ont un coût de système plus élevé. Les techniques SFM utilisent un traitement minimal et peuvent être utilisées pour les caméras frontales sans miroir, mais les caméras frontales ont des mouvements longitudinaux compliqués par rapport à la scène.

- **Suivi d'état :**

Le suivi d'état est utilisé à de nombreux endroits différents tout au long du pipeline de traitement d'image. Cette technologie améliore la robustesse du système et évite les détections manquées ou perdues et le bruit, ainsi que pour gérer les erreurs de calcul et prédire l'état dans l'intervalle suivant. Le traçage d'état peut également être utilisé pour lisser temporairement la sortie de certaines parties du fil de traitement et améliorer les performances globales du système.

- **Détection et classification d'objets :**

La détection et la classification d'objets impliquent deux types de classification : supervisée et non supervisée. La classification supervisée consiste à générer un modèle de prédiction à l'aide de techniques d'extraction de caractéristiques et d'échantillons classés manuellement, tandis que la classification non supervisée consiste à trouver une structure dans les données sans connaissances préalables. Les techniques courantes de classification supervisée comprennent Soutenir la machine vectorielle (SVM) et les réseaux neuronaux convolutifs (CNN), tandis que le regroupement K-means et ISODATA sont couramment utilisés pour la classification non supervisée. La classification non supervisée est souvent utilisée dans l'agrégation du trafic, tandis que la classification supervisée est utilisée dans des applications telles que la détection des piétons, des visages et des véhicules, et la reconnaissance des panneaux de signalisation.

2.4.3 Types d'ADAS basés sur la vision

Récemment, dans la recherche automobile et dans le domaine des transports intelligents, plusieurs projets intéressants ont été menés afin de diminuer le nombre d'accidents. Lors du développement de ces projets, de nouveaux types de systèmes d'aide à la conduite ont été proposés :

- A. **La détection de fatigue** : Le détecteur de fatigue est une nouvelle technologie utilisée dans les voitures qui scanne le visage du conducteur pour détecter les signes de somnolence et de fatigue, notamment les yeux fermés, les hochements de tête et la fatigue extrême. Le système avertit le conducteur par des avertissements, des vibrations et d'autres alertes afin d'éviter les accidents. En outre, le système peut suivre les mouvements du conducteur et l'inviter à faire une pause s'il présente des réactions inhabituelles ou lentes [9].

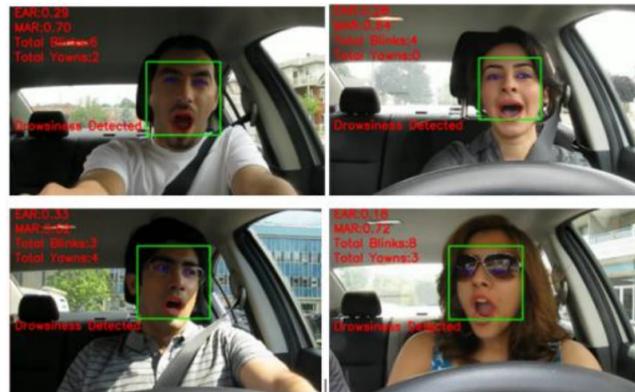


FIGURE 2.5 – exemple de détection de fatigue [26]

- B. **Régulateur de vitesse adaptatif** : Le régulateur de vitesse adaptatif (ACC) est une composante de ce système particulièrement utile sur les autoroutes, où les conducteurs ont de la difficulté à surveiller leur vitesse et celle des autres véhicules pendant de longues périodes. À l'aide du système de régulateur de vitesse avancé, la voiture peut accélérer, ralentir et parfois s'arrêter, au besoin et en fonction d'autres objets en mouvement à proximité [10].



FIGURE 2.6 – Exemples Régulateur de vitesse adaptatif [2]

C. **Le système de freinage d'urgence** : Le système de freinage d'urgence (AEB) fait partie du système ADAS qui utilise des caméras de vision et des capteurs pour détecter les dangers. Le système reconnaît les objets qui passent devant le véhicule et évalue s'ils présentent ou non un danger, comme la présence d'autres véhicules, de personnes ou d'obstacles. Dans ce cas, le système alerte le conducteur par un son, une vibration ou un feu rouge et active automatiquement le système de freinage. Cela permet de réduire la gravité de la collision et de limiter les dommages et les blessures causés par les accidents [11].

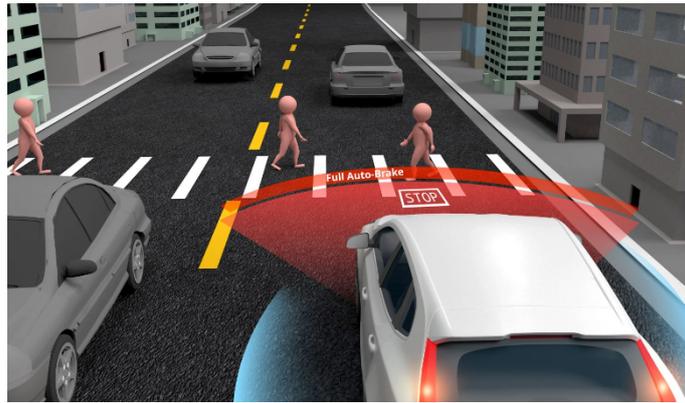


FIGURE 2.7 – exemple Le système de freinage d'urgence [3]

D. **Avertissement de sortie de voie** : Avertissement de sortie de voie (LDW) utilise des caméras avancées pour détecter les sorties de voie involontaires et avertit le conducteur par divers signaux, tels que des vibrations ou des sons. Certaines voitures corrigent même automatiquement la trajectoire. Ce système est très utile pour améliorer la sécurité routière en réduisant le nombre d'accidents causés par des changements de voie involontaires [12].

E. **La reconnaissance des panneaux de signalisation** : Le système de reconnaissance des panneaux routiers est l'un des composants les plus importants des ADAS basés sur la vision, et il aide à conduire en toute sécurité sur la route. Ce système identifie les panneaux sur la route, y compris les panneaux permanents et temporaires (tels que les limites de vitesse et les restrictions de dépassement), et envoie ces informations directement au tableau de bord de la voiture pour être affichées devant le conducteur. Le système de reconnaissance des panneaux routiers s'appuie sur la vision par ordinateur et les technologies d'apprentissage automatique pour lui permettre de reconnaître et d'interpréter

avec précision les panneaux. Cette fonctionnalité aide les conducteurs à conduire de manière plus sûre et à réduire le risque d'infractions au code de la route ou d'accidents causés par le non-respect des panneaux de signalisation [13].

2.5 Les avantages des ADAS

Le système ADAS comprend de nombreux avantages notamment : [11]

- Il améliore la sécurité des conducteurs et des piétons et aide à éviter les collisions et les accidents mortels.
- Avertit le conducteur des dangers potentiels sur la route en évaluant le comportement des autres véhicules et l'attention du conducteur.
- Contribue à une sécurité accrue sur la route grâce à des fonctions de navigation modernes, qui comprennent des fonctions d'éclairage automatique, de régulateur de vitesse, d'angle mort et de sortie de voie.
- Le système ADAS contribue à rendre la conduite plus confortable, facilite le stationnement automatique et permet une meilleure reconnaissance des panneaux de signalisation.

2.6 Limites des systèmes ADAS

Malgré les avantages des ADAS, ces derniers présentent des limites à savoir : [11]

- Les conducteurs peuvent devenir trop dépendants de ces systèmes et se laisser distraire de leur conduite.
- Une formation et une sensibilisation insuffisantes aux fonctions avancées des systèmes d'aide à la conduite font qu'il est difficile pour les conducteurs de tirer pleinement parti de leurs avantages.
- le coût élevé de l'installation des systèmes d'aide à la conduite peut dissuader les conducteurs de les choisir.
- Les systèmes d'aide à la conduite se trompent parfois sur les véhicules en pente en supposant qu'ils se dirigent directement.

2.7 Etude de l'existant

Les algorithmes de détection des objets dédiés ADAS peuvent être classés en deux classes à savoir : les algorithmes traditionnels et les algorithmes à bases de méthodes modernes. Les algorithmes traditionnels de détection d'objets reposent sur l'utilisation de fonctionnalités sélectionnées manuellement pour distinguer et identifier les objets de l'arrière-plan. Alors que les méthodes modernes qui utilisent des réseaux de neurones artificiels profonds dépendent de la formation de modèles de réseau sur de grands échantillons d'images et de leur nommage, y compris des objets, et des réseaux d'apprentissage pour distinguer différents modèles et prédire l'emplacement et le type d'objets [23].

En 2001, Le détecteur d'objets Viola-Jones [50] à été introduit pour la détection des visages, utilise diverses méthodes, notamment les caractéristiques de type Haar, les images intégrales, Adaboost et les classificateurs en cascade. Il commence par détecter les caractéristiques de type Haar dans une image à l'aide d'une image intégrale, puis utilise un algorithme Adaboost entraîné pour trouver le classificateur pour chaque caractéristique Haar et les mettre en cascade. Cet algorithme est encore largement utilisé dans les petits appareils en raison de sa précision et de sa rapidité.

En 2005, Dalal et Triggs [50] ont présenté le HOG, un descripteur de caractéristiques qui utilise le gradient et l'orientation des bords pour créer une table de caractéristiques pour la détection d'objets. Le HOG génère des histogrammes pour chaque cellule d'une grille appliquée à une image. Il extrait les caractéristiques d'une région d'intérêt et les introduit dans un classificateur SVM linéaire pour la détection. Le HOG a été développé à l'origine pour la détection des piétons, mais il peut être entraîné à reconnaître diverses autres classes.

Le modèle des pièces déformables (DPM) est un algorithme de détection développé par Felzenszwalb et al.[50] Il utilise des pièces individuelles d'un objet pour la détection, avec une plus grande précision que le HOG. Le modèle sépare un objet en plusieurs parties pendant le temps d'inférence et identifie une disposition probable de ces parties pour produire une détection. Par exemple, un corps humain est décomposé en plusieurs parties, telles que la tête, les bras, les jambes et le torse, chacune étant associée à un modèle de détection. Le modèle élimine les modèles improbables de combinaisons de pièces pour produire une détection. Le DPM était un algorithme très performant avant

l'arrivée des modèles d'apprentissage profond.

Afin d'augmenter d'avantage la performance des systèmes de reconnaissance d'objets dédiés aux ADAS, les chercheurs se sont orientés vers les nouveaux algorithmes qui sont plus performants, comme les algorithmes CNN.

Les recherches menées sur ce contexte ont essayé de développer ce type de réseau d'avantage afin qu'il soit adapté au contexte de la conduite. Dans les algorithmes appelés "Détecteurs à un étage", le processus d'extraction du retour sur investissement est supprimé et les boîtes de fixation candidates sont directement classées.

[4]. Et dans l'algorithme appelés "détecteurs à deux étapes" par contre, la tâche de détection d'objets est divisée en deux phases, où la première phase consiste à extraire les régions d'intérêt (RoI) de l'image, et la deuxième phase consiste à classer les RoI et à redéfinir les boîtes englobantes des objets au sein de celles-ci (Figure 2.8) [4].

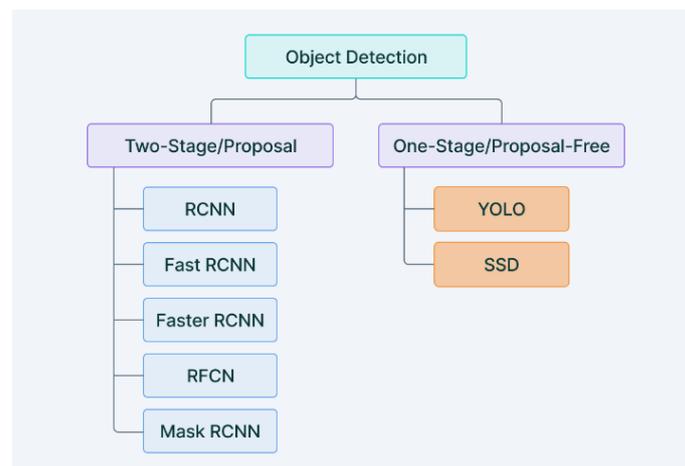


FIGURE 2.8 – Détecteurs d'objets à une et deux étapes [4].

Les boîtes de délimitation dans le réseau R-CNN sont proposées par l'algorithme de "recherche sélective", chacune d'entre elles est étirée et les caractéristiques sont extraites via un réseau neuronal convolutionnel profond, tel qu'AlexNet, avant qu'un ensemble final de classifications d'objets ne soit réalisé avec des SVM linéaires [14].

La conception Fast R-CNN est simplifiée, les boîtes de délimitation sont toujours fournies en entrée dans un modèle unique, mais après le CNN profond, une couche de mise en commun des régions d'intérêt est utilisée pour combiner les régions, et le modèle prédit directement les étiquettes de classes ainsi que les zones d'intérêt. [14].

Le Faster R-CNN, ajout un réseau de proposition de régions, qui apprend à proposer directement des régions d'intérêt sur la base de l'interprétation des caractéristiques du

CNN [14].

L'extension Mask R-CNN du Faster R-CNN, inclut un modèle de sortie prévoyant un masque pour chaque objet détecté [14]. Le réseau de neurone convolutionnels CNN profond yolo (acronyme de "You Only Look Once"), vise à extraire des caractéristiques pour détecter et reconnaître des objets dans l'image en temps réel et a largement surpassé d'autres algorithmes de détection d'objets, en effectuant toutes les prédictions à l'aide d'une seule couche entièrement connectée, ce qui permet d'obtenir des résultats précis plus rapidement et plus efficacement [15][16].

Nous allons donner plus de détails sur ce réseau dans le chapitre suivant.

2.8 Conclusion

Dans ce chapitre, nous avons présenté le concept de base des systèmes avancés d'aide à la conduite. Nous avons également appris à connaître les types d'algorithmes de détection d'objets spécifiques aux ADAS, y compris les algorithmes traditionnels et les algorithmes basés sur des méthodes modernes.

Dans le prochain chapitre, nous décrirons l'un des algorithmes modernes proposés.

Chapitre 3

Réseau de neurones convolutionnel

YOLO

3.1 Introduction

Après avoir étudié les systèmes d'aide à la conduite et les méthodes utilisées pour la reconnaissance et la détection d'objets dans le chapitre précédent, ce chapitre se concentre sur l'algorithme yolov5 et utilisé pour construire notre système.

3.2 Définition du réseau de neurones YOLO

Le réseau de neurones convolutionnels CNN profond yolo (acronyme de "You Only Look Once"), vise à extraire des caractéristiques pour détecter et reconnaître des objets dans l'image en temps réel (figure 3.1) et a largement surpassé d'autres algorithmes de détection d'objets, en effectuant toutes les prédictions à l'aide d'une seule couche entièrement connectée, ce qui permet d'obtenir des résultats précis plus rapidement et plus efficacement [15][16].

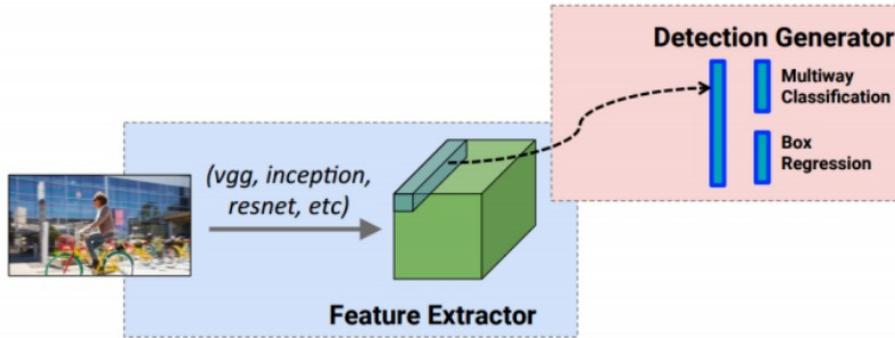


FIGURE 3.1 – Architecture et composition de YOLO [26]

YOLO détecte les objets figurant dans l’image en une seule propagation vers l’avant (figure 3.2). Cela signifie que la prédiction de l’image entière est effectuée en une seule exécution de l’algorithme [15].

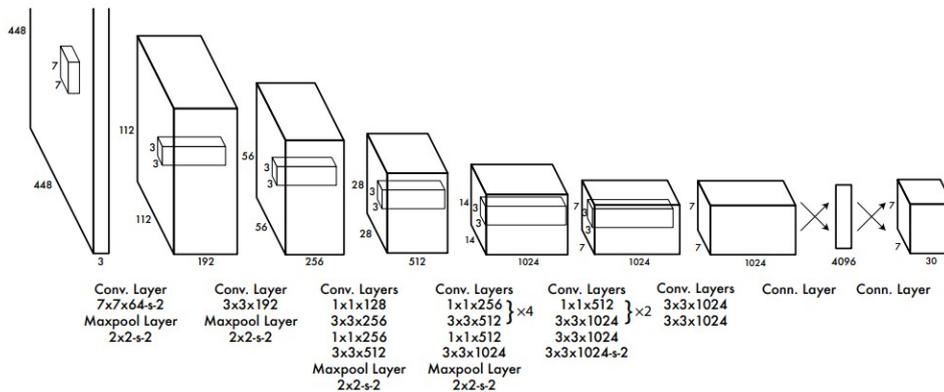


FIGURE 3.2 – L’architecture. Notre réseau de détection comporte 24 couches convolutionnelles suivies de 2 couches entièrement connectées. Alternance 1×1 les couches convolutionnelles réduisent l’espace des caractéristiques des couches précédentes. Nous pré-entraînons les couches convolutives sur la classification ImageNet tâche à la moitié de la résolution (image d’entrée 224×224), puis doublez la résolution pour la détection [37].

Yolo est le résultat de longues années de recherches comme le montre la frise chronologique de l’algorithme de détection d’objets :

Dans ce qui suit, nous allons nous concentrer sur la cinquième version de l’algorithme : YOLOv5

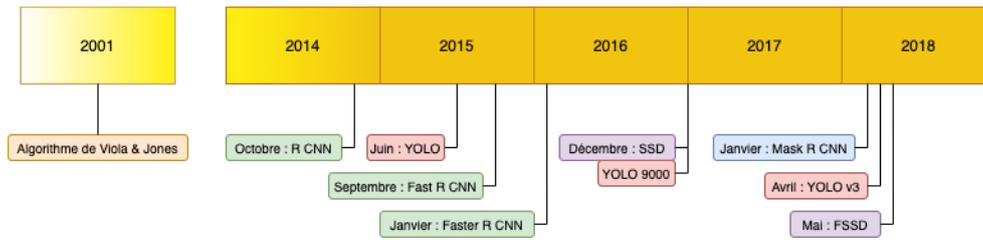


FIGURE 3.3 – Frise chronologique de l’algorithmes de détection d’objets yolo

3.3 Architecture du réseau profond Yolov5

Le réseau profond yolov5 est composé de trois couches à savoir : backbone, neck et prediction [47], comme indiqué dans la figure ci-dessous :

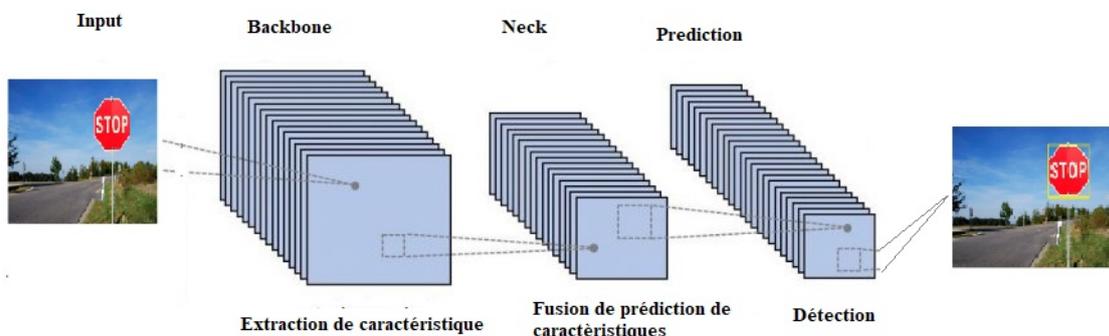


FIGURE 3.4 – Architecture réseau Yolov5

1. La couche backbone (extraction de caractéristiques)

L’image originale est segmentée en une petite section, puis la carte des caractéristiques de 32 noyaux convolutifs est formée à l’aide du processus convolutif. L’opération de mise au point peut réduire les dimensions d’entrée sans utiliser de paramètres et préserver autant que possible les informations d’image d’origine [51][33].

2. La couche Neck (fusion de prédiction de caractéristiques)

Neck est une couche réseau qui combine les caractéristiques de l’image et les transmet à la couche de prédiction [51]. Le problème avec la couche de fusion d’entités est que la carte d’entités profondes contient des informations sémantiques plus fortes et des informations de position plus faibles, tandis que la carte d’entités peu profonde contient des informations de position plus fortes et des informations sémantiques plus faibles. Pour résoudre cette contradiction, YOLOv5

utilise la structure de FPN (Pyramid of Feature Networks) et PAN (Perceptual Adversary Networks) [33], FPN collecte, connecte et intègre des informations de haut niveau sur les caractéristiques de manière descendante pour obtenir une carte des caractéristiques pour la prédiction. PAN est la pyramide de base, qui transmet de fortes propriétés de positionnement de manière ascendante [51].

3. La couche prediction (Détection)

Dans la couche de détection, trois feature maps sont produites à différentes échelles, à savoir 7676, 3838 et 1919, et ce, pour détecter des objets petits, moyens et grands. Le réseau contient trois bounding box pour prédire la catégorie, et chaque bounding box contient des informations de confiance d'objet et de position de bounding box. Enfin, bounding box en excès est éliminé grâce à la technologie NMS (suppression non maximale) [33].

3.4 Fonctionnement de l'algorithme YOLOv5

Les principales étapes pour faire fonctionner l'algorithme YOLOv5 :

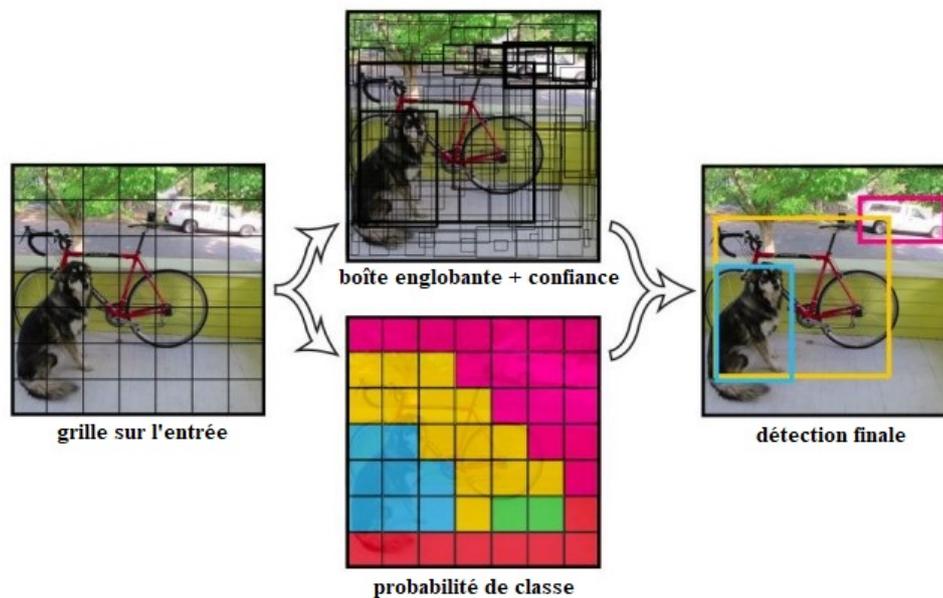


FIGURE 3.5 – Le modèle YOLO avec une cellule de grille 7x7 a été appliqué sur l'image d'entrée [44]

a) Cellules de la grille

Le réseau de neurones yolov5 vise à unifier les composants discrets de la détection d'objets. Il s'appuie sur les caractéristiques de l'image entière pour prédire la boîte englobante et prédit simultanément toutes les boîtes englobantes et les différentes classes de l'image, permettant une gestion globale de l'image entière et tous les objets qu'il contient. YOLOv5 est conçu pour un entraînement complet et pour atteindre des vitesses en temps réel tout en maintenant une précision moyenne élevée. Le système divise l'image entrante en une grille de taille $SS(77\text{pardfaut})$, où chaque cellule de la grille est responsable de la détection de l'objet s'il trouve son centre dans les limites de la cellule [37].

Pour chaque cellule de la grille, notre modèle prédit B boîtes englobantes ainsi que leurs scores de confiance associés (Comme le montre la Figure 3.5). Ces scores indiquent à la fois la confiance du modèle dans la présence d'un objet dans la boîte et la précision de sa boîte englobante prédite. Nous définissons la confiance comme : [37]

$$\text{confidencescore} = Pr(\text{Object}) * IOU_{pred}^{truth} \quad (3.1)$$

$Pr(\text{Object})$ est la probabilité qu'il y ait un objet à l'intérieur de la cellule.

IOU_{pred}^{truth} est l'intersection sur l'union de la boîte de prédiction et de la boîte de vérité terrain.

En l'absence d'un objet dans une cellule donnée, les scores de confiance doivent être nuls, alors qu'en présence d'un objet, le score de confiance doit refléter à quel point la boîte prédite s'aligne sur la vérité terrain [37].

b) Régression de boîte englobante (Bounding box)

Bounding box est un rectangle qui entoure un objet et qui spécifie sa position, sa classe (par exemple : voiture, personne) et son degré de confiance (probabilité qu'il se trouve à cet endroit). Les boîtes de délimitation sont principalement utilisées dans la tâche de détection d'objets, où l'objectif est d'identifier la position et le type de plusieurs objets dans l'image [17].

Chaque Bounding box se compose de 5 prédictions : x, y, w, h, p_c où : [37]

- (x, y) : Centre de la Bounding box, la plage de ces deux valeurs doit toujours

l'étiquette de vecteur y et le vecteur prédit \hat{y} , où la cellule violette ne contient aucun objet et le score de confiance des boîtes englobantes dans cette cellule est égal à 0. Par conséquent, tous les autres paramètres de cette cellule seront ignorés [44].

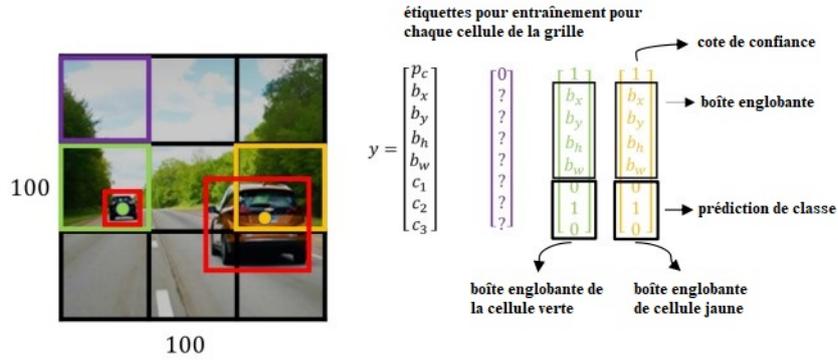


FIGURE 3.7 – La spécification du vecteur d'étiquette Y dans un modèle YOLO a des cellules de grille 3×3 et prédit l'objet pour 3 Des classes [44]

c) **Anchor boxes [18]**

Malgré l'efficacité de l'algorithme yolov5 dans la détection d'objets, il a toujours une limite. Si plusieurs objets sont présents dans la même cellule de grille, l'algorithme a du mal à les classer. Par exemple, si un piéton et une voiture sont présents dans la même cellule, il devient difficile de déterminer s'il s'agit d'un piéton ou d'une voiture. Cependant, nous avons une solution à ce problème appelée anchor boxes.

anchor boxes introduit une dimension supplémentaire dans les étiquettes de sortie, permettant à l'algorithme yolo de détecter plusieurs objets dans une seule cellule de grille. Cette dimension supplémentaire représente le nombre de anchor boxes attribuées à chaque cellule de la grille. En définissant le nombre de anchor boxes par cellule, nous pouvons maintenant attribuer un seul objet à chaque anchor boxes, surmontant ainsi le problème de classification.

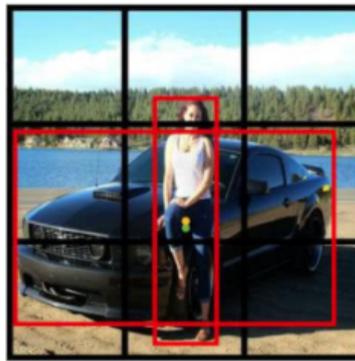


FIGURE 3.8 – Un exemple d'une voiture et d'un piéton au centre de la cellule

Pour augmenter la précision de la détection d'objets, il suffit d'ajuster le nombre de anchor boxes à 2. Cela permettra à chaque cellule de la grille d'identifier un piéton et une voiture séparément. Lorsqu'un piéton et une voiture sont présents dans la même cellule, chaque objet sera affecté à sa anchor boxes correspondante en fonction de la valeur d'intersection sur l'union (IoU) la plus élevée entre l'objet et anchor boxes. En utilisant des anchor boxes spécialisées conçues pour détecter certaines formes, nous pouvons améliorer la précision de notre modèle.

Exemple : Une anchor boxes peut être formée pour identifier des objets fins tandis qu'une autre peut être formée pour détecter des objets plus larges. Dans notre cas spécifique, une anchor boxes peut être conçue pour apprendre les ca-

ractéristiques d'objets hauts et maigres comme des piétons, tandis qu'une autre anchor boxes peut apprendre à détecter des objets courts et larges comme des voitures. En conséquence, notre nouvelle étiquette contiendra désormais des informations sur les piétons et les voitures dans la même cellule de grille.

$$y = \begin{bmatrix} p_c \\ b_x \\ b_y \\ b_h \\ b_w \\ c_1 \\ c_2 \\ c_3 \\ p_c \\ b_x \\ b_y \\ b_h \\ b_w \\ c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 1 \\ b_x \\ b_y \\ b_h \\ b_w \\ 1 \\ 0 \\ 0 \\ 1 \\ b_x \\ b_y \\ b_h \\ b_w \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

Anchor box 1
 Pedestrian
 Anchor box 2
 Car

FIGURE 3.9 – Caption

d) **Intersection over union (IoU)**

$$IoU = \frac{Areaof(BoxABoxB)}{Areaof(BoxA \cup BoxB)} \quad (3.2)$$

L'équation (3.2) (IoU) est utilisée pour calculer le score de confiance d'une boîte prédite par rapport à une vérité terrain. Si aucune boîte n'existe dans une cellule de la grille, le score de confiance est de zéro. Sinon, le score de confiance est déterminé par le rapport entre l'intersection et l'union entre la boîte prédite et la vérité terrain. Les cases de vérité terrain sont préalablement définies manuellement par l'utilisateur. Ainsi, un IoU plus élevé indique un score de confiance plus élevé, ce qui traduit une plus grande précision de prédiction de l'algorithme. Les boîtes ne contenant pas d'objet sont filtrées en fonction de la probabilité d'objets présents dans la boîte [35].

e) **Non-Maximum Suppression (NMS)**

YOLOv5 utilise la technologie de suppression non maximale (NMS) pour nettoyer toutes les boîtes englobantes qui ne contiennent aucun objet ou qui contiennent le même objet dans d'autres boîtes englobantes. Lorsqu'une valeur de seuil est spécifiée (0,5 par défaut), toutes les boîtes englobantes qui se chevauchent lors-

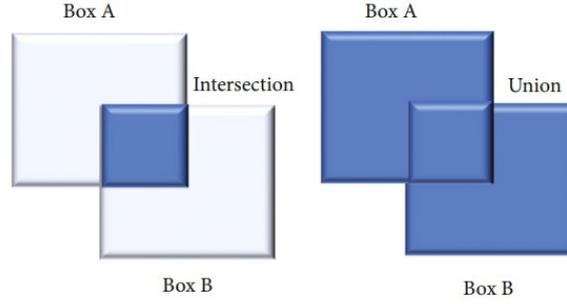


FIGURE 3.10 – (a) L’intersection des boîtes englobantes. (b) L’union des boîtes englobantes [35]

qu’une valeur d’intersection sur l’union (IOU) est supérieure à la valeur de seuil sont supprimées [44].

Dans la phase de détection finale, l’algorithme détecte les probabilités d’objet et de classe avec des scores de confiance.

3.5 La fonction Loss (Loss function)

L’erreur de somme au carré est l’épine dorsale de la fonction de perte de YOLO. Il existe plusieurs cellules de grille qui ne contiennent aucun objet dont le score de confiance est égal à zéro. Ils submergent les dégradés des cellules qui contiennent les objets. Pour éviter une telle surcharge entraînant une divergence d’apprentissage et une instabilité du modèle, YOLO applique la pénalité la plus élevée pour les prédictions à partir de boîtes englobantes contenant des objets ($\lambda_{coord} = 5$) et la plus faible pour les prédictions lorsqu’aucun objet n’est présent ($\lambda_{coord} = 0.5$). La fonction de perte de YOLO est calculée en prenant la somme de la fonction de perte de tous les paramètres de la boîte englobante, y compris $(x, y, w, h, confidencescore, classprobability)$ [44].

$$L = L1 + L2 + L3 + L4 \quad (3.3)$$

Tel que :

$$L1 = \lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^B \ell_{ij}^{obj} [(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2] \quad (3.4)$$

$$L2 = \lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^B \ell_{ij}^{obj} \left[\left(\sqrt{w_i} - \sqrt{\hat{w}_i} \right)^2 + \left(\sqrt{h_i} - \sqrt{\hat{h}_i} \right)^2 \right] \quad (3.5)$$

$$L3 = \sum_{i=0}^{S^2} \sum_{j=0}^B \ell_{ij}^{obj} (c_i - \hat{c}_i)^2 + \lambda_{noobj} \sum_{i=0}^{S^2} \sum_{j=0}^B \ell_{ij}^{noobj} (c_i - \hat{c}_i)^2 \quad (3.6)$$

$$L4 = \sum_{i=0}^{S^2} \ell_{ij}^{obj} \sum (p_i(c) - \hat{p}_i(c))^2 \quad (3.7)$$

L1 calcule la perte liée à la position prédite de la boîte englobante et à la position de la boîte englobante de vérité terrain en fonction des coordonnées (x_{center}, y_{center}) . ℓ_{ij}^{obj} est défini comme 1 si l'objet est présent à l'intérieur de j^{th} cadre englobant prédit dans la cellule i^{th} , et 0 sinon. La boîte englobante prédite sera « responsable » de la prédiction d'un objet en fonction de la prédiction qui a l'IOU actuelle la plus élevée avec la vérité terrain.

L2 calcule l'erreur de prédiction de la largeur et de la hauteur de la boîte englobante similaire à la première partie de l'équation. Cependant, l'ampleur de l'erreur dans les grandes cases affecte moins l'équation que dans les petites cases. Comme la largeur et la hauteur sont normalisées entre 0 et 1, leurs racines carrées augmentent les différences pour les petites valeurs plus que les grandes valeurs. Par conséquent, la racine carrée de la largeur et de la hauteur de la boîte englobante est utilisée au lieu de la largeur et de la hauteur directement.

Le score de perte de confiance est calculé dans les deux cas, que l'objet soit présent ou non dans la boîte englobante. La fonction de perte (L3) ne pénalise l'erreur de confiance d'objet que si ce prédicteur est responsable de la boîte de vérité terrain. ℓ_{ij}^{obj} vaut 1 lorsqu'il y a un objet dans la cellule, et 0 sinon. ℓ_{ij}^{noobj} est le contraire.

L4 est similaire à la perte de classification normale qui calcule la probabilité de perte de classe, à l'exception du terme ℓ_{ij}^{obj} . Ce terme est utilisé car YOLOv5 ne pénalise pas les erreurs de classification même lorsqu'il n'y a pas d'objets présents dans la cellule [44].

3.6 Conclusion

Dans ce chapitre, nous avons expliqué en détail l'algorithme YOLOv5 pour la détection d'objets, qui est considéré comme l'un des meilleurs algorithmes appliqués en temps réel en termes de précision et d'efficacité.

Dans le chapitre suivant, nous décrivons la conception et l'implémentation de notre

système et examinons les résultats obtenus.

Chapitre 4

Conception, Implementation et Discussion des resultats

4.1 Introduction

Nous commençons ce chapitre par décrire la mise en œuvre des différentes phases du système proposé qui consiste à reconnaître les panneaux de signalisation ainsi que les obstacles que peut un conducteur trouvé dans un milieu urbain. Nous traiterons et discuterons par la suite des différentes expérimentations effectuées sur l’algorithme de détection et reconnaissance des panneaux de signalisation en milieu urbain.

4.2 Environnements et outils de développement

4.2.1 Matériel utilisé

- **Processeur** :Intel(R) Core(TM) i5-4200M CPU @ 2.50GHz.
- **Mémoire** :8.00 Go.
- **Système d’exploitation** :Windows 10, 64 bits, processeur x64

4.2.2 Langages de programmation

4.2.2.1 python

GUIDO VAN ROSSUM a lancé le langage de programmation Python, qui est maintenant le langage de programmation open source le plus populaire parmi les program-

meurs. Python est un langage de programmation orienté objet (POO) polyvalent et facile à apprendre qui est très populaire parmi les développeurs et les programmeurs. Les packages Python (bibliothèques) encouragent la modularité et la réutilisation du code, une fonctionnalité très importante pour accélérer le processus de développement logiciel [29].

<https://www.python.org/>



FIGURE 4.1 – Logo python

4.2.3 Environnements de développement

4.2.3.1 Visual Studio Code (VS Code)

VS Code est un éditeur de code libre et gratuit développé par Microsoft. Très populaire parmi les développeurs, il permet d'écrire, d'éditer et de déboguer du code. Hautement personnalisable, il offre une large gamme d'extensions pour des fonctionnalités supplémentaires. Il prend en charge plusieurs langages de programmation tels que JavaScript, Python, C++ et Java. Parmi ses nombreuses fonctionnalités, on trouve des outils de débogage intégrés, l'autocomplétion, la coloration syntaxique et le formatage du code [19].

<https://code.visualstudio.com/>

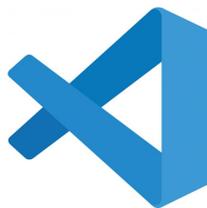


FIGURE 4.2 – Logo Visual Studio Code

4.2.3.2 Roboflow

Roboflow est une plateforme de développement d'applications de vision par ordinateur qui améliore les méthodes de collecte de données, de prétraitement et d'entraînement de modèles. Les utilisateurs de Roboflow ont facilement accès à des ensembles de données publiques et ont la possibilité de soumettre leurs propres données. Roboflow prend en charge de nombreux formats d'annotation [20].

<https://roboflow.com>



FIGURE 4.3 – Logo Roboflow

4.2.3.3 Google colab

Colab, abréviation de Colaboratoire, est une plateforme en ligne qui permet aux utilisateurs d'écrire et d'exécuter des scripts Python dans leur navigateur Web sans avoir besoin de configurer un environnement de développement local. Colab offre un accès gratuit à un GPU, ce qui est particulièrement utile pour accélérer les calculs intensifs liés à l'apprentissage automatique et à l'analyse de données [29][22].

<https://colab.research.google.com/>



FIGURE 4.4 – Logo Google colab

4.2.4 bibliothèque utilisé

1. **OpenCV** : OpenCV est une bibliothèque gratuite qui se concentre principalement sur le traitement d'images en temps réel, mais aussi d'une bibliothèque de liaison de logiciels visant à développer la vision par ordinateur. Développé à

l'origine par Intel, il a ensuite été pris en charge par Willow Garage for Robotics et Itseez [22].

<https://opencv.org/>



FIGURE 4.5 – Logo OpenCV

2. **Kivy** : Kivy est un framework open source pour les applications pour smartphones, tablettes, PC, etc. basé sur le langage de programmation Python. Kivy se concentre particulièrement sur l'interface graphique, l'interactivité et permet aux développeurs de créer des applications interopérables sur plusieurs plateformes [43].

<https://kivy.org/>



FIGURE 4.6 – Logo Kivy

3. **Pygame** : La lecture audio est prise en charge par le package Python Pygame Mixer. Dans n'importe quel programme Python ordinaire, vous pouvez lire de l'audio et de la musique en utilisant la bibliothèque indépendante Mixer [21].
4. **PIL** : Pillow est une bibliothèque Python pour le traitement d'images, basée sur la bibliothèque d'imagerie Python (PIL). Elle permet d'ouvrir, de manipuler et d'enregistrer des images. Pillow prend en charge de nombreux formats d'image et offre un support limité pour l'écriture, en se concentrant principalement sur les formats d'échange et de présentation les plus couramment utilisés [22].

<https://pypi.org/project/Pillow/>



FIGURE 4.7 – Logo Pillow

5. **PyTorch** : PyTorch est une bibliothèque, développée par Facebook écrite en Python. Le package PyTorch contient des structures de données spécifiques aux tenseurs multidimensionnels et définit des opérations arithmétiques sur ces tenseurs. De plus, il fournit plusieurs utilitaires pour faciliter la concaténation des Tenseurs et des Types Arbitraires [46].

<https://PyTorch.org/>



FIGURE 4.8 – Logo PyTorch

6. **TensorFlow** : tensorflow est une bibliothèque libre logicielle open source. Elle est couramment utilisée pour des applications d'apprentissage automatique telles que les réseaux de neurones, et est considérée comme une bibliothèque mathématique symbolique [26].

<https://www.tensorflow.org/>



FIGURE 4.9 – Logo tensorflow

4.3 Etude de conception

Après avoir découvert les systèmes d'aide à la conduite et les méthodes utilisées pour reconnaître et détecter les objets dans le chapitre précédent, nous allons présenter dans cette section les différentes étapes utilisées dans la reconnaissance des panneaux de signalisation routière et des dangers.

4.3.1 Présentation de l'architecture générale

L'architecture générale du système, telle que illustrée par la figure 4.1, comprend trois étapes principales : La Perception, l'étape de Détection et l'étape d'Identification pour pouvoir donner suite à la signalisation du danger.

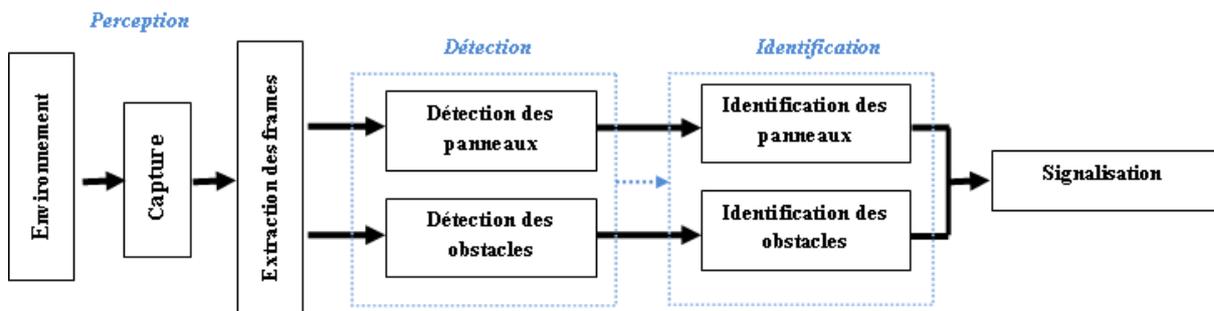


FIGURE 4.10 – Représentation générale du système.

4.3.2 Architecture détaillée

4.3.2.1 Conception de la base des images format yolov5s

Le réseau yolov5s nécessite en entrée une base de données portant le format yolov5s décrite comme suit : Les images de ces deux bases sont accompagnées de données :

- Une image comportant les panneaux au format **jpg**.
- Une annotation, c'est-à-dire un fichier texte (**txt**) portant le même nom que l'image elle-même et décrivant la position des panneaux, leur taille, et leur type.

Lorsque l'image est introduite dans le modèle yolov5s, celui-ci va la redimensionner en 640 X 640 pixels afin d'unifier la taille de toutes les images et d'obtenir une plus grande précision pour effectuer l'entraînement sur l'ensemble d'images.

4.3.2.2 Apprentissage du réseau yolov5

L'étape cruciale dans la conception de notre système est de construire et d'apprendre à notre réseau profond à détecter les panneaux et les dangers présents sur la route. L'apprentissage consiste à présenter au classifieur plusieurs exemples de panneaux de signalisation et de lui indiquer à chaque fois de quel type de panneau qu'il s'agit. Pour la réalisation de cette étape, nous avons fait recours au réseau profond CNN modèle yolov5s (détaillé dans le troisième chapitre) qui est la plus petite version des modèles de détection d'objets Yolov5. Il a été conçu pour être plus rapide et plus léger tout en conservant une bonne précision de détection.

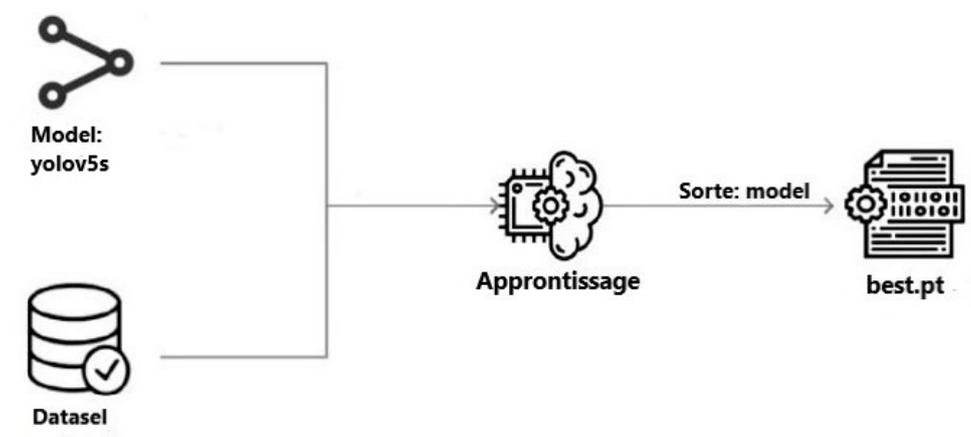


FIGURE 4.11 – l'apprentissage d'un model.

Une fois l'étape d'apprentissage du modèle yolov5s achevée, nous pouvons passer au test et validation de notre système :

4.3.2.3 Test et validation

Cette étape consiste à percevoir l'environnement à l'aide d'un capteur embarqué sur une voiture en mouvement. Les données récupérées à partir de ce capteur ou caméra seront une scène vidéo qui nécessite d'être fragmentée en un ensemble de frames. Ces frames présenteront dans la suite les entrées du système conçu. Cette étape est réalisée à l'aide des fonctionnalités du OpenCV.

La détection et la reconnaissance des objets Pour assurer la sécurité des conducteur, nous avons réalisé un système capable de détecter et de reconnaître les panneaux de signalisation à l'aide de techniques d'apprentissage en profondeur(réseau yolov5s).

```
capture = cv2.VideoCapture(0)
ret, frame = capture.read()
```

FIGURE 4.12 – ouvrir la camera

Le conducteur, en mouvement dans un milieu urbain, est soumis à d'autres dangers pouvant entraîner des risques d'accidents tels que les piétons et les animaux qui traversent la route.

De ce fait, nous avons appuyé notre système yolov5s entraîné sur une base de panneaux de signalisations, par un autre réseau yolov5s entraîné préalablement sur des êtres humains, les deux types d'animaux qu'on peut trouver sur le milieu urbain : le chat et le chien et ainsi que différentes automobiles à savoir : bus, vélo, moto ou camion.

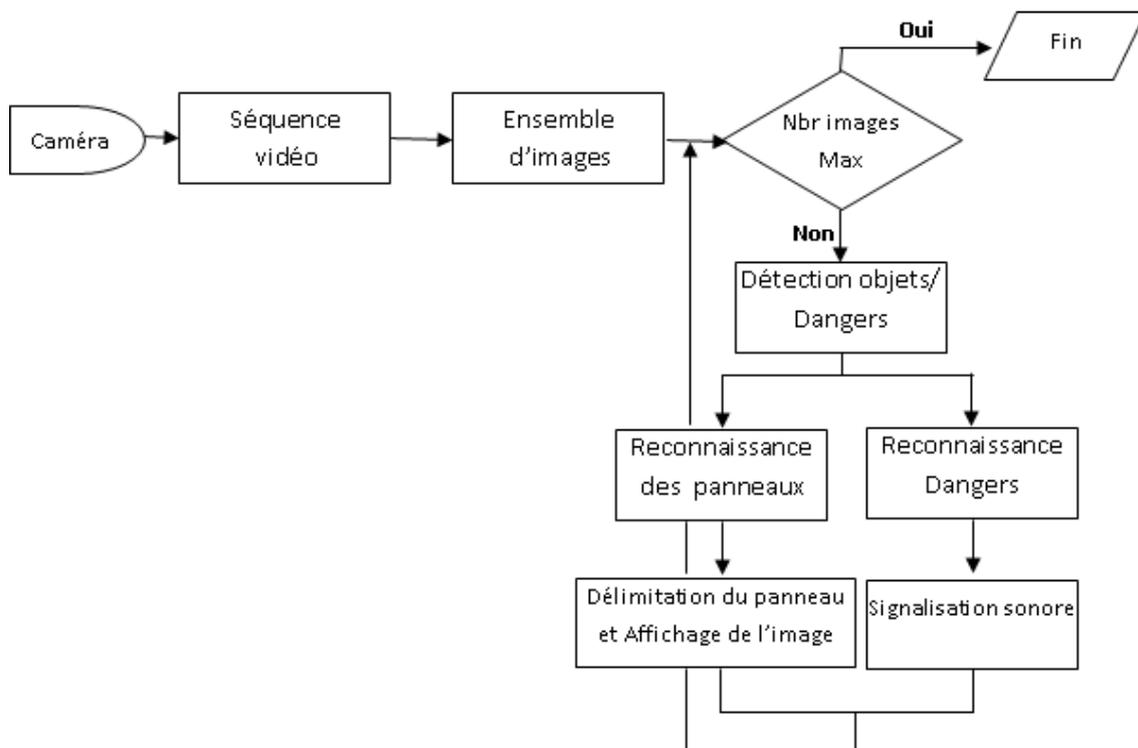


FIGURE 4.13 – Processus de détection par le modèle yolov5s

Le choix de combiner deux systèmes yolov5s, chacun spécialisé dans un type d'image est justifié par :

- Premièrement, en augmentant le nombre de classes à apprendre au réseau nécessite une machine puissante et un grand débit d'internet, chose qui n'a pas été disponible.

- deuxièmement, et la raison la plus importante est que notre but est de réaliser une application en temps réel puisqu'elle est destinée à un conducteur en mouvement. Nous avons donc proposé de réaliser un système à deux réseaux profond qui fonctionnent en parallèle. Cela nous a conduit à entraîner des réseaux profonds spécialisés, c'est-à-dire, chacun d'entre eux ne reconnaît qu'un seul type d'objet de fait que les panneaux de signalisation ne présentent pas les mêmes caractéristiques que et les êtres humains et les animaux.

Les objets (panneaux de signalisation et/ ou obstacles) sont détectés par le réseau profond puis classés pour être identifiés. L'étape de la détection est illustrée par l'algorithme ci-dessous :

Algorithme 1 : Détection d'objet

```

model ← best.pt;
classee ← model.names;
capture ← captureVideo(0);
while true do
    | frame ← readCapture();
    | result ← model(frame);
    | if inds ≠ 0 then
    |     dessine rectangle sur l'objet ;
    |     création d'un nom de l'objet affecte par le fichier name classes.names ;
    | end
end

```

4.3.2.4 Sorties du système

Le système proposé dans ce travail vise à détecter et reconnaître les panneaux de signalisation et les obstacles rencontrés par un conducteur en mouvement dans un milieu urbain. De ce fait les sorties de ce système vont être une image de l'environnement aperçu avec délimitation de la plaque de signalisation ou de l'objet danger détecté et reconnu. Et puisque cette application est destinée à un conducteur en mouvement, c'est-à-dire qu'il ne peut pas consulter constamment les images fournies par ce système, nous avons doté notre système par des sorties sonores indiquant le type d'objet ou de danger reconnus.

Dans ce qui suit, nous allons présenter les résultats obtenus par le test de notre système.

4.4 Implémentation et discussion des résultats

Dans l'objectif d'évaluer les performances de l'approche présentées dans la section précédente, nous avons procédé à des expérimentations et évaluations du système proposé.

4.4.1 Choix de la base d'images

Pour évaluer notre algorithme, et après une recherche sur internet, nous avons opté pour les deux bases compatibles au modèle yolov5s, et qui sont largement utilisées dans des recherches à savoir : **GTSRB** (German Traffic Sign Recognition Benchmark) et **China Traffic Signs Dataset** (Mapillary and DFG) décrites dans la section 4.3.2.1. La description des bases d'images utilisées est donnée par le tableau ci-dessous :

Titre	Nb d'images	Nb de classes	Taille d'images	Source
GTSRB	50,000	40	varie entre 15*15 et 250*250 pixele	Allemagne
Traffic Signs Dataset	5998	58	varie entre 28 et 402 pixels en Longueur et 26 et 491 pixele en Largeur	Chine

TABLE 4.1 – Description des deux bases d'images

Dans notre travail, nous nous intéressons au milieu urbain. Or dans l'urbain on ne trouve pas toutes les catégories de panneaux de signalisation, comme par exemple les limites de vitesse et les alertes de virages. De ce fait, nous avons opté pour ne prendre en considération que les classes existantes dans ce milieu.

Après avoir fait une recherche sur le type de panneaux de signalisation qu'on peut trouvé en urbain. De ce fait, nous avons procédé à la formation de notre propre base d'images à partir des deux bases précédentes et qui soit adapté à nos besoins. Nous avons donc sélectionné 15 classes à savoir : *'Stop'*, *'no-left-turn'*, *'no-right-turn'*, *'regulatory-go-straight'*, *'regulatory-go-straight-or-turn-left'*, *'regulatory-go-straight-or-turn-right'*, *'regulatory-keep-left'*, *'regulatory-keep-right'*,

'regulatory-no-u-turn', 'regulatory-roundabout', 'regulatory-turn-left', 'regulatory-turn-right', 'regulatory-turn-right-or-left', 'regulatory-yield', 'warning-reod-bump' .

Nous avons choisit à partir des deux bases précédente les images qui sont les plus descriptives et qui présentent les différents états des plaques de signalisation sur la route comme l'inclinaison par exemple. Notre base d'images est composée de 1600 images où chaque classe est formée de 100 à 110 images de tailles variantes.



FIGURE 4.14 – Différentes classes de panneaux de signalisation en Urbain

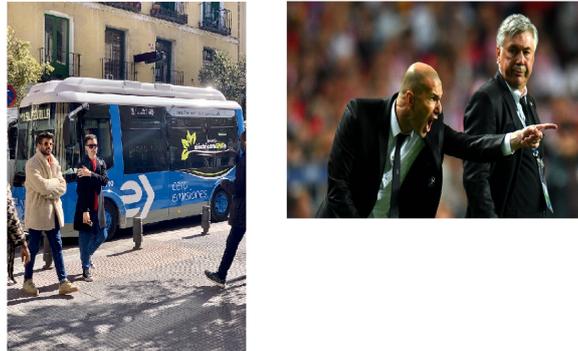


FIGURE 4.15 – Exemples des images de la base utilisée pour l'apprentissage pour la reconnaissance des obstacles

4.4.2 Discussion des résultats obtenus

Nous allons présenter dans ce qui suit les résultats des tests effectués sur différentes images.

4.4.2.1 Apprentissage du réseau profond yolov5s

Le tableau ci-dessous illustre la précision et le taux de reconnaissance obtenus par l'apprentissage du modèle yolov5s et ce, en variant le taux des itérations.

Nbr itérations	20	50	150	200
Précision	50%	70%	83%	83%
taux de reconnaissance	70%	95%	93%	93%

TABLE 4.2 – Résultats de l'apprentissage du réseau yolov5s

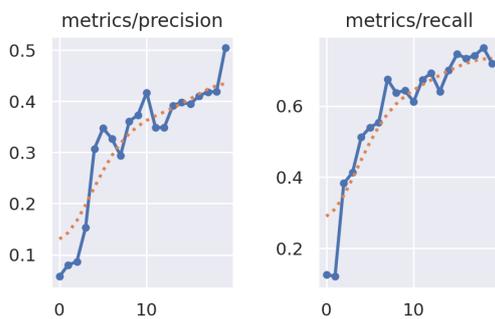


FIGURE 4.16 – 20 itérations

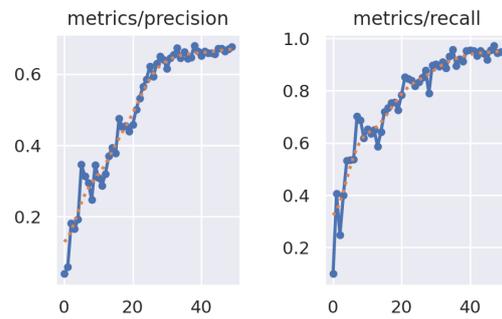


FIGURE 4.17 – 50 itérations

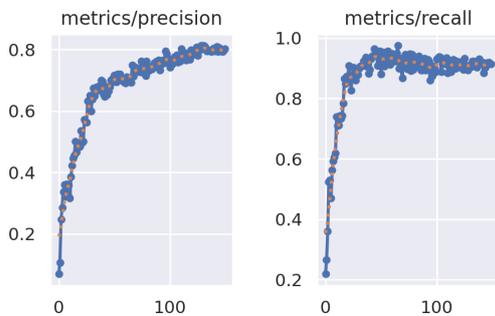


FIGURE 4.18 – 150 itérations

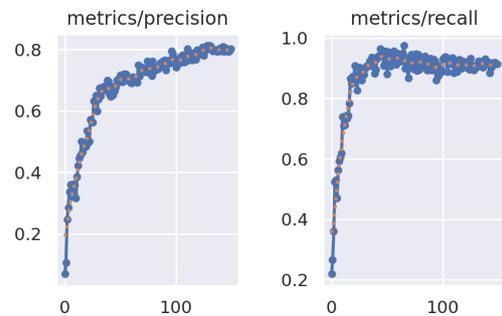


FIGURE 4.19 – 200 itérations

Nous remarquons que quand on augmente le nombre des itérations, le taux de reconnaissance augmente avec une meilleure précision du modelé. Mais à partir de 150 itérations, cette précision ne varie presque pas, alors que le taux de précision recule un peu. Nous avons opté pour un compromis entre la précision et le taux de reconnaissance, et vu que ce dernier, et malgré qu'il a diminué un peu, il reste toujours élevé. C'est pourquoi nous avons opté pour le nombre d'itérations de 150. La matrice de confusion relative à la reconnaissance des plaques de signalisation est illustrée par la figure ci-dessus :

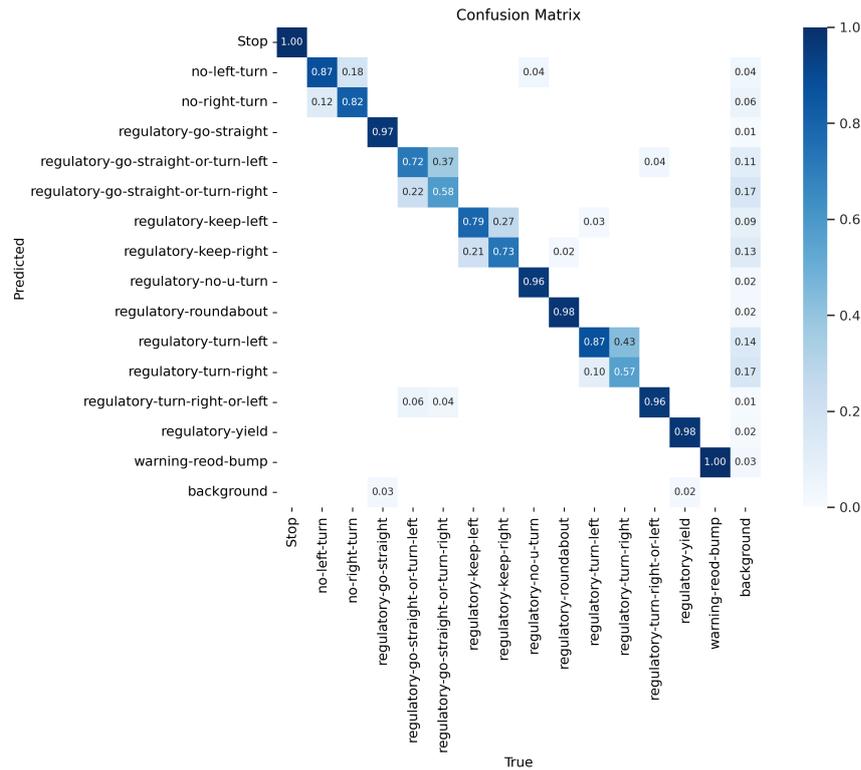


FIGURE 4.20 – Matrice de confusion relatives aux panneaux de signalisation

4.4.3 Détection des objets par le réseau yolov5s

1. Chargement du modèle

```
self.model_trf = torch.hub.load('ultralytics/yolov5', 'custom', 'best.pt', force_reload=True)
self.cla=self.model_trf.names
```

FIGURE 4.21 – Chargement du modèle

2. Détection et délimitation de l'objet L'objet détecté est délimité par un rectangle accompagné de son nom.

```
results = self.model(frame)
labels, cord = results.xyxy[0][:, -1], results.xyxy[0][:, :-1]
cv2.rectangle(frame, (x, y), (w, h), bgr, 4)
```

FIGURE 4.22 – Détection des objets

3. Reconnaissance de l'objet détecté

La reconnaissance de l'objet détecté induit l'affectation d'un nom à ce dernier (panneaux de signalisation ou obstacle).

```
cv2.putText(frame, f'{{label}} {{confidence:.2f}}', (x+5, y+30),  
            cv2.FONT_HERSHEY_COMPLEX, 1.0, (0,0,255), 3)
```

FIGURE 4.23 – Affectation du nom de l'objet

4.4.4 Tests et résultats obtenus

Nous avons testé notre système sur des séquences d'images réelles et sur des images présentant différents panneaux de signalisation ainsi que des obstacles, et ce, en utilisant la caméra de notre ordinateur. Les résultats obtenus sont présentés dans ce qui suit :

1. Détection des panneaux de signalisation

Dans un premier temps, nous avons présenté au système des images relatives aux différents panneaux de signalisation. Les tests ont été menés à détecter et reconnaître en premier lieu un seul panneau, et deux panneaux de signalisation en deuxième lieu. Notre système arrive à détecter et à reconnaître toutes les plaques présentées en les délimitant sur les images et en alertant le conducteur par le nom de la ou les plaques détectées. Le système proposé est donc capable de détecter les plaques de signalisation quelle que soit sa forme, couleur ou nombre.

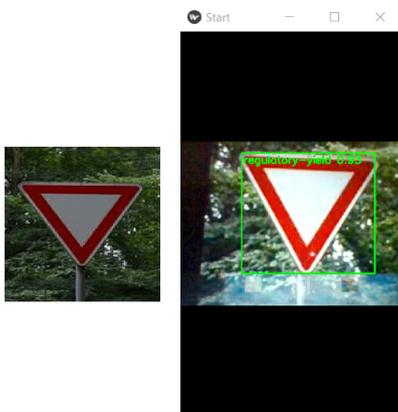


FIGURE 4.24 – Taux.Recon=93%

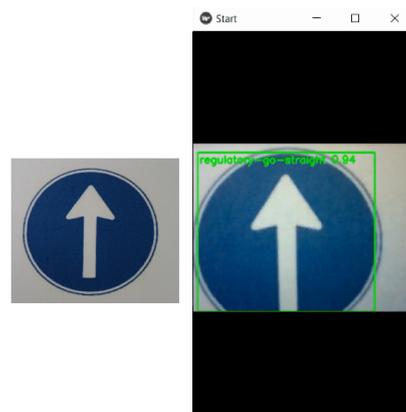


FIGURE 4.25 – Taux.Recon=94%

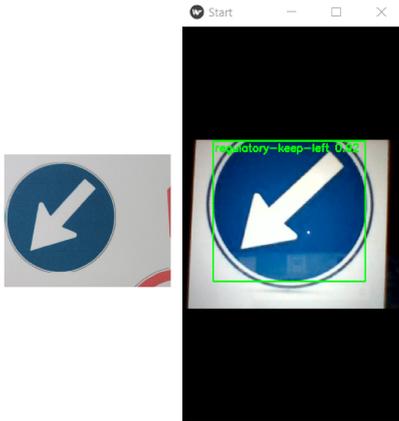


FIGURE 4.26 – Taux.Recon=92%

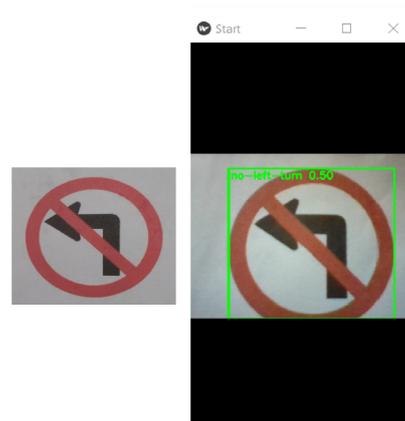


FIGURE 4.27 – Taux.Recon=50%

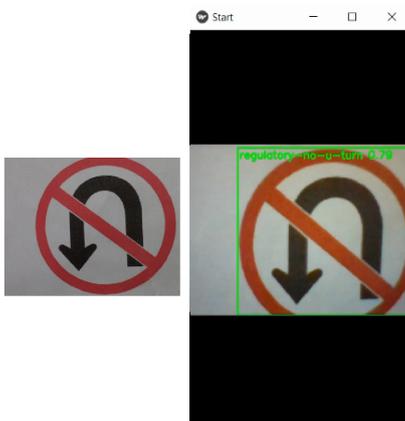


FIGURE 4.28 – Taux.Recon=79%



FIGURE 4.29 – Taux.Recon=96%

Le système proposé arrive à détecter et à reconnaître toutes les plaques dont il a été entraîné avec un taux de reconnaissance variables. Ceci peut être expliqué par le nombre d'images utilisées pour chaque type dans la phase d'apprentissage. La figure 4.35 illustre les taux de reconnaissance relatif à chaque type de plaques de signalisation :

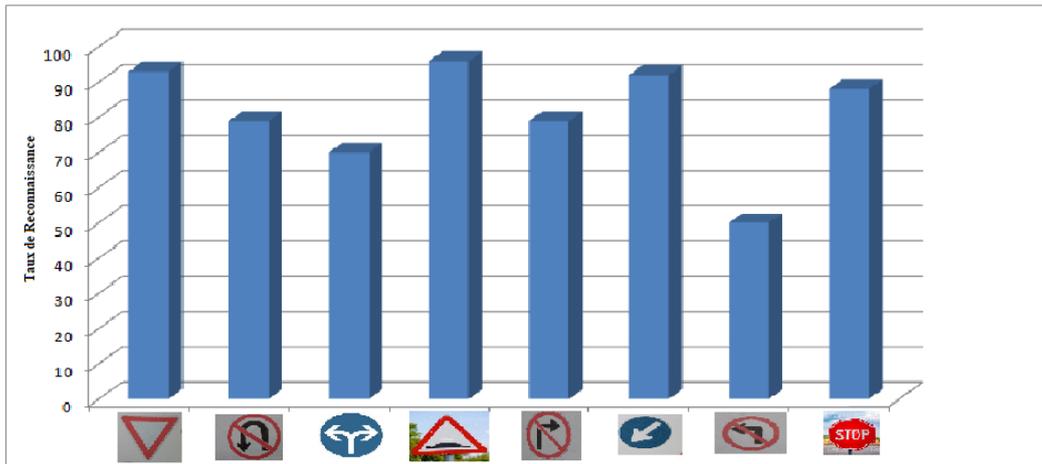


FIGURE 4.30 – Taux de reconnaissance des plaques de signalisation

Notons que nous pas pris en considération les plaques de signalisation comportant des écritures car cela est plus compliqué car il nécessite un système de reconnaissance de l'écriture.



FIGURE 4.31 – Taux de reconnaissance des plaques de signalisation

2. Détection des obstacles

Nous avons procédé en deuxième lieu au test de notre système sur des images relative au milieu urbain, ces images contiennent des personnes, des automobiles ou des animaux. Le système arrive à détecter et à reconnaître les objets présents dans les images de test. Cette reconnaissance est illustrée par une délimitation de l'objet suivie d'un bip sonore alertant le conducteur de la présence d'un obstacle.

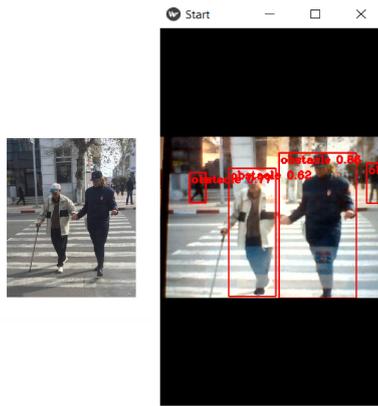


FIGURE 4.32 – Taux.Recon=96%

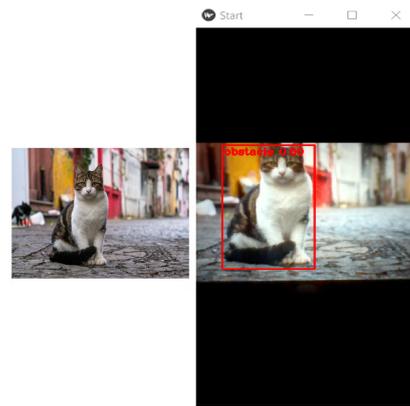


FIGURE 4.33 – Taux.Recon=50%

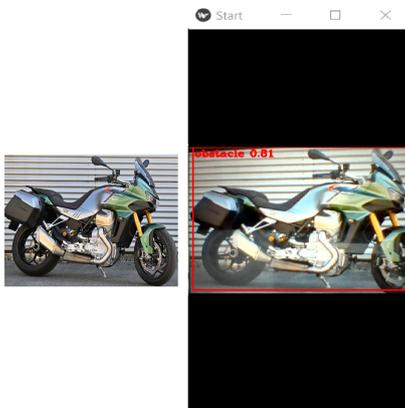


FIGURE 4.34 – Taux.Recon=81%



FIGURE 4.35 – Taux.Recon=83%

Nous pouvons constaté que le système proposé arrive à détecter les objets qu'il a appris en apprentissage.

Nous avons présenté à notre système en deuxième lieu des images de personnes, d'animaux ou d'objets dans le même angle de vue d'un conducteur au volons. Les résultats sont satisfaisant. En effet, le système proposé arrive à détecter et à reconnaître tous les obstacles qui se trouvent dans les images. Les résultats sont illustrés par les figures suivantes :

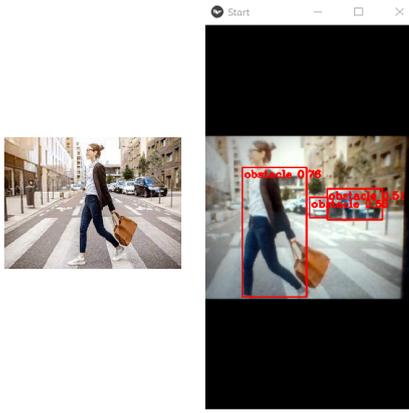


FIGURE 4.36 – Taux.Recon=76%



FIGURE 4.37 – Taux.Recon=83%

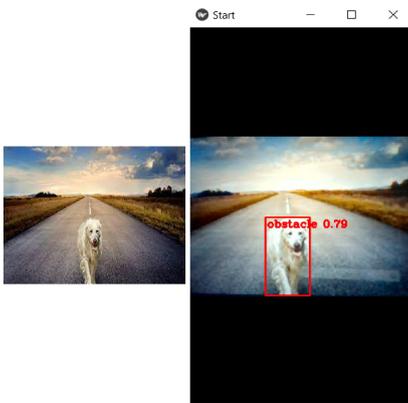


FIGURE 4.38 – Taux.Recon=79%

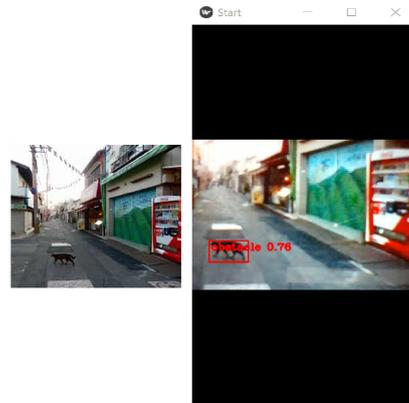


FIGURE 4.39 – Taux.Recon=76%

voici en diagramme, le taux de reconnaissance des différents obstacles par notre système :

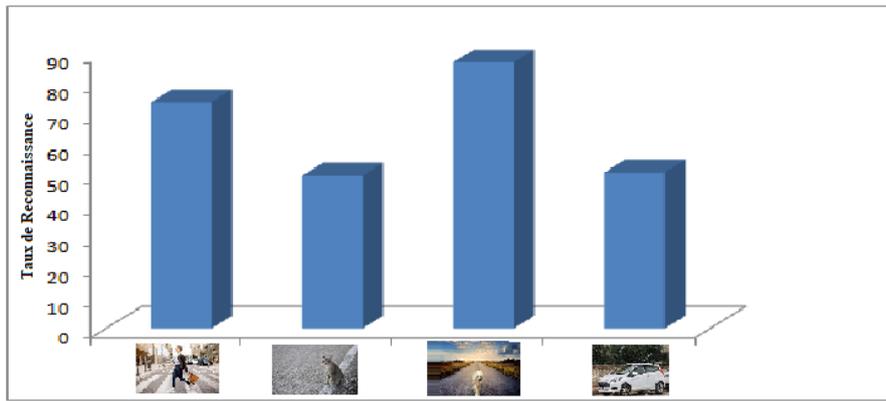


FIGURE 4.40 – Taux de reconnaissance des plaques de signalisation

3. Détection de plusieurs types d’objets

Le système proposé est capable de détecter, les panneaux de signalisation, et les dangers que peut un automobiliste rencontrer dans un milieu urbain. Cela est réalisé en délimitant tous les objets détectés et en alertant le conducteur par le nom de la plaque et un bip sonore.

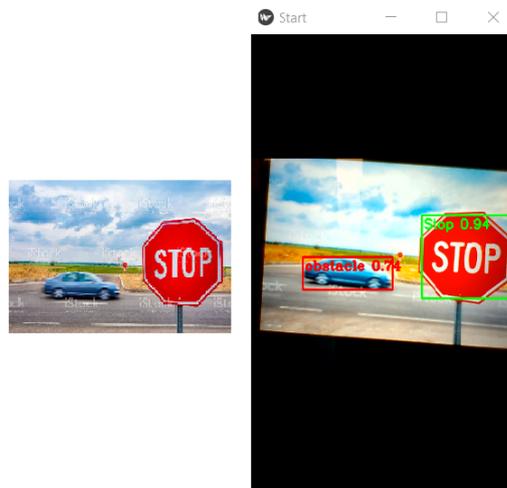


FIGURE 4.41 – Détection de panneaux de signalisation et une voiture

Analysons maintenant les figures suivantes :

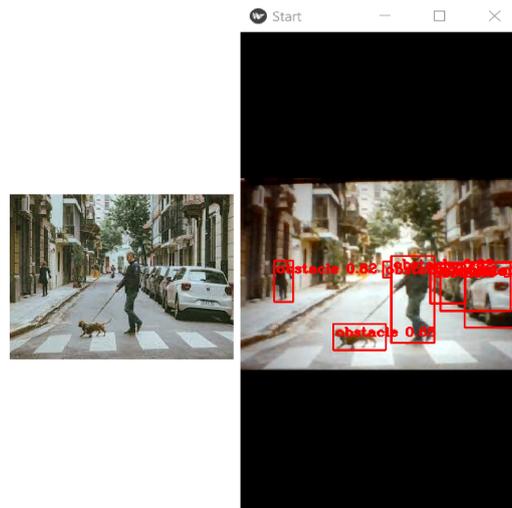


FIGURE 4.42 – Détection des obstacles

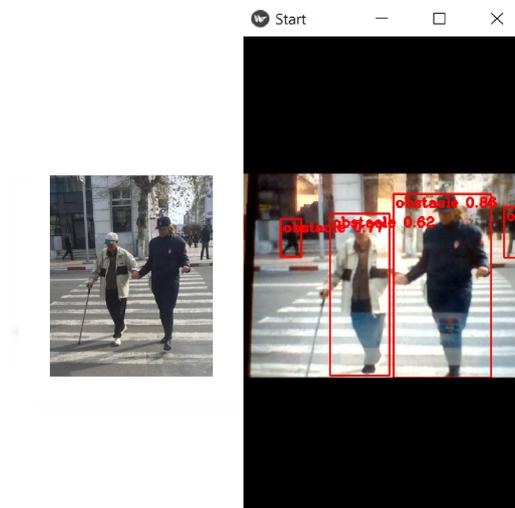


FIGURE 4.43 – Détection des obstacles

Nous pouvons constaté que le système a bien détecté et reconnu les obstacles se trouvant sur la figure. Néanmoins, il a détecter même les personnes et les objets qui ne présentent pas un danger pour le conducteur comme les personnes se trouvant sur le trottoir ou les voiture qui sont garées.

Cette situation peut entraîner un cas de perturbation pour le conducteur en lui signalent un obstacle qui ne présente pas un danger pour lui.

Pour cela, il est nécessaire d'améliorer d'avantage ce système afin qu'il ne détecte que les objet qui se trouvent sur l'axe de la route.

Dans le cadre d'évaluer notre système sur des objets mobiles, nous avons procédé

à des tests sur des scènes en utilisant la caméra de notre ordinateur. Les résultats obtenus sont encourageants du fait que la détection des objets mobiles est réalisée avec une précision qui dépasse les 50% en fonction de la luminosité et la position de l'objet détecté par rapport la caméra.

4.5 conclusion

Ce chapitre nous a permis de présenter les résultats obtenus par la méthode proposée sur plusieurs images présentant des panneaux de circulation, des personnes, des animaux et des automobiles.

Notre approche constitue un outil d'aide à la conduite, destiné aux conducteur d'automobiles en mouvement et dotés de caméra. De ce fait, nous avons proposé une méthode de détection et de reconnaissance des plaques de la circulation et des obstacles que peut rencontrer un conducteur en milieu urbain (personne, animal, automobile) en utilisant un réseau de neurone profond à savoir le modèle yolov5s.

Les tests effectués sur plusieurs images relatives aux panneaux de circulations ainsi que des obstacles vous ont révélé la puissance du système proposé dans la détection et la reconnaissance des objets rencontrés mais avec une exactitude variante selon le nombre d'images correspondant en phase d'apprentissage.

Nous avons constaté également que le système proposé détecte même les objets qui ne présentent pas un danger pour le conducteur comme les personnes et les animaux se trouvant sur le trottoir, ou des automobiles qui sont garés. Pour soulever ce problème. Le traitement de ce problème peut faire la problématique d'un autre sujet.

Le système proposé est fiable en milieu urbain, il pourrait être efficace pour les conducteurs imprudents et les piétons ayant un comportement à risque. Il alerte le conducteur en temps réel en fonction de la gravité du risque d'accident.

Chapitre 5

conclusion générale

5.1 Conclusion

Au cours de la dernière décennie, les systèmes de transport intelligents ont connu un énorme progrès, principalement au niveau des systèmes d'aide à la conduite avancés (ADAS). Cependant, les accidents mortels sont toujours en hausse, conservant ainsi la sécurité routière comme un enjeu important pour la recherche scientifique. Les accidents de la route en milieu urbain ne sont pas moins importants que celle du milieu rurale. De ce fait, nous avons proposé un système d'aide à la conduite dédié au milieu urbain.

5.2 Bilan des travaux réalisés et des résultats obtenus

Dans le cadre de cette étude, nous avons fait appel aux outils de l'intelligence artificielle et plus particulièrement les réseaux de neurones profonds pour détecter et reconnaître les plaques de signalisation ainsi que les obstacles que peut rencontrer un conducteur en milieu urbain.

Dans ce contexte, nous avons apporté une vue d'ensemble des algorithmes de vision existants pour la détection des panneaux de circulation, allant des descripteurs classiques aux méthodes d'apprentissage automatique. Le système proposé est dédié à un conducteur en mouvement, il est donc important qu'il donne des résultats en temps réel. De ce fait, et afin d'augmenter également la précision d'avantage, Nous nous somme intéressé à la troisième version de l'algorithme yolo à savoir yolov5s qui vise

à extraire des caractéristiques pour détecter et reconnaître des objets dans l'image en temps réel et qui a largement surpassé d'autres algorithmes de détection d'objets.

Nous avons donc implémenté deux réseaux profonds modèles yolov5s qui sont spécialisés et travaillent conjointement. C'est-à-dire, que chaque réseau n'apprend qu'un seul type d'images : un pour les plaques de signalisation et le deuxième les obstacles rencontrés en milieu urbain (être humain, chat ou chien traversant la route, des automobiles). Ce choix a été fait afin d'accélérer le temps d'apprentissage et de détection d'un côté, et d'augmenter le taux de précision d'un autre côté.

Afin d'entraîner le réseau susmentionné, nous avons créé une nouvelle base d'images formée à partir de deux bases (GTSRB et China Traffic Signs Dataset) de telle façon que ne prendre en considération que les panneaux de signalisation rencontrés en milieu urbain. Nous avons de choisir les images qui sont les plus descriptives et qui présentent les différents états des plaques de signalisation sur la route comme l'inclinaison par exemple. Nous avons tout d'abord pu constater que notre application est capable de s'exécuter en temps réel.

L'analyse des résultats obtenus sur des images de panneaux de la signalisation nous ont révélé sur différents types d'images que le système arrive à détecter et reconnaître toutes les plaques mais avec un taux variant. Ceci est expliqué par le fait que le nombre d'images, de chaque catégorie, appris au réseau n'était pas le même.

Le système proposé arrive à détecter et reconnaître des objets considérés comme obstacles sur des images portant des personnes, des chats, chiens et des automobiles. Néanmoins, il arrive à détecter tous les objets figurants sur l'image même s'ils ne présentent pas un obstacle pour le conducteur (personnes et animaux sur le trottoir, et voiture garée). Cette fausse alerte peut perturber le conducteur.

Dans le cadre d'évaluer notre système sur des objets mobiles, nous avons procédé à des tests sur des scènes en utilisant la caméra de notre ordinateur. Les résultats obtenus sont encourageants du fait que la détection des objets mobiles est réalisée avec une précision qui dépasse les 50% en fonction de la luminosité et la position de l'objet détecté par rapport la caméra.

Notons que n'avons pas pris les panneaux de signalisation qui contiennent des écritures car cela nécessite un système de reconnaissance de l'écriture.

5.3 Perspectives

Les résultats atteints par notre système sur plusieurs images et scènes vidéo sont encourageants. Néanmoins, Ce dernier présente des limites qui ont été soulevé dans le chapitre précédent. Des améliorations seraient donc sans doute souhaitables dans le futur.

Pour cela, nous proposons d'enrichir d'avantage notre base d'apprentissage afin d'augmenter le taux de reconnaissances d'un coté, et qu'il soit capable de reconnaître des personnes et objets dans des différents angles de vue. Nous proposons également d'améliorer ce système afin qu'il ne détecte que les objets qui présentent un état de danger pour le conducteur en ajoutant un module de reconnaissance pour éliminer les fausses alarmes constituées par des objets or champ de danger.

Nous proposons également d'introduire les paramètres de luminosité et conditions météorologiques pour augmenter le taux de reconnaissance sous différentes contraintes. Afin d'améliorer d'avantage le système proposé, nous proposons également d'appuyer notre système par un autre qui peut reconnaître les plaques manuscrites.

Une autre perspective est de transformer ce système en une application mobile afin qu'elle soit facilement utilisable par les conducteur.

Bibliographie

- [1] Nakul Maan, Vision and Image Processing in Automotive ADAS/AD Development, Disponible sur : <https://www.linkedin.com/pulse/vision-image-processing-automotive-adasad-development-nakul-maan>. Publiée le 23 mai 2021, consultée le 06 mai 2023.
- [2] UN RÉGULATEUR DE VITESSE ADAPTATIF, ÇA SERT À QUOI?, Disponible sur : <https://mobile.guideautoweb.com/articles/66285/un-regulateur-de-vitesse-adaptatif-ca-sert-a-quoi/>. Publiée le 15 juin 2022, consultée le 04 mai 2023.
- [3] L'aide au freinage d'urgence, Disponible sur : <https://www.codedelaroute.fr/articles/aide-au-freinage-urgence/>. Publiée le 29 décembre 2021, consultée le 05 mai 2023.
- [4] Alberto Rizzoli, The Ultimate Guide to Object Detection, Disponible sur : <https://www.v7labs.com/blog/object-detection-guide>. Publiée le 10 June 2021, consultée le 12 mars 2023.
- [5] Qu'est-ce que la Computer Vision ?, Disponible sur : <https://www.ibm.com/fr-fr/topics/computer-vision>. consultée le 04 avril 2023.
- [6] présentation de la vision par ordinateur, Disponible sur : <https://azure.microsoft.com/fr-fr/resources/cloud-computing-dictionary/what-is-computer-vision/>, consultée le 28 mars 2023.
- [7] Réseaux neuronaux, Disponible sur : <https://www.ibm.com/fr-fr/topics/neural-networks>. consultée le 17 mars 2023.
- [8] Chiradeep BasuMallick, What Is ADAS (Advanced Driver Assistance Systems) ? Meaning, Working, Types, Importance, and Applications, Disponible sur : <https://www.spiceworks.com/tech/iot/articles/what-is-adas/>. Publiée July 15, 2022, consultée le 20 avril 2023.

- [9] ADAS systems : definition, advantages and training, Disponible sur : <https://www.sneci.com/blog/adas-systems-definition-advantages-and-training/>. Publiée le 03 February 2022, consultée le 04 mai 2023.
- [10] What is ADAS ?, Disponible sur : <https://www.synopsys.com/automotive/what-is-adas.html>. consultée le 04 mai 2023.
- [11] Qu'est-ce qu'ADAS? Importance des données de formation pour former des modèles ADAS, Disponible sur : <https://fr.shaip.com/blog/training-data-for-adas/>. Publiée 26 Avril 2022, consultée le 04 mai 2023.
- [12] Quels sont les différents types d'ADAS ou d'aide à la conduite?, Disponible sur : <https://www.carglass.lu/bris-de-glace/calibrage-systemes-de-aide-a-la-conduite/types-adas>. consultée le 07 mai 2023.
- [13] Top 10 des dispositifs d'aide à la conduite (parfois oubliés) dans nos voitures, Disponible sur : <https://media.roole.fr/quotidien/equipement/top-10-des-dispositifs-daide-a-la-conduite>. Publiée le 02 mars 2021, consultée le 05 mai 2023.
- [14] Jason Brownlee,How to Use Mask R-CNN in Keras for Object Detection in Photographs, Disponible sur : <https://machinelearningmastery.com/how-to-perform-object-detection-in-photographs-with-mask-r-cnn-in-keras/>. Publiée le 24 may 2019, consultée le 04 Juin 2023.
- [15] Grace Karimi,Introduction to YOLO Algorithm for Object Detection, Disponible sur : <https://www.section.io/engineering-education/introduction-to-yolo-algorithm-for-object-detection/>. Publiée le 15 April 2021, consultée le 01 avril 2023.
- [16] Rohit Kundu,YOLO : Algorithm for Object Detection Explained [+Examples], Disponible sur : <https://www.v7labs.com/blog/yolo-object-detection>. Publiée le 17 January 2023, consultée le 01 avril 2023.
- [17] Analytics Vidhya,Basics of Bounding Boxes, Disponible sur : <https://medium.com/analytics-vidhya/basics-of-bounding-boxes-94e583b5e16c>. Publiée le 16 Jan 2021, consultée le 23 Juin 2023.
- [18] YOLO Algorithm, Disponible sur : <https://dkhara-razi.github.io/notes/ml/cnn/yolo>. consultée le 05 avril 2023.

- [19] Martin Heller, What is Visual Studio Code? Microsoft's extensible code editor, Disponible sur : <https://www.infoworld.com/article/3666488/what-is-visual-studio-code-microsofts-extensible-code-editor.html>. Publiée le 8 JUL 2022, consultée le 08 mai 2023.
- [20] Jayita Bhattacharyya, Step by Step Guide To Object Detection Using Roboflow, Disponible sur : <https://analyticsindiamag.com/step-by-step-guide-to-object-detection-using-roboflow/>. Publiée le 11 October 2020, consultée le 04 Juin 2023.
- [21] Disponible sur : <https://coderslegacy.com/python/pygame-mixer/>. consultée le 04 Juin 2023.
- [22] MAAMOULI ABDELOUAHEB. *Utilisation de méthodes de Deep learning pour l'extraction de texte dans les images*. PhD thesis.
- [23] Pape Demba Beye. *Faisabilité technique des systèmes avancés d'aide à la conduite (ADAS) pour la sécurité routière*. PhD thesis, Université du Québec à Trois-Rivières, 2021.
- [24] Abdelkarim Chohra. Exploitation automatique de la signalisation routière. 2016.
- [25] Tausif Diwan, G Anirudh, and Jitendra V Tembhurne. Object detection using yolo : Challenges, architectural successors, datasets and applications. *Multimedia Tools and Applications*, pages 1–33, 2022.
- [26] Lamine Djenaihi. *Un système de détection des objets de la circulation routière et d'estimation de leur distance*. PhD thesis, Mémoire Master INFORMATIQUE, Université Mohamed Khider-BIKSRA-Département de l'informatique, 2019-2020.
- [27] Jon L Grossman. Thermal infrared vs. active infrared : A new technology begins to be commercialized. *Cell*, 305 :546–9538, 2019.
- [28] Ammar Hamadeh. الكشف عن الأغراض باستخدام مكتبة في بايثون object detection imageai, février 2020. consultée le 25 février 2023.
- [29] ABDELAZIZ HAMADI. Real time multi-object tracking using deep learning. 2021.
- [30] Abdelmalek HANNACHI, Imed KHELIFI, and BELKACEM Soundes. *Détection des panneaux de signalisation routière*. PhD thesis, 2020.

- [31] Jonathan Horgan, Ciarán Hughes, John McDonald, and Senthil Yogamani. Vision-based driver assistance systems : Survey, taxonomy and advances. In *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, pages 2032–2039. IEEE, 2015.
- [32] Vipin Kumar Kukkala, Jordan Tunnell, Sudeep Pasricha, and Thomas Bradley. Advanced driver-assistance systems : A path toward autonomous vehicles. *IEEE Consumer Electronics Magazine*, 7(5) :18–25, 2018.
- [33] Zhenzhe Li. Road aerial object detection based on improved yolov5. In *Journal of Physics : Conference Series*, volume 2171, page 012039. IOP Publishing, 2022.
- [34] M .Shatha Moaz. *تقنيات الرؤية الحاسوبية* computer vision techniques, septembre 2023. consultée le 25 février 2023.
- [35] Jamuna S Murthy, GM Siddesh, Wen-Cheng Lai, BD Parameshachari, Sujata N Patil, and KL Hemalatha. Objectdetect : A real-time object detection framework for advanced driver assistant systems using yolov5. *Wireless Communications and Mobile Computing*, 2022, 2022.
- [36] William Puech. *Analyse et Traitement des images*. PhD thesis, Université Montpellier II - Nîmes.
- [37] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once : Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 779–788, 2016.
- [38] Mohamed Sandeli. Traitement d’images par des approches bio-inspirées application à la segmentation d’images. *Université Constantine*, 2 :s1, 2014.
- [39] Khalfallah Sara. Introduction à l’imagerie médicale. 2020.
- [40] Amel Sidahmed and Karima Rabhi. *La prédiction du diabete en utilisant les algorithmes de machine learning*. PhD thesis, Université Akli Mohand Oulhadje-Bouira, 2020.
- [41] Bahman Soheilian, Arlicot Aurore, and Nicolas Papanoditis. Extraction de panneaux de signalisation routière dans des images couleurs. 01 2010.
- [42] Sanghyun Son and Yunju Baek. Design and implementation of real-time vehicular camera for driver assistance and traffic congestion estimation. *Sensors*, 15(8) :20204–20231, 2015.

- [43] Khaled Taha. Python! إنشاء تطبيقات الهاتف المحمول في kivy python-, 2021. consultée le 23 mai 2023.
- [44] Do Thuan. Evolution of yolo algorithm and yolov5 : The state-of-the-art object detection algorithm. 2021.
- [45] Jordan Tunnell, Zachary D Asher, Sudeep Pasricha, and Thomas H Bradley. Toward improving vehicle fuel economy with adas. *SAE International Journal of Connected and Automated Vehicles*, 1(12-01-02-0005) :81–92, 2018.
- [46] MILOUDI WIEM. *Classification multispectrale des images satellites*. PhD thesis, Mémoire Master informatique ,Université Mohamed Khider-BIKSRA-Département de l’informatique, 2021-2022.
- [47] Renjie Xu, Haifeng Lin, Kangjie Lu, Lin Cao, and Yunfei Liu. A forest fire detection system based on ensemble learning. *Forests*, 12(2) :217, 2021.
- [48] yasmīna saadna. *method d’apprentissage*. PhD thesis, these de doctorat, Université batna 2.
- [49] zineb yazid and ouafaa yahi. Contours actifs paramétriques pour la segmentation d’images. 2017.
- [50] Syed Sahil Abbas Zaidi, Mohammad Samar Ansari, Asra Aslam, Nadia Kanwal, Mamoona Asghar, and Brian Lee. A survey of modern deep learning based object detection models. *Digital Signal Processing*, page 103514, 2022.
- [51] Yu Zhang, Zhongyin Guo, Jianqing Wu, Yuan Tian, Haotian Tang, and Xinming Guo. Real-time vehicle detection based on improved yolo v5. *Sustainability*, 14(19) :12274, 2022.