



N° d'ordre :
N° de série :

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique



UNIVERSITE ECHAHID HAMMA LAKHDAR - EL OUED FACULTÉ DES
SCIENCES EXACTES Département D'Informatique

Mémoire de Fin D'étude
Présenté pour l'obtention du Diplôme de

MASTER ACADEMIQUE

Domaine : **Mathématique et Informatique**
Filière : **Informatique**
Spécialité : **Systèmes Distribués et Intelligence Artificielle**

Présenté par :

- **Rimi Chouaib**
- **Meniai Bachir**

Thème

**La Détection des objets par dégradation
dans les images**

Soutenue le : 07/10/2020 devant le jury:

Dr. Hamoud Meriem	MCA	Président
Dr. Bellila Khaoula	MAA	Rapporteur
Dr. Laouid Abdelkader	MAA	Encadreur

Année Universitaire: 2019/2020

REMERCIEMENT

Nous tenons remercier avant tout DIEU qui nous a donné la force, la volonté , le courage, et la patience de pouvoir réaliser ce modeste travail.

On tient aussi remercier Dr Laouid Abdelkader , notre encadreur qui a su orienter notre travail, aussi pour sa disponibilité nous prodiguer des conseils, pour sa confiance et pour sa précieuse aide, on le remercie du fond du cœur.

Nous remercions aussi notre encadreur aux ses efforts et ses aides, en lui espérant une meilleure vie pleine de joie, de bonheur et de succès en différentts domaines nous lui dédions ce modeste travail

Nous sinc res remerciements sont adress s tous les membres du jury qui ont accept de juger notre modeste travail, et tous ceux qui, sans avoir t'impliqués directement dans ce travail, ont toujours t d'un grand support : nos enseignants, nos familles, nos amis, nos collègues chacun son nom.

Dédicace

À mes très chers parents

*pour tous les soins et le suivi dont ils font preuve depuis ma naissance et au long de mes tudes
pour leurs soutien et surtout leurs conseils et amour*

À mes très chères sœurs

À mes très chers frères

À ma grande famille

À tous mes amis

À tous ceux qui m'aiment et que j'aime

Rimi

Dédicace

A mes très chers parents

*pour tous les soins et le suivi dont ils font preuve depuis ma naissance et au long de mes tudes
pour leurs soutien et surtout leurs conseils et amour*

A mes très chères sœurs

A mes très chers frères

A ma grande famille

A tous mes amis

A tous ceux qui m'aiment et que j'aime

Meniai

Résumé

Aujourd'hui, les progrès du multimédia et de l'informatique visuelle dans la technologie numérique apportent un changement potentiel dans la vie humaine. De nombreuses applications exploitent les images capturées à partir d'entités autonomes comme sources de données pour plusieurs objectifs. En effet, ces images capturées doivent être interprétées pour extraire leur environnement externe.

Les chercheurs de ce domaine devront relever des défis comme la détection et l'interprétation du contexte des images.

Cet memoire est de proposer une technique efficace de détection des objets d'une image donnée en fonction de la divergence des couleurs Les résultats montrent clairement la précision et la vitesse de calcul de la technique proposée par rapport aux autres méthodes.

Mots clés : Analyse des données, Détection des objets, Traitement des images.

ملخص

في الوقت الحاضر، تُحدث تطورات الوسائط المتعددة والحوسبة المرئية في التكنولوجيا الرقمية تغييراً محتملاً في حياة الإنسان ، حيث تستغل العديد من تطبيقات الصور الملتقطة من الكيانات المستقلة كمصادر بيانات لعدة أهداف. في الواقع ، يجب تفسير هذه الصور الملتقطة من أجل استخراج بيئتها الخارجية.

سيواجه الباحثون في هذا المجال بعض التحديات مثل كيفية اكتشاف وتفسير سياق الصور ، ويهدف هذا البحث إلى اقتراح تقنية فعالة للكشف عن الأشياء لصورة معينة بناءً على اختلاف اللون. تظهر النتائج بوضوح دقة و سرعة حساب التقنية المقترحة مقارنة بالطرق الأخرى.

الكلمات المفتاحية: معالجة الصور، تحديد الكائنات، معالجة البيانات.

Abstract

Nowadays, multimedia and visual computing advances in digital technology make a potential change in human life. Many applications exploit the captured images from autonomous entities as data sources for several goals. In fact, these captured images need to be interpreted in order to extract their external environment.

The researchers of this domain will meet some challenges such as how to detect and interpret the images context.

This paper is to propose an efficient technique that detects objects of a given image based on the color divergence. The results clearly show the accuracy and the computation speed of the proposed technique compared with other methods.

Key words : Image Processing, Object Detection, Data Analysis.

TABLE DES MATIÈRES

Table des matières	i
Table des figures	iv
Liste des tableaux	vi
Introduction générale	1
1 Généralité sur l'image numérique et l'intelligence artificielle	3
1.1 Introduction	4
1.2 Concepts de base	4
1.2.1 L'image numérique	4
1.2.2 Les types d'image numérique	4
1.2.3 Caractéristiques de l'image	6
1.2.4 Les différents formats d'images	7
1.3 L'intelligence artificielle (IA)	8
1.3.1 La vision par ordinateur	9
1.3.2 Le traitement d'image	10
1.3.3 Différents composants d'un système de détection automatique	10
1.3.4 La vision humaine	11
1.4 Conclusion	11
2 Généralité sur la détection des objets saillants et l'état de l'art	12
2.1 Introduction	13
2.2 Segmentation et extraction de caractéristiques	13
2.2.1 Segmentation d'images	13

2.2.2	Extraction de primitives	14
2.3	Représentation d'objet	16
2.4	Densités de probabilités d'apparence d'objet	17
2.5	Les caractéristiques visuelles d'objet	17
2.5.1	Contours (Bords)	17
2.5.2	Flux optique	17
2.5.3	Texture	18
2.5.4	Couleur	18
2.6	Détection d'objets	18
2.6.1	Les détections basées sur la différence inter-images	18
2.6.2	Les détections basées sur la modélisation du fond	19
2.6.3	Les détections extrayant des couches de mouvement	19
2.6.4	Les détections utilisant la notion de cohérence	19
2.7	Catégories des détecteurs	20
2.7.1	Méthodes basées sur les contours	20
2.7.2	Méthodes basées sur l'intensité	20
2.7.3	Les détecteurs utilisant des modèles paramétriques	20
2.8	Les travaux connexes	20
2.8.1	Technique de contour actif	21
2.8.2	Détection robuste d'objets couleur	23
2.8.3	Soustraction de l'arrière-plan	23
2.8.4	Détection d'objets RVB-D inspirée de l'incertitude	24
2.9	Conclusion	24
3	L'approche proposée	25
3.1	Introduction	26
3.2	Contexte et contribution	26
3.3	L'hypothèse dans un contexte philosophique	26
3.4	La technique de divergence de couleur proposée (CDT) pour détecter les objets saillants	27
3.4.1	Classification de cube RVB	28
3.4.2	Segmentation et détection d'objets saillants	30
3.5	Conclusion	31
4	implémentation et analyse	32
4.1	Introduction	33
4.2	Environnement de développement	33
4.2.1	Environnement logiciel	33

4.2.2 Environnement matériel	33
4.3 L'évaluation des performances	33
4.4 Les tests et résultats	34
4.5 Discussion	34
4.6 Conclusion	37
Conclusion générale	38
Bibliographie	39

TABLE DES FIGURES

1.1	Représentation des images numériques.	4
1.2	Représentation numérique d'une image noir et blanc.	5
1.3	Représentation numérique d'une image niveau de gris.	5
1.4	Représentation numérique d'une image en couleur.	6
1.5	Le relation entre IA et ML et DL.	9
1.6	Schéma d'un système de traitement d'images [23]	10
1.7	Schéma général d'une chaîne de détection d'objet en image.	11
2.1	Différentes types de représentation d'objet.	16
2.2	Evolution du contour actif.	21
2.3	Schéma globale de technique du contour actif.	22
2.4	Initialisation automatique du snake dans la deuxième image.	23
3.1	Un cube 3D représentant géométrique le RVB espace de couleur.	28
3.2	Un cube 3D régulier et irrégulier géométrique représentant les écarts types des pixels.	30
4.1	L'image originale pour détecter les objets rouges saillants.	35
4.2	L'image de sortie après détection des objets rouges saillants.	35
4.3	L'image d'entrée pour la détection d'objets saillants en noir et blanc.	35
4.4	L'image de sortie après la détection d'objets saillants en noir et blanc.	35
4.5	L'image d'entrée pour la détection d'objets saillants en image haute résolution(1200*1859).	36
4.6	L'image de sortie après la détection d'objets saillants en image haute résolution(1200*1859).	36
4.7	L'image d'entrée pour la détection d'objets saillants dans un environnement complexe.	36
4.8	L'image de sortie après la détection d'objets saillants dans un environnement complexe.	36
4.9	L'image d'entrée pour la détection d'objets saillants en image avec des couleurs imbriquées.	36
4.10	L'image de sortie après la détection de tous objets saillants (non- sélective).	36

4.11 L'image de sortie après détection des objets rouges saillants.	37
---	----

LISTE DES TABLEAUX

4.1		34
-----	-------	----

INTRODUCTION GÉNÉRALE

Récemment, il y a eu beaucoup d'améliorations dans le secteur de l'intelligence artificielle grâce au Deep Learning et au traitement d'image. Il est maintenant possible de reconnaître des images ou même de détecter des objets dans une image.

La détection d'objets est une technologie informatique liée à la vision par ordinateur et au traitement d'image qui traite de la détection des instances d'objets sémantiques d'une certaine classe (tels que les humains, les bâtiments ou les voitures) dans les images et vidéos numériques. Les domaines bien documentés de la détection d'objets comprennent la détection des visages et la détection des piétons. La détection d'objets a des applications dans de nombreux domaines de la vision par ordinateur, y compris la récupération d'images et la vidéosurveillance. Le problème posé est comment détecter l'objet saillant et déterminer quel groupe connu des objets peut être présent et fournir des informations sur leur position dans l'image avec précision ?

Le contenu de ce mémoire est organisé comme suit :

Dans **le premier chapitre**, nous décrirons des généralités sur les images et l'intelligence artificielle, où nous définissons chacun des l'image numérique, les types d'image numérique et leur caractéristiques c'est en termes de l'image numérique. En ce qui concerne à l'intelligence artificielle : La vision par ordinateur, le traitement d'image ...etc.

Dans **le deuxième chapitre**, nous parlerons des bases de la détection des objets saillants et l'état de l'art et nous avons parlé de l'élément le plus important en détection qui est la segmentation et extraction de caractéristiques, puis des certaines méthodes modernes pour la détection des objets et expliquerons les principes de leur fonctionnement.

Dans **le troisième chapitre**, nous allons proposer une autre nouvelle approche, en utilise une nouvelle méthodologie de regroupement qui classifera la distribution pixels d'image dans un groupe, afin de montrer l'objet mis en évidence sur cette image.

Dans **le dernier chapitre**, nous entamons maintenant sur ce dernier chapitre la partie de la

réalisation et de l'analyse, nous présentons notre application réalisée, et les différents outils nécessaires utilisés dans son développement. Nous avons aussi terminé par un ensemble d'analyses et tests pour les résultats expérimentaux qui montrent la bonne détection des objets par la dégradation de couleurs dans les images.

Puis on va terminer par une conclusion générale et quelques perspectives pouvant aider nous vérifierons ses performances on exécutant sur la reconnaissance des objets à l'aide de différentes méthodes d'apprentissage en profondeur du système dans le futur.

CHAPITRE 1

GÉNÉRALITÉ SUR L'IMAGE NUMÉRIQUE ET L'INTELLIGENCE ARTIFICIELLE

1.1 Introduction

Le XX^e siècle est celui de l'image numérique, Le traitement d'images est un domaine récente afin d'offrir aux utilisateurs, des outils de manipulation de ces données digitales issues du monde réel

Dans ce chapitre, nous abordons quelques concepts de base pour comprendre les techniques de traitement d'image. Ensuite, nous définissons chacun des l'image numérique, les types d'image numérique et leur caractéristiques ,aussi nous donnerons un aperçu de l'intelligence artificielle. Nous citons quelques objectifs dans le domaine du traitement d'image : interpréter différentes informations d'image traitées et détection des objets saillants de l'image, c'est-à-dire séparer les objets saillants du fond de l'image .

1.2 Concepts de base

Dans cette section, on va présenter certains concepts de base concernant l'image et ses caractéristiques

1.2.1 L'image numérique

L'image numérique est un code qui est acquis, stocké et interprété.Cette abstraction de l'image en un code apporte un grand nombre d'avantages., cela correspond au nombre de pixels qui composent l'image en hauteur (axe vertical) et en largeur (axe horizontal) : 200 pixels par 450 pixels par exemple, abrégé en « 200 450 ».Une image numérique est une matrice de pixels repérés par leur coordonnées (x,y) Chaque pixel représente un point de l'image (dans un espace colorimétrique prédéfini).[1].

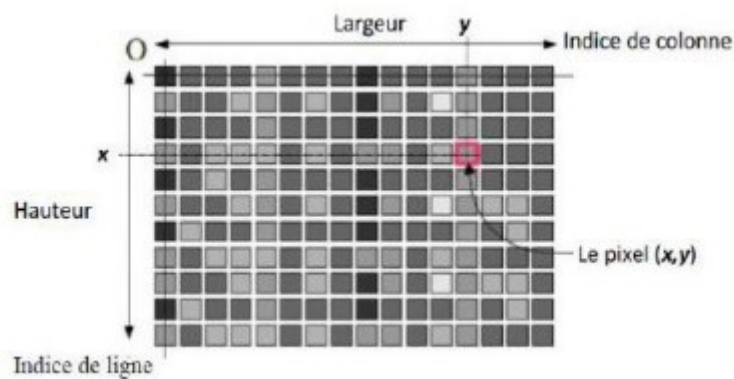


FIGURE 1.1 – Représentation des images numériques.

1.2.2 Les types d'image numérique

Nous avons vu qu'une image numérique est un tableau de points représentant les couleurs (pixels). Il existe plusieurs formats de codage de l'information de couleur des pixels .

Image binaire

Un seul bit suffit pour coder l'information, par exemple 0 pour le noir et 1 pour le blanc.

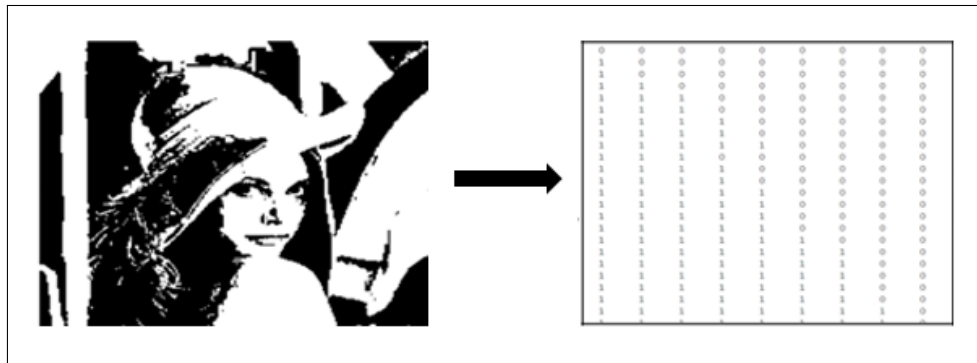


FIGURE 1.2 – Représentation numérique d'une image noir et blanc.

Image niveau de gris

On utilise en générale 8 bits, ce qui donne 256 nuances de gris possibles pour le pixel, de 0 (noir) à 255 (blanc).

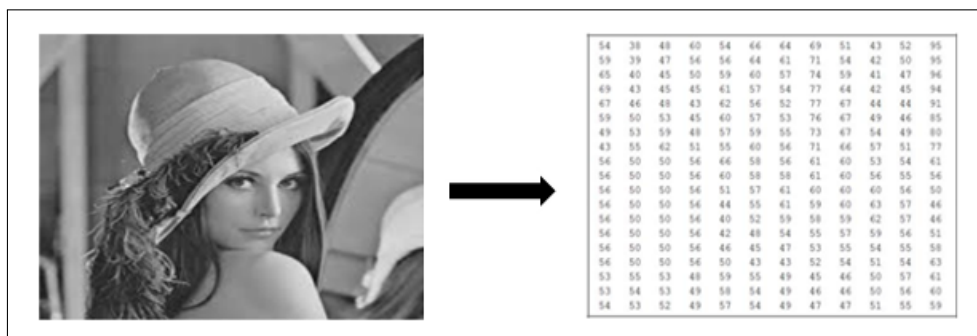


FIGURE 1.3 – Représentation numérique d'une image niveau de gris.

Image couleur (RGB)

La couleur de chaque pixel est définie par 3 composantes : Rouge, Vert et Bleu (système RVB où RGB en anglais).

L'intensité de chaque composante est codée sur 8 bits, donc chaque composante a une valeur comprise entre 0 (absence de couleur) et 255 (intensité maximale de la couleur).

Ainsi la couleur d'un pixel nécessite 24 bits (3 octets) pour être codée.

La couleur du pixel est obtenue par synthèse additive (RGB), en particulier :

- Si les 3 composantes sont à 0, on obtient du noir.
- Si les 3 composantes sont identiques on obtient une nuance de gris.
- Si les 3 composantes sont à 255, on obtient le blanc.

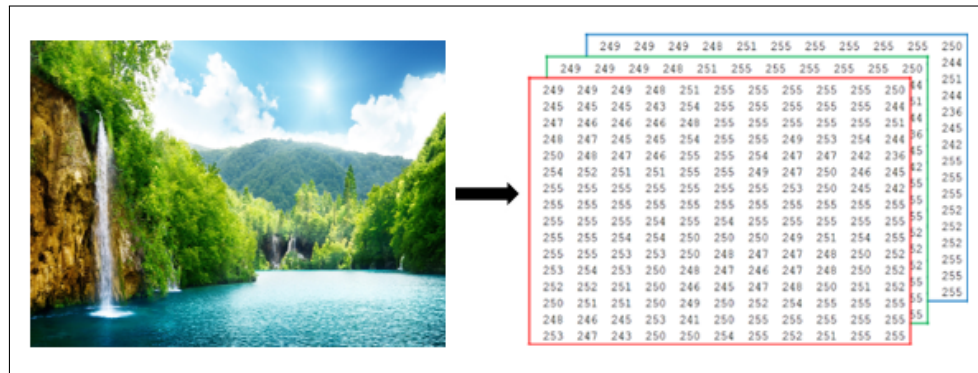


FIGURE 1.4 – Représentation numérique d'une image en couleur.

1.2.3 Caractéristiques de l'image

L'image est un ensemble structuré d'informations caractérisé par les paramètres suivants :

Pixel

Le pixel est une unité de mesure de la définition d'une image numérique. Abrégée en px , elle est souvent présentée comme un petit carré de couleur. Lorsqu'on zoome sur une image numérique matricielle (les images vectorielles ne sont pas basées sur les pixels), on aperçoit en effet que celle-ci se compose d'une multitude de petits carrés colorés. Il s'agit des pixels. À noter qu'il n'existe pas de plus petit élément d'une surface d'affichage que le pixel [2].

Codage de la couleur

Pour représenter informatiquement une image, il suffit de créer un tableau de pixels dont chaque case contient une valeur. La valeur stockée dans une case est codée sur un certain nombre de bits déterminant la couleur ou l'intensité du pixel, on l'appelle profondeur de codage (parfois profondeur de couleur). Il existe plusieurs standards de codage de la profondeur : bitmap noir et blanc, bitmap 16 couleurs ou 16 niveaux de gris, bitmap 256 couleurs ou 256 niveaux de gris, palette de couleurs (color-map) et Couleurs vraies (true color) [3].

Luminance

C'est le degré de luminosité des points de l'image. Elle est définie aussi comme étant le quotient de l'intensité lumineuse d'une surface par l'aire apparente de cette surface. Pour un observateur lointain, le mot luminance est substitué au mot brillance, qui correspond à l'éclat d'un objet [4].

Contraste

C'est l'opposition marqué entre deux régions d'une image, plus précisément entre les régions sombres et les régions claires de cette image [5].

Bruit

Un bruit (parasite) dans une image est considéré comme un phénomène de brusque variation de l'intensité d'un pixel par rapport à ses voisins, il provient de l'éclairage des dispositifs optiques et électroniques du capteur [6].

Format

Un format d'image est la représentation informatique de l'image, associée à des informations sur la façon dont l'image est codé et fournissant éventuellement des indications sur la manière de la décoder et la manipuler [6]. Parmi les formats d'image le plus couramment utilisé et certainement celle la plus répandue image bitmap, elle est tout simplement constituée d'une grille de milliers de pixels représentant les points de couleur successifs de l'image. On pourrait comparer cette technique à celle du pointillisme. Chaque pixel est alors un tout petit carré ayant une place bien définie avec sa propre couleur et l'image devient alors un quadrillage de pixels qui placés les uns à côté des autres, restituent le rendu visuel [7].

1.2.4 Les différents formats d'images

On peut classer les images en deux formats :

Format vectorielle

Les données d'une image vectorielle sont représentées par des formes géométriques simples qui sont décrites d'un point de vue mathématique. Par exemple, un cercle est décrit par une information du type (cercle, position du centre, rayon). Ces images sont essentiellement utilisées pour réaliser des schémas ou des plans.

Format matricielle

Une image matricielle est constitué d'un tableau de points ou pixels. Plus la densité des points sont élevée, plus le nombre d'informations est grand et plus la résolution de l'image est élevée. Corrélativement la place occupée en mémoire et la durée de traitement seront d'autant plus grandes. Les images vues sur un écran de télévision ou une photographie sont des images matricielles. On obtient également des images matricielles à l'aide d'un appareil photo numérique, d'un scanner ou d'une caméra vidéo numérique. Parmi ces formats on peut citer [7] :

- **BMP (BitMap)**

C'est le format par défaut du logiciel Windows. C'est un format matriciel. Les images ne sont pas compressées. Son logiciel d'origine.

matriciel n'est pas très différent du EPS vectoriel. En fait seules les données contenues dans le fichier sont différentes. Ainsi un logiciel de retouche de photos tel que Photoshop permet l'importation, la modification et l'exportation de fichiers en format EPS.

- **GIF (Graphical Interchange Format)**

C'est un format qui a ouvert la voie à l'image sur le World Wide Web. C'est un format de compression qui n'accepte que les images en couleurs indexés codé sur 8 bits, C'est un format qui perd beaucoup de son marché suite à une bataille juridique concernant les droits d'utilisation sur Internet.

- **JPEG (Joint Photographique Experts Group)**

JPEG sont des images de 24 bits. C'est-à dire qu'elles peuvent afficher un spectre de 16 millions de couleurs. C'est la meilleure qualité d'images disponible.

- **Le format PNG (Portable Network Graphics, ou format Ping)**

C'est un format de fichier graphique bitmap. Il a été mis au point en 1995 afin de fournir une alternative libre au format GIF. Le format PNG permet de stocker des images en niveaux de gris(jusqu'à 16 bits par pixel de profondeur de codage), en couleurs réelles (True color, jusqu'à 48 bits par pixel de profondeur de codage), ainsi que des images indexées, faisant usage d'une palette de 256 couleurs.

- **Le format TIFF**

Le format TIFF ou TIF (Tagged Image File Format) est un format de fichier graphique bitmap. Il a été mis au point en 1987. Le format TIFF est un ancien format graphique, permettant de stocker des images bitmap de tailles importantes (plus de 4 Go compressées), sans perte de qualité et indépendamment des plateformes ou des périphériques utilisés.

Le format TIFF permet de stocker des images en noir et blanc (niveaux de gris), en couleurs réelles (jusqu'à 32 bits par pixels) ainsi que des images indexées, utilisant une palette de couleurs. De plus le format TIFF permet l'usage de plusieurs espaces de couleurs.

1.3 L'intelligence artificielle (IA)

Largement définition de l'intelligence artificielle est concerné par un comportement intelligent dans les artefacts. L'IA est un domaine d'informatique qui vise à construire des machines intelligentes capables d'exécuter des tâches qui requièrent généralement l'intelligence humaine. L'IA est une science interdisciplinaire aux multiples approches .Ils analysent et traitent des images et des vidéos, agissent en conséquence et reconnaissent des objets [8].

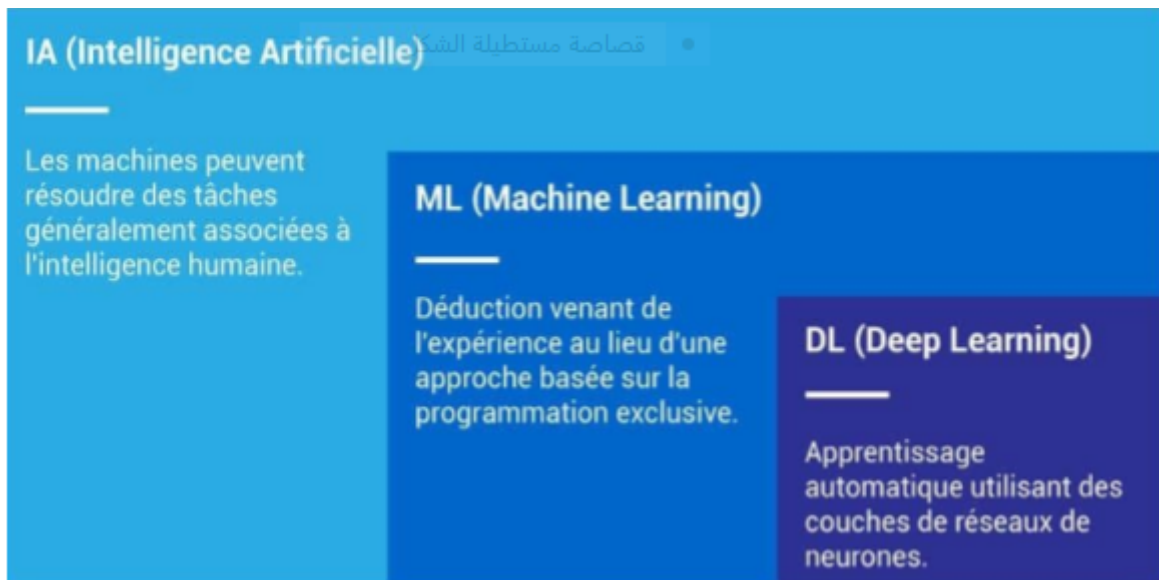


FIGURE 1.5 – Le relation entre IA et ML et DL.

1.3.1 La vision par ordinateur

La vision par ordinateur s'agit d'une sous-catégorie d'intelligence artificielle. Il est les différentes techniques permettant aux ordinateur de voir et de comprendre le contenu d'images. La vision par ordinateur peut être mieux comprise en considérant différents types d'applications qui sont utilisées dans de nombreux et divers types d'environnements [9], parmi ses applications traitement d'image.

Les domaines d'application de la vision par ordinateur sont :

- Imagerie médicale
- Manufacturing
- Sécurité
- Recherche scientifique
- Informatique de gestion

On peut citer quelques applications exemplaires pour illustrer les enjeux des travaux de recherche sur ce problème

- ✓ La video-surveillance : la reconnaissance et la détection automatique des personnes, des visages et des objets d'intérêt sur la très grande quantité d'images video utilisées pour la surveillance.
- ✓ La recherche dans des bases d'images numériques : les systèmes actuels effectuent en général des requêtes à partir des mots clés associés aux images. On voudrait exploiter le contenu visuel des bases d'images sans avoir besoin de les indexer textuellement.
- ✓ La navigation des robots mobiles et des véhicules : on a besoin de reconnaître les obstacles, les objets divers dans l'environnement ou les autres véhicules pour la navigation automatique.

La différence entre classification, détection et segmentation

La classification, c'est identifier à quelle catégorie prédéterminée appartient une image. C'est à dire identifier le sujet principal ou si l'image a une composition particulière.

La détection utilise la classification et la localisation. La localisation détermine combien il y a d'objet connus (entrés dans la base de données) sur l'image. Donc, la détection, c'est rechercher des objets ou caractéristiques dans des images et déterminer leurs emplacements.

La segmentation est la découpe de différents éléments qui composent une image. Elle transforme les pixels dans des polygones, permettant ainsi une description plus condensée de l'image [10].

1.3.2 Le traitement d'image

- **Système de traitement d'images :** Un système de traitement numérique d'images est composé de [11] :

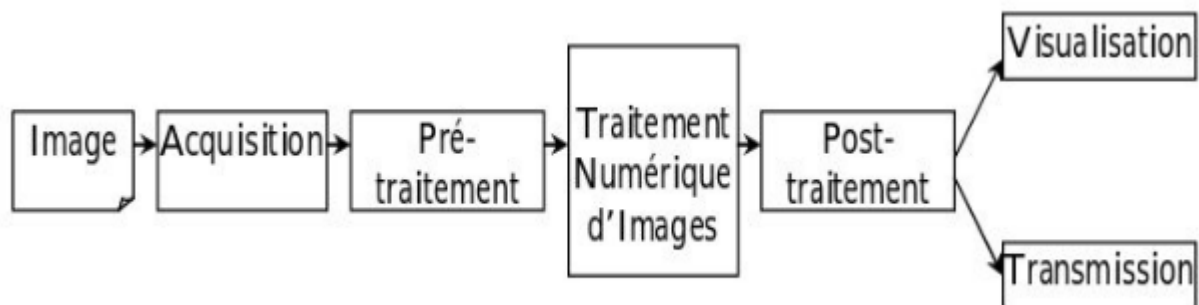


FIGURE 1.6 – Schéma d'un système de traitement d'images [23]

1.3.3 Différents composants d'un système de détection automatique

De manière globale, un système de détection automatique des objets se décompose en une étape de localisation suivie d'une prise de décision. La localisation peut selon les cas être composée d'un prétraitement (observation visuelle ou détection rapide) suivie d'une segmentation. La prise de décision se compose quant à elle d'une étape d'extraction de paramètres, mesures suivie d'une étape d'analyse (Identification, classification). Certains de ces éléments peuvent apparaître de manière plus ou moins implicite. La figure 1.7 illustre la décomposition de haut niveau des processus de détection automatique d'objet. Dans certains cas, un prétraitement des données permettant de mettre en évidence les signes recherchés peut être utilisé. Néanmoins, cette étape montre en général, rapidement ses limitations dans la mesure où, pour vraiment mettre en évidence un motif, il faut être capable de le détecter, or c'est le but de cette étape de prétraitement [12].

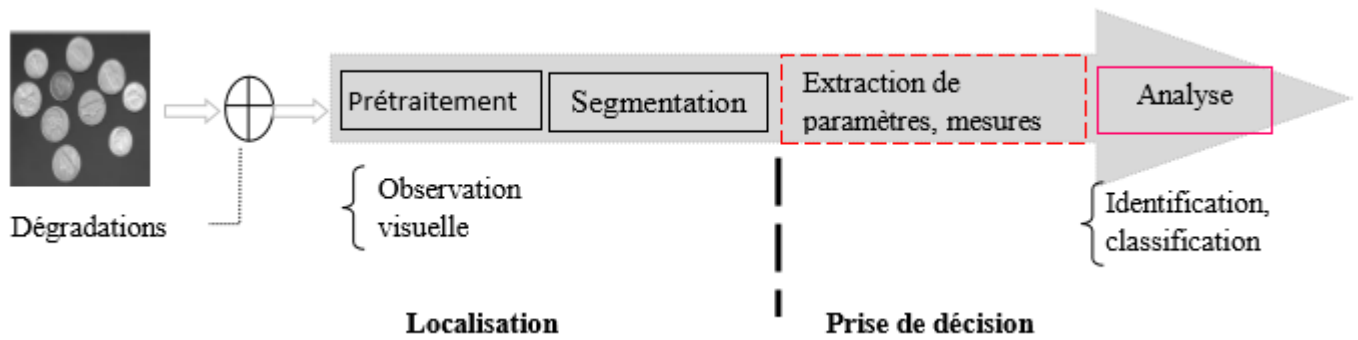


FIGURE 1.7 – Schéma général d'une chaîne de détection d'objet en image.

1.3.4 La vision humaine

L'œil humain est capable de reconnaître rapidement une multitude d'objets différents avec des variations de point de vue, d'illumination et de forme, même dans les contextes les plus inhabituels. On reconnaît les catégories d'objets aussi bien que les instances d'objets uniques, et cela avec une capacité de généralisation sur plusieurs échelles et sous plusieurs sortes de transformations. On peut ajouter à cela l'apprentissage rapide d'objets qu'on voit pour la première fois.

On est loin de connaître le modèle complet de la perception humaine qui réunit des mécanismes cognitifs complexes et de l'imiter pour développer un système automatique capable d'approcher la performance humaine. Pourtant, Les expériences psycho-cognitives sur les étapes primales (la détection rapide sans utiliser les processus cognitifs de haut niveau) de la vision humaine et sur la vision des pigeons ont révélé des indices qui peuvent être exploités par les systèmes automatiques. Il nous semble pertinent de citer deux travaux expérimentaux importants.

1.4 Conclusion

De nombreuses idées et technique pertinentes pour mécaniser l'intelligence ont été développées par de chercheurs en IA. L'analyse d'images numériques est un aspect qui peut traiter par l'IA. Dans ce chapitre, Nous avons introduit les notions de base qui servent de fondement à la compréhension de différentes techniques de traitement d'images.

Dans le chapitre suivant nous exposerons quelques concepts sur la détection des objets dans les images et on va présenter un état de l'art.

CHAPITRE 2

GÉNÉRALITÉ SUR LA DÉTECTION DES
OBJETS SAILLANTS ET L'ÉTAT DE L'ART

2.1 Introduction

En vision par ordinateur, La détection d'objets fait référence à une méthode qui permet de découvrir l'existence d'une instance (reconnaissance d'un objet) ou d'une classe d'objets sous forme numérique. Une attention particulière est accordée à la détection des visages et à la détection des personnes. Ces méthodes impliquent souvent un apprentissage supervisé et ont des applications dans de multiples domaines, Comme la recherche sur le contenu d'image ou la vidéo surveillance.

2.2 Segmentation et extraction de caractéristiques

Le but de l'analyse de l'image est d'extraire les informations distinctes contenues dans l'image. Le résultat de cette analyse est souvent appelé une description structurelle. Fondamentalement, l'analyse d'image implique une segmentation où nous allons essayer de relier une étiquette à chaque pixel de l'image en fonction des informations fournies (niveaux de gris), de la distribution spatiale au centre de l'image, des modèles simples (souvent des modèles géométriques) [13].

2.2.1 Segmentation d'images

La segmentation est une étape essentielle du traitement d'image dans la mesure où l'interprétation qui sera faite sur ces images est requise.

Par conséquent, plusieurs algorithmes ont été proposés et sont basés sur différentes approches : **contour, région, texture** [14].

Segmentation par approche régions

La segmentation d'une image par rapport à la norme d'homogénéité H (par exemple, "tous les pixels sont de même niveau de gris" ou "le contraste du niveau de gris ne dépasse pas n niveaux") peut être comme suit :

La segmentation d'une image I en regard du critère H est une partition de l'image I en n régions homogènes $X_1, \dots, X_n$ telles que :

- $\bigcup_{i=1}^n X_i = I$
- Pour tout i , X_i est connexe.
- Pour tout i , $H[X_i]$ est connexe.
- Pour tout couple (X_i, X_j) des régions voisines, $H[X_i, X_j]$ est faux.

Cette définition fait deux observations très importantes. Tout d'abord, la segmentation dépend du critère utilisé. Par conséquent, le choix de la norme est essentiel.

Ensuite, la décomposition obtenue n'est pas unique. Pour un certain critère, il existe plusieurs solutions.

Segmentation par approche frontière

Cette approche cherche à exploiter le fait qu'il existe une transition détectable entre deux régions liées. Les méthodes plus anciennes utilisent des agents de traitement d'image, tels que le filtre **Canny**, pour distinguer les pixels qui apparaissent au contour. Souvent, créer le résultat est difficile. On peut également utiliser des modèles déformables utilisant des courbes paramétriques (courbe de Bézier, Spline, ...etc.) ou des polygones (par exemple l'algorithme à bulles). Pour lors du démarrage du processus, nous recherchons les points observés dans l'image, tels que les points à l'intersection d'au moins trois parties. Ces points sont appelés **graines**.

Le principal avantage des méthodes de segmentation selon l'approche frontières est de réduire le nombre d'opérations nécessaires si l'opération est répétée sur une série d'images peu différentes les unes des autres (dans le cas des images vidéo notamment). En effet, une fois que vous avez trouvé les diagrammes de zones dans la première image, appliquer le motif déformable à l'image suivante est plus efficace que de tout recalculer, si la différence entre les images est faible.

Segmentation par classification ou seuillage

Les techniques de classification ont été beaucoup utilisées dans le domaine de la segmentation d'images. Plusieurs algorithmes ont été développés depuis. On part ici d'une relation que chaque pixel entretient individuellement avec les informations calculées sur l'ensemble de l'image, par exemple des niveaux de gris moyens pour tous les pixels, ou moyenne, ce qui permet de générer n classes d'intensités. Lorsque des catégories sont définies en choisissant un seuil, on parle d'un seuil. Ils forment des pixels appartenant à la même classe et reliés par des régions.

Le seuillage a pour objectif de segmenter une image en plusieurs classes en n'utilisant que l'histogramme. On suppose donc que l'information associée à l'image permet à elle seule la segmentation, qu'une classe est caractérisée par sa distribution de niveaux de gris.

2.2.2 Extraction de primitives

Les méthodes d'extraction de primitives graphiques basés sur une approche hiérarchique pour comprendre la description d'une forme sont tels que cette description est une représentation du contenu de l'image.

Les indices visuels

Les repères visuels sont des objets d'image extraits qui contiennent brièvement des informations importantes pour leur analyse. Habituellement, ces indices sont soit des contours, soit des régions de l'image.

Les contours et les régions sont complémentaires parce que les régions sont qualitativement définies

comme des régions d'une image homogène au sens d'une norme spécifique, tandis que les contours sont des régions de transitions entre des régions homogènes [15].

Les caractéristiques requises par les repères visuels dépendent de l'application et constituent généralement un sous-ensemble de la liste ci-dessous :

- ✓ **Compacts** : la représentation de l'image doit être aussi concise que possible pour réduire la complexité des algorithmes ultérieurs.
- ✓ **Intrinsèques** : les indices visuels doivent correspondre à la projection dans l'image d'objets physiques ; en particulier, ils doivent être invariants par changement de point de vue.
- ✓ **Robustes** : la représentation doit être peu sensible aux petites variations d'intensité dans l'image provoquées par des bruits divers tels que ceux liés à l'acquisition, la digitalisation.
- ✓ **Discriminants** : les indices visuels doivent posséder des propriétés qui permettent de les discriminer, afin de faciliter la mise en correspondance entre deux descriptions.
- ✓ **Précis** : la position des indices visuels doit pouvoir se calculer avec précision car la qualité de la localisation des objets physiques en dépend.
- ✓ **Denses** : la densité des indices visuels doit être suffisante pour représenter tous les objets intéressants de la scène.

Méthodes basées primitives

Les caractéristiques de l'image à extraire pour détecter les indicateurs de bas niveau, pour suivre l'objet sont nombreuses, il existe des descripteurs connus, qui partagent les deux caractéristiques de nombreuses méthodes primitives, ces dernières sont discutées en termes de propriétés fixes, de reconstructions, de distorsions prédites. La performance des résultats obtenus prend en compte les formats des propriétés extraites [16]. Parmi les méthodes basées sur le primitivisme, on peut citer :

✓ Attributs géométriques

L'objectif est de découvrir des descripteurs associés à des points spécifiques et de décrire un objet à travers un ensemble de caractéristiques géométriques (points, segments, courbes paramétriques, arêtes, contours) ou zones d'image. Ces méthodes ont un avantage bonne durabilité de la dissimulation car les liens qui ne peuvent pas être créés à des points spécifiques de l'objet caché dans l'image n'annulent pas le traçage à tous les points.

✓ Caractéristiques de l'image

Ces méthodes, parfois sans modèle, dépendent de la région selon la couleur, la texture, les priorités et les contours, il s'agit de processus de correspondance d'une image à l'autre. Les primitives sont extraits pour une correspondance en séquence, via la distance euclidienne ou une approche basée sur la corrélation. La voie peut ensuite être évaluée en regroupant ces primitives tout au long de la séquence. Cette approche ne définit que quelques points sur l'objet suivi et non l'objet entier, ce que font les approches de contour et de surface.

2.3 Représentation d'objet

Il existe différentes représentations de forme de l'objet utilisé pour le suivi :

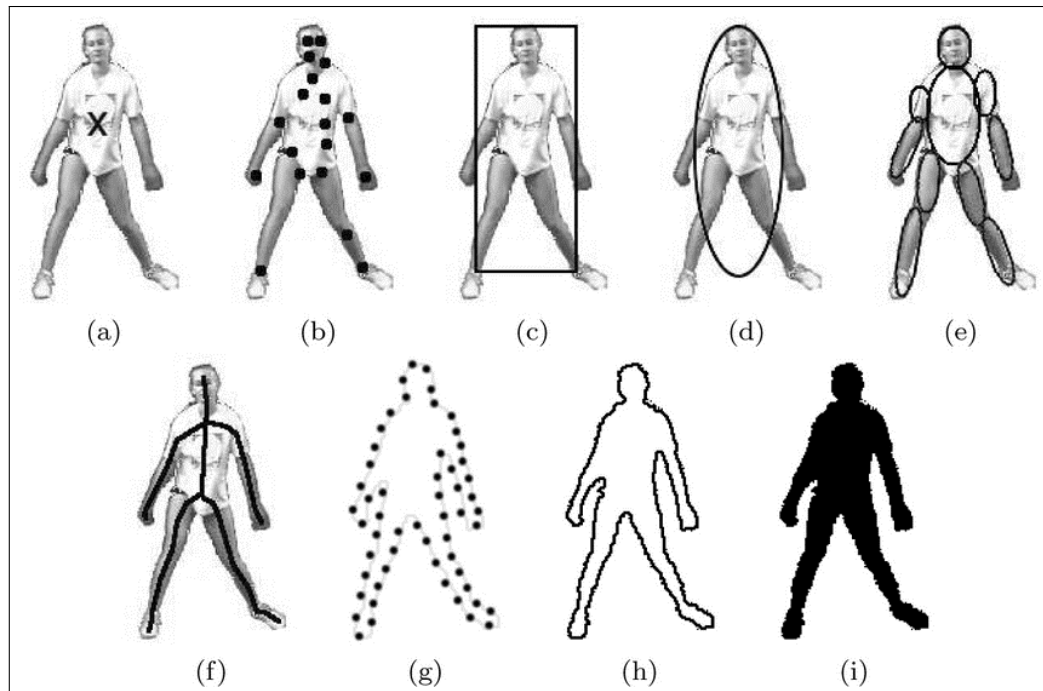


FIGURE 2.1 – Différents types de représentation d'objet.

✓ Point

L'objet est représenté par un point, c'est le centre de gravité, ou par un ensemble de points (a).

Un objet est représenté par un ensemble de points (b). En général, la représentation ponctuelle convient au suivi d'objets qui occupent de petites zones de l'image [17].

✓ Formes géométriques primitives (Noyaux)

La Forme de l'objet est représentée par un rectangle, ellipse (c) et (d) dans l'objet à suivre est représenté par un histogramme pondéré calculé sur une région elliptique contenant l'objet. Et dans l'approche la plus évidente pour un objet repose sur l'utilisation d'un gabarit (en anglais template).

Bien que des formes géométriques primitives soient les plus appropriées pour représenter des objets rigides simples, ils sont également utilisés pour le suivi d'objets non rigides [18].

✓ Silhouette de l'objet et le contour

La représentation du contour définit la limite d'un objet (g), (h) et (i). La région à l'intérieur du contour est appelée la silhouette de l'objet.

Les représentations de Silhouette et contour sont adaptés pour le suivi des formes complexes non rigides.

Dans la représentation d'un tel objet par une silhouette permet de tenir compte précisément de la forme de l'objet. Le but des méthodes de suivi fondées sur l'utilisation de silhouettes est d'estimer la

silhouette des objets d'intérêt pour chaque image de la vidéo [18].

Cette approche est également connue sous le nom de segmentation d'images. La segmentation d'images est une opération de traitement d'images qui a pour but de rassembler des pixels entre eux suivant des critères prédéfinis.

✓ Formes articulées

Les objets articulés sont composés de parties du corps qui sont liés avec des joints. Par exemple le corps humain est un objet articulé avec torse, jambes, mains, tête et pieds reliés par des joints (e). On peut modéliser ces parties en utilisant des cylindres ou des ellipses, les rapports entre eux sont régis par des modèles de cinématique de mouvement, par exemple l'angle commun [19].

✓ Modèles squelettiques

Le squelette d'objet peut être extrait en appliquant la transformation de l'axe médiane à la silhouette d'objet. (f) Ce modèle est généralement utilisé pour la reconnaissance de l'objet et du suivi des modèles cinématiques complexes et la reconnaissance de mouvement [19].

2.4 Densités de probabilités d'apparence d'objet

La densité de probabilité décrit une couleur et une texture observées, fournit une estimation raisonnable de l'objet, décrit les occurrences communes de plusieurs traits d'objet et des estimations de la densité de probabilité d'un objet peuvent être paramétrées, telles que gaussienne, un mélange de gaussiens, ou non paramétriques comme l'histogramme. La densité de probabilité des caractéristiques d'apparence d'un objet (couleur, texture) peut être calculée en spécifiant des zones de l'image à l'aide d'une représentation d'une forme connue (rectangle, ellipse).

2.5 Les caractéristiques visuelles d'objet

2.5.1 Contours (Bords)

En règle générale, les limites d'un objet génèrent de grandes différences d'intensité d'image. La détection des bords est utilisée pour identifier ces changements. Une caractéristique importante des bords est qu'ils sont moins sensibles aux différences d'éclairement par rapport aux caractéristiques de couleur. Les algorithmes qui tracent les limites des bords des objets l'utilisent généralement pour représenter la fonction. En raison de sa simplicité et de sa précision, la méthode la plus populaire pour détecter les bords est le détecteur de Canny [20].

2.5.2 Flux optique

Le flux optique est un champ dense des vecteurs de déplacement qui définit la traduction de chaque pixel dans une région [20].

2.5.3 Texture

La texture est une composante essentielle de la perception humaine, ce qui en fait une propriété essentielle à considérer dans les algorithmes de détection ou de reconnaissance. La texture se refuse sur une zone entière et pas à un moment donné.

La texture présente de nombreuses qualités parmi lesquelles on peut citer son faible sensibilité aux changements d'illumination [20].

2.5.4 Couleur

La couleur est utilisée comme une caractéristique pour des représentations d'apparence basée sur un histogramme, alors que pour la représentation à base de contour, les bords des objets sont en général utilisés en tant que caractéristiques [17].

2.6 Détection d'objets

Dans le domaine de la vision par ordinateur, la découverte consiste à percevoir une scène statique ou dynamique, à connaître le changement de mouvement approprié, et ainsi à déduire tous les objets à travers des images. La découverte est une tâche difficile, et elle dépend généralement de la nature de la scène et de l'application correspondante [21].

Les méthodes de détection peuvent être classées en général en quatre catégories :

- ✓ Les détections basées sur la différence inter-images.
- ✓ Les détections basées sur la modélisation du fond.
- ✓ Les détections extrayant des couches de mouvement.
- ✓ Les détections utilisant la notion de cohérence.

2.6.1 Les détections basées sur la différence inter-images

La méthode de détection de mouvement la plus simple consiste à soustraire deux images consécutives. Elle est également appelée méthode de différence temporelle dans d'autres publications. Les pixels résultants proches de zéro sont traités comme des pixels d'arrière-plan.

Les méthodes qui reposent sur la différence entre les images s'adaptent bien aux environnements dynamiques mais laissent des «trous» dans les objets en mouvement. Par exemple, le décalage horaire permet de détecter des objets se déplaçant dans une plage séquence vidéo. L'amélioration de cette méthode est d'extraire trois images différentes au lieu de deux ce qui permet de déterminer si un pixel a été transféré ou non [22].

2.6.2 Les détections basées sur la modélisation du fond

Une autre classe de techniques, très courante dans la détection de mouvement, est basée sur la modélisation d'arrière-plan. Plusieurs conditions sont nécessaires pour utiliser ces technologies. Tout d'abord, la caméra doit rester immobile et masquer des parties d'un fichier l'arrière-plan doit être conservé en déplaçant temporairement les choses en minorité. De plus, il est préférable d'avoir un certain nombre d'images au début de la séquence sans déplacer les objets afin de pouvoir reconnaître correctement l'arrière-plan.

Ces techniques peuvent être soit prédictives, soit non prédictives, ces dernières faisant l'objet de recherches complémentaires. La modélisation en arrière-plan, tout comme les méthodes basées sur la détection entre images, permettent de déterminer si un pixel est en mouvement ou non [19].

- **Méthodes non prédictives**

Les techniques non prédictives de modélisation du fond se décomposent en deux étapes principales : l'apprentissage du fond [23].

- **Méthodes prédictives**

Les méthodes prédictives utilisent un modèle dynamique pour prédire la valeur d'un pixel à l'aide des observations précédentes. Elles emploient généralement un filtre de Kalman pour modéliser la dynamique de l'état d'un pixel [19].

2.6.3 Les détections extrayant des couches de mouvement

Le but de ces techniques est de fournir une description de la scène en termes d'objets et de leur mouvement dans le plan de l'image ou, en d'autres termes, en termes de couches de mouvement. Ainsi, la segmentation du mouvement peut être considérée comme une décomposition de Scène en différentes couches, chacune étant liée au modèle de mouvement [24].

En fait, la plupart des méthodes d'extraction de couche de mouvement sont basées sur le calcul d'un certain nombre de modèles paramétriques et supposent que le mouvement 2D d'un objet 3D montré dans l'image suit un modèle paramétrique.

L'avantage de ces méthodes est qu'elles ne nécessitent pas de caméra fixe ou ne donnent pas de résultats dans des moments plus anciens. Il est basé sur une estimation préliminaire du mouvement dans l'image.

2.6.4 Les détections utilisant la notion de cohérence

La dernière méthode consiste à définir un objet en mouvement comme une zone de mouvement cohérent. Dans ce contexte, un mouvement cohérent peut être défini comme un mouvement qui a de fortes chances de provenir d'un «but» classique (personne, véhicule).

Cette définition a été utilisée pour segmenter le mouvement en ajoutant l'hypothèse suivante : Un

objet se déplace en mouvement cohérent dans une direction presque constante pendant une courte période (pratiquement quelques images).

Les modèles de mouvement dérivés du flux optique sont utilisés comme base pour la détection ou le suivi. Le rôle du flux optique est de décrire le mouvement cohérent des points entre les images successives, et dans le sens directionnel, les mesures de flux optique cohérent de quelques pas de temps s'accumulent. Une estimation du déplacement de chaque pixel est obtenue avec une série d'images [25].

2.7 Catégories des détecteurs

On distingue trois catégories de détecteurs [26] :

- Les détecteurs s'appuyant sur une détection de contours.
- Les détecteurs exploitant les niveaux de gris.
- Les détecteurs utilisant des modèles paramétriques.

2.7.1 Méthodes basées sur les contours

On commence par appliquer un détecteur de contour dans l'image puis on regarde ces contours des points d'inflexion, ou des points où la courbure est un maximum localement ou les points d'intersection de plusieurs lignes (intersections).

2.7.2 Méthodes basées sur l'intensité

Ces méthodes reposent sur la dérivée des niveaux de gris pour identifier les points où les intensités varient considérablement dans une ou plusieurs directions. L'un des premiers détecteurs de cette classe était le Moravic. Lowe utilise les différences gaussiennes pour trouver les points clés.

2.7.3 Les détecteurs utilisant des modèles paramétriques

Ces méthodes s'appuient sur la déformation d'un modèle paramétrique de coin pour qu'il se rapproche des niveaux de gris au voisinage d'un coin. Ce détecteur est précis à condition d'avoir de bonnes valeurs initiales pour les paramètres du modèle. On peut citer par exemple, le détecteur de Baker.

2.8 Les travaux connexes

Durant ce travail du mémoire, on a eu l'occasion d'aborder plusieurs projets liés à notre travail de recherche. Nous présentons ici quelques-uns de ces travaux.

2.8.1 Technique de contour actif

Les contours actifs sont définis par une courbe paramétrique pouvant être fermée ou non. Un snake consiste à placer aux alentours de la forme à détecter une ligne initiale de contour. Cette ligne va se déformer progressivement selon l'action de plusieurs forces qui vont la tirer ou la pousser vers la forme. Ces forces sont représentées par trois énergies associées au snake, donc la courbe est placée dans la zone d'intérêt de l'image ou autour d'un objet. Des contraintes permettent de conserver une courbe lisse avec des points équidistants tout en laissant un certain champ libre pour les déformations [27].

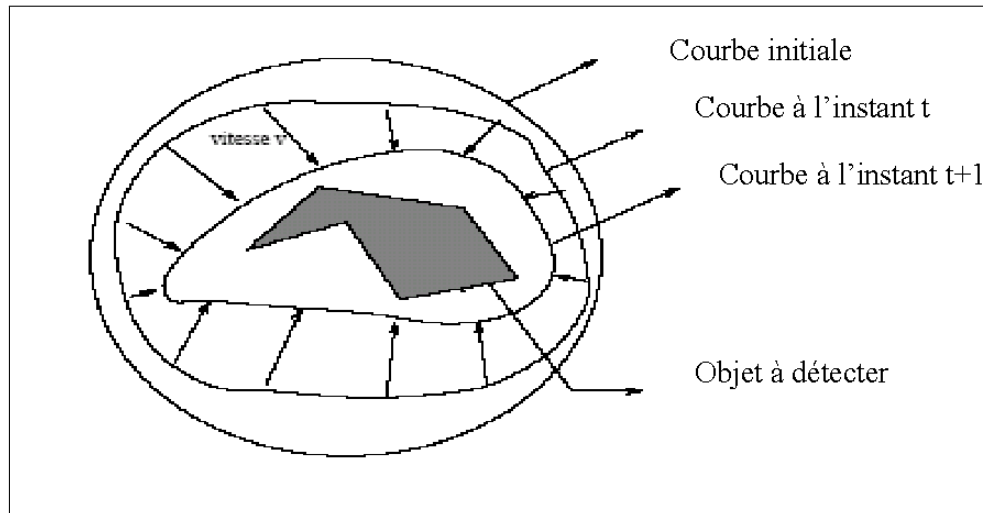


FIGURE 2.2 – Evolution du contour actif.

Les étapes de la méthode proposée peuvent être schématisées par la figure suivante.

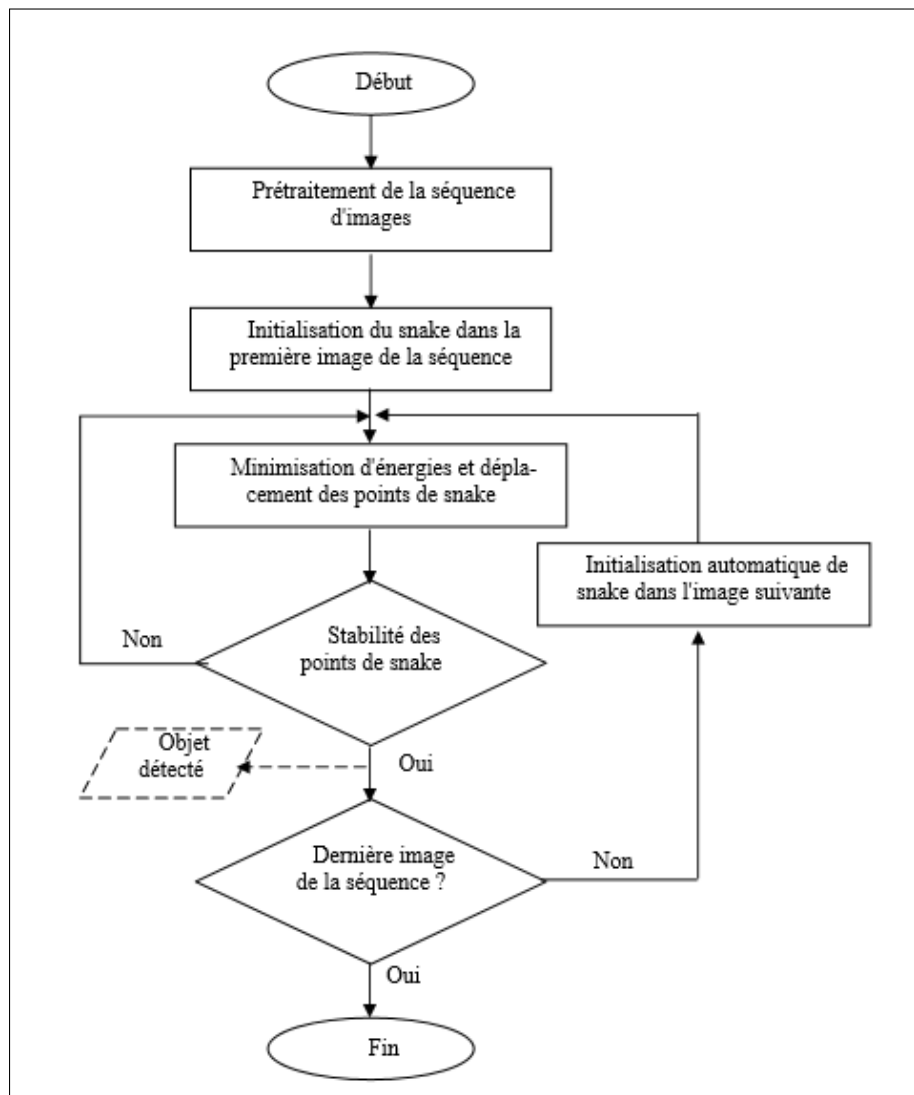


FIGURE 2.3 – Schéma globale de technique du contour actif.

✓ **Prétraitement de la séquence** : C'est une étape nécessaire qui permet de réduire l'effet du bruit présent dans la plupart des images médicales. Ce prétraitement dépend du type de la séquence utilisé et de l'effet de bruit.

✓ **du snake dans la première image** : On initialise notre contour actif à l'intérieur de l'objet ou de telle sorte qu'il englobe l'objet dont on veut détecter ses contours. Ce contour actif est composé d'un ensemble des points, et il est initialisé sous forme d'un cercle. Pour cela, on doit initialiser seulement le centre du cercle et son rayon.

✓ **Initialisation automatique de snake dans l'image suivante** : Après la détection des contours de l'objet dans la première image, une étape de suivi de cet objet est lancée dans les images ultérieures.

Cette initialisation manuelle est remplacée par une initialisation automatique qui utilise des informations obtenues à partir des traitements précédents. Il représente, dans l'image courante, le centre du cercle d'initialisation du snake.

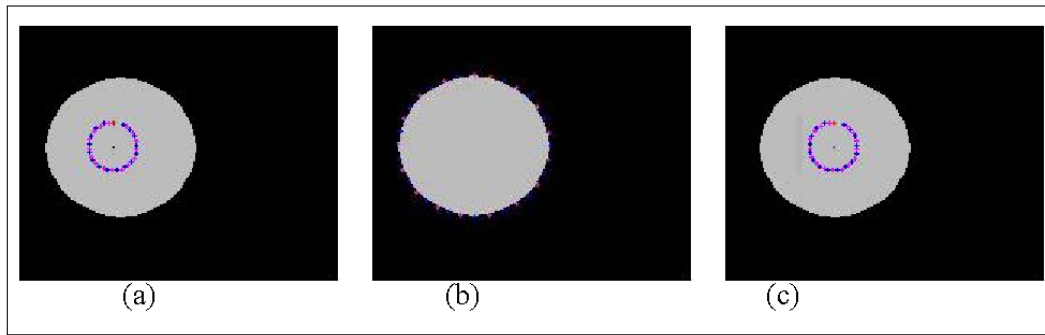


FIGURE 2.4 – Initialisation automatique du snake dans la deuxième image.

2.8.2 Détection robuste d'objets couleur

La méthode de détection et de reconnaissance d'objets couleur est composée des étapes suivantes [28] :

- ✓ Segmentation couleur afin de séparer l'objet du fond complexe.
- ✓ Etiquetage de l'image pour distinguer chaque objet.

Segmentation couleur

Afin de segmenter l'image couleur, ils ont utilisé un algorithme efficace proposé et décrit par D. Il est basé sur l'analyse d'espace de caractéristiques.

Dans ce paradigme, les pixels sont groupés et exprimés dans un espace de couleur. Chaque groupe représente une région homogène dans l'image. L'algorithme utilise la procédure «Mean Shift» ou déplacement de la moyenne pour palier aux problèmes issus des techniques de clustering. L'algorithme de segmentation proposé ici utilise l'espace des couleurs L^*u^*v (au lieu de l'espace RGB) plus robuste pour calculer les distances dans l'hypersphère qui constitue la fenêtre de «Mean shift».

Etiquetage de composantes connexes

L'extraction des composantes connexes est largement utilisée en reconnaissance de formes pour étiqueter les régions détectées par segmentation. La technique consiste à regrouper les pixels voisins dans un ensemble appelé composante connexe. Chaque ensemble est disjoint des autres et peut ensuite être aisément isolé.

2.8.3 Soustraction de l'arrière-plan

L'image de l'arrière-plan d'une scène contient les éléments statiques de cette scène, c'est la représentation de l'environnement dans lequel les objets évoluent. Cette image est considérée comme une image de référence qui se soustrait de chaque image de la séquence afin d'extraire les objets en mouvement [29].

Cette méthode dans sa version simple, consiste à faire la différence des intensités des pixels $I(x,$

y, t) d'une image I prise à l'instant t , par rapport aux intensités des pixels $B(x,y)$ d'une image de référence d'arrière-plan B , où il n'y a pas de mouvement.

Elle permet de détecter les pixels (x, y) ayant subi un changement d'intensité significatif.

Ces pixels forment les objets en mouvement. Cela est modélisé par :

Si $|I(x, y, t) - B(x, y)| > \tau$

Alors le pixel (x, y) de I à l'instant t est en mouvement

Sinon le pixel (x, y) appartient à l'arrière-plan

Avec τ un seuil fixé manuellement.

2.8.4 Détection d'objets RVB-D inspirée de l'incertitude

La détection de saillance RVB-D, qui apprend la distribution conditionnelle sous-jacente des cartes de saillance plutôt qu'une fonction de cartographie de RGB-D entrée dans une seule carte de saillance. Soit $D = N_i, Y_i, N, i = 1$ soit l'ensemble de données d'entraînement, où X_i désigne l'entrée RVB-D, Y_i désigne la carte de saillance GT(ground-truth), et N désigne le nombre total d'images dans l'ensemble de données. Nous avons l'intention de modéliser $P_\omega(Y | X, z)$, où z est un latent variable représentant l'incertitude inhérente aux régions saillantes ce qui peut également être vu dans la façon dont un humain annote les objets saillants [30].

Notre cadre utilise deux composants principaux pendant la formation :

- 1) a modèle de générateur, qui mappe l'entrée RGB-D X et la variable latente z à la prédiction conditionnelle $P_\omega(Y | X, z)$.
- 2) une inférence modèle, qui déduit la variable latente z . Pendant les tests, nous pouvons échantillonner plusieurs variables latentes de la distribution antérieure apprise $P_\theta(z | X)$ pour produire une prédiction de saillance stochastique. Plus précisément, pendant l'entraînement, le modèle apprend la saillance du «modèle de générateur» et se met à jour la variable latente avec le «modèle d'inférence». Pendant les tests, nous échantillon de la distribution «a priori» de la variable latente pour obtenir prédictions de saillance stochastiques.

2.9 Conclusion

Dans ce chapitre on a parlé sur les notions de bases de détection des objets et quelques méthodes modernes pour la détection et expliquez les principes de leur fonctionnement. Dans le prochain chapitre, nous allons parler de notre proposition, nous l'avons fait a technique de divergence des couleurs (CDT) proposée pour détecter les objets saillants.

CHAPITRE 3

L'APPROCHE PROPOSÉE

3.1 Introduction

En ce chapitre concerne la proposition en générale, en utilise une nouvelle méthodologie de regroupement qui classifera la distribution pixels d'image dans un groupe, afin de montrer l'objet mis en évidence sur cette image. Tel que cette proposition est basée sur une technique de clustering se concentre sur l'écart type pour détecter les objets saillants en utilisant un processus de clustering optimisé.

3.2 Contexte et contribution

La segmentation d'image est un processus qui divisé une image en des parties ou des objets constitutifs. Plusieurs algorithmes à usage général et des techniques ont été développées pour la segmentation d'image car il n'y a pas de solution de premier plan pour la segmentation d'image de ce problème. Les régions d'une image numérique dont les pixels sont caractérisés par l'homogénéité des couleurs peut être interprétée comme constituant des parties des objets présents dans l'image. Cependant, si une image en couleurs vraies 24 bits est considéré, où le nombre de couleurs différentes possibles peut atteindre 16 millions, le nombre de régions parfaitement homogènes dans l'image serait le plus probablement sensiblement plus grand que le nombre d'objets perçus dans l'image par un observateur humain. Dans fait, le système visuel humain est capable de distinguer une grand nombre de couleurs. En informatique visuelle, seul un groupe de couleurs avec une tonalité similaire pourrait définir les objets car peu de couleurs sont souvent suffisant pour comprendre et interpréter l'image.

3.3 L'hypothèse dans un contexte philosophique

Le regroupement d'images couleur de scènes naturelles est largement étudié pour de nombreuses applications en vision par ordinateur comme la segmentation d'image et détection des objets saillants.

La combinaison de différentes fonctionnalités pour la détection des régions saillantes pourrait être une bonne solution car manuellement modéliser une méthode d'interaction pour intégrer des caractéristiques saillantes est un problème difficile. Une bonne intégration d'une méthode pourrait être développée pour résoudre ce problème d'adoption le mécanisme CNN Convolutional Neural Network pour former une saillance modèle d'intégration de carte. Ce mémoire vise à introduire une nouvelle méthodologie de clustering qui classera la distribution des pixels d'une image en ensembles de clusters pour montrer efficacement l'objet saillant sur cette image. La proposition l'hypothèse doit être conceptuellement simple en termes de complexité afin de fournir un temps de calcul rapide. De plus, la technique proposée doit tenir compte des limites des algorithmes précédemment proposés, car de nombreux algorithmes sont utilisé pour détecter les objets saillants. La technique de clustering proposée se concentre sur l'écart type pour détecter les objets saillants à l'aide d'un processus de clustering

optimisé. Le regroupement d'images couleur de scènes naturelles est largement étudié pour de nombreuses applications en vision par ordinateur comme la segmentation d'image et détection des objet saillant. La combinaison de différentes fonctionnalités pour la détection des régions saillantes pourrait être une bonne solution car manuellement modéliser une méthode d'interaction pour intégrer des caractéristiques saillantes est un problème difficile. Une bonne intégration d'une méthode pourrait être développée pour résoudre ce problème d'adoption le mécanisme CNN Convolutional Neural Network pour former une saillance modèle d'intégration de carte.

Ce mémoire vise à introduire une nouvelle méthodologie de clustering qui classera la distribution des pixels d'une image en ensembles de clusters pour montrer efficacement l'objet saillant sur cette image. La proposition de l'hypothèse doit être conceptuellement simple en termes de complexité afin de fournir un temps de calcul rapide. De plus, la technique proposée doit tenir compte des limites des algorithmes précédemment proposés, car de nombreux algorithmes sont utilisés pour détecter les objets saillants. La technique de clustering proposée se concentre sur l'écart type pour détecter les objets saillants à l'aide d'un processus de clustering optimisé.

3.4 La technique de divergence de couleur proposée (CDT) pour détecter les objets saillants

Ce travail présente une technique de détection d'objets saillants basée sur la divergence de couleur à l'aide d'une représentation de cube RVB. Un cube RVB représente les couleurs d'une image en trois dimensions, rouge, vert et Bleu. L'écart type est utilisé dans le but de découvrir des classes d'homogénéité des couleurs dans un endroit proche, c'est-à-dire qui sont très susceptibles de représenter des objets saillants. Plus précisément, nous utilisons d'abord la matrice euclidienne de distance entre chaque pixel de couleur RVB de l'image d'origine et les huit sommets du cube RVB.

Pour chaque pixel, huit valeurs (les distances entre ce point et huit sommets du cube) seront calculées. La neuvième valeur est l'écart type des huit valeurs calculées. Ce calcul réduira trois valeurs de pixels d'une image à une seule valeur *sdv*. En fonction de la couleur souhaitée des objets saillants, nous classons *sdv* pour extraire les objets saillants.

Ensuite, nous appliquons différents types de coefficients pour obtenir un ensemble de matrices, où pour chaque matrice, nous créons le vecteur d'écart type dans le but de l'utiliser pour la classification d'images. Enfin, nous obtenons une carte de saillance approximative après la fusion des résultats de toutes les classifications de coefficients.

On observe que la difficulté à détecter les objets saillants ou à dessiner la limite d'un objet dépend de la qualité et de la complexité de l'image.

Dans ce travail, nous allons redistribuer les couleurs RVB de chaque pixel d'une image donnée dans un cube 3D en suivant les trois valeurs (Rouge, Vert et Bleu). Dans le but d'extraire les objets saillants,

la technique proposée regroupe ces points de cube 3D dispersés. Le principe est qu'au lieu d'explorer des millions de points pour détecter les objets saillants, le processus de regroupement minimise ce nombre en fonction du nombre de clusters proposés.

Par conséquent, la segmentation d'image dans cette contribution sera atteinte en utilisant le processus de clustering. Chaque segment présentera un objet et les objets qui résident dans la région éloignée seront présentés comme des objets saillants.

3.4.1 Classification de cube RVB

Le but de la représentation du cube RVB est de montrer les classes de couleur homogénéité dans un endroit proche qui sont très susceptibles de représenter objets. La technique proposée se concentre sur l'idée de représenter le cube de couleur RVB d'une image fournie en entrée. Nous supposons que chaque couleur d'objet réside dans une région distincte du cube 3-D en fonction de la couleur de l'objet. La technique proposée se concentre sur le principe selon lequel la détection d'un objet est un processus qui part de unifier les couleurs de chaque objet. Cette procédure réduira la nombre de couleurs similaires à une couleur de manière à fournir une nouvelle version de l'image d'origine tout en préservant les caractéristiques de l'image.

Segmentation d'image basée sur un vecteur de distance

L'espace de couleur RVB est caractérisé par un cube avec les axes R, G, B , et les couleurs sont interprétées comme des vecteurs tridimensionnels (3D), chaque élément vectoriel ayant une plage dynamique de 8 bits. Chacun des trois bords du cube a une longueur de 256, car chaque composante de couleur peut prendre n'importe quelle valeur dans la plage $[0, 255]$ comme le montre la figure suivante :

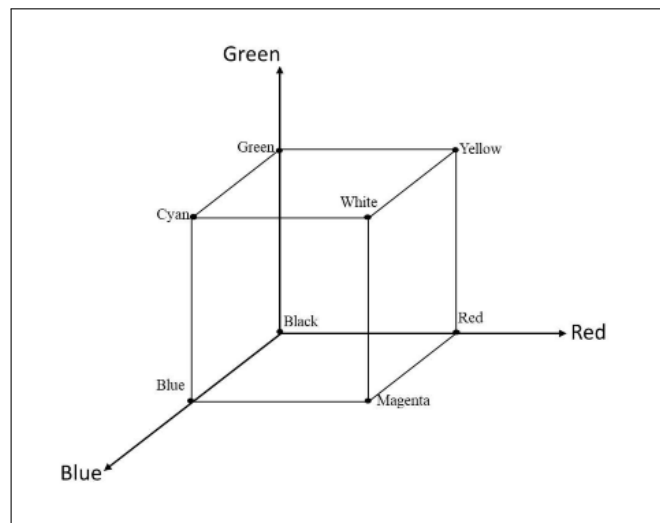


FIGURE 3.1 – Un cube 3D représentant géométrique le RVB espace de couleur.

La représentation d'une image sous forme de cube de couleur RVB est utilisée pour unifier des pixels similaires, c'est-à-dire regrouper les points similaires qui représenteront un objet. À cette fin, les distances entre chaque pixel et le huit sommets de cube sont calculés à l'aide de l'équation 1. Le vecteur de distance d'un pixel donné p est défini comme suit : $D_p = [d_0, d_1, d_2, d_3, d_4, d_5, d_6, d_7]$

Où d correspond à la distance entre le pixel p et le sommet i ($i = 0...7$) cube qui se calcule comme suit :

$$d_i = \sqrt{r_i^2 + g_i^2 + b_i^2} \quad (1)$$

où r_i, g_i et b_i représentent respectivement le rouge, le vert et le bleu valeurs du pixel p dépendant du sommet i .

Les valeurs obtenues D_p de tous les pixels seront présentées dans un tableau dans le but de segmenter l'image d'entrée en représentant des couleurs par une seule couleur.

Par conséquent, le processus de segmentation classe les lignes de la table de telle manière que chaque couleur similaire apparaît sur les lignes consécutives.

Tri du tableau obtenu en suivant les distances d_p peut offrir un résultat préliminaire.

Cet ordre prendra en compte tous les pixels points qui ont la même valeur d_i que des pixels similaires, par exemple le bleu et des pixels rouges de même valeur par rapport à d_0 . Par conséquent, l'écart type est utilisé pour distinguer les pixels ayant le même d_i .

On définit le vecteur d'écart type sdv qui sera utilisé pour ordonner les lignes du tableau. Une colonne sera insérée dans le tableau qui contient l'écart type du vecteur distance D_p .

$$sdv_p = \frac{\sqrt{\sum (d_i - \text{average} D_p)^2}}{8} \quad (2)$$

Où d_i correspond à la distance du pixel p et du les sommets i ($i = 0...7$) du cube. $\text{average} D$ est la moyenne du vecteur distance D_p .

En fait, l'écart type peut offrir plus de précision en distinguer d'autres points de pixel de l'image d'entrée. Cependant, nous peut observer que certains pixels symétriques ont le même sdv , c'est-à-dire les pixels qui ont le même écart type avec des valeurs d_i différentes. Mathématiquement, deux pixels sont symétriques si la condition suivante est vérifiée :

$$\exists i \neq j \mid d_i = d_j \mid \quad \forall i, j \in \{0, \dots, 7\} \quad (3)$$

Nous définissons le coefficient f pour séparer ces pixels symétriques. Le cube qui présente les couleurs sera présenté dans un nouveau cube avec bords irréguliers. Par conséquent, nous attribuons une valeur différente de f pour chaque sommet du cube. Les valeurs du d_i sont calculés comme suit :

où f_i est utilisé pour garantir que les pixels symétriques ne ont le même écart-type et n'appartient pas au même classe, séparant ainsi clairement les différents objets. On remarque que le coefficient f ne sépare les valeurs symétriques que lorsque la condition suivante est vérifiée : $i \neq j \Rightarrow f_i \neq f_j$.

La figure 3.2 montre un exemple de cube irrégulier avec $f_i =$

Les huit valeurs de f doivent être attribuées avec soin car leur rôle est de séparer les objets de l'image et de détecter les objets saillants.

Par conséquent, nous attribuons à la couleur d'arrière-plan de l'image le minimum valeur de f et de l'objet le plus saillant la valeur maximale. Par défaut, on attribue la valeur minimale à la couleur noire et la valeur maximale à la couleur blanche. Le potentiel de ces coefficients pourrait être observé lors de la détection d'objets saillants ayant une couleur dominante, donc on peut attribuer la valeur maximale au coefficient de la dominante couleur de l'objet saillant.

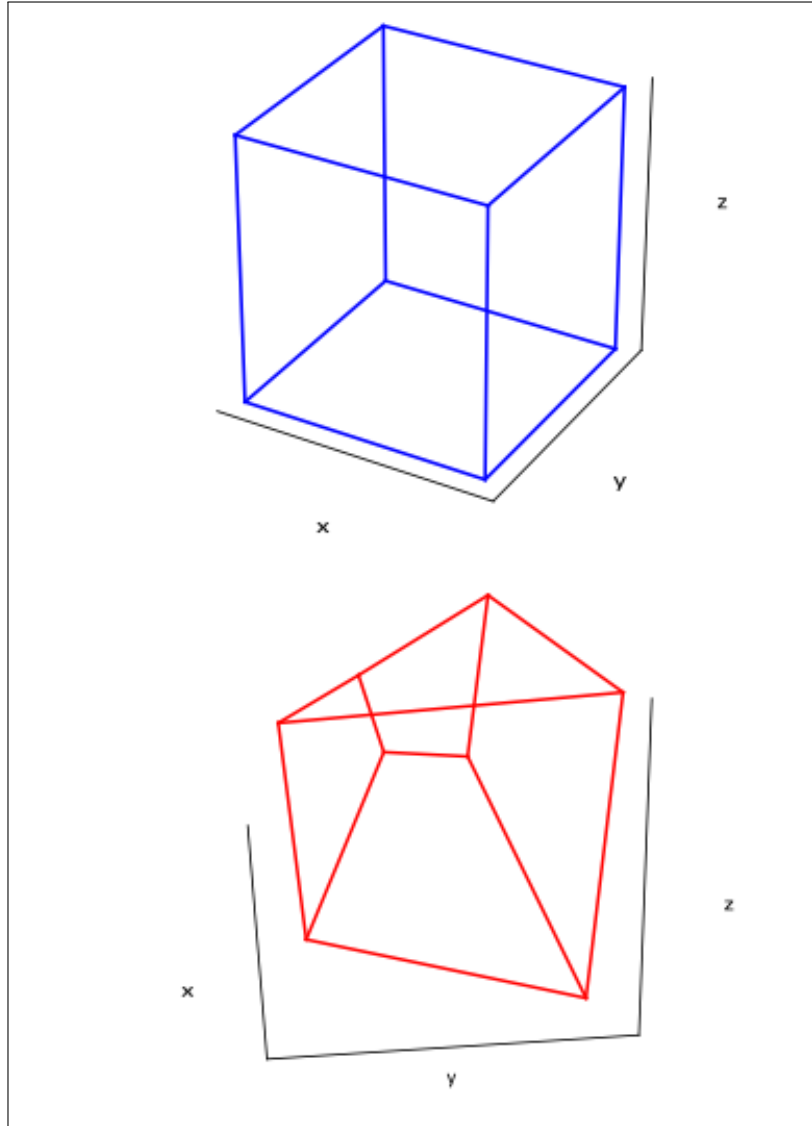


FIGURE 3.2 – Un cube 3D régulier et irrégulier géométrique représentant les écarts types des pixels.

3.4.2 Segmentation et détection d'objets saillants

Pour détecter les objets saillants, nous utilisons une table qui contient la valeur de l'écart type sdv et des valeurs de pixel de l'image d'entrée. Le tableau sera trié par ordre croissant suivant l'écart

type *sdv*. Ensuite, nous le divisons en classes où les pixels de la même classe représenteront une classe d'objets. La dernière classe représente les objets les plus saillants qui existent dans l'image. Dans cet travail, nous segmentons l'image en deux classes en fin d'obtenir plusieurs cartes de saillance en appliquant des valeurs différentes de coefficients à la matrice de distance euclidienne. Par conséquent, nous allons créer pour chaque type de coefficients un tableau des écarts-types et chaque tableau d'écart-type est divisé en deux classes qui représentent la carte de saillance.

3.5 Conclusion

Dans ce chapitre, on a présenté notre proposition pour objectif détection des objets par la dégradation de couleurs dans les images. Nous avons utilisé une technique de clustering se concentre sur l'écart type en utilisant un processus de clustering optimisé.

CHAPITRE 4

IMPLÉMENTATION ET ANALYSE

4.1 Introduction

On a présenté dans le chapitre précédent notre proposition pour objectif détection des objets par la dégradation de couleurs dans les images.

Nous entamons maintenant sur ce dernier chapitre la partie de la réalisation et de l'analyse, nous présentons notre application réalisée et les différentes mesures d'évaluation pour montrer la simplicité en termes de complexité afin de fournir un temps de calcul rapide et l'efficacité de l'algorithme proposé.

4.2 Environnement de développement

Dans cette partie nous allons citer l'environnement logiciel (Software) et matériel (Hardware) utilisés.

4.2.1 Environnement logiciel

On a utilisé l'environnement Python pour l'implémentation de notre approche.

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.

On a utilisé Python pour notre projet de traitement d'image parce qu'il est riche en bibliothèques prêtes à l'emploi qui facilitent le processus de programmation pour que nous soyons plus concentrés sur l'idée du projet. Voici quelques-unes des bibliothèques de traitement d'image utiles et disponibles gratuitement en Python nous l'avons utilisé dans notre projet : Numpy, PIL/Pillow, Math et Tkinter.

4.2.2 Environnement matériel

L'application a été développée sur un PC(Laptop HP modèle : Latitude E5450) ayant les caractéristiques suivantes :

- Processeur : Intel®Core(TM) i3-5300U CPU 2.30GHz 2.29 GHz.
- Mémoire installée (RAM) :8,00 Go.
- Disque Dure :HDD 500 Go.
- Carte graphique :Intel HD graphics5500.
- Système d'exploitation :Windows 10 Professionnel 64 bit.

4.3 L'évaluation des performances

L'évaluation des performances de l'algorithme de clustering proposé est pertinent pour montrer ses performances principalement en termes de rapidité d'exécution et de précision. Pour cela, nous allons

tester sur des images réelles et voir ce qu'il extrait. En fait, il est difficile de juger de la qualité pour obtenir des résultats sans comparer avec d'autres algorithmes. Dans cette section, nous présentons les expériences réalisées afin de les voir et évaluer notre technique de classification par rapport aux autres.

L'objectif de ces expériences est notamment d'évaluer les capacités de l'algorithme à détecter et à extraire correctement le plus objets saillants basés sur ces couleurs déjà prédéfinies.

Les prochaines images montre des exemples de la technique de divergence de couleur pour détecter des objets saillants. L'évaluation du CDT proposé se fait souvent en appliquant la technique à des sous-ensembles de repères. Nous avons mené nos expériences avec les différents algorithmes contenant un vision. Pour toutes les expériences que nous avons menées, nous avons varié le temps de traitement des paramètres et la précision.

4.4 Les tests et résultats

On a utilisé des images aléatoires avec différents niveaux de difficulté pour détecter les objets saillants car notre objectif dans cette étude est de comparer la technique proposée dans différents cas. Le tableau 1 montre les premiers résultats obtenus. (Nous avons pris ce tableau à travers d'étude de l'encadreur)

Le temps de traitement et la précision :

Technique	Accuracy %	Processing time (ms)
Our CDT	0.785	4.202
SVR	0.662	6.342
DSR_D	0.801	8.127

TABLE 4.1 –

4.5 Discussion

L'ensemble de données (dataset) MSRA-B traité [31] contient une diversité de 5000 images largement utilisées dans la communauté de détection d'objets saillants. Cette ensemble de données est traité par un programme Python pour visualiser le traitement des résultats de temps et de précision. Les résultats obtenus en appliquant plusieurs d'exécutions sont présentées dans le tableau 4.1. Pour la précision des pixels évalue la précision des objets saillants détectés, nous avons évalué $accuracy = TP + TN / (TP + TN + FP + FN)$ qui est le pourcentage de pixels de l'image correctement classés. Un vrai positif représente un pixel dont il est prédit correctement qu'il appartient au saillant objet (selon la vérité terrain) alors qu'un vrai négatif représente un pixel correctement identifié comme n'appartenant pas au objet saillant.



FIGURE 4.1 – L'image originale pour détecter les objets rouges saillants.

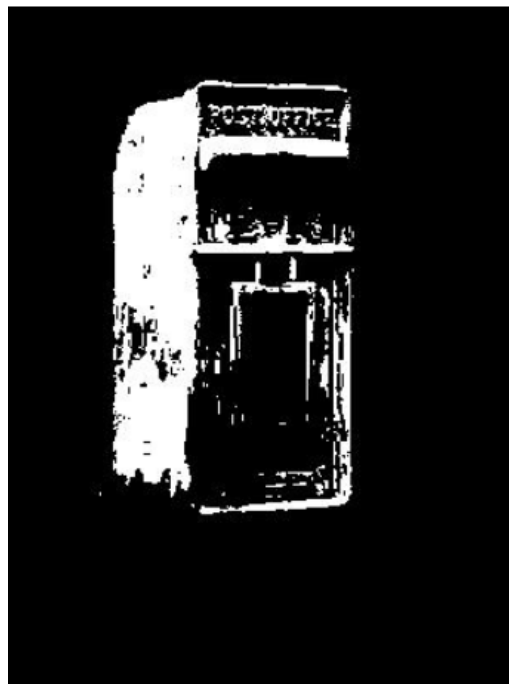


FIGURE 4.2 – L'image de sortie après détection des objets rouges saillants.



FIGURE 4.3 – L'image d'entrée pour la détection d'objets saillants en noir et blanc.



FIGURE 4.4 – L'image de sortie après la détection d'objets saillants en noir et blanc.



FIGURE 4.5 – L'image d'entrée pour la détection d'objets saillants en image haute résolution(1200*1859).

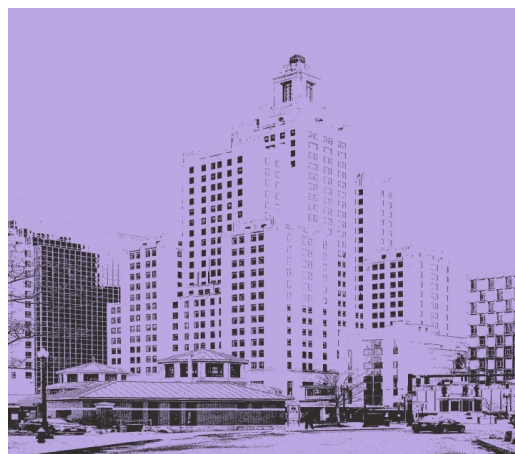


FIGURE 4.6 – L'image de sortie après la détection d'objets saillants en image haute résolution(1200*1859).

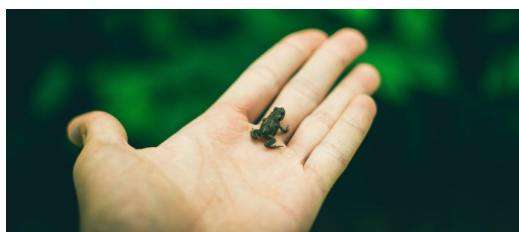


FIGURE 4.7 – L'image d'entrée pour la détection d'objets saillants dans un environnement complexe.

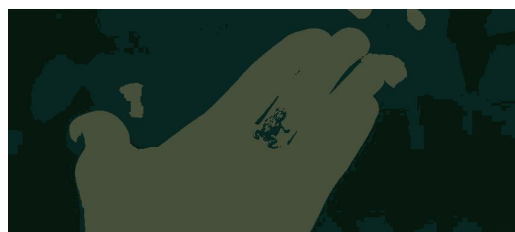


FIGURE 4.8 – L'image de sortie après la détection d'objets saillants dans un environnement complexe.



FIGURE 4.9 – L'image d'entrée pour la détection d'objets saillants en image avec des couleurs imbriquées.



FIGURE 4.10 – L'image de sortie après la détection de tous objets saillants (non-sélective).

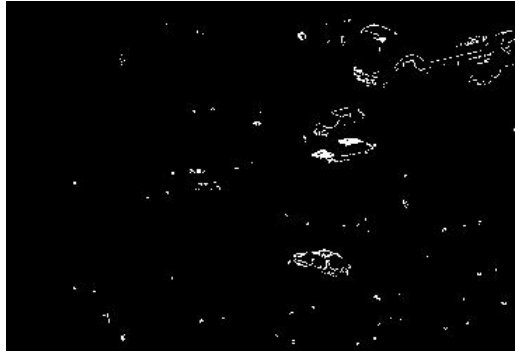


FIGURE 4.11 – L'image de sortie après détection des objets rouges saillants.

4.6 Conclusion

Nous avons présenté dans ce chapitre notre application réalisée, et les différents outils nécessaires utilisés dans son développement. Nous avons aussi terminé par un ensemble d'analyses et tests pour les résultats expérimentaux qui montrent la bonne détection des objets par la dégradation de couleurs dans les images. Tel que par l'analyse de notre méthode, obtenir le temps de traitement et la précision efficacité.

CONCLUSION GÉNÉRALE

La détection des objets est une technologie informatique dépende à la vision par ordinateur et le traitement d'images qui traite de la détection des instances des objets d'une certaine classe (tels que les humains, les bâtiments ou les voitures) dans les images et vidéos numériques.

En regardant dans une image ou un flux vidéo, comment détecter l'objet saillant et déterminer quel groupe connu des objets peut être présent et fournir des informations sur leur position dans l'image avec précision ?

Une nouvelle technique d'étalonnage des couleurs a été proposée dans ce memoire de manière efficace de détection des objets à l'aide de la divergence de couleur et la représentation du cube RVB.

Nous avons appliqué le concentration de classification du cube RVB Sur les valeurs du vecteur de distance et de l'écart type de clustering des pixels RVB.

Notre méthode atteint une performance comparable en durée de traitement avec les systèmes compacts de détection d'objets saillants.

Notre technologie est efficacement capable d'effectuer des tâches exceptionnelles avec les objets colorés dans des situations simples et difficiles, ce qui contribue à améliorer le Processus d'apprentissage en profondeur pour l'étape d'identification des objets.

De plus, le processus de détection des objets saillants est simple et rapide La vitesse de détection peut offrir des avantages supplémentaires au traitement des images dans les domaines de la vision par ordinateur et de l'intelligence artificielle. Comme futur de Travail, nous vérifierons ses performances on exécutant sur la Reconnaissance des objets à l'aide de différentes méthodes d'apprentissage en profondeur.

BIBLIOGRAPHIE

- [1] D. Ameisen, “Qu’est-ce qu’une image numérique ?”
- [2] D. Salles, “Éducation à l’image et aux médias : la liberté de la presse,” *Centre de Ressources en éducation aux médias CREM*, 2005.
- [3] M. Rabbani and P. W. Jones, *Digital image compression techniques*. SPIE press, 1991, vol. 7.
- [4] A. Habri and K. Houari, “Un système d’indexation et recherche d’images par le contenu basée sur la classification,” 2011.
- [5] K. Tabari and S. Tagma, “Compression d’images animées à très faible débit par la géométrie des fractales,” *PFE USTHB*, 1996.
- [6] R. Gonzales and R. Woods, “Digital image processing addison-wesley,” *Reading, MA*, 1992.
- [7] (2017, jul. 15) Les formats d’image. [Online]. Available : <http://tvaira.free.fr/bts-sn/activites/activite-webcam-reseau/formats-image.html#:~:text=Un%20format%20d’image%20est,essentiellement%20deux%20types%20d’images%20%3A&text=les%20images%20vectorielles>
- [8] S. Dourlens, “Intelligence artificielle et automatique,” 06 2015.
- [9] R. Horaud and O. Monga, *Vision par ordinateur : outils fondamentaux*, 1995.
- [10] (2016, jan. 03) La vision par ordinateur. [Online]. Available : <https://www.numericube.com/la-vision-par-ordinateur>
- [11] D. Boukhlof, “Résolution de problèmes par écosystèmes : Application au traitement d’images,” Master’s thesis, Université Mohamed Khider-Biskra, 2005.
- [12] Z. BELDJILALI, “Intitulé du sujet : La détection des objets dans des images,” Master’s thesis, 2019.
- [13] N. Voisine, “Approche adaptative de coopération hiérarchique de méthodes de segmentation : application aux images multicomposantes,” Ph.D. dissertation, Rennes 1, 2002.

- [14] Y. J. Zhang, "A survey on evaluation methods for image segmentation," *Pattern recognition*, vol. 29, no. 8, pp. 1335–1346, 1996.
- [15] R. Deriche, "Using canny's criteria to derive a recursively implemented optimal edge detector," *International journal of computer vision*, vol. 1, no. 2, pp. 167–187, 1987.
- [16] A. J. Lipton, H. Fujiyoshi, and R. S. Patil, "Moving target classification and tracking from real-time video," in *Proceedings fourth IEEE workshop on applications of computer vision. WACV'98 (Cat. No. 98EX201)*. IEEE, 1998, pp. 8–14.
- [17] A. Yilmaz, O. Javed, and M. Shah, "Object tracking : A survey," *Acm computing surveys (CSUR)*, vol. 38, no. 4, pp. 13–es, 2006.
- [18] G. Vincent, "Suivi d'objets d'intérêt dans une séquence d'images : des points saillants aux mesures statistiques," Ph.D. dissertation, 2008.
- [19] H. T. Nguyen, M. Worring, and R. Van Den Boomgaard, "Occlusion robust adaptive template tracking," in *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*, vol. 1. IEEE, 2001, pp. 678–683.
- [20] T. Penne, "Développement d'un système de tracking vidéo sur caméra robotisée," Ph.D. dissertation, 2011.
- [21] A. Bugeau, "Détection et suivi d'objets en mouvement dans des scenes complexes, application a la surveillance des conducteurs." Ph.D. dissertation, 2007.
- [22] T. Chateau, F. Jurie, R. Marc, and M. Dhome, "Suivi et reconnaissance de gestes par vision monoculaire en temps réel : application à la formation des chargés de manoeuvres pour la conduite des ponts polaires," 2004.
- [23] D. W. Jacobs, D. Weinshall, and Y. Gdalyahu, "Classification with nonmetric distances : Image retrieval and class representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 6, pp. 583–600, 2000.
- [24] M. Oren, C. Papageorgiou, P. Sinha, E. Osuna, and T. Poggio, "Pedestrian detection using wavelet templates," in *Proceedings of IEEE computer society Conference on computer vision and pattern recognition*. IEEE, 1997, pp. 193–199.
- [25] S. Jehan-Besson, "Modèles de contours actifs basés régions pour la segmentation d'images et de vidéos," Ph.D. dissertation, 2003.
- [26] L. Mousaoui, F. Z. Fares, and S. Kadi, "Suivi de cible," Master's thesis, 2015.
- [27] A. Fekir, N. Benamrane, and A. Taleb-Ahmed, "Détection et suivi d'objets dans une séquence d'images par contours actifs." in *CIIA*, 2009.
- [28] C. Maaoui, H. Laurent, and B. Emile, "Reconnaissance et détection robuste d'objets couleur," in *20 Colloque sur le traitement du signal et des images, FRA, 2005*. GRETSI, Groupe d'Etudes du Traitement du Signal et des Images, 2005.

- [29] B. Asma and B. Sarah, “Détection, segmentation et suivi d’un individu en mouvement dans une vidéo de surveillance,” 2011.
- [30] J. Zhang, D.-P. Fan, Y. Dai, S. Anwar, F. Saleh, S. Aliakbarian, and N. Barnes, “Uncertainty inspired rgb-d saliency detection,” *arXiv preprint arXiv :2009.03075*, 2020.
- [31] T. Liu, Z. Yuan, J. Sun, J. Wang, N. Zheng, X. Tang, and H.-Y. Shum, “Learning to detect a salient object,” *IEEE Transactions on Pattern analysis and machine intelligence*, vol. 33, no. 2, pp. 353–367, 2010.