Ministère De L'enseignement Supérieur Et De La Recherche Scientifique
Université « Yahia Fares » De Médéa



Faculté Des Sciences
Département Mathématique Et Informatique

Projet de fin d'études

Présenté pour l'obtention du Diplôme de Master en Informatique

Option : ***Ingénierie des Systèmes et Technologie du Web***

Thème :

Détection des panneaux de signalisation routière

Dirigé par :
Dr. Soundes BELKACEM
Préparé et présenté par :
HANNACHI Abdelmalek
KHELIFI Imed

Année Universitaire 2019-2020

Remerciment

Nous aimerions d'abord remercier **ALLAh** le Tout-Puissant et notre Créateur de nous avoir donné santé, opportunités et conseils à chaque étape de notre vie et en particulier dans nos études jusqu'à ce moment où nous sommes à la fin de notre carrière d'études.

Nous continuerons à remercier nos chers professeurs de l'Université **Dr. Yaia Fares Médéa** qui sont toujours à notre service et ils nous aident à réaliser nos rêves malgré les difficultés et les défis auxquels ils sont confrontés dans leur travail.

En particulier, nous tenons à remercier notre cher professeur **Dr. Belkacem Soundes** qui a toujours été à notre disposition sans réserve lors de la préparation de cette thèse de fin d'études. Nous avons fait ce travail grâce à ses conseils, son aide et ses encouragements constants pour nous tout le temps 'Merci Infiniment'.

A la fin nous remercions tous ceux qui leur service nos a aidé à réaliser ce travail mais nous ne peux pas mentionner leurs noms car ils sont très nombreux.

Dedicace

"Je dédie ce mémoire à : Mes parents :

Ma mère, qui a œuvré pour ma réussite, de par son amour, son soutien, tous les sacrifices consentis et ses précieux conseils, pour toute son assistance et sa présence dans ma vie, reçois à travers ce travail aussi modeste soit-il, l'expression de mes sentiments et de mon éternelle gratitude.

Mon père, qui peut être fier et trouver ici le résultat de longues années de sacrifices et de privations pour m'aider à avancer dans la vie. Puisse Dieu faire en sorte que ce travail porte son fruit ; Merci pour les valeurs nobles, l'éducation et le soutient permanent venu de toi.

Mes frères et sœurs qui n'ont cessé d'être pour moi des exemples de persévérance, de courage et de générosité.

Ma promotrice Dr. Belkacem Soundes. Vraiment, elle est un exemple de courage et de patience.

Mon partenaire dans le projet KHELIFI Imed.

Mes amis avec qui je partage de grands moments de ma vie au fil du temps. "

HANNACHI Abdelmalek

Dedicace

"Je remercier Mon Dieu de me donner la capacité de finir le projet.
Je dédie ce modeste travail a :

- *Mes parents qui ont fait le tout pour moi.*
- *Mes frères.*
- *Ma famille.*
- *Mon promoteur BELKACEM soundes.*
- *Mon partenaire dans le projet HANNACHI Abdlemalek.*
- *Tous mes amis qui ont encouragé.*

"

KHELIFI Imed

Résumé

Les systèmes de détection et de reconnaissance des panneaux de signalisation sont considérés comme un secteur d'excellence dans de nombreux pays développés, mais ils restent un secteur de développement timide en Algérie.

En générale, le système de détection et de reconnaissance des panneaux de signalisation consiste de deux phases principales, détection des panneaux puis reconnaissance. L'objectif de ce projet est la mise en œuvre de tel système pour véhicules intelligents. Les panneaux présentent plusieurs dégradations, généralement relatives aux changements environnementaux ce qui rend la tâche de détection et reconnaissance complexe et difficile.

Dans ce travail, les deux phases du système ont été étudiées et des solutions ont été proposées. La détection se base sur le classificateur Haar Cascade en utilisent l'espace colorimétrique HSV pour la détection du panneau sur l'image. Puis, la reconnaissance via les architectures GoogLeNet, AlexNet et LeNet. Pour l'évaluation, deux bases de données ont été utilisés, l'étape de détection des panneaux de signalisation est menée sur la base de données GTSDB la bases de détection des panneaux de signalisation allemand, et l'étape de classification et de reconnaissance est menée sur la base de données GTSRB la base de reconnaissance des panneaux de signalisation allemand.

Une détection favorable et une reconnaissance précise des panneaux de signalisation sont obtenues grâce à une formation continue et à des tests du modèle. Pour la phase de détection, nous avons obtenu une précision de 100 % et un rappel de 79%, et pour la phase de reconnaissance nous avons obtenu une précision de 98%. L'expérimentation a prouvé que l'espace colorimétrique HSV est le plus efficace pour la phase de détection, et l'architecture GoogLeNet est la plus précise pour la phase de reconnaissance. Les résultats obtenus montrent l'amélioration de ce système, qui est d'une importance considérable pour réduire le taux d'accidents et améliorer la situation de la sécurité routière, offrant une bonne garantie technique pour le développement régulier de l'aide à la conduite intelligente des véhicules.

Mots clé : Détection des panneaux de signalisation ; Reconnaissance des panneaux de signalisation ; Reconnaissance des panneaux de limitation de vitesse ; Haar-Cascade ; GoogLeNet ; AlexNet ; LeNet.

Abstract

Traffic sign detection and recognition systems are considered a sector of excellence in many developed countries, but they remain a weak development sector in Algeria.

In general, the traffic sign detection and recognition system consist of two main phases, sign detection and then recognition. The objective of this project is the implementation of such a system for intelligent vehicles. Traffic signs present several degradations, generally relating to environmental changes, which makes the task of detection and recognition complex and difficult.

In this work, the two phases of the system were studied and solutions were proposed. Detection is based on the Haar Cascade classifier using the HSV color space for traffic sign detection in the image. Then, recognition via GoogLeNet, AlexNet and LeNet architectures. For the evaluation, two databases were used, the traffic sign detection step is carried out on the GTSDB database German Traffic Sign Detection Benchmarks, and the classification and recognition step are conducted on the GTSRB database the German Traffic Sign Recognition Benchmarks.

Favorable detection and accurate recognition of traffic signs is achieved through continuous training and model testing. For the detection phase, we obtained an accuracy of 100% and a recall of 79%, and for the recognition phase we obtained an accuracy of 98%. Experimentation has proven that the HSV color space is the most efficient for the detection phase, and the GoogLeNet architecture is the most accurate for the recognition phase. The results obtained show the improvement of this system, which is of considerable importance for reducing the accident rate and improving the road safety situation, offering a good technical guarantee for the regular development of driving assistance. intelligent vehicles.

Keywords: Traffic signs detection, Traffic signs Recognition, Speed limit signs Recognition, Haar-Cascade, GoogLeNet, AlexNet, LeNet.

Table des matières

Remerciment **i**

Résumé **i**

Liste des figures **iii**

Liste des tableaux **v**

1 Introduction générale **1**
1.1 Contexte et problématique 1
1.2 But et Objectifs 2
1.3 Structure du mémoire 2

2 Les systèmes de détections et de reconnaissance des panneaux de signalisation **5**
2.1 Introduction 5
2.2 Architecture générale de système 5
2.2.1 Acquisition 6
2.2.2 Detection 6
2.2.3 Reconaissance 7
2.2.4 Suivi temporel 7
2.2.5 Affichage 7
2.3 Les panneaux de signalisation 7
2.3.1 Classification des panneaux de signalisation 7
2.3.2 Les difficultés pour la détection et la reconnaissance des panneaux 9
2.4 Les bases de données existantes 10
2.5 Métriques d'évaluation 14
2.6 Utilité et Domaines d'application 15
2.6.1 Utilité 15
2.6.2 Domaines d'application 15
2.7 Conclusion 15

3 Méthodes de détection et de classification des panneaux 17
3.1 Introduction . . . 17
3.2 Les méthodes de détection des panneaux de signalisation . . . 17
3.2.1 Méthodes colorimétriques . . . 18
RGB . . . 18
HSV . . . 19
Niveau de gris . . . 20
YCbCr . . . 21
3.2.2 Méthodes géométriques . . . 21
3.2.3 Les méthodes basées sur l'apprentissage automatique . . . 22
Les méthodes supervisées . . . 22
Les méthodes non supervisées . . . 23
3.2.4 Le classificateur en cascade de Haar . . . 23
3.3 Les méthodes d'apprentissage profond pour la classification des panneaux de signalisation . . . 26
3.3.1 Réseaux de neurones convolutifs CNN . . . 26
3.3.2 Architectures CNN . . . 29
Architecture Lenet . . . 29
Architecture AlexNet . . . 30
Architecture GoogleNet . . . 31
3.4 Conclusion . . . 32

4 Approche Proposée 33
4.1 Introduction . . . 33
4.2 Architecture générale . . . 33
4.3 Déscription détaillée . . . 33
4.3.1 Prétraitement . . . 33
4.3.2 Détection . . . 34
4.3.3 Reconaissance . . . 35
L'architecture GoogLeNet . . . 37
L'architecture AlexNet . . . 37
L'architecture LeNet . . . 38
4.4 Conclusion . . . 38

5 Expérimentation et Evaluation 39
5.1 Introduction . . . 39
5.2 Présentation des outils . . . 39
5.2.1 Matériel . . . 39
5.2.2 Logiciel et environnement . . . 39
5.3 Les bases de données utilisées . . . 40
5.3.1 La base de données allemande GTSDB . . . 40
5.3.2 La base de données allemande GTSRB . . . 41
5.4 Paramètres de configuration . . . 41

La phase de reconnaissance 41
La phase de détection 42
5.5 Résultats et Discussions: Detection 42
5.5.1 La phase de detetion 42
5.5.2 Comparaison entre les resultats sur differents espaces de couleurs 43
5.6 Résultats et Discussions: Reconnaissance 45
5.6.1 Métriques d'évaluation 45
5.6.2 Performances de l'architecture LeNet 46
5.6.3 Performances de l'architecture AlexNet 46
5.6.4 Performances de l'architecture GoogLeNet 47
5.6.5 Comparaison entre les résultats des trois architectures. 50
5.7 Evaluation Globale 50
5.8 Conclusion 51

6 Conclusion et perspectives **53**

Bibliographie **54**

Liste des figures

2.1 Exemple de système DRPS dans une voiture 6
2.2 Architecture générale d'un système DRPS 6
2.3 Panneaux d'interdiction ou de restriction 8
2.4 Panneaux d'avertissement de danger 8
2.5 Panneaux des obligations 8
2.6 Panneaux de priorité 8
2.7 Panneaux de fin de l'interdiction ou de la restriction 9
2.8 Ensemble d'images montre des difficultés pour la detection et reconaissance des panneaux: (a) Flou de mouvement, (b) Lumière indésirable, (c) Décoloration, (d) et (e) Neige, (f) Occlusion partielle.[35] 10
2.9 Exempele de deux boîtes englobantes utilisées pour le calcul de IoU. . 14

3.1 Modèle de couleur RGB [23]. 18
3.2 Modèle de couleur HSV [5]. 19
3.3 Modèle de couleur niveaux de gris [21]. 21
3.4 Modèle de couleur YCbCr [5]. 22
3.5 Les étapes de la méthode du classificateur en cascade Haar 24
3.6 Une vue générale de CNN. 27
3.7 La différence entre Average Pooling et Max Pooling. 28
3.8 Deux fonctions d'activations, (a) ReLU et (b) SoftMax. 29
3.9 Architecture LeNet 29
3.10 Architecture AlexNet 30
3.11 Module d'Inception Naïf. 31
3.12 Module d'Inception avec réduction de dimension. 32
3.13 Architecture GoogleNet 32

4.1 Archiecture Générale du systeme 34
4.2 Un exemple de transformation de l'espace colorimétrique: **RGB** vers **HSV** 35
4.3 Un exemple des résultats de la phase de détection 35

5.1 Un exemple d'images de panneaux de limitation de vitesse de la base BTSDB 40
5.2 Exemples d'images de la base GTSRB 41
5.3 Evaluation des performances sur les quatre espaces de couleurs: **RGB**, **niveau de gris**, **HSV** et **YCbCr**. 43
5.4 Exemples de resultats de detection sur differents espaces de couleurs.Sur la première colonne les images originaux respectivement les résultats de détection sur les espaces de couleurs : **RGB**, **niveau de gris**, **HSV** et **YCbCr**. 44
5.5 Performances de l'architecture LeNet : Accuracy. 46
5.6 Performances de l'architecture LeNet : Fonction de perte. 46
5.7 Performances de l'architecture ALexNet : Accuracy. 47
5.8 Performances de l'architecture AlexNet : Fonction de perte. 47
5.9 Performances de l'architecture GoogleNet : Accuracy. 48
5.10 Performances de l'architecture GoogleNet : Fonction de perte. 48
5.11 Resultats de detection sur differents espaces de couleurs.Sur la première ligne les résultats obtenus via **AlexNet**, a deuxième les résultats obtenus via**leNet** et la troisième les résultats obtenus via **GoogLeNet**. . 49
5.12 Exemple de résultats de bout en bout. 51

Liste des tableaux

2.1 Informations sur des bases de données pour les panneaux de signalisation. 13

4.1 Les détails de l'architecture GoogLeNet 36
4.2 Les détails de l'architecture AlexNet 37
4.3 Les détails de l'architecture LeNet 38

5.1 Les détails des paramètres de configuration de la base de données GTSRB. 42
5.2 Les détails des paramètres de configuration de la base de données GTSDB. 42
5.3 Comparaison des résultats obtenus sur les quatre espaces couleurs. . . 43
5.4 Performances des trois architectures. 50
5.5 Evaluation des performances du système de bout en bout en utilisant l'architecture GoogleNet sur l'espace HSV. 50

Liste d'abréviations

IA: Intelligence artificielle
TSDR: Système de détection et de reconnaissance des panneaux de signalisation
GTSRB: La base de données de Benchmark Allemand de détection
GTSRB: La base de données de benchmark Allemand de reconnaissance
KUL BTS: La base de données des panneaux de signalisation KUL Belgique
BTSC: La base de données de classification des panneaux de signalisation de Belgique
STS: La base de données des panneaux de signalisation Suédois
CTS: La base de données des panneaux de signalisation Chinois
LISA TS: La base de données des panneaux de signalisation aux États-Unis
ITS: La base de données des panneaux de signalisation Italiens
CTSC: La base de données de classification des panneaux de signalisation du Croate
IOU: Intersection sur union
ROI: Région d'intérêt
ML: Apprentissage automatique
DL: Apprentissage en profondeur
CNNs: Les réseaux de neurones convolutifs **ANNs:** Réseaux de neurones artificiels
GPU: Unité de traitement
R-CNN: Réseaux de neurones convolutifs basés sur la région
Fast R-CNN: Réseaux de neurones convolutifs basés sur des régions rapides
Faster R-CNN: Réseaux de neurones convolutionnels basés sur des régions plus rapides
Mask R-CNN: Masque réseaux de neurones convolutionnels basés sur des régions masque
RPN: Réseau de proposition de région

Chapitre 1

Introduction générale

Avec le développement rapide de l'économie et de la technologie dans la société moderne, les automobiles sont devenues un moyen de transport indispensable dans les déplacements quotidiens des personnes. Bien que la popularité des automobiles ait introduit une commodité considérable pour les gens, elle a également causé de nombreux problèmes de sécurité routière qui ne peuvent être ignorés, tels que les embouteillages et les accidents de la route fréquents. Les problèmes de sécurité routière sont en grande partie causés par des raisons subjectives liées au conducteur, telles que l'inattention, une conduite inappropriée et le non-respect des règles de circulation [30].

1.1 Contexte et problématique

Les voitures intelligentes sont devenues un moyen efficace d'éliminer ces facteurs humains [28]. La technologie de conduite autonome peut aider, voire compléter indépendamment l'opération de conduite, ce qui est d'une importance remarquable pour libérer le corps humain et réduire considérablement l'incidence des accidents. Les voitures intelligentes peuvent obtenir des informations réelles et efficaces sur le trafic routier à partir d'une caméra montée sur le véhicule, ils peuvent également reconnaître et comprendre les panneaux de signalisation en temps réel dans des scènes de route réelles pour fournir une sortie de commande et un contrôle des mouvements corrects pour les véhicules intelligents ou pour alerter les conducteurs afin d'éviter d'éventuels dangers.

Les systèmes de détection et de localisation est composé de plusieurs étapes : la détection et la reconnaissance. La phase de détection couvre la localisation des régions d'intérêt ROI qui contiennent les panneaux et la détection des panneaux sur l'image. La reconnaissance identifie le type du panneau.

Les panneaux utilisent les couleurs pour attirer l'attention du conducteur et distinguer les différents types de panneaux. Dans le cas des panneaux de limitations de vitesse, la couleur rouge est utilisée pour marquer le panneau et le distinguer des

autres objets de la scène. La couleur représente les informations nécessaires pour détecter les panneaux.

Dans un environnement routier réel, la forme et la couleur des panneaux de signalisation ne seront pas toujours stables en raison de nombreux facteurs défavorables, tels que les variations de point de vue, les dommages physiques, le mauvais temps, etc. Cela rend la détection et la reconnaissance despanneaux de signalisation difficiles pour les humains et les systèmes de détections [35]. En particulier, il y a des panneaux de grande importance tels que les panneaux de limitation de vitesse qui ne peuvent être négligés, en raison de l'importance de leurs informations. Manque d'attention à ces panneaux et non-respect de leurs limites de vitesse met la vie de conducteur et leurs occupants en danger[22].

Pour mettre en œuvre un système de détection et de reconnaissance des panneaux de signalisation, il existe plusieurs méthodes, telles que : les méthodes basées sur la couleur, la forme et l'apprentissage automatique [27].

Dans ce projet, nous nous intéressons aux problématiques a la détection des panneaux de signalisation routière sur des images à scène naturelle ? Et une fois détecter, comment reconnaitre a qu'elle classe appartient-il ?

1.2 But et Objectifs

L'objectif de ce travail est de proposer des méthodes pour la détection et classification des panneaux de signalisation routiers, particulièrement les panneaux de limitation de vitesse et reconnaître leurs significations.

Afin d'atteindre notre objectif principal, d'autres sous objectifs spécifiques ont étés spécifiés :

- Présenter un état de l'art des méthodes existantes de détection et de reconnaissance d'objets ;
- Proposée un approche pour la détection et reconnaissances des panneaux routiers en se basent sur un choix justifier des méthodes efficaces de détection et de la reconnaissance d'objets ;
- Implémenter, présenter et discuter les résultats de l'approche proposée.

1.3 Structure du mémoire

Ce mémoire s'organise en quatre chapitres : Dans *le premier chapitre*, nous présenterons un aperçu des systèmes de détection et de reconnaissance des panneaux de signalisation, leur utilisation et leur architecture en général. ; Puis, dans *le deuxième chapitre*, nous verrons les différentes méthodes et techniques utilisées pour mettre en œuvre un système de détection et de reconnaissance des panneaux de signalisation ; Ensuite, dans *le troisième chapitre* nous allons présenter l'architecture du système

de détection et de reconnaissance des panneaux de signalisation que nous avons proposé, ainsi que les méthodes et techniques utilisées. Dans *le quatrième chapitre*, l'implémentation des méthodes, les détails matériels, logiciels, paramètres de configuration et description des bases de données que nous avons utilisées. Aussi, nous présenterons en détails les résultats obtenus avec discutions. Enfin, nous *conclurons* notre ce mémoire, par une synthèse des résultats et mise en perspectives du travail réalisait au cours de ce projet.

Chapitre 2

Les systèmes de détections et de reconnaissance des panneaux de signalisation

2.1 Introduction

Les systèmes de détection et de reconnaissance des panneaux de signalisation (Traffic Signs Detection and Recognition system TSDR en anglais) joue un rôle très important dans la conduite, ils aident les conducteurs en fournissant plus d'informations sur les panneaux de signalisation qui apparaissent sur la route pour protéger les occupants du véhicule.

Dans ce chapitre, nous présentons ce qu'est un système TSDR, pourquoi nous l'utilisons, les différents types de ce système et nous verrons également son architecture générale et les différentes phases qui le constituent. Nous présentons les différents types de panneaux de signalisation et les bases de données existantes.

2.2 Architecture générale de système

Un systeme de détection et de reconnaissance des panneaux de signalisation: est un système qui détecte et reconnaître les panneaux de signalisation[25]. Il prend comme une entrée une séquence d'images provenant d'une caméra multifonctions positionnée contre le pare-brise du véhicule, ces images seront ensuite traitées par une application spécialisée dans la détection et la reconnaissance des panneaux de signalisation. Enfin, ce système informer le conducteur au cas où il n'aurait pu les voir. La figure suivante explique le fonctionnement de ce système.

(a) Où installer la caméra, (b) Détecter et reconnaître le panneau de signalisation et (c) Affichage du panneau.

Le système TSDR est basé sur deux étapes principales : la détection et la reconnaissance (aussi appelée classification). La première étape localise les panneaux

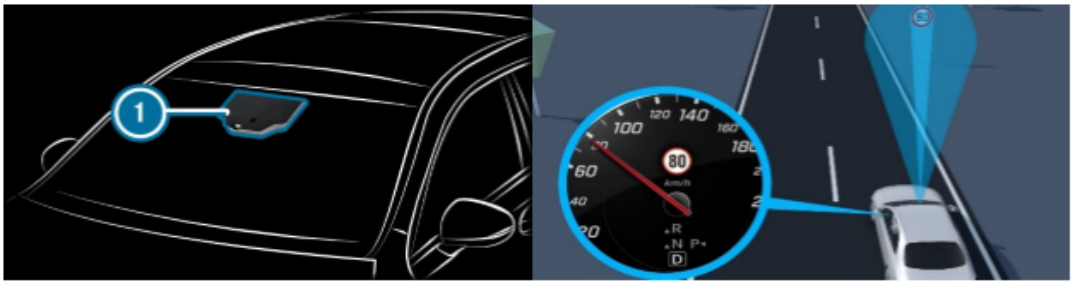


FIGURE 2.1: Exemple de système DRPS dans une voiture

dans les images en se basant sur leurs caractéristiques visuelles, comme la couleur et la forme. La deuxième étape reconnait le type de chaque panneau détecté. Il est possible d'ajouter une troisième étape qui est le suivi temporel afin de ne pas resignaler au conducteur les mêmes panneaux reconnus.

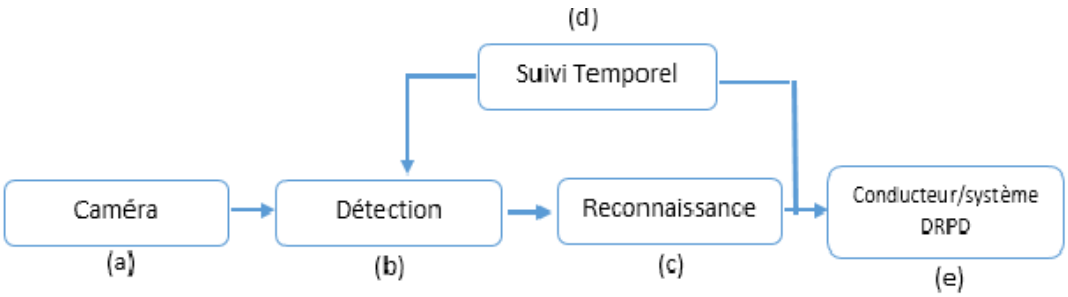


FIGURE 2.2: Architecture générale d'un système DRPS

2.2.1 Acquisition

Une caméra multifonctions qui positionnée contre le pare-brise du véhicule transmet une séquence d'images pour faire l'étape suivante qu'est la détection

2.2.2 Detection

Elle est effectuée pour identifier les régions d'intérêt (Regions of Interest en anglais ROIs), Cette étape exploite les caractéristiques visuelles des panneaux, comme la couleur et la forme, ou-bien en utilisant des techniques d'apprentissage automatique. La détection doit prendre en considération différents facteurs comme la diversité des catégories de panneaux (panneaux d'interdiction, panneaux de danger, panneaux d'obligation, panneaux de priorité et panneaux de la restriction). Les ROIs détectés sont ensuite soit reconnus ou rejetés pendant la phase de reconnaissance.

2.2.3 Reconaissance

Elle est réalisée sur les ROIs, où certains ROIs peuvent ne contenir pas des panneaux de signalisation. Cette étape est réalisée avec certaines techniques d'apprentissage automatique telles que le réseau neuronal convolutif (Convolutional Neural Network en anglais CNN).

2.2.4 Suivi temporel

Il identifie les panneaux déjà reconnus afin de ne pas les reclassifies ce qui réduit le temps de traitement. Cette identification est assurée par une étape d'association de données qui met en relation les panneaux reconnus précédemment avec les panneaux détectés à l'instant courant. Cette association permet également de ne pas resignaler les mêmes panneaux au conducteur afin de ne pas le perturber[15]. À l'aide d'une étape de filtrage, le suivi temporel peut également procurer un a priori sur les positions et les résolutions des panneaux dans les images suivantes facilitant ainsi leur détection[20].

2.2.5 Affichage

Le rôle de cette dernière étape est d'informer le conducteur en affichant le panneau ainsi reconnu sur le tableau de bord. La figure suivante représente les étapes du système DRPS

2.3 Les panneaux de signalisation

Les panneaux de signalisation constituent un élément fondamental de tout système de circulation, ils sont généralement implantés de part et d'autre sur nos routes, ils indiquent les règles de la circulation établies pour permettre aux véhicules de se déplacer en toute sécurité sur les routes [3].

Les usagers de la route doivent d'être bien sensé des panneaux de signalisation qui demeure aujourd'hui un des éléments phares de la prévention routière. Grâce aux panneaux de signalisation les conducteurs sont informés des règles d'avertissements sur les dangers pouvant apparaître sur la route ou tout type d'informations intéressant le conducteur.

2.3.1 Classification des panneaux de signalisation

Les panneaux de signalisation sont classés en différents types, nous avons présenté les types les plus importants :

1. 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. Ces signes sont circulaires et avec un bord rouge[2, 25]. Comme montre la figure suivante



FIGURE 2.3: Panneaux d'interdiction ou de restriction

2. Panneaux d'avertissement de danger Les signes de danger ont pour mission d'indiquer la nature d'un danger, leur objectif se conformer aux règles de comportement et d'éviter les chocs éventuels lors de la conduite. Sa forme est triangulaire avec un bord rouge, comme indique la figure suivante [2, 25]. Comme montre la figure suivante



FIGURE 2.4: Panneaux d'avertissement de danger

3. Panneaux des obligations Ils sont destinés à obliger les usagers à respecter certaines prescriptions et certaines règles à partir du niveau du panneau ou devant celui-ci. Ces panneaux peuvent donc indiquer non seulement une direction que l'usager devra suivre, mais aussi une vitesse minimum à respecter, ou encore une voir obligatoire pour certains types d'usagers. Ils sont ronds avec un couleur bleu. Comme montre la figure suivante:



FIGURE 2.5: Panneaux des obligations

4. Panneaux de priorité Ce 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[2]. Comme montre la figure suivante



FIGURE 2.6: Panneaux de priorité

5. Panneaux de fin de l'interdiction ou de la restriction

Ces panneaux de fin d'interdiction ou de la restriction signalent que l'interdiction ou de la limitation est terminée. Ils sont ronds et blancs avec une ligne diagonale noire. Comme montre la figure suivante



FIGURE 2.7: Panneaux de fin de l'interdiction ou de la restriction

2.3.2 Les difficultés pour la détection et la reconnaissance des panneaux

Dans un environnement routier réel, la forme et la couleur des panneaux de signalisation ne seront pas toujours stables en raison de nombreux facteurs défavorables, tels que les variations de point de vue, les dommages physiques, le mauvais temps, etc. Cela rend la détection et la reconnaissance des panneaux de signalisation difficiles pour les humains et les systèmes informatiques [35]. Les défis rencontrés sont les suivantes :

1. Mouvement Les images des scènes de circulation sont souvent sujettes au flou de mouvement car les images sont prises à partir d'une caméra de véhicule à grande vitesse.

2.Décoloration Les images peuvent être affectées par la pluie, la neige, le brouillard et d'autres facteurs, ce qui entraîne un flou des images et une distorsion des couleurs.

3. Occlusion partielle Certains panneaux de signalisation sont souvent masqués par d'autres objets sur la route, tels que des arbres, des piétons ou d'autres véhicules. Par conséquent, il est nécessaire de détecter les panneaux de signalisation avec seulement une partie des informations d'image de panneau de signalisation.

4. Arrière-plan complexe Des panneaux qui ressemblent à des panneaux de signalisation et génèrent du bruit lors de la détection des panneaux de signalisation, tels que des pancartes commerciales et des panneaux d'affichage du bien-être public.

5. Déformation des panneaux Les caméras montées sur véhicule ne sont pas toujours perpendiculaires aux panneaux de signalisation, la forme des panneaux de signalisation est généralement déformée dans les images capturées et la forme des panneaux de signalisation dans les scènes de circulation n'est pas toujours fiable.

6. Les panneaux de signalisation endommagés Parfois, les panneaux de signalisation sont partiellement cassés ou bien endommagés, cela rend la détection et la reconnaissance plus difficiles.

La figure suivante montre un ensemble d'images sous défis :



FIGURE 2.8: Ensemble d'images montre des difficultés pour la detection et reconaissance des panneaux: (a) Flou de mouvement, (b) Lumière indésirable, (c) Décoloration, (d) et (e) Neige, (f) Occlusion partielle.[35]

Il est clair que chaque type des panneaux de signalisation a sont role et peut servir a avertir ou informer le conducteur afin d'éviter les accidents et les problemes de circulations. Mais,certains des autres signes peuvent être négligés, du fait que le risque de leurs absence n'est pas assez dangereux.

Cependant, ce n'est pas le cas des panneaux de limitation de vitesse. En effet, si le conducteur ne fait pas attention à un panneau, il peut mettre sa vie ainsi que les autres en question. C'est pour cela que dans notre travail on se concentre sur ce type spécifique du panneau.

2.4 Les bases de données existantes

La base de données de benchmark Allemand de reconnaissance : (German Traffic Signs Recognition Benchmark - GTSRB dataset) Le Benchmark allemand de reconnaissance des panneaux de signalisation - GTSRB dataset[9, 11] est un défi de classification multi-classe à image unique organisé lors de la Conférence conjointe internationale sur les réseaux de neurones (IJCNN) 2011.

Cette base de données publiée en 2011 par l'université de la Ruhr Bochum, Allemagne, de la part de "Johannes Stallkamp, Marc Schlipsing et Jan Salmen". Cette base de données est très grande avec plus de 50 000 images de panneaux de signalisation dans 43 classes avec une taille de 300 Mo au total, elle est disponible sur ce site officielle "INI Benchmark Website" (INI BW) [1] .

La base de données de Benchmark Allemand de détection :(German Traffic Signs Detection Benchmark - GTSDB dataset) Le Benchmark allemand de détection des panneaux de signalisation - GTSDB dataset [12] est une évaluation unique de la détection d'images pour les chercheurs intéressés par le domaine de la vision par ordinateur, la reconnaissance des formes et l'aide à la conduite assistée par image. Il a été présenté à la conférence internationale conjointe de l'IEEE 2013 sur les réseaux

[1] http://benchmark.ini.rub.de/?section=gtsrbsubsection=news

de neurones. Cette base de données publiée en 2013 de la part de "Houben.S, Stallkamp.J, Salmen.J, Schlipsing.M et Igel.C". Cette base de données contient 900 images de panneaux de signalisation dans 43 classes avec une taille plus de 1.5 Go, elle est disponible sur ce site officielle “INI Benchmark Website” (INI BW) .

La base de données de Belgique KUL Belgique (KUL Belgium Traffic Signs - KUL BTS dataset) La base de données des panneaux de signalisation KUL Belgique KUL BTS dataset[12, 13] est très grande avec plus de 10000 annotations de panneaux de signalisation, des milliers de panneaux de signalisation physiquement distincts. Cette base de données est disponible sur ce site officiel “Traffic Signs Recognition” (TSR) , avec une taille plus de 50 Gb au total.

La base de données de classification de la Belgique : (Blegium Traffic Signs for Classification - BTSC dataset) La base de données de classification des panneaux de signalisation de Belgique BTSC dataset [14] est construit à des fins de classification des panneaux de signalisation, il contient des images recadrées autour des annotations pour les différentes classes de panneaux de signalisation. Cette base de données est disponible sur ce site officiel “Traffic Signs Recognition” (TSR) , avec une taille de 248 Gb au total.

La base de données de la Suède : (Swedish Traffic Signs - STS dataset) La base de données des panneaux de signalisation Suédois STS dataset [8] a été créé en enregistrant des séquences de plus de 350 km d’autoroutes et de routes urbaines suédoises. Cette base de données est disponible sur ce site officiel “Computer Vision Laboratory” (CVL) .

La base de données de la Chine :(Chinese Traffic Signs - CTS dataset) La base de données des panneaux de signalisation Chinois CTS dataset [20] contient des images capturées dans différentes conditions météorologiques (ensoleillées, nuageuses et pluvieuses) et des scénarios (autoroute, urbain, rural) sur les routes de Pékin et de Xiamen, en Chine. Cette base de données est disponible sur ce site officiel “Chinese Traffic Signs Dataset” (CTSD) .

La base de données des États-Unis :(LISA Traffic Signs - LISA TS dataset) La base de données des panneaux de signalisation aux États-Unis LISA TS dataset [11] est un ensemble de vidéos et de cadres annotés contenant des panneaux de signalisation américains. Il est publié en deux étapes, l’une avec uniquement les images et l’autre avec les images et les vidéos. Cette base de données est disponible sur ce site officiel “LISA Traffic Sign Dataset” (LISA TSD) , avec une taille de 7,7 Gb au total.

La base de données de détection de l’Italie (Italian Traffic Signs - ITS dataset) La base de données des panneaux de signalisation Italiens ITS dataset[**16**, 16] a été généré à l’origine à partir de 43289 images extraites de 14h de vidéos (1280 x720 à 10 fps) capturées en Italie à la nuit. Il est composé des images d’apprentissage et de test, il contient également un fichier texte avec des annotations, et trois super-classes basées sur la forme sont définies : Interdit, indication et avertissement. Cette base de données est disponible sur ce site officiel “Dataset of Italian Traffic Signs”. **La base de données de classification de la Croatie (rMASTIF) :** (Croatian Traffic Signs

Classification - CTSC dataset) La base de données de classification des panneaux de signalisation du Croate CTSC dataset [**18**] a été préparé en révisant la base de données collectés lors du projet de recherche MASTIF (Mapping and Assessing the State of Traffic InFrastructure). Cette base peut être librement utilisé dans la recherche universitaire. Cette base de données est disponible sur ce site officiel rMASTIF datasets (rM D). Le tableau suivant montre plus de détails 2.4

TABLE 2.1: Informations sur des bases de données pour les panneaux de signalisation.

Dataset	**Nbrclasses**	**Nbr d'images**	**Nbr d'annotations**	**Taille d'images**	**Taille du signe**	**Format d'image**	**Pays d'origine**
GTSRB	43	+50.000	+50.000	15x15 a 250x250 px	15x15 a 250x250 px	PNG	Allemand
GTSDB	43	900	910	1360 x800px	de 16x16a128x128px	PPM	Allemand
KUL BTS	+100	9.006	13.444	1628x1236px	100x100 a1628x1628	PPM	Belgique
BTSC	62	7.125	7.125	784x784 px	16x16 a 128x128 px	PPM	Belgique
STS	7	20.000	3.488	1280x960 px	3x5 a 263x248 px	/	Suède
CTS	48	1.100	1.574	1224x768a1270x800 px	20x20a573x557 px	/	Chine
LISA TS	49	6.610	7.855	640x480a1024x522px	6x6a167x168 px	/	États Unis

2.5 Métriques d'évaluation

Les performances d'un système TSDR sont mesurées à l'aide des deux métriques: **rappel** et **précision** qui sont calculées avec les formules ci-dessous :

$$Rappel = TP/(TP + FN) \tag{2.1}$$

$$prcision = TP/(TP + FP) \tag{2.2}$$

Tel que :

- **TP**: le nombre de vrais positifs
- **FP**: le nombre de faux positifs
- **FN**: le nombre de faux négatifs

intersection sur union (IoU) : est calculer a partir de l'intersection sur l'union des deux boîtes englobantes : la boîte englobante de la vérité terrain (Ground Truth) et c'elle prédite par le systeme. Lorsque l'IoU est égal à 1, cela signifie que les boîtes



FIGURE 2.9: Exempele de deux boîtes englobantes utilisées pour le calcul de IoU.

englobantes de la vérité prédite et du terrain se chevauchent parfaitement. Il faut donc définir une valeur seuil pour l'IoU afin de déterminer si la détection d'objet est valide ou non, dans le cas général, la valeur de l'IoU est affectée à 0.5, dans ce cas :

- Lorsqu'il y a un panneau de signalisation dans l'image et que le modèle le détecte et (IoU>=0,5) alors c'est une bonne détection et le classer comme vrai positif (VP).

- Lorsqu'il y a un panneau de signalisation dans l'image et que le modèle le détecte et (IoU<0,5), alors c'est une mauvaise détection et classez-le comme vrai positif (FP).
- Lorsqu'il y a un panneau de signalisation dans l'image et que le modèle ne l'a pas détecté, il est classé comme un faux négatif (FN).
- Vrai négatif (TN) est chaque partie de l'image où nous n'avons pas prédit un objet. Cette métrique n'est pas utile pour la détection d'objets, c'est pourquoi nous ignorons.

2.6 Utilité et Domaines d'application

2.6.1 Utilité

Les systèmes de detction et reconnaissance despanneaux de signalisation présentent plusieurs utilités et avantages :

- Fournir aux conducteurs des informations précieuses, afin de rendre la conduite plus sûre et plus facile
- Éviter l'apparition d'une situation dangereuse risquant d'aboutir à un accident
- Libérer le conducteur d'un certain nombre de taches qui pourraient atténuer sa vigilance

2.6.2 Domaines d'application

Il existe deux types d'applications de la technologie des systèmes DRPS.

- **Application passive:** Consiste à informer le conducteur par le biais de pictogrammes ou des sons qu'il entre dans une zone qui a une nouvelle limitation indiquée pour un panneau de signalisation. Dans ce cas, ce sera le conducteur qui prendra la décision de respecter le signal ou non
- **Application active:** Consiste à intervenir automatiquement sur la voiture quand il détecte un panneau de signalisation. Par exemple, si le conducteur conduit à une vitesse excessive et le système DRPS détecte un signal d'arrêt et la voiture n'interprète pas que le conducteur a l'intention de s'arrêter, l'ordre de freiner sera directement dirigé vers la voiture pour éviter un éventuel accident.

2.7 Conclusion

Dans ce chapitre, nous avons vu en générale les systèmes de détection et de reconnaissances des panneaux de signalisation. En plus, nous avons présentés les panneaux de signalisations, leurs caractéristiques et défis que présentent.

Pour cela dans le deuxième chapitre nous présenterons des techniques et méthodes qui permettent de réaliser ces systèmes : les techniques de détection et de reconnaissance.

Chapitre 3

Méthodes de détection et de classification des panneaux

3.1 Introduction

Les systèmes de détections et de localisation est composé de plusieurs étapes : la détection et la reconnaissance. La phase de détection couvre la localisation des régions d'intérêt ROIs qui contiennent les panneaux et la détection des panneaux sur l'image. La reconnaissance identifie le type du panneau.

La phase de détection doit trouver tous et que les panneaux figurants sur l'image. Ainsi en doit trouvez maximum de panneaux et minimum d'autres objets considérer comme arrière-plan. Les méthodes de détection sont souvent groupées en trois catégories : les méthodes basées sur les couleurs, les méthodes basées sur les formes géométriques et les méthodes basées sur l'apprentissage automatique. La phase de reconnaissance ou bien classification reconnaît le type de panneau détecté et l'assigne à une classe. Les méthodes de classification sont groupées en deux catégories : les méthodes qui utilisent des caractéristiques extraites par un expert et les méthodes de l'apprentissage profond (Deep Learning).

Dans ce chapitre, nous allons présenter les différentes méthodes et technique de détection et de reconnaissance des panneaux de signalisation a base d'apprentissage approffondis.

3.2 Les méthodes de détection des panneaux de signalisation

Il existe de nombreuses méthodes de détection des panneaux de signalisation, qui sont généralement groupées en méthodes basées sur la couleur, sur la forme et sur l'apprentissage [36].

3.2.1 Méthodes colorimétriques

Les méthodes coloimétriques effectuent une segmantation basée sur la couleur dominante pour détecter les régions d'intérêts IOUs dans une image.

La couleure des panneaux de signalisation est une indication et porte une information tres importantes. Pour chaque types il y'a une couleur caractéristique spécifique : rouge, bleu et jaune, ces caractéristiques indiquent une sensibilité à divers facteurs.Neanmoins, elle présente une sensibilité à divers facteurs tels que : la variation de la lumière et l'âge des signes, cela rend la segmentation un processus très difficile

Les méthodes de détection travaillent sur différents espaces colorimétriques parmi lesquels :

RGB

Est l'espace colorimétrique le plus élémentaire pour les vidéos et les images capturées par les caméras.

L'espace colorimétrique RVB se compose de trois canaux, Rouge (R), Vert (V) et Bleu (B), qui sont additionnés de différentes manières pour reproduire un large éventail de couleurs. Cet espace colorimétrique est couramment utilisé dans les images numériques. Il est visualisé sous la forme d'un cube 3D où R, V et B sont les trois axes perpendiculaires, comme illustré à la figure suivante:

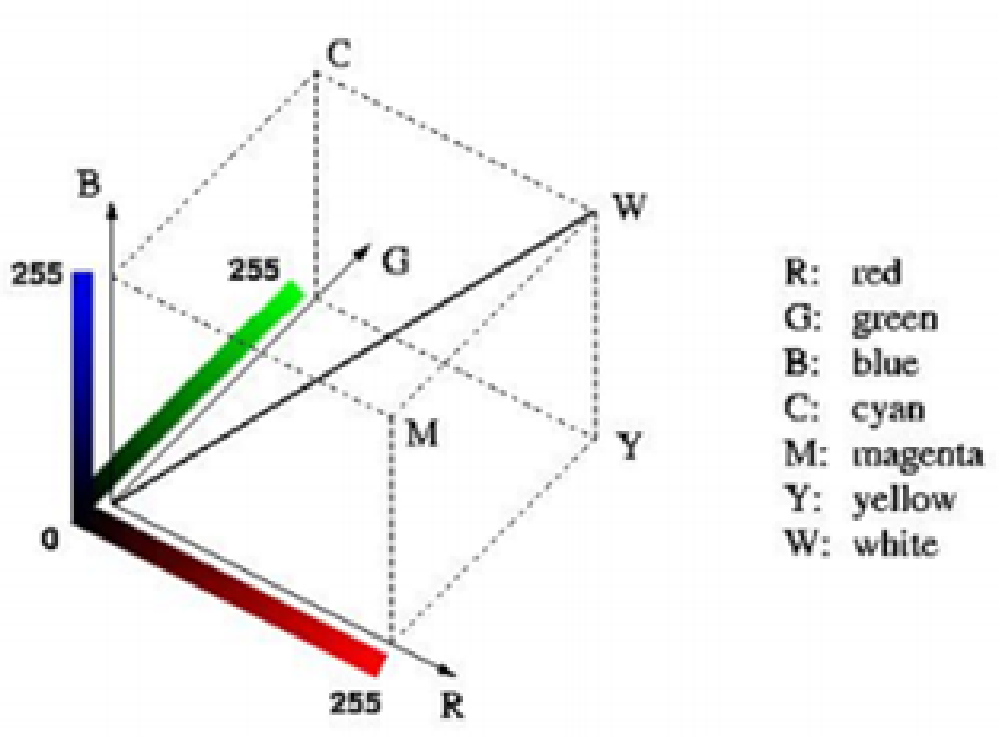


FIGURE 3.1: Modèle de couleur RGB [23].

Bien que RGB puisse être utilisé sans transformation, les canaux R, V et B ont une corrélation élevée et sont sensibles aux changements d'éclairage[34].

HSV

L'espace colorimétrique **HSV** représente la couleur sous la forme d'un triplé : teinte H (Hue), saturation S (Saturation) et luminosité V (Value).

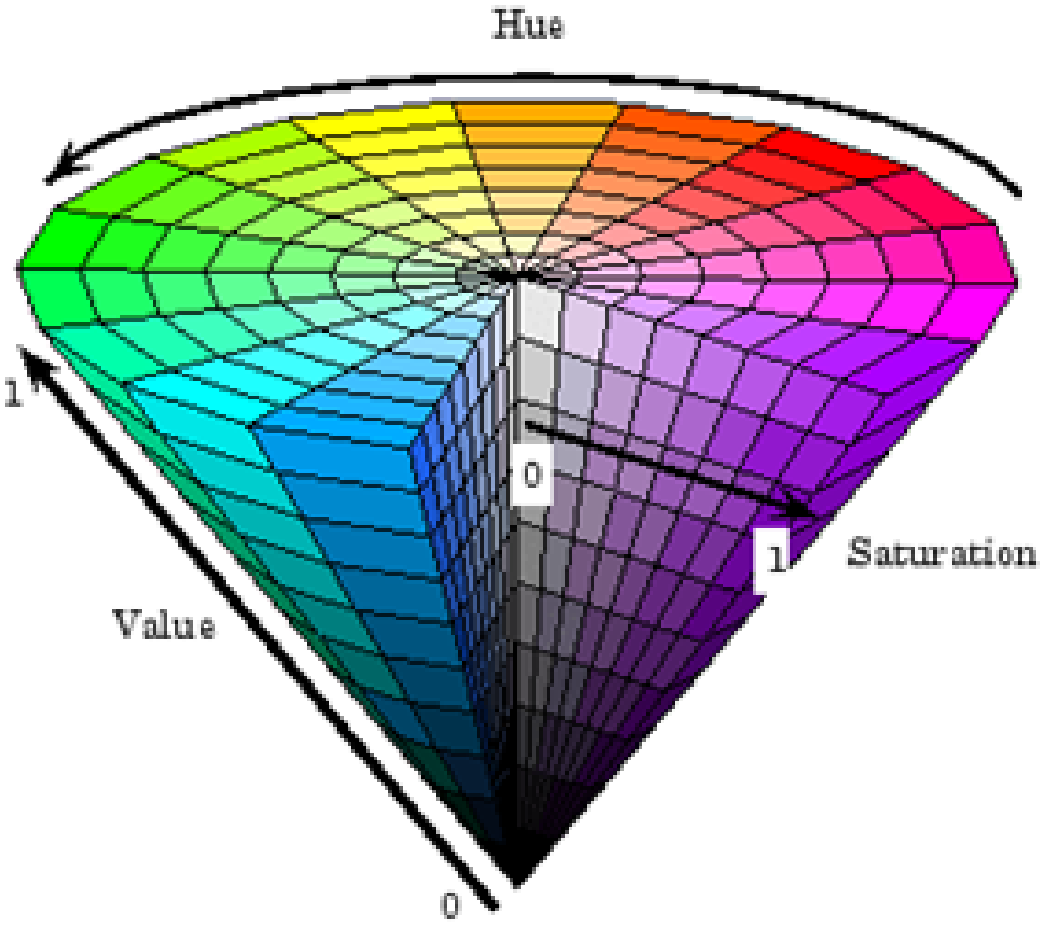


FIGURE 3.2: Modèle de couleur HSV [5].

- La valeur représente l'intensité d'une couleur, qui est découplée des informations de couleur dans l'image représentée.
- Les composantes de teinte (Hue) et de saturation sont intimement liées à la façon dont l'œil humain perçoit les couleurs.
- La teinte varie de 0 à 1.0, les couleurs correspondantes varient du rouge au jaune, le vert, le cyan, le bleu et le magenta de sorte qu'il y ait réellement des valeurs rouges à la fois à 0 et à 1.0.
- La saturation varie de 0 à 1.0, les couleurs correspondantes (Hue) varient de non saturé (nuances de gris) à totalement saturé (pas de composant blanc).

Les méthodes basées sur l'espace HSV sont partiellement insensibles aux changements d'éclairage, il est plus immunisé contre les changements d'éclairage que RVB.

L'espace HSV peut être calculé à l'aide de l'espace RVB, ce qui augmente le temps de traitement[24].

Yakimov [19] a estimé qu'il n'est pas possible de détecter les panneaux de signalisation en images réelles en appliquant un seuil simple directement dans l'espace RVB en raison des variations d'éclairage, c'est ce qui lui a poussés à choisir l'espace HSV. Wang et son équipe [17] choisissent également l'espace HSV, et ils ont constaté que la méthode classique de seuillage donne de bons résultats dans de nombreuses conditions d'éclairage différentes.

RVB à HSV est une transformation non-linéaire, mais réversible, l'équation suivante montre cette transformation [26]:

$$R' = R/255 \tag{3.1}$$

$$G' = G/255 \tag{3.2}$$

$$B' = B/255 \tag{3.3}$$

$$Cmax = max(R', G', B') \tag{3.4}$$

$$Cmin = min(R', G', B') \tag{3.5}$$

$$\Delta = Cmax - Cmin \tag{3.6}$$

$$H = \begin{cases} 0x & \Delta = 0; \\ 60 \circ x((G' - B')/\Delta) mod 6 & \Delta = R'; \\ 60 \circ x((B' - R')/\Delta) mod 6 & \Delta = G'; \\ 60 \circ x((R' - G')/\Delta) mod 6 & \Delta = B'. \end{cases}$$

(3.7)

$$S = \begin{cases} 0 & \text{Cmax=0;} \\ \Delta/Cmax & \text{Cmax}\neq 0. \end{cases} \tag{3.8}$$

Niveau de gris

L'espace de couleur niveaux de gris se compose d'un seul canal, il utilise différentes nuances de gris pour représenter une image. Chaque pixel d'une image en niveaux de gris a une valeur de luminosité comprise entre 0 (noir) et 255 (blanc). Les valeurs de niveaux de gris peuvent également être mesurées sous forme de pourcentages de couverture d'encre noire (0% est égal au blanc, 100% au noir)[21]. La conversion de couleur niveaux de gris effectue une opération de réduction de dimension en transformant généralement les valeurs trichromatiques RVB de R3 en valeurs d'intensité de R1. Les méthodes de conversion couleur-niveaux de gris effectuent une opération de réduction de dimension en transformant généralement les valeurs

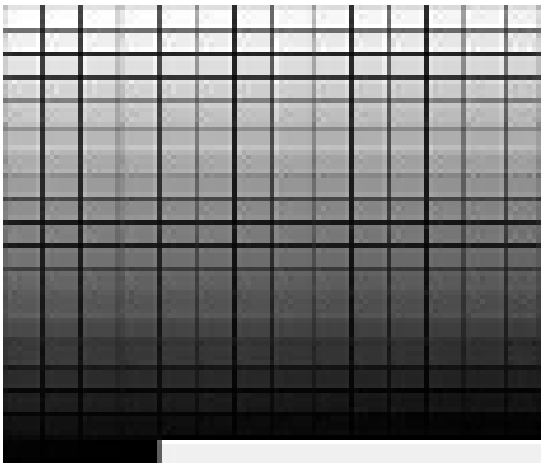


FIGURE 3.3: Modèle de couleur niveaux de gris [21].

trichromatiques RVB (R, V, B) R3 en valeurs d'intensité (I) R1. Il existe deux équations de conversion pour l'espace niveaux de gris : méthode moyenne, méthode pondérée,calculées par les equations suivantes: La méthode moyenne:

(3.9)

I= 0.33 *R +0.33 *B +0.33 *G La méthode pondérée :

$$I = 0.299 * R + 0.587 * B + 0.114 * G \tag{3.10}$$

YCbCr

L'espace colorimétrique YCrCb comporte deux éléments, la luminance Ŷ) qui représente l'intensité lumineuse et la chrominance (Cb et Cr) qui représente les composantes de chrominance de différence de bleu et de différence de rouge. La représentation de l'espace colorimétrique YCbCr peut être vue sur la figure 3.4:

La transformation de RGB en YCbCr est formulée comme l'équation suivante :

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} \begin{bmatrix} 65.481 & 128.553 & 24.966 \\ 37.797 & 74.203 & 112.000 \\ 112.000 & 93.786 & 18.214 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{3.11}$$

3.2.2 Méthodes géométriques

Il s'agit des méthodes populaires utilisées pour la détection des panneaux de signalisations. En effet, les formes standard courantes des panneaux de signalisation sont:

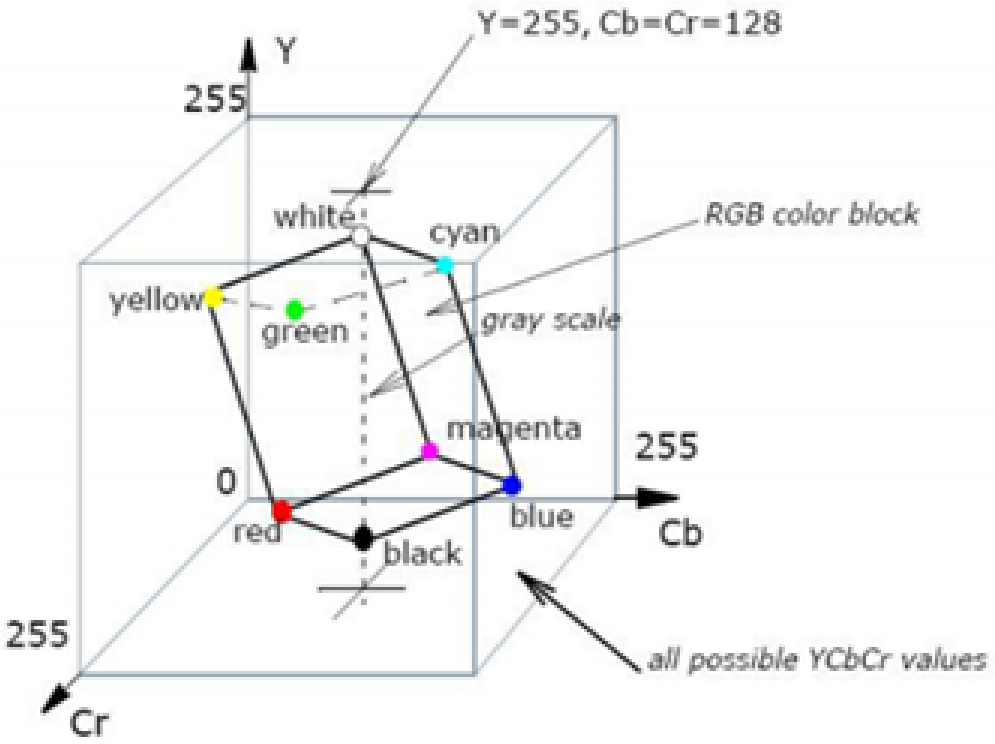


FIGURE 3.4: Modèle de couleur YCbCr [5].

le triangle, le cercle, le rectangle et l'octogone [34]. Les caractéristiques de forme utilisées pour la détection comprennent : les formes standard, les contours, la texture, etc.

Les algorithmes de détection basés sur la forme sont largement étudiés. Par exemple la transformation de Hough qu'est appliquée dans [7] pour détecter les signes triangulaires, rectangulaires et fléchés. Une autre méthode comme le vote à symétrie radiale [4] utilise la caractéristique de l'apparence des panneaux de signalisation.

Dans les scènes complexes, les méthodes basées sur la couleur ou bien sur la forme sont sensibles aux changements d'éclairage, aux ombres et aux différentes conditions d'acquisitions.

3.2.3 Les méthodes basées sur l'apprentissage automatique

Les méthodes supervisées

L'apprentissage supervisé est le paradigme d'apprentissage le plus populaire en ML et en DL et aussi le plus utilisé, l'apprentissage supervisé a besoin des donnés de la tâche qu'elle doit réaliser[32, 33]. Avec l'apprentissage supervisé, la machine peut apprendre à faire une certaine tâche via l'étude des données de cette tâche ; Par exemple la détection des panneaux après avoir vu des millions d'exemples des panneaux

Les méthodes non supervisées

Contrairement à l'apprentissage supervisé, L'apprentissage non supervisé est un système de ML qui apprend et crée des modèles à partir d'un ensemble de données d'entraînement non étiquetées. Il s'agit d'un type d'apprentissage hebbien auto-organisé qui aide à trouver des modèles auparavant inconnus dans un ensemble de données sans étiquettes préexistantes. Il est également connu sous le nom d'auto-organisation et permet de modéliser les densités de probabilité d'entrées données. Les méthodes non supervisées sont utilisées dans des domaines très variés, comme la sociologie, le marketing[6].

3.2.4 Le classificateur en cascade de Haar

Algorithme de Viola et Jones, connue aussi par Le classificateur en cascade de Haar est l'une des premières et des plus connues des méthodes de détection d'objet. Publié en 2001 par Paul Viola et Michael Jones [1] est devenu une référence dans le domaine de la détection d'objet. Il appartient aux méthodes de classification supervisées.LE classificaeur est trainé sur un ensemble de donnée ensuite utilisé pour la detection d'objet sur l'image. Inventée à l'origine pour détecter des visages, elle peut également être utilisée pour détecter d'autres types d'objets comme des panneaux de signalisations. Les principaux avantages de ce classificateur est le fait qu'il permet une detection efficace, rapide et dans n'importe la position ou la taille de l'objet sur l'image.

Principe : La méthode de Viola et Jones consiste à balayer une image à l'aide d'une fenêtre de détection de taille initiale 24x24px (Article Original) et de déterminer la présence d'unobjet. Lorsque l'image a été parcourue entièrement, la taille de la fenêtre est augmentée et le balayage recommence, jusqu'à ce que la fenêtre fasse la taillede l'image. L'augmentation de la taille de la fenêtre se fait par un facteur multiplicatif de 1.25. Le balayage, quant à lui, consiste simplement à décaler la fenêtre d'un pixel. Ce décalage peut être changé afin d'accélérer le processus, mais un décalage d'un pixel assure une précision maximale[1].

Cette méthode est une approche basée sur l'apparence, qui consiste à parcourir l'ensemble de l'image en calculant un certain nombre de caractéristiques dans des zones rectangulaires qui se chevauchent. Elle a la particularité d'utiliser des caractéristiques très simples mais très nombreuses.

La méthode de Viola et Jones se caractérisent par :

- Utilisation d'images intégrales qui permettent de calculer plus rapidement les caractéristiques.
- La combinaison en cascade de classifieurs boostés, apportant un net gain de temps d'exécution.

- La sélection par Boosting des caractéristiques.

Etapes de la méthode :

La méthode se décompose en quatre étapes:

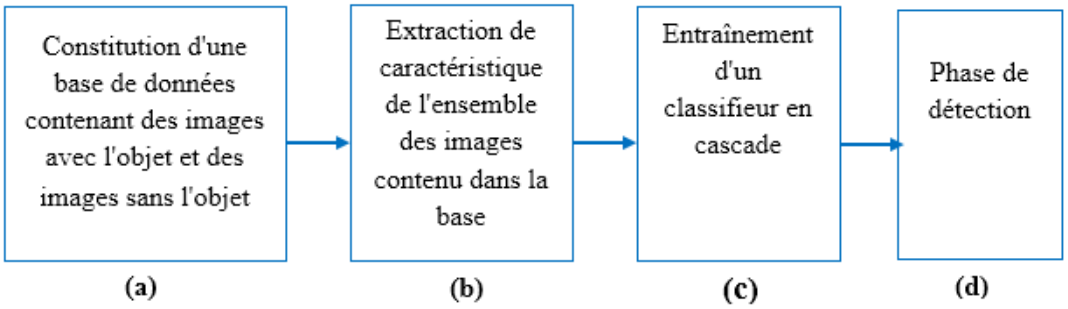


FIGURE 3.5: Les étapes de la méthode du classificateur en cascade Haar

a) Constitution d'une base de données contenant des images avec l'objet et des images sans l'objet: Pour pouvoir utiliser les cascades de Haar, il faut dans un premier temps constituer un jeu de données composé d'images contenant l'objet et des images ne contenant l'objet que nous souhaitons détecter.

b) Extraction de caractéristique de l'ensemble des images contenu dans la base: Une fois la base d'images constitué la deuxieme étape consiste à extraire les caractéristiques ne définis pas Viola et Jones dans leurs articles. Les cascade de Haar crée les caractéristiques qu'il utilise en appliquant sur l'image un filtre carré de 24x24. Il s'agit des filtres présentés sur l'image ci-dessous. Les caractéristiques sont obtenues en soustrayant la somme des pixels de l'image couvert par la zone blanche du filtre à la somme des pixels couverts par la zone bleue. Ce processus génère énormément de caractéristique étant donné que le filtre est appliqué sur toute l'image. Viola et Jones [33] évoque un nombre de caractéristique supérieur à 180 000 dans leur article pour une image de 384 par 288.

c) Entraînement d'un classifieur en cascade: Une étape très importante est l'apprentissage du classifieur. Il s'agit d'entraîner le classifieur afin de le sensibiliser à ce que l'on veut détecter, ici des panneaux. Pour cela, il est mis dans deux situations : La première où une énorme quantité de cas positifs lui sont présentés et la deuxième où, à l'inverse, une énorme quantité de cas négatifs lui sont présentés. Ainsi, on obtient un classifieur sensibles aux panneaux de signalisation.

Algorithme d'apprentissage basé sur Adaboost:Le deuxième élément clé de la méthode de Viola et Jones est l'utilisation d'une méthode de Boosting afin de sélectionner les meilleures caractéristiques. Le Boosting est un principe qui consiste à construire un classifieur fort à partir d'une combinaison pondérée de classifieurs faibles, c'est-à-dire, donnant en moyenne une

réponse meilleure qu'un tirage aléatoire. Viola et Jones adaptent ce principe en assimilant une caractéristique à unclassifieur faible, en construisant un classifieur faible qui n'utilise qu'une seule caractéristique. L'apprentissage du classifieur faible consiste alors à trouver la valeur seuil de la caractéristique qui permet de mieux séparer les exemples positifs des exemples négatifs. Le classifieur se réduit alors à un couple (caractéristique, seuil). le classifieur AdaBoost (Adaptive Boosting) présente l'avantage d'avoir une bonne vitesse d'exécution pour l'application tout en conservant un taux de reconnaissance élevé.

L'algorithme de Boosting utilisé est en pratique une version modifiée d'AdaBoost, qui est utilisée à la fois pour la sélection et pour l'apprentissage d'un classifieur fort. Les classifieurs faibles utilisés sont souvent des arbres de décision. Un cas remarquable, fréquemment rencontré, est celui de l'arbre de profondeur 1, qui réduit l'opération de classification à un simple seuillage.

L'algorithme est de type itératif, à nombre d'itérations déterminé. À chaque itération, l'algorithme sélectionne une caractéristique, qui sera ajoutée à la liste des caractéristiques sélectionnées aux itérations précédentes, et le tout va contribuer à la construction du classifieur fort final. Cette sélection se fait en entraînant un classifieur faible pour toutesles caractéristiques et en sélectionnant celui avec l'erreur la plus faible sur l'ensemble d'apprentissage. L'algorithme tient également à jour une distribution de probabilité sur l'ensemble d'apprentissage, réévaluée à chaque itération en fonction des résultats de classification. En particulier, plus de poids est attribué aux exemples difficiles à classer, c'està-dire ceux dont l'erreur est élevée. Le classifieur fort final construit par AdaBoost est composé de la somme pondérée des classifieurs sélectionnés.

Cascade de classifieurs: La méthode de Viola et Jones est basée sur une approche par recherche exhaustive sur l'ensemble de l'image, qui teste la présence de l'objet dans une fenêtre à toutes les positions et à plusieurs échelles. Cette approche est cependant extrêmement coûteuse en calcul. L'une des idées-clés de la méthode pour réduire ce coût réside dans l'organisation de l'algorithme de détection en une cascade de classifieurs. Appliqués séquentiellement, ces classifieurs prennent une décision d'acceptation ; la fenêtre contient l'objet et l'exemple est alors passé au classifieur suivant, ou de rejet ; la fenêtre ne contient pas l'objet et dans ce cas l'exemple est définitivement écarté. L'idée est que l'immense majorité des fenêtres testées étant négatives (c.-à-d. ne contiennent pas l'objet), il est avantageux de pouvoir les rejeter avec le moins possible de calculs. Ici, les classifieurs les plus simples, donc les plus rapides, sont situés au début de la cascade, et rejettent très rapidement la grande majorité des exemples négatifs. Cette structure en cascade peut également s'interpréter comme un arbre de décision dégénéré, puisque

chaque nœud ne comporte qu'une seule branche

d) Phase de détection C'est l'étape finale, Une fois le classifieur est entraîné, nous pouvons passer à la phase de détection. en fournissant une image en entrée et en recoit en soortie l'objet déttecter sur l'images representer par ces cordonées.

3.3 Les méthodes d'apprentissage profond pour la classification des panneaux de signalisation

L'apprentissage approfondis (DL) est un type d'intelligence artificielle (IA) et le sous-domaine de l'apprentissage automatique (ML), où la machine est capable d'apprendre par elle-même. Le développement du DL fut motivé en partie par l'échec des algorithmes traditionnels

Les différences entre les Méthodes d'apprentissagetraditionnelles et les méthodes d'apprentissage approfondis sont :

- plus la quantité de données fournie est grande plus les performances sont meilleures;
- pas de "plateau de performance": ils ne possèdent pas une borne supérieure à la quantité de données qu' ils peuvent recevoir
- l'étape de l'extraction de caractéristiques:

Les méthodes d'apprentissage approfondis sont utilisées dans plusieurs domaines:

- Résoudre les problèmes réels des scènes de circulation (occlusion partielle, encombrement d'arrière-plan, etc.)
- Résoudre d'autres problèmes liés aux méthodes de détection basées sur les couleurs et les formes (changements d'éclairage, panneaux endommagés, etc.).
- Résoudre les problèmes liés aux méthodes de détection basées sur la ML (conception manuelle, limitation des fonctionnalités de conception artificielle, etc.).

3.3.1 Réseaux de neurones convolutifs CNN

Un réseau de neurone convolutifs (Convolutional Neural Network CNN en anglais) est un type particulier de réseau de neurone qui se base sur l'opération de convolution, les CNN sont particulièrement adaptés à la reconnaissance d'images. Un réseau de neurones est un système composé de neurones,

généralement répartis en plusieurs couches connectées entre eux, il est construit en empilant les couches : la sortie d'une couche correspond à l'entrée de la suivante. Les CNNs se composent de deux parties, la première est appelée la partie convolutif et la deuxième est appelée la partie la classification, Ils se composent également de quatre types de couches : la couche de convolution, la couche de pooling, la couche de correction (ReLU) et la couche fully-connected. La figure suivante montre une vue générale de CNN: **La**

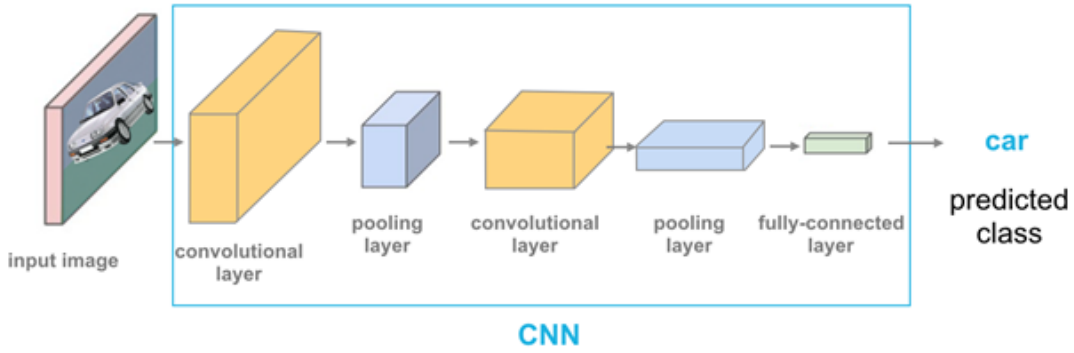


FIGURE 3.6: Une vue générale de CNN.

couche de Convolution: est la composante clé des CNNs, constitue toujours elle constitue au moins leur première couche, cette couche est l'extracteur de caractéristiques des images. Son but est de repérer la présence d'un ensemble de caractéristiques (Features en anglais) dans les images reçues en entrée. Pour cela, on réalise un filtrage par convolution, le principe est de faire "glisser" une fenêtre représentant la feature sur l'image, et de calculer le produit de convolution entre la feature et chaque portion de l'image balayée. Trois hypers paramètrent permettant de dimensionner le volume de la couche de convolution :

- 1.Profondeur de la couche K : nombre de noyaux de convolution;
- 2.Le pas S : contrôle le chevauchement des champs récepteurs.
- 3.La marge (à 0) ou zéro padding P : Cette marge permet de contrôler la dimension spatiale du volume de sortie.

La couche de Polling: C'est parmi les outils puisant de CNN, cette couche permet de réduire la taille des images avec la préservation des caractéristiques.

Pooling est une fonction de sous-échantillonnage spatial. Elle prend toujours en entrée des feature maps de taille (nx * ny * D), et réduit les deux premières dimensions spatiales. Pour appliquer cette opération il ya deux méthodes, soit on fait la moyenne des valeurs de la zone (Average Pooling

en anglais), soit on extrait uniquement la valeur la plus élevée (Max Pooling en anglais). Et nous utilisons celle qui applique l'opération max au résultat de chaque filtre. Pour cela, on découpe l'image en cellules régulière, puis on garde au sein de chaque cellule la valeur maximale. En pratique, on utilise souvent des cellules carrées de petite taille pour ne pas perdre trop d'informations. Enfin, la couche de pooling permet de réduire le nombre de paramètres et de calculs dans le réseau, on améliore ainsi l'efficacité du réseau et on éviter le sur-apprentissage. La figure suivante montre la différence entre Average Pooling et Max Pooling :

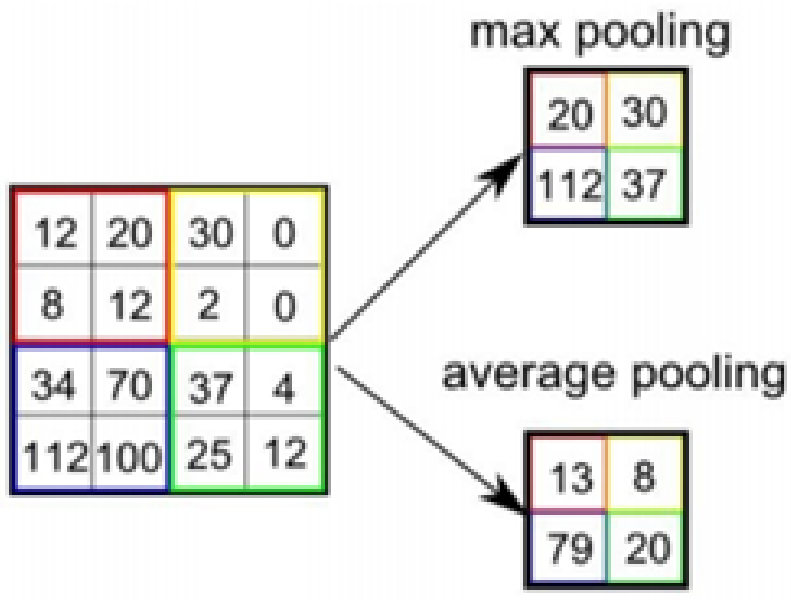


FIGURE 3.7: La différence entre Average Pooling et Max Pooling.

La couche de ReLY: Cette couche améliore l'efficacité du traitement en intercalant entre les couches de traitement une couche qui va opérer une fonction mathématique (fonction d'activation) sur les signaux de sortie. Il existe beaucoup de fonctions d'activations comme : 1. ReLU. L'unité linéaire rectifiée (Rectified Linear Unit en anglais), est une fonction d'activation très couramment utilisée, elle permet tout simplement de remplacer les résultats négatifs par zéro. 2. SoftMax. Utilisé pour de la multi classification en couche de sortie. Intervalle de sortie (- +). La figure suivante montre ces deux fonctions:

La couche de Fully-Connected: Cette couche constitue toujours la dernière couche d'un réseau de neurones convolutif ou non – elle n'est donc pas caractéristique d'un CNN, ce type de couche reçoit un vecteur en entrée et produit un nouveau vecteur en sortie. Pour cela, elle applique une combinaison linéaire puis éventuellement une fonction d'activation aux valeurs reçues en

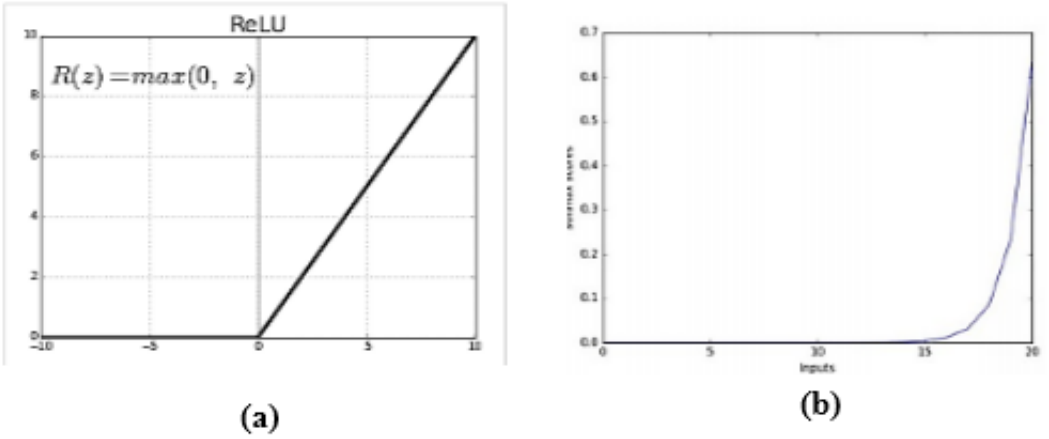


FIGURE 3.8: Deux fonctions d'activations, (a) ReLU et (b) SoftMax.

entrée. La couche Fully-Connected permet de classifier l'image en entrée du réseau, elle renvoie un vecteur de taille N, où N est le nombre de classes dans notre problème de classification d'images. Chaque élément du vecteur indique la probabilité pour que l'image en entrée appartient à une classe.

3.3.2 Architectures CNN

Architecture Lenet

LeNet est un réseau convolutif de LeCun [29]conçu pour la reconnaissance de caractères manuscrits et imprimés. Il s'agit d'un réseau neuronal convolutif avec suffisamment d'entrée pour recevoir plusieurs objets et plusieurs sorties appelés réseaux neuronaux à déplacement spatial (SDNN), capable de reconnaître des chaînes en une seule passe sans segmentation préalable. Les couches convolutives clairsemées et la mise en commun maximale sont au cœur de la famille de modèles LeNet [31-32]. La figure suivante montre l'architecture LeNet:

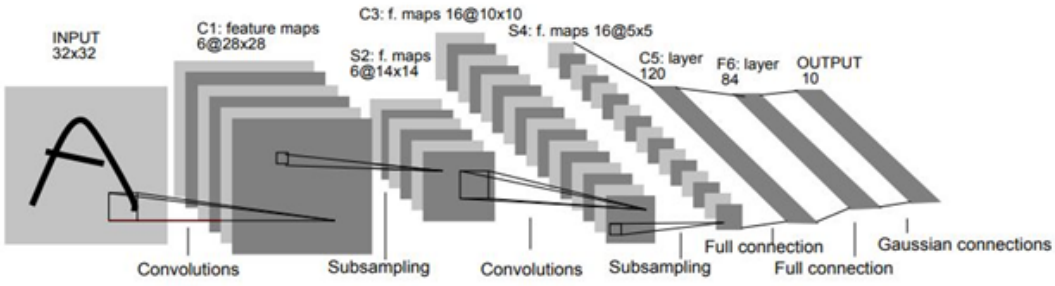


FIGURE 3.9: Architecture LeNet

L'architecture LeNet-5 se compose de deux ensembles de couche convolutive et pooling, suivie d'une couche flatten, puis de deux couches fully-connected et enfin d'un classificateur softmax.

Première couche L'entrée pour LeNet-5 est une image 32 32 pixels qui traverse la première couche convolutive avec 6 filtres de taille 5 5 et une foulée (en anglais Stride) d'un. Les dimensions de l'image passent de 32x32x1 à 28x28x6 pixels.

Deuxième couche: Ensuite, le LeNet-5 applique une couche de pooling moyenne (en anglais Average Pooling Layer) avec une taille de filtre 2 2 et une foulée de deux. Les dimensions de l'image résultante seront réduites à 14x14x6 pixels.

Troisième couche: ext, il y a une deuxième couche convolutive avec 16 filtres de taille 5x5 et une foulée de 1. Les dimensions de l'image passent de 14x14x16 à 10x10x16 pixels.

Quatrième couche: Le LeNet-5 applique à nouveau une couche de pooling moyenne avec une taille de filtre 2 2 et une foulée de deux. Cette couche est la même que la deuxième couche (S2) sauf qu'elle comporte 16 filtres, donc la sortie sera réduite à 5x5x16 pixels.

Cinquième couche: Est une couche convolutive entièrement connectée (en anglais Fully Connected) avec 120 filtres chacune de taille 1 1. Chacune des 120 unités de cinquième couche est connectée à tous les 400 nœuds (5x5x16) de la quatrième couche.

Sixième couche: Est une couche convolutive entièrement connectée avec 84 unités.

Couche de sortie: Enfin, il existe une couche de sortie softmax entièrement (en anglais Fully Connected Softmax) connectée $\hat{y}$ avec 10 valeurs possibles correspondant aux chiffres de 0 à 9.

Architecture AlexNet

Alexnet est un réseau de neurones à convolution profonde (CNN) pour la classification des images qui a remporté le concours ILSVRC-2012 et a obtenu un taux d'erreur de test parmi les 5 premiers de 15,3%, contre 26,2% obtenu par la deuxième meilleure entrée [37].

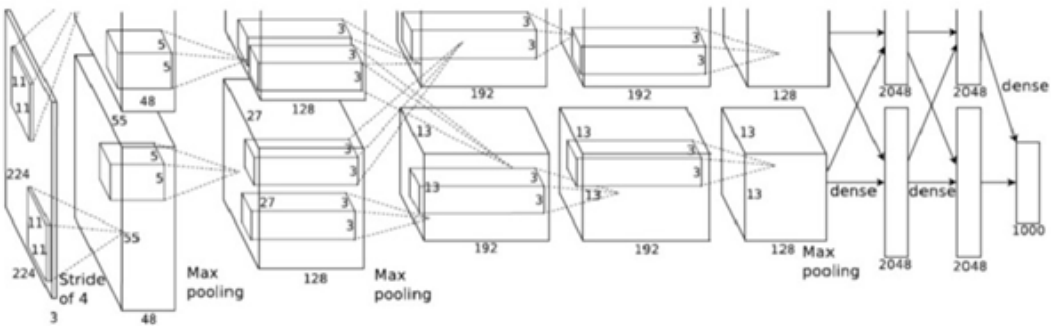


FIGURE 3.10: Architecture AlexNet

Alexnet a 8 couches[10]. Les 5 premiers sont convolutifs et les 3 derniers sont des couches entièrement connectées. Entre les deux, nous avons également des couches appelées regroupement (en anglais Pooling) et activation.

Architecture GoogleNet

GoogLeNet (ou Inception V1) a été proposé par la recherche de Google (avec la collaboration de diverses universités) en 2014 dans le document de recherche intitulé 'Going Deeper with Convolutions'[18]. Cette architecture a été lauréate du défi de classification d'images ILSVRC 2014. Il a permis une diminution significative du taux d'erreur par rapport aux précédents.

L'architecture de GoogLeNet est très différente des architectures précédentes, elle est basé sur l'architecture Inception, car il contient plusieurs blocs d'Inception. L'architecture Inception est différente des architectures précédentes. Dans cette architecture, il existe une taille de convolution fixe pour chaque couche. Dans cette architecture (1 1, 3 3, 5 5) convolution et (3 3) max pooling réalisés de manière parallèle à l'entrée et la sortie de ceux-ci sont empilés ensemble pour générer la sortie finale. L'idée derrière que les filtres de convolution de différentes tailles gèrent mieux les objets à plusieurs échelles.

Si un réseau est construit avec de nombreuses couches profondes, il peut être confronté au problème de surajustement. Pour résoudre ce problème, les auteurs [18] ont proposé l'architecture GoogLeNet avec l'idée d'avoir des filtres de tailles multiples pouvant fonctionner au même niveau. Avec cette idée, le réseau devient en fait plus large que plus profond. Ci-dessous, une figure montrant un module d'Inception Naïf

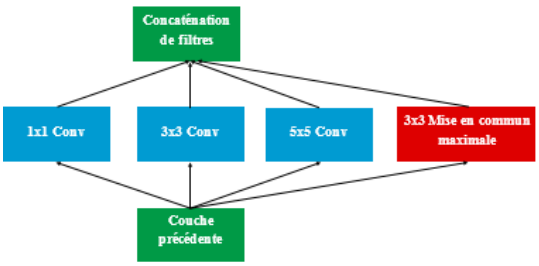


FIGURE 3.11: Module d'Inception Naïf.

Comme on peut le voir dans le diagramme ci-dessus, l'opération de convolution est effectuée sur des entrées avec trois tailles de filtre : (1 * 1), (3 * 3) et (5 * 5). Une opération de pooling max est également effectuée avec les convolutions et est ensuite envoyée dans le module d'Inception suivant.

Généralement, les réseaux de neurones sont longs et coûteux à entraîner, Dans GoogleNet le nombre de canaux d'entrée est limiter par l'ajout d'une convolution supplémentaire (1 1) avant les convolutions (3 3) et (5 5) pour

réduire les dimensions du réseau et effectuez des calculs plus rapides. La figure suivante montre un module d'Inception avec réduction de dimension

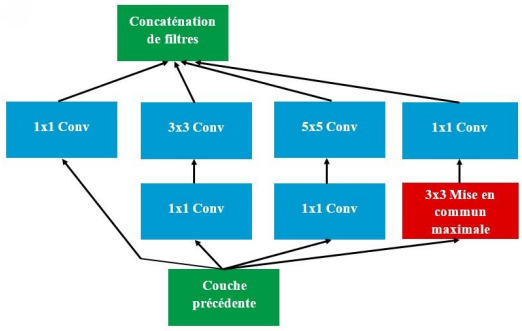


FIGURE 3.12: Module d'Inception avec réduction de dimension.

L'architecture GoogleNet comprend 22 couches de profondeur, avec 27 couches de pooling incluses[31]. Il y a 9 modules d'inceptions empilés linéairement au total. Les extrémités des modules d'inceptions sont connectées à la couche de pooling moyenne globale. La figure suivante montre l'architecture complète de GoogleNet:

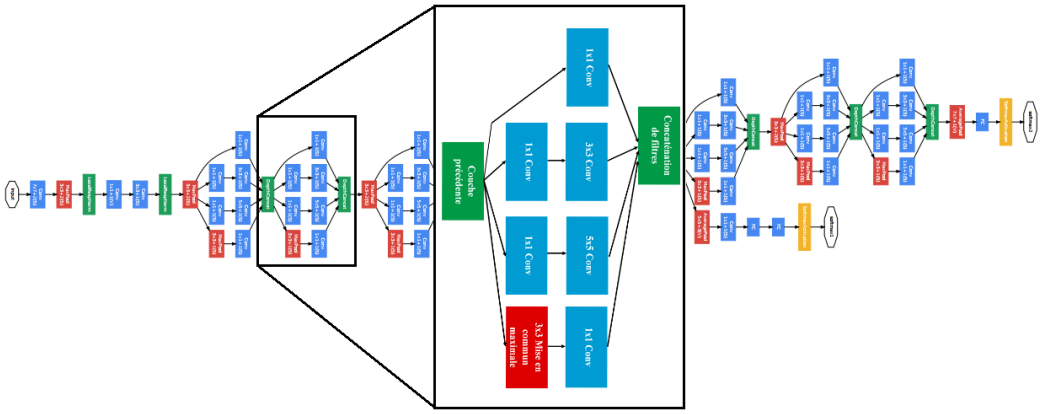


FIGURE 3.13: Architecture GoogleNet

3.4 Conclusion

Dans ce chapitre, nous avons présenté les méthodes de détection et localisation des panneaux de signalisations. Les différents espaces de couleur présentent des caractéristiques utiles pour la segmentation en ce qui suit les espaces RGB et HSV sont utilisés avec le classificateur en cascade de Haar qui présente des taux de détection élevés. Pour l'étape de classification les méthodes basées sur l'apprentissage approfondis ont été choisis.

Chapitre 4

Approche Proposée

4.1 Introduction

Dans ce chapitre, nous proposerons une architecture d'un système de détection et reconnaissance des panneaux de signalisation basée sur trois étapes: le prétraitement, la détection et la reconnaissance.

Pour chaque étape nous exploitant les méthodes et techniques vues précédemment. Dans l'étape de prétraitement, on applique une transformation d'espace colorimétrique Ensuite, nous présentons la méthode de détection utilisée basée sur l'apprentissage automatique. Et enfin, nous présentons la méthode de classification utilisée basée sur l'apprentissage en profondeur, en utilisant trois architectures.

4.2 Architecture générale

Notre systeme, est composé de trois tapes, commençant par le prétraitement, puis la détection, Et enfin la reconnaissance. L'architecture genérale du systeme est donnée par la figure 4.1 :

4.3 Déscription détaillée

4.3.1 Prétraitement

L'étape de prétaitement consiste à la préparation de l'image pour l'etape de détection cette eteape comporte la transformation d'espace de couleurs. En effet la detection repose sur l'extarction de l'objet a base des informations de la couleurs. Quatres espaces on étatit utilisés : **RGB**, **niveau de gris**, **HSV** et **YCbCr**. les résultats de chaque espace et donné est comparer dans le chapitre suivant.

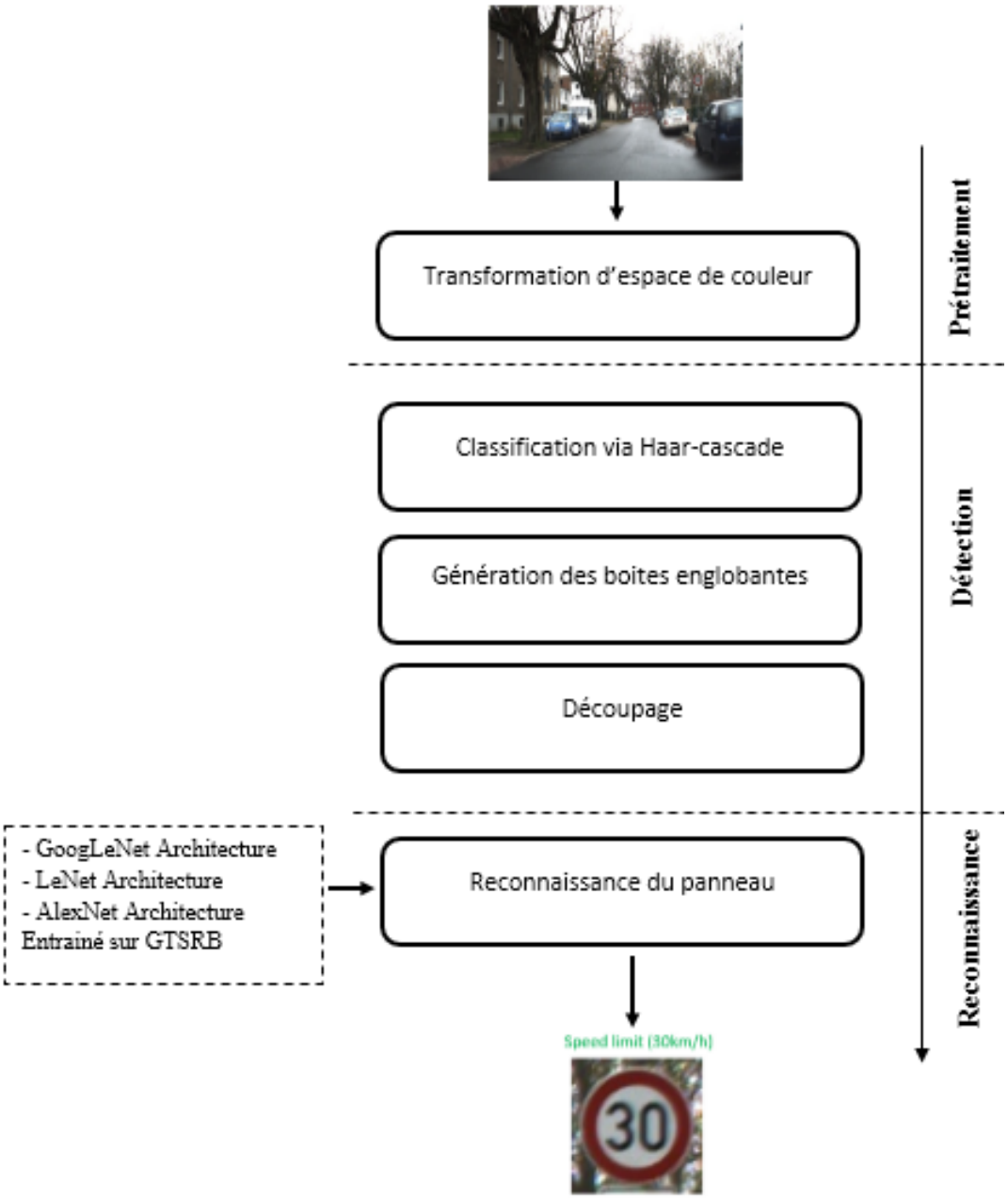


FIGURE 4.1: Archiecture Générale du systeme

HSV est l'espace colorimétrique le plus approprié, il permet de séparer les composants de couleur de l'intensité. Cette sparation garantie une robustesse aux changements d'éclairage et la suppression des ombres.La figure suivante montre un exemple de conversion d'espace de couleur de RGB vers L'HSV.

4.3.2 Détection

La phase de détéction permet de detecter le pannaeux sur l'image convertie et l'encadrer par une boite englobant sur enfin l'image est découpée. L'objectif

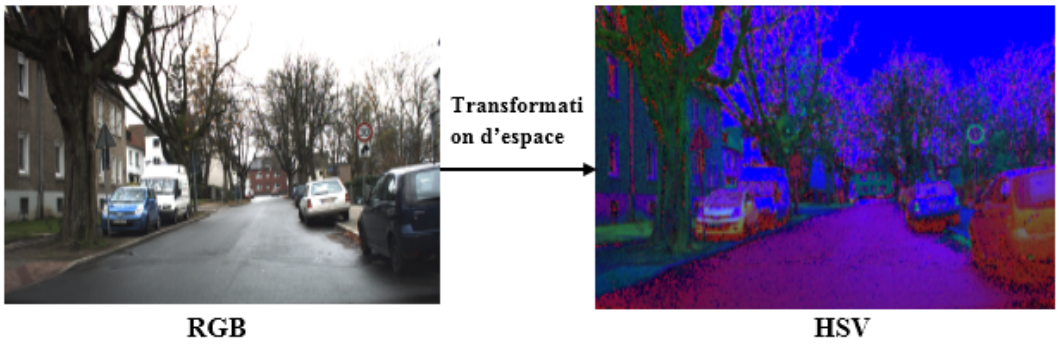


FIGURE 4.2: Un exemple de transformation de l'espace colorimétrique: **RGB** vers **HSV**

de cette partie est la détection des panneaux de limitation de vitesse en utilisant un classificateur Haar Cascade entrainé sur plus 2000 images positives et 1000 images négatives :

- Les images positives: Contiennent les images des panneaux de limitation de vitesse que nous voulons que notre classificateur identifie.
- Les images négatives: Contiennent les Images de tout le reste, qui ne contiennent pas les panneaux de limitation vitesse que nous voulons détecter.

La figure suivante montre un exemple des résultats des différentes étapes de la phase de détection:

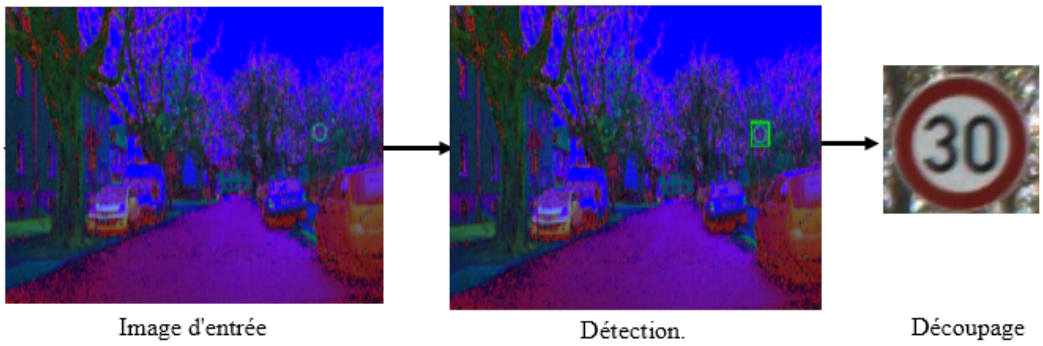


FIGURE 4.3: Un exemple des résultats de la phase de détection

4.3.3 Reconaissance

La reconnaissance consiste a classer le panneau. Pour cela, nous avons utilisé trois architectures des réseaux neuronaux convolutifs CNN : **GoogLeNet**, **AlexNet** et **LeNet**.

Nous avons formé et validé un modèle pour chaque architecture afin qu'il puisse classer ces panneaux à l'aide de la base de données GTSRB.

couche	kernel/Stride/Padding	Entrée	Sortie	Paramètres
Convolution	3x3 / 2x2 / padding : valid	32x32x3	28x28x64	1,792
Convolution	3x3 / 2x2 / padding : same	28x28x64	28x28x192	110,784
Inception(2a)	/	28x28x192	16x16x256	128K
Inception(2b)	/	16x16x256	28x28x480	380K
Max pooling	3x3/2x2	28x28x480	14x14x480	/
Inception(3a)	/	14x14x480	14x14x512	364K
Inception(3b)	/	14x14x512	14x14x512	437K
Inception(3c)	/	14x14x512	14x14x512	463K
Inception(3d)	/	14x14x512	14x14x528	580K
Inception(3e)	/	14x14x528	14x14x832	840K
Max pooling	3x3/2x2	14x14x832	7x7x832	/
Inception(4a)	/	7x7x832	7x7x832	1072K
Inception(4b)	/	7x7x832	7x7x1024	1388K
Avg pooling	7x7/1x1	7x7x1024	1x1x1024	/
Flatten	/	1x1x1024	1024	/
Softmax	/	1024	43	/

TABLE 4.1: Les détails de l'architecture GoogLeNet

L'architecture GoogLeNet

Nous avons construit cette architecture en utilisant deux couches convolutives, suivie de deux blocs-d'inceptions et d'une couche de pooling, puis de cinq blocs-d'inceptions et d'une couche pooling, puis de deux blocs-d'inceptions et d'une couche pooling et enfin une couche flatten suivie d'un classificateur softmax.

L'architecture GoogLeNet a beaucoup de couches (22 couches), elle a donc besoin de performances GPU très élevées pour l'entraîner. Nous avons apporté des modifications sur les dimensions d'entrée de la base de données GTRSB et ajouter une couche convolutive pour répondre aux exigences.

Les détails du réseau sont indiqués dans le tableau 5.5

L'architecture AlexNet

Cette architecture a été construit en utilisant deux ensembles de couches convolutives et de pooling, suivie par trois couches convolutives et une couche pooling, puis une couche de flatten, ensuite deux couches de fully-connected et enfin un classificateur softmax.

Étant donné que les dimensions d'entrée et de sortie de la base de données GTRSB sont respectivement 32x32 avec 43 classes, qui sont différentes des dimensions d'origine de l'architecture Alex Net, nous avons apporté des modifications pour répondre aux exigences.

L'architecture que nous avons réalisée est présentée dans le tableau suivant.

couche	kernel/Stride/Padding	Entrée	Sortie	Paramètres
Convolution	5x5 / 1x1 / padding : valid	32x32x3	28x28x9	684
Max pooling	2x2 / 2x2	28x28x9	14x14x9	/
Convolution	3x3 / 1x1 / padding : valid	14x14x9	12x12x32 3	,776
Max pooling	2x2 / 2x2	12x12x32	6x6x32	/
Convolution	3x3 / 1x1 / padding : same	6x6x32	6x6x48	13,872
Convolution	3x3 / 1x1 / padding : same	6x6x48	6x6x64	27,712
Convolution	3x3 / 1x1 / padding : same	6x6x64	6x6x96	55,392
Max pooling	2x2 / 2x2	6x6x96	3x3x96	/
Flatten	/	3x3x96	864	/
Fully connected	/	864	400	346,000
Fully connected	/	400	160	64,160
Softmax	/	160	43	6,923

TABLE 4.2: Les détails de l'architecture AlexNet

L'architecture LeNet

Nous avons construit cette architecture en utilisant deux ensembles de couche convolutive et pooling, suivie d'une couche flatten, puis de deux couches fully-connected et enfin d'un classificateur softmax.

Les dimensions d'entrée et de sortie de cette architecture sont compatibles avec les dimensions de la base de données GTSRB, nous n'avons donc eu aucun problème de mise en œuvre.

Les détails de cette architecture sont présentés dans le tableau suivant.

couche	kernel/Stride/Padding	Entrée	Sortie	Paramètres
Convolution	5x5 / 1x1 / padding : valid	32x32x3	28x28x6	465
Max pooling	2x2 / 2x2	28x28x6	14x14x6	/
Convolution	3x3 / 1x1 / padding : valid	14x14x6	10x10x16	880
Max pooling	2x2 / 2x2	10x10x16	5x5x16	/
Flatten	/	5x5x16	400	/
Fully connected	/	400	120	48,120
Fully connected	/	120	80	9,680
Fully connected	/	80	43	3,483

TABLE 4.3: Les détails de l'architecture LeNet

4.4 Conclusion

Dans ce chapitre, nous avons présentées en détails l'approche proposée. Pour chaque étape nous spécifions les méthodes et techniques utilisées ainsi que la démarche à suivre pour aboutir aux résultats. Les différentes étapes seront implémentées et discutées dans le chapitre suivant.

Chapitre 5

Expérimentation et Evaluation

5.1 Introduction

Dans ce chapitre nous présentons l'implémentation de l'approche proposée. Nous commençons par la présentation des enivrements de développement matériels et logiciels ainsi que les paramètres de configuration et les bases de données utilisées.

Les résultats obtenus sont présentés en détails, comparés et discutés avec des exemples sur des images des bases utilisées.

5.2 Présentation des outils

5.2.1 Matériel

Nous avons utilisé plusieurs méthodes (Pré traitement, Détection et Reconnaissance), ces dernières nécessitent une haute performance c'est pourquoi nous avons utilisé Google Colab. Notre système est basé sur le traitement d'image, il nécessite le GPU (Unité de Traitement Graphique) que Google Colab nous permet d'utiliser dans un environnement de notebook 'Jupyter' en ligne et gratuit.

5.2.2 Logiciel et environnement

Pour réaliser ce travail, nous avons utilisé plusieurs modules, parmi lesquels :

- **Module TensorFlow:**Cela nous permet d'implémenter et de manipuler toutes les architectures de manière simple et facile[1].

[1]https://github.com/tensorflow/tensorflow

- **Module Scikit-Learn:**Cela nous permet de mélanger l'ensemble de données sans affecter les matrices d'origine en utilisant la fonction *shuffle ()*.[2]
- **Module Mathplotlib:**Cela nous permet de créer des visualisations statiques, telles que des graphiques.[3]
- **Module OpenCv:**Cela nous permet de créer des visualisations statiques, telles que des graphiques[4].
- **Module Numpy:**Cela nous permet de manipuler les données pour répondre aux exigences spécifiées[5].
- **Module Pickle:**Cela nous permet de sérialiser les objets afin qu'ils puissent être enregistrés dans un fichier et chargés ultérieurement dans un programme.

5.3 Les bases de données utilisées

Dans ce travail, deux bases de données ont été utilisées, à savoir GTSDB et GTSRB :

5.3.1 La base de données allemande GTSDB

Cette base de données a été utilisée dans la phase de détection, elle se caractérise par la taille uniforme de toutes les images avec une résolution de 1360800 pixels, dans lequel la taille des panneaux de signalisation varie de 16 à 128 pixels.



FIGURE 5.1: Un exemple d'images de panneaux de limitation de vitesse de la base BTSDB

[2]https://github.com/scikit-learn/scikit-learn
[3]https://github.com/matplotlib/matplotlib
[4]https://github.com/opencv/opencv
[5]https://github.com/numpy/numpy

5.3.2 La base de données allemande GTSRB

Nous avons utilisé cette base de données pour faire l'étape de reconnaissance, elle se caractérise par son grand nombre d'images (51000), différents types d'images (43 types de panneaux de signalisation) et la taille de l'image qui n'est pas uniforme (de 15 15 à 250 250 pixels). Cette base de données est divisée en 3 sous-bases de données (Train, Test et Valid). La figure suivante montre quelques images de cette base de données (Figure



FIGURE 5.2: Exemples d'images de la base GTSRB

5.4 Paramètres de configuration

La phase de reconnaissance

Dans cette étape, nous avons utilisé la base de données de reconnaissance allemande GTSRB, elle contient plus de 51 000 images. Toutes ces images ont été utilisées avec leurs différents types de panneaux de signalisation (43 types). Nous avons utilisé 67% pour l'apprentissage, 24% pour les tests et 9% pour la validation. Nous avons traité l'ensemble de données afin que la taille de toutes les images devienne uniforme avec une dimension de 32x32px. Pour l'espace colorimétrique, nous avons utilisé l'espace RVB d'origine des images.

Le tableau suivant présente tous les détails des paramètres de configuration de cette base de données :

/	Totales	Apprentissage	Test	Validation	Nbr Classes
GTSRB	51839 (100 %)	34799(67 %)	12630(24 %)	4410(9 %)	43

TABLE 5.1: Les détails des paramètres de configuration de la base de données GTSRB.

La phase de détection

Dans cette étape, nous avons utilisé la base de données de détection allemande GTSDB qui contient deux sous-ensembles de données, un sous-ensemble d'apprentissage avec 600 images et un autre test avec 300 images, soit 900 images au total.

Nous n'avons utilisé que les panneaux de limitation de vitesse trouvés dans le sous-ensemble de test, nous avons trouvé 101 images avec 124 annotations pour les 7 différentes classes.

Le tableau suivant présente les détails des paramètres de configuration de cette base de données :

Classe de panneaux	NTI	NTA
Speed Limit (30 km/h)	31	31
Speed Limit (50 km/h)	18	22
Speed Limit (60 km/h)	6	8
Speed Limit (70 km/h)	31	35
Speed Limit (80 km/h)	8	16
Speed Limit (100 km/h)	2	2
Speed Limit (120 km/h)	5	10

TABLE 5.2: Les détails des paramètres de configuration de la base de données GTSDB.

Tel que : **NTI** : Nombre total d'images. **NTA** : Nombre total de panneaux (Annotations).

5.5 Résultats et Discussions: Detection

5.5.1 La phase de detetion

Après l'évaluation de l'étape de detection, le système a donné une précision de 1.0, un rappel de 0.79, un taux d'erreur de 0.21 et un temps d'exécution de 0.62 s pour l'espace de couleur HSV, pour l'espace RGB une précision de 0.85, un rappel de 0.4, un taux d'erreur de 0.64 et un temps d'exécution de 0.51, et pour l'espace GRAY le système a donné une précision de 0.87, un rappel de 0.48, un taux d'erreur de 0.55 et un temps d'exécution de 0.52 et

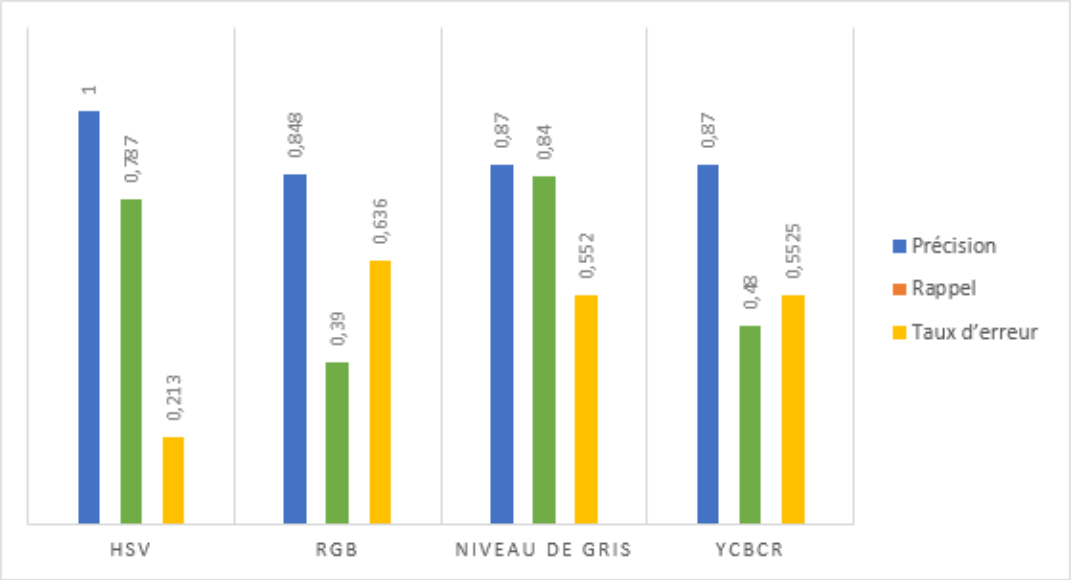


FIGURE 5.3: Evaluation des performances sur les quatre espaces de couleurs: **RGB**, **niveau de gris**, **HSV** et **YCbCr**.

enfin pour l'espace YCbCr une précision de 0.87, un rappel de 0.48, un taux d'erreur 0.55 et un temps d'exécution de 0.38.

5.5.2 Comparaison entre les resultats sur differents espaces de couleurs

La comparaison entre les résultats montre que HSV présente meilleurs résultats que les autres espaces, bien que son temps d'exécution soit le plus long parmi les autres. En effet, cet espace sépare la Hue de la saturation et diminue donc l'influence des changements environnementaux. Contrairement au RVB qui présente plus de bruit. Une autre différence est le taux

	Précision	Rappel	Taux d'erreur	Temps d'exé sec
HSV	1	0.79	0.21	0.62 /
RGB	0.85	0.4	0.64	0.52
Niveau de gris	0.87	0.84	0.55	0.51
YCbCr	0.87	0.84	0.55	0.38

TABLE 5.3: Comparaison des résultats obtenus sur les quatre espaces couleurs.

d'erreur élevé pour les espaces Niveaux de gris et YCbCr. Nous observons les débits inférieurs dans les espaces à trois canaux par rapport aux espaces à un seul canal en raison de la perte d'informations des espaces à trois canaux.

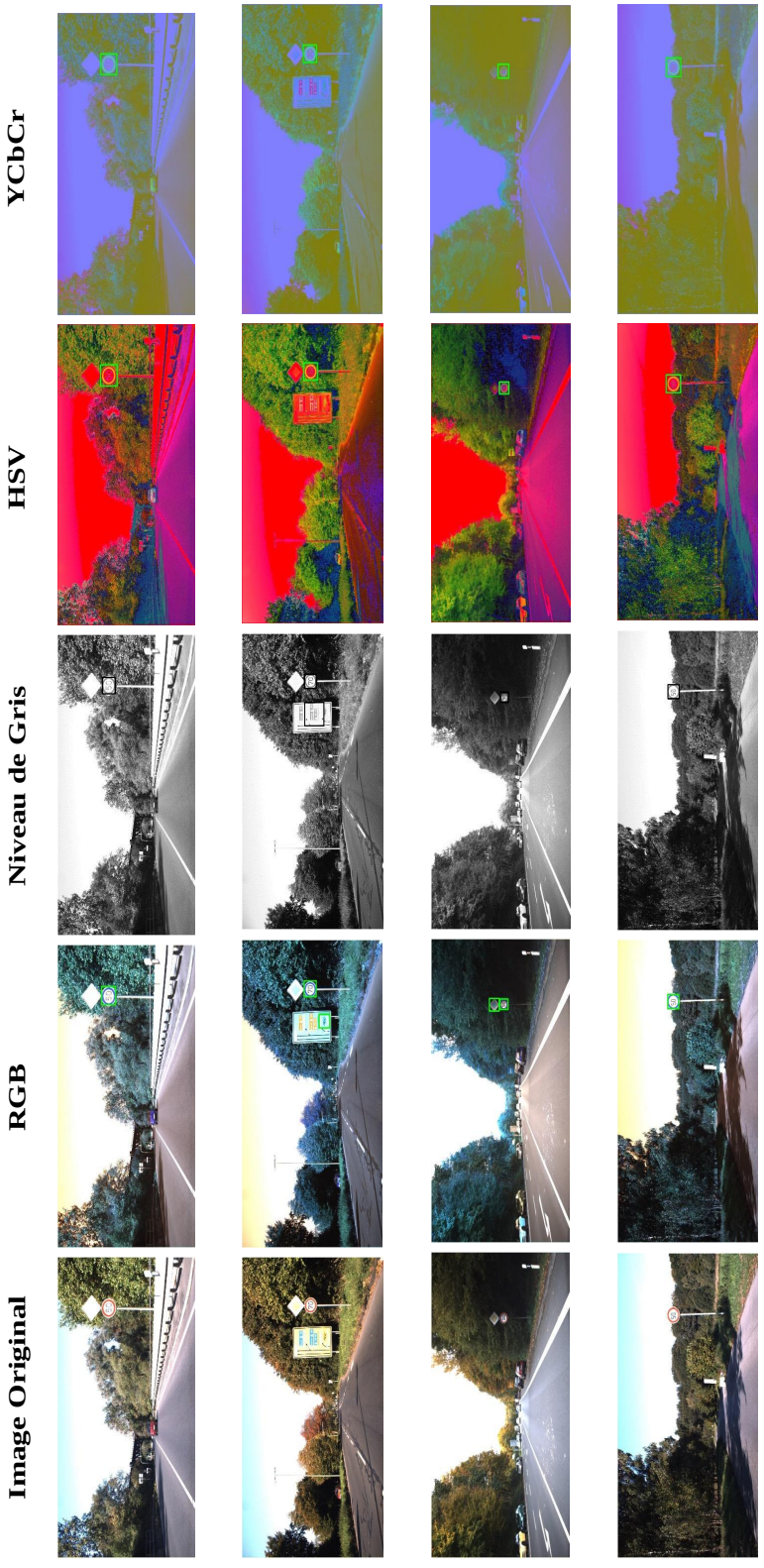


FIGURE 5.4: Exemples de resultats de detection sur differents espaces de couleurs.Sur la première colonne les images originaux respectivement les résultats de détection sur les espaces de couleurs : **RGB**, **niveau de gris**, **HSV** et **YCbCr**.

Le rappel dépend de la capacité de la méthode à détecter tous les panneaux. Par conséquent, le classificateur Haar a besoin de plus d'entrainement sur un grand nombre d'images pour obtenir de meilleurs résultats. Sans compter que les résultats pour l'espace HSV sont toujours bons et acceptables par rapport au rappel, c'est ce qui confirme que cet espace est le meilleur parmi les autres.

5.6 Résultats et Discussions: Reconnaissance

5.6.1 Métriques d'évaluation

précision(Accuracy) de la classification : est un moyen de mesurer la fréquence à laquelle le modèle classe correctement les panneaux de signalisation. La précision est le nombre de panneaux correctement prédits parmi tous les panneaux. Plus formellement, il est défini comme le nombre de vrais positifs et de vrais négatifs divisé par le nombre de vrais positifs, de vrais négatifs, de faux positifs et de faux négatifs. Un vrai positif ou un vrai négatif est un panneau de signalisation que le modèle a correctement classé comme vrai ou faux, respectivement. En revanche, Un faux positif ou un faux négatif est un panneau de signalisation que le modèle a mal classé.

$$Prcision = (Nombredeprdictioncorrecte)/(Nombretotaldeprdictionseffectues) \tag{5.1}$$

La fonction de perte par entropie croisée sigmoïde: est une méthode d'évaluation de la capacité d'un algorithme à modéliser un ensemble de données. Si les prédictions sont totalement fausses, la fonction de perte affichera un nombre plus élevé. S'ils sont plutôt bons, un nombre inférieur sera affiché.

Il existe de nombreuses fonctions de perte, parmi lesquelles la fonction de perte d'entropie croisée (Cross Entropy Loss ou Log Loss) calculer par l'équation suivante:

$$E(y', y) = ylog(y') + (y1)log(1y') \tag{5.2}$$

tel que :y': probabilité prédite par le modele et y: probablité attendue.

La fonction de perte d'entropie croisée mesure les performances d'un modèle de classification dont la sortie est une valeur de probabilité comprise entre 0 et 1. La perte d'entropie croisée augmente lorsque la probabilité prédite diverge de l'étiquette réelle. Lorsque l'étiquette d'observation réelle est 1, ce serait mauvais et entraînerait une valeur de perte élevée. Un modèle parfait aurait une perte de journal de 0.

5.6.2 Performances de l'architecture LeNet

La figure 5.6.2 montre les résultats de l'accuracy et de perte pour les trois ensembles de données d'entraînement, de test et de validation obtenus via l'architecture LeNet:

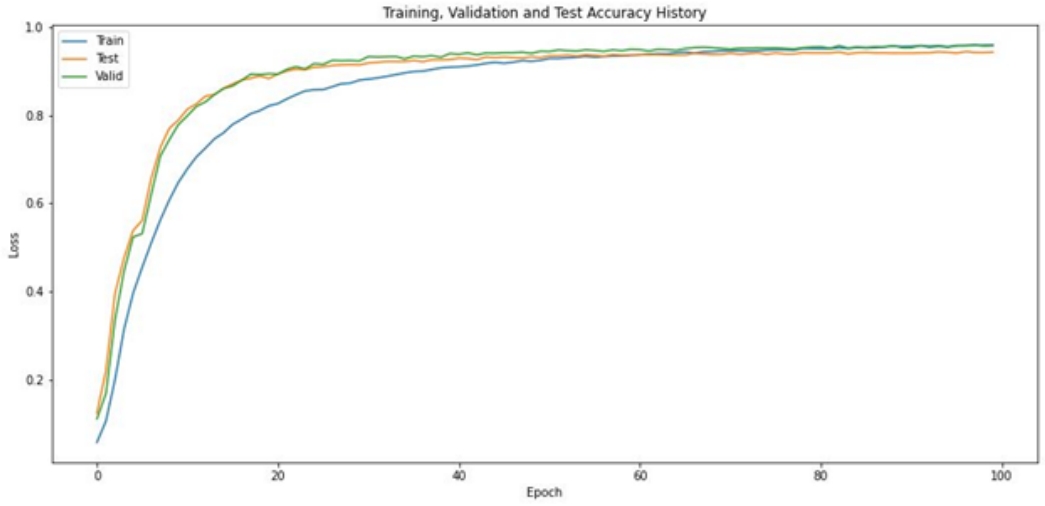


FIGURE 5.5: Performances de l'architecture LeNet : Accuracy.

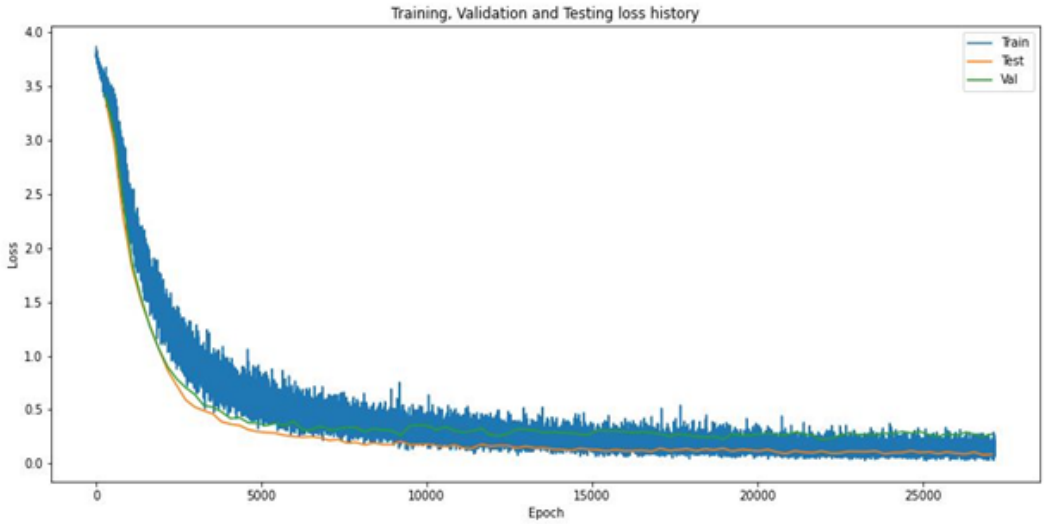


FIGURE 5.6: Performances de l'architecture LeNet : Fonction de perte.

5.6.3 Performances de l'architecture AlexNet

La figure 5.6.3 montre les résultats de l'accuracy et de perte pour les trois ensembles de données d'entraînement, de test et de validation obtenus via l'architecture ALexNet:

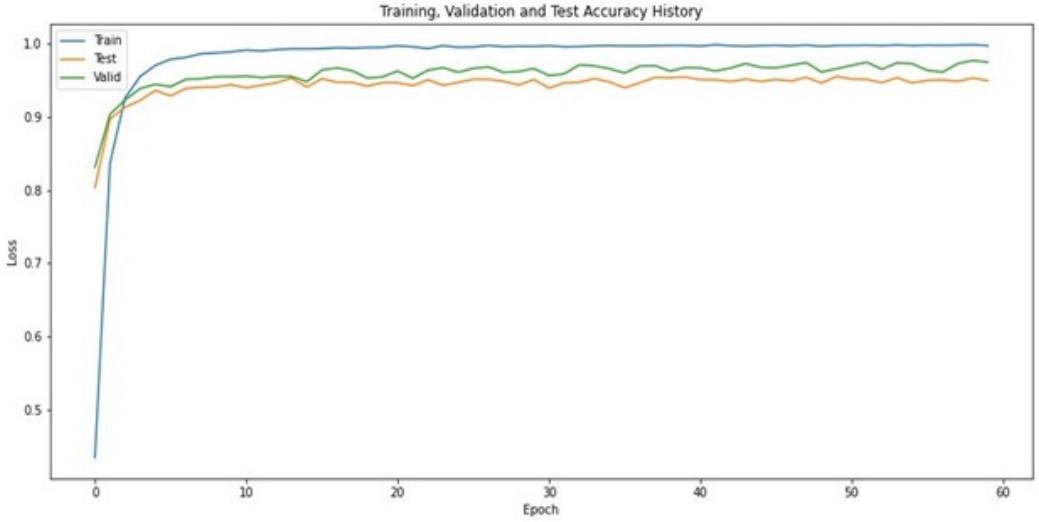


FIGURE 5.7: Performances de l'architecture ALexNet : Accuracy.

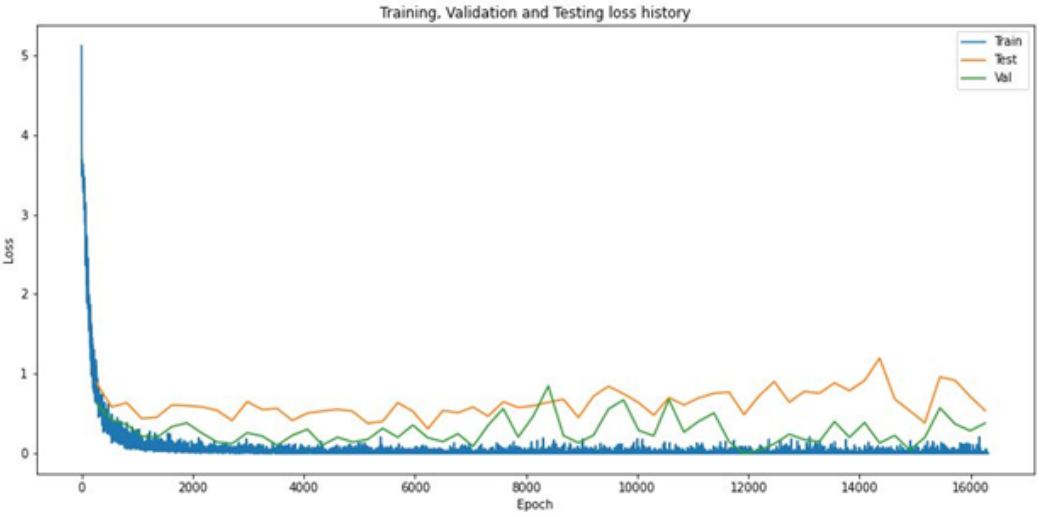


FIGURE 5.8: Performances de l'architecture AlexNet : Fonction de perte.

5.6.4 Performances de l'architecture GoogLeNet

La figure 5.6.4 montre les résultats de l'accuracy et de perte pour les trois ensembles de données d'entraînement, de test et de validation obtenus via l'architecture GoogLeNet:

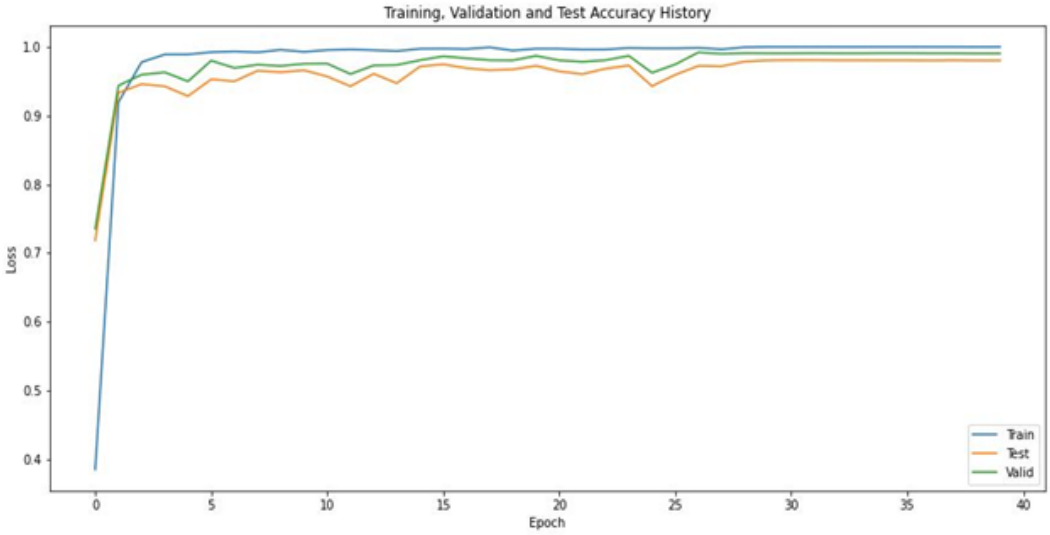


FIGURE 5.9: Performances de l'architecture GoogleNet : Accuracy.

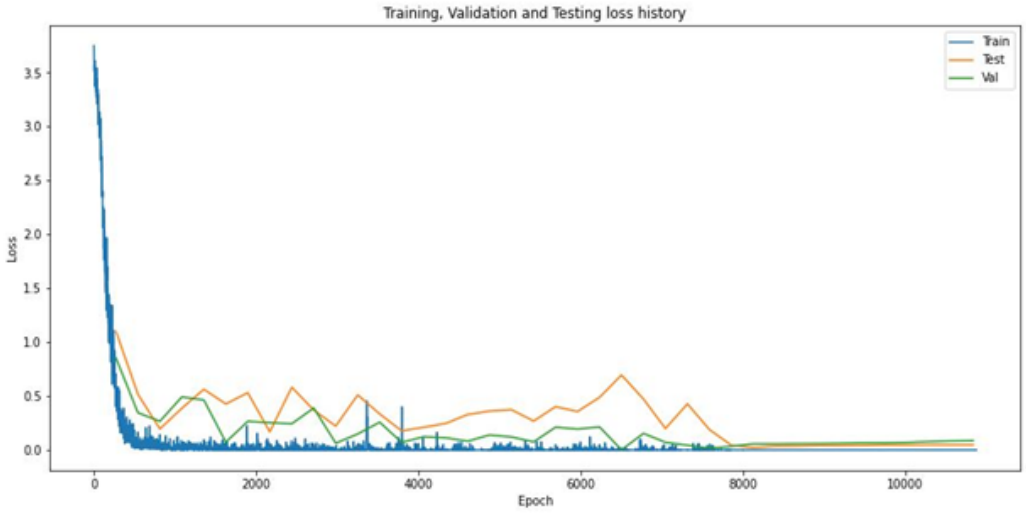


FIGURE 5.10: Performances de l'architecture GoogleNet : Fonction de perte.



FIGURE 5.11: Resultats de detection sur differents espaces de couleurs.Sur la première ligne les résultats obtenus via **AlexNet**, a deuxième les résultats obtenus via**leNet** et la troisième les résultats obtenus via **GoogLeNet**.

5.6.5 Comparaison entre les résultats des trois architectures.

La figure 5.6.4 presente les résultats de reconnaissance produis par les trois architectures. Le tableau 5.4 montre une comparaison avec les résultats des trois architectures: **LeNet**, **AlexNet** et **GoogLeNet**:

	Epoches	Temps d'exécution	Précision	Taux d'erreur
LeNet	100	3min 22s	0.944	/
AlexNet	100	5min 25s	0949	/
GoogLeNet	50	1h 38min 49s	0.980	0.048

TABLE 5.4: Performances des trois architectures.

5.7 Evaluation Globale

Les sections précédentes montrent une évaluation complète de chaque phase du système avec comparaisons des résultats. On conclue que pour la détection l'espace HSV présente des meilleurs résultats par rapport aux autres espaces. Quand a la reconnaissance l'architecture GoogleNet est la plus efficaces.

Dans cette section nous présenterons l'évaluation des deux phases d'un système de DRS. Il s'agit d'une évaluation de bout-en-bout en utilisant les méthodes les plus efficaces pour chaque phase. Le tableau suivant montre les résultats obtenus :

Classe de panneaux	NTP	NTD	NDR	NDNR	Detec.Acc	Recogn.Acc
Speed Limit (30 km/h)	31	26	26	0	84%	100%
Speed Limit (50 km/h)	22	19	18	1	86%	95%
Speed Limit (60 km/h)	8	6	6	0	75%	100%
Speed Limit (70 km/h)	35	21	21	0	60%	100%
Speed Limit (80 km/h)	16	14	13	1	88%	93%
Speed Limit (100 km/h)	2	2	1	1	100%	50%
Speed Limit (120 km/h)	10	7	7	0	70%	100%

TABLE 5.5: Evaluation des performances du système de bout en bout en utilisant l'architecture GoogleNet sur l'espace HSV.

Tel que :

- **NTP**: le nombre totale de panneaux;
- **NTD**: le nombre totale de panneau détectés;
- **NDR**: le nombre total de panneaux détecter et correctement reconnus;
- **NDNR**: Le nombre total de panneaux détecter et non reconnus.



(A) (B)

FIGURE 5.12: Exemple de résultats de bout en bout.

Discusions:

L'évaluation globale de ce système, pour la detection des panneaux le système a donné une précision de 60 à 100% pour les 7 classes différentes de panneaux de limitation de vitesse, avec une moyenne de 80%. Et pour la reconnaissance une précision de 50% à 100%, avec une moyenne de 91%.

Ces variations de précision sont dues au changement de luminance, à la taille des panneaux de signalisation sur les images, au flou de mouvement, etc. Et la raison la plus influente est la grande différence entre le nombre d'annotations pour les 7 classes différentes, certaines classes ne contiennent pas suffisamment d'annotations comme les panneaux de limitation de vitesse à 100 km/ h.

5.8 Conclusion

Dans ce chapitre, nous avons vu l'implémentation de l'approche proposée.Les résultats pour chaque phase ont etés présentés, comparer et discuter.

Chapitre 6

Conclusion et perspectives

Dans ce mémoire, nous avons présentés les systèmes de détection et de reconnaissance des panneaux de signalisation. Nous avons décrits les différentes phases constituant ce système : détection et reconnaissances. Pour une meilleure compréhension du problème et des difficultés nous avons étudiés les différents panneaux de signalisation, leurs classifications ainsi que défis que présent.

Après avoir cerné les problèmes, on est à fin de proposé une solution efficace nous avons présentés les méthodes et techniques les plus utilisés et les plus efficaces qui permettent la détection et la reconnaissance d'objets pour choisi c'elles répondent aux exigences du système. La méthode Haar Cascade a était sélectionnée pour la détection et trois architectures :GoogLeNet, AlexNet et LeNet pour la reconnaissance.

Les résultats qu'on a obtenus nous permettes de conclure que : pour la détection Haar permet une grande précision sur l'espace HSV. La reconnaissance via l'architecture GoogleNet et meilleures que c'elle des autres architectures utilisées.

Perspectives

Les méthodes proposées dans cette thèse ont montré de très bonnes performances. En perspective, on prévoit l'amélioration du système de détection et l'etendre sur de tous les types de panneaux de signalisation. On envisage aussi travailler sur des séquences videos plutot que des images fixes. Une autre perspective serait la mise en œuvre des méthodes sur des véhicules pour la validation dans un environnement réel.

Bibliographie

[1] Paul Viola, Michael Jones, et al. "Robust real-time object detection". In: *International journal of computer vision* 4.34-47 (2001), p. 4.

[2] Françoise Anceaux, Bako Rajaonah, and Paris V LEI. "PROJET SARI-PREDIT 3". In: (2006).

[3] Philippe Da Costa. *Sécurité routière et circulation: la responsabilité des différents acteurs*. Les éditions des journaux officiels, 2007.

[4] Nick Barnes, Alexander Zelinsky, and Luke S Fletcher. "Real-time speed sign detection using the radial symmetry detector". In: *IEEE Transactions on Intelligent Transportation Systems* 9.2 (2008), pp. 322–332.

[5] Muhammad Riaz et al. "Efficient image retrieval using adaptive segmentation of HSV color space". In: *2008 International Conference on Computational Sciences and Its Applications*. IEEE. 2008, pp. 491–496.

[6] Thomas Laloë. "Sur quelques problèmes d'apprentissage supervisé et non supervisé". PhD thesis. 2009.

[7] Álvaro González et al. "Automatic traffic signs and panels inspection system using computer vision". In: *IEEE Transactions on intelligent transportation systems* 12.2 (2011), pp. 485–499.

[8] Fredrik Larsson and Michael Felsberg. "Using Fourier descriptors and spatial models for traffic sign recognition". In: *Scandinavian conference on image analysis*. Springer. 2011, pp. 238–249.

[9] Johannes Stallkamp et al. "The German traffic sign recognition benchmark: a multi-class classification competition". In: *The 2011 international joint conference on neural networks*. IEEE. 2011, pp. 1453–1460.

[10] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. "Imagenet classification with deep convolutional neural networks". In: *Advances in neural information processing systems*. 2012, pp. 1097–1105.

[11] Andreas Mogelmose, Mohan Manubhai Trivedi, and Thomas B Moeslund. "Vision-based traffic sign detection and analysis for intelligent driver assistance systems: Perspectives and survey". In: *IEEE Transactions on Intelligent Transportation Systems* 13.4 (2012), pp. 1484–1497.

[12] Sebastian Houben et al. "Detection of traffic signs in real-world images: The German Traffic Sign Detection Benchmark". In: *The 2013 international joint conference on neural networks (IJCNN)*. IEEE. 2013, pp. 1–8.

[13] Markus Mathias et al. "Traffic sign recognition—How far are we from the solution?" In: *The 2013 international joint conference on Neural networks (IJCNN)*. IEEE. 2013, pp. 1–8.

[14] Markus Mathias et al. "Traffic sign recognition—How far are we from the solution?" In: *The 2013 international joint conference on Neural networks (IJCNN)*. IEEE. 2013, pp. 1–8.

[15] M Boumediene et al. "Détection, association et suivi de pistes pour la reconnaissance de panneaux routiers". In: 2014.

[16] Siniša Šegvić et al. "Exploiting temporal and spatial constraints in traffic sign detection from a moving vehicle". In: *Machine vision and applications* 25.3 (2014), pp. 649–665.

[17] Gangyi Wang et al. "Hole-based traffic sign detection method for traffic signs with red rim". In: *The Visual Computer* 30.5 (2014), pp. 539–551.

[18] Christian Szegedy et al. "Going deeper with convolutions". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015, pp. 1–9.

[19] Pavel Yakimov. "Traffic signs detection using tracking with prediction". In: *International Conference on E-Business and Telecommunications*. Springer. 2015, pp. 454–467.

[20] Yi Yang et al. "Towards real-time traffic sign detection and classification". In: *IEEE Transactions on Intelligent transportation systems* 17.7 (2015), pp. 2022–2031.

[21] Ali Güneş, Habil Kalkan, and Efkan Durmuş. "Optimizing the color-to-grayscale conversion for image classification". In: *Signal, Image and Video Processing* 10.5 (2016), pp. 853–860.

[22] Ilse M Harms and Karel A Brookhuis. "Dynamic traffic management on a familiar road: Failing to detect changes in variable speed limits". In: *Transportation research part F: traffic psychology and behaviour* 38 (2016), pp. 37–46.

[23] Romi Fadillah Rahmat et al. "Skin color segmentation using multi-color space threshold". In: *2016 3rd International Conference on Computer and Information Sciences (ICCOINS)*. IEEE. 2016, pp. 391–396.

[24] Yassmina Saadna and Ali Behloul. "An overview of traffic sign detection and classification methods". In: *International journal of multimedia information retrieval* 6.3 (2017), pp. 193–210.

[25] Jianming Zhang et al. "A real-time chinese traffic sign detection algorithm based on modified YOLOv2". In: *Algorithms* 10.4 (2017), p. 127.

[26] Pocholo James M Loresco, Ira C Valenzuela, and Elmer P Dadios. "Color space analysis using KNN for lettuce crop stages identification in smart farm setup". In: *TENCON 2018-2018 IEEE Region 10 Conference*. IEEE. 2018, pp. 2040–2044.

[27] Dip Nandi et al. "Traffic Sign Detection based on Color Segmentation of Obscure Image Candidates: A Comprehensive Study." In: *International Journal of Modern Education & Computer Science* 10.6 (2018).

[28] Sergey Andreev et al. "Dense moving fog for intelligent IoT: Key challenges and opportunities". In: *IEEE Communications Magazine* 57.5 (2019), pp. 34–41.

[29] Amal Bouti et al. "A robust system for road sign detection and classification using LeNet architecture based on convolutional neural network". In: *Soft Computing* (2019), pp. 1–13.

[30] Jingwei Cao et al. "Improved traffic sign detection and recognition algorithm for intelligent vehicles". In: *Sensors* 19.18 (2019), p. 4021.

[31] Zhenli He et al. "Traffic sign recognition by combining global and local features based on semi-supervised classification". In: *IET Intelligent Transport Systems* 14.5 (2019), pp. 323–330.

[32] Sarah Marie-Solveig Leclerc. "Automatisation de la segmentation sémantique de structures cardiaques en imagerie ultrasonore par apprentissage supervisé". PhD thesis. Université de Sherbrooke, 2019.

[33] Sarah Marie-Solveig Leclerc. "Automatisation de la segmentation sémantique de structures cardiaques en imagerie ultrasonore par apprentissage supervisé". PhD thesis. Université de Sherbrooke, 2019.

[34] Chunsheng Liu et al. "Machine vision based traffic sign detection methods: Review, analyses and perspectives". In: *IEEE Access* 7 (2019), pp. 86578–86596.

[35] Faming Shao et al. "Improved faster R-CNN traffic sign detection based on a second region of interest and highly possible regions proposal network". In: *Sensors* 19.10 (2019), p. 2288.

[36] Sai Subhakar Tirumaladasu and Shirdi Manjunath Adigarla. *Autonomous Driving: Traffic Sign Classification*. 2019.

[37] Hamidah Ibrahim, Nor Fazlida Mohd Sani, Razali Yaakob, et al. "An Indexed Non-probability Skyline Query Processing Framework for Uncertain Data". In: *International Conference on Advanced Machine Learning Technologies and Applications*. Springer. 2020, pp. 289–301.