

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieure et de la Recherche
Scientifique

UNIVERSITE KASDI MERBAH OUARGLA
Faculté des Nouvelles Technologies 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

Filière : Informatique

Spécialité : Informatique Industrielle

Présenté par : ABDERRAHMANI Ahmed et SEDRAT Salah

Thème

**Interpretation and semantic indexing of
an image content**

Soutenu le : 20/062022

| | | |
|-------------|----------------------|----------------------------------|
| Président | Dr. Boukhamla Akram. | Maître assistant A, UKM Ouargla. |
| Examineur | Dr. Khaldi Amine. | Maître assistant A, UKM Ouargla. |
| Superviseur | Dr. Zga Adel. | Maître assistant A, UKM Ouargla. |

Année universitaire : 2021/2022

Résumé

Le domaine de la détection d'objets a suscité beaucoup d'intérêt de la part des chercheurs au cours des dernières décennies. Comme il est un estuaire pour de nombreuses applications intéressantes, dans de nombreux domaines de la vie qui nécessitent à la reconnaissance d'objets dans les images.

Notre méthodologie fournir des informations supplémentaires pour l'analyse sémantique et l'interprétation du contenu des images, ce qui est le résultat souhaité. Nous avons utilisé le modèle HCVRD pour vérifier la présence d'humains dans les images, et nous avons ensuite déterminons les relations entre les objets en utilisant les ontologies avec les réseaux nuerons convolutifs CNN utilisons le modèle Vgg16. Le résultat obtenu est acceptable et prometteur, mais elle a besoin de plus de développement.

Mots-clés : détection d'objets, les ontologies, CNN, HCVRD, VGG16.

Remerciements

Ce mémoire a été réalisé dans le cadre du projet de fin d'étude au
DEPARTEMENT D'INFORMATIQUE, FACULTE DES NOUVELLES
TECHNOLOGIES DE L'INFORMATION ET DE LA COMMUNICATION,
l'université KASDI MERBAH - Ouargla.

Avec un grand plaisir, nous remercions avant tout notre dieu ALLAH qui
nos donné la patience, le courage, la capacité et la santé de terminer ce
travail.

Nous tenons à remercier en cette occasion, toutes les personnes proches
ou lointaines qui nous aider, même si avec un mot de motivation.

Nous ne pouvons pas continuer à remercier sans mentionner l'encadreur,
Mr. ZGA ADEL maître assistant de classe A, nous le remercions beaucoup
pour sa confiance, ses encouragements, son soutien continu et les
conseils précieux tout au long de la réalisation de ce travail.

Nous remercions également les membres du jury, qui nous ont honorons
en acceptant de juger et d'évaluer ce modeste travail.

Nous remercions tous les enseignants de département des nouvelle
technologies de l'information et de la communication à filière
d'informatique de l'université kasdi merbah, et nous remercions
également tous les travailleurs et le personnel administratif pour l'activité
constante et le suivi pédagogique.

ABDERRAHMANI Ahmed

SEDRAT Salah

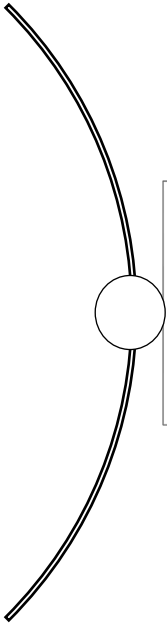
Dédicaces

Nous dédions notre travail à

Les deux êtres les plus chères au monde : les parents.

A nos frères et sœurs, A toute la famille.

A toutes les amies, et tous les camarades.



*Table
des matières*

Table des matières

Résumé

Remerciements

Dédicaces

Table des matières i

Liste des figures iv

Liste des tableaux v

Liste des abréviations vi

Introduction Générale..... 1

I. Chapitre I : Vision par ordinateur et la recherche d'information en images 2

 I.1. Introduction 2

 I.2. Vision par ordinateur et l'analyse d'image. 2

 I.3. Les descripteurs d'image 4

 I.3.1. Les descripteurs Local et les descripteurs globale 4

 I.3.2. Les descripteurs visuels 5

 I.4. Segmentation des images 6

 I.4.1. Segmentation basée sur les régions..... 7

 I.4.2. Segmentation basée sur le contour..... 7

 I.4.3. Segmentation basée sur la classification ou le seuillage..... 7

 I.4.4. Segmentation sémantique 7

 I.5. La recherche des images 8

 I.5.1. Les Techniques de recherche des images 9

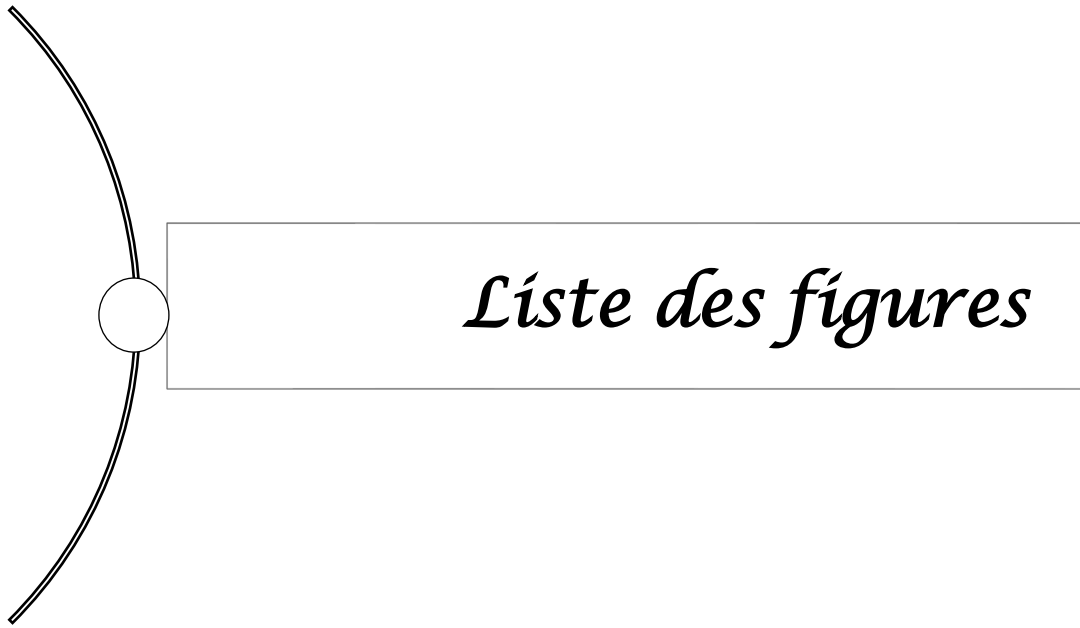
 I.5.2. Les aspects de la recherche des images 11

 I.5.3. Architecture générale d'un système d'indexation..... 16

 I.6. Apprentissage automatique (ML) 17

| | |
|--|----|
| I.6.1. Les classes d'apprentissage automatique..... | 18 |
| I.7. Les techniques de classification et de l'apprentissage..... | 19 |
| I.7.1. KNN..... | 19 |
| I.7.2. FCM..... | 19 |
| I.7.3. K-Means..... | 19 |
| I.7.4. KMS..... | 20 |
| I.7.5. SVM..... | 20 |
| I.8. Les modèle d'ensemble d'apprentissage | 21 |
| I.8.1. Le Voting | 21 |
| I.8.2. Le Bagging..... | 21 |
| I.8.3. Le Boosting | 21 |
| I.8.4. Le Stacking | 22 |
| I.9. Contexte sur des réseaux de neurones | 22 |
| I.9.1. Réseaux neuronaux convolutionnels..... | 22 |
| I.9.2. Réseaux neuronaux artificiels | 23 |
| I.9.3. Les couche de réseaux neuronaux..... | 23 |
| I.10. Apprentissage par transfert | 26 |
| I.11. Apprentissage profond | 26 |
| I.11.1. Fonctionnement d'apprentissage profond | 27 |
| I.11.2. L'apprentissage en profondeur et la détection d'objets | 27 |
| I.12. Conclusion | 31 |
| II. Chapitre II : Les modèle d'apprentissage ontologiques..... | 33 |
| II.1. Introduction | 33 |
| II.2. Interdépendance entre la connaissance et la langue | 33 |
| II.3. Préliminaires :..... | 34 |
| II.4. Apparition d'ontologie | 35 |
| II.5. La notion d'ontologie | 35 |
| II.5.1. Définitions | 35 |

| | |
|---|-----|
| II.5.2. Les différents types d'ontologies :..... | 37 |
| II.5.3. Les composants de l'ontologie | 37 |
| II.5.4. Classifications d'ontologies :..... | 39 |
| II.5.5. Domaines d'application des ontologies | 42 |
| II.5.6. Les ressources lexicales..... | 43 |
| II.6. Langages et formalismes d'ontologie | 46 |
| II.6.1. XML | 46 |
| II.6.2. RDF | 46 |
| II.6.3. OWL | 47 |
| II.6.4. Les logiques descriptives (DL)..... | 47 |
| II.7. Conclusion..... | 48 |
| III. Chapitre III : Interprétation sémantique du contenu d'images. | 49 |
| III.1. Introduction | 49 |
| III.2. Présentation des outils de développement..... | 49 |
| III.2.1. Le matériel..... | 49 |
| III.2.2. Anaconda..... | 49 |
| III.2.3. Python..... | 52 |
| III.3. Présentation de l'application..... | 53 |
| III.3.1. L'application | 53 |
| III.4. Les tests et les résultats | 55 |
| III.5. Discussion des résultats..... | 56 |
| III.6. Conclusion..... | 57 |
| Conclusion générale | 58 |
| Références bibliographiques | vii |



Liste des figures

Liste des figures

| | |
|---|----|
| Figure I-1: Vision par ordinateur. | 3 |
| Figure I-2: Construction d'un descripteur SIFT. | 6 |
| Figure I-3: Exemple de segmentation semantique. | 8 |
| Figure I-4: Diagramme typique pour un système de recherche d'images par le contenu. | 9 |
| Figure I-5: Processus d'annotation manuelle avec les éléments de l'image. | 14 |
| Figure I-6: Architecture de l'annotation semi-automatique dans le système ayant été annotés avec les mots-clés de la requête soumise. | 15 |
| Figure I-7: Processus d'annotation automatique. | 16 |
| Figure I-8: Le fonctionnement d'un système de recherche et d'indexation d'images. | 17 |
| Figure I-9: Exemple de operation max-pooling. | 25 |
| Figure I-10: Architecture R-CNN. | 28 |
| Figure I-11: Architecture FAST R-CNN. | 29 |
| Figure I-12: Architecture FASTER R-CNN. | 30 |
| Figure II-1: le triangle sémiotique générique | 34 |
| Figure II-2: La pyramide de la sagesse DIKW. | 34 |
| Figure II-3: Hiérarchie de concepts d'une ontologie pour les recettes de cuisine. | 36 |
| Figure II-4: Fragment du modèle de domaine | 44 |
| Figure II-5: Les couches de L'OWL | 47 |
| Figure III-1: symbole de anaconda | 49 |
| Figure III-2: Interface du navigateur anaconda | 50 |
| Figure III-3: le symbole spyder | 51 |
| Figure III-4: Interface du spyder. | 52 |
| Figure III-5: Symbole du python | 52 |
| Figure III-6: Interface d'application | 53 |
| Figure III-7: Selection de type d'objet | 54 |
| Figure III-8: Selection d'image | 54 |
| Figure III-9: Detection d'objet, exemple 1 | 55 |
| Figure III-10: Detection d'objet, exemple 2. | 55 |
| Figure III-11: Detection d'objet, exemple 3. | 56 |
| Figure III-12: Detection d'objet, exemple 4. | 57 |



Liste des tableaux

Liste des tableaux

Tableau I-1: Comparaison des méthodes d'apprentissage automatique. 18

Tableau II-1: Les propriétés permettant de distinguer la donnée, l'information et la connaissance..... 35

Tableau III-1: le taux d'erreur par rapport la nombre d'objet. 56



*Liste des
abréviations*

Liste des abréviations

AI: Artificial Intelligence.

ANN: Artificial Neural Network.

CBIR: Content Based Image Retrieval.

CNN: Convolutional Neural Network.

CSS: Curvative Scale Space Descriptors.

DIKW: Data-Information-Knowledge-Wisdom.

DL: Deep Learning.

OWL: Ontology Web Language.

DNN: Deep Neural Network.

FC: Fully Connected.

IR: Information Retrieval.

ISO: International Organization for Standardization.

KMS: Kernel Methods.

KNN: K-Nearest Neighbors.

LU: Lexical Units.

ReLU: Rectified Linear Unit.

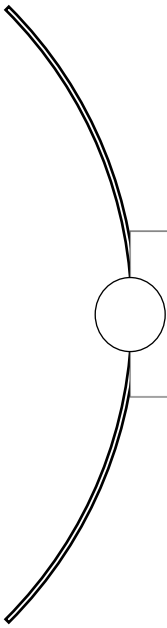
RGB: Red Green Blue.

SEI: Semantic Extraction and Interpretation.

SVM: Support Vector Machine.

VGG: Pretrained model.

XML: Extensible Markup Language.



*Introduction
générale*

Introduction Générale

A cause du développement numérique rapide de cette ère, la quantité de bases d'images a explosé. La mise à la portée de chacun d'appareils photos numériques, scanners, webcams et téléphones portables, la montée en puissance des capacités de calcul et de stockage des ordinateurs, ont contribué à cette explosion. Cette disponibilité d'information a cependant donné naissance à un nouveau besoin que l'on ne connaissait pas auparavant celui de localiser les images voulues en un temps raisonnable. Ceci a stimulé l'émergence d'un nouveau domaine de recherche qui est aujourd'hui en plein essor, en l'occurrence la recherche d'images. L'objectif principal de ce domaine de recherche est de développer des outils capables d'aider l'utilisateur à localiser les images recherchées en un temps raisonnable.

Notre mémoire est organisé en trois chapitres. Il commence par cette introduction générale et se termine par une conclusion générale.

Dans le chapitre 1, nous décrit la vision par ordinateur et la recherche d'informations dans les images. Ce chapitre contient deux parties. La première partie est consacrée à la présentation des représentations d'images, les descripteurs et les techniques de segmentation, les techniques d'annotation, et les techniques de recherche. La deuxième partie définit l'apprentissage automatique, l'apprentissage profond, et la classification.

Le chapitre 2, est sur les modèles d'apprentissage ontologiques, consacrée à donner un aperçu sur les ontologies. Ce chapitre est compris également la définition d'ontologie, ces types, ces composants, et ces classifications. Aussi, le domaine d'application des ontologies, ces ressources, et ces formalismes.

Le chapitre 3, est un état de l'art sur l'Interprétation sémantique du contenu des images, détection des objets, application du méthode et discussion des résultats.

Dans la conclusion générale, nous présenterons les principaux points de ce travail et quelques perspectives qui peuvent en découler



Chapitre 01:



*Vision par ordinateur et la
recherche d'information en
images*

I. Chapitre I : Vision par ordinateur et la recherche d'information en images

I.1. Introduction

La recherche d'images par le contenu 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 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.

Dans ce chapitre, nous présentons les techniques de représentation d'images, les techniques d'annotation, les techniques de recherche et les techniques de recommandation.

I.2. Vision par ordinateur et l'analyse d'image.

La représentation d'une image consiste à décrire une image via des descripteurs ou des caractéristiques. Deux types de représentation sont détaillés dans cette section :

La représentation par les caractéristiques physiques (les caractéristiques de bas niveau) et la représentation par les caractéristiques sémantiques (les caractéristiques de haut niveau).[1]

La représentation physique, c'est une approche traditionnelle pour rechercher les images est de les représenter avec des caractéristiques de bas niveau. 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. Les caractéristiques locales se sont avérées plus proche de la perception et sont basées sur la segmentation des images en des régions.[2]

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.[3]

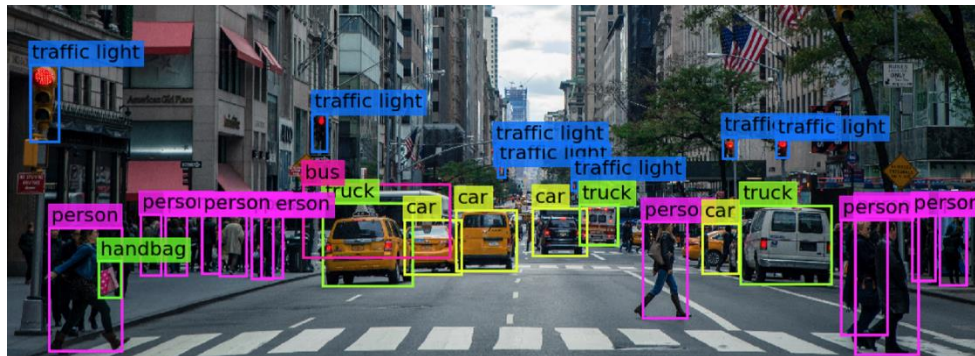


Figure I-1: Vision par ordinateur.

a) 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 [3]. En effet, cette donnée varie considérablement avec l'orientation des surfaces, la caméra et l'illumination, 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 RVB (RGB en anglais) qui code la couleur d'un pixel sur un vecteur en trois dimensions : rouge, vert, bleu.

b) 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.

c) 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. Les formes sont beaucoup utiles dans certains domaines tels que dans les objets fabriqués par l'homme [3]. Les formes sont difficilement applicables sur des images en couleur utilisées dans la plupart des documents.

d) La localisation spatiale

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. 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.

I.3. Les descripteurs d'image

I.3.1. Les descripteurs Local et les descripteurs globale

On peut utiliser un descripteur global qui caractérise la totalité de l'image 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 seule descripteur d'écrit la totalité de l'image, cela les rend robustes au bruit qui peut affecter le signal, les histogrammes de couleur et des niveaux de gris. L'inconvénient de ces descripteurs est qu'ils ne permettent pas de

distinguer des parties de l'images. 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. [4]

I.3.2. Les descripteurs visuels

Les descripteurs peuvent également être classés selon le type de modalité qu'ils représentent.

I.3.2-1 Les descripteurs de couleur

Les descripteurs de couleur sont largement utilisés dans le domaine de la récupération d'images/vidéos basée sur le contenu. La détection d'objets abordé lors de la comparaison entre plusieurs descripteurs de couleur en images et vidéos. L'histogramme est le descripteur le plus simple à calculer, il se compose de compter le nombre d'événements de valeurs d'intensité de pixel dans l'image. Nous pouvons distinguer plusieurs catégories d'histogrammes et on peut les classer. [4]

I.3.2-2 Les descripteurs de texture

La texture représente également un descripteur bas niveau efficace utilise dans le Framework de l'indexation et la recherche par le contenu. Plusieurs techniques ont été d'enveloppées 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. D'autres méthodes d'analyse de texture pour la recherche d'images incluent l'utilisation de filtres de Gabo et les fractales. [5]

I.3.2-3 Les descripteurs de formes

Les descripteurs de formes permettent, comme leur nom l'indique, de présenter une information pertinente sur le contenu de l'image et précisément sur la forme. Il existe différents types de descripteurs de formes qui se différencient par leur simplicité/complexité. Il existe plusieurs descripteurs de formes comme CSS, les filtres de convolution, les descripteurs de fourrier, les moments de Hu et de Zernike. [5]

I.3.2-4 Les 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 dans l'image. A 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 (**voir Figure I-2**). Une manière de les déterminer est de prendre en compte les zones où le signal change.

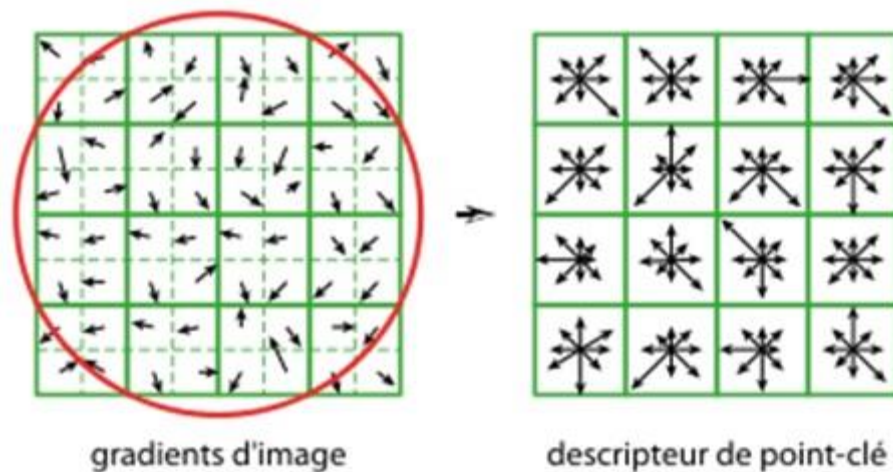


Figure I-2: Construction d'un descripteur SIFT.

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, 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. [6]

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

I.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...). 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 comme le 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.

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 (algorithme à bulle...).

I.4.3. Segmentation basée sur la classification ou le seuillage

Le seuillage a pour objectif de segmenter une image en plusieurs classes via des histogrammes. On suppose donc que l'information associée à l'image permet à elle seule la segmentation, c-à-dire qu'une classe est caractérisée par sa distribution de niveaux de gris. A chaque pic de l'histogramme est associée une classe. Il existe de très nombreuses méthodes de seuillage.

La plupart de ces méthodes s'appliquent correctement si l'histogramme contient réellement des pics séparés. De plus, ces méthodes ont été développées pour traiter le cas particulier de la segmentation en deux classes et leur généralité face aux cas multi-classes n'est que très rarement garantie.

I.4.4. Segmentation sémantique

La segmentation d'images est une tâche de vision par ordinateur de longue date qui consiste à diviser une image en plusieurs parties, normalement en suivant certains critères, mais pas nécessairement en cherchant à comprendre le contenu de l'image.

Une façon très courante de réaliser la segmentation sémantique consiste à étiqueter chaque pixel d'une image en fonction de la catégorie ou de l'objet auquel il appartient, parfois également appelée compréhension de scène ou classification par pixel.

D'autre part, la segmentation sémantique est une tâche différente qui va plus loin que la segmentation d'image en essayant de diviser une image en parties sémantiquement significatives. Plus précisément, la segmentation sémantique s'intéresse au découpage de l'image en régions de sens différents ou appartenant à des catégories différentes. Très souvent, ces catégories sont un ensemble prédéfini d'objets qui permettent une segmentation totale de l'image. Un exemple de segmentation sémantique vu dans (**Figure I- 3**).



Figure I-3: Exemple de segmentation sémantique.

Différentes métriques d'évaluation pour la segmentation sémantique peuvent afficher des résultats divergents car il n'est pas clair comment définir une bonne performance sur cette tâche. Trois des métriques les plus couramment utilisées, et celles prises en considération dans cette thèse de master, sont la précision des pixels, l'intersection moyenne sur l'union et la précision moyenne par classe.

I.5. La recherche des images

Les informations textuelles sont extrêmement limitées pour la représentation des données. De cette manière, mélange l'innovation médiatique, les images principales, sont progressivement utilisées. L'ensemble de données d'images s'adresse désormais un énorme volume de données. La remise en question des jeux de données d'images devient un test majeur dans le monde du génie logiciel. Pour superviser et utiliser avec succès ces ensembles de données d'images, un framework de recherche d'image est nécessaire. C'est la raison pour laquelle c'est une région extrêmement dynamique pour un bout de temps. Un framework de recherche d'image supervise l'accès à quelques images, les images sont adressées et commandées par la méthodologie d'occasion et contrastées en utilisant différentes méthodologies avec des enquêtes produites utilisant des procédures explicites. Dans les grandes lignes, Le framework de recherche d'image

traite de la représentation de l'image, du développement de l'enquête et de la détermination des images.

I.5.1. Les Techniques de recherche des images

Les principaux Framework de la recherche d'image dépendaient d'une représentation imprimée de l'image, les slogans sont une source extérieure de données physiquement annexé aux images. Le système de commande dépendant des slogans est fastidieux puisqu'il est manuel, précaire au motif que la nature du framework de recherche repose dans une large mesure sur la signification des termes relégués aux images et non normalisés à la lumière du fait que la source de ces mots d'ordre n'est en général pas spéciale et s'appuie sur l'individu qui associe les mots d'ordre aux images. De plus, une liste d'attente des mots d'ordre ne peut couvrir totalement la sémantique souvent riche véhiculée par une image.

La deuxième ère des framework de la recherche d'image dépend du contenu. La substance à base la méthodologie prévoit d'extraire directement les données de l'image réelle pour avoir l'option pour le caractériser, cela vient comme une option contrairement à la méthodologie basée sur le texte encore les problèmes sémantiques identifiés avec le traitement programmé des images sont immédiatement notés. Nous pouvons classer les approches de recherche d'image en deux flux principaux : le courant de recherche d'image et courant de recherche d'image sémantique dépendant d'une information formalisme de représentation.

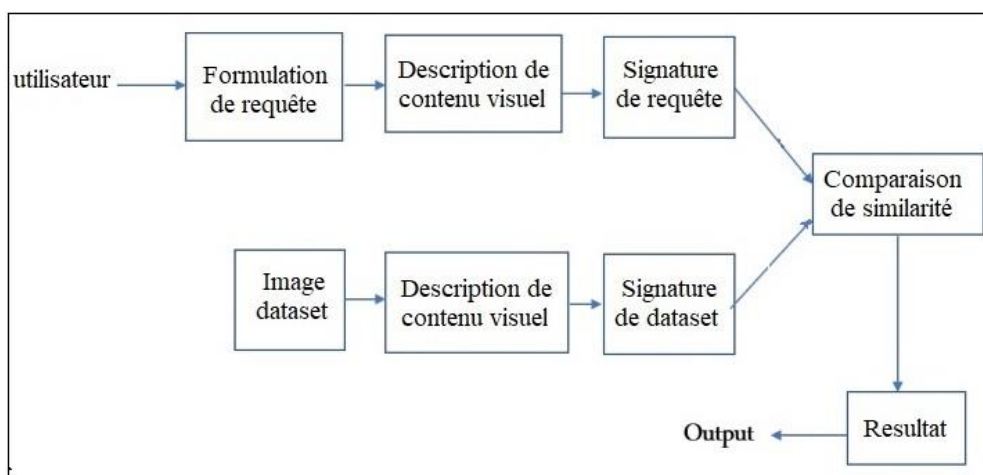


Figure I-4: Diagramme typique pour un système de recherche d'images par le contenu.

I.5.1-1 La recherche des images basée sur le contenu

Les normes de CBIR approches doivent répondre à deux questions, comment représenter une image ? de plus, comment jauger la ressemblance en fonction de la théorie de représentation ? Ces méthodologies appliquent pour la plupart des mesures et des stratégies d'IA.

Pour représenter les images, la plupart des structures CBIR utilisent des attributs naturellement séparés des images comme le ton, la surface et la forme. Ces qualités sont utilisées dans divers cycles, comme la similitude calcul, la structure du modèle ou même des explications.

L'un des problèmes les plus connus des méthodologies basées sur le contenu est le fossé entre attributs extraits des images pour la représenter qui sont des qualités de bas niveau, et la représentation humaine de la qui est d'un niveau indéniable. En tout, le trou entre la représentation picturale assemblée en conséquence, et la traduction humaine de l'image. Sans aucun doute, l'humain utilise plus de données et de qualités que le mécanisé cycles pour déchiffrer l'image, il existe des différentes stratégies proposées qui se caractérisent en cinq grandes catégories :

- ❖ **Utiliser une ontologie** : pour caractériser les articles de niveau significatif et mieux décrypter les qualités extraites de bas niveau.
- ❖ **Utiliser une stratégie d'IA** : prévoir à partir d'une mesure d'information la valeur d'un résultat proportion d'une interaction ou décrire l'association de la mesure d'information.
- ❖ **Utiliser un feedback client** : essayer de se familiariser avec les objectifs du client à tous les plus susceptibles de comprendre leurs besoins en utilisant la gestion basée sur le Web.
- ❖ **Génération d'un modèle sémantique** : pour prendre en charge la recherche d'image de niveau significatif en utiliser des formats pour traiter des idées déterminées à partir d'un assortiment d'images de référence.
- ❖ **Utilisation d'un contexte Web** : pour la recherche d'image Web. Il comprend l'utilisation de données qui peuvent être organisés sous forme d'URL ou de pages HTML pour travailler sur la sémantique des images.

I.5.1-2 La recherche des images basée sur la sémantique

Les approches sémantiques de la recherche d'images utilisent couramment un formalisme de représentation de l'information comme des logiques de représentation ou des organisations

sémantiques. L'objectif est de traquer un modèle de représentation d'image. Cette représentation doit être effectivement similaire et décrire l'image de la manière la plus idéale. La recherche sémantique dépend de la signification des slogans, et non de leur grammaire. La corrélation entre les idées dépend de procédures plus développées que celles de recherche syntaxique. Il considère en outre la communication entre diverses idées de la question de développer davantage la signification contrastée avec une recherche grammaticale. La recherche d'image basée sur des mots clés est simple et facile à appliquer avec une précision exacte, et l'image basée sur la cosmologie la recherche développe davantage la précision mais nécessite une représentation totale des images.

I.5.2. Les aspects de la recherche des images

Pour concevoir un système de recherche d'images, nous devons répondre à trois questions clés : comment construire la requête ? Comment représente-t-on une image ? Et comment faire correspondre la requête et la description ?

I.5.2-1 La modalité de requête

Une variable importante dans un framework de recherche d'image est la méthodologie d'enquête. Cette méthodologie caractérise le langage des questions soutenu par le framework en ce qui concerne l'expressivité de l'enquête client créée. Les modalités de question les plus connues dans le framework de la recherche d'images sont :

- **Les mots-clés** : cette méthodologie est la plus utilisée. L'enquête est introduite sous forme de mots d'ordre. La disposition des slogans approuvés n'est pas globalement limitée comme les framework de recherche d'images sur le Web, mais peut être prédéfinie, ce qui limitera l'expressivité des questions.
- **Free-Text** : pour cette méthodologie, le client tente de caractériser son besoin par des phrases. Ces phrases sont développées sans complexe, il peut très bien s'agir de questions, histoires, ou différentes articulations.
- **Image** : comme son nom l'indique, la question présentée par le client est une image et la structure renvoie toutes les images comme l'image du client. Cette méthodologie est entièrement appropriée pour les structures CBIR.
- **Graphiques** : cette méthodologie repose sur une représentation graphique de la question, le client caractérise graphiquement ses besoins en dessinant une image. L'image peut également être créé par un PC.

- **Composite** : il comprend un mélange de différentes modalités. Dans l'ensemble, le client peut choisir une méthodologie pour caractériser sa question ou caractériser une partie de son enquête en utilisant quelques modalités. La mixité des modalités est fascinante du fait du questionnement intelligent, le client peut à chaque fois donner plus d'éclaircissements sur sa demande en changeant de méthodologie.

I.5.2-2 La similitude des images

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 généralement pas numérique mais booléen et ne permet pas forcément le classement des images réponses. La similarité hybride utilise une combinaison de similarités syntaxiques et sémantiques. Il utilise la base de connaissances pour enrichir et pour une meilleure comparaison, et le calcul numérique pour le classement.

a) La similitude syntaxique

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

- **Signature visuelle utilisant** : Ces approches sont principalement utilisées dans les systèmes CBIR. Le principe est de formuler mathématiquement des caractéristiques appelées signatures. Ces signatures sont extraites des images pour utiliser une technique de calcul à distance pour trouver une ou un ensemble de valeurs représentant la similarité.
- **Modèles de recherche d'informations** : Le principe est simple, il consiste à utiliser du texte sous forme de mots-clés, de balises, de paragraphes ou d'autres formes pour indexer des images. L'idée est de transformer la recherche d'images en recherche d'informations pour utiliser des techniques et des modèles de recherche d'informations.

b) La similitude sémantique :

La similarité dans ce cas n'est pas une distance à calculer habituellement mais un processus de comparaison d'éléments d'une langue donnée. Un formalisme de représentation des connaissances est utilisé pour définir formellement une image d'annotation et une requête. Ce formalisme a une sémantique bien définie pour interpréter les annotations et les requêtes. La gestion de l'interprétation s'appelle le raisonnement. Le raisonnement sert à comparer une annotation et une requête. La comparaison est généralement logique et basée sur la théorie des ensembles, elle peut utiliser un raisonnement standard ou non standard.

- **Raisonnement standard** : Ce sont des raisonnements standardisés classiques qui appartiennent à la définition standard du langage. Ces raisonnements sont généralement maîtrisés, nécessaires à la meilleure exploitation du langage et incluent sur ses outils techniques tel que l'outil qui permet de raisonner appelé raisonneur. Ils sont également à la base de la définition d'autres raisonnements dits non standards.
- **Raisonnement non-standard** : Il s'agit d'un raisonnement non classique utilisant souvent le calcul numérique ou la théorie des probabilités. Ils sont souvent définis pour des applications spécifiques et non disponibles sur des outils techniques tels que les raisonneurs. Ces raisonnements sont généralement basés sur des raisonnements standards mais peuvent également utiliser une sémantique non standard.

I.5.2-3 L'annotation des images

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 semi-automatique et l'annotation 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 recherche d'images basées sur la sémantique mais peut également être utilisé par les 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.

a) L'annotation manuelle

Ce sont des opérateurs humains, généralement experts d'un ou plusieurs domaines, qui se chargent de caractériser, selon leurs connaissances propres, le contenu des images. Cette méthode est nécessaire dans le cas d'images provenant de collections de photographies personnelles, par exemple, car les images ne sont pas accompagnées de texte. La collection

d'images est alors annotée manuellement, c'est-à-dire que chaque image est associée à un commentaire, ou un ensemble de mots-clés. Dans le cas d'une annotation sémantique, l'annotation manuelle consiste simplement à mettre en place une interface utilisateur dans laquelle l'utilisateur humain peut sélectionner la ressource à annoter, choisir le modèle formel servant à la création des annotations sémantiques, tout en respectant les contraintes imposées par le modèle formel, créer les annotations voulues sur la ressource sélectionnée (**Figure I.5**).

L'annotation manuelle est risqué de devenir obsolète sur le long terme, car on doit faire face à des corpus d'images de plus en plus volumineux et dont le contenu évolue régulièrement. Néanmoins, l'annotation manuelle est plus performante que l'annotation automatique, car les annotateurs organisent les données et choisissent leurs termes d'annotation de façon à retrouver facilement les images.

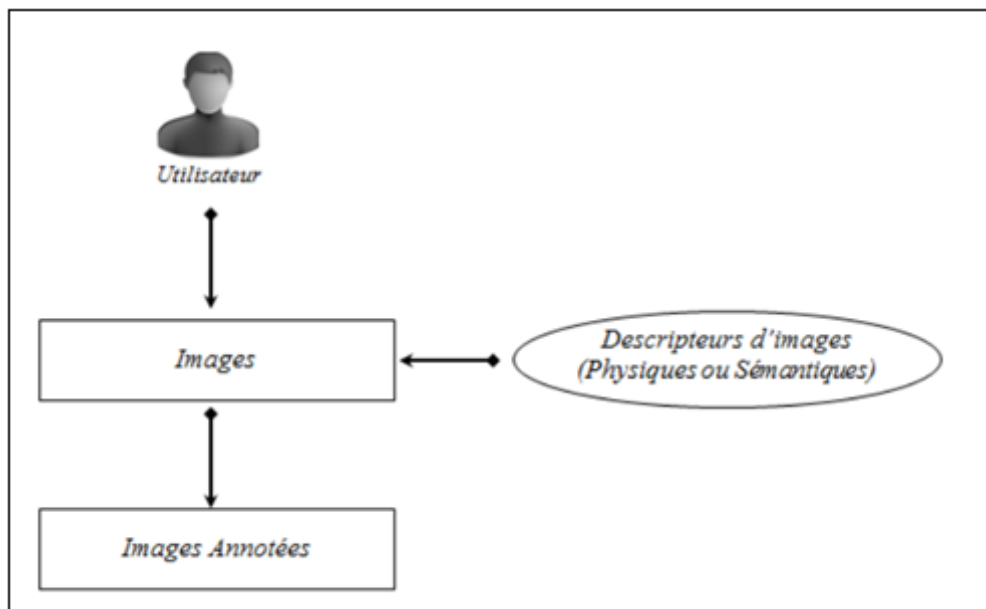


Figure I-5: Processus d'annotation manuelle avec les éléments de l'image.

b) L'annotation semi-automatique

L'annotation semi-automatique résulte de la combinaison des deux annotations précédentes. Elle consiste à générer l'annotation des images à partir de deux phases, La phase manuelle, qui consiste à faire une annotation manuelle d'un échantillon d'informations seulement. Puis on procède dans la deuxième phase à une propagation de l'annotation.

L'annotation semi-automatique s'appuie généralement sur un moteur d'extraction d'informations qui est intégré dans un outil d'annotation. Il suggère des annotations à l'utilisateur qui peut ensuite les valider manuellement, voir (**Figure I.6**).

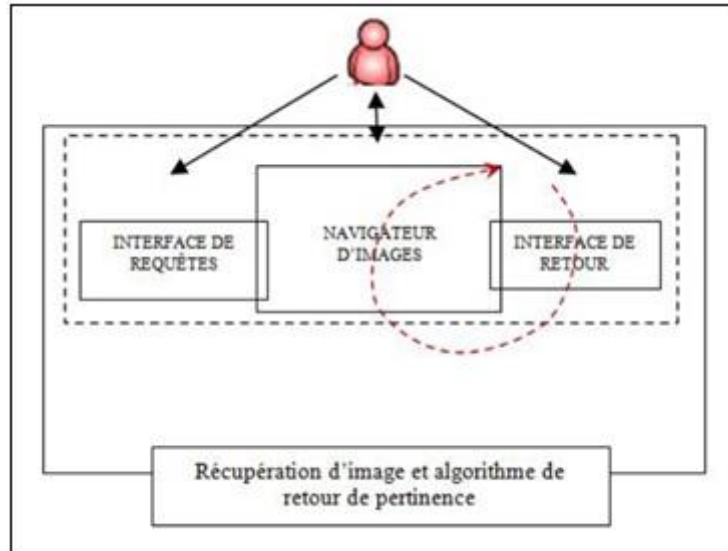


Figure I-6: Architecture de l'annotation semi-automatique dans le système ayant été annotés avec les mots-clés de la requête soumise.

c) L'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. Le système utilise automatiquement chaque nouvelle image comme une requête et effectue une recherche d'images par le contenu. Pour un nombre n d'images similaires à une requête, les caractéristiques ayant servi à l'annotation de chacune d'elles sont analysées. Une liste de caractéristiques des images similaires à la nouvelle image est fournie et attribuée à la nouvelle image. La nouvelle image est ainsi annotée.

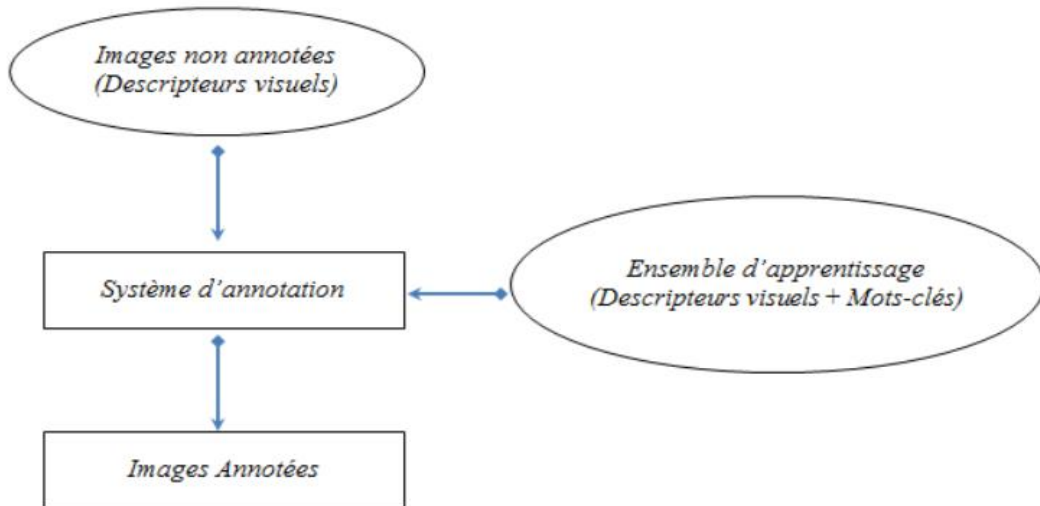


Figure I-7: Processus d'annotation automatique

I.5.3. Architecture générale d'un système d'indexation

Deux aspects indissociables coexistent dans les systèmes de recherche d'images par le contenu, l'indexation et la recherche.

- La phase d'indexation (hors-Ligne) : Dans cette phase, des caractéristiques sont automatiquement extraites à partir de l'image et stockées dans un vecteur numérique appelé descripteur visuel. Grâce aux techniques de la base de données, on peut stocker ces caractéristiques et les récupérer rapidement et efficacement.
- La phase recherche (On-line) : Dans cette étape, le système analyse une ou plusieurs requêtes émises par l'utilisateur et lui donne le résultat correspond en une liste d'images ordonnées, en fonction de la similarité entre leur descripteur visuel et celui de l'image requête en utilisant une mesure de distance.

La figure 2 schématise le fonctionnement d'un système de recherche et d'indexation d'images.

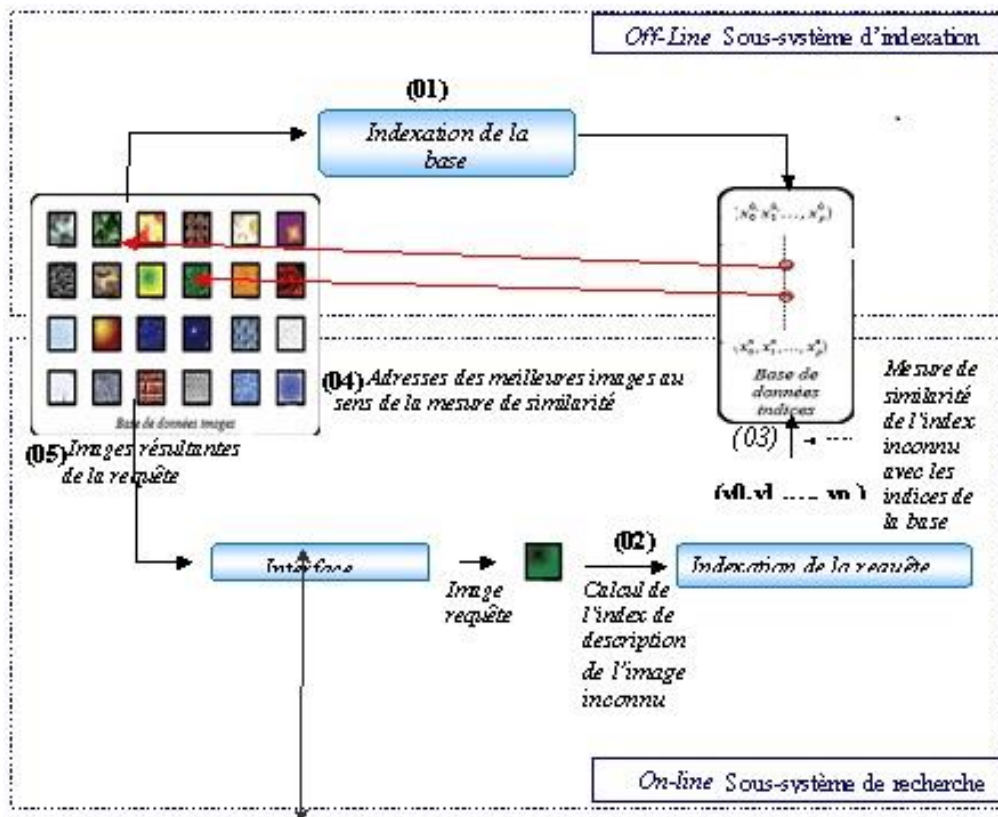


Figure I-8:Le fonctionnement d'un système de recherche et d'indexation d'images.

I.6. Apprentissage automatique (ML)

L'apprentissage automatique est appliqué dans les cas où un programmeur ne peut pas indiquer explicitement à la machine ce qu'il faut faire et quelles sont les mesures à prendre. De nos jours, et depuis plusieurs années, plusieurs applications tirent profit de l'apprentissage automatique. L'exemple typique de ce genre d'application est celui dans lequel on a besoin de retourner une liste de résultats ordonnés en termes de pertinence par rapport à une requête donnée. Parmi les plus connues et plus utilisées par la majorité des utilisateurs, on peut en citer les applications de traduction automatique, la reconnaissance de visages qui peut faire partie de systèmes de sécurité ou d'accès contrôlés à des endroits ou services, la reconnaissance vocale, la reconnaissance des empreintes digitales, etc.

L'objectif de l'apprentissage automatique est d'apprendre à partir d'un ensemble de données dit "ensemble d'apprentissage", des informations de manière à pouvoir les obtenir ces mêmes types d'informations sur des données non encore vues.

I.6.1. Les classes d'apprentissage automatique

On distingue trois classes d'apprentissage automatique (**Tableau I.1**) :

a. Apprentissage supervisé

A partir d'une donnée d'entrée, par exemple une photographie avec un panneau de signalisation, la tâche consiste à prédire le résultat ou l'étiquetage correct, par exemple le panneau de signalisation qui se trouve dans l'image (limitation de vitesse, panneau stop, etc.) ou dans les cas les plus simples, les réponses se présentent sous la forme de (oui/non).

b. Apprentissage non-supervisé

Il n'y a pas d'étiquetage ou de résultats corrects, la tâche consiste à découvrir la structure des données, par exemple le regroupement des éléments similaires pour former des grappes, ou réduire les données à un petit nombre de dimensions importantes. La visualisation des données peut aussi être considérée comme un apprentissage non supervisé.

c. Apprentissage par renforcement

Communément utilisé dans les situations où un agent d'IA, comme une voiture autonome, doit fonctionner dans un environnement et où un retour d'information sur les bons ou mauvais choix est disponible avec un certain retard. Également utilisé dans les jeux dont le résultat ne peut être décidé qu'à la fin de la partie.

Tableau I-1: Comparaison des méthodes d'apprentissage automatique.

| | APPRENTISSAGE SUPERVISÉ | APPRENTISSAGE NON-SUPERVISÉ | APPRENTISSAGE PAR RENFORCEMENT |
|-------------------|---|--|---|
| DÉFINITION | L'algorithme apprend à partir de données labellisées | L'algorithme est entraîné à partir de données non labellisées sans indications particulières | L'algorithme interagit avec son environnement en réalisant des actions et en apprenant de ses erreurs et succès |
| TYPE DE PROBLÈMES | Régression et classification | Association et Clustering | Basés sur un système de récompense |
| TYPE DE DONNÉES | Données labellisées | Données non labellisées | Pas de données fournies au préalable |
| APPROCHE | Étudie les relations sous-jacentes qui lient les données en entrée aux labels | Découvre les motifs communs au sein des données d'entrée | Apprend une stratégie de comportement en fonction d'expériences passées et des récompenses perçues |

I.7. Les techniques de classification et de l'apprentissage

Indépendamment des problèmes de calculs et de l'absence de données, le consensus qui régnait stipule que les méthodes discriminantes sont généralement préférées aux méthodes génératives.

I.7.1. KNN

KNN ou en français "K-plus proche voisins" est une méthode bien connue dans le domaine d'apprentissage et la vision. Contrairement à certaines approches, la phase d'apprentissage dans K-NN est transparente. En effet, K-NN nécessite une mémorisation de l'ensemble des données d'apprentissage parce qu'elle ne fait aucune généralisation. Elle consiste à classer un objet par vote majoritaire de ses voisins. La classe la plus dominante parmi ses K- plus proches voisins (K est généralement de petite taille) lui sera attribuée. [18]

I.7.2. FCM

Fuzzy C-Means est un algorithme de classification non-supervisée floue. Il introduit la notion d'ensemble flou dans la définition des classes. Comme les autres algorithmes de classification non supervisée, il utilise un critère de minimisation des distances intra-classe et de maximisation des distances interclasse, mais en donnant un certain degré d'appartenance à chaque classe pour chaque pixel. Cet algorithme nécessite la connaissance préalable du nombre de clusters et génère les classes par un processus itératif en minimisant une fonction objective. Ainsi, il permet d'obtenir une partition floue de l'image en donnant à chaque pixel un degré d'appartenance à une classe donnée. Le cluster auquel est associé un pixel est celui dont le degré d'appartenance sera le plus élevé. Les principales étapes de cet algorithme sont respectivement, la fixation arbitraire d'une matrice d'appartenance, le calcul des centroïdes des classes, le réajustement de la matrice d'appartenance suivant la position des centroïdes, et le calcul du critère de minimisation et retour à la deuxième étape s'il y a non convergence de critère.

I.7.3. K-Means

L'algorithme K-Means est l'algorithme de regroupement le plus connu et le plus utilisé, du fait de sa simplicité de mise en œuvre. Il partitionne les données d'une image en K clusters. Contrairement à d'autres méthodes dites hiérarchiques, qui créent un arbre de clusters pour décrire les groupements en un seul niveau de clusters. L'algorithme renvoie une partition des données, dans laquelle les objets à l'intérieur de chaque cluster sont aussi proches que possible les uns des autres et aussi loin que possible des objets des autres clusters. Chaque cluster de la partition est défini par ses objets et son centroïde. Le K-Means est un algorithme itératif qui

minimise la somme des distances entre chaque objet et le centroïde de son cluster. La position initiale des centroïdes conditionne le résultat final, de sorte que les centroïdes doivent être initialement placés le plus loin possible les uns des autres de façon à optimiser l'algorithme. K-Means change les objets de cluster jusqu'à ce que la somme ne puisse plus diminuer. Le résultat est un ensemble de clusters compacts et clairement séparés, sous réserve qu'on ait choisi la bonne valeur K du nombre de clusters. Les principales étapes de cet algorithme sont respectivement, le choix aléatoire de la position initiale des K clusters, la réaffectation des objets à un cluster suivant un critère de minimisation des distances, recalculer les K centroïdes, et réitérer les étapes 2 et 3 jusqu'à ce que plus aucune réaffectation ne soit faite. [19]

I.7.4. KMS

Les méthodes à noyaux sont surtout sollicitées dans le cas de données non linéairement réparables. Ce genre d'approches se basent sur le théorème de Mercer, qui dit que chaque fonction noyau continue symétrique semi-définie positive peut être exprimée en un produit scalaire dans un espace de haute dimension. La première application des noyaux dans l'apprentissage automatique remonte à l'années 1964 avec le travail de Aizerman et al. En effet, il n'y a que récemment que les chercheurs ont reconnu la grande importance et la large applicabilité des noyaux. Avec le travail de Boser et al. SVM "machines `a vecteurs de supports", devient la plus illustre technique à base de noyau. Grâce à l'utilisation des fonctions noyau, il devient ainsi possible d'avoir le meilleur des deux mondes, utiliser des techniques simples et rigoureusement garanties, et traiter des problèmes non linéaires. C'est pourquoi ces méthodes sont devenues très populaires récemment.

I.7.5. SVM

Les machines à vecteurs de support sont un ensemble de techniques d'apprentissage supervisé destinées à résoudre des problèmes de classification. Les SVM sont une généralisation des classifiées linéaires, ils rapidement adoptés pour leur capacité à travailler avec des données de grandes dimensions, le faible nombre d'hyper paramètres, leurs garanties théoriques, et leurs bons résultats en pratique. Les SVM ont été appliqués à de très nombreux domaines (bio-informatique, recherche d'information, vision par ordinateur, finance...). Selon les données, la performance des machines à vecteurs de support est de même ordre, ou même supérieure, à celle d'un réseau de neurones ou d'un modèle de mixture gaussienne. [7]

I.8. Les modèle d'ensemble d'apprentissage

Généralement, il n'existe pas un algorithme d'apprentissage individuel qui dans n'importe quel domaine toujours induit le modèle le plus précis. En effet, chaque algorithme d'apprentissage se base sur un ensemble d'hypothèses, dans le cas où ces dernières ne conviennent pas aux données considérées, cela mené à des erreurs et des performances très basses en termes de précision. Pour remédier à ce problème, certains chercheurs ont proposé des méthodes qui utilisent non seulement un seul mais un ensemble d'algorithmes d'apprentissage, qui peuvent être de même type ou de types différents. Ces techniques sont appelées "techniques d'apprentissage par ensemble". L'idée de ce type de méthodes est de construire un groupe d'apprenants de base qui, une fois combinés, génèrent un "méta- modèle" qui a une meilleure précision que les apprenants individuels.

I.8.1. Le Voting

Cette stratégie consiste à combiner les décisions (prédictions) de l'ensemble des apprenants individuels. La plus simple technique de vote est le vote majoritaire. On peut aussi pondérer les votes des apprenants individuels par des degrés de confiance ou d'importance (poids). Cette méthode est appelée "vote pondéré".

I.8.2. Le Bagging

Le "Bagging" est l'abréviation de "Bootstrap aggregating". Il a été proposé afin d'améliorer la performance de classification en combinant les classifications d'un ensemble de données générées aléatoirement. C'est une méthode dans laquelle, les apprenants sont entraînés sur des ensembles de données légèrement différents. En effet, pour chaque apprenant A, un nouvel ensemble de données d'apprentissage est généré par tirage aléatoire avec remise (bootstrap) des exemples depuis l'ensemble de données d'apprentissage d'origine. Les décisions des apprenants de base sont ensuite combinées suivant un vote majoritaire. N'importe quel type de modèle de classification peut être utilisé comme apprenant de base.

I.8.3. Le Boosting

La technique du "boosting" consiste à améliorer des apprenants ayant une performance faible mais supérieure à celle du hasard. Il est possible de transformer de tels apprenants en bons apprenants pouvant bien classer des exemples non annotés. La similarité entre le "bagging" et le "boosting" se résume uniquement dans la construction d'un ensemble de classificateurs en

échantillonnant l'ensemble de données d'apprentissage, et en combinant les décisions des différents classificateurs suivant un vote majoritaire.

L'échantillonnage des sous-ensembles d'apprentissage se fait dans le "boosting" de manière à fournir au prochain apprenant un ensemble de données le plus informatif possible.

I.8.4. Le Stacking

Le "stacking" est une technique d'apprentissage par ensemble consiste à combiner plusieurs classificateurs. Premièrement, un certain nombre de classificateurs sont entraînés sur l'ensemble de données d'apprentissage, leurs sorties sont ensuite combinées en utilisant un nouveau méta-classificateur pour apprendre un modèle permettant de faire une correspondance entre les décisions des classificateurs individuels de base et les classes correctes des exemples. Dans l'indexation multimédia, la technique de "stacking" est largement utilisée comme une méthode de fusion tardive, parce que c'est une bonne stratégie pour fusionner les scores des classificateurs, qui sont obtenus de différentes modalités.

I.9. Contexte sur des réseaux de neurones

Dernièrement, profitant de l'amélioration rapide de l'innovation en matière d'apprentissage profond, le domaine de vision de l'ordinateur a remporté un succès sans précédent. Il existe de nombreux algorithmes et techniques développés dans ce domaine. Dans ce qui suit, nous détaillerons chaque méthode d'apprentissage.

I.9.1. Réseaux neuronaux convolutionnels

En tant que réseau neuronal le plus largement utilisé, les organisations neuronales convolutives (CNN) sont au centre des calculs d'apprentissage pour la reconnaissance visuelle des exemples. Ils ont été créés à partir de perceptrons, des calculs de planification vectorielle animés par des apprentissages affiliés des neurones. Des chercheurs ont employé les CNN et leurs variantes pour relever divers défis tels que la classification des images, la reconnaissance d'objets, la reconnaissance d'actions, l'estimation de pose, le transfert de style neuronal, etc.

Les CNN sont constitués de diverses unités neurales, qui peuvent être pour la plupart séparées en trois sortes, en particulier, la couche d'informations, la couche secrète et la couche de rendement. Le couche d'information d'une organisation neuronale convolutive est principalement utilisée pour acquérir des données d'entrée, capable de traiter des informations multidimensionnelles.

La couche de rendement d'une organisation neuronale convolutive utilise généralement une capacité constante ou une capacité softmax pour produire les noms d'ordre. L'application utile fluctue selon le type d'affectation.

En général, chaque unité neuronale de la couche d'entrée se connecte directement aux données d'origine et fournit des informations sur les fonctionnalités à la couche masquée. Chaque unité neuronale dans la couche cachée représente différents poids pour différentes unités neuronales dans la couche d'entrée, il a donc tendance à être sensible à un certain schéma de reconnaissance. Les valeurs dans la couche de sortie varient en fonction au degré d'activation des couches cachées, qui est le résultat final de la reconnaissance du modèle.

I.9.2. Réseaux neuronaux artificiels

Les réseaux neuronaux artificiels (ANN) sont des cadres de calcul destinés à imiter la composante de traitement des données de l'esprit humain. De tels cadres "apprennent" à exécuter des projets en considérant des modèles sans être modifiés avec toutes les règles d'affectation et ils ont des capacités d'auto-apprentissage qui les rendent produire de meilleurs résultats lorsque davantage d'informations deviennent accessibles.

Un ANN comprend un assortiment de hubs associés appelés neurones contrefaits, qui modéliser les neurones dans un esprit humain. Chaque association peut communiquer un signe commençant par un faux neurone puis sur le suivant. Le faux neurone (récepteur du signal) peut le gérer et ensuite le communiquer aux faux neurones associés. De la même manière pour les exécutions ANN, le signe d'une association entre de faux neurones est un nombre réel, et le rendement de chacun faux neurone est enregistré par une certaine capacité non linéaire de la quantité de ses bits de rétroactions

I.9.3. Les couche de réseaux neuronaux

Régulièrement, les neurones sont totalisés en couches. Différentes couches peuvent effectuer différents types de changements sur leurs morceaux de rétroactions. Les signes voyagent de la couche principale (la couche d'information) à la dernière couche (la couche de rendement), peut-être après l'intersection de différentes couches. [17]

I.9.3-1 La couche de convolution

Les couches convolutives (CONV) sont les blocs de structure essentiels utilisés dans les organisations neuronales convolutives. La convolution en tant que canal permet à l'organisation

neuronale d'extraire faits saillants de niveau indéniable convaincants. La carte des éléments, également appelée l'initiation carte, peut être créé en appliquant encore et encore un canal similaire, ce qui démontre le domaine et la force des faits saillants reconnus dans l'image d'information. Le canal contient les charges à maîtriser lors de la préparation de la couche. En outre, la taille de canal ou la taille de bit influencera essentiellement l'état de la carte de surbrillance de rendement. C'est important que la connexion du canal avec la ligne de l'image puisse inciter impacts aux limites, en particulier pour l'image d'entrée de petite taille et exceptionnellement profonde organisme. Nous-partenaire, nous pouvons résoudre le problème d'impact sur les limites en ajoutant des pixels supplémentaires au bord de l'image, ce qu'on appelle l'amorti. De plus, la mesure du développement entre les applications de canal à l'image d'information est évoquée comme l'étape, et c'est assez souvent équilibré dans les aspects de stature et de largeur.

I.9.3-2 La couche de pooling

L'opération pooling (POOL) joue un rôle essentiel dans la structure des réseaux de neurones convolutives. Première de toutes, les couches de regroupement améliorent l'invariance spatiale dans une certaine mesure, comme l'invariance de translation, l'invariance d'échelle et l'invariance de déformation. A savoir, même si l'entrée d'image est légèrement transformée, la couche de regroupement peut toujours produire des fonctionnalités de regroupement similaires, ce qui rend le système d'apprentissage plus robuste. Deuxièmement, l'opération de mise en commun équivaut à un sous-échantillonnage des caractéristiques, ce qui augmente la taille du champ réceptif. Pour certaines tâches visuelles, un grand réceptif aide à apprendre les relations spatiales à longue portée et les modèles spatiaux implicites. De plus, l'opération de mise en commun réduit considérablement les paramètres du modèle, ce qui réduit le risque de surajustement (**Figure I-9**).

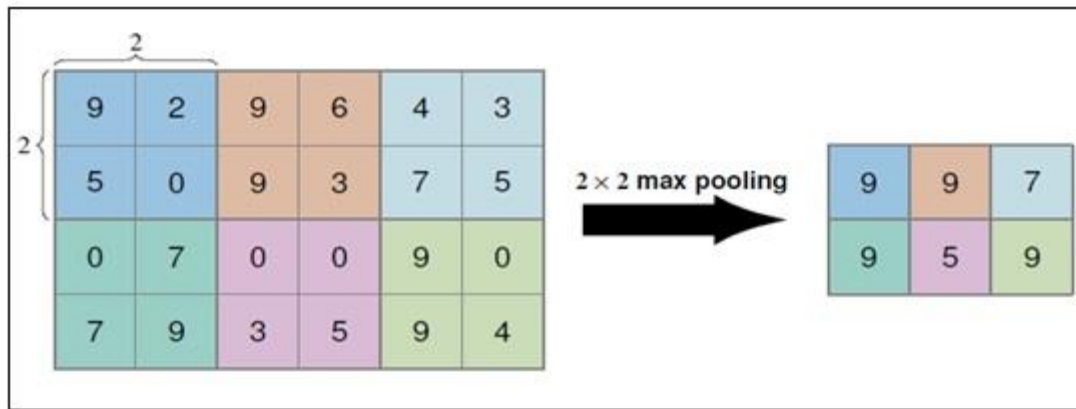


Figure I-9: Exemple de operation max-pooling.

I.9.3-3 La couche de correction

Dans les réseaux de neurones au niveau de la couche (ReLU), la fonction d'activation effectue la transformation non linéaire de l'entrée, la rendant capable d'apprendre et d'effectuer des tâches plus complexes. Afin d'augmenter la non-linéarité des réseaux de neurones, certaines fonctions non linéaires sont introduites. De toute évidence, l'accumulation de plusieurs fonctions linéaires est toujours linéaire, tandis que les fonctions linéaires ont une expression limitée. L'utilisation de fonctions non linéaires rend le réseau plus expressif et donc mieux adapté à la fonction cible. Deux fonctions non-linéaires courantes utilisées dans les réseaux de neurones convolutionnels sont la fonction sigmoïde et l'unité linéaire rectifiée (ReLU).

I.9.3-4 La couche entièrement connectée

Le raisonnement de haut niveau dans le réseau neuronal se fait via des couches entièrement connectées (FC), après plusieurs couches de convolution et de pooling. Les neurones dans une couche entièrement connectée ont des connexions vers toutes les sorties de la couche précédente. Leurs fonctions d'activations peuvent donc être calculées avec une multiplication matricielle suivie d'un décalage de polarisation.

I.9.3-5 La couche de perte

La couche de perte (LOSS), est la dernière couche dans le réseau, spécifie comment l'entraînement du réseau pénalise l'écart entre le signal prévu et réel. Diverses fonctions de perte adaptées à différentes tâches peuvent y être utilisées. La perte "Soft max" est utilisée pour prédire une seule classe parmi plusieurs classes mutuellement exclusives. La perte par entropie croisée sigmoïde est utilisée pour prédire plusieurs valeurs de probabilité indépendante.

I.10. Apprentissage par transfert

De nombreuses méthodes d'apprentissage automatique fonctionnent bien sous une hypothèse commune dite que les données d'apprentissage et de test sont tirées du même espace de caractéristiques et de la même distribution. Lorsque la distribution change, la plupart des modèles statistiques doivent être reconstruits à partir de zéro en utilisant des données d'entraînement nouvellement collectées. Dans de nombreuses applications du monde réel, il est coûteux, voire impossible, de collecter à nouveau les données de formation nécessaires et de reconstruire les modèles. Ce serait bien de réduire le besoin et l'effort de recollecter les données de formation. Dans de tels cas, le transfert de connaissances ou l'apprentissage par transfert entre les domaines de tâches serait souhaitable.

L'apprentissage par transfert est une méthode d'apprentissage automatique où un modèle développé pour une tâche est réutilisé comme point de départ pour un modèle sur une deuxième tâche. Il s'agit d'une approche populaire dans l'apprentissage en profondeur où des modèles préformés sont utilisés comme point de départ pour les tâches de vision par ordinateur et de traitement du langage naturel, compte tenu des vastes ressources de calcul et de temps nécessaires pour développer des modèles de réseaux neuronaux sur ces problèmes et des énormes sauts de compétences qu'ils fournissent sur des problèmes connexes.

I.11. Apprentissage profond

L'apprentissage profond (DL) et spécialement les réseaux de neurones à convolution (CNN) ont attiré beaucoup d'intérêts surtout dans le domaine de la vision par ordinateur et notamment la classification d'images. L'apprentissage profond a été proposé initialement pour la représentation de données (image, audio, texte, etc) en imitant le mécanisme d'abstraction multicouches du cerveau humain. Ils ont décrit un système combinant l'apprentissage des caractéristiques et la classification dans un unique processus d'apprentissage. Les CNNs peuvent être utilisés pour l'apprentissage non supervisé des caractéristiques et peuvent servir donc, d'outils pour générer de descripteurs de bas niveau. En effet, avec leur architecture multicouche hiérarchique, les CNNs sont capables d'apprendre et reconnaître des motifs visuels directement à partir des pixels des images. Ils peuvent être également utilisés comme des apprenants renvoyant en sortie les classes des échantillons tout en faisant un apprentissage combiné des caractéristiques et de la séparation entre les différentes classes. De plus, les CNNs sont bien connus pour leur adaptation au pré-traitement minimal et leur robustesse à la distorsion.

Les réseaux de neurones profonds ont été largement étudiés et appliqués pour la classification des images, l'analyse des scènes, etc. et ont démontré leur efficacité en atteignant de bonnes performances comparables à celles des meilleures approches de l'état de l'art.

Dans l'apprentissage profond, chaque couche apprend à transformer ses données d'entrée en une représentation plus abstraite.

I.11.1. Fonctionnement d'apprentissage profond

L'apprentissage profond basée sur un réseau de neurones artificiels inspirés du cerveau humain. Ce réseau est constitué de dizaines ou même de centaines de couches de neurones, chacune recevant et interpréter les informations de la couche précédente.

A chaque étape, les mauvaises réponses sont éliminées et renvoyées aux niveaux amont pour ajuster le modèle mathématique. Par mesure, le programme réorganise les informations en plus complexes blocs. Lorsque ce modèle est ensuite appliqué à d'autres cas, il est normalement capable de reconnaître les nouveaux concepts, avec autre manière, plus le système accumule différentes expériences mieux il effectuer, c-à-dire les performances des algorithmes d'apprentissage en profondeur sont améliorées lorsque la quantité de données augmente

. L'architecture de l'apprentissage en profondeur est flexible pour être modifiée par de nouveaux problèmes dans l'avenir, et même la robustesse aux variations naturelles des données est automatiquement apprise.

I.11.2. L'apprentissage en profondeur et la détection d'objets

Il existe deux types principaux de détecteurs d'objets. D'une part, nous avons des détecteurs à deux-étages comme "Faster R-CNN" ou "Mask R-CNN", tel que la première utilise un réseau de proposition de région pour générer des régions d'intérêts dans la première étape et la deuxième envoyer à la région des propositions dans le pipeline pour la classification des objets et régression de la boîte englobante.

D'autre part, nous avons des détecteurs à un seul étage comme YOLO et SSD, qui traitent la détection d'objets comme un simple problème de régression en prenant une image d'entrée et en apprenant les probabilités de classe et la délimitation coordonnées. Ces modèles atteignent des taux de précision inférieurs, mais sont beaucoup plus rapides que les modèles de détecteurs d'objets à deux-étages.

Il existe de nombreuses méthodes utilisées dans le domaine de la détection humaine-objet.

I.11.2-1 Le modèle R-CNN

Pour extraire des fonctionnalités de haut niveau, il est important d'améliorer la qualité des candidats boîtes englobantes et pour obtenir une structure profonde. Pour résoudre ces problèmes, R-CNN a été proposé par Ross Girshick en 2014 comme une alternative exhaustive pour capturer l'emplacement des objets dans les images. Ce commence à initialiser de petites régions dans l'image d'entrée puis les fusionne en fonction de la variété de métriques de couleur et de similarité à l'aide d'un groupement hiérarchique comme dans (Figure I-10). [10]

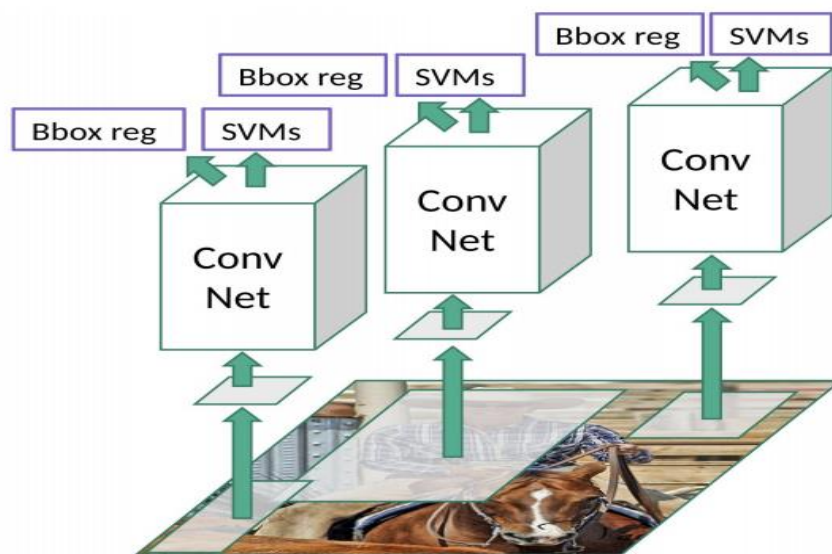


Figure I-10: Architecture R-CNN.

Ainsi, le groupe final est une boîte contenant toute l'image. La sortie est quelques chiffres de régions pouvant contenir un objet. Les R-CNN, consistait en trois étapes simples :

- ♦ Scannez l'image d'entrée pour d'éventuels objets à l'aide d'un algorithme appelé "recherche sélective" générant environ 2000 propositions de région
- ♦ Exécutez un réseau de neurones convolutifs à la tête de chacune de ces propositions de région.
- ♦ Prenez la sortie de chaque CNN et introduisez-la dans un SVM pour classer la région et dans une régression linéaire pour resserrer la boîte englobante de l'objet.

Tout le processus de détection d'objets à l'aide de R-CNN a trois modèles, CNN pour l'extraction des caractéristiques, classificateur SVM pour l'identification d'objets, et modèle de régression pour resserrer les boîtes englobantes. [9]

I.11.2-2 Le modèle FAST R-CNN

Le but de cette méthode est de réduire le temps, qui est lié au nombre élevé de modèles nécessaires pour analyser toutes les propositions de régions [11]. Dans Fast R-CNN, nous présentons l'image d'entrée au CNN, qui à son tour génère les cartes de caractéristiques convolutionnelles. En utilisant de ces cartes, les régions de propositions sont extraites. Nous utilisons ensuite la couche "pooling" pour remodeler toutes les régions proposées dans une taille fixe, de sorte qu'elles puissent être introduites dans un réseau entièrement connecté.[12]

Fast R-CNN résout deux problèmes majeurs de R-CNN, c'est-à-dire en passant une seule région par image au lieu de 2 000 au ConvNet, et en utilisant un au lieu de trois modèles différents pour extraire caractéristiques, classification et génération de boîtes englobantes (**Figure I-11**). [12]

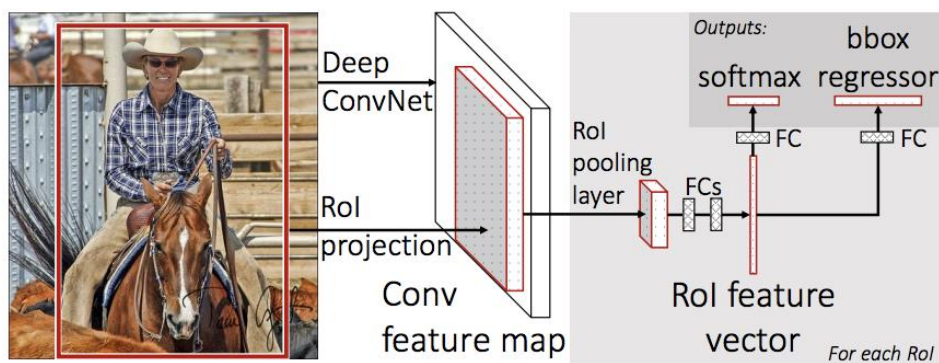


Figure I-11: Architecture FAST R-CNN.

Les étapes ce modèle comme suite :

- ✓ Transmettre l'image d'entrée à un ConvNet, qui à son tour génère les régions d'intérêt.
- ✓ La couche de pooling est appliquée sur toutes ces régions pour les remodeler selon l'entrée du ConvNet. Ensuite, chaque région est transmise à un réseau entièrement connecté.
- ✓ La couche softmax est utilisée au-dessus du réseau entièrement connecté pour les classes de sortie, une couche de régression linéaire est également utilisée parallèlement à la limite de sortie coordonnées de boîte pour les classes prédites.

I.11.2-3 Le modèle FASTER R-CNN

Cette méthode est une combinaison entre le R-CNN et le modèle Fast R-CNN. L'image entière servira comme entrée au modèle CNN, après le modèle CNN produit les cartes de caractéristique. Toutes les cartes de caractéristiques seront glissées par une fenêtre de taille 3x3 et génère un vecteur de caractéristiques lié à deux couches entièrement connectées, une pour la classification et l'autre pour la régression (**Figure I-12**). [13]

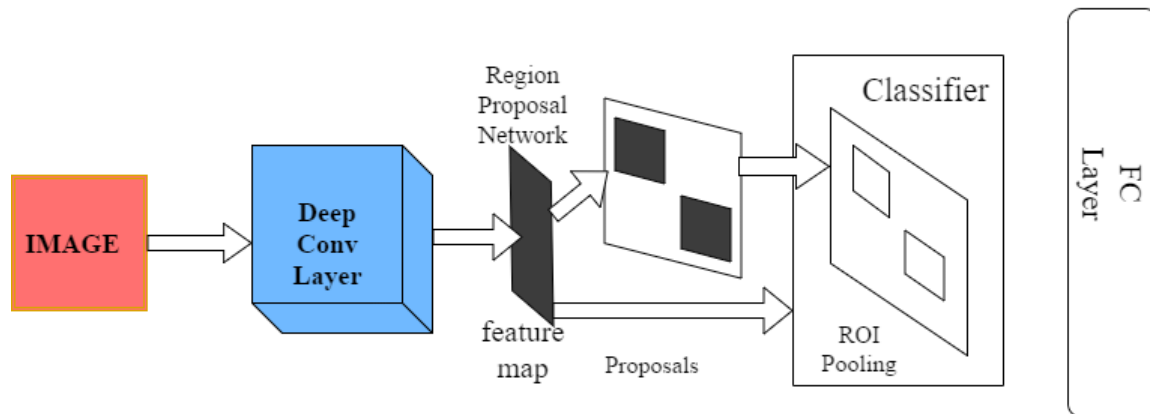


Figure I-12: Architecture FASTER R-CNN.

Les étapes ci-dessous sont généralement suivies dans une approche FASTER R-CNN :

- Prenez une image en entrée et transmettez-la au ConvNet, qui renvoie la carte des caractéristiques de cette image.
- Le réseau de proposition de région est appliqué sur ces cartes de caractéristiques. Cela renvoie les propositions l'objet avec leur score d'objectivité.
- La couche de pooling est appliquée sur ces propositions pour faire descendre toutes les propositions même taille.
- Les propositions sont transmises à une couche entièrement connectée qui a une couche soft max et une couche de régression linéaire, pour classer et produire les boîtes englobantes pour objets.

I.11.2-4 Le modèle YOLO

Cette méthode consiste essentiellement à obtenir l'image et à la diviser en une grille $S \times S$, chacune contenant m boîtes englobantes. Pour ce dernier, le réseau génère une probabilité de classe et des valeurs de décalage. Les boîtes englobantes dont la probabilité de classe est supérieure à une valeur seuil sont sélectionnées et utilisées pour localiser l'objet dans l'image.

La limite de cette méthode est qu'elle rencontre des difficultés avec les petits objets dans l'image. [14]

Vous trouverez ci-dessous les étapes exactes suivies par l'algorithme YOLO :

- Prend une image d'entrée de forme (608, 608, 3).
- Passe cette image à un réseau de neurones convolutifs (CNN), qui renvoie l'image un (19, 19, 5,85) sortie dimensionnelle.
- Les deux dernières dimensions de la sortie sont aplaties pour obtenir un volume de sortie de (19, 19, 425).
- Évitez de sélectionner des cases qui se chevauchent en faisant "IoU" et "Non-Max suppression".

I.11.2-5 Le modèle SSD

Le modèle prend une image en entrée qui traverse plusieurs couches convolutionnelles avec différentes tailles de filtre (10x10, 5x5 et 3x3). Pour prédire les boîtes englobantes cartes de caractéristique à partir de couches convolutionnelles à différentes positions du réseau sont utilisées. Ils sont traités par des couches convolutives spécifiques avec des filtres 3x3 appelées couches de fonctionnalités supplémentaires pour produire un ensemble de limites boîtes similaires aux boîtes d'ancrage du Fast R-CNN. [13]

I.12. Conclusion

Dans ce chapitre, nous avons introduit deux parties. Nous avons parlé dans la première partie de la représentation des images, des descripteurs et des techniques de segmentation, des techniques d'annotation, des techniques de recherche et des techniques de recommandation. Nous avons introduit dans la deuxième partie les techniques d'apprentissage automatique, de classification et d'optimisation. Sur la base des techniques d'annotation de pointe présentées, nous pouvons dire que choisir ou savoir combiner deux métriques afin d'obtenir un système de recherche d'images hybride basé sur le contenu est la clé pour obtenir un résultat efficace et robuste. Pour cette raison, nous sommes motivés par les avantages offerts par l'utilisation de techniques d'apprentissage en profondeur et les utilisons pour développer un extracteur hybride et un interprète sémantique de modèle d'images basé.



Chapitre 02:

*Les modèle d'apprentissage
ontologiques*

II. Chapitre II : Les modèle d'apprentissage ontologiques

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 métadonnées 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 dessus.

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 (**Figure II-1**), tel que, les sommets (signifié, signifiant, référent) sont interprétés respectivement comme (concept, terme, instance).

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", dit DIKW (**Figure II-2**).

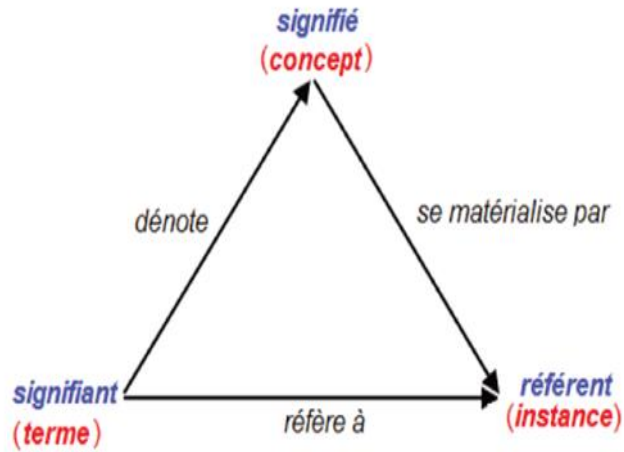


Figure II-1: le triangle sémiotique générique

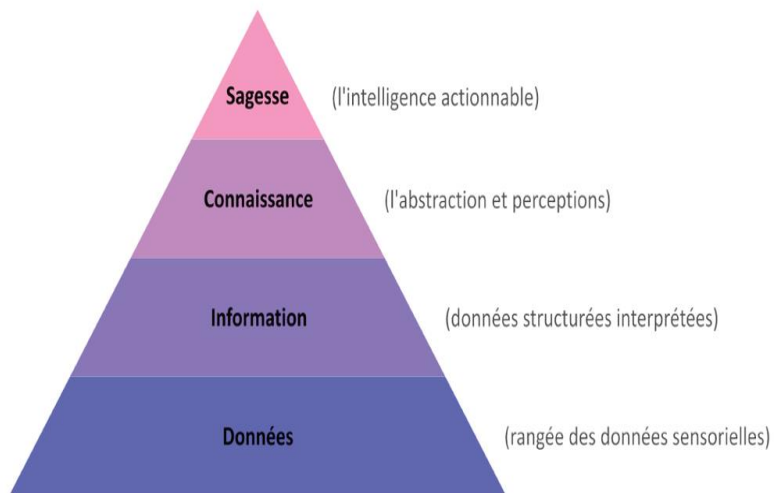


Figure II-2:La pyramide de la sagesse DIKW.

II.3. Préliminaires :

Le tableau ci-dessus (**Tableau II.1**), permet de visualiser le chevauchement des propriétés partagées par les notions de données, d’information et de connaissance. Ainsi, les données peuvent être enregistrées ou reconnues par un système d’information (reconnaissables). L’information pertinente peut être distinguée du bruit donc elle est interprétable. Enfin, la connaissance, dans le contexte d’un système d’information, suppose la capacité d’utiliser l’information, d’en déduire de nouveaux faits non explicitement présents, on dit que la connaissance est prédictive.

Tableau II-1: Les propriétés permettant de distinguer la donnée, l'information et la connaissance.

| | Reconnaisable | Interprétable | Prédicative |
|--------------|---------------|---------------|-------------|
| Données | x | x | |
| Information | x | x | |
| Connaissance | | x | x |

II.4. 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 dessus.

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. Pour permettre un traitement automatique efficace, les représentations à utiliser par les machines (les modèles) doivent être chargées de sens et cela en reliant les informations ainsi collectées et représentées à d'autres types d'informations vouées essentiellement à la sémantique sous-jacente. Cela a donné naissance à l'ingénierie ontologique.

II.5. La notion d'ontologie

Le terme "Ontologie" est un terme philosophique qui signifie une Partie de la métaphysique qui s'applique à l'être en tant qu'être, indépendamment de ses déterminations particulières. Avec l'émergence de l'ingénierie des connaissances, 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.5.1. Définitions

Il est difficile d'attribuer une unique définition définitive à la notion "ontologie" puisque celle-ci a été employée dans des contextes différents.

- ✓ N'èches et al (1991) : « 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) : « une spécification explicite d'une conceptualisation ».
- ✓ Borst (1997) : « spécification formelle d'une conceptualisation partagée ».
- ✓ Studer (1998) : « spécification formelle et explicite d'une conceptualisation partagée »
- ✓ Swartout et al., (1997) : « 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 ».

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.

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.

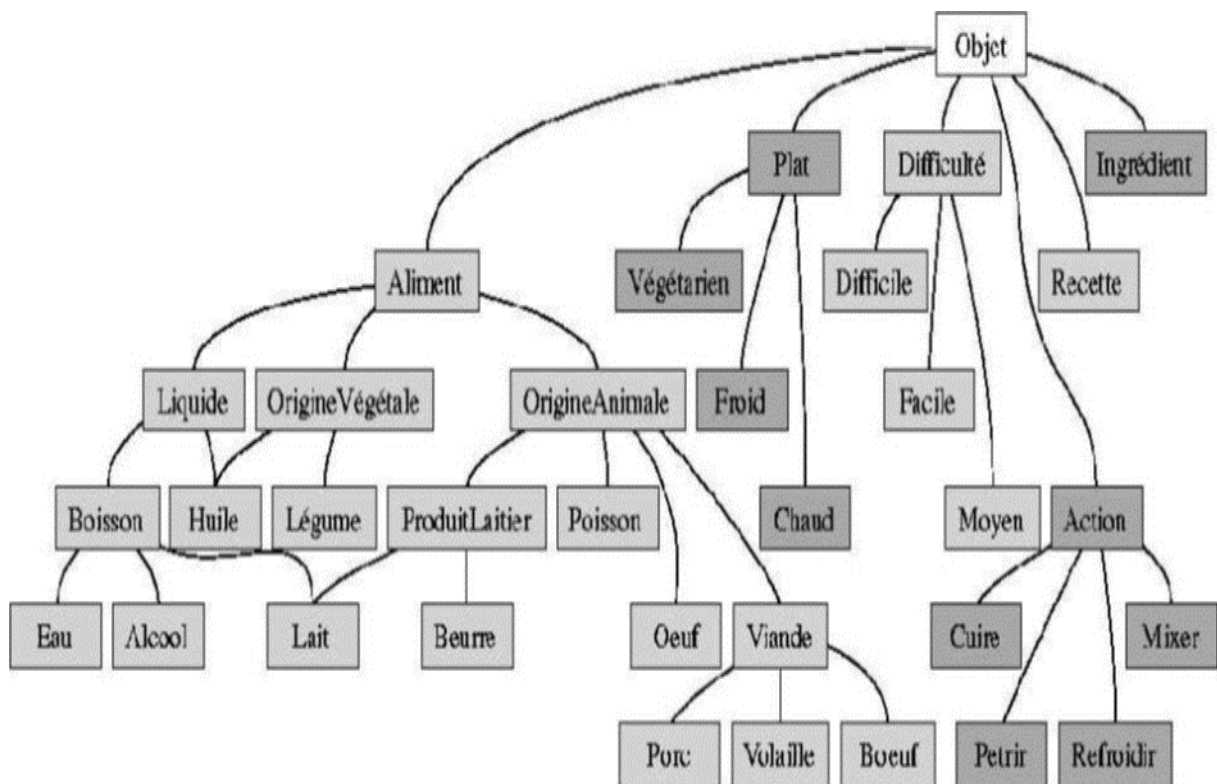


Figure II-3: Hiérarchie de concepts d'une ontologie pour les recettes de cuisine.

II.5.2. Les différents types d'ontologies :

On peut définir deux classes d'ontologies.

La première catégorie basée sur la structure de la conceptualisation, on distingue trois sous catégories :

- Les ontologies terminologiques (lexiques, glossaires...).
- Les ontologies d'information (schéma d'une BD).
- Les ontologies des modèles de connaissances.
- La deuxième catégorie basée sur le sujet de la conceptualisation, qui distingue quatre sous catégories :

Dans la deuxième catégorie

- 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 : 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.
- Les ontologies de représentation : 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.

II.5.3. Les composants de l'ontologie

Une ontologie est constituée de plusieurs composants à savoir : les concepts souvent représentés par des termes, les relations entre ces concepts, les axiomes et les instances.

II.5.3-1 Les concepts

Un concept représente un objet matériel, une notion ou une idée. Il est composé de trois parties, un ou plusieurs termes, une notion et un ensemble d'objets [15] :

- Le terme, appelé aussi label, permet de désigner le concept.
- La notion, également appelée intension du concept, correspond à la sémantique du concept, exprimée en termes de propriétés et d'attributs.

- L'ensemble d'objets, également appelé extension du concept, regroupe les objets manipulés à travers le concept, ces objets sont appelés instances du concept.

On peut identifier les propriétés que peut avoir un concept de l'ontologie comme suit [16] :

- La généralité : un concept est générique s'il ne possède pas d'extension.
- L'identité : un concept porte une propriété d'identité si cette propriété permet de conclure quant à l'identité de deux instances de ce concept.

Cette propriété peut porter sur des attributs du concept ou sur d'autres concepts.

- ❖ La rigidité : un concept est dit rigide si toute instance de ce concept en reste instance dans tous les mondes possibles.
- ❖ L'anti-rigidité : un concept est anti-rigide si toute instance de ce concept est essentiellement définie par son appartenance à l'extension d'un autre concept.
- ❖ L'unité : un concept est un concept unité, si pour chacune de ses instances, les différentes parties de l'instance sont liées par une relation qui ne lie pas d'autres instances du concept.

Aussi, deux concepts peuvent avoir les propriétés suivantes :

- ❖ L'équivalence : deux concepts sont équivalents s'ils ont la même extension.
- ❖ La disjonction : deux concepts sont disjoints si leurs extensions sont disjointes.
- ❖ La dépendance : Un concept 'A' est dépendant d'un concept 'B' si pour toute instance de 'A' il existe une instance de 'B' qui ne soit ni partie ni constituant de l'instance de 'A', et vice-versa.

II.5.3-2 Les relations

Les relations représentent un type d'interaction entre les concepts d'un domaine. Elles se décomposent généralement en deux catégories : les relations taxonomiques et les relations associatives ou non taxonomiques.

- Les relations taxonomiques organisent l'ensemble des concepts de l'ontologie dans une structure arborescente. Par exemple, un concept 'A' subsume un concept 'B' si toute propriété sémantique de 'A' est aussi une propriété sémantique de 'B'.

- Les relations associatives sont des relations d'interaction entre deux concepts qui sont non taxonomiques. Comme un exemple de relations non taxonomiques, 'Associé-a', 'est-situé-dans', 'adjacent à', 'déconnecté-à', etc.

II.5.3-3 Les axiomes

Les axiomes ont pour but de définir dans un langage logique la description des concepts et des relations permettant de représenter leur sémantique. Ils représentent les intentions des concepts et des relations du domaine et, de manière générale, les connaissances n'ayant pas un caractère strictement terminologique [17].

L'intégration des axiomes dans une ontologie peut avoir plusieurs objectifs :

- Définir la signification des composants.
- Définir des restrictions sur la valeur des attributs.
- Définir les arguments d'une relation.
- Vérifier la validité des informations spécifiées ou en déduire de nouvelles.

II.5.3-4 Les instances

Les instances constituent la définition extensionnelle de l'ontologie, et véhiculent les connaissances statiques ou factuelles, à propos du domaine du problème. L'individu est une instance de concept, autrement dit, c'est l'élément décrit par le concept.

II.5.3-5 Les fonctions

Elles constituent des cas particuliers de relation dans laquelle un élément de la relation est défini en fonction des éléments précédents

II.5.4. Classifications d'ontologies :

On peut distinguer quatre typologies selon plusieurs critères :

II.5.4-1 Classification selon l'objet de conceptualisation

Les ontologies peuvent être subdivisées en plusieurs niveaux qui sont :

- Les ontologies de type thesaurus : Sont aussi appelées taxonomie, elles servent à définir un vocabulaire de référence.

- Les ontologies du domaine : Ces ontologies expriment des conceptualisations spécifiques à un domaine. Elles sont réutilisables pour plusieurs applications de ce domaine. L'ontologie du domaine caractérise la connaissance du domaine où la tâche est réalisée.
- Les ontologies applicatives : Ces ontologies contiennent des connaissances du domaine nécessaire à une application donnée, elles sont spécifiques et non réutilisable.
- Les ontologies génériques : Ces ontologies expriment des conceptualisations valables dans différents domaines. Son sujet est l'étude des catégories des choses qui existent dans le monde. Comme les concepts de haute abstraction tels que les entités, les évènements, les états, les actions, le temps, l'espace, les relations, etc.
- Les ontologies de représentation (méta-ontologies) : Ces ontologies conceptualisent les primitives des langages de représentation des connaissances.
- Les ontologies géographiques : les ontologies de l'espace plus spécifiquement dédiées à la description des concepts qui caractérisent l'espace comme le point, la ligne, etc... Ces ontologies sont typiquement élaborées par de grands organismes de normalisation
- Les ontologies spatialisées : sont des ontologies dont les concepts sont localisés dans l'espace. Une composante temporelle est souvent nécessaire en complément pour la modélisation de l'information géographique.

II.5.4-2 Classification selon le niveau de complétude

On peut définir trois niveaux de complétude : Niveau sémantique : Ces principes correspondent à l'engagement sémantique et assurent que chaque concept aura un sens univoque et non contextuel associé. Deux concepts sont identiques si l'interprétation du terme/libellé à travers les quatre principes différentiels aboutis à un sens équivalent :

- ✚ Communauté avec l'ancêtre.
- ✚ Différence, spécification, par rapport à l'ancêtre.
- ✚ Communauté avec les concepts frères, situés au même niveau.
- ✚ Différence par rapport aux concepts frères.

Tous les concepts, caractérisés par un terme/libellé, doivent respecter les quatre principes différentiels précédents.

- ✓ **Niveau-Référentiel** : Les concepts référentiels ou formels, se caractérisent par un terme/libellé dont la sémantique est définie par une extension d'objets. L'engagement

ontologique spécifie les objets du domaine qui peuvent être associés au concept, conformément à sa signification formelle. Deux concepts formels seront identiques s'ils possèdent la même extension.

- ✓ **Niveau-Opérationnel** : Les concepts du niveau opérationnel ou computationnel sont caractérisés par les opérations qu'il est possible de leur appliquer pour générer des interfaces ou engagement computationnel.

II.5.4-3 Classification selon le niveau de détail

On peut différencier les ontologies selon le niveau de description utilisé :

- ♦ **Granularité fine** : Ce niveau correspond à des ontologies très détaillées, possédant ainsi un vocabulaire plus riche capable d'assurer une description détaillée des concepts pertinents d'un domaine ou d'une tâche.
- ♦ **Granularité large** : Ce niveau correspond à des vocabulaires moins détaillés. Par exemple les scénarios d'utilisation spécifique ou les utilisateurs sont déjà préalablement d'accord à propos d'une conceptualisation sous-jacente. Les ontologies de haut niveau possèdent une granularité large, compte tenu que les concepts qu'elles traduisent sont normalement raffinés ultérieurement dans d'autres ontologies de domaine ou d'application.

II.5.4-4 Classification selon le formalisme utilisé

On peut distinguer les ontologies selon le formalisme utilisé pour les exprimer.

- a) **Formelle** : l'ontologie est exprimée dans un langage artificiel disposant d'une sémantique formelle, permettant de prouver des propriétés de cette ontologie. L'intérêt d'une ontologie formelle est la possibilité d'effectuer des vérifications sur l'ontologie.
- b) **Semi-formelle** : l'ontologie est exprimée dans un langage artificiel définit formellement.
- c) **Informelle** : l'ontologie est exprimée en langage naturelle. Cela peut permettre de rendre plus compréhensible l'ontologie pour l'utilisateur, mais cela peut rendre plus difficile la vérification de l'absence de redondances ou de contradiction.
- d) **Semi-informelle** : l'ontologie est exprimée dans une forme restreinte et structurée de la langue naturelle ; cela permet d'augmenter la clarté de l'ontologie tout en réduisant l'ambiguïté.

II.5.5. Domaines d'application des ontologies

Dans cette section 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.5.5-1 Web Sémantique

Aujourd'hui, le Web Sémantique est le plus grand domaine d'application des technologies sémantiques. Son idée fondamentale, est d'accompagner l'extension et la croissance au terme du Web actuel dans le cadre des recommandations de W3C. Afin de garantir l'interopérabilité sur la toile on vise à compléter les ressources textuelles par des informations permettant leur interprétation univoque par des personnes et des ordinateurs.

II.5.5-2 Recherche d'information (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 un 'Sac de mots'.
- ♦ Classification de documents, C-à-dire, l'attribution de chaque document à l'une des catégories prédéfinies.
- ♦ Regroupement sémantique, C-à-dire, 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. L'extraction des rubriques est une des tâches actuelles de l'ingénierie des connaissances ayant des méthodes communes avec l'apprentissage des ontologies.
- ♦ Production de résumés automatiques.

II.5.5-3 Systèmes du type Question-Réponse

Dans ce cas, il s'agit du développement de systèmes interactifs du type Question-Réponse pour que l'utilisateur obtienne un résultat concret, et pas seulement une liste de références aux documents correspondant plus ou moins à sa requête-question.

II.5.5-4 Intégration de bases de données hétérogènes

L'intégration de bases de données hétérogènes est un problème complexe qui est devenu crucial pour fournir aux utilisateurs une interface unifiée permettant l'accès à des ressources hétérogènes. Dans ce cas, les ontologies sont utilisées pour spécifier le contenu des ressources hétérogènes.

II.5.5-5 Ingénierie logicielle

Le principe de conceptualisation et de distinction des objets selon leurs propriétés est universel, il est aussi utilisé en ingénierie logicielle (SE). La tendance est à l'unification et la spécification des processus pendant tout le cycle de vie d'un Système d'Information. Il s'agit de la stratégie MDD et de l'architecture de construction de logiciel basée sur modèles MDA. La stratégie MDD permet d'économiser du temps et des ressources, de réduire les charges et de garantir la flexibilité des processus de mise en œuvre, de la maintenance, des tests et simulation et l'interopérabilité des SI grâce aux dessins de modules et modèles de données lisibles automatiquement.

II.5.6. Les ressources lexicales

Trois types des ressources lexicales sont de plus en plus utilisées dans les méthodes de NLP proposant des solutions au problème de l'ambiguïté des mots, les dictionnaires, les corpus annotés et les bases d'information lexicales. Dans cette section, nous allons d'écrire brièvement les particularités de chaque type d'outil et présenter plusieurs projets à la fois typiques et signifiants.

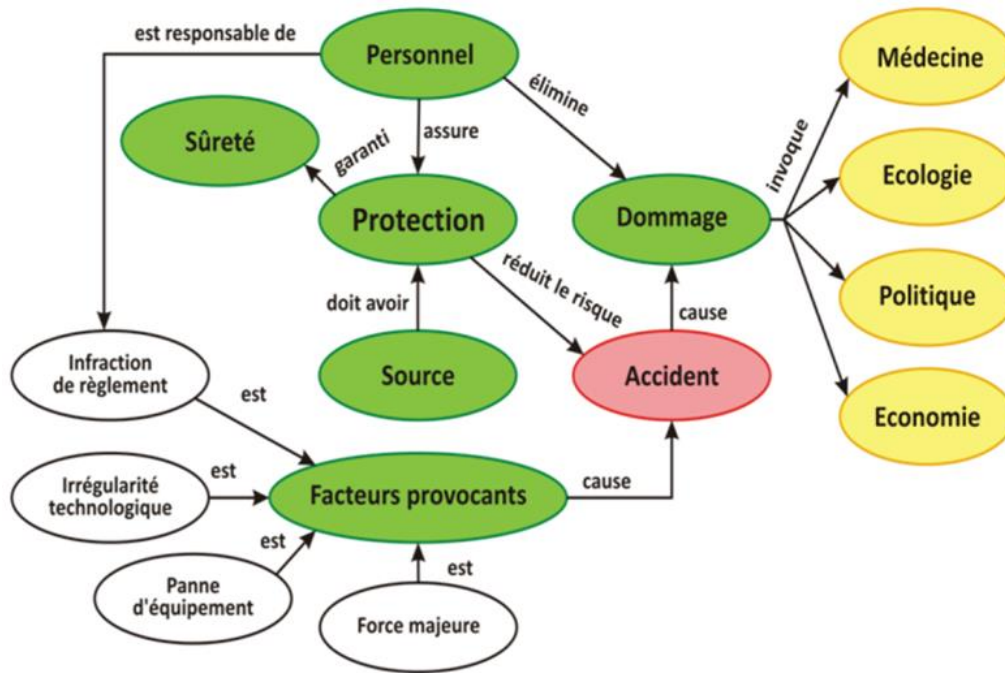


Figure II-4: Fragment du modèle de domaine

II.5.6-1 FrameNet

FrameNet est une ressource linguistique en anglais, lisible par la machine et par l'homme. Sa création a débuté en 1997 dans le cadre d'un grand projet de lexique sémantique, créé sous la conduite de Ch. Fillmore. FrameNet vise à la description de la compatibilité sémantique et syntaxique des mots en fonction de leurs valences qui, à leur tour, dépendent du sens contextuel du mot. Le contenu de FrameNet ne cesse d'augmenter. On peut consulter sa statistique courante sur le site de Berkeley ⁷. Nous présentons plusieurs chiffres récents afin de montrer son ampleur : il y a plus de 1 100 cadres lexicaux hiérarchisés. L'index contient plus de 13 400 unités lexicales (LU) illustrées par des exemples textuels annotés. Au total, il y a plus de 28 000 sets de textes complets annotés et plus de 227 000 de sets textuels.

II.5.6-2 WordNet

Le thesaurus anglais WordNet est apparu sur l'Internet en 1995 mais son développement a été lancé à l'Université de Princeton dès 1984 sous la direction du psycholinguiste George Miller. Sa version 3.0 comprend environ 155 000 lexèmes avec les exemples organisées en 117 000 ensembles de synonymes, pour la langue anglaise. Chaque ensemble de synonymes peut être envisagé comme la présentation lexicalisée d'une notion générale, ou concept.

II.5.6-3 RuTez

Le thésaurus russe RuTez est construit depuis 1994 au centre de recherches du traitement de l'information de l'Université Lomonosov de Moscou. Actuellement il comprend plus de 51 500 concepts, plus de 155 millions d'entrées lexicales (mots ou phrases) et plus de 200 000 relations entre les concepts. Au total, compte tenu de la hiérarchie des liens, le thésaurus comprend plus de 2 millions de relations entre les concepts. Les concepts du thésaurus sont également assortis d'entrées lexicales en anglais avec plus de 125 000 mots et phrases.

II.5.6-4 BabelNet

Les concepteurs définissent BabelNet comme un " dictionnaire encyclopédique" fournissant des concepts et des entités-nommées lexicalisés qui sont liées par des relations sémantiques variées, et ceci en plusieurs langues.

BabelNet encode les connaissances sous forme d'un graphe orienté et labellisé où chaque nœud dans l'ensemble des nœuds correspondant à un concept ou à une entité nommée, et chaque arc dans l'ensemble des arcs liant les paires de concepts correspond une relation sémantique qui peut être spécifiée en WordNet ou non.

II.5.6-5 VerbNet et VerbOcean

VerbNet et VerbOcean sont deux ressources lexicales dédiées uniquement aux verbes. Elles ont à peu près la même taille de quatre mille lemmes de verbes. VerbNet est une base lexicale de verbes compatibles avec WordNet et avec FrameNet et construite sur le principe du cadre.

VerbeOcean est un réseau sémantique de relations entre les verbes qui sont proches sémantiquement, construit à partir de la classification des verbes de B.Levin, mais sa classification est plus détaillée ou les classes initiales ont été réorganisées en sous-classes complémentaires afin de garantir la cohérence sémantique et syntaxique des verbes-membres. Ensuite, les classes de verbes peuvent être examinées à différents niveaux.

De son côté, Verbe Océan distingue cinq types de relations, la similitude, l'antonymie, l'inclusion, la force et la relation temporelle.

II.6. Langages et formalismes d'ontologie

II.6.1. XML

Le langage XML est un méta-langage permettant de représenter un document texte de manière arborescente en utilisant un système de balisage. XML fournit ainsi un cadre de structuration des données qui peut être employé pour la définition rapide et sans ambiguïté syntaxique d'un format de document. Ce langage a été élaboré pour faciliter l'échange, le partage et la publication des données à travers le web. Ainsi, la majorité des langages/modèles proposés pour le web sémantique sont exprimés en XML.

XML permet de structurer un document en définissant ses propres balises en fonction des besoins et sans tenir compte ni de la signification de cette structure ni des systèmes informatiques qui vont l'exploiter. Des standards comme XPath et XQuery ont été développés afin de parcourir et d'interroger l'arborescence XML des documents.

II.6.2. RDF

RDF est un modèle de données pour les objets et les relations entre eux, fournissant des sémantiques simples pour ce modèle de données qui peuvent être représentés en XML.

RDF est, aussi, un langage pour les métadonnées sur le web. La syntaxe de RDF est basée sur celle de XML. Le modèle de base de ce langage est conçu pour permettre d'associer des attributs aux ressources du Web en utilisant la description de métadonnées sémantiques. Il est la spécification d'un système d'expression d'assertions sémantiques simples. Ainsi, RDF permet de voir le Web comme un ensemble de ressources reliées par les liens étiquetés, et permet, alors d'accroître la facilité de traitement automatique des ressources Web.

La structure fondamentale de toute expression en RDF est une collection de triplets, chacun composé d'un sujet, un prédicat et un objet.

Ressource (Sujet) : une entité d'informations pouvant être référencée par un identificateur. Cet identificateur doit être une URI. Les sujet ou les ressources sont variés, une page Web, une partie d'une page Web, un site Web complet, un objet non accessible par le Web comme un livre, etc.

Propriété (prédicat) : un aspect spécifique, un attribut, une caractéristique ou relation utilisée pour décrire une ressource.

Valeur : Un littéral (simple chaîne de caractères) ou une ressource.

II.6.3. OWL

Le langage d'ontologie Web OWL est un langage Web sémantique conçu pour représenter des connaissances riches et complexes sur des objets, des groupes d'objets et des relations entre des objets. OWL est un langage basé sur la logique de calcul tel que les connaissances exprimées dans OWL peuvent être exploitées par des programmes informatiques.

La figure (**Figure II-5**) montre Les trois couches OWL **OWL Lite**, est le sous langage de OWL le plus simple. Il est destiné aux utilisateurs qui ont besoin d'une hiérarchie de concepts simple. **OWL DL**, est plus complexe que OWL Lite, permettant une expressivité bien plus importante. OWL DL est fondé sur la logique descriptive, un domaine de recherche étudiant la logique, et conférant donc à OWL DL son adaptation au raisonnement automatisé. Malgré sa complexité, il garantit la complétude des raisonnements et leur décidabilité. **OWL Full**, est la version la plus complexe d'OWL, mais également celle qui permet le plus haut niveau d'expressivité. Il est destiné aux situations où il est plus important d'avoir un haut niveau de capacité de description, quitte à ne pas pouvoir garantir la complétude et la décidabilité des calculs liés à l'ontologie.

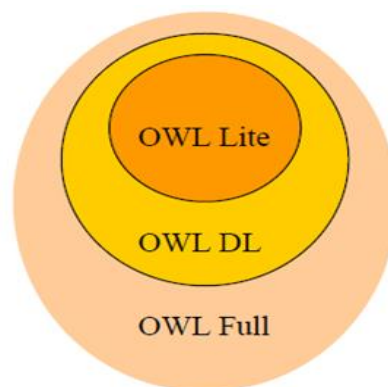


Figure II-5:Les couches de L'OWL

II.6.4. Les logiques descriptives (DL)

Les logiques descriptives sont une famille de langages formels de représentation des connaissances, de nombreux DL sont plus expressifs que la logique propositionnelle mais moins expressifs que la logique du premier ordre. Contrairement à ces derniers, les problèmes de raisonnement de base pour les DL sont généralement décidables, et des procédures de décision efficaces ont été conçues et mises en œuvre pour ces problèmes. Il existe des logiques

de description générales, spatiales, temporelles, spatio-temporelles et floues, et chaque logique de description présente un équilibre différent entre la puissance expressive et la complexité du raisonnement en prenant en charge différents ensembles de constructeurs mathématiques.

Les DLs sont utilisés en intelligence artificielle pour décrire et raisonner sur les concepts pertinents d'un domaine d'application Il est particulièrement important pour fournir un formalisme logique aux ontologies et au Web sémantique. L'application la plus notable des DL et OWL est en informatique biomédicale où DL aide à la codification des connaissances biomédicales.

II.7. Conclusion

L'objectif principal d'une ontologie est de modéliser un ensemble de connaissances dans un domaine donné. Ainsi, ils jouent un rôle important emploi de la représentation formelle des connaissances dans le monde réel et traduire ainsi un consensus explicite. Dans le CBIR, les ontologies peuvent être utilisées pour, d'une part, modéliser des ressources Web à partir de représentations conceptuelles des domaines concernés et, d'autre part, dans le but d'en permettre un traitement automatique. La partie intelligente est de mélanger deux grands champs ensemble de telle manière que, les deux champs affectent chacun autres et surmonter les problèmes et améliorer leur robustesse et leur efficacité. Dans ce chapitre, nous avons détaillé tout ce qui concernait les modèles d'apprentissage de l'ontologie.



Chapitre 03:

*Interpretation semantique
du contenu d'image*

III. Chapitre III : Interprétation sémantique du contenu d'images.

III.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 avec anaconda. 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.

III.2. Présentation des outils de développement

III.2.1. Le matériel

Le matériel réalisé est PC personnel HP I3 avec un 4GB capacité mémoire, et un processeur Intel® coré™ 2 Duo 2.50 GHz, avec Windows 10 édition intégral, service pack 1 64 bits type système.

III.2.2. Anaconda

Anaconda est un outil en distribution libre et open source, destinée à la programmation Python et R. Il est véritablement utilisé en science de données, apprentissage automatique et l'intelligence artificielle car il contient plusieurs packages nécessaires dans ce domaine notamment Python, Numpy, Panda, Jupyter, etc. Et comme le langage Python, il est multiplateforme.

❖ La distribution Anaconda



Figure III-1: symbole de anaconda

La distribution Anaconda est utilisée par plus de 6 millions d'utilisateurs. La version d'installation comprend plus de 250 paquets populaires en science des données adaptés pour Windows, Linux et MacOS. Plus de 7 500 paquets open-source supplémentaires peuvent être

installés à partir de PyPI ainsi que du gestionnaire de paquets et d'environnements virtuels conda.

Elle comprend également une interface graphique, Anaconda Navigator, qui est une alternative graphique à l'interface de ligne de commande.

La grande différence entre Conda et le gestionnaire de paquets pip consiste dans la gestion des dépendances des paquets.

❖ Le Navigateur Anaconda

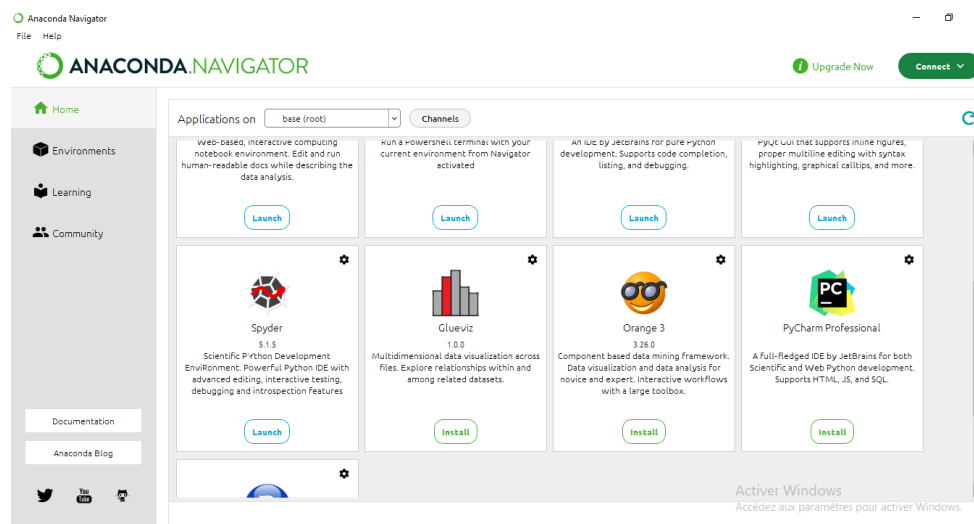


Figure III-2: Interface du navigateur anaconda

Le Navigateur Anaconda est une interface graphique (GUI) incluse dans la distribution Anaconda, et qui permet aux utilisateurs de lancer des applications, mais aussi de gérer les librairies conda, les environnements et les canaux sans utiliser la moindre ligne de commande.

Le Navigateur peut également accéder à des librairies présentes sur le Cloud Anaconda ou dans un Repository Anaconda local, afin de les installer dans un environnement, les exécuter et les mettre à jour. Il est disponible pour Windows, macOS et Linux.

Les applications suivantes sont disponibles par défaut dans le navigateur :

- JupyterLab
- Jupyter Notebook
- QtConsole⁵
- Spyder

- Glue
- Orange
- RStudio
- Visual Studio Code.

❖ **Spyder**



Figure III-3: le symbole spyder

C'est un environnement de développement interactif puissant pour le langage Python. Il a des fonctionnalités d'édition avancées, des tests interactifs, le débogage et l'introspection, ainsi qu'un environnement de calcul numérique. Grâce au support d'IPython (interpréteur Python interactif amélioré) et les bibliothèques Python populaires telles que NumPy, SciPy ou matplotlib (Tracé interactif 2D / 3D). Spyder peut également être utilisé comme bibliothèque qui fournit des widgets puissants liés à la console pour nos applications basées sur PyQt. Il peut être utilisé pour intégrer une console de débogage directement dans la conception de votre interface utilisateur graphique.

Spyder Créé et développé par Pierre Raybaut en 2008, Spyder est maintenu, depuis 2012, par une communauté de développeurs qui ont pour point commun d'appartenir à la communauté Python scientifique.

En comparaison avec d'autres IDE pour le développement scientifique, Spyder a un ensemble unique de fonctionnalités - multiplateforme, open-source, écrit en Python et disponible sous une licence non-copyleft. Spyder est extensible avec des plugins, comprend le support d'outils interactifs pour l'inspection des données et incorpore des instruments d'assurance de la qualité et d'introspection spécifiques au code Python, tels que Pyflakes, Pylint et Rope.

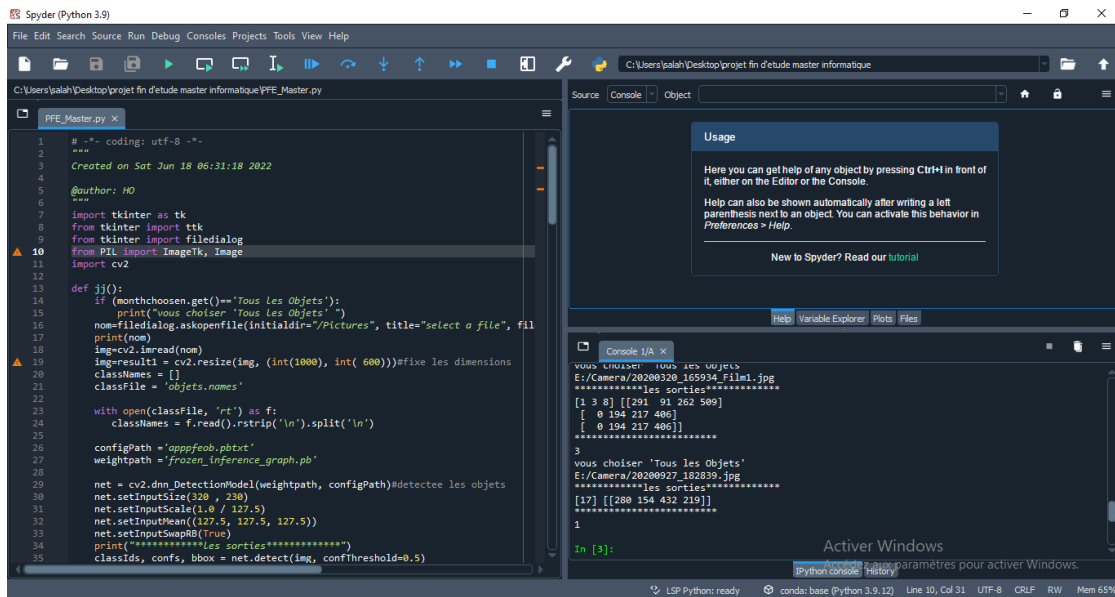


Figure III-4: Interface du spyder

III.2.3. Python



Figure III-5: Symbole du python

Python est un langage de programmation interprété, multi-paradigme et multiplateformes. Il favorise la programmation impérative structurée, fonctionnelle et orientée objet. Il est doté d'un typage dynamique fort, d'une gestion automatique de la mémoire par ramasse-miettes et d'un système de gestion d'exceptions il est ainsi similaire à Perl, Ruby, Scheme, Smalltalk et Tcl.

Le langage Python est placé sous une licence libre proche de la licence BSD et fonctionne sur la plupart des plates-formes informatiques, des smartphones aux ordinateurs centraux, de Windows à Unix avec notamment GNU/Linux en passant par macOS, ou encore Android, iOS, et peut aussi être traduit en Java ou .NET. Il est conçu pour optimiser la productivité des programmeurs en offrant des outils de haut niveau et une syntaxe simple à utiliser.

Il est également apprécié par certains pédagogues qui y trouvent un langage où la syntaxe, clairement séparée des mécanismes de bas niveau, permet une initiation aisée aux concepts de base de la programmation.

III.3. Présentation de l'application

Nous allons maintenant présenter les résultats obtenus grâce à des expériences réalisées par l'application de notre approche d'interprétation sémantique du contenu d'une images.

Le schéma de notre système est le suivant :

1. Choisir le type d'objet à détecter.
2. La sélection d'image.
3. Affichage des résultats, c-a-dire, les types des objets détecter et le nombre total des objets.

III.3.1. L'application

La fenêtre principale de notre application est :

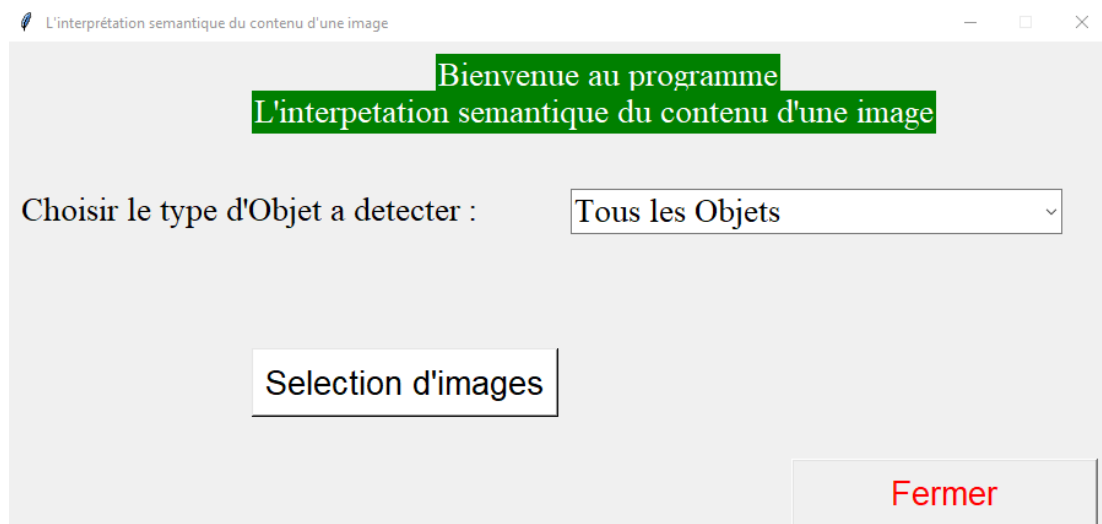


Figure III-6: Interface d'application

Tout d'abord, il faut choisir le type d'objet a détecter (ou tous les objets).

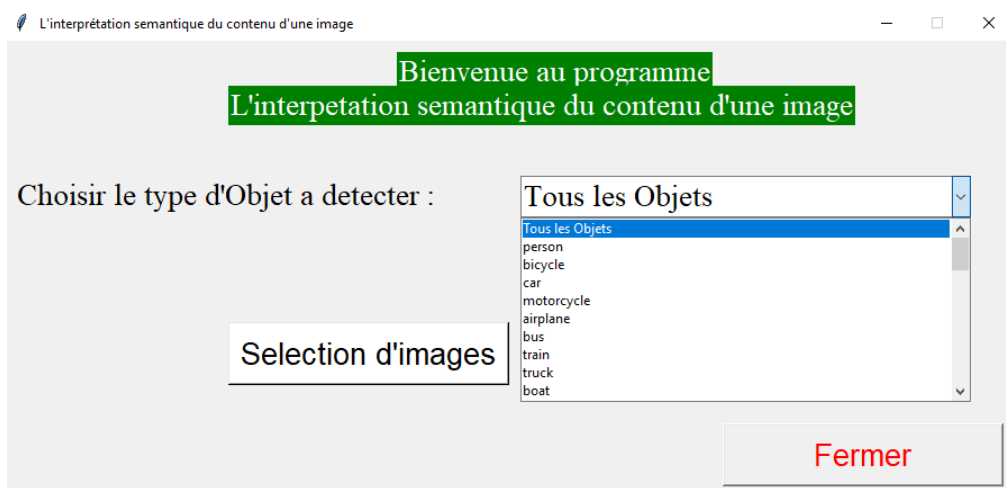


Figure III-7: Selection de type d'objet

Puis, cliquer sur le bouton "Selection d'image" dans la fenêtre (l'interface) pour ouvrir une boite de dialogue pour choisir un fichier d'images à partir d'un emplacement local.

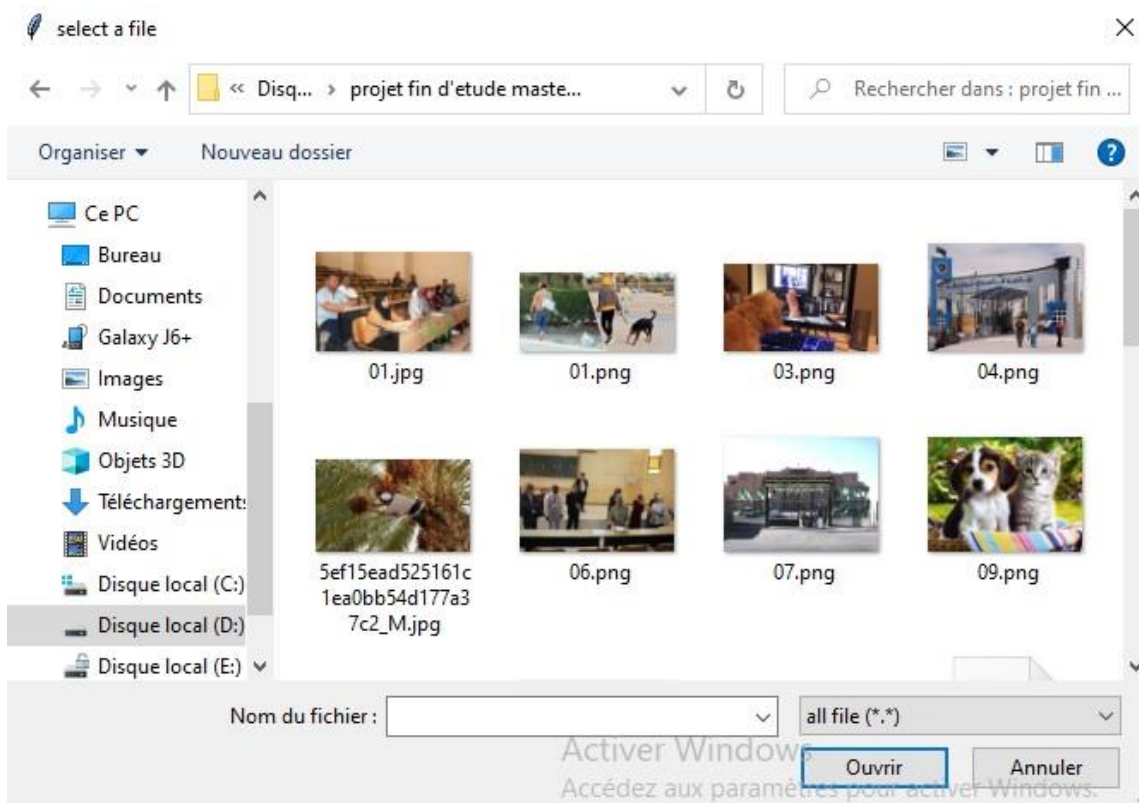


Figure III-8: Selection d'image

III.4. Les tests et les résultats

Pour bien tester notre application et juger sa performance nous tester plusieurs échantillons d'images pour le but de faire l'interprétation sémantique de ces contenus, et voici quelques résultats d'exécution :

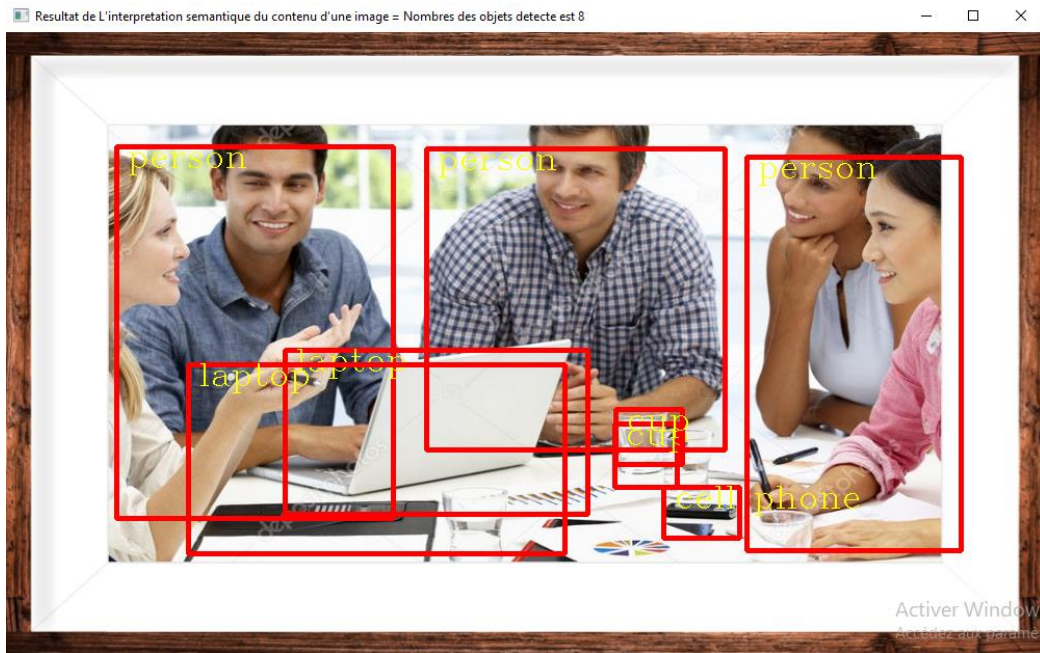


Figure III-9: Detection d'objet, exemple 1

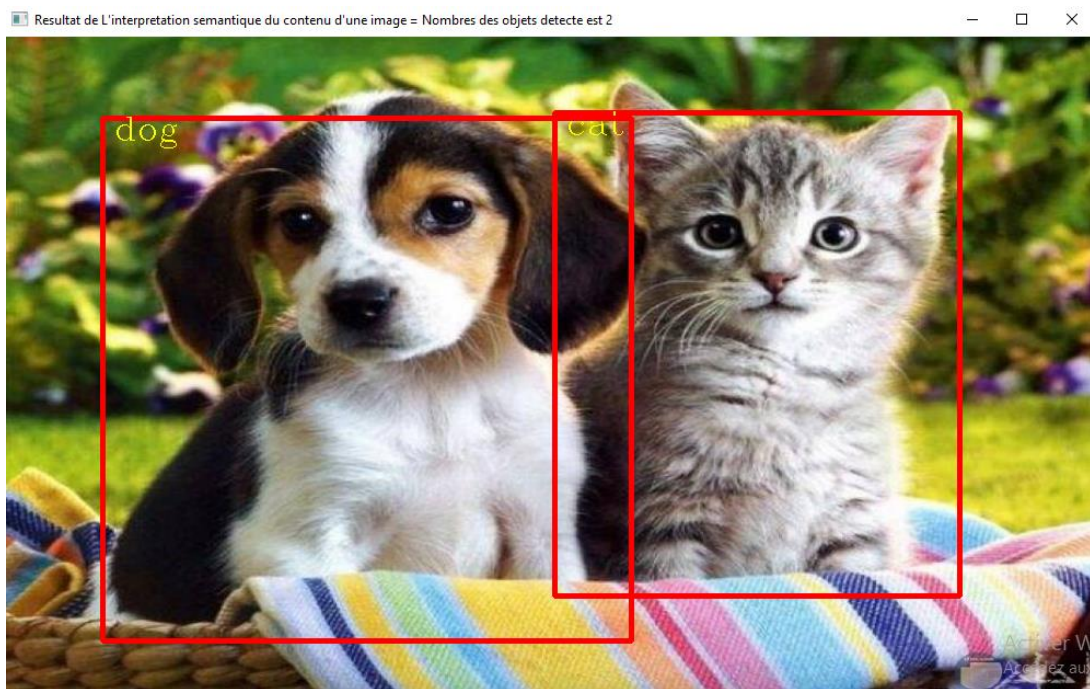


Figure III-10: Detection d'objet, exemple 2



Figure III-11: Detection d'objet, exemple 3

III.5. Discussion des résultats

Les résultats de notre application sont présentés sur le tableau (Tableau III-1), qui discute le taux d'erreur par rapport la nombre d'objet pour une échantillon fixe (de 20 images).

Tableau III-1: le taux d'erreur par rapport la nombre d'objet.

| | | | | | | |
|--------------------------|----|----|----|----|----|----|
| Nbr d'objet dans l'image | 1 | 2 | 3 | 5 | 6 | 7 |
| Nmbr d'échantillon | 20 | 20 | 20 | 20 | 20 | 20 |
| Erreue % | 0 | 4 | 9 | 12 | 16 | 20 |

On remarque que malgré le fonctionnement parfait de notre programme il y a des erreurs concerne la détection des objets, nous avons obtenu un résultat de 80.17% comme meilleur résultat d'exécution.

On remarque, que le taux d'erreur augment lorsque le nombre d'objet augment, a cause du chevauchement des objet dans l'image, par exemple (Figure III-12).

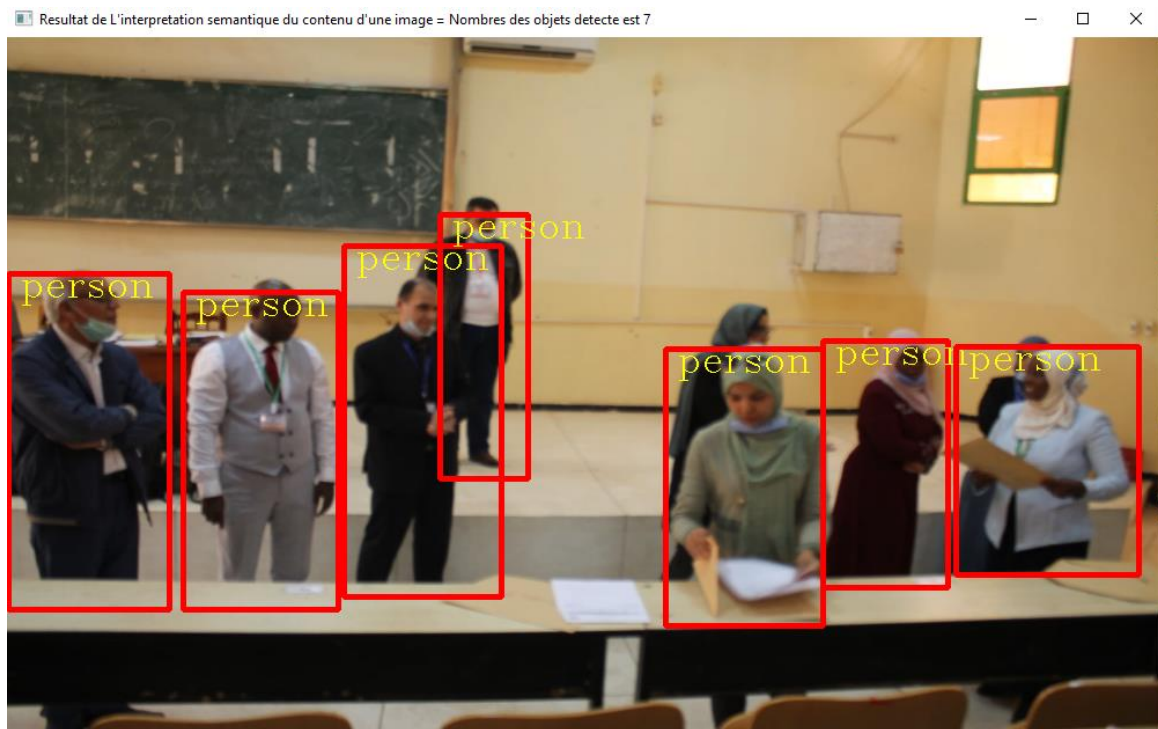


Figure III-12: Detection d'objet, exemple 4

L'application indique qu'il y a 7 objets, alors que réellement il y a 9 objets.

III.6. Conclusion

Après avoir achevé notre conception nous avons donné les outils nécessaires pour la réalisation de notre travail. Nous avons présenté aussi l'environnement de développement. A la fin nous avons présenté notre application en donnant quelques captures d'écran qui expliquent le déroulement et le fonctionnement de notre travail, ainsi que les résultats obtenus. On a donné un exemple d'interprétation sémantique par notre application d'un ensemble d'images et la détection de type d'objet et le nombre total d'objet. Nous avons montré l'impact de la méthode d'interprétation sémantique sur les résultats de la classification et la recherche d'images.



*Conclusion
générale*

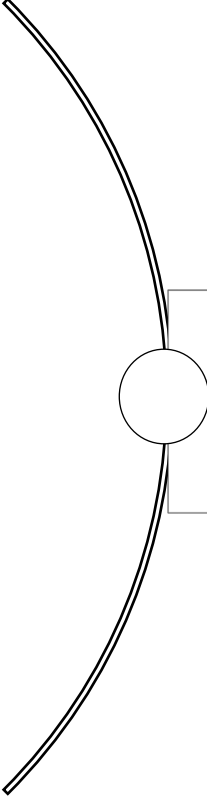
Conclusion générale

La recherche d'images est devenue un domaine de recherche actif qui se développe très rapidement. L'explosion du nombre de collections d'images personnelles et professionnelles ainsi que sur le Web a rendu le développement des outils qui organisent ces données une nécessité. L'objectif principal des moteurs de recherche d'images est de localiser les images pertinentes à l'utilisateur avec une bonne précision et avec un minimum de temps.

Dans ce mémoire, nous avons abordé le problème d'indexation et de recherche d'images. Plus précisément, nous sommes focalisés sur l'indexation et la recherche d'images basée sur la sémantique, qui permet d'une part d'améliorer les performances du système en réduisant considérablement le temps d'accès lors de la recherche et d'autre part cela améliora la qualité des résultats obtenus lors de la recherche aussi. Le travail réalisé s'inscrit dans ce contexte particulier, son objectif est d'implémenter un système d'interprétation du contenu d'images par la sémantique.

Nous avons réalisé comme application un système d'interprétation sémantique du contenu d'images. Les résultats obtenus étaient acceptables, nos méthodes ont atteint une précision de 80.17 % mais pas satisfaisante.

La perspective de notre travail est d'enrichir notre système d'interprétation sémantique du contenu d'images par d'autres concepts et relations sémantiques pour couvrir une variété d'images avec une sémantique plus riche.



*Références
bibliographiques*

Références bibliographiques

- [1] David A Forsyth and Jean Ponce. Computer vision: a modern approach. Pearson, 2012.
- [2] Shiv Ram Dubey. A Decade Survey of Content Based Image Retrieval using Deep Learning. arXiv preprint arXiv:2012.00641, 2020.
- [3] Abdelkader Hamadi. Utilisation du contexte pour l’indexation sémantique des images et vidéos. Intelligence artificielle. Université de Grenoble, 2014.
- [4] Koen EA Van de Sande, Theo Gevers, and Cees GM Snoek. A comparison of color features for visual concept classification. Proceedings of the 2008 international conference on Content-based image and video retrieval, pages 141–150, 2008.
- [5] what is the semantic gap in computer vision, https://en.wikipedia.org/wiki/Semantic_gap, seen in 11/01/2020.
- [6] Wei-Ying Ma and Bangalore S Manjunath. Netra: A toolbox for navigating large image databases. Multimedia systems, 7(3):184–198, 2010.
- [7] Support vector machine, <https://www.analyticsvidhya.com/blog/2017/09/understaing-support-vector-machine-example-code>, seen 17/03/2020.
- [8] deep learning, <https://www.tutorialandexample.com/deep-learning-tutorial>, seen 3/08/2020.
- [9] Review of Deep Learning Algorithms for Object Detection, <https://medium.com/comet-app/review-of-deep-learning-algorithms-for-object-detection>, seen in 8/08/2020
- [10] R. Girshick, J. Donahue, T. Darrell, and J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation,” in CVPR, 2014.
- [11] Object detection, <https://www.analyticsvidhya.com/blog/2018/10/a-step-bystep-introduction-to-the-basic-object-detection-algorithms-part-1/>, seen in 10/08/2020.
- [12] Review of Deep Learning Algorithms for Object Detection, <https://medium.com/comet-app/review-of-deep-learning-algorithms-for-object-detection-c1f3d437b852>, seen in 10/08/2020.
- [13] Deep Learning for Object Detection: A Comprehensive Review, -73930816d8d9, seen in 11/08/2020.
- [14] M Manoj krishna¹, M Neelima², M Harshali³, M Venu Gopala Rao⁴, Image classification using Deep learning (pdf).
- [15] Olena Orobinska, Jean-Hugues Chauchat, and Natalya Sharonova. Methods and models of automatic ontology construction for specialized domains. In Computational linguistics and intelligent systems (COLINS 2017). National Technical University «KhPI», 2017.
- [16] Wilson Wong, Wei Liu, and Mohammed Bennamoun. Ontology learning and knowledge discovery using the web: challenges and recent advances. Information Science Reference Hershey, PA, 2011.
- [17] A. Krizhevsky, I. Sutskever et G. E. Hinton. ImageNet Classification with Deep Convolutional Neural Networks, Advances in neural Processing Systems de traitement. 2012.
- [18] K nearest neighbors, <https://mrmint.fr/introduction-k-nearest-neighbors>, seen in 17/03/2020.
- [19] K-means machine learning, <https://web.cse.msu.edu/~cse802/notes/ConstrainedKmeans.pdf>, seen in 18/03/2020.