

Généralités sur la classification d'images

Chapitre I : Généralités sur la classification d'images

I.1 Introduction

La classification des images consiste à attribuer une classe à une image à l'aide d'un système de classification en se basant sur ce qu'on appelle l'apprentissage. On retrouve ainsi la classification d'objets, de scènes, de textures, la reconnaissance de visages, d'empreintes digitales et de caractères. Il existe deux principaux types d'apprentissage : l'apprentissage supervisé et l'apprentissage non-supervisé. Dans l'approche supervisée, chaque image est associée à une étiquette qui décrit sa classe d'appartenance. Dans l'approche non supervisée les données disponibles ne possèdent pas d'étiquettes. Dans notre travail on s'intéresse de l'approche supervisée.

I.2 notions de base

I.2.1 Définition d'une image

Une image est une représentation planaire d'une scène ou d'un objet situé en général dans un espace tridimensionnel, elle est issue du contact des rayons lumineux provenant des objets formant la scène avec un capteur (caméra, scanner, rayons X, ...). Il ne s'agit en réalité que d'une représentation spatiale de la lumière.

L'image est considérée comme un ensemble de points auquel est affectée une grandeur physique (luminance, couleur). Ces grandeurs peuvent être continues (image analogique) ou bien discrètes (images digitales). Mathématiquement, l'image représente une fonction continue IF , appelée fonction image, de deux variables spatiales représentée par $IF(x, y)$ mesurant la nuance du niveau de gris de l'image aux coordonnées (x, y) . [1]

I.2.2 Caractéristiques de l'image

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

I.2.2.1 Pixel

Le pixel est l'abréviation du mot « Picture élément », est une unité de surface permettant de définir la base d'une image numérique. Il matérialise un point donné (x, y) du plan de l'image. L'information présentée par le pixel est le niveau de gris (ou la couleur) prélevée à l'emplacement correspondant dans l'image réelle. La différence entre image monochrome et image couleur réside dans la quantité d'informations contenue dans chaque pixel, par exemple

Chapitre I : Généralités sur la classification d'images

dans une image couleur (RVB : Rouge, Vert, Bleu) la valeur d'un pixel est représentée sur trois octets pour chaque couleur.

I.2.2.2 Dimension & Résolution

La dimension est la taille de l'image. Elle se présente sous forme d'une matrice dont les éléments sont des valeurs numériques représentatives des intensités lumineuses (pixels). Le nombre de lignes de cette matrice multiplié par le nombre de colonnes nous donne le nombre total de pixels dans une image.

Par contre, la résolution est la clarté ou la finesse de détails atteinte par un moniteur ou une imprimante dans la production d'images. Sur les moniteurs d'ordinateur, la résolution est exprimée en nombre de pixels par unité de mesure (pouce ou centimètre). On utilise aussi le mot résolution pour désigner le nombre total de pixels horizontaux et verticaux sur un moniteur. Plus ce nombre est grand, plus la résolution est meilleure

I.2.2.3 Voisinage

Le plan de l'image est divisé en termes de formes rectangulaires ou hexagonales permettant ainsi l'exploitation de la notion de voisinage (voir figure I.1). Le voisinage d'un pixel est formé par l'ensemble des pixels qui se situent autour de ce même pixel. On définit aussi l'assiette comme étant l'ensemble de pixels définissant le voisinage pris en compte autour d'un pixel.

On distingue deux types de voisinage :

1-Voisinage à 4: On ne prend en considération que les pixels qui ont un coté commun avec le pixel considéré.

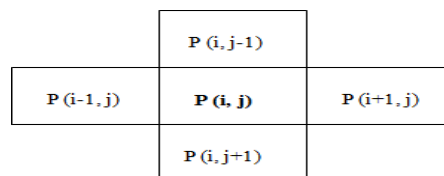


Figure I.1 Voisinage à 4

Chapitre I : Généralités sur la classification d'images

2-Voisinage à 8 : On prend en compte tous les pixels qui ont au moins un point en liaison avec le pixel considéré.

$P(i-1, j-1)$	$P(i, j-1)$	$P(i+1, j-1)$
$P(i-1, j)$	$P(i, j)$	$P(i+1, j)$
$P(i-1, j+1)$	$P(i, j+1)$	$P(i+1, j+1)$

Figure I.2 Voisinage à 8

I.2.2.4 Niveau de gris

C'est la valeur d'intensité lumineuse d'un pixel. Cette valeur peut aller du noir (0) jusqu'au blanc (255) en passant par les nuances qui sont contenues dans l'intervalle $[0, 255]$. Elle correspond en fait à la quantité de la lumière réfléchie.

Pour 8 bits, on dispose de 256 niveaux de gris dont 40 sont reconnus à l'œil nue. Plus le nombre de bit est grand plus les niveaux sont nombreux et plus la représentation est fidèle. [02]

I.2.2.5 Contraste

C'est l'opposition marquée entre deux régions d'une image. Une image contrastée présente une bonne dynamique de la distribution des valeurs de gris sur tout l'intervalle des valeurs possibles, avec des blancs bien clairs et des noirs profonds. Au contraire une image peu contrastée a une faible dynamique, la plupart des pixels ayant des valeurs de gris très proches.

Si $L1$ et $L2$ sont les degrés de luminosité respectivement de deux zones voisines $A1$ et $A2$ d'une image, le contraste est défini par le rapport : $C=L1-L2/L1+L2$

I.2.2.6 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.

Une bonne luminance se caractérise par :

- Des images lumineuses (brillantes) .
- Un bon contraste: il faut éviter les images où la gamme de contraste tend vers le blanc ou le noir; ces images entraînent des pertes de détails dans les zones sombres ou lumineuses.

Chapitre I : Généralités sur la classification d'images

- L'absence de parasites .

I.2.2.7 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. C'est un parasite qui représente certains défauts (poussière, petits nuages, baisse momentanée de l'intensité électrique sur les capteurs, ...etc.). Il se traduit par des taches de faible dimension et dont la distribution sur l'image est aléatoire. [03]

I.2.2.8 Contour

Les contours représentent la frontière entre les objets de l'image, ou la limite entre deux pixels dont les niveaux de gris représentant une différence significative.

Dans une image numérique, les contours se situent entre les pixels appartenant à des régions ayant des intensités moyennes différentes ; il s'agit de contours de type « saut d'amplitude ». Un contour peut également correspondre à une variation locale d'intensité présentant un maximum ou un minimum ; il s'agit alors de contour « en toit »

I.2.3 Les Types d'image

Il y a 3 types d'image :

Image couleur RVB : L'œil humain analyse la couleur à l'aide de trois types de cellules photo 'les cônes'. Ces cellules sont sensibles aux basses, moyennes, ou hautes fréquences (rouge, vert, bleu). Pour représenter la couleur d'un pixel, il faut donc donner trois nombres, qui correspondent au dosage de trois couleurs de base : Rouge, Vert, Bleu. On peut ainsi représenter une image couleur par trois matrices chacune correspondant à une couleur de base.



Figure I.3 couleur RVB

Chapitre I : Généralités sur la classification d'images

Image d'intensités :C'est une matrice dans laquelle chaque élément est un réel compris entre 0 (noir) et 1 (blanc). On parle aussi d'image en niveaux de gris, car les valeurs comprises entre 0 et 1 représentent les différents niveaux de gris.



Figure I.4 Image en niveaux de gris

Image binaire :Une image binaire est une matrice rectangulaire dans l'élément valent 0 ou 1. Lorsque l'on visualise une telle image, les 0 sont représentés par du noir et les 1 par du blanc.
[01]



Figure I.5 Image binaire

I.3 Méthodes de classification d'images

I.3.1 Définition de la classification

La classification est une tâche qui tente à identifier les objets en se basant sur certaines de leurs caractéristiques. Par exemple, pour pouvoir utiliser les images pour les analyses complémentaires ou pour la cartographie, il est souvent important de traduire l'information de fréquence contenue dans les images en information thématique portant sur l'occupation du sol ou

Chapitre I : Généralités sur la classification d'images

la couverture végétale. On a généralement le choix entre deux approches : la classification supervisée et non supervisée

I.3.2 Les différentes méthodes de la classification

I.3.2.1 Méthodes hiérarchiques

En classification hiérarchique ascendante. Le procédé consiste à grouper les observations individuelles en classes par a partie de la même classe.

Les méthodes se distinguent par le choix de la distance entre les observations et la définition de la stratégie d'agrégation .Dans l'algorithme de base, le calcul de la distance (il s'agit plus exactement d'une quantité critère que l'on appelle distance par abus de langage) fait par récurrence à partir de la matrice des distances entre observations.

I.3.2.2 Méthodes non hiérarchiques

La classification non hiérarchique ou partitionnement, aboutissant à la décomposition de l'ensemble de tous les individus en m ensemble disjoints ou classes d'équivalence, le nombre m de classes est fixé [04]. Le résultat obtenu est alors une partition de l'ensemble des individus, un ensemble de parties, ou classes de l'ensemble I des individus telles que :

- Toute classe soit non vide.
- Deux classes distinctes sont disjointes.
- Tout individu appartient à une classe.

Cet algorithme porte le nom de "agrégation autour de centres variables". Une version égèremment différente, connue sous le nom de "nuées dynamiques" consiste à représenter chaque groupe non pas par son centre, mais par un ensemble de points (noyau) choisis aléatoirement à l'intérieur de chaque groupe. [04]On calcule alors une distance "moyenne" entre chaque observation et ces noyaux et l'on procède à l'affectation.

I.3.2.3 Méthodes paramétriques

Un classifieur est dit paramétrique s'il associe à la signature spectrale (ou profil) une distribution statistique connue, le plus fréquemment pour le traitement d'images, la loi normale ou multi

Chapitre I : Généralités sur la classification d'images

normale. Cette association offre la possibilité d'affecter à chaque pixel une probabilité d'appartenance à une classe donnée

I.3.2.4 Méthodes non paramétriques

Un classifié est dit non paramétrique si aucune distribution statistique paramétrique n'est exploitée, seule la distance spectrale sera alors prise en compte. Cette catégorie comprend notamment les méthodes fondées sur la minimisation de distance (hyper boîte , la distance minimale et la distance de Mahalanobis, K plus proches voisins, K-means, ISODATA, etc.), de nouvelles méthodes apparues récemment s'ajoutent à cette catégorie comme les réseaux neuronaux et les Machines à Support Vecteurs (SVM)

I.3.2.5 Méthodes structurelles

Ce type de méthodes exploite des informations structurelles et contextuelle d'un objet, elles analysent l'objet en termes de ses composantes (primitives) et de leurs propriétés, on trouve par exemple l'analyse syntaxique d'une forme ou un objet à partir d'une grammaire, la distance d'arbres, la distance de graphes (isomorphismes de graphes, de sous-graphes, avec correction d'erreurs, etc.). Dans la méthode structurelle la classe se présente principalement sous la forme de petites régions rondes. [05]

I.3.2.6 Méthodes supervisées

Dans le cas de l'apprentissage supervisé, on dispose d'un ensemble de données étiquetées, ou d'exemples qui se sont vus associés une classe par un professeur ou un expert. Cet ensemble d'exemples constitue la base d'apprentissage .Les méthodes d'apprentissage supervisé se donnent alors comme objectif général déconstruire à partir de la base d'apprentissage, ou fonctions de classement. Une telle fonction permet, à partir de la description d'un objet, de reconnaître un attribut particulier, la classe (Figure I.6) [06]

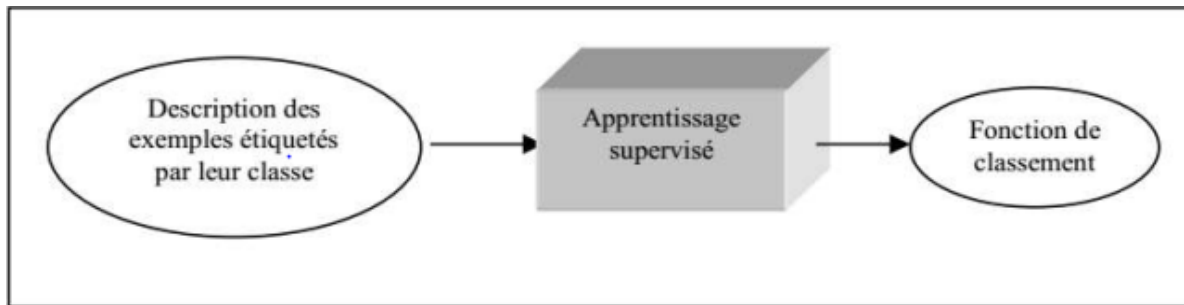


Figure I.6-L'apprentissage supervisé

l'inférence inductive est définie comme un processus qui à partir d'une connaissance spécifique observée sur certains objets et d'une hypothèse inductive initiale ,permet d'obtenir une assertion inductive impliquant ou rendant compte fortement ou faiblement des observations. Dans le cas de l'apprentissage inductif supervisé, qui est un sous domaine de l'inférence inductive, la connaissance spécifique consiste en un ensemble d'objets appartenant à des classes connues. L'assertion inductive est exprimée par une règle de classification qui assigne une classe à chaque objet. L'implication forte est satisfaite si la règle classe correctement tous les objets connus [07].

I.3.2.7 Méthodes non-supervisées

L'apprentissage non-supervisé, encore appelé apprentissage à partir d'observations ou découverte, consiste à déterminer une classification« sensée » à partir d'un ensemble d'objets ou de situations données (des exemples non étiquetés).On dispose d'une masse de données indifférenciées, et l'on désire savoir si elles possèdent une quelconque structure de groupes. Il s'agit d'identifier une éventuelle tendance des données à être regroupées en classes. Ce type d'apprentissage, encore appelé Cluster ING ou Cluster Analysais, se trouve en classification automatique et en taxinomie numérique. Cette forme de classification existe depuis des temps immémoriaux. Elle concerne notamment les sciences de la nature (Figure I.7), les classifications des documents et des livres mais également la classification des sciences élaborées au cours des siècles par les philosophes [08]

Chapitre I : Généralités sur la classification d'images

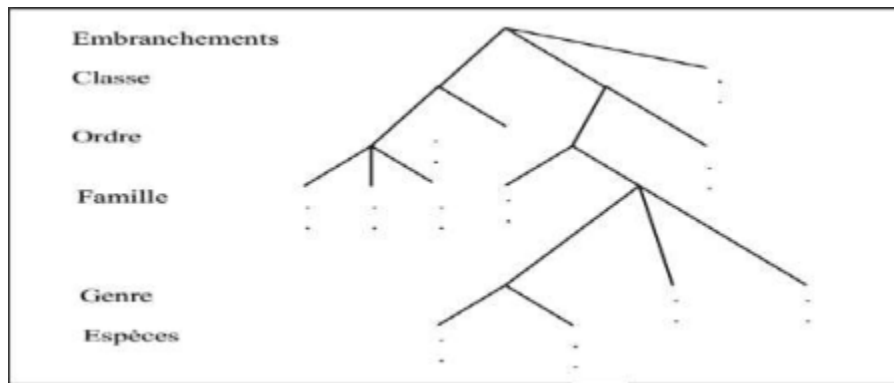


Figure I.7 Extrait de la classification taxinomique de Linné

L'automatisation de la construction de classification constitue aujourd'hui un véritable domaine de recherche. La notion clé utilisée pour créer des classes d'objets est une mesure de la similarité entre les objets. Les classes ou concepts sont construits de façon à maximiser la similarité intra-classes et à minimiser la similarité interclasses.

L'apprentissage non supervisé correspond également à la classification conceptuelle, où une collection d'objets forme une classe si cette classe peut être décrite par un concept, compte tenu d'un ensemble de concepts prédéfinis [09].

I.3.3 L'objectif de la classification

L'objectif de la classification d'images est d'élaborer un système capable d'affecter un classement automatique d'images. Ainsi, ce système permet d'effectuer une tâche d'expertise qui peut s'avérer coûteuse à acquérir pour un être humain en raison notamment de contraintes physiques comme la concentration, la fatigue et le temps nécessaire pour un volume important de données images.

I.3.4 Domaines d'application de la classification

La classification joue un rôle important dans toutes les sciences et techniques qui font appel à la statistique multidimensionnelle. Citons tout d'abord les sciences biologiques : botanique, zoologie, écologie,... Ces sciences utilisent également le terme de "taxinomie" pour désigner l'art de la classification. De même les sciences de la terre et des eaux : géologie, pédologie, géographie, étude des pollutions, font grand usage de classifications

Chapitre I : Généralités sur la classification d'images

I.3.5 Présentation de certaines techniques de la classification

I.3.5.1 k plus proches voisin

L'algorithme KNN figure parmi les plus simples algorithmes d'apprentissage artificiel. Dans un contexte de classification d'une nouvelle observation x , l'idée fondatrice simple est de faire voter les plus proches voisins de cette observation. La classe de x est déterminée en fonction de la classe majoritaire parmi les k plus proches voisins de l'observation x . Donc la méthode du plus proche voisin est une méthode non paramétrique où une nouvelle observation est classée dans la classe d'appartenance de l'observation de l'échantillon d'apprentissage qui lui est la plus proche, au regard des covariables utilisées. La détermination de leur similarité est basée sur des mesures de distance.

I.3.5.2 k-means

L'algorithme k-means est l'algorithme de regroupement le plus connu et le plus utilisé, du fait de sa simplicité de mise en œuvre. Il partitionne les données d'une image en K clusters. Contrairement à d'autres méthodes dites hiérarchiques, qui créent une structure en « arbre de clusters » pour décrire les groupements, k-means ne crée qu'un seul niveau de clusters. L'algorithme renvoie une partition des données, dans laquelle les objets à l'intérieur de chaque cluster sont aussi proches que possible les uns des autres et aussi loin que possible des objets des autres clusters. Chaque cluster de la partition est défini par ses objets et son centroïde. Le k-means est un algorithme itératif qui minimise la somme des distances entre chaque objet et le centroïde de son cluster. La position initiale des centroïdes conditionne le résultat final, de sorte que les centroïdes doivent être initialement placés le plus loin possible les uns des autres de façon à optimiser l'algorithme. K-means change les objets de cluster jusqu'à ce que la somme ne puisse plus diminuer. Le résultat est un ensemble de clusters compacts et clairement séparés, sous réserve qu'on ait choisi la bonne valeur K du nombre de clusters. Les principales étapes de l'algorithme k-means sont :

1. Choix aléatoire de la position initiale des K clusters.
2. (Ré-Affecter les objets à un cluster suivant un critère de minimisation des distances (généralement selon une mesure de distance euclidienne).

Chapitre I : Généralités sur la classification d'images

3. Une fois tous les objets placés, recalculer les K centroïdes.
4. Réitérer les étapes 2 et 3 jusqu'à ce que plus aucune réaffectation ne soit faite.

I.3.5.3 Fuzzy c-means

Fuzzy C-Means(FCM) est un algorithme de classification non-supervisée floue. Issu de l'algorithme des C-moyennes (C-means), il introduit la notion d'ensemble flou dans la définition des classes : chaque point dans l'ensemble des données appartient à chaque cluster avec un certain degré, et tous les clusters sont caractérisés par leur centre de gravité. Comme les autres algorithmes de classification non supervisée, il utilise un critère de minimisation des distances intra-classe et de maximisation des distances interclasse, mais en donnant un certain degré d'appartenance à chaque classe pour chaque pixel. Cet algorithme nécessite la connaissance préalable du nombre de clusters et génère les classes par un processus itératif en minimisant une fonction objective. Ainsi, il permet d'obtenir une partition floue de l'image en donnant à chaque pixel un degré d'appartenance (compris entre 0 et 1) à une classe donnée. Le cluster auquel est associé un pixel est celui dont le degré d'appartenance sera le plus élevé. Les principales étapes de l'algorithme Fuzzy C-means sont :

1. La fixation arbitraire d'une matrice d'appartenance.
2. Le calcul des centroïdes des classes.
3. Le réajustement de la matrice d'appartenance suivant la position des centroïdes.
4. Calcul du critère de minimisation et retour à l'étape 2 s'il y a non convergence de critère.

I.3.5.4 Machine à vecteurs de support

Les machines à vecteurs de support ou séparateurs à vaste marge (en anglais Support Vector Machine, SVM) sont un ensemble de techniques d'apprentissage supervisé destinées à résoudre des problèmes de classification, en cherchant à séparer les classes d'une façon optimale. Les SVM ont été développés dans les années 1990 à partir des considérations théoriques de Vladimir Vapnik sur le développement d'une théorie statistique de l'apprentissage : la Théorie de Vapnik Chervonenkis. Les SVM ont rapidement été adoptés pour leur capacité à travailler avec des données de grandes dimensions, le faible nombre d'hyper paramètres, leurs garanties théoriques, et leurs bons résultats en pratique. Les SVM ont été appliqués à de très nombreux domaines (bioinformatique, recherche d'information, vision par ordinateur, finance...). Selon les données,

Chapitre I : Généralités sur la classification d'images

la performance des machines à vecteurs de support est de même ordre, ou même supérieure, à celle d'un réseau de neurones ou d'un modèle de mixture gaussienne.

L'hyperplan optimal qui sépare les deux classes est le lieu des points x satisfaisant $\langle w, x \rangle + b = 0$ où w et b sont des paramètres à déterminer. On voit que le vecteur w définit la pente de l'hyperplan (w est perpendiculaire à l'hyperplan), le terme b quant à lui permet de translater l'hyperplan parallèlement à lui-même. L'obtention de cet hyperplan est basée sur la recherche de deux autres hyperplans canoniques passant par les vecteurs supports. La règle de décision $h(x)$ consiste à déterminer l'appartenance d'une forme x à une classe en observant de quel côté de l'hyperplan se trouve x .

I.4 conclusion

Nous avons consacré ce chapitre à la présentation des notions des bases d'images et leur classification et on a présenté aussi certaines techniques de la classification. Dans le deuxième chapitre on va bien détailler les réseaux de neurones et plus précisément l'utilisation des réseaux de neurones convolutionnels dans la classification des images.

Chapitre II

Les réseaux de neurones

Chapitre II : Les réseaux de neurones

II.1 Introduction

Les réseaux de neurones sont des structures cellulaires artificielles et constituent une approche permettant d'aborder sous des angles nouveaux les problèmes de perception, de mémoire, d'apprentissage et de raisonnement (en d'autres termes... d'intelligence artificielle ou abrégée "I.A.") au même titre que les algorithmes génétiques. Ils s'avèrent aussi des alternatives très prometteuses pour contourner certaines des limitations des ordinateurs classiques. Grâce à leur traitement parallèle de l'information et à leurs mécanismes inspirés des cellules nerveuses (neurones), ils infèrent des Propriétés émergentes permettant de solutionner des problèmes jadis qualifiés de complexes. [10]

De nombreux termes sont utilisés pour désigner le domaine des réseaux de neurones artificiels, comme connexionnisme ou neuromimétique. Connexionnisme et neuromimétique sont tous deux des domaines de recherche à part entière, qui manipulent chacun des modèles de réseaux de neurones artificiels, mais avec des objectifs différents [11]

II.2 Le neurone formel

Un neurone formel est un modèle mathématique. Il est utilisé dans le cadre des recherches sur l'intelligence artificielle, principalement en association avec d'autres neurones formels pour former un réseau de neurones. [12]

Il contient plusieurs formules mathématiques afin de reproduire le plus fidèlement possible le fonctionnement d'un neurone avec ses différentes entrées (dendrites), leurs importances (coefficient de pondération) et sa sortie (axone) et ainsi comprendre son potentiel d'interaction avec les autres neurones.

Il est caractérisé par :

- Des entrées e_i et leur poids synaptique w_i
- Une fonction d'activation (de transfert) f
- Une sortie y