

UNIVERSITE KASDI MERBAH OUARGLA
**Faculté des Nouvelles Technologies de l'Information et de
la Communication**
Département d'Informatique et des Technologies de l'information



Mémoire Master Professionnel

Domaine : Informatique et Technologie de l'Information

Filière : Informatique

Spécialité : fondamentale

Présenté par : Boughaba Mohammed et Boukhris Brahim

Thème

***L'apprentissage profond (Deep Learning)
pour la classification et la recherche
d'images par le contenu***

Président	Mr Bekkari Fouad	Maître assistant A, UKM Ouargla
Examineur	Mr Zitouni Farouk	Maître assistant B, UKM Ouargla
Rapporteur	Mr Meflah Mohammed Salim	Maître assistant A, UKM Ouargla

Année universitaire 2016/2017

Remerciements

-Nous tenons à remercier tout d'abord notre encadreur : M^r **Meflah mohammed salim**, pour tous ses conseils, son soutien, son aide, et son orientation.

-Nous remercions les membres de jury M^r **Bekkari Fouad** et M^r **Zitouni Farouk** d'avoir accepté de juger notre travail.

-Nos remerciements s'adressent aussi à notre ami égyptien **Makhlouf Fadi Younis** pour son aide précieuse durant notre projet,

-Nous remercions tous les enseignants du département d'informatique et des technologies de l'information, aussi notre promotion 2016-2017 Master informatique fondamentale.

-Nous remercions toutes les personnes qui nous ont soutenues de près ou de loin dans la réalisation de ce travail.

Dédicaces

Nous dédions notre travail à :

Les deux êtres les plus chères au monde, les bougies
qui m'ont toujours guidé vers le bon chemin :

Ma mère, la lumière de notre vie, qui a tout fait pour
notre réussite et notre bonheur.

Mon père, à lui, nous devons offrir tout le respect et
l'amour pour son soutien et sa tendresse.

A nos sœurs pour leur aide et courage
A nos frères, sœurs
À mon ami MOHAMMED,
À mon ami Brahim

A tous les membres de la famille et à toutes les amies.

BRAHIM & MOHAMED

Table desmatières

Introduction générale	7
A. Contexte et objectif :	7
B. Organisation du mémoire :	8
II. Chapitre I : Recherche et classification d'images : Notions, Méthodes	10
I .1 - Introduction :	10
I .2-Principe de la recherche d'image :	10
I .3- approches de représentation et de recherche d'images :	10
I .3-1 Approches Basées sur le contexte :	11
I .3-2 Approches Basée sur le contenu :	11
I .4 –Définition de la classification :	11
I .5-L'objectif de la classification :	11
I .6- Domaines d'application de la classification :	12
I .7-Les différentes méthodes de la classification et l'apprentissage :	12
I .7-1 Méthodes supervisées :	12
I .7-2 Méthodes non-supervisées :	13
I .7-3 Méthodes hiérarchiques :	13
I .7-4 Méthodes non hiérarchiques :	14
I .7-5 Méthodes non paramétriques :	14
I .7-6 Méthodes paramétriques :	14
I .8- Présentation de certaines techniques de la classification :	15
I .8-1 k plus proches voisin :	15
I .8-2 Fuzzy c-means:	15
I .8-3 k-means :	16
I .8.4 Machine à vecteurs de support :	16
I .9- Présentation de quelques Types de réseaux neuronaux :	17
I .9-1 Neurone Formel :	17
I .9-2 Neurones multicouche :	18
I .9-3 Neurones récurrents :	18
I .9-4 Réseaux de Hopfield :	18
I .9-5-Réseaux Neurones Convolutionnels :	19
I .10-Conclusion :	20
III. Chapitre II : Deep learning (L'apprentissage profond).....	22
II .1- Introduction :	22
II .2- Définition de l'apprentissage profond (deep learning) :	22
II.2- Domaines d'application de l'apprentissage profonde :	22

II.3- Architectures de réseaux de neurones profonds :.....	23
II.4- Quelques explications sur les CNNs :.....	23
II.4.1- Principe d'architecture d'un CNN :	23
II.4.2-L'entraînement d'un nouveau CNN est difficile :.....	24
II.4.3-Les blocs de construction :.....	24
II.4.3.1-La couche de convolution (CONV) :	24
II.4.3.2-Couche de pooling (POOL) :	25
II.4.3.3-Couche de correction (ReLU):	26
II.4.3.3-Couche entièrement connectée(FC):	27
II.4.3.3-Couche de perte (LOSS) :	27
II.4.4-Choix des hyperparamètres :.....	27
II.4.4.1-Nombre de filtres :	27
II.4.4.2-Forme de filtres :	28
II.4.4.3-Forme de Max pooling :.....	28
II.5-Avantages de CNNs :.....	28
II.6-Conclusion :	29
Chapitre III: Conception et implémentation:	31
III.1- Introduction :	31
III.2-Présentation des outils de développement :	31
III.2.1- Matériel :.....	31
III.2.2- Matlab :.....	31
III.2.3- Visual Studio.NET :.....	32
III.2.4- MatConvNet :	32
III.3- Présentation de l'application :.....	33
III.3.1- Création des classes :.....	33
III.3.2- l'apprentissage :.....	33
III.3.3- Application :	35
III.3.4- La Classification et recherche :.....	36
III.3.5- Présentation des méthodes comparaison :.....	37
III.3.6- Tests et résultats :.....	37
III.3.7- Discussion des résultats :.....	39
III.4-Conclusion :.....	39
Conclusion générale :.....	40
Bibliographie.....	41

Liste des figures :

Figure 1-Le principe général de la recherche d'image par le contenu.....	11
Figure 2-L'apprentissage supervisé.	12
Figure 3-Extrait de la classification taxinomique de Linné.	13
Figure 4-perceptron multicouche.	18
Figure 5-Les réseaux de neurones convolutifs.....	23
Figure 6-Ensemble de neurones (cercles) créant la profondeur d'une couche de convolution (bleu). Ils sont liés à un même champ récepteur (rouge).....	25
Figure 7-Pooling avec un filtre 2x2 et un pas de 2.....	26
Figure 8-exemple de max pooling (2x2).....	28
Figure 9-Matlab environnement.....	32
Figure 10-création de classe exemples animaux.....	33
Figure 11-redimensionne toutes les images.....	34
Figure 12-pourcentage d'erreur 500 epoche.....	34
Figure 13-interface d'application.....	35
Figure 14- sélectionnez un fichier.....	35
Figure 15-Affichage de l'image sélectionnée.....	36
Figure 16-Rechercher les images.....	36
Figure 17-Recherche par la classification.....	37
Figure 18-Comparaison des résultats obtenus par chaque méthode.....	38

Introduction générale

A. Contexte et objectif :

L'homme fait recours toujours à la classification dans sa vie, est-il essaie de répondre aux problèmes et questions sur les catégories des objets, c'est-à-dire réalisé l'affectation d'objets à leurs classes (formats, couleurs, tailles . . . etc.).

Généralement des bases d'observations caractérisent un domaine particulier (animaux, fruit, malades, gènes, . . . etc.), où on les regroupe en plusieurs classes.

La classification automatique d'images est une application de la reconnaissance de formes, qui consiste à attribuer automatiquement une classe à une image à l'aide d'un système de classification.

On retrouve ainsi, la classification d'objets, de textures, la reconnaissance de visages, d'empreintes digitales, de caractères parmi les applications courantes.

La classification des images, permet de classifier une base d'images en plusieurs classes comme par exemple (classe animaux, classe humaine, classe transport,...etc.), cette catégorisation permet une meilleure exploitation de cette base par des applications comme la recherche d'images à titre indicatif.

Les systèmes de recherche d'images cherchent à reconnaître les images similaires à une image requête, et cherche à obtenir de bonnes performances en moyenne sur toutes les images.

Parmi les problèmes rencontrés lors de la manipulation de grandes quantités d'images est la structuration, et la recherche. De ce fait l'utilisation de la classification peut contribuer à réduire la taille de ces problèmes.

Le deep learning est une technique d'apprentissage permettant à un programme, par exemple, de reconnaître le contenu d'une image ou de comprendre le langage parlé « *La technologie du deep learning apprend à représenter le monde. C'est-à-dire comment la machine va représenter la parole ou l'image par exemple* ».

Ce système d'apprentissage et de classification, basé sur des « réseaux de neurones artificiels » numériques est une technique courante en IA, permettant aux machines d'apprendre et reconnaître des objets. Il approfondit sa compréhension de l'image avec des concepts de plus en plus précis.

Notre travail s'inscrit dans le cadre de la classification des images en utilisant le Deep Learning, on vise à classer les images en plusieurs classes. Le résultat de la classification sera exploité pour rechercher des images similaires à une image requête. Ce travail a comme objectif de simplifier la tâche aux non spécialistes du domaine, ainsi simplifier la recherche d'*images*.

B. Organisation du mémoire :

Notre travail est divisé en deux parties, dans **la première partie** on fait un état de l'art qui donne une vision générale sur les méthodes de classification, ainsi dans le **premier chapitre** on présente la recherche d'images et la classification. Le **deuxième chapitre** est consacré, au deep learning, une description plus détaillée sur les réseaux de neurones convolutifs (CNNs) qui est la méthode choisie dans notre projet.

Dans la **deuxième partie**, on représente la conception de notre application, ainsi que la méthode d'implémentation de notre travail, on expliquant l'ensemble des choix techniques, (langage de programmation Matlab) utilisés pour la réalisation de cette application.

Chapitre I:

Recherche et Classification d'images : Notions, Méthodes

...

II. Chapitre I : Recherche et classification d'images : Notions, Méthodes ...

I.1 - Introduction :

La recherche d'images est un domaine informatique pour la navigation, recherche et extraction des images d'une grande base d'images numériques, c'est une recherche de données spécialisées utilisées pour trouver des images.

Le recours aux méthodes de la classification, pour regrouper des images en plusieurs classes est important dans notre l'étude.

Dans ce chapitre nous allons présenter tout d'abord les notions en relation avec la recherche et la classification d'images : types d'apprentissages, méthodes de classification, domaine d'application,...etc.

I.2-Principe de la recherche d'image :

La recherche d'images est un domaine informatique pour la navigation, recherche et extraction des images d'une grande base de données d'images numériques, c'est une recherche de données spécialisées utilisées pour trouver des images. Pour rechercher des images, un utilisateur peut fournir des termes de recherche tels que les mots-clés, cliquez sur une image...etc. Le système retourne les images "similaires" à la requête. La similitude utilisée pour les critères de recherche pourrait être entre les chaînes de caractères, la distribution des couleurs dans les images, la région d'attributs de forme et des concepts sémantique...etc. (1)

La recherche d'images est extrêmement active qui connaît depuis la dernière décennie un développement majeur tant au niveau académique qu'industriel. (1)

I.3- approches de représentation et de recherche d'images :

La recherche d'images est un domaine de recherche très actif au sein de deux grandes communautés de recherche : gestion de bases de données et la vision et infographie. Ces deux communautés de recherche étudient la recherche des images sous deux approches différentes : l'une emploie des annotations textuelles et l'autre emploie des descripteurs de bas ou haut niveaux extraits à partir des images.

La première approche a été proposée en 1970s, elle est basée sur l'annotation textuelle manuelle d'images, aujourd'hui la plus employée.

La deuxième approche a été introduite dans le début des années 1980s. Elle a deux directions, la première est la recherche d'image basée sur le contenu symbolique et la deuxième² est basée sur le contenu sémantique.

I.3-1 Approches Basées sur le contexte :

Le contexte concerne l'ensemble des informations autour d'une image permettant d'orienter sa signification, [Richard. C, 2001] la recherche basé sur le contexte consiste à chercher dans l'environnement global de l'image sans aborder le contenu, l'image est considérée comme une boîte noire.

I. 3-2 Approches Basée sur le contenu :

Le principe général de la recherche d'image par le contenu se déroule en deux phases. Lors d'une première phase le système décrit le contenu des images. Lors de la seconde phase, l'utilisateur interroge la base à l'aide d'une requête. Le système recherche les images de la base qui corresponde à la demande de l'utilisateur.

Elle est constituée de deux directions selon leurs niveaux de représentation du contenu des images. (Figure.1)

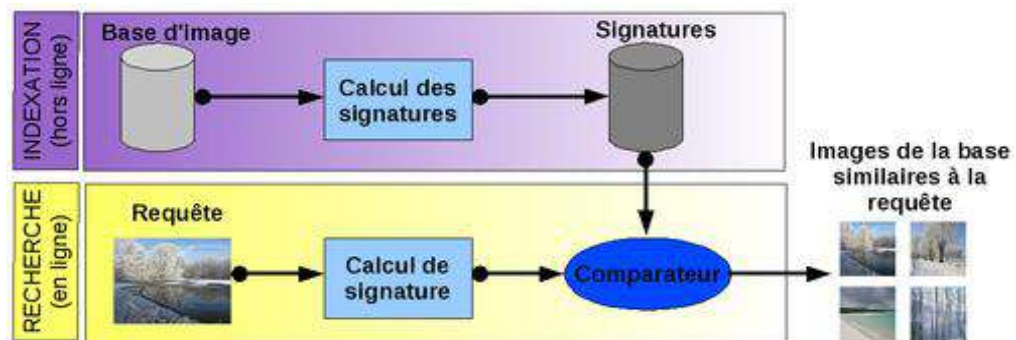


Figure 1-Le principe général de la recherche d'image par le contenu

I.4 –Définition de la classification :

Classifier une image est une tâche ou une série de méthodes qu'une théorie unifiée 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 la couverture végétale.

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

I.5-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.6- 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.

I.7-Les différentes méthodes de la classification et l'apprentissage :

I.7-1 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 de construire 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.2) (2)

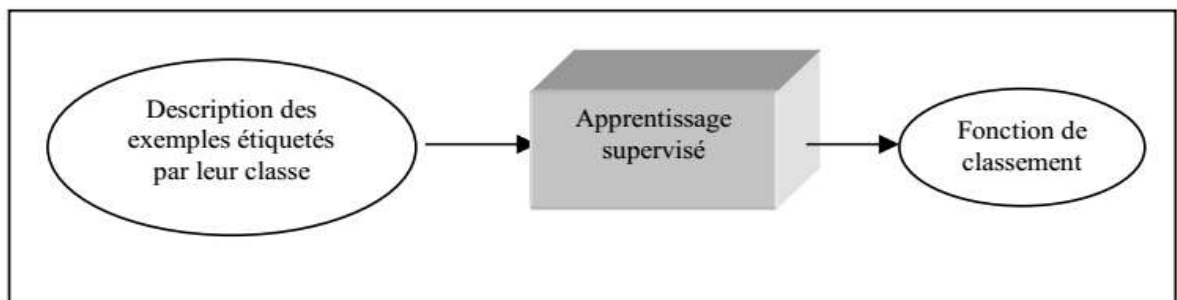


Figure 2-L'apprentissage supervisé.

Dans [3], 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.

I.7-2 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 Clustering ou Cluster Analysis, 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 3), les classifications des documents et des livres mais également la classification des sciences élaborées au cours des siècles par les philosophes (4).

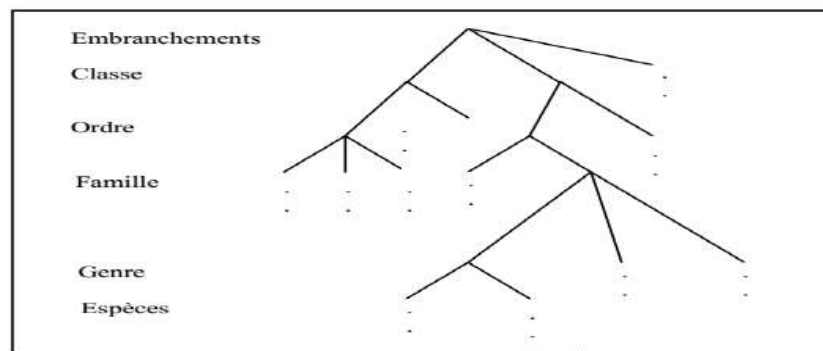


Figure 3-Extrait de la classification taxinomique de Linné.

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 (5).

I.7-3 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.7-4 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 (6). 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 légèrement 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. (6)

On calcule alors une distance "moyenne" entre chaque observation et ces noyaux et l'on procède à l'affectation.

I.7-5 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 ou parallélépipédique, 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.7-6 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 normale. Cette association offre la possibilité d'affecter à chaque pixel une probabilité d'appartenance à une classe donnée (7).

I.7-7 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 structurale la classe se présente principalement sous la forme de petites régions rondes. (8)

I.8- Présentation de certaines techniques de la classification :

I.8-1 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. [kpv15].

I.8-2 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.

[EMC15].

I .8-3 k-means :

L'algorithme *k-means* est l'algorithme de regroupement le plus connu et le plus utilisé, du fait de sa simplicité de mise en œuvre. Il partitionne les données d'une image en K clusters.

Contrairement à d'autres méthodes dites hiérarchiques, qui créent 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).
3. Une fois tous les objets placés, recalculer les K centroïdes.
4. Répéter les étapes 2 et 3 jusqu'à ce que plus aucune réaffectation ne soit faite. [EMC 15]

I .8.4 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.

les SVM sont une généralisation des classifiées linéaires. [1] 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 (bio-informatique, recherche d'information, vision par ordinateur, finance...). Selon les données, la performance des

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

Hyperplan qui est le lieu des points x satisfaisant $\langle w, x \rangle + b = 0$. En orientant l'hyperplan, la règle de décision correspond à observer de quel côté de l'hyperplan se trouve l'exemple x . 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.

;_Décision $h(s)$. La classe de tous les hyperplans qui en découle sera notée H .

I.8-5 Les Réseaux de Neurones :

Les réseaux de neurones proposent une simulation du fonctionnement de la cellule nerveuse à l'aide d'un automate : le neurone formel. Les réseaux neuronaux sont constitués d'un ensemble de neurones (nœuds) connectés entre eux par des liens qui permettent de propager les signaux de neurone à neurone.

Grâce à leur capacité d'apprentissage, les réseaux neuronaux permettent de découvrir des relations complexes non-linéaires entre un grand nombre de variables, sans intervention externe. De ce fait, ils sont largement utilisés dans de nombreux problèmes de classification (ciblage marketing, reconnaissance de formes, traitement de signal,...) d'estimation (modélisation de phénomènes complexes,...) et prévision (bourse, ventes,...). Il existe un compromis entre clarté du modèle et pouvoir prédictif. Plus un modèle est simple, plus il sera facile à comprendre, mais moins il sera capable de prendre en compte des dépendances trop variées.

I.9- Présentation de quelques Types de réseaux neuronaux :

Il existe beaucoup de types de réseaux neurones, chaque type étant développé pour un objectif particulier.

I.9-1 Neurone Formel :

Un neurone formel est une représentation mathématique et informatique d'un neurone biologique. Le neurone formel possède généralement plusieurs entrées et une sortie qui correspondent respectivement aux dendrites et au cône d'émergence du neurone biologique (point de départ de l'axone). Les actions excitatrices et inhibitrices des synapses sont représentées, la plupart du temps, par des coefficients numériques (les poids synaptiques) associés aux entrées. Les valeurs numériques de ces coefficients sont ajustées dans une phase d'apprentissage. Dans sa version la plus simple, un neurone formel calcule la somme pondérée des entrées reçues, puis applique à cette valeur une fonction d'activation, généralement non linéaire. La valeur finale obtenue est la sortie du neurone.

Le neurone formel est l'unité élémentaire des réseaux de neurones artificiels dans lesquels il est associé à ses semblables pour calculer des fonctions arbitrairement complexes, utilisées pour diverses applications en intelligence artificielle. [(9)

I.9-2 Neurones multicouche :

Le perceptron multicouche (*multi layer perceptron* MLP) est un classifieur linéaire de type réseau neuronal formel organisé en plusieurs couches (Figure .1.4) au sein desquelles une information circule de la couche d'entrée vers la couche de sortie uniquement ; il s'agit donc d'un réseau de type feedforward (en). Chaque couche est constituée d'un nombre variable de neurones, les neurones de la couche de sortie correspondant toujours aux sorties du système (10).

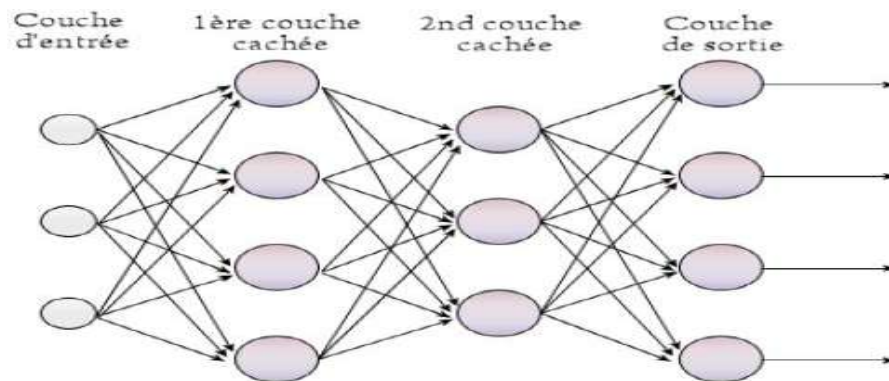


Figure 4-Perceptron multicouche.

I.9-3 Neurones récurrents :

Les réseaux de Neurones récurrents (RNNs) permettent d'analyser les séquences de vecteurs tout comme les modèles de Markov cachés .Le temps entre ici en ligne de compte car les sorties (de la coche de sortie et/ou de la coche cachées) calculées à l'instant t sont ré-injectées en entrée du réseau et/ou en entrée de la couche cachée. On peut en théorie conserver dans le réseau la mémoire de ce qui s'y est passé depuis le début.

I.9-4 Réseaux de Hopfield :

Le réseau de neurones d'Hopfield est un modèle de réseau de neurones récurrents à temps discret dont la matrice des connexions est symétrique et nulle sur la diagonale et où la dynamique est asynchrone (un seul neurone est mis à jour à chaque unité de temps). Il a été découvert par le physicien John Hopfield en 1982. Sa découverte a permis de relancer l'intérêt dans les réseaux de neurones qui s'était essoufflé durant les années 1970 à la suite d'un article de Marvin Minsky et Seymour Papert.

Un réseau de Hopfield est une mémoire adressable par son contenu : une forme mémorisée est retrouvée par une stabilisation du réseau, s'il a été stimulé par une partie adéquate de cette forme.

I.9-5-Réseaux Neurones Convolutionnels :

En apprentissage automatique, un réseau de neurone convolutifs (ou réseau de neurones à convolution, ou CNN ou ConvNet) est un type de réseau de neurones artificiels acycliques dans lequel le motif de connexion entre les neurones est inspiré par le cortex visuel des animaux. Les neurones de cette région du cerveau sont arrangés de sorte à ce qu'ils correspondent à des régions qui se chevauchent lors du pavage du champ visuel. Leur fonctionnement est inspiré par les processus biologiques, ils consistent en un empilage multicouche de perceptrons, dont le but est de prétraiter³ de petites quantités d'informations. Les réseaux neuronaux convolutifs ont de larges applications dans la reconnaissance d'image et vidéo, les systèmes de recommandation et le traitement du langage naturel.

I .10-Conclusion :

Dans ce chapitre, on a introduit les notions de recherche d`images et on a donné une vision générale sur les méthodes ou approches de la classification (supervisées non supervisées ...etc). Dans ce chapitre on a présenté aussi quelques algorithmes ou méthodes de classification (kNN, Machine à Vecteur Support, Fuzzy c-means, K-means, Réseaux de neurones). Dans le prochain chapitre, nous allons voir le principe de l`apprentissage profond (Deep learning) et les réseaux de neurones convolutifs (CNNs).



Chapitre II:

Deep learning (L'apprentissage profond).

III. Chapitre II : Deep learning (L'apprentissage profond).

II .1- Introduction :

L'intelligence artificielle est une discipline scientifique recherchant des méthodes de solution de problèmes à forte complexité logique ou algorithmique. L'apprentissage automatique champ d'étude de l'intelligence artificielle. Par conséquent, L'apprentissage profond (en anglais deep learning, deep structured learning, hierarchical learning) est un ensemble de méthodes d'apprentissage automatique tentant de modéliser avec un haut niveau d'abstraction des données grâce à des architectures articulées de différentes transformations non linéaires.

Dans ce chapitre nous allons présenter tout d'abord les notions en relation avec l'apprentissage profond.

II .2- Définition de l'apprentissage profond (deep learning) :

L'apprentissage profond (« *deep learning* ») est un ensemble de techniques d'apprentissage automatique qui a permis des avancées importantes en intelligence artificielle dans les dernières années.

Dans l'apprentissage automatique, un programme analyse un ensemble de données afin de tirer des règles qui permettront de tirer des conclusions sur de nouvelles données.

L'apprentissage profond est basé sur ce qui a été appelé, par analogie, des « réseaux de neurones artificiels », composés de milliers d'unités (les « neurones ») qui effectuent chacune de petites opérations simples. Les résultats d'une première couche de « neurones » servent d'entrée aux calculs d'une deuxième couche et ainsi de suite.

Par exemple, pour la reconnaissance visuelle, des premières couches d'unités identifient des lignes, des courbes, des angles... des couches supérieures identifient des formes, des combinaisons de formes, des objets, des contextes...

Les progrès de l'apprentissage profond ont été possibles notamment grâce à l'augmentation de la puissance des ordinateurs et au développement de grandes bases de données (« *big data* »).

II.2- Domaines d'application de l'apprentissage profonde :

Ces techniques se développent dans le domaine de l'informatique appliquée aux NTIC (reconnaissance visuelle — par exemple d'un panneau de signalisation par un robot ou une voiture autonome — et vocale notamment) à la robotique, à la bio-informatique, la reconnaissance ou comparaison de formes, la sécurité, la santé, etc..., la pédagogie assistée par l'informatique, et plus généralement à l'intelligence artificielle. L'apprentissage profond peut par exemple permettre à un ordinateur de mieux reconnaître des objets hautement déformables et/ou analyser par exemple les émotions révélées par un visage photographié ou filmé, ou analyser les mouvements et position des doigts d'une main, ce qui peut être utile pour traduire le langage des signes, améliorer le positionnement automatique d'une caméra, etc... Elles sont utilisées pour certaines formes d'aide au diagnostic médical (ex. : reconnaissance automatique d'un cancer en imagerie médicale), ou de prospective ou de prédiction (ex. : prédiction des propriétés d'un sol filmé par un robot).

II.3- Architectures de réseaux de neurones profonds :

Il existe un grand nombre de variables d'architectures profondes. La plupart d'entre eux sont dérivés de certaines architectures originales. Nous allons choisir les réseaux de neurones convolutifs (CNNs).

II.4- Quelques explications sur les CNNs :

II.4.1- Principe d'architecture d'un CNN :

Les réseaux de neurones convolutifs sont à ce jour les modèles les plus performants pour classer des images. Désignés par l'acronyme CNN, de l'anglais Convolutional Neural Network, ils comportent deux parties bien distinctes. En entrée, une image est fournie sous la forme d'une matrice de pixels. Elle a deux dimensions pour une image aux niveaux de gris. La couleur est représentée par une troisième dimension, de profondeur 3 pour représenter les couleurs fondamentales [Rouge, Vert, Bleu].

La première partie d'un CNN est la partie convolutive à proprement parler. Elle fonctionne comme un extracteur de caractéristiques des images. Une image est passée à travers d'une succession de filtres, ou noyaux de convolution, créant de nouvelles images appelées cartes de convolutions (Figure 2.5). certains filtres intermédiaires réduisent la résolution de l'image par une opération de maximum local. En fin, les cartes de convolutions sont mises à plat et concaténées en un vecteur de caractéristiques, appelé code CNN.

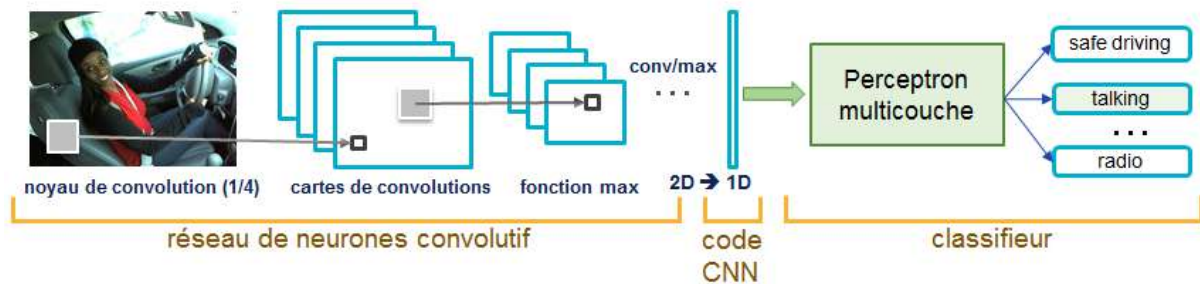


Figure 5-Les réseaux de neurones convolutifs.

Ce code CNN en sortie de la partie convolutive est ensuite branché en entrée d'une deuxième partie, constituée de couches entièrement connectées (perceptron multicouche page 10). Le rôle de cette partie est de combiner les caractéristiques du code CNN pour classer l'image.

La sortie est une dernière couche comportant un neurone par catégorie. Les valeurs numériques obtenues sont généralement normalisées entre 0 et 1, de somme 1, pour produire une distribution de probabilité sur les catégories.

II.4.2-L'entraînement d'un nouveau CNN est difficile :

Créer un nouveau réseau de neurones convolutif est coûteux en termes d'expertise, de matériel et de quantité de données annotées nécessaires.

Il s'agit d'abord de fixer l'architecture du réseau, c'est-à-dire le nombre de couches, leurs tailles et les opérations matricielles qui les connectent. L'entraînement consiste alors à optimiser les coefficients du réseau pour minimiser l'erreur de classification en sortie. Cet entraînement peut prendre plusieurs semaines pour les meilleurs CNN, avec de nombreux GPU travaillant sur des centaines de milliers d'images annotées.

Des équipes de recherche se spécialisent dans l'amélioration des CNN. Elles publient leurs innovations techniques, ainsi que le détail des réseaux entraînés sur des bases de données de références.

II.4.3-Les blocs de construction :

Une architecture CNN est formée par un empilement de couches de traitement indépendantes :

II.4.3.1-La couche de convolution (CONV) :

Trois hyper paramètres permettent de dimensionner le volume de la couche de convolution (aussi appelé volume de sortie) : la 'profondeur', le 'pas' et la 'marge'.

1. 'Profondeur' de la couche : nombre de noyaux de convolution (ou nombre de neurones associés à un même champ récepteur).
2. 'Le pas' contrôle le chevauchement des champs récepteurs. Plus le pas est petit, plus les champs récepteurs se chevauchent et plus le volume de sortie sera grand.
3. 'La marge (à 0)' ou 'zero padding' : parfois, il est commode de mettre des zéros à la frontière du volume d'entrée. La taille de ce 'zero-padding' est le troisième hyperparamètre. Cette marge permet de contrôler la dimension spatiale du volume de sortie. En particulier, il est parfois souhaitable de conserver la même surface que celle du volume d'entrée (11).

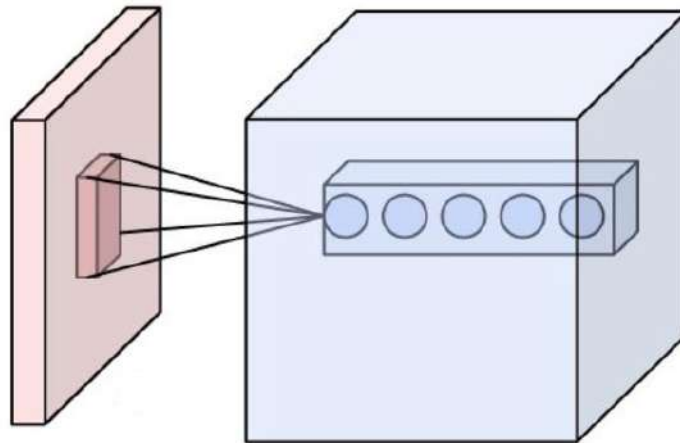


Figure 6-Ensemble de neurones (cercles) créant la profondeur d'une couche de convolution (bleu). Ils sont liés à un même champ récepteur (rouge).

II.4.3.2-Couche de pooling (POOL) :

Un autre concept important des CNNs est le pooling (« mise en commun »), ce qui est une forme de sous-échantillonnage de l'image. L'image d'entrée est découpée en une série de rectangles de n pixels de côté ne se chevauchant pas (pooling). Chaque rectangle peut être vu comme une tuile. Le signal en sortie de tuile est défini en fonction des valeurs prises par les différents pixels de la tuile.

Le pooling réduit la taille spatiale d'une image intermédiaire, réduisant ainsi la quantité de paramètres et de calcul dans le réseau. Il est donc fréquent d'insérer périodiquement une couche de pooling entre deux couches convolutives successives d'une architecture CNN pour contrôler l'overfitting (sur-apprentissage). L'opération de pooling créait aussi une forme d'invariance par translation.

La couche de pooling fonctionne indépendamment sur chaque tranche de profondeur de l'entrée et la redimensionne uniquement au niveau de la surface. La forme la plus courante est une couche de mise en commun avec des tuiles de taille 2×2 (largeur/hauteur) et comme valeur de sortie la valeur maximale en entrée (cf. schéma). On parle dans ce cas de « Max-Pool 2×2 » (compression d'un facteur 4).

Il est possible d'utiliser d'autres fonctions de pooling que le maximum. On peut utiliser un « average pooling » (la sortie est la moyenne des valeurs du patch d'entrée), du « L2-norm pooling ». Dans les faits, même si initialement l'average pooling était souvent utilisé il s'est avéré que le max-pooling était plus efficace car celui-ci augmente plus significativement

l'importance des activations fortes. En d'autres circonstances, on pourra utiliser un pooling stochastique (voir « Méthodes de régularisation » plus bas dans ce document).

Le pooling permet de gros gains en puissance de calcul. Cependant, en raison de la réduction agressive de la taille de la représentation (et donc de la perte d'information associée), la tendance actuelle est d'utiliser de petits filtres [en] (type 2x2) (figure (2.3)). Il est aussi possible d'éviter la couche de pooling [en] mais cela implique un risque sur-apprentissage plus important. (12)

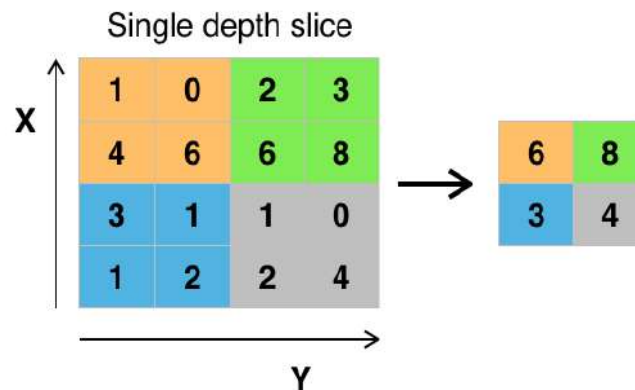


Figure 7-Pooling avec un filtre 2x2 et un pas de 2.

II.4.3.3-Couche de correction (ReLU):

Souvent, il est possible d'améliorer l'efficacité du traitement en intercalant entre les couches de traitement une couche qui va opérer une fonction mathématique (fonction d'activation) sur les signaux de sortie. On a notamment :

- La correction ReLU (abréviation de Unités Rectifié linéaires) : $f(x)=\max(0,x)$. Cette fonction, appelée aussi « fonction d'activation non saturante », augmente les propriétés non linéaires de la fonction de décision et de l'ensemble du réseau sans affecter les champs récepteurs de la couche de convolution.
- La correction par tangente hyperbolique $f(x)=\tanh(x)$.
- La correction par la tangente hyperbolique saturante : $f(x)=|\tanh(x)|$.
- La correction par la fonction sigmoïde.

Souvent, la correction Relu est préférable, car il en résulte la formation de réseau neuronal plusieurs fois plus rapide, sans faire une différence significative à la généralisation de précision.

II.4.3.3-Couche entièrement connectée(FC):

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

II.4.3.3-Couche de perte (LOSS) :

La couche de perte spécifie comment l'entraînement du réseau pénalise l'écart entre le signal prévu et réel. Elle est normalement la dernière couche dans le réseau. Diverses fonctions de perte adaptées à différentes tâches peuvent y être utilisées. La perte « Soft max » est utilisée pour prédire une seule classe parmi K classes mutuellement exclusives. La perte par entropie croisée sigmoïde est utilisée pour prédire K valeurs de probabilité indépendante dans [0,1]. La perte euclidienne est utilisée pour régresser vers des valeurs réelles.

II.4.4-Choix des hyperparamètres :

Les CNNs utilisent plus d'hyperparamètres qu'un MLP standard. Même si les règles habituelles pour les taux d'apprentissage et des constantes de régularisation s'appliquent toujours, il faut prendre en considération les notions de nombre de filtres, leur forme et la forme du max pooling.

II.4.4.1-Nombre de filtres :

Comme la taille des images intermédiaires diminue avec la profondeur du traitement, les couches proches de l'entrée ont tendance à avoir moins de filtres tandis que les couches plus proches de la sortie peuvent en avoir davantage. Pour égaliser le calcul à chaque couche, le produit du nombre de caractéristiques et le nombre de pixels traités est généralement choisi pour être à peu près constant à travers les couches. Pour préserver l'information en entrée, il faudrait maintenir le nombre de sorties intermédiaires (nombre d'images intermédiaire multiplié par le nombre de positions de pixel) pour être croissante (au sens large) d'une couche à l'autre.

Le nombre d'images intermédiaires contrôle directement la puissance du système, dépend du nombre d'exemples disponibles et la complexité du traitement.

II.4.4.2-Forme de filtres :

Les formes de filtre varient grandement dans la littérature. Ils sont généralement choisis en fonction de l'ensemble de données. Les meilleurs résultats sur les images de MNIST (28x28) sont habituellement dans la gamme de 5x5 sur la première couche, tandis que les ensembles de données d'images naturelles (souvent avec des centaines de pixels dans chaque dimension) ont tendance à utiliser de plus grands filtres de première couche de 12x12, voire 15x15.

Le défi est donc de trouver le bon niveau de granularité de manière à créer des abstractions à l'échelle appropriée et adaptée à chaque cas.

II.4.4.3-Forme de Max pooling :

Les valeurs typiques sont 2x2 (figure 8). De très grands volumes d'entrée peuvent justifier un pooling 4x4 dans les premières couches. Cependant, le choix de formes plus grandes va considérablement réduire la dimension du signal, et peut entraîner la perte de trop d'information.

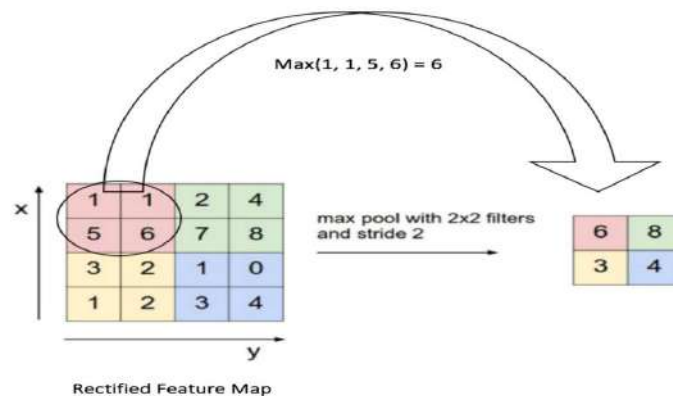


Figure 8-Exemple de max pooling (2x2).

II.5-Avantages de CNNs :

Un avantage majeur des réseaux convolutifs est l'utilisation d'un poids unique associé aux signaux entrant dans tous les neurones d'un même noyau de convolution. Cette méthode réduit l'empreinte mémoire, améliore les performances (13) et permet une invariance du traitement par translation. C'est le principal avantage du CNN par rapport au MLP, qui lui considère chaque neurone indépendant et donc affecte un poids différent à chaque signal entrant.

Lorsque le volume d'entrée varie dans le temps (vidéo ou son), il devient intéressant de rajouter un paramètre de temporisation (delay) dans le paramétrage des neurones. On parlera dans ce cas de réseau neuronal à retard temporel (TDNN).

Comparés à d'autres algorithmes de classification de l'image, les réseaux de neurones convolutifs utilisent relativement peu de pré-traitement. Cela signifie que le réseau est responsable de faire évoluer tout seul ses propres filtres (apprentissage sans supervision), ce qui n'est pas le cas d'autres algorithmes plus traditionnels. L'absence de paramétrage initial et d'intervention humaine est un atout majeur des CNN.

II.6-Conclusion :

Dans ce chapitre on a présenté les notions importantes qui sont en relation avec l'apprentissage profond (définition, Architectures....etc). Aussi qu'une vision générale sur l'apprentissage profond, toute on donnant en détail la méthode choisie dans notre travail de recherche qui est le CNNs. Le prochain chapitre, traite les détails de la conception, ainsi que la méthode et les outils utilisés pour la réalisation de notre application.



**Chapitre III:
Conception et
implémentation**

Chapitre III: Conception et implémentation:

III.1- Introduction :

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

III.2-Présentation des outils de développement :

III.2.1- Matériel :

Le matériel réalisé est PC personnel Lenovo I5 avec un 8GB capacité mémoire, et un processeur Intel® coré™ 2 Duo 2.53 GHz, avec Windows 10 édition intégral, service pack 1 64 bit type système.

III.2.2- Matlab :

MATLAB_ (Matrix LABORatory) est un logiciel interactif basé sur le calcul matriciel. Il est utilisé dans les calculs scientifiques et les problèmes d'ingénierie parce qu'il permet de résoudre des problèmes numériques complexes en moins de temps requis par les langages de programmation courant, et ce grâce à une multitude de fonctions intégrées et à plusieurs programmes outils testés et regroupés selon usage dans des dossiers appelés boîtes à outils ou "toolbox". (14)

Son objectif, par rapport aux autres langages, est de simplifier au maximum la transcription en langage informatique d'un problème mathématique, en utilisant une écriture la plus proche possible du langage naturel scientifique.

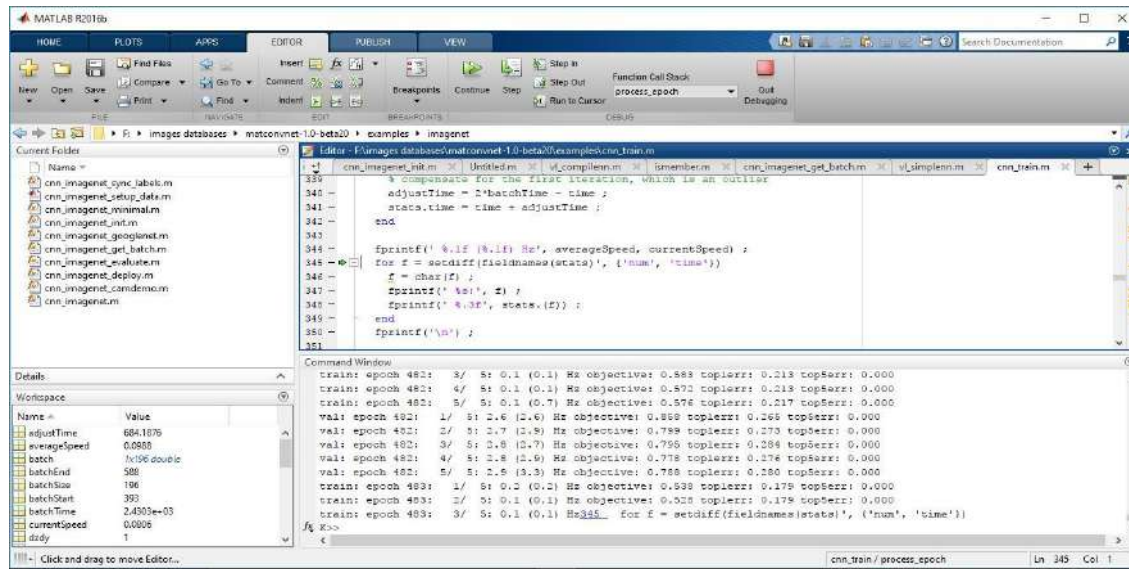


Figure 9-L` environnement.Matlab.

III.2.3- Visual Studio.NET :

Visual Studio.NET est un ensemble complet d'outils de développement pour créer des applications ASP Web, des services Web XML, des applications bureautiques et les applications mobiles. Visual Basic .NET, Visual C ++ .NET, Visual C # .NET et Visual J # .NET utilisent tous le même environnement de développement intégré (IDE), qui leur permet de partager des outils et facilite la création de solutions de langue mixte. En outre, ces langues tirées parti de la fonctionnalité du .NET Framework, qui permet d'accéder à des technologies clés qui simplifient le développement d'applications Web ASP et des services Web XML. (15)

III.2.4- MatConvNet :

MatConvNet est une boîte à **Toolbox** MATLAB implémentant des réseaux neuronaux convolutionnels (CNN) Pour les applications de vision par ordinateur. C'est simple, efficace et peut fonctionner et apprendre des CNN à la fine pointe de la technologie. CNN préformés pour la classification d'image, la segmentation, la reconnaissance de visage et la détection de texte sont disponibles.

III.3- Présentation de l'application :

Nous allons maintenant présenter les résultats obtenus grâce à des expériences réalisées par l'application de notre approche de classification des images en plusieurs classes comme exemple (class animaux, class fruit, class informatique, transport,...etc)

On utilise deep learning et la méthode de classification CNN.

Le schéma de notre système est le suivant :

1. Création des classes.
2. La classification et l'apprentissage.
3. Réalisation du moteur de recherche d'images.

L'objectif est la classification des images et réduire les taux d'erreurs de la recherche.

III.3.1- Création des classes :

Nous allons rechercher les images du même classe par exemple (classe animaux, classe fruit, classe informatique, transport,...etc),

Est-on classe les fichiers (figure.1) pour la préparation de l'apprentissage.

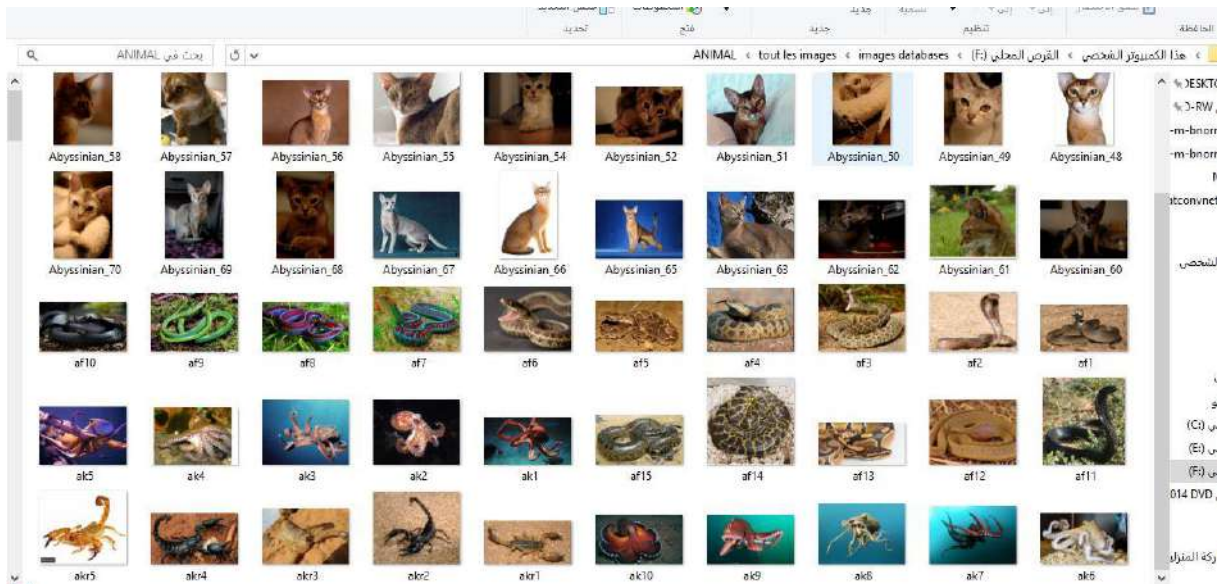


Figure 10-Création de classes.

III.3.2- l'apprentissage :

Nous préparons les classes (classe animaux, classe fruit, classe informatique, transport) que nous choisissons est exécuter le code de l'apprentissage en Matlab.

Le lancement de l'apprentissage crée un fichier pour Redimensionner toutes les images à la même taille. (figure.10)



Figure 11-Redimensionner les images.

L'apprentissage des systèmes crée un fichier net-epoch-01.mat à chaque fois jusqu'à obtenir le nombre demandé, nous avons terminé 500 époques chaque époque a durée (voir tableau) approximativement 2h, et le fichier net-deployed.mat a été créé pour la classification.

Nombre d'époques	20	80	120	200	220	300	500
erreur %	38	36	33	32	31	28	23

Le pourcentage d'erreur comme suit :

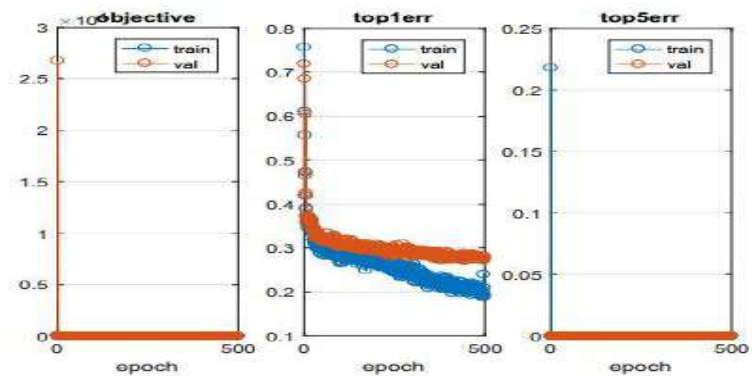


Figure 12-Taux d'erreur durant l'apprentissage de 500 époques.

III.3.3- Application :

La fenêtre principale de notre application est :

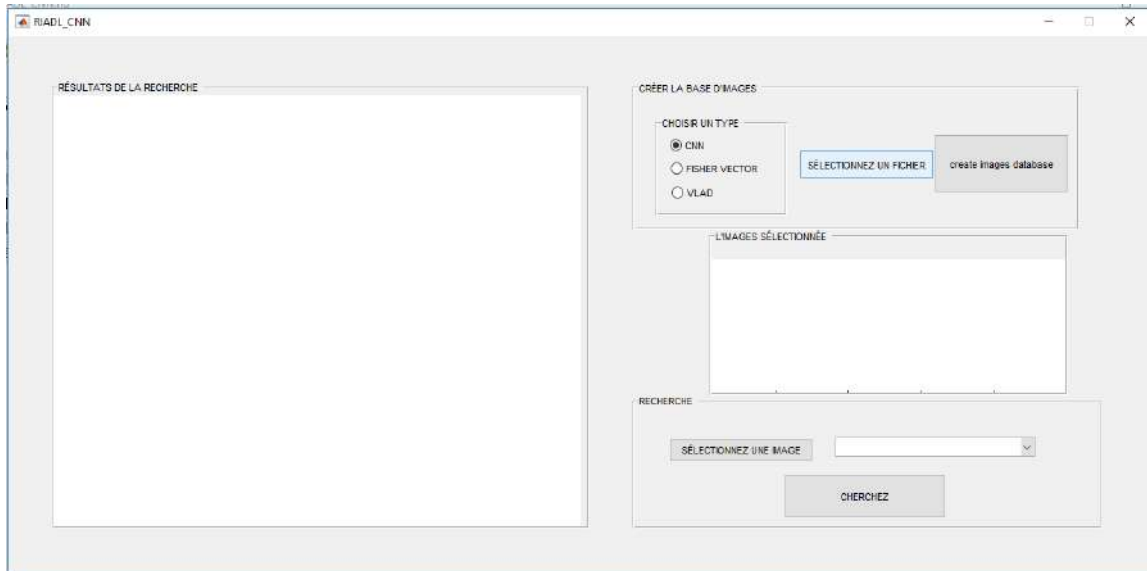


Figure 13-Interface d'application.

Cliquer sur le bouton "sélectionnez un fichier " dans la fenêtre (interface) pour ouvrir une boîte de dialogue pour choisir un fichier d'images à partir d'un emplacement local.

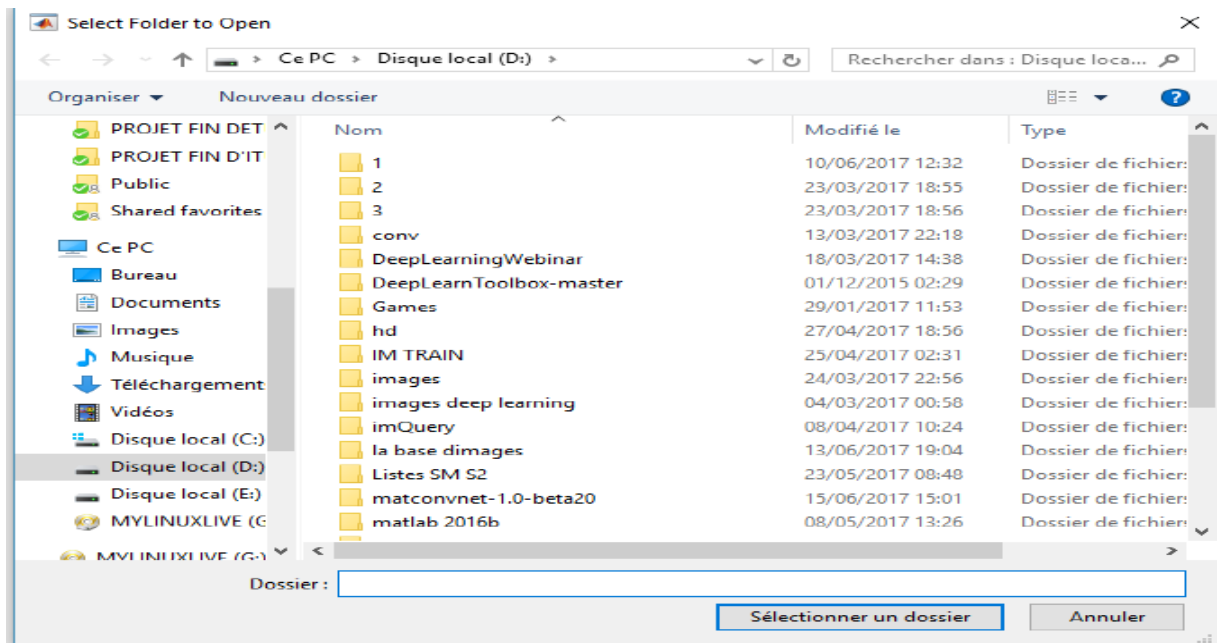


Figure 14- Sélectionnez un fichier.

cliquer sur le bouton "sélectionnez une image ou choisir une image comme suit :

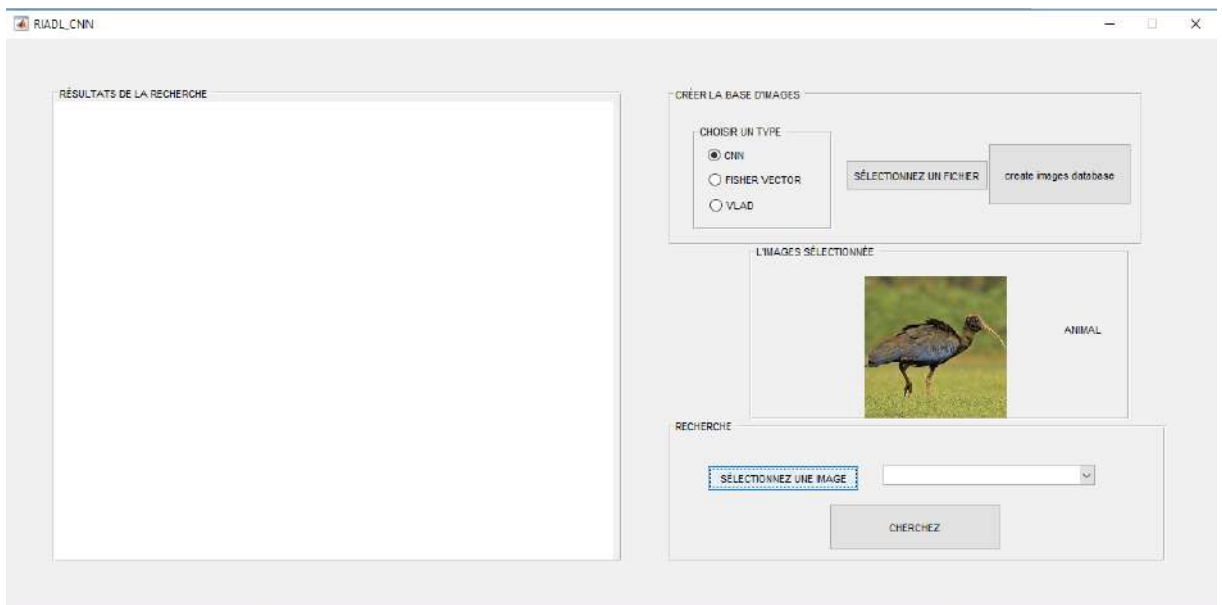


Figure 15-Affichage de l'image sélectionnée.

III.3.4- La Classification et recherche :

Utiliser le bouton "Recherchez" pour la recherche normal sur les images de la même classe donne le résultat comme suit : (figure 16)

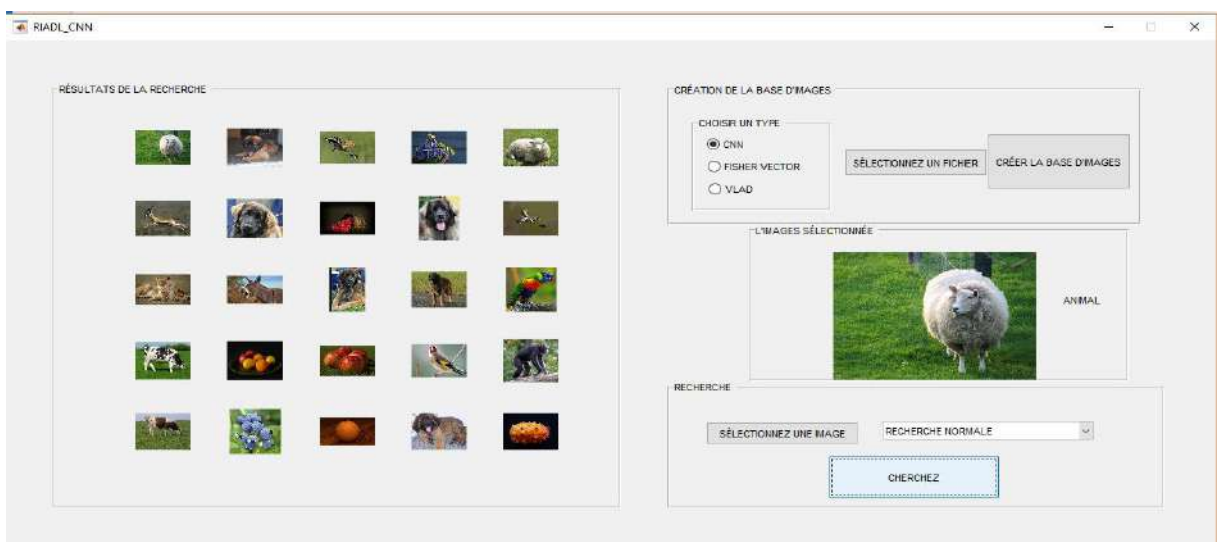


Figure 16-Rechercher les images

Si on choisit une recherche par classification CNN, elle élimine les images qui ne font pas partie de la class (figure 17).

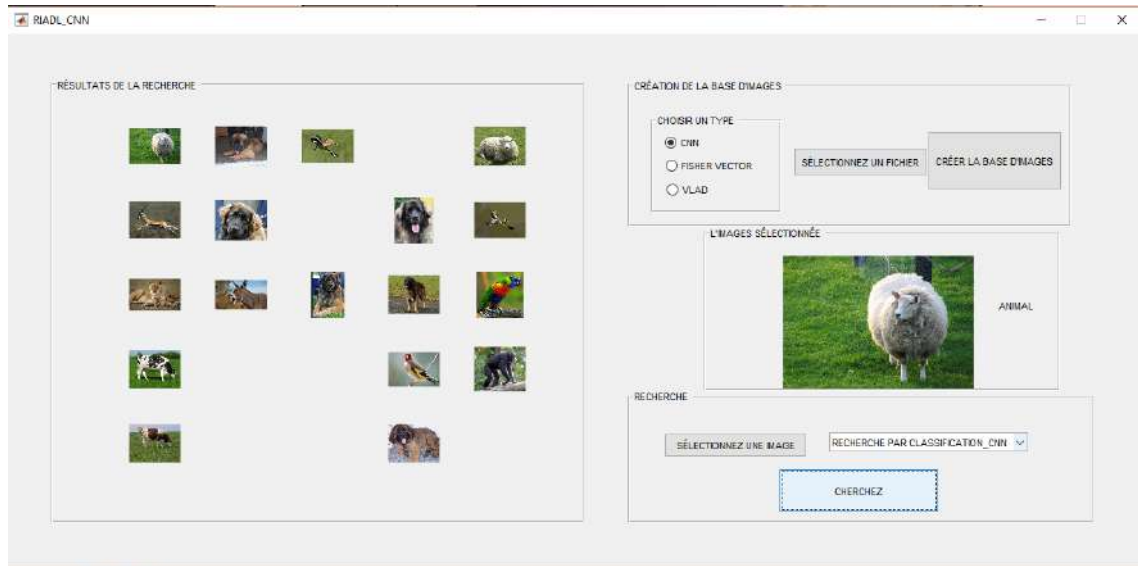


Figure 17-Recherche par la classification.

III.3.5- Présentation des méthodes comparaison :

Nous avons ajouté deux méthodes de recherche et de classification d'images pour tester notre système.

1. **Le Fisher Vector (FV)** : est une représentation d'image obtenue par la mise en commun des caractéristiques locales de l'image. Il est souvent utilisé comme descripteur global de l'image dans la classification visuelle. Alors que le FV peut être dérivé comme un cas particulier, approximative, et l'amélioration du cadre général Fisher noyau, il est facile de décrire directement. (16)
2. **VLAD (Vector of locally aggregated descriptors)** : est une méthode de codage de fonction et mise en commun, semblable à des vecteurs Fisher. VLAD code pour un ensemble de descripteurs de caractéristiques locales $I = (x_1, \dots, x_n)$ extraite d'une image en utilisant un dictionnaire intégré en utilisant une méthode de classification tel que modèles de mélange gaussien (GMM) ou K-means. (17)

III.3.6- Tests et résultats :

Pour tester notre application et juger sa performance nous avons ajouté deux méthodes de recherche et de classification d'images et testé sur trois bases d'images .La relation que nous avons utilisée pour calculer le pourcentage de notre classification est : $P=R/T$.

P : Le pourcentage de classification.

R : Nombre d'image reconnue.

T : Nombre total d'image de test.

Nous avons obtenu les résultats suivants :

Les bases d'images	Le type de la classification	Nombre d'images reconnues	Nombre d'image inconnue	Nombre totale d'images de test	Taux de reconnaissance
Bases d'images n°1	CNN	173	27	200	86.5%
	FISHER VECTOR	141	59	200	70.5%
	VLAD	123	77	200	61.5%
Bases d'images n°2	CNN	231	49	280	82.5%
	FISHER VECTOR	228	52	280	81.42%
	VLAD	199	81	280	71.07%
Bases d'images n°3	CNN	252	68	320	78.75%
	FISHER VECTOR	241	79	320	75.31%
	VLAD	195	125	320	60.93%

En comparant les résultats, nous concluons que la méthode CNN donne de meilleurs résultats par rapport à l'autre méthode (fisher_vector et Vlad).

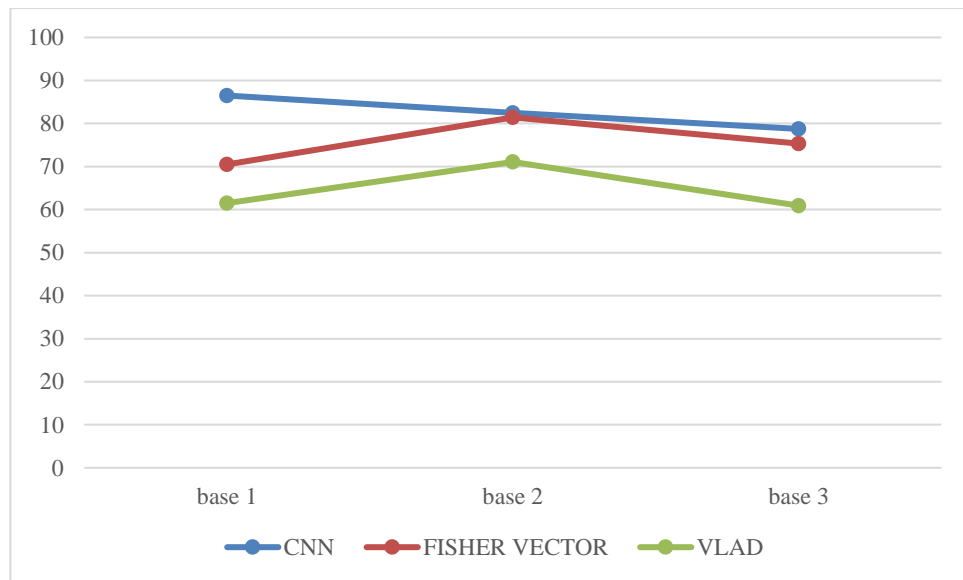


Figure 18-Comparaison des résultats obtenus par chaque méthode.

III.3.7- Discussion des résultats :

Les résultats de notre application sont présentés sur la courbe précédente (figure 18) pour les trois méthodes de classification.

On remarque que malgré l'utilisation de la méthode CNN qui présente le problème de l'initialisation aléatoire du centre au début de la classification, nous avons obtenu un résultat de 82.52% comme meilleur résultat d'exécution, par rapport aux méthodes (fisher_vector et Vlad).

III.4-Conclusion :

Après avoir achevé notre conception nous avons donné les outils nécessaires pour la réalisation de notre travail. Nous avons présenté aussi l'environnement de développement. A la fin nous avons présenté notre application en donnant quelques captures d'écran qui expliquent le déroulement et le fonctionnement de notre travail, ainsi que les résultats obtenus.

On a donné un exemple de classification par notre application d'un ensemble d'images qui sont déjà classifiées au paravent. Nous avons montré l'impact de la méthode choisie (CNNs) sur les résultats de la classification et la recherche d'images.

Conclusion générale :

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

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

C'est dans ce cadre que s'inscrit notre travail, qui a pour objectif de proposer une application qui réalise une classification d'une base d'images en un ensemble de classes (informatique – transport-animaux-fruit).

Pour réaliser notre travail de classification on a utilisé le deep learning, la méthode d'apprentissage qui a montré ses performances ces dernières années et nous avons choisi la méthode CNNs comme méthode de classification, ce choix est justifié par la simplicité et l'efficacité de la méthode.

Le résultat obtenu lors de la phase de test confirme l'efficacité de notre approche.

Notre travail n'est que dans sa version initiale, on peut dire que ce travail reste ouvert pour des travaux de comparaison et/ou d'hybridation avec d'autres méthodes de classification.

Bibliographie

1. **BOUCHER, ALAIN.** *OUTIL D'AIDE A L'ANNOTATION.* 22 janvier 2007.
2. **BORGI A., AKDAG H.** *Supervised Learning and Approximate.* 2001.
3. **R. S. Michalski, R.L. Chilausky.** *Knowledge acquisition by encoding expert rules versus computer induction from examples : a case study involving.* 1981.
4. **Parrochia.** *Classifications, histoire et problèmes formels. Cinquièmes Rencontres de la société Francophone de classification SFC'97.* Lyon : s.n., Septembre 1997.
5. **R.E. Stepp, R.S. Michalski.** *Learning from Observation: Conceptual Clustering, Machine Learning, An Artificial Intelligence Approach.* s.l. : Morgan Kaufman Publishers, 1983.
6. **Nasrin, Abdllaoui Nezha.** *la classification on hiérarchique axendante.* 2013/2014.
7. **complément-de-cours-sur-la-classification.**
8. **Hirotk Suzuki, Pascal Matsakis.** *Exploitation de connaissances structurelles en classification on d'image:Utilisation de méthodes heuristiques d'optimisation combinatoire.*
9. **McCulloch, W. S., Pitts, W.** *A Logical Calculus of the Ideas Immanent in Nervous Activity Bulletin of Mathematical Biophysics.* 1943.
10. **Parizeau, Marc.** *Réseaux de Neurons (Le perceptron multicouche et son algorithme de retropropagation des erreurs).* s.l. : Université Laval, 2004.
11. **A. Krizhevsky, I. Sutskever et G. E. Hinton.** *ImageNet Classification with Deep Convolutional Neural Networks », Advances in neural Processing Systems de traitement.* 2012.
12. **Graham, Benjamin.** *Fractional Max-Pooling.* 18 décembre 2014.
13. **LeCun, Yann.** *L'apprentissage prédictif est le grand défi scientifique de l'intelligence artificielle .*
14. **A.Seghir.** *méthode des éléments finis.* s.l. : Département de génie civil, mémoire magister, Université A.Mira de Bejaia.
15. **présentation, Visual Studio.Net.**
16. **Dance, F. Perronnin and C. Fisher** *kenrels on visual vocabularies for image.* In Proc. CVPR : s.n., 2006.
17. **Lowe., M. Muja and D. G.** *Fast approximate nearest neighbors with automatic algorithmic.* In Proc. CVPR : s.n., 2009.
18. **Tahe-ISTIA, Serge.** *Apprentissage du langage C# Université d'angers.* Mai 2002.
19. **A. Krizhevsky, I. Sutskever, and G. E. Hinton.** *Imagenet classification with deep convolutional neural networks.* In Proc. NIPS : s.n., 2012.

20. **Fulkerson, A. Vedaldi and B.** *VLFeat { An open and portable library of computer vision algorithms.* In Proc. ACM Int. Conf. on Multimedia, : s.n., 2010.

21. **Lowe., M. Muja and D. G.** *Fast approximate nearest neighbors with automatic algorithmic.* In Proc. VISAPP : s.n., 2009.

Résumé : Dans notre mémoire on a utilisé l'apprentissage profond (Deep Learning) et les réseaux de neurones convolutifs (CNNs) pour classier une base d'images en un ensemble de classes, on a utilisé Matlab et la bibliothèque Matconv Vfeat....etc . On a implémenté un moteur de recherche d'images basé sur cette classification pour montrer l'efficacité de la méthode et on a comparé notre approche avec VLAD et Fisher.

Mots clés : recherche d'images, classification, CNN : Convolutional Neural Networks, Deep Learning : apprentissage profond,

ملخص : في مذكرتنا قمنا بإستخدام التعلم العميق والشبكات التلافيفية الإصطناعية (CNNs) لتصنيف قواعد الصور إلى مجموعة من الأصناف، بحيث استعملنا Matlab وكل من مكتبات Matconvnet, Vfeat ... إلخ . وأنجزنا محرك بحث عن الصور يرتكز على هذا التصنيف بغرض إظهار فعالية هذا الأسلوب ومقارنته مع نوعين آخرين من تقنيات التصنيف (VLAD et Fisher Vector).

كلمات البحث : البحث عن الصور، التصنيف الصور، CNN : الشبكات العصبية التلافيفية، التعلم العميق.

Abstract: In our paper we have used deep learning (Deep Learning) and convolutional neural networks (CNNs) to classify a database of images into a set of classes, we used Matlab and the Matconv library and Vfat ... etc. We have implemented an image search engine based on this classification to show the efficiency of the method and we compared our approach with VLAD and Fisher..

Keyword: Image search, classification, CNN: Convolutional Neural Networks, Deep Learning