

الجمهورية الجزائرية الديمقراطية الشعبية
وزارة التعليم العالي والبحث العلمي

جامعة سعيدة د. مولاي الطاهر
كلية التكنولوجيا
قسم: الإعلام الآلي



Mémoire de Master

Spécialité : **Modélisation Informatique des connaissances et du raisonnement**

Thème

**RECHERCHE D'IMAGES PAR LE CONTENU :
UNE APPROCHE BASÉE SUR LE DEEP LEARNING**

Présenté par :

BOUBTAIMA Marwa

LAREDJ Nor El Houda

Dirigé par :

M. MOSTEFAI Abdelkader



Année universitaire 2022-2023

Remerciement

Nous remercions avant tout, Allah de nous a prodigué la force morale et physique et nous a permis d'achever ce travail

Nous tenons tout d'abord à remercier Mr MOSTEFAI Abdelkader pour nous avoir encadrées tout au long de ce présent projet, pour leur disponibilité, leurs critiques constructives, et leurs suggestions pertinentes.

Nous remercions les membres du jury d'avoir accepté d'examiner ce modeste travail

Nous remercions toutes les personnes qui nous ont aidées et entourées durant notre mémoire de près ou de loin, surtout Dr. CHAIBI Hassene et nos collègues TAZI Manar Dounia et KAID Khouloud

Nous remercions tous les enseignants du département d'informatique que nous les respectons beaucoup.

Enfin, nous remercions toutes nos familles et nos amis.

Marwa & Nor El Houda

Dédicace

*Je dédie cet humble travail à la lumière de mes yeux et à la source de mes efforts, ma chère maman *ZIDOURI FATI MA* et à la personne qui m'a le plus soutenu dans mon cheminement académique, mon père *MOUHMED* et je les remercie beaucoup.*

*A mes belles sœurs : *RIM* et *NACIRA* et *SOUNDOUS**

*A mon chère frère : *RADWAN* et *AYMENE* et *MOUSTAPHA**

A Toute ma famille du grand au petit, mes grands-parents et a tous mes amis qui sont restés près de moi dans les moments difficiles et quand j'en avais besoin.

*A mon binôme de travail * NOR EL HOUDA **



BOUBTAIMA MARWA

Dédicace

Je dédie ce modeste travail à :

À l'homme, mon précieux offre du dieu, qui doit ma vie, ma réussite et tout mon respect : mon cher père ;

À la femme qui a souffrit sans me laisser souffrir, qui n'a jamais dit non à mes exigences et qui n'a épargné aucun effort pour me rendre heureuse : mon adorable mère ;

À mes précieuses sœurs « Asma, Amina, Douaa »

À mes très chers frères « Ahmed Abderrahmane, Rayene » ;

À mon mari qui a partagé avec moi tous les moments d'émotion lors de la réalisation de ce travail, qui n'a pas cessé de me conseiller, qui il m'a chaleureusement supporté et encouragé ;

À mon cœur et ma vie mon adorable fils qui je fais tout ça pour lui "Taha Ahmed",

À toute ma grande famille

À mon binôme de ce travail MARWA**

À toute personne qui occupe une place dans mon cœur

À Tous ceux que j'aime et je respecte.



LAREDJ NOR EL HOUDA

RÉSUMÉ

Ce mémoire aborde la problématique de la recherche d'images par contenu dans des bases de données d'images.

Nous avons réalisé un CBIR en utilisant des réseaux de neurones convolutifs.

Notre application reçoit une image de requête, la convertit en niveaux de gris, puis la fusionne avec une image de la base de données qu'elle est convertie en niveaux de gris et la transmet au réseau neuronal convolutif, et ce, pour voir si les deux images sont similaires ou non.

ABSTRACT

This dissertation addresses the issue of image search by content in image databases.

We first perform a CBIR using convolutional neural networks.

Our application receives a query image, converts it to grayscale, then merges it with an image from the database that it converts to grayscale and passes it to the convolutional neural network. To see if the two images are similar or not.

ملخص

تتناول هذه الرسالة مسألة البحث عن الصور حسب المحتوى في قواعد البيانات.

أجرينا CBIR باستخدام الشبكات العصبية التلافيفية.

يتلقى تطبيقنا صورة طلب، ويحولها إلى التدرج الرمادي، ثم يدمجها مع صورة من قاعدة البيانات تم تحويلها إلى التدرج الرمادي، ويرسلها إلى الشبكة العصبية التلافيفية لمعرفة ما إذا كانت الصورتان متشابهتان أم لا.

SOMMAIRE



SOMMAIRE

Remerciement	i
Dédicace.....	ii
Dédicace.....	iii
Résumé	iv
Sommaire	v
Liste des Figures.....	ix
Liste des Tableaux.....	x
Liste des Abréviations	xi
Introduction Générale.....	1
1. Contexte générale.....	2
2. Objectifs de l'étude	3
3. Organisation du mémoire	3
CHAPITRE1 : Généralités sur le Traitement d'Images	4
1- Introduction	5
2- C'est quoi le traitement d'images :	5
3- C'est quoi une image :	5
3.1 Définitions :	5
3.2 Diverses catégories d'images :	5
4- Représentation des images :	6
4.1 Types d'images.....	6
4.2 Formats d'images :	13
5- Caractéristiques d'images :	14
5.1. Pixel :	14
5.2. La dimension :	14
5.3. La résolution :	15
5.4. La couleur :	15
5.5. Contours et Textures :	15
5.6. Luminance et Contraste :	16
5.7. Bruit :	16
6. Traitement d'image numérique :	17
7. Filtrage :	17
7.1. Définition :	17
7.2. Les types de filtrage :	17

8. Segmentation :	20
9. Quelques applications concrètes de traitement d'images :	20
10. Conclusion :	20
CHAPITRE 2 : Systèmes de Recherche d'Image par Contenu (CBIR)	21
1. Introduction :	22
2. Différents domaines d'application :	23
3. Composants d'un CBIR :	23
3.1. La base d'image :	24
3.2. L'index :	24
3.3. La gestion des index :	25
3.4. Les requêtes :	25
3.5. L'analyse de la requête :	26
3.6. Mise en correspondance requête/base :	26
3.7. La présentation des résultats :	27
4. Représentation des images dans un CBIR :	27
5. Mesures pour évaluer un système :	27
5.1. Rappel et précision :	27
5.2. La courbe de rappel et précision :	28
6. Conclusion :	28
CHAPITRE 3 : Descripteurs d'Images & Mesures de Similarité	29
1. Introduction	30
2. Les descripteurs d'image :	30
2.1. Les descripteurs de la couleur :	30
2.2. Descripteurs des textures :	33
2.3. Les descripteurs de la forme :	35
3. Mesures de similarité	38
3.1. Les méthodes de calcul	39
4. Conclusion	41
CHAPITRE 4 : Le Travail Effectué	42
1. Introduction :	43
2. Outils de développement :	43
- Python	43
- Google Colab	43
- Tensorflow	43
- Keras	43

3. La base d'image utilisée :	43
4. Les descripteurs d'image utilisés :	44
4.1. Les modèles de couleur utilisés.....	44
4.2. Grayscale Images.....	44
5. Architecture de l'application :	45
6. La structure du modèle :	47
7. Représentation graphique du modèle CNN :	48
8. Conclusion :	48
CHAPITRE 5 : Étude Expérimentale.....	49
1. Introduction :	50
2. Les méthodes d'évaluation :	50
2.1 : Fonction de perte (Loss Function) :	50
2.2. Accuracy (exactitude) :	50
3. Résultats du modèle réalisé :	50
3.1. Loss function	50
3.2. CNN train	51
3.3. Matrix de confusion :	52
4. Conclusion :	52
Conclusion Générale.....	53
Bibliographie & Sitographie	55
Bibliographie	56
Sitographie.....	57

LISTE DES FIGURES

Figure 1. Le spectre électromagnétique.....	7
Figure 2. Les trois couleurs primaires additives de lumière : Rouge, Vert, et Bleu	8
Figure 3. Les couleurs primaires de lumière et leur complémentaires (couleurs secondaires) ..	9
Figure 4. Le modèle TSL.....	9
Figure 5. La teinte, la saturation, et luminosité représentée sur une roue de couleurs	10
Figure 6. Représentation du modèle CIE XYZ par le triangle de Maxwell	11
Figure 7. Le modèle CIE LAB.	11
Figure 8. La dimension d'une image.	15
Figure 9. Exemples de textures.	16
Figure 10. Système de traitement d'image.	17
Figure 11. Principe du filtre médian.....	19
Figure 12. Principaux composants d'un système de recherche par le contenu.....	23
Figure 13. Les trois façons pour donner une requête dans CBIR (Deselaers, 2003)	26
Figure 14. Une courbe de rappel et précision	28
Figure 15. L'histogramme d'une image.	31
Figure 16. Limite des histogrammes et des couleurs dominantes.....	32
Figure 17. Représentation d'un corrélogramme de couleur	33
Figure 18. Exemples de matrices de cooccurrences.....	34
Figure 19. Exemple de décomposition par ondelettes.	35
Figure 20. Exemple de base à classer en familles d'images similaires	38
Figure 21. Quelques images de la base d'image.....	44
Figure 22. Schéma de représentation de l'application.....	45
Figure 23. Architecture du modèle CNN.....	48
Figure 24. « Loss function » du modèle réalisé	51
Figure 25. « CNN train » du modèle réalisé	51
Figure 26. Matrix de confusion du modèle réalisé.....	52

LISTE DES TABLEAUX

Tableau 1. RGB.....	13
Tableau 2. Tableau comparatif de différents formats d'images.....	14

LISTE DES ABRÉVIATIONS

RGB : Red Green Blue

RVB : Rouge Vert Bleu

CBIR : Recherche d'images par le contenu (Content Based Image Retrieval)

CNN : Convolutional Neural Networks

INTRODUCTION GÉNÉRALE



1. Contexte générale

Grâce aux avancées technologiques récentes, notamment dans les domaines du multimédia et de l'informatique, l'information numérique est devenue centrale dans tous les domaines d'activité : l'industrie, la médecine, la science, le monde du droit, le monde de la géographie. Le coût du matériel informatique a diminué, facilitant la diffusion et l'échange de données multimédias numérisées pour le grand public. Cette quantité de données n'a aucun sens si vous ne trouvez pas facilement les informations pertinentes à votre intérêt particulier. Ceci a créé un besoin pour le développement de techniques de recherche d'informations multimédia, en particulier la recherche d'images.

L'indexation et la recherche d'images par le contenu (CBIR) est une piste prometteuse. Elle offre la possibilité aux utilisateurs d'accéder, d'interroger et d'exploiter directement ces bases d'images en utilisant leur contenu ; ceci explique l'activité de recherche consacrée à ce domaine.

Le « *Deep Learning* » ou apprentissage profond est un type d'intelligence artificielle dérivé de la *machine Learning* (apprentissage automatique) où la machine est capable d'apprendre par elle-même, contrairement à la programmation où elle se contente d'exécuter à la lettre des règles prédéterminées. Une des technologies dans ce domaine est les réseaux de neurones convolutifs.

Un réseau de neurones convolutifs, en anglais : Convolutional Neural Networks (CNN) est une sorte d'architecture de réseau pour les algorithmes d'apprentissage en profondeur et est spécifiquement utilisé pour la reconnaissance d'images et les tâches impliquant le traitement de données de pixels.

2. Objectifs de l'étude

L'objectif principal de ce travail est la réalisation d'un CBIR en utilisant le modèle CNN.

Dans notre cas, nous nous intéressons aux images en couleur. Les attributs de couleur font partie des attributs de bas niveau les plus couramment utilisés pour rechercher des images dans des bases de données d'images. Ils ont été utilisés dans de nombreux systèmes de recherche de contenu. L'utilisation de la couleur dans ce domaine est motivée par deux facteurs. Premièrement, la couleur est un descripteur puissant qui facilite l'identification et l'extraction d'objets d'une scène. De plus, les humains peuvent distinguer des milliers de nuances et d'intensités de couleurs différentes. A travers le présent travail, nous avons essayé de construire une implémentation du système de recherche d'images couleur basé sur le contenu.

3. Organisation du mémoire

Notre manuscrit comprend les parties suivantes :

- Le premier chapitre représente une généralité sur le traitement d'image. Ce chapitre décrit la définition d'une image, ses propriétés, les principes des systèmes de traitement d'image, le filtrage, la segmentation et certaines applications spécifiques de traitement d'image.
- Le chapitre 2 concerne la définition des domaines et la recherche d'images en fonction de leur contenu. Ce chapitre décrit le fonctionnement du système de récupération image-contenu, ses composants et les métriques utilisées pour évaluer le système de récupération image-contenu.
- Le troisième chapitre décrit les descripteurs d'image tels que la couleur, la texture et la forme. Ce chapitre fournit également un aperçu des mesures de similarité entre les images de la base de données.
- Le quatrième chapitre présente les différentes étapes de notre travail.
- Le cinquième chapitre c'est l'étude expérimentale qui représenté les résultats de Notre application.

[illegible]

1- Introduction

Le traitement d'images est la branche des mathématiques appliquées qui étudie les images numériques et leurs transformations, dans le but d'améliorer la qualité des images ou d'extraire des informations des images.

Ce chapitre, va présenter les notions de base de l'image et du traitement de l'information. On donnera quelques concepts de l'image et leurs types, et par la suite nous présenterons quelques applications de traitement d'image.

2- C'est quoi le traitement d'images :

Le traitement d'image est un ensemble d'opérations effectuées sur une image pour améliorer sa lisibilité et sa facilité d'interprétation. C'est le cas, par exemple, pour les opérations d'amélioration du contraste, de réduction du bruit et de suppression du flou. C'est aussi un ensemble d'opérations effectuées pour extraire des "informations" d'une image p.ex. Segmentation et extraction des bords. [01]

Vous pouvez également effectuer des opérations de prétraitement avant le traitement de l'image. Toutes ces techniques visent à améliorer la qualité de l'image. Ainsi, les données de sortie sont l'image initiale et le résultat est également une image.

3- C'est quoi une image :

3.1 Définitions :

Une image est une représentation plate d'une scène ou d'un objet, généralement disposée dans un espace tridimensionnel. Mathématiquement, il s'agit d'une fonction bidimensionnelle de la forme $f(x, y)$. Où $f(x_0, y_0)$ est la valeur de f aux coordonnées (x_0, y_0) . Il s'agit donc d'un processus 2D continu issu de mesures physiques. L'amplitude de f est exprimée comme l'intensité de l'image (ou niveau de gris) au point de coordonnées (x, y) . Si l'image a été créée par un processus physique, $f(x, y)$ correspond à l'énergie émise par ce processus. Exemples de processus physiques : rayonnement et réflexion d'ondes lumineuses, infrarouges, rayons X, échos ultrasonores, etc...

3.2 Diverses catégories d'images :

- Images morphologiques : forme et placement des objets en 2D ou 3D
- Image volumétrique : Rendu volumétrique d'un objet ou d'une scène.
- Imagerie dynamique : Une série d'images (vidéo) en fonction du temps.
- Images fonctionnelles ou paramétriques : comportement des phénomènes Physiquement.
- Images multi spectrales : Acquises à différentes bandes de fréquence du rayonnement source.

4- Représentation des images :

L'image est une fonction d'au moins deux variables réelles :

Image : $f(x, y)$	image 2D
Volume : $f(x, y, z)$	image 3D
Séquence d'image : $f(x, y, t)$	image 4D
Séquence de volumes : $f(x, y, z, t)$	image 5D

La valeur de $f(.)$ est :

- Scalaire (intensité lumineuse)
- Images vectorielles (couleur RGB(Red Green Blue), ...), multi spectrales, images paramétriques...)
- Réel ou complexe

4.1 Types d'images

On peut classer les images numériques selon deux critères :

4.1.1. Selon la représentation numérique :

- **Images matricielles** : Jusqu'à présent, nous avons utilisé des matrices pour décrire des images. Cette image est appelée raster ou bitmap. Ce type d'image convient à l'affichage sur écran, mais est souvent de faible résolution et donc peu adapté à l'impression.
- **Images vectorielles** : Le principe des images vectorielles est de représenter des données d'image avec des formules mathématiques. Cela permet un agrandissement illimité de l'image sans perte de qualité et un faible encombrement.
- **Représentation spatiale** : Il est créé directement à partir d'un échantillon d'images dans le domaine spatial. Une véritable image scalaire 2D $f(x, y)$ peut être vue comme une surface 3D. De plus, cette expression est utilisée indépendamment des paramètres représentés par $f(.)$ (Température, pression, etc.). Ce que vous voyez lorsque vous regardez une image est la correspondance entre les niveaux de gris et les tailles physiques.

4.1.2. Selon la représentation des couleurs :

Il y'a plusieurs modes de codage des couleurs, L'espace colorimétrique le plus couramment utilisé pour le traitement d'image est le rouge, le vert et le bleu (RVB ou RGB).

4.1.2.1. LES MODELES DE COULEUR :

La couleur est l'information visuelle la plus utilisée pour la représentation d'image. Elle est le descripteur le plus utilisé dans les systèmes CBIR en raison de son invariance par rapport à l'échelle et sa richesse d'information. Il est possible d'extraire ce descripteur de l'image entière, ou juste d'une partie de l'image. Généralement, la couleur est représentée par un espace colorimétrique de trois composantes, il existe plusieurs manières de caractériser la couleur tels que l'histogramme, les moments couleur. [02]

C'est aussi l'une des composantes principales pour la description et l'analyse des images dans le domaine de la recherche d'images par le contenu.

Comme le démontre la figure 1 du spectre électromagnétique, les couleurs visibles par l'œil humain s'étendent entre les longueurs d'ondes allant de 4000 à 7000 angströms, représentant respectivement les couleurs : violet et rouge et toutes les couleurs entre les deux. Toutes les autres ondes s'étendant des rayons cosmiques des étoiles jusqu'aux ondes FM de nos radios ne peuvent pas être perçues par l'œil humain. Cette petite gamme du spectre est l'espace de couleur perceptible par l'œil humain. [03]

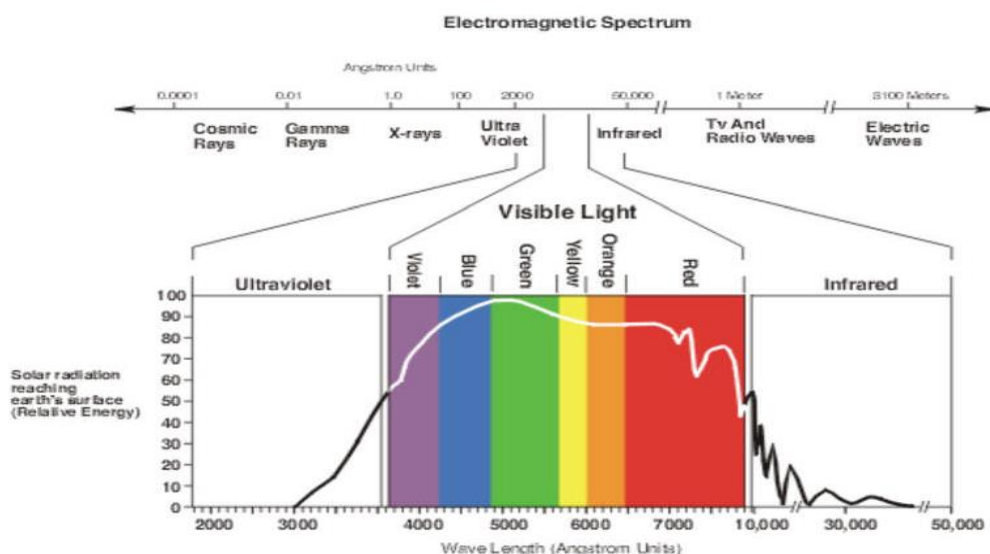


Figure 1. Le spectre électromagnétique

4.1.2.1.1. Définition Du Modèle De Couleur :

Un modèle de couleur, également connu sous le nom de modèle chromatique, est un modèle mathématique abstrait qui décrit comment les couleurs sont représentées sous forme de tuples de nombres, généralement composés de trois ou quatre valeurs ou composants de couleur. Le but du modèle de couleur est de faciliter la spécification des couleurs de manière standardisée. [04]

Chaque pixel de l'image peut être représenté comme un point dans l'espace 3D. Les espaces les plus couramment utilisés dans CBIR sont l'espace RVB (Rouge Vert Bleu), Cie $L^*a^*b^*$, CIE $L^*u^*v^*$, HSV (ou HSL, HSB). Les images sont généralement encodées avec trois canaux, mais un seul en monochrome. Son efficacité semble être liée au fait que les humains peuvent distinguer des milliers de couleurs et seulement 24 nuances de gris.

Certaines recherches ont été menées pour identifier des espaces colorimétriques plus sophistiqués. Par exemple, en projetant une image dans l'espace HSV, les informations de teinte, de saturation et d'intensité peuvent être séparées. Il a été démontré que la teinte est indépendante de la lumière et des conditions de prise de vue. D'autres espaces également courants dans le domaine, tels que CIE, XYZ et CIE-LUV, prétendent être perceptuellement

uniformes et indépendants de l'intensité. Encore une fois, ce sont des modèles d'affichage, il n'y a pas d'espace colorimétrique idéal. Comparez les espaces colorimétriques et leurs propriétés et analysez leurs avantages. [05]

4.1.2.1.2. Dans ce qui suit nous allons présenter quelques modèles de couleur :

A) LE MODÈLE RVB (en anglais RGB) :

Le système le plus couramment utilisé est le système RVB, qui est un système à trois couleurs. Attribuez à chaque couleur trois composants (ou canaux) correspondant à l'intensité de chacune des trois couleurs primaires du composite Additif. Le blanc correspond à la valeur maximale de chaque canal et le noir correspond aux trois composantes nulles. (Fig. 02)

L'objectif principal du modèle de couleur RVB est la détection, la représentation et l'affichage d'images dans des systèmes électroniques, tels que les téléviseurs et les ordinateurs, bien qu'il ait également été utilisé dans la photographie conventionnelle. Avant l'ère électronique...

Le modèle RVB attribue à chaque couleur une valeur numérique comprise entre 0 et 255, qui représentent l'intensité de chaque couleur de la combinaison. Par exemple, la combinaison du rouge, du vert et du bleu à leur intensité maximale (255, 255, 255) produit du blanc, tandis que la combinaison du rouge, du vert et du bleu à leur intensité minimale (0, 0, 0) produit du noir.

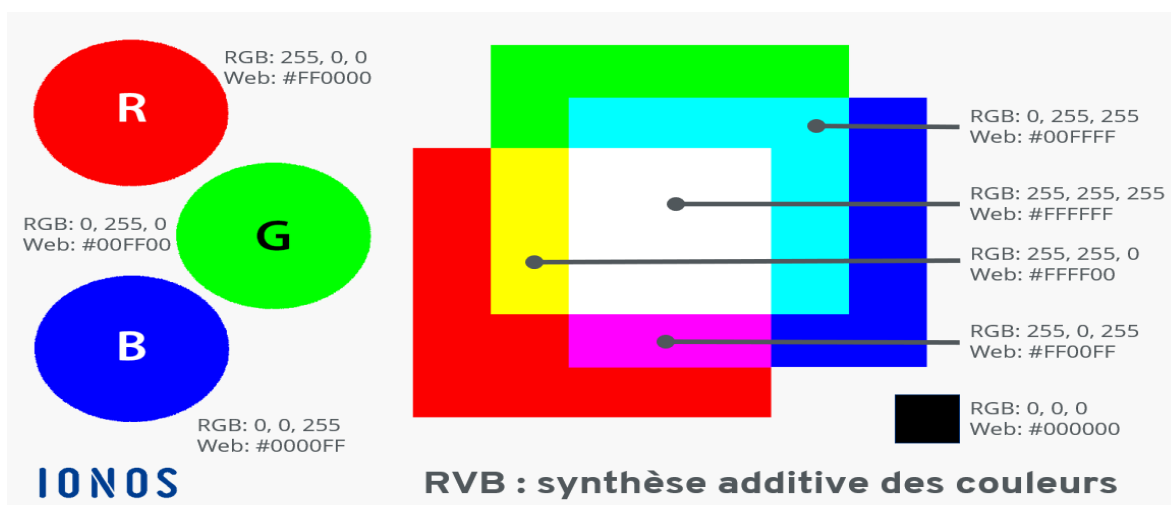


Figure 2. Les trois couleurs primaires additives de lumière : Rouge, Vert, et Bleu

B) LE MODÈLE CMJ (CYAN, MAGENTA, JAUNE) :

Ce modèle est basé sur les couleurs secondaires du modèle RVB. Le modèle de couleur RVB (RVB) est un modèle de couleur additif, de sorte que les modèles de couleur CMJ (CMJ français) et CMJN (CMJN français) sont basés sur un mélange de couleurs soustractif. Les lettres représentent les composants cyan (C), magenta (M), jaune (Y = Yalou) et clé (K, noir/clé). Le cyan est un type de bleu et le magenta est un type de rouge. Voir Noms des couleurs pour plus de détails. Cet espace colorimétrique est la base de l'impression moderne

en quadrichromie et est donc utilisé, par exemple, dans le domaine des imprimantes PC et pour l'impression de magazines et de livres de couleurs.



Figure 3. Les couleurs primaires de lumière et leur complémentaires (couleurs secondaires)

C) LE MODÈLE TSL (TEINTE, SATURATION, LUMINOSITÉ) :

Le modèle TSL (HSL en anglais pour Hue, Saturation et Lightness) est le plus intuitif de tous les modèles de couleurs que nous allons considérer. C'est ce qu'on appelle un modèle perceptif parce qu'il est basé sur la perception humaine. Chaque norme de couleur est clairement séparée, ce qui en fait le modèle le plus pratique pour l'édition d'images et la correspondance des couleurs.

Les trois critères qui caractérisent le TSL mesurée par un angle de 0 à 360° autour de la roue chromatique, la saturation, qui reflète bien la notion intuitive de coloration, car elle va des couleurs vives vers le gris. Et enfin, la luminosité, mesurée entre le noir (pas de lumière ou valeur 0) et le blanc (lumière maximum ou valeur 1).

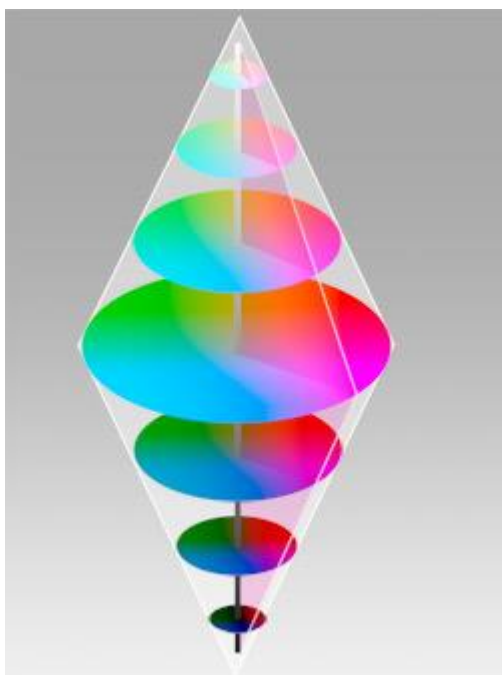


Figure 4. Le modèle TSL

Le modèle TSL a l'avantage de définir Clairement l'échelle des luminosités entre le noir et le blanc.

Au-dessus du cercle chromatique, les couleurs sont éclaircies (tramage CMJN), ce qui correspond à l'ajout de blanc.

Au-dessous du cercle chromatique, les couleurs s'assombrissent par le mélange avec d'autres couleurs.

D) LES MODÈLES DE LA CIE (COMMISSION INTERNATIONALE DE L'ÉCLAIRAGE) :

Pour approfondir notre compréhension de la couleur, la CIE (Commission Internationale de l'Eclairage), une institution dédiée à tout ce qui concerne la lumière, l'illumination, la couleur et l'espace colorimétrique, a créé plusieurs espaces colorimétriques à partir des années 1930 pour représenter le spectre apparaître. Cela nous a permis de comparer les différents espaces colorimétriques des différents écrans et périphériques du marché.

CIE a mené une série de tests sur un grand nombre de sujets afin de déterminer la L'observateur humain moyen hypothétique et sa réaction à la couleur, étant un profil donné Comme un 'observateur référencé'.

L'œil humain possède trois types de capteurs de couleur qui réagissent à des gammes de longueurs d'onde différentes, un diagramme en trois dimensions serait nécessaire pour obtenir une représentation complète de toutes les couleurs visibles.

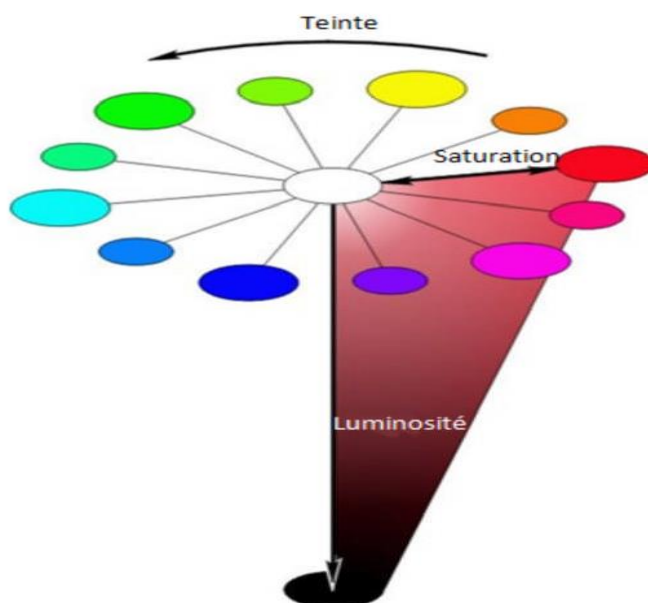


Figure 5. La teinte, la saturation, et luminosité représentée sur une roue de couleurs sous d'un cône.

E) LE MODELE CIE XYZ :

Le système XYZ a été établi par la CIE afin de pallier certains inconvénients du système RVB. Ce correspond à un changement de couleurs primaires et s'obtient simplement à partir du Système RVB utilisant une matrice de passage. Les coefficients de cette matrice sont déterminés par rapport à un blanc de référence appelé illuminant. Celui que nous utilisons s'appelle illuminant standard. Il existe ainsi plusieurs codages XYZ, qui dépendent de l'illuminant choisi.

Le passage du système RGB que nous utilisons (tous les canaux sont codés par des entiers entre 0 et 255) au système XYZ se fait par le calcul matriciel suivant :

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0.430574 & 0.341550 & 0.178325 \\ 0.222015 & 0.706655 & 0.071330 \\ 0.020183 & 0.129553 & 0.939180 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

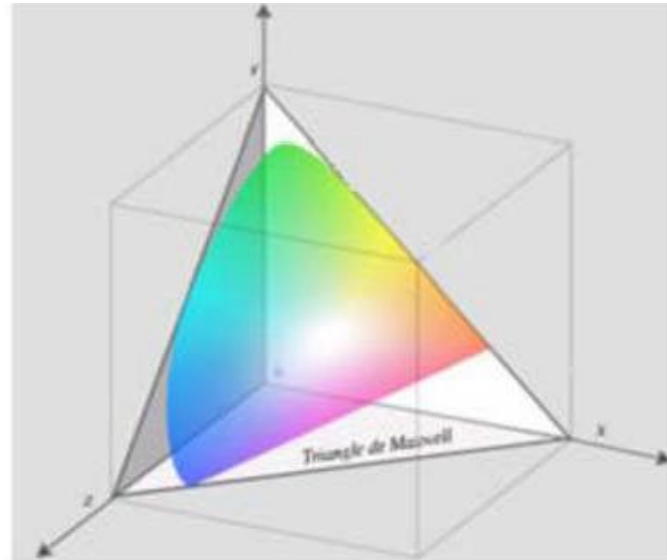


Figure 6. Représentation du modèle CIE XYZ par le triangle de Maxwell

F) LE MODÈLE LAB :

Le modèle chromatique Lab (**Figure 7**) a été développé par la CIE en 1976 dans le but d'améliorer la représentation des couleurs, elle est décrite mathématiquement toutes les couleurs perçues dans les trois dimensions L pour la luminance et a et b pour les composantes de couleur verte, rouge, bleue et jaune. "Lab" est fréquemment utilisé comme abréviation pour l'espace colorimétrique CIE L *a*b*1976 (également CIELAB).

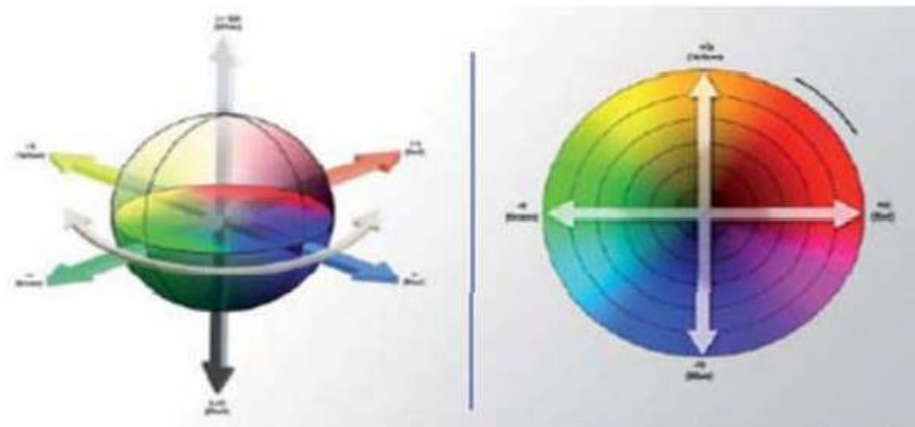


Figure 7. Le modèle CIE LAB.

4.1.2.2. LES TYPE D'IMAGES :

On distingue généralement les différents types d'images suivants :

A) Images en teintes (ou niveaux) de gris :

Une matrice dans laquelle chaque élément est un réel, A ce niveau, l'intensité de la lumière est codée, généralement sur un octet (256 valeurs). Par convention, une valeur de zéro représente le noir (intensité lumineuse nulle) et la Valeur 255 le blanc (intensité lumineuse maximale) :

000	008	016	024	032	040	048	056	064	072	080	088	096	104	112	120	128
255	248	240	232	224	216	208	200	192	184	176	168	160	152	144	136	

Ce procédé est utilisé pour reproduire des images ou du texte en noir et blanc sous certaines conditions.

B) Images à palettes, images en 256 couleurs (8 bits) :

Pour réduire l'espace occupé par les informations de couleur, une palette de couleurs "connectée" à l'image est utilisée. On parle alors de couleurs indexées : la valeur associée à un pixel ne véhicule plus la couleur réelle du pixel, mais pointe sur l'entrée correspondant à cette valeur dans une table de couleurs appelée table de correspondance ou LUT en Anglais, où nous avons à l'esprit la représentation complète de la couleur.

Selon le nombre de couleurs présentes dans l'image, il est ainsi possible de gagner un espace non négligeable : en pratique on considère que 256 des 16 millions de couleurs 24 bits suffisent. Pour l'encoder, on aura donc une carte qui occupe 24 bits x 256 entrées, soit 3 x 256 octets, et les pixels de l'image seront associés à des indices encodés sur un seul octet. Ainsi, l'occupation d'une telle image est de 1 octet par pixel.

C) Images 24 bits (ou « couleurs vraies ») :

C'est un nom trompeur car le monde numérique (fini, limité) ne peut pas expliquer entièrement la réalité (infinie). Le codage couleur s'effectue sur trois octets, chaque octet représentant la valeur de la composante couleur par un entier de 0 à 255. Ces trois valeurs symbolisent généralement la couleur dans l'espace RVB. Le nombre de couleurs différentes pouvant ainsi être représentées est de 256 x 256 x 256 possibilités, soit environ 16 millions de couleurs. Etant donné que la différence de teinte entre deux couleurs très proches mais différentes dans ce mode de représentation est quasiment imperceptible à

l'œil humain, il convient de considérer que ce système permet une reproduction fidèle des couleurs, c'est pourquoi on parle de "vraies couleurs".

R	V	B	Couleur
0	0	0	noir
0	0	1	nuance de noir
255	0	0	Rouge
0	255	0	Vert
0	0	255	Bleu
128	128	128	Gris
255	255	255	Blanc

Tableau 1. RGB

Les bitmaps basés sur cette représentation peuvent rapidement occuper beaucoup d'espace de stockage, car chaque pixel nécessite trois octets pour coder sa couleur.

4.2 Formats d'images :

Le terme « image numérique » fait référence à toute image (graphique, icône, photographie, etc.) capturée, créée, manipulée et stockée au format binaire. Un format d'image est donc une représentation informatique d'une image associée à des informations qui donnent une idée de la façon dont l'image est encodée et éventuellement comment elle est décodée et traitée (métadonnées). Le format consiste en un en-tête contenant des attributs (taille de l'image, type d'encodage, LUT, etc.) suivis de données (l'image elle-même). Les attributs et la structure des données sont différents pour chaque format d'image.

De plus, les formats existants incluent souvent une zone de métadonnées utilisée pour spécifier des informations sur l'image. Par exemple. :

- Date, heure et lieu de tournage.
- Les caractéristiques physiques de l'image (sensibilité ISO, vitesse d'obturation, utiliser le flash...)

	Compression des données	Nombre de couleurs supportées	Affichage progressif	Animation	Transparence
JPEG	Oui, réglable (avec perte)	16 millions	Oui	Non	Non
JPEG2000	Oui, avec ou sans perte	16 millions	Oui	Non	Non
GIF	Oui, Sans perte	256 maxi (palette)	Oui	Oui	Oui
PNG	Oui, sans perte	Palettisé (256 couleurs ou moins) ou 16 millions	Oui	Non	Oui (couche Alpha)
TIFF	Compression ou pas avec ou sans pertes	de monochrome à 16 millions	Non	Non	Non
SVG	compression possible	16 millions	* ne s'applique pas *	Oui	Oui (par nature)

Tableau 2. Tableau comparatif de différents formats d'images

[S1]

5- Caractéristiques d'images :

Les images numériques sont également définies par des propriétés, Les propriétés sont dites globales, permettant une meilleure Pour calculer certaines des caractéristiques visuelles de l'image, il est utilisé Pour le post-traitement dans la gamme d'applications telles que la détection d'objet ou de la recherche d'images par contenu par exemple.

Caractéristiques générales d'une image numérique :

5.1. Pixel :

Pixel vient d'un acronyme pour l'expression britannique PICture élément (souvent abrégé pixel). C'est l'unité de mesure de base définition de l'image numérique bitmap.

Tous ces pixels sont dans une matrice bidimensionnelle l'image finalement obtenue forme une couleur associée à chacune d'elles pixels. Les pixels sont si petits que vous pouvez à peine les voir à l'œil nu, vous pouvez donc beaucoup d'exposition et une image nette.

5.2. La dimension :

La dimension ou la définition d'une image numérique est un nombre de la somme des pixels qui composent cette image. Calculez simplement le produit : la largeur [px] * hauteur [px] pour spécifier la dimension. Par exemple : une image Largeur = Hauteur = 8px, Dimension = $8 * 8 = 64px$.

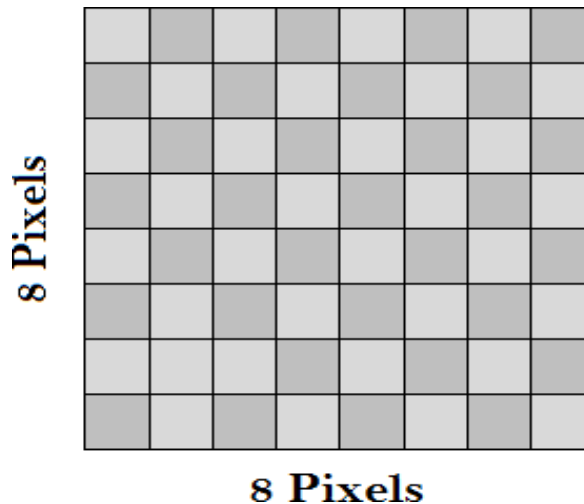


Figure 8. La dimension d'une image.

5.3. La résolution :

La résolution est la densité d'une image en pixels une fois qu'elle a été reproduite, Mesuré en "ppi" (pixels par inch) ou en français "ppp" (pixels par pouce). Décrit la clarté ou la finesse des détails d'un bitmap. Plus le nombre de pixels par pouce est élevé, plus la résolution est élevée. En général, l'image haute résolution fait bonne impression de meilleure qualité.

5.4. La couleur :

La couleur est la perception visuelle que nous avons de la différence Les longueurs d'onde qui composent la lumière visible. Cette partie du spectre La lumière va du violet (4000 angströms) au rouge (7000 angströms), D'autres ne voient pas.

Dans les systèmes de recherche d'images basés sur le contenu, la couleur Informations visuelles les plus couramment utilisées en raison de leur stabilité par rapport à largement. La couleur est généralement représentée par trois composantes. Ces composantes définissent l'espace colorimétrique. Il y a beaucoup Les espaces colorimétriques ont chacun des propriétés spécifiques Fait intéressant, le plus largement utilisé est le système RVB.

5.5. Contours et Textures :

Un contour représente la limite entre les objets de l'image ou la limite entre deux pixels. [06] La texture est caractérisée par un motif répétitif ou certains articles. Plus précisément, la texture peut être considérée comme une bobine, Une collection de pixels (niveaux de gris) disposés spatialement les uns par rapport aux autres nombres de relations spatiales, créant ainsi une zone homogène. [07]



Figure 9. Exemples de textures.

5.6. Luminance et Contraste :

➤ **Luminance :**

C'est la luminosité des points de l'image. Il est également défini : il est obtenu en divisant l'intensité d'éclairement d'une surface par la surface claire de cette surface. Pour un observateur éloigné, le mot *luminance* Le mot *brillance*, qui correspond à la brillance d'un objet, a été remplacé. [08]

➤ **Contraste :**

C'est l'écart observé entre deux zones de l'image, plus précisément entre les zones sombres et les zones claires de cette image. Le contraste est déterminé en fonction de la luminosité de deux zones d'image (distinguer les couleurs claires des couleurs foncées).

5.7. Bruit :

Le bruit de l'image correspond à des variations indésirables de couleur ou de luminosité qui masquent les détails de l'image et dégradent la qualité de l'image.

Le bruit a de nombreuses causes :

- ✓ L'environnement capturé.
- ✓ Qualité du capteur.
- ✓ Qualité de numérisation.

6. Traitement d'image numérique :

Le traitement d'image est un ensemble d'opérations qui permettent Rehaussement (filtrage, rehaussement du contraste), réglage (rotation, analogie) et extraire des informations à partir d'images numériques. [09]

D'autres étapes sont nécessaires avant de commencer le traitement d'image tel que :

- L'acquisition et la numérisation de l'image par des dispositifs comme les scanners, les appareils photo qui permettent d'effectuer l'échantillonnage et la quantification d'une image.
- Il est facile d'effectuer les différentes opérations de traitement d'images grâce à des outils informatiques qui sont installés sur une unité centrale.
- Le stockage de l'image numérisée sur un support informatique (disque dur, CD-ROM, Clé USB, etc.).
- L'affichage grâce à des dispositifs de visualisation qui permet l'affichage des images traitées.

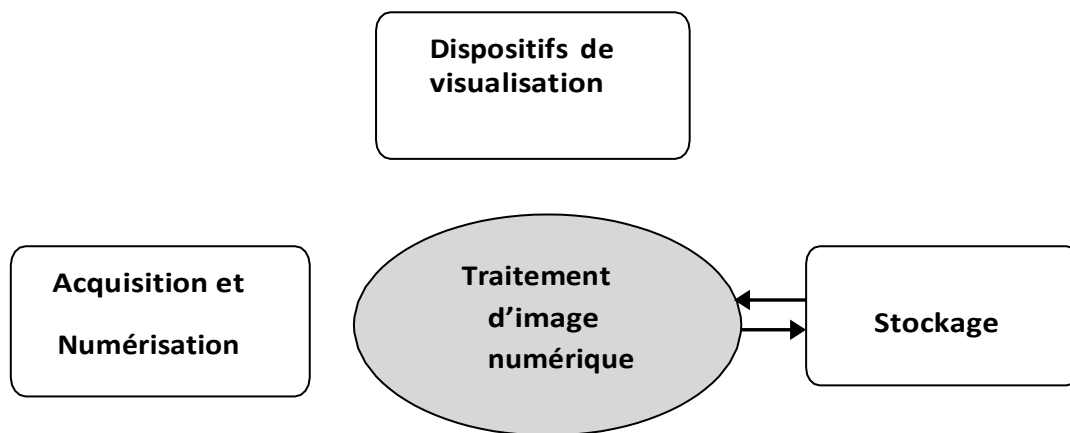


Figure 10. Système de traitement d'image.

7. Filtrage :

Présenté ici le Principe du filtre et différents types de filtres.

7.1. Définition :

Les images numériques capturées sont souvent inutilisables pour le traitement d'image. Ils contiennent des signaux bruyants. Par conséquent, il est souvent nécessaire d'améliorer sa qualité.

7.2. Les types de filtrage :

Il existe deux manières de concevoir un filtre d'image. La première consiste à construire un travailleur qui fonctionne d'une manière ou d'une autre linéaire et fixe dans le décalage. Ces filtres au sens littéral du terme, Ils sont appelés filtres linéaires.

L'alternative est de ne pas forcer le candidat à procéder de manière linéaire ou statique dans le quart de travail. On peut intégrer un très grand nombre de facteurs dans cette hypothèse que nous regroupons sous le terme de filtres non linéaires [10]

7.2.1. Filtres linéaires :

Les premières et les plus simples méthodes de filtrage sont basées sur le filtrage linéaire, Le filtre linéaire convertit l'ensemble de données d'entrée dans un fichier Un ensemble de données de sortie selon une opération arithmétique appelée convolution. Lorsqu'il s'agit de données numériques comme dans le cas de traitement d'image, la relation entre les valeurs de pixel de sortie et celle les pixels d'entrée sont décrits par un tableau de nombres, généralement Carré appelé matrice de convolution. [02]

Les différents types des filtres linéaires sont :

➤ Filtre gaussien :

L'expression gaussienne en deux dimensions est donnée par :

$$G_0(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (1.1)$$

Où x et y sont les distances sur les deux axes, et le biais Critère de distribution gaussien. L'intérêt de ce candidat est que nous Contrôle facilement le degré de filtrage à travers le paramètre σ . Dans Numérique, le filtre est représenté mathématiquement par une matrice. La convolution qui permet de manipuler l'image est mise en œuvre de manière très simple. Le pixel à traiter est placé au centre du filtre, on multiplie Filtrer les coefficients en fonction des valeurs de pixel correspondantes et les calculer moyenne pondérée. Appréciation pour ce candidat Pour σ égal à 0,6 le masque suivant donne :

	1	2	1
1/16*	2	4	2
	1	2	1

Masque du filtre

➤ Filtre passe-bas :

En traitement d'image, des filtres passe-bas sont utilisés pour supprimer ces hautes fréquences (grands changements d'intensité dans l'image).

Le filtre de moyenne est une sorte de filtre passe-bas, son principe est de faire la moyenne des valeurs des pixels adjacents. [02]

$1/9^*$

1	1	1
1	1	1
1	1	1

Masque du filtre

➤ Filtre passe-haut :

C'est l'opposé des filtres passe-bas, et le but de ce filtre est de ressortir caractéristiques et les extraire en amplifiant les détails haute fréquence (Changements rapides d'intensité de l'image) laissant les basses fréquences (zones uniformes inchangées). Filtre amélioré sur les bords c'est une sorte de filtre passe-haut dont le principe est de faire ressortir les détails et antithèse. [02]

-1	-1	-1
-1	9	-1
-1	-1	-1

Masque du filtre

7.2.2. Filtres non linéaires :

Il a été développé pour ajuster les filtres linéaires insuffisants Les contours sont particulièrement mal conservés mais ils ont un inconvénient Pour imposer des distorsions irréversibles à l'image. Leur principe est le même Ce sont des filtres linéaires.

➤ Filtre médian :

Il appartient à la famille des filtres d'ordre. Principe de ce filtre remplacez chaque valeur de pixel central par la valeur de pixel central observé dans leur voisinage.

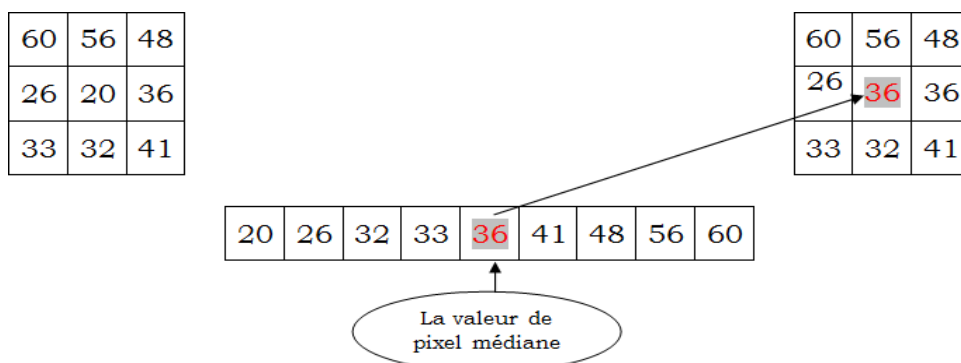


Figure 11. Principe du filtre médian

➤ Filtre maximum :

On applique le même principe que celui du filtre médian sauf que la valeur du pixel du centre va être remplacée par la valeur maximale.

➤ Filtre minimum :

Même principe mais 'on va remplacer la valeur du pixel du centre par la valeur minimale.

8. Segmentation :

La segmentation est un processus et une étape importante du traitement d'image visant à regrouper les pixels selon des critères prédéfinis. Les pixels sont donc regroupés en régions qui forment des pavages ou des partitions de l'image.

On considère principalement de nombreux types de segmentation, que l'on peut regrouper en quatre principales classes :

- Segmentation par régions.
- Segmentation par détection de contour.
- Segmentation par seuillage.
- Segmentation fondée sur la combinaison entre les trois premières segmentations.

9. Quelques applications concrètes de traitement d'images :

- Contrôle de présence/absence.
- Examinez le degré de maturité des fruits sur la chaîne de conditionnement.
- Construire et corriger des cartes géographiques à partir de satellites ou d'images photos aériennes.
- Reconnaissance de l'écriture.

10. Conclusion :

Les images sont le moyen le plus efficace de communiquer le traitement numérique des images et toutes les méthodes et techniques qui y sont appliquées, rendant cette opération possible, plus facile, plus efficace et plus efficace.

Ce chapitre a introduit la recherche sur le traitement d'images en général. Ce chapitre a d'abord traité des représentations d'images et des différents types d'images. Ensuite, les propriétés de l'image, le traitement numérique de l'image, le filtrage, la segmentation, etc. A la fin de ce chapitre, nous avons présenté quelques applications spécifiques du traitement d'images. Comme vous pouvez le voir, le traitement d'image s'est imposé dans tous les domaines de notre vie quotidienne. Par conséquent, il attire beaucoup d'attention et devient une direction de recherche très active dans tous ses aspects.

Dans le chapitre suivant nous présentons les notions de bases sur les systèmes CBIR.

CHAPITRE 2 : SYSTÈMES DE RECHERCHE D'IMAGE PAR CONTENU (CBIR)



1. Introduction :

La **recherche d'images par le contenu** (Content Based Image Retrieval) (CBIR), est une technique visant à effectuer des recherches d'images à l'aide de requêtes portant sur les caractéristiques visuelles d'une image telle que la texture, la couleur, la forme... Typiquement, ces systèmes sont utilisés lorsque l'on dispose d'une image pour laquelle on souhaiterait obtenir des images similaires visuellement. Il s'oppose à la recherche d'images par mots clés, qui est le cas typique proposé actuellement par les moteurs de recherche tels que Google ou Yahoo.

Cette technologie se distingue des moteurs de recherche traditionnels d'image. Ces derniers utilisent généralement deux techniques :

-Une recherche contextuelle par mots clés (tels que Google ou Altavista) : l'image est retrouvée à partir des mots clés entourant cette image dans les pages web.

Une recherche par mots clés manuellement associés aux images indexées, par exemple dans les musées.

Les CBIR tentent, à l'inverse, de permettre une indexation et une recherche de l'image portant sur les caractéristiques de l'image :

- La texture (filtre de Gabor, transformée en ondelette discrète, ...)
- La couleur (histogramme de couleurs, histogrammes dans l'espace RGB, HSV, ...),
- Les formes (descripteurs de Fourier, ...),
- Une combinaison de plusieurs de ces caractéristiques.

Ces caractéristiques sont dites de bas-niveau, car elles sont très proches du signal, et ne véhiculent pas de sémantique particulière sur l'image. Une fois ces caractéristiques extraites, la suite consiste généralement à définir diverses distances entre ces caractéristiques, et de définir une mesure de similarité globale entre deux images. Armés de cette mesure de similarité et d'une image requête, on peut alors calculer l'ensemble des mesures de similarité entre cette image requête et l'ensemble des images de la base d'images. On peut alors ordonner les images de la base suivant leur score, et présenter le résultat à l'utilisateur, les images de plus grand score étant considérées comme les plus similaires.

Du fait des caractéristiques calculées, qui sont de *bas-niveau*, ces techniques obtiennent des résultats satisfaisant pour certains types de requêtes et certains types de base d'images. Par exemple rechercher des images de paysages enneigés, parmi une base d'images de paysages.

Toutefois ces systèmes rendent souvent des réponses extravagantes, et souvent éloignées de l'idée qu'avait l'utilisateur lorsqu'il a soumis sa requête.

2. Différents domaines d'application :

- Archivage, gestion des bases de données de contenu (personnelles ou professionnelles).
- Reconnaissance automatique du visage, d'empreinte humaine, d'objets, caractères, ...
- Détection de plagiat.
- Protection du copyright.
- Applications judiciaires : recherche de suspects.
- Applications médicales.
- Applications industrielles.
- Commerce électronique.
- Systèmes d'informations géographiques.
- Moteur de recherche ...

3. Composants d'un CBIR :

Nous décrivons brièvement ici les caractéristiques communes à la plupart des approches : le traitement de la base d'images, les requêtes puis la mise en correspondance et la présentation des résultats. La Figure 12 illustre l'ordonnancement de ces étapes.

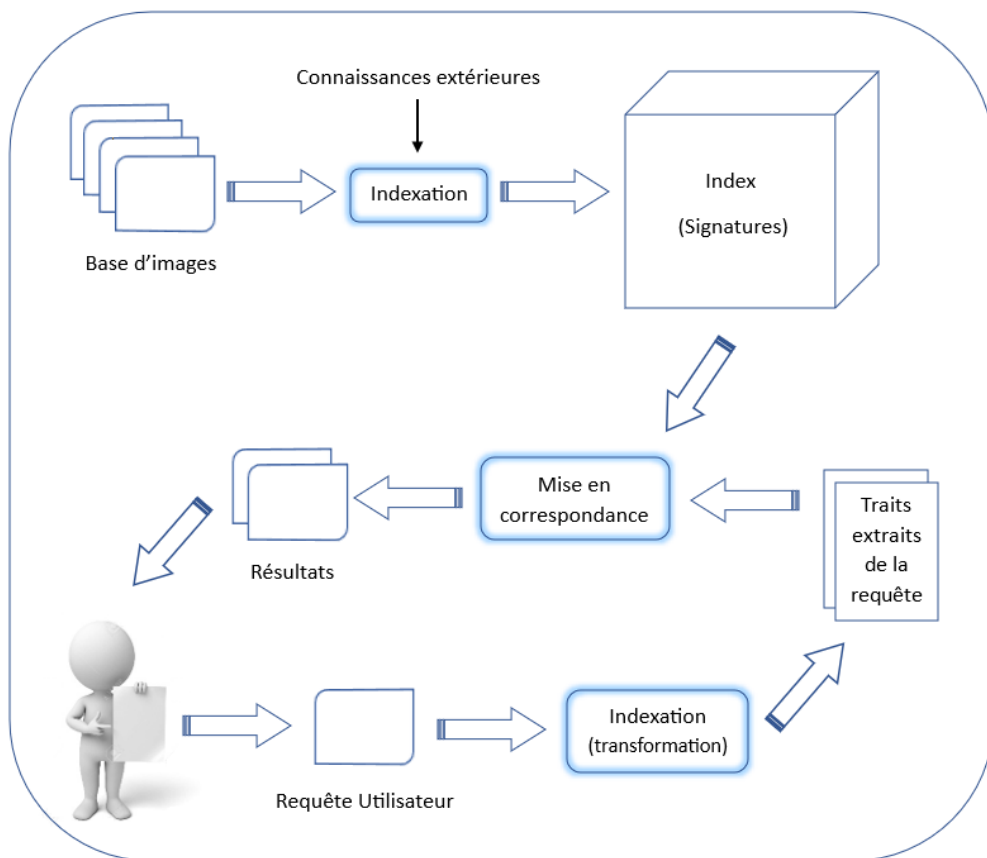


Figure 12. Principaux composants d'un système de recherche par le contenu

Dans un premier temps, des descripteurs sont calculés à partir de chaque image de la collection, ils peuvent être de type signal ou/et symbolique (le vocabulaire d'indexation). Les données extraites (à présent représentatives du contenu de l'image du point de vue du système) constituent la base d'index. Les requêtes de l'utilisateur sont alors transformées afin d'être comparables avec la base d'index ; une mise en correspondance entre la requête transformée et la base d'index permet ensuite de produire le résultat de la requête. Il se peut également que le système possède des composantes liées à la personnalisation, comme par exemple l'extraction, le stockage et l'utilisation d'un profil d'utilisateur.

3.1. La base d'image :

La collection (ou base) d'images est la donnée principale du système. Les bases d'images varient d'abord par leur taille : la majorité des systèmes est conçue pour des bases de quelques centaines ou milliers d'images ([11, 12, 13]). Ce nombre peut cependant s'approcher du milliard (880 millions d'images) lorsque la base est constituée par les images collectées par des robots sur Internet. La taille de la base d'image impose des contraintes sur la complexité des traitements effectués sur chaque image. Il en résulte que la comparaison qualitative entre des systèmes travaillant sur des bases de tailles très différentes est peu pertinente.

Le type d'image composant la base varie également : des portraits en noir et blanc [14], des peintures chinoises anciennes [13], des images personnelles [11], des images de tissus humains [15], etc. Le type d'image influe fortement sur la conception globale du système, particulièrement sur les descripteurs de bas niveau calculés. D'une manière générale, plus la variabilité intra et inter images est importante, plus le système doit être riche et précis (et plus le problème d'indexer/rechercher ces images est difficile).

Finalement, les collections diffèrent par leur stabilité, c'est-à-dire le taux de changements (ajouts d'images, retraits, etc.) en fonction du temps. Faible pour une collection d'images représentant les œuvres d'un peintre ne créant plus, elle peut être très forte lorsque, par exemple, on s'intéresse aux images de la Toile ou à l'actualité.

3.2. L'index :

L'indexation a pour but de substituer à une image un représentant (ou descripteur) moins encombrant qui la caractérise le mieux possible et de ne travailler que sur ce modèle lors de la recherche. Cela permettra une meilleure organisation des données, de limiter la quantité de données examinées durant une recherche, d'y accéder rapidement et de confiner la recherche au maximum.

Un système d'indexation comprend généralement deux phases de traitement :

Indexation logique : L'indexation logique consiste à extraire et à modéliser les caractéristiques de l'image qui sont principalement la forme, la couleur et la texture. Chacune de ces caractéristiques pouvant être considérée pour une image entière ou pour une région de l'image.

Indexation physique : L'indexation physique consiste à déterminer une structure efficace d'accès aux données pour trouver rapidement une information. De nombreuses techniques basées sur des arbres (arbre-B, arbre-R, arbre quaternaire, ...) ont été proposées.

Pour qu'un système de recherche d'images soit performant, il faut que l'indexation logique soit pertinente et que l'indexation physique permette un accès rapide aux documents recherchés.

3.3. La gestion des index :

Elle concerne la manière dont sont gérés les index des images : stockage et accès. La gestion des index, anecdotique pour une collection de taille modeste, devient une préoccupation essentielle lorsque l'on travaille sur une base de taille conséquente. La manière la plus basique de stocker les index est la liste séquentielle, que ce soit en mémoire ou dans un fichier. Cependant, lorsque le nombre d'images augmente, le temps d'accès à une image augmente linéairement et il est souvent nécessaire d'organiser les index de manière hiérarchique, sous forme d'arbres (organisés selon les descripteurs), ou de tables de « hash-code » par exemple, afin d'accélérer l'accès à l'information.

3.4. Les requêtes :

Le type de requête proposé découle de choix fait en amont, au niveau de l'indexation. Dans des systèmes où seuls des descripteurs de bas niveau sont extraits, les requêtes ne peuvent être que de bas niveau : requête par « image exemple », par croquis ou par manipulation directe des traits de bas niveau. Dans ces systèmes, des descripteurs sont extraits à partir de la requête (une image, un croquis...) et sont comparés aux descripteurs calculés à partir des images de la base (les index des images).

Contrairement aux systèmes qui proposent plus d'abstraction, les requêtes peuvent être sémantiques (textuelles par exemple). Ce qui permet à l'utilisateur de formuler des requêtes sémantiques (« Je veux des images prises à l'extérieur. »).

3.4.1. Les types des requêtes :

Il existe 3 façons de faire une requête dans un système d'indexation et recherche des images : soit une requête par mots clés, soit une requête par esquisse, soit une requête par exemple. La figure 13 donne une démonstration pour les trois façons.

3.4.1.1. Requête par exemple : Pour les systèmes de recherche d'images à base d'exemples, l'utilisateur, pour représenter ses besoins, utilise une image (ou une partie d'image) qu'il considère similaire aux images qu'il recherche. Cette image est appelée image exemple ou requête. L'image exemple peut soit être fournie par l'utilisateur, soit être choisie par ce dernier dans la base d'images utilisée. Cette technique est simple et ne nécessite pas de connaissances approfondies pour manipuler le système.

3.4.1.2. Requête par mot clé : le système doit extraire des mots clés à partir d'une image. Afin de faire ce travail le système se base sur l'annotation textuelle manuelle d'images. Cette façon est aujourd'hui la plus employée. Beaucoup de moteurs de recherche d'images tels que Google, Yahoo...utilisent cette façon. Mais cette façon n'est pas parfaite parce que quelques mots n'expriment pas le sens d'une image.

3.4.1.3. Requête par esquisse : Dans ce cas, le système fournit à l'utilisateur des outils lui permettant de constituer une esquisse (dessin) correspondant à ses besoins. L'esquisse fournie sera utilisée comme exemple pour la recherche. L'esquisse peut être une ébauche de forme ou contour d'une image entière ou une ébauche des couleurs ou textures des régions d'une image. L'utilisateur choisira, en fonction de la base d'images utilisée, de ses besoins et préférences, l'une ou l'autre de ces représentations. Cette technique présente l'inconvénient majeur qu'il est parfois difficile pour l'utilisateur de fournir une esquisse, malgré les outils qui lui sont fournis.

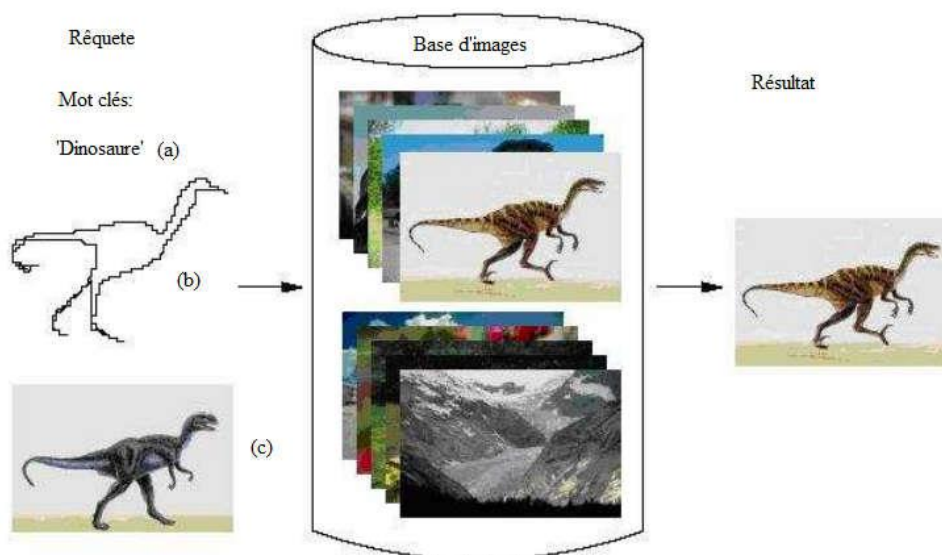


Figure 13. Les trois façons pour donner une requête dans CBIR (Deselaers, 2003)

3.5. L'analyse de la requête :

Cette étape a pour but de transformer la requête utilisateur pour la rendre comparable avec les index de la base d'images ; elle consiste donc généralement à extraire les mêmes types de descripteurs que ceux extraits de la base d'image lors de l'indexation.

3.6. Mise en correspondance requête/base :

Il s'agit d'estimer dans quelle mesure une image (son index) satisfait une requête donnée.

Dans le contexte de la recherche d'images, cela s'attribue souvent à calculer la similarité entre Traits extraits de la requête et les caractéristiques de chaque image de la base. Cela renvoie une valeur de correspondance qui caractérise la pertinence (du point de vue du système) d'une image par rapport à la requête.

La phase de mise en correspondance peut également inclure une pondération des descripteurs (comme dans [11] où chaque descripteur est pondéré par rapport à son pouvoir discriminant dans la base). Pondérer les descripteurs permet d'éliminer une partie du bruit dans la mesure où les descripteurs les moins pertinents voient leur influence diminuer dans l'évaluation de la similarité requête/image.

3.7. La présentation des résultats :

Généralement et dans la majorité des systèmes disponibles, le résultat d'une requête est présenté sous forme d'une liste d'image (avec petite taille), ordonnées par ordre décroissant de la pertinence.

4. Représentation des images dans un CBIR :

Dans la majeure partie des systèmes existant, les images sont représentées avec des descripteurs de bas niveau, i.e., en termes de couleur, texture, formes, points d'intérêt ou par des descripteurs hybrides.

5. Mesures pour évaluer un système :

La recherche d'images par le contenu est un sous-domaine de la recherche d'informations (RI). Les mesures d'évaluation adoptées ont donc logiquement été reprises de la RI classique. Les deux informations largement utilisées sont le rappel et la précision. Ces paramètres mesurent la pertinence d'un système RI. Autrement dit, ils mesurent la concordance des informations retournées relatives à la requête. La figure 14 montre une illustration de la précision et du rappel.

5.1. Rappel et précision :

Deux valeurs pour évaluer le système de recherche par contenu :

5.1.1. Le rappel (Recall) :

Le rappel est le rapport entre le nombre d'images pertinentes dans l'ensemble des images trouvées et le nombre d'images pertinentes dans la base d'images.

Rappel = (Nombre des images pertinents sélectionnés) / (Nombre total des images pertinents)
[S5]

5.1.2. La précision (Precision) :

La précision est le rapport entre le nombre d'images pertinentes dans l'ensemble des images trouvées et le nombre d'images trouvées.

Précision = (Nombre des images pertinents sélectionnés) / (Nombre total des images sélectionnés)

Les deux métriques rappels et précision s'utilisent conjointement pour l'évaluation des performances des systèmes de recherche d'information et varient inversement : lorsque la

précision diminue, le rappel augmente et réciproquement. Les valeurs de ces deux métriques reflètent le point de vue de l'utilisateur :

- si le rappel est faible, une partie de l'information pertinente ne lui sera pas accessible ;
- si la précision est faible, l'utilisateur ne sera pas satisfait à cause de la forte concentration des informations non-pertinentes fournies dans les résultats.

Dans les deux cas, le système ne répond pas aux attentes des utilisateurs à retourner l'information utile et pertinente et par la suite, il est non performant. Le cas idéal est d'avoir la valeur de précision et rappel respectivement égale à un, chose qui n'est pas atteinte en réalité. Une des manières de tenir compte à la fois du rappel et de la précision d'un système est d'exprimer les valeurs de précision en fonction des différents niveaux de rappel selon la courbe rappel/précision.

5.2. La courbe de rappel et précision :

Les deux valeurs du rappel et précision sont très importantes pour l'évaluation du système mais pour la performance du système cette paire n'est pas suffisante. Pour cette raison on représente cette paire de valeur sous forme d'une courbe rappel/précision.

On le dessine avec un calcul de plusieurs paires rappel/précision et les lie ensemble. La figure ci-dessous montre un exemple de courbe rappel/précision.

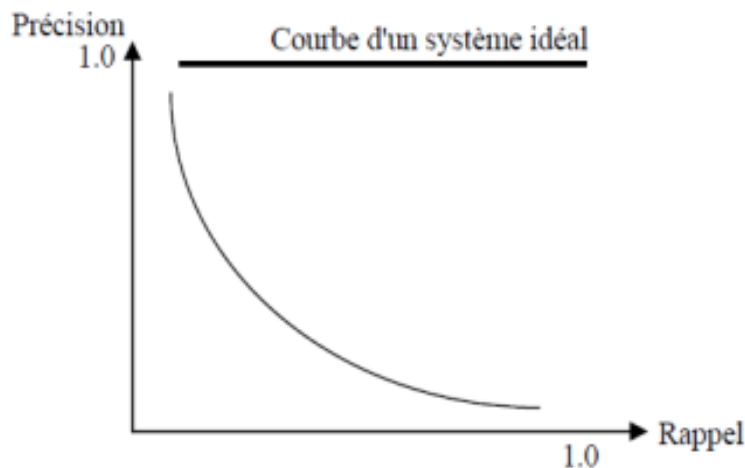
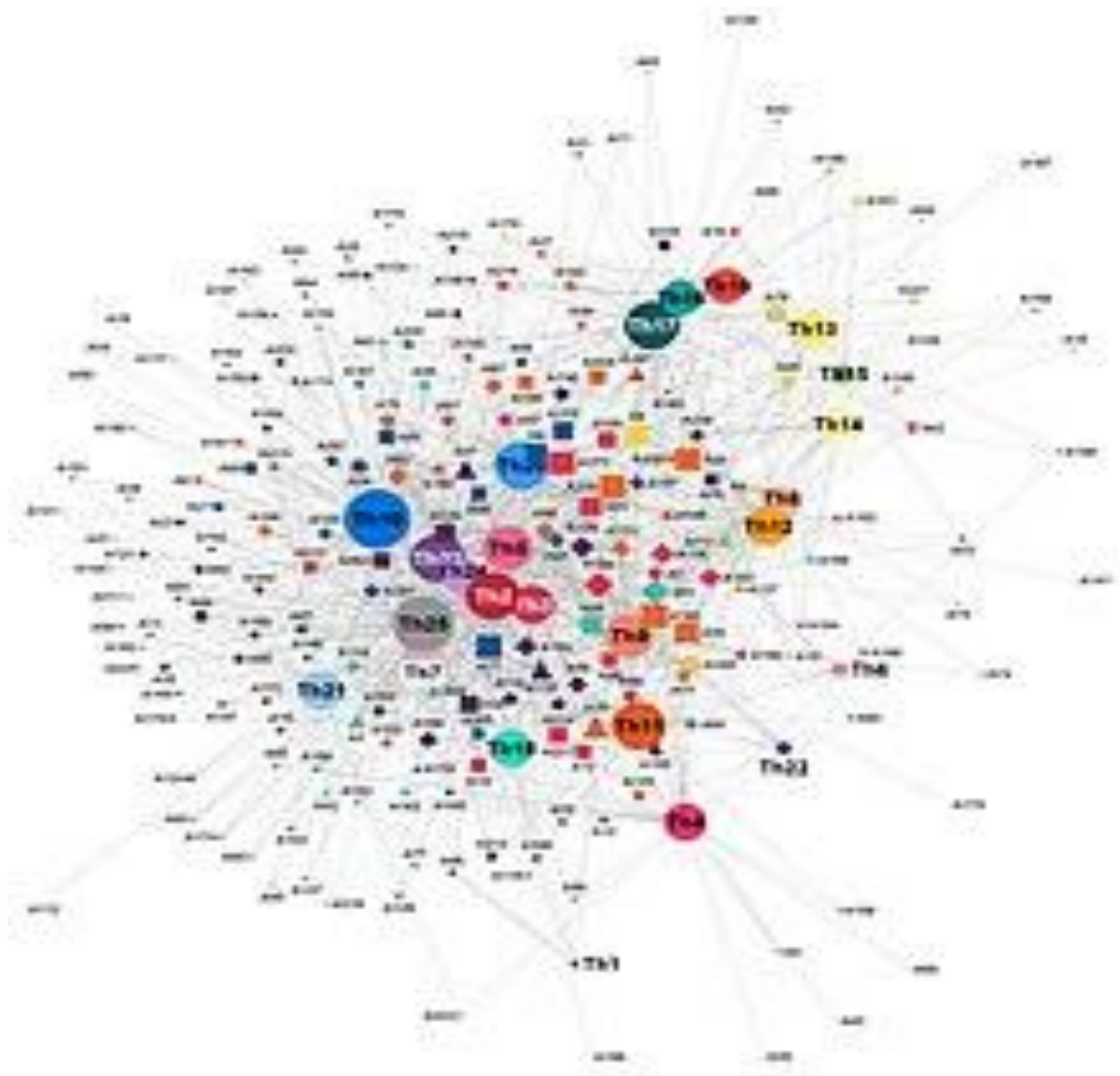


Figure 14. Une courbe de rappel et précision

6. Conclusion :

Dans ce chapitre, nous avons tenté de décrire les différents concepts pour mettre en cadre un système CBIR. Nous avons abordé au début les différents composants d'un CBIR, puis nous avons parlé de la représentation des images dans un CBIR et les mesures pour l'évaluer, et nous avons également mis en évidence ces domaines d'application.

CHAPITRE 3 : DESCRIPTEURS D'IMAGES & MESURES DE SIMILARITÉ



1. Introduction

Dans notre domaine, le premier but de n'importe quel système de recherche d'images est de fournir des résultats totalement satisfaits par l'utilisateur, ce qui nécessite d'utiliser les meilleures méthodes d'extraction de caractéristiques et les représenter d'une manière réduite afin d'atteindre un temps de réponse naturaliste pour un système de recherche d'images.

De nombreux descripteurs d'images et métriques de similarité sont utilisés dans les systèmes CBIR, et certaines fonctionnalités fonctionnent bien lorsqu'elles sont utilisées en combinaison avec des métriques spécifiques.

L'objectif de ce chapitre est de décrire les différents descripteurs utilisés pour représenter des images d'une base afin de commencer la mesure de similarité entre ces images.

2. Les descripteurs d'image :

La description du contenu des images est une étape importante dans le CBIR, car les performances du CBIR dépendent principalement du choix des descripteurs utilisés et des techniques impliquées dans leur extraction. Un descripteur est défini comme la connaissance utilisée pour caractériser l'information contenue dans une image.

2.1. Les descripteurs de la couleur :

La couleur est l'information visuelle la plus utilisée pour représentation d'images. C'est le descripteur le plus utilisé dans les systèmes CBIR en raison de son invariance par rapport l'échelle et sa richesse d'information, Avant de sélectionner le descripteur de couleur approprié, la couleur doit être déterminée d'abord par son modèle (voir le chapitre 1).

2.1.1. L'histogramme

L'histogramme est un graphique montrant le nombre de pixels par intensité lumineuse (couleur). Il consiste de calculer le nombre de répétitions de chaque couleur sur l'ensemble du pixel. Pour les images en couleur, plusieurs histogrammes sont nécessaires. Par exemple pour une image codée en RVB, un histogramme représente la distribution de la luminance et trois histogrammes représentant respectivement la distribution des composantes "rouge, bleu et vert". [16]

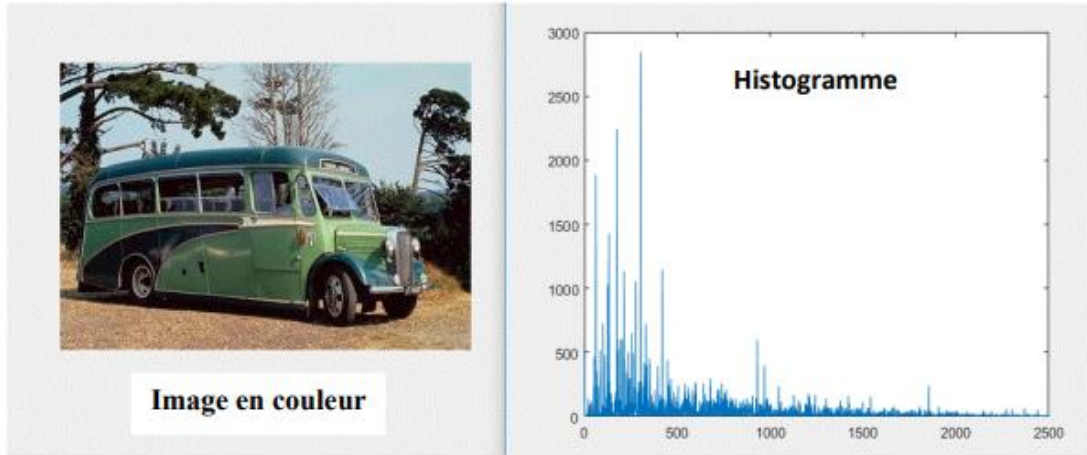


Figure 15. L'histogramme d'une image.

2.1.2. Les moments de couleur

La méthode de l'histogramme des couleurs utilise la distribution complète des couleurs. Il nécessite un grand espace pour le stockage des données. Au lieu d'un compte la distribution complète, seuls certains systèmes de récupération d'images sont utilisés. Caractéristiques de couleur dominantes telles que : μ_c (noté μ_c), asymétrie type (σ_c) et asymétrie (θ_c). Ces propriétés sont calculées pour chaque composant colorie les formules suivantes : Équation (1.2), Équation (1.3) et Équation (1.4), respectivement. [17]

$$\mu_c = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N p_{ij}^c \quad (3.1)$$

$$\sigma_c = \left[\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (p_{ij}^c - \mu_c)^2 \right]^{\frac{1}{2}} \quad (3.2)$$

$$\theta_c = \left[\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (p_{ij}^c - \mu_c)^3 \right]^{\frac{1}{3}} \quad (3.3)$$

Où M et N sont les tailles horizontale et verticale de l'image. Et p_{ij}^c : est la valeur de la couleur c dans la ligne i et la colonne j de l'image I .

2.1.3. Couleurs dominantes

L'utilisation d'un histogramme pour montrer la distribution des couleurs présente certains inconvénients. Du point de vue du stockage de la mémoire, un histogramme multidimensionnel est ((creux)). En d'autres termes, la plupart des cellules n'ont pas de pixels. Beaucoup d'espace de stockage est utilisé inutilement. Aussi, bien que tous les bacs aient la même taille, il serait intéressant d'avoir des bacs plus petits pour les zones contenant des couleurs très courantes et des bacs plus grands pour les couleurs moins courantes. En termes de mesures de similarité utilisées, les mesures conventionnelles n'effectuent que des comparaisons de cellule à cellule. Même si l'histogramme est ordonné, les voisins de la cellule ne sont pas pris en compte si les valeurs sont différentes.

Pour dénouer ces différents problèmes, Rubner et Tomasi [18] Proposent une signature de couleur dominante. La signature $s = \{s_i = (m_i, W_i)\}$ est l'ensemble des nuages de points. Chaque nuage est décrit par son mode m_i (le mode du nuage de points correspond au maximum de sa densité de probabilité) et le nombre de pixels appartenant au nuage W_i . [11]

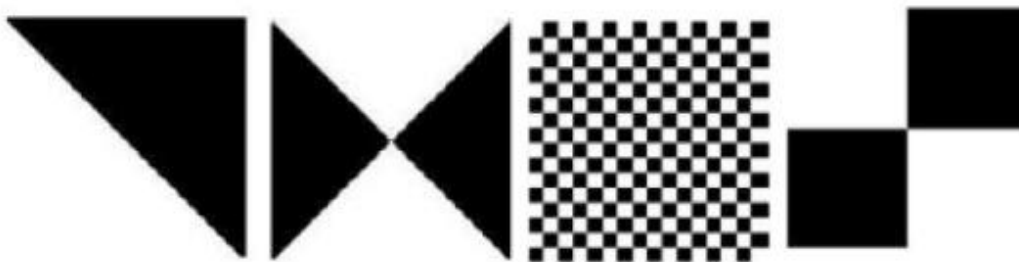


Figure 16. Limite des histogrammes et des couleurs dominantes.

Ces quatre images ont le même histogramme et les mêmes couleurs dominantes.

2.1.4. Le Corrélogramme

Pour améliorer l'utilisation d'histogramme, nous avons ajouté un corrélogramme de couleurs puisqu'il peut fournir la relation entre la distribution de couleurs des pixels et la corrélation spatiale entre les paires de couleurs. Les corrélogramme créent une représentation visuelle de l'image similaire. En utilisant ce diagramme **Figure (17)**, Une voie simpliste de noter comment l'information spatiale est incluse dans un corrélogramme ; est de noter que le pixel bleu foncé cerclé est 3 Pixel à gauche et un Pixel au-dessous du cadre rouge moyen carré. Ainsi tandis que l'histogramme notera le nombre de couleurs et leurs intensités, un corrélogramme pourra noter l'information spatiale en indiquant la distance entre les différentes couleurs. [03]

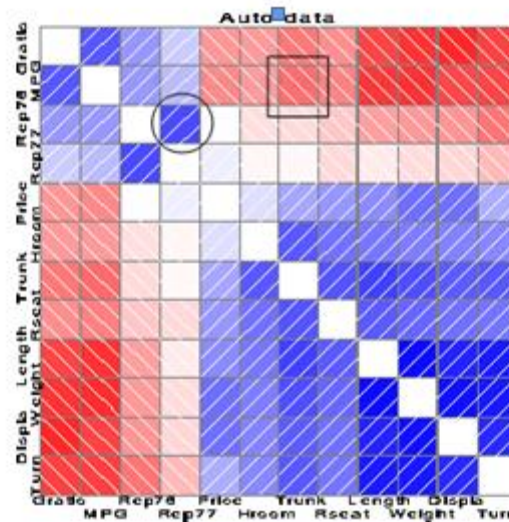


Figure 17. Représentation d'un corrélogramme de couleur

Résumé :

La couleur est un descripteur très puissant pour la description des images couleurs. Les histogrammes de couleur, étant le descripteur de couleur le plus utilisé dans les projets, académiques ainsi que commerciaux, de recherche d'images couleur par le contenu, ont été utilisés.

2.2. Descripteurs des textures :

La texture est une information de plus en plus utilisée pour le catalogage et la recherche d'images par contenu, car il permet de s'affranchir de certains des problèmes posés par l'indexation par Couleur, surtout lorsque les distributions de couleurs sont très proches.

La majorité des méthodes d'analyse de texture sont basées sur des méthodes statistiques, comme les matrices de cooccurrences, les filtres de Gabor et les ondelettes.

La notion de texture est liée à trois concepts principaux :

- 1- Un certain ordre local qui se répète dans une région de taille assez grande,
- 2- Cet agencement est déterminé par un agencement ordonné de ses composants de base,
- 3- Ces composants primaires représentent des entités unifiées qui les distinguent. Des proportions similaires sont observées dans toute la région.

Nous introduisons dans ce qui suit quelques représentations de la texture qui sont utilisées dans le domaine de la recherche d'images par le contenu.

2.2.1. Matrice de cooccurrence

La matrice de cooccurrences est définie comme une relation géométrique entre deux pixels. En 1973, Haralick. [19]

La matrice de cooccurrence contient les fréquences spatiales relatives d'occurrence des niveaux de gris dans quatre directions ($\theta=0$, $\theta=\pi/2$, $\theta=\pi/4$ et $\theta=3\pi/4$). La matrice de cooccurrence est une matrice carrée $n \times n$. ici, n est le nombre de niveaux de gris dans l'image. Par conséquent, chaque matrice doit choisir la distance et l'angle de décalage séparant deux pixels du motif. Une classification appropriée des textures est basée sur une sélection appropriée à partir de 14 critères qui peuvent être extraits de n'importe quelle matrice de cooccurrence, tels que l'homogénéité, l'entropie, le contraste et la corrélation.

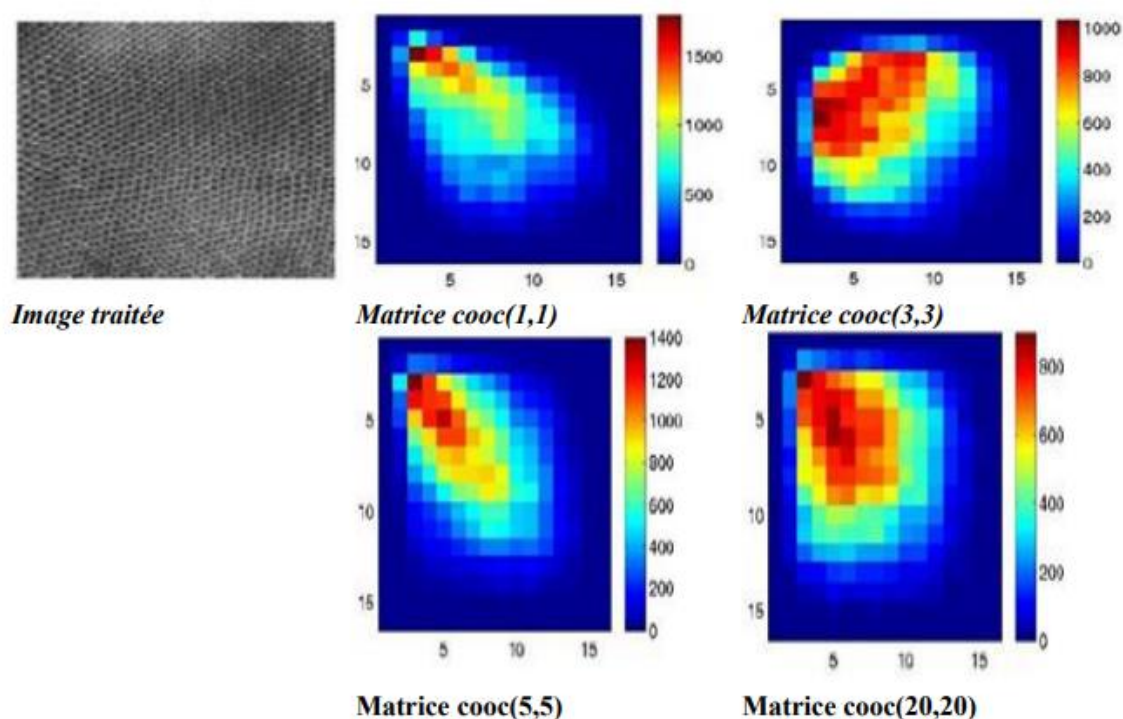


Figure 18. Exemples de matrices de cooccurrences.

2.2.2. Transformée en ondelettes

Le terme " ondelette " en anglais (wavelet) a été utilisé pour la première fois en 1984 par J. Morlet et A. Grassmann pour résoudre des problèmes de traitement des signaux pour la prospection pétrolière.

Une ondelette est la fonction sous-jacente de la décomposition en ondelettes. Chaque ondelette a une fréquence spécifique pendant un temps limité, comme une note de musique. L'expression suivante représente l'ondelette fille :

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi \left(\frac{t-b}{a} \right) \quad (3.4)$$

$$a > 0, b \in \mathbb{R}$$

Le paramètre a est le facteur d'échelle, il détermine la dilatation de l'atome de base (ondelette fille). Le paramètre b est le facteur de translation, il permet de traduire l'atome de base à gauche ou à droite. Le paramètre $1/\sqrt{a}$ est un facteur de normalisation à travers les différentes échelles.

À partir de la transformée en ondelettes on peut extraire des attributs de différents types et à différents niveaux de résolution. Donc les ondelettes permettent de caractériser la texture en décrivant les primitives et les règles d'arrangement qui les relient. [02]

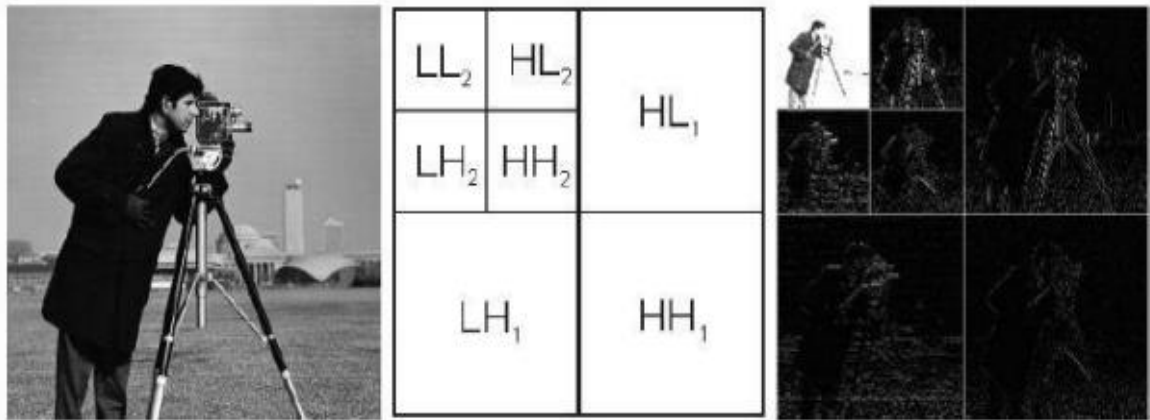


Figure 19. Exemple de décomposition par ondelettes.

Image d'origine (gauche), Schéma de décomposition de l'image par DWT avec 2 échelles de décomposition (centre), Résultat de la décomposition (droite).

2.3. Les descripteurs de la forme :

Comme les descripteurs/caractéristiques de texture, les attributs de forme sont complémentaires aux descriptions de couleur. [20] La forme est un attribut visuel essentiel et l'une des caractéristiques/descripteurs de base pour afficher le contenu de l'image. Ci-dessous, nous introduisons quelques méthodes pour décrire les formes :

2.3.1. Les descripteurs géométriques de région

Les descripteurs géométriques des formes vous permettent de distinguer les différents types de formes que peuvent prendre les objets de votre scène. Ils doivent être divisés dans les zones devant la photo. Ensuite, calculez-les dans différentes régions de l'image. Ensuite, calculez-les dans différentes régions de l'image. L'aire relative (ou normalisée) d'une région $\mathcal{R}_k$ dans l'image I est le nombre de pixels contenus dans la région par rapport au nombre total de pixels de l'image. [21]

$$Sk = \text{card}(\mathcal{R}k) / \text{hauteur}(I) * \text{largeur}(I) \quad (3.5)$$

Le centre de masse des pixels de la région est défini par :

$$P = (Pi, Pj) = (\sum_{i \in \mathcal{R}} i / \text{card}(\mathcal{R}k) / \text{largeur}(I), \sum_{j \in \mathcal{R}} j / \text{card}(\mathcal{R}k) / \text{hauteur}(I)) \quad (3.6)$$

La longueur du contour de la région est le nombre de pixels en bordure de la région :

$$lk = \text{card}(\text{contour}(\mathcal{R}k)) \quad (3.7)$$

2.3.2. Les moments géométriques

Les moments géométriques sont utilisés pour décrire des formes, propriétés statistiques. Le processus est simple, mais le calcul prend beaucoup de temps. [07]

Nous avons :

I (M, N) : image.

Les moments (p, q) sont calculés par la formule suivante :

$$M_{pq} = \sum_{i=1}^M \sum_{j=1}^N i^p j^q I(i, j) \quad (3.8)$$

La surface est calculée par la formule suivante :

$$M_{00} = \sum_{i=1}^M \sum_{j=1}^N I(i, j) \quad (3.9)$$

Le centre de gravité est calculé par la formule suivante :

$$M_{01} = \sum_{i=1}^M \sum_{j=1}^N j^q I(i, j) \quad (3.10)$$

$$M_{10} = \sum_{i=1}^M \sum_{j=1}^N i^p I(i, j) \quad (3.11)$$

$$\bar{i} = \frac{M_{10}}{M_{00}} \text{ and } \bar{j} = \frac{M_{01}}{M_{00}}$$

Le moment central invariant en translation est calculé par la formule suivante :

$$\mu_{pq} = \sum_{i=1}^M \sum_{j=1}^N (i - \bar{i})^p (j - \bar{j})^q I(i, j) \quad (3.12)$$

Le moment normalisé invariant à mettre à l'échelle est calculé par la formule suivante :

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{1+\frac{p+q}{2}}} \quad (3.13)$$

2.3.3. Transformée de Hough.

Cette transformation a été proposée par P.V.C, **Hough** dans un brevet déposé en 1960 afin de détecter des alignements à l'aide d'un oscilloscope et de deux caméras vidéo. Inaperçue pendant plusieurs années.

On regroupe généralement le nom de la conversion "TH", ce qui rend détecter dans une image la présence d'une courbe paramétrique appartenant à une série connue, À partir d'un ensemble de points sélectionnés, appelés points caractéristiques. Utilisation TH est essentiellement l'information spatiale des points caractéristiques (leur position dans l'image), cependant, Parfois, les informations contenues dans le signal d'image lui-même (Valeur de luminosité d'un point donné). On suppose que ce signal est une fonction scalaire, mais rien n'empêche qu'il soit un vecteur ; c'est le cas des images en couleur, ou Imagerie multi spectrale. Enfin, bien que la plupart des images soient en deux dimensions, nous Nous pourrons appliquer TH aux images 3D. [07]

Considérons l'espace image, et un ensemble de N points sélectionnés par prétraitement :

$$\xi = \{M_i, i = 1 \dots \bar{N}\} \in \mathbb{R}^n \quad (3.14)$$

Un point M d'est identifié par sa coordonnée x.

Soient p l'espace des paramètres et F une famille de courbes paramétrées par n passer :

$$F = \{ \{x: f(x, a) = 0, x \in \mathbb{R}^n\}, a \in \Omega \} \quad (3.15)$$

La transformée de Hough liée à la famille F correspondant à l'ensemble ζ des fonctions g définies sur Ω .

3. Mesures de similarité

Il est difficile de définir la notion de similarité entre deux images. Ça prend Indique les critères sur lesquels repose cette notion de similarité. La figure 20 montre un exemple où l'utilisateur est invité à classer quatre images suggérées en familles Image similaire.

Exemple :

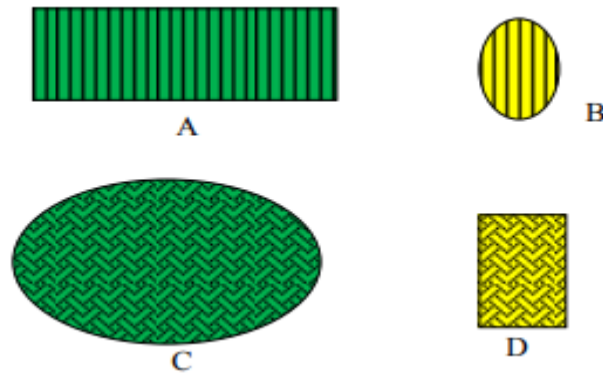


Figure 20. Exemple de base à classer en familles d'images similaires

La mesure de la similarité est une étape fondamentale de la recherche Images par contenu. Dépend de la mesure de similarité Entre l'image de la requête et l'image de la base de données. C'est Le système utilise le descripteur de demande et descripteurs pour toutes les images de la base dans l'espace attributaire Si la distance est petite, les images sont considérées comme similaires. Le résultat est Au lieu d'afficher une seule image, il affiche plusieurs images classées en conséquence Similitude avec la requête.

Une requête est généralement exprimée sous la forme d'un ou plusieurs vecteurs. Pour les grandes dimensions, recherchez les voisins les plus proches de la requête. C'est, Une phase en ligne dans laquelle le système calcule les correspondances entre les descripteurs d'image de requête et chaque descripteur d'image de base de données. [21]

En résumé, les mesures de similarité vérifient généralement les propriétés [07] :

- **Perception** : de petites distances dans l'espace des caractéristiques montrent deux images ressemblées.
- **Calculs** : Les mesures de distance sont calculées rapidement pour une faible latence.
- **La scabilité** : les calculs de distance ne doivent pas être affectés par les changements de taille. Quartier général
- **Robustesse** : Les mesures doivent être robustes aux conditions d'acquisition changeantes d'image.

3.1. Les méthodes de calcul

Vous trouverez ci-dessous les distances les plus couramment utilisées pour la comparaison d'images.

3.1.1. Distance de Mahalanobis

Cette distance tient compte des corrélations entre les distributions de classe. Elle est définie par :

$$D_M = \sqrt{(f_1 - f_2)^T C^{-1} (f_1 - f_2)} \quad (3.16)$$

Dans les cas où les dimensions des caractéristiques sont indépendantes, C n'a que la variance, la distance de Mahalanobis simplifie sous la forme suivante :

$$D_M = \frac{\sum (f_1(i) - f_2(i))^2}{c_i} \quad (3.17)$$

3.1.2. Distances entre histogrammes

Distance euclidienne, applicable aux histogrammes. Cependant, les histogrammes ont certains indicateurs. Cette métrique est l'une des premières distances utilisées dans la recherche d'images par le contenu, qui mesure l'intersection entre deux histogrammes. Étant donné deux histogrammes h1 et h2.

$$D_{Intersec} = \frac{\sum_{i=1}^n \min(h_1(i), h_2(i))}{\sum_{i=1}^n h_2(i)} \quad (3.18)$$

3.1.3. Earth Mover's Distance (EMD)

La mesure de la distance du **Earth Mover** (EMD) est basée sur la minimisation du coût nécessaire pour transformer une distribution en une autre. Elle peut Appliquer pour calculer la similarité entre deux distributions ou entre deux ensembles Distribution.

Voici sa définition au cas où vous voudriez calculer la distance entre les deux ensembles de distributions. Soit $x = \{(x^1, \omega x, 1), (x^2, \omega x, 2), \dots, (x^n, \omega x, n)\}$ et $y = \{(y^1, \omega y, 1), (y^2, \omega y, 2), \dots, (y^m, \omega y, m)\}$ deux ensembles de distributions où chaque $k^{ième}$ distributions z^k est associé à un poids $w_{z,k}$, alors la distance EMD est définie par :

$$EMD(x, y) = \frac{\sum_{i=1}^n \sum_{j=1}^m \delta(\vec{x}_i, \vec{y}_j) W_{i,j}}{\sum_{i=1}^n \sum_{j=1}^m W_{i,j}} \quad (3.19)$$

EMD peut gérer les histogrammes, il n'est donc pas nécessaire d'effectuer une quantification vectorielle. Cette mesure est utile pour calculer la distance entre deux images segmentées. Ici, chaque blob est représenté par une distribution visuelle. Mais, L'EMD est très coûteux en calcul. [21]

3.1.4. Distances de Minkowski

La façon la plus simple de mesurer la similarité entre deux images est la distance de Minkowski. Cette distance est calculée entre descripteurs/vecteurs de caractéristiques et définie par la formule.

$$L_r(V_1, V_2) = \left[\sum_{i=0}^n |V_1(i) - V_2(i)|^r \right]^{1/r} \quad (3.20)$$

Où V_1, V_2 : vecteurs descripteurs/caractéristiques, $r \geq 1$ est le facteur de Minkowski, n est la dimension de vecteur.

3.1.5. Distance quadratique

Un inconvénient des métriques telles que la distance de Minkowski et l'intersection d'histogramme est qu'elles traitent équitablement les éléments des vecteurs de caractéristiques (en comparant les composants vectoriels individuellement sans tenir compte des autres composants). Pour résoudre ce problème on utilise la distance quadratique. Cette distance favorise les éléments les plus similaires. Il est défini comme :

$$D_Q = \sqrt{(f_1 - f_2)^T A (f_1 - f_2)} \quad (3.21)$$

Cette propriété de distance se rapproche de la perception humaine des couleurs, ce qui en fait une mesure attrayante pour les systèmes de récupération d'images couleur basés sur le contenu.

3.1.6. Similarité multi-vectorielle

Lorsqu'une image est indexée non pas par un seul vecteur, mais par une série de vecteurs liés à différents espaces de caractéristiques (couleurs, textures, formes, etc.). Il y a un problème de fusion des informations de différents modèles.

La première consiste à concaténer les différentes signatures et à utiliser un des moyens classiques de comparaison entre vecteurs. Les sous-vecteurs étaient auparavant centrés et réduits pour revenir à une échelle de valeurs commune pour divers attributs.

La deuxième méthode consiste à traiter chaque espace attributaire séparément. La similarité globale est calculée en fusionnant les valeurs de similarité pour chaque espace de caractéristiques. C'est ce qu'on appelle un "modèle hiérarchique". [21]

3.1.7. Distances entre distributions

La similarité entre les distributions consiste à déterminer si deux distributions peuvent être dérivées de la même distribution de probabilité.

Notez que les mesures suivantes (test χ^2 , divergence de **Kullback-Leibler** et divergence de différence de Jensen) ne satisfont pas toutes les conditions et doivent être qualifiées de mesures de similarité au lieu de mesures de distance.

- a. La distance de χ^2
- b. La divergence de Kullback-Leibler
- c. Divergence de Jeffrey (JD)
- d. Distance de Kolmogorov Smirnov
- e. Distance de Cramer Von Mises
- f. Distance de Bhattacharya

4. Conclusion

Le choix des descripteurs pour un système de recherche d'images par contenu est important, dans le sens où, ce choix influe sur les résultats attendus. Et Nous avons consacré ce chapitre à la présentation de notion et les différentes mesures de similarité.

CHAPITRE 4 : LE TRAVAIL EFFECTUÉ



1. Introduction :

Après avoir étudié le domaine de CBIR et leur principe de fonctionnement. Nous allons appliquer ces règles et ces approches pour créer notre propre système.

Nous présenterons alors, dans ce chapitre l'implémentation de l'application, les descripteurs utilisés et la base d'images choisi afin de pouvoir évaluer notre système.

2. Outils de développement :

- Python

Python est un langage de programmation interprété. Il est très sollicité par une large communauté de programmeurs. Python est un langage simple, facile à apprendre et permet une bonne réduction du coût de la maintenance des codes.

- Google Colab

Colab est un environnement de développement Python qui s'exécute dans le navigateur WEB. Colab offre gratuitement un niveau minimum de capacité de calcul et de stockage sans configuration requise.

- Tensorflow

Tensorflow est un Framework de programmation pour le calcul numérique. Développé par Google et publié sous la licence Apache 2.0 fin 2015. Il est basé sur l'infrastructure DistBelief développée par Mountain View pour entraîner des modèles d'apprentissage sur des clusters de centaines de serveurs et est reconnu pour ses performances.

- Keras

Keras est une API de réseaux de neurones de haut niveau, écrite en Python et capable de fonctionner sur Tensorflow ou Theano. Elle a été développée en mettant l'accent sur l'expérimentation rapide. Elle permet d'aller de l'idée aux résultats avec le moins de délai possible ce qui représente une aide considérable à la recherche scientifique.

3. La base d'image utilisée :

Notre travail est fait sur un échantillon d'images en couleur composé de 299 images constituant notre base de données. Ces images ont été divisées en 03 catégories : papillons, voitures et fleurs.

Pour la réalisation de notre modèle, on a divisé la base de données comme suite :

- 70 % pour l'entraînement (train).
- 20% pour le test.
- 10% pour la validation.



Figure 21. Quelques images de la base d'image.

4. Les descripteurs d'image utilisés :

Nous avons choisi les caractéristiques suivantes

4.1. Les modèles de couleur utilisés

Dans notre application nous n'avons intégré que l'espace RVB. Il se compose des trois couleurs primaires rouge, vert et bleu.

4.2. Grayscale Images

Une image en niveaux de gris est simplement une image dans laquelle les seules couleurs sont des nuances de gris. La raison pour différencier ces images de tout autre type d'image couleur est que moins d'informations doivent être fournies pour chaque pixel. En fait, une couleur « grise » est une couleur dans laquelle les composantes rouge, verte et bleue ont toutes la même intensité dans l'espace RVB. [S2]

5. Architecture de l'application :

La méthode proposée est présente par la suivante (figure 22) :

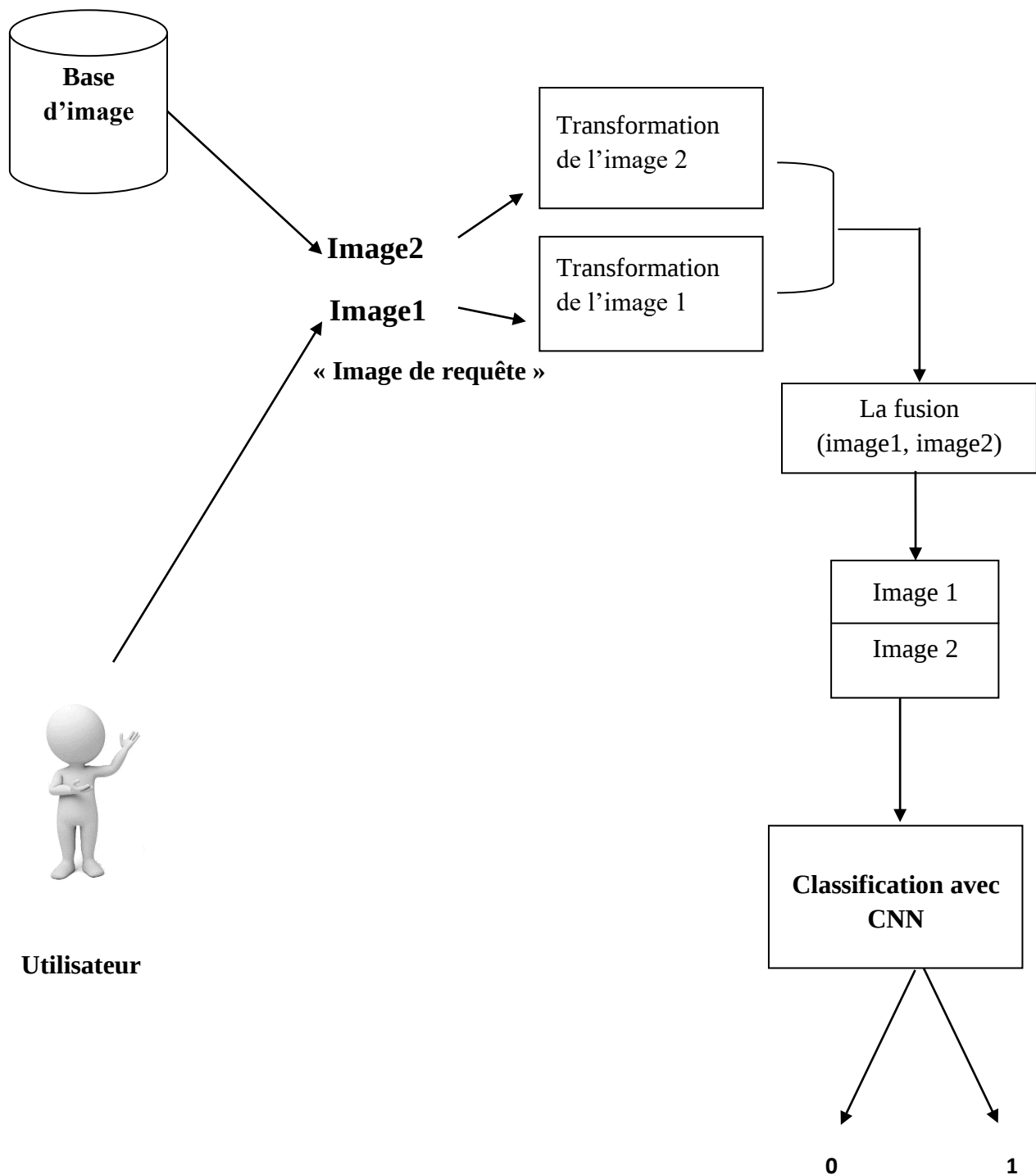
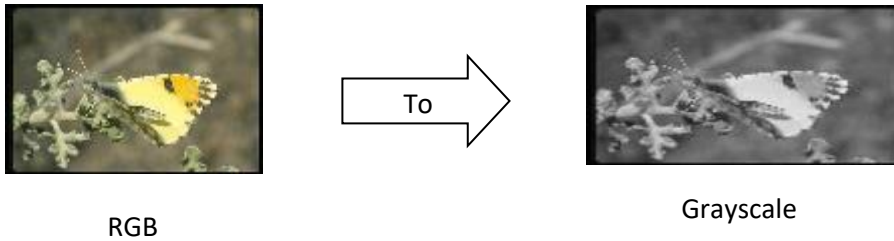


Figure 22. Schéma de représentation de l'application

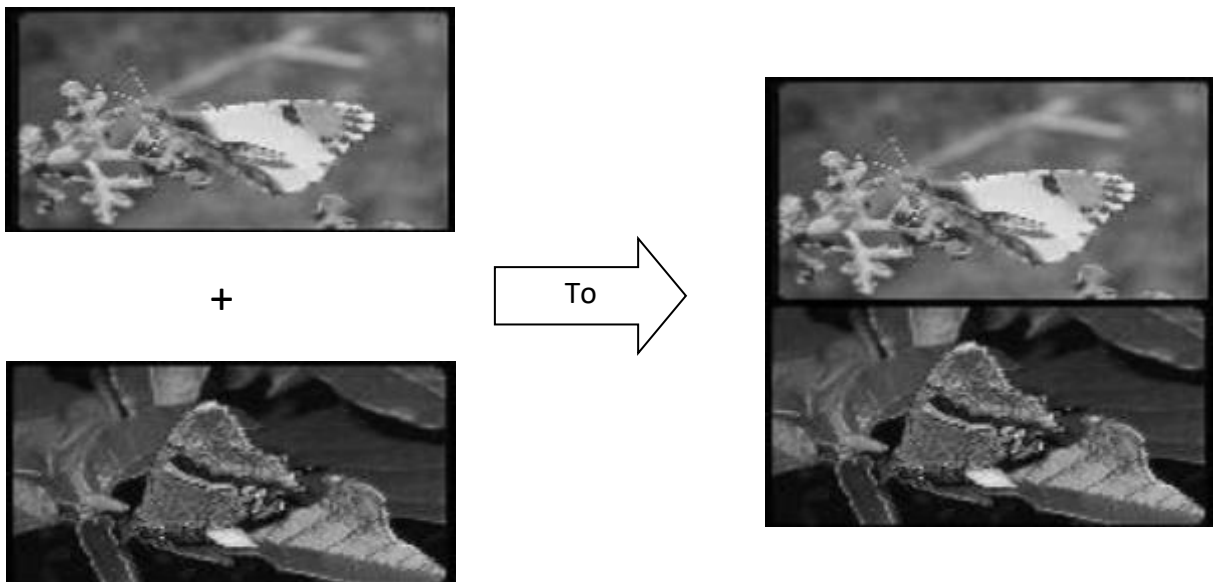
ETAPE 01 : Nous avons converti les images RGB en niveaux de gris. En effet, l'échelle de gris comprime l'image jusqu'à un nombre minimum de pixels. Plus facile à visualiser. Étant donné que l'image est principalement un espace bidimensionnel (2D) et non un espace 3D, elle aide à distinguer les détails d'ombre et de surbrillance dans l'image.

Exemple :



Etape 02 : nous fusionons les deux images, pour obtenir une seule image

Exemple :



Etape 03 : Nous donnons le résultat de la fusion des deux images au réseau de neurones convolutif pour faire la classification.

6. La structure du modèle :

Model : " my_model/conv1/ "

Layer (type)	Output Shape	Param #
=====		
Img_input (Input Layer)	[(None, 224, 224, 3)]	0
layer_1 (Conv2D)	(None, 224, 224, 32)	896
layer_2 (Conv2D)	(None, 224, 224, 64)	18496
layers_3 (MaxPooling2D)	(None, 112, 112, 64)	0
dropout (Dropout)	(None, 112, 112, 64)	0
layer_4 (Conv2D)	(None, 112, 112, 64)	36928
layers_5 (MaxPooling2D)	(None, 56, 56, 64)	0
dropout_1 (Dropout)	(None, 56, 56, 64)	0
layer_6 (Conv2D)	(None, 56, 56, 128)	73856
layer_7 (MaxPooling2D)	(None, 28, 28, 128)	0
dropout_2 (Dropout)	(None, 28, 28, 128)	0
fc_1 (Flatten)	(None, 100352)	0
layer_8 (Dense)	(None, 64)	6422592
dropout_3 (Dropout)	(None, 64)	0
Prediction (Dense)	(None, 1)	130
=====		

Total params : 6, 552,898
Trainable params : 6, 552,898
Non-trainable params : 0

7. Représentation graphique du modèle CNN :

On va représenter le réseau de neurones convolutifs avec la figure (23) qui montre le nombre de couches utilisées et comment elle interconnectées et séquencées les unes avec les autres, ainsi que les fonctions d'activation.

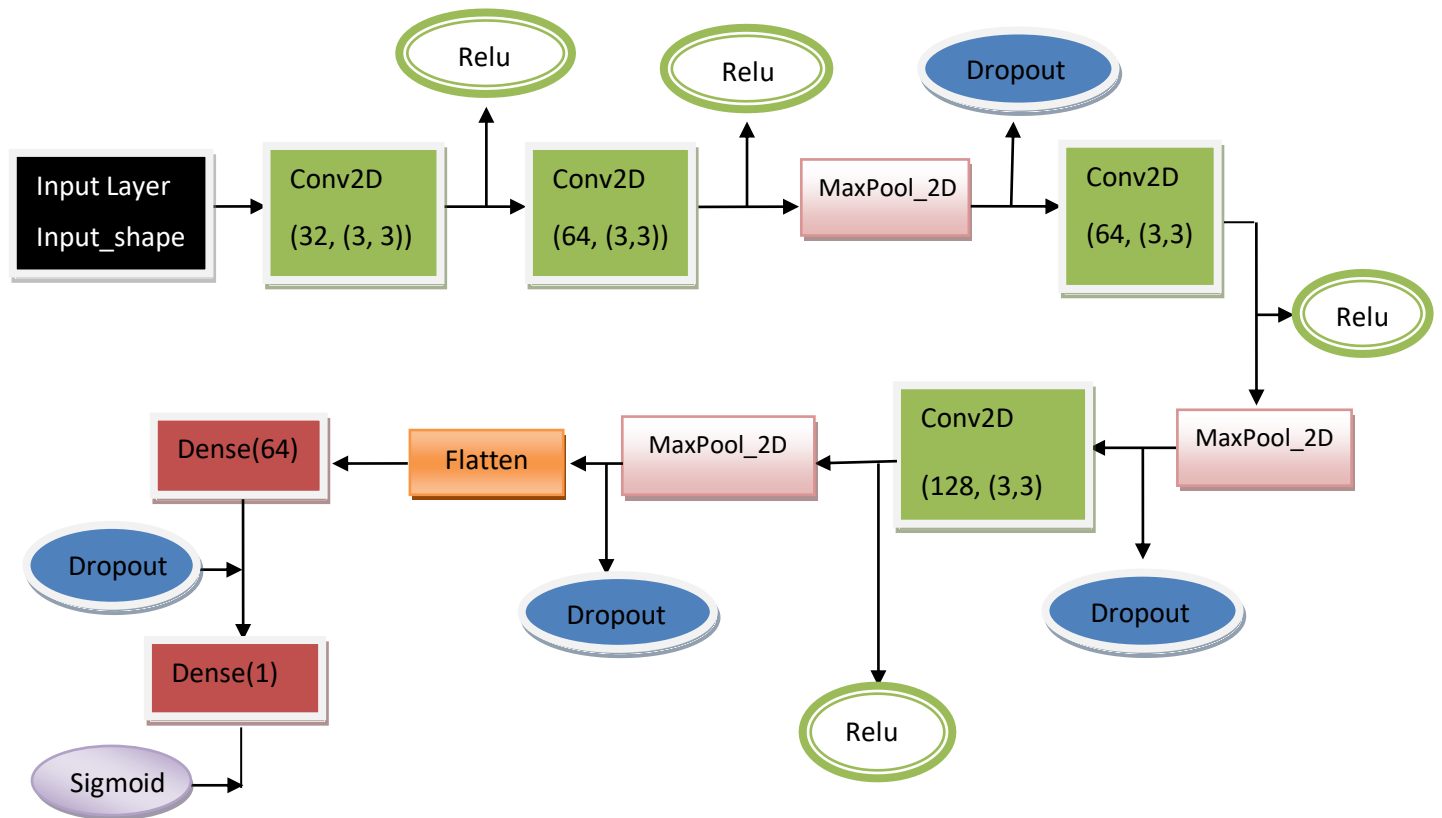


Figure 23. Architecture du modèle CNN

8. Conclusion :

Les réseaux de neurones convolutifs ont un rôle très important dans la classification des images, ce que nous avons consacré ce chapitre à expliquer l'architecture de notre application.

CHAPITRE 5 : ÉTUDE EXPÉRIMENTALE



1. Introduction :

Le but général de cette méthode expérimentale est de pouvoir établir une relation causale entre deux paramètres. Il s'agit donc d'une méthode qui permet d'expliquer réellement le phénomène étudié en termes de causalité. Dans ce chapitre nous avons présenté les méthodes utilisées pour évaluer ce système et leurs résultats.

2. Les méthodes d'évaluation :

Dans un modèle CNN, la fonction de perte (loss) et l'exactitude (accuracy) sont des métriques couramment utilisées pour évaluer les performances du modèle lors de l'entraînement et de l'évaluation.

2.1 : Fonction de perte (Loss Function) :

La fonction de perte mesure à quel point les prédictions du modèle diffèrent des valeurs réelles (étiquettes). L'objectif est de minimiser cette fonction afin d'obtenir des prédictions plus précises. Dans un modèle de classification binaire, la fonction de perte typiquement utilisée est l'entropie croisée binaire (binary cross-entropy). Pour un modèle de classification multiclasse, l'entropie croisée catégorielle (categorical cross-entropy) est couramment utilisée.

2.2. Accuracy (exactitude) :

L'exactitude est une mesure qui indique la proportion de prédictions correctes par rapport au nombre total d'échantillons. C'est une métrique couramment utilisée pour évaluer les performances d'un modèle de classification. Dans un modèle CNN, l'exactitude est calculée en comparant les prédictions du modèle avec les valeurs réelles (étiquettes) et en déterminant le pourcentage de prédictions correctes.

3. Résultats du modèle réalisé :

Le système nous a donné les résultats suivants :

3.1. Loss function

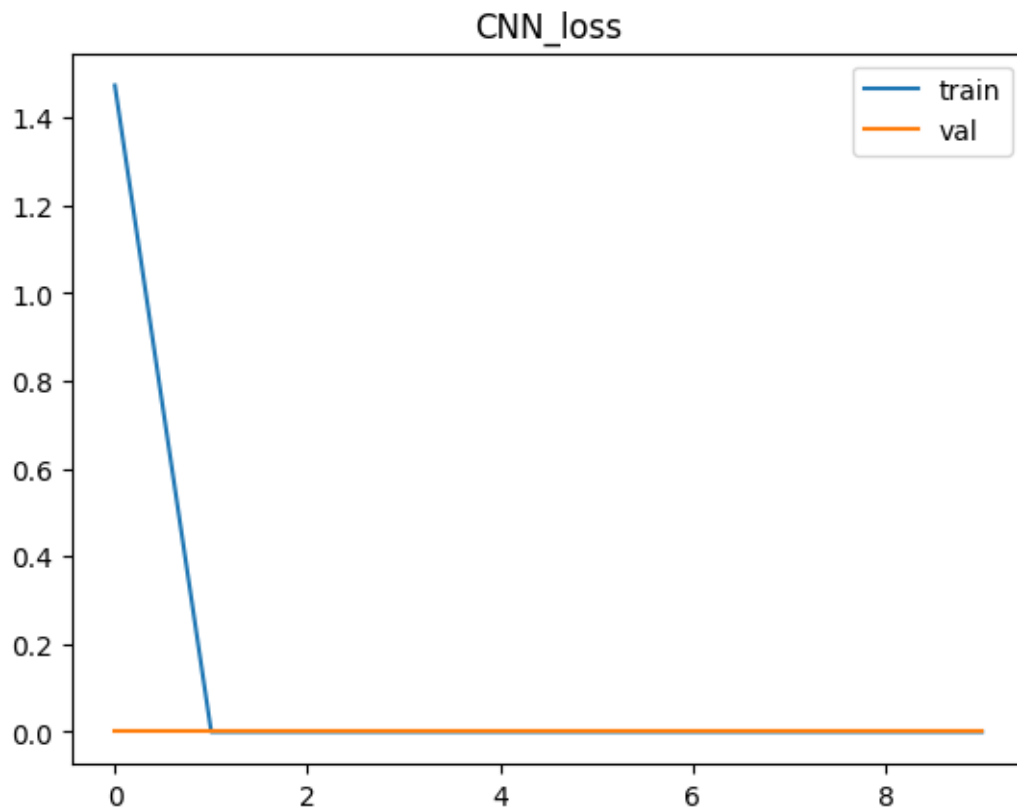


Figure 24. « Loss function » du modèle réalisé

3.2. CNN train

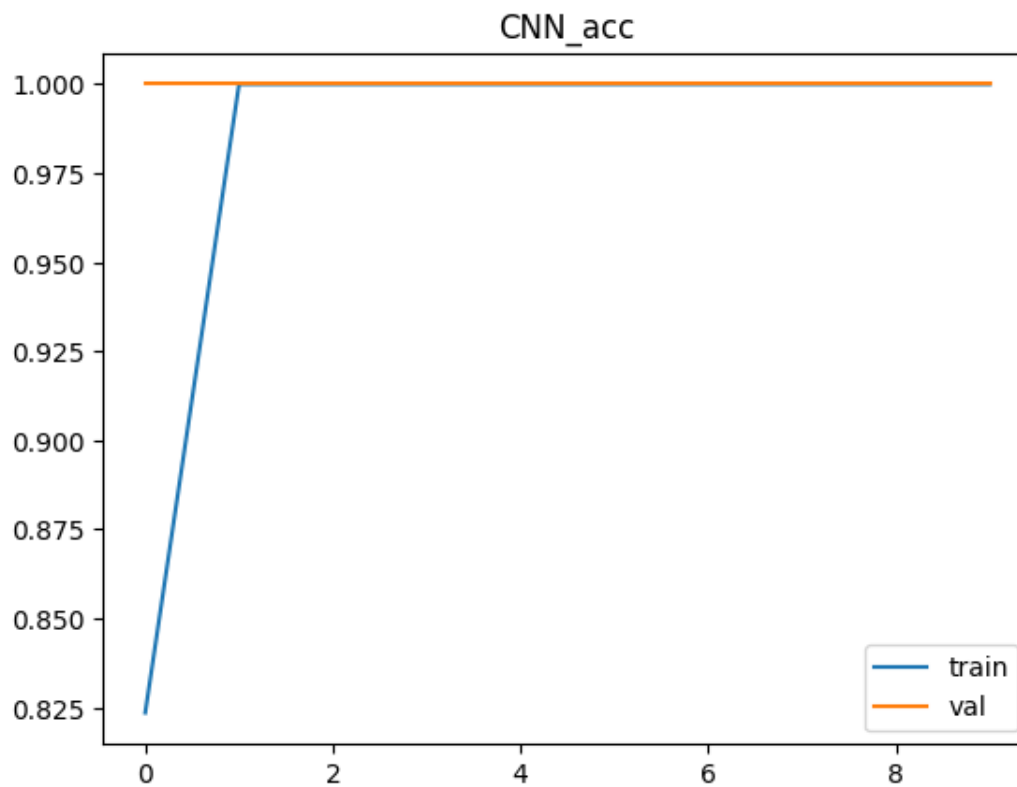


Figure 25. « CNN train » du modèle réalisé

Avec ce modèle et cette base de données on a obtenu les résultats représentés dans les graphes ci-dessus.

Où :

- Après la 1^{ère} itération des epochs on a trouvée le bon résultat du train (=1).
- À l'opposé on a obtenu le bon résultat de loss (=0) de la même façon.

3.3. Matrix de confusion :

Le vrai négative 0.

Le faux positive 30.

Le faux négative 0.

Le vrai positive 0.

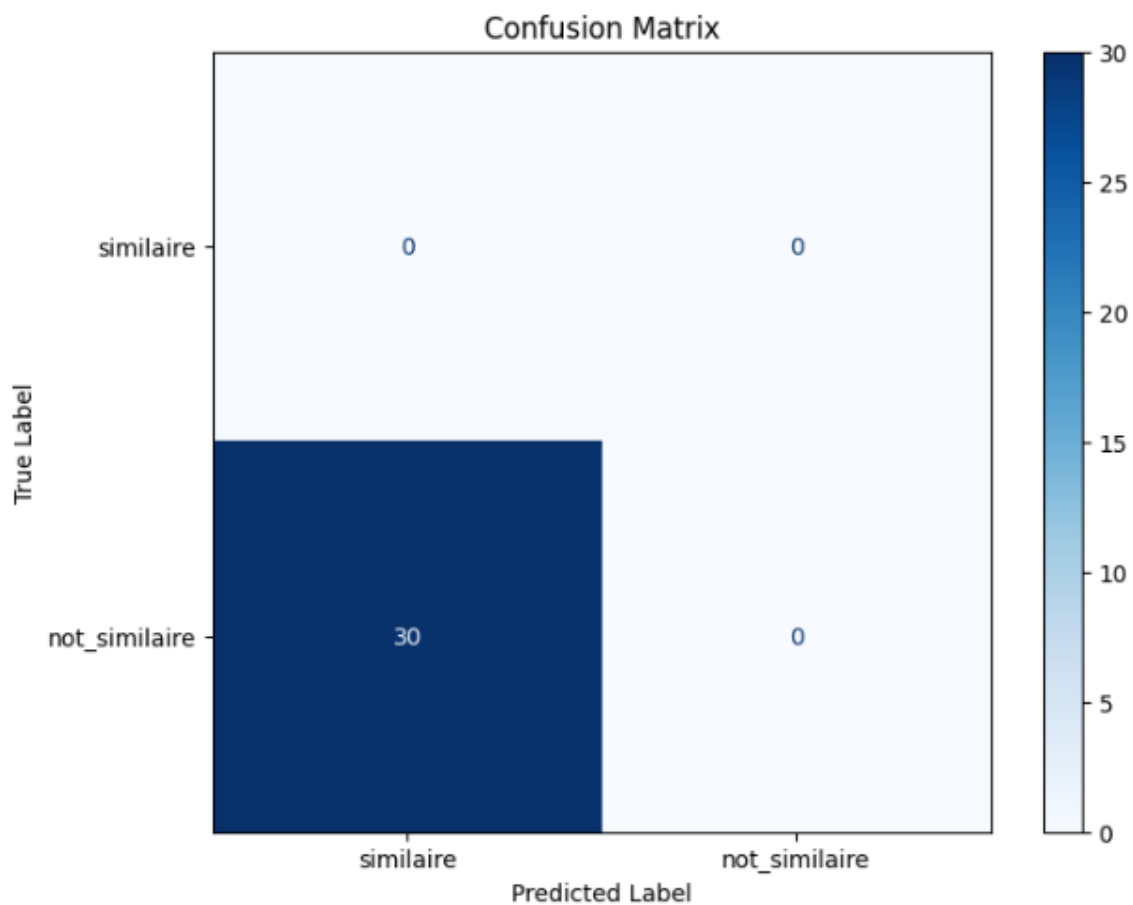


Figure 26. Matrix de confusion du modèle réalisé

4. Conclusion :

A la fin de ce chapitre nous avons atteint l'étude du modèle et l'explication des fonctions d'évaluation « loss & train ».

CONCLUSION GÉNÉRALE



CONCLUSION GÉNÉRALE

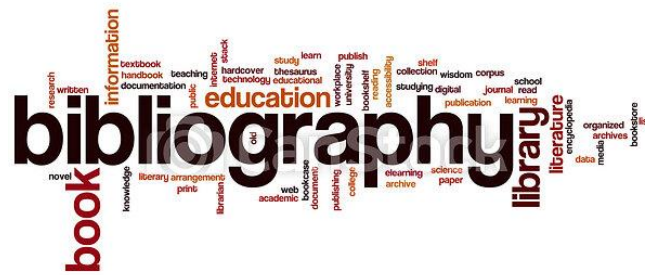
La recherche d'images par le contenu devient un problème. Compte tenu de la place qu'occupent les images numériques dans notre vie quotidienne, cela est complexe et inévitable.

Nous avons visé dans ce travail la réalisation d'un système de recherche des images similaires à une image requête. A cet effet, nous avons proposé un système de classification d'images en utilisant CNN comme un modèle de classification.

Les domaines de recherche d'images basés sur le contenu sont de très grands domaines et les techniques utilisées sont différentes. Il n'y a pas de loi qui l'exige. Sélectionnez une technique spécifique pour extraire des images en fonction de leur contenu.

Nous espérons que ce document a atteint notre objectif, mais ce n'est vraiment qu'une goutte d'eau dans le seau du sujet.

BIBLIOGRAPHIE & SITOGRAFIE



BIBLIOGRAPHIE

- [01]** : Mohammed KHAMADJA et Said BENIERBAH, Traitement d'images. Université des Frères Mentouri (Constantine).
- [02]** : ZERROUKI Houcemeddin et ZENNAKI Abderrahmane, Développement d'un système de recherche d'images par le contenu, Université Aboubakr Belkaïd– Tlemcen ,2018.
- [03]** : Khouloud MESKALDJI, Extraction et traitement de l'information : Un prototype d'un système de recherche d'images couleurs par le contenu magistère, Université Mentouri de Constantine, 2009.
- [04]** : Gonzales, R. C et Woods, R. E. Digital image processing. 2e édition. s.l. Prentice-Hall, Inc., 2002.
- [05]** : Kamel Houari Dr. Khoulladi Mohamed-Khireddine, Recherche d'images par le contenu, Université Mentouri Constantine, 2010.
- [06]** : Kaddour Chakib et Aissa Brahim Sali., La compression des images fixes par les approximations fractales basées sur la triangulation de Delaunay et la quantification vectorielle, Mémoire d'ingénieur. Université d'Alger. 1999.
- [07]** : Amira LAOUBI. Recherche d'images sémantique basée sur la sélection automatique des concepts. Mémoire en master. Université d'Ouargla. 2014.
- [08]** : BOUKHLOUF Djemaa. Résolution de problèmes par écosystèmes : Application au traitement d'images. Mémoire de Magister. Université de Biskra. 2005.
- [09]** : Chikh Mohammed Tahar. Amélioration des images par un modèle de réseau de neurones (Comparaison avec les filtres de base). Mémoire de master. Université de Tlemcen. 2011.
- [10]** : Bernard Gibaud, Olivier Hélénou, Joël Chabriaux et Denis Mariano Goulart. Traitement de l'image de la numérisation à l'archivage et la communication. Edition Elsevier Masson. 2013.
- [11]** : Aditya Vailaya, Mário Figueiredo, Anil Jain, Hong Jiang Zhang. A Bayesian Framework for Semantic Classification of Outdoor Vacation Images, IEEE Trans. Image Processing, Vol. 10, No. 1, pp. 157-172, Jan. 2001.
- [12]** : T. P. Minka An Image Database Browser that Learns from User Interaction, Master of Engineering Thesis, 1996.
- [13]** : Jia Li, James Z. Wang, "Studying digital imagery of ancient paintings by mixtures of stochastic models," IEEE Transactions on Image Processing, vol.

- [14] : Michael S. Lew, D. P. Huijsmans, and Dee Denteneer. Content based image retrieval: KLT, projections, or templates. In A. W. M. Smeulders and R. Jain, editors. Image Databases and Multi-Media Search, proceedings of the First International Workshop IDB-MMS'96, Amsterdam, The Netherlands. Amsterdam University Press, pages 27-34. , August 1996.
- [15] : Francis R. Bach & Michael I. Jordan. Learning Spectral Clustering. EECS Department, University of California, Berkeley. Technical Report No. UCB/CSD-03-1249. June 2003.
- [16] : BEN CHEIKH Noura et BEN BEZZIANE Rima. La recherche d'images par la sémantique. Mémoire de master. Université d'Ouargla. 2011.
- [17] : DUBEY, Rajshree S., CHOUBEY, Rajnish, et BHATTACHARJEE, Joy. Multi feature content-based image retrieval. International Journal on Computer Science and Engineering, 2010, vol. 2, no 6, p. 2145-2149.
- [18] : RUBNER, Yossi et TOMASI, Carlo. Perceptual metrics for image database navigation. Springer Science & Business Media, 2013.
- [19] : Haralik, R.M., Shanhu, K., et Dinstein, I. Textural features for images Classification. IEEE Transaction on System, Man, Cybernetics, 3,610-621, 1973.
- [20] : FOURNIER, Jérôme. Indexation d'images par le contenu et recherche interactive dans les bases généralistes. 2002. Thèse de doctorat. Cergy Pontoise.
- [21] : HADI Fairouz. Indexation des bases de données images. Thèse de Doctorat. Université Ferhat Abbas de Sétif 1.2021.

SITOGRAPHIE

- [S1] : http://fr.wikipedia.org/wiki/Image_num%C3%A9rique
- [S2] : <https://homepages.inf.ed.ac.uk/rbf/HIPR2/gryimage.htm>
- [S3] : <https://www.technoscience.net/definition/11202.html> 02/04/2023 (15:50)
- [S4] : <https://bu.umc.edu.dz/theses/informatique/HOU5734.pdf> 02/04/2023 (23:01)
- [S5] : https://www-sop.inria.fr/orion/personnel/Lan.Le_Thi/MemoireMaster/Indexation%20et%20recherche%20d'images%20par%20le%20contenu.pdf 04/04/2023 (03:45)