



UNIVERSITE SULTAN MOULAY SLIMANE
Faculté des Sciences et Techniques
Béni-Mellal



Centre d'Études Doctorales : Sciences et Techniques
Formation Doctorale : Mathématique et Physique Appliquées (MPA)

THÈSE

Présentée par

SALHI ABDERRAHIM

Pour l'obtention du grade de

DOCTEUR

Spécialité : Electronique et Traitement du Signal

**Systeme de détection et reconnaissance automatiques des
panneaux de signalisation routière**

Soutenue le 28/12/2017 à 9h devant la commission d'examen composée de :

Président	: Pr. Saïd MELLIANI	PES	FST - Béni Mellal
Rapporteurs	: Pr. Belaid BOUIKHALENE	PH	FP – Béni Mellal
	Pr. Rachid EL AYACHI	PH	FST - Béni Mellal
Directeur de thèse	: Pr. Brahim MINAOUI	PES	FST - Béni Mellal
Co-directeur de thèse	: Pr. Mohamed FAKIR	PES	FST - Béni Mellal

Remerciement

Je tenais tout d'abord à remercier mon directeur de thèse, M. Brahim MINAOUI, qui m'a prodigué ses conseils tout au long de ma thèse. Il a fait preuve d'une grande écoute et de compréhension lors de nos discussions portant sur mes travaux de recherche ou le travail au sein de la faculté. Son soutien et son enthousiasme m'ont beaucoup aidé à surmonter les difficultés inhérentes à tout travail de thèse.

Je souhaiterais exprimer ma gratitude à mon co-encadrant, M. Mohamed FAKIR. Toujours disponible, ouvert et compréhensif, il m'a permis d'aborder mon travail sous d'autres perspectives et d'éviter de perdre de vue mes objectifs.

Je remercie également tous les membres de mon jury qui ont accepté d'évaluer mon travail et notamment Pr Said MELLIANI, président du jury, et mes rapporteurs : Pr Belaid BOUIKHALENE, Pr Rachid EL AYACHI et Pr Jilali ANTARI. Leur remarques et questions m'ont été d'une grande aide pour améliorer mon rapport. Ils m'ont permis d'aborder mes travaux avec un œil neuf, nécessaire après trois années, et de concilier les points de vue académiques, industriels et utilisateur final.

Je saisi l'occasion pour remercier mes collègues à la faculté des sciences et techniques de Béni Mellal, administratifs et enseignants sans exception, et plus particulièrement, le doyen de la faculté M. Ahmed ZEGHAL, et le Staff du département de physique, qui m'ont prodigué leurs conseils et leurs aides morales et physiques tout au long de ma thèse.

En dernier lieu, je souhaite vivement remercier tous mes collègues du Laboratoire de Traitement de l'Information et l'Aide à la Décision (LTIAD) pour tous les bons moments partagés pendant ma thèse. Une bonne ambiance et des personnes toujours prêtes à aider et faire partager leur connaissance sont des éléments indispensables à tout thésard.

Résumé

Dans ce travail, nous avons proposé un système d'identification des panneaux de signalisation routière. Ce système intègre des techniques d'analyse sémantique appropriées permettant une détection et reconnaissance prometteuses des panneaux routiers. La détection enchaine en cascade deux types de critères discriminants : la couleur et la forme. La couleur est utilisée dans un premier temps, pour extraire les zones d'intérêt contenant éventuellement des panneaux : ceci a permis de réduire le temps de traitement pour, déterminer par la suite, la forme de ces panneaux en se basant sur la modélisation de leurs contours par l'approximation polygonale. Les panneaux détectés selon ces deux critères sont répartis en quatre classes : panneaux triangulaires rouges, circulaires rouges, circulaires bleus et rectangulaires bleus. La méthode de détection ainsi proposée est appliquée avec succès sur plusieurs images, de la scène routière aussi bien en terme de précision que de rapidité de détection.

Quant à la reconnaissance, elle consiste à identifier les panneaux, détectés précédemment, pour le classificateur « Support Vector Machine » SVM à partir des descripteurs « Histogram Projection » HP.

Une évaluation expérimentale du système complet (module de détection + module de reconnaissance) a été effectuée sur des images de la scène routière en mode (offline), les résultats obtenus prouvent la fiabilité du système comme étant un système d'aide à la conduite. Ceci est confirmé par des tests en temps réel (mode online) en embarquant un ordinateur avec un caméscope à bord d'une voiture et aussi en utilisant une carte Raspberry Pi B+.

Mots clefs:

Panneaux de signalisation routière ; Détection de panneaux ; Reconnaissance de panneaux ; Classificateur SVM ; Descripteur HP ; Vision par ordinateur ; Système embarqué ; Système d'aide à la conduite.

abstract

In this work, we proposed a system of identification of road signs. This system incorporates appropriate semantic analysis techniques for promising detection and recognition of road signs. Detection cascades two types of discriminating criteria: color and shape. The color is used initially, to extract the areas of interest possibly containing road signs: this made it possible to reduce the processing time to determine, later, the shape of these road signs based on the modulation of their outlines by the polygonal approximation. The road signs detected according to these two criteria are divided into four classes: red triangular signs, red circular, circular blue and rectangular blue. The detection method thus proposed is successfully applied to several images of the traffic scene, with good results both in terms of accuracy and speed of detection.

As for the recognition, it consists in identifying the previously detected panels for the SVM " Support Vector Machine " classifier from the HP "Histogram Projection" descriptors.

An experimental evaluation of the complete system (detection module + recognition module) was performed on images of the traffic scene in (offline) mode, the results obtained prove the reliability of the system as a driver assistance system. This is confirmed by real-time tests (online mode) by embedding a computer with a camcorder in a car and by using a Raspberry Pi B + card.

Keywords:

Traffic signs; Panel detection; Panel recognition; SVM classifier; HP descriptor; Computer vision; Embedded system; Driving assistance system.

ملخص

في هذا العمل، اقترحنا نظاما لتحديد العلامات الطرفية. ويشمل هذا النظام تقنيات التحليل الدلالي المناسبة التي تمكن من تحقيق كشف وتعرف واعدن في تحديد العلامات الطرفية. عملية الكشف تطبق على التوالي نوعين من المعايير المميزة: اللون والشكل. يستخدم اللون في البداية لاستخراج المساحات المهمة والتي قد تحتوي على علامات طرفية. ويقلل هذا من الوقت المستغرق في معالجة المعلومة، ليتيح مزيدا من الوقت بعد ذلك لتحديد شكل هذه العلامات، اعتمادا على تصميم خاص لمحيطاتها باستعمال التقريب المضلعي للخطوط المنحنية. وتقسّم العلامات المكشوفة بعد ذلك إلى أربع فئات هي: علامات مثلثية حمراء، دائرية حمراء، دائرية زرقاء ومستطيلة زرقاء. ومن ثم يتم تطبيق طريقة الكشف المقترحة بنجاح على عدة صور لمشهد الطريق سواء من حيث الدقة وسرعة الكشف.

أما بالنسبة لطريقة التعرف على العلامات والمكتشفة في المرحلة السابقة، فنستعمل المصنف SVM على المميز HP لصور هذه العلامات.

أجري تقييم تجريبي للنظام بأكمله (وحدة الكشف + وحدة التعرف) على صور من المشهد الطريقي وأثبت النتائج المحصل عليها موثوقية النظام كنظام مساعدة السائق. وهذا ما تؤكدته التجارب في الوقت الحقيقي عن طريق دمج جهاز كمبيوتر مع كاميرا على متن سيارة وأيضا باستخدام بطاقة Raspberry Pi B+.

كلمات مفتاح:

علامات الطريق؛ كشف العلامات؛ التعرف على العلامات؛ المصنف SVM؛ المميز HP؛ الرؤية بالكمبيوتر؛ نظام مضمن؛ نظام مساعدة السائق.

Table des matières

Remerciement	i
Résumé	ii
abstract	iii
ملخص	iv
Table des matières	v
Liste des tableaux	ix
Liste des figures	ix
Introduction	1
1. Contexte et motivation	1
2. Détection et reconnaissance des panneaux routiers	2
3. Sommaire des chapitres de la thèse	3
Chapitre I . Etat de l'art	5
I.1. Introduction	6
I.2. Acquisition des images	6
I.3. Détection des panneaux de signalisation routière.	8
I.3.1. Utilisation de l'information couleur	8
I.3.1.1. Espace de couleur RGB	9
I.3.1.2. Espace de couleur HSV	10
I.3.1.3. Conclusions	12
I.3.2. Utilisation de la forme	12
I.3.2.1. Contraintes sur l'entièreté de l'objet	12
I.3.2.2. Contraintes sur les contours de l'objet	14
I.3.3. Le détecteur de Viola et Jones	15
I.4. Reconnaissance des signaux	17

I.4.1. Réseaux de neurones.....	17
I.4.2. Corrélation normalisée.....	20
I.4.3. Classeurs polynomiaux.....	20
I.4.4. Machine à vecteurs supports.....	21
I.4.5. Modèle génératif.....	22
I.4.6. Ensembles d'arbres.....	23
I.5. Conclusion.....	24
Chapitre II . Méthodologie, outils et environnement de travail -----	25
II.1. Introduction.....	26
II.2. Méthodologie de travail.....	26
II.3. Structure générale du programme.....	26
II.4. Signalisation routière marocaine.....	28
II.5. Ressources matérielles et logicielles utilisées.....	29
II.5.1.1. Caméscope.....	29
II.5.1.2. Machines.....	31
II.5.1.3. Logiciels.....	32
II.6. Conclusion.....	32
Chapitre III . Module de détection des panneaux routiers -----	33
III.1. Introduction.....	34
III.2. Méthode de détection.....	34
III.3. Détection de la couleur des panneaux.....	34
III.4. Détection des formes.....	37
III.4.1. Extraction des contours.....	37
III.4.2. Estimation de la forme des panneaux.....	37
III.5. Evaluation du module de détection.....	41

III.5.1.1. Premier test	41
III.5.1.2. Deuxième test	42
III.5.1.3. Analyse des résultats.....	43
III.6. Conclusion.....	43
Chapitre IV . Module de reconnaissance des panneaux routiers-----	44
IV.1. Introduction	45
IV.2. Classification.....	45
IV.3. Méthode de la reconnaissance	46
IV.4. Descripteur HP	46
IV.5. Descripteur HOG	47
IV.6. Classificateur MLP.....	49
IV.7. Classificateur SVM.....	51
IV.8. Préparation de la base des images de test	54
IV.9. Evaluation de module de reconnaissance.....	55
IV.9.1. Test sur les panneaux triangulaires rouges	56
IV.9.2. Test sur les panneaux circulaires rouges.....	57
IV.9.3. Test sur les panneaux circulaires bleus et rectangulaires bleus	58
IV.9.4. Analyses et conclusions	59
IV.10. Conclusion.....	62
Chapitre V . Système complet-----	63
V.1. Introduction.....	64
V.2. Structure générale du système complet	64
V.3. Test hors ligne	67
V.3.1. Test sur ordinateur.....	67
V.3.2. Test sur Raspberry Pi B+	69

V.4. Test en ligne	70
V.4.1. Test sur ordinateur.....	70
V.4.2. Test sur Raspberry Pi B+	71
Conclusions et perspectives-----	73
Publications -----	75
Communications -----	76
Bibliographie -----	77

Liste des tableaux

Tableau 1: Résultat d'évaluation du module de détection sur des images avec arrière-plans blancs	42
Tableau 2: Résultat d'évaluation du module de détection sur des images avec arrière-plans réels	42
Tableau 3: Distances de visibilité des panneaux routiers en fonction de la vitesse d'approche ..	61
Tableau 4: résultats expérimentaux en mode hors ligne du système complet TSDS	68
Tableau 5: Nombre de fois de traitement d'un panneaux routiers par le TSDS en fonction de la vitesse d'approche.....	71

Liste des figures

Figure 1: Image de scène routière.	2
Figure 2: Schéma synoptique du système utilisé dans [7].....	7
Figure 3: Lookup tables sur la teinte (à gauche) et la saturation (à droite) utilisées.....	11
Figure 4: Modèles utilisés pour le Template Matching dans [13].....	14
Figure 5: Une image binaire et sa transformée en distance (distance euclidienne).....	14
Figure 6: Exemples de chiffres de taille 7x5 utilisés pour lors de la reconnaissance des signaux de limitation de vitesse.....	18
Figure 7: Architecture du réseau de neurones adaptatif à temps de retard utilisé dans [6].	20
Figure 8: Architecture d'un classeur polynomial quadratique utilisé dans [6].	21
Figure 9: Principe d'une SVM.	22
Figure 10: Architecture du système de reconnaissance utilisé par Bahlmann et al.....	22
Figure 11: Structure générale du programme.....	27
Figure 12: Plateforme de test du Système TSDS embarqué.....	28
Figure 13: Exemple d'images de signaux routiers.....	29
Figure 14: Caméscope utilisé pour l'acquisition des séquences vidéo.	30
Figure 15: Ordinateur HP i3 utilisé dans le projet.....	31
Figure 16: Carte de Raspberry Pi B+	31

Figure 17: Structure du programme de détection des panneaux routiers.	35
Figure 18: Résultats de détection de couleur. a) Image initiale, b) Masque de couleur bleu et c) Masque de couleur rouge	36
Figure 19: Approximation polygonale du contour.	39
Figure 20: Résultats de détection des panneaux. a) Image initiale, b) imagerie des panneaux détectés et rognés, c) imagerie redimensionnées en 32X32 pixels	41
Figure 21: Schéma synoptique du module de reconnaissance.	46
Figure 22: structure de la multi couche perceptron.....	49
Figure 23: Structure de programme de génération des bases de données de test.....	55
Figure 24: Organigramme générique d'entraînement et de test des quatre combinaisons « descripteur-classificateur ».....	56
Figure 25: Taux de reconnaissance pour les panneaux triangulaires rouges.....	57
Figure 26: Taux de reconnaissance pour les panneaux circulaires rouges.....	58
Figure 27: Taux de reconnaissance pour les panneaux circulaires et rectangulaires bleus.....	59
Figure 28: Taux de reconnaissance globales.....	60
Figure 29: Temps de reconnaissance globales	61
Figure 30: Schéma Bloc du système complet TSDS.....	65
Figure 31: Organigramme général de système complet TSDS.	66
Figure 32: Echantillon des images de la scène routière.	67
Figure 33: Capture d'écran du système complet TSDS.....	68
Figure 34: Capture d'écran du système complet TSDS sur la carte Raspberry Pi B+.....	70

Introduction

1. Contexte et motivation

La détection et la reconnaissance automatiques des panneaux de signalisation routière (PSR) représentent un enjeu important en analyse de scènes routières. Leurs applications sont nombreuses. Citons par exemple les systèmes d'aide à la conduite, la gestion du patrimoine routier, la sécurité routière, ou encore la mise au point des outils multimédia sur le web pour la navigation 3D géographique, ou l'amélioration des logiciels de calcul d'itinéraire.

Sur le plan méthodologique, elles présentent d'importantes difficultés, liées à la fois au caractère non contrôlé des prises de vues employées, à l'intensité de la luminosité ambiante, aux conditions climatiques, et à la variabilité d'aspect des objets recherchés. Ces problématiques ont donc, naturellement, donné lieu à une grande variété de contributions au cours de ces deux dernières décennies. Dans la littérature liée à ce domaine, on trouve des techniques (souvent motivées par le contexte temps-réel des applications d'aide à la conduite) exploitant judicieusement les informations de forme et de couleurs extraites des images par des algorithmes "bas-niveau", ainsi que des techniques plus sophistiquées de reconnaissance des formes, mettant en jeu des techniques d'apprentissage statistique ou des problèmes complexes d'optimisation.

L'objectif de ce travail est la conception et la réalisation d'un système de détection et de reconnaissance de panneaux automatique des panneaux de signalisation routière en temps réel, en vue de construire un outil d'aide à la conduite performant et efficace.

Sur le plan technique, les images numériques acquises par des véhicules en circulation sur le réseau routier présentent une masse importante d'informations à traiter. Dans le but de faciliter l'extraction de l'information pertinente, cette thèse propose des méthodes d'analyse automatique d'images de scènes routières. L'information recherchée concerne plus particulièrement la couleur et la géométrie des panneaux de signalisation routière, qui doivent être détectés, puis reconnus. La stratégie générale de recherche consiste à effectuer une détection des objets d'intérêt pouvant inclure des fausses alarmes mais si possible sans oublis par un module de détection. Ces résultats sont affinés et reconnus ultérieurement par un module de reconnaissance.

Des séquences d'images couleur sont acquises par une caméra montée sur une voiture en circulation sur le réseau routier. Les scènes à analyser sont des scènes dynamiques extérieures

prises de jour et de nuit sous un éclairage naturel, et sous des conditions climatiques différents. La nature des scènes routières peut varier d'un environnement urbain ou rural, provenant de routes provinciales ou nationales. La Figure 1 présente un exemple d'image de scène routière à analyser ainsi que quelques objets d'intérêt à localiser.



Figure 1: Image de scène routière.

2. Détection et reconnaissance des panneaux routiers

La détection et la reconnaissance automatiques d'objets sont des tâches essentielles dans plusieurs applications telles que la vidéo-surveillance ou l'indexation par le contenu. Les méthodes classiques en vision par ordinateur reposent sur la comparaison de caractéristiques visuelles d'une observation avec des références préalablement apprises par le système. Prendre en compte toutes les variations d'apparence d'un objet dans une image est un défi important. Celles-ci dépendent essentiellement dans le cas des panneaux de signalisation routière de trois facteurs : la couleur et la géométrie extérieure du panneau, la scène et le système d'acquisition.

Les panneaux de signalisation routière peuvent être classés selon leur couleur en deux catégories : les panneaux rouges et les panneaux bleus, et selon leur géométrie extérieure en trois catégories : les panneaux circulaires, les panneaux triangulaires et les panneaux rectangulaires.

Les panneaux de signalisation routière varient selon les conditions de prises de vue qui ne sont pas contrôlées : l'illumination incidente sur la scène dépend des conditions météorologiques et horaires. Les changements de position du soleil et de son intensité impliquent des changements d'apparence des couleurs des panneaux de signalisation routière dans les images.

Les panneaux de signalisation routière appartiennent selon différents points de vue et à différentes échelles, dépendamment de la position du système d'acquisition dans la scène, Par ailleurs, ils peuvent être partiellement cachés ou dans l'ombre d'un autre objet. En plus des effets d'occultations et d'ombres portées, les reflets et les contre-jours liés à la position de la caméra par rapport au soleil, introduisent aussi des variabilités dans l'apparence d'un panneau de signalisation routière. Enfin, notre système d'acquisition n'est pas calibré photométriquement, et l'on ne dispose pas de référence de couleur fiable.

3. Sommaire des chapitres de la thèse

La détection et la reconnaissance automatiques des objets est un thème de recherche scientifique qui intègre les techniques d'analyse permettant l'accès rapide à des images ou des objets particuliers, dans les bases d'images et de vidéos, généralistes ou spécialisées. Les systèmes de détection et de reconnaissance des panneaux de signalisation routière adoptent généralement l'approche de deux étapes. La première, appelé étape de détection qui exploite la particularité des couleurs et de la géométrie de ces panneaux, et met en œuvre des méthodes de transformation et de segmentation en couleur des images de la scène routière, ainsi qu'un ensemble de techniques de détection de contours et de formes pour l'extraction des régions d'intérêt susceptibles de contenir des panneaux de signalisation routière. Ceux-ci font l'objet de la deuxième étape dite de classification qui utilise des algorithmes de classification pour identifier leurs types et leurs significations.

Dans le premier chapitre nous présenterons l'état de l'art sur les systèmes d'acquisition d'image ; la matière première de ce projet de recherche, et sur la détection des panneaux de signalisation routières tout en présentant les différentes méthodes et techniques utilisées dans ce domaine, nous aborderons aussi dans ce chapitre les techniques de reconnaissances des panneaux routiers en discutant les différents descripteurs et classificateurs utilisés par les auteurs dans la littérature.

Introduction

Dans le deuxième chapitre qui suit, nous étalerons la méthodologie de travail suivie et la structure générale de programme à réaliser pour atteindre l'objectif de ce projet, nous présenterons aussi les diverses ressources matérielles et logicielles requises lors du développement du système de détection et de reconnaissance de panneaux de signalisation routière.

Le module de détection des panneaux de signalisation routière est le sujet du troisième chapitre, où nous aborderons en détails la méthode de détection développée pour ce projet, de la détection des couleurs et des formes particulières des panneaux, jusqu'à leur extraction. Nous trouverons aussi dans ce chapitre l'ensemble des tests menés pour évaluer les performances de ce module que ce soit sur des images en mode offline ou en temps réel en mode online.

Nous présenterons dans le quatrième chapitre les descripteurs et les classificateurs utilisés lors de la conception du module de reconnaissance des panneaux de signalisation routière, ainsi que les comparaisons entre les quatre combinaisons descripteur-classificateur. Les résultats des tests expérimentaux sont commentés à la fin du chapitre.

Le cinquième chapitre est consacré à la présentation de l'association des deux modules de détection et de reconnaissance des panneaux routiers dans un système complet. Nous discuterons ainsi les différentes associations effectuées, tout en analysant l'ensemble des résultats des tests expérimentaux faits sur des images fixes de la scène routière, ou en temps réel.

Finalement nous terminerons ce rapport de thèse par une conclusion, sans oublier de citer quelques perspectives de cet ambitieux projet dans le domaine de l'intelligence artificielle.

Chapitre I . Etat de l'art

Sommaire

I.1. Introduction.....	6
I.2. Acquisition des images.....	6
I.3. Détection des panneaux de signalisation routière.....	8
I.3.1. Utilisation de l'information couleur	8
I.3.2. Utilisation de la forme	12
I.3.3. Le détecteur de Viola et Jones.....	15
I.4. Reconnaissance des signaux.....	17
I.4.1. Réseaux de neurones	17
I.4.2. Corrélation normalisée	20
I.4.3. Classeurs polynomiaux.....	20
I.4.4. Machine à vecteurs supports	21
I.4.5. Modèle génératif.....	22
I.4.6. Ensembles d'arbres.....	23
I.5. Conclusion	24

I.1. Introduction

De nombreux travaux ont été réalisés au sujet de la détection et la reconnaissance automatiques des panneaux de signalisation routière. La plupart travaillent image par image et adoptent une approche en deux phases. Une première phase de détection consiste à détecter et à localiser les régions d'intérêt susceptibles de contenir les panneaux routiers. Ensuite, une deuxième phase de reconnaissance permet de déterminer le type de panneau contenu dans chaque région d'intérêt. L'efficacité de la méthode de détection est très importante. On doit tenter de détecter tous les signaux présents dans l'image et de rejeter au maximum les détections d'autres objets, car celles-ci rendraient le travail de reconnaissance plus délicat. Certains travaux ajoutent une étape entre la détection et la reconnaissance : le suivi des signaux au travers d'un flux vidéo. Le suivi permet de faciliter grandement la reconnaissance puisque grâce à lui, celle-ci peut se baser sur une série d'images pour un même panneau routier.

Ce chapitre donnera un aperçu des méthodes utilisées dans les récents travaux de recherche. Premièrement, la section I.2 présentera brièvement les moyens mis en place pour l'acquisition des images. Ensuite, dans la section I.3, nous examinerons les techniques utilisées pour la détection et la localisation des panneaux routiers. Enfin dans la section I.4, nous traiterons des techniques utilisées pour la reconnaissance des panneaux de signalisation routière. Les méthodes dont nous nous servirons seront décrites plus en détails dans les chapitres suivants.

I.2. Acquisition des images

Généralement les travaux de recherche en détection et reconnaissance automatiques de signaux routiers ont pour objectif la construction d'un outil d'assistance à la conduite. Cet outil devrait être capable de fonctionner en temps réel et devrait comporter un module dédié à l'acquisition du flux vidéo. Cependant, les recherches se concentrent essentiellement sur les parties détection et reconnaissance des signaux. Les travaux effectués se basent soit sur des images fixes (off-line), soit sur des séquences vidéo acquises en conditions réelles (on-line).

Étant donné la complexité du problème de détection et de reconnaissance automatiques des signaux, certains auteurs se limitent à travailler sur des images fixes. Les

images peuvent être obtenues en se rendant directement sur le terrain pour prendre des photographies des panneaux désirés [1]. Ce travail est certainement fastidieux et long. D'autres auteurs utilisent des images fixes obtenues à partir de séquences vidéo [2] suite de certains travaux, des bases de données d'images de panneaux sont mises en ligne à disposition du public [3]. Elles ne sont malheureusement pas toujours représentatives et de taille importante.

Pour faire l'acquisition des séquences vidéo, plusieurs méthodes sont envisageables. Les auteurs du travail [4] utilisent un caméscope placé au niveau du pare-brise. Ils travaillent sur autoroute à des vitesses supérieures à 90 km/h. Ils utilisent le mode entrelacé, qui donne de meilleurs résultats à ces vitesses. D'autres placent la caméra sur le toit du véhicule [5]. Certains auteurs travaillent avec un zoom et une orientation fixe [6] et [7], ils travaillent avec deux caméras : une caméra grande angle et une caméra à téléobjectif. La caméra grand angle détecte et localise les signaux. Lorsqu'un signal est suivi dans la séquence vidéo depuis un certain temps, on focalise la caméra à téléobjectif sur lui. Grâce à cette caméra, on continue le suivi et on réalise la reconnaissance. Une illustration du système complet est représentée sur la Figure 2.

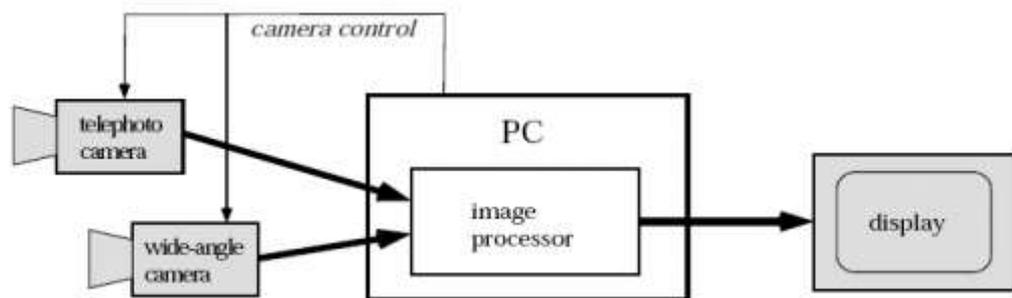


Figure 2: Schéma synoptique du système utilisé dans [7].

Certains auteurs travaillent directement sur des séquences vidéo, plusieurs d'entre eux utilisent l'information temporelle disponible pour effectuer un suivi des signaux au travers des séquences [8]. La plupart des auteurs travaillent off-line, peu de travaux comportent des tests temps réel comme ceux réalisés dans le travail [9].

L'acquisition sur un véhicule en mouvement (parfois à grande vitesse), la qualité des séquences obtenues est variable. On peut obtenir des images floues. Certains travaux visent

à régler ce problème, par exemple en se servant des marques longitudinales qui délimitent les bandes de circulation [10].

Pour notre travail, dans un premier temps travaillons off-line sur des bases de données d'images et de vidéos de la scène routière, pour tester l'efficacité de nos algorithmes, Après nous réalisons l'acquisition de séquences vidéo à partir d'un caméscope monté dans un véhicule. Cela nous permet de tester et optimiser nos algorithmes en temps réel, et pouvoir ainsi faire un suivi des panneaux routiers.

I.3. Détection des panneaux de signalisation routière.

Lors de la conduite de la voiture, nous n'avons aucun mal à déterminer où se situent les panneaux routiers dans notre champ de vision. Il est difficile de comprendre comment notre cerveau réalise cette tâche. Néanmoins, il paraît évident que plusieurs éléments jouent un rôle : les panneaux de signalisation routière ont des couleurs et des formes géométriques particulières qui dénotent du reste du paysage. Aussi les méthodes classiques de détection des panneaux de signalisation routière utilisent deux types d'information : les couleurs et la forme géométrique des panneaux. La section I.3.1 Présente des méthodes liées à l'information couleur et la section I.3.2 traite des techniques utilisant la forme géométrique.

I.3.1. Utilisation de l'information couleur

Les panneaux de signalisation routière ont des couleurs caractéristiques normalisés même si celles-ci se dégradent à cause de plusieurs facteurs, par exemple le vieillissement du panneau, les conditions climatiques etc. Plusieurs auteurs travaillent sur les panneaux à bords rouges. Ils utilisent cette information de couleur pour segmenter les images de la scène routière.

Certains isolent directement les pixels à forte composante rouge, d'autres segmentent les images en différentes régions et labellisent ces régions selon leur couleur afin de se servir de la disposition spatiale entre régions de couleurs différentes pour une étape ultérieure. Pour la segmentation, certains auteurs utilisent des réseaux de neurones ou des classeurs polynomiaux [9], mais la plupart optent pour une approche plus intuitive et relativement simple.

I.3.1.1. Espace de couleur RGB

En raison des variations de l'illumination suivant l'heure, la saison ou les conditions météorologiques, il n'est pas conseillé d'utiliser les canaux classiques RGB. Certains auteurs se basent quand même sur l'espace RGB, principalement pour économiser le temps de calcul lié à la conversion entre espaces de couleur. Ces auteurs doivent alors se prémunir des effets des variations de luminosité. Dans la plupart des cas, les auteurs fixent des seuils α , β et γ et des conditions à respecter pour les valeurs : rouge R, verte G et bleue B d'un pixel pour que celui-ci soit considéré comme rouge.

Les auteurs fixent les valeurs des seuils le plus souvent après un échantillonnage des valeurs des pixels des signaux dans différentes images. Voici les conditions utilisées par différents auteurs :

– Shneier [5] :

$$\frac{R}{G} > \alpha_m \text{ et } \frac{R}{B} > \beta_m \text{ et } \frac{G}{B} > \gamma_m \quad (1)$$

– Escalera et al. [11] et Shojania [12] :

$$\alpha_M > R > \alpha_m \text{ et } \beta_M > \frac{G}{R} > \beta_m \text{ et } \gamma_M > \frac{B}{R} > \gamma_m \quad (2)$$

– Torresen et Sekanina [13] ainsi que Benallal et Meunier [14] :

$$R > \alpha_m \text{ et } (R - G) > \beta_m \text{ et } (R - B) > \gamma_m \quad (3)$$

Pour des luminosités extrêmes (faible ou forte), les seuils sont adaptés dans [13] pour plus de robustesse.

Dans le même ordre d'idée, Zadeh et al. [1] définissent une région cylindrique dans l'espace RGB pour effectuer la segmentation. Yang [15], pour sa part, utilise un seuillage très léger défini par : $2R > G + B$. Ce seuillage lui permet néanmoins d'éliminer une grande partie des pixels de l'image.

I.3.1.2. Espace de couleur HSV

Certains auteurs utilisent l'espace HSV¹, espace de couleurs moins sensible aux variations d'illumination. Dans cet espace, la teinte H permet de caractériser la couleur d'un pixel, la saturation S sa pureté, et la valeur V sa luminosité. L'avantage est que la teinte est peu dépendante de la luminosité, mais elle est instable lorsque la saturation est faible.

Vitabile et al. [6] utilisent une technique d'agrégation dynamique des pixels. Ils filtrent les pixels de l'image qui ont des coordonnées H, S et V proches de la coordonnée standard pour les pixels des signaux. Ils divisent l'image en régions rectangulaires et choisissent, dans chaque région, un pixel source² parmi les pixels filtrés. Dans chaque région, un pixel est considéré comme similaire au pixel source si la distance euclidienne dans l'espace cylindrique HSV entre les deux pixels est inférieure à un seuil t , non linéaire avec la saturation et défini par :

$$t = k - \sin(S_{Seed}) \quad (4)$$

Où k est un paramètre de normalisation et S_{Seed} est la saturation du pixel source, comprise entre 0 et 1. Le seuil devient donc plus grand pour les faibles valeurs de saturation. Ceci permet de réduire les problèmes liés à l'instabilité de la teinte lorsque la saturation est faible. Les pixels similaires aux pixels sources sont conservés. Dans chaque région, la segmentation de l'image s'effectue donc par agrégation des pixels à partir du pixel source. Escalera et al. [11] travaillent uniquement sur les composantes H et S.

Ils appliquent aux images deux lookup tables (voir Figure 3) et multiplient ensuite les images obtenues. Par rapport au classique ET logique entre images obtenues, la multiplication permet aux deux composantes (teinte et saturation) de « s'entraider » lorsqu'une des deux a une faible valeur, afin de passer le seuillage appliqué après la

¹ Hue Saturation Value.

² Seed Pixel

multiplication.

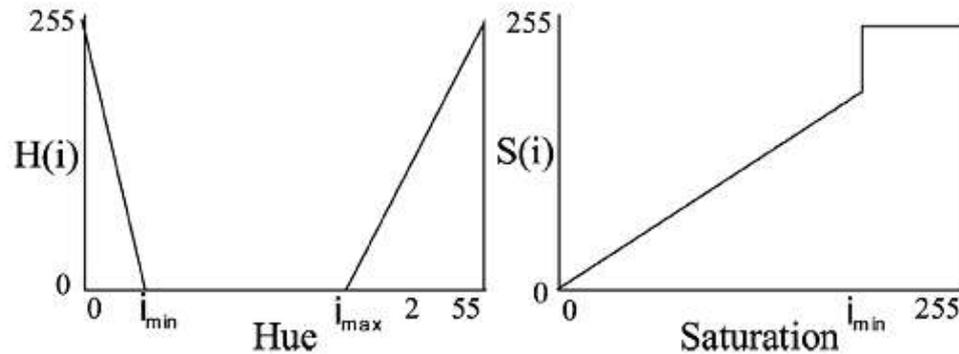


Figure 3: Lookup tables sur la teinte (à gauche) et la saturation (à droite) utilisées.

Shadeed [16] allie les informations de l'espace HSV et de l'espace YUV³. Il effectue un seuillage combiné sur la teinte H et sur les composantes U et V. Ce seuillage est précédé par une égalisation d'histogramme et un contrôle de la luminosité effectué sur le canal luminance Y.

Fang et al. [4] supposent de leur côté que chaque couleur particulière d'un signal peut être représentée par une valeur de teinte distribuée de manière gaussienne avec une variance σ^2 . L'ensemble de toutes ces valeurs de teinte est noté $\{h_1, h_2, h_q\}$. Les auteurs calculent un degré de similarité z entre la teinte h d'un pixel et les teintes h_k des couleurs de signaux :

$$z = \max_{k=1,2,\dots,q} z_k \text{ avec } z_k = \frac{1}{\sqrt{2\pi}} e^{-\frac{(h-h_k)^2}{2\sigma^2}} \quad (5)$$

Ainsi, ils obtiennent, non pas une image segmentée, mais une image où chaque valeur de pixel représente la similarité entre la couleur du pixel et les couleurs standard des pixels de signaux.

³ Espace colorimétrique notamment utilisé dans les systèmes de transmission télévisuelle PAL et NTSC. L'information couleur est contenue dans les composantes u et v et l'information luminance est contenue dans la composante Y

I.3.1.3. Conclusions

Dans la plupart des cas, les traitements qui viennent d'être décrits donnent en sortie une image segmentée. On ne peut pas vraiment comparer les performances des méthodes, elles sont liées le plus souvent aux seuils choisis. Ceux-ci ne doivent d'ailleurs pas être fixés précisément. Le but est d'éliminer une grande partie des pixels de l'image tout en conservant presque tous les pixels des signaux, ceci afin de faciliter les opérations qui suivront. La proportion de pixels que l'on souhaite supprimer est d'ailleurs fonction de ces opérations. D'autres objets que les signaux sont segmentés, des phares de voitures ou des enseignes par exemple. Ceci n'est pas grave tant que l'on se donne les moyens d'éliminer ces objets par d'autres traitements.

I.3.2. Utilisation de la forme

Les signaux routiers ont des formes de figures géométriques remarquables : cercle, triangle, carré, rectangle, losange, octogone. La forme particulière d'un signal permet de le distinguer de nombreux autres objets présents sur une image. L'information donnée par la forme est généralement utilisée à la suite d'une première segmentation basée sur l'information couleur. Généralement, on diminue d'abord le bruit dans l'image segmentée par des opérations morphologiques. La plupart des auteurs calculent ensuite les composantes connexes de l'image et labellisent les objets obtenus. Certains auteurs travaillent sur la totalité des pixels de l'objet. D'autres se contentent de l'information fournie par les contours.

Beaucoup de méthodes sont intéressantes, mais les auteurs sont rarement précis sur les résultats obtenus. Nous présentons ici, parfois très brièvement, certaines méthodes. Même s'il n'existe pas de technique universelle, nous nous forgerons ainsi une idée sur l'éventail de techniques disponibles.

I.3.2.1. Contraintes sur l'entièreté de l'objet

Pour trouver les signaux routiers dans l'ensemble d'objets segmentés, certains auteurs dans [5] [2] utilisent des contraintes sur des caractéristiques géométriques simples des objets :

- L'aire ;
- La hauteur et la largeur ;

- Le rapport hauteur sur largeur ;
- Le rapport entre l'aire de l'objet et l'aire du rectangle dans lequel il est inscrit ;
- ...

Malheureusement, ces contraintes ne suffisent pas à dissocier les signaux routiers de certains autres objets comme les feux de voiture par exemple. Franke et al. [9] travaillent avec des régions labellisées suivant 6 classes de couleurs. En plus des contraintes imposées ci-dessus, ils ajoutent des contraintes sur les relations de couleur entre :

- Les régions adjacentes ;
- Les régions incluses dans une autre région.

Dans une autre approche, Vitabile et al. [6] travaillent sur des objets considérés comme rouges. Ils calculent sur chacun des objets une mesure de similarité avec un ensemble d'images binaires représentant chacune une forme de signal routier.

Les modèles étant de taille fixe (36×36 pixels), les régions d'intérêt sont remises à la bonne échelle avant la mesure de similarité. La fonction de similarité utilisée est le coefficient de Tanimoto : si χ et Υ sont les ensembles de pixels des images binaires comparées, le coefficient de Tanimoto s est défini par :

$$s = \frac{\chi \cap \Upsilon}{\chi \cup \Upsilon} \quad (6)$$

Remarquons que $0 \leq s \leq 1$ et que s augmente avec la similarité entre le modèle et la région d'intérêt testée. Les auteurs ont effectué des tests sur 620 images de 24 classes. Ils obtiennent plus de 86 % de détections. Torresen et Sekanina [13] travaillent sur des signaux de limitation de vitesse, à partir d'images dans lesquelles les objets sont labellisés suivant 3 classes : rouge, noir et blanc. Ils appliquent un algorithme de Template Matching. Ils déterminent si au moins 50 % des pixels rouges correspondent au cercle du modèle. Pour que la méthode soit valable pour plusieurs tailles de signaux comme le montre la Figure 4, plusieurs modèles sont utilisés.

Des contraintes supplémentaires sont imposées sur la zone circulaire intérieure du modèle, afin d'éviter les fausses détections. Les pixels rouges doivent être distribués

symétriquement et ne doivent pas se situer au centre du cercle. Des pixels noirs et blancs doivent se trouver à l'intérieur du cercle dans une certaine proportion. Ceci sert notamment à éliminer les signaux circulaires autres que les signaux de limitation de vitesse.



Figure 4: Modèles utilisés pour le Template Matching dans [13].

La taille des modèles dans la Figure 4 varie de 32×32 à 78×78 pixels.

La méthode utilisée a détecté et localisé correctement les 115 images de signaux parmi les 198 images de l'ensemble de test. Cinq détections de signaux ont été relevées dans des images n'en contenant pas. Les résultats obtenus sont intéressants, cependant la méthode utilisée ne s'applique qu'aux signaux de limitation de vitesse.

I.3.2.2. Contraintes sur les contours de l'objet

De nombreux auteurs se basent sur une détection de contours effectuée sur une image déjà segmentée grâce à l'information couleur.

La méthode utilisée par Gavrilu [17] est basée sur une transformée en distance. Une transformée en distance convertit une image binaire, composée de pixels caractéristiques et de pixels non caractéristiques, en une image où chaque valeur de pixel représente la distance du pixel au plus proche pixel caractéristique (voir Figure 5). La distance utilisée peut être la distance euclidienne, mais d'autres distances existent. Les caractéristiques utilisées par Gavrilu sont les contours.

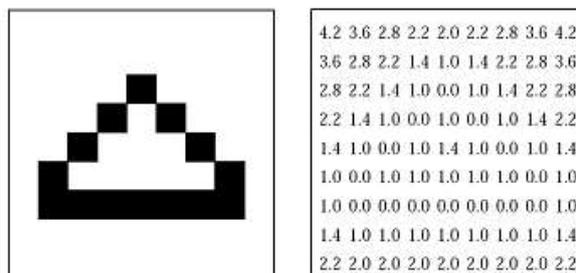


Figure 5: Une image binaire et sa transformée en distance (distance euclidienne).

À partir d'un modèle binaire de signal routier T , où les pixels « on » (pixels caractéristiques) représentent les contours, Gavrilu calcule la distance de chanfrein entre T et l'image I , définie par :

$$D_{chanfrein}(T, I) = \frac{1}{|T|} \sum_{t \in T} d_I(t) \quad (7)$$

Où $|T|$ est le nombre de pixels « on » dans le modèle et $d_I(t)$ la valeur de la transformée en distance de l'image I au pixel t . La distance de chanfrein évalue la ressemblance entre le modèle et l'image⁴.

Pour rendre la méthode robuste, l'auteur répartit les caractéristiques en différentes catégories et effectue le calcul de correspondance pour chacune d'entre elles. Les contours sont répartis en bords de même orientation. Ainsi, les correspondances entre le modèle et l'image ne se font plus sur des distances par rapport au plus proche contour, mais sur des distances par rapport au plus proche bord de même orientation.

Plusieurs modèles sont utilisés. Il faut un modèle pour chaque forme de signal : triangulaire sur base, triangulaire sur pointe et circulaire. De plus, pour être capable de détecter les signaux à différentes échelles, il faut utiliser différents modèles, ce qui augmente le temps de calcul. Pour diminuer ce temps de calcul, l'auteur adopte une approche pyramidale avec une hiérarchie de modèles.

I.3.3. Le détecteur de Viola et Jones

Viola et Jones apportent dans [18] trois contributions importantes dans le domaine de la détection d'objets. Premièrement, les auteurs introduisent « l'image intégrale », qui permet de calculer très rapidement les attributs utilisés par leur détecteur. Ces attributs évoquent les ondelettes de Haar. Ensuite, un algorithme d'apprentissage basé sur l'algorithme AdaBoost est présenté. Il sélectionne les attributs les plus intéressants parmi un grand nombre d'attributs disponibles, ce qui mène à des classeurs extrêmement efficaces. Enfin, une méthode pour combiner des classeurs de complexité croissante sous forme de

⁴ La distance de chanfrein est sensible aux occlusions. D'autres mesures permettent de réduire les effets des occlusions, comme la distance de Hausdorff

CHAPITRE I. Etat de l'art

cascade est introduite. Celle-ci permet d'éliminer rapidement les régions de l'image ne contenant pas d'objet pour passer plus de temps de calcul sur les régions intéressantes

L'article de Viola et Jones a inspiré de nombreuses recherches dans le domaine de la détection d'objets, entre autres pour la détection des signaux routiers.

Dans [8], les auteurs reprennent l'idée de Viola et Jones, mais à la différence de ces derniers, ne travaillent pas sur des images en niveaux de gris. Ils utilisent 7 canaux couleurs :

- Les canaux classiques r , g et b ;
- Les canaux normalisés :

$$R = r/s, G = g/s \text{ et } B = b/s, \text{ avec } s = r + g + b \quad (8)$$

- Le canal en niveaux de gris $s/3$.

L'apport de la couleur leur permet de diminuer le taux de fausses détections par un facteur 10 par rapport à l'utilisation en niveaux de gris de la méthode de Viola et Jones.

Dans [19], les auteurs se basent également sur l'article de Viola et Jones [18]. Ils soulignent le problème de la technique Adaboost. Celle-ci nécessite de construire de nombreux étages pour que la cascade de classeurs possède un taux de fausse alarme raisonnable. Ceci entraîne un surapprentissage. La cascade classe presque parfaitement les échantillons d'apprentissage, mais n'a aucun pouvoir de généralisation sur le classement d'échantillons inconnus. La solution adoptée par les auteurs a été de combiner une cascade avec moins d'étages avec d'autres méthodes pour éliminer les fausses détections. Le schéma utilisé est représenté Après la cascade, le traitement est séparé selon la forme des signaux (triangulaire ou circulaire). Les signaux circulaires étant difficilement descriptibles par des attributs rectangulaires, la cascade fournit en sortie un grand nombre de fausses détections. Les auteurs emploient alors une méthode rapide qui utilise la symétrie radiale locale pour extraire des points d'intérêt [20]. Cette méthode est sensible aux relations entre l'intensité du signal et l'intensité du fond de l'image. Cela oblige donc les auteurs à utiliser un critère de convergence basé sur la convergence du gradient.

Pour les signaux triangulaires, les auteurs se basent sur le fait que la direction du gradient est constante sur les trois côtés des signaux et qu'elle change d'une manière connue

aux pointes du triangle. Ils classent les valeurs de la direction du gradient en 8 groupes. Ils cherchent ensuite les variations dans la direction du gradient pour trouver les pointes du triangle et utilisent les relations géométriques pour trouver la forme complète.

Les auteurs terminent par une approche globale (valable pour les signaux triangulaires et circulaires) pour éliminer les fausses détections. Ils appliquent une analyse en composantes principales. Celle-ci permet d'éliminer les fausses détections, car en tenant compte uniquement des premières composantes principales, l'erreur de reconstruction d'une fausse détection est plus grande que celle d'une image contenant un vrai signal.

I.4. Reconnaissance des signaux

La phase de reconnaissance consiste à assigner une classe à chaque signal détecté. Généralement, les signaux de formes différentes ont été dissociés lors de la phase de détection. Les réseaux de neurones artificiels tiennent une place importante dans les travaux réalisés sur la reconnaissance de signaux routiers. À côté de cela, de nombreuses autres méthodes de classification existent. Nous en présentons certaines, la plupart ayant déjà été utilisées pour la reconnaissance des signaux.

I.4.1. Réseaux de neurones

Yang et al. [15] utilisent des réseaux de neurones pour la reconnaissance. Les meilleurs taux de détection obtenus sont proches de 70 %. Comme entrée aux réseaux de neurones, ils fournissent les valeurs des pixels et les coefficients de la transformée en cosinus discrète DCT⁵ des images. Comme traitement préalable à l'image, les auteurs sélectionnent le pictogramme et font une normalisation en luminance.

Torresen et Sekanina [13] utilisent un perceptron multicouche MLP⁶. Ils se concentrent sur les signaux de limitation de vitesse. Pour déterminer la vitesse indiquée par le signal, les auteurs ne prennent en considération que le premier chiffre du signal (le deuxième étant toujours 0). L'intérieur du signal a été segmenté et les pixels des chiffres sont dissociés des pixels du fond. Il est donc facile de trouver la région d'intérêt contenant

⁵ *Discrete Cosine Transform*

⁶ *Multilayer perceptron*

le premier chiffre. Cette région d'intérêt est remise à une échelle 7×5 pixels (voir la Figure 6). Ces petites images sont présentées à l'entrée d'un réseau de neurones feed-forward entraîné avec l'algorithme de rétropropagation du gradient. Le réseau de neurones possède 3 couches. La couche d'entrée contient 35 (7×5) neurones, la couche cachée en contient 35 et 6 neurones forment la couche de sortie (un pour chaque chiffre présent sur les signaux).



Figure 6: Exemples de chiffres de taille 7×5 utilisés pour lors de la reconnaissance des signaux de limitation de vitesse

Ces images présentées dans la Figure 6 de 35 pixels sont fournies à l'entrée du réseau de neurones. [13]

Franke et al. [9] utilisent des fonctions à base radiale RBF⁷. On utilise les réseaux RBF à simple couche, car ils ne nécessitent pas les longs apprentissages des MLP. Les neurones des RBF utilisent généralement des fonctions d'activation gaussiennes.

Vitabile et al. [6] utilisent un MLP. Les signaux de même forme ont été précédemment groupés. Ils travaillent en RGB. Pour chaque canal, l'image est divisée en blocs de 3×3 pixels. Les entrées du réseau sont les valeurs moyennes des blocs. Les taux de reconnaissance obtenus sont supérieurs à 84 %.

Escalera et al. [11] testent également un MLP (à 3 ou 4 couches). Ils fournissent directement les valeurs des pixels au MLP. Les images sont remises à une résolution de 30×30 pixels qui correspond à la taille de la couche d'entrée du MLP. La couche de sortie comporte 1 neurone de plus que le nombre de classes des signaux, pour pouvoir indiquer que l'image ne contient pas de signal connu. Pour entraîner le MLP, les auteurs utilisent un ensemble d'apprentissage basé sur des images de signaux idéaux (schémas). Ils construisent leur ensemble d'apprentissage en modifiant ces images par :

⁷ *Radial Basis Function*

- Des rotations ;
- Ajout de bruit gaussien ;
- Seuillage afin d'obtenir une information localisée dans la partie intérieure du signal.

Dans d'autres travaux [11], les mêmes auteurs utilisent un réseau de neurones ART⁸. Ce réseau est capable de développer un clustering stable à partir de séquences d'entrées arbitraires en s'autoorganisant. Grâce à cela, lorsque de nouvelles classes de signaux lui sont présentées, il ne nécessite pas de nouvel entraînement.

Les réseaux ART sont des réseaux à apprentissage par compétition, dont le problème majeur est le dilemme « stabilité/plasticité ». Afin d'assurer la stabilité, on doit faire tendre le coefficient d'apprentissage vers zéro, mais le réseau perd alors toute sa plasticité. Les réseaux ART permettent de contourner ce problème. Les vecteurs de poids n'y seront adaptés que si l'entrée fournie est semblable à un prototype déjà connu par le réseau. À ce moment, on utilisera le terme : résonance. Par contre, si l'entrée est trop différente des prototypes existants, une nouvelle catégorie va se créer, avec pour prototype, l'entrée qui a engendré sa création. Il y a deux principaux types de réseaux ART : les ART-1 pour des entrées binaires et les ART-2 pour des entrées continues.

Fang et al. [4] utilisent deux types de réseaux de neurones : un réseau CART⁹ et un réseau CHAM¹⁰. Le premier permet de séparer les signaux en catégories et la seconde classe les signaux à l'intérieur des différentes catégories.

Franke et al. [9] emploient un réseau ATDNN¹¹ qui permet de classer un objet non pas sur base d'une image, mais sur base d'une série temporelle d'image de cet objet. L'architecture de ce type de réseaux est montrée sur la Figure 7.

⁸ *Adaptive Resonance Theory* [21]

⁹ *Configurable Adaptive Resonance Theory*

¹⁰ *Configurable Heteroassociative Memory*

¹¹ *aptive Time Delay Neural Network*

Pour conclure cette section sur les réseaux de neurones, remarquons que ceux-ci sont couramment utilisés pour la reconnaissance des signaux. Les réseaux de neurones sont un outil puissant pour la classification des objets. Cependant, il faut une certaine expérience pour les construire et les entraîner.

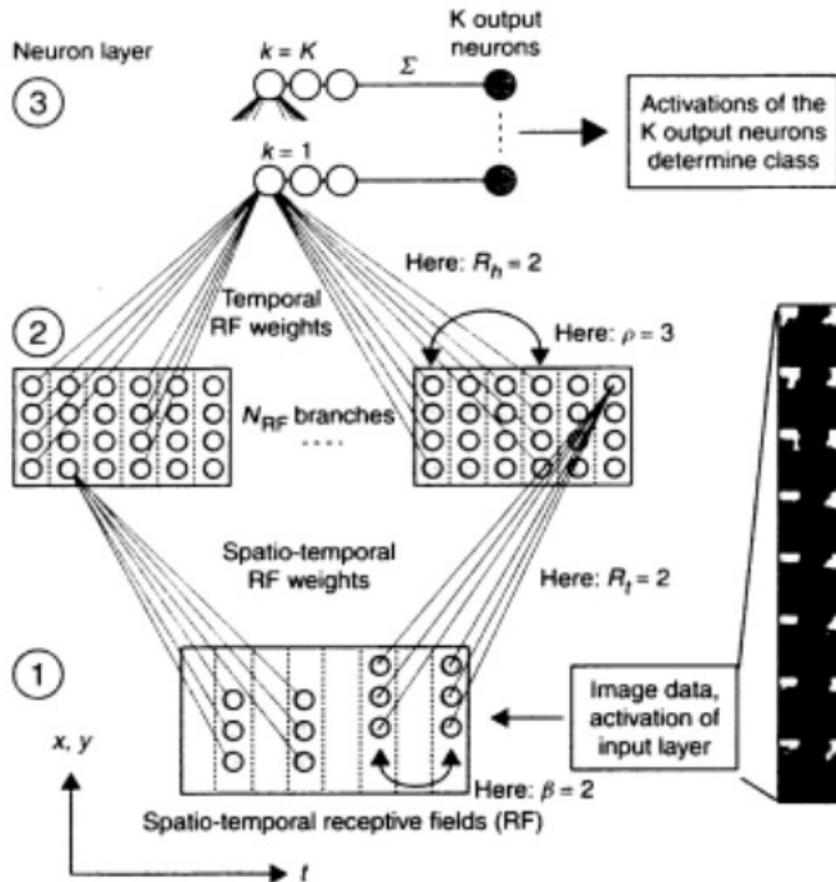


Figure 7: Architecture du réseau de neurones adaptatif à temps de retard utilisé dans [6].

I.4.2. Corrélation normalisée

Miura et al. [7] utilisent une méthode de pattern Matching basée sur la corrélation normalisée. La corrélation normalisée possède l'avantage d'être robuste par rapport aux variations d'illumination, ce qui est intéressant lorsque l'on travaille en extérieur.

I.4.3. Classeurs polynomiaux

Franke et al. [9] proposent de nombreuses solutions pour la reconnaissance de signaux. Par exemple l'utilisation de classeurs polynomiaux dont le schéma représenté sur

la Figure 8. La structure fait penser aux réseaux de neurones. Le principe est le suivant. Le classeur transforme l'espace d'attributs d'entrée. Il crée une liste de structure polynomiale en effectuant des produits des attributs d'entrées. La seconde couche est une combinaison linéaire de ces nouveaux attributs définie par les coefficients de la matrice W (voir Figure 8). La liste de structure polynomiale est choisie par l'utilisateur et les coefficients de la matrice W sont ajustés lors d'un apprentissage.

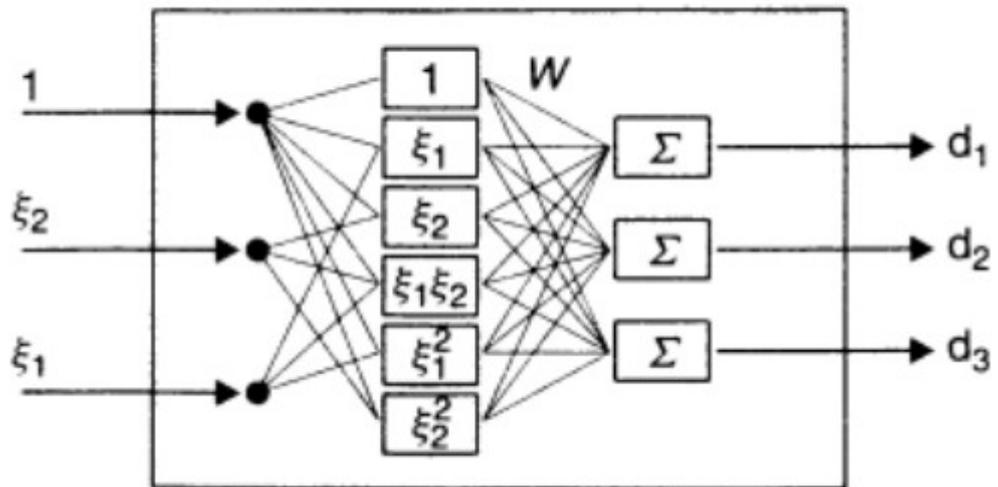


Figure 8: Architecture d'un classeur polynomial quadratique utilisé dans [6].

I.4.4. Machine à vecteurs supports

Une autre méthode de classification proposée par Franke et al. [9] : les machines à vecteurs supports SVM¹². Le principe des SVM est de déterminer une frontière entre les classes, dans un espace d'attributs, en utilisant uniquement les échantillons frontières qui sont en fait les seuls échantillons déterminants (voir Figure 9). La frontière est fixée de manière à maximiser la marge, c'est-à-dire l'écart entre les vecteurs supports et la frontière

¹² Support Vectors Machine.

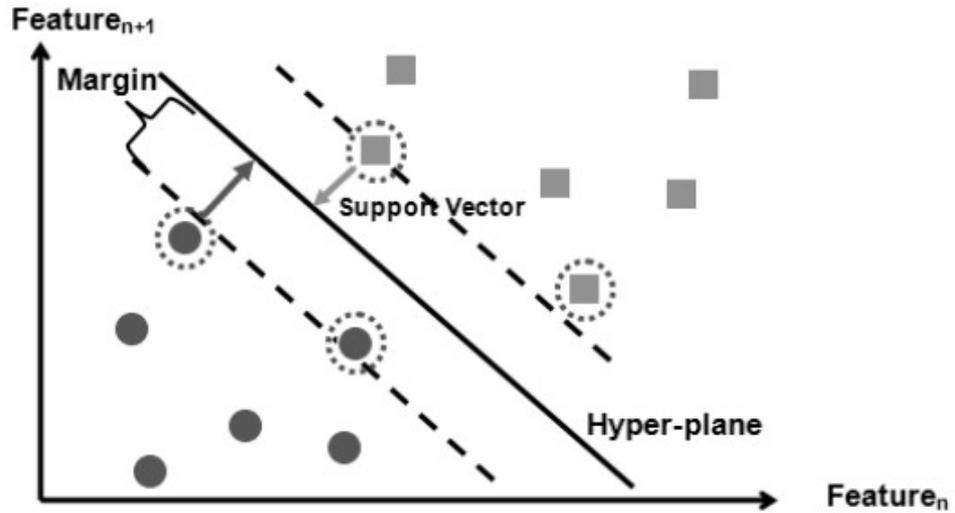


Figure 9: Principe d'une SVM.

Un faible nombre d'échantillon détermine l'allure de la frontière entre les classes. [22]

I.4.5. Modèle génératif

Bahlmann et al. [8] utilisent un modèle génératif, c'est-à-dire qu'ils supposent que les objets de l'ensemble d'apprentissage qu'ils possèdent ont été générés par une certaine distribution de probabilité. L'architecture globale du système est illustrée sur la Figure 10.

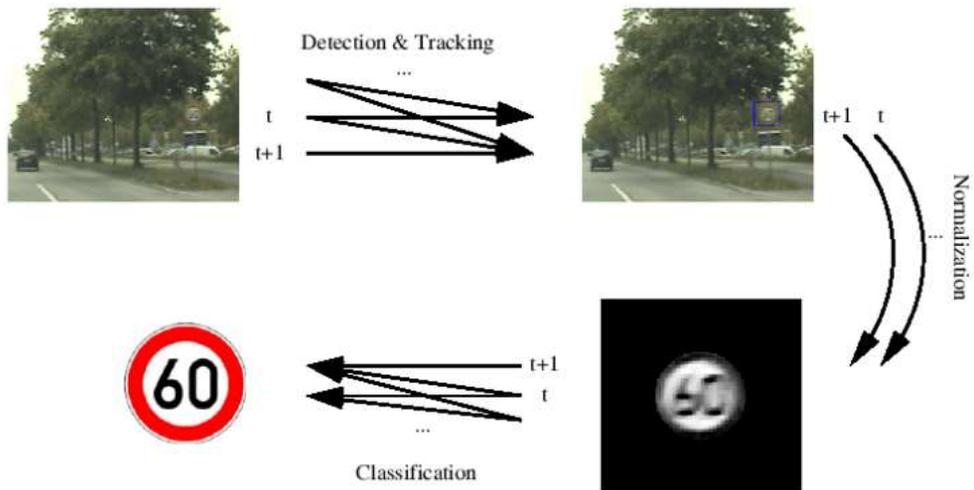


Figure 10: Architecture du système de reconnaissance utilisé par Bahlmann et al.

[8]. Une cascade de classeurs, entraînés sur des filtres de Haar avec une variante de l'algorithme AdaBoost, détecte les signaux dans chaque image t . Une fois détectés, les objets

sont suivis, et les détections individuelles des images $\{1, \dots, t_0\}$ sont fusionnées pour une détection globale plus robuste. Sur la figure, ceci est représenté par des flèches entrelacées. Ensuite, le signal est masqué par une forme circulaire et normalisée selon la position, l'échelle et la luminosité. Finalement, le signal est classé en maximisant une fonction de vraisemblance en tenant compte de dépendances temporelles.

Entraînement. Avant la modélisation, l'espace des attributs (les pixels en niveaux de gris) est transformé par une analyse discriminante linéaire (LDA). Le nouveau vecteur d'attributs $a \in \mathbb{R}^{25}$ est constitué des 25 premiers facteurs les plus discriminants de LDA. Ensuite, pour chaque classe $j \in \{1, 2, \dots, M\}$, ils estiment les paramètres d'une loi normale multivariée et unimodale

$$p(a|y) = \mathcal{N}(a; \mu_y; \Sigma_y) \quad (9)$$

I.4.6. Ensembles d'arbres

Marée [23] introduit, dans sa thèse sur la classification automatique d'images, une méthode d'apprentissage supervisé capable de s'appliquer à de nombreux types de problèmes de classification. Elle est en outre très rapide et présente une précision comparable aux meilleurs résultats obtenus dans la littérature. Partant d'un ensemble d'apprentissage composé d'une série d'images des différentes classes, il sélectionne aléatoirement dans les images des sous-fenêtres afin de se créer un nouvel ensemble d'apprentissage. Sur base de celui-ci, il construit un ensemble d'arbres extrêmement aléatoires [24] basé uniquement sur les valeurs de pixels des sous-fenêtres. Cette méthode possède plusieurs avantages. L'extraction de sous-fenêtres implique une approche locale. Celle-ci permet notamment d'augmenter la taille de l'ensemble, ce qui permet une meilleure généralisation du classifieur lorsque l'ensemble de départ est trop petit. De plus, l'approche locale possède une meilleure robustesse qu'une approche globale par rapport aux occlusions et aux changements de point de vue. D'autre part, l'utilisation directe des valeurs des pixels lors la construction de l'ensemble d'arbres permet une conservation de toute l'information des sous-fenêtres sans nécessiter les calculs inhérents à d'autres techniques. Enfin, la construction d'arbres extrêmement aléatoire est très peu gourmande en temps de calcul.

Marée [25] précise qu'un meilleur taux d'erreur peut être obtenu en utilisant du boosting si l'on accepte de sacrifier la rapidité.

I.5. Conclusion

Dans ce chapitre nous avons présenté quelques systèmes et plateforme d'acquisition d'image, utilisé dans la littérature, puis nous exploré les différentes méthodes est techniques adoptées par les auteurs, pour faire la détection des panneaux de signalisation routière, que ce soit en utilisant l'information de couleur, ou l'information des de forme. En ce qui concerne la reconnaissance de ces panneaux, nous avons discuté le fonctionnement de quelques classificateurs utilisés à cette fin.

Chapitre II . Méthodologie, outils et environnement de travail

Sommaire

II.1. Introduction	26
II.2. Méthodologie de travail	26
II.3. Structure générale du programme	26
II.4. Signalisation routière marocaine	28
II.5. Ressources matérielles et logicielles utilisées	29
II.6. Conclusion.....	32

II.1. Introduction

Avant de commencer tout travail, il faut une planification et une définition de cahier de charges, et une préparation des ressources nécessaires pour l'accomplissement de ce travail, dans les normes précédemment fixées dans le cahier de charges et dans les délais. C'est dans cette perspective vient ce chapitre, pour présenter la méthodologie de travail suivi pour aborder ce projet, et discuter la structure générale du programme à réaliser, en suite nous présentons la signalisation marocaine sujet et matière première de notre projet, et la fin nous étalons les différentes ressources matérielles et logiciels nécessaires au développement et l'implémentation de nos algorithmes et programmes.

II.2. Méthodologie de travail

Pour déterminer la méthodologie de travail, il fallait répondre à la première question qui se pose « comment va-t-on procéder ? ».

La technique adoptée pour répondre à cette question est de séparer la phase de détection de la phase de reconnaissance des panneaux, et par la suite développer et tester les deux algorithmes séparément, cette technique permet d'évaluer et améliorer chaque algorithme à part.

La première étape à suivre est de faire la détection des panneaux de signalisation routière sur des images entièrement en off-line avant de passer au temps réel, pour bien évaluer et améliorer les algorithmes de détection d'objet

L'étape suivante est de développer et tester les méthodes de reconnaissance sur les panneaux de signalisation routière en mode off-line.

La troisième étape est de combiner les algorithmes de détection et de reconnaissance des panneaux routier, et le faire fonctionner en mode off-line, puis en en temps réel.

II.3. Structure générale du programme

La structure générale du programme comme il est montré dans la Figure 11 comportera six grandes phases, chacun fait appelle un algorithme spécifique développé et testé d'une manière unitaire, pour réponde aux objectifs ci-dessous :

CHAPITRE II. Méthodologie, outils et environnement de travail

- Initialisation : a pour objectif d'initialiser les périphériques, charger les bibliothèques, et préparer l'environnement physique et logique de la machine ;
- Acquisition : a pour but de charger des images ou des vidéos à partir d'un support de stockage, ou de prendre de capture à partir d'une caméra, connectée à la machine. (Voir Figure 12) ;
- Prétraitement : est principalement destinée à préparer l'image, en appliquant des techniques de filtrage, de convertirons et de segmentations de couleur, de binarisation et d'extraction de contours ;
- Détection : permet d'identifier et d'extraire les régions de la scène routière susceptibles de contenir des panneaux de signalisation routière en se basant sur leurs couleur et géométrie particulières ;

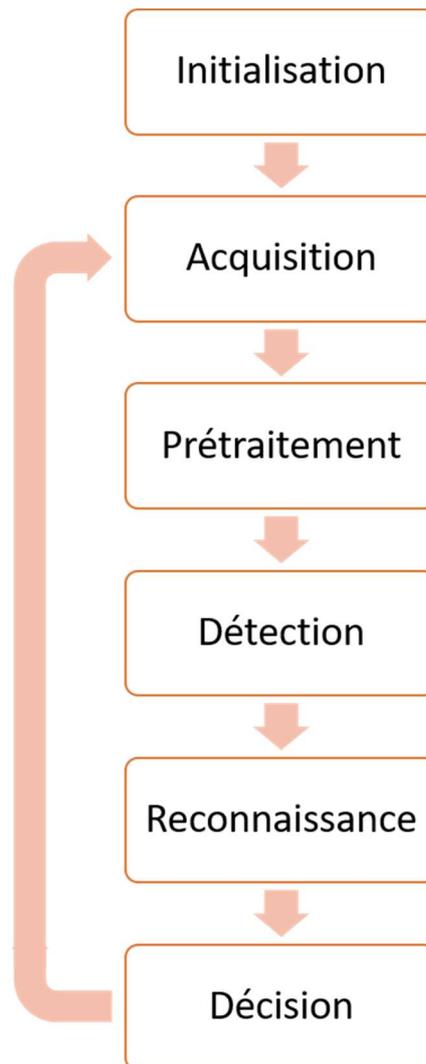


Figure 11: Structure générale du programme

- Reconnaissance : a pour rôle de classer chacun des panneaux de signalisation routière extrait dans la phase de détection, et afficher une Template correspondante au panneaux reconnu.



Figure 12: Plateforme de test du Système TSDS embarqué.

II.4. Signalisation routière marocaine

La signalisation routière marocaine est riche. Elle est composée d'environ 200 signaux routiers. Dans le dahir n° 1-83-353 du 11 rabii I 1407 (14 novembre 1986) portant publication de la convention sur la signalisation routière faite à Vienne le 8 novembre 1968; et Arrêté conjoint du ministre de l'équipement et des transports et du ministre de l'intérieur n° 1918-11 du 15 chaabane 1432 (17 juillet 2011) relatif à la signalisation routière (Ministère de justice - Maroc, 2012) (Ministère de justice - Maroc, 2012), les signaux routiers sont répartis en 6 catégories :

- A. Signaux de danger ;
- B. Signaux relatifs à la priorité ;

- C. Signaux d'interdiction ;
- D. Signaux d'obligation ;
- E. Signaux relatifs à l'arrêt et au stationnement ;
- F. Signaux d'indication.



Figure 13: Exemple d'images de signaux routiers.

Pour désigner un signal routier en particulier, nous utiliserons le plus souvent le code de la route. Quelques exemples de signaux routiers sont repris sur la Figure 13. De gauche à droite, on peut voir les signaux A14, B1, C43 et E1.

Nous avons travaillé sur l'ensemble des signaux routiers. Nous avons tenté d'obtenir un maximum de l'ensemble de ces signaux routiers lors de l'acquisition de nos images de la scène routière et de nos séquences vidéo.

II.5. Ressources matérielles et logicielles utilisées

II.5.1.1. Caméscope

Pour l'acquisition, nous avons disposé d'un caméscope DV présenté dans la Figure 14. Le format DV utilise certains éléments du standard JPEG pour le codage de la vidéo. La compression DV ne joue que sur les redondances spatiales à l'intérieur de l'image, sans chercher à réduire les redondances temporelles comme le fait le MPEG. Le facteur de compression est de 5 : 1.

Le signal vidéo est compressé de manière telle qu'il ne nécessite « plus que » 3.2 Mo/s pour être lu. Il faut donc environ 11 Go pour stocker 1 heure de vidéo. Nous avons installé le caméscope dans une voiture (voir Figure 12) et enregistré environ 10min de

CHAPITRE II. Méthodologie, outils et environnement de travail

séquences vidéo prises dans différentes régions au Maroc et à l'Espagne sur des routes locales et nationales.



Figure 14: Caméscope utilisé pour l'acquisition des séquences vidéo.

Les conditions météorologiques sont variées : ensoleillé, temps couvert et sec, couvert et humide, nuit. Nous avons enregistré des séquences vidéo avec différents zooms afin de déterminer lequel serait le plus adéquat pour notre travail. Avec un zoom, l'avantage est que les signaux apparaissent avec une taille plus grande sur l'image, ce qui facilite la détection. D'un autre côté, et la reconnaissance devient un peu difficile, car la vitesse de déplacement du panneau entre deux images successives est plus importante. Ceci est aggravé par plusieurs problèmes : la stabilité de l'installation est précaire et l'état des routes dans notre région n'est pas irréprochable. Pour ces raisons, nous travaillerons le plus souvent sans zoom. Les séquences obtenues sont des séquences DV entrelacées. De plus, le flux de données du format DV est très important. Ceci rend les séquences originales difficilement utilisables. Nous désentrelaçons le flux en ne conservant qu'une seule trame sur deux et en dupliquant chaque ligne. Ensuite, nous convertissons les séquences DV au format MPEG-4. La taille des fichiers est ainsi divisée par 7 sans trop de perte de qualité visible. La qualité de nos séquences vidéo est limitée par un certain nombre de facteurs. La compression MPEG-4 diminue la qualité. Le désentrelacement diminue la résolution : nous travaillons avec séquences 640×480 , mais le désentrelacement diminue la résolution verticale réelle par 2. La vitesse du véhicule est parfois importante ce qui amène du flou dans les images,

celles-ci n'étant pas stabilisées. Enfin, les conditions météorologiques variables entraînent une illumination non maîtrisée.

II.5.1.2. Machines

La machine utilisée pour le développement et le test de l'ensemble des algorithmes, est un ordinateur Marque HP, de processeur Intel i3, quadcore, de vitesse 2.4Gz par unité. Il est aussi doté d'une mémoire 4Go de type DDR3. Ce sont les caractéristiques minimales qu'on trouve pour les ordinateurs de bord de la majorité des nouveaux véhicules doté d'une confortabilité un peu augmentée.



Figure 15: Ordinateur HP i3 utilisé dans le projet.

Nous avons aussi mené des tests à bord d'un véhicule, en utilisant une carte Raspberry Pi B+ [27], de processeur Intel 4040, de 700MHz de fréquence. Sa mémoire RAM est un SDRAM de 1Go.

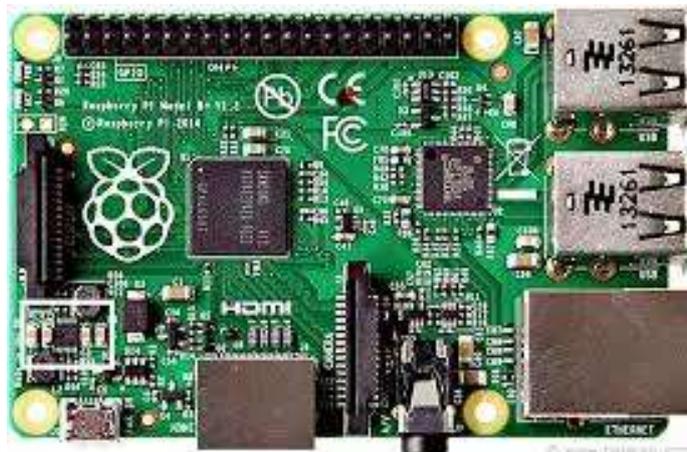


Figure 16: Carte de Raspberry Pi B+

II.5.1.3. Logiciels

Microsoft Windows 10 x64 : est le système d'exploration installé sur l'ordinateur HP utilisé dans le projet.

Linux Raspbian : est le système d'exploration installé sur la carte Raspberry Pi B+, utilisée pour le test à bord du véhicule, et c'est est une distribution de Linux Debian adaptée pour ce type de carte.

Microsoft Visual Studio 10 x64 : est un outil de développement logiciel, intégrant tous les langages de programmation conçu par Microsoft. Le projet est développé sous le langage C/C++.

Open CV v2.44 : est une bibliothèque libre, comprenant les célèbres algorithmes de traitement de l'information dans le domaine de la vision par ordinateur. Distribuée avec licence libre (GPL) sous plusieurs langages de programmation (C/C++, Python, Java, ...).

Python v2.7 : est le langage de programmation utilisé pour faire fonctionner les algorithmes sous la carte Raspberry Pi B+.

Notepad++ : est l'utilitaire utilisé pour écrire le code python, il permet de faire des collaborations syntaxiques du code, et facilite son aménagement.

II.6. Conclusion

Dans ce chapitre nous avons présenté la méthodologie de travail suivie dans ce projet, la structure générale du programme, la signalisation marocaine, et l'ensemble des ressources matérielles et logicielles utilisées pour développer et implémenter les programmes de ce projet.

Chapitre III . Module de détection des panneaux routiers

Sommaire

III.1. Introduction.....	34
III.2. Méthode de détection.....	34
III.3. Détection de la couleur des panneaux.....	34
III.4. Détection des formes.....	37
III.4.1. Extraction des contours.....	37
III.4.2. Estimation de la forme des panneaux.....	37
III.5. Evaluation du module de détection.....	41
III.6. Conclusion.....	43

III.1. Introduction

La détection des panneaux routiers dans l'image consiste à y trouver et segmenter des régions d'intérêt susceptibles de correspondre à un panneau. Pour ce faire, il faut pouvoir distinguer leurs signes sémantiques particuliers qui sont principalement de deux ordres : la couleur et la forme. Dans ce travail, nous avons adopté une méthodologie qui consiste en une segmentation par couleur poursuivi par une segmentation de formes pour détecter les panneaux.

III.2. Méthode de détection

La méthodologie adoptée pour la détection des panneaux routiers, est schématisée par la Figure 17. Elle est composée d'un enchaînement en cascade de deux grandes étapes. La première d'entre elles détecte les pixels d'une couleur donnée (bleu ou rouge). Elle permet d'obtenir des régions d'intérêt dans l'image. La seconde étape consiste à détecter des formes géométriques connues de panneaux (cercles, triangles, carrés) dans les régions d'intérêts précédemment obtenues. Cette étape propose des candidats potentiels de panneaux. La catégorie de chaque candidat est déterminée en fonction de sa couleur et de sa forme.

III.3. Détection de la couleur des panneaux

La première étape, requise pour arriver à détecter des signaux routiers, consiste à rechercher leurs couleurs dans les images. Pour ce faire, nous avons adopté une approche qui consiste à transformer les images RGB en images HSV (Teinte-Saturation-Valeur) et à se restreindre au plan HS pour minimiser l'effet de la variation de la lumière, puis à segmenter ces dernières par une technique de seuillage binaire afin de localiser séparément les zones bleues et les zones rouges susceptibles de correspondre à des panneaux routiers. Ainsi, deux images binaires sont obtenues après la segmentation de chaque image HSV : une image binaire, appelée masque de couleur bleu, correspondant aux zones bleues, et une autre, appelée masque de couleur rouge, correspondant aux zones rouges. Le seuil de la segmentation est déterminé empiriquement après le traitement d'un grand nombre d'images en faisant un compromis entre sur-détections et sous-détections. La Figure 18 montre un exemple d'image RGB et les masques de couleur bleu et rouge ainsi obtenus.

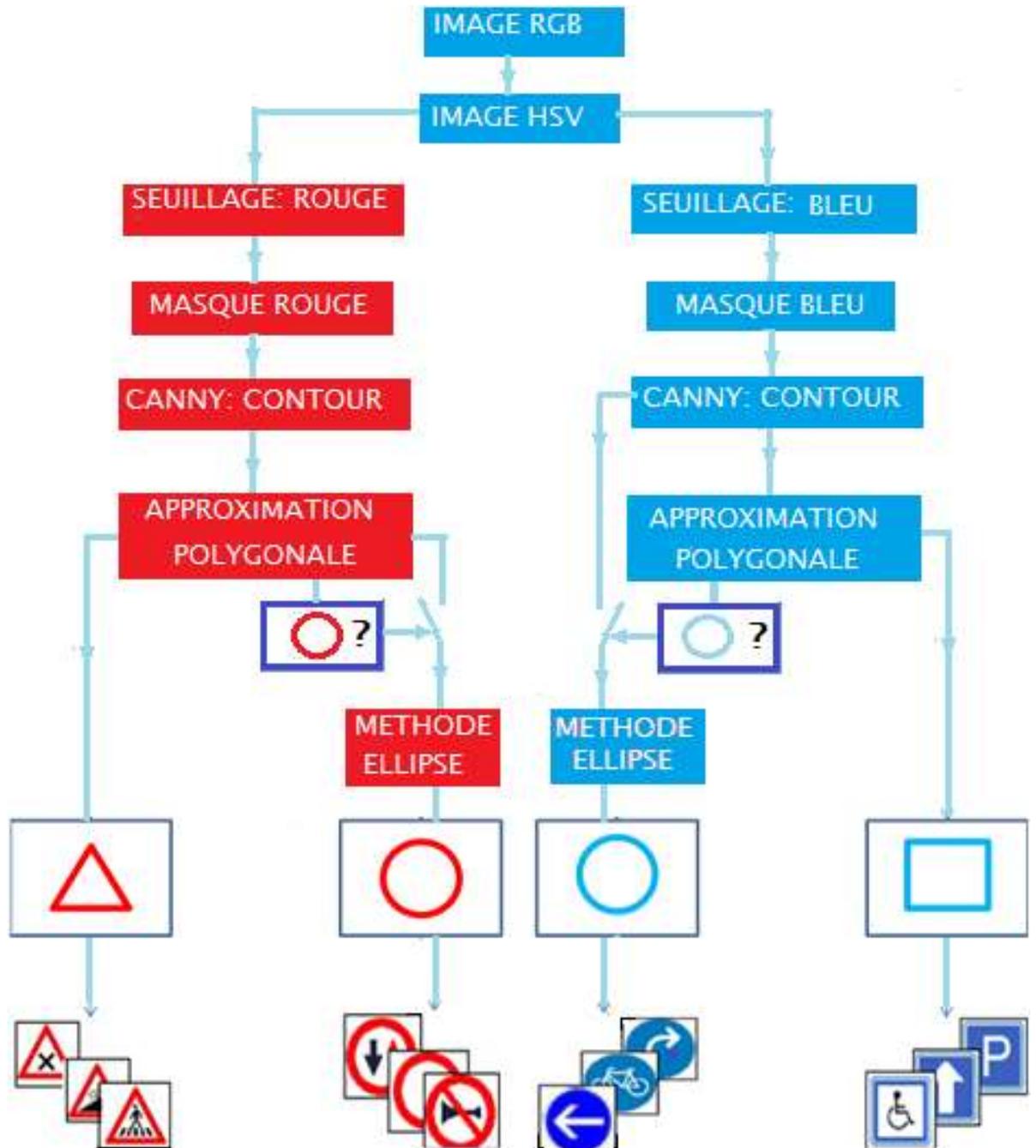


Figure 17: Structure du programme de détection des panneaux routiers.



(a)



(b)



(c)

Figure 18: Résultats de détection de couleur. a) Image initiale, b) Masque de couleur bleu et c) Masque de couleur rouge

III.4. Détection des formes

L'objectif de cette partie est de détecter, à partir des masques obtenus par la segmentation en couleur de chaque image, certaines formes caractéristiques des panneaux de signalisation routière (cercles, triangles, carrés). Afin d'atteindre cet objectif, nous avons utilisé une méthode basée sur l'estimation des formes à partir de l'extraction des contours des régions d'intérêt présents dans chaque masque.

III.4.1. Extraction des contours

Pour extraire les contours, nous avons utilisé le filtre de canny [27] qui est un filtre de Sobel suivi par un seuillage et précédé par un lissage gaussien pour atténuer les bruits. Ce filtre est conçu pour être optimal, au sens de trois critères :

- Bonne détection : faible taux d'erreur dans la signalisation des contours ;
- Bonne localisation : minimisation des distances entre les contours détectés et les contours réels ;
- Clarté de la réponse : une seule réponse par contour et pas de faux positifs.

Les contours extraits par ce filtre, pour chaque masque, constituent des caractéristiques bas niveau distinctifs permettant l'estimation des formes des panneaux routiers.

III.4.2. Estimation de la forme des panneaux

Plusieurs méthodes d'estimation de la forme des panneaux à partir des contours sont proposées dans la littérature :

- La transformée de Hough [28,29] pour l'estimation de la forme triangulaire et/ou circulaire des panneaux ;
- La transformation en Symétrie Radiale [30,31] qui est une variante de la transformée de Hough pour la détection de cercles ;
- Le détecteur de Polygones Réguliers [32] qui est une adaptation de la transformation en Symétrie Radiale pour la détection de polygones réguliers à n coins comme les triangles, les rectangles, et les octogones ;
- La transformée en Sommet et Bissectrice [33] pour l'estimation du triangle ;

- La transformée Chinoise Bilatérale [33] pour l'estimation de la forme circulaire et rectangulaire.

Bien qu'elles soient simples à implémenter, ces méthodes basées sur des transformations, en particulier la transformée de Hough, demeurent lourdes en calcul. C'est pourquoi, dans notre travail dont la finalité est l'identification des panneaux en temps réel, nous avons utilisé, pour l'estimation des formes des panneaux, une méthode, basée sur l'approximation polygonale [34,35,36], facile à mettre en œuvre et rapide.

Le principe de cette méthode dit approximation linéaire par morceaux, consiste à trouver successivement les segments d'un polygone approximant le plus fidèlement possible la forme du contour original, de façon à minimiser un critère d'erreur. L'intérêt est d'obtenir une description plus compacte et plus adaptée pour des traitements ultérieurs.

Pour mieux comprendre le fonctionnement de la méthode, prenons un exemple de la Figure 19 (a) où le critère d'erreur d'approximation utilisé est la distance maximale entre les points du contour et le segment du polygone initial. Pour simplifier l'explication, notons \widehat{AB} (resp. \overline{AB}) la portion du contour (resp. Le segment du polygone initial d'approximation) comprise entre le point A et le point B. Pour savoir si l'approximation \overline{AB} de \widehat{AB} est valide, l'algorithme calcule la distance de tous les points de \widehat{AB} au segment \overline{AB} . Dans l'exemple de la Figure 19 (a), la distance entre le point C de \widehat{AB} et \overline{AB} est indiquée par $d(C, \overline{AB})$. Supposons que $d(C, \overline{AB}) = \max d(P, \overline{AB})$ avec P un point de \widehat{AB} .

Si $d(C, \overline{AB}) < \varepsilon$, on peut décider que \widehat{AB} peut être remplacé (approximé) par \overline{AB} . Il est important de noter, concernant le seuil d'erreur ε du paramètre entré de l'algorithme, que plus ε est grand, plus le niveau d'approximation est grossier et le nombre de sommets et de segments du polygone est relativement faible. A l'inverse si ε est petit, le polygone résultant est très proche du contour origine mais le nombre de sommets et de segments du polygone est relativement élevé.

Dans le cas contraire, si $d(C, \overline{AB}) \geq \varepsilon$ alors la portion du contour \widehat{AB} ne peut pas être représentée par le segment de droite \overline{AB} car la plus grande distance entre un point C du contour au segment \overline{AB} est supérieure à l'erreur ε autorisée. Dans ce cas, on doit partitionner le contour \widehat{AB} en deux. Bien qu'il y ait plusieurs façons de partitionner le contour \widehat{AB} en deux, le partitionnement le plus naturel est le partitionnement au point C, c'est à dire qu'il

faut partitionner au point où sa distance au segment \overline{AB} est la plus grande. Le résultat du partitionnement de la Figure 19 (a) est montré dans la Figure 19 (b) où le contour \widehat{AB} est approximé par deux segments polygonaux \overline{AC} et \overline{CB} . L'étape suivante de l'algorithme consiste à appliquer les mêmes traitements aux contours \widehat{AC} et \widehat{CB} (Figure 19 (b)). La répétition de ce processus continue tant que la distance maximale, entre le contour et le segment de polygone correspondant, est supérieure à ε et il s'arrête dans le cas contraire (Figure 19 (c)).

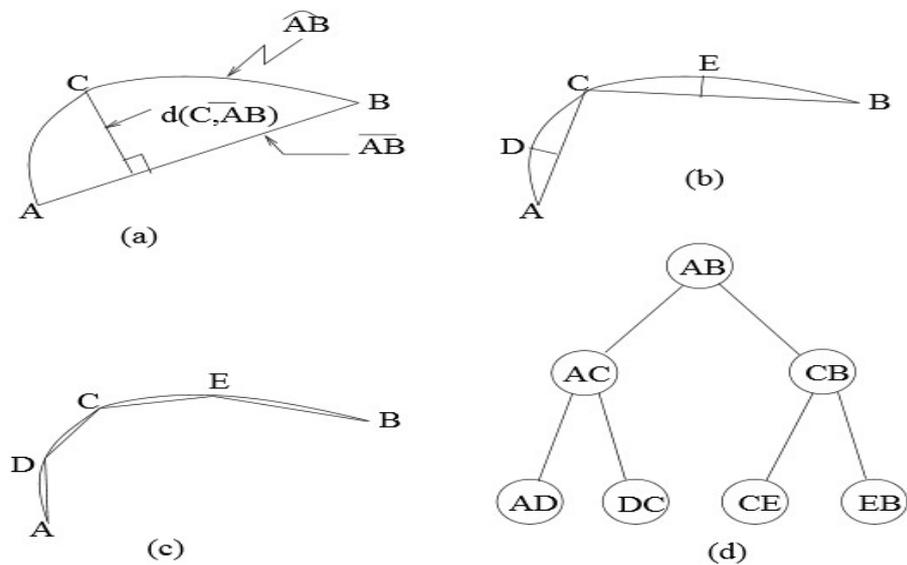


Figure 19: Approximation polygonale du contour.

Dans ce travail, l'Approximation polygonale est utilisée pour approximer un contour fermé par un polygone de N sommets dont les coordonnées sont bien déterminées. La valeur de ε est déterminée empiriquement, elle dépend de la taille de l'image de la scène de trafic routier, et de la taille des panneaux prise en considération :

$$\varepsilon = 0.02 \times \text{contour} \quad (10)$$

Ce polygone est utilisé pour estimer la forme du contour, extrait à partir du masque de couleur rouge ou bleu, et de vérifier s'il s'agit d'un panneau routier ou non selon les critères suivants :

- **Pour le masque de couleur rouge**

Si le polygone a 3 sommets et s'il est équilatéral avec une précision inférieure à 5%, tel que sa base est horizontale avec une erreur de tolérance limitée à 2%, et sa surface est comprise entre 0,07% et 1,5% de la surface de l'image originale, alors le contour approché par ce polygone représente le contour d'un panneau triangulaire rouge,

- **Pour le masque de couleur bleu**

Si le polygone a 4 sommets, est convexe, sa tolérance de pente des côtes horizontales et verticales inférieur à 2%, sa surface est comprise entre 0,07% et 1,5% de la surface de l'image originale et si l'inclinaison de ce polygone ne dépasse pas une erreur fixée, alors le contour approché représente un panneau rectangulaire.

- **Pour les deux masques de couleur bleu et rouge**

Si le polygone a plus que 6 sommets, est convexe et si sa surface est comprise entre 0,07% et 1,5% de la surface de l'image originale, alors c'est éventuellement un panneau circulaire bleu (pour le masque de couleur bleu) ou rouge (pour le masque de couleur rouge). Dans ce cas, une méthode, conçue spécialement pour la détection du cercle [34], est utilisée pour valider ou rejeter ce résultat.

Les panneaux, détectés selon ces critères, sont rognés et extraits de l'image couleur RGB initiale en se basant sur les coordonnées des sommets des polygones approximant les contours de ces panneaux.

Les imageries des panneaux ainsi obtenus, sont redimensionnées en 32x32 pixels et réparties, en fonction de leurs formes et de leurs couleurs, en quatre classes : panneaux triangulaires rouges, circulaires rouges, circulaires bleus et rectangulaires bleus.

La Figure 20 illustre la détection des panneaux routiers selon la méthode de détection proposée.



(a)



(b)



(c)

Figure 20: Résultats de détection des panneaux. a) Image initiale, b) imagerie des panneaux détectés et rognés, c) imagerie redimensionnée en 32X32 pixels

III.5. Evaluation du module de détection

Actuellement, il n'existe aucune base de données publique qui permettrait de comparer notre approche à la littérature. Nous avons donc testé le module de détection sur des images que nous avons préparées.

Pour évaluer les performances de ce module en termes de précision et de rapidité de détection des panneaux routiers, nous avons utilisé des images contenant plusieurs panneaux avec des arrière plans différents.

III.5.1.1. Premier test

Dans ce test, l'évaluation du module de la détection est effectuée sur des images avec arrière-plan blanc. Les résultats de cette évaluation sont résumés dans le Tableau 1.

Tableau 1: Résultat d'évaluation du module de détection sur des images avec arrière-plans blancs

Classe des panneaux	Nombre de panneaux dans l'image	Nombre de panneaux détectés	Taux de détection (%)	Temps globaux de détection (ms)
Circulaire rouge	4	4	100	9.63
	8	8	100	12.72
	12	11	92	14.27
Triangulaire rouge	4	4	100	11.38
	8	7	88	13.03
	12	11	92	14.71
Circulaire bleue	4	4	100	10.22
	8	8	100	12.27
	12	11	100	14.34
Rectangulaire bleue	4	4	100	9.11
	8	8	100	11.82
	12	10	83	13.53

III.5.1.2. Deuxième test

Dans ce test, l'évaluation du module de la détection est effectuée sur des images avec arrière-plan réel. Les résultats de cette évaluation sont résumés dans le Tableau 2.

Tableau 2: Résultat d'évaluation du module de détection sur des images avec arrière-plans réels

Classe des panneaux	Nombre de panneaux dans l'image	Nombre de panneaux détectés	Taux de détection (%)	Temps global de détection (ms)
Circulaire rouge	4	4	100	10.39
	8	7	88	12.88
	12	10	83	14.53
Triangulaire rouge	4	3	75	10.82
	8	7	88	13.27
	12	10	83	15.02
Circulaire bleue	4	4	100	11.71
	8	8	100	13.16
	12	10	83	14.96
Rectangulaire bleue	4	4	100	10.38
	8	8	100	12.20
	12	10	83	13.87

III.5.1.3. Analyse des résultats

D'après le Tableau 1 et Tableau 2, nous constatons que la méthode de détection proposée est performante aussi bien en terme du taux de détection globale (96 % dans le cas de l'arrière plan blanc et 90 % dans le cas de l'arrière plan réel) qu'en terme du temps global nécessaire pour la détection des panneaux. En effet, Ce temps ne dépasse pas 16 ms pour la détection de douze panneaux dans une image. Pour une vitesse de 120km / h soit 33,33m / s, la distance parcourue pour afficher le résultat de la détection de ces panneaux est $0.54 \text{ m} \approx 16 \text{ ms} \times 33.33 \text{ m/s}$. Ce qui signifie que, avec cette vitesse, notre méthode permettrait la détection de 12 panneaux tous les 0.54 m. Or ce nombre de panneaux dans une image de scène routière ne peut exister que dans un environnement urbain où la vitesse est limité à 60 Km/h. Dans ce cas, notre méthode permettrait la détection de douze panneaux tous les $0.27 \text{ m} \approx 16 \text{ ms} \times 16,67 \text{ m/s}$. Ceci met bien en évidence que le temps nécessaire pour la détection des panneaux routiers par la méthode proposée est largement suffisant pour la coupler avec une méthode de reconnaissance afin d'identifier les panneaux en temps réel.

III.6. Conclusion

Dans ce chapitre nous avons présenté le module de détection des panneaux de signalisation routière, qui consiste à extraire les zones rouges et bleus dans l'image de la scène de trafic routier, puis chercher les formes circulaire et triangulaire dans le masque rouge, et circulaire et rectangulaire dans le masque bleu, ceci est effectué en utilisant l'approximation polygonale des courbe numérique, en suite nous avons présenté les résultats de test de ce module, qui ont montré ses performances, avec un taux de détection qui dépasse 96% pour les panneaux avec un arrière plan blanc, et 90% pour les panneaux à arrière plan réel, et avec un temps d'exécution inférieur à 12 ms dans le cas où l'image contient 4 panneaux de signalisation routière.

Ces résultats sont largement suffisants pour coupler ce module avec un module de reconnaissance afin d'identifier les panneaux en temps réel.

Chapitre IV . Module de reconnaissance des panneaux routiers

Sommaire

IV.1. Introduction.....	45
IV.2. Classification	45
IV.3. Méthode de la reconnaissance.....	46
IV.4. Descripteur HP	46
IV.5. Descripteur HOG	47
IV.6. Classificateur MLP	49
IV.7. Classificateur SVM	51
IV.8. Préparation de la base des images de test.....	54
IV.9. Evaluation de module de reconnaissance	55
IV.9.1. Test sur les panneaux triangulaires rouges.....	56
IV.9.2. Test sur les panneaux circulaires rouges	57
IV.9.3. Test sur les panneaux circulaires bleus et rectangulaires bleus.....	58
IV.9.4. Analyses et conclusions.....	59
IV.10. Conclusion	62

IV.1. Introduction

Le module de détection des panneaux de signalisation routière a permis d'extraire de l'image les régions s'approchant le plus à un panneau routier. Toutefois, cet algorithme est loin d'être parfait. Ces candidats doivent maintenant être filtrés pour ne retenir que les panneaux réels et supprimer toutes les fausses alarmes qui pourraient alerter inutilement le conducteur. L'objectif de la reconnaissance est ainsi double, éliminer les négatifs et classifier les panneaux routiers pour déchiffrer l'information qui y est contenue.

Dans ce chapitre, nous commençons par présenter de manière générale les deux types classifieurs et les autres de descripteurs utilisés dans ce travail. Ensuite, nous décrivons nos méthodes de reconnaissance basées sur les combinaisons descripteur-classificateur. Enfin, nous évaluons les performances de notre algorithme tout en justifiant nos choix.

IV.2. Classification

L'objectif de la classification est de définir une règle permettant d'assigner tout motif inconnu \mathbf{x} d'un certain espace X à une classe donnée l à partir de connaissances *a priori*. Ces dernières sont fournies sous forme d'une base d'apprentissage, représentative de la tâche de classification. Elle représente la mémoire, ou l'expérience, du système. Elle doit notamment contenir un ensemble significatif de N motifs $\mathbf{x}_i, i \in [1, N]$ de chacune des catégories à distinguer. Si chaque motif \mathbf{x}_i a été au préalable associé à classe, étiquetée l_i , l'algorithme de classification vise à séparer au mieux l'espace entre les L classes. L'apprentissage est dit *supervisé*. En revanche, si cette distribution est inconnue, il est dit *non-supervisé* et l'objectif est de trouver le nombre optimal de classes permettant la meilleure séparation au sens de critères donnés.

La qualité du classifieur peut être appréciée selon trois aspects particulièrement importants :

- Une représentation des motifs alliant **simplicité** et **discriminabilité**. Un motif sera représenté par un ensemble de k caractéristiques qui contiennent toute son information intrinsèque ;
- Une séparation de l'espace entre les classes telle qu'elle maximise la distance interclasse et minimise la distance intra classe. Plus une classe sera compacte

dans l'espace de représentation, plus l'estimation de sa frontière sera facilitée. De même, si deux classes se superposent, il sera plus difficile, voire impossible, d'assigner un motif à une classe plutôt qu'une autre sans risque d'erreur ;

- Une bonne capacité à estimer les classes des exemples inconnus avec le minimum d'erreur, ou **capacité de généralisation**.

Parmi les célèbres algorithmes de classification, nous les réseaux de neurones (NN) ou le perceptron multi couche (MLP) et la machine à vecteurs supports (SVM) utilisés dans ce travail.

IV.3. Méthode de la reconnaissance

Le processus de reconnaissance des panneaux routier est illustré dans la Figure 21 suivante, il est composé deux principales étapes, la première est l'extraction de descripteur des panneaux, et la deuxième est de le reconnaître par un classificateur entraîné.

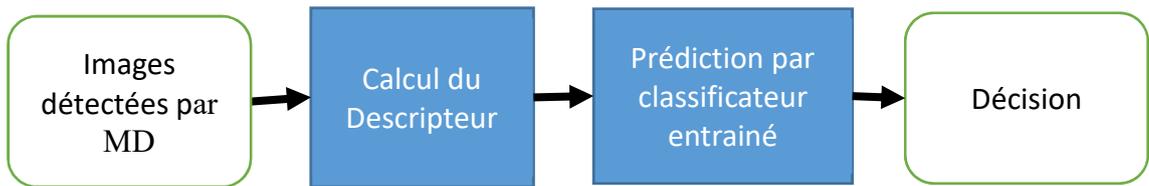


Figure 21: Schéma synoptique du module de reconnaissance.

Le descripteur est choisi de telle sorte que son temps de calculs et sa taille de donnée soient réduits pour assurer temps de traitement compétitif sans perdre la distinguabilité entre les panneaux.

Le classificateur est entraîné en utilisant par les descripteurs extraits d'une partie des images de panneaux détectés par le module de détection, et le test se fera sur une autre partie. Ces images sont sélectionnées catégorisés manuellement.

IV.4. Descripteur HP

Avant de passer les images de panneaux détectés par le module de détection, aux classificateur, l'étape d'extraction du descripteur est primordiale pour réduire la quantité de données à traiter, et par suite réduire le temps de calcul du programme.

On procède par une extraction de descripteur HP de chaque image de panneaux routiers qui sont présentés au module de reconnaissance sous la taille de 32x32 pixels ce qui est 1024 pixels. Pour les trois couches de couleur rouge, verte et bleue, la taille du vecteur d'entrée du classificateur aura une taille de 3072, et cela va certainement ralentir le programme.

Pour chaque canal une projection des pixels est effectuée sur les deux axes vertical et horizontal.

Alors pour chaque point CX_i sur l'axe horizontal :

$$CX_i = \frac{1}{255} \times \frac{1}{32} \sum_{j=1}^{32} C_{i,j} \quad i = 1,2,3,\dots,32 \quad (11)$$

Alors pour chaque point CY_j sur l'axe vertical :

$$CY_j = \frac{1}{255} \times \frac{1}{32} \sum_{i=1}^{32} C_{i,j} \quad j = 1,2,3,\dots,32 \quad (12)$$

$C_{i,j}$ est l'intensité du pixel de coordonnées (i, j) de la couche de couleur C ;

C est la couche de couleur (Rouge, Verte ou Bleue) ;

CX_i et CY_j sont des valeurs normalisées entre 0 et 1.

Le nouveau descripteur est composé de 192 éléments, les 32 éléments de CX_i , et les 32 autres de CY_j pour les trois couches de couleur RVB.

IV.5. Descripteur HOG

L'histogramme de gradient orienté (HOG) est un descripteur utilisé au domaine de la vision par ordinateur pour caractériser les objets. Son principe est de calculer des histogrammes locaux de l'orientation du gradient sur des zones régulièrement réparties sur l'image. Ce classificateur est largement utilisé dans la reconnaissance des personnes suite à son efficacité particulière notée dans de type de problèmes.

Le calcul du descripteur se fait en suivant quatre étapes :

- Calcule du gradient : consiste à appliquer un filtre dérivatif centré d'une seule dimension, dans les directions horizontales et verticales. Les masques suivants sont utilisés sont : $[-1,0,1]$ et $[-1,0,1]^T$. Le gradient est calculé pour chacune des couches RVB, et retient le gradient ayant la plus grande norme pour chaque pixel.
- Construction de l'histogramme : consiste à la création des histogrammes de l'orientation des gradients. Ceci est fait dans des cellules carrées de petite taille (de 4x4 à 12x12 pixels). Chaque pixel de la cellule vote alors pour une classe de l'histogramme, en fonction de l'orientation du gradient à ce point. Le vote du pixel est pondéré par l'intensité du gradient en ce point. Les histogrammes sont uniformes de 0 à 180° (cas non signé) ou de 0 à 360° (cas signé). On obtient de meilleurs résultats avec un histogramme quantifié à 9 classes
- Formation des blocs : consiste à regrouper plusieurs cellules dans un bloc, qui est l'unité sur laquelle est effectuée la normalisation. Les blocs se recouvrent, donc une même cellule participe plusieurs fois au descripteur final, comme membre de blocs différents. Les blocs rectangulaire (R-HOG) contenant 3x3 cellules de 6x6 pixels chacune, ont de meilleure performance.
- Normalisation des blocs : consiste à normaliser le descripteur fin d'éviter les disparités dues aux variations d'illumination, quatre types de normalisation sont proposés. Le vecteur non normalisé contenant tous les histogrammes d'un seul bloc est désigné par v , sa k-norme par $\|v\|_k$ et ε est une constante de faible valeur. Le facteur de normalisation est alors défini par :

- L2-norme : $f = \frac{v}{\sqrt{\|v\|_2^2 + \varepsilon^2}}$

- L1-norme : $f = \frac{v}{\|v\|_1 + \varepsilon}$

- L1-racine : $f = \sqrt{\frac{v}{\|v\|_1 + \varepsilon}}$

- L2-hys : consistant à calculer v tout d'abord par la L2-norme, puis limiter les valeurs maximales de v à 0.2.

Le HOG utilisé dans note cas est appliquée sur des objets de taille 64x64 pixels. Nous avons pris des cellules de taille 8x8 pixels, et des blocs de taille 16x16 pixels (4x4 cellules) et c'est la configuration par défauts (recommandés par les auteurs de l'algorithme) dans la bibliothèque OpenCV 2.44.

IV.6. Classificateur MLP

Le module de reconnaissance est implémenté par l'utilisation des réseaux de neurones artificiels avancés, ou plus particulièrement un perceptron multicouche (MLP), le type des réseaux de neurones artificiels le plus utilisé dans résolution des problèmes de classification.

Ce sont un modèle mathématique inspiré du fonctionnement du cerveau des êtres vivants, et qui est le plus utilisé dans l'apprentissage automatique « *Machine Learning* », ils étaient initialement proposés dans les années quarante, ils attiraient peu d'intérêt au départ, mais peu de temps après, ils devenaient les plus utilisés suite à leurs algorithmes d'entraînement performants, et leur compétitive puissance de calcul.

Ils sont devenus très utilisés à nouveau, depuis l'introduction des auto-encodeurs, des réseaux conventionnels et les autres techniques qui améliorent significativement leur performance.

Le MLP contient au moins trois couches, la première est la couche d'entrée, la dernière est la couche de sortie, et une ou plusieurs couches cachées au milieu, chaque couche du MLP contient un ou plusieurs neurones directement liés aux neurones de la couche précédente, at à ceux de la couche suivante

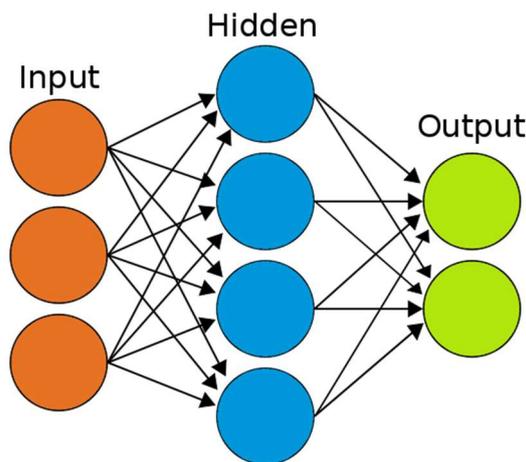


Figure 22: structure de la multi couche perceptron

La Figure 22 représente un exemple d'un perceptron à trois couches avec trois entrées, deux sorties et la couche cachée contenant quatre neurones. Tous les neurones de la

CHAPITRE IV. Module de reconnaissance des signaux

MLP sont similaires, chacun d'eux a plusieurs liens d'entrée (il prend les valeurs de sortie de plusieurs neurones dans la couche précédente comme des entrées) et plusieurs liens de sortie (il passe la réponse à plusieurs neurones de la couche suivante). Les valeurs reçues depuis la couche précédente sont additionnées avec certains poids individuels pour chaque neurone, plus le terme de biais, la somme est transformée en utilisant la fonction d'activation f qui peut être aussi différente pour différents neurones, (figure), autrement dit, les entrées qui sont les sorties x_j de couche n , et les sorties y_i de la couche $n+1$ sont calculées comme suit :

$$u_i = \sum_j (w_{i,j}^{n+1} \times x_j) + w_{i,bias}^{n+1} \quad (13)$$

$$y_i = f(u_i) \quad (14)$$

La fonction d'activation qui est utilisée dans ce travail est la fonction sigmoïde binaire qui est définie comme suit :

$$f(x) = \beta \times \frac{1 - e^{-\alpha x}}{1 + e^{-\alpha x}} \quad (15)$$

Avec $\alpha = 1$ et $\beta = 1$, comme il est indiqué dans la figure, la fonction sigmoïde basiquement a quelques propriétés mathématiques très utiles, la monotonie et la continuité.

La monotonie signifie que $f(x)$ reste toujours soit croissante, soit décroissante : son sens de variation est constant, et la continuité veut dire qu'elle n'est pas brisée de quel que soit x , elle est lisse.

Ces paramètres sont des propriétés intrinsèques, ils assistent éventuellement la puissance des réseaux de neurones à estimer et généraliser sur les fonctions, par apprentissage.

Le module de la reconnaissance utilisé dans ce travail est divisé en quatre classificateurs MLP, chacun est utilisé pour un type de panneaux, (circulaire bleu, rectangulaire bleu, circulaire rouge et triangulaire rouge).

IV.7. Classificateur SVM

Les Machine à Vecteurs de support o les Séparateurs à Vastes Marges « Support Vector Machine » SVM, sont des classificateurs qui reposent sur deux idées clés, qui permettent de traiter des problèmes de discrimination non linéaire, et de reformuler le problème de classement comme un problème d'optimisation quadratique.

La première idée clé est la notion de marge maximale. La marge est la distance entre la frontière de séparation et les échantillons les plus proches. Ces derniers sont appelés vecteurs supports. Dans les SVM, la frontière de séparation est choisie comme celle qui maximise la marge. Ce choix est justifié par la théorie statistique de l'apprentissage, qui montre que la frontière de séparation de marge maximale possède la plus petite capacité. Le problème est de trouver cette frontière séparatrice optimale, à partir d'un ensemble d'apprentissage. Ceci est fait en formulant le problème comme un problème d'optimisation quadratique, pour lequel il existe des algorithmes connus.

Afin de pouvoir traiter des cas où les données ne sont pas linéairement séparables, la deuxième idée clé des SVM est de transformer l'espace de représentation des données d'entrées en un espace de plus grande dimension (possiblement de dimension infinie), dans lequel il est probable qu'il existe une séparation linéaire. Ceci est réalisé grâce à une fonction noyau, et qui a l'avantage de ne pas nécessiter la connaissance explicite de la transformation à appliquer pour le changement d'espace. Les fonctions noyau permettent de transformer un produit scalaire dans un espace de grande dimension, ce qui est coûteux, en une simple évaluation ponctuelle d'une fonction. Cette technique est connue sous le nom de kernel trick.

Les SVM peuvent être utilisés pour résoudre des problèmes de discrimination, c'est-à-dire décider à quelle classe appartient un échantillon, ou de régression, c'est-à-dire prédire la valeur numérique d'une variable

Les techniques SVM (non linéaires) font appel à une fonction implicite Φ transformant l'espace d'entrée $X \subset R^d$ en un espace hilbertien $H, \langle \cdot, \cdot \rangle$ de plus grande dimension. L'apprentissage s'effectue alors à partir du modèle $(\Phi(X), Y)$ dans l'espace H , de dimension plus grande certes, mais dans lequel on espère que les données soient "davantage linéairement séparables". Du point de vue pratique, il convient de noter que le calcul des projections $\Phi(X)$ n'est pas utilisé dans la méthode, seuls les produits scalaires

$\langle \Phi(x), \Phi(x') \rangle$ et $(x, x') \in \mathcal{X}^2$, sont requis. Or, ceux-ci sont donnés par un noyau K , via la relation ("kernel trick") :

$$K(x, x') = \langle \Phi(x), \Phi(x') \rangle \quad (16)$$

La méthode requiert donc de sélectionner un noyau (ainsi que d'autres paramètres). Parmi les choix possibles, on compte en particulier :

- Le noyau linéaire : $K(x, x') = \langle x, x' \rangle$ (correspondant aux SVM linéaires)
- Le noyau Gaussien radial (Gaussien RBF) $K(x, x') = e^{-\sigma \|x-x'\|^2}$
- Les noyaux polynomiaux $K(x, x') = (\alpha + \beta \langle x, x' \rangle)^\delta$
- Le noyau radial de Laplace (Laplace RBF) $K(x, x') = e^{-\sigma \|x-x'\|}$
- Le noyau tangent hyperbolique (sigmoïde) $K(x, x') = \tanh(\alpha + \beta \langle x, x' \rangle)$

Un classifieur SVM est de la forme :

$$C(X) = \text{sign}(\langle \omega, \Phi(X) \rangle + b), \quad (17)$$

où $\omega \in H$ et $b \in \mathbb{R}$ sont des paramètres ajustés lors de la phase d'apprentissage à partir d'un échantillon d'exemples $\{(x_i, y_i) : 1 \leq i \leq n\}$. La frontière associée à cette règle de décision a pour équation : $\langle \omega, \Phi(x) \rangle + b = 0$. Elle correspond à un hyperplan dans l'espace H , mais est beaucoup plus complexe dans X (selon la forme du noyau choisi). Dans H , l'hyperplan est obtenu en maximisant la marge séparant les deux classes, ce qui revient à résoudre un problème d'optimisation sous contraintes linéaires :

$$\text{minimiser } \frac{1}{2} \|\omega\|^2 + \frac{C}{m} \sum_{i=1}^n \xi_i \quad (18)$$

sous les contraintes :

$$\forall i \in \{1, \dots, n\}, \xi_i \geq 0 \text{ et } y_i (\langle \omega, \Phi(x_i) \rangle + b) \geq 1 - \xi_i \quad (19)$$

On peut montrer que la solution ω peut s'exprimer de la façon suivante :

$$\omega = \sum_{i=1}^n \alpha_i y_i \Phi(x_i), \quad (20)$$

les indices i pour lesquels $\alpha_i \neq 0$ sont ceux pour lesquels l'égalité est réalisée dans la contrainte, les points x_i correspondants sont appelés vecteurs supports (de la décision). Les coefficients α_i désignent les solutions du problème quadratique dual :

$$\text{maximiser } \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{1 \leq i, j \leq n} \alpha_i \alpha_j y_i y_j K(x_i x_j) \quad (21)$$

sous les contraintes :

$$\forall i \in \{1, \dots, n\}, 0 \leq \alpha_i \leq \frac{C}{n} \text{ et } \sum_{i=1}^n \alpha_i y_i = 0 \quad (22)$$

Le paramètre C contrôle la complexité du classifieur dans la mesure où il détermine le coût d'une mauvaise classification : plus C est grand, plus la règle obtenue est complexe (le nombre de points pour lesquels on veut minimiser l'erreur de classification croît). Cette approche est appelée *C-classification*.

Une autre façon de contrôler la complexité (i.e. le nombre de vecteurs supports), appelée *v-classification*, revient à considérer, à la place du problème dual décrit ci-dessus, le problème suivant :

$$\text{minimiser } \frac{1}{2} \sum_{1 \leq i, j \leq n} \alpha_i \alpha_j y_i y_j K(x_i x_j) \quad (23)$$

sous les contraintes :

$$\forall i \in \{1, \dots, n\}, 0 \leq \alpha_i \leq \frac{1}{n} \text{ et } \sum_{i=1}^n \alpha_i \geq v \text{ et } \sum_{i=1}^n \alpha_i y_i \geq v, \quad (24)$$

où $v \in [0, 1]$ est un paramètre approchant le pourcentage de vecteurs supports parmi les données d'apprentissage.

Dans le cas où la variable de sortie Y compte plus de deux modalités, il existe plusieurs façons d'étendre directement les méthodes du cas binaire.

"Un contre un". Dans le cas où l'on cherche à prédire un label pouvant $K \geq 3$ modalités, on peut considérer toutes les paires de labels (k, l) possibles, $1 \leq k < l \leq K$ (il y en a C_K^2) et ajuster un classifieur $C_{k,l}(X)$ pour chacune d'entre elles. La prédiction correspond alors au label qui a gagné le plus de "duels".

"Un contre tous". Pour chaque modalité k , on apprend un classifieur permettant de discriminer entre les populations $Y = k$ et $Y \neq k$. A partir des estimations des probabilités a posteriori, on affecte le label estimé le plus probable.

IV.8. Préparation de la base des images de test

Pour entraîner les quatre combinaisons descripteur-classificateur proposées, la base de données des images réelles de trafic routier utilisée contient 300 images avec un arrière-plan complexe sous des conditions variables, et avec des tailles de 1300x800 pixels, nous utilisons le module de détection des panneaux routiers présenté dans le chapitre précédent, pour extraire les objets ressemblants aux panneaux de signalisation routière comme il est montré dans la Figure 23, redimensionné à 64x64 pixels, et après une classification manuelle en rassemblant les images de chaque type de panneaux apart, même les objets non-significatifs. Puis les trier selon les quatre grandes catégories forme-couleur, à savoir : les triangles-rouges, les cercles-rouges, les cercles-bleues et les rectangles-bleus.

Ces bases de données d'images de panneaux seront utilisées pour entraîner nos classificateurs.

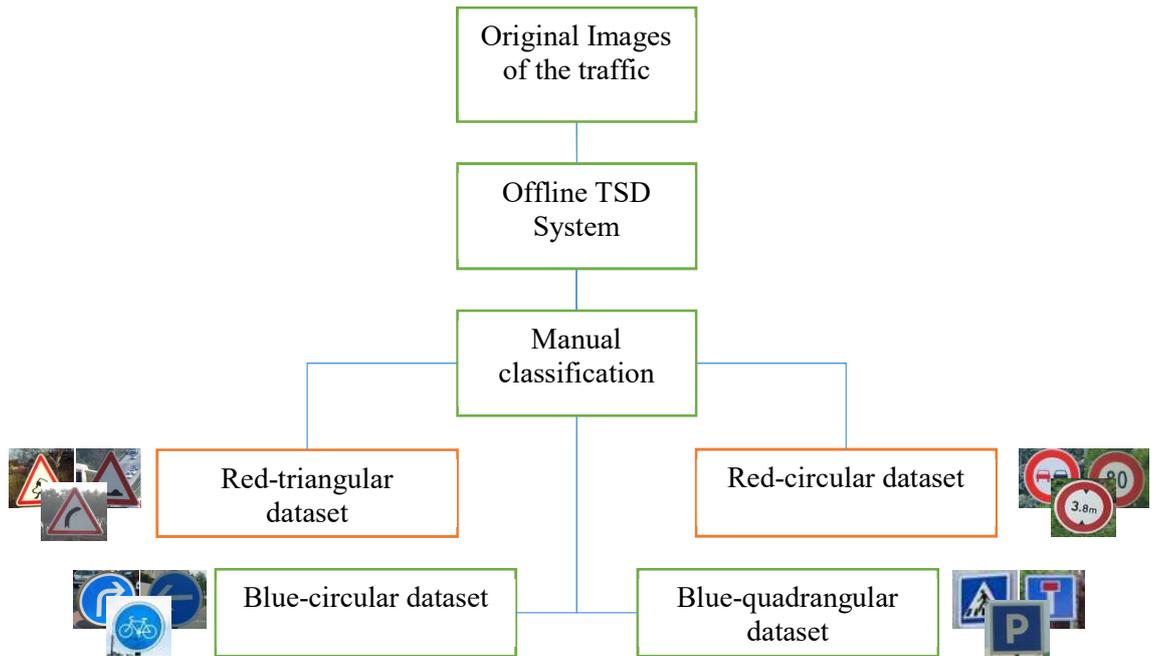


Figure 23: Structure de programme de génération des bases de données de test.

La base de données globale générée contient 2535 images de panneaux routiers, classifiés en 49 classes + une classe des panneaux non significatifs, 1724 images seront utilisées pour l'apprentissage, 811 autres seront réservées pour les tests.

IV.9. Evaluation de module de reconnaissance

Le module de reconnaissance est une combinaison de « descripteur-classificateur », pour les deux descripteurs HP et HOG, et les deux classificateurs MLP et SVM, quatre combinaisons sont possibles : HP-MLP, HP-SVM, HOG-MLP et HOG-SVM.

Une étude comparative sera menée dans cette section pour identifier parmi les quatre méthodes celle qui a les meilleures performances.

Pour tester les performances de ces méthodes, nous avons calculer les descripteurs HP et HOG des 1724 images en couleur de panneaux routiers de la base de données d'apprentissage, puis les passer en entrée à l'algorithme d'apprentissage des classificateurs MLP et SVM.

Nous extrayons les descripteur HP et HOG de chaque panneau routier de la base de test, pour le passer en suite au classificateurs entraînés qui vont identifier sa signification.

Les tests des performances des quatre combinaisons « descripteur-classificateur » est effectué séparément sur les base de données des quatre grandes catégories forme-couleur, à savoir : les triangles-rouges, les cercles-rouges, les cercles-bleues et les rectangles-bleus.

Les algorithmes d'entraînement et de test sont décrits dans l'organigramme générique, présenté dans la Figure 24.

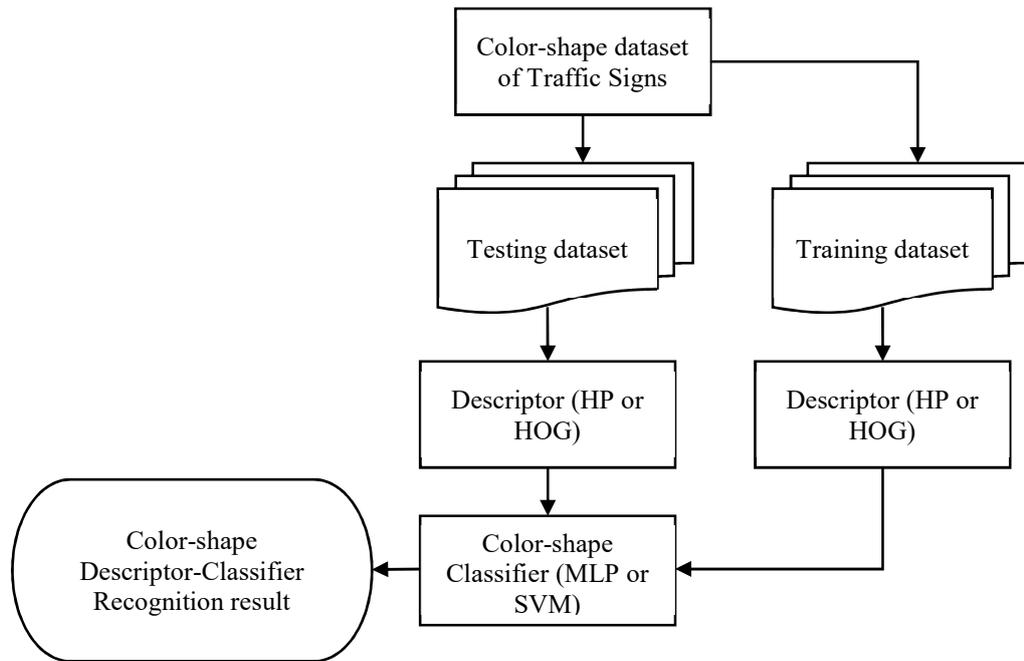


Figure 24: Organigramme générique d'entraînement et de test des quatre combinaisons « descripteur-classificateur »

IV.9.1. Test sur les panneaux triangulaires rouges

Les résultats de test des quatre méthodes de reconnaissance sur les panneaux routiers triangulaires rouges sont présentés dans graphique de la Figure 25.

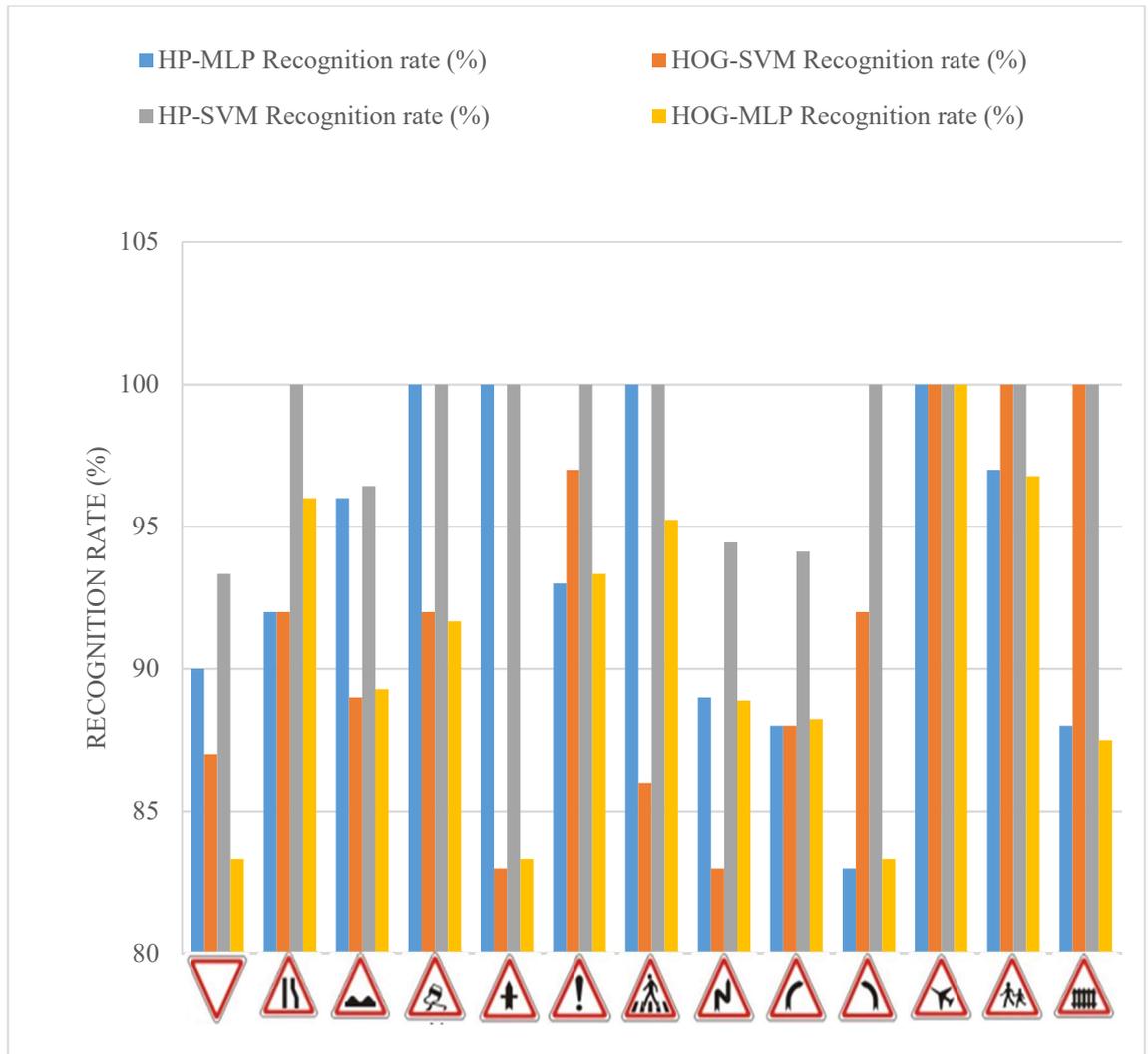


Figure 25: Taux de reconnaissance pour les panneaux triangulaires rouges.

On constate que pour ce type de panneaux routier la méthode HP-SVM (couleur grise) présente un taux de reconnaissance le plus élevé, HP-MLP vient en deuxième suivie de HOG-SVM puis HOG-MLP. Cela est à la capacité de caractérisation du descripteur HP, est à l'efficacité de discrimination de l'hyperplan calculée de SVM.

IV.9.2. Test sur les panneaux circulaires rouges

Le test effectué sur les panneaux circulaires rouges a confirmé les résultats du test précédent, en mettant au premier rang les performances de la méthode HP-SVM comme le montre la Figure 26 suivante.

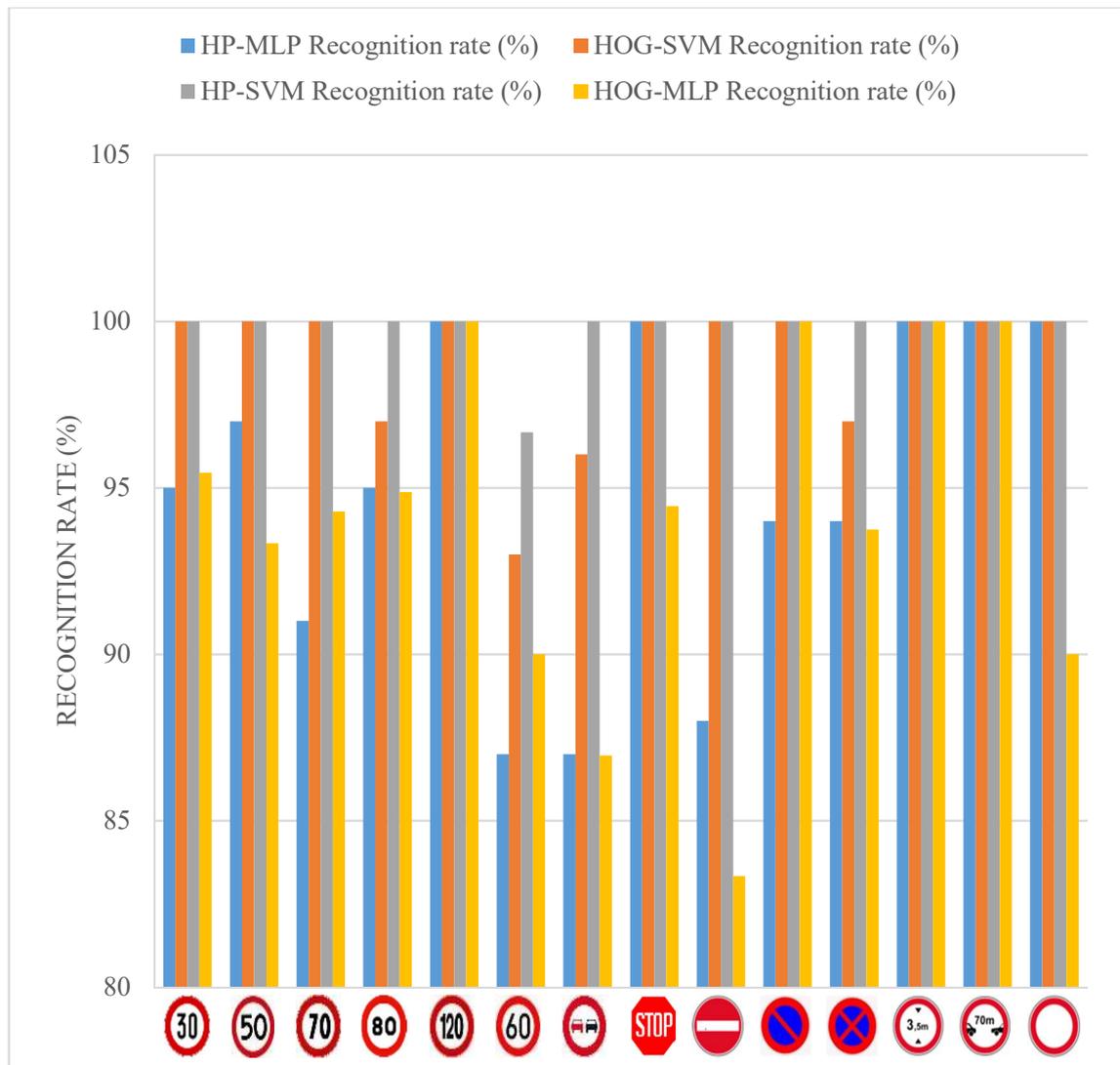


Figure 26: Taux de reconnaissance pour les panneaux circulaires rouges.

IV.9.3. Test sur les panneaux circulaires bleus et rectangulaires bleus

Les panneaux bleus de formes circulaires et rectangulaire sont testés ensemble, les résultats obtenus ; présentés dans la Figure 27 ; s'alignent avec les tests précédents, en mettant en évidence les performances de la méthode HP-SVM comme étant la meilleure méthode adaptée à la reconnaissance de tous les types de panneaux de signalisation routière.

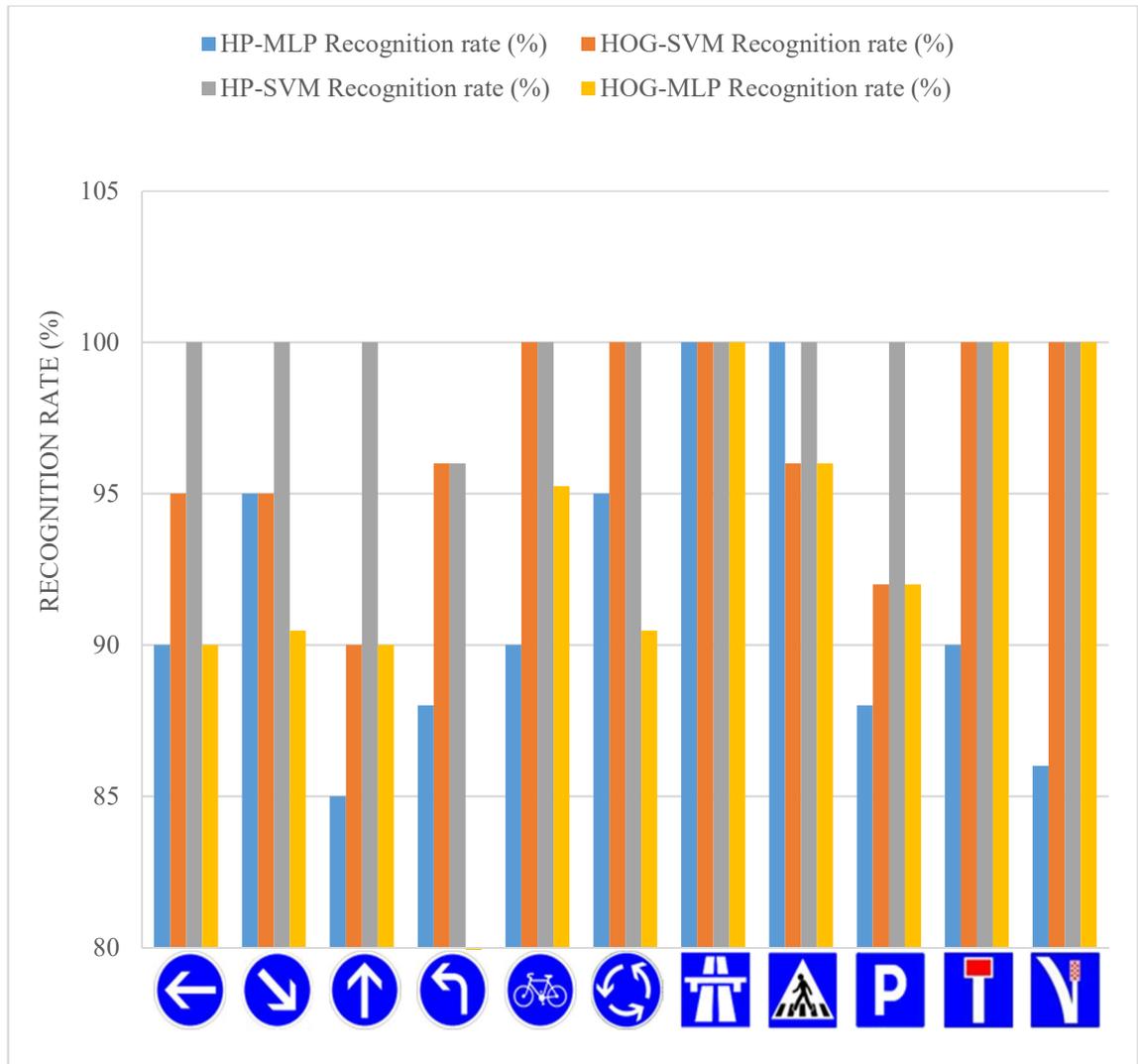


Figure 27: Taux de reconnaissance pour les panneaux circulaires et rectangulaires bleus.

IV.9.4. Analyses et conclusions

En résumé, les taux de reconnaissance globales des quatre méthodes (HP-MLP, HP-SVM, HOG-MLP et HOG-SVM) groupés dans la Figure 28, montrent que la méthode HP-SVM donne un taux de reconnaissance globale plus important que les autres méthodes. Et cela est due ; comme nous l'avons cité précédemment, d'une part à la capacité du descripteur HP de caractériser les types de panneaux routiers et d'autre part au pouvoir de discrimination des classes, présenté par l'hyperplan séparateur calculé par la SVM.

Les résultats présentés par la méthode HP-SVM, ne s'arrêtent pas à un taux élevé de reconnaissance de panneaux routiers, elle présente aussi un temps global de reconnaissance

CHAPITRE IV. Module de reconnaissance des signaux

très compétitive par rapport aux autres techniques (voir Figure 29), et qui ne dépasse pas 352ms, ce qui est nécessaire pour une implémentation en temps réel.

Ce temps de reconnaissance globale de la méthode HP-SVM, permet l'identification de plus de trois panneaux routiers par seconde.

A une vitesse d'approche de 120 km/h, la distance de visibilité des panneaux routiers (la distance à partir de laquelle le panneau devient visible) est au moins 200m (voir Tableau 3), ceci laisse le panneau visible au chauffeur durant un temps de visibilité de plus 6 seconde.

En divisant ce temps de visibilité par le temps globale de reconnaissance 0.352ms nous aurons plus de 17 fois que le panneau sera traité par le système de reconnaissance, cela veut dire que aucun panneau ne passera inaperçus au système de reconnaissance.

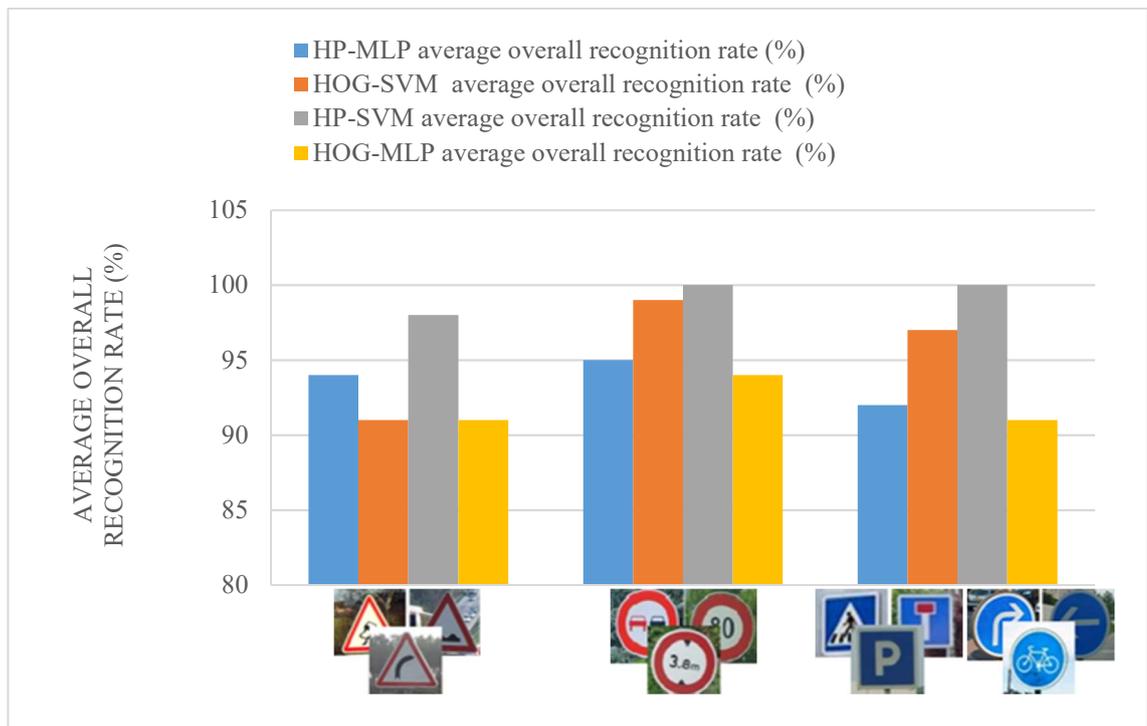


Figure 28: Taux de reconnaissance globales.

CHAPITRE IV. Module de reconnaissance des signaux

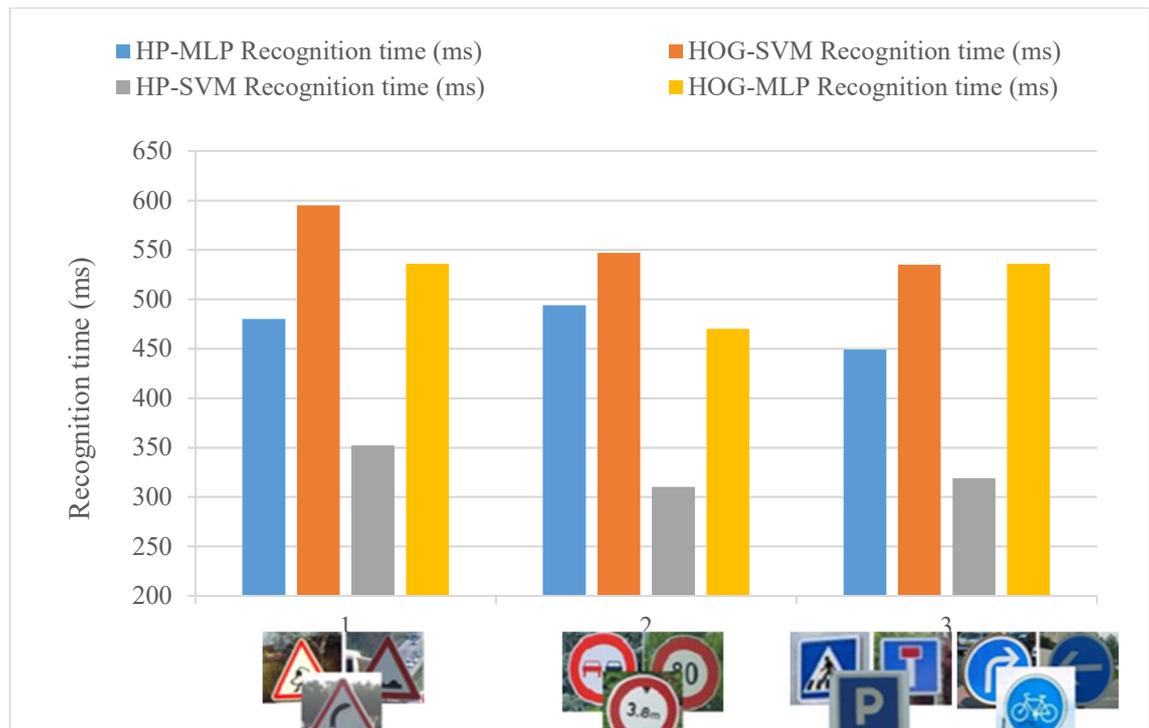


Figure 29: Temps de reconnaissance globales

Tableau 3: Distances de visibilité des panneaux routiers en fonction de la vitesse d'approche¹³

Vitesse d'approche en Km/h	Distance de visibilité normale en m	Distance de visibilité minimum en m
120	360	200
110	300	180
100	250	160
80	160	120
60	90	80

La méthode HP-SVM sera utilisée comme noyau du système de reconnaissance des panneaux de signalisation routière, qui sera combiné avec le système de détection traité dans le chapitre précédent, pour former un system complet de détection et de reconnaissance de panneaux de signalisation routière « Traffic Sign Detection and recognition System (TSDS) » dont nous réservons le chapitre suivant.

¹³ Extrait du « SIGNALISATION ROUTIERE : Guide technique des marques sur chaussées », Ministère des Travaux Publics De la Formation Professionnelle Et de la Formation des Cadres, Section C.1.2 page 31.

IV.10. Conclusion

Dans ce chapitre nous avons présenté les méthodes de reconnaissance adoptées dans ce travail, qui consistent à combiner un classificateur MLP ou SVM avec un descripteur de l'image HP ou HOG.

Les quatre combinaisons sont développées et tester pour comparer leurs performances. Les résultats de test, étalés dans ce chapitre, montrent que la combinaison HP-SVM a réalisé le meilleur taux de reconnaissance qui atteint 99,33%, cela est dû à la grande capacité de discrimination du descripteur HP, et à la généralisation minutieuse des SVM pour les différentes classes reconnues.

Les performances de la HP-SVM ne se limite pas à un taux de reconnaissance élevé, mais elle présente un temps d'exécution record par rapport aux autres méthodes, grâce à la simplicité de calcul de HP, et la faible complexité des SVM.

La méthode HP-SVM est adopté tant que noyau du module de reconnaissance des panneaux de signalisation routière, ce dernier qui va être mis en cascade avec le module de détection pour former le système complet de détection et de reconnaissance des panneaux de signalisation routière, sujet du chapitre suivant.

Chapitre V . Système complet

Sommaire

V.1. Introduction	64
V.2. Structure générale du système complet.....	64
V.3. Test hors ligne	67
V.3.1. Test sur ordinateur	67
V.3.2. Test sur Raspberry Pi B+	69
V.4. Test en ligne.....	70
V.4.1. Test sur ordinateur	70
V.4.2. Test sur Raspberry Pi B+	71

V.1. Introduction

Ce consacré à au système complet de détection et de reconnaissance des panneaux routiers « Traffic Sign Detection and recognition System (TSDS) », d'une part, pour étaler les techniques d'intégration mises en ouvres pour mettre en cascade les deux modules de détection et de reconnaissance conçus et développés précédemment dans les chapitres III et IV, et d'autre part pour présenter les processus de test de ce système TSDS, en mode hors ligne :sur les images fixes de la scène de trafic routier, et en mode en ligne : en temps réel, en embarquant, le caméscope et l'ordinateur sur un véhicule.

Nous analysons et discutons aussi les résultats obtenus des différents tests effectués sur ce système. Une évaluation sur la carte électronique Raspberry Pi B+, est présentée pour valider la fiabilité du système comme étant un système d'aide à la conduite.

V.2. Structure générale du système complet

Après avoir amélioré de manière indépendante de module de détection et le module de reconnaissance, le système complet qui intègre ces deux modules, est l'association en cascade de ces deux derniers.

Les images en couleur de panneaux routiers extraites par le module de détection sont passées en entrée au module de reconnaissance entraîné pour identifier le type de panneau en question.

La Schéma Bloc du système complet TSDS. Figure 30 présente le schéma bloc du système complet TSDS.

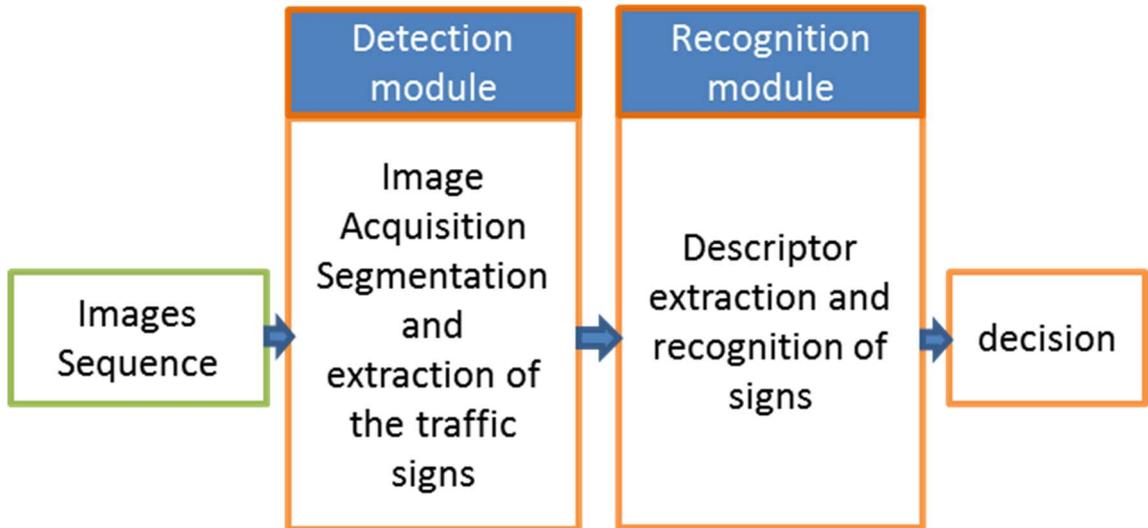


Figure 30: Schéma Bloc du système complet TSDS.

Le module de détection ne fait pas seulement de la détection de zone susceptible de contenir des panneaux routiers, mais il fait aussi une première classification selon la forme et la couleur du panneau, que nous allons en profiter de réduire le nombre de classe à identifier par le classificateur, ce qui augmente visiblement la capacité de discrimination des classes.

Quatre copies du classificateur sont entraînées individuellement sur chacune des bases de données de panneaux routiers (triangulaire rouge, circulaire rouge, circulaire bleue et rectangulaire bleue), et exécutés en parallèle et indépendamment en utilisant la programmation en multitâche sous le Language C/C++.

L'organigramme de la Figure 31 présente le déroulement des différents étapes et processus du système TSDS.

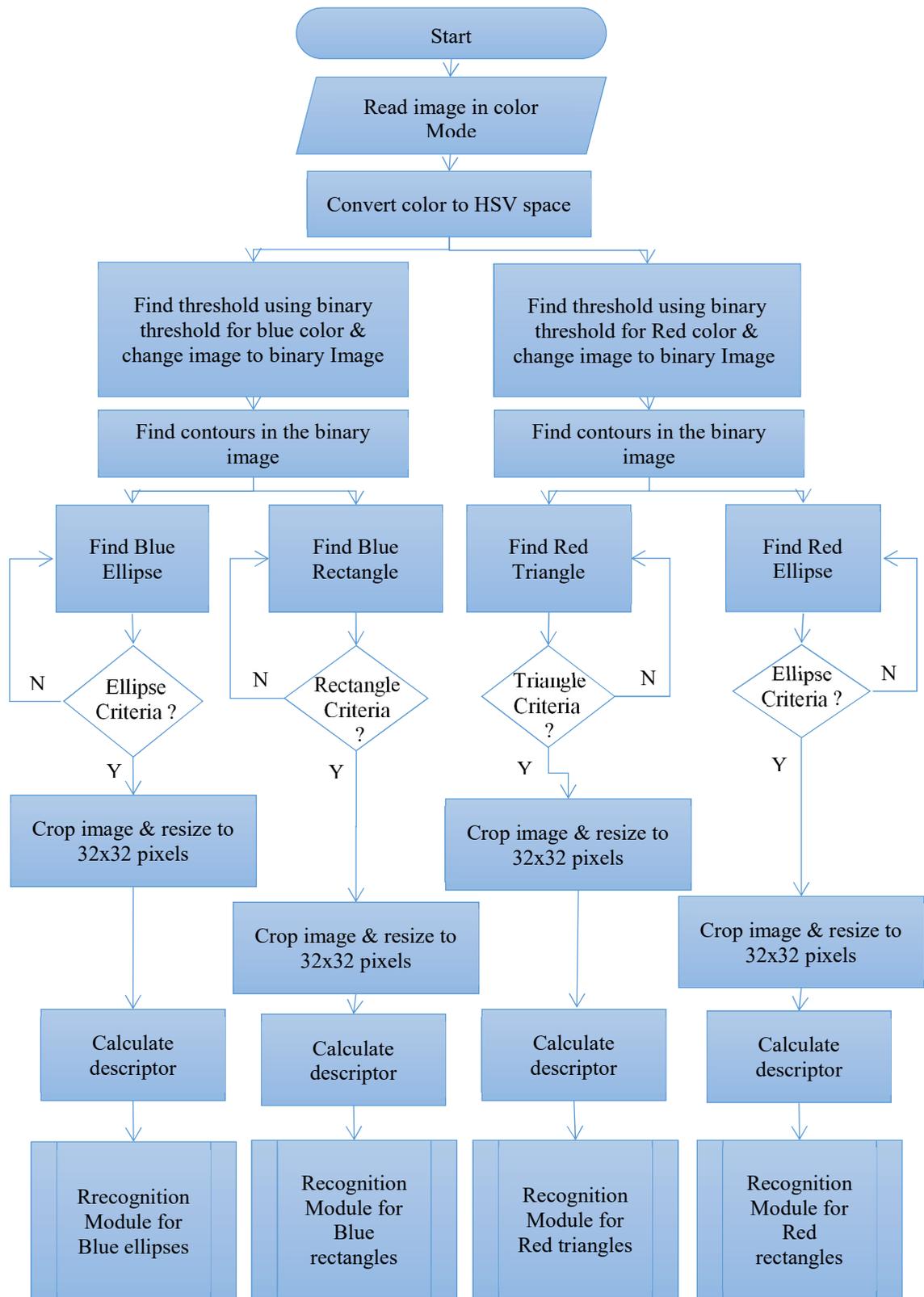


Figure 31: Organigramme général de système complet TSDS.

V.3. Test hors ligne

Le test hors ligne est effectué sur une large base de données d’image de trafic routier sous différentes conditions de météo et de luminosité, contenant plus que 3000 images de taille 1300x800 pixels, collectée de plusieurs sites sur internet, et une bonne partie prise avec ma caméra sur les routes lors de mes voyages nationaux et internationaux.

Les figures suivantes montrent un échantillon des images de la scène routière utilisées dans le test hors ligne.



Figure 32: Echantillon des images de la scène routière.

V.3.1. Test sur ordinateur

Le programme du système TSDS est codé sur le langage C/C++ en utilisant le Microsoft Visual Studio 2012. Nous avons développé, en plus des programmes backgrounds, une interface graphique (voir Figure 1Figure 33) avec des menus permettant l’accès aux différentes fonctionnalités et réglages du TSDS, en partant de l’ouverture des fichiers images et flux vidéo, en passant par le réglage du module de détection (choix des critères des formes, réglage seuils de couleur...) jusqu’aux réglages de du module de reconnaissance (choix du descripteur et du classificateur...).

CHAPITRES V. Système complet TSDS

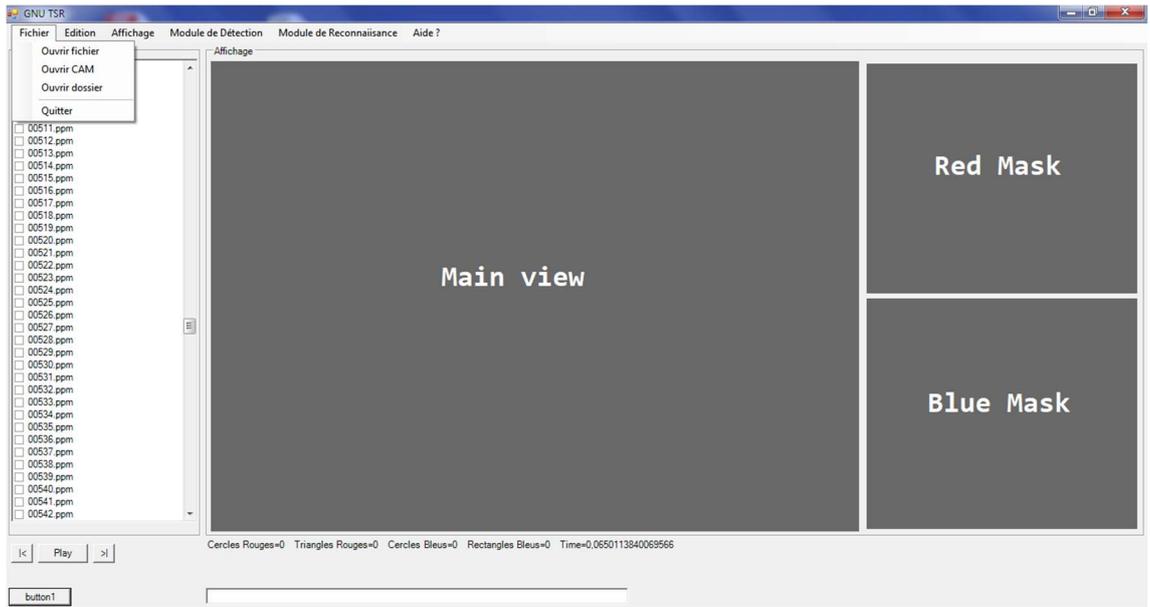


Figure 33: Capture d'écran du système complet TSDS.

Le test est effectué sur un nombre des images fixes de la scène du trafic routier contenant 1000 panneaux routiers de différentes formes et couleurs.

Les résultats obtenus sont présentés dans le Tableau 4 suivant.

Tableau 4: résultats expérimentaux en mode hors line du système complet TSDS

Number of Real Signs NRS	Number of Signs Detected NSD	Number of Failed Signs in Detection NFSD	Number of Signs Recognized NSR	Number of Failed Signs in Recognition NFSR
1000	1007	32	993	14

On constate que pour la détection, le nombre de panneaux détectés NSD (y compris les faux panneaux) est 1007 supérieur au nombre des vrais panneaux NRS présente dans les images de trafic routiers, ce qui tout à fait normale. Le nombre des panneaux ratés par la détection NFSD (parmi les vrais panneaux) est 32 panneaux.

Le nombre des vrais panneaux détecté (NRSD) par le système de détection est donc :

$$NRSD = NRS - NFSD = 1000 - 32 = 968$$

Le taux global de la détection (GDR) sera donc donné par :

$$GDR = \frac{NRSD}{NRS} \times 100 = \frac{968}{1000} \times 100 = 96.8\%$$

Ce qui s'aligne avec les résultats obtenus pour le module de détection testé tout seul.

Quant à la reconnaissance on constate que le nombre de panneaux reconnus NSR (y compris le faux panneaux) est 993 panneaux. Le taux global de la reconnaissance (GRR) sera calculé comme le suivant :

$$GRR = \frac{NSR}{NSD} \times 100 = \frac{993}{1007} \times 100 = 98.6\%$$

Le taux de détection et de reconnaissance du système complet TSDS « Global Rate of Detection and Recognition (GRDR) » sera le résultat du produit du GDR et GRR, car les deux modules de détection de et reconnaissance sont monté en cascade.

$$GRDR = GDR \times GRR = 0.968 \times 0.986 = 0.9544 = 95.44\%$$

V.3.2. Test sur Raspberry Pi B+

La carte raspberry Pi B+ est tourne sous le système d'exploitation Linux Raspbian, une distribution de Linux Debian adaptée aux cartes Raspberry.

Pour pouvoir implémenter notre système TSDS, sur cette carte, il nous a fallu de changer le langage de programmation, cette fois au lieu de C/C++, nous utilisons Python car il est plus adapté et maniable sur Linux, de plus la bibliothèque OpenCV exit sous une version open source, codé sous le langage Python.

Un travail colossal d'installation et de configuration de la carte Raspberry Pi B+ a été fait avant la tâche de convertir les code C/C++ en python.

Les tests sont effectués sur la même base donnée d'image de trafic routier utilisé sur l'ordinateur. La Figure 34 présente une capture d'écran de déroulement des tests sur la carte Raspberry Pi B+.

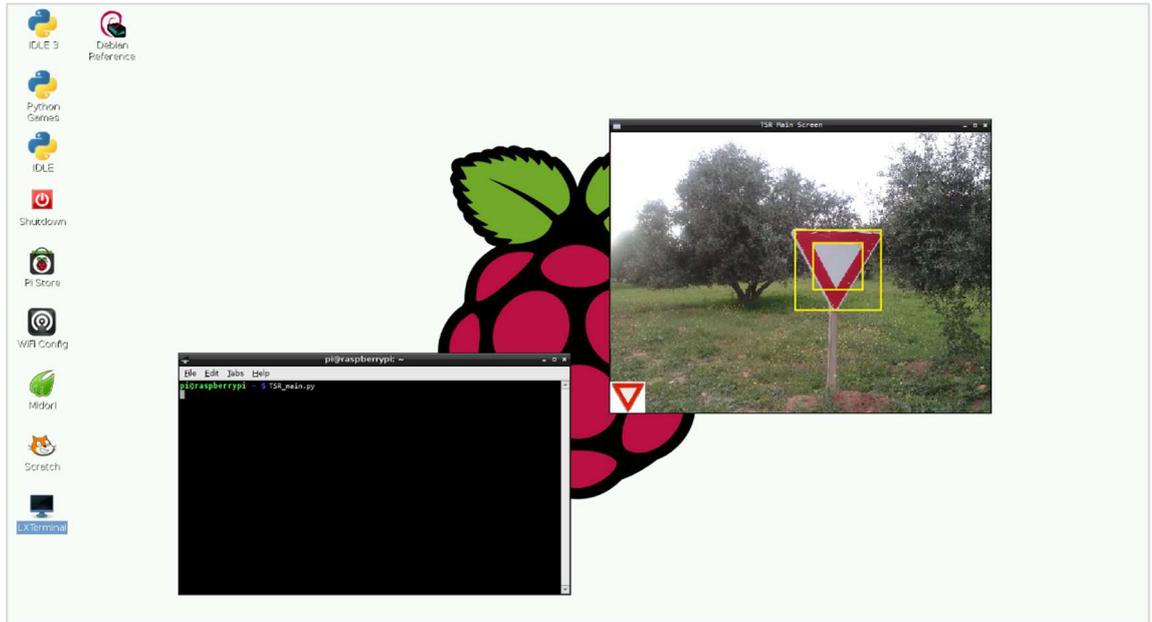


Figure 34: Capture d'écran du système complet TSDS sur la carte Raspberry Pi B+.

Les résultats obtenus sur Raspberry sont les mêmes que ceux obtenus sur ordinateur, avec un retard remarquable en ce qui concerne de durée de test. Cela est justifié par le fait que la configuration en terme de vitesse de microprocesseur et capacité de mémoire de la Raspberry est moins performante que celle de l'ordinateur.

V.4. Test en ligne

Le test en ligne est effectué en temps réel sur des vidéos de trafic routier, capturées sur la route, sous différentes conditions de météo et de luminosité.

Nous avons aussi embarqué le caméscope avec le PC sur une voiture, pour faire des tests directs sur la route. La figure présente la plateforme de test montée à bord de la voiture.

V.4.1. Test sur ordinateur

En temps réel, ou sur les séquences vidéo utilisées dans ce test, le nombre de panneaux visible à la fois sur une capture ne dépasse pas quatre panneaux par frame.

En faisant un calcul de temps de traitement par capture on trouve que le temps de traitement globale qui est le temps maximal mis dans le traitement (Détection + Reconnaissance + Décision) est inférieur à 0,5 second.

Tableau 5: Nombre de fois de traitement d'un panneaux routiers par le TSDS en fonction de la vitesse d'approche

Vitesse d'approche		Cas normal			Cas de minimum		
En Km/h	En m/s	Distance de visibilité en m	Durée de visibilité en s	Nombre de fois de traitement du panneau	Distance de visibilité en m	Durée de visibilité en s	Nombre de fois de traitement du panneau
120	33,33	360	10,8	21,6	200	6	12
110	30,56	300	9,82	19,64	180	5,89	11,78
100	27,78	250	9	18	160	5,76	11,52
80	22,22	160	7,2	14,4	120	5,4	10,8
60	16,67	90	5,4	10,8	80	4,8	9,6

D'après le Tableau 5, le nombre de fois de traitement d'un panneau dépasse 21 fois dans le cas d'une vitesse d'approche de 120km/h, cela permet d'éviter les ratés de détection. Le minimum de fois est réalisé dans le cas de visibilité minimale à une vitesse d'approche de 60km/h, et qui dépasse 9 fois. Ce qui est largement suffisant une bonne détection et reconnaissance de panneaux routiers en temps réel.

V.4.2. Test sur Raspberry Pi B+

Les temps de traitement mis par les algorithmes de test sur Raspberry Pi B+, ne s'élèvent pas à la hauteur de ceux enregistrés par les tests sur le PC.

Le temps de traitement global dépasse les 4 second, ce qui donne 2 traitements par panneau dans le cas d'une vitesse d'approche de 120km/h dans les conditions de visibilité normale. Et un seul traitement par panneau dans le cas d'une vitesse d'approche de 60km/h dans les conditions visibilité minimum.

Ce temps global de traitement peut être accepté pour une détection et reconnaissance en réel mais avec un risque de ratés de détection de panneau.

Pour améliorer ce résultat, nous avons arrêté les processus et les services intitulé qui tourne sous Raspberry. Et nous aussi éliminer tous ce qui est affichage graphique dans les

CHAPITRES V. Système complet TSDS

algorithmes du système TSDS, pour en arrêter l'affichage graphique qui bouffe les ressources en mémoire et en microprocesseur.

L'affichage est réalisé par un afficheur LCD, et un adapter SPI qui fait l'interfaçage entre Raspberry Pi B+ et le LCD. Au lieu d'afficher des templates de panneaux routiers sur l'écran, nous n'affichons que sa catégorie et son numéro sur l'afficheur LCD.

Le temps de traitement global s'est un peu amélioré, et il s'est baissé à 2,3 seconde, ce qui donne un peu près de 5 traitements par panneau dans le cas d'une vitesse d'approche de 120km/h dans les conditions de visibilité normale. Et 2 traitements par panneau dans le cas d'une vitesse d'approche de 60km/h dans les conditions visibilité minimum.

Cela peut réaliser une détection et une reconnaissance en temps réel avec moins de risque de ratés de détection.

Avec une carte électronique plus évoluée, de configurations plus sophistiquées, on atteindre ou dépasser même, les résultats obtenus par l'ordinateur I3, chose qui peut rendre de notre système complet TSDS, une bonne fonctionnalité de confort d'automobile, et bon système d'aide à la conduite.

Conclusions et perspectives

La sécurité routière est un problème de santé publique très préoccupant vu les conséquences des accidents de la route. Outre les préjudices physiques, les accidents entraînent un coût économique important aux victimes et à la société. Ces dernières décennies la sécurité routière a eu un engouement particulier auprès des constructeurs automobiles pour le développement des systèmes avancés d'aide à la conduite communément appelés "ADAS" (*Advanced Driver Assistance System*). L'objectif est de garantir la sécurité des passagers du véhicules et des usagers de la route. Dans ce contexte, les ADAS assistent le conducteur durant son trajet en détectant les situations à risque afin d'éviter l'accident. Le système de détection et reconnaissance de panneaux routiers (*TSDR-Traffic Sign Detection and Recognition*) est l'un de ces ADAS qui tient informé le conducteur de la signalisation liée au trajet parcouru. Durant le processus de conduite le conducteur peut avoir des moments d'inattention ce qui peut présenter un danger. En effet, manquer un panneau de limitation de vitesse, à cause d'un appel téléphonique juste avant de négocier un virage, engendre une situation à risque qu'il ne faut pas négliger. Ainsi, avec un système TSDR le conducteur est informé en temps-réel de la signalisation via des messages visuels et/ou sonores ce qui attire son attention sur son environnement routier. Les systèmes TSDR sont basés sur un capteur optique. La scène routière est perçue via une caméra embarquée, sur l'avant du véhicule, dont les images seront interprétées par le système en reconnaissant la signalisation présente. Pour cela, deux principales étapes sont nécessaires : *Détection* et *Reconnaissance* (appelée aussi classification). La première étape détecte la présence des panneaux dans les images. Ce processus consiste à déterminer les régions d'intérêt (*ROI - Region Of Interest*) susceptibles de contenir un panneau et cela en se basant sur leurs caractéristiques visuelles comme la couleur et la forme. Quant à la deuxième étape, elle reconnaît les panneaux détectés en identifiant leur descripteur.

Dans cette thèse nous avons étudié la détection des panneaux de signalisation routière. Cette partie consiste à effectuer un filtrage et d'une transformation de l'espace de couleurs des images de la scène routière acquises pour éliminer les effets des variations de luminosité, puis appliquer une segmentation de couleurs et une extraction des contours afin d'exploiter les propriétés de couleur et de forme des panneaux routiers, ces contours sont affinés par une approximation polygonale des contours numériques, pour identifier les objets d'intérêt pouvant inclure des fausses panneaux routiers mais si possible sans oublier. La

Conclusion et perspectives

vérification expérimentale de ce module de détection proposé a montré qu'il peut détecter 99% de formes présents dans la scène routière ayant la même géométrie et la même couleur que les panneaux.

Nous nous sommes intéressés ensuite à la reconnaissance des panneaux de signalisation routière, cette étape qui vient pour affiner les résultats de la détection, et identifier les panneaux, elle est réalisée par l'application de deux méthodes de classification, « Multi Layer Perceptron » MLP et « Support Vector Machine » SVM, ces deux classificateurs sont combinés avec deux descripteurs d'image, « Histogram Projection » HP et « Histogram Oriented Gradient » HOG. Pour mieux accélérer le processus de reconnaissance, nous avons utilisé la méthode de Multithreading qui consiste à exécuter en parallèle et en même temps plusieurs programmes, en effet nous avons divisé le classificateur en quatre sous-programmes selon la couleur et la géométrie du panneau (rouge-triangle, rouge-cercle, bleu-triangle et bleu-rectangle) qui s'exécutent en parallèle. Pour faire un choix judicieux de la combinaison descripteur-classificateur, nous avons fait une comparaison expérimentale des performances des quatre combinaisons, HP-MLP, HOG-MLP, HP-SVM et HOG-SVM. L'expérience a montré que la méthode HP-SVM est la plus performante en terme de temps d'exécution qui ne dépasse pas 360 ms et de taux de reconnaissance qui dépasse 95%.

En fin, un système complet a été développé mettant en œuvre le module de détection et le module de reconnaissance. Une vérification expérimentale de ces performances a été menée sur des images de la scène routière en mode (offline), les résultats obtenus prouvent la fiabilité du système comme étant un système d'aide à la conduite, avec un temps de global réflexion inférieure à 500 ms. Ces résultats sont confirmés par des tests en temps réel (mode online) en embarquant un ordinateur avec un caméscope à bord d'une voiture, et aussi en utilisant une carte embarquée Raspberry Pi B+.

De nombreuses améliorations sont toutefois possibles, Ajouter la détection et la reconnaissance des panneaux, des Textes et des Pictogrammes, à condition d'avoir une caméra ayant une résolution suffisante, et sans affecter le temps de calcul afin d'assurer un fonctionnement en temps réel. Une fois correctement détectées, les informations sémantiques présentes sur les panneaux et les pictogrammes pourraient permettre de savoir la durée pendant laquelle la vitesse s'applique ou si la catégorie représentée concerne le véhicule.

Publications

1. Salhi, Abderrahim; Minaoui, Brahim; Fakir, Mohamed; Robust Automatic Traffic Signs Detection using fast polygonal approximation of digital curves, » Multimedia Computing and Systems (ICMCS), 2014 International Conference on",433-437,2014, IEEE
2. SALHI, A; MINAOUI, B; FAKIR, M; Robust Automatic Traffic Signs Recognition Using Fast Polygonal Approximation of Digital Curves and Neural Network, International Journal of Advanced Computer Science and Applications(IJACSA),2014,
3. Salhi, Abderrahim; Minaoui, Brahim; Fakir, Mohamed; Sajieddine, Mohammed; Real-Time Detection of Road Signs, Journal of Electronic Commerce in Organizations (JECO),13,3,36-46,2015, IGI Global
4. Chakib, Houda; Minaoui, Brahim; Fakir, Mohamed; Salhi, Abderrahim; Badi, Imad; A Proposed Approach for Image Compression based on Wavelet Transform and Neural Network, International Journal of Advanced Computer Science and Applications(IJACSA),8,9,2017
5. A. Salhi, B. Minaoui, M. Fakir, H. Chakib and H. Grimech, "Traffic Signs Recognition using HP and HOG Descriptors Combined to MLP and SVM Classifiers" International Journal of Advanced Computer Science and Applications(IJACSA), 8(11), 2017. <http://dx.doi.org/10.14569/IJACSA.2017.081165>

Communications

- The first international conference on business intelligence (CBI'14), April 29-30, 2014 in Beni Mellal, Morocco.
- The international conference on business intelligence (CBI'15), April 23-25, 2015 in Beni Mellal, Morocco.
- The international conference on multimedia computing and systems (ICMCS'14), April 14-16, 2014 in Marrakech, Morocco.
- 1ère Edition des Journées Doctorales sous thème “Recherche Scientifique Innovation et Développement”, organisées du 10 au 11 juin 2014 à la Faculté des Sciences et Techniques de Béni Mellal, Maroc.

Bibliographie

- [1] Mahmoud M. Zadeh, T. Kasvand, and Ching Y. Suen. Localization and recognition of traffic signs for automated vehicle control systems. In Marten J. de Vries, Pushkin Kachroo, Kaan Ozbay, and Alan C. Chachich, editors, *SPIE Intelligent Transportation Systems*, volume 3207, pages 272–282. SPIE Publishing, January 1998.
- [2] M.V. Shirvaikar. Automatic detection and interpretation of road signs. In *Proceedings of the Thirty-Sixth Southeastern Symposium on System Theory*, pages 413–416, 2004.
- [3] DATASET, THE GERMAN TRAFFIC SIGN DETECTION BENCHMARK, <http://benchmark.ini.rub.de/?section=gtsdb&subsection=dataset> 2013
- [4] C. Y. Fang, C. S. Fuh, P. S. Yen, S. Cherng, and S. W. Chen. An automatic road sign recognition system based on a computational model of human recognition processing. *Computer Vision and Image Understanding*, 96(2) :237–268, November 2004.
- [5] Michael Shneier. Road sign detection and recognition. Submitted to the IEEE Computer Society International Conference on Computer Vision and Pattern Recognition, June 2005
- [6] Salvatore Vitabile, Giorgio Pollaccia, Giovanni Pilato, and Filippo Sorbello. Road signs recognition using a dynamic pixel aggregation technique in the HSV color space. In *ICIAP*, pages 572–577. IEEE Computer Society, 2001.
- [7] J. Miura, T. Kanda, and Y. Shirai. An active vision system for real-time traffic sign recognition. In *IEEE Intelligent Transportation Systems*, pages 52–57, October 2000.
- [8] C. Bahlmann, Ying Zhu, V. Ramesh, M. Pellkofer, and T. Koehler. A system for traffic sign detection, tracking, and recognition using color, shape, and motion information. In *IEEE Intelligent Vehicles Symposium*, pages 255–260, June 2005
- [9] L. Vlacic, M. Parent, and F. Harashima. *Intelligent Vehicle Technologies*. Butterworth-Heinemann, 2001.
- [10] Yu-Ming Liang, Hsiao-Rong Tyan, Shyang-Lih Chang, H.-Y.M. Liao, and Sei-Wang Chen. Video stabilization for a camcorder mounted on a moving vehicle. *IEEE Transactions on Vehicular Technology*, 53(6) :1636–1648, November 2004.
- [11] A. de la Escalera and L. Moreno. Road traffic sign detection and classification. *IEEE Transactions on Industrial Electronics*, 44(6) :848–859, December 1997.
- [12] Hassan Shojania. Real-time traffic sign detection, 2002. Student project for Computer Vision course (ELEC824) in the Department of Electrical and Computer Engineering at the Queen’s University.
- [13] J. Torresen, J.W. Bakke, and L. Sekanina. Efficient recognition of speed limit signs. In *The 7th International IEEE Conference on Intelligent Transportation Systems*, pages 652–656, October 2004.
- [14] M. Bénallal and J. Meunier. Real-time color segmentation of road signs. In *Canadian Conference on Electrical and Computer Engineering*, volume 3, pages 1823–1826, May 2003.

Bibliographie

- [15] Hsiu-Ming Yang, Chao-Lin Liu, Kun-Hao Liu, and Shang-Ming Huang. Traffic sign recognition in disturbing environments. In Ning Zhong, Zbigniew W. Ras, Shusaku Tsumoto, and Einoshin Suzuki, editors, Foundations of Intelligent Systems, 14th International Symposium, ISMIS 2003, Maebashi City, Japan, October 28-31, 2003, Proceedings, volume 2871 of Lecture Notes in Computer Science, pages 252–261. Springer, 2003.
- [16] W.G. Shadeed, D.I. Abu-Al-Nadi, and M.J. Mismar. Road traffic sign detection in color images. In Proceedings of the 2003 10th IEEE International Conference on Electronics, Circuits and Systems, volume 2, pages 890–893, December 2003
- [17] D. M. Gavrilă. Multi-feature hierarchical template matching using distance transforms. In International Conference on Pattern Recognition, pages Vol I : 439–444, 1998
- [18] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. In Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, volume 1, pages 511–518, 2001.
- [19] Xavier Baró and Jordi Vitrià. Fast traffic sign detection on greyscale images. In Catalan Conference on Artificial Intelligence, 2004.
- [20] Gareth Loy and Alexander Zelinsky. Fast radial symmetry for detecting points of interest. In IEEE Transactions on Pattern Analysis and Machine Intelligence, volume 25, August 2003.
- [21] G. A. Carpenter and S. Grossberg. A massively parallel architecture for a self-organizing neural pattern recognition machine. *Computer Vision*, 37 :54–115, 1983.
- [22] Equbits - accurate, interpretable, automated support vector machine. <http://www.equbits.com>.
- [23] Raphaël Marée. Classification automatique d'images par arbres de décision. PhD thesis, Department of Electrical Engineering and Computer Science, University of Liège, Belgium, February 2005. <http://www.montefiore.ulg.ac.be/services/stochastic/pubs/2005/Mar05/>.
- [24] Pierre Geurts, Damien Ernst, and Louis Wehenkel. Extremely randomized trees. *Machine Learning Journal*, 2006.
- [25] Raphaël Marée, Pierre Geurts, Justus Piater, and Louis Wehenkel. Decision trees and random subwindows for object recognition. In ICML workshop on Machine Learning Techniques for Processing Multimedia Content, 2005.
- [26] RASPBERRY PI FOUNDATION site web, Raspberry <https://www.raspberrypi.org/>
- [27] Canny, J., A Computational Approach To Edge Detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8:679-698 (1986).
- [28] Duda, R. O. et P. E. Hart, « Use of the Hough Transformation to Detect Lines and Curves in Pictures », *Comm. ACM*, Vol. 15, pp. 11 - 15 (janvier 1972).
- [29] TARSHA-KURDI, F., LANDES, T. and GRUSSENMEYER, P., 2007. « Hough-Transform and Extended RANSAC Algorithms for Automatic Detection of 3D Building

Bibliographie

- Roof Planes from Lidar Data », *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36(3/W52): 407-412.
- [30] Barnes N., Zelinsky A., Fletcher L., « Traffic sign recognition and analysis for intelligent vehicles », *Image and Vision Computing*, vol. 21, n° 3, p. 247-258, 2003.
- [31] Baro X., Escalera S., Vitria J., Pujol O., P.Radeva, « Traffic Sign Recognition Using Evolutionary Adaboost Detection and Forest-ECOC Classification », *IEEE Transactions on Intelligent Transportation Systems*, vol. 10, p. 113-126, 2009.
- [32] Nick Barnes, Gareth Loy, David Shaw, The regular polygon detector, In *Pattern Recognition*, Volume 43, Issue 3, 2010, Pages 592-602, ISSN 0031-3203, <https://doi.org/10.1016/j.patcog.2009.09.008>.(<http://www.sciencedirect.com/science/article/pii/S003132030900346X>)
- [33] Rachid Belaroussi and Jean-Philippe Tarel. Modèle géométrique du sommet et de la bissectrice d'un angle pour la détection de panneaux triangulaires. In à paraître dans les Actes du 17ème congrès francophone de Reconnaissance des Formes et Intelligence Artificielle (RFIA), Caen, France, Jan 2010. AFRIFAFIA.
- [34] A. SALHI, B. MINAOUI & M. FAKIR (2015). Real-Time Detection of Road Signs. *Journal of Electronic Commerce in Organisations (JECO)*, IGI, pages 36-46.
- [35] A. SALHI, B. MINAOUI & M. FAKIR (2014). Robust Automatic Traffic Signs Recognition Using Fast Polygonal Approximation of Digital Curves and Neural Network. *Int. Journal of Adv Comp Sc and Applications(IJACSA)*, Special Issue on Advances in Vehicular Ad hoc Networking and Applications, IEEE, pages 433-437.
- [36] A. SALHI, B. MINAOUI & M. FAKIR (2014). Robust Automatic Traffic Signs Detection using fast polygonal approximation of digital curves. *Multimedia Computing and Systems (ICMCS)*, 2014 International Conference on, IEEE, pages 14-16.
- [37] A. Lorsakul and J. Suthakorn, "Traffic Sign Recognition for Intelligent Vehicle/Driver Assistance System Using Neural Network on OpenCV", *The 4th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI 2007)*
- [38] F.A. Aly and A.E. Alaa. Detection, categorization and recognition of road signs for autonomous navigation. In *Proceeding of Advanced Concepts for Intelligent Vision System*, Brussels, Belgium, Aug 2004.
- [39] H. X. Liu, and B. Ran, "Vision-Based Stop Sign Detection and Recognition System for Intelligent Vehicle", *Transportation Research Board (TRB) Annual Meeting 2001*, Washington, D.C., USA, January 11, 2001.
- [40] H. Fleyeh, and M. Dougherty, "Road And Traffic Sign Detection And Recognition", *Proceedings of the 16th Mini - EURO Conference and 10th Meeting of EWGT*, pp. 644-653.
- [41] Y. Kameda et M. Minoh. A human motion estimation method using 3- successive video frames. In *International Conference on Virtual Systems and Multimedia*, pages 135_140, 1996.
- [42] S. Karaman, J. Benois-Pineau, R. Mégret, V. Dovgalecs, J.F. Dartigues et Y. Gaëstel. Human daily activities indexing in videos from wearable cameras for monitoring of

Bibliographie

- patients with dementia diseases. In Pattern Recognition (ICPR), 2010 20th International Conference on, pages 4113_4116. IEEE, 2010.
- [43] S. Karaman, J. Benois-Pineau, V. Dovgalecs, R. Mégret, J. Piquier, R. André-Obrecht, Y. Gaëstel et J.F. Dartigues. Hierarchical Hidden Markov Model in Detecting Activities of Daily Living in Wearable Videos for Studies of Dementia. Arxiv preprint arXiv :1111.1817, 2011.
- [44] C. Kas et H. Nicolas. Rough compressed domain camera pose estimation through object motion. In Image Processing (ICIP), 2009 16th IEEE International Conference on, pages 3481_3484. IEEE, 2009.
- [45] V. Kastinaki, M. Zervakis et K. Kalaitzakis. A survey of video processing techniques for tra_c applications. Image and Vision Computing, vol. 21, no. 4, pages 359_381, 2003.
- [46] K. Kim, T.H. Chalidabhongse, D. Harwood et L. Davis. Background modeling and subtraction by codebook construction. In Image Processing, 2004. ICIP'04. 2004 International Conference on, volume 5, pages 3061_3064. IEEE, 2004.
- [47] B. Krausz et R. Herpers. MetroSurv : detecting events in subway stations. Multimedia Tools and Applications, vol. 50, no. 1, pages 123_147, 2010.
- [48] T. Krishnan et GJ McLachlan. The EM algorithm and extensions, 1997.
- [49] H.W. Kuhn. The Hungarian method for the assignment problem. Naval research logistics quarterly, vol. 2, no. 1-2, pages 83_97, 1955.
- [50] G. Lavee, E. Rivlin et M. Rudzsky. Understanding video events : a survey of methods for automatic interpretation of semantic occurrences in video. Systems, Man, and Cybernetics, Part C : Applications and Reviews, IEEE Transactions on, vol. 39, no. 5, pages 489_504, 2009.
- [51] F. Legland. Filtrage particulière. In Proceedings 19eme Colloque GRETSI sur le Traitement du Signal et des Images, volume 1, pages 1_8, 2003.
- [52] F. Lei et X. Zhao. Adaptive background estimation of underwater using Kalman-Filtering. In Image and Signal Processing (CISP), 2010 3rd International Congress on, volume 1, pages 64_67. IEEE, 2010.
- [53] Y. Li, L.Q. Xu, J. Morphet et R. Jacobs. An integrated algorithm of incremental and robust pca. In Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on, volume 1, pages I_245. IEEE, 2003.
- [54] Y. Li, F. Chen, W. Xu et Y. Du. Gaussian-Based Codebook Model for Video Background Subtraction. Lecture notes in computer science, 2006.
- [55] J. Lou, Q. Liu, T. Tan et W. Hu. Semantic interpretation of object activities in a surveillance system. In Pattern Recognition, 2002. Proceedings. 16th International Conference on, volume 3, pages 777_780. IEEE, 2002.
- [56] D.G. Lowe. Object recognition from local scale-invariant features. In Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on, volume 2, pages 1150_1157. Ieee, 1999.

Bibliographie

- [57] D.G. Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, vol. 60, no. 2, pages 91_110, 2004.
- [58] W. Lu et Y.P. Tan. A color histogram based people tracking system. In *Circuits and Systems, 2001. ISCAS 2001. The 2001 IEEE International Symposium on*, volume 2, pages 137_140. IEEE, 2001.
- [59] B. Luvison, T. Chateau, J.T. Lapreste, P. Sayd et Q.C. Pham. *Automatic Detection of Unexpected Events in Dense Areas for Videosurveillance Applications*. 2012.
- [60] J. Makhoul. Linear prediction : A tutorial review. *Proceedings of the IEEE*, vol. 63, no. 4, pages 561_580, 1975.
- [61] D. Makris et T. Ellis. Learning semantic scene models from observing activity in visual surveillance. *Systems, Man, and Cybernetics, Part B : Cybernetics, IEEE Transactions on*, vol. 35, no. 3, pages 397_408, 2005.
- [62] Y. Malinovskiy, Y.Wang et Y.J.Wu. Video-based vehicle detection and tracking using spatio-temporal maps. *Proceedings of 88th Annual Transportation Research Board Meeting, 2009. DVD-ROM, Washington, D.C.*
- [63] K.V. Mardia et P.E. Jupp. *Directional statistics*. John Wiley & Sons Inc, 2000.
- [64] A. Matessi et L. Lombardi. Vanishing point detection in the hough transform space. *Euro-Par99 Parallel Processing*, pages 987_994, 1999.
- [65] N.J.B. McFarlane et C.P. Schoeld. Segmentation and tracking of piglets in images. *Machine Vision and Applications*, vol. 8, no. 3, pages 187_193, 1995.
- [66] A.M. McIvor. Background subtraction techniques. *Proc. of Image and Vision Computing*, vol. 1, no. 3, pages 155_163, 2000.
- [67] S.J. McKenna, Y. Raja et S. Gong. Tracking colour objects using adaptive mixture models. *Image and vision computing*, vol. 17, no. 3-4, pages 225_231, 1999.
- [68] S.J. McKenna, S. Jabri, Z. Duric, A. Rosenfeld et H. Wechsler. Tracking groups of people. *Computer Vision and Image Understanding*, vol. 80, no. 1, pages 42_56, 2000.
- [69] G.J. McLachlan et T. Krishnan. *The em algorithm and extensions*, volume 382. LibreDigital, 2008.
- [70] R. Melli, A. Prati, R. Cucchiara, L. de Cock et NV Tracon. Predictive and probabilistic tracking to detect stopped vehicles. In *Application of Computer Vision, 2005. WACV/MOTIONS.05 Volume 1. Seventh IEEE Workshops on*, volume 1, pages 388_393, 2005.
- [71] D.A. Migliore, M. Matteucci et M. Naccari. A revaluation of frame difference in fast and robust motion detection. In *Proceedings of the 4th ACM international workshop on Video surveillance and sensor networks*, pages 215_218. ACM, 2006.
- [72] K. Mikolajczyk et C. Schmid. A performance evaluation of local descriptors. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 27, no. 10, pages 1615_1630, 2005.

Bibliographie

- [73] T.B. Moeslund, A. Hilton et V. Kruger. A survey of advances in visionbased human motion capture and analysis. *Computer vision and image understanding*, vol. 104, no. 2-3, pages 90_126, 2006.
- [74] G. Monteiro, M. Ribeiro, J. Marcos et J. Batista. Wrongway drivers detection based on optical flow. In *Image Processing, 2007. ICIP 2007. IEEE International Conference on*, volume 5, pages V_141. IEEE, 2007.
- [75] H.P. Moravec. Obstacle avoidance and navigation in the real world by a seeing robot rover. tech report CMURITR8003 Robotics Institute Carnegie Mellon University doctoral dissertation Stanford University, 1980.
- [76] J.C. Nascimento, M.A.T. Figueiredo et JS Marques. Segmentation and classification of human activities. In *Proceedings of International Workshop on Human Activity Recognition and Modelling*, 2005.
- [77] A.T. Nghiem, F. Bremond, M. Thonnat et V. Valentin. ETISEO, performance evaluation for video surveillance systems. In *Advanced Video and Signal Based Surveillance, 2007. AVSS 2007. IEEE Conference on*, pages 476_481. IEEE, 2007.
- [78] N.M. Oliver, B. Rosario et A.P. Pentland. A Bayesian computer vision system for modeling human interactions. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 22, no. 8, pages 831_843, 2000.
- [79] G. Paschos. Perceptually uniform color spaces for color texture analysis : an empirical evaluation. *Image Processing, IEEE Transactions on*, vol. 10, no. 6, pages 932_937, 2001.
- [80] P. Perez, C. Hue, J. Vermaak et M. Gangnet. Color-based probabilistic tracking. *Computer Vision-ECCV 2002*, pages 661_675, 2002.
- [81] M. Piccardi. Background subtraction techniques : a review. In *Systems, Man and Cybernetics, 2004 IEEE International Conference on*, volume 4, pages 3099_3104. IEEE, 2004.
- [82] F. Porikli. Detection of temporarily static regions by processing video at different frame rates. In *Advanced Video and Signal Based Surveillance, 2007 AVSS 2007. IEEE Conference on*, pages 236_241. IEEE, 2007.
- [83] P.W. Power et J.A. Schoonees. Understanding background mixture models for foreground segmentation. In *Proceedings Image and Vision Computing New Zealand*, volume 2002, 2002.
- [84] M. Pressigout et E. Marchand. Suivi temps-réel d'objet plan : approche hybride contour/texture. 2005.
- [85] M. Pucher, D. Schabus, P. Schallauer, Y. Lypetsky, F. Graf, H. Rainer, M. Stadtschnitzer, S. Sternig, J. Birchbauer, W. Schneider et al. Multimodal highway monitoring for robust incident detection. In *13th International IEEE Conference on Intelligent Transportation Systems*, pages 837_842, 2010.
- [86] GW Pulford. Taxonomy of multiple target tracking methods. In *Radar, Sonar and Navigation, IEE Proceedings-*, volume 152, pages 291_304. IET, 2005.

Bibliographie

- [87] D. Reid. An algorithm for tracking multiple targets. *Automatic Control, IEEE Transactions on*, vol. 24, no. 6, pages 843_854, 1979.
- [88] M.I. Ribeiro. Kalman and extended kalman filters : Concept, derivation and properties. *Institute for Systems and Robotics*, page 43, 2004.
- [89] P.C. Ribeiro et J. Santos-Victor. Human activity recognition from video : Modeling, feature selection and classification architecture. In *Proceedings of International Workshop on Human Activity Recognition and Modelling*. Citeseer, 2005.
- [90] Christof Ridder, Olaf Munkelt et Harald Kirchner. Adaptive Background Estimation and Foreground Detection using Kalman-Filtering. pages 193_199, 1995.
- [91] B. Rosenhahn, U. Kersting, S. Andrew, T. Brox, R. Klette et H.P. Seidel. A silhouette based human motion tracking system. *Rapport technique, CITR, The University of Auckland, New Zealand*, 2005.
- [92] A. Roy, S. Kumar Parui, A. Paul et U. Roy. A color based image segmentation and its application to text segmentation. In *Computer Vision, Graphics & Image Processing, 2008. ICVGIP'08. Sixth Indian Conference on*, pages 313_319. IEEE, 2008.
- [93] M. Seibert, B.J. Rhodes, N.A. Bomberger, P.O. Beane, J.J. Sroka, W. Kogel, W. Kreamer, C. Stauer, L. Kirschner, E. Chalomet al. SeeCoast port surveillance. In *Proceedings of SPIE*, volume 6204, page 62040B. Spie, 2006.
- [94] M. Shah. *Motion-based recognition*, volume 9. Springer, 1997.
- [95] M. Shah, J. Deng et B. Woodford. Enhanced Codebook Model for Real-Time Background Subtraction. In *Neural Information Processing*, pages 449_458. Springer, 2011.
- [96] J. Shi et C. Tomasi. Good features to track. In *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR'94., 1994 IEEE Computer Society Conference on*, pages 593_600. IEEE, 1994.
- [97] R. Sicre et H. Nicolas. Human behaviour analysis and event recognition at a point of sale. In *Image and Video Technology (PSIVT), 2010 Fourth Pacific-Rim Symposium on*, pages 127_132. IEEE, 2010.
- [98] C. Tomasi et T. Kanade. Detection and tracking of point features. *School of Computer Science, Carnegie Mellon Univ.*, 1991.
- [99] S. Torkan et A. Behrad. A new contour based tracking algorithm using improved greedy snake. In *Electrical Engineering (ICEE), 2010 18th Iranian Conference on*, pages 150_155. IEEE, 2010.
- [100] Y. Wang, E.K. Teoh et D. Shen. Lane detection and tracking using B-Snake. *Image and Vision computing*, vol. 22, no. 4, pages 269_280, 2004.
- [101] X. Wang et X. Wang. Efficient online appearance models for object tracking. In *Proceedings of the 10th WSEAS international conference on Multimedia systems & signal processing*, pages 217_221. World Scientific and Engineering Academy and Society (WSEAS), 2010.

Bibliographie

- [102] L.F. Wang et C.H. Pan. Effective multi-resolution background subtraction. In Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on, pages 909_912. IEEE, 2011.
- [103] S.C. Wang, T.F. Su et S.H. Lai. Detecting moving objects from dynamic background with shadow removal. In Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on, pages 925_928. IEEE, 2011.
- [104] X. Xie. A review of recent advances in surface defect detection using texture analysis techniques. ELCVIA : electronic letters on computer vision and image analysis, vol. 7, no. 3, pages 1_22, 2010.
- [105] Arturo de la Escalera, Luis E. Moreno, Miguel Angel Salichs, and José Maria Armingol. Road traffic sign detection and classification. IEEE Transactions on Industrial Electronics, 44(6) :848–859, Dec 1997.
- [106] S. Belongie, “Shape Matching and Object Recognition Using Shape Contexts”, IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL 24, NO 24, APRIL 2002.
- [107] Intel Corporation, “Open Source Computer Vision Library,” Reference Manual, Copyright © 1999-2001, Available: www.developer.intel.com
- [108] Urs Ramer, "An iterative procedure for the polygonal approximation of plane curves", Computer Graphics and Image Processing, 1(3), 244–256 (1972)
- [109] Laurence Fausett, “Fundamentals of Neural Networks Architectures, Algorithms, and Applications”, Prentice Hall Upper Saddle River, New Jersey 1994.
- [110] Y. Zeng, X. Xu, D. Shen & all (2016), Traffic Sign Recognition Using Kernel Extreme Learning Machines With Deep Perceptual Features, *IEEE Transactions on Intelligent Transportation Systems*, IEEE, pages 1-7.
- [111] M. Xing, M. Chunyang, W. Yan & all (2016), Traffic sign detection and recognition using color standardization and Zernike moments, *Control and Decision Conference (CCDC), Chinese*, IEEE, pages 28-30.
- [112] N. Dalal & B. Triggs (2005), Histograms of oriented gradients for human detection, *Computer Vision and Pattern Recognition. IEEE Computer Society Conference on*. IEEE, pages: 886-893.
- [113] C. Cortes & V. Vapnik (1995), Support-vector Networks. *Machine Learning*, springer link pages 273–297
- [114] Bahlmann, C., Pellkofer, M., Giebel, J. and Baratoff, G., 2008. Multi-Modal Speed Limit Assistants : Combining Camera and GPS Maps. In : IEEE Intelligent Vehicles Symposium, Eindhoven, Netherlands, pp. 132–137.
- [115] Bahlmann, C., Zhu, Y., Ramesh, V., Pellkofer, M. and Koehler, T., 2005. A System for Traffic Sign Detection, Tracking, and Recognition Using Color, Shape, and Motion Information. In : IEEE Intelligent Vehicles Symposium, Las Vegas, NV, USA, pp. 255–260.

Bibliographie

- [116] Bay, H., Ess, A., Tuytelaars, T. and Van Gool, L., 2008. Speeded-Up Robust Features (SURF). *Computer Vision and Image Understanding* 110(3), pp. 346–359.
- [117] Bloch, I. and Maitre, H., 1998. Fusion of Image Information under Imprecision. *Aggregation and Fusion of Imperfect Information* pp. 189–213.
- [118] Bosch, A., Zisserman, A. and Munoz, X., 2007b. Representing Shape with a Spatial Pyramid Kernel. In : *ACM International Conference on Image and Video Retrieval*, New York, NY USA, pp. 401–408.
- [119] Boser, B. E., Guyon, I. M. and Vapnik, V. N., 1992. A Training Algorithm for Optimal Margin Classifiers. In : *ACM Workshop on Computational Learning Theory*, New York, NY, USA, pp. 144–152.
- [120] Breiman, L., 2001. Random Forests. *Machine Learning* 45(1), pp. 5–32.
- [121] Bunke, H., 1990. *Syntactic and Structural Pattern Recognition : Theory and Applications*. World Scientific.
- [122] Chan, T. F. and Vese, L. A., 2001. Active Contours without Edges. *Transactions on Image Processing* 10(2), pp. 266–277.
- [123] Chanda, G. and Dellaert, F., 2004. *Grammatical Methods in Computer Vision : An Overview*. Technical report.
- [124] Chang, C.-C. and Lin, C.-J., 2011. LIBSVM : A Library for Support Vector Machines. *Transactions on Intelligent Systems and Technology* 2(3), pp. 1–27.
- [125] Daniel, J. and Lauffenburger, J.-P., 2011. Conflict Management in Multi-sensor Dempster-Shafer Fusion for Speed Limit Determination. In : *IEEE Intelligent Vehicles Symposium*, Baden-Baden, Germany, pp. 985–990.
- [126] Dasarathy, B. V., 1997. Sensor Fusion Potential Exploitation - Innovative Architectures and Illustrative Applications. *Proceedings of the IEEE* 85(1), pp. 24–38.
- [127] de La Escalera, A., Armingol, J. M. and Mata, M., 2003. Traffic Sign Recognition and Analysis for Intelligent Vehicles. *Image and Vision Computing* 21(3), pp. 247–258.
- [128] de la Escalera, A., Moreno, L. E., Salichs, M. A. and Armingol, J. M., 1997. Road Traffic Sign Detection and Classification. *Transactions on Industrial Electronics* 44(6), pp. 848–859.
- [129] Dempster, A. P., 1967. Upper and Lower Probabilities Induced by a Multivalued Mapping. *The Annals of Mathematical Statistics* 38(2), pp. 325–359.
- [130] Dougherty, E. R., 1992. *An Introduction to Morphological Image Processing*. Society of Photo Optical.
- [131] Dubois, D. and Prade, H., 1988. *Possibility Theory : An Approach to Computerized Processing of Uncertainty*. Plenum Press.
- [132] Ehrlich, J., Marchi, M., Jarri, P., Salesse, L., Guichon, D., Dominois, D. and Leverger, C., 2003. LAVIA , The French ISA Project : Main Issues and First Results on Technical Tests. In : *IEEE Intelligent Transport Systems*, pp. 1–11.

Bibliographie

- [133] Escalera, S. and Radeva, P., 2004. Fast Greyscale Road Sign Model Matching and Recognition. *Recent Advances in Artificial Intelligence Research and Development* pp. 69–76.
- [134] Felzenszwalb, P. F. and Huttenlocher, D. P., 2004. Efficient Graph-Based Image Segmentation. *International Journal of Computer Vision* 59(2), pp. 167–181.
- [135] Freund, Y. and Schapire, R. E., 1997. A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. *Journal of Computer and System Sciences* 55(1), pp. 119–139.
- [136] Garcia-Garrido, M. A., Sotelo, M. A. and Martin-Gorostiza, E., 2006. Fast Traffic Sign Detection and Recognition under Changing Lighting Conditions. In : *IEEE International Conference on Intelligent Transportation Systems*, Toronto, Ont., Canada, pp. 811–816.
- [137] Garza-Jinich, M., Meer, P. and Medina, V., 1999. Robust Retrieval of Three-Dimensional Structures from Image Stacks. *Medical Image Analysis* 3(1), pp. 21–35.
- [138] Gavrilu, D. M., 1999. Traffic Sign Recognition Revisited. In : *DAGM Symposium für Mustererkennung*, Bonn, Germany, pp. 86–93.
- [139] Gdalyahu, Y. and Weinshall, D., 1999. Flexible Syntactic Matching of Curves and its Application to Automatic Hierarchical Classification of Silhouettes. *Transactions on Pattern Analysis and Machine Intelligence* 21(12), pp. 1312–1328.
- [140] Hamdoun, O., Bargeton, A., Moutarde, F., Bradai, B. and Chanussot, L., 2008. Detection and Recognition of End-of-Speed-Limit and Supplementary Signs for Improved European Speed Limit Support. In : *World Congress on Intelligent Transport Systems*, New York, NY, USA, pp. 7–10.
- [141] Harris, C. and Stephens, M., 1988. A Combined Corner and Edge Detector. In : *Alvey Vision Conference*, Manchester, UK, pp. 147–152.
- [142] Herbin-Sahler, A., Chanussot, L. and Moutarde, F., 2007. Procédé de Détection d'un Objet Cible. HPLabs, 2005. Tesseract-OCR.
- [143] Hu, M.-K., 1962. Visual Pattern Recognition by Moment Invariants. *Transactions on Information Theory* 8(2), pp. 179–187.
- [144] Huang, H., Gu, M. and Chao, H., 2008. An Efficient Method of License Plate Location in Natural-Scene Image. In : *IEEE International Conference on Fuzzy Systems and Knowledge Discovery*, Jinan, Shandong, China, pp. 15–19.
- [145] Idrissa, M. and Acheroy, M., 2002. Texture Classification using Gabor Filters. *Pattern Recognition Letters* 23(9), pp. 1095–1102.
- [146] Jain, A. K., Duin, R. P. W. and Mao, J., 2000. Statistical Pattern Recognition : A Review. *Transactions on Pattern Analysis and Machine Intelligence* 22(1), pp. 4–37.
- [147] Jung, C. R. and Schramm, R., 2004. Rectangle Detection based on a Windowed Hough Transform. In : *IEEE Brazilian Symposium on Computer Graphics and Image Processing*, Curitiba, Brazil, pp. 113–120.

Bibliographie

- [148] Kang, D. S., Griswold, N. C. and Kehtarnavaz, N., 1994. An Invariant Traffic Sign Recognition System Based on Sequential Color Processing and Geometrical. In : IEEE Southwest Symposium on Image Analysis and Interpretation, pp. 88–93.
- [149] Keller, C. G., Sprunk, C., Bahlmann, C., Giebel, J. and Baratoff, G., 2008. Real-Time Recognition of U.S. Speed Signs. In : IEEE Intelligent Vehicles Symposium, Eindhoven, Netherlands, pp. 518–523.
- [150] Kokar, M. M. and Kim, K., 1994. Preface to the Special Section on Data Fusion : Architectures and Issues. *Control Engineering Practice* 2(5), pp. 803–809.
- [151] Kröse, B. and van der Smagt, P., 1996. *Introduction to Neural Networks*.
- [152] Lagunovsky, D. and Ablameyko, S., 1997. Fast Line and Rectangle Detection by Clustering and Grouping. In : International Conference on Computer Analysis of Images and Patterns, Kiel, Germany, pp. 503–510.
- [153] Lampert, C. H., Blaschko, M. B. and Hofmann, T., 2008. Beyond Sliding Windows : Object Localization by Efficient Subwindow Search. In : IEEE Conference on Computer Vision and Pattern Recognition, Anchorage, AK, USA, pp. 1–8.
- [154] Lauffenburger, J.-P., Bradai, B., Basset, M. and Nashashibi, F., 2008. Navigation and Speed Signs Recognition Fusion for Enhanced Vehicle Location. In : World Congress of the International Federation of Automatic Control, Seoul, Republic of Korea, pp. 2069–2074.
- [155] Li, S. and Shawe-Taylor, J., 2005. Comparison and Fusion of Multiresolution Features for Texture Classification. *Pattern Recognition Letters* 26(5), pp. 633–638.
- [156] Lin, W.-C. and Fu, K.-S., 1986. A Syntactic Approach to Three-Dimensional Object Recognition. *Transactions on Systems, Man, and Cybernetics* 16(3), pp. 405–422.
- [157] Liu, W., Lv, J., Gao, H., Duan, B., Yuan, H. and Zhao, H., 2011. An Efficient Real-Time Speed Limit Signs Recognition Based on Rotation Invariant Feature. In : IEEE Intelligent Vehicles Symposium, Baden-Baden, Germany, pp. 998–1003.
- [158] Lowe, D. G., 2004. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision* 60(2), pp. 91–110.
- [159] Loy, G. and Barnes, N., 2004. Fast Shape-based Road Sign Detection for a Driver Assistance System. In : IEEE/RSJ International Conference on Intelligent Robots and Systems, Sendai, Japan, pp. 70–75.
- [160] McCulloch, W. S. and Pitts, W., 1943. A Logical Calculus of the Ideas Immanent in Nervous Activity. *Bulletin of Mathematical Biophysics* 5(4), pp. 115–133.
- [161] Meyer, D., Leisch, F. and Hornik, K., 2003. The Support Vector Machine under Test. *Neurocomputing* 55(55), pp. 169–186.
- [162] Mikolajczyk, K. and Schmid, C., 2001. Indexing based on Scale Invariant Interest Points. In : IEEE International Conference on Computer Vision, Vancouver, BC, Canada, pp. 525–531.

Bibliographie

- [163] Miura, J., Kanda, T. and Shirai, Y., 2000. An Active Vision System for Real-Time Traffic Sign Recognition. In : IEEE International Conference on Intelligent Transportation Systems, Dearborn, MI, USA, pp. 52–57.
- [164] Müller, K.-R., Mika, S., Rätsch, G., Tsuda, K. and Schölkopf, B., 2001. An Introduction to Kernel-based Learning Algorithms. *Transactions on Neural Networks* 12(2), pp. 181–202.
- [165] Mumford, D. and Shah, J., 1985. Boundary Detection by Minimizing Functionals. In : IEEE Conference on Computer Vision and Pattern Recognition, San Fransisco, CA, USA, pp. 137–154.
- [166] Nienhüser, D., Gumpp, T., Zöllner, J. M. and Natroshvili, K., 2010. Fast and Reliable Recognition of Supplementary Traffic Signs. In : IEEE Intelligent Vehicles Symposium, San Diego, CA, USA, pp. 896–901.
- [167] Paclik, P., Novovicova, J. and Duin, R. P. W., 2006. Building Road Sign Classifiers using a Trainable Similarity Measure. *Transactions on Intelligent Transportation Systems* 7(3), pp. 309–321.
- [168] Piccioli, G., de Micheli, E., Parodi, P. and Campani, M., 1996. Robust Method for Road Sign Detection and Recognition. *Image and Vision Computing* 14(3), pp. 209–223.
- [169] Puthon, A.-S., Nashashibi, F. and Bradai, B., 2010. Improvement of Multisensor Fusion in Speed Limit Determination by Quantifying Navigation Reliability. In : IEEE International Conference on Intelligent Transportation Systems, Funchal, Portugal, pp. 855– 860.
- [170] Quinlan, J. R., 1986. Induction of Decision Trees. *Machine Learning* 1(1), pp. 81–106.
- [171] Revol, C. and Jourlin, M., 1997. A New Minimum Variance Region Growing Algorithm for Image Segmentation. *Pattern Recognition Letters* 18(3), pp. 249–258.
- [172] Rombaut, M., 1998. Decision in Multi-Obstacle Matching Process using Dempster-Shafer's Theory. *Advances in Vehicle Control and Safety* pp. 63–68.
- [173] Rosenblatt, F., 1959. *Principles on Neurodynamics*. Spartan Books.
- [174] Roth, P. M. and Winter, M., 2008. Survey of Appearance-based Methods for Object Recognition. Technical report.
- [175] Rumelhart, D. E., Hinton, G. E. and Williams, R. J., 1986. Learning Representations by Back-Propagating Errors. *Nature* 323(6088), pp. 533–536.
- [176] Ruta, A., Li, Y. and Liu, X., 2010. Real-Time Traffic Sign Recognition from Video by Class-Specific Discriminative Features. *Pattern Recognition* 43(1), pp. 416–430.
- [177] Sekiguchi, H., Sano, K. and Yokoyama, T., 1994. Interactive 3-Dimensional Segmentation Method based on Region Growing Method. *Systems and Computers in Japan* 25(1), pp. 88–97.
- [178] Sentz, K. and Ferson, S., 2002. Combination of Evidence in Dempster- Shafer Theory. Technical report.

Bibliographie

- [179] Sethian, J. A., 1997. Level Set Methods : An Act of Violence.
- [180] Shafer, G., 1976. A Mathematical Theory of Evidence. Princeton University Press.
- [181] Steinberg, A. N. and Bowman, C. L., 2004. Rethinking the JDL Data Fusion Levels. In : MSS National Symposium on Sensor and Data Fusion, SENSIAC, Columbia, SC, USA.
- [182] Steinberg, A. N., Bowman, C. L. and White, F. E., 1998. Revisions to the JDL Data Fusion Model. In : Society of Photo-Optical Instrumentation Engineers – Sensor Fusion : Architectures, Algorithms, and Applications III, SPIE, Orlando, FL, USA, pp. 430–441.
- [183] Teague, M. R., 1980. Image Analysis via the General Theory of Moments. Journal of Optical Society of America 70(8), pp. 920–930.
- [184] Tuytelaars, T. and Mikolajczyk, K., 2007. Local Invariant Feature Detectors : A Survey. Foundations and Trends in Computer Graphics and Vision 3(3), pp. 177–280.
- [185] Vincent, L., 1993. Morphological Grayscale Reconstruction in Image Analysis : Applications and Efficient Algorithms. Transactions on Image Processing 2(2), pp. 176–201.
- [186] Viola, P. and Jones, M., 2001. Robust Real-Time Object Detection. In : IEEE Workshop on Statistical and Computational Theories of Vision - Modeling, Learning, Computing and Sampling, Vancouver, BC, Canada, pp. 137–154.
- [187] Vlassenroot, S., Broekx, S., De Mol, J., Int Panis, L., Brijs, T. and Wets, G., 2007. Driving with Intelligent Speed Adaptation : Final Results of the Belgian ISA-Trial. Transportation Research 41(3), pp. 267–279.
- [188] Wu, T.-F., Lin, C.-J. and Weng, R. C., 2004. Probability Estimates for Multiclass Classification by Pairwise Coupling. Machine Learning Research 5, pp. 975–1005.
- [189] Zaklouta, F., Stanculescu, B. and Hamdoun, O., 2011. Traffic Sign Classification using K-d Trees and Random Forests. In : IEEE International Joint Conference on Neural Networks, San Jose, CA, USA, pp. 2151–2155.
- [190] Zhang, Z., Chen, C., Sun, J. and Luk Chan, K., 2003. EM Algorithms for Gaussian Mixtures with Split-and-Merge Operation. Pattern Recognition 36(9), pp. 1973–1983.
- [191] Zucker, S. W., 1976. Region Growing : Childhood and Adolescence. Computer Graphics and Image Processing 5(3), pp. 382–399.
- [192] M. Benallal and J. Meunier, “Real-time color segmentation of road signs,” Electrical and Computer Engineering, 2003. IEEE CCECE 2003. Canadian Conference on, vol. 3, pp. 1823–1826 vol.3, May 2003.
- [193] L. Estevez and N. Kehtarnavaz, “A real-time histogrammic approach to road sign recognition,” Image Analysis and Interpretation, 1996., Proceedings of the IEEE Southwest Symposium on, pp. 95–100, Apr 1996.
- [194] S. Varun, S. Singh, R. S. Kunte, R. D. S. Samuel, and B. Philip, “A road traffic signal recognition system based on template matching employing tree classifier,” in ICCIMA '07: Proceedings of the International Conference on Computational Intelligence and

Bibliographie

- Multimedia Applications (ICCIMA 2007). Washington, DC, USA: IEEE Computer Society, 2007, pp. 360–365.
- [195] A. Broggi, P. Cerri, P. Medici, P. Porta, and G. Ghisio, “Real time road signs recognition,” Intelligent Vehicles Symposium, 2007 IEEE, pp. 981–986, June 2007.
- [196] A. Ruta, Y. Li, and X. Liu, “Detection, tracking and recognition of traffic signs from video input,” Oct. 2008, pp. 55–60.
- [197] W.-J. Kuo and C.-C. Lin, “Two-stage road sign detection and recognition,” Multimedia and Expo, 2007 IEEE International Conference on, pp. 1427–1430, July 2007.
- [198] G. Piccioli, E. D. Micheli, P. Parodi, and M. Campani, “Robust method for road sign detection and recognition,” Image and Vision Computing, vol. 14, no. 3, pp. 209–223, 1996. [Online]. Available: <http://www.sciencedirect.com/science/article/B6V09-3VVCMCX-4/2/0f2793e7828195ecb68735a80a9ef904>
- [199] P. Paclík, J. Novovicova, P. Pudil, and P. Somol, “Road sign classification using laplace kernel classifier,” Pattern Recogn. Lett., vol. 21, no. 13-14, pp. 1165–1173, 2000.
- [200] C.-Y. Fang, S.-W. Chen, and C.-S. Fuh, “Road-sign detection and tracking,” vol. 52, no. 5, pp. 1329–1341, Sep. 2003.