



UNIVERSITE SULTAN MOULAY SLIMANE

Faculté des Sciences et Techniques

Béni-Mellal



N° d'ordre : 80 / 2015

Centre d'Etudes Doctorales « Sciences et Techniques »  
Formation Doctorale « Mathématiques et Physique Appliquées »

## THESE

Présentée par

**BRAHIM MINAOUI**

Pour obtention du grade de

**Docteur**

**Discipline** : Electronique

**Spécialité** : Traitement du signal

# Analyse Sémantique d'Images

Soutenue le 20 novembre 2015 devant la commission d'examen:

<b>Pr. Mohammed SAHLAOUI</b> , Professeur à la Faculté des Sciences et Techniques, Béni-Mellal	<b>Président</b>
<b>Pr. Said MELLIANI</b> , Professeur à la Faculté des Sciences et Techniques, Béni-Mellal	<b>Rapporteur</b>
<b>Pr. Abdelouahed ABOUNADA</b> , Professeur à la Faculté des Sciences et Techniques, Béni-Mellal	<b>Rapporteur</b>
<b>Pr. El Hocine AGOURIANE</b> , Professeur à la Faculté des Sciences et Techniques, Béni-Mellal	<b>Examineur</b>
<b>Pr. Mohammed SAJIEDDINE</b> , Professeur à la Faculté des Sciences et Techniques, Béni-Mellal	<b>Directeur de thèse</b>
<b>Pr. Mohamed FAKIR</b> , Professeur à la Faculté des Sciences et Techniques, Béni-Mellal	<b>Co-directeur de thèse</b>

## Remerciements

Je tiens tout d'abord à remercier les professeurs Mohammed Sajieddine et Mohamed Fakir d'avoir accepté de diriger mon travail de thèse et de contribuer à l'aboutissement de ce mémoire.

J'adresse également mes remerciements à Messieurs les professeurs Said Melliani, Benaceur Outtaj et Abdelouahed Abounada qui m'ont fait l'honneur d'être rapporteurs de cette thèse.

Mes sincères remerciements vont aussi aux membres de jury, les professeurs Mohammed Sahlaoui et El Hocine Agouriane pour l'intérêt qu'ils ont porté à mon travail et d'avoir mobilisé leurs temps pour siéger au jury.

Enfin, je ne peux manquer de remercier toutes les personnes qui ont pris part à ce travail aussi bien sur le plan scientifique qu'humain.

## RESUME

L'objectif ultime de la recherche en vision par ordinateur est de créer un système qui puisse générer une description de l'image de la même qualité que celle d'un être humain. Un tel système doit être capable de franchir le fossé sémantique qui est un problème complexe dû à l'écart sémantique existant entre les caractéristiques visuels bas-niveau et les concepts sémantiques haut-niveau. Notre travail s'intègre dans les travaux de recherche qui ont naturellement émergé pour résoudre ce problème.

Dans ce travail, nous nous sommes intéressés à l'analyse sémantique d'images pour l'annotation d'images et pour la reconnaissance de formes qui sont les thèmes de la vision artificielle les plus impliqués dans le problème du fossé sémantique.

Le travail sur l'analyse sémantique d'images pour l'annotation, nous a permis de réaliser un système d'annotation automatique d'images intégrant un processus d'analyse sémantique que nous avons élaboré pour réduire le fossé sémantique en apportant un soin particulier à l'amélioration de :

- La segmentation par croissance des régions en regroupant les régions adjacentes pour avoir des objets sémantiques ;
- La représentation des caractéristiques visuelles en combinant différents types de descripteurs ;
- La classification, en combinant le pouvoir discriminatif, des réseaux de neurones avec le caractère génératif des réseaux bayésiens naïfs pour tirer bénéfice de leur complémentarité.

Une évaluation du système d'annotation proposé, réalisée sur un corpus d'images différent de celui utilisé au cours de l'apprentissage, a montré que notre système permet d'obtenir des résultats probants.

En ce qui concerne la reconnaissance de formes, nous nous sommes intéressés à l'analyse sémantique d'images pour l'identification des panneaux de signalisation routière et des maladies des agrumes.

Dans le cas de l'identification des panneaux de la signalisation routière, nous avons développé une chaîne d'analyse sémantique qui consiste à détecter et reconnaître automatiquement les panneaux routiers. La détection consiste à segmenter en couleur les images pour localiser et isoler séparément les zones bleues et les zones rouges susceptibles de

correspondre à des panneaux routiers ; puis d'extraire les contours de ces éventuels panneaux pour déterminer, à l'aide en particulier de l'approximation Polygonale, leurs formes et leurs classes d'appartenance : panneaux triangulaires rouges, circulaires rouges, circulaires bleus ou rectangulaires bleus. Quant à la reconnaissance, elle consiste à identifier les panneaux détectés à l'aide des réseaux de neurones entraînés à mettre en correspondance, pour chaque classe, les moyennes normalisées des valeurs des pixels de chaque ligne et de chaque colonne des zones détectées avec des modèles de panneaux routiers appris à priori. Des tests effectués sur un ensemble d'images réelles du trafic routier, ont mis en évidence les performances prometteuses du système actuellement développé.

Concernant l'identification des maladies des agrumes, nous avons élaboré une démarche expérimentale pour la détection et la reconnaissance automatique des maladies dont les symptômes apparaissent sur les feuilles. Le principe de cette démarche est basé sur l'analyse statistique de la texture des images des feuilles par les matrices de cooccurrence couplée à une classification des différentes maladies par un réseau de neurones multicouches. La mise en œuvre de la démarche proposée, en utilisant des images de feuilles malades identifiées par des experts, a mis en évidence que l'analyse de la texture des feuilles permet une interprétation sémantique intéressante des maladies des agrumes.

Nous avons donc développé, dans ce travail de thèse, des procédés d'analyse sémantique d'images intégrant, selon l'application envisagée, des signatures à base de descripteurs de détails sémantiques, ayant un meilleur pouvoir informatif et permettant d'améliorer de manière conséquente les performances de la classification.

**Titre** : Analyse sémantique d'images

**Mots clés** : *Image, analyse, sémantique, segmentation, caractéristiques bas-niveau, concepts haut-niveau, détection, apprentissage, classification, interprétation, reconnaissance, annotation automatique, panneaux routiers, maladies des agrumes.*

## ABSTARCT

The ultimate goal of research in computer vision is to create a system that can generate a description of the image of the same quality as that of a human. Such a system should be able to bridge the semantic gap that is a complex problem due to the existing semantic gap between low-level visual features and semantic high level concepts. Our work is part the research that have naturally emerged to solve this problem.

In this work, we are interested to the semantic images analysis for images annotation and pattern recognition that are the artificial vision themes that are the most confronted to the problem of semantic gap.

The work on semantic images analysis for annotation, allowed us to realize a system of automatic image annotation integrating a semantic analysis process that we have developed to reduce the semantic gap by providing special care to improve the:

- region growing segmentation by grouping adjacent regions in order to have semantic objects,
- representation of visual characteristics by combining different types of descriptors,
- classification by combining the discriminative power of the neural networks and the generative nature of Bayesian networks in order to gain the benefit of their complementarity.

The evaluation of the proposed annotation system, performed on a corpus of images different from that used during learning, has shown some success in automatic image annotation.

For pattern recognition, we are interested to the images semantic analysis for identification of road panels and citrus diseases.

In the case of identifying of road signs, we have developed a analysis semantic chain able to automatically detect and recognize the traffic signs. The detection consists to segment color images in order to locate and isolate separately the blue areas and red areas which may correspond to road panels; then to extract the contours of these panels to determine , using in particular the Polygonal approximation, their forms and their classes: Red triangular panels, circular red, blue circular or blue rectangular. The recognition phase consists to identify signs detected using neural network trained to match, for each class, the average normalized pixel values of each row and of each column of the detected areas with models road signs learned a

priori. The tests on a set of real images of road traffic, have shown promising performance of the system being developed.

Concerning the identification of citrus diseases, we developed an experimental approach for the automatic detection and recognition of diseases whose symptoms appear on the leaves. The principle of this approach is based on statistical analysis of the images texture of the leaves by the co-occurrence matrices coupled to a classification of various diseases by a multilayer neural network. The evaluation of the proposed approach, using images of diseased leaves identified by experts, revealed that the analysis of the texture leaves allows an interesting semantic interpretation of citrus diseases.

In this thesis work, we have developed an image semantic analysis processes including, depending on the envisaged application, discriminative descriptors allowing to improve consequently the classification performance.

**Title:** Image Semantic Analysis

**Key-words:** *Image, Analysis, segmentation, features low-level, high-level concepts, semantic, detection, learning, classification, interpretation, automatic annotation, recognition, road panels, citrus diseases.*

## ملخص

ان الهدف النهائي للبحث في الرؤية بواسطة الحاسوب هو خلق نظام يمكن أن يولد شرحا للصورة بنفس الجودة التي تم القيام بها من قبل الإنسان. وينبغي أن يكون هذا النظام قادرا على اجتياز الفجوة الدلالية التي تعتبر مسألة جد معقدة بسبب الفارق الدلالي القائم بين الميزات المرئية المنخفضة المستوى والمفاهيم الدلالية العالية المستوى. عملنا هذا يندرج ضمن البحوث التي برزت بشكل طبيعي من اجل حل هذه المشكلة.

في هذا العمل، ركزنا على التحليل الدلالي للصور من اجل وصف الصور والتعرف على الأنماط داخلها وهي مواضيع الرؤية الاصطناعية الأكثر تداخلا مع مشكلة الفجوة الدلالية.

وأتاح لنا العمل على التحليل الدلالي للصور الخاص بالشرح، التوصل لنظام أوتوماتيكي لشرح الصور، يشمل عملية التحليل الدلالي التي قمنا بتطويرها لتخفيف الفجوة الدلالية من خلال تقديم اهتمام خاص لتحسين:

- تقسيم الصور بواسطة تنمية الاجزاء ومن خلال تجميع المناطق المتجاورة من أجل الحصول على أشياء دلالية،
  - تمثيل الخصائص البصرية من خلال الجمع بين أنواع مختلفة من الواصفات،
  - التصنيف من خلال الجمع بين القوة التمييزية للشبكات العصبية مع الطبيعة التوليدية لشبكات النظرية الافتراضية السانجة من اجل الاستفادة من التكامل فيما بينها.
- تقييم نظام الشرح المقترح، المنجز اعتمادا على قاعدة بيانات صور مختلفة عن تلك المستخدمة أثناء التعلم، أظهر أن نظامنا يوفر نتائج مقنعة.

أما بالنسبة للتعرف على الأنماط، اهتمنا بالتحليل الدلالي للصور الخاصة بتحديد علامات الطرق وامراض أشجار الحوامض. في حالة التعرف على علامات الطرق، وضعنا سلسلة تحليل دلالي للكشف والتعرف التلقائي على علامات المرور. يتجلى الكشف في تقسيم الصور بواسطة اللون لتحديد وعزل على حدة المناطق الزرقاء والمناطق الحمراء التي قد تتوافق مع علامات الطريق. ثم استخراج معالم هذه الألواح المفترضة لتحديد أشكالها وفئاتها: ألواح ثلاثية حمراء، دائرية حمراء، دائرية زرقاء أو مستطيلة زرقاء، وذلك باستخدام تقريب المضلع. اما بالنسبة للتعرف، فهو يتجلى في تحديد العلامات المكشوفة باستخدام الشبكة العصبية المدربة لمطابقة، كل فئة، مع متوسط قيم بيكسل كل صف وكل عمود من المناطق المكشوفة مع نماذج لوحات الطرق المكتسبة مسبقا، وقد أظهرت الاختبارات التي أجريت على مجموعة من الصور الحقيقية لحركة المرور على الطرق أداء واعد من النظام الجديد بعد تطويره.

وفيما يتعلق بتحديد أمراض أشجار الحمضيات، قمنا بتطوير المنهج التجريبي للكشف والتعرف التلقائي عن الأمراض التي تظهر أعراضها على الأوراق. ويستند مبدأ هذا النهج على التحليل الإحصائي لنسيج صور اوراق أشجار الحمضيات بواسطة المصفوفات التشاركية بالإضافة إلى تصنيف الأمراض المختلفة عن طريق الشبكة العصبية متعدد الطبقات. تنفيذ النهج المقترح من خلال الجمع بين الأوراق المريضة التي تم تحديدها مسبقا من قبل الخبراء، بين أن تحليل نسيج صور اوراق أشجار الحمضيات يتيح تفسيراً دلاليا مهما لأمراض أشجار الحمضيات.

في هذا العمل، قمنا بوضع أساليب التحليل الدلالية التي تدمج، اعتمادا على التطبيق المقصود، توقيعات على أساس واصفات مفصلة من الدلالات، مع قوة أفضل غنية بالمعلومات تمكن من استمرار تطوير أداء.

## العنوان: التحليل الدلالي للصور

كلمات البحث: صورة، تحليل، تجزئة، الميزات المرئية المنخفضة المستوى، المفاهيم العالية المستوى، الدلالة، الكشف، التعلم، التصنيف، التأويل، التعرف، الشرح، الالية، علامات الطرق، امراض الحوامض.



# Table des matières

## Chapitre 1 : INTRODUCTION GENERALE

1. Contexte et motivation.....	1
2. Bref état de l'art.....	2
3. Contribution.....	3
4. Organisation du manuscrit.....	4

## Chapitre 2 : ANALYSE SEMANTIQUE D'IMAGES POUR L'ANNOTATION AUTOMATIQUE D'IMAGES

1. Contexte et motivation.....	6
2. Difficulté et défi.....	7
3. Objectif.....	7
4. Etat de l'art.....	7
4.1. Les modèles graphiques.....	8
4.2. Les modèles probabilistes ou génératifs.....	8
4.3. Les modèles discriminatifs.....	9
5. Réalisation d'un système d'annotation automatique d'images.....	9
5.1. Démarche expérimentale.....	10
5.1.1. Elaboration de la démarche expérimentale.....	12
5.1.1.1. Segmentation d'images.....	12
5.1.1.2. Extraction des descripteurs.....	13
5.1.1.3. La classification.....	14
5.1.2. Mise en œuvre de la démarche expérimentale.....	16
5.1.2.1. Expérience 1.....	17

<b>5.1.2.1.1. Résultats</b> .....	18
<b>5.1.2.1.2. Analyse des résultats et conclusion</b> .....	21
<b>5.1.2.2. Expérience 2 : Combinaison des descripteurs</b> .....	22
<b>5.1.2.2.1. La première approche</b> .....	23
<b>5.1.2.2.1.1. Résultats</b> .....	23
<b>5.1.2.2.2. Deuxième approche</b> .....	25
<b>5.1.2.2.2.1. Résultats</b> .....	25
<b>5.1.2.2.2.2. Analyse des résultats et conclusion</b> .....	27
<b>5.1.2.3. Expérience 3 : Combinaison des classificateurs</b> .....	27
<b>5.1.2.3.1. Résultats</b> .....	28
<b>5.1.2.3.2. Analyse des résultats et conclusion</b> .....	29
<b>5.1.2.4. Expérience 4 : regroupement des régions des objets</b> .....	30
<b>5.1.2.4.1. Résultats</b> .....	32
<b>5.1.2.4.2. Analyse des résultats et conclusion</b> .....	32
<b>5.2. Structure du système d’annotation réalisé</b> .....	33
<b>6. Conclusion générale et perspectives</b> .....	37
<b>6.1. Conclusion générale</b> .....	37
<b>6.2. Perspectives</b> .....	38

**Chapitre 3 : ANALYSE SEMANTIQUE D’IMAGES  
POUR L’IDENTIFICATION DES PANNEAUX DE SIGNALISATION ROUTIERE**

<b>1. Contexte et motivation</b> .....	40
<b>2. Etat de l’art</b> .....	40
<b>3. Contribution</b> .....	42
<b>3.1. Module de détection</b> .....	43
<b>3.1.1. Méthode de détection</b> .....	43

<b>3.1.1.1. Détection de la couleur des panneaux.....</b>	<b>43</b>
<b>3.1.1.2. Détection des formes.....</b>	<b>46</b>
<b>3.1.1.2.1. Extraction des contours.....</b>	<b>46</b>
<b>3.1.1.2.2. Estimation de la forme des panneaux.....</b>	<b>46</b>
<b>3.1.2. Evaluation du module de détection.....</b>	<b>50</b>
<b>3.1.2.1. Premier test.....</b>	<b>50</b>
<b>3.1.2.2. Deuxième test.....</b>	<b>50</b>
<b>3.2. Module de la reconnaissance.....</b>	<b>52</b>
<b>3.2.1. Classification.....</b>	<b>52</b>
<b>3.2.2. Evaluation.....</b>	<b>53</b>
<b>4. Conclusion et perspectives.....</b>	<b>54</b>

**Chapitre 4:** ANALYSE SEMANTIQUE D'IMAGES  
POUR L'IDENTIFICATION DES MALADIES DES AGRUMES

<b>1. Contexte et motivation.....</b>	<b>56</b>
<b>2. Etat de l'art.....</b>	<b>56</b>
<b>3. Travail réalisé.....</b>	<b>59</b>
<b>3.1. Description de la démarche expérimentale développée.....</b>	<b>59</b>
<b>3.1.1. Acquisition.....</b>	<b>60</b>
<b>3.1.2. Filtrage de l'image.....</b>	<b>62</b>
<b>3.1.3. Suppression de l'arrière plan.....</b>	<b>63</b>

<b>3.1.4. Conversion de l'espace RVB en HSI</b> .....	64
<b>3.1.5. Extraction des descripteurs</b> .....	65
<b>3.1.6. Classification</b> .....	67
<b>3.2. Expériences</b> .....	67
<b>3.2.1. Expérience 1:</b> .....	68
<b>3.2.2. Expérience 2:</b> .....	68
<b>3.2.3. Expérience 3:</b> .....	68
<b>3.2.4. Expérience 4:</b> .....	69
<b>3.2.5. Analyse des résultats</b> .....	70
<b>4. Conclusion et perspective</b> .....	70

**Chapitre 5 : CONCLUSION GENERALE ET PERSPECTIVES**

<b>1. Conclusion générale</b> .....	72
<b>1.1. Annotation automatique d'images</b> .....	72
<b>1.2. Reconnaissance de formes</b> .....	72
<b>1.2.1. Identification des panneaux de signalisation routière</b> .....	73
<b>1.2.2. Identification des maladies des agrumes</b> .....	73
<b>2. Perspectives</b> .....	74
<b>2.1 . Amélioration des techniques d'analyse sémantique d'images</b> .....	74
<b>2.2. Utilisation d'autres approches d'analyse sémantiques</b> .....	74
<b>2.3. Développement d'autres processus d'analyse sémantique d'images</b> .....	74
<b>2.4. Analyse sémantique d'autres types d'images</b> .....	74
<b>Bibliographie</b> .....	75

## INTRODUCTION GENERALE

### **5. Contexte et motivation**

Le besoin en application de traitement numérique d'images se fait de plus en plus pressant à mesure que l'image numérique s'impose comme un support et une source d'information privilégiés. Grâce à l'émergence des nouvelles technologies de l'information, cette discipline s'est rapidement développée. Les applications de traitements envisageables sur les images sont : la restauration, le codage avec compression, la synthèse et enfin l'analyse qui correspond au cadre dans lequel notre travail s'est déroulé.

La démarche souvent adoptée en analyse d'image, consiste à segmenter l'image en régions d'intérêt et à utiliser ensuite les statistiques (caractéristiques bas-niveau) de ces régions pour en extraire l'information contenue dans l'image par des mesures de similarité visuelle. La difficulté rencontrée dans ce genre d'approche réside dans la correspondance qu'il faut trouver entre les caractéristiques numériques de bas-niveau liées au signal image et les concepts sémantiques représentant les objets effectivement présents dans l'image. Malgré de nombreuses et réelles avancées, le problème du fossé séparant ces deux niveaux n'est toujours pas résolu aujourd'hui, et reste l'un des principaux défis dans le domaine de la vision par ordinateur. Pour faire face à ce défi, des travaux intégrant l'information sémantique dans le processus d'analyse d'images, ont rapidement fleuri. La démarche générale adoptée consiste à commencer d'abord par identifier, pour une tâche donnée, les catégories sémantiques ayant un intérêt visuel pour l'observateur et de calculer les caractéristiques bas-niveau décrivant chacune de ces catégories, ensuite de relier, par apprentissage, ces deux niveaux par un modèle sémantique permettant la prédiction des concepts, a priori identifiés, à partir du contenu visuel de nouvelles images. L'analyse d'image se fait, dans ce cas, en termes de concept sémantique et non pas seulement en termes de caractéristiques de bas-niveau de l'image ; ce qui permet de créer une description de haut niveau de l'image et de réduire par conséquent le fossé sémantique.

Cette émergence de l'analyse sémantique d'images pour combler le fossé sémantique, nous a motivé à mener des travaux de recherche sur l'analyse sémantique d'images pour l'annotation d'images et la reconnaissance de formes.

L'objectif escompté consiste à mettre en œuvre des modèles capables d'extraire automatiquement les concepts sémantiques présents dans les images. Pour atteindre cet objectif, il faut chercher à améliorer le processus d'analyse en particulier au niveau de la segmentation, la représentation des caractéristiques visuelles et la classification.

## **6. Bref état de l'art**

Les travaux de recherche, réalisés sur l'analyse sémantique des images, ont proposé plusieurs approches pour l'interprétation sémantique du contenu visuel des images. Parmi les approches proposées, nous citons :

- Les approches directes [1-7] : qui, en général, ont pour objectif de résoudre le problème par des méthodes statistiques. Ces méthodes cherchent directement des liens entre les caractéristiques images bas-niveau et les caractéristiques sémantiques haut-niveau.
- Les approches linguistiques [8-12] : elles sont basées sur la construction d'un vocabulaire visuel comme intermédiaire entre les données numériques et sémantiques. Nous classons parmi les approches linguistiques les méthodes de type sacs de mots dont le principe est inspiré du traitement de documents textuels. Ce principe consiste à extraire les caractéristiques de l'image et à les grouper par rapport à des clusters constituant un vocabulaire visuel. Des histogrammes d'occurrences de ces mots visuels sont calculés, permettant une description très efficace de l'image.
- Les approches compositionnelles [13-18]: nous appelons "approches compositionnelles" les approches cherchant à reconnaître des parties de l'image avant d'étiqueter le tout (éventuellement). Dans ce schéma, l'image est considérée comme une composition d'objets, ou plus généralement de concepts sémantiques.
- Les approches structurelles [19-21]: ces approches intègrent une notion de géométrie par rapport aux approches linguistiques et compositionnelles. Des liens géométriques peuvent être intégrés à différents niveaux : entre les parties d'un objet, ou entre des objets d'une scène.

D'après la lecture de ce bref rétrospectif, il ressort que l'analyse sémantique des images est un champ de recherche très actif et qu'un effort de recherche scientifique considérable a été effectué pour résoudre le problème du fossé sémantique. Pour chaque

approche, de nombreuses méthodes sont développées en adoptant différentes manières d'extraction et de représentation des caractéristiques visuelles, et aussi différentes techniques d'apprentissage et de classification. Cependant, malgré la pluralité des approches d'analyse sémantique d'images et la diversité des méthodes proposées, le fossé sémantique demeure un problème ouvert qui suscite un intérêt croissant.

## 7. Contribution

L'objectif de notre contribution, à l'analyse sémantique d'images, est la réduction du fossé sémantique observé au niveau des deux thèmes principaux de la vision par ordinateur à savoir l'annotation automatique d'image et la reconnaissance de formes. Comme les formes à reconnaître dépendent des applications envisagées, nous nous sommes intéressés, dans ce travail, à l'identification des panneaux de signalisation routière et des maladies des agrumes.

Afin de chercher à réduire le fossé sémantique, nous avons apporté, au cours de l'analyse sémantique des thèmes considérés, un soin particulier à l'extraction et la représentation visuelle des caractéristiques, et aussi à l'établissement de la correspondance entre les caractéristiques bas-niveau et les concepts haut-niveau par des modèles statistiques.

Notre contribution dans cette thématique de recherche se résume comme suit:

- ✓ concernat l'analyse sémantique des images pour l'annotation automatique d'images
  - Proposition d'une méthode de combinaison des caractéristiques visuelles qui a fait l'objet d'une publication :
    - **B. Minaoui**, M. Oujouara, M. Fakir and M. Sajieddine.: *Toward an Effective Combination of multiple Visual Features for Semantic Image Annotation* : TELKOMNIKA Indonesian Journal of Electrical Engineering, Vol. 15, N°. 3, p: 401-408, 2015.
  - Proposition d'un classificateur intégrant la complémentarité des classificateurs discriminatifs et génératifs qui a fait l'objet d'une publication :
    - **B. Minaoui**, M. Oujouara, M. Fakir: "*Combining Generative and discriminative Classifiers for Semantic Automatic Image Annotation*". *International Journal of Image Processing (IJIP)*, Vol. 8, N°5, p:225- 244, 2014.
  - Proposition d'une méthode de segmentation donnant lieu à une publication :
    - M. Oujouara, **B. Minaoui**, M. Fakir : *A semantic hybrid approach based on regrouping adjacent regions for automatic image annotation* : acceptée pour publication au journal *Pattern Recognition and Image Analysis - Advances in Mathematical Theory and Applications*, Springer.

- ✓ concernant l'analyse sémantique des images de la signalisation verticale routière
  - Proposition d'une méthode de détection des panneaux de signalisation routière. Celle-ci a donné lieu à une communication:
    - A. Salhi, B. **Minaoui**, M. Fakir: *Robust Automatic Traffic Signs Detection Using Fast Polygonal Approximation of Digital Curves*: International Conference on Multimedia computing and system (ICMCS) , p:433 - 437, Marrakech, April 2014: Publié dans IEEE xplore Digital Library, DOI:10.1109/ICMCS.2014.6911185
  - Proposition d'un système de reconnaissance des panneaux de circulation routière. Ce travail a abouti à une publication :
    - A. Salhi, B. **Minaoui**, M. Fakir : *Robust Automatic Traffic Signs Recognition Using Fast Polygonal Approximation of Digital Curves and Neural Network* : International Journal of Advanced Computer Science and Applications(IJACSA), Special Issue on Advances in Vehicular Ad Hoc Networking and Applications p: 1-7, 2014 DOI: 10.14569/SpecialIssue.2014.040201
  
- ✓ En ce qui concerne l'analyse sémantique des images des feuilles des agrumes
  - Proposition d'une méthodologie de détection et de reconnaissance des maladies des agrumes basée sur l'utilisation de la texture des feuilles. Ce travail a fait l'objet d'une communication:
    - E. Moustamsik, B. **Minaoui**, M. Fakir: *Contribution to the automatic detection of citrus disease*: International conference on Business Intelligence (CBI'14) April 29-30, 2014, Beni Mellal Morocco

## 8. Organisation du manuscrit

Le manuscrit est organisé de la manière suivante :

Le chapitre 2 est consacré à l'analyse sémantique des images pour l'annotation des images. Il définit le contexte du travail et décrit la démarche expérimentale et les techniques d'analyse mises en œuvre pour réaliser un système d'annotation automatique des images.

Dans le chapitre 3, nous présentons le travail concernant l'analyse sémantique des images de la signalisation verticale routière. Après avoir défini le contexte de ce travail et donné un bref état de l'art, nous avons présenté les méthodes d'analyse sémantiques développées pour la détection et la reconnaissance automatique des panneaux de signalisation routières.

Le chapitre 4 présente le travail réalisé sur l'analyse sémantique des images des feuilles des agrumes. Il définit le contexte du travail et décrit la méthodologie d'analyse proposée pour la détection et la reconnaissance automatique des maladies des agrumes.



En fin, le chapitre 5 donne une synthèse générale de la thèse et dégage les perspectives importantes qui font suite à ce travail.

## *Chapitre 2*

### ANALYSE SEMANTIQUE D'IMAGES POUR L'ANNOTATION AUTOMATIQUE D'IMAGES

#### **5. Contexte et motivation**

La croissance rapide d'Internet et de l'information multimédia a engendré un besoin en technique d'indexation et de recherche d'information multimédia, et plus particulièrement en recherche d'images. Des systèmes de recherche d'images ont été développés pour permettre de faire des recherches dans des bases de données images.

Les premiers systèmes de recherche d'image réalisés, généralement utilisés par le grand public, sont basés sur des index textuels, dans lesquels les images sont annotées manuellement ou à travers un processus d'extraction d'annotation à partir du texte qui entoure les images dans les documents (web, pdf, doc, ...). Les moteurs de recherche d'images classiques (Google, Yahoo!) s'appuient sur les noms des fichiers ou sur les informations textuelles qui entourent les images, dans les pages web ou les documents, pour indexer les images. Cependant, ces informations textuelles ne décrivent pas toujours la sémantique des images d'une façon précise.

La nécessité de chercher les images indépendamment des annotations textuelles mène à la proposition de la deuxième famille de systèmes de recherche d'image appelés systèmes de recherche d'images par le contenu (CBIR). Ces systèmes reposent sur une description numérique du contenu visuel pour mesurer la similarité entre les requêtes et les images. Cependant, la performance de ces systèmes de recherche d'images par similarité est loin d'être satisfaisante à cause de ce que l'on appelle le fossé sémantique séparant les concepts sémantiques et les caractéristiques numériques de bas niveau.

Afin de réduire cet écart sémantique, la communauté scientifique s'est orientée vers l'annotation automatique des images en vue de développer des systèmes de recherche d'images par des mots clés décrivant leur contenu sémantique. De tels systèmes permettraient aux utilisateurs d'utiliser un langage expressif simple pour chercher une image dans une base d'images annotées.

Cependant, les techniques d'annotation automatique d'images utilisées actuellement ne sont pas assez efficaces et assez développées pour permettre une bonne description des images, ce qui fait du problème d'annotation automatique un champ de recherche actif

largement ouvert à de nouvelles propositions. C'est dans ce contexte qu'a émergé notre motivation de mener un travail de recherche sur l'analyse sémantique d'images afin de proposer un système d'annotation automatique d'images à des fins d'amélioration de la recherche d'images par mots clés.

## **6. Difficulté et défi**

La principale difficulté dans l'élaboration d'un système d'annotation d'images performant est la variation visuelle des objets. En effet, l'annotation automatique des images consiste à décrire automatiquement le contenu sémantique des images par des mots clés à travers une analyse sémantique de leur contenu visuel. Cette analyse est confrontée par le problème de la variabilité de forme, de couleur propre à chaque type d'objets et des variations qui peuvent parvenir de changements d'échelle, de changements de luminosité, de la rotation et de l'occultation. Ceci accentue le problème de la variation intra-classes et de la ressemblance inter-classes. Cette variation visuelle des objets crée des relations compliquées entre le niveau du contenu visuel et le niveau des classes d'objets, ce qui rend le problème du fossé sémantique très difficile à résoudre.

Ainsi, le franchissement du fossé sémantique constitue le défi majeur des systèmes d'annotation automatique d'image et des systèmes de recherche d'images basés sur eux.

## **7. Objectif**

L'objectif de notre travail, mené sur l'analyse sémantique pour l'annotation automatique d'images, est la réduction du fossé sémantique.

Dans le but de franchir ce fossé sémantique, nous allons chercher à réduire l'effet des variations visuelles des objets sur la qualité de l'annotation d'images par l'amélioration, d'une part de la description visuel pour augmenter la corrélation entre le niveau du contenu visuel et le niveau sémantique et, d'autre part, de la classification automatique pour faciliter l'association des caractéristiques de bas niveau avec celles de haut niveau.

## **4. Etat de l'art**

Plusieurs modèles d'annotation automatique d'images par des mots-clefs ont été proposés dans la littérature. On peut distinguer trois catégories : les modèles graphiques, les modèles probabilistes ou génératifs et les modèles discriminatifs

#### **4.1. Les modèles graphiques**

Ces modèles sont basés sur des méthodes d'annotation à base de graphes [22-27]. La plupart de ces méthodes consistent à construire un graphe dit de voisinage ou de similarité. La plupart des méthodes existantes utilisent des méthodes de construction de graphes simples, telles que les graphes entièrement connectés, les méthodes  $\varphi$ -ball, Minimum Spanning Tree (MST) ou les  $k$  plus proches voisins.

Ces méthodes ont l'avantage d'être indépendantes du domaine et simples à paramétrer. Mais, dans le problème d'annotation d'images, il est nécessaire de représenter une large quantité de données dans le graphe, d'autant plus que les caractéristiques de bas niveau utilisées sont de grande dimension. Ainsi les graphes  $\varphi$ -ball ou les méthodes basées sur les  $k$  plus proches voisins, sont plus adaptées que les méthodes à base de graphes entièrement connectés. Cependant, ces méthodes dépendent des paramètres  $k$  et  $\varphi$  utilisés. Un mauvais réglage de ces paramètres peut entraîner des oublis d'arêtes ou au contraire de mauvaises connections entre les données. Les méthodes à base de MST semblent donc plus appropriées, mais les arbres couvrants minimaux présentent souvent beaucoup de branches et de connections entre un sommet et plusieurs autres sommets, ce qui engendre une plus grande complexité.

#### **4.2. Les modèles probabilistes ou génératifs**

Le principe de ces modèles est basé sur l'estimation de la distribution conjointe des mots-clés du vocabulaire d'annotation conditionnellement aux caractéristiques visuelles des images. Pour les nouvelles images non ou partiellement annotées, le calcul de cette distribution permet la prédiction des mots-clés manquants. Parmi ces modèles, on trouve le modèle bayésien [28], le modèle de co-occurrence [29], le modèle de traduction statistique [30], le modèle de Markov caché [31] et les modèles introduisant une variable aléatoire latente (cachée) pour faire le lien entre les caractéristiques visuelles et les mots-clés [32-34] tels que le modèle de l'allocation de Dirichlet latente (LDA), le modèle de mélange de distributions Gaussiennes - multinomiales (modèle GM-Mixture), le modèle Gaussian-Multinomial LDA (GM-LDA) et le modèle correspondance LDA (CORR-LDA).

Les modèles probabilistes présentent les avantages:

- Ils peuvent rendre en compte des données manquantes ou partiellement annotées,
- Ils peuvent ajouter facilement de nouvelles classes, il suffit d'en apprendre le modèle

- Ils prennent en compte la composition (rajout de lunettes, d'un chapeau ou de moustaches à un visage) avec le minimum d'exemples possibles durant l'entraînement.

En fin, l'inconvénient majeur de la plupart des méthodes probabilistes est qu'elles ont souvent besoin d'un certain nombre d'itérations pour classer une nouvelle image.

### **4.3. Les modèles discriminatifs**

Ces modèles sont basés sur des méthodes d'annotation qui consistent à extraire les caractéristiques visuelles de bas niveau du contenu des images et de les relier aux mots clés à l'aide des classificateurs discriminatifs tels que les réseaux de neurones (RN) [35], les machines à vecteurs de support ou séparateurs à vaste marge (SVM) [36, 37] et les arbre de décision (DT) [37,38]. Ces classificateurs permettent, après une phase d'apprentissage, de prédire et associer des mots clés à de nouvelles images. L'apprentissage dans ce cas a pour objectif de tracer des frontières précises entre les différentes classes d'images à annoter.

Ces modèles ont les avantages suivants :

- Ils ne nécessitent pas d'hypothèse sur la distribution des caractéristiques,
- Ils se concentrent sur les aspects qui distinguent une classe des autres ce qui est potentiellement plus efficace pour prendre une décision,
- Ils sont rapides pour classer une nouvelle image,

Cependant leurs inconvénients sont :

- Ils doivent recevoir tous les exemples possibles durant l'entraînement pour prendre en compte la composition.
- Ils ont besoin d'un grand nombre d'exemples au cours de la phase de l'apprentissage

## **6. Réalisation d'un système d'annotation automatique d'images**

Afin de contribuer à la résolution de la problématique du fossé sémantique, nous nous sommes intéressés à la réalisation d'un système d'annotation automatique d'images et à l'amélioration de ses performances, en terme de précision, pour réduire le fossé sémantique par une analyse sémantique d'images approfondie.

Afin d'atteindre cet objectif, nous nous sommes penchés, au début de ce travail, sur l'étude bibliographique concernant l'annotation d'images. Cette étude nous a permis, d'une

part, de concevoir la structure du système à réaliser (Figure 1) et, d'autre part, d'instancier et d'étudier les différents modèles de l'annotation automatique d'images de l'état d'art pour définir les éventuelles pistes à explorer en vue d'améliorer le processus d'analyse.

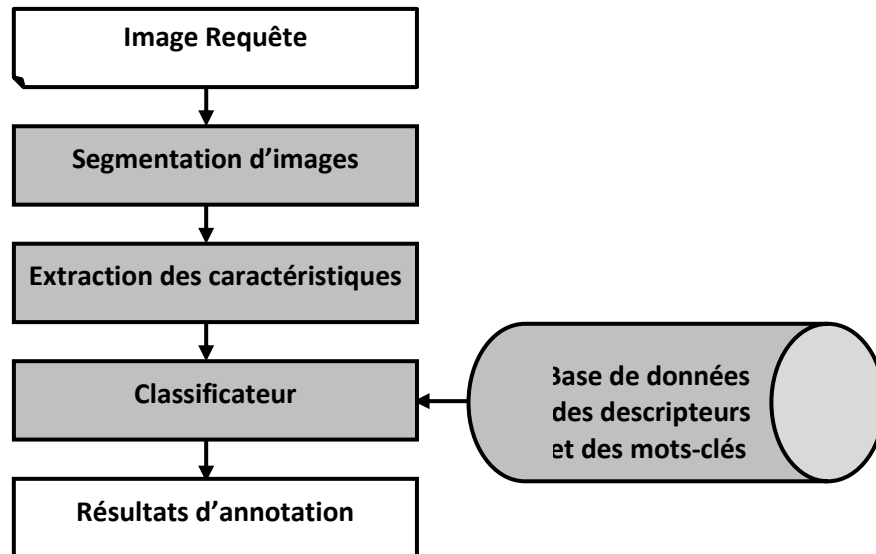


Figure 1 : Structure générale d'un système d'annotation

Dans les parties subséquentes de ce chapitre, nous allons décrire la démarche expérimentale d'analyse sémantique d'images adoptée et les expériences réalisées pour mettre en œuvre et évaluer le système d'annotation automatique d'images ainsi conçu. Les résultats obtenus pour chaque expérience seront présentés et discutés afin de tirer des conclusions sur l'efficacité du système d'annotation automatique d'images réalisé, ainsi que les difficultés et les défis qui doivent être relevés.

### 6.1. Démarche expérimentale

La démarche expérimentale adoptée pour réaliser le système d'annotation automatique d'images conçu, est représentée par le schéma fonctionnel donné par la figure 2. Cette démarche expérimentale se déroule en deux phases:

- La phase d'apprentissage durant laquelle un classificateur est entraîné à établir une correspondance entre les vecteurs descripteurs des régions d'une image segmentée et les concepts sémantiques, exprimés par des mots clés, représentant ces régions.
- La phase d'annotation automatique des images de test pour évaluer le système.

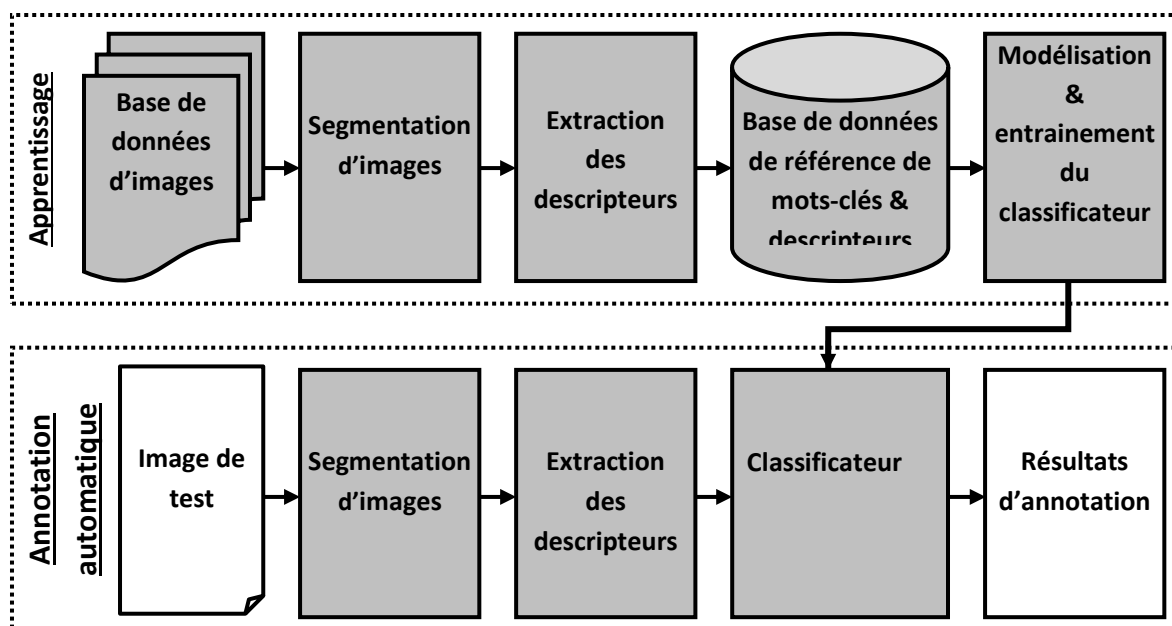
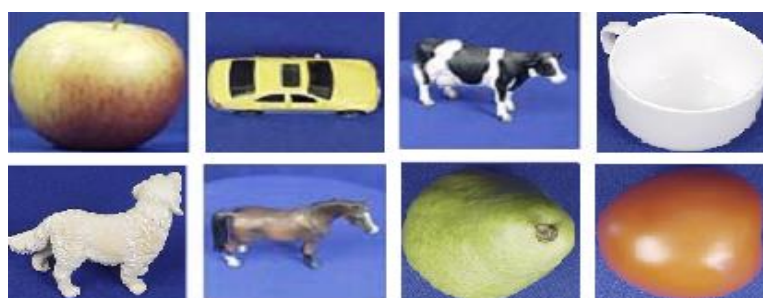


Figure 2 : Schéma fonctionnel du système d'annotation automatique d'images proposé.

Les bases de données images utilisées au cours de cette démarche expérimentale, sont : la base de données images ETH-80 contenant un ensemble d'images en couleur de 8 objets différents [39], la base de données COIL-100 qui contient des images en couleur de 100 objets différents [40] et la base de données NATURE [41] contenant des images en couleur de 6 éléments de la nature. La figure 3 montre quelques exemples d'images de ces bases.



ETH-80



COIL-100



*Figure 3: Exemples d'objets des bases de données d'images utilisées*

### 6.1.1. Elaboration de la démarche expérimentale

Pour mettre en œuvre la démarche expérimentale adoptée, nous avons élaboré les différentes parties du procédé d'analyse sémantique d'images, utilisé lors de ces phases, en y intégrant les techniques d'analyse et les méthodes d'apprentissage suivantes :

#### 6.1.1.1. Segmentation d'images

Idéalement la segmentation consiste à partitionner l'image en un ensemble de régions connexes représentant les concepts sémantiques présents dans l'image. Jusqu'à présent, il n'existe pas de méthode universelle pour segmenter les images correctement. C'est pourquoi nous avons implémenté l'algorithme de segmentation par croissance de région [42] qui est un algorithme de segmentation simple à mettre en œuvre, rapide et il peut fournir une bonne approximation de la segmentation recherchée. Le principe de cet algorithme consiste à faire croître chaque région autour d'un pixel de départ,

La figure 4 représente deux exemples de segmentation d'image par ces algorithmes.



*Figure 4: Exemples de segmentation d'images par croissance de région*



### 6.1.1.2. Extraction des descripteurs

Après avoir segmenté l'image d'entrée en un ensemble de régions représentant les objets contenus dans l'image, les descripteurs représentant le contenu visuel des objets de l'image sont extraits par les algorithmes suivants :

- Algorithmes d'extraction des descripteurs de forme [43] qui permettent d'extraire pour chaque plan de l'image couleur :
  - 7 éléments pour les moments invariants de Hu (Hu1, Hu2, Hu3, Hu4, Hu5, Hu6, Hu7).
  - 9 éléments en utilisant l'ordre 4 pour les moments de Zernike (Z00, Z11, Z20, Z22, Z31, Z33, Z40, Z42, Z44).
  - 10 éléments en utilisant l'ordre 3 pour les moments de Legendre (L00, L01, L02, L03, L10, L11, L12, L20, L21, L30).
- L'algorithme d'extraction des descripteurs de couleur qui consiste à extraire 16 éléments par plan d'images pour les histogrammes de couleur d'images [44].
- L'algorithme d'extraction des descripteurs de la texture qui permet de calculer un vecteur contenant 56 composantes des caractéristiques résultant des 14 indices d'Haralick extraits des matrices de cooccurrences pour chacune des 4 directions de proximité de chaque plan d'image [45].

Ces descripteurs sont calculés pour chaque plan de couleurs d'images et ils sont regroupés et fusionnés par concaténation en un seul vecteur de façon à ce qu'ils représentent une entrée unique aux classificateurs. La figure 5 représente le processus de calcul de ces descripteurs pour les trois canaux de couleur R, G et B d'une image couleur. Ainsi, le vecteur des moments de Hu contient 21 éléments, le vecteur des moments de Zernike est formé par 27 éléments, le vecteur des moments de Legendre est constitué par 30 éléments, le vecteur des histogrammes de couleur contient 48 éléments et le vecteur de la texture 168 éléments.

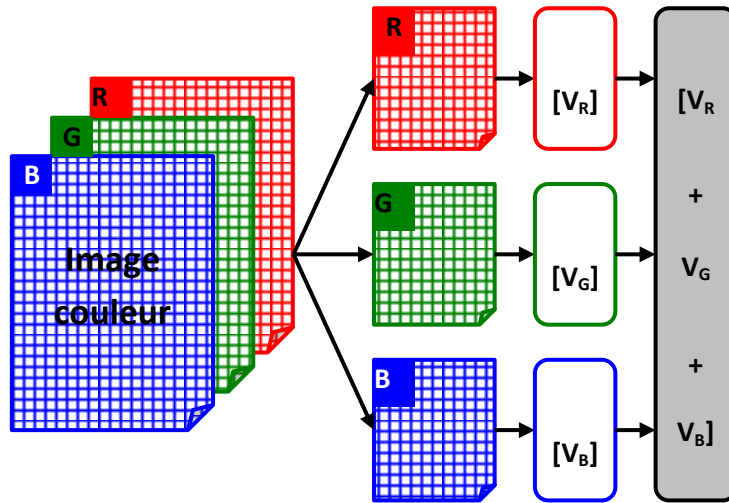


Figure 2.5 : Processus de calcul des descripteurs d'une image couleur pour les trois canaux de couleur R, G et B.

En plus de ces descripteurs, nous avons utilisé le descripteur global GIST [46] construit en combinant les amplitudes obtenues à la sortie de K filtres de Gabor [47] à différentes échelles E et orientations O. Chaque image est subdivisée en N \* N blocs (N compris entre 2 et 16). Ce qui donne un vecteur de dimension N \* N \* K \* E \* O = 128 (dans notre cas, N=4, K=4, E=4, O=2).

### 6.1.1.3. La classification

Afin d'automatiser l'annotation des images, nous avons implémenté des classificateurs permettant d'associer à une image les mots-clés convenables à partir des descripteurs extraits des différentes régions de cette image. Les classificateurs implémentés sont :

- Un classificateur réseau de neurones de trois couches (RNM) à apprentissage par rétro propagation [48]. la fonction de transfert utilisée par ce classificateur discriminatif est la tangente hyperbolique défini par:

$$f(x) = \frac{2}{1 + \exp(-2x)} - 1 \quad (2.1)$$

Le nombre M de nœuds de la couche d'entrée est fixé par le nombre de composante du vecteur descripteur et celui de la couche de sortie est fixé par le nombre N de classes utilisées

lors de l'apprentissage du réseau de neurones. Quant au nombre  $L$  de neurones dans la couche cachée, il est approximativement égale à:

$$L = E\left(1 + \sqrt{M(N+2)}\right) \quad (2.2)$$

Avec  $E(x)$  désigne la partie entière de  $x$ .

- Deux classificateur Séparateurs à Vaste Marge (SVM) multi classes basés sur deux approches différentes [49] :
  - L'approche un contre tous (SVM All) qui consiste à entraîner un SVM binaire en utilisant les éléments d'une classe contre tous les autres.
  - L'approche un contre un (SVM one) qui consiste, pour  $N$  classes, à entraîner chacun des  $N(N-1)/2$  classificateurs SVM binaires sur chacun des couples de classes, puis à décider la classe gagnante par un vote majoritaire.

Les 2 classificateurs discriminateurs SVM multi classes, implémentés dans ce travail, combinent des classificateurs SVM binaires à noyau gaussien définie par:

$$k(x, y) = \exp\left(-\frac{\|x - y\|^2}{2\sigma^2}\right) \quad (2.3)$$

Où  $\sigma > 0$

- Un classificateur K plus proche voisin (K-ppv) [50] :
 

Ce classificateur est basé sur la mémoire, qui contrairement à d'autres méthodes statistiques, ne nécessite aucun apprentissage. Il consiste à observer les  $k$  plus proches voisins d'une observation afin de décider de la classe d'appartenance de cette observation. Pour une nouvelle observation à classer, cet algorithme calcule la distance de cette nouvelle observation à chaque observation présente dans l'ensemble d'apprentissage. Cette distance est souvent la distance euclidienne alors que d'autres distances sont utilisées telle la distance tangente, distance de Manhattan,..... etc.

- Un classificateur réseau bayésien naïf (RBN) [51]

La structure de ce classificateur génératif est composée de deux niveaux : un premier ne comportant qu'un seul nœud  $C$ , représentant les classes  $C_i$  d'objets possibles, et le second plusieurs nœuds  $X_1...X_n$  représentant des descripteurs ayant pour seul parent la classe  $C_i$ . Ces modèles sont dits naïfs car ils font l'hypothèse que tous les descripteurs  $X_i$  sont indépendants entre eux. La loi de probabilité jointe d'un classificateur bayésien s'écrit :

$$P(C_i|X) = P(C_i) \prod_{j=1}^n P(X_j|C_i) \quad (2.4)$$

Avec 
$$P(C_i) = \frac{N(C_i)}{N} \quad (2.5)$$

où  $N(C_i)$  représente le nombre d'image ou objet appartenant à la classe  $C_i$ ,  $N$  le nombre total d'images d'apprentissage

et 
$$P(X_j|C_i) = NG(\mu_i, \sigma_i|C_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left\{-\frac{1}{2\sigma_i^2}(x_i - \mu_i)^2\right\} \quad (2.6)$$

où NG est la loi normale gaussienne.  $\mu_j$  et  $\sigma_j$  sont les moyennes et les variances des attributs  $X_j$

Ainsi, l'apprentissage de paramètres, dans ce cas, revient à calculer les moyennes et les variances de tous les attributs de chaque classe. Après ce calcul, le réseau sera quantifié et il peut être utilisé pour classer de nouveaux objets étant données leurs nouvelles valeurs d'attributs en appliquant la règle de décision de Bayes d comme suit:

$$d(X) = \arg \max_{C_i} P(C_i|X) = \arg \max_{C_i} P(C_i) \prod_{j=1}^n P(X_j|C_i) \quad (2.7)$$

Les classes choisies seront donc celles dont la probabilité est la plus grande.

### 6.1.2. Mise en œuvre de la démarche expérimentale

La mise en œuvre de la démarche expérimentale consiste à faire tourner les algorithmes intégrés dans le procédé d'analyse sémantiques tout en cherchant parmi ces algorithmes les quels permettent d'obtenir le meilleur taux d'annotation automatique d'images. Pour ce faire, nous avons réalisé les expériences suivantes :

### 6.1.2.1. Expérience 1

Afin de de mettre en lumière le comportement de chaque type de descripteur et de chaque type de classificateur et de faire une étude comparative sur leur performance, nous les avons utilisé séparément pour annoter les images des trois bases ETH-80, COIL-100 et NATURE.

Le principe utilisé est représenté par le schéma bloc illustré sur la figure 6.

Dans cette expérience, nous avons utilisé 40 images pour l'apprentissage et 40 autres images pour les tests pour chacune des bases de données images ETH-80, COIL-100 et NATURE,

#### 6.1.2.1.1. Résultats

Les résultats d'annotation d'images obtenus par cette expérience sont regroupés dans les Tableaux 2.1 et 2.2, et illustrés par les histogrammes des figures 7 et 8.

*Tableau 2.1 : Taux d'annotation d'images obtenus par l'expérience1*

Base de données	Descripteur	Classificateur				
		Discriminatif				Génératif
		K-PPV	SVM-One	SVM-All	RNM	RBN
ETH-80	Hu	60%	54%	58%	62%	70%
	Zernike	68%	63%	65%	70%	76%
	Legendre	78%	74%	75%	80%	83%
	RGB	56%	58%	55%	65%	72%
	HSV	50%	46%	41%	55%	66%
	Texture	48%	44%	43%	54%	58%
	GIST	59%	55%	53%	60%	66%
COIL-100	Hu	50%	44%	44%	61%	65%
	Zernike	71%	60%	65%	74%	76%
	Legendre	74%	70%	75%	78%	80%
	RGB	63%	60%	59%	65%	68%
	HSV	61%	60%	59%	64%	73%
	Texture	45%	43%	41%	50%	56%
	GIST	55%	53%	50%	59%	65%

Tableau 2.2: Taux d'annotation d'images obtenus par l'expérience 1

Base de données	Descripteur	Classificateur			
		Discriminatif			Génératif
		SVM-One	SVM-All	RNM	RBN
NATURE	Legendre	67%	67%	67%	73%
	RGB	66%	65%	60%	67%
	Texture	68%	69%	70%	80%

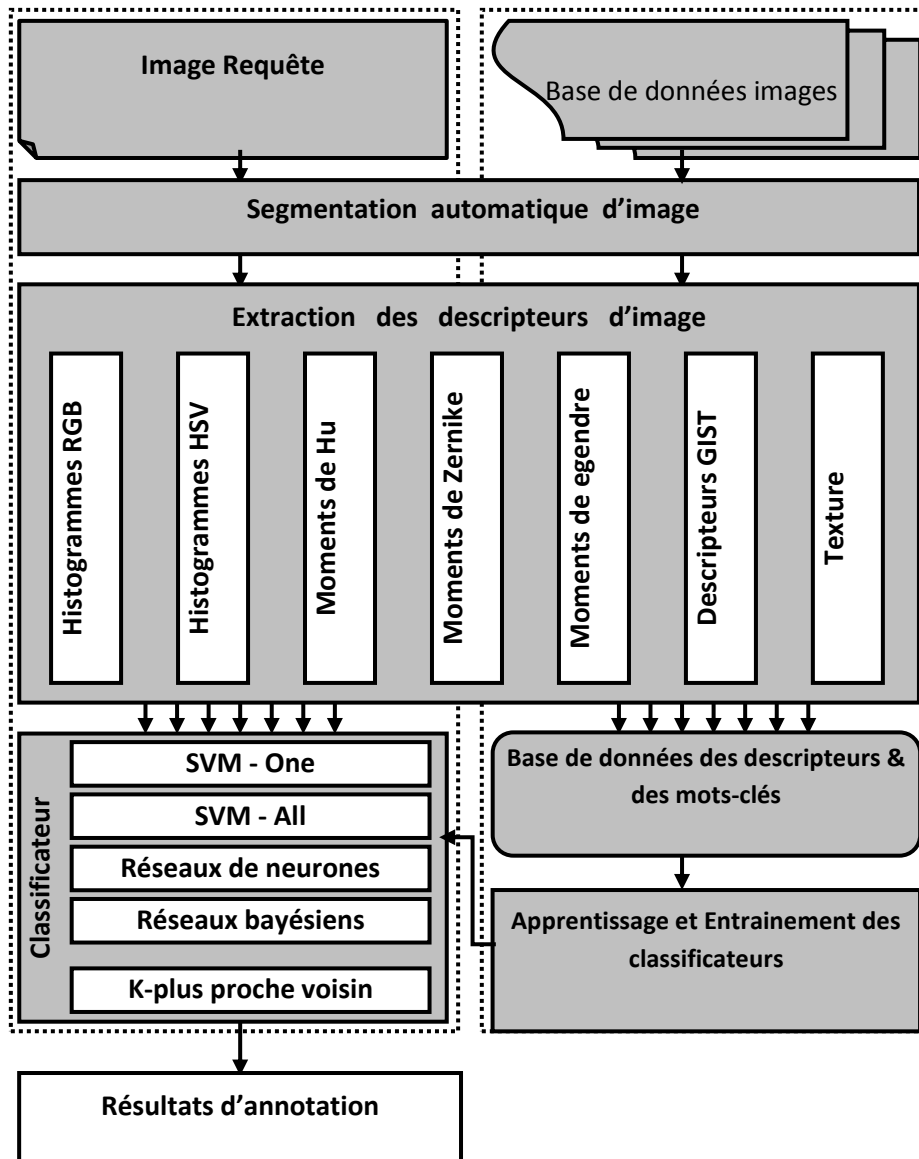


Figure 6: Principe du système d'annotation automatique d'image basé sur l'utilisation séparée des descripteurs et classificateurs.

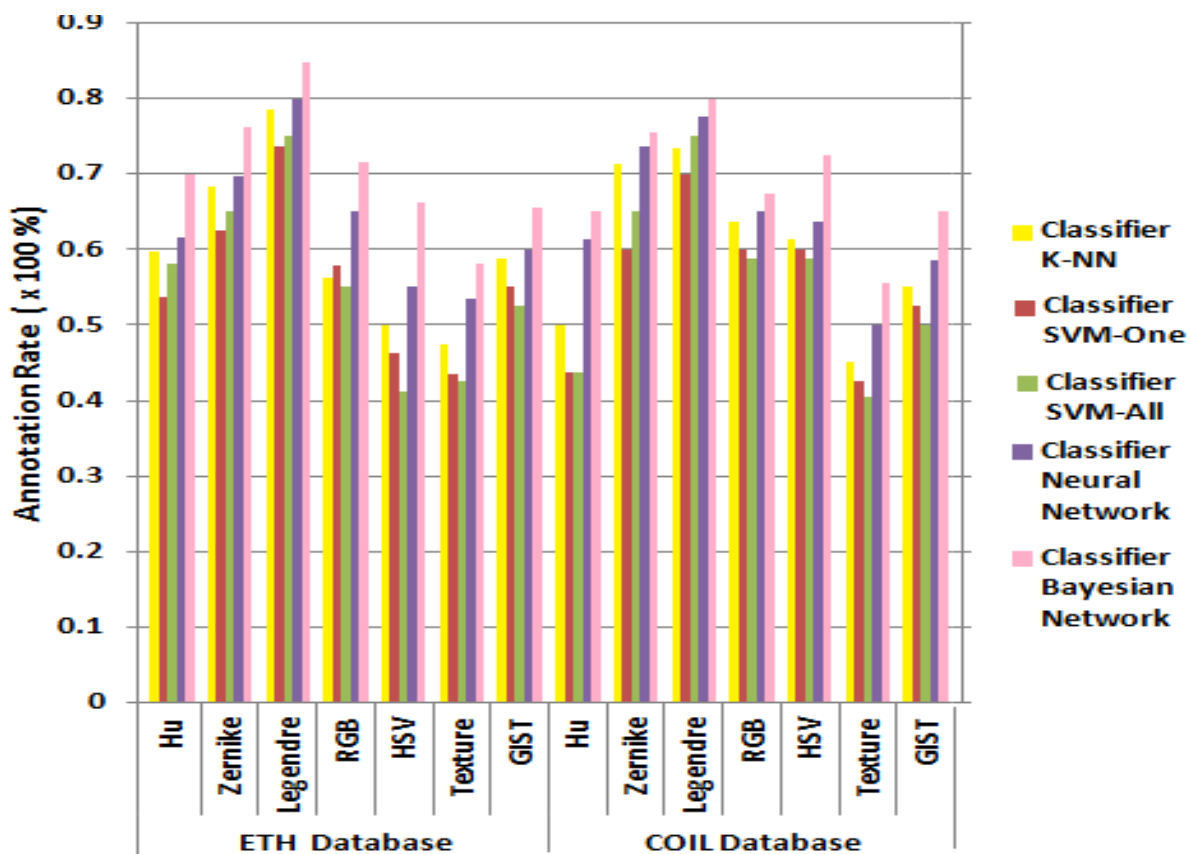


Figure 7: Taux d'annotation dans le cas d'usage séparé des classificateurs et descripteurs pour les bases de données d'images ETH-80 et COIL-100.

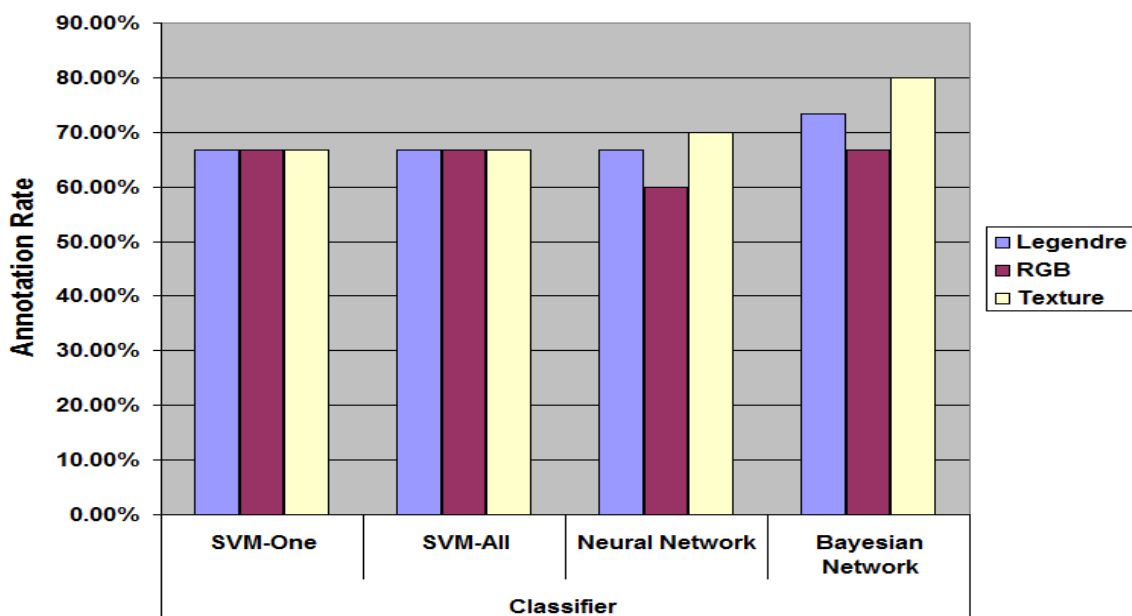


Figure 8: Taux d'annotation dans le cas d'usage séparé des classificateurs et descripteurs pour la base de données d'images NATURE.

Les matrices de confusion données par les figures 9, 10 montrent les objets mal annotés (indiqués par la couleur rouge) dans le cas d'utilisation des moments de Legendre comme descripteur et le réseau bayésien naïf comme classificateur pour des images des deux bases ETH-80 et NATURE.

	Mots clés prédits							
	apple	car	cow	cup	dog	horse	pear	tomato
apple	5	0	0	0	0	0	0	0
car	0	3	1	0	1	0	0	0
cow	0	2	3	0	0	0	0	0
cup	0	1	0	4	0	0	0	0
dog	0	1	0	0	4	0	0	0
horse	0	1	0	0	0	4	0	0
pear	0	0	0	0	0	0	5	0
tomato	2	0	0	0	0	0	0	3

Figure 9: Matrice de confusion dans le cas d'utilisation des réseaux bayésiens et des moments de Legendre pour des images de la base de données ETH-80.

	Predicted keyword					
	forest	gazon	ground	sahara	sky	water
forest	4	1	0	0	0	0
gazon	0	4	1	0	0	0
ground	0	2	3	0	0	0
sahara	2	0	0	3	0	0
sky	0	0	0	0	5	0
water	0	2	0	0	0	3

Figure 10: Matrice de confusion dans le cas d'utilisation des moments de Legendre et le réseau bayésien naïf pour des images à partir de la base de données d'images NATURE.

La figure 11 montre la matrice de confusion dans le cas d'utilisation de la texture (indices d'Haralik extraits des matrices de cooccurrences chromatiques) comme descripteur et le réseau bayésien naïf comme classificateur pour des images à partir de la base de données d'images NATURE.



	Mots clés prédits					
	forest	gazon	ground	sahara	sky	water
forest	5	0	0	0	0	0
gazon	0	4	1	0	0	0
ground	0	1	4	0	0	0
sahara	0	0	0	5	0	0
sky	0	0	1	1	3	0
water	0	1	1	0	0	3

Figure 11: Matrice de confusion dans le cas d'utilisation de la texture et les réseaux bayésien pour des images à partir de la base de données d'images NATURE.

### 6.1.2.1.2. Analyse des résultats et conclusion

L'analyse des Figures 7,8, 9, 10 et 11, montre bien que le taux d'annotation automatique d'images dépend du type de classificateur et de la nature du descripteur. En effet, nous avons constaté que :

- Pour chaque descripteur, le meilleur taux d'annotation est obtenu par le classificateur génératif qui est le réseau bayésien naïf.
- Le meilleur taux d'annotation obtenu par les classificateurs discriminatifs est celui qui est réalisé par les réseaux de neurones. Ce taux s'approche bien de celui obtenu par le réseau bayésien naïf, en particulier dans le cas de la base de données d'images COIL-100 contenant plus d'objets que la base de données d'images ETH-80.
- Pour les deux bases de données d'images ETH-80 et COIL-100, contenant des objets ayant en général des formes bien définies, les descripteurs de Legendre permettent d'obtenir, pour chaque classificateur, le meilleur taux d'annotation.
- Pour la base de données d'images NATURE, le meilleur taux d'annotation est obtenu, pour chaque classificateur, par le descripteur de texture (indices d'Haralik extraits des matrices de cooccurrences chromatiques).
- Certains descripteurs réalisent avec les réseaux de neurones des taux d'annotation meilleurs que ceux obtenus avec le réseau bayésien naïf ; et l'inverse pour d'autres descripteurs.

- Certains objets sont confondus avec d'autres objets (voir les matrices de confusion sur les Figures 9,10 et 11 ).

En se basant sur l'ensemble des constatations et remarques citées ci-dessus, nous pouvons conclure que pour mieux annoter les images d'une large base de données d'images contenant des objets qui ne peuvent être se distinguer que, soit par leurs formes, soit par leurs textures ou soit par leurs couleurs, il faut alors :

- Combiner aux moins trois descripteurs de nature différente : un descripteur de forme, un descripteur de texture et un autre descripteur de couleur.
- Combiner le classificateur discriminatif réseau de neurones, qui a permis d'obtenir les meilleurs résultats d'annotation parmi les classificateurs discriminatifs utilisés, avec le classificateur génératif réseau bayésien, réalisant le meilleur taux d'annotation, afin d'exploiter leur complémentarité.
- Améliorer la segmentation d'images.

Ceci, constitue l'objet des expériences 2, 3 et 4. Les résultats seront présentés et analysés pour enfin ressortir la structure finale du système d'annotation automatique d'images le plus performant.

#### **5.1.2.2. Expérience 2 : *Combinaison des descripteurs***

Récemment, afin d'améliorer la qualité de l'annotation d'images, plusieurs travaux de recherche [52-64] se sont attachés à la résolution du problème d'extraction, à partir des caractéristiques visuelles de bas niveau, des concepts sémantiques qui permettent une interprétation correcte du contenu d'une image (objets, thèmes, événements). Ils ont montré qu'aucun descripteur n'est suffisant pour réduire la diversité intra-classe et la corrélation inter-classes d'une manière efficace, et que pour avoir une annotation d'images performante, les procédés d'analyse doivent intégrer différents types de descripteurs. Par conséquent, une combinaison de divers descripteurs est devenue une perspective désirable et prometteuse pour réduire d'avantage le fossé sémantique. Ceci corrobore nos remarques issues de l'analyse des résultats obtenus au niveau de la première expérience.

Motivé par ce fait, nous nous sommes intéressés à chercher une méthode efficace permettant la combinaison des descripteurs complémentaires pour améliorer l'annotation

sémantique d'images. Afin d'atteindre cet objectif, nous avons étudié, dans cette deuxième expérience, deux approches qui consistent à combiner trois différents types de caractéristiques visuelles complémentaires à savoir : les moments de Legendre, les indices de Haralick et les histogrammes de couleur.

### 5.1.2.2.1. La première approche

Dans cette approche expérimentale, les descripteurs considérés sont concaténés en un seul vecteur descripteur. Ce vecteur sera l'entrée de l'un des classificateurs adoptés: le réseau bayésien naïf ou le réseau de neurones. Le principe de la combinaison par concaténation des descripteurs est illustré sur la figure 12.

#### 5.1.2.2.1.1. Résultats

Les taux d'annotation obtenus, par concaténation des descripteurs considérés, sont donnés dans le Tableau 2.3.

Les matrices de confusion d'annotation automatique d'images, basée sur la concaténation des descripteurs en utilisant le réseau bayésien naïf, sont illustrées par les figures 13 et 14 pour des images de la base de données ETH-80 et NATURE.

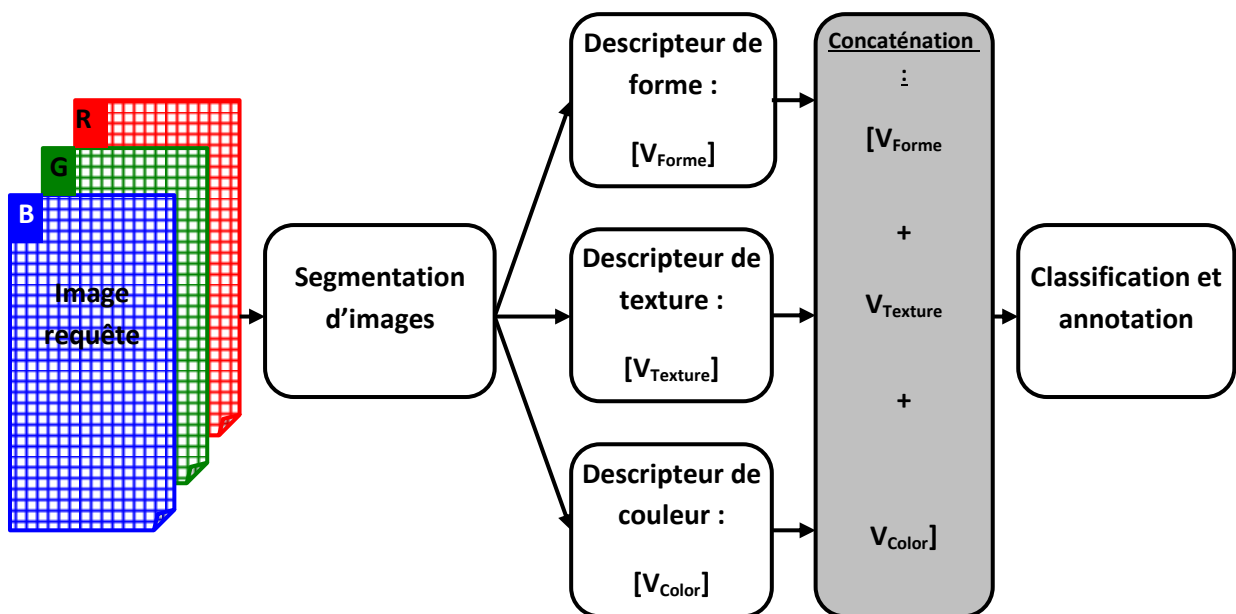


Figure 12 : Principe de la concaténation des descripteurs considérés

Tableau 2.3: Taux d'annotation obtenus en concaténant les descripteurs.

Base de données	Approche de classification	Taux d'annotation	Taux d'Erreur
ETH-80	réseaux de neurones	85%	15%
	réseaux bayésiens	88%	12%
COIL-100	réseaux de neurones	80%	20%
	réseaux bayésiens	83%	17%
NATURE	réseaux de neurones	87%	13%
	réseaux bayésiens	90%	10%

	Mots clés prédits							
	apple	car	cow	cup	dog	horse	pear	tomato
apple	5	0	0	0	0	0	0	0
car	0	4	1	0	0	0	0	0
cow	0	2	3	0	0	0	0	0
cup	1	0	0	4	0	0	0	0
dog	0	1	0	0	4	0	0	0
horse	0	0	0	0	0	5	0	0
pear	0	0	0	0	0	0	5	0
tomato	1	0	0	0	0	0	0	4

Figure 13: Matrice de confusion d'annotation automatique d'images basée sur la concaténation des descripteurs en utilisant le réseau bayésien naïf et la base ETH-80.

	Mots clés prédits					
	forest	gazon	ground	sahara	sky	water
forest	5	0	0	0	0	0
gazon	0	4	1	0	0	0
ground	0	1	4	0	0	0
sahara	1	0	0	4	0	0
sky	0	0	0	0	5	0
water	0	0	0	0	0	5

Figure 14: Matrice de confusion du système d'annotation automatique d'images basé sur la concaténation des descripteurs en utilisant le réseau bayésien naïf et la base de données NATURE.

### 5.1.2.2.2. Deuxième approche

Dans cette approche expérimentale, les trois types de descripteurs sont combinés en parallèle avec trois classificateurs de même nature, c'est-à-dire que chaque type de descripteur constitue l'entrée unique d'un classificateur qui vote pour un mot clés. La décision de l'annotation est prise en choisissant le mot-clé ayant le maximum de votes. Les classificateurs utilisés sont les réseaux bayésiens naïfs et les réseaux de neurones.

Le principe de combinaison des descripteurs est illustré sur la figure 15.

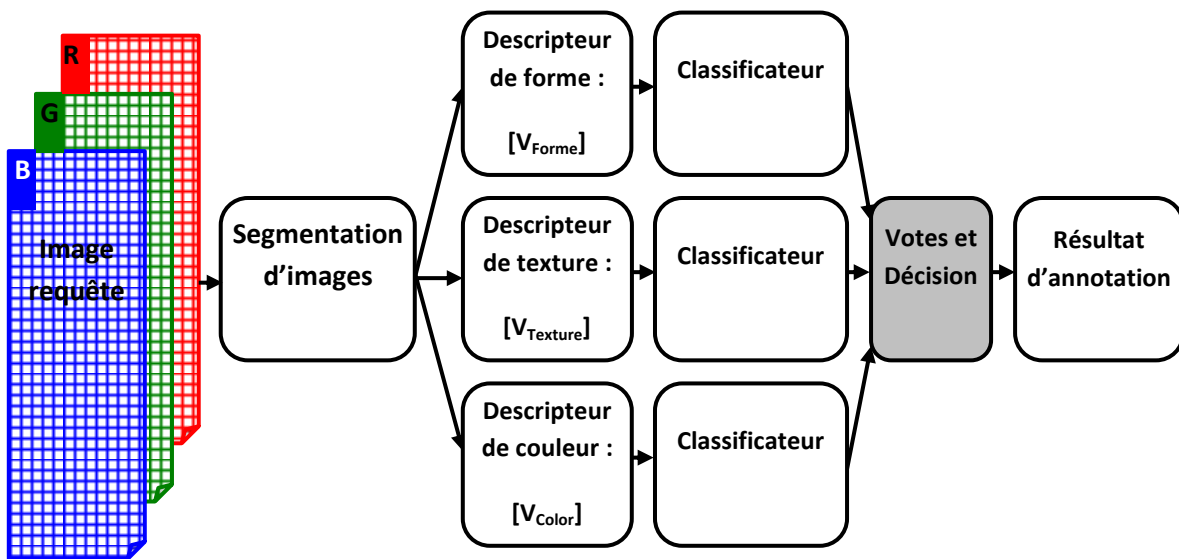


Figure 15: Principe de combinaison des descripteurs en parallèle

#### 5.1.2.2.2.1. Résultats

Les résultats d'annotation obtenus en combinant les descripteurs en parallèle sont donnés par le Tableau 2.2.

Les matrices de confusion de l'annotation automatique d'images, obtenues dans cette expérience en utilisant les réseaux bayésiens naïfs comme classificateurs, sont illustrées par les figures 16 et 17 pour les images des bases de données ETH-80 et NATURE.

Tableau 2.4 : Résultats d'annotation en combinant en parallèle les descripteurs.

Base de données	Approche de classification	Taux d'annotation	Taux d'Erreur
ETH-80	réseaux de neurones	88%	12%
	réseaux bayésiens	90%	10%
COIL-100	réseaux de neurones	83%	17%
	réseaux bayésiens	85%	15%
NATURE	réseaux de neurones	90%	10%
	réseaux bayésiens	93%	7%

	Mots clés prédits							
	apple	car	cow	cup	dog	horse	pear	tomato
apple	5	0	0	0	0	0	0	0
car	0	4	1	0	0	0	0	0
cow	0	2	3	0	0	0	0	0
cup	0	0	0	4	0	0	0	1
dog	0	0	0	0	5	0	0	0
horse	0	0	0	0	0	5	0	0
pear	0	0	0	0	0	0	5	0
tomato	0	0	0	0	0	0	0	5

Figure 16: Matrice de confusion du système d'annotation automatique d'images basé sur la combinaison en parallèle des descripteurs en utilisant les réseaux bayésiens naïfs et la base de données ETH-80.

	Mots clés prédits					
	forest	gazon	ground	sahara	sky	water
forest	5	0	0	0	0	0
gazon	0	4	0	0	1	0
ground	0	0	5	0	0	0
sahara	0	0	0	4	1	0
sky	0	0	0	0	5	0
water	0	0	0	0	0	5

Figure 17: Matrice de confusion d'annotation automatique d'images basée sur la combinaison en parallèle des descripteurs en utilisant les réseaux bayésiens naïfs et la base de données NATURE.

#### **5.1.2.2.2. Analyse des résultats et conclusion**

L'examen des résultats donnés par les Tableaux 2.3 et 2.4 et les Figures 13,14 et 16, 17 montre que :

- La combinaison, par concaténation ou en parallèle, des descripteurs permet d'améliorer le taux d'annotation d'images pour chacun des classificateurs utilisés. Ce résultat a été prévisible, car les classificateurs ont saisi plus d'informations pour distinguer les objets et par la suite les classifier.
- Le taux d'annotation obtenu par la combinaison en parallèle des descripteurs est meilleur que celui obtenu par la concaténation. Ce taux peut être amélioré davantage en augmentant le nombre de descripteurs combinés en parallèle sans rendre les résultats inexploitable en raison d'un vecteur de caractéristiques trop important. Dans le cas de la combinaison par concaténation, ceci n'est pas évident, car la taille du vecteur des descripteurs concaténés devient très grande.

D'après ces remarques, il apparaît que la combinaison en parallèle des descripteurs est la structure la plus performante pour étudier, dans l'expérience 3, la combinaison, signalée dans l'expérience 1, des réseaux de neurones avec les réseaux bayésiens.

#### **5.1.2.3. Expérience 3 : *Combinaison des classificateurs***

Partant des remarques dégagées au niveau des deux expériences précédentes, et de celles inspirées de l'état d'art, nous avons combiné, dans cette expérience, le pouvoir discriminatif des réseaux de neurones et le caractère génératif des réseaux bayésiens afin de tirer bénéfice de leur complémentarité.

Le principe de cette combinaison est illustré par le schéma bloc représenté par la figure 18. Ainsi, avec la combinaison en parallèle des trois types de descripteurs utilisés dans l'expérience précédente et des deux types de classificateurs considérés, on aura un maximum de vote égal à  $3 \times 2 = 6$ . Chaque classificateur vote pour un mot-clé. Le mot-clé ayant un maximum de votes sera considéré comme le mot convenable pour l'annotation d'un objet contenu dans une image requête.

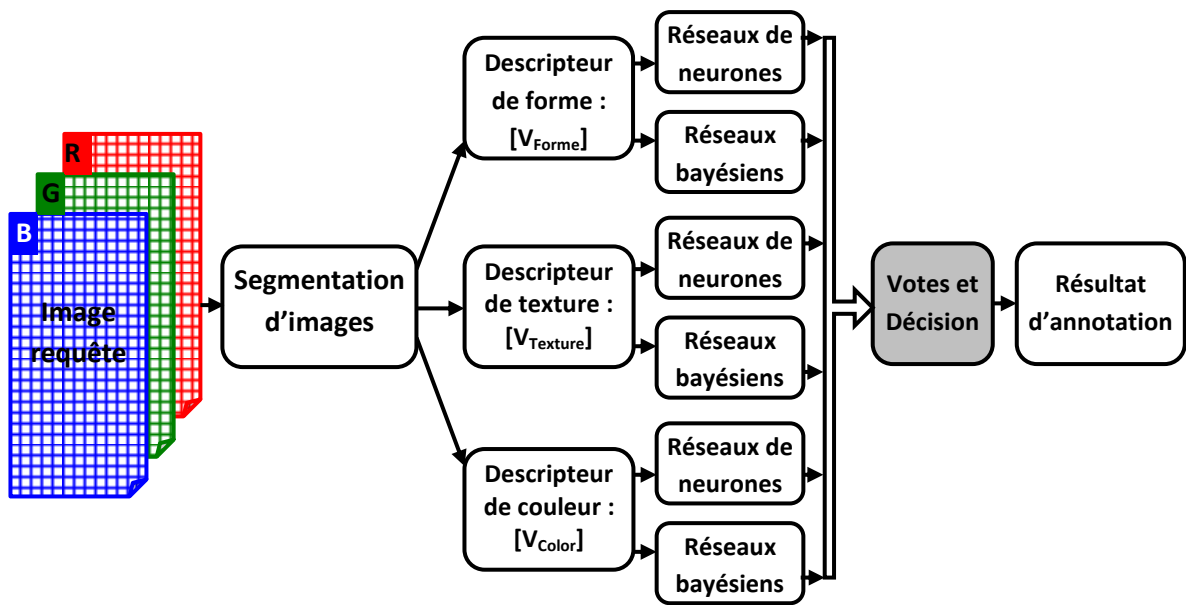


Figure 18: Schéma bloc du système d'annotation automatique d'images basé sur la combinaison des classificateurs et descripteurs en parallèle.

### 5.1.2.3.1. Résultats

Les taux et les matrices de confusion d'annotation d'images des trois bases de données d'images ETH-80 et NATURE, obtenus dans cette expérience, sont donnés respectivement par le Tableau 2.5, et les figures 19 et 20.

Tableau 2.5: Résultats de la combinaison des descripteurs et classificateurs en parallèle.

Base de données	Taux d'annotation	Taux d'Erreur
ETH-80	93%	7%
COIL-100	88%	12%
NATURE	97%	3%



		Mots clés prédits						
		apple	car	cow	cup	dog	horse	pear
apple	5	0	0	0	0	0	0	0
car	0	4	1	0	0	0	0	0
cow	0	2	3	0	0	0	0	0
cup	0	0	0	5	0	0	0	0
dog	0	0	0	0	5	0	0	0
horse	0	0	0	0	0	5	0	0
pear	0	0	0	0	0	0	5	0
tomato	0	0	0	0	0	0	0	5

Figure 19: Matrice de confusion d'annotation des images de la base de données ETH-80.

		Mots clés prédits					
		forest	gazon	ground	sahara	sky	water
forest	5	0	0	0	0	0	
gazon	0	5	0	0	0	0	
ground	0	0	5	0	0	0	
sahara	0	0	0	5	0	0	
sky	0	0	0	0	5	0	
water	0	0	1	0	0	4	

Figure 20: Matrice de confusion d'annotation des images de la base de données NATURE.

### 5.1.2.3.2. Analyse des résultats et conclusion

Nous constatons, d'après les résultats présentés par le Tableau 2.5 et les Figures 19 et 20, que la combinaison des réseaux de neurones avec les réseaux bayésiens, a permis d'améliorer nettement la qualité de l'annotation d'images. Ceci montre que l'exploitation de la complémentarité des approches discriminatives et génératives de la classification, permet effectivement d'augmenter la précision de l'annotation automatique des images.

Cependant, nous avons constaté qu'il y'a des erreurs d'annotation qui persistent à savoir en particulier la confusion entre "car" et "cow" (Figure 19). Partant du fait que cette confusion est due à la ressemblance entre les régions obtenues par la segmentation de ces objets, nous nous sommes intéressés au développement d'une méthode basée sur le regroupement des régions adjacentes afin d'avoir des objets sémantiquement compacts en vue

de faciliter leur classification et augmenter ainsi le taux d'annotation automatique d'images. Ceci fera l'objet de l'expérience 4.

#### **5.1.2.4. Expérience 4 : regroupement des régions des objets**

La segmentation doit permettre d'obtenir, à partir d'une image originale, une partition de régions homogènes, qui servira de base pour l'annotation automatique d'images. Cependant, les régions résultantes d'une segmentation automatique d'images ne représentent pas tout à fait les objets réels contenus dans les images. En effet, les méthodes de segmentation automatiques d'images utilisent souvent des agrégats et prédicats de bas niveau qui ne permettent pas de garder la compacité des objets composés de plusieurs couleurs. Ceci peut conduire à des éventuelles confusions entre les objets et par conséquent à des annotations erronées. Afin de remédier à cette problématique, nous avons proposé et mis en œuvre, dans cette quatrième expérience, une approche permettant de regrouper et fusionner les régions adjacentes, obtenues par la segmentation, afin de trouver des objets compacts (Figure 21). Cette approche consiste à chercher, à partir d'un ensemble de régions, toutes les combinaisons possibles de regroupement et les classer par ordre de probabilité d'annotation. Les regroupements, ayant les plus grandes probabilités d'annotation, représentent probablement des objets compacts dans l'image. Ces regroupements sont gardés dans la liste des mots-clés d'annotation puisqu'ils constituent les meilleurs candidats pour l'annotation automatique de l'image requête. Il ne s'agit donc pas de trouver la meilleure méthode de segmentation possible, mais de mettre en place une solution simple et efficace qui permet de trouver les différents objets sémantiques de l'image.

La Figure 23 représente la structure du système d'annotation automatique d'images intégrant l'approche proposée pour le regroupement des régions adjacentes.

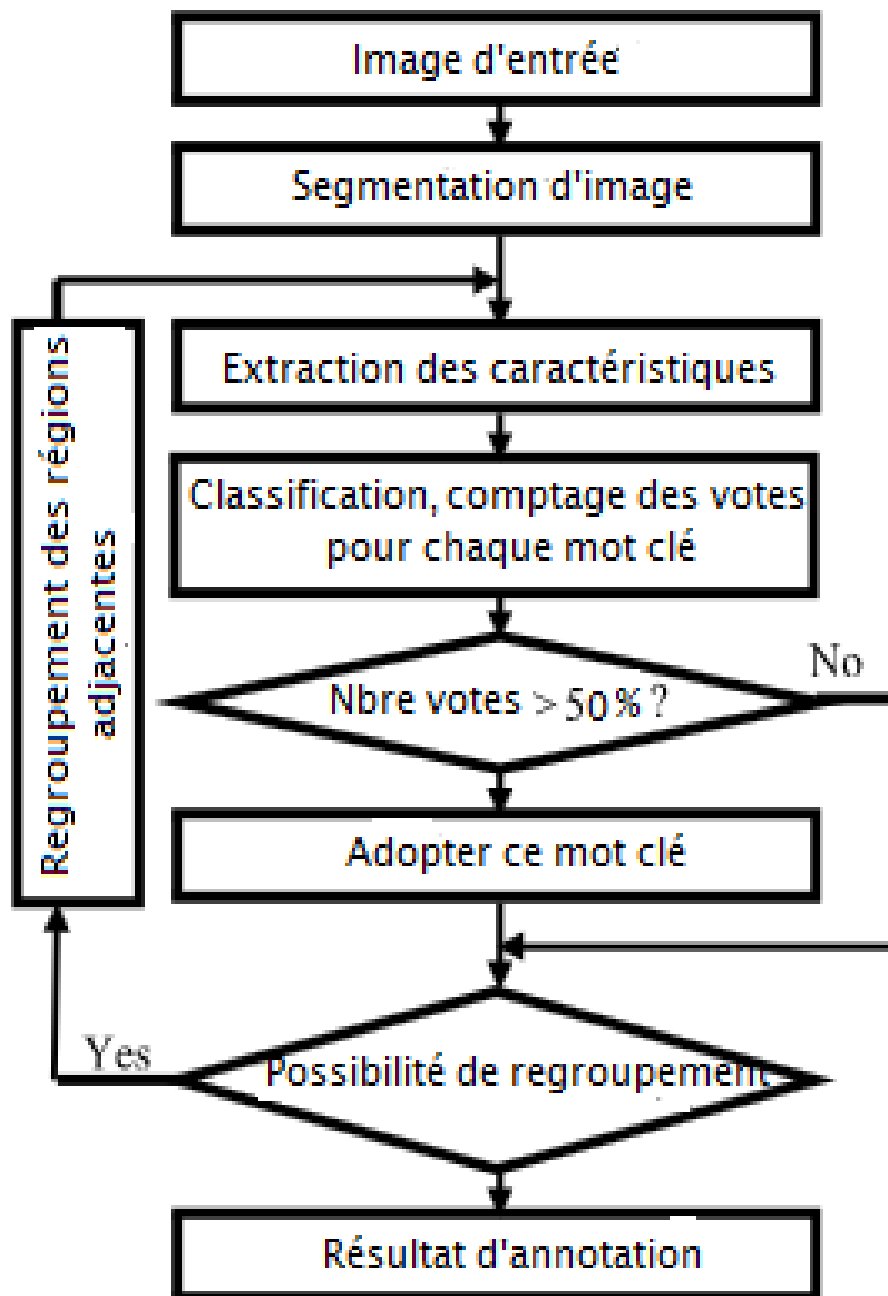


Figure 21: Exemple de segmentation et regroupement des clusters d'une image.

### 5.1.2.4.1. Résultats

Les taux et les matrices de confusion d'annotation d'images des trois bases de données d'images ETH-80, obtenus dans cette expérience, sont donnés respectivement par le Tableau 2.6 et la figure 22.

Tableau 2.6: Résultats obtenu en utilisant le regroupement des régions adjacentes

Base de données	Taux d'annotation	Taux d'Erreur
ETH-80	98%	2%
COIL-100	93%	7%

	Mots clés prédits							
	apple	car	cow	cup	dog	horse	pear	tomato
apple	5	0	0	0	0	0	0	0
car	0	4	1	0	0	0	0	0
cow	0	0	5	0	0	0	0	0
cup	0	0	0	5	0	0	0	0
dog	0	0	0	0	5	0	0	0
horse	0	0	0	0	0	5	0	0
pear	0	0	0	0	0	0	5	0
tomato	0	0	0	0	0	0	0	5

Figure 22: Matrice de confusion dans le cas d'utilisation du regroupement des régions adjacentes pour ETH-80.

### 5.1.2.4.2. Analyse des résultats et conclusion

Les résultats obtenus, présentés par le Tableau 2.6 et la figure 22, montrent l'utilité de regroupement des régions adjacentes pour augmenter le taux d'annotation. Nous pouvons voir en particulier à partir de la figure 22 que la confusion, entre les objets « car » et « cow », est réduite. Ceci montre bien l'efficacité du regroupement de régions adjacentes lors du processus d'annotation automatique d'images. En effet, ce regroupement permet d'avoir des régions compactes qui peuvent être annotées correctement, d'une manière plus efficace que l'annotation des régions résultantes directement d'une segmentation automatique.

## 5.2. Structure du système d'annotation réalisé

La structure finale adoptée pour le système d'annotation automatique d'images est représentée par la figure 23. Cette structure, finalisée par les études expérimentales réalisées au cours de ce travail, intègre, en plus de la combinaison des descripteurs et des classificateurs en parallèle, les regroupements des régions.

Des exemples d'annotation automatique d'images sont présentés par les figures 24 et 30 en utilisant l'interface graphique. Nous constatons que les objets, qui se trouvent dans les images de ces figures, sont bien annotés par leurs mots clés.

Il faut noter que nous n'avons pas comparé les performances prometteuses du système d'annotation automatique d'images que nous avons réalisé avec celles d'autres systèmes de l'état d'art, car les images, sur lesquels les expériences sont réalisées, sont pré segmentées et différentes de celles utilisées dans notre travail.

Néanmoins, la structure modulaire, du système réalisé, permet l'amélioration de ses performances, en particulier lors du passage à l'échelle, par l'intégration d'autres outils d'analyse et/ou de classification.

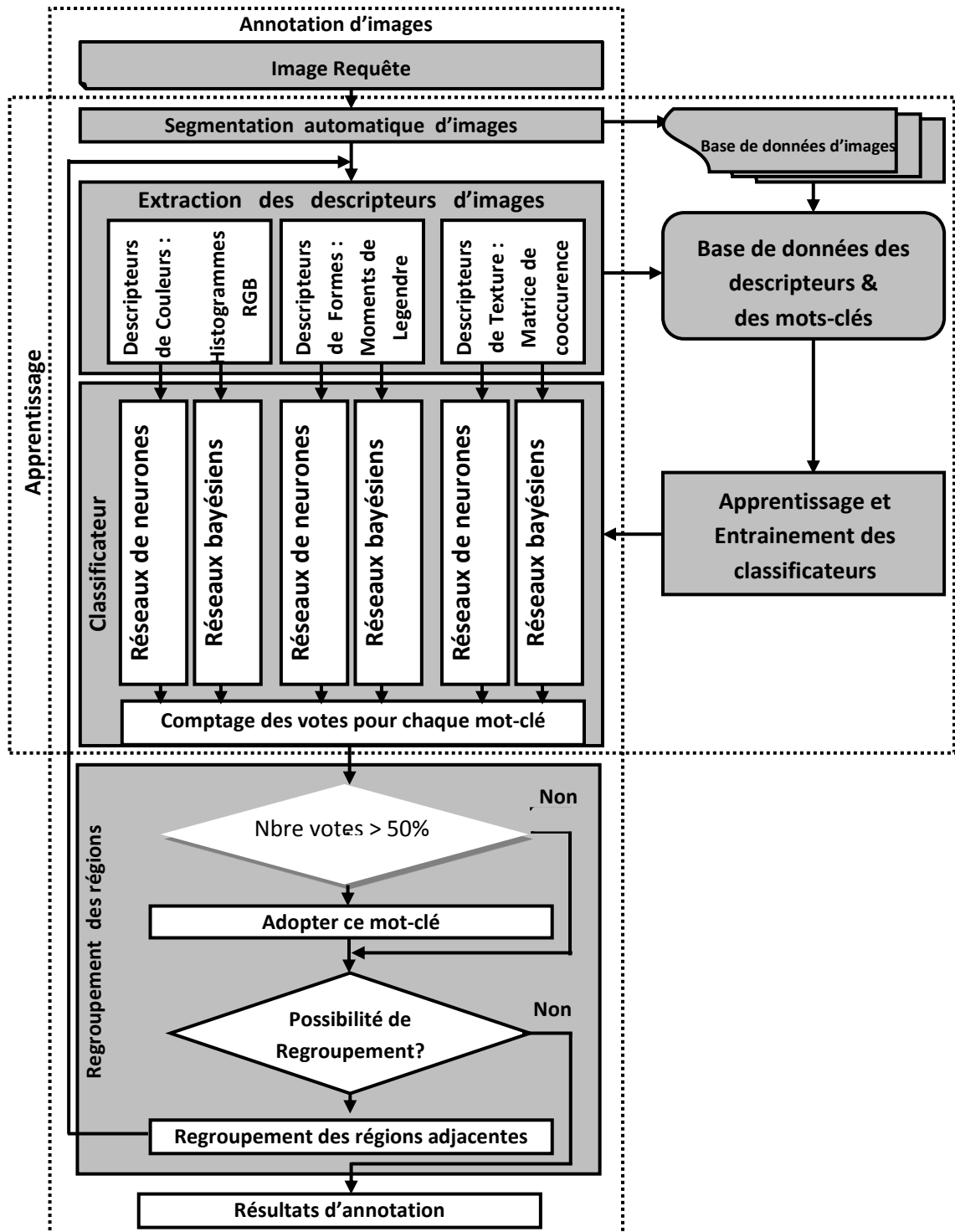
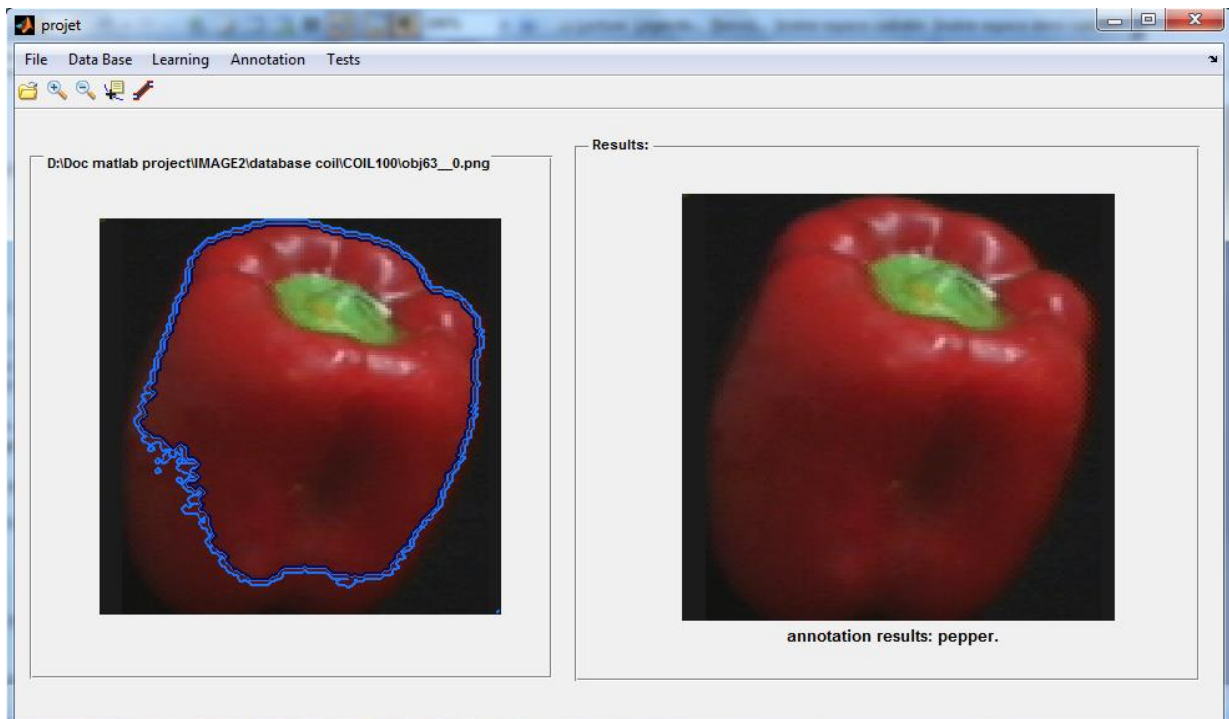
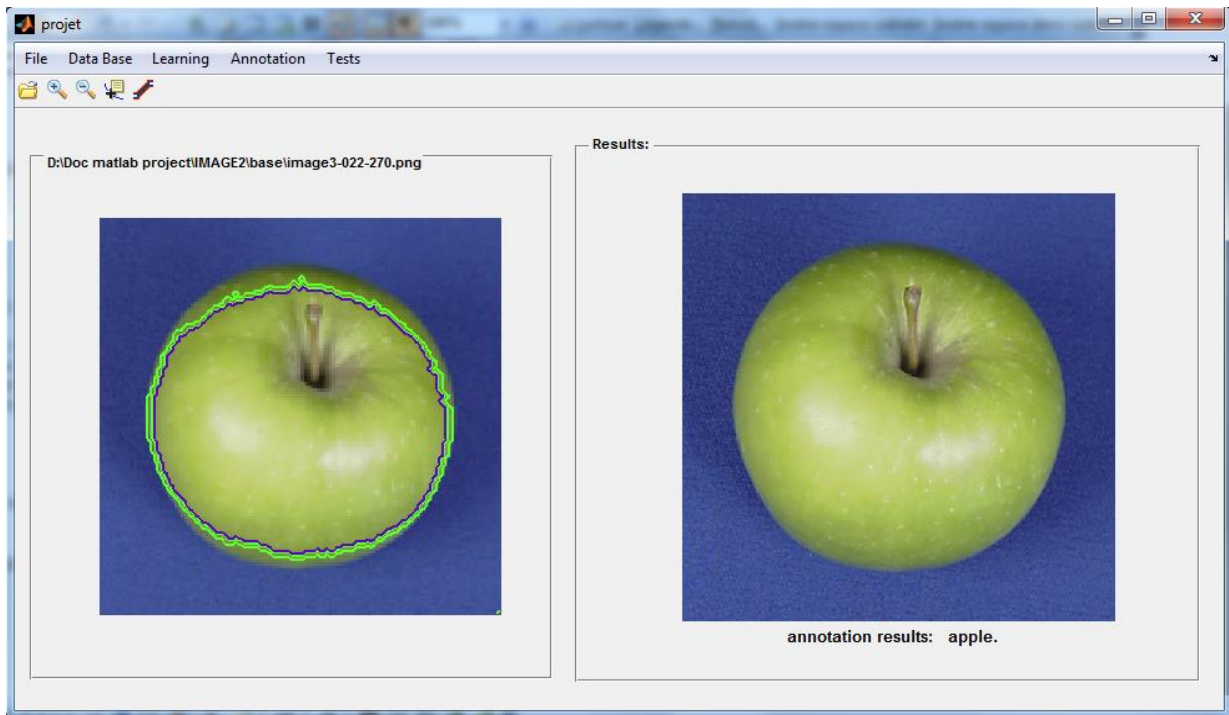


Figure 23: Structure finale du système d'annotation automatique d'images réalisé



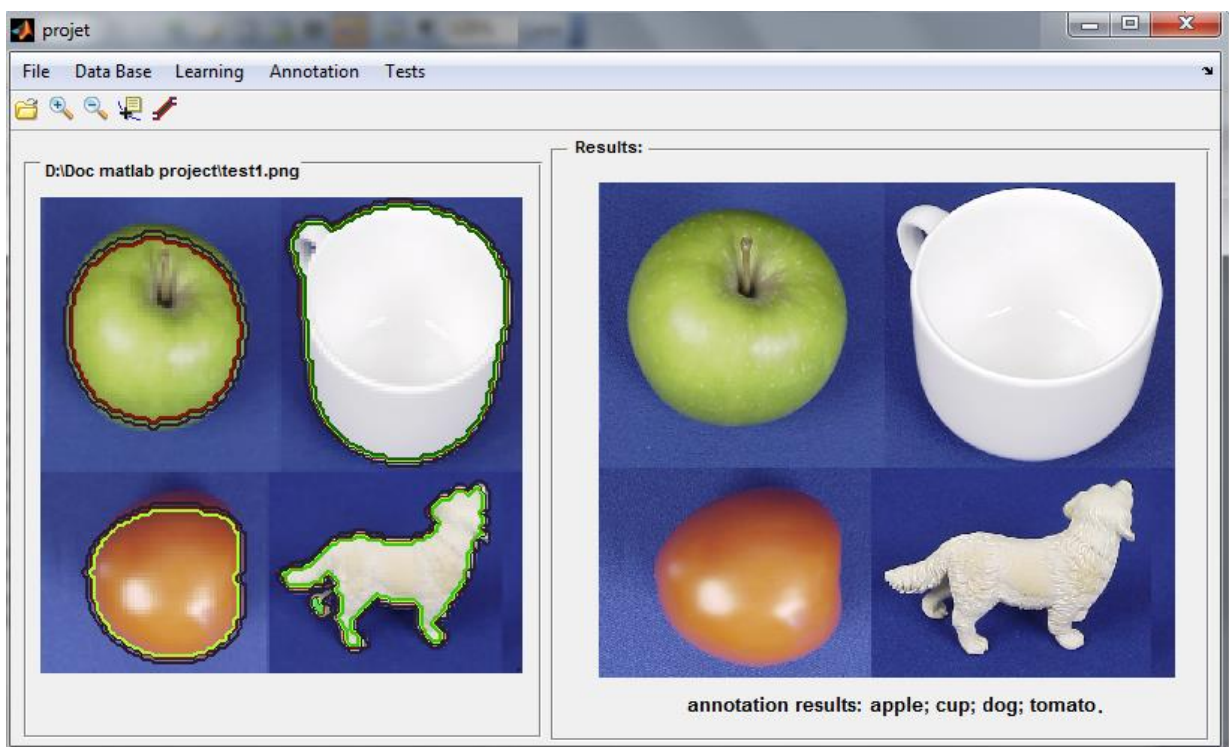
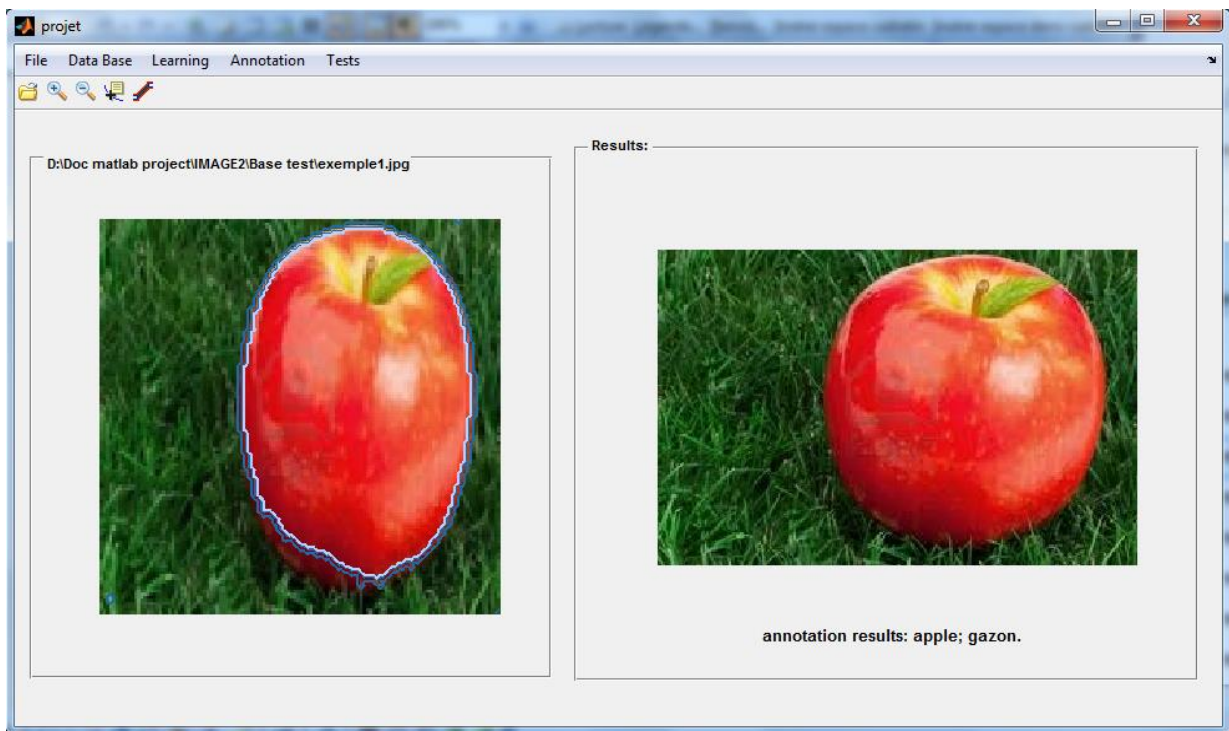


Figure 24: Exemple de résultats d'annotation d'images.



## **7. Conclusion générale et perspectives**

### **6.1 Conclusion générale**

Dans ce travail, nous nous sommes intéressés à la conception et la réalisation d'un système d'annotation automatique d'images permettant de réduire le fossé sémantique existant entre la description visuelle et la sémantique associée à une image. Ainsi, nous nous sommes penchés, au début de ce travail, sur l'étude bibliographique concernant l'annotation d'images. Cette étude nous a permis de concevoir la structure du système à réaliser. Pour rendre ce système opérationnel, nous avons implémenté plusieurs algorithmes d'extraction de paramètres descriptifs du contenu visuel des images, de classification et de segmentation. Ces algorithmes sont appliqués, dans un premier temps, à la reconnaissance des caractères Tifinagh. Cette application nous a permis de mettre en lumière le comportement de chaque descripteur et de chaque classificateur ainsi que leur combinaison. Ensuite, afin de mettre en œuvre le système d'annotation, nous les avons utilisés pour réaliser plusieurs expériences d'annotation d'images de trois bases de données images ETH-80, COIL-1000 et NATURE. L'analyse des résultats obtenus, a mis en évidence la possibilité d'améliorer le taux d'annotation en combinant par concaténation ou en parallèle les descripteurs de forme, de texture et de couleur, et en combinant les réseaux de neurones avec les réseaux bayésiens. Pour explorer cette piste, nous avons commencé par réaliser des expériences d'annotation en combinant par concaténation ou en parallèle les descripteurs pour chacun des classificateurs mentionnés ci-dessus. L'étude des résultats émanant de ces expériences, a montré que les deux procédures, c'est-à-dire la combinaison par concaténation ou en parallèle des descripteurs, permettent effectivement d'améliorer le taux d'annotation et que la combinaison en parallèle surpasse la combinaison par concaténation. Ainsi, nous avons opté par la suite à la combinaison des descripteurs en parallèle pour réaliser d'autres expériences d'annotation en combinant les deux classificateurs considérés. Les résultats de ces expériences ont révélé une nette amélioration du taux d'annotation et que l'un des facteurs qui engendre les erreurs qui persistent, réside dans la segmentation. Pour remédier à ce problème, nous avons recouru au regroupement des régions de couleurs obtenues après la segmentation. Les expériences faites en appliquant cette technique ont amélioré davantage le taux d'annotation.

Ces études, c'est-à-dire ces expériences et ces analyses, que nous avons menées au cours de ce travail, nous ont conduits à réaliser un système doté d'une architecture modulaire et performante en terme d'annotation d'images. Cette architecture originale s'articule autour :

- de la combinaison du pouvoir discriminatif des réseaux de neurones avec le caractère génératif des réseaux bayésiens. Ce qui a permis de tirer bénéfice de leur complémentarité. La prise de la décision d'annotation est réalisée par le vote des différents classificateurs combinés.
- de la combinaison des descripteurs de forme (les moments de Legendre), de texture (indices de Haralick) et de couleur (les histogrammes de couleur RGB).
- du regroupement des régions d'un objet après la segmentation par croissance des régions.

Le système ainsi réalisé est un système modulaire, évolutif et non figé. Il permet l'intégration d'autres descripteurs du contenu visuel des images et d'autres méthodes de segmentation pour réduire davantage le fossé sémantique en particulier lorsqu'on passe à l'annotation de larges bases de données d'images.

Une évaluation du système proposé est réalisée à partir d'un corpus d'images différent de celui utilisé au cours de l'apprentissage. L'évaluation a montré que notre système permet d'obtenir des résultats probants.

En conclusion, la principale contribution de cette thèse réside dans la proposition d'une architecture originale permettant la réalisation d'un système d'annotation automatique d'images performant.

## **6.2. Perspectives**

A cours terme, nous envisageons passer à l'échelle, c'est-à-dire utiliser notre système pour annoter les images de larges bases. Dans ce cas, nous aurons une large variabilité visuelle intra-concept et une grande similarité visuelle inter-concept, qui conduisent souvent à des annotations imparfaites. Ainsi, pour maintenir une bonne performance en terme du taux d'annotation, nous comptons explorer les pistes d'amélioration suivantes :

- Intégrer d'autres descripteurs qui pourraient s'avérer utiles afin de perfectionner la description du contenu visuel des images,
- Améliorer la segmentation afin d'extraire correctement les objets présents dans une image complexe,
- Inclure une méthode de désambiguïsation basée sur une mesure de similarité sémantique qui intègre plusieurs sources d'informations : visuelle, contextuelle et spatiale.

Après cette étape d'amélioration dudit système, nous procéderons à:

- Sa mise en ligne sur le Web en utilisant la plateforme Java pour qu'il soit à la disposition d'une large communauté ;
- Son utilisation pour développer un système de recherche d'image par mots clés.

## **Chapitre 3**

### ANALYSE SEMANTIQUE D'IMAGES POUR L'IDENTIFICATION DES PANNEAUX DE SIGNALISATION ROUTIERE

#### **1. Contexte et motivation**

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. C'est pourquoi, les gouvernements multiplient stratégies et méthodes pour réduire les accidents de la circulation en sensibilisant les conducteurs aux bienfaits du respect du code de la route. Cependant, malgré ces efforts, les accidents de la route persistent. Pour faire face à cette problématique, des équipes de recherche se sont penchées sur le développement des systèmes avancés d'aide à la conduite. L'objectif applicatif précis pour l'industriel est le prototypage d'un système de détection et reconnaissance des panneaux ("Traffic Sign Recognition", TSR), afin de proposer un système capable d'aider le conducteur à respecter le code de la route en l'informant en temps réel de la signalisation via des messages visuels et/ou sonores ce qui attire son attention sur son environnement routier et augmente sa sécurité et la sécurité des usagers de la route. Le grand défi lié à cette application est l'extrême variabilité des conditions d'éclairage et de contraste ; à ceci vient s'ajouter le potentiel vieillissement et dégradation des panneaux, ainsi que leur occlusion partielle possible. Ainsi, obtenir un système de détection et reconnaissance en temps-réel robuste et fiable en toute circonstance s'avère un problème difficile qui requiert l'usage des techniques d'analyse sémantique d'images avancées pour la détection et la reconnaissance de formes. La recherche et la mise en œuvre de telles techniques pour réaliser un système d'aide à la conduite robuste est rapide, nous a motivé à mener un travail de recherche sur la détection et la reconnaissance des panneaux de signalisation routière.

#### **2. Etat de l'art**

La majorité des travaux sur l'identification des panneaux de signalisation routière, se focalisent sur la détection. Les méthodes de détection, proposées dans la littérature, peuvent être classées en deux catégories selon la caractéristique visuelle utilisée pour détecter les panneaux: couleur ou forme.

L'information couleur est souvent exploitée afin de localiser la position des panneaux dans l'image. Dans [65], et [66] les régions d'intérêt, censées contenir un panneau, sont extraites par un simple seuillage de l'image couleur RGB. Cependant, l'espace RGB est considéré comme un espace sensible aux variations de lumière et donc déconseillé pour l'environnement à l'extérieur. [66] proposent un seuillage adaptatif aux conditions d'éclairage pour remédier à la sensibilité de RGB. Dans d'autres travaux, cet espace a été remplacé par des espaces plus robustes comme HSI (Hue-Saturation-Intensity , Teinte-Saturation-Intensité) ou HSV (Hue-Saturation-Value , Teinte-Saturation-Valeur) [67-69]. Le seuillage peut être appliqué sur l'ensemble des composantes de ces espaces comme dans [67] ou seulement sur les composantes chromatiques, teinte et saturation. Des espaces plus complexes comme LCH (Lightness, Chroma, Hue) [70] et CIE Lab [71] sont aussi utilisés.

D'autres approches, proposées dans la littérature, pour détecter les panneaux à travers la détection des régions des couleurs, rouges et bleues dans l'image, utilisent le résultat d'une classification de la couleur des pixels par un réseau de neurones [72] ou par un séparateur à vastes marges [73].

Cependant, bien que la couleur demeure un des aspects pertinents pour la détection des panneaux routiers, elle reste sensible aussi bien aux conditions d'éclairage, qu'au changement d'apparence dû à une altération de la couleur. Avec ces facteurs, la détection n'est jamais parfaite et sans erreur. C'est pourquoi, plusieurs équipes de recherche se sont intéressées à la détection des panneaux à l'aide de la forme des contours.

Le contour est considéré comme étant la caractéristique de bas niveau la plus utilisée après la couleur pour la détection des panneaux. Les panneaux sont souvent circulaires, rectangulaires ou triangulaires. La détection de ces formes dans les images peut donner de très bons indices pour la détection de panneaux. Cela est utilisé pour détecter les panneaux circulaires [74-77], triangulaires [74,78-81] , rectangulaire [79] et polygonales [82-83] Cependant, les occultations partielles rendent la détection des formes assez difficile, voir impossible dans certains cas. C'est pourquoi, dans d'autres travaux, les deux méthodes de détection par couleur et par contour de forme sont combinées en parallèle [84] ou en cascade [65].

Les zones d'intérêt, détectées dans chaque images, sont exploitées au niveau de la phase de la reconnaissance pour identifier le type de panneaux ou à supprimer les fausses réponses. Dans la littérature, La reconnaissance des panneaux est réalisée soit par

appareillement [85-88] de leurs illustrations avec une base de données de référence, soit par une classification [8,89-94] à base des réseaux de neurones ou les séparateurs à vastes marges.

Le présent état de l'art de la détection met en évidence le dynamisme de ce domaine. En effet, les nombreuses récentes publications témoignent de l'intérêt que suscite la détection auprès des chercheurs. Les récentes méthodes, proposées dans ce contexte, tentent d'augmenter le taux de détection tout en diminuant le nombre des fausses réponses. Cependant, dans les situations réelles la détection n'est jamais parfaite. Les mauvaises conditions de visibilité ainsi que les problèmes d'occultation rendent la tâche de la détection non-triviale. Toutefois, à ce jour, peu de travaux ont présenté un algorithme complet pour la détection et la reconnaissance de la signalisation verticale. Lorsque l'un d'eux atteint un taux de reconnaissance et de fausse détection acceptable, c'est souvent au prix d'un temps de calcul qui ne permet pas une identification en temps réel. C'est pourquoi, le développement des systèmes d'identification des panneaux routiers, rapide et ayant un fort taux de détection, est jusqu'à présent un sujet d'actualité.

### 3. Contribution

Dans ce travail, nous avons développé un système d'identification automatique de quatre catégories de panneaux de signalisation routière verticale par vision (Figure 3.1) : panneaux triangulaires rouges (Danger), circulaires rouges (Interdiction), circulaires bleus (Obligation) et rectangulaires bleus (Indication). L'approche adoptée dans ce travail est constituée de deux modules principaux: un module de détection qui se base sur l'attention visuelle pour repérer des zones de la scène susceptibles de contenir des panneaux routiers et un module de reconnaissance dont le rôle est de classifier les panneaux en fonction de l'information visuelle détectée.

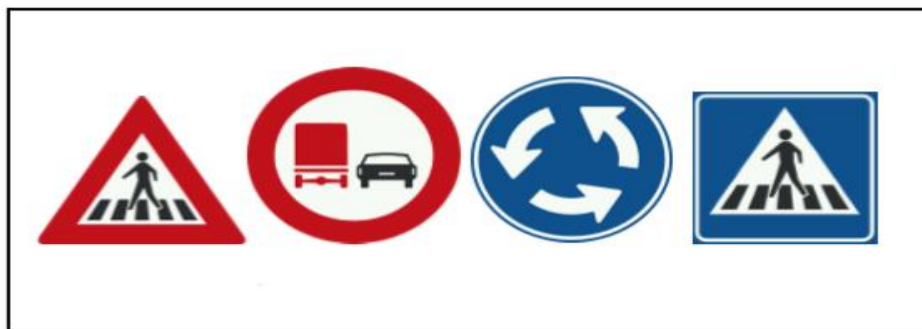


Figure 3.1 : Exemple des quatre catégories de panneaux considérés dans ce travail

### **3.1. Module de détection**

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. C'est pourquoi, les approches de l'état d'art, proposées dans ce contexte, se scindent en deux grandes catégories : segmentation par couleur ou segmentation par contours de 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.

#### **3.1.1. Méthode de détection**

La méthodologie adoptée pour la détection des panneaux routiers, est schématisée par la figure 3.2. 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.

##### **3.1.1.1. 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 fixé empiriquement après le traitement d'un grand nombre d'images en faisant un compromis entre sur-détections et sous-détections. La figure 3.3 montre un exemple d'image RGB et les masques de couleur bleu et rouge ainsi obtenus.







*Figure 3.3 : Résultats de détection de couleur. a) Image initiale, b) Masque de couleur bleu et c) Masque de couleur rouge*

### **3.1.1.2. 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.

#### **3.1.1.2.1. Extraction des contours**

Pour extraire les contours, nous avons utilisé le filtre de canny [95] 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.

#### **3.1.1.2.2. 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 [81,82] pour l'estimation de la forme triangulaire et/ou circulaire des panneaux,
- la transformation en Symétrie Radiale [76,77] qui est une variante de la transformée de Hough pour la détection de cercles,
- le détecteur de Polygones Réguliers [84] 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 [79,80] pour l'estimation du triangle,

- la transformée Chinoise Bilatérale [78] 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 [96,97], 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 3.4 (a) où le critère d'erreur d'approximation utilisé est la distance maximale entre les points du contour et le segment du polygone initiale. Pour simplifier l'explication, notons  $\widehat{AB}$  (resp.  $\overline{AB}$ ) la portion du contour (resp. le segment du polygone initiale 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 3.4 (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  $\varepsilon$  du paramètre entré de l'algorithme, que plus  $\varepsilon$  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  $\varepsilon$  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  $\varepsilon$  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 2.1 (a) est montré dans la figure 2.1 (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 2.1(b)). La répétition de ce processus continue tant que la distance maximale, entre le contour et le segment de polygone correspondant, est supérieure à  $\epsilon$  et il s'arrête dans le cas contraire (figure 3.4(c)).

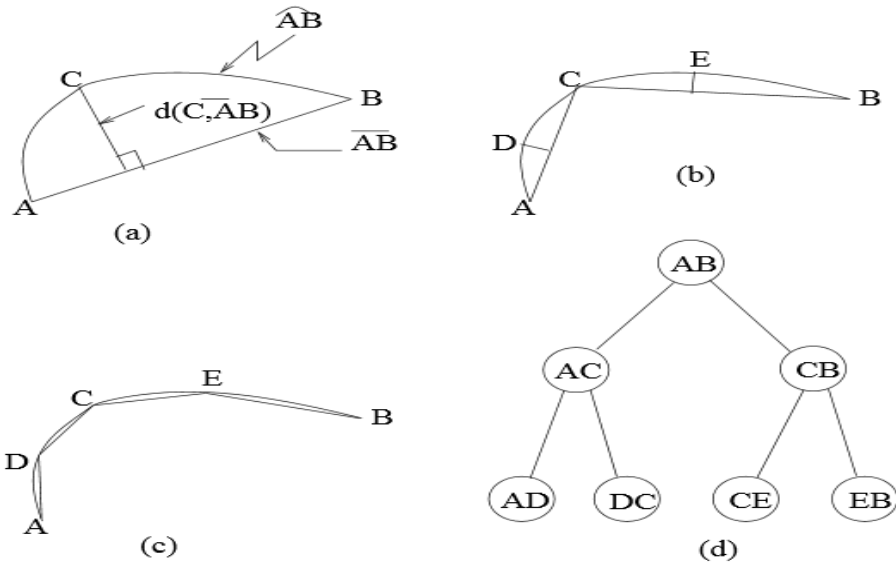


Figure 3.4 : Approximation polygonale du contour

Dans ce travail, l'Approximation polygonale est utilisée pour approximer un contour fermé par un polygone de N sommets dont les coordonnées sont bien déterminés. 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 tel que sa base est horizontale avec une erreur de tolérance limitée, 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 et ses segments opposés sont égaux, 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, 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 [98], 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 imagerie des panneaux ainsi obtenues, 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 3.5 illustre la détection des panneaux routiers selon la méthode de détection proposée.

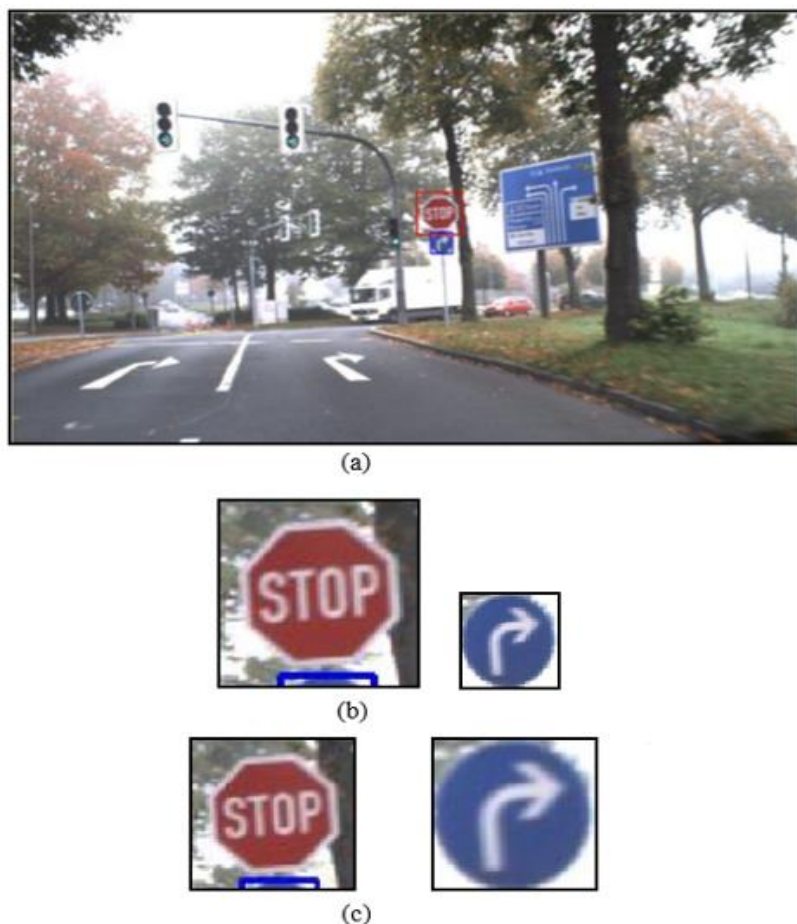


Figure 3.5 : Résultats de détection des panneaux. a) Image initiale, b) imagerie des panneaux détectés et rognés, c) imagerie redimensionnées en 32X32 pixels

### 3.1.2. 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 terme de précision et de rapidité de détection des panneaux routiers, nous avons utilisé des images contenant plusieurs panneaux avec des arrières plans différents.

#### 3.1.2.1. 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 3.1.

Tableau 3.1 : Résultat d'évaluation du module de détection

Classe des panneaux	Nombre de panneaux dans l'image	Nombre de panneaux détectés	Taux de détection (%)	Temps globale 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

#### 3.1.2.2. 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 3.2.

Tableau 3.2 : Résultat d'évaluation du module de détection

Classe des panneaux	Nombre de panneaux dans l'image	Nombre de panneaux détectés	Taux de détection (%)	Temps globale 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

### Analyse des résultats

D'après les Tableau 3.1 et 3.2, 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 m  $\approx$  16 ms x 33.33 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 m  $\approx$  16 ms x 16,67 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.

### 3.2. Module de la reconnaissance

Pour chaque classe des panneaux détectés au niveau du module de détection, le module de reconnaissance consiste à reconnaître le type de panneau, s'il s'agit d'un panneau, ou à le rejeter dans le cas inverse selon un intervalle de tolérance donné.

#### 3.2.1. Classification

Ce module est constitué par quatre réseaux de neurones permettant la classification des panneaux routiers. Chaque réseau de neurone [48] est utilisé pour identifier les panneaux de l'une des quatre classes des panneaux détectées en se basant sur les caractéristiques obtenues à partir des moyennes normalisées des valeurs des pixels de chaque ligne ( $CX_i$ ) et de chaque colonne ( $CY_j$ ) des imagerie de taille 32X32 des panneaux détectées. Ces caractéristiques sont ainsi calculées :

$$CX_i = \frac{1}{255} \frac{1}{32} \sum_{j=1}^{32} C_{i,j} \quad i=1,2,\dots,32 \quad (3.1)$$

$$CY_j = \frac{1}{255} \frac{1}{32} \sum_{i=1}^{32} C_{i,j} \quad j=1,2,\dots,32 \quad (3.2)$$

Avec  $C_{i,j}$  : intensité du pixel de coordonnées (i,j) de l'un des plans de couleur de l'imagerie RVB.

Le nombre de caractéristiques extraites pour chaque imagerie RVB, est donc 192.

La structure des réseaux de neurones utilisés est présentée par la figure 3.6.

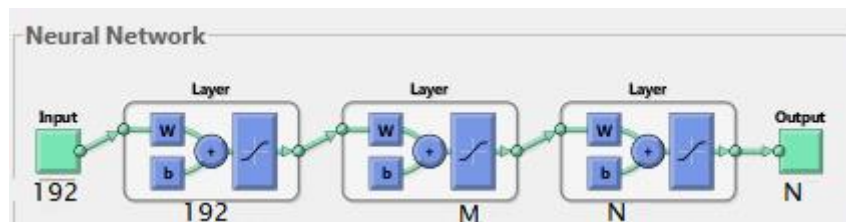


Figure 3.6 : Architecture du réseau de neurones utilisé

Avec M : le nombre de neurones de la couche cachée estimé à partir de l'équation (2.2)








et N : le nombre de neurones de la couche de sortie, égal au nombre des types de panneaux dans chaque classe.

La fonction d'activation utilisée par ce classificateur est la tangente hyperbolique définie par l'équation (2.1).









### 3.2.2. Classification

Les performances de ce module sont évaluées, pour chaque classe, sur les imagerie des panneaux fournies par le module de détection. La moitié du nombre de ces panneaux est utilisée pour entrainer les réseaux de neurones. Les Tableaux 3.3, 3.4 et 3.5 résument les résultats de cette évaluation.









**Tableau 3.3 : Reconnaissance des panneaux triangulaires rouges**

					
Nombre de panneaux de test	30	5	5	7	12
Nombre de panneaux reconnus	26	4	4	6	10
Taux de reconnaissance (%)	87	80	80	85	83

**Tableau 3.4 : Reconnaissance des panneaux circulaires rouges**

								
Nombre de panneaux de test	33	18	36	14	10	9	23	11
Nombre de panneaux reconnus	30	15	32	10	10	8	20	11
Taux de reconnaissance (%)	90	83	89	71	100	89	87	100

**Tableau 3.5 Reconnaissance des panneaux bleus**

								
Nombre de panneaux de test	29	6	5	6	7	11	18	6
Nombre de panneaux reconnus	25	5	4	5	6	9	16	4
Taux de reconnaissance (%)	86	83	80	83	86	82	89	67

D'après les résultats présentés dans ces Tableaux, il apparaît bel et bien que la méthode proposée pour la reconnaissance des panneaux routiers est probante. Elle permet d'atteindre un taux moyen de reconnaissance des panneaux des différentes classes de l'ordre de 85 %. Ce taux peut être amélioré en procédant à une suppression automatique de l'arrière plan des imagerie rognées (Figure 3.5 .c) à la sortie du module de détection avant d'en extraire les caractéristiques visuelles utilisées aussi bien dans l'apprentissage que dans la reconnaissance.

#### **4. Conclusion et perspectives**

Dans ce travail, nous avons proposé un système d'identification des panneaux de la signalisation verticale routière. Ce système intègre des techniques d'analyse sémantique appropriées permettant une détection et une reconnaissance prometteuses des panneaux routiers. La détection enchaîne 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.

Pour chaque classe des panneaux détectés, la reconnaissance du type des panneaux détectés, est réalisée par un réseau de neurones multicouches couplé à des descripteurs correspondants aux moyennes normalisées des valeurs des pixels de chaque ligne et de chaque colonne des images des pictogrammes des panneaux. Les résultats des tests des performances de cette méthode de reconnaissance, sont prometteurs.

Bien que les contributions apportées par ce travail présentent des résultats satisfaisants pour l'identification des panneaux routiers, plusieurs améliorations sont envisageables :

- ✓ Amélioration des performances de la méthode de détection proposée en présence d'occultation. En effet, en présence d'occultation les contours extraits par canny ne sont pas fermés et par conséquent ne sont pas pris en considération lors de la détection.

- ✓ Amélioration du taux de la reconnaissance des panneaux en utilisant :
  - d'autres signatures locales proposées dans l'état de l'art, comme la texture, pour discriminer les pictogrammes des panneaux d'une même classe.
  - d'autres classificateurs tels que les séparateurs à vastes marges, les arbres de décision et les réseaux bayésiens.
  
- ✓ Ajouter une étape de suivi des panneaux reconnus afin de ne pas les resigner au conducteur. En effet, les panneaux apparaissent sur plusieurs images avant leur disparition du champ de vision de la caméra ce qui permet leur suivi pour ne pas les reclassifier. Ceci entraîne une réduction du temps de calcul.

Un fois ces améliorations seront apportées au système d'identification proposé, nous envisageons l'implémenter sur un processeur spécialisé en traitement du signal (DSP) qui sera relié à une caméra de haute résolution et à un système de synthèse de la parole par phrase. L'embarquement d'un tel système sur les véhicules, permettrait de maintenir les conducteurs en information, en temps réel, de l'existence des panneaux dans la scène routière.

## **Chapitre 4**

### ANALYSE SEMANTIQUE D'IMAGES POUR L'IDENTIFICATION DES MALADIES DES AGRUMES

#### **1. Contexte et motivation**

Dès les années 90, l'agrumiculture s'est imposée comme le fleuron des exportations agricoles marocaines. Elle joue au Maroc un rôle socio-économique très important. Economiquement, les exportations d'agrumes constituent pour l'état une source importante de devises. Socialement, ce secteur constitue la source principale de revenus pour des milliers de familles de producteurs et procure, soit directement (vergers), soit indirectement (stations de conditionnement, ports et autres) plusieurs million de journées de travail.

C'est pourquoi le Plan Maroc Vert a fixé comme objectif de doubler la production d'agrumes, avec des perspectives de production aux alentours de 3,19 millions de tonnes annuellement à partir de 2020.

Ce secteur d'une importance socio- économique incontestable souffre d'un ensemble de problèmes dont le plus important réside dans les maladies ravageurs qui attaquent et détruisent les agrumes. La lutte contre ces maladies demeure un défi majeur pour les disciplines scientifiques de l'agronomie. Dans ce contexte, nous nous sommes proposés de mener un travail de recherche sur la détection et la reconnaissance automatique des maladies des agrumes en faisant appel à des techniques de traitement d'image et de l'intelligence artificielle.

#### **2. Etat de l'art**

Dans cette partie, nous présentons un bref aperçu sur les méthodes, de l'état de l'art, qui utilisent les techniques d'analyse d'image pour détecter et reconnaître les maladies de différentes plantes à partir d'images numériques. Bien que les symptômes d'une maladie peuvent se manifester dans toutes les parties de la plante, seules les méthodes qui explorent les symptômes visibles sur les feuilles pour identifier les maladies des plantes, sont présentées dans cette partie. Parmi Les méthodes récentes proposées dans ce contexte, nous citons :

La méthode proposée par Elham Omrani, et al [99]. Cette méthode consiste à convertir l'image RVB des feuilles malades des pommes en image CIE LAB pour en extraire la zone infectée par une segmentation basée sur le regroupement par K-moyennes (K-means clustering). A partir de cette zone, les caractéristiques de couleur, forme et texture sont déterminées et présentées à un classificateur à régression par les machines à vecteurs supports pour classer les maladies des pommes. Dans ce travail, les caractéristiques de texture utilisées sont les indices de Haralick, extrais à partir de la matrice de co occurrence, et les coefficients de la transformation en ondelettes.

Pugoy et Mariano [100] ont proposé une méthode pour identifier deux différents types de maladies qui attaquent les feuilles de riz. L'algorithme convertit d'abord l'image RVB en image HSI. La technique K-means clustering est appliquée pour regrouper les pixels en un certain nombre de groupes. Ces groupes sont ensuite comparés à une bibliothèque contenant les couleurs correspondant respectivement à ces maladies pour déterminer les régions affectées par chacune des maladies.

La méthode proposée par Abdullah et al. [101] permet de distinguer une maladie donnée (*Corynespora*) d'autres pathologies qui affectent les feuilles des arbres à caoutchouc. Une analyse en composantes principales est appliquée directement, sans segmentation, sur les valeurs des pixels d'une image RVB des feuilles. Les deux premières composantes principales sont ensuite envoyées à un réseau de neurones multicouches pour identifier le type de maladie.

Sanyal et al. [102], a abordé le problème de la détection et la classification de six types de carences en minéraux dans la culture de riz. Tout d'abord, l'algorithme extrait les caractéristiques de texture et de couleur à partir des images des feuilles de la plante de riz. Ensuite, chaque type de caractéristiques est soumis à son propre réseau de neurones spécifiques. Les deux réseaux ont une couche cachée, mais le nombre de neurones dans la couche cachée est différente (40 pour la texture et 70 pour la couleur). Les résultats renvoyés par les deux réseaux sont ensuite combinés, ce qui donne le classement final.

Kai et al. [103] ont proposé une méthode pour identifier trois types de maladies du maïs. Tout d'abord, les images RVB sont converties en image YCbCr. Puis, un seuillage est appliqué afin de segmenter les régions malades. Les auteurs extraient ensuite un certain nombre de caractéristiques de texture de la matrice niveau de gris de co-occurrence. Enfin, les

caractéristiques sont soumises à un réseau de neurones avec une couche cachée pour identifier les trois différentes maladies du maïs.

Wang et al. [104] ont proposé une méthode pour discriminer entre des paires de maladies du blé et de la vigne. Les images sont segmentées par un algorithme K-means clustering, puis 50 caractéristiques de couleur, forme et texture sont extraites. Au niveau de la classification, les auteurs ont testé quatre différents types de réseaux de neurones: Perceptron multicouches, fonction Radial de base, régression généralisée et probabilistes.

Huang [105] a proposé une méthode pour détecter et classifier trois types de maladies qui affectent les jeunes plantes d'orchidées Phalaenopsis. Après la segmentation de l'image des feuilles, les caractéristiques de couleur et de texture sont ensuite extraites de la matrice de cooccurrence, et sont soumises à un réseau neuronal artificiel, avec une couche cachée, pour la classification des maladies étudiées.

Pydipati et al.[106] ont proposé et comparé deux approches différentes pour détecter et classer trois types de maladies d'agrumes en utilisant les caractéristiques de texture de feuilles. La première approche est basée sur le classificateur K plus proches voisins en utilisant la distance de Mahalanobis. La seconde approche est basée sur un réseau de neurones à fonction radiale de base.

Jian et Wei [107] ont proposé une méthode pour reconnaître trois types de maladies foliaires du concombre. La séparation entre les régions saines et malades des feuilles est faite par une technique de seuillage. Par la suite, les caractéristiques de couleur, forme et texture sont extraites et soumises à un SVM, avec noyau à fonction radiale de base, pour effectuer le classement final.

Alham F. et Al. [108] ont proposé une méthode d'analyse d'images pour identifier trois maladies du palmier à huile. Cette méthode consiste à extraire six types de caractéristiques, à partir des images des feuilles des palmiers, et à les utiliser par un réseau de neurones pour la classification des maladies considérées.

En résumé, la stratégie générale, adoptée par les méthode proposées dans la littérature pour identifier les maladies dont les symptômes sont visible sur les feuilles des plantes, consistent à localiser la région infectée dans un premier temps, puis d'en extraire différentes caractéristiques de couleur, texture et/ou forme, permettant l'identification des maladies par

différentes techniques de classification telles que les réseaux de neurones, les séparateurs à vastes marges...etc .

### 3. Travail réalisé

La finalité ultime de ce sujet de recherche est la réalisation d'un système d'identification automatique des maladies les plus ravageurs des agrumes au Maroc. Un tel système permettrait aux agricultures de reconnaître et traiter les agrumes sans faire appel à un expert.

Afin d'atteindre cet objectif, nous avons élaboré et mis en œuvre une démarche expérimentale dont le principe est basé sur l'analyse de la texture des images des feuilles des agrumes. Ce principe est inspiré du procédé adopté par les experts pour identifier les différentes maladies des agrumes à partir de l'analyse des symptômes apparaissant sur les feuilles.

#### 3.1. Description de la démarche expérimentale développée

La structure générale, de la démarche développée pour la reconnaissance automatique des maladies des agrumes, est représentée par la figure 4.1.

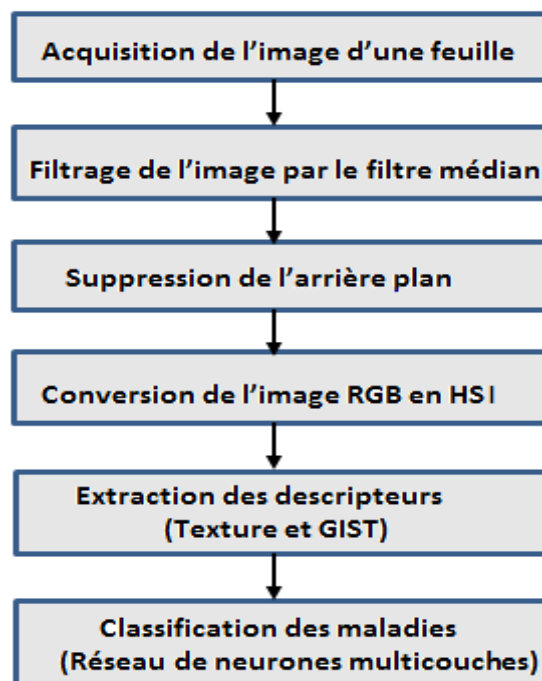


Figure 4.1 : structure générale de la démarche expérimentale développée

### 3.1.1. Acquisition

Pour évaluer la faisabilité et la capacité de cette démarche expérimentale à identifier les maladies des agrumes, nous l'avons testé sur une base de données images que nous avons conçue et développée en y répertoriant, à partir du sites web, 40 images des feuilles normales et 40 images pour chacune des maladies suivantes :

**Le Scab** : C'est une maladie fongique qui se localise sur les rameaux, les feuilles et les fruits (Figure 4.2). Elle se présente sous la forme de pustules liégeuses irrégulières. Cette maladie affecte les limettiers, pomelos et orangers.



*Figure 4.2 : Feuille et fruit atteints de Scab*

Une humidité excessive de la parcelle favorise le Scab. Les feuilles sont sensibles lors des poussées végétatives.

**La mélanose**: Cette maladie, originaire d'Amérique, attaque les feuilles à partir de juillet jusqu'à septembre, en formant de nombreuses petites taches visibles sur les deux faces du limbe (Figure 4.3.). D'abord jaunes et irrégulièrement distribuées, ces taches deviennent confluentes; elles prennent une coloration d'un brun fauve tout en conservant une auréole d'un vert jaunâtre. La maladie restreint le travail des feuilles qui se dessèchent et tombent prématurément. Les sarments des souches infectées mûrissent difficilement et le bois devient impropre au greffage.





*Figure 4.3 : Feuille et fruit atteints de Mélanose*

**Greasy spot :** Pendant longtemps, les symptômes causés par cette maladie ont été attribués à des dégâts d'acariens ou de nature physiologique. Cette maladie a été attribuée plus tard à un champignon dont les symptômes touchent les feuilles et les fruits (figure 4.4). Les symptômes foliaires débutent par l'apparition d'une marbrure jaune à la surface de la feuille. Plus tard les surfaces touchées deviennent un brun plus sombre à noir et prennent un aspect graisseux. Les feuilles atteintes chutent souvent avant que la couleur sombre noire des taches ne se développe. Des petites taches en forme de tête d'épingle apparaissent entre les glandes à huiles sur les fruits. Ces taches en général n'apparaissent pas avant 3 à 6 mois après l'infection. Chez la plupart des variétés, ces taches restent petites et n'entraînent pas de défaut notable sur le fruit. Mais les cellules autour de ces taches gardent souvent une couleur verte beaucoup plus longtemps que la normale. Ces taches de teinte rosée au début prennent plus tard une couleur brune à noir.



*Figure 4.4 : Feuille et fruit atteints de Greasy spot*

Cette dernière image représente une feuille normale



Figure 4.5 : Feuille normale

### - 3.1.2. Filtrage de l'image

Le filtrage est une étape de pré-traitement classique des images visant à réduire le bruit pour améliorer les résultats de traitements futurs. Dans notre travail, nous avons utilisé le filtre médian car il permet de réduire le bruit tout en conservant les contours, le contraste et la luminosité de l'image. C'est un filtre numérique non linéaire qui permet d'éliminer les valeurs aberrantes sans se limiter à faire un calcul de moyenne qui aura tendance à contaminer les valeurs voisines avec cette valeur aberrante et flouter l'image. Dans les zones où l'intensité est monotone (uniquement croissante ou uniquement décroissante) le filtre laisse l'image inchangée. Il respecte les contours, et élimine les valeurs extrêmes.

Pour le filtrage des images RVB par le filtre médian, nous avons suivi la stratégie donnée par la figure 4.6. Un exemple de filtrage d'image par ce filtre est présenté par la figure 4.7.

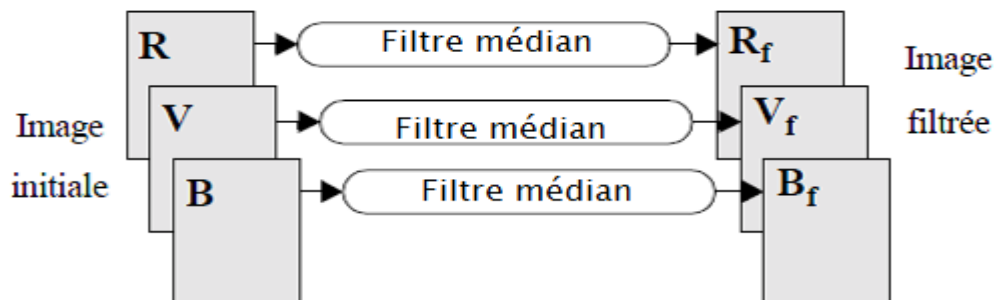


Figure 4.6



*Figure 4.7 : Filtrage d'une image*

### **3.1.3. Suppression de l'arrière plan**

La suppression de l'arrière plan consiste à isoler la région de l'image qui contient effectivement la feuille afin d'en extraire les caractéristiques visuelles bas niveau. Pour cela, une copie de chaque image RVB est convertie, dans un premier temps, en niveau de gris pour créer un masque binaire (figure 4.8). Ensuite, Ce masque est appliqué à chaque plan de couleur (rouge, vert et bleu) de l'image originale pour obtenir une image sans arrière plan (figure 4.9).



*Figure 4.8 : Création d'un masque*



Figure 4.9: Suppression de l'arrière plan

- 
- 

### - 3.1.4. Conversion de l'espace RVB en HSI

Nous avons travaillé dans l'espace de couleur HSI car les composantes Teinte  $H$  et Saturation  $S$  sont indépendantes de toute variation de luminosité. Plus particulièrement,  $H$  est peu sensible aux phénomènes d'ombre.

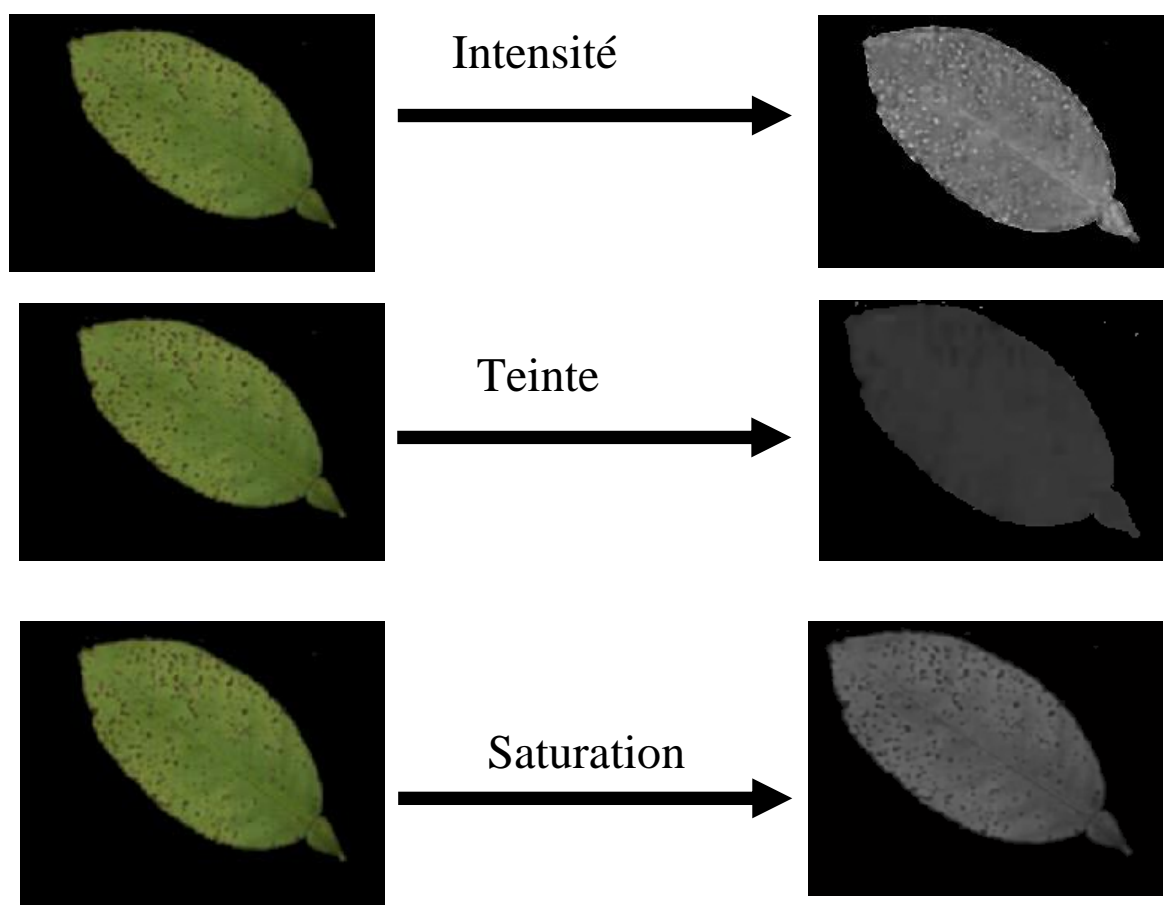
La figure 4.10 présente un exemple de la conversion l'espace couleur RVB en HSI on utilisant les formules suivantes :

$$\text{Intensité : } I = \left( \frac{R+V+B}{3} \right)$$

$$\text{Saturation : } S = 1 - \left( \frac{3 \times \min(R,V,B)}{(R+V+B)} \right)$$

$$\text{Hue(Teinte) : } H = 2 - \text{ACOS} \left( \frac{(R-V)+(R-B)}{2\sqrt{(R-V)^2+(R-V)(V-B)}} \right) \quad \text{Si } B > G$$

$$H = \text{ACOS} \left( \frac{(R-V)+(R-B)}{2\sqrt{(R-V)^2+(R-V)(V-B)}} \right) \quad \text{Si non}$$



*Figure 4.10 : Conversion de RVB en HSI*

### **3.1.5. Extraction des descripteurs**

L'examen des feuilles malades des agrumes, par observation visuelle, a mis en évidence que les symptômes, qui apparaissent aux surfaces de ces feuilles, sont caractérisés par des propriétés texturales spécifiques à chaque maladie. Ces propriétés texturales représente, pour les experts, des traits distinctifs permettant d'identifier plusieurs maladies des agrumes. C'est pourquoi, dans ce travail, nous avons utilisé des descripteurs de texture pour automatiser l'identification des maladies des agrumes. Les descripteurs utilisés sont les indices d'Haralick extraits des matrices de cooccurrences [45]. L'algorithme adopté pour l'extraction de ces indices, permet de calculer un vecteur contenant 39 composantes des caractéristiques résultantes des 13 indices d'Haralick extraits des matrices de cooccurrences pour chaque plan des images HSI. Des exemples d'indices d'Haralick, extraits des images de feuilles atteintes par différentes maladies, sont présentés dans les tableaux 4.1, 4.2, 4.3 et 4.4

pour les plans intensité (ligne 1), teinte (ligne2) et saturation (ligne 3).

	Contraste	Corrélation	Energie	entropie	Homogenité	variance	dif variance	dif entropie	info correlation 1	info correlation 2
1	64.6728	0.9957	0.4857	3.0156	0.7550	4.2162e+04	2.2574	-0.5028	0.9317	2.2574
2	168.3589	0.9824	0.4857	2.8132	0.7639	1.9365e+05	2.0689	-0.4702	0.9070	2.0689
3	7.9112	0.9584	0.4879	2.0290	0.8897	2.6219e+04	1.7421	-0.6723	0.9337	1.7421

*Tableau 4.1 : caractéristiques de texture pour une feuille atteinte par la maladie Mélanose*

	Contraste	Corrélation	Energie	entropie	Homogenité	variance	dif variance	dif entropie	info correlation 1	info correlation 2
1	42.1227	0.9969	0.3634	3.8192	0.7044	4.2613e+04	2.9770	-0.5627	0.9745	2.9770
2	284.7773	0.9684	0.3635	3.6054	0.7061	1.7903e+05	2.7032	-0.5076	0.9560	2.7032
3	26.9834	0.8795	0.3667	2.6241	0.8365	1.6800e+04	2.1630	-0.6016	0.9463	2.1630

*Tableau 4.2 : caractéristiques de texture pour une feuille atteinte par la maladie Scab*

	Contraste	Corrélation	Energie	entropie	Homogenité	variance	dif variance	dif entropie	info correlation 1	info correlation 2
1	36.2260	0.9971	0.2453	4.4707	0.6314	4.7891e+04	3.3267	-0.4808	0.9701	3.3267
2	147.6590	0.9866	0.2454	4.1946	0.6355	1.5144e+05	3.1057	-0.4764	0.9631	3.1057
3	8.8157	0.9753	0.2470	3.6118	0.7416	3.9421e+04	2.9235	-0.5877	0.9750	2.9235

*Tableau 4.3 : caractéristiques de texture pour une feuille atteinte par la maladie Greasy spot*

	Contraste	Corrélation	Energie	entropie	Homogenité	variance	dif variance	dif entropie	info correlation 1	info correlation 2
1	61.9109	0.9969	0.3245	3.6003	0.7200	6.8098e+04	2.8493	-0.5739	0.9720	2.8493
2	272.7937	0.9763	0.3247	3.4557	0.7430	1.6532e+05	2.7268	-0.5609	0.9656	2.7268
3	31.6485	0.6758	0.3420	2.3239	0.8939	2.1060e+04	2.0059	-0.6829	0.9540	2.0059

*Tableau 4.4 : Tableau de caractéristiques d'une feuille normal*

En plus de ces descripteurs, nous avons utilisé le descripteur global GIST [46] construit en combinant les amplitudes obtenues à la sortie de  $K$  filtres de Gabor [47] à différentes échelles  $E$  et orientations  $O$ . Chaque image est subdivisée en  $N * N$  blocs ( $N$  compris entre 2 et 16). Ce qui donne un vecteur de dimension  $N * N * K * E * O = 128$  pour chaque plan de l'image HSI (dans notre cas,  $N=4$ ,  $K=4$ ,  $E=4$ ,  $O=2$ ).

### 3.1.6. Classification

Afin d'automatiser l'identification des différentes maladies qui se manifestent par des symptômes sur les feuilles des agrumes, nous avons utilisé un réseau de neurones multicouches [48] pour l'apprentissage et la classification de ces maladies à partir des indices d'Haralick et des descripteurs de GIST. Pour chaque maladie, 20 images sont utilisées pour l'apprentissage et 20 pour le test. La figure 4.11 présente l'architecture de ce classificateur.

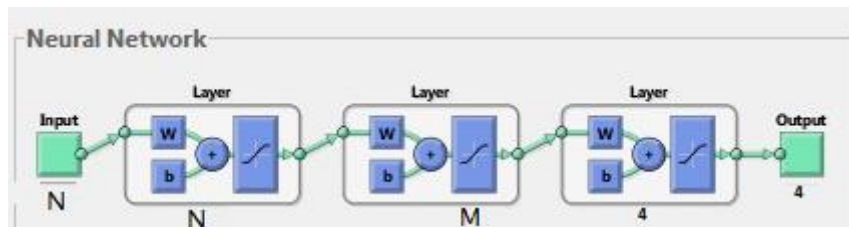


Figure 4.11 : Architecture du réseau de neurones utilisé

Avec  $N$  : le nombre de neurones de la couche d'entrée égal au nombre des composantes du vecteur d'entrée.

et  $M$  : le nombre de neurones de la couche cachée estimé à partir de l'équation (2.2)

La fonction d'activation utilisée par ce classificateur est la tangente hyperbolique définie par l'équation (2.1).

### 3.3. Expériences

Pour mettre en œuvre le procédé d'analyse sémantique d'identification des maladies des agrumes selon la démarche expérimentale présentée dans le paragraphe précédent, nous avons réalisé les expériences suivantes :

### 3.2.1. Expérience 1:

Dans cette expérience, nous avons utilisé 39 caractéristiques statistiques de la texture (13 de la teinte, 13 de la saturation et 13 de l'intensité) pour étudier la classification des maladies considérées.

Les résultats d'identification obtenus sont regroupés dans le tableau 4.5 ci-dessous :

Classes ----->	Greasy spot	Mélanose	Feuille Normal	Scab
Résultats de classement de 20 images de test.	18	17	19	17
Précision d'identification	90%	85%	95%	85%

Tableau 4.5

### 3.2.2. Expérience 2:

Dans cette expérience, nous avons utilisé 26 caractéristiques statistiques de la texture (13 de la teinte, 13 de la saturation) pour étudier la classification des maladies considérées.

Les résultats de d'identification obtenus sont regroupés dans le tableau 4.6 ci-dessous :

Classes----->	Greasy spot	Mélanose	Feuille Normal	Scab
Résultats de classement de 20 images de test.	18	18	20	18
Précision d'identification	90%	90%	100%	90%

Tableau 4.6

### 3.2.3. Expérience 3:

Dans cette expérience, nous avons utilisé 384 caractéristiques de GIST (128 de la teinte, 128 de la saturation et 128 de l'intensité) pour étudier la classification des maladies considérées.

Les résultats d'identification obtenus sont regroupés dans le tableau 4.7 ci-dessous :

Classes ----->	Greasy spot	Mélanose	Feuille Normal	Scab
Résultats de classement de 20 images de test.	3	4	6	3
Précision d'identification	15%	20%	30%	15%

Tableau 4.7



### 3.2.4. Expérience 4:

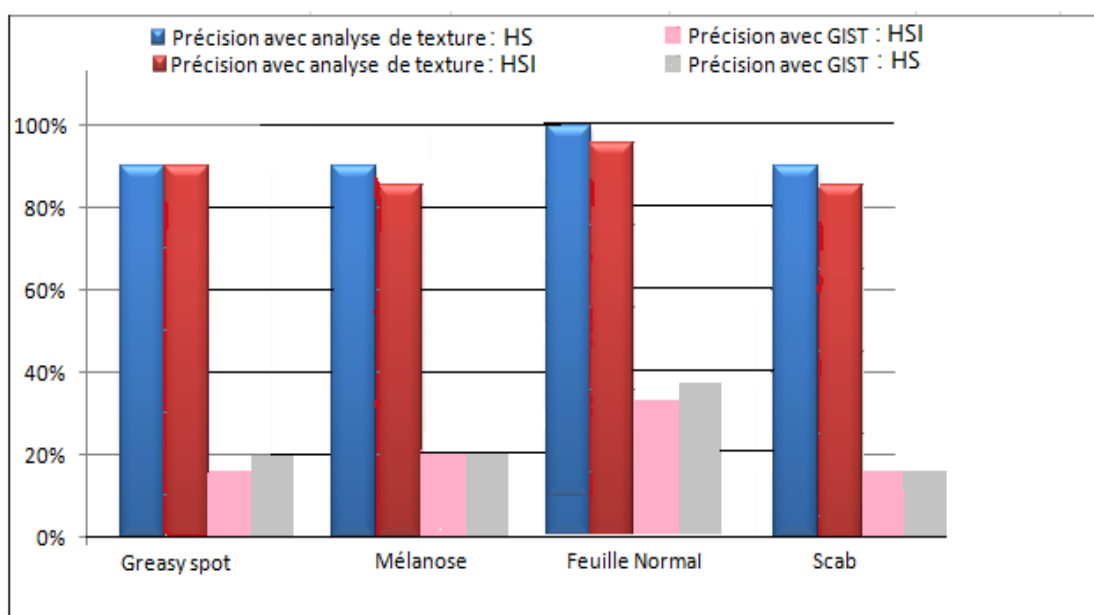
Dans cette expérience, nous avons utilisé 256 caractéristiques de GIST ( 128 de la teinte et 128 de la saturation) pour étudier la classification des maladies considérées.

Les résultats d'identification obtenus sont regroupés dans le tableau 4.8 ci-dessous :

Classes ----->	Greasy spot	Mélanose	Feuille Normal	Scab
Résultats de classement de 20 images de test.	4	4	7	3
Précision d'identification	20%	20%	35%	15%

Tableau 4.8

La figure 4.12 permet de comparer les résultats, obtenus au cours de ce travail, en terme de la précision d'identification des différentes maladies d'agrumes étudiées.



La figure 4.12 : Résultats d'identification des différentes maladies d'agrumes en terme de précision.

### **3.2.5. Analyse des résultats**

D'après la figure 4.12, nous constatons que les résultats, d'identification des maladies des agrumes, obtenus en utilisant les descripteurs de texture sont probants, en particulier lorsque ces descripteurs ne sont extraits qu'à partir des deux plans HS, car l'élimination de l'intensité annule l'effet des variations d'intensité. Cependant, pour le descripteur de GIST, les résultats obtenus ne sont pas encourageants.

Les résultats de ce travail mettent ainsi en évidence l'utilité des descripteurs de texture pour l'identification des maladies des agrumes dont les symptômes apparaissent sur leurs feuilles. Ceci nous a incité d'intégrer d'autres types de descripteurs de texture dans le procédé d'analyse sémantique des images pour améliorer davantage l'identification des maladies des agrumes.

## **4. Conclusion et perspectives**

Devant la recrudescence des problèmes posés par les maladies des agrumes, nous avons mené un travail de recherche sur l'analyse sémantique des images pour l'identification automatique de ces maladies à des fins de traitement précoce et efficace sans faire appel aux experts. Ce travail, nous a permis de développer un système de reconnaissance de trois maladies : la mélanose, le scab et le greasy spot. La technique d'analyse adoptée pour ce système, après évaluation de ces performances pour différents types de descripteurs, est l'analyse de la texture des images des feuilles d'agrumes malades.

Ce travail représente une plate forme pour la réalisation d'un système d'identification automatique des maladies des agrumes, et un pas vers le développement d'un système de reconnaissances précoce des pathologies des différentes plantes. C'est pourquoi à la suite de ce travail, il faut en particulier à court terme:

- élargir la base de données des images en y intégrant les images des feuilles d'agrumes touchées par d'autres maladies dues soit à d'autres ravageurs ou soit aux carences en fer, zinc, ... ;

- utiliser d'autres méthodes d'extraction des caractéristiques de la texture telle que les méthodes fréquentielles : les filtres de Gabor, la transformée de Fourier, les transformées en ondelettes ;
- utiliser les descripteurs de couleur ;
- utiliser autres classificateurs tels que les séparateurs à vastes marges, les réseaux bayésien, les arbres de décision, ...

## CONCLUSION GENERALE ET PERSPECTIVES

### **3. Conclusion générale**

Dans ce manuscrit, nous avons présenté le travail de recherche que nous avons mené sur l'analyse sémantique d'images permettant l'extraction automatique des informations visuelles pour l'annotation sémantique et la reconnaissance de formes. Ainsi, nous avons étudié et mis en œuvre les différentes parties du procédé d'analyse sémantique telles que la segmentation, la représentation des caractéristiques visuelles et la classification en y intégrant des techniques d'analyse et des méthodes d'apprentissage spécifiques à chacun des thèmes traités .

#### **3.1. Annotation automatique d'images**

Dans ce travail, après avoir défini la structure du système d'annotation automatique d'images, nous avons implémenté différents algorithmes de segmentation, d'extraction des caractéristiques visuelles et de classification. Ensuite, nous avons réalisé plusieurs expériences d'annotation d'images de trois bases de données images ETH-80, COIL-1000 et NATURE. Ces expériences nous ont conduits à élaborer les différentes parties du procédé d'analyse sémantique d'images le plus approprié pour l'annotation sémantique d'images en terme de réduction du fossé sémantique. Le procédé d'analyse ainsi obtenu consiste à segmenter l'image par croissance de région et à regrouper les régions adjacentes pour produire des régions de tailles importantes à partir desquelles les descripteurs de forme, de texture et de couleur, qui se sont avérés efficaces pour la caractérisation du contenu visuel, sont extraits. La correspondance, entre ces descripteurs et les mots clés définis pour annoter les images, est réalisée par un classificateur combinant les réseaux de neurones avec les réseaux bayésiens.

#### **3.2. Reconnaissance de formes**

Partant du fait que les formes à reconnaître varient en fonction de l'application envisagée, nous avons restreint notre travail à l'identification des panneaux de signalisation routière et des maladies des agrumes.

### **3.2.1. Identification des panneaux de signalisation routière**

L'analyse sémantique des images de la scène routière nous a permis de développer un système d'identification des panneaux routiers. Ce système est composé de deux modules : la détection et la reconnaissance. Le principe de la détection est basé sur l'analyse de la couleur et de la forme intrinsèques aux panneaux. La couleur est utilisée pour détecter les zones rouges et bleues de l'image susceptibles de contenir les panneaux routiers. Quant à la forme, déterminée par l'approximation polygonale et l'approche elliptique à partir des contours détectés par le filtre de canny, elle est utilisée pour finaliser la détection des éventuels panneaux et les classer en fonction de leurs couleurs et de leurs formes: les rouges triangulaires, les rouges circulaires, les bleus circulaires et les bleus rectangulaires. La reconnaissance de leur type est réalisée, pour chaque classe, par des réseaux de neurones à partir d'informations visuelles reflétant la répartition des pixels dans les images des panneaux détectés.

Actuellement, la mise en œuvre de ce système, pour une identification des panneaux routiers à partir des images vidéo en temps réel, est en cours.

### **1.2.2. Identification des maladies des agrumes**

Dans cette partie, nous avons automatisé l'identification des maladies des agrumes par une procédure d'analyse sémantique d'image intégrant le savoir-faire d'experts en reconnaissance des maladies des agrumes à partir de l'examen des symptômes apparaissant sur les feuilles. Le principe de cette procédure consiste à identifier les maladies des agrumes par un réseau de neurones multicouches entraîné à priori à discriminer les différentes maladies à partir d'attributs de texture extraits des images des feuilles. Les attributs utilisés sont les indices d'Haralick extraits de matrices de cooccurrences.

Les résultats de l'évaluation des performances des systèmes ainsi développés dans cette partie, sont prometteurs en particulier en terme de précision.

Le travail de cette thèse représente une contribution à l'analyse sémantique d'images. Il porte sur le procédé d'extraction de descripteurs, sur les informations extraites de l'image, ainsi que sur leur utilisation pour accéder à un verdict de classification.

## **4. Perspectives**

Le travail que nous avons présenté dans ce manuscrit peut être poursuivi selon quatre orientations :

### ***2.2. Amélioration des techniques d'analyse sémantique d'images***

Bien que les résultats obtenus au cours de ce travail sont prometteurs, Il existe plusieurs axes d'amélioration des techniques d'analyse et qui consistent en l'amélioration :

- de la description visuelle par l'utilisation de nouvelles caractéristiques robustes aux variations visuelles ;
- de la segmentation en introduisant l'apprentissage actif pour développer des méthodes de segmentation sémantique afin de se rapprocher d'un processus de segmentation humain ;
- des techniques de la classification pour qu'elles soient efficaces selon plusieurs critères: le coût de calcul à l'apprentissage et au test, les besoins en stockage mémoire pour les représentations d'images, le nombre d'images d'apprentissage nécessaires à la généralisation et à la discrimination.

### ***2.2. Utilisation d'autres approches d'analyse sémantiques***

Etant donné que nous avons utilisé une approche d'analyse sémantique directe pour tous les thèmes traités. L'utilisation d'autres approches d'analyse telles que les approches linguistiques, les approches compositionnelles et les approches structurelles (voir introduction générale), serait un autre axe d'amélioration distinctif.

### ***2.3. Développement d'autres processus d'analyse sémantique d'images***

Nous envisageons mener d'autres travaux d'analyse sémantique pour d'autres thèmes tels que l'analyse sémantique des images médicales et l'analyse sémantique pour la reconnaissance d'autres formes telles les visages, les plantes, .....

### ***2.4. Analyse sémantique d'autres types d'images***

Les images traitées dans ce travail, sont des images couleurs RGB, il serait intéressant, dans une suite logique, de mener des travaux de recherche sur l'analyse sémantique des images multispectrales et hyperspectrales telles que les images satellitaires, les images astronomiques et les images médicales.

## Bibliographie

- [1] R. Zhang : Combining visual features and contextual information for image retrieval and annotation: Theses and dissertations, 2011.
- [2] A. PUJOL : Contributions à la Classification Sémantique d'Images : Thèse de Doctorat, L'Ecole Centrale de Lyon, France, 2009.
- [3] A. Makadia, V. Pavlovic, et S. Kumar : A new baseline for image annotation: Proceedings of the 10th European Conference on Computer Vision Berlin, Heidelberg, Springer-Verlag, , p:316–329, 2008.
- [4] J. Li et J.Z. Wang : Real-time computerized annotation of pictures : Pattern Analysis and Machine Intelligence, IEEE Transactions on, 30(6), p:985–1002, 2008.
- [5] G. Lowe David: Distinctive image features from scale-invariant keypoints : IJCV, 60(2),p :91–110, 2004.
- [6] A. Barla, F. Odone, A. Verri : Histogram intersection kernel for image classification: Dans ICIP03, volume 3, p : 513–516, 2003.
- [7] S. Boughorbel, J.P. Tarel, N. Boujemaa : Generalized histogram intersection kernel for image recognition: Dans ICIP05, p : 161–164, 2005.
- [8] R. Albatal : Annotation automatique d'images à base de Phrases Visuelles : Thèse de Doctorat. Université de Grenoble, France, 2010.
- [9] A. Bosch, A. Zisserman, X. Munoz : Image classification using random forests and ferns: Dans ICCV'07, Proceedings of the 11th International Conference on Computer Vision, Rio de Janeiro, Brazil, 2007.
- [10] C. Dance, J. Willamowski, L. Fan, C. Bray, G. Csurka : Visual categorization with bags of keypoints: Dans ECCV'04, International Workshop on Statistical Learning in Computer Vision, 2004.
- [11] G. Griffin, A. Holub, P. Perona : Caltech-256 object category dataset: Rapport Technique 7694, California Institute of Technology, 2007.
- [12] S. Lazebnik, C. Schmid, J. Ponce : Beyond bags of features, Spatial pyramid matching for recognizing natural scene categories: Dans CVPR06, p: 2169–2178, 2006.
- [13] J. Fan, Yuli Gao, H. Luo, R. Jain : Mining multilevel image semantics via hierarchical classification: IEEE Transactions on Multimedia, 10(2), p:167–187, 2008.
- [14] J. Vogel, B. Schiele : Semantic modeling of natural scenes for content-based image retrieval: IJCV, 72(2), p:133–157, 2007.
- [15] J. Jeon, V. Lavrenko, R. Manmatha : Automatic image annotation and retrieval using cross-media relevance models: In Proceedings of the 26th conference on Research and development in informaion retrieval, p:119–126, New York, USA, 2003.

- [16] P. Duygulu, K. Barnard, J. Freitas, A. Forsyth : Object recognition as machine translation Learning a lexicon for a fixed image vocabulary: In Proceedings of the 7th European Conference on Computer Vision-Part IV, p:97–112, London, UK, 2002.
- [17] K. Barnard, D. Forsyth : Learning the semantics of words and pictures: In International Conference on Computer Vision, volume 2, p: 408–415. Vancouver : IEEE, 2001.
- [18] S.L. Feng, R. Manmatha, V. Lavrenko : Multiple bernoulli relevance models for image and video annotation. In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, volume 2, p:1002-1009,2004.
- [19] R. Fergus, P. Perona, A. Zisserman : Object class recognition by unsuper vised scale-invariant learning. CVPR, 02, p :264, 2003.
- [20] A. Kushal, C. Schmid, J. Ponce : Flexible object models for category-level 3D object recognition: In CVPR07, p: 1–8, 2007.
- [21] B. Leibe, A. Ettlín, B. Schiele : Learning semantic object parts for object categorization. IVC, 26(1) p:15–26, 2008.
- [22] J. Tang, H Li, G. Qi, T. Chua : Image annotation by graph-based inference with integrated multiple/single instance representations: IEEE Trans. on Multimedia, Vol 12, N°2, 2010.
- [23] J. Liu, M. Li, Q. Liu, H. Lu, S. Ma : Image annotation via graph learning: Pattern Recognition, vol. 42, N° 2, p: 218–228, 2009.
- [24] J. Liu, M. Li, W. Y. Ma, Q. Liu, H. Lu : An adaptive graph model for automatic image annotation: In MIR' 06, p: 61–70, 2006.
- [25] H. Wang, H. Huang ; C. Ding : Image annotation using bi-relational graph of images and semantic labels: In Computer Vision and Pattern Recognition (CVPR), p: 793 – 800, 2011.
- [26] J. Liu, B. Wang, H. Lu, S. Ma : A graph-based image annotation framework: Pattern Recognition Letters, Vol. 29, p: 407-415, 2008.
- [27] X. Rui, Z. Li, W.Y. Ma, N. Yu :Bipartite graph reinforcement model for web image annotation : In ACM Multimedia ' 07, p :585–594, 2007.
- [28] S. Barrat, S. Tabbone : Classification and Automatic Annotation Extension of Images Using Bayesian Network: Chapitre d'ouvrage, Structural, Syntactic, and Statistical Pattern Recognition, 5342, Springer Verlag, p:937-946, 2008.
- [29] Y. Mori, H. Takahashi, R. Oka : Image-to-word transformation based on dividing and vector quantizing images with words: In MISRM' 99, 1999.
- [30] P. Duygulu, K. Barnard, J. F. G. de Freitas, D. A. Forsyth : Object Recognition as Machine Translation - Learning a Lexicon for a Fixed Image Vocabulary: In ECCV' 02, p: 97–112, 2002.
- [31] H. Gracidas, L. E. Sucar : Markov random fields and spatial information to improve automatic image annotation: In Volume 4872 of Lecture Notes in Computer Science, p: 879–892. Springer,2007.



- [33] D. Blei, A. Ng, M. Jordan : Latent dirichlet allocation : Journal of Machine Learning Research, p:993–1022, 2003.
- [34] D. Blei, M. Jordan : Modeling annotated data: In SIGIR '03, p: 127–134, Toronto, Canada,2003.
- [35] Y. Zhao,Y. Zhao, Z. Zhu, J. S. Pan : A Novel Image Annotation Scheme Based on Neural Network: Eighth International Conference on Intelligent Systems Design and Applications ( ISDA '08), Vol.3, p:644-647, Kaohsiung, Taiwan, 2008.
- [36] C. Yang, M. Dong, J. Hua : Region-based image annotation using asymmetrical support vector machine-based multiple instance learning: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '06), Vol. 2, p: 2057–2063, 2006.
- [37] S. Manochitra , E. J. Janice, S. Suganya : Hybrid based Semantic Image Annotation using SVM and DT: International Journal of Computer Applications, Vol. 65, No.21, 2013.
- [38] S. Wan : Image Annotation Using the Simple Decision Tree: Fifth International Conference on Management of e-Commerce and e-Government (ICMeCG), p:141-146, Hubei , China, 2011
- [39] ETH-80 database image. Available Online:  
<http://www.d2.mpi-inf.mpg.de/Datasets/ETH80>
- [40] COIL-100 database image. Available Online:  
<http://www.cs.columbia.edu/CAVE/software/softlib/coil-100.php>.
- [41] B. Minaoui, M. Oujaoura, M. Fakir :Combined generative and discriminative Classifiers for Semantic Automatic Image Annotation: International Journal of Image Processing (IJIP),Vol.8, N° 5, p:225-244, 2014.
- [42] Y. S. Frank, S. Cheng : Automatic seeded region growing for color image segmentation: Image and Vision Computing 23, p:877–886, 2005.
- [43] M. Oujaoura, B. Minaoui, M. Fakir : Image Annotation by Moments: Chapter edited in the Book ‘‘ Moments and Moment Invariants - Theory and Applications’’, Chapter10, Vo.11, p: 227-252, Science Gate Publishing 2014.
- [44] S. Ryszard : Image Feature Extraction Techniques and Their Applications for CBIR and Biometrics Systems: International Journal Of Biology And Biomedical Engineering,Vol. 1, Issue 1, p:6-16, 2007.
- [45] R. Haralick : Statistical and structural approaches to texture: IEEE Transactions on SMC, Vol.67, N° 5, p:786–804, 1979.
- [46] A. Oliva , A. Torralba : Building the gist of a scene: the role of global image features in recognition: Progress in Brain Research, 2006.
- [47] H. G. Feichtinger, T. Strohmer : Gabor Analysis and Algorithms – theory and applications: Birkhäuser, Boston, 1998.

- [48] M. Oujaoura, B. Minaoui, M. Fakir : Image Annotation using Moments and Multilayer Neural Networks: Special Issue of International Journal of Computer Applications (IJCA), p:46-55, 2012.
- [49] C. W Hsu : A comparison of methods for multiclass support vector machines and Neural Networks, IEEE Transactions, Vol.13 , Issue.2, p :415 – 425, 2002.
- [50] M. Oujaoura, B. Minaoui, M. Fakir : Multilayer neural networks and nearest neighbor classifier Performances for image annotation: International Journal of Advanced Computer Science and Applications, Vol. 3, No. 11,p:165-171, 2012
- [51] M. Oujaoura, B. Minaoui, M. Fakir :Color, texture and shape descriptor fusion with Bayesian network classifier for automatic image annotation: International Journal of Advanced Computer Science and Applications (IJACSA). Vol. 4, No. 12, p:22-29, 2013
- [52] A. Ajimi , K. Sree : Efficient Automatic Image Annotation using Weighted Feature Fusion and its Optimization using Genetic Algorithm: Communications on Applied Electronics,Vol.1, N°6, p:15-19, 2015.
- [53] D.Ivan, S. Luciano : Mixing Low-Level and Semantic Features for Image Interpretation: Proceedings of Computer Science. Computer Vision - Workshops. Zurich, Switzerland, Vol.8926, p: 283-298, 2015.
- [54] J. Cong : Automatic image annotation using feature selection based on improving quantum particle swarm optimization: Signal Processing journal, Vol.109, p:172-181, 2015.
- [55] D. Hengam, A. Eskandari : A Novel semantic statistical model for automatic image annotation Using the Relationship between the Regions Based on multi-criteria Decision Making: Electrical and computer Engineering, Vol.4, N°1, p: 37-51, 2014.
- [56] J. Zhang, Da Li, W. Hu, Z. Chen, Y. Yuan : Multilabel Image Annotation Based on Double-Layer PLSA Model: The Scientific World Journal, p:1-9, 2014.
- [57] Z. Dongping, Y. Li , p. Huailiang, Y. Lu : Image Annotation Based on Joint Feature Selection with Sparsity: Information Technology Journal, Vol.13, p : 102-109, 2014.
- [58] P. Dong,Tian : A Review on Image Feature Extraction and Representation: Techniques Multimedia and Ubiquitous Engineering journal, Vol.8, N° 4, p :385-395, 2013.
- [59] B. Fernando, E. Fromont, D. Muselet, M. Sebban : Discriminative Feature Fusion for Image Classification: IEEE Conference Computer on Vision and Pattern Recognition (CVPR), p: 3434-3441, 2012.
- [60] A. Krichna, B. Prasad : Automated image annotation for retrieval of medical images: Computer Application journal. Vol. 55, N°(3), p: 0975-8887, 2012.

- [61] R. Zhang : Combining visual features and contextual information for image retrieval and annotation: Theses and dissertations, Ryerson University, Toronto, Ontario, Canada. 2011.
- [62] R. Zhang, L. Guan, L. Zhang, W. Xin-Jing: Multi-Feature pLSA for Combining Visual Features in Image Annotation: Proceedings of the 19th ACM international conference on Multimedia, Scottsdale, Arizona, USA, p:1513-1516, 2011.
- [63] H. Jian H, B. Zhang, Nai-Ming Qi, Y. Yang : Evaluating Feature Combination in Object Classification: Advances in Visual Computing journal; p : 597-606, 2011
- [64] Wang, T. Mei , S. Gong, X. Hua : Combining global, regional and contextual features for automatic image annotation: Pattern Recognition journal.; Vol.42, N°2, p : 259–266, 2009.
- [65] A. Broggi, P. Cerri, P. Medici, P.P. Porta, G. Ghisio : Real Time Road Signs Recognition: Proc. of IEEE Intelligent Vehicles symposium (IV'2007). Istanbul, Turke), 2007.
- [66] R. Timofte, K. Zimmermann, L. Van Gool : Multi-view traffic sign detection, recognition, and 3D localisation. Machine Vision and Applications, Vol.25, N°3, p:633–647, 2004.
- [67] Y. Nguwi, A. Z. Kouzani : Detection and classification of road signs in natural environments: Neural Computing and Applications, Vol.17, N°3, p:265–289, 2008.
- [68] X. Qingsong, S. Juan, L. Tiantian : A detection and recognition method for prohibition traffic signs: In International Conference on Image Analysis and Signal Processing, Zhejiang, China, p:583–586, 2010.
- [69] S. Vitabile, G. Pollacia, G. Pilato, F. Sorbello : Road signs recognition using a dynamical pixel aggregation technique in the HSV color space: Proceedings of 11th Int. conf. on Image Analysis and Processing (ICIAP), p:572-577, 2001.
- [70] D.G. Shaposhnikov, L.N. Podladchikova, A.V. Golovan, N.A. Shevtsova: Road sign recognition by single positioning of space-variant sensor window: In Proc. 15th International Conference on Vision Interface, p:213–217, Calgary, Canada, 2002.
- [71] A. Vázquez Reina, R. J. López Sastre, S. Lafuente Arroyo, P. Gil Jiménez : Adaptive traffic road sign panels text extraction: In Proceedings of the 5th WSEAS International Conference on Signal Processing, Robotics and Automation (ISPR'06), p:295–300, Wisconsin, USA, 2006.
- [72] H. Ohara, I. Nishikawa, S. Miki, N. Yabuki : Detection and recognition of road signs using simple layered neural networks: Proceedings of the 9th International Conference on Neural Information Processing (ICONIP), vol. 2, p: 626-630, 2002.
- [73] S. Zhu, L. Liu L : Traffic sign recognition based on color standardization: Proceedings of the IEEE International Conference on Information Acquisition, p: 951-955, 2006.
- [74] M. Boumediene : Contributions aux systèmes d'aide à la conduite automobile : Thèse de Doctorat, Université des Sciences et de la Technologie d'Oran, 2015.
- [75] Y. Ishizuka, Y. Hirai : Segmentation of road sign symbols using opponent-color filters: In ITSWC, Nagoya, Japon, October 2004.

- [76] N. Barnes, A. Zelinsky : Real-time radial symmetry for speed sign detection: In IEEE Intelligent Vehicles Symposium, p:556–571, Parma, Italy, 2004.
- [77] N. Barnes, A. Zelinsky, L. Fletcher : Real-Time Speed Sign Detection Using the Radial Symmetry Detector: IEEE Transactions on Intelligent Transportation Systems, Vol.9, N°2, p:322–332, 2008
- [78] R. Belaroussi, J. P. Tarel : Détection des panneaux de signalisation routière par accumulation bivariée : Traitement du signal, Vol.27, N°3, p:265–298, 2010.
- [79] R. Belaroussi, J. P. Tarel : Modèle géométrique du sommet et de la bissectrice d'un angle pour la détection de panneaux triangulaires : Actes du 17ème congrès franco- phone de Reconnaissance des Formes et Intelligence Artificielle (RFIA), Caen, France, Jan 2010.
- [80] R. Belaroussi, J. P. Tarel : Angle Vertex and Bisector Geometric Model for Triangular Road Sign Detection: In IEEE Workshop on Applications of Computer Vision (WACV'09), p:577–583, Snowbird, Utah, USA, 2009.
- [81] G. Garrido and al. : Robust Traffic Signs Detection by means of Vision and V2I Communications: In IEEE Conference on Intelligent Transportation Systems, Washington, USA, p:1003–1008, 2011.
- [82] G. Garrido and al : Fast Traffic Sign Detection and Recognition Under Changing Lighting Conditions. In IEEE Conference on Intelligent Transportation Systems, Toronto, Canada, p: 811–816, 2006.
- [83] A.F. Habib, M.N. Jha : Hypothesis generation of instances of road signs in color imagery captured by mobile mapping systems: International Archives of the Photogrammetry, Remote Sensing and Spatial In- formation Sciences 36, p:159–165, 2007.
- [84] G. Loy, N. Barnes : Fast Shape-based Road Sign Detection for a Driver Assistance System. In International Conference on Intelligent Robots and Systems, Sendai, Japan, p:70–75, 2004.
- [85] Y. Chiung, C. Fang : Road Sign Detection and Tracking: IEEE transactions on vehicular technology, vol. 52, N° 5, p: 1329-1340, 2003.
- [86] A. Behloul, Y. Saadna : A Fast and Robust Traffic Sign Recognition: International Journal of Innovation and Applied Studies, vol. 5, N°2, p:139–149, 2014.
- [87] F.A. Aly, A.E. Alaa : Detection, categorization and recognition of road signs for autonomous navigation: In Proceeding of Advanced Concepts for Intelligent Vision System, Brussels, Belgium, 2004.
- [88] A. Escalera, J.M. Armingol, J.M. Pastor, F.J. Rodriguez : Visual sign information extraction and identification by deformable models for intelligent vehicles: IEEE Transactions on Intelligent Transportation Systems, Vol.5 N°2, p :57–68, 2004.
- [89] P. Anne-Sophie : Détermination de vitesse limite par fusion de données vision et cartographiques temps réel embarquées : Thèse de Doctorat, MINES ParisTech , France, Avril 2013
- [90] A. Salhi, B. Minaoui, M. Fakir : Robust Automatic Traffic Signs Recognition Using Fast Polygonal Approximation of Digital Curves and Neural Network: International Journal of

- Advanced Computer Science and Applications(IJACSA), Special Issue on Advances in Vehicular Ad Hoc Networking and Applications p: 1-7, 2014
- [91] J. Greenhalgh, M. Mirmehdi,: Real-Time Detection and Recognition of Road Traffic Signs: IEEE transactions on Intelligent transportation systems, Vol. 13, N° 4, p: 1498 – 1506, 2012
- [92] N. Yok-Yen, A. Kouzani : Automatic road sign recognition using neural networks: International Joint Conference on Neural Networks, p:3955–3962, 2006.
- [93] M. Paolo, C. Claudio, C. Elena, P. Pier : Real Time Road Signs Classification : Proceedings of the IEEE International Conference on Vehicular Electronics and Safety Columbus, OH, USA, 2008
- [94] S. Bascon and al.: Road-Sign Detection and Recognition Based on Support Vector Machines : Intelligent Transportation Systems, IEEE Transactions on Vol.8 , Issue. 2, p : 264 – 278, 2007
- [95] A. Canny : Computational approach to edge detection: IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 8, p: 679-698, 1986.
- [96] D. P. Ramer : An iterative procedure for the polygonal approximation of plane curves: Computer Graphics and Image Processing, Vol.1, N°3, p:244–256, 1972.
- [97] D. Douglas, T. Peucker: Algorithms for the reduction of the number of points required to represent a digitized line or its caricature: The Canadian Cartographer, Vol.10, N°2, p:112–122, 1973.
- [98] A. Lorsakul, 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 , 2007.
- [99] E. Omrani et Al. : Potential of radial basis function based support vector regression for apple disease detection: Measurement Journal, vol.55, p:512–519, 2014.
- [100] R. Pugoy, VY. Mariano : Automated rice leaf disease detection using color image analysis: 3rd international conference on digital image processing, Vol.8009, p:1–7, 2011.
- [101] NE. Abdullah, AA. Rahim, H. Hashim, MM. Kamal : Classification of rubber tree leaf diseases using multilayer perceptron neural network: 5th student conference on research and development. Selangor: IEEE, p : 1–6, 2007.
- [102] P. Sanyal et Al. : Color texture analysis of rice leaves diagnosing deficiency in the balance of mineral levels towards improvement of crop productivity: 10th International Conference on Information Technology (ICIT) Orissa: IEEE, p: 85–90, 2007.
- [103] S. Kai S et Al. : A research of maize disease image recognition of corn based on BP networks: third international conference on measuring technology and mechatronics automation. Shangshai, IEEE, p: 246–249. 2011.

- [104] H. Wang et Al. : Application of neural networks to image recognition of plant diseases: Proceedings of the International Conference on Systems and Informatics (ICSAI) Yantai: IEEE, p:2159–2164, 2012.
- [105] KY. Huang : Application of artificial neural network for detecting Phalaenopsis seedling diseases using color and texture features: Comput Electron Agric., Vol.57, p:3–11, 2007.
- [106] R. Pydipati et Al. : Statistical and neural network classifiers for citrus disease detection using machine vision: Trans ASAE, Vol. 48, N°5, p:2007–2014, 2005.
- [107] Z. Jian , Z. Wei : Support vector machine for recognition of cucumber leaf diseases: 2nd international conference on advanced computer control. Shenyang: IEEE, p: 264–266, 2010.
- [108] F. Alham et Al. : Detection of Palm Oil Leaf Disease with Image Processing and Neural Network Classification on Mobile Device: International Journal of Computer Theory and Engineering, Vol. 5, No. 3, p:528-532, 2013.