



**MINISTRE DE L'ENSEIGNEMENT SUPERIEUR  
ET DE LA RECHERCHE SCIENTIFIQUE  
UNIVERSITE ABBES LAGHROUR DE KHENCHELA  
FACULTE DES SCIENCES ET DE LA TECHNOLOGIE**



**Département mathématique et informatique**

N° de série : .....

## **Mémoire de fin d'études**

*Pour l'obtention du diplôme de Master (L.M.D)*

**Spécialité : Sécurité et technologie web**

***Thème : Les réseaux convolutifs  
pour la détection et la reconnaissance  
d'objets dans une image***

*Présenté et soutenu publiquement par :  
BOULEKOUIRAT Sabrina et KENZARI Zouleykha*

*Encadreur : BEKHOUCHE Abdelali.  
Co-encadreur : ABBAS Fayçal.*

*Année universitaire : 2017/2018.*

# *Dédicace*

*Je dédie ce mémoire :*

*A mes très chers parents Pour leurs soutient, leurs patience, leurs amour et encouragement durant toutes mes années d'études et sans lesquels je n'aurais jamais réussi, Que dieu leur procure bonne santé et longue vie.*

*A celui que j'aime beaucoup et qui m'a soutenue tout au long de ce projet mon fiancé Ayoub, qui nous à aidé à notre travail,*

*Ma sœur, leurs enfants et mes beaux frères,*

*Mon binôme et chère amies Fatima et kamília,*

*A toute la famille Boulekouirat et Boucha.*

*Et à tous ceux que j'aime et à toutes les personnes qui m'ont prodigué des Encouragements et se sont donné la peine de me soutenir durant cette formation.*

*Boulekouirat Sabrina*

# *Dédicace*

*A l'homme de ma vie, mon exemple éternel, mon soutien moral et source de joie et de bonheur, celui qui s'est toujours sacrifié pour me voir réussir, que dieu te garde dans son vaste paradis, à toi mon père.*

*A la lumière de mes jours, la source de mes efforts, la flamme de mon cœur, ma vie et mon bonheur ; maman que j'adore.*

*Aux personnes dont j'ai bien aimé la présence dans ce jour, à tous mes frères et mes sœurs, particulièrement ma tante A et ma chérie Sabrina, je dédie ce travail dont le grand plaisir leurs revient en premier lieu pour leurs conseils, aides, et encouragements.*

*Aux personnes qui m'ont toujours aidé et encouragé, qui étaient toujours à mes côtés, et qui m'ont accompagnaient durant mon chemin d'études supérieures, mes aimables amis, collègues d'étude.*

*Kenzari Zouleykha*

# *Remerciements*

Nous remercions tout d'abord notre dieu (Allah) pour la patience qui nous offre pour finaliser ce travail dans des bonnes conditions.

Nous sommes très reconnaissant à nos encadrant :

Monsieur Bekhouche Abdelali et Monsieur Abbas Fayçal pour l'aide compétente qu'ils nous ont apportée, pour leur patience et leur encouragement. Leurs critiques nous ont été très précieuses pour structurer ce travail et pour améliorer la qualité des différentes sections.

Nous remercions chaleureusement Graït Ayoub qui nous a aidé durant notre travail et par sa patience et ses Précieux conseils.

Nous tenons à remercier très sincèrement l'ensemble des membres du jury Hioual Ouassila et M.H Souïdi qui nous font le grand honneur d'accepter de juger notre travail.

Nous voulons remercier aussi tous nos professeurs du master informatique, Sécurité et technologie web et tous ceux qui nous ont offert la possibilité de mener à bien ce travail.

**TABLE DES MATIERES**

**Résumé**

**Introduction générale** 01

**Chapitre I : Traitement d'images** 02-10

1	Introduction.....	02
2	Définition de l'image.....	02
2.1	Définition de l'image numérique.....	02
2.2	Représentation des images numériques.....	03
2.3	Types de l'image.....	03
2.3.1	Les images de type bitmap.....	03
2.3.2	Les images de type vectoriel.....	04
2.3.3	Les types de format d'image.....	05
2.4	Caractéristiques images numérique.....	05
3.3.1	Pixel.....	05
3.3.2	Dimension.....	06
3.3.3	Résolution.....	06
3.3.4	Contours.....	06
3.3.5	Bruit.....	07
3.3.6	Luminance.....	07
3.3.7	Histogramme.....	07
3.3.8	Contraste.....	08
3.3.9	Images à niveaux de gris.....	08
3	Traitement d'images.....	09
3.1	Définition.....	09
3.2	Prétraitement d'image.....	09
3.3	Conversion des images couleurs en niveau de gris.....	09
3.4	Binarisation.....	10
3.5	Normalisation.....	10

**TABLE DES MATIERES**

---

4	Conclusion.....	10
	<b>Chapitre I : L'apprentissage automatique</b>	<b>11-37</b>
1	Introduction.....	12
2	Intelligence Artificielle.....	12
2.1	Une courte histoire de l'intelligence artificielle.....	13
2.2	Les sous-domaines de l'intelligence artificielle.....	13
2.3	Les grands acteurs de l'intelligence artificielle.....	14
3	L'apprentissage Automatique.....	15
3.1	Définition.....	16
3.2	Domaines d'applications de l'apprentissage automatique.....	16
3.3	Types d'apprentissage.....	16
3.3.1	Apprentissage par renforcement.....	17
3.3.2	L'apprentissage non supervisé.....	17
3.3.3	Apprentissage semi-supervisé.....	18
3.3.4	Apprentissage supervisé.....	18
3.4	Les approches d'apprentissage.....	19
3.4.1	Les Réseaux Neuronaux artificiels.....	19
3.4.1.1	Le neurone biologique.....	19
3.4.1.2	Le neurone formel.....	20
3.4.1.3	Les différents types de réseaux de neurones.....	21
3.4.1.4	L'apprentissage des réseaux de neurones.....	24
3.4.1.5	Les algorithmes.....	24
4	L'apprentissage profond .....	26
4.1	La différence entre l'apprentissage automatique et l'apprentissage en profondeur.....	27
4.2	Les algorithmes d'apprentissage profond.....	28
4.2.1	Réseau de neurones convolutifs (CNN).....	29
4.2.1.1	Couche Convolution.....	31
4.2.1.2	Couche de pooling.....	33

## *TABLE DES MATIERES*

---

4.2.1.3	Les fonctions d'activation.....	34
4.2.1.4	Couche entièrement connectée.....	36
4.2.1.5	Couche de perte (LOSS).....	36
4.2.1.6	Méthode de régularisation.....	36
5	Conclusion.....	37
	<b>Chapitre III : Reconnaissance, classification d'images</b>	<b>38-50</b>
1	Introduction.....	39
2	La classification.....	39
2.1	Définition de classification.....	39
2.2	Les motivations de la classification d'images.....	39
2.3	Méthodes de classification.....	40
2.4	L'objectif de la classification.....	41
2.5	Domaines d'application de la classification.....	41
2.6	Indicateurs de performance en classification.....	42
3	La reconnaissance.....	43
3.1	La reconnaissance de formes.....	43
3.2	Les principaux domaines d'application.....	43
3.3	La reconnaissance spécifique versus générique.....	44
3.4	La difficulté majeure de la reconnaissance d'objets.....	44
3.5	Les méthodes de reconnaissance d'objets.....	45
3.6	Les méthodes de reconnaissance dans le traitement d'images...	45
3.7	Les traitements de bas niveau: extraction des descripteurs.....	46
3.8	Les traitements d'haut niveau: reconnaissance d'objets.....	46
4	La reconnaissance, classification des panneaux de signalisation.....	46
5	Conclusion.....	50
	<b>Chapitre IV: Implémentation</b>	<b>51-60</b>
1	Introduction.....	52
2	La configuration du matériel utilise.....	52
3	L'environnement de travail.....	52

## TABLE DES MATIERES

---

3.1 Python.....	52
3.2 notebook Jupyter.....	52
3.3 Scikit-Learn.....	53
3.4 Tensorflow.....	53
4 Notre modèle.....	54
4.1 Prétraitements des données.....	54
4.1.1 Image en niveaux de gris.....	55
4.1.2 Normalisation.....	55
4.1.3 Augmentation.....	56
4.2 Architecture.....	57
<b>Chapitre V: Résultat</b>	<b>61-66</b>
1 Introduction.....	62
2 Le jeu de données.....	62
3 Evaluation.....	63
4 Conclusion.....	66
<b>Conclusion et perspectives</b>	<b>67</b>
<b>Bibliographie</b>	

**LISTE DES FIGURES**

Fig. I.1 Exemple d'une image numérique.....	03
Fig. I.2 Image bitmap.....	04
Fig. I.3 Image vectoriel.....	04
Fig. I.4 Groupe de pixel formant la lettre A.....	06
Fig. I.5 Les contours d'image.....	07
Fig. I.6 L'histogramme.....	08
Fig. I.7 Niveaux de gris.....	08
Fig. II.1 L'apprentissage non-supervisé.....	17
Fig. II.2 L'Apprentissage semi-supervisé.....	18
Fig. II.3 L'apprentissage supervisé .....	19
Fig. II.4 Neurones biologiques .....	20
Fig. II.5 Représentations d'un neurone formel .....	20
Fig. II.6 Les fonctions de transition (a) linéaire, (b) sigmoïde, (c) seuil.....	21
Fig. II.7 Classification des réseaux de neurones.....	22
Fig. II.8 Exemple de perceptron multicouche avec une couche cachée et une couche de sortie.....	23
Fig. II.9 Un réseau de hopfield.....	23
Fig. II.10 D'une carte de Kohonen.....	24
Fig. II.11 Schéma du modèle de la rétropropagation de l'erreur.....	25
Fig. II.12 Représentation schématique du l'intelligence artificielle.....	27
Fig. II.13 Comparaison de méthodes de catégorisation de plaque d'apprentissage automatique et d'apprentissage profond.....	28
Fig. II.14 Spécialisation des filtres d'un CNN .....	30
Fig. II.15 Réseau de neurones LeNet 5.....	31
Fig. II.16 Principe d'une convolution.....	32
Fig. II.17 La couche de pooling.....	33
Fig. II.18 La tangente hyperbolique.....	34
Fig. II.19 La sigmoïde.....	35

## *LISTE DES FIGURES*

---

Fig. II.20 La fonction ReLU.....	35
Fig. II.21 Dropout (a) : avant l'application dropout (b) : après l'application dropout.....	37
Fig. III.1 Courbe ROC.....	42
Fig. III.2 La classification MPPS (max pooling position).....	47
Fig. III.3 Résultat de la méthode proposée pour l'ensemble de test Allemagne.....	48
Fig. III.4 Résultat de la méthode proposée pour l'ensemble de test Corée.....	48
Fig. III.5 Le temps moyen de reconnaissance de HP-MLP, HP-SVM, HOG-MLP et HOG-SVM.....	49
Fig. III.6 Reconnaissance.....	49
Fig. IV.1 L'architecture globale de notre modèle.....	54
Fig. IV.2 L'étape de prétraitements des données.....	55
Fig. IV.3 Conversion d'une image RVB en niveaux de gris. ....	55
Fig. IV.4 Normalisation des images.....	56
Fig. IV.5 Augmentation des données.....	56
Fig. IV.6 (a) jeu d'entraînement avant l'augmentation, (b) jeu d'entraînement après l'augmentation.....	57
Fig. IV.7 Les étapes de l'architecture CNN.....	57
Fig. IV.8 Calcul de la taille de volume de sortie dans une couche de convolution.....	58
Fig. IV.9 Max pooling.....	59
Fig. IV.10 Architecture CNN.....	60
Fig. V.1 Ensemble des images du jeu de données.....	62
Fig. V.2 La précision de l'entraînement et de la validation.....	63
Fig. V.3 L'erreur de l'entraînement et de la validation.....	63
Fig. V.4 Résultats de tests sur des nouvelles images.....	64
Fig. V.5 Les résultats de prédictions.....	65

**LISTE DES ÉQUATIONS**

Éq. I.1	Formule de conversion des images couleurs en niveau de gris.....	10
Éq. II.1	Formule de transition d'état.....	20
Éq. II.2	Formule de fonction sigmoïde.....	26
Éq. II.3	Formule de changement de poids.....	26
Éq. II.4	Formule de mesure d'erreur.....	26
Éq. II.5	Formule d'erreur aux unités cachées.....	26
Éq. II.6	Formule de couche convolution (largeur).....	31
Éq. II.7	Formule de couche convolution (longueur).....	31
Éq. II.8	Formule de taille de pooling (largeur).....	33
Éq. II.9	Formule de taille de pooling (longueur).....	33
Éq. II.10	Formule de tangente hyperbolique.....	34
Éq. II.11	Formule de sigmoïde.....	34
Éq. II.12	Formule de fonction ReLU.....	35
Éq. IV.1	Formule de normalisation de données.....	55

## ***Résumé***

Dans le domaine de vision par ordinateur le problème de la reconnaissance et de la classification est imposé, désormais les réseaux de neurones convolutifs (CNN) sont un nouveau concept pour l'apprentissage machine en particulier la reconnaissance et la classification d'images et vidéos, traitement du langage naturel et d'autres applications. Dans ce mémoire, on va travailler sur la reconnaissance et la classification de panneaux de signalisation, Notre modèle opère en deux étapes, une étape de traitement des données afin de simplifier l'étape d'extraction de caractéristiques, la deuxième étape consiste à appliquer un réseau de neurone convolutif afin de reconnaître et de classifier les panneaux de signalisation, notre méthode a obtenu un faible taux d'erreur et atteint une précision de 99,6%.

## ***Abstract***

In the field of computer vision the problem of recognition and classification is imposed, the Convolutional Neural Networks (CNN) are a new concept for the deep learning way especially recognition and classification of images and videos, natural language processing and other applications. In this memory we focus on the recognition and classification of traffic sign, our model operates in two stages, a data processing step to simplify the feature extraction step, the second step is to apply a Convolutional Neural Network to recognize and classify the traffic signs, our method has achieved a low error rate and reached an accuracy of 99.6%.

## ملخص

في مجال رؤية الكمبيوتر ، يتم فرض مشكلة التعرف والتصنيف ، فإن الشبكات العصبية التلافيفية هي مفهوم جديد للتعلم الآلي وخاصة لتعرف وتصنيف الصور ومقاطع الفيديو ، ومعالجة اللغة الطبيعية وغيرها من التطبيقات. في هذه الأطروحة ، سوف نعمل على التعرف على علامات الطريق وتصنيفها ، يعمل نموذجنا على مرحلتين ، مرحلة لمعالجة البيانات لتبسيط خطوة استخلاص الميزة ، والمرحلة الثانية هي تطبيق الشبكات العصبية التلافيفية للتعرف وتصنيف علامات الطريق ، وقد حققت طريقتنا معدل أخطاء منخفض وبلغت دقة 99.6%.

## *Introduction générale*

La reconnaissance des panneaux de signalisation est une opération très importante pour l'assistance et la sécurité des conducteurs, en particulier pour les voitures autonomes, le rôle d'un humain serait réduit à fournir simplement les points limites de voyage et quelques exigences supplémentaires, Il est très difficile de reconnaître les panneaux de signalisation en raison d'un manque de visibilité tels que l'illumination, le chevauchement d'objets ainsi que les conditions météorologiques difficile.

Notre démarche se focalise sur l'utilisation de technique de réseaux de neurones convolutifs noté CNNs qui sont à ce jour les modèles les plus performants pour la reconnaissance et la classification des objets dans le domaine de vision par ordinateur.

Nous proposons dans ce travail une nouvelle approche basée sur la reconnaissance et la classification des panneaux de signalisation routière, elle va nous permettre d'identifier les panneaux de signalisation routière dans différentes conditions météorologiques. Notre modèle a prouvé son efficacité car il atteint une précision supérieure avec un faible taux d'erreur.

Ce mémoire est organisé en Cinq chapitres.

- Dans le premier chapitre nous traiterons les caractéristiques, les différents types et les méthodes de traitement d'images, pour améliorer les performances de notre modèle.
- Dans le deuxième chapitre, nous présenterons le principe de l'apprentissage en intelligence artificielle, ainsi que les caractéristiques d'apprentissage, aussi les différentes types d'apprentissage automatique, nous présenterons quelques approches d'apprentissages, qui sont les plus connues dans le domaine de l'intelligence artificielle et les réseaux de neurones convolutifs qui seront utilisés tout au long de notre travail.
- Dans le troisième chapitre nous présenterons les notions de base de la reconnaissance et la classification des images et finalement nous parlerons de la reconnaissance et la classification des panneaux de signalisation.
- Dans le quatrième chapitre, nous allons montrer la partie implémentation de notre travail.
- le Cinquième chapitre, on va discuter les différents résultats obtenus.

Enfin, nous terminerons ce travail par une conclusion générale et nous proposerons quelques perspectives pour des futurs travaux.

## ***1 Introduction***

Le traitement d'images est un domaine très vaste qui a connu, et qui connaît encore, un développement important depuis quelques dizaines d'années. Le traitement d'images c'est une discipline qui étudie les images numériques et leurs transformations, dont l'objectif est l'amélioration du rendu des images en vue d'en extraire de l'information.

Dans ce chapitre, nous abordons les notions de base nécessaires à la compréhension des techniques de traitement d'images. Ensuite, nous allons donner un aperçu sur les différentes techniques.

## ***2 Définition de l'image***

L'image est définie comme étant la reproduction exacte, ou la représentation analogique d'une scène réelle, ou bien, une représentation d'une scène tridimensionnelle sur un support en deux dimensions [36].

Visuellement, c'est une représentation bidimensionnelle. Elle contient en chaque point l'intensité lumineuse, l'image peut être décrite sous forme d'une fonction  $f(x, y)$ , tel que  $x$  et  $y$  sont les coordonnées linéaire d'un point de l'image et  $f$  une fonction d'intensité lumineuse définie sur un domaine borné. Sous cet aspect, l'image est inexploitable par ordinateur, ce qui nécessite sa numérisation.

## ***3 Définition de l'image numérique***

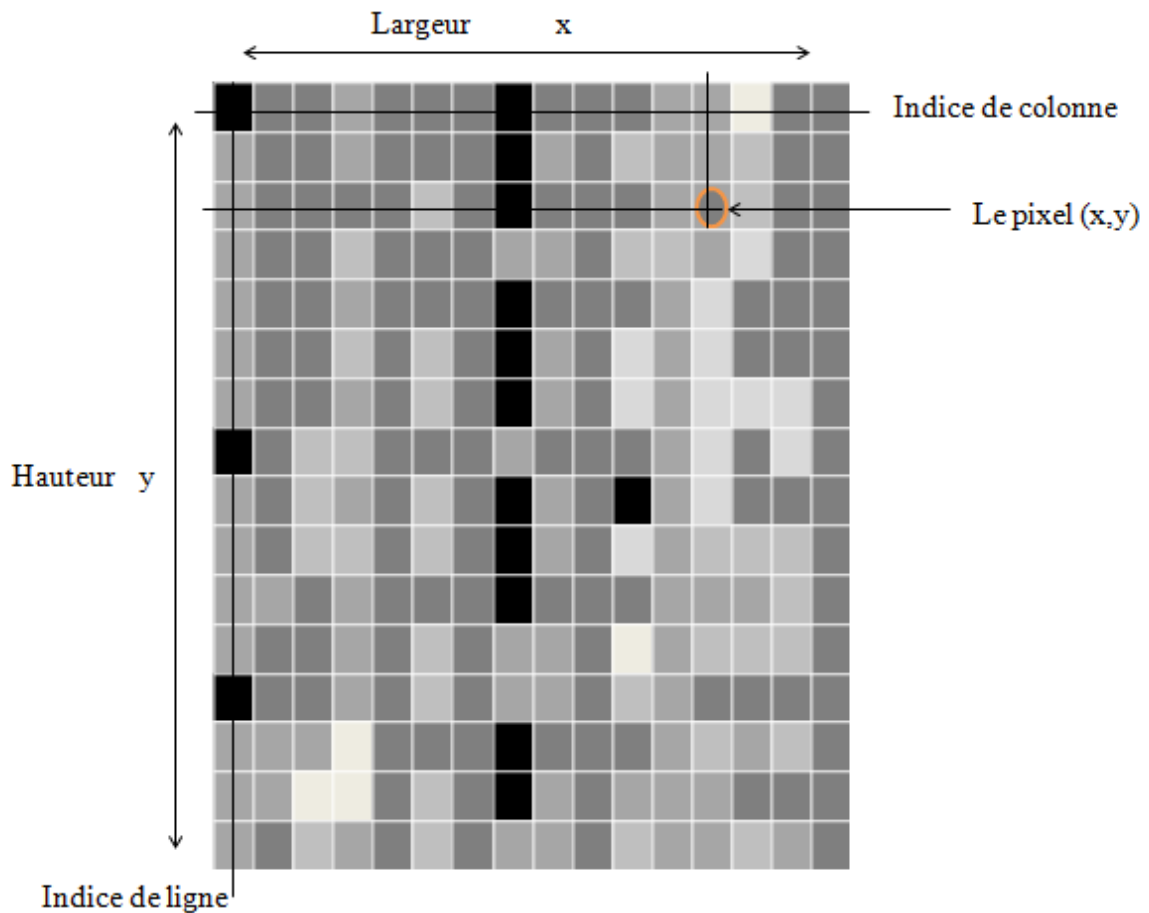
L'image numérique est l'image dont la surface est divisée en éléments de taille fixe appelés pixels, ayant chacun comme caractéristique un niveau de gris ou de couleurs [37].

La numérisation d'une image est la conversion de celle-ci de son état analogique en une image numérique représentée par une matrice bidimensionnelle de valeurs numériques  $f(x, y)$ , comme la montre la Fig. I.1 :

$x, y$ : coordonnées cartésiennes d'un point de l'image.

$f(x, y)$ : niveau d'intensité.

La valeur en chaque point exprime la mesure d'intensité lumineuse perçue par le capteur.



**Fig. I.1** Exemple d'une image numérique.

### 3.1 Représentation des images numériques

Une image numérique est une matrice de pixels repérés par leur coordonnées  $(x, y)$ .

S'il s'agit d'une image couleur, un pixel est codé par 3 composantes  $(r, v, b)$  chacune comprise au sens large entre  $(0 \text{ et } 255)$ , représentant respectivement les "doses" de rouge, vert et bleu qui caractérisent la couleur du pixel.

S'il s'agit d'une image en niveau de gris, il est codé par 1 composante comprise au sens large entre  $0 \text{ et } 255$ , représentant la luminosité du pixel.

### 3.2 Types de l'image

Il existe différents types d'image numérique

#### 3.2.1 Les images de type bitmap

Une image bitmap est composée en mode point (Pixels). Le système de codage le plus universel consiste en effet à décomposer la représentation graphique, l'image, en un certain nombre de points élémentaires caractérisés par leurs coordonnées spatiales et leur

couleur. Il s'agit donc d'une représentation graphique définie par l'ensemble des points qui la composent [38].

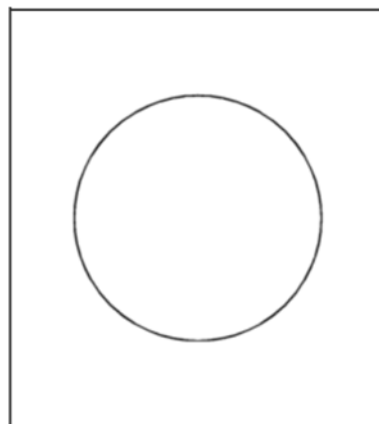
Format d'images bitmap : BMP, PCX, GIF, JPEG, TIFF, Raw, etc. Les photos numériques et les images scannées sont de ce type.



**Fig. I.2** Image bitmap.

### 3.2.2 *Les images de type vectoriel*

La description vectorielle d'une image consiste en une description géométrique: le document numérisé prend donc la forme d'une suite de formules mathématiques décrivant les formes élémentaires constituant l'image (carrés, rectangles, ellipses, cercles, courbes, etc.). Chaque forme élémentaire constitue un objet et se voit assigné un certain nombre d'attributs tels que la couleur, la transparence, l'épaisseur du trait, le type de trait (pointillé, etc.). Le stockage d'une image de type vectoriel est donc très différent de celui d'une image de type bitmap: il consiste en la mémorisation de la représentation des coordonnées des points caractéristiques des formes qui constituent l'image. Il s'agit dès lors d'une représentation relative que l'on peut qualifier de "symbolique".



**Fig. I.3** Image vectoriel.

### 3.2.3 Les types de format d'image

- **Image couleur RVB:** L'œil humain analyse la couleur à l'aide de trois types de cellules photo réceptrices '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.
- **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
- **Image binaire :** Une image binaire est une matrice rectangulaire dans laquelle l'élément vaut 0 ou 1. Lorsque l'on visualise une telle image, les 0 sont représentés par du noir et les 1 par du blanc [39].

### 3.3 Caractéristiques images numérique

Les caractéristiques images sont représentatives des différentes propriétés visuelles d'un objet dans une image ou une vidéo : couleur, forme, texture, apparence, mouvement, ainsi que des contraintes spatiales qui peuvent lier l'objet à son environnement, ou définir des interactions entre les régions le composant [40].

#### 3.3.1 Pixel

Le pixel est le plus petit élément que peuvent manipuler les matériels et logiciels d'affichage ou d'impression.

La quantité d'information que véhicule chaque pixel donne des nuances entre images monochromes et images couleur. Dans le cas d'une image monochrome, chaque pixel est codé sur un octet, et la taille mémoire nécessaire pour afficher une telle image est directement liée à la taille de l'image.

Dans une image couleur (R.V.B.), un pixel peut être représenté sur trois octets : un octet pour chacune des couleurs [41].

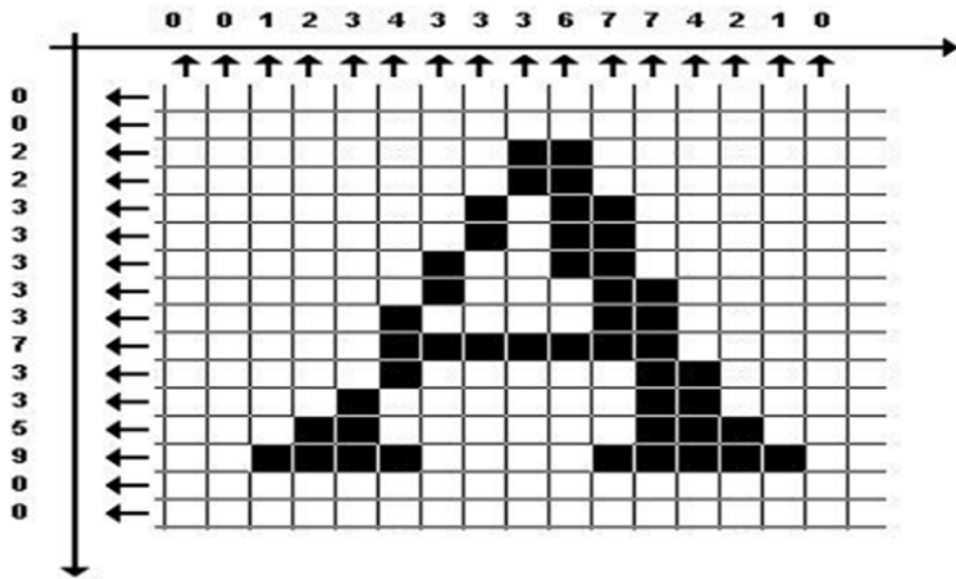


Fig. I.4 Groupe de pixel formant la lettre A.

### 3.3.2 Dimension

C'est la taille de l'image, cette dernière qui est représentée sous forme de matrice. Par conséquent, le nombre de lignes de cette matrice multiplié par le nombre de colonnes donne le nombre total de pixels dans l'image.

### 3.3.3 Resolution

C'est la clarté ou la fitness de détails atteinte par un moniteur dans la production de l'image, sur lequel la résolution est exprimée en nombre de pixel par unité de mesure (pouce ou centimètre). Elle désigne aussi le nombre de pixels affichables horizontalement ou verticalement sur un moniteur.

### 3.3.4 Contours

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 [42].

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 » [43].



**Fig. I.5** Les contours d'image.

### 3.3.5 *Bruit*

Un bruit (parasite) dans une image est considéré comme un phénomène de brusque variation de l'intensité d'un pixel par rapport à ses voisins, il provient de l'éclairage des dispositifs optiques et électroniques du capteur [44]. C'est un bruit 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.

### 3.3.6 *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 [45].

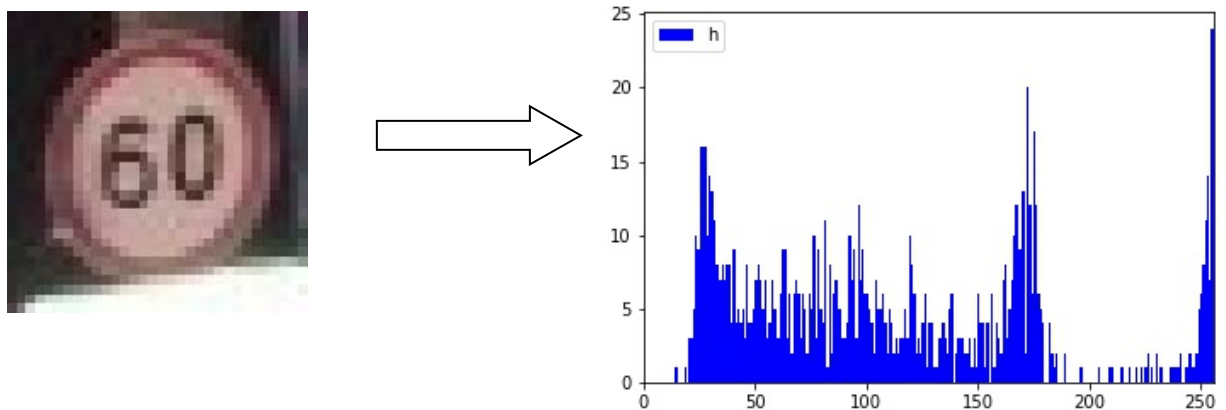
Une bonne luminance se caractérise par :

- 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.
- L'absence de parasites (bruit).

### 3.3.7 *Histogramme*

L'histogramme des niveaux de gris ou des couleurs d'une image est une fonction qui donne la fréquence d'apparition de chaque niveau de gris (couleur) dans l'image. Pour

diminuer l'erreur de quantification pour comparer deux images obtenues sous des éclairages différents, ou encore pour mesurer certaines propriétés sur une image, on modifie souvent l'histogramme correspondant [44]. Il permet de donner un grand nombre d'information sur la distribution des niveaux de gris (couleur) et de voir entre quelles bornes est répartie la majorité des niveaux de gris (couleur) dans les cas d'une image trop claire ou d'une image trop foncée. Il peut être utilisé pour améliorer la qualité d'une image (Rehaussement d'image) en introduisant quelques modifications pour pouvoir extraire les informations utiles de celle-ci.



**Fig. I.6** L'histogramme.

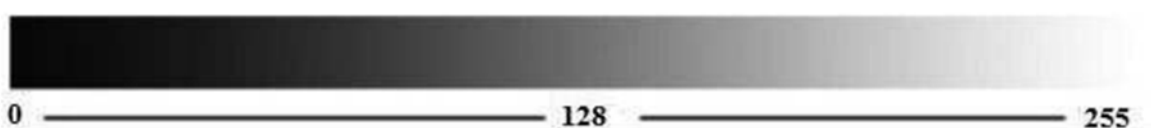
### 3.3.8 *Contraste*

C'est l'opposition marquée entre deux régions d'une image, plus précisément entre les régions sombres et les régions claires de cette image. Le contraste est défini en fonction des luminances de deux zones d'images [46].

### 3.3.9 *Images à niveaux 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. [42]



**Fig. I.7** Niveaux de gris.

## 4 Traitement d'images

Le traitement d'image est l'ensemble d'opérations qui permettent l'amélioration (filtrage, rehaussement de contraste), la modification (rotation, translation, ...) et l'extraction de l'information à partir des images (segmentation).

D'un point de vue syntaxique, les opérations de traitement d'images se scindent en trois catégories :

- Les opérations ponctuelles : l'image d'un pixel donné n'est influencée que par la valeur de ce dernier. Exemple : le négatif d'une image, l'égalisation de l'histogramme...
- Les opérations locales : l'image d'un pixel donné dépend de la valeur du pixel en question et de son voisinage. Exemple : le filtrage...
- Les opérations globales : l'image d'un pixel donné dépend de tous les pixels de la matrice qui représente l'image. Exemple : la transformée de Fourier d'une image [36].

### 4.1 Définition

C'est un ensemble des opérations relatives à la collecte, à l'enregistrement, à l'élaboration, à la modification, à l'édition, ... de données. Le principe général du traitement d'image est donc à quelques détails près, un système qui reçoit des images, y applique un traitement et produit une information de nature liée à l'application visée [47]. Une source de rayonnement envoie des ondes sur un objet qui sont ensuite réfléchies et collectées par un capteur. Le capteur transforme ces ondes en un ensemble de points. Ces points sont traités et une information est produite en sortie du système.

### 4.2 Prétraitement d'image

Prétraitement est une étape importante dans tout système de reconnaissance [49]. Le but de cette étape dans la reconnaissance d'objet est d'améliorer la lisibilité de l'image et de supprimer les détails qui n'ont pas de pouvoir discriminatif dans le processus de reconnaissance [50, 60].

L'étape de prétraitement comprend habituellement plusieurs tâches: binarisation, augmentation et normalisation.

#### 4.2.1 Conversion des images couleurs en niveau de gris

Pour les images couleurs, un pixel dispose généralement des trois composantes RVB (Rouge, Vert, Bleu). Un pixel gris a ses trois valeurs RVB identiques. Une méthode

simple pour convertir une image couleur en niveaux de gris pourrait être: calculer la moyenne des trois composantes RVB et utiliser cette valeur moyenne pour chacune des composantes.

$$Gris = \frac{(Rouge + Vert + Bleu)}{3} \quad \text{Éq. I. 1}$$

#### 4.2.2 *Binarisation*

La binarisation a pour but de transformer une image en niveau de gris en une image binaire. Cette étape est effectuée pour réduire le bruit et réduire la quantité d'information dans une image. Dépendant de la binarisation utilisée et de l'image, cette étape peut être source d'une perte d'information. Il est important d'identifier la binarisation la plus adaptée au document traité. La binarisation peut s'effectuer de manière simple par un seuillage statique, les valeurs supérieures à un seuil  $\theta$  donné prennent une valeur binaire, les valeurs inférieures prennent l'autre valeur

#### 4.2.3 *Normalisation*

La normalisation de la taille des images cherche à réduire les variations entre images dues à la taille des images afin d'améliorer les performances du reconnaiseur.

En général, la normalisation consiste à forcer les images à avoir une hauteur identique ([52,53]). Certains systèmes ([54]) proposent en plus de repropotion les images avec l'aide de leurs lignes de base. Les proportions de la normalisation ont pour effet d'accentuer l'importance de la zone entre les lignes de base par rapport à sa taille initiale.

### 5 *Conclusion*

Les techniques de traitement d'image sont des techniques très diverses et le choix de l'une parmi elle est un choix qui dépend essentiellement de la nature de l'application et des résultats qui peuvent être obtenus par l'application de l'une ou de l'autre.

Nous avons introduit dans ce chapitre les notions de base de l'image qui sert de fondement à la compréhension de différentes techniques de traitement d'images. Plusieurs méthodes classiques de traitement ont été proposées dans la littérature. Dans le chapitre suivant, nous effectuons une étude d'apprentissage automatique.

***CHAPITRE II :***  
***L'APPRENTISSAGE AUTOMATIQUE***

## ***1 Introduction***

L'intelligence artificielle (IA) est l'un des domaines les plus récents de la science et de l'ingénierie. IA comprend actuellement une grande variété de sous-domaines, comme jeux d'échecs, démonstration des théorèmes mathématiques, reconnaissance des panneaux de signalisation, diagnostic des maladies...

L'apprentissage automatique est l'une des extensions de l'IA. L'apprentissage automatique fait référence à la capacité d'un système à acquérir et intégrer de façon autonome des connaissances. Cette notion englobe toute méthode permettant de construire un modèle de la réalité à partir de données, soit en améliorant un modèle partiel ou moins général, soit en créant complètement le modèle.

L'apprentissage automatique fait référence au développement, l'analyse et l'implémentation de méthodes qui permettent à une machine d'évoluer et de remplir des tâches associées à une intelligence artificielle grâce à un processus d'apprentissage. Cet apprentissage permet d'avoir un système qui s'optimise en fonction de l'environnement, les expériences et les résultats observés [1].

L'apprentissage profond dérivé de l'apprentissage automatique, qui permet de construire des modèles qui ont montré des performances supérieures pour un large éventail d'applications, en particulier la vision par ordinateur et le traitement du langage naturel.

Dans ce chapitre nous allons présenter le principe de l'apprentissage en IA pour cela nous avons présenté quelques définitions de base associées à l'apprentissage, ainsi que les caractéristiques d'apprentissage et aussi les différentes classifications d'apprentissage automatique, pour les approches d'apprentissage en IA, il existe plusieurs types, nous avons présenté quelques approches d'apprentissages, qui sont les plus connues dans le domaine de l'IA, et nous avons présenté les réseaux de neurones convolutifs.

## ***2 Intelligence artificielle***

L'intelligence artificielle est une discipline qui vise à comprendre la nature de l'intelligence chez les humains ou dans la nature et reproduire cette intelligence sous forme de comportements réalisables par des machines et des ordinateurs. Bien qu'il est difficile de définir exactement l'intelligence, il est possible de définir quelques éléments entrant dans sa définition comme : la capacité d'exploiter les savoirs, la capacité de généralisation, la capacité de détecter des analogies entre des situations similaires et reproduire les mêmes raisonnements, la capacité d'omission sélective,...etc.

### 2.1 Une courte histoire de l'intelligence artificielle

L'intelligence artificielle a connu plusieurs étapes dans l'histoire, voici quelques étapes importantes dans son histoire :

- 1943 : McCulloch et Pitts créent le modèle du neurone formel.
- 1948 : Création de la cybernétique (science des systèmes) par Norbert Wiener.
- 1949 : Hebb établit la première règle d'apprentissage neuronal.
- 1950 Shannon, 1952 Samuel, 1953 Turing : machine pour jouer aux échecs.
- 1956 Workshop où est né le terme "intelligence artificielle".
- 1959 Rochester : Geometry Theorem Prover.
- 1958 McCarthy au MIT crée le LISP et le "time sharing". Créé DIGITAL.
- 1960 John McCarthy, Allen Bewell & Herbert Simon: L'ordinateur peut être utilisé pour autre chose que des calculs « manipuler des symboles » (idée proposée par Ada Lovelace amie de Babbage 1842).
- 1969 arrêt des RNs (Minsky & Paper 1969) limitations des perceptrons.
- 1969-1979 : systèmes experts.
- Depuis 1986 : retour des réseaux de neurone [2].

### 2.2 Les sous-domaines de l'intelligence artificielle

On s'en serait douté, créer des agents intelligents n'est pas si simple. Pour cette raison, l'IA s'est divisée en de nombreuses sous-disciplines qui essaient chacune de traiter une partie du problème [2]. Voici les principales :

- **Représentation des connaissances et raisonnement automatique** Comme son nom le suggère, cette branche de l'IA traite le problème de la représentation des connaissances (qui peuvent être incomplètes, incertaines, ou incohérentes) et de la mise en œuvre du raisonnement.
- **Résolution de problèmes généraux** L'objectif est de créer des algorithmes généraux pour résoudre des problèmes concrets.
- **Traitement du langage naturel** Ce sous-domaine vise à la compréhension, la traduction, ou la production du langage (écrit ou parlé).
- **Vision artificielle** Le but de cette discipline est de permettre aux ordinateurs de comprendre les images et la vidéo (par exemple, de reconnaître des panneaux routiers ou des chiffres).

- Robotique cette discipline vise à réaliser des agents physiques qui peuvent agir dans le monde
- **Apprentissage automatique** Dans cette branche de l'IA, on essaie de concevoir des programmes qui peuvent s'auto-modifier en fonction de leur expérience.

Il existe bien entendu des liens très forts entre ces sous-domaines. Par exemple, les langages développés dans la représentation des connaissances peuvent servir de base à des systèmes experts. Ou encore, beaucoup d'algorithmes pour la reconnaissance des formes sont développés en utilisant des méthodes d'apprentissage.

Il y a aussi de forts liens entre l'IA et d'autres domaines tels que la philosophie, la psychologie, les neurosciences, les sciences cognitives, la linguistique, et l'économie [2].

### 2.3 Les grands acteurs de l'intelligence artificielle

Google, Apple, Facebook et Amazon ne pouvaient pas ne pas s'investir fortement dans l'intelligence artificielle alors qu'ils possèdent des données en énorme quantité. En voici quelques exemples.



Google, très impliqué dans l'intelligence artificielle, procède habituellement par rachat. En 2014, Google a acheté la société anglaise DeepMind, qui avait développé des réseaux de neurones pour jouer aux jeux vidéo. Mais l'objectif avoué de DeepMind est actuellement de « comprendre ce qu'est l'intelligence ». DeepMind est célèbre pour son programme AlphaGo, qui a battu le champion du monde de go. En octobre 2017, le programme a franchi une étape supplémentaire : en jouant contre lui-même, non seulement son apprentissage a été plus court, mais surtout il est devenu plus fort que la version précédente. Nous avons ici un premier exemple d'apprentissage non supervisé, facilité par le fait que le contexte, à savoir les règles du jeu de go, est parfaitement mathématisable. Google a aussi son propre moteur de recommandation nommé Google Home, un haut-parleur et un assistant vocal disponible en trois versions différentes.



Amazon utilise de l'intelligence artificielle dans son moteur de recommandation, nommé Echo, et dans ses assistants basés sur son système de reconnaissance vocale, Alexa, disponible

en sept versions différentes. Via son offre de services dans le cloud, Amazon propose également des services fondés sur l'intelligence artificielle, comme la reconnaissance de la parole ou des robots de discussion, les fameux chatbot.

## facebook

Facebook est un énorme utilisateur d'intelligence artificielle. Il choisit les messages qu'il affiche en utilisant un moteur de type moteur de recommandation. Récemment, Facebook a mis en place un moteur d'intelligence artificielle pour détecter les tendances suicidaires. Comme le dit Joaquin Candela, directeur du département d'intelligence artificielle appliquée, « Facebook n'existerait pas sans intelligence artificielle ».

## Apple

Apple investit énormément dans l'intelligence artificielle et propose même un blog où sont expliquées ses recherches. Siri en est l'expression la plus évidente : ce système de reconnaissance de la parole très performant utilise des réseaux de neurones de dernière génération qui permettent de comprendre des commandes même dans des environnements très bruyants, comme dans la rue. Apple va également bientôt lancer son propre moteur de recommandation, nommé HomePod, un haut-parleur connecté avec Apple Music.

### *3 L'apprentissage automatique*

L'apprentissage automatique (machine learning en anglais), un des champs d'étude de l'intelligence artificielle, est la discipline scientifique concernée par le développement, l'analyse et l'implémentation de méthodes automatisables qui permettent à une machine d'évoluer grâce à un processus d'apprentissage, et ainsi de remplir des tâches qu'il est difficile ou impossible de remplir par des moyens algorithmiques plus classiques.

La faculté d'apprendre de ses expériences passées et de s'adapter est une caractéristique essentielle des formes de vies supérieures. Elle est essentielle à l'être humain dans les premières étapes de la vie pour apprendre des choses aussi fondamentales que reconnaître une voix, un visage familier.

L'apprentissage automatique est une tentative de comprendre et reproduire cette faculté d'apprentissage dans des systèmes artificiels. Il s'agit, très schématiquement, de concevoir des algorithmes capables, à partir d'un nombre important d'exemples, d'en assimiler la nature afin de pouvoir appliquer ce qu'ils ont ainsi appris aux cas futurs [3].

### **3.1 Définition**

Une définition de l'apprentissage automatique est celle donné par H. Simon [4]: "L'apprentissage dans un système est indiqué par les changements qu'il subit. Ces changements sont adaptatifs dans le sens où ils rendent possible au système de réaliser une même tâche, ou des tâches tirées d'une même population, d'une façon plus efficace et plus efficiente la prochaine fois qu'elle sera réalisée".

### **3.2 Domaines d'applications de l'apprentissage automatique**

L'apprentissage automatique s'applique à un grand nombre d'activités humaines et convient en particulier au problème de la prise de décision automatisée. Il s'agira, par exemple:

- Analyse et traitement automatique des images.
- De la reconnaissance des formes.
- Analyse et traitement automatique des images.
- Interprétation et Reconnaissance de formes et de la parole humaine.
- Recherche d'information (moteur internet...).
- Découvrir les relations entre les données dans les grandes bases de données.
- Faciliter l'exploitation des connaissances qui ne puissent pas être explicitées par le cerveau humain.
- aider les experts humains dans la prise de décisions complexes.
- De la reconnaissance de la parole et du texte écrit.
- Etc...

### **3.3 Types d'apprentissage**

L'apprentissage automatique offre une grande variété d'approches qui peuvent être choisies pour capturer des tendances cachées à partir des données d'apprentissage [5]. Nous classons ces techniques en quatre principales catégories :

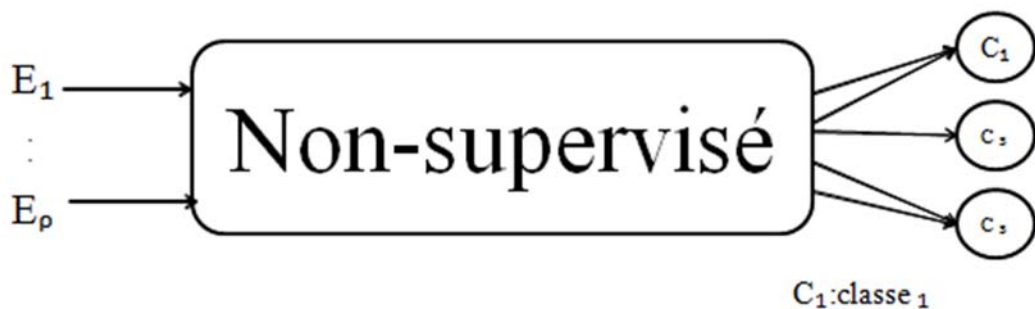
- Apprentissage par renforcement,
- Apprentissage non supervisé,
- Apprentissage semi-supervisé,
- Apprentissage supervisé.

### 3.3.1 Apprentissage par renforcement

L'apprentissage par renforcement est une bête très différente. Le système d'apprentissage, appelé agent dans ce contexte, peut observer l'environnement, sélectionner et effectuer des actions, et obtenir des récompenses en retour. Il doit ensuite apprendre par lui-même quelle est la meilleure stratégie, appelée politique, pour obtenir le plus de récompense possible. Une politique définit quelle action l'agent doit choisir lorsqu'il se trouve dans une situation donnée [6].

### 3.3.2 Apprentissage non-supervisé

Il est aussi appelé apprentissage par corrélation. Ce type d'apprentissage est utilisé dans les cas où nous avons une base d'apprentissage dont les exemples de données ne sont pas étiquetés à l'avance. L'apprentissage non supervisé consiste à regrouper les différentes formes en classes sur la base d'un critère de similarité choisi dans la technique. La Fig. II.1 suivante montre l'apprentissage non-supervisé.



**Fig. II.1** L'apprentissage non-supervisé.

Ce type d'apprentissage permet la construction automatique des classes sans aucune intervention. Cependant, cette approche nécessite une bonne estimation du nombre de classes [6] :

- **Estimation de densité**

Dans un problème d'estimation de densité, on cherche à modéliser convenablement la distribution des données. L'estimateur obtenu doit pouvoir donner une bonne estimation de la densité de probabilité à un point de test issu de la même distribution (inconnue) que les données d'apprentissage.

- **Partitionnement**

Un algorithme de partitionnement divise l'espace d'entrée en un certain nombre de "classes" en se basant sur un ensemble de données d'apprentissage ne contenant aucune information sur

la classe de sortie. Les critères utilisés pour décider si deux points devraient appartenir à la même classe ou à des classes différentes sont spécifiques à chaque algorithme, mais sont très souvent liés à une mesure de distance entre points.

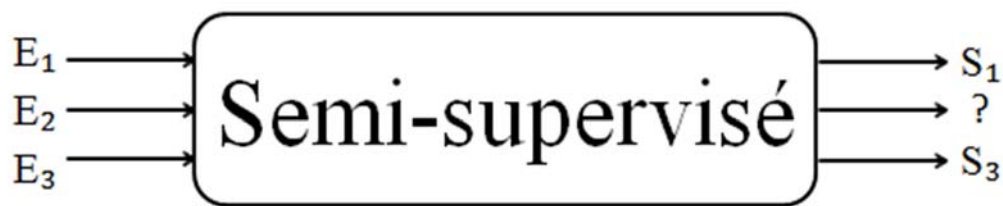
### - Réduction de dimensionnalité

Le but d'un algorithme de réduction de dimensionnalité est de parvenir à "résumer" l'information présente dans les coordonnées d'un point en haute dimension par un nombre plus réduit de caractéristiques. Le but espéré est de préserver l'information "importante", de la mettre en évidence en la dissociant du bruit, et possiblement de révéler une structure sous-jacente qui ne serait pas immédiatement apparente dans les données d'origine en haute dimension.

### 3.3.3 Apprentissage semi-supervisé

L'Apprentissage semi-supervisé est un paradigme de l'apprentissage automatique concerné par l'étude de la façon dont les ordinateurs et les systèmes naturels tels que les humains apprennent, en présence des exemples de données étiquetées et non étiquetées. L'Apprentissage semi-supervisé a un grand intérêt car il peut utiliser des exemples de données non étiquetées pour améliorer les tâches d'apprentissage supervisé lorsque les données étiquetées sont rares ou coûteux dans le but de comprendre comment la combinaison de données étiquetées et non étiquetées peuvent modifier le comportement de l'apprentissage [7].

La Fig. II.2 suivant montre l'apprentissage semi-supervisé.



**Fig. II.2** L'Apprentissage semi-supervisé.

### 3.3.4 Apprentissage supervisé

L'apprentissage supervisé est tout simplement une formalisation de l'idée d'apprendre à partir d'exemples. L'objectif de ce paradigme consiste à déterminer des règles typiques qui peuvent être utilisées pour différencier entre les données de différentes classes. Les approches de l'apprentissage supervisé utilisent un ensemble d'exemples de données pour construire un modèle, où chaque exemple de données a une étiquette de classe déterminée. D'où le travail d'apprentissage supervisé consiste à analyser les ressemblances entre les formes d'une même

classe et les dissemblances entre les formes de classes différentes pour déduire les meilleures séparations possibles afin qu'il puisse identifier des nouveaux exemples avec une grande précision [8]. La Fig. II.3 suivante montre l'apprentissage supervisé.



Fig. II.3 L'apprentissage supervisé.

### 3.4 Les approches d'apprentissage

Dans ce qui suit, nous allons présenter quelques approches d'apprentissage automatique, qui sont connues dans le domaine de l'intelligence artificielle :

#### 3.4.1 Les réseaux neuronaux artificiels

##### 3.4.1.1 Le neurone biologique

**Le neurone :** Un neurone est une cellule qui est à la base du fonctionnement du système nerveux. Le neurone est principalement composé de trois parties

- le corps cellulaire,
- les axones qui transigent l'information,
- les dendrites qui la reçoivent d'autres neurones.

Un neurone n'a qu'un seul axone mais peut posséder plusieurs dendrites. La communication se fait au niveau des synapses qui agissent de manière chimique par sécrétion ou réception d'hormones. Les neurones ont donc la possibilité de pouvoir communiquer des informations. C'est ce qui fait la force et la flexibilité du système nerveux et suscité l'intérêt des mathématiciens pour les formaliser en un modèle utilisable [9].

La Fig. II.4 représente un neurone biologique et ses composantes.

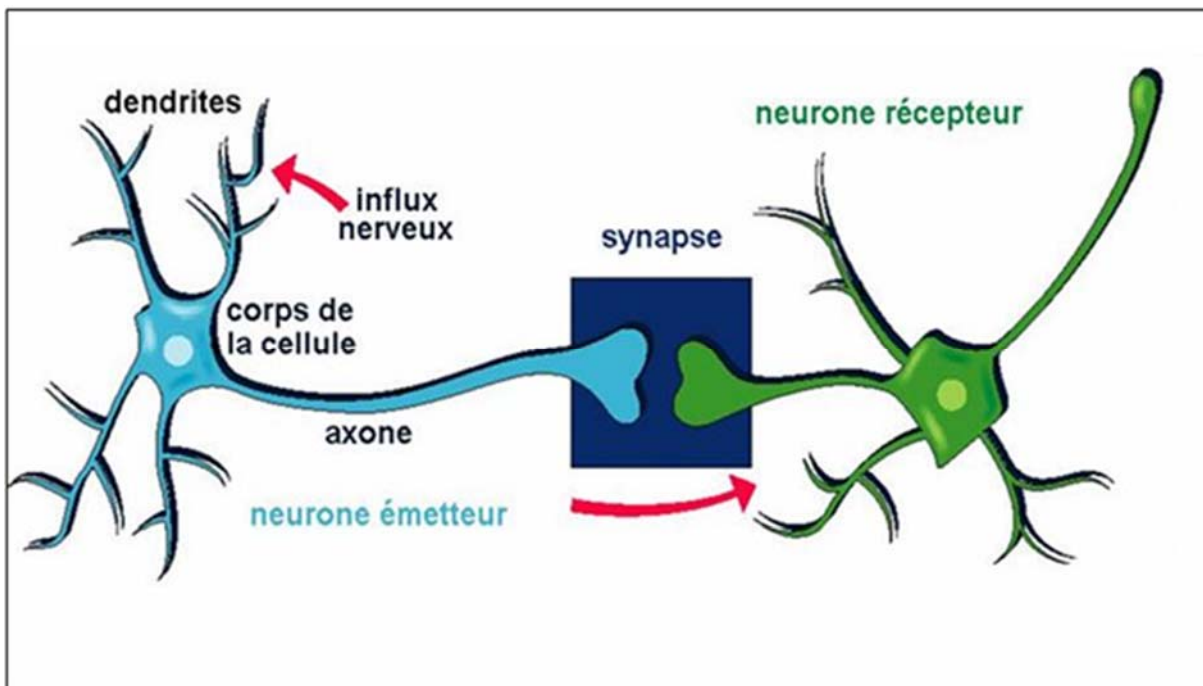


Fig. II.4 Neurones biologiques.

### 3.4.1.2 Le neurone formel

Le neurone formel est un modèle qui se caractérise par un état interne  $s \in S$ , des signaux d'entrée  $x_1, x_2, \dots, x_p$  et une fonction de transition d'état [10].

$$s = h(x_1, \dots, x_p) = f(\beta_0 + \sum_{j=1}^p \beta_j x_j) \quad \text{Éq. II.1}$$

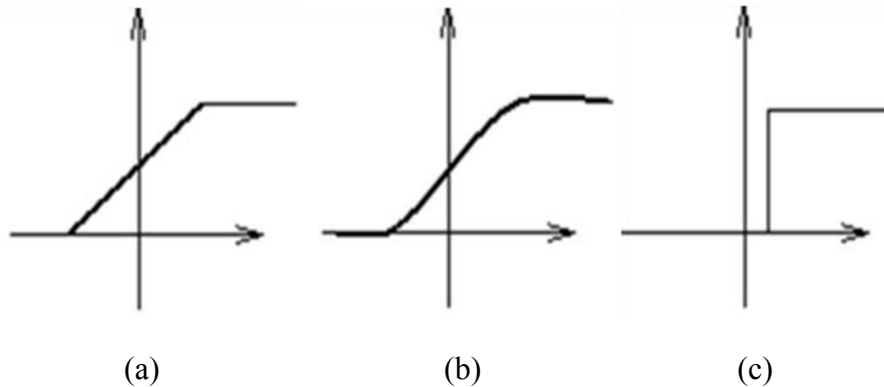


Fig. II.5 Représentations d'un neurone formel.

La fonction de transition opère une transformation d'une combinaison affine des signaux d'entrée,  $\beta_0$  étant appelé le biais du neurone. Cette combinaison affine est déterminée par un vecteur de poids  $[\beta_0, \beta_1, \dots, \beta_p]$  associée à chaque neurone et dont les valeurs sont estimées dans la phase d'apprentissage. Ils constituent "la mémoire" ou "connaissance répartie" du réseau.

Les différents types de neurones se distinguent par la nature  $f$  de leur fonction de transition. Les principaux types sont :

- linéaire,
- sigmoïde,
- seuil,
- ...



**Fig. II.6** Les fonctions de transition (a) linéaire, (b) sigmoïde, (c) seuil.

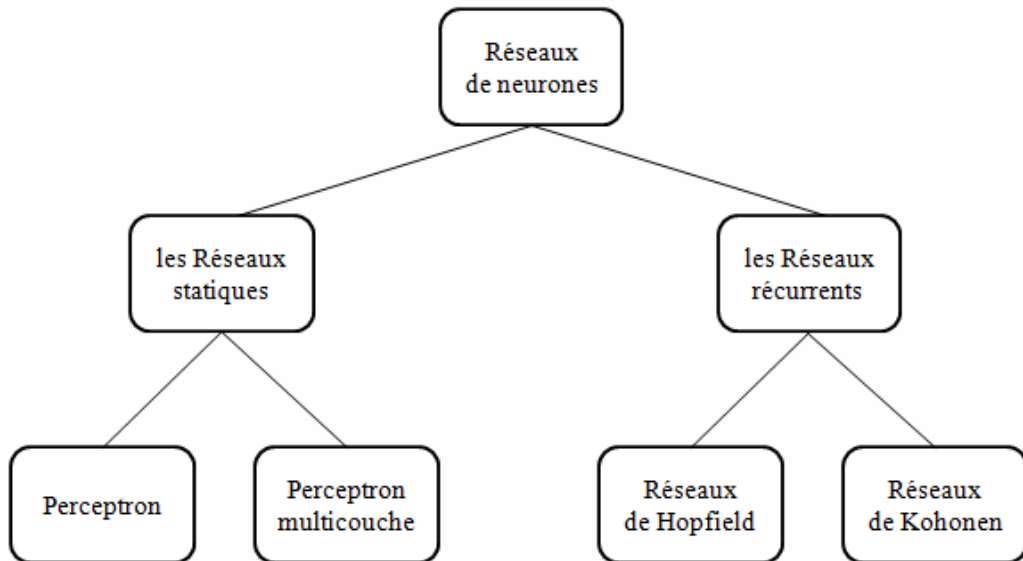
Les modèles linéaires et sigmoïdaux sont bien adaptés aux algorithmes d'apprentissage comme celui de rétro propagation du gradient car leur fonction de transition est différentiable. Ce sont les plus utilisés.

### 3.4.1.3 Les différents types de réseaux de neurones

Les réseaux de neurones sont classés selon deux principales topologies :

- Les réseaux statiques «non bouclé» : où les données circulent de l'entrée vers la sortie à travers les liens sans possibilité de retour en utilisant ces mêmes liens (transmission unidirectionnelle). Dans ce cas, le réseau est similaire à un graphe orienté.
- Les réseaux récurrents: avec existence des retours entre l'entrée et la sortie. Dans ce cas, le réseau est similaire à un graphe non orienté.

La Fig. II.7 montre les différentes classes de réseau de neurones ainsi que quelques exemples de réseaux pour chaque classe.



**Fig. II.7** Classification des réseaux de neurones.

- *Les réseaux statiques*

- *Perceptron simple :*

Un perceptron simple ou perceptron monocouche est un réseau de neurones développé par Rosenblatt [11]. Ce réseau est constitué d'un seul neurone.

- *Perceptron Multicouche :*

Un perceptron multicouche est constitué de trois types de couches [12]:

- Une couche d'entrée qui correspond aux données d'entrée  $x = [x_1, \dots, x_n]^T$ . Cette couche ne contient pas de neurones.
- Une couche de sortie constituée de  $K$  neurones et produisant les sorties du réseau  $y = [y_1, \dots, y_n]^T$  c'est-à-dire les valeurs de sortie associées aux données d'entrée  $x$ .
- Des couches cachées constituées chacune de plusieurs neurones. Ces couches permettent la transformation non-linéaire du signal d'entrée vers le signal de sortie.

Dans le perceptron multicouche, tous les neurones d'une couche sont connectés aux neurones de la couche précédente. La Fig. II.8 illustre un perceptron multicouche constitué une couche cachée.

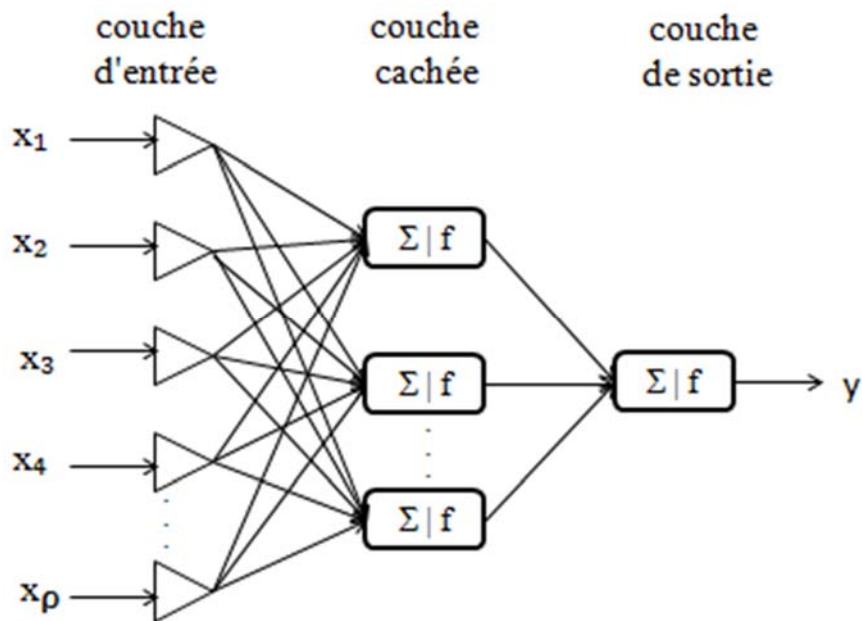


Fig. II.8 Exemple de perceptron multicouche avec une couche cachée et une couche de sortie.

- *Les réseaux récurrents*
  - *Réseaux de Hopfield (HNN) :*

Les réseaux de hopfield [13] sont des réseaux dans lesquels tous les neurones sont interconnectés par des connexions à poids réels. Ici, il n'y a qu'une seule couche et la couche de sortie est assimilée à la couche d'entrée. De plus, chaque neurone possède une valeur d'activation binaire (0/1 ou -1/1). L'ensemble des valeurs d'activation des neurones d'un réseau de hopfield représente son état et sa sortie.

La Fig. II.9 montre un réseau de hopfield contenant 6 neurones.

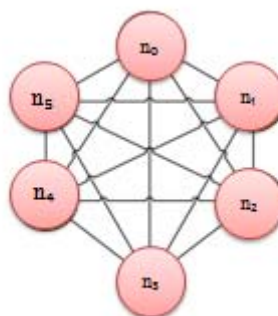


Fig. II.9 Un réseau de hopfield.

- *Réseaux de Kohonen :*

La carte de Kohonen [14] est une carte auto-organisatrice en deux couches (voir la Fig. II.10), une couche d'entrée et une autre de sortie. La couche de sortie appelée aussi couche compétitive est en deux dimensions. Chaque neurone d'entrée est connecté à l'ensemble des neurones de sortie par des poids  $W_{ij}$ . La carte de Kohonen est une carte à apprentissage non supervisé, à partir d'une base d'apprentissage constitué seulement des éléments d'entrées, elle permet de regrouper ces entrées en classes.

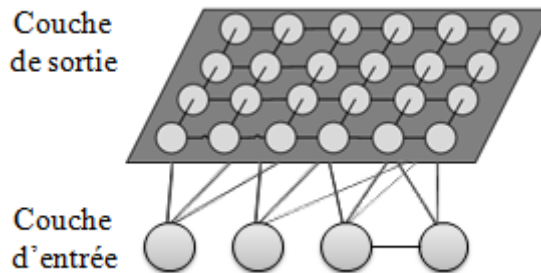


Fig. II.10 D'une carte de Kohonen.

#### 3.4.1.4 L'apprentissage des réseaux de neurones

La recherche des algorithmes d'apprentissage amené à d'importantes découvertes, qui ont changé la manière de concevoir le problème de rendre les ordinateurs plus intelligents. Un agent est intelligent parce qu'il a des connaissances opérationnelles qui lui permettent d'effectuer certaines tâches ou de répondre à certaines questions sur un certain domaine.

L'objectif de l'apprentissage est de fournir une méthode au réseau, afin qu'il puisse ajuster ses paramètres lorsqu'on lui présente des exemples à traiter. En effet les réseaux de neurones ont été appliqués avec succès à l'apprentissage de tâches de classification. L'apprentissage est le processus d'adaptation des paramètres d'un système pour donner une réponse désirée à une entrée ou stimulation quelconque.

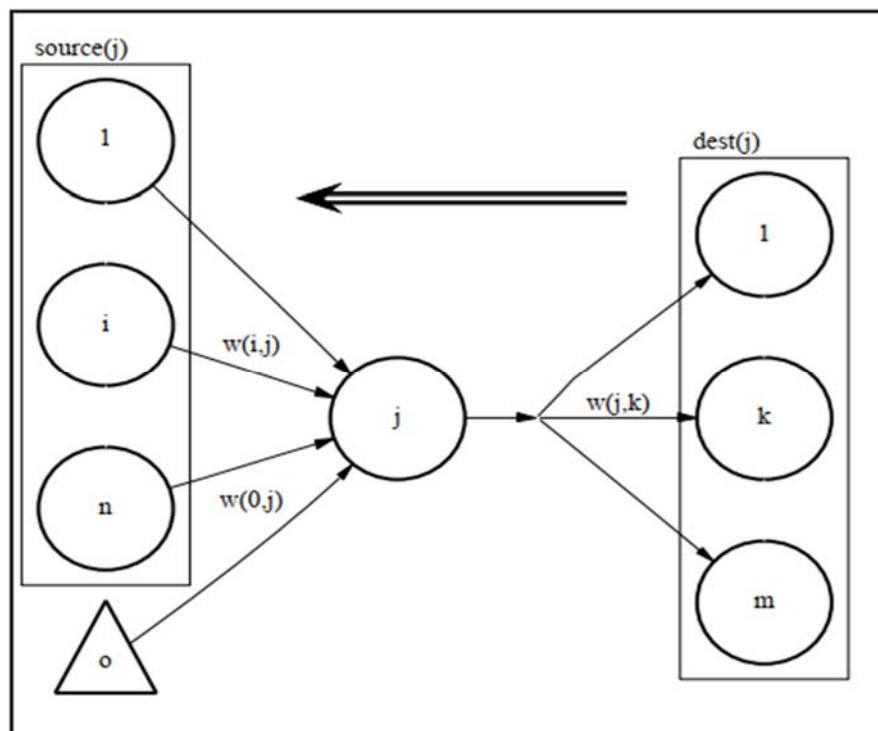
##### 3.4.1.4.1 Les algorithmes

Les algorithmes d'apprentissage pour les réseaux de neurones artificiels (réseaux de neurones) ont été développés pour la première fois dans [15] sous le terme perceptrons. Toutefois, l'optimisation de tels réseaux était ardue, et ce n'est que dans les années 80, lorsque Rumelhart et al. [16] découvrirent la rétropropagation du gradient, que la popularité des réseaux de neurones grandit. Ils ont depuis lors été utilisés dans toutes sortes de problèmes, de la reconnaissance de caractères.

- *Rétropropagation*

*L'apprentissage par rétropropagation du gradient de l'erreur :*

C'est seulement en 1986 que la généralisation de la règle delta aux réseaux à couches cachées a été formulée. Cette généralisation, la règle de la rétropropagation du gradient de l'erreur, consiste à propager l'erreur obtenue à une unité de sortie d'un réseau à couches comportant une ou plusieurs couches cachées à travers le réseau par descente du gradient dans le sens inverse de la propagation des activations[1]. La Fig. II.11 montre une illustration du principe.



**Fig. II.11** Schéma du modèle de la rétropropagation de l'erreur. La modification à apporter aux poids entre la couche source(j) et le neurone formel j ne peut être calculée que si on connaît déjà la modification qu'il faut apporter aux poids  $w(j; k)$  entre j et les éléments de dest(j).

Un réseau à couches est composé d'un ensemble de neurones formels groupés en sous-ensembles distincts (les couches) de telle sorte qu'il n'y ait aucune connexion entre deux neurones d'une même couche.

À la fin de l'apprentissage, lorsque le réseau a appris à modéliser son environnement, le comportement souhaité du réseau est le suivant : on présente un vecteur d'entrée au réseau, celui-ci propage vers la sortie les valeurs d'activation correspondantes (en utilisant une règle de propagation), afin de générer, par l'intermédiaire des neurones de sortie, un vecteur de

sortie. Celui-ci devrait correspondre à la sortie désirée, telle qu'apprise lors de la phase d'apprentissage.

La généralisation de la règle delta aux réseaux multicouches utilise une méthode de descente du gradient, permettant de calculer la modification des poids des connexions entre les couches cachées. Afin de pouvoir calculer le gradient de l'erreur par rapport aux poids du réseau, la fonction de sortie d'un neurone doit être différentiable et non linéaire (sinon, on pourrait réduire le réseau à un perceptron). La fonction la plus souvent utilisée est, la sigmoïde :

$$y_j = f(\sigma_j) = \frac{1}{1+e^{-\sigma_j}} \quad \text{Éq. II.2}$$

La règle delta généralisée dicte alors le changement de poids entre le neurone  $i$  et le neurone  $j$  de la façon suivante :

$$\Delta w(i, j) = \alpha \delta_j y_i \quad \text{Éq. II.3}$$

C'est-à-dire de façon proportionnelle à une mesure d'erreur  $\delta_j$  caractéristique du neurone  $j$  et à la valeur d'entrée noté ici  $y_i$ . Pour les connexions aboutissant aux neurones de sortie, cette mesure d'erreur est évidemment calculée ainsi :

$$\delta_j = (u_j - y_j) y_j (1 - y_j) \quad \text{Éq. II.4}$$

Le calcul de l'erreur aux unités cachées se fait ensuite récursivement par la descente du gradient. Soit  $\text{dest}(j)$  l'ensemble des neurones auxquels  $j$  se connecte :

$$\delta_j = y_j (1 - y_j) \sum_{k \in \text{dest}(j)} \delta_k w(j, k) \quad \text{Éq. II.5}$$

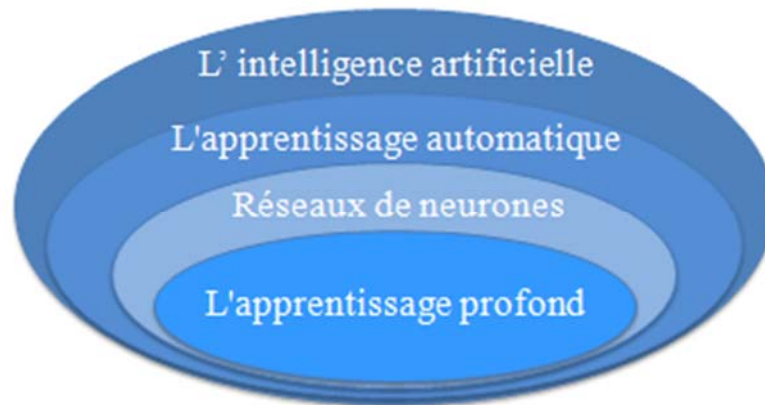
Lorsque l'on applique la règle delta généralisée sur le réseau de façon itérative pour un ensemble de vecteurs d'entrées, le réseau tentera de minimiser l'erreur obtenue à la sortie, et donc de modéliser le mieux possible la fonction désirée entre les entrées et les sorties.

#### **4 L'apprentissage profond**

L'apprentissage profond (ou Deep Learning) est un ensemble de méthodes d'apprentissage automatique tentant de modéliser des connaissances avec un haut niveau d'abstraction grâce à des architectures hiérarchiques de neurones artificiels. Ces techniques ont permis des progrès importants et rapides dans les domaines de l'analyse du signal sonore ou visuel et notamment de la reconnaissance faciale, de la reconnaissance vocale, de la vision par ordinateur, du traitement automatisé du langage.

Est une technique d'apprentissage dit « profond » permettant à un programme informatique d'apprendre à identifier le contenu d'une image, ou de comprendre le langage naturel. C'est une forme d'intelligence artificielle qui tente grosso modo d'imiter le fonctionnement neuronal du cerveau humain, par le biais d'un réseau de neurones artificiels composé de milliers d'unités de calculs [17].

L'apprentissage profond apprend à un modèle informatique comment réaliser des tâches de classification directement à partir d'images, de textes ou d'audio. Les modèles d'apprentissage profond peuvent atteindre un niveau de précision exceptionnel, parfois supérieur aux performances humaines. L'entraînement des modèles s'effectue via un vaste ensemble de données labellisées et d'architectures de réseaux de neurones qui contiennent de nombreuses couches



**Fig. II.12** Représentation schématique de l'intelligence artificielle.

#### ***4.1 La différence entre l'apprentissage automatique et l'apprentissage profond***

L'apprentissage profond est une branche particulière de l'apprentissage automatique. Un processus de l'apprentissage automatique commence par l'extraction manuelle de caractéristiques pertinentes à partir d'images. En s'appuyant sur ces caractéristiques, un modèle qui catégorise les objets de l'image est ensuite créé. Dans un processus d'apprentissage profond, l'extraction de caractéristiques pertinentes à partir d'images est automatique. En outre, l'apprentissage profond effectue un apprentissage « de bout en bout » : à partir de données brutes, un réseau se voit assigner des tâches à accomplir (une classification, par exemple) et apprend comment les automatiser.

Une autre différence majeure est le fait que les algorithmes d'apprentissage profond évoluent avec les données, tandis que l'apprentissage peu profond converge. L'apprentissage peu profond désigne les méthodes d'apprentissage automatique dont la progression s'arrête à partir d'un certain niveau de performance après l'alimentation du réseau en exemples supplémentaires et en données d'apprentissage.

Un des avantages majeurs des réseaux d'apprentissage profond réside dans leur capacité à continuer à s'améliorer en même temps que le volume de vos données augmente.

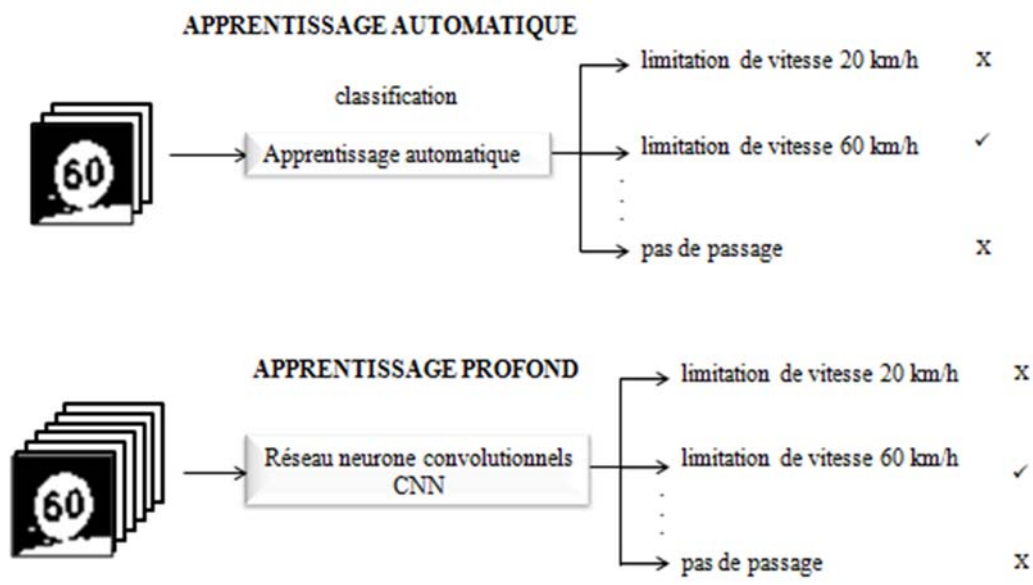


Fig. II.13 Comparaison de méthodes de catégorisation de plaque d'apprentissage automatique et d'apprentissage profond.

Pour classer des images avec l'apprentissage automatique, les choix de caractéristiques et de classificateur doivent être effectués manuellement. Avec l'apprentissage profond, l'extraction de caractéristiques et le processus de modélisation sont automatiques [17].

4.2 Les algorithmes de l'apprentissage profond

Il existe différents algorithmes de l'apprentissage profond. Nous pouvons ainsi citer [17]:

- **Les réseaux de neurones profonds** (Deep Neural Networks): Ces réseaux sont similaires aux réseaux de perceptron multicouche mais avec plus de couches cachées. L'augmentation du nombre de couches, permet à un réseau de neurones de détecter de légères variations du modèle d'apprentissage, favorisant le sur-apprentissage ou sur-ajustement (« overfitting »).

- **Les réseaux de neurones convolutifs** (CNN ou Convolutional Neural Networks). Le problème est divisé en sous parties, et pour chaque partie, un « cluster » de neurones sera créer afin d'étudier cette portion spécifique. Par exemple, pour une image en couleur, il est possible de diviser l'image sur la largeur, la hauteur et la profondeur (les couleurs).

#### 4.2.1 Réseaux de neurones convolutifs (cnn)

Le réseau de neurone convolutif CNN est un type de réseaux neuronaux très prometteur dans le domaine du traitement de l'image [18].

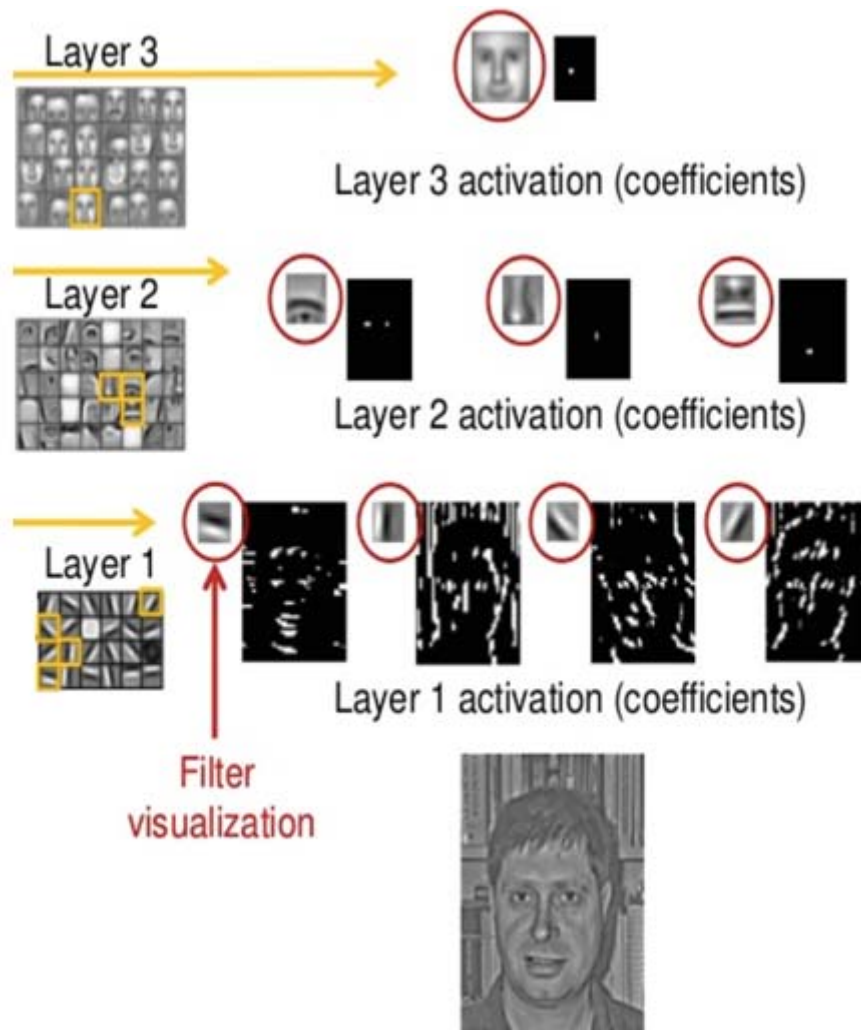
Un CNN est un réseau de neurones dont les couches sont liées non pas par une opération matricielle mais par une convolution en 2 dimensions. Les CNN permettent de traiter directement de grandes quantités de données que sont les images pour plusieurs raisons :

- Les images sont corrélées spatialement (les valeurs des pixels adjacents sont généralement très proches) et les couches de convolutions permettent de créer des liens entre ces données corrélées spatialement.
- Les couches de convolutions sont généralement suivies de couches de sous échantillonnage (pooling) qui diminuent grandement la taille des données.

Cela implique que les entrées des CNN sont des données de dimensions 2 ou 3. Chaque donnée de dimension 2 est analogue à une image alors que la 3ieme dimension représente le nombre de canaux. On peut citer comme exemple de donnée d'entrée :

- Une image noir et blanc de dimension L\*H : dimension 2 (L\*H\*1) car un seul canal
- Une image en couleur de dimension L\*H : dimension 3 (L\*H\*3) car plusieurs canaux d'entrées (Rouge, Vert et Bleu)

Suite à l'apprentissage, les filtres de convolutions arrivent à repérer des formes particulières sur les images. Les formes que les filtres détectent changent par rapport à la couche où se trouve les filtres de convolution. Comme on peut le voir sur La Fig. II.14, les filtres des différentes couches se spécialisent dans des domaines différents : ceux de la première couche seront des détecteurs de contours, ceux de la deuxième des détecteurs de parties de panneau routière, et ceux de la dernière des détecteurs de panneau routier entière. Ainsi, on comprend facilement que les CNN se sont fait une place de choix dans la Vision par Ordinateur, et en particulier la reconnaissance d'objets.



**Fig. II14** Spécialisation des filtres d'un CNN [19].

Historiquement, les premiers CNN ayant eu du succès ont été inventés en 1989 par LeCun [20] et appliqués à la reconnaissance d'écriture manuscrite. Ce système a été perfectionné par son auteur en 1998 [21], pour être maintenant utilisé dans les machines de dépôt automatique de chèques bancaires. Ces réseaux composés de couches de convolutions suivi par des couches de type perceptron multicouche sont appelés réseaux de type LeNet.

#### - *LeNet-5*

LeCun et al ont proposé une version moderne de CNN: c'est LeNet-5 qui a obtenu la performance sur plusieurs ensembles de données de reconnaissance de motifs [22]. LeNet-5 contient plusieurs couches et il est formé avec l'algorithme de rétropropagation dans une formulation end to end, C'est-à-dire il se base sur la classification des motifs visuels directement en utilisant des images brutes. Cependant, le Net5 souffre de certaines limites telles que l'échelle des données limitée et du pouvoir de calcul petit. Pour mieux former des réseaux profonds, Hinton et al ont introduit Deep Belief Networks (DBN) pour former chaque

couche de DBN de manière non supervisée [23]. Et les chercheurs ont développé plus de méthodes pour surmonter les difficultés d'apprentissage des architectures CNN. En particulier, AlexNet [24] a été proposé Par Krizhevsky et al en 2012 et il a été appliqué avec succès à la classification d'image à grande échelle dans le défi d'ImageNet. AlexNet contient cinq couches convolutives suivies de 3 couches entièrement connectées. Inspiré par AlexNet, plusieurs variantes, dont VGGNet [25], GoogLeNet [26] et ResNet [27], ont été proposés pour améliorer encore la performance des CNN sur les tâches de reconnaissance visuelle.

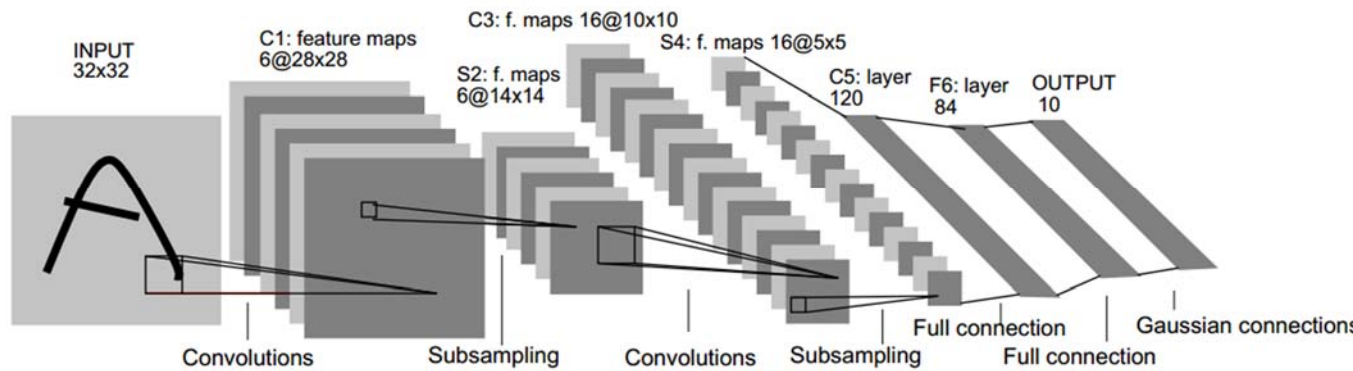


Fig. II.15 Réseau de neurones LeNet-5.

Nous allons décrire plus en détails les différentes couches ci-dessous.

#### 4.2.1.1 La couche Convolution

La convolution est énormément utilisée dans le traitement de l'image, une image ayant plusieurs représentations est de dimension 3 = longueur, largeur, profondeur ou de dimension 2 = longueur, largeur.

Mathématiquement, La formule de couche de convolution est définie comme suite [28]:

$$W_x^n = \frac{W_x^{n-1} - F_x^n + 2P}{S_x^n} + 1 \quad \text{Éq. II.6}$$

$$H_y^n = \frac{H_y^{n-1} - F_y^n + 2P}{S_y^n} + 1 \quad \text{Éq. II.7}$$

- $W_x^{n-1}, H_y^{n-1}$  la taille du volume d'entrée.
- $F_x^n$ : La taille spatiale du volume de sortie.
- P: la taille de la marge (zero padding).
- $S_y^n$ : le pas (stride).

- $W_x^n$ ,  $H_y^n$  : la taille du volume de sortie et L'indice n indique la couche.

Cette dernière sera représenter une image « cartes caractéristiques » de taille inférieur à la taille de la couche de convolution. Suite à cette opération, on va faire passer chacun des neurones résultants à travers une fonction d'activation.

Il y a plusieurs types de convolutions possible, en effet la convolution peut être effectué avec ou sans padding :

- Sans padding, la convolution ne dépasse pas les bords de l'image initiale, le résultat sera donc plus petit que ce qu'on avait à l'origine. Pour une image de dimension  $(l, m)$  et un filtre de dimension  $(r, s)$  on obtient une convolution de dimension  $l - r, +1, m - s + 1$ , c'est le cas de l'exemple de La Fig. II.16.
- Avec padding, le principe est de rajouter des valeurs au bord de notre image d'entrée afin d'avoir une image de même taille en sortie.

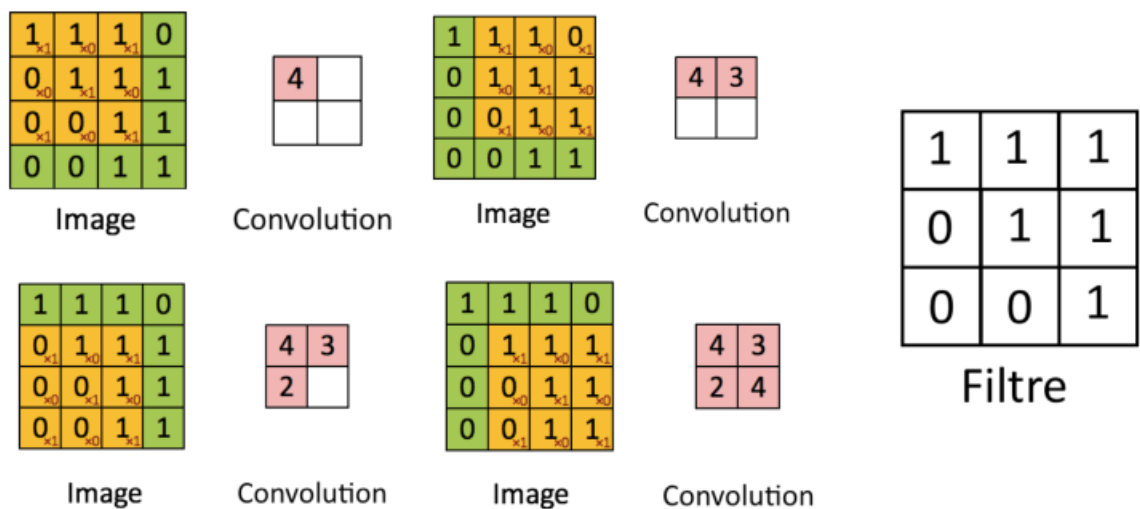


Fig. II.16 Principe d'une convolution.

**Précisons ici les différences par rapport au perceptron multicouche :**

- Les paramètres que l'on souhaite optimiser ici sont les poids des différents filtres de convolution, ce ne sont plus les poids qui liaient simplement un neurone à un autre dans les couches du perceptron multicouche.
- Chaque « pixel » de chaque carte caractéristique est un neurone.
- $f$  est une fonction d'activation similaire.

4.2.1.2 Couche de pooling

La couche de pooling (couche d'agrégation) permet de diminuer la taille comme décrit dans La Fig. II.17.

Les formules ci-dessous illustrent le calcul de la taille de pooling [28]:

$$W_x^n = \frac{W_x^{n-1} - F_x^n}{S_x^n} \quad \text{Éq. II. 8}$$

$$H_y^n = \frac{H_y^{n-1} - F_y^n}{S_y^n} \quad \text{Éq. II. 9}$$

- $W_x^{n-1}, H_y^{n-1}$ : la taille du volume d'entrée.
- $F_y^n$ : La taille spatiale du volume de sortie.
- $S_y^n$ : le pas (stride).
- $W_x^n, H_y^n$ : la taille du volume de sortie et L'indice n indique la couche.

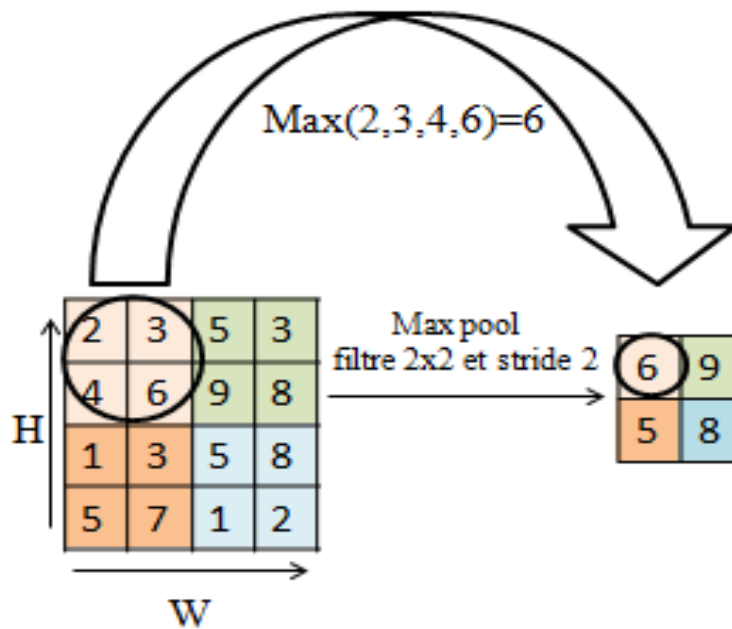


Fig. II.17 La couche de pooling.

Les paramètres du pooling, sont :

- Son type car le pixel sortant peut être issu d'une fonction max, moyenne ou min (ici : max) ;
- Le pas (stride) qui correspond au déplacement du filtre (ici : 2 en hauteur et largeur) ;
- La taille des filtres (ici : 2x2).

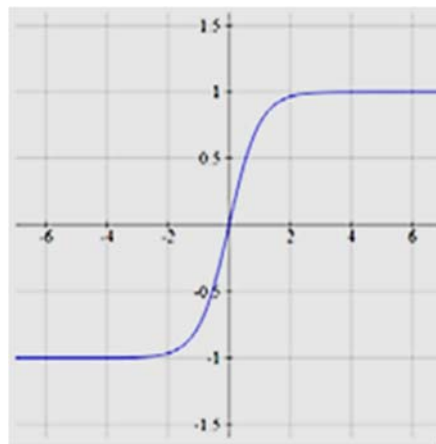
#### 4.2.1.3 Les fonctions d'activation

Il existe différentes fonctions d'activation permettant la non-linéarité dans les différentes couches des CNN. Parmi les plus connues :

##### - La tangente hyperbolique :

L'une des fonctions d'activation les plus populaires est la fonction tangente hyperbolique (Equation). L'entrée est une combinaison linéaire pondérée des entrées du nœud. Cette fonction fonctionne le plus efficacement sur les entrées dans la plage (0; 1), produisant des sorties en intervalle (-1; 1).

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad \text{Éq. II.10}$$

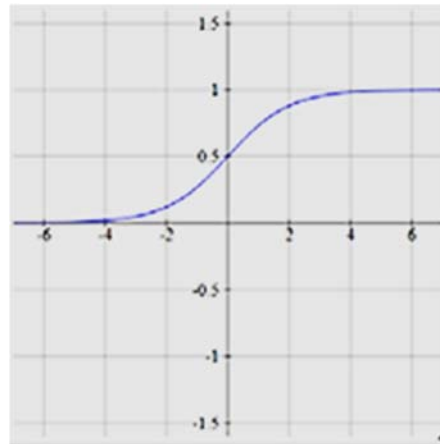


**Fig. II.18** La tangente hyperbolique.

##### - La fonction sigmoïde

Est une fonction d'activation largement utilisée biologiquement plus plausible que la tangente hyperbolique [29]. Utilisée notamment dans le perceptron multicouche original. La formule de La fonction sigmoïde est définie comme suite :

$$f(x) = \frac{1}{1 + e^{-x}} \quad \text{Éq. II.11}$$

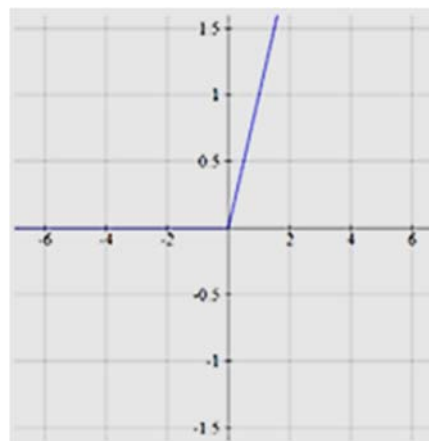


**Fig. II.19** La sigmoïde.

- **La fonction d'Unité de Rectification Linéaire (ReLU)**

Est utilisée dans le but d'augmenter la non-linéarité du réseau. Ils bénéficient de leur simplicité, entraînant une formation plus rapide et des améliorations de performance dans des cas particuliers [30], et donc souvent utilisés dans les réseaux de neurones convolutifs. ReLU est donné par l'équation:

$$f(x) = \max(0, x) \quad \text{Éq. II.12}$$



**Fig. II.20** La fonction ReLU.

Les fonctions d'activation est certainement la plus utilisée dans les CNN profonds car elle permet une optimisation plus facile et permet de réduire les problèmes de disparition de gradient. En effet, la sigmoïde et la tangente hyperbolique ont pour inconvénient de renvoyer des gradients très petits lorsque la valeur absolue de l'entrée est grande. Le réseau a alors du mal à mettre à jour ses paramètres lors de la phase d'apprentissage. La fonction ReLU renvoie quant à elle, un gradient constant pour une entrée grande permettant ainsi d'apprendre plus rapidement (en particulier les réseaux d'une certaine profondeur). Il existe d'autres fonctions

d'activation de la même famille que les ReLU comme les LReLU [31], les PReLU [32] et les RELU [33].

#### **4.2.1.4 Couche entièrement connectée (FC)**

Après plusieurs couches de convolution et de maxpooling, 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. Leurs fonctions d'activations peuvent donc être calculées avec une multiplication matricielle suivie d'un décalage de polarisation.

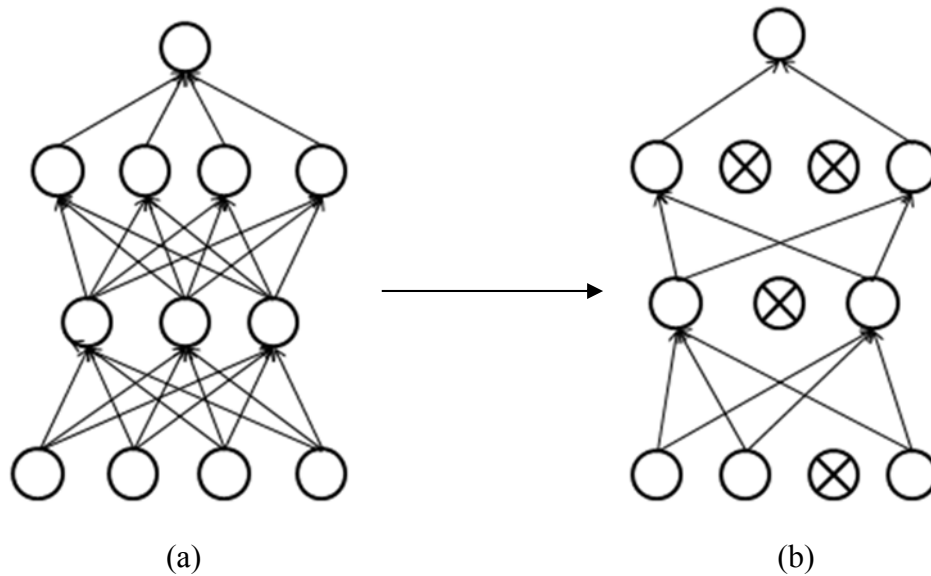
#### **4.2.1.5 Couche de perte (LOSS)**

La couche de perte spécifie comment l'entraînement du réseau pénalise l'écart entre le signal prévu et réel. Elle est la dernière couche dans le réseau. Diverses fonctions de perte adaptées à différentes tâches peuvent y être utilisées. La fonction « Softmax » permet de calculer la distribution de probabilités sur les classes de sortie.

#### **4.2.1.6 Méthode de régularisation**

- **Le dropout**

Pour éviter le sur-ajustement (overfitting) [34], Cette couche est utilisée pendant l'apprentissage. Elle permet de désactiver aléatoirement des neurones durant les différentes itérations de l'apprentissage. En d'autres termes, le dropout permet au réseau d'apprendre des sous-réseaux contenant moins de paramètres et donc moins sujets au sur-ajustement. Cette manière de faire permet d'apprendre des paramètres plus génériques qui ne se focalisent pas sur des détails de la base d'apprentissage. Une fois l'apprentissage terminé, tous les neurones sont réactivés.



**Fig. II.21** Dropout (a) : avant l'application dropout (b) : après l'application dropout.

## 5 Conclusion

L'intelligence automatique est un domaine très vaste. Dans ce chapitre, nous avons vu les concepts importants en apprentissage automatique; évidemment, il reste d'autres concepts et d'autres techniques d'apprentissage, nous avons essayé de faire un balayage sur tout ce qui concerne l'apprentissage automatique, nous avons mis l'accent sur ce qui est essentiel pour notre travail. Nous avons présenté une étude détaillée sur les réseaux de neurones convolutifs.

Dans le chapitre suivant, nous effectuons une étude de la classification et reconnaissance d'images.

***CHAPITRE III :***  
***RECONNAISSANCE, CLASSIFICATION***  
**D'IMAGES**

## ***1 Introduction***

La reconnaissance et la classification d'objets dans les images est un des problèmes les plus difficiles en vision par ordinateur. Elle est en même temps une étape primordiale pour la mise en œuvre de plusieurs applications actuelles qui nécessitent une interprétation de haut niveau d'images. Par conséquent, il existe un intérêt croissant sur ce domaine de recherche dans les dernières années et une vaste littérature.

Dans ce chapitre nous présentons les notions de base de la reconnaissance et la classification des images et finalement nous parlerons de la reconnaissance et la classification des panneaux de signalisation.

## ***2 Classification***

### ***2.1 Définition de classification***

La classification consiste à organiser un ensemble de données multidimensionnelles en un ensemble fini de classes selon un ou plusieurs critères de classification à l'aide d'un classifieur. La classification consiste à regrouper des pixels en catégories les plus homogènes possibles, c'est-à-dire créer des classes contenant le plus de pixels similaires.

Les méthodes de classification permettent de regrouper des objets en groupes ou classes d'objets plus homogènes. Les objets regroupés ont des caractéristiques communes, ils sont similaires mais se distinguent clairement des objets des autres classes, Nous avons déjà présenté les méthodes de classification dans le chapitre précédent.

### ***2.2 Les motivations de la classification d'images***

L'objectif de la classification d'images est d'élaborer un système capable d'affecter une classe automatiquement à une image. 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 ou le temps nécessité par un volume important de données images.

Les applications de la classification automatique d'images sont nombreuses et vont de l'analyse de documents à la médecine en passant par le domaine militaire. Ainsi on retrouve des applications dans le domaine médical comme la reconnaissance de cellules [74], dans le domaine du document comme la reconnaissance d'écriture manuscrite pour les chèques, les codes postaux [75], les cartes [76]; dans le domaine urbain comme la reconnaissance de panneaux de signalisation [77], la reconnaissance de piétons [78], la détection de véhicules

[79]. Le point commun à toutes ces applications est qu'elles nécessitent la mise en place d'une chaîne de traitement à partir des images disponibles composée de plusieurs étapes afin de fournir en sortie une décision. Chaque étape de la mise en place d'un tel système de classification nécessite la recherche de méthodes appropriées pour une performance globale optimale; à savoir la phase d'extraction de caractéristiques et la phase d'apprentissage. Typiquement, nous disposons de données images desquelles il nous faut extraire des informations pertinentes traduites sous formes de vecteurs numériques. Cette phase d'extraction nous permet de travailler dans un espace numérique. Il s'agit ensuite d'élaborer, dans la phase d'apprentissage, à partir de ces données initiales, une fonction de décision pour décider de l'appartenance d'une donnée nouvelle à l'une des classes en présence.

La phase d'extraction de caractéristiques peut être précédée d'une phase dite de prétraitement. Cette phase a pour but de nettoyer l'image. On retrouve ainsi des techniques d'atténuation de bruits, de renforcement de contours, des techniques d'amélioration de l'image comme le rehaussement de contraste, la réduction de la dimension de l'image par la binarisation [80], la réduction de l'image à ses primitives visuelles comme la squelettisation [81].

### ***2.3 Méthodes de classification***

Nous nous intéressons dans cette section aux méthodes d'apprentissage. Le but de cette étape est de construire une fonction de décision à partir de données d'apprentissage projetées dans cet espace de description. On distingue principalement deux tendances dans les méthodes d'apprentissage : l'approche à classifieur unique et l'approche avec une combinaison de classifieurs agrégeant les réponses de plusieurs classifieurs [69].

#### ***2.3.1 Les classifieurs***

Dans cette section Nous présentons l'algorithme classique des K-Plus Proches Voisins (KPPV), l'algorithme standard d'arbre de décision « Classification And Regression Trees » (CART) proposée par [70], « l'algorithme "Support Vector Machine" ou « Séparateur à Vaste Marge » (SVM) [71], et nous avons déjà présenté les réseaux de neurones dans le chapitre précédent [72].

##### ***2.3.1.1 K-Plus Proches Voisins***

L'algorithme des K-Plus Proches Voisins [73] est une approche basée sur l'estimation de distance. Il n'a pas besoin d'apprentissage car le modèle de classification est directement le set d'apprentissage.

### 2.3.1.2 *Machine à vecteurs de support (Support Vector Machine)*

Le "Support Vector Machine" ou "Séparateur à Vaste Marge" (SVM) [71] est un classifieur discriminant paramétrique établissant un hyperplan séparateur de marge maximale entre les exemples représentants de chacune des classes des données d'apprentissage. Le Support Vector Machine est des techniques d'apprentissage supervisées destinées à résoudre des problèmes de classification.

## 2.4 *L'objectif de la classification*

- Obtenir une représentation schématique simple d'un tableau de données complexe à partir d'une typologie (segmentation).
- On utilise la classification pour mieux agir, face à un ensemble d'états qu'ils ont besoin d'un traitement (par exemple un ensemble des cas d'intervention), la classification nous aide pour donner le traitement qui convient à chaque cas.
- Simplifier la recherche: si on a un grand ensemble d'objets, la recherche d'un objet peut être une tâche coûteuse (temps, matériel, personnel...), on peut réduire la taille de l'espace de recherche si on commence par une classification des objets.
- La classification d'images est d'élaborer un système capable d'affecter une classe automatiquement à une image. 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 ou le temps nécessaire pour un volume important de données dans les images.

## 2.5 *Domaines d'application de la classification*

La classification a un rôle à jouer dans toutes les sciences et les techniques qui font appel à la statistique multidimensionnelle.

Le domaine médical : les classifications des maladies, de traitements curatifs de ces maladies ou de symptômes pour ces maladies peuvent mener à des typologies très utiles.

Le domaine du marketing : pour le lancement d'un nouveau produit, le responsable du service de marketing d'une entreprise pourra chercher à constituer des groupes de villes semblables vis-à-vis d'un certain nombre de critères, et choisir dans chaque groupe une ville type utilisée comme marché-test.

Le domaine politique : un candidat aux élections municipales fixera sa stratégie électorale en fonction de différents types d'électeurs construits à partir de différentes caractéristiques.

En psychiatrie : un diagnostic correct de groupes de symptômes comme paranoïa, la schizophrénie, etc. est indispensable à la réussite de la thérapie.

En archéologie: les chercheurs ont tenté d'établir des typologies d'outils en pierre, objets funéraire, etc. En utilisant des techniques de classification.

## 2.6 Indicateurs de performance en classification

### 2.6.1 Courbe ROC (Received Operating Characteristic):

La courbe ROC (Receiver Operating Characteristics) [55] offre à la fois une vision graphique et une mesure pertinentes de la performance d'un classifieur. Elle possède de nombreux avantages par rapport aux mesures de rappel et précision par classe : la performance est synthétisée par une unique mesure qui ne dépend pas des proportions de classe. Cet avantage se transforme néanmoins en inconvénient lorsqu'il s'agit de revenir rapidement aux mesures de rappel et de précision par classe ou d'estimer le comportement du classifieur selon les classes. La courbe ROC est une caractéristique de la performance indépendante de la classe, cependant elle représente le point de vue de la classe positive, généralement la classe majoritaire. Des mesures concernant la classe minoritaire sont donc difficiles à obtenir graphiquement

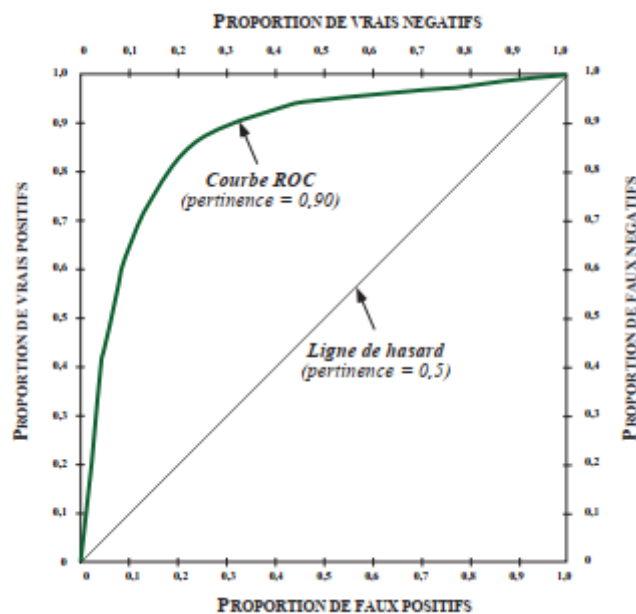


Fig. III.1 Courbe ROC.

### 3 *La reconnaissance*

#### 3.1 *La reconnaissance de formes*

Parallèlement à l'apprentissage artificiel, la reconnaissance de formes s'est fortement développée au cours des vingt dernières années, en raison notamment des progrès de l'informatique [56]. L'une des vocations de la reconnaissance de formes est le traitement de données. Elle se fonde sur des connaissances a priori du problème ou sur de l'information tirée d'un ensemble de mesures et d'observations.

La reconnaissance de formes souhaite au final obtenir un outil de régression, de classification ou de discrimination. La nature de cet outil dépend avant tout de la forme du résultat à obtenir. En effet, dans le cas où la sortie est inconnue, l'objectif est de regrouper des sous-ensembles de données en classes, ou plus précisément en clusters : on parle alors de classification, ou de classification automatique. Lorsque la sortie est connue, le processus permet d'obtenir un outil de régression si la sortie est quantitative (numérique), ou un outil de discrimination si la sortie est qualitative. Notons que dans la littérature, le terme de « discrimination » est parfois remplacé par celui de « classification », et dans ce cas, celui de « classification » par « classification automatique ».

#### 3.2 *Les principaux domaines d'application*

La reconnaissance d'objets est une étape indispensable pour de nombreuses applications spécifiques. On peut citer, parmi d'autres, quelques applications exemplaires pour illustrer les enjeux des travaux de recherche sur ce problème [57]:

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

### 3.3 La reconnaissance spécifique versus générique

On peut faire la distinction entre deux problèmes de reconnaissance différents : la reconnaissance spécifique on cherche à détecter un objet précis qui a éventuellement subi des déformations perspectives ou de la transformation d'illumination. Les exemples sont la reconnaissance d'un jouet d'enfant [58], la reconnaissance d'un produit spécifique d'une marque [59], la reconnaissance des panneaux de signalisation ou l'identification d'une personne.

### 3.4 La difficulté majeure de la reconnaissance d'objets

La difficulté majeure du problème de la reconnaissance d'objets repose sur les variations d'un objet ou des objets d'une même catégorie sur des images différentes. On peut grouper les sources de variations dans sept catégories principales [57]:

- **Les variations de point de vue** : les objets réels sont des entiers physiques tridimensionnels. Leur apparence change de façon significative en fonction du point de vue.
- **Les variations d'illumination** : en fonction de changements de l'environnement et des conditions d'acquisition d'une image, un objet peut subir des variations d'illumination importantes. Les objets dans les images des scènes naturelles sont très sensibles aux variations d'illumination qui modifient les couleurs et les valeurs radiométriques des objets considérablement.
- **Les occlusions** : dans une image les objets peuvent être partiellement masqués par d'autres objets ou par leurs propres composants.
- **Les variations d'échelle** : les objets d'une même catégorie peuvent avoir des tailles physiques différentes ou les images d'un même objet peuvent avoir des tailles différentes en fonction de la résolution de l'image.
- **Les déformations** : la plupart des objets réels n'ont pas une forme rigide. Particulièrement les objets articulés subissent généralement des déformations linéaires et non-linéaires de leurs composants.
- **Complexité du fond** : en général, dans une image les objets ne sont pas isolés de leur environnement. Sur un fond complexe il est difficile de distinguer les régions appartenant à l'objet et au fond.
- **Variations intra-classe** : c'est le plus important des problèmes envisagés pour la reconnaissance générique d'objets. Les objets d'une même classe peuvent avoir des apparences très variables, même quand on se limite aux catégories à un niveau

élémentaire. Dans le cas extrême, les catégorise fonctionnelles comme la classe « chaise » est définie par un haut niveau d'abstraction et les indices visuels seuls ne sont pas suffisants pour les reconnaître.

### 3.5 Les méthodes de reconnaissance d'objets

Les traitements de bas niveau et les traitements de haut niveau qui correspondent aux étapes suivantes [57] :

- **Extraction** : l'extraction d'un ensemble de descripteurs à partir de l'image, qui reflètent les informations pertinentes pour la reconnaissance de l'objet .
- **Représentation** : la construction d'un modèle d'objet pour représenter une catégorie d'objet.
- **Apprentissage** : l'apprentissage des paramètres du modèle à partir des images d'apprentissage étiquetées.
- **Reconnaissance** : l'appariement avec le modèle ou la classification en utilisant le modèle pour reconnaître l'objet cible dans une image de test.

Les méthodes de reconnaissance d'objets se différencient principalement au niveau de:

- **La définition du problème de reconnaissance.**
- **Le nombre d'images utilisées pour l'apprentissage** : une extrémité étant l'apprentissage à partir d'une seule image, on peut aussi utiliser un très grand nombre d'images pour l'apprentissage.
- **La modélisation** : nous utilisons le terme modèle dans le sens général de "la représentation d'un objet". Un très grand nombre de modèles d'objets différents ont été proposés pour représenter les objets dans les images.

### 3.6 Méthodes de reconnaissance dans le traitement d'images

La reconnaissance d'images est le processus d'identification et de détection d'un objet ou d'une caractéristique dans une vidéo ou une image numérique. Ce concept est utilisé dans de nombreuses applications. Les algorithmes standards de reconnaissance d'images incluent :

- La reconnaissance optique de caractères.
- La correspondance de formes et de gradients.
- La reconnaissance faciale.
- L'identification des plaques d'immatriculation.
- La détection des changements de scène.

### *3.7 Les traitements de bas niveau: extraction des descripteurs*

L'utilisation directe de l'image pour la reconnaissance n'est pas envisageable par ce qu'elle contient une grande quantité d'information impertinente ou redondante. C'est pourquoi, la première étape de presque toutes les méthodes de reconnaissance est d'extraire un ensemble de descripteurs qui représentent les propriétés discriminantes de l'objet cible. Les descripteurs sont un ensemble de nombres scalaires générés pour décrire un objet (ou une forme). Un objet ne peut pas être reconstruit en général à partir des descripteurs, mais les descripteurs pour des objets distincts doivent avoir des valeurs suffisamment différentes pour les discriminer, et les descripteurs pour les objets similaires doivent avoir des valeurs proches. On peut extraire un nombre infini de descripteurs à partir d'une image, pourtant l'objectif est d'en sélectionner un nombre moins important que le nombre de pixels dans l'image. Pour une reconnaissance robuste il est souhaitable que les descripteurs soient invariants aux changements d'échelle et d'illumination et aux transformations affines.

### *3.8 Les traitements de haut niveau: reconnaissance d'objets*

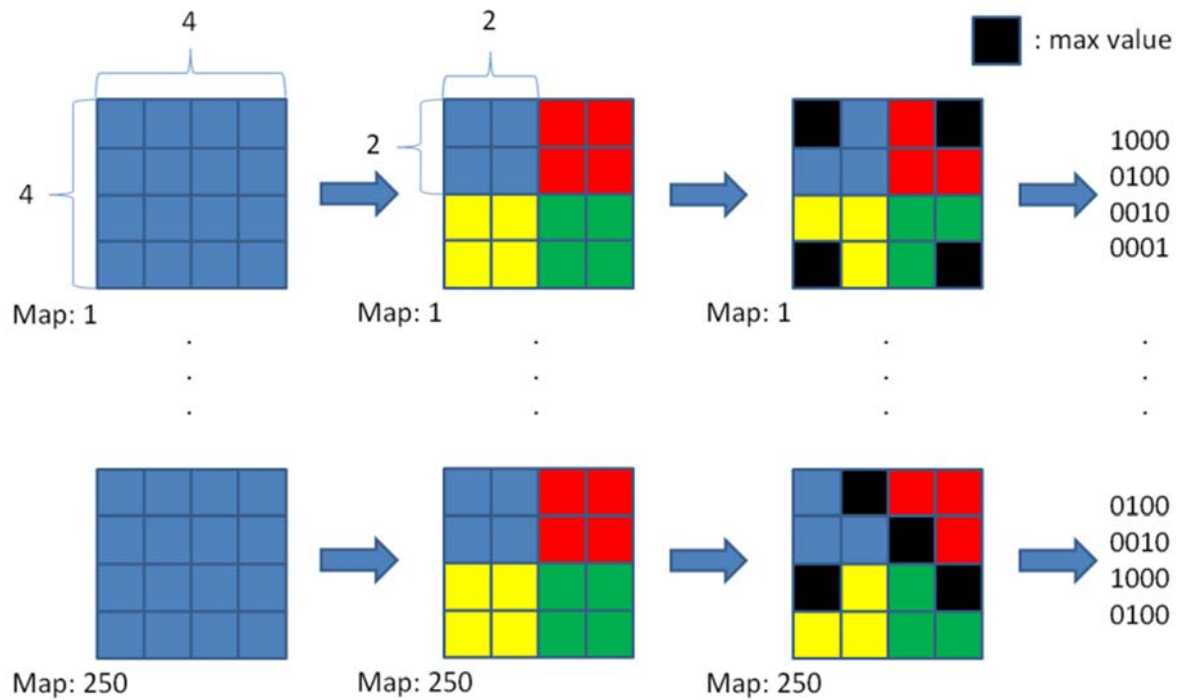
Au début des travaux de recherche en reconnaissance d'objets dans les années 1970 les représentations géométriques ont joué un rôle important. On a souvent ciblé la reconnaissance d'objets rigides, et la géométrie, extraite à partir des contours de l'image, a fourni une base théorique solide pour gérer les variations d'illumination et de perspective de ces objets. L'approche centrée sur l'objet de Marr [35] au début des années 1980 a eu une influence majeure sur les travaux de recherche pendant deux décennies. Selon Marr la reconnaissance passe par la reconstruction de la structure tridimensionnelle de l'objet physique à partir des différentes vues bidimensionnelles de l'objet. Cette approche a donné lieu à des méthodes fondées sur l'identification de la structure de l'objet par la correspondance entre les régions et les contours extraits de l'image et un modèle hypothétique d'objet tridimensionnel.

## *4 La reconnaissance, la classification des panneaux de signalisation*

Les réseaux convolutifs ont été introduits pour la première fois par Fukushima [60], il a dérivé une architecture de réseau nerveux hiérarchique inspirée par le travail de recherche de Hubel [61]. Lecun [62] a proposé les architectures généralisées pour classer les chiffres avec succès et pour reconnaître les numéros de contrôle manuscrit par LeNet-5.

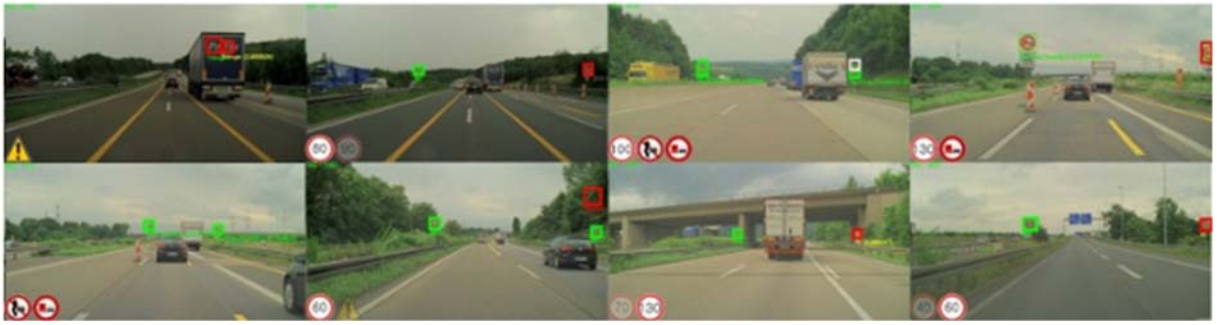
Ciresan [63] a utilisé les réseaux convolutifs et il a réalisé les meilleures performances dans la littérature pour la reconnaissance d'objets multiples des bases de données d'images

multiples: MNIST, NORB, HWDB1.0, CIFAR10 et l'ensemble de données ImageNet. Grâce à l'application du CNN des nouvelles méthodes ont été développées par Qian, R. et al [64], Pour effectuer une reconnaissance MPPS (max pooling position) comme une fonctionnalité pour prédire les classifications des classes, MPPS explique les propriétés entre les couches basée sur la base de donnée (GTSRB), Et améliore la classification et la vitesse, cela a conduit à une augmentation de la précision.



**Fig. III.2** La classification MPPS (max pooling position).

Jorge Enrique Zafra et al [65] présente deux algorithmes de réseaux de neurones pour la reconnaissance des panneaux de signalisation en utilisant Backpropagation et réseaux de neurones convolutifs. Dans les temps d'entraînement le réseau Backpropagation est beaucoup plus rapide par rapport aux réseaux de neurones convolutifs mais la précision est très élevée qui est égale à 80% et la précision aléatoires de Backpropagation qui est entre 50 % et 98%. Lim K, et al [66] présente une méthode de reconnaissance des panneaux de signalisation en temps réel basée sur une unité de traitement graphique appliquée sur une base de données d'Allemagne et Corée du Sud, Ils ont proposé une méthode puissante contre les changements d'illumination, et ils ont effectué par la suite la reconnaissance de région en utilisant un modèle hiérarchique. Cette méthode a produit des résultats stables dans des environnements à faible illumination. La hiérarchie du modèle a été effectuée en temps réel, la méthode proposée a obtenu un score de 0,97 de précision.

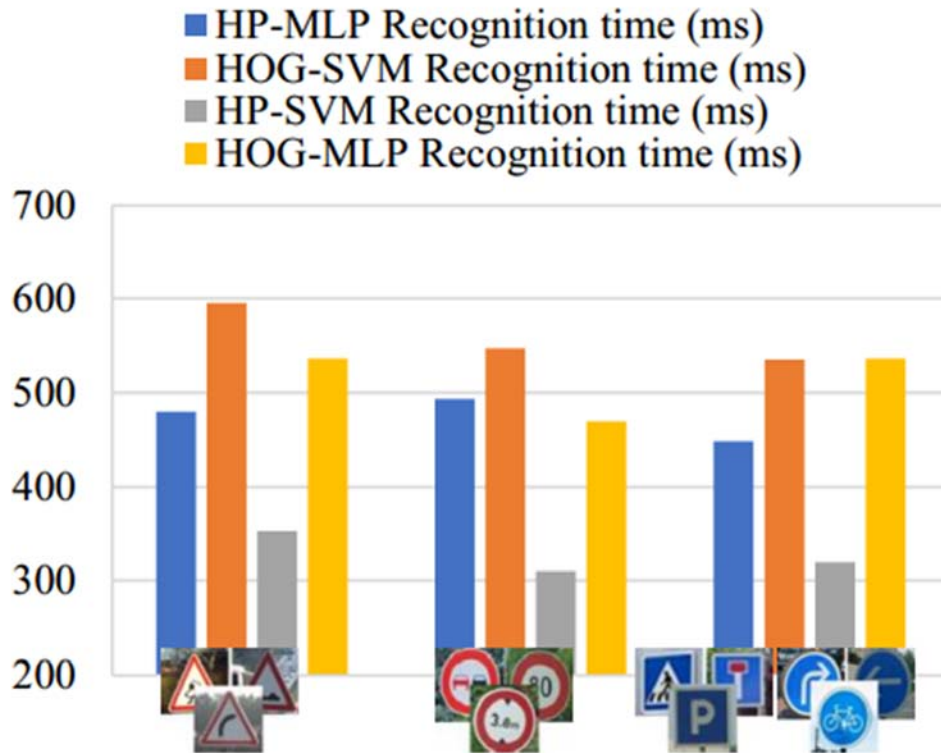


**Fig. III.3** Résultat de la méthode proposée pour l'ensemble de test Allemagne.



**Fig. III.4** Résultats de la méthode proposée pour l'ensemble de test Corée.

Une étude a été présentée par Qian R et al [67] pour comparer quatre méthodes, Ces méthodes sont divisées en deux catégories ; la première catégorie est constituée sous forme de deux descripteurs HP et HOG, la deuxième catégorie est constituée sous forme de deux classificateurs MLP et SVM. L'étude a conclu que la méthode HP-SVM offre des performances compétitives en termes de la précision et du temps de traitement de la reconnaissance des panneaux de signalisation.



**Fig. III.5** Le temps moyen de reconnaissance de HP-MLP, HP-SVM, HOG-MLP et HOG-SVM.

Un système de reconnaissance a été proposé par Reinders C et al [68] où les avantages des réseaux de neurones convolutifs et les forêts aléatoires ont été combinés pour construire un réseau entièrement convolutifs pour la prédiction des boîtes englobantes.



**Fig. III.6** Reconnaissance.

## **5 Conclusion**

Dans ce chapitre, nous avons présenté les notions de base de la reconnaissance et la classification des images, Aussi quelques concepts de la reconnaissance et la classification et nous avons parlé de la reconnaissance et la classification des panneaux de signalisation.

Dans notre projet, nous avons utilisé les réseaux des neurones pour classifier les images. Nous détaillerons notre model dans le chapitre suivant.

***CHAPITRE IV:***  
***IMPLEMENTATION***

## ***1 Introduction***

Dans ce chapitre, on va définir l'architecture du modèle qu'on a créé, ensuite on va appliquer ce modèle sur un jeu de donnée qui s'appelle GTRSB. Pour cela, nous avons utilisé le langage Python ainsi qu'un ensemble de bibliothèque insérés pour rendre notre travail plus souple.

## ***2 La configuration du matériel utilise***

Afin de mener à bien ce projet, nous avons utilisé un ensemble de matériels dont les principales caractéristiques sont les suivantes :

- Processeur: Intel(R) Core (TM) i3-3120M CPU @ 2.50 GHz x 2.
- RAM: 4 GO.
- Système d'exploitation : Windows 7 64bits.

## ***3 L'environnement de travail***

### ***3.1 Python***

Python est un langage de programmation de haut niveau interprété (il n'y a pas d'étape de compilation) et orienté objet avec une sémantique dynamique. Il est très sollicité par une large communauté de développeurs et de programmeurs. Python est un langage simple, facile à apprendre et permet une bonne réduction du cout de la maintenance des codes. Les bibliothèques (packages) python encouragent la modularité et la réutilisabilité des codes. Python et ses bibliothèques sont disponibles (en source ou en binaires) sans charges pour la majorité des plateformes et peuvent être redistribués gratuitement [82].

### ***3.2 Notebook jupyter***

est une interface web permettant d'exécuter du code Python (ou autres langages) de façon graphique et interactive.

### 3.3 Scikit-Learn

Scikit-Learn est une bibliothèque libre Python dédiée à l'apprentissage automatique [83]. Elle est développée par de nombreux contributeurs notamment dans le monde académique par des instituts français d'enseignement supérieur et de recherche comme Inria et Télécom ParisTech

Elle comprend notamment divers utilitaires :

- Des classifieurs tels que les forêts aléatoires, des régressions logistiques, les machines à vecteurs de support, des algorithmes de Boosting...
- Des méthodes non supervisées telles que les KMeans, Mean Shift, ...
- Des fonctions de prétraitements de données telles que la normalisation, la discrétisation, un one hot encodé, et bien d'autres
- Elle est conçue pour s'intégrer avec les autres bibliothèques libres Python, notamment NumPy et SciPy.

### 3.4 TensorFlow

TensorFlow [84] est en simplifiant le pendant de Scikit-Learn pour l'apprentissage profond : cette librairie a été développée afin de simplifier la programmation et l'utilisation de réseaux profonds. La librairie est majoritairement maintenue par Google. Elle propose une interface simple et concise pour exprimer des algorithmes d'apprentissage et une implémentation efficace pour l'exécution de tels algorithmes. Un calcul exprimé à l'aide de TensorFlow peut être effectué avec peu ou pas de changement sur une grande variété de systèmes hétérogènes, allant des appareils mobiles tels que les téléphones et les tablettes aux systèmes distribués à grande échelle de plusieurs machines et des composants informatiques tels que les cartes GPU. Le système est flexible et est utilisé pour exprimer une grande variété d'algorithmes, y compris pour concevoir des modèles de réseaux de neurones profonds. Elle a été utilisée pour la recherche et le déploiement de l'apprentissage automatique dans une grande variété de domaines, contenant la reconnaissance vocale, la vision par ordinateur, la robotique, la récupération d'informations, le traitement du langage naturel, ...

4 Notre modèle

Les réseaux de neurones convolutifs repose sur une architecture à base de couches, celles-ci vont subir une succession de différents traitement afin d’extraire les caractéristiques les plus importantes de l’image. Au fur et à mesure que l’on traverse le réseau; les couches seront compresser, chaque couche de convolution est suivi d’une couche d’activation(RELU) ensuite ces caractéristiques seront transmises à un réseau de neurone intégralement connecté afin d’effectuer une phase de reconnaissance ou de classification d’images.

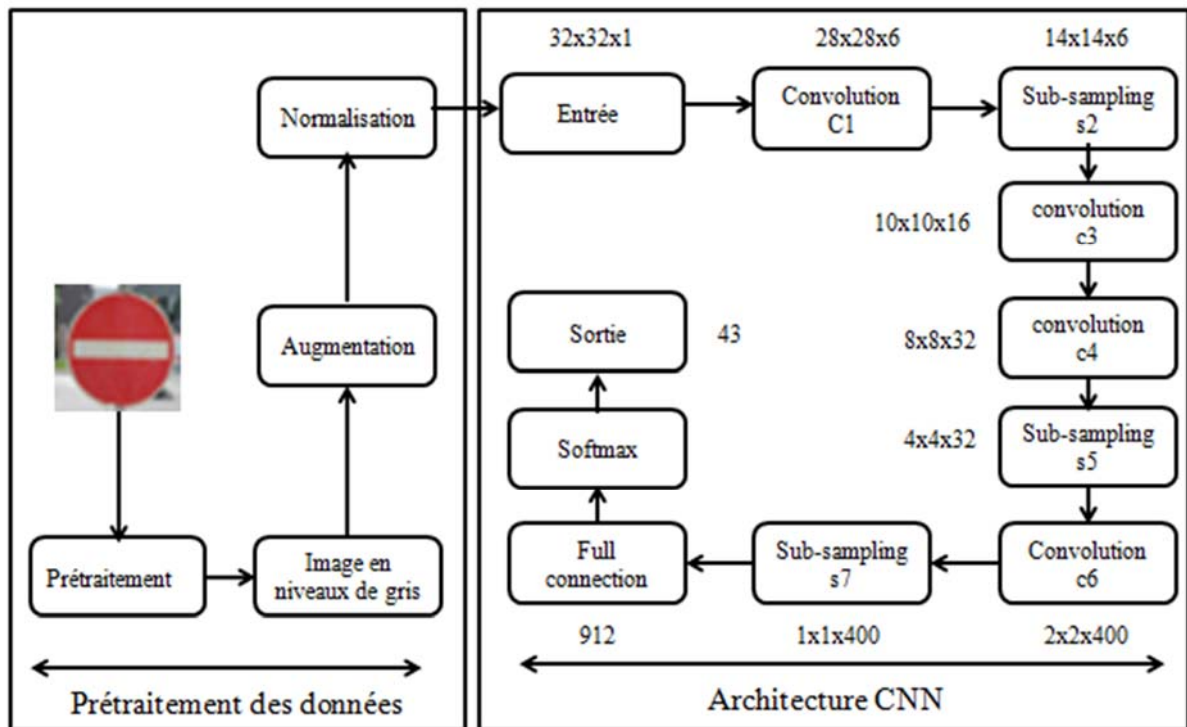


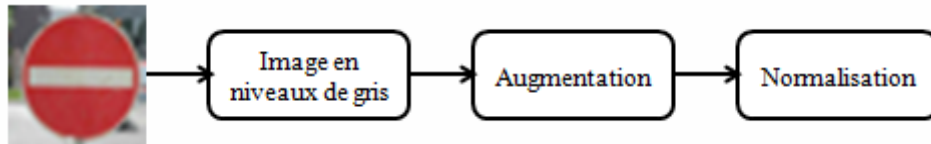
Fig. IV.1 L’architecture globale de notre modèle.

4.1 Prétraitements des données

Cette étape est primordiale pour améliorer les performances de notre modèle.

La Fig. IV.2 montre les prétraitements appliqués sur nos images.

L’étape de prétraitement d’images est primordiale elle est effectuées avant l’étape d’apprentissage. Dans notre modèle nous avons tous d’abord appliqué trois opérations de prétraitement.



**Fig. IV.2** L'étape de prétraitements des données.

#### 4.1.1 *Image en niveaux de gris*

Une image est représentée sous la forme d'une matrice de pixels. L'image numérique en niveaux de gris se représente sous la forme d'un tableau à deux dimensions, tandis qu'une image couleur est représentée sous la forme d'un tableau à trois dimensions. On transforme une image couleur RVB en niveaux de gris pour élever sa précision et minimiser son temps d'exécution.

Dans notre modèle on va utiliser une image RVB (32, 32,3) par la suite on va la transformer en une image en niveaux de gris (32, 32,1).



**Fig. IV.3** Conversion d'une image RVB en niveaux de gris.

#### 4.1.1 *Normalisation*

La normalisation consiste à mettre le jeu d'apprentissage en même échelle.

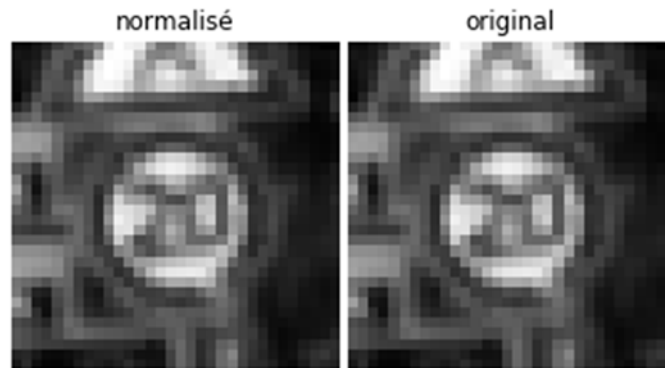
La méthode utilisée pour la normalisation de données est basée sur la formule suivante :

$$X = \frac{\text{Pixels} - 128}{128} \quad \text{Éq. IV. 1}$$

- X : les variantes normalisées.

- 128 : est une constante pour normaliser l'intensité des pixels pour les images dont la valeur de l'intensité des pixels est fixée dans le domaine [0.255].

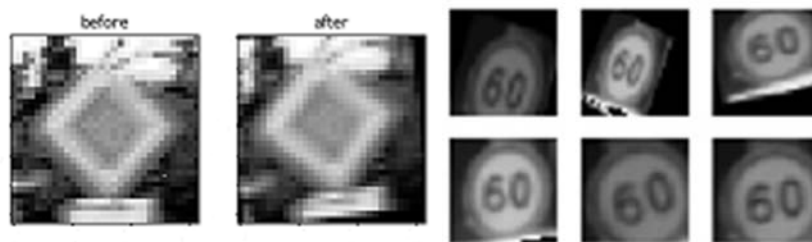
La Fig. IV.4 montre quelques images générées avec normalisation.



**Fig. IV.4** Normalisation des images.

#### 4.1.2 *Augmentation*

Consiste à faire augmenter le nombre d'images contenues dans le jeu d'apprentissage en effectuant un ensemble d'opérations qui sont : translation, échelle, rotation et contraste, tous cela permet à notre modèle de se généraliser et donner de bonnes prédictions vis-à-vis aux problèmes dus aux opérations appliquées sur nos images (les panneaux de signalisation), la Fig. IV.5 montre quelques images générées avec augmentation.



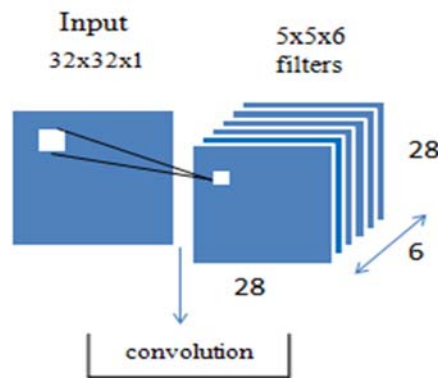
**Fig. IV.5** Augmentation des données.

Nous avons utilisé le graphique de distribution dans notre modèle pour montrer le nombre d'images contenues dans notre jeu d'entraînement avant et après l'augmentation afin de vérifier que notre modèle peut se généraliser et donner de bonnes prédictions. La Fig. IV.7 montre la représentation graphique de distribution du jeu d'entraînement avant et après l'étape de prétraitement.



**La couche de convolution C1 :** après le prétraitement de l'image, la première couche de convolution C1 est paramétrée par la taille, le nombre de cartes, les tailles de noyau et la table de connexion,

- la couche de convolution adopte 6 noyaux de convolution (kernels).
- la taille de chaque noyau de convolution est 5x5.
- il peut produire six cartes de caractéristiques (feature map), chaque carte de caractéristiques contient 784 neurones c'est-à-dire  $(32-5+1)*(32-5+1)=28*28$ .
- Au final un total est égale à 156 de paramètres à entraîner c'est-à-dire  $6*(5*5+1)$ .
- Une fois l'étape de convolutions est effectuée, une fonction d'activation RELU doit être appliquée.



**Fig. IV.8** Calcul de la taille de volume de sortie dans une couche de convolution.

**La couche de Max Pooling S2 :** Une couche de Max Pooling est utilisée pour réduire la taille des images et pour obtenir une convergence plus rapide pendant l'entraînement,

- elle contient six cartes de caractéristiques (feature map).
- chaque carte de caractéristiques contient 196 neurones c'est-à-dire  $(14*14)$ .
- La fenêtre de maxpooling est une matrice de 2\* 2 dimension
- la taille de pas de maxpooling est 1, donc la couche S2 contient 5880 connexions c'est-à-dire  $6* 196*(2 *2+1)$ .

- Chaque carte de caractéristiques dans la couche S2 contient un poids et un biais, donc un total de 12 paramètres peut être formé dans la couche S2. Cela nous permet de réduire le temps de calculs et l'espace mémoire et le risque de sur-ajustement.

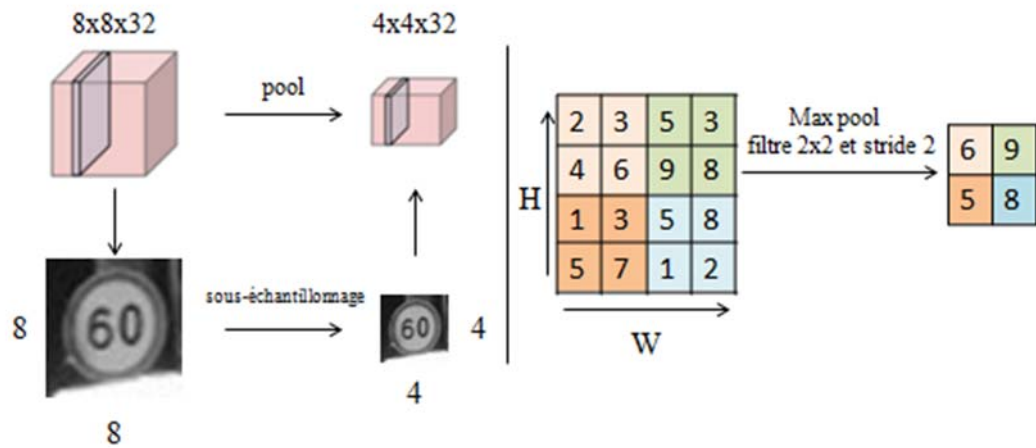


Fig. IV.9 Max pooling.

Les couches de convolution C3, C4, C6 et les couches de maxpooling S5, S7 sont exécutées de la même façon.

- **La couche Fully Connected** est exécutée après les couches mentionnées précédemment.
- La particularité de cette couche est que ses neurones ne sont pas connectés seulement avec les neurones de la couche précédente mais aussi avec tous les neurones de la couche suivante.
- **la fonction softmax** normalise les résultats, elle est connectée avec la couche fully connected.
- Cette fonction est utilisée pour produire une distribution de probabilité entre les différentes classes.

Le tableau ci-dessous résume les étapes de construction du réseau convolutifs

Table 1 : Description de CNN utilisé.

layer	Description	Input	Output
1	Conv (5x5,1x1 stride, valid padding) RELU activation	32x32x1	28x28x6
2	max_Pool (2x2, stride 2)	28x28x6	14x14x6
3	Conv (5x5,1x1 stride, valid padding) RELU activation	14x14x6	10x10x16
4	Conv (3x3,1x1 stride, valid padding) RELU activation	10x10x16	8x8x32
5	Max_Pool(2x2, stride 2)	8x8x32	4x4x32
6	Conv(3x3 ,1x1 stride, valid padding) RELU activation	4x4x32	2x2x400
7	Max_Pool(2x2, stride 2)	2x2x400	1x1x400
8	Flatten	4x4x32	512
9	Concat	512+400	912
10	Fully_conn	912	43

La Fig. IV.10 montre l'architecture CNN utilisé :

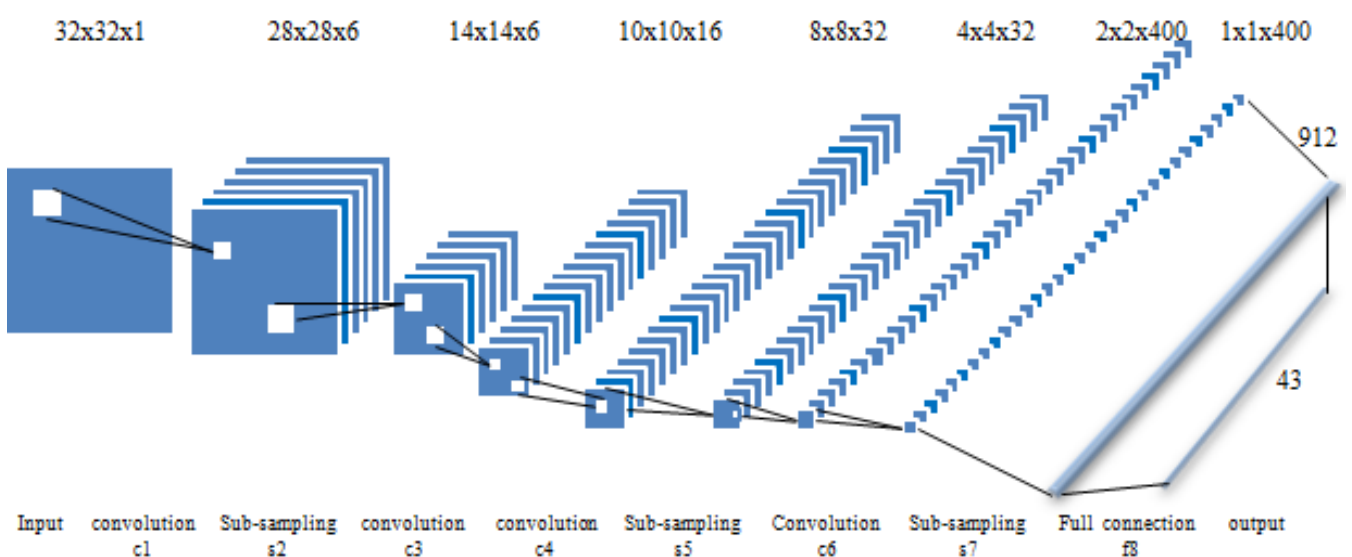


Fig. IV.10 Architecture CNN.

***CHAPITRE V :***

***RESULTAT***

## 1 Introduction

Une fois notre modèle construit, il est temps d'évaluer les résultats obtenus et de mesurer les performances de notre modèle prédictif. C'est à ce moment que le jeu de test est crucial. Les prédictions sont faites sur ce dernier, en faisant comme s'il s'agissait de nouvelles données dont on ne connaît pas la classe, pour ensuite comparé les prédictions avec les classes réelles. Si nous faisons cela sur le jeu d'entraînement, les résultats seraient alors biaisés car l'évaluation serait faite sur des données qui ont servi à l'apprentissage. Le fait de garder un jeu de test pour lequel les données n'ont pas servi à l'apprentissage permet de mesurer la capacité de généralisation du modèle créé.

## 2 Le jeu de données

Le jeu de données a été créé à partir d'environ 10 heures de vidéo enregistrées pendant la conduite sur différents types de routes en Allemagne. Les séquences ont été enregistrées en mars, octobre et novembre 2010. Pour la collecte des données, L'ensemble contient des images de plus de 1700 instances de panneaux de signalisation. La taille des panneaux de signalisation varie entre 15×15 et 222×193 pixels, Cet ensemble de données sont réparties dans 43 différents types des panneaux de signalisation ; il contient un total de 39 209 d'images consacrées pour l'entraînement, un total 12 630 d'images destinées pour les tests et finalement 9902 images utilisées pour la validation. Chaque image a une taille de 32 × 32 pixel.

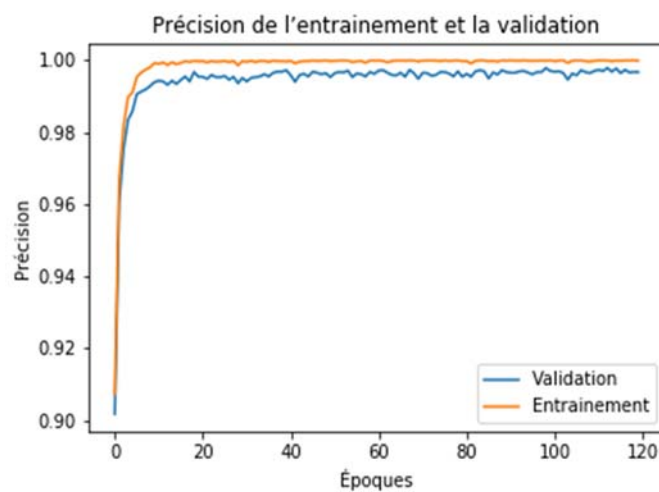


Fig. V.1 Ensemble des images du jeu de données.

### 3 Evaluation

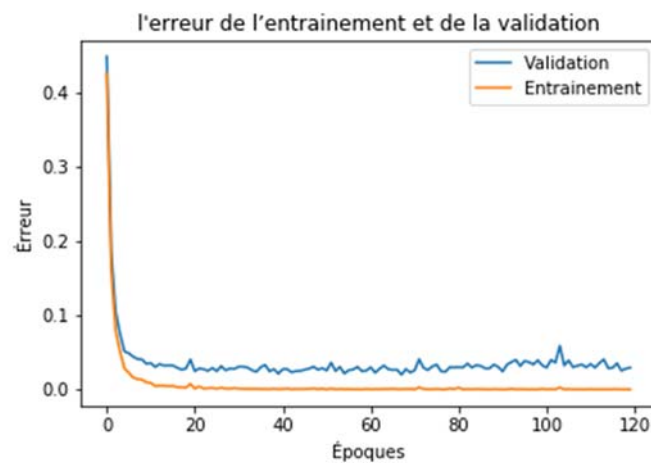
Dans cette phase, nous allons présenter les résultats de la précision, de l'erreur, de testes et de prédictions.

D'après la Fig. V.2, les courbes de précision de l'entraînement et de la validation augmentent en fonction du nombre d'époques, notre modèle offre de meilleurs résultats en terme de précision de l'entraînement (résultats est égale à 100%) et en terme de précision de la validation (résultats est égale à 99,6%).



**Fig. V.2** La précision de l'entraînement et de la validation.

Dans la Fig. V.3, les courbes d'erreur d'entraînement et de la validation diminuent d'une valeur de 120 d'époque.



**Fig. V.3** L'erreur de l'entraînement et de la validation.

Nous avons évalué notre modèle en utilisant des panneaux de signalisation, La Fig. V.4 ci-dessous représente quelques résultats de test générés par notre modèle.

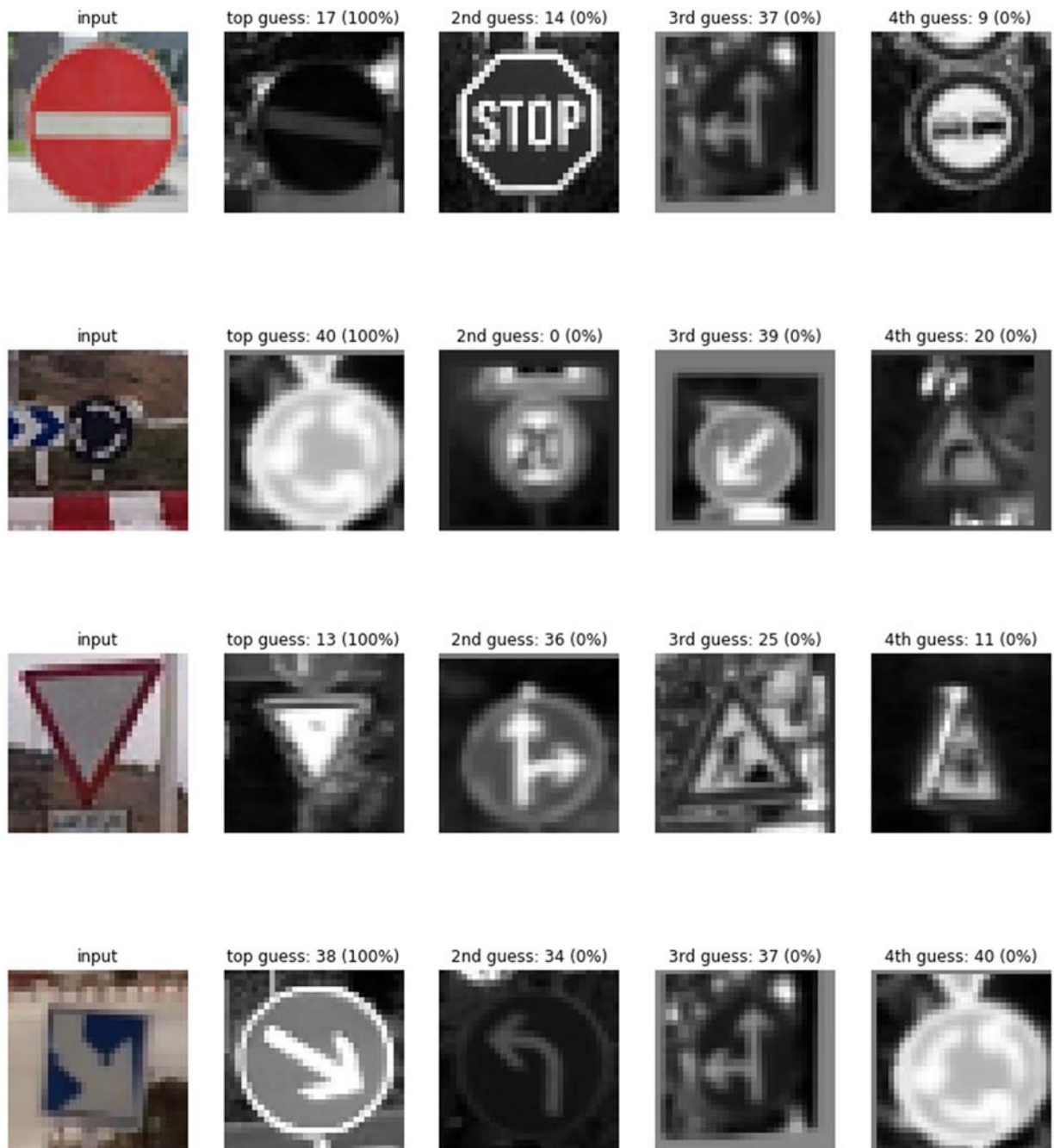


Fig. V.4 Résultats de tests sur des nouvelles images.

La Fig. V.5 montre les prédictions sur des nouvelles images, nous utilisons les probabilités softmax du modèle pour La certitude d'une prédiction, Les résultats de prédictions du notre modèle sont très satisfaisants.

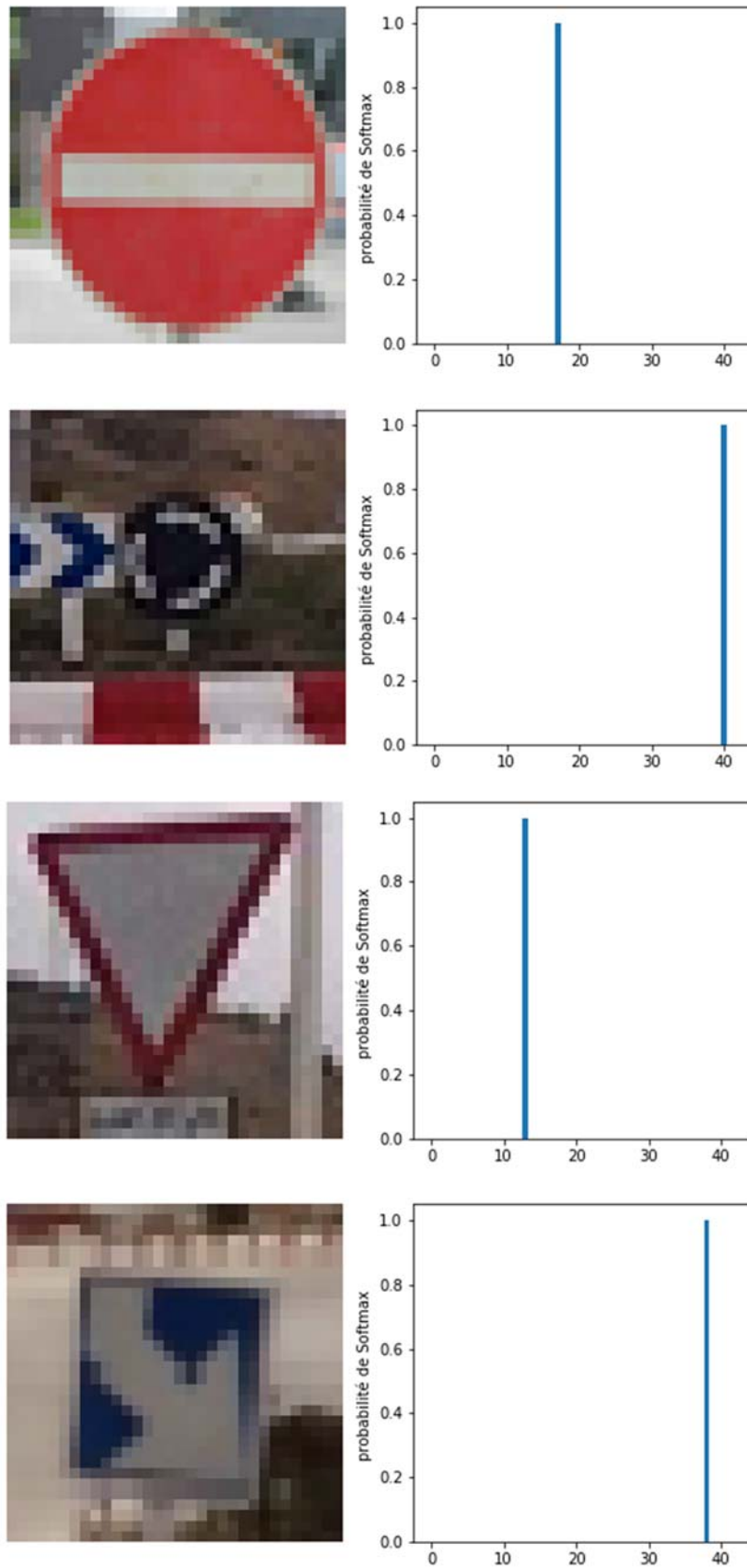


Fig. V.5 Les résultats de prédictions.

#### *4 Conclusion*

Nous avons présenté dans ce chapitre une approche de classification et reconnaissance des panneaux de signalisation routière basée sur les réseaux de neurones convolutifs, pour cela on a utilisé une architecture CNN et on a montré les différents résultats obtenus en matière de précision et d'erreur. La comparaison des résultats trouvés a montré que le nombre d'époques, la taille des jeux de données et la profondeur de réseaux, sont des facteurs importants pour l'obtention de meilleurs résultats.

## *Conclusion et perspectives*

Dans le domaine de vision par ordinateur en particulier la conduite autonome, la reconnaissance et la classification d'objets est un problème crucial. Dans ce projet nous avons tendance à relever le défi de reconnaissance et la classification des panneaux de signalisation en proposant un modèle à base de réseaux de neurones convolutifs.

Dans ce projet nous avons présenté des notions générales sur le traitement d'images, aussi nous avons discuté des notions fondamentales d'apprentissage automatique en générale et des réseaux de neurones convolutifs en particulier. Nous avons introduit ces réseaux de neurones convolutifs en présentant les différents types de couches utilisées dans la classification.

Nous constatons que cette approche améliore fortement le taux de précision et diminue le taux d'erreur, Cependant l'utilisation des techniques de régularisation, de normalisation, et d'optimisation, le temps d'entraînement de notre modèle reste un problème à soulever. Comme extension futur de notre modèle nous envisagerons l'utilisation des réseaux de neurones pré-entraîné cela va nous permettre d'augmenter les performances de notre modèle, Ainsi nous envisagerons tirer profit de l'immense puissance de calcul des processeurs graphique GPU en distribuant les calculs sur plusieurs GPU.

## *Bibliographie*

- [1] Cornuéjols, A., & Miclet, L. (2002). Apprentissage artificiel, concepts et algorithms, Eyrolles. ISBN 2-212-11020-0.
- [2] Russell, S. (2009). Artificial intelligence: A modern approach author: Stuart russell, peter norvig, publisher: Prentice hall pa.
- [3] Vincent, P. (2003). Modeles a noyaux a structure locale (French text).
- [4] Simon, H. A. (1983). Why should machines learn?. In Machine Learning, Volume I (pp. 25-37).
- [5] Zighed, D. A., & Rakotomalala, R. (2000). Graphes d'induction: apprentissage et data mining. Paris: Hermes.
- [6] Huang, G., Song, S., Gupta, J. N., & Wu, C. (2014). Semi-supervised and unsupervised extreme learning machines. IEEE transactions on cybernetics, 44(12), 2405-2417.
- [7] Zhu, X., & Goldberg, A. B. (2009). Introduction to semi-supervised learning. Synthesis lectures on artificial intelligence and machine learning, 3(1), 1-130.
- [8] Chen, D., & Stow, D. (2002). The effect of training strategies on supervised classification at different spatial resolutions. Photogrammetric Engineering and Remote Sensing, 68(11), 1155-1162.
- [9] Touzet, C. (1992). Les réseaux de neurones artificiels, introduction au connexionnisme. EC2.
- [10] PHILIPPE, B. (2006). Apprentissage Statistique & Data mining. Institut national des sciences appliquées (INSA) de toulouse, Version Octobre.
- [11] Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. Psychological review, 65(6), 386.
- [12] Parizeau, M. (2004). Réseaux de neurones. GIF-21140 et GIF-64326, 124.
- [13] Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. Proceedings of the national academy of sciences, 79(8), 2554-2558.
- [14] Kohonen, T. (1982). "Self-organized formation of topologically correct feature maps", Biological Cybernetics, Vol. 43, pp. 59-69.
- [15] Rosenblatt, F. (1962). Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms Spartan Books. New York.
- [16] Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. nature, 323(6088), 533.
- [17] <https://fr.mathworks.com/discovery/deep-learning.html#howitworks>.

- [18] Valentin. B. (2015) APPROCHES « DEEP LEARNING » APPLIQUEES AUX SIGNAUX AUDIO : PAROLE ET MUSIQUE (Soutenance le 04/09/15).
- [19] Lee, H., Grosse, R., Ranganath, R., & Ng, A. Y. (2009, June). Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations. In Proceedings of the 26th annual international conference on machine learning (pp. 609-616). ACM.
- [20] LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural computation*, 1(4), 541-551.
- [21] LeCun, Y., Haffner, P., Bottou, L., & Bengio, Y. (1999). Object recognition with gradient-based learning. In *Shape, contour and grouping in computer vision* (pp. 319-345). Springer, Berlin, Heidelberg.
- [22] LeCun, Y., Boser, B. E., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W. E., & Jackel, L. D. (1990). Handwritten digit recognition with a back-propagation network. In Advances in neural information processing systems (pp. 396-404).
- [23] Hinton, G. E., Osindero, S., & Teh, Y. W. (2006). A fast learning algorithm for deep belief nets. *Neural computation*, 18(7), 1527-1554.
- [24] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105).
- [25] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.
- [26] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015, June). Going deeper with convolutions. *Cvpr*.
- [27] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).
- [28] LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278-2324.
- [29] Glorot, X., Bordes, A., & Bengio, Y. (2011, June). Deep sparse rectifier neural networks. In Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics (pp. 315-323).
- [30] Maas, A. L., Hannun, A. Y., & Ng, A. Y. (2013, June). Rectifier nonlinearities improve neural network acoustic models. In Proc. icml (Vol. 30, No. 1, p. 3).

- [31] Maas, A. L., Hannun, A. Y., & Ng, A. Y. (2013, June). Rectifier nonlinearities improve neural network acoustic models. In Proc. icml (Vol. 30, No. 1, p. 3).
- [32] He, K., Zhang, X., Ren, S., & Sun, J. (2015). Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In Proceedings of the IEEE international conference on computer vision (pp. 1026-1034).
- [33] Clevert, D. A., Unterthiner, T., & Hochreiter, S. (2016). Fast and accurate deep network learning by exponential linear units (elus). *arXiv preprint arXiv:1511.07289*.
- [34] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, 15(1), 1929-1958.
- [35] Marr, D., & Hildreth, E. (1980). Theory of edge detection. *Proc. R. Soc. Lond. B*, 207(1167), 187-217.
- [36] Rousselle, J. J. (2003). Les contours actifs, une méthode de segmentation : application à l'imagerie médicale (Doctoral dissertation, Tours).
- [37] M Sandeli, traitement d'images par des approches bio-inspirées application à la segmentation d'images, université constantine 2. 2014.
- [38] D. Azerradj, D. Kessoum. "Application des contours actifs géodésiques sur des séquences d'images", Mémoire d'ingénieur. INI 2007.
- [39] Naciri H., Chaoui N, Conception et Réalisation d'un système automatique d'identification des empreintes digitales, Mémoire de PFE, Université de Tlemcen, 2003.
- [40] A. Del BIMBO. Visual Information Retrieval. Academic Press Morgan Kaufmann Publishers, 1999.
- [41] MADJAOUI, A., & FARES, F. (2014). *Segmentation des Images par Contours Actifs: Application sur les Images Satellitaires à Haute Résolutions* (Doctoral dissertation).
- [42] Hadjila F. & Bouabdellah R., « Reconnaissance des visages par les réseaux de neurones », Mémoire de PFE, Université de Tlemcen, 2003.
- [43] Cocquerez J.-P., Philipp S., « Analyse d'images : filtrage et segmentation », Edition Masson, 1998.
- [44] Gonzales R., Wintz P. « Digital Image Processing », Addison Wesley, 1977.
- [45] Tabari K., Tagma S., « Compression d'images animées à très faible débit par la géométrie des fractales », Projet de fin d'études, USTBH, 1996.
- [46] Ayadi hakim, Bouanoun bilal, "Système d'aide à la détection de maladies sur image médicale basé sur les réseaux de neurones", mémoire d'ingénieur d'état en informatique;

2010.

- [47] Raffi ENFICIAUD, “Algorithmes multidimensionnels et multi spectraux en Morphologie Mathématique : Approche par méta-programmation”, Thèse pour obtenir le grade de Docteur de l'école des Mines de Paris Spécialité « Morphologie Mathématique » le 26 février 2007.
- [48] CHIKH, M. T. (2011). Amélioration des images par un modèle de réseau de neurones (Comparaison avec les filtres de base.
- [49] El Abed, H., & Margner, V. (2007, September). Comparison of different preprocessing and feature extraction methods for offline recognition of handwritten arabic words. In Document Analysis and Recognition, 2007. ICDAR 2007. Ninth International Conference on (Vol. 2, pp. 974-978). IEEE.
- [50] Al-Rashaideh, H. (2006). Preprocessing phase for Arabic word handwritten recognition. Information Process (Russian), 6(1).
- [51] Farooq, F., Govindaraju, V., & Perrone, M. (2005, August). Pre-processing methods for handwritten Arabic documents. In Document Analysis and Recognition, 2005. Proceedings. Eighth International Conference on (pp. 267-271). IEEE.
- [52] Doménech, J., Toselli, A. H., Juan, A., Vidal, E., & Casacuberta, F. (2001). An off-line HTK-based OCR system for isolated handwritten lowercase letters.
- [53] Dreuw, P., Jonas, S., & Ney, H. (2008, December). White-space models for offline Arabic handwriting recognition. In Pattern Recognition, 2008. ICPR 2008. 19th International Conference on (pp. 1-4). IEEE.
- [54] Deufemia, V., & Paolino, L. (2013, July). Combining unsupervised clustering with a non-linear deformation model for efficient petroglyph recognition. In International Symposium on Visual Computing (pp. 128-137). Springer, Berlin, Heidelberg.
- [55] Fawcett, T. (2003). ROC graphs: notes and practical considerations for data mining researchers Technical Report HPL-2003-4. HP Labs.
- [56] Bishop, C. M. (2006). Machine learning and pattern recognition. Information Science and Statistics. Springer, Heidelberg.
- [57] Erus, G. (2008). Reconnaissance d'objets cartographiques dans les images satellitaires à haute résolution (Doctoral dissertation, Paris 5).
- [58] Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. International journal of computer vision, 60(2), 91-110.
- [59] Auclair, A., Cohen, L. D., & Vincent, N. (2007, July). How to use SIFT vectors to analyze an image with database templates. In International Workshop on Adaptive

- Multimedia Retrieval (pp. 224-236). Springer, Berlin, Heidelberg.
- [60] Fukushima, K. (1980). Neocognitron: a self organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological cybernetics*, 36(4), 193.
- [61] Hubel, D. H., & Wiesel, T. N. (1977). Ferrier lecture-Functional architecture of macaque monkey visual cortex. *Proc. R. Soc. Lond. B*, 198(1130), 1-59.
- [62] LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278-2324.
- [63] Ciresan, D. C., Meier, U., Masci, J., Maria Gambardella, L., & Schmidhuber, J. (2011, July). Flexible, high performance convolutional neural networks for image classification. In *IJCAI Proceedings-International Joint Conference on Artificial Intelligence (Vol. 22, No. 1, p. 1237)*.
- [64] Qian, R., Yue, Y., Coenen, F., & Zhang, B. (2016, August). Traffic sign recognition with convolutional neural network based on max pooling positions. In *Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), 2016 12th International Conference on (pp. 578-582)*. IEEE.
- [65] Zafra, J. E., Moreno, R. J., & Hernández, R. D. (2017). Comparison between Backpropagation and CNN for the Recognition of Traffic Signs. *International Journal of Applied Engineering Research*, 12(17), 6814-6820.
- [66] Lim, K., Hong, Y., Choi, Y., & Byun, H. (2017). Real-time traffic sign recognition based on a general purpose GPU and deep-learning. *PLoS one*, 12(3), e0173317.
- [67] Salhi, A., Minaoui, B., Fakir, M., Chakib, H., & Grimech, H. (2017). Traffic Signs Recognition using HP and HOG Descriptors Combined to MLP and SVM Classifiers. *Traffic*, 8(11).
- [68] Reinders, C., Ackermann, H., Yang, M. Y., & Rosenhahn, B. (2017). Learning a Fully Convolutional Network for Object Recognition using very few Data. *arXiv preprint arXiv:1709.05910*.
- [69] Friedman, J., Hastie, T., & Tibshirani, R. (2001). *The elements of statistical learning (Vol. 1, pp. 337-387)*. New York: Springer series in statistics.
- [70] Breiman, L., Friedman, J., Olshen, R., & Stone, C. (1984). *Classification and Regression Trees (Wadsworth & Brooks Cole, Monterey, CA)*. A powerful yet simple technique for ecological data analysis. *Ecology*, 81, 3178-3192.
- [71] Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20(3), 273-297.

- [72] Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20(3), 273-297.
- [73] Blum, A. L., & Langley, P. (1997). Selection of relevant features and examples in machine learning. *Artificial intelligence*, 97(1-2), 245-271.
- [74] Keysers, D., Dahmen, J., & Ney, H. (2001). Invariant classification of red blood cells. In *Bildverarbeitung für die Medizin 2001* (pp. 367-371). Springer, Berlin, Heidelberg.
- [75] LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural computation*, 1(4), 541-551.
- [76] Trier, Ø. D., Jain, A. K., & Taxt, T. (1996). Feature extraction methods for character recognition-a survey. *Pattern recognition*, 29(4), 641-662.
- [77] Lauziere, Y. B., Gingras, D., & Ferrie, F. P. (2001). A model-based road sign identification system. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on* (Vol. 1, pp. I-I). IEEE.
- [78] Suard, F. (2006). Méthodes à noyaux pour la détection de piétons (Doctoral dissertation, INSA de Rouen).
- [79] Negri, P., Clady, X., Hanif, S. M., & Prevost, L. (2008). A cascade of boosted generative and discriminative classifiers for vehicle detection. *EURASIP Journal on Advances in Signal Processing*, 2008, 136.
- [80] Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics*, 9(1), 62-66.
- [81] Di Baja, G. S., & Thiel, E. (1996). Skeletonization algorithm running on path-based distance maps. *Image and vision Computing*, 14(1), 47-57.
- [82] <https://www.python.org>.
- [83] <http://scikit-learn.org>.
- [84] Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., ... & Ghemawat, S. (2016). Tensorflow: Large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv:1603.04467.) .