

République Algérienne Démocratique et Populaire
Ministère de l'enseignement supérieur et de la recherche scientifique
Université KASDI Merbah - Ouargla
Faculté des Nouvelles Technologie de l'information et de la communication
Département d'informatique et des technologies de l'information



Mémoire Master Académique

Domaine : Informatique et Technologie de l'information

Spécialité : Administration et Sécurité de réseau

Présenté par : ZAIR Nassima

Encadré par : Dr. Adel ZGA

THEME

**L'INTERPRETATION SEMANTIQUE DES
EXTRACTIONS DU CONTENU D'UNE IMAGE**

Soutenu le 19 Juin 2022

Membres de jury :

M. BEKKRRI.F

UKMO

Examineur

Mme. KHELILI F

UKMO

Président

M. ZGA Adel

UKMO

Superviseur

Année universitaire: 2021/2022

Remerciements

En premier lieu, nous remercions ALLAH, le Tout-Puissant pour ses faveurs et ses grâces, de nous avoir donné le courage et la patience de mener ce travail durant toute cette année.

Nos remerciements et nos profondes gratitudes vont à notre encadreur Monsieur ZGA Adel qui nous a honoré par son encadrement, ses orientations, sa modestie, ses conseils et toutes ses constructives pour le bon déroulement de notre mémoire.

Nous tenons aussi à remercier les membres du jury pour leur précieux temps accordé à l'étude de notre mémoire.

Nous adressons aussi nos remerciements à tous les Enseignants surtout, les enseignants du département d'Informatique et des Technologies de l'Information.

En fin, nous tenons à remercier tous ceux qui nous ont aidé et encouragé de près ou de loin à la réalisation de ce travail

DÉDICACES

Nous dédions notre travail à :

Les deux êtres les plus chers au monde, ma mère et mon père.

*Mon mari, Farid, je le remercie pour sa patience et m'aide tout au long des
années d'étude.*

Mes enfants : Ines, Imen, Anouar.

ABSTRACT

Nowadays, three challenges of relationship-detection should be considered in order to build a strong model namely; long-tail problem, large intra-class divergence, and the semantic dependency or semantic gap. The aim of content-based image retrieval systems must provide maximum support in bridging the semantic gap between the simplicity of available visual features and the richness of the user semantics. Another issue is the long-tail problem where there is a low infrequent appearance of some objects (i.e. predicates) versus to the high occurrence of others. For that, an adequate scaling is demanded. The third problem that must be solved to build a strong CBIR system is the Intra/inter-class divergence. For the Intra-class divergence, objects are belonging to the same class but can't be represented with the same visual characteristics, whereas the inter-class divergence is where the similar visual descriptors can relate to two objects that are not related to each other.

Keywords

Content based image retrieval; Semantic Gap; Ontologies; large Intra/inter class divergence; Long tail problem;

RÉSUMÉ

De nos jours, trois défis de détection de relation doivent être prises en considération afin de construire un modèle fiable qui sont ; le problème de long-tail, la large divergence intra-classe et la dépendance sémantique ou le trou sémantique. L'objectif des systèmes de recherche basée sur le contenu doit fournir un support maximal pour combler le trou sémantique entre la simplicité des caractéristiques visuelles disponibles et la richesse des sémantiques de l'utilisateur. Un autre problème qui est le long-tail problème où il y a une faible apparition peu fréquente de certains objets versus à l'occurrence élevée des autres. Pour cela, une mise à l'échelle est exigée. Le troisième problème qui doit être résolu pour construire un système CBIR fiable est bien la divergence Intra/inter-classe. Pour la divergence Intra-classe, les objets appartiennent à la même classe mais ils ne peuvent pas être représentés avec les mêmes caractéristiques visuelles, tandis que la divergence inter-classe est celle où les descripteurs visuels similaires peuvent se rapporter à deux objets qui ne sont pas liés entre eux.

Mots clés :

recherche d'image basée sur le contenu. ; fossé sémantique ; les ontologies ; la divergence Intra/inter classe, le problème de long-tail

Table des matières

Remerciements.....	i
Dédicace.....	ii
Abstract.....	iii
Résumé.....	iv
Table des matières.....	vi
Abréviation.....	vii
Liste des figures.....	x
Introduction général.....	1
Chapitre I. La vision informatique et la recherche documentaire dans les images.	
I.1. Introduction.....	4
I.2. Représentation d'image.....	4
I.2.1. Représentation physique.....	4
I.2.1.1. La couleur.....	5
I.2.1.2. La texture.....	5
I.2.1.3. la forme.....	7
I.2.1.4. La localisation spatial.....	7
I.3. description Image.....	8
I.3.1. Descripteur locaux et descripteur globaux.....	8
I.3.2. Descripteur de couleur.....	9
I.3.3. descripteur de texture.....	9
I.3.4. Descripteur de forme.....	10
I.3.5. Descripteur basés sur les points d'intérêts.....	10
I.4. Segmentation des images.....	11
I.4.1. Segmentation basée sur les régions.....	12
I.4.2. Segmentation basée sur le contour.....	13
I.5. Technique recouvrement d'image.....	14
I.5.1. Recouvrement d'image base de contenu.....	14
I.5.2. Recouvrement d'image sémantique.....	15

I.6.Facette de la récupération d'image.....	16
I.6.1. Modalité de requête.....	16
I.6.2.Similté des images.....	17
I.6.3.Similitud syntaxique.....	18
I.7.Architecture général d'un système d'indexation.....	19
I.7.1.L'étape d'indexation.....	19
I.7.2.L'étape de recherche.....	20
I.8. annotation Image	22
I.8.1. Annotation manuel.....	22
I.8.2.Annotation automatique.....	24
I.9.Les différents types de l'apprentissage.....	26
I.9.1.Apprentissage supervisé.....	26
I.9.2.Apprentissage non supervisé.....	26
I.9.3.Apprentissage par renforcement.....	26
Chapitre II. Présentation des nouvelles d'apprentissage de l'ontologie	
II.1.Introduction.....	28
II.2.Interdépendance entre la connaissance et la langue.....	29
II.3.Apparition d'ontologie.....	29
II.4.La notion d'ontologie.....	30
II.5.Définition.....	30
II.6.Les différents type d'ontologies.....	31
II.7.Les constituants de l'ontologie.....	32
II.8.Classification d'ontologies.....	33
II.9.Domaine d'application des ontologies.....	34
II.9.1.Recherche d'information.....	35
II.9.2. Système d'information.....	36
Chapitre. III. Outils et résultats expérimentaux	
III.1. Introduction.....	39
III.2. Le problème de longue taille.....	39
III.3. La large divergence intra-classe.....	40
III.4. Quelque exemple d'une ontologie générée à partir de l'utilisation des objets détectés dans une image.....	49

Chapitre IV. Conception et implémentation	
IV.1. Introduction.....	45
IV.2.Présentation des outils de développement.....	45
IV.2.1.Matériel.....	52
IV2.2..Langage de programmation.....	52
Conclusion et perspective.....	52
Référence.....	54

ABREVIATIONS

CBIR : Content Based Image Retrieval.

CNN : Convolutoinal Neural Network.

HCVRD: Human Centric Visual Relationship Fetection Dataset.

VGG : Pretrained Model VGG.

RGB : Red Green Bleu.

LISTE DES FIGURES

Figure I.1. Un exemple d'espace couleur.....	5
Figure I.2. Quelques exemples de texture.....	6
Figure I.3. Quelques exemples de formes.....	7
Figure I.4. Différents types de point d'intérêt.....	10
Figure I.5. le lissage et l'analyse des gradients locaux.....	11
Figure I.6. Un exemple d'utilisation de segmentation par région.....	12
Figure I.7. Un exemple segmentation basée sur le contour.....	13
Figure I.8. Diagramme pour CBIR typique.....	21
Figure I.9. processus d'annotation manuelle avec les éléments de l'image.....	24
Figure I.10. Architecture générale d'une approche d'annotation automatique d'image.....	25
Figure II.1. Triangle sémiotique générique.....	29
Figure II.2. Fragment du model de domaine.....	37
Figure III.1. La distribution étiquette à longue queue de l'ensemble de donnée HCVRD.....	40
Figure III.2. Ambiguité visuelle.....	47
Figure III.3 Exemple1 représente une ontologie générée à partir de l'utilisation des objets détectés dans l'image	42
Figure III.4 Exemple2 représente une ontologie générée à partir de l'utilisation des objets détectés dans l'image.....	42
Figure IV.1. Interface d'application.....	48
Figure IV.2. Rechercher de l'image.....	49
Figure IV.3. Affichage de l'image sélectionnée.....	50
Figure IV.4. Insertion l'image avec model CNN.....	51

INTRODUCTION

GENERAL

INTRODUCTION GENERAL

Cet mémoire est dédiée à l'extraction sémantique et à l'interprétation de contenu d'image. Le but principal est de fournir des approches qui exploitent les informations contextuelles sémantiques aussi bien que le rapport d'inter-concept. Cela ne peut être fait en définissant le contexte de chaque objet appartenant à l'image en question. La recherche du contexte de chaque objet et de leur rapport visuel avec d'autres objets appartenant à la même image était la méthode adoptée par beaucoup de types de recherche faite jusqu'à maintenant. Mais il est avant tout nécessaire de développer un système d'interprétariat sémantique automatique. Dans notre travail, nous profitons des avancées dans l'ontologie. Les ontologies sont "une spécification explicite d'une conceptualisation". Ils garantissent une compréhension partagée d'un domaine particulier, aussi bien qu'un modèle formel qui est responsable de la machine non supervisée traitant.

Nous proposons l'utilisation d'approches qui définissent le contexte de chaque objet basé sur la collection d'objets appartenant à l'image en question.

Notre travail est organisé comme suit :

- Chapitre 1: dédié pour présenter la représentation d'image, la recherche la segmentation, techniques d'annotation, recherchent des techniques et une recommandation de techniques et définit.
- Chapitre 2 : est un état de l'art de l'ontologie en apprenant le modèle est décrit la présentation des nouveaux apprentissages de l'ontologie. et nous exposerons en détail la stratégie d'apprentissage utilisée.
- le Chapitre 3 : cette partie est dédiée à l'offre les outils et les résultats expérimentaux.
- le Chapitre 4 : on représente la conception de notre application, ainsi que la méthode d'implémentation de notre travail, on expliquant l'ensemble des choix techniques, (langage de programmation python) utilisés pour la réalisation de cette application.

CHAPITRE I

La Vision informatique et la recherche documentaire dans les images

I.1. INTRODUCTION

La recherche d'images par le contenu (en anglais : *Content Based Image Retrieval-CBIR*) est une technique permettant de rechercher des images à partir de leurs caractéristiques visuelles. Avant de pouvoir rechercher les images, il faut d'abord les extraire leurs caractéristiques. L'extraction des caractéristiques contenues dans une image s'appelle la description structurelle. Celle-ci peut prendre la forme d'une image ou de toute structure de données permettant une description des entités contenues dans l'image. Essentiellement, l'analyse de l'image fait appel à la segmentation où l'on va tenter d'associer à chaque région de l'image un label en s'appuyant sur l'information portée et la distribution spatiale. Les images sont classiquement décrites comme rendant compte de leurs caractéristiques de bas niveau telles que la texture, couleur, forme, etc. Un cas typique d'utilisation des caractéristiques de bas niveau est la recherche d'images visuellement similaires à un exemple donné en requête (image).

Cette technique s'oppose à la recherche d'images par des mots-clés, qui fut historiquement proposée par les moteurs de recherche où les images sont retrouvées en utilisant le texte qui les accompagne plutôt que le contenu de l'image elle-même.

I.2. Représentation d'image

La représentation d'une image consiste à décrire une image via des descripteurs (ou caractéristiques). Deux types de représentation sont détaillés dans ce chapitre : la représentation par les caractéristiques physiques (les caractéristiques de bas niveau) et la représentation par les caractéristiques sémantiques (les caractéristiques de haut niveau).

1.2.1. Représentation physique

Les caractéristiques physiques sont la base des systèmes de recherche d'images par le contenu. Les caractéristiques extraites de l'image entière sont appelées des caractéristiques globales et les caractéristiques obtenues à partir d'une région de l'image sont appelées des caractéristiques locales. Ces dernières caractéristiques (locale) se sont avérées plus proches de la perception sont basées sur la segmentation des images. Les systèmes de recherche d'images basés sur les caractéristiques de bas niveau permettent aux utilisateurs de représenter leurs besoins par une ébauche de la forme, du

contour de la couleur, de la texture des images qu'ils recherchent ou par la proposition d'une image similaire aux images qu'ils recherchent. L'image peut être fournie par l'utilisateur ou choisie dans une base de données image. Nous avons quatre caractéristiques de bas niveau : la couleur, la texture, la forme et la localisation spatiale comme le présentent Ying Liu et al. dans [1].

1.2.1.1. La couleur

Les couleurs sont l'une des caractéristiques les plus utilisées dans les systèmes de recherche d'images par le contenu. Elles sont riches en information et très utilisées pour la représentation des images. Elles forment une partie significative de la vision humaine. La perception de la couleur dans une image par l'être humain est un processus complexe et subjectif [2]. En effet, cette donnée varie considérablement avec l'orientation des surfaces, la caméra et l'illumination (positions et longueur d'onde des sources lumineuses), par exemple. Il est possible de représenter la couleur dans différents espaces à savoir RGB, HVS, LAB, etc. Le plus répandu est sans aucun doute l'espace R, V, B (RGB en anglais) qui code la couleur d'un pixel sur un vecteur en trois dimensions : rouge, vert, bleu.

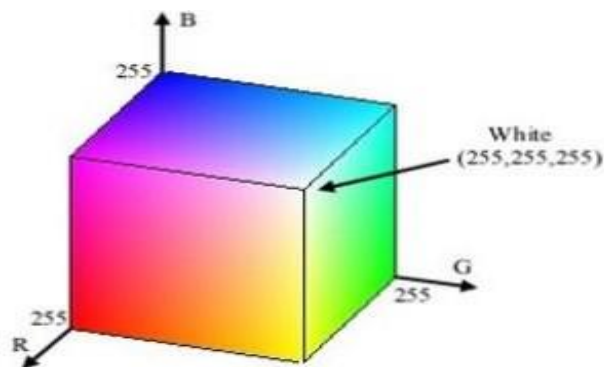


Figure I.1 . Un exemple d'espace couleur

1.2.1.2. La texture

La texture est la seconde caractéristique visuelle largement utilisée dans la recherche d'image par le contenu. Elle permet de combler un vide que la couleur est incapable de faire, notamment lorsque les distributions de couleurs sont très proches. La texture est généralement définie comme la répétition d'un motif créant une image visuellement homogène. Elle peut être vue comme un ensemble de pixels spatialement agencés, créant ainsi une région homogène.

Dans les travaux consacrés aux caractéristiques de bas niveau, la texture joue un rôle très important parce qu'elle fournit des informations importantes dans la classification des images puisqu'elle décrit le contenu de nombreuses images comme par exemple la peau des fruits, des arbres, des briques et le tissu, etc. Par conséquent, la texture est un élément important dans la définition de sémantique de haut niveau. De nombreux approches et modèles sont proposés dans [3] pour la caractérisation de la texture. Parmi les méthodes les plus connues, nous pouvons citer :

_ les méthodes statistiques basées sur les évaluations quantitatives de la distribution du niveau de gris. Elles sont focalisées sur l'étude des relations entre un pixel et ses voisins et sont utilisées pour caractériser des structures sans une régularité apparente.

_ les méthodes fréquentielles concernent les propriétés fréquentielles et s'appuient sur la transformée de Fourier [4], le filtre de Gabor [5], les ondelettes [6].

_ les méthodes géométriques permettent de décrire une forme à l'aide de propriétés statistiques [7]. Elles représentent les propriétés spatiales de la distribution des pixels dans l'image. La figure 1.2 présente quelques exemple de textures utilisées pour décrire les images.



Figure I.2. Quelques exemples de textures

1.2.1.3.La forme

La forme permet d'extraire des caractéristiques particulièrement robustes et discriminantes. Nous les appellerons caractéristiques d'objets car elles présupposent une segmentation de l'image et manipulent donc des informations liées aux objets réels. Au même titre que la texture, la forme est une caractéristique complémentaire de la couleur. Les systèmes de recherche d'images par le contenu se sont focalisés sur l'extraction d'attributs géométriques ces dernières années. Deux méthodes ont été proposées pour représenter les formes : les descripteurs basés sur les régions et les descripteurs basés sur les contours (frontières). Les descripteurs basés sur les régions sont utilisés pour caractériser l'intégralité de la forme d'une région et les descripteurs basés sur les frontières se focalisent sur les contours des formes [2].

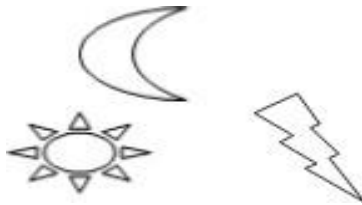


Figure I.3 . Quelques exemples de formes

1.2.1.4.La localisation spatiale

Outre les caractéristiques précitées, il est également possible de considérer l'organisation spatiale des différentes primitives comme une description en tant que telle. Il est évident qu'une telle description constitue un niveau intermédiaire entre les caractéristiques de bas niveau brutes et l'interprétation des images, et que ce niveau peut se révéler très expressif. Ainsi, lorsqu'un utilisateur recherche une image représentant un objet complexe, par exemple une voiture, la reconnaissance de chacune de ses parties (roues, carrosserie...) est une étape indispensable. La localisation spatiale n'est rien d'autre que la position des régions dans l'image. La position des régions est définie par rapport à la région du centre de l'image. Elle est généralement représentée par des expressions comme *Supérieur*, *Inférieur*, *Gauche*, *Bas* etc.

En résumé, nous pouvons dire que les caractéristiques de bas niveau de l'image peuvent être soit extraites de la totalité de l'image ou des régions de l'image. La plupart des systèmes actuels sont focalisés sur les régions d'images parce que les recherches basées sur les caractéristiques globales de l'image sont relativement simples mais les recherches basées sur les régions sont efficaces car elles sont plus proches de la perception humaine.

I.3. Description Image

I.3.1. descripteurs locaux et descripteurs globaux

On peut utiliser des descripteurs caractérisant la totalité de l'image (descripteur global) ou plusieurs descripteurs locaux caractérisant chacun une partie de l'image. Les techniques modernes en imagerie tendent à privilégier les descripteurs locaux aux globaux car les descripteurs locaux sont plus efficaces et ils permettent une recherche plus fine et absorbent mieux certaines variations. Dans le cas de descripteurs globaux, un seul descripteur décrit la totalité de l'image, cela les rend robustes au bruit qui peut affecter le signal, les histogrammes en couleur et des niveaux de gris en sont des exemples classiques, mais d'autres descripteurs existent comme le corrélogramme et les angles de couleurs. L'inconvénient de ces descripteurs est qu'ils ne permettent pas de distinguer des parties de l'image

ils ne distinguent pas, par exemple, les objets dans l'image, sauf dans le cas où l'image ne contient qu'un seul objet dans un fond uni. Par opposition, les descripteurs locaux s'associent à une partie/région de l'image qu'on commence par détecter avant de calculer le descripteur, cette partie peut concerner un objet par exemple, la détection se fait indépendamment de la position dans l'image, ce qui assure l'invariance par translation. On peut également catégoriser les descripteurs selon le type de modalité qu'ils représentent : descripteurs visuels, descripteurs de l'audio, descripteurs de mouvement, etc.

1.3.2.Descripteurs de couleur

Les descripteurs de couleurs sont les plus utilisés dans le domaine de la recherche des images et des vidéos par le contenu. L'histogramme s'avère le descripteur le plus simple à calculer, son calcul consiste à compter le nombre d'occurrences des différentes valeurs possibles d'intensité des pixels dans l'image. On peut distinguer plusieurs catégories d'histogrammes, on peut les classer, par exemple selon l'espace de couleur considéré lors du calcul : "histogramme RGB", "histogramme HSV", "histogramme Opponent, associés respectivement aux espaces de couleurs : "RGB", "HSV", "Opponent color space". "L'histogramme rg " est décrit dans le modèle de couleur RGB normalisé ($r+g+b=1$), les composantes r et g représentent l'information sur les couleurs dans une image (b est redondant du fait que $r+g+b=1$) : $r = r/(r+g+b)$, $g = g/(r+g+b)$, $b = b/(r+g+b)$. À cause de la normalisation, les composantes r et g sont invariantes à l'échelle et ainsi invariantes par rapport aux changements de l'intensité de lumière et ombres.

1.3.3.Descripteurs de texture

La texture représente également un descripteur bas niveau efficace utilisé dans le cadre de l'indexation et la recherche par le contenu. Plusieurs techniques ont été développées pour mesurer la similarité de textures. La majorité des techniques comparent les valeurs de ce qui est connu par les statistiques du second ordre, calculées à partir des images requêtes. Ces méthodes calculent les mesures de textures d'images comme étant le degré de contraste, la grossièreté, la directivité et la régularité; ou de la périodicité, la directivité et l'aspect aléatoire.

Une texture se caractérise par la répétition d'un motif de base dans différentes directions.

- On peut les séparer en deux classes: les textures structurées (macro textures) et les textures aléatoires (micro textures).
- Deux types de méthodes
 - Statistiques : distribution spatiales des intensités de l'image
 - Géométriques : segmentation de motifs qui se répètent

1.3.4.Descripteurs de formes

Un descripteur est un ensemble de nombres organisés de diverses manières : valeurs simples, vecteurs, matrices ou même des graphes [8]. Ces nombres sont produits pour décrire une forme donnée, cette dernière peut ne pas être entièrement restructurable à partir de ce descripteur mais les descripteurs de formes différentes devront être suffisamment différents pour pouvoir dissocier ces formes.

Un descripteur de forme suit le cheminement schématisé dans la Figure 1

1.3.4.Descripteurs basés sur les points d'intérêts

L'extraction des descripteurs visuels sur l'image entière (descripteurs globaux) permet de réduire le nombre de calculs nécessaires, la taille de la base de données ainsi que le coût de recherche des images les plus similaires. Cependant, l'approche globale ne permet pas une recherche efficace d'objets (au sens large) dans l'image. À l'inverse, les descripteurs extraits d'une partie de l'image (descripteurs locaux) sont efficaces, mais coûteux. Les descripteurs locaux peuvent être des régions de l'image obtenues soit par segmentation de l'image entière (par recherche de régions d'intérêt) ou par recherche des points d'intérêt. Les points d'intérêt d'une image sont les points qui seront trouvés similaires dans les images similaires. Une manière de les déterminer est de prendre en compte les zones où le signal change.

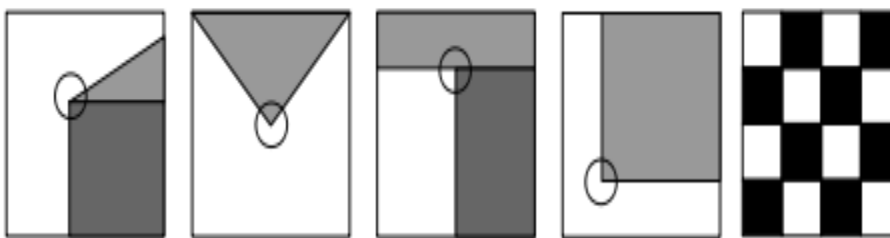


Figure I.4. Différents types de points d'intérêt : coin simple, jonction en 'V', jonction en 'T', jonction en 'L', jonction en 'damier'.

L'étape suivante consiste à concaténer et normaliser les seize histogrammes ainsi Obtenus. Afin de limiter la sensibilité du descripteur aux changements de luminosité, les Valeurs inférieurs à 0,2 sont remplacées par 0 et l'histogramme est de nouveau normalise. La figure I.8 illustre la construction de ce descripteur.

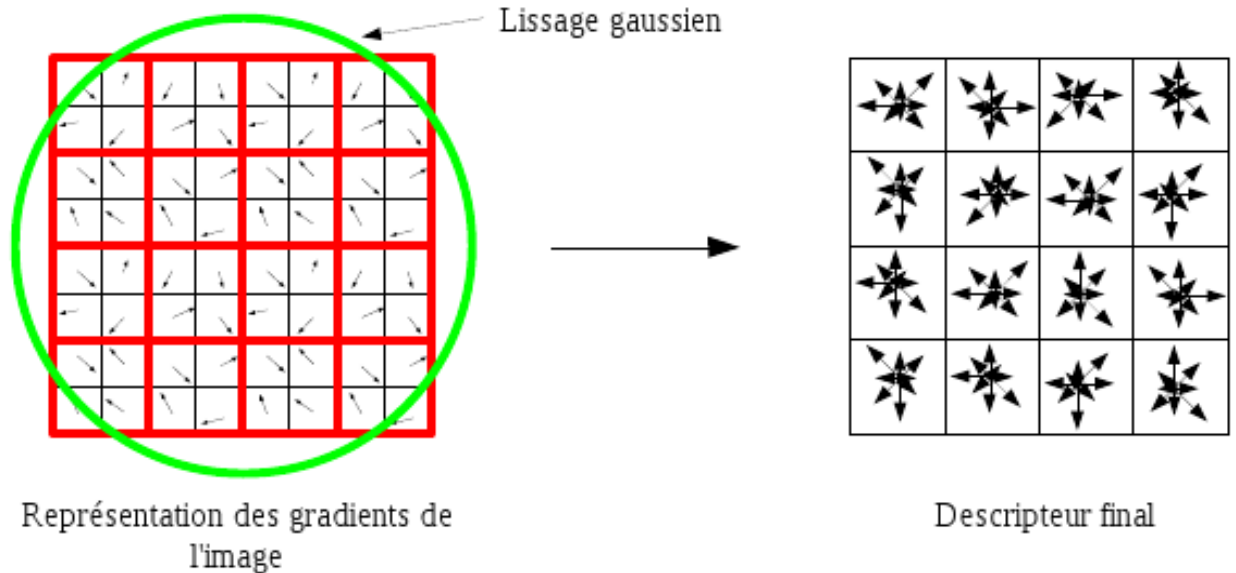


Figure I.5. (Gauche) : le lissage et l'analyse des gradients locaux, (droite) : la concaténation des histogrammes de zones

I.4.Segmentation des images

Afin de localiser automatiquement les objets dans les images, les chercheurs ont d'abord proposé la segmentation des images en plusieurs régions. De manière générale, segmenter une image consiste à délimiter dans l'image des plages régulières ou cohérentes, c'est-à-dire des régions dans lesquelles les informations de l'image suivent un modèle d'organisation : zones relativement homogènes en intensité, en texture ou en couleur, zones relativement planes, lisses (de variations lentes), etc. De nombreux systèmes d'interprétation d'images s'appuient sur la segmentation des images en des régions, qui permet d'extraire des constituants élémentaires, ou primitives, qui serviront de base pour l'identification ou la reconstruction des images [9].

De nombreuses techniques de segmentation des images sont proposées dans la littérature. Les algorithmes de segmentation se basent généralement sur les caractéristiques de bas niveau.

Ces techniques fonctionnent bien pour les images contenant uniquement des zones de couleur homogènes, tels que les méthodes de classification directes dans l'espace.

Les deux approches utilisées en segmentation d'image sont principalement l'approche région qui est basée sur la notion d'homogénéité d'une région et l'approche contour qui suppose que toute région est délimitée par un contour fermé ou bord.

1.4.1. Segmentation basée sur les régions

Les algorithmes de type décomposition/fusion exploitent les caractéristiques propres de chaque région (surface, intensité lumineuse, colorimétrie, texture, etc.). On cherche des couples de régions candidates à une fusion et on les note en fonction de l'impact que cette fusion aurait sur l'apparence générale de l'image. On fusionne alors les couples de régions les mieux notés, et on réitère jusqu'à ce que les caractéristiques de l'image remplissent une condition prédéfinie : nombre de régions, luminosité, contraste ou texture, ou alors jusqu'à ce que les meilleures notes attribuées aux couples de régions atteignent un certain seuil.

Les algorithmes de croissance de régions partent d'un premier ensemble de régions qui peuvent être calculées automatiquement (par exemple, les minima de l'image), ou fournies par un utilisateur de manière interactive. Les régions grandissent ensuite par incorporation.

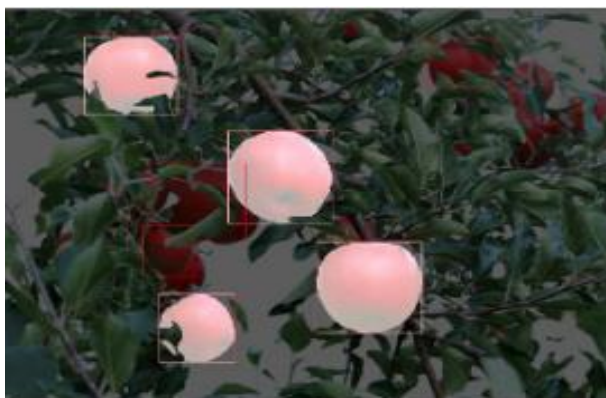


Figure I.10 : un exemple d'utilisation de segmentation par régions : étiquetage de pièces, d'objets

I.4.2.Segmentation basée sur le contour

Cette approche cherche à exploiter le fait qu'il existe une transition détectable entre deux régions connexes. Les méthodes les plus anciennes utilisent des opérateurs de traitement d'images, tels que le filtre de Canny, pour mettre en évidence les pixels qui semblent appartenir à un contour. On peut aussi faire intervenir des modèles déformables à l'aide de courbes paramétriques (courbe de Bézier, spline...) ou de polygones (par exemple algorithme à bulle).

Pour commencer le processus, on recherche les points remarquables de l'image, tels que les points à l'intersection de trois segments au moins. De tels points sont appelés des graines (seeds en anglais). L'intérêt principal des méthodes de segmentation selon l'approche frontières est de minimiser le nombre d'opérations nécessaires en cas d'itération du processus sur des séries d'images peu différentes les unes des autres (cas des images vidéo notamment). En effet, une fois que les contours des régions aient été trouvés dans la première image, l'application du modèle déformable à l'image suivante est plus efficace que de tout recalculer, si la différence entre les images est peu importante.



Figure I.6. Exemple segmentation basée sur le contour.

I.5. Techniques de recouvrement d'image

Les premiers systèmes de recouvrement d'image ont été fondés sur une description textuelle de l'image (les mots clé associés aux images) [5 ,10,11], les mots clé sont la source d'information externe manuellement attachée aux images. Le processus établissant un index basé sur les mots clé est le temps en consommant puisque c'est manuel, délicat, parce que la qualité du système de recouvrement dépend grandement de la pertinence des termes alloués aux images et non standardisée parce que la source de ces mots clé est généralement pas unique et compte sur la personne qui associe des mots clé aux images. En plus, une liste courte de mots clé ne peut pas complètement couvrir la sémantique souvent riche portée par une image.

La deuxième génération de systèmes de recouvrement d'image est fondée sur le contenu. Le contenu a basé des buts d'approche de directement extraire des informations de l'image lui-même pour être en mesure de le classifier, cela vient comme une alternative à l'approche textuelle mais les problèmes sémantiques rattachés au traitement automatique d'images sont vite notés. Les techniques pour résoudre ces problèmes ont été proposées dans la littérature [12,13,14]. Nous pouvons classifier les approches de recouvrement d'images dans deux courants principaux : le courant de contenu le recouvrement d'image basé sur le contenu et le courant de recouvrement d'image sémantique basé sur un formalisme de représentation de connaissance.

1.5.1. Recouvrement d'image à base de Contenu

Les principes de contenu le recouvrement d'image basé (CBIR) les approches doivent répondre à deux questions [28] : comment mathématiquement décrire une image ? et comment mesurer la similarité basée sur la description abstraite ? Ces approches appliquent d'habitude la statistique et la machine en apprenant des techniques. Pour décrire des images, la plupart des systèmes CBIR utilisent des caractéristiques automatiquement extraites des images comme la couleur, la texture et la forme. Ces caractéristiques sont utilisés dans de différents processus, comme le compte de similarité, le bâtiment modèle, ou même les annotations. nous étudions et décrivons les différentes méthodes proposées qui sont classifiés dans cinq catégories importantes [58] :

- Utilisation d'une ontologie : définir des objets de haut niveau et interpréter mieux les caractéristiques extraites de niveau bas.
- Utilisation d'un feed-back d'utilisateur : essayer d'apprendre les intentions de l'utilisateur de mieux comprendre leurs besoins en utilisant un traitement en ligne.

- le fait de Produire un gabarit sémantique : soutenir le haut niveau reflète le recouvrement en utilisant des gabarits pour représenter des concepts calculés d'une collection d'images de référence.
- Utilisation d'un contexte de web : car le web reflète le recouvrement. Il se compose à l'utilisation de renseignements qui peuvent être structurés comme l'url ou les pages HTML pour améliorer la sémantique d'images.

1.5.2. Recouvrement d'image sémantique

Dans les approches de recouvrement d'image classiques, les images sont décrites avec un ensemble de mots clé [5], [10], [11]. La qualité de ce processus dépend grandement de la pertinence de ces mots clé, les mots clé utilisés dans la requête d'utilisateur et l'adéquation entre eux. Généralement, ce processus exécute une comparaison syntaxique simple entre les mots clé d'annotation et les mots clé de requête d'utilisateur et si l'utilisateur n'utilise pas les mêmes mots clé, il peut ne pas recevoir qu'il cherche. Parmi les problèmes de ces approches est que les mots clé peuvent ne pas couvrir la sémantique riche portée par l'image.

Le recouvrement d'image sémantique s'approche de l'utilisation généralement un formalisme de représentation de connaissance comme les logiques de description ou les réseaux sémantiques [15]. Le but est de trouver un modèle de représentation d'image. Ces riches de représentation pour être facilement comparable et caractériser l'image de la meilleure façon. Le recouvrement sémantique est fondé sur le sens de mots clé . La comparaison entre les concepts (généralement renvoyé au raisonnement) est fondée sur les techniques plus avancées que ceux de recouvrement syntaxique. Il tient compte aussi de l'action réciproque entre de différents concepts de la requête pour améliorer la pertinence comparée à un recouvrement syntaxique. Wang et d'autres. [16] proposent une comparaison entre les approches de recouvrement d'image à base d'ontologie et les approches de recouvrement d'image à base de mots clé. Les mots clé - le recouvrement d'image basé est simple et facile de faire une demande avec une précision acceptable et le recouvrement d'image à base d'ontologie améliore la précision, mais exige une description complète de l'image.

1.6.facettes de la récupération d'image

Pour concevoir un système de récupération d'image, nous devons répondre à trois questions clés: comment construire la requête? Comment représenter une image? Et comment correspondre entre la requête et la représentation de l'image (description)?

Ces trois questions sont fortement liées les unes aux autres et sont considérées comme les trois fondements d'un système de récupération d'image. nous aurons un bref aperçu des réponses possibles aux questions ci-dessus. Les première et troisième questions concernent les modalités de requête et le calcul de similitude respectivement. Nous sommes intéressés par ce travail à une récupération basée sur le sémantique, la question deux est traitée dans une vue de techniques basée sur sémantique, qui représentent généralement l'image par annotations. D'autres approches, en particulier dans le cas des systèmes CBIR, utilisent d'autres techniques ne seront pas mentionnées ici.

1.6.1 .Modalités de requête

Un facteur important dans un système de récupération d'image est la modalité de requête. Cette modalité définit le langage de requête pris en charge par le système en termes d'expressivité de la requête utilisateur générée. Les modalités de requête les plus connues du système de récupération d'images sont [28]:

- **Mots-clés:** cette modalité est la plus utilisée. La requête est présentée sous forme de mots clés. L'ensemble des mots clés autorisés n'est pas limité en général comme les systèmes de récupération d'image sur le Web, mais peut être prédéfini, ce qui limitera l'expressivité des requêtes.
- **Texte libre:** Pour cette modalité, l'utilisateur essaie de définir ses besoins par des phrases. Ces phrases sont construites librement, ce peuvent être des questions, des histoires ou d'autres expressions.

- Image: comme son nom l'indique, la requête introduite par l'utilisateur est une image et le système renvoie toutes les images similaires à l'image de l'utilisateur. Cette modalité est très adaptée aux systèmes CBIR.
- Graphiques: cette modalité est basée sur une représentation graphique de la requête, l'utilisateur définit graphiquement ses besoins en dessinant une image. L'image peut également être générée automatiquement par ordinateur.
- Composite: il se compose d'une combinaison des autres modalités. En général, l'utilisateur peut choisir une modalité pour définir sa requête ou définir une partie de sa requête en utilisant plusieurs modalités. La combinaison de modalités est intéressante dans le cas d'une requête interactive, l'utilisateur peut à chaque fois donner plus de détails sur sa requête en modifiant la modalité. Il est important de faire la différence entre une technique de récupération d'image basée sur les mots clés et la modalité de requête en fonction des mots clés. Les modalités de requête sont utilisées comme entrée du processus de construction de requête, qui fournit après une requête au processus de récupération d'image, quelle que soit la technique utilisée.

En général, le processus de construction de requête n'est pas nécessaire pour les techniques de récupération syntaxique basées sur des mots clés, l'ensemble des mots clés introduits par l'utilisateur est utilisé sans prétraitement. Par inconvénient, il est très important pour les techniques de récupération d'image sémantique pour donner une sémantique à la requête.

1.6.1. Similité des images

Le calcul de similitude est un processus utilisé pour récupérer l'ensemble des réponses d'images. Il définit une relation d'affinité qui peut être quantifiée ou non entre une requête et une image. La requête est générée en utilisant la sortie de la modalité de requête, il peut s'agir d'une autre image, texte ou combinaison des deux. Le calcul de similitude permet également généralement le classement des réponses des images même si d'autres approches utilisent d'autres paramètres en plus de faire le classement [17, 18]. La similitude peut être syntaxique, sémantique ou hybride. La similitude syntaxique est basée sur une comparaison numérique entre les représentations des aspects visuels des images et la requête. Le résultat de cette comparaison est numérique et permet le classement des réponses. Par contre, la similitude sémantique est

basée sur la comparaison entre l'interprétation des représentations des aspects visuels des images et l'interprétation de la requête, il utilise généralement un formalisme de représentation des connaissances et une connaissance

Cette base de connaissances est utilisée pour mieux interpréter les représentations des images et la requête pour une meilleure comparaison. Le résultat de la comparaison n'est pas généralement numérique mais booléen et ne permet pas nécessairement le classement des réponses des images. La similitude hybride utilise une combinaison de similitudes syntaxiques et sémantiques. Il utilise la base de connaissances pour enrichir et pour une meilleure comparaison et un calcul numérique pour le classement.

1.6.2.similitude syntaxique

Le principe général de la similitude syntaxique est le calcul des distances numériques entre la requête et une réponse potentielle. La distance est définie selon le modèle utilisé, il peut être par exemple un ou un ensemble de valeurs, calculé numériquement en utilisant la représentation de la requête et la représentation d'une réponse potentielle. La sélection des réponses est basée sur les valeurs de la distance, par exemple, en définissant un seuil de distance ou la sélection de N premières réponses après le classement. Nous pouvons distinguer deux approches principales pour calculer la similitude syntaxique:

Signature visuelle Utilisation: Ces approches sont principalement utilisées dans les systèmes CBIR. Le principe est de formuler mathématiquement les caractéristiques appelées signatures. Ces signatures sont extraites des images pour utiliser une technique de calcul à distance pour en trouver un ou un ensemble de valeurs représentant la similitude.

Modèles de récupération des informations: le principe est simple, il consiste à utiliser du texte sous forme de mots clés, de balises, de paragraphes ou d'autres formulaires pour indexer les images. L'idée est de transformer la récupération d'image en une recherche d'informations pour utiliser des techniques et des modèles de récupération d'informations.

1.7. Architecture générale d'un système d'indexation

Classiquement, un système de recherche d'images par contenu visuel comporte une phase hors ligne d'indexation de la base d'images et une phase en ligne de recherche, que nous décrivons dans les deux sections suivantes. La figure 1.8, représente l'architecture générale d'un système d'indexation et recherche d'images par le contenu, ce système s'exécute avec les deux étapes : l'étape d'indexation et l'étape de recherche

1.7.1. L'étape d'indexation

Le but l'organisation est la préparation de la base de données. Cette phase est dite hors-ligne, car elle s'effectue avant toute recherche. Elle englobe les traitements suivants :

- L'extraction de descripteurs caractéristiques des images.
- La construction des index à partir des descripteurs. Le but est de mettre en place des techniques permettant d'accéder le plus rapidement possible à n'importe quel descripteur lors de la recherche.

1.7.2. L'étape de recherche

Prend un vecteur requête en entrée et utilise un algorithme qui profite pleinement de l'étape d'indexation, confinant ainsi la recherche de données similaires. En d'autres termes ; pendant la consultation de la base, l'utilisateur sélectionne une image par le biais d'une interface graphique. Les index ou signatures de la requête sont confrontés aux index des images de références, ainsi le système sélectionne et présente à l'utilisateur les images les plus similaires à la requête.

- **Recherche d'image par croquis (query by sketch)** : l'utilisateur dessine sa requête à l'aide d'une interface graphique pour chercher les images qui lui ressemblent dans la base. On distingue deux types de dessin :

- **Le croquis:** l'utilisateur décrit ce qu'il désire en représentant précisément les contours des objets, généralement en une seule et unique couleur, seul l'aspect forme porte l'information.

- **Recherche d'images par objets** : l'utilisateur décrit les caractéristiques d'un objet d'une image plutôt que l'image entière, le but est de rechercher un objet précis dans une série d'images. Un exemple de ce type de requête est de trouver les images contenant une personne. On peut aussi combiner plusieurs objets et spécifier les relations spatiales entre les différents objets. Un exemple de ce type d'interrogation est de trouver les images où une personne est assise à côté d'un arbre.

- **Recherche par l'exemple** : Dans ce cas, le système a besoin de comparer un exemple de même type avec la base pour produire les images similaires. Cette méthode est simple naturelle et ne nécessite pas de connaissances approfondies pour manipuler le système. Elle est donc bien adaptée à un utilisateur non spécialiste. Après la formulation, le système représente la requête par sa signature. Les mesures de similarités/dissimilarités entre la signature de la requête et celle de toutes les images de la base sont calculées et comparées. Le résultat est le plus souvent présenté sous forme d'une liste des images de similarité descendante. Comme toutes les images de la base doivent être examinées pour retrouver des images similaires à la requête.

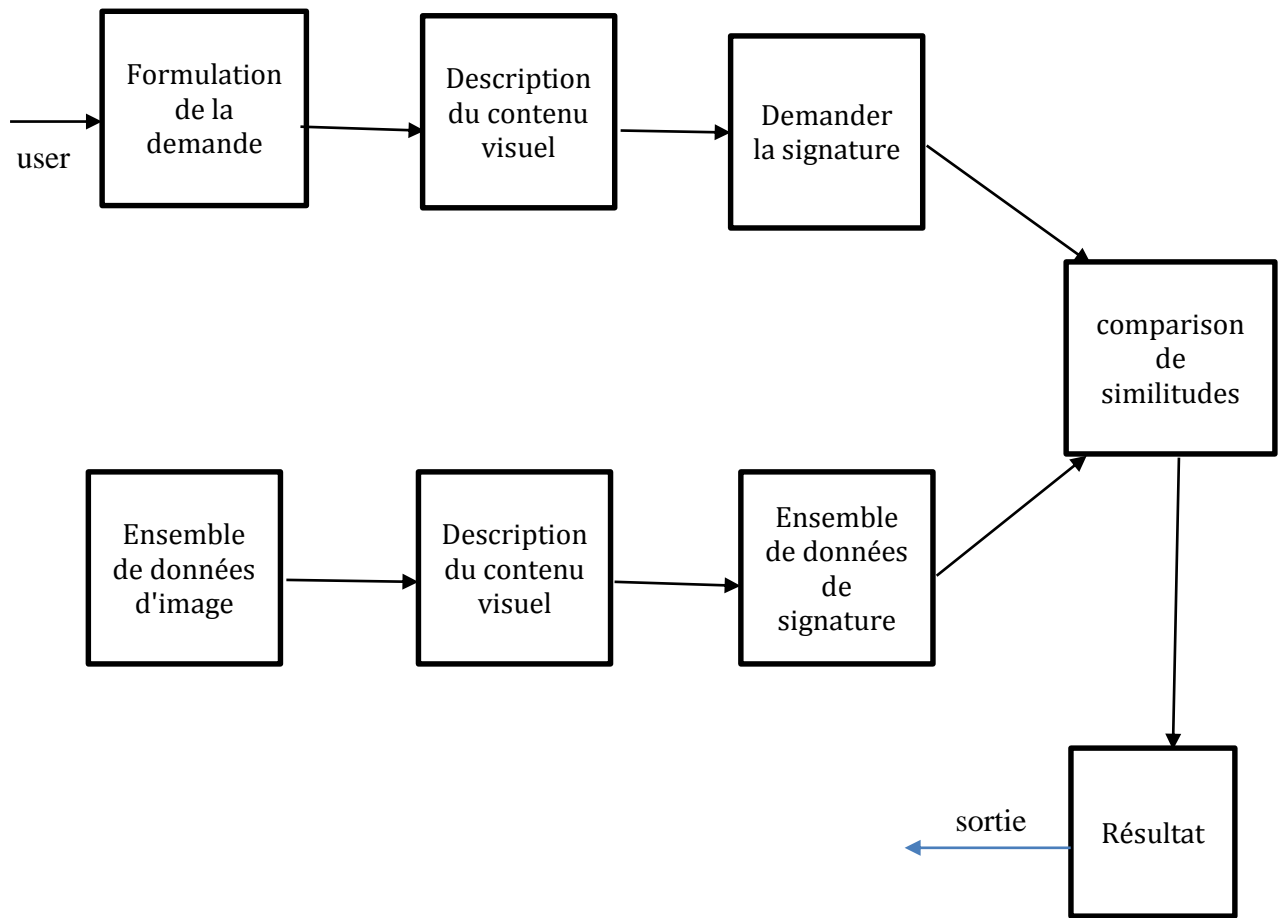


Figure I.7. Diagramme pour CBIR typique

I.8. annotation d' Image

La description des images se fait grâce à des techniques d'annotation. Cette section sera consacrée à ces techniques d'annotation. Trois techniques d'annotation sont proposées dans la littérature : l'annotation manuelle, l'annotation automatique et l'annotation semi-automatique. L'annotation d'image est un processus de description d'image qui facilite l'accès aux images. Il est principalement utilisé par les techniques de récupération d'images basés sur sémantique mais peut également être utilisé par des techniques syntaxiques. La différence est que les techniques sémantiques utilisent un formalisme de représentation des connaissances pour définir une annotation et une base de connaissances pour enrichir et mieux interpréter les annotations. Pour décrire une image, nous pouvons y associer plusieurs types d'informations. Ces informations sont [11, 19]:

- **Meta-données** : indépendantes du contenu: ces informations ne décrivent pas le contenu des images. Ils décrivent le contexte comme l'auteur de l'image, la caméra utilisée, la résolution d'image, etc....
- **Contenu visuel Meta-Data**: ces informations décrivent le contenu visuel des images. Il peut être divisé en deux types:
 - **Meta-Data dépendant du contenu**: décrit les caractéristiques de faible niveau des images. Ce sont généralement des informations qui peuvent être extraites automatiquement telles que la couleur, la forme, la texture, etc...
 - **Meta-Data contenu descriptif**: décrit le contenu sémantique des images. Ces informations représentent les interprétations possibles qu'un homme peut donner aux images.

1.8.1. Annotation Manuelle

L'annotation manuelle d'images consiste à faire annoter des bases d'images par un ou plusieurs annotateurs. Lors d'une annotation manuelle, l'annotateur assigne une description à chaque image en fonction de sa perception. Plusieurs systèmes d'annotation manuelle des images ont été présentés dans la littérature [20], [21] et [22].

L'annotation des images basée sur les mots-clés est un processus qui permet de décrire les images avec des termes dans le but de faciliter leur récupération. Agosti

et al. dans [23] et Bradley dans [60] ont fourni une étude exhaustive sur les contours et la complexité des annotations.

Un mot-clé est un mot associé à une image qui, une fois indexé, permet d'identifier l'image dans une base de données. Il permet de faire une description textuelle du contenu des images. Comme tous les autres descripteurs, les mots-clés permettent d'une part de décrire les images, et d'autre part d'accéder aux images par des requêtes textuelles. Comparativement aux descripteurs de bas niveau (couleur, texture, forme, etc.), les mots-clés sont considérés comme des descripteurs efficaces. Grâce aux mots-clés, l'annotation, la récupération et même la comparaison des images deviennent possible. Plusieurs approches se sont focalisées sur la description et la récupération des images via les mots-clés. Les premiers systèmes apparus dans les années 90 sont basés sur le texte [24], [25], [26]. Ces derniers adoptent une approche qui consiste à décrire le contenu visuel sous forme textuelle. Les mots-clés utilisés servent d'index pour accéder aux données visuelles associées. L'avantage de cette approche est qu'elle permet de consulter les bases de données en utilisant les langages d'interrogation standard, par exemple SQL. Plusieurs approches sont basées sur la description du contenu visuel des images.

Une autre méthode consiste à combiner le contenu visuel et le contenu textuel. Contrairement au contenu visuel décrit par les mots-clés, le contenu textuel représente l'ensemble des attributs spécifiques aux images, à savoir : les titres, les légendes, les commentaires, etc. qui entourent l'image. La combinaison du contenu visuel et textuel a été proposée par plusieurs chercheurs parmi lesquels nous pouvons citer Jin et al. dans [27], Gao et al. dans [28] et Ameesh et al. dans [29]. Jin et al. ont réalisé l'annotation des images en utilisant un modèle de langage cohérent parce que les mots-clés seuls seraient indépendants. Pour eux, il y a nécessité de prendre en compte d'autres informations comme les textes qui entourent les images, afin de produire une bonne description de l'image. Nous avons également les annotations à plusieurs niveaux proposées dans [28] qui visent non seulement à identifier les objets spécifiques dans une image, mais intègrent également les concepts pour regrouper des éléments similaires. Par exemple Ameesh et al. proposent que l'image d'un bureau peut être annotée non seulement

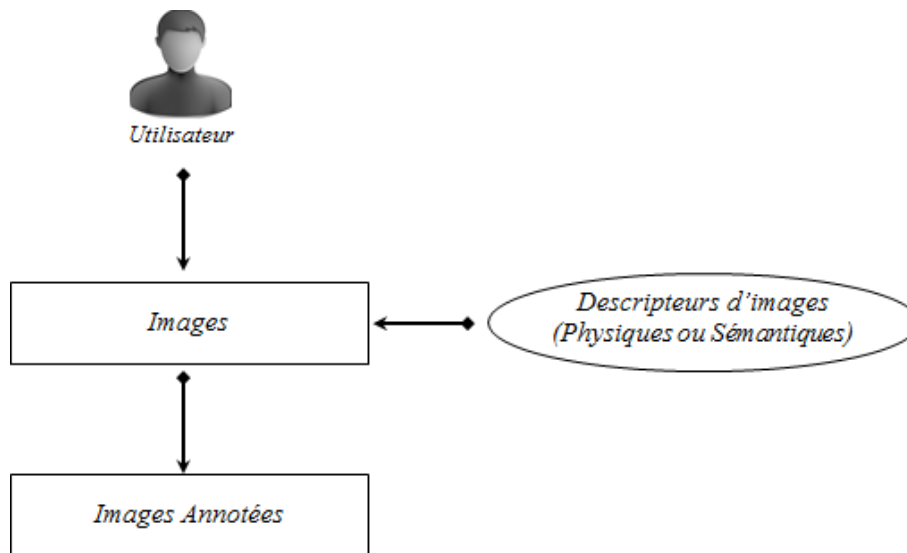


Figure 1.8 – Processus d’annotation manuelle avec les éléments de l’image, comme un écran d’ordinateur et tapis de souris, mais aussi le bureau et l’intérieur du bureau.

Généralement, le processus de l’annotation manuelle comporte l’annotateur (ou les annotateurs) qui décrit l’image, l’image à décrire et les descripteurs fournis pour mener à bien la description comme le montre la figure 1.8

I.8.2. Annotation automatique

L’annotation automatique est un processus qui consiste à faire annoter les images par un système d’annotation. Le processus d’annotation automatique a généralement lieu lorsqu’une ou plusieurs nouvelles images sont ajoutées dans la base de données.

Notre objectif étant d’associer des symboles à des images en se basant sur des exemples fournis d’images annotées, nous adoptons les approches basées sur l’apprentissage supervisé.

Plusieurs travaux ont été déjà menés dans le domaine de l’annotation automatique. Parmi ces travaux, nous pouvons citer les méthodes proposant

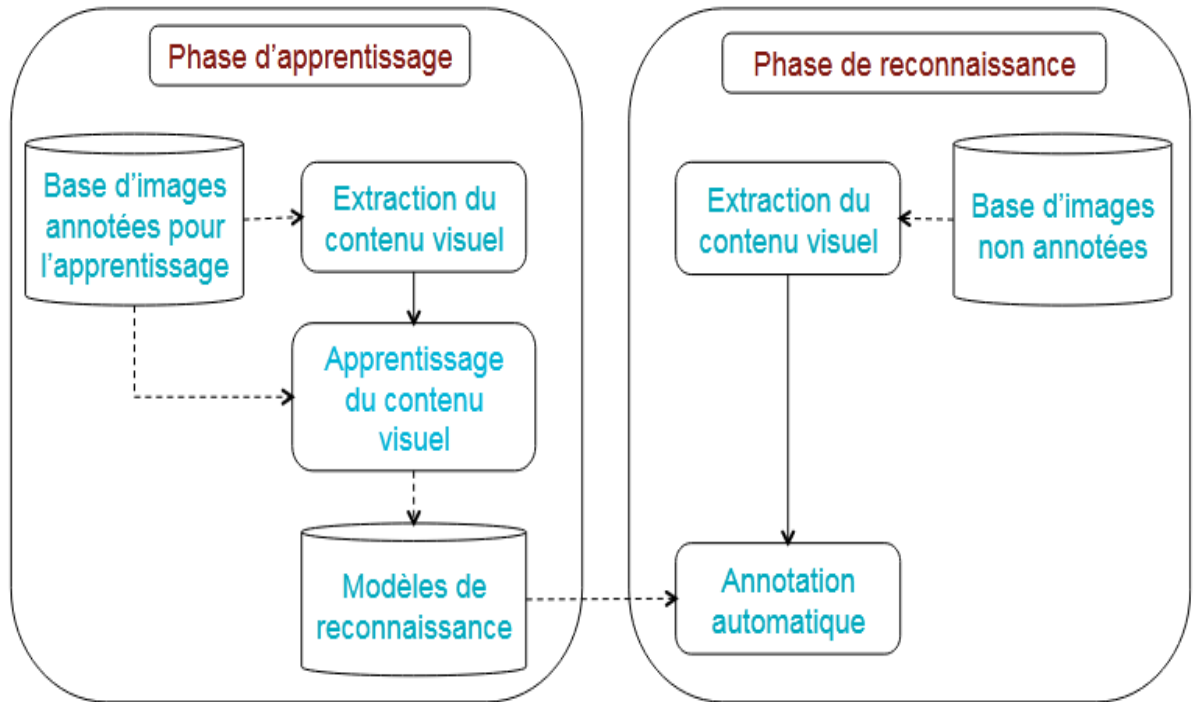


Figure I.9. Architecture générale d'une approche d'annotation automatique d'image.

Une approche d'annotation automatique d'image comprend deux phases principales :

- *Une phase d'apprentissage* : chargée d'apprendre les relations entre des descripteurs visuels (signaux visuels de bas niveau) et des annotations symboliques associées à ces descripteurs. L'apprentissage est effectué en s'appuyant sur des collections d'images annotées sur lesquelles une extraction de contenus visuels est effectuée. Cette phase génère des modèles de reconnaissance des symboles d'annotation. Un tel modèle attribue des scores de reconnaissance aux descripteurs du contenu visuel, ces scores expriment des relations entre les descripteurs et les symboles d'annotation.

- *Une phase de reconnaissance* : prenant comme entrée des images non annotées, et propose des annotations symboliques en s'appuyant sur les contenus visuels des images d'entrée et les modèles de reconnaissance appris dans la phase d'apprentissage.

I.9. Les différents types de l'apprentissage

L'apprentissage est un processus visant à améliorer les performances d'un système en se basant sur ses expériences passées. Trois grandes familles de techniques d'apprentissage sont répertoriées [30] :

I.9.1. *l'apprentissage supervisé* : les variables du phénomène à étudier peuvent être divisées en deux groupes : les variables observées (les descripteurs visuels) et des variables d'annotation (les symboles d'annotation). Le but de l'apprentissage est de déterminer les relations entre les variables observées et les variables d'annotation. . Pour appliquer un apprentissage supervisé il faut avoir un ensemble de vérités terrain (considérées correctes) contenant des variables d'annotation associées à des variables observées. Et en analysant cet ensemble (appelé ensemble d'apprentissage) ; la méthode d'apprentissage supervisé génère des modèles de reconnaissance. Les modèles de reconnaissance estiment des relations entre des variables observées non vues précédemment et des variables d'annotation.

I.9.2. *L'apprentissage non supervisé* : ce paradigme est plus proche de l'esprit de la fouille de données. Les variables d'annotation ne sont pas connues, et toutes les variables observées sont traitées de la même façon. L'objectif de cet apprentissage est de retrouver des phénomènes que se répètent dans les variables observées, afin de les regrouper dans des catégories inconnues a priori.

- **1.9.3. *L'apprentissage par renforcement*** : C'est un apprentissage qui nécessite un superviseur qui dit à l'agent quelle action est correcte dans telle situation [31]. Dans l'apprentissage par renforcement, l'agent interagit avec l'environnement qui lui donne un retour quantitatif sur les valeurs de ses actions. L'objectif de l'apprentissage par renforcement est alors de générer à partir d'expériences (état courant, action, état suivant) une politique maximisant le rendement au cours d'une période donnée

CHAPITRE II

**Présentation des modèles
d'apprentissage de l'ontologie**

II.1.Introduction

Récemment, plusieurs travaux se sont intéressés à l'utilisation des avancées scientifiques dans le domaine de l'intelligence artificielle pour franchir le problème de fossé sémantique. Ces systèmes se basent sur les connaissances de domaine représentées par des ontologies. Ces dernières offrent une compréhension commune, structurée et partagée d'un domaine ou d'une tâche, qui peut être utilisée pour la communication entre les personnes et les machines. Les données sur le Web constituent une grande source d'informations représentées sous des formes utiles pour l'être humain mais difficiles pour des traitements automatiques. Le contenu des ressources du Web peut être décrit en utilisant les méta-données formelles. Ces dernières doivent s'appuyer sur des ontologies afin de pouvoir les exploiter par des machines et les munir de sémantique.

L'objectif principal d'une ontologie est de modéliser un ensemble de connaissances dans un domaine donné. Elles jouent ainsi un rôle important dans le Web sémantique et sont employées comme une forme de représentation formelle de la connaissance sur un sujet du monde réel et traduire, ainsi, un consensus explicite. Dans le Web sémantique, les ontologies sont utilisées pour, d'une part, modéliser des ressources du Web à partir des représentations conceptuelles des domaines concernés et, d'autre part, dans l'objectif de permettre un traitement automatiquement.

II.2. Interdépendance entre la connaissance et la langue

Une ontologie, en tant que système formel de signes, s'inscrit parfaitement dans cette présentation où elle « joue le rôle » de langue naturelle dans les systèmes d'information. Ainsi le triangle sémiotique peut être facilement transformé en triangle « ontologique » : les sommets (signifié, signifiant, référent) sont interprétés comme (concept, terme, instance), respectivement. Les relations entre les sommets ne changent pas. En informatique les sommets du triangle sémiotique peuvent être associés aux notions de données, d'information et de connaissances. Ici leur connexité se présente autrement, sous forme de la « pyramide de sagesse » (cf. figure II.2), dit DIKW (Data- Information-Knowledge-Wisdom)

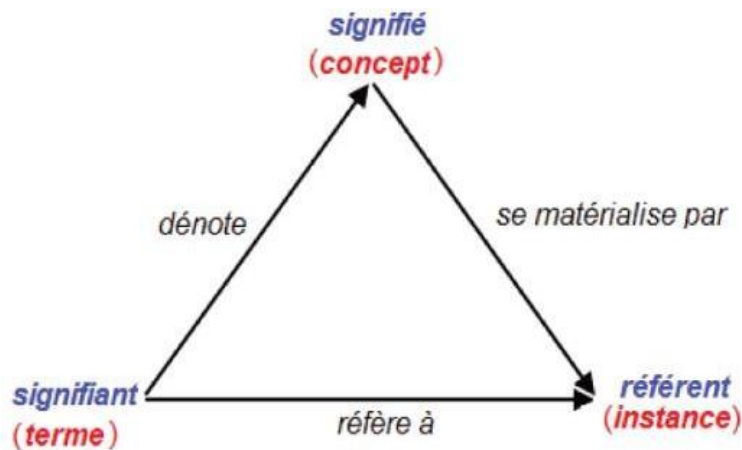


Figure II.1. le triangle sémiotique générique

II.3. Apparition d'ontologie

L'ingénierie des connaissances (IC) a pour but la résolution automatique des problèmes, alors que les Systèmes à Base de Connaissances (SBC) devraient permettre le stockage et la consultation et la modification des connaissances, ainsi que le raisonnement automatique. Le partage de connaissances entre systèmes informatiques permettra, de plus, une interaction et une coopération entre ces derniers et l'utilisateur humain. Cela se manifeste

II.4. La notion d'ontologie

Le terme "Ontologie" est un terme philosophique qui signifie « Partie de la métaphysique qui s'applique à l'être en tant qu'être, indépendamment de ses déterminations particulières » (Le Petit Robert). Avec l'émergence de l'ingénierie des connaissances (IC), l'ontologie est introduite en Intelligence Artificielle (IA) comme réponse aux problématiques de représentation et de manipulation des connaissances au sein des systèmes informatiques.

II.5. Définitions

Il est difficile d'attribuer à la notion d'ontologie une unique définition définitive puisque celle-ci a été employée dans des contextes différents. Neeches et al., (1991) furent les premiers à proposer une définition de l'ontologie à savoir : « une ontologie définit les termes et les relations de base du vocabulaire d'un domaine ainsi que les règles qui indiquent comment combiner les termes et les relations de façon à pouvoir étendre le vocabulaire »

Gruber (1993) a défini l'ontologie comme « une spécification explicite d'une conceptualisation ». Cette définition a été modifiée légèrement par Borst (1997) comme « spécification formelle d'une conceptualisation partagée ».

Ces deux définitions ont été combinées par Studer (1998) comme « spécification formelle et explicite d'une conceptualisation partagée » :

« Conceptualisation » fait référence à un modèle abstrait de certains phénomènes du monde, modèle qui identifie les concepts pertinents de ce phénomène.

« Explicite » signifie que le type des concepts utilisés et les contraintes sur leur utilisation sont explicitement définis.

« Formelle » se réfère au fait que l'ontologie doit être compréhensible par les machines.

« Partagée » reflète la notion de connaissance consensuelle décrite par l'ontologie, c'est-à-dire qu'elle n'est pas restreinte au point de vue de certains individus seulement, mais reflète un point de vue plus général, partagé et accepté par un groupe.

CHAPITRE II Présentation des modèles d'apprentissage de l'ontologie

Selon Swartout et al., (1997), une ontologie est « un ensemble de termes structurés de façon hiérarchique, conçue afin de décrire un domaine et qui peut servir de charpente à une base de connaissances ».

Sowa (1995) a mentionné que le but d'une ontologie est l'étude des catégories de concepts qui existent ou peuvent exister dans certains domaines. Le résultat de cette étude, appelé une ontologie, est un catalogue de types de choses qui existent dans un domaine d'intérêt dans la perspective d'utiliser un langage formel pour parler de domaine. On remarque que les définitions, malgré leur diversité, offrent des points de vue complémentaires.

Ainsi, une ontologie offre les moyens pour présenter les concepts d'un domaine en les organisant hiérarchiquement et en définissant leurs propriétés sémantiques dans un langage formel de représentation des connaissances.

II.6. Les différents types d'ontologies

On peut définir deux classes d'ontologies : (i) une catégorie basée sur la structure de la conceptualisation et (ii) une deuxième catégorie basée sur le sujet de la conceptualisation

Dans la première catégorie, on distingue trois sous catégories :

- Les ontologies terminologiques (lexiques, glossaires...).
- Les ontologies d'information (schéma d'une BD).
- Les ontologies des modèles de connaissances.

Dans la deuxième catégorie, on distingue quatre sous catégories:

- Les ontologies d'application : elles contiennent toutes les informations nécessaires pour modéliser les connaissances pour une application particulière.
- Les ontologies de domaine : elles fournissent un ensemble de concepts et de relation décrivant les connaissances d'un domaine spécifique.
- Les ontologies génériques (dites aussi de haut niveau) : elles sont similaires aux ontologies de domaine, mais les concepts qui y sont définis sont plus génériques et décrivent des connaissances tels que l'état, l'action, l'espace et les composants.
- Généralement, les concepts d'une ontologie de domaine sont des spécialisations des concepts d'une ontologie de haut niveau.

CHAPITRE II Présentation des modèles d'apprentissage de l'ontologie

- Les ontologies de représentation (méta-ontologies) : elles fournissent des primitives de formalisation pour la représentation des connaissances. Elles sont généralement utilisées pour écrire les ontologies de domaine et les ontologies de haut niveau. Exemples : Frame Ontology et RDF Schema Ontologie

II.7.les Constituants de l'ontologie

Comme nous l'avons abordé, les ontologies fournissent un vocabulaire commun d'un domaine et définissent la signification des termes et des relations entre elles. La connaissance dans les ontologies est principalement formalisée en utilisant les cinq types de composants [12] à savoir : **concepts** (ou classes), **relations** (ou propriétés), **fonctions**, **axiomes** (ou règles) et **instances** (ou individus).

- Les concepts, aussi appelés termes ou classe de l'ontologie, correspondent aux abstractions pertinentes d'un segment de la réalité (le domaine du problème) retenus en fonction des objectifs qu'on se donne et de l'application envisagée pour l'ontologie
- Les relations traduisent les associations (pertinentes) existant entre les concepts présents dans le segment analysé de la réalité. Ces relations incluent les associations suivantes :
 - Sous classes de (généralisation-spécialisation) ;
 - Partie de (agrégation ou composition) ;
 - Associe à ;
 - Instance de, etc.

Ces relations nous permettent d'apercevoir la structuration et l'interrelation des concepts, les uns par rapport aux autres ;

- Les fonctions constituent des cas particuliers de relations, dans laquelle un élément de la relation, (le nième) est défini en fonction des N-1 éléments précédents ;
- Les axiomes constituent des assertions, acceptées comme vraies, à propos des abstractions du domaine traduites par l'ontologie.
- Les instances constituant la définition extensionnelle de l'ontologie ; ces objets véhiculent les connaissances (statiques, factuelles) à propos du domaine du problème.

II.8. Classifications d'ontologies

Plusieurs critères de classification ont été proposés pour catégoriser une ontologie. Parmi ces classifications, on trouve l'objet de conceptualisation, le niveau de détail et le niveau de formalisation (informel, semi-formel, ou formel). La classification de [2] est basée sur le niveau de généralité et de dépendance de l'ontologie.

Nous distinguons quatre types d'ontologies (voir Figure II.3) :

Les ontologies de haut niveau, les ontologies de domaine, les ontologies de tâche et les ontologies d'application.

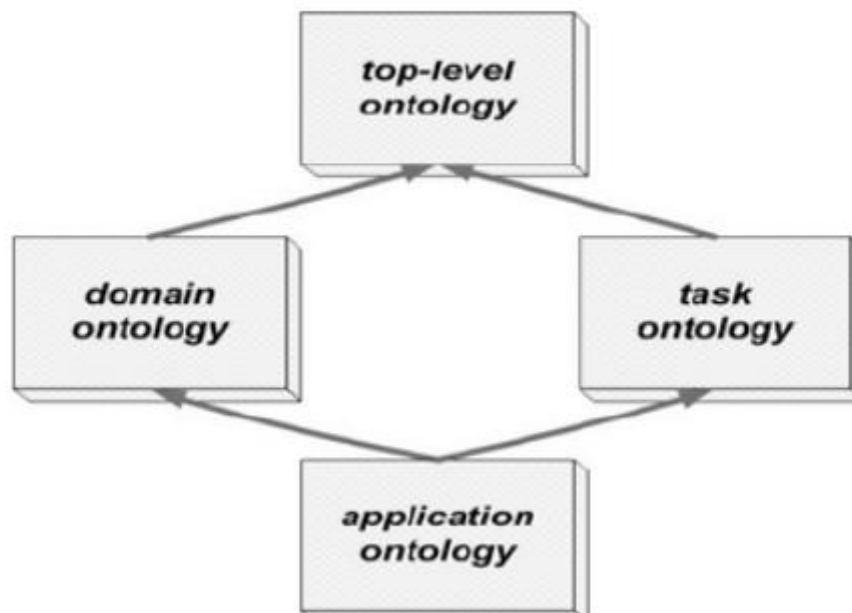


Figure 2.3. classification des ontologies.

CHAPITRE II Présentation des modèles d'apprentissage de l'ontologie

- **Les ontologies de haut niveau (ou « top-level ontology »)** : Ce type concerne une ontologie générique, globale et de plus haut niveau. Il est fondé principalement sur la théorie de l'identité et la théorie de la dépendance. De plus, il traite des sujets et des catégories universels tels que les entités, les concepts, les états, les événements les actions, le temps, les processus, les relations, et les propriétés. SUMO (Standard Upper Ontology) et DOLCE (Descriptive Ontology for Linguistic and Cognitive Engineering) sont les ontologies de haut niveau les plus connues [6] [7].
- **Les ontologies de domaine (ou « domain ontology »)** : Cette ontologie s'intéresse à la création des concepts relatifs à un domaine d'application cible tels qu'ils sont manipulés par les experts. Elle s'agit d'un méta-modèle de connaissance où les concepts et les propriétés dans ce cas sont de type déclaratif. La majorité des ontologies existantes sont des ontologies de domaine. Parmi les ontologies de domaine, nous pouvons citer TOVE (Ontologie sur la notion d'activité)
- **Les ontologies de tâche (ou « task ontology »)** : Ce type d'ontologie traite des tâches spécifiques à savoir les tâches de planification, de diagnostic, d'éducation et de configuration. Une ontologie de tâche crée un ensemble de vocabulaires servant à la résolution des problèmes liés aux tâches et indépendante du domaine. Cette catégorie d'ontologie est basée sur une architecture computationnelle d'un système qui exerce une tâche. L'ontologie de formation par ordinateur (« Computer Based Training Ontology ») et l'ontologie des objectifs d'apprentissage (« Learning Goal Ontology ») sont deux exemples d'ontologies de tâches.
- **Les ontologies d'application (ou « application ontology »)** : Elle s'intéresse non seulement aux concepts du domaine mais aussi aux tâches. Elle caractérise l'ontologie la plus spécifique permettant d'appliquer les concepts pour l'exécution d'une certaine tâche ou une activité

II.9. Domaines d'application des ontologies

Nous allons aborder brièvement les domaines où sont requises les ontologies, en tant que telles, et les domaines appliquant les principes méthodologiques fondant leur construction. La diversité des applications possibles montre l'actualité du problème de la construction automatisée des ontologies.

II.9.1. Recherche d'information (Information Retrieval, IR)

Ce domaine couvre des activités telles que :

— Recherche des documents qui contiennent les informations pertinentes, correspondant aux requêtes de l'utilisateur. La plupart des moteurs de recherche réalisent l'indexation des textes à l'aide du modèle vectoriel où chaque texte est présenté comme . Sac de mots . (*bag of words*). Les principaux inconvénients de cette approche sont :

1. la redondance des index, les mêmes notions étant d'nommées par les mots différents.
 - 2-les mots d'un document sont considérés comme indépendants, ce qui ne correspond évidemment pas à la réalité.
 - 3-.les mots sont polysémiques, ce qui induit des ambiguïtés et aboutit à résultats non pertinents pour l'utilisateur. On peut pallier ces inconvénients par l'utilisation d'une indexation conceptuelle à l'aide des ontologies de domaines ; les concepts sont alors associés aux termes correspondants et liés par des relations prédéfinies.
- Classification de documents, i.e. l'attribution de chaque document à l'une des catégories prédéfinies..
- Regroupement sémantique, i.e. rassemblement des documents dont les sujets sont proches. Dans ce cas, le travail principal est la définition des rubriques autour desquelles les documents doivent être réunis.
- Production de résumés automatiques.

II.9.2. Système d'information

L'intégration d'une ontologie dans un système d'information vise à réduire, voire éliminer, la confusion conceptuelle et terminologique à des points clés du système, et à tendre vers une compréhension partagée pour améliorer la communication, le partage, l'interopérabilité et le degré de réutilisation possible, ce qui permet de déclarer formellement un certain nombre de connaissances utilisées pour caractériser les informations gérées par le système, et de se baser sur ces caractérisations et la formalisation de leur signification pour automatiser des tâches de traitement de l'information.

CHAPITRE II Présentation des modèles d'apprentissage de l'ontologie

L'ontologie retrouve maintenant dans une large famille de systèmes d'information. Elle est utilisée pour :

- Décrire et traiter des ressources multimédia ;
- Assurer l'interopérabilité d'applications en réseaux .
- Piloter des traitements automatiques de la langue naturelle.
- Construire des solutions multilingues et interculturelles.
- Permettre l'intégration des ressources hétérogènes d'information.
- Vérifier la cohérence de modèles.
- Permettre les raisonnements temporel et spatial.
- Faire des approximations logiques , etc.. ..

Ces utilisations des ontologies se retrouvent dans de nombreux domaines d'applications tel que :

- Intégration d'information géographique.
- Gestion de ressource humaine .
- Aide à l'analyse en biologie, suivi médicale informatisé .
- Commerce électronique .
- Enseignement assisté par ordinateur .
- Bibliothèque numériques .

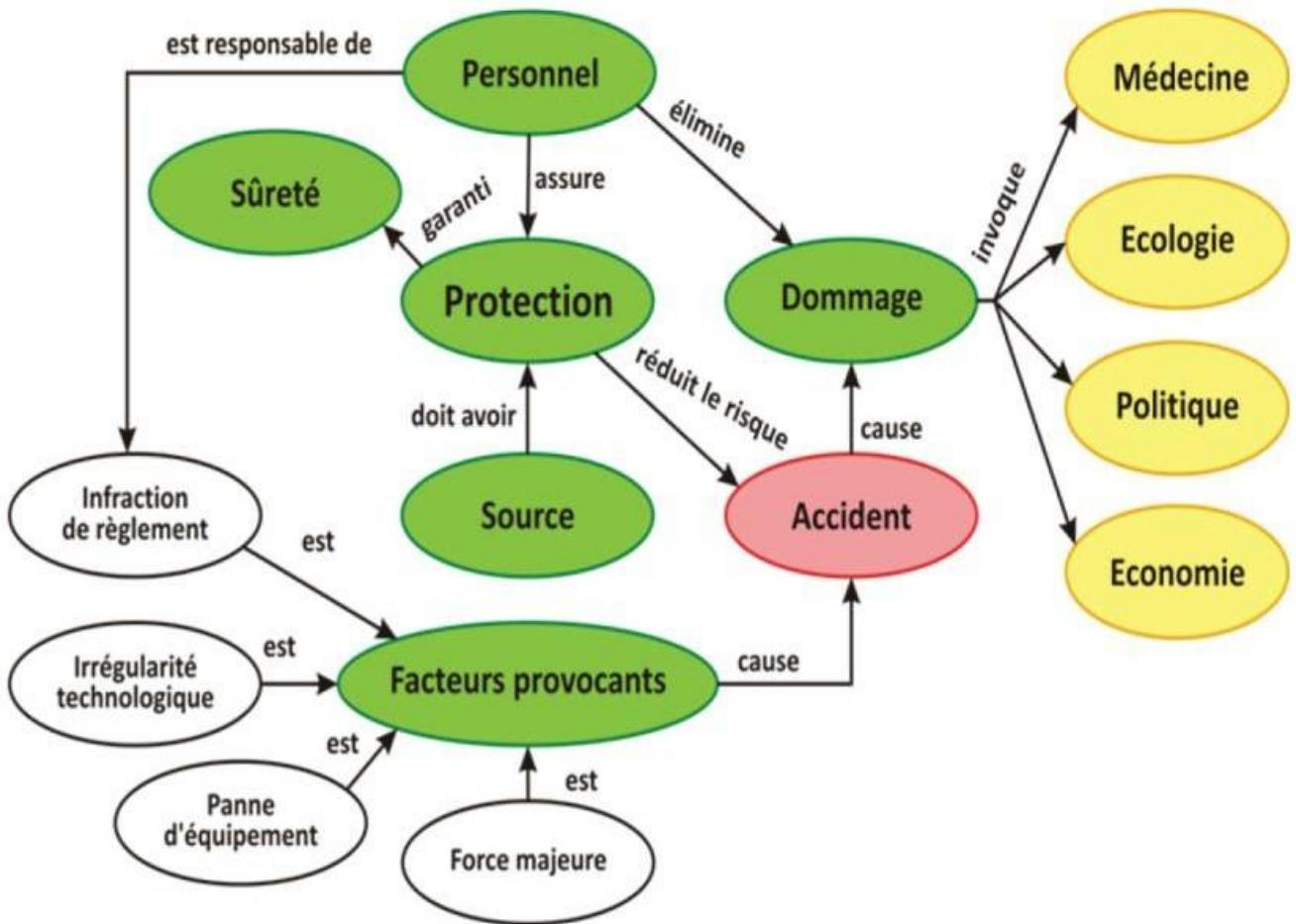


Figure II.2. Fragment du modèle de domaine.

CHAPITRE III

**Outils et résultats
expérimentaux**

III.1.Introduction

Dans ce chapitre, nous avons résolu deux principaux problèmes de détection des relations (une grande divergence intra-classe et la dépendance sémantique ou un écart sémantique), et utiliser un ensemble de données HCVRD qui a déjà résolu le troisième problème (c'est-à-dire un problème de longue queue).

Nous avons proposé un modèle sémantique ontologique pour filtrer les faux négatifs /positifs prenant l'avantage des connaissances de base riches et fortes de l'ontologie, qui a aidé à filtrer les propositions de classe d'objets inappropriées, et également du module de relation sémantique qui a pénalisé la relation inappropriée des objets humains et garantit des affectations profondes et strictes Il a été offert par l'ontologie sémantique qui aide à améliorer la qualité des affectations relationnelles.

III.2. Le problème de long-tail

HCVRD est un nouvel ensemble de données qui considère la relation entre les humains et les objets dans les images. Le principal avantage de l'utilisation du HCVRD est qu'il est construit avec certaines contraintes qui visent à éviter de tomber dans le problème de distribution à longue queue et le problème de tirs zéro. HCVRD a été collecté à partir du grand ensemble de données du génome visuel (c'est-à-dire la recherche d'images contenant des humains). En outre, un composant supplémentaire de 788 160 images tiré des 100 premiers résultats de recherche d'images pour chaque triple relation. Au total, il y a 52 855 images avec 1 824 catégories d'objets et 927 prédicats. 256 550 Instances de relations avec 9 852 types de relations non tissées et 18 471 types de relations avec des coups zéro. Il existe en moyenne 10,63 prédicats par catégorie d'objets. 31 586 images sont utilisées pour l'entraînement et deux divisions de test sont utilisées. Le premier test de test contient 10 000 images où toutes les relations se produisent dans l'ensemble de formation. Une autre scission de test comprend toutes les relations zéro-shot.

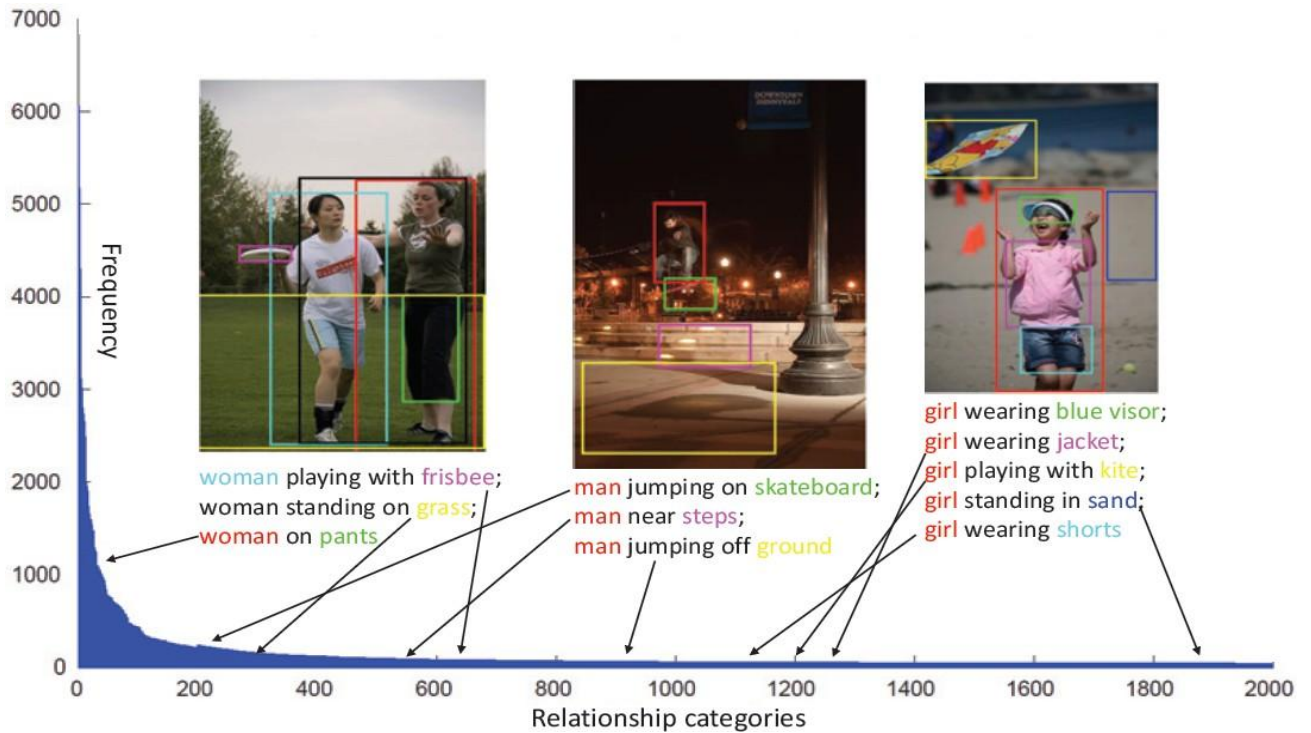


Figure III.1. La distribution d'étiquette à longue queue de l'ensemble de données HCVRD

III.3. La large divergence intra-classe

La divergence intra / interclasse est deux autres problèmes qui ont un impact énorme sur l'extraction d'informations sémantiques dans les images. Pour la divergence intra-classe, les objets (c'est-à-dire les prédicats) appartiennent à la même classe mais ne peuvent pas être représentés avec les mêmes caractéristiques visuelles; Dans un autre terme, ce sont les variations visuelles ou plusieurs représentations du même objet. Un exemple est illustré dans la figure suivante (figure III.2), l'avion volant est différent d'un avion cassé!



Figure III.2. Ambiguïté visuelle. Le même contenu visuel ou deux contenus visuels similaires peut se référer à deux significations différentes.

2. VRD: La détection des relations visuelles avec les périeurs de langage (VRD) (298) est un ensemble de données qui a récemment atteint les progrès de la teinte dans la détection des relations visuelles. Il contient 5000 images avec 37 993 mille relations, 100 catégories d'objets et 70 catégories de prédicat reliant ces objets ensemble. Les prédicats sont classés en 5 types suivants: action, spatial, préposition, comparatif, verbe.

3. VG: le génome visuel (VG) (299) est un grand ensemble de données. Il contient 99 658 images avec 200 catégories d'objets et 100 prédicats. Il y a totalement 1 174 692 instances de relation entre 19 237 triplets uniques. La scission par défaut contient 73 801 images pour la formation et 25 857 images pour les tests.



Figure III.3. Différence entre les caractéristiques visuelles du même objet (exemple de la classe d'avion)

4. VGG-16 et plus rapide-RCNN: Le module d'extraction de caractéristique est une pile de couches de convolution et de couches de mise en commun max - qui ont la même configuration que le VGG - 16 (174). Le module de détection est dans le style de RCNN plus rapide (292), qui est utilisé pour détecter l'objet et le sujet humain (dans sa sous-catégorie).

5. WordNet: WordNet Ontology (291, 300) est une base de données lexicale de relations sémantiques entre les mots dans plus de 200 langues. Le pouvoir de WordNet dépasse sa capacité à trouver le sens des phrases en utilisant des relations sémantiques (c'est-à-dire les synonymes, les hyponymes, les hypernymes, les meronymes, les holonymes, les impliqués, etc.). À l'aide de WordNet, un programme d'ordinateur sera en mesure d'identifier les classes de relations et d'objets.

WordNet a été utilisé à plusieurs fins dans les systèmes d'information, notamment la désambiguïsation des mots, la récupération de l'information, la classification des relations automatique. Pour nous, WordNet Ontology est utilisé comme arrière-plan pour soutenir les informations visuelles extraites du processus d'apprentissage en profondeur. Il nous fournit une solide connaissance des relations entre les objets détectés pour la compréhension automatique de l'image.

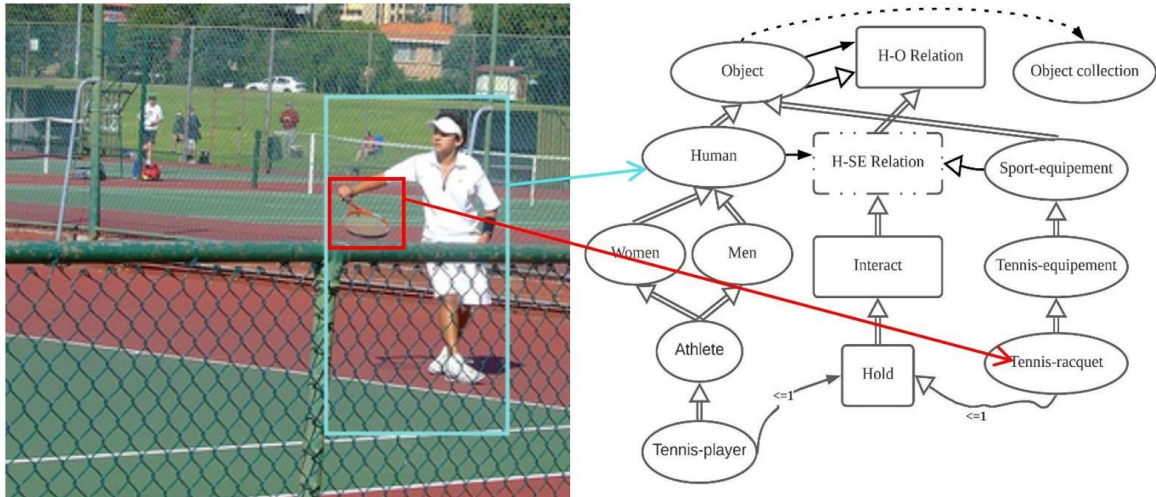


Figure III.4 Exemple d'une ontologie générée à partir de l'utilisation des objets détectés dans l'image.

L'image comme «raquette», tandis que le module d'ontologie statistique proposé l'a qualifié de «tennis-joueur» et de «tennis-racquet».

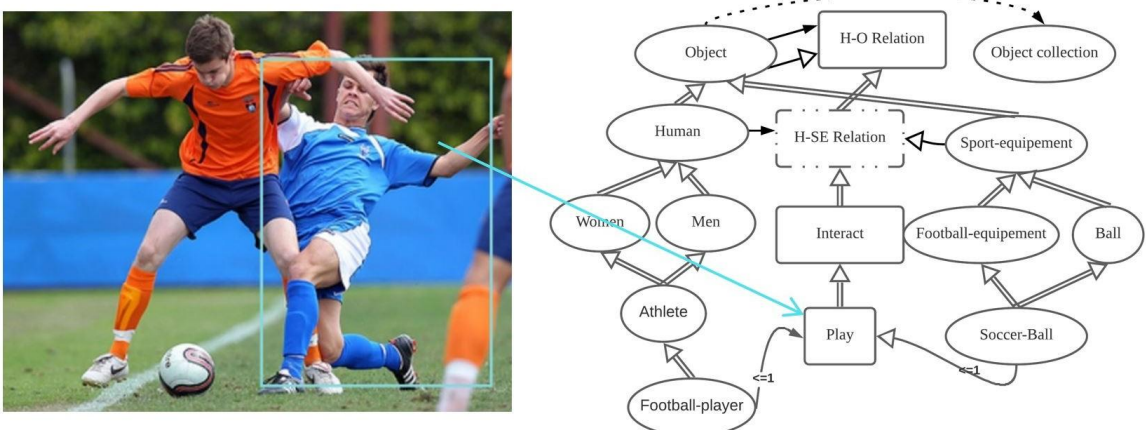


Figure III.5. Exemple Les données explorées du génome l'ont étiquetée comme «humain joue» et «une balle», tandis que l'utilisation du module statistique de l'ontologie avec l'avantage du riche anté Ballon de football

CHAPITRE IV

Conception et implémentation

CHAPITRE IV- Conception et implémentation

IV.1.Introduction

Dans ce chapitre, nous allons faire la conception de notre application. Nous allons présenter Aussi la mise en œuvre de notre application en utilisant le langage PYTHON . en commençant Tout d'abord par une présentation du langage de programmation choisi. Ensuite nous présentons des captures d'écran de l'exécution de notre application.

IV.2. Présentation des outils de développement

IV.2.1.Matériel

Le matériel réalisé est pc personnel TOSHIBA avec un 8GB capacité mémoire, et un processus INTEL CORE i5, avec windows 13 édition intégral, service pack1 64 bit type système.

IV.2.2. Langage de programmation PYTHON

Le langage Python est devenu si célèbre que chaque domaine et secteur en est désormais un utilisateur. Même si les autres langages de programmation ne perdent pas leurs fans, Python augmente sa base de fans. Par conséquent, de plus en plus de personnes aspirent maintenant à apprendre le python. Certaines des raisons pour lesquelles une certification en Python peut être utile sont décrites ci-dessous:

➤ **Simplicité**

Enfin, la principale raison pour laquelle les programmeurs utilisent Python est qu'il s'agit d'un langage de programmation simple. C'est un langage convivial pour débutant, car il ne nécessite pas beaucoup de codes et de syntaxes complexes qui ne sont pas compréhensibles. Python a une syntaxe et un codage faciles et lisibles, ce qui facilite beaucoup sa configuration et son utilisation

➤ **Apprentissage automatique (machine learning)**

Aujourd'hui, presque tout se passe dans les algorithmes, qu'il s'agisse d'un moteur de recherche, de médias sociaux, de robots de chat, d'assistants personnels virtuels, etc. Ces algorithmes sophistiqués sont le résultat d'un apprentissage automatique et ont transformé l'ensemble du scénario technologique. Avec l'apprentissage automatique, le principal langage de programmation utilisé est Python. On peut trouver de nombreuses bibliothèques.

➤ Bibliothèques

Les bibliothèques sont vraiment utiles pour le développement d'applications et de sites Web. On peut trouver n'importe quel type de code. Python a un grand nombre de frameworks et de bibliothèques comme Flask, Django, NumPy, Scipy, Pandas, Tensorflow, Keras, etc. Il faut se concentrer sur la logique et l'objectif, les codes sont facilement disponibles dans les bibliothèques.

➤ développement web

De nos jours, comme Reddit, de nombreux sites Internet sont développés avec le langage Python. La principale raison pour laquelle le langage de programmation Python est utilisé dans le développement Web est sa rapidité et son efficacité. L'utilisation de PHP pour développer un site Web peut prendre des heures, tandis que l'utilisation de Python ne prend que quelques minutes. En outre, il existe des frameworks et des bibliothèques comme Django et Flask qui facilitent grandement le travail.



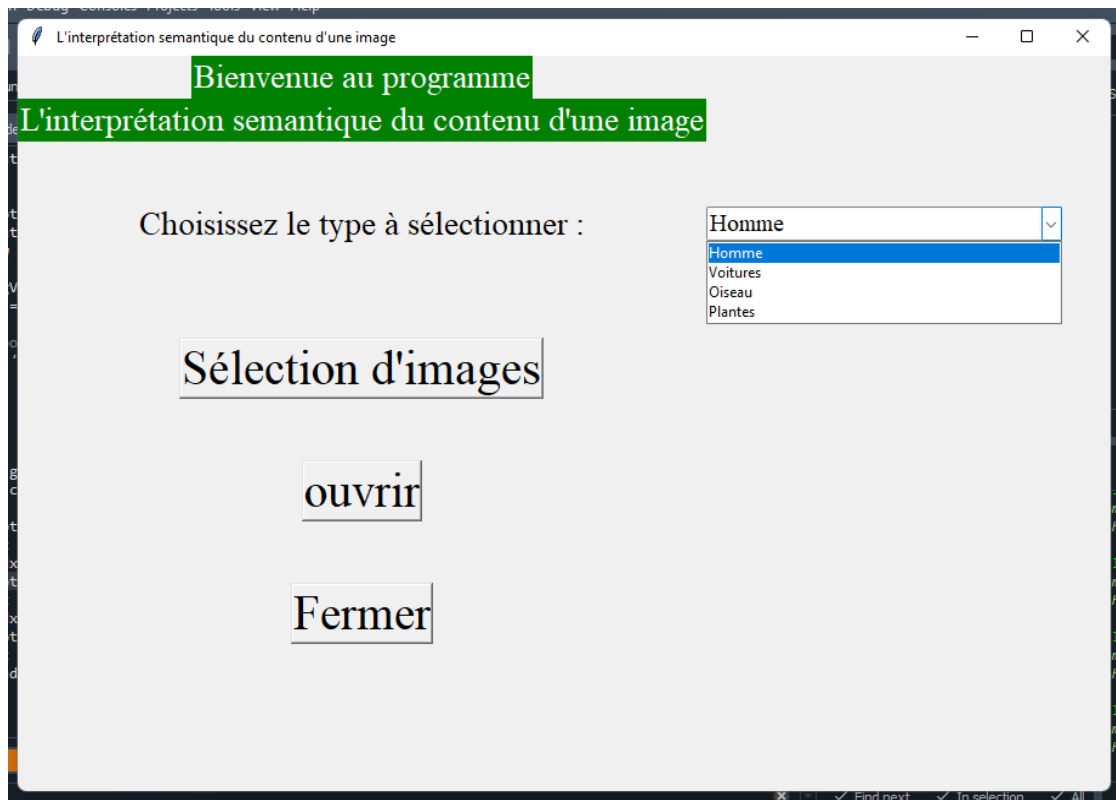


Figure. IV.1. Interface d'application

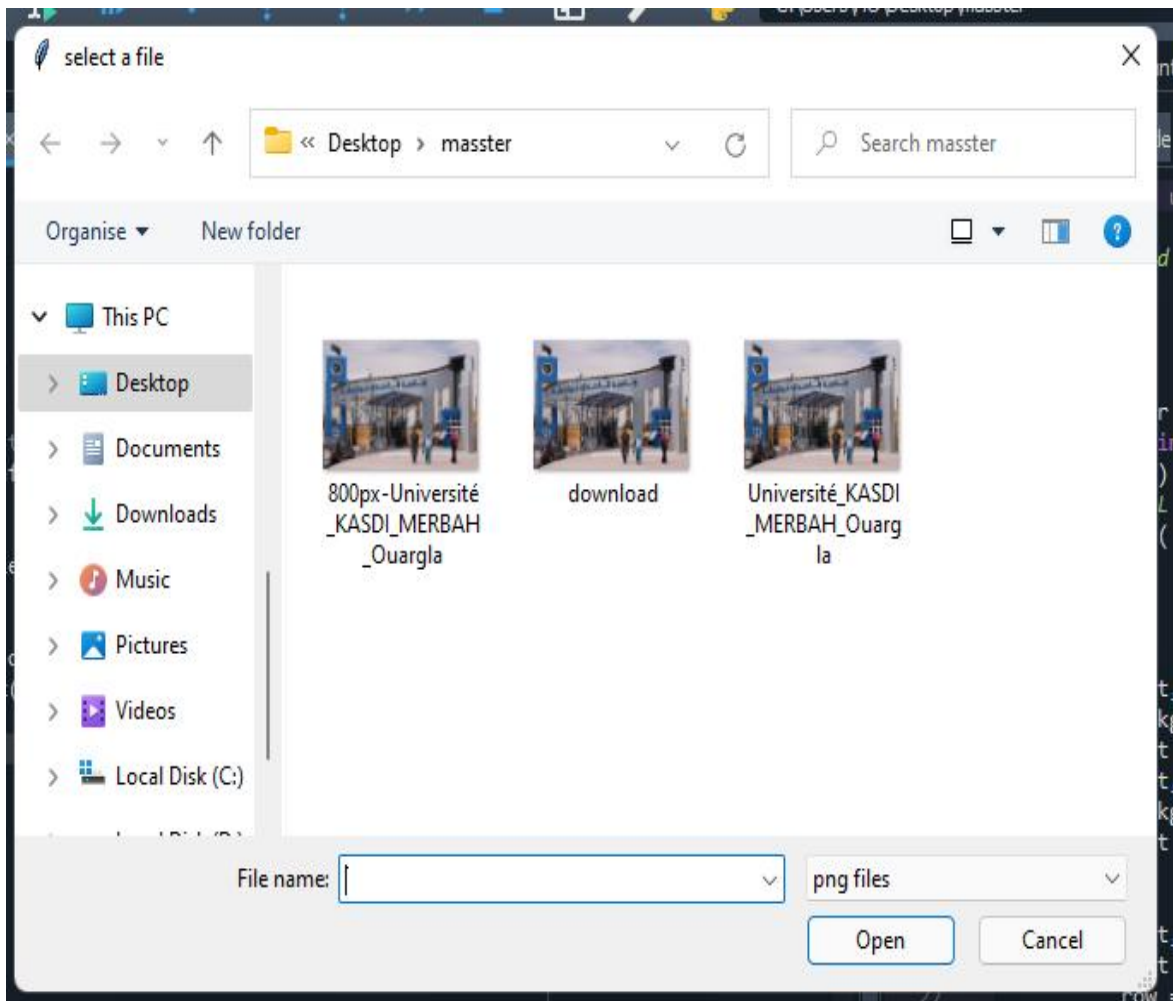


Figure IV.2. Recherche de l'image.



Figure IV.3. Affichage de l'image sélectionnée

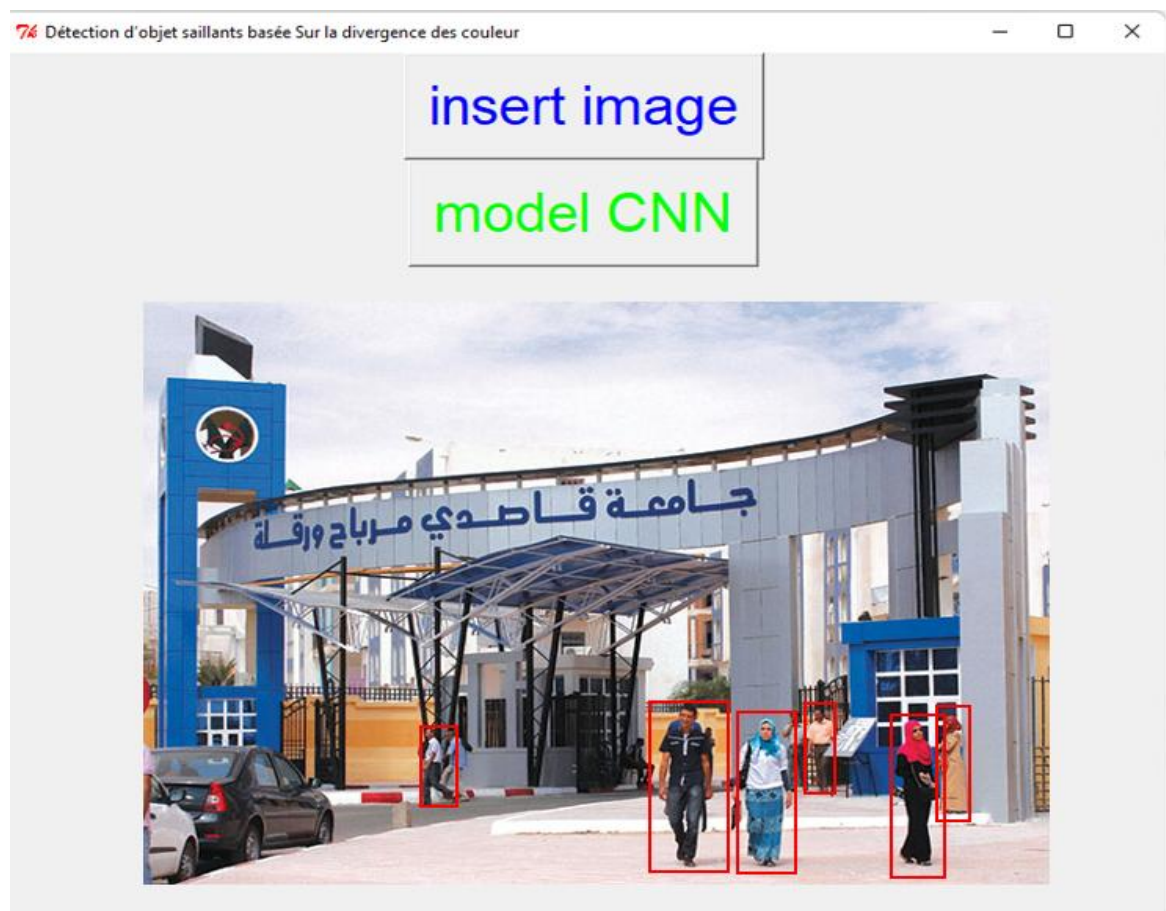


Figure.IV.4. Insertion l'image avec model CNN

CONCLUSION

ET

PERSPECTIVE

La classification d'image est une tâche importante dans le domaine de la vision par ordinateur, la reconnaissance d'objets et l'apprentissage automatique.

Bien que les capacités des activités réalisées dans le domaine de classification des images soient nombreuses, aucune méthode n'est jugée faible à 100%, mais au fur et à mesure les nouveaux travaux essayent d'améliorer les scores pour des meilleurs résultats.

Dans notre travail, nous avons résolu deux principaux problèmes de détection des relations (une grande divergence intra-classe et la dépendance sémantique ou un écart sémantique), et utiliser un ensemble de données HCVRD qui a déjà résolu le troisième problème (long-tail problème). Nous avons proposé un modèle sémantique ontologique pour filtrer les faux négatifs positifs en utilisant. Prenant l'avantage des connaissances de base riches et fortes de l'ontologie, qui a aidé à filtrer les propositions de classe d'objets inappropriées. Et également du module de relation sémantique qui a pénalisé la relation inappropriée des objets humains et garantit des affectations profondes et strictes. Le résultat obtenu lors de la phase de test confirme l'efficacité de notre approche.

En tant que perspectives, nous visons dans les travaux futurs pour proposer l'utilisation du même modèle ontologique et nous développer de telle manière que, un modèle peut dire si

- 1) les piétons traversent la rue,
- 2) combien ils sont loin d'un conducteur, et
- 3) Est-il sûr ou de ne pas continuer à conduire dans un état. Toutes les perspectives décrivent la sécurité des voitures autonomes.

REFERENCES

.RÉFÉRENCES

- [1] ISHWAR K SETHI, IOANA L COMAN, AND DANIELA STAN. **Mining association rules between low-level image features and high-level concepts.** *Data mining and knowledge discovery: theory, tools, and technology III*, 4384:279–290, 2001.
- [2] NAJLAE IDRISSE, JOSÉ MARTINEZ, AND DRISS ABOUTAJDINE. **Bridging the semantic gap for texture-based image retrieval and navigation.** *Journal of Multimedia*, 4(5), 2009.
- [3] THOMAS R GRUBER. **A translation approach to portable ontology specifications.** *Knowledge acquisition*, 5(2):199–220, 1993.
- [4] ZHILIANG MA, ZHE LIU, AND ZHENHUA WEI. **Formalized representation of specifications for construction cost estimation by using ontology.** *Computer-Aided Civil and Infrastructure Engineering*, 31(1):4–17, 2016.
- [5] THOMAS HUANG. **Computer vision: Evolution and promise.** 1996.
- [6] HAIHUA FENG, DAVID A CASTANON, AND WILLIAM CLEMENT KARL. **A curve evolution approach for image segmentation using adaptive flows.** *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*, 2:494–499, 2001.
- [7] TIM MORRIS. *Computer vision and image processing.* Palgrave Macmillan Ltd, 2004
- [8] RANJAY KRISHNA, YUKE ZHU, OLIVER GROTH, JUSTIN JOHNSON, KENJI HATA, JOSHUA KRAVITZ, STEPHANIE CHEN, YANNIS KALANTIDIS, LI-JIA LI, DAVID A SHAMMA, MICHAEL BERNSTEIN, AND LI FEI-FEI. **Visual Genome: Connecting Language and Vision Using Crowdsourced Dense Image Annotations.** 2016.
- [9] HIDEYUKI TAMURA, SHUNJI MORI, AND TAKASHI YAMAWAKI. **Textural features corresponding to visual perception.** *IEEE Transactions on Systems, man, and cybernetics*, 8(6):460–473, 1978.
- [10] ALLAN HANBURY. **A survey of methods for image annotation.** *Journal of Visual Languages & Computing*, 19(5):617–627, 2008.
- [11] MANISH BHATTARAI, D. OYEN, JUAN CASTORENA, L. YANG, AND B. WOHLBERG. **Diagram Image Retrieval using Sketch-Based Deep Learning and Transfer Learning.** *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 663–672, 2020.
- [12] LAURENT GUIGUES, ROGER TRIAS-SANZ, NESRINE CHEHATA, FRANCK TAILLANDIER, AND MATTHIEU DEVEAU. **B. 4 Segmentation.**
- [13] LUIS VON AHN AND LAURA DABBISH. **Labeling images with a computer game.** *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 319–326, 2004.
- [14] JOHN DOUGLAS BRADLEY. **Pliny: A model for digital support of scholarship.** *Journal of Digital Information*, 9(1), 2008.
- [15] AMARNATH GUPTA AND RAMESH JAIN. **Visual information retrieval.** *Communications of the ACM*, 40(5):70–79, 1997.
- [16] DIANE LINGRAND. *Introduction au Traitement d’images.* Vuibert, 2008.
- [17] MARISTELLA AGOSTI, GIORGETTA BONFIGLIO-DOSIO, AND NICOLA FERRO. **A historical and contemporary study on annotations to derive key features for systems design.** *International Journal on Digital Libraries*, 8(1):1–19, 2007.
- [18] RONG JIN, JOYCE Y CHAI, AND LUO SI. **Effective automatic image annotation via a coherent language model and active learning.** *Proceedings of the 12th annual ACM international conference on Multimedia*, pages 892–899, 2004.
- [19] B.SRINIVASA REDDY AND BISWANATH N CHATTERJI. **An FFT-based technique for translation, rotation, and scale-invariant image registration.** *IEEE transactions on image processing*, 5(8):1266–1271, 1996.
- [20] WEI-YING MA AND BANGALORE S MANJUNATH. **Netra: A toolbox for navigating large image databases.** *Multimedia systems*, 7(3):184–198, 1999.
- [21] YUQING SONG, WEI WANG, AND AIDONG ZHANG. **Automatic annotation and retrieval of images.** *World Wide Web*, 6(2):209–231, 2003.
- [22] ALEKSANDRA MOJSILOVIC, JOSÉ GOMES, AND BERNICE E ROGOWITZ. **Isee: Perceptual features for image library navigation.** *Human Vision and Electronic Imaging VII*, 4662:266–277, 2002.

REFERENCES

- [23] WOLFGANG FÖRSTNER. **A framework for low level feature extraction.** *European Conference on Computer Vision*, pages 383–394, 1994.
- [24] BOHAN ZHUANG, QI WU, CHUNHUA SHEN, IAN REID, AND ANTON VAN DEN HENGEL. **Hcvrd: a benchmark for large-scale human-centered visual relationship detection.** 2018.
- [25] JONATHON S HARE, PAUL H LEWIS, PETER GB ENSER, AND CHRISTINE J SANDOM. **Mind the Gap: Another look at the problem of the semantic gap in image retrieval.** *Multimedia Content Analysis, Management, and Retrieval 2006*, 6073:607309, 2006.
- [26] MILAN SONKA, VACLAV HLAVAC, AND ROGER BOYLE. *Image processing, analysis, and machine vision.* Cengage Learning, 2014.
- [27] MICHELE TREVISIOL, LUCA CHIARANDINI, LUCA MARIA AIELLO, AND ALEJANDRO JAIMES. **Image ranking based on user browsing behavior.** *Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval*, pages 445–454, 2012.
- [28] BIN XU, JIAJUN BU, CHUN CHEN, DENG CAI, XIAOFEI HE, WEI LIU, AND JIEBO LUO. **Efficient manifold ranking for image retrieval.** *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval*, pages 525–534, 2011.
- [29] Y. ALP ASLANDOGAN AND CLEMENT T. YU. **Techniques and systems for image and video retrieval.** *IEEE transactions on Knowledge and Data Engineering*, 11(1):56–63, 1999.
- [30] STÉPHANE BRES, JEAN-MICHEL JOLION, AND FRANK LEBOURGEOIS. *Traitement et analyse des images numériques.* Hermes Science Publications, 2003.
- [31] YUL GAO AND JIANPING FAN. **Incorporating concept ontology to enable probabilistic concept reasoning for multi-level image annotation.** *Proceedings of the 8th ACM international workshop on Multimedia information retrieval*, pages 79–88, 2006.
- [32] PHILIPPE AIGRAIN, HONGJIANG ZHANG, AND DRAGUTIN PETKOVIC. **Content-based representation and retrieval of visual media: A state-of-the-art review.** *Multimedia tools and applications*, 3(3):179–202, 1996.
- [33] PHILIPPE AIGRAIN, HONGJIANG ZHANG, AND DRAGUTIN PETKOVIC. **Content-based representation and retrieval of visual media: A state-of-the-art review.** *Multimedia tools and applications*, 3(3):179–202, 1996.