

République Tunisienne  
Ministère de l'Enseignement Supérieur  
et de la recherche scientifique



Université de Gafsa  
Institut Supérieur des Sciences  
Appliquées et de la Technologie de Gafsa

Cycle de Formation en Mastère  
Professionnelle dans la  
Discipline Science de données

Mémoire de **MASTERE**  
**Sciences de Données**  
N° d'ordre : 01-MPSD

# MEMOIRE

*Présenté à*

**L'Institut Supérieur des Sciences Appliquées et  
de Technologie de Gafsa**

**(Département Informatique et télécommunication)**  
*En vue de l'obtention Diplôme en Sciences de données*

## MASTERE

Dans la discipline Sciences de Données

*Par*

Saifeddine OUNI

---

### CREATION D'UN MODELE DE DETECTION AUTOMATIQUE DES COMPORTEMENTS DES CITOYENS BASE SUR L'APPRENTISSAGE AUTOMATIQUE

---

*Soutenu devant le jury composé de :*

M.

*Président*

M.

*Rapporteur*

Mm.

Hayfa CHORFI

*Encadreur*

M.

Ahmed KHLIFI

*Co-Encadreur*

A.U : 2022 – 2023

## Dédicaces

---

*Tous les mots ne sauraient exprimer la gratitude, l'amour, le respect, la reconnaissance, c'est tout simplement que : Je dédie ce travail à : à ma tendre Mère : Tu représentes pour moi la source de tendresse et l'exemple de dévouement qui n'a pas cessé de m'encourager. Tu as fait plus qu'une mère puisse faire pour que ses enfants suivent le bon chemin dans leur vie et leurs études. . À l'esprit de mon père : aucune dédicace ne saurait exprimer l'Amour, l'estime, le dévouement et le respect que j'ai toujours pour vous. Rien au monde ne vaut les efforts fournis jour et nuit pour mon éducation et mon bien-être. Ce travail est le fruit de tes sacrifices que tu as consentis pour mon éducation et ma formation le long de ces années. À mes chères sœurs et mon frère : pour vous exprimer mon amour éternel, mes respects et ma fierté d'être membre de cette famille.*

## Remerciements

---

*Je voudrais premièrement remercier à Dieu pour toutes les bénédictions dont il a comblé ma vie .Ce mémoire n'aurait pas été possible sans l'intervention, consciente, d'un grand nombre de personnes. Je remercie **M me CHORFI HAYFA & Mr KHLIFI AHMED** d'avoir accepté de diriger ce travail de recherche .Je tiens également à exprimer ma gratitude à ma famille, en particulier à mes frères et sœurs qui ont toujours soutenu et poussé à poursuivre mes études. Grâce à leur soutien, ce travail actuel devient possible. Je tiens également à remercier tous les professionnels et amis qui ont travaillé dans le domaine d'expert sciences de données dans le cadre de cette étude qui ont gentiment répondu à mes différentes questions dans ce domaine pour favoriser le développement de ce travail*

*Enfin, je remercier tous ceux qui, de près ou de loin, ont contribué à la réalisation de ce travail.*

*Merci*

## RÉSUMÉ :

J'étais très intéressé par l'avancement de l'intelligence artificielle, et je voulais enrichir mes connaissances dans ce domaine. L'intelligence artificielle étant un vaste sujet, j'ai choisi de me concentrer sur la détection automatique des comportements des citoyens.

La prédiction des comportements chez les humains nous permet d'éviter de nombreuses catastrophes et crimes, et nous aide à sauver davantage de vies et à atténuer la souffrance. Dans ce contexte, nous avons appliqué la technique de l'apprentissage profond, qui est l'un des mécanismes de l'intelligence artificielle, plus précisément les réseaux neuronaux convolutionnels (CNN), pour concevoir un système de prédiction des comportements des citoyens. Nous avons utilisé une base de données nommé UCF101 contenant de nombreuses vidéos représentant différents comportements humaines, et nous avons obtenu de bons résultats.

**Mots clés :** l'intelligence artificielle, comportements, l'apprentissage profond, CNN, UCF101

## ABSTRACT:

I was very interested in the advancement of artificial intelligence, and I wanted to enrich my knowledge in this field. Artificial intelligence being a vast subject, I chose to focus on the automatic detection of citizen behavior.

Predicting behaviors in humans allows us to avert many disasters and crimes, and helps us save more lives and alleviate suffering. In this context, we applied the technique of deep learning, which is one of the mechanisms of artificial intelligence, more precisely convolutional neural networks (CNN), to design a system for predicting citizen behaviors. We used a database named UCF101 containing many videos representing different human behaviors, and we got good results.

**Keywords:** Artificial intelligence , abnormal behaviors, deep learning, CNN, UCF101

## التلخيص:

كنت مهتمًا جدًا بتقدم الذكاء الاصطناعي ، وأردت إثراء معرفتي في هذا المجال. نظرًا لكون الذكاء الاصطناعي موضوعًا واسعًا ، فقد اخترت التركيز على الاكتشاف التلقائي لسلوك المواطن. يتيح لنا التنبؤ بالسلوكيات البشرية تجنب العديد من الكوارث والجرائم ، ويساعدنا في إنقاذ المزيد من الأرواح وتخفيف المعاناة. في هذا السياق ، قمنا بتطبيق تقنية التعلم العميق ، وهي إحدى آليات الذكاء الاصطناعي ، وبشكل أكثر دقة الشبكات العصبية التلافيفية (CNN) ، لتصميم نظام للتنبؤ بسلوكيات المواطنين. استخدمنا قاعدة بيانات تسمى UCF101 تحتوي على العديد من مقاطع الفيديو التي تمثل سلوكيات بشرية مختلفة ، وحصلنا على نتائج جيدة.

الكلمات المفتاحية: الذكاء الاصطناعي، السلوكيات الشاذة، التعلم العميق، CNN، UCF101

# *Abréviations*

- **IA** Intelligence Artificielle
- **ML** Machine Learning
- **DL** Deep Learning
- **CNN** Convolutional neural network
- **SVM** Support Vector Machines
- **GPU** Graphics processing unit
- **TOC** Trouble obsessionnel compulsif
- **RNN** Recurrent Neural Network
- **TALN (NLP)** Natural Language Processing
- **ANN** Artificial Neural Network

# *Sommaire*

Introduction générale .....	1
Chapitre 1 : .....	2
<b>Etat du l'art</b> .....	2
1. Introduction .....	3
2. Cadre général du projet.....	3
2.1. Présentation du projet.....	3
2.2. Problématique .....	3
2.3. Solution proposée et déficit.....	4
3. Etude technologique .....	5
3.1. Intelligence artificielle .....	5
3.2. Machine Learning .....	5
3.3. Deep Learning .....	6
3.4. La Computer Vision .....	6
4. Le Comportement.....	7
4.1. Définition.....	7
4.2. Les différents types de troubles du comportement.....	8
4.3. Les causes des troubles du comportement .....	9
5. la détection d'anomalies .....	9
5.1 Etude de quelques approches pour la détection d'anomalies .....	9
5.1.1 K-Means Clustering.....	10
5.1.2 Forêt d'isolement.....	10
5.1.3 Machines vecteurs de support (SVM) .....	10
5.1.4 Réseaux de neurones.....	11
5.1.5 Convolutional neural network (CNN) .....	11
6. Utilisation de la détection d'anomalies basée sur l'apprentissage automatique .....	12
6.1. Jeux de données utilisés .....	12
6.2. Méthodologies appliquées.....	12
7. Conclusion .....	13
Chapitre2 : .....	14

<b>Les réseaux de neurones convolutifs</b>	14
1. Introduction	15
2. Neurone formel	15
3. L'architecture des réseaux de neurones convolutifs (CNN)	16
3.1. Couche entièrement connecté (Fully Connected)	17
3.2. Couches convolutives	18
3.3. Couches de pooling	18
3.4. Couches totalement connectées	18
4. Avantage du CNNs	19
5. L'application des réseaux neuronaux convolutifs	19
6. Principe général de CNN	19
7. Différents types de filtres à convolutions	20
7.1. Filtres linéaires	21
7.2. filtres non linéaires	21
8. Apprentissage des réseaux de neurones convolutifs	21
8.1. Données d'entraînement	21
8.2. Fonctionnement de détection	22
9. Apprentissage automatique	22
9.1.1. L'apprentissage supervisé	23
9.1.2. Apprentissage non supervisé	23
10. Conclusion	23
Chapitre3 :	24
<b>Validation</b>	24
1. Introduction	25
2. Les outils de développement	25
2.1 . L'environnement logiciel	25
2.1.1 Python	25
2.1.2 Google Colab	25
2.2 Les différentes bibliothèques	26
2.2.1 Numpy	26
2.2.2 TensorFlow	26
2.2.3 OpenCV	27
2.3 L'environnement matériel	28
3. Base d'apprentissage	28
4. Model proposé	31

4.1.	Reconnaissance et détection d'actions humaines .....	32
4.1.1.	Gestes .....	33
4.1.2.	Actions .....	33
4.1.3.	Activités .....	34
4.2.	Étapes de création du modèle proposé.....	34
5.	Résultats obtenus et discussion .....	35
5.1.	Teste1 : .....	36
5.2.	Teste2 : .....	36
5.3.	Teste3 : .....	37
6.	Conclusion .....	37
Conclusion Générale.....		38
& perspectives .....		38
Références/bibliographiques .....		39



# *Liste des Figures*

Figure 1. 1: les différences entre IA, ML et DL[4] .....	7
Figure 1. 2: Les troubles du comportement alimentaire[5] .....	8
Figure 1. 3: Les troubles de l'humeur[6].....	9
Figure 1. 4: Means Clustering[7] .....	10
Figure 1. 5: SVM [8] .....	11
Figure 1. 6: CNN [9].....	12
Figure 2. 1:Les couches de CNN[12] .....	16
Figure 2. 2:Figure représentant les deux parties d'un CNN [13].....	17
Figure 2. 3: Principe de la couche entièrement connectée.....	17
Figure 2. 4: Max pooling .....	18
Figure 2. 5: Reconnaissance d'une image avec CNN [17].....	19
Figure 2. 6: Fonctionnement d'un CNN .....	20
Figure 2. 7 : Schéma de l'architecture de permettant de faire de la fusion de données multi-sources dans un CNN [14] .....	20
Figure 2. 8: Schéma des différentes techniques issues de l'apprentissage automatique .....	23
Figure3. 1: google Colab.....	26
Figure3. 2: bibliothèque Nump .....	26
Figure3. 3: Logo de tensorFlow .....	27
Figure3. 4: bibliothèque OpenCV .....	28
Figure3. 5: Catégories de clips vidéo dans l'ensemble de données UCF101 .....	29
Figure3. 6: - QUELQUES ILLUSTRATIONS.....	29
Figure3. 7: Catégories vs durée de la vidéo dans chaque catégorie.....	30
Figure3. 8: liste de fichiers vidéo UCF101 .....	30
Figure3. 9: fichier étiquettes .....	31
Figure3. 10: QUELQUES ILLUSTRATIONS .....	32
Figure3. 11: organigrammes le fonctionnement du Model .....	35
Figure3. 12: Test1 de validation.....	36
Figure3. 13: Test2 de validation.....	36
Figure3. 14: Test3 de validation.....	37

## *Liste des Tableaux*

Tableau 1: Architecture de machine.....	28
Tableau 2: Principaux moyens de communiquer les intentions.....	33

# Introduction générale

La détection d'anomalies traite le problème de la recherche d'éléments de données qui ne suivent pas les modèles de la majorité des données. La tâche consiste à distinguer les bons éléments des éléments anormaux. Cela peut être défini comme un problème de multiclassés et en tant que tel résolu avec des techniques d'apprentissage supervisé.

Avec la demande croissante de sécurité d'aujourd'hui, en particulier dans les lieux publics tels que les aéroports, les gares, les supermarchés, les écoles et les rues bondées, les caméras de surveillance sont utilisées pour surveiller les activités quotidiennes et détecter les événements anormaux. Cette tâche se concentre sur la localisation des anomalies en utilisant à la fois des informations temporelles et partielles dans des vidéos. Les anomalies peuvent être définies comme des événements s'écartant du comportement normal [1], par exemple, se battre, se faufiler ou porter des sacs sans surveillance dans un aéroport. Le but de l'utilisation des caméras de surveillance est la détection précoce des comportements humains anormaux. Il s'agit d'une tâche critique dans de nombreux cas où une intervention humaine est nécessaire, par exemple pour la prévention du crime ou la lutte contre le terrorisme. Afin de détecter ces anomalies, qu'elle soit simples ou complexes, il est possible d'utiliser des méthodes récentes basées sur de l'Intelligence Artificielle (I.A.)

L'objectif du travail présenté dans ce mémoire consiste à Création d'un model de détection automatique des comportements anormaux des citoyens basé sur l'apprentissage automatique Pour cela nous avons partagé notre travail en trois grandes parties organisées comme suit :

- Le premier chapitre est consacré de la préparation d'une étude bibliographique sur des exemples existants et les méthodes d'apprentissage automatique ainsi une étude technologique sur les approches de détection d'anomalies
- Dans le deuxième chapitre, nous allons déployer la technique de Deep Learning et les réseaux de neurones profonds a fin les bien utiliser.
- Le troisième chapitre nous présenterons les différents résultats obtenus
- nous terminerons par une conclusion générale et les perspectives ouvertes par ce travail

# **Chapitre 1 :**

## **Etat du l'art**

## **1. Introduction**

Une anomalie est une déviation ou un changement inattendu par rapport à un modèle attendu dans un ensemble de données. La détection des anomalies est utilisée pour signaler un comportement anormal, car les anomalies montrent que quelque chose de différent de ce qui était attendu se produit.

Dans ce chapitre, nous allons préparer une étude bibliographique sur des exemples existants et les méthodes d'apprentissage automatique ainsi une étude technologique sur les approches de la création de système de détection d'anomalie

## **2. Cadre général du projet**

De nos jours, le système de détection les comportements anormaux des citoyens sont très important car ils sont très appréciés des utilisateurs, La détection d'anomalies traite le problème de la recherche d'éléments de données qui ne suivent pas les modèles de la majorité des données. Problème de distinguer les bons éléments des éléments anormaux. Cela peut être défini comme un problème de classification binaire et en tant que tel résolu avec des techniques d'apprentissage supervisé.. Le domaine de l'intelligence artificielle étant un domaine assez vaste, avec diverses possibilités d'extension et avantages pouvant offrir, et puisque les agents conventionnels sont aujourd'hui au centre de l'attention en matière de nouvelles technologies informatiques dans ce travail, nous attèlerons pour créer notre modèle à la deep learning et plus précisément le réseau de neurones convolutifs.

### **2.1.Présentation du projet**

Dans le cadre de notre étude en Mastère Professionnelle en Sciences des Données à l'ISSAT de Gafsa, nous avons développé un projet qui s'intitule : Création d'un modèle de détection de comportements anormaux des citoyens basé sur l'apprentissage automatique qui a comme but de comprendre maximum possible des images. Donc, nous allons par la suite essayer de comprendre les notions fondamentales de l'apprentissage profond basé sur l'intelligence artificielle (IA) pour arriver à créer ce modèle.

### **2.2.Problématique**

Comportement individuel au quotidien en inter-relation avec son environnement est soumis à des limites intrinsèques et extrinsèques personnelles sociales et culturelles ; plus précisément les lois en général.

Et à nos recherches, nous trouvons que l'intelligence artificielle est un phénomène d'actualité d'aujourd'hui qui ne cesse de développer et d'évoluer puisque on trouve des nouveautés dans ce domaine presque chaque jour, c'est pour cela nous trouvons plusieurs entreprises investissent dans ce domaine afin de sécuriser.

Comme toute solution informatisée, bien qu'elle offre des avantages pour gérer les espaces publics. Mais il y a encore plusieurs inconvénients, qui nous mènent à créer notre projet, notamment :

- Manque de signaler toute activité, agissement ou comportement suspects, car inhabituels, inappropriés, anormaux...
- Manque de Signaler tout objet suspect inhabituel ou abandonné dont le propriétaire ne peut être identifié
- Manque de respecter les consignes, notamment dans les lieux et bâtiments publics (gares, aéroports, administrations, musées, restaurants, cafés, commerces...)
- Difficultés d'actions des autorités, tels les fouilles, les inspections de paquets, de bagages,...

### **2.3.Solution proposée et défit**

Pour faire face aux problèmes qui existent, nous sommes besoin d'un système qui automatise toutes les tâches d'un bon système de Détection des comportements des citoyens. A ce niveau nous proposons une solution améliorée et fiable qui prend en compte tous les exigences actuels de la technologie et qui permettra de répondre aux futures besoins des utilisateurs. les comportements sont également de rigueur pour les responsables de lieux ou de bâtiments publics (administrations) ou privés (entreprises) qui mettent en oeuvre des mesures de protection spécifiques sur base du niveau d'alerte respectif.

Donc, nous avons réalisé une modèle de détection anormaux des citoyens qui a pour but de :

- Détecté tout objet suspect inhabituel ou abandonné dont le propriétaire ne peut être identifié
- Respecter les consignes, notamment dans les lieux et bâtiments publics (gares, aéroports, administrations, musées, restaurants, cafés, commerces...)

- se soumettre et faciliter les actions des autorités, tels les fouilles, les inspections de paquets, de bagages,...

### **3. Etude technologique**

Une étude du projet commence toujours par la collecte d'informations qui pourraient faciliter l'imagination de la solution technologique. L'utilité de créer un système de détection et de reconnaissance des comportements anormaux des citoyens. C'est une discussion entre humain et machine, le créateur toujours aimerait rendre la machine aussi vivante qu'un être humain, c'est le concept de l'intelligence Artificielle. Et Pour parler d'un système intelligent basé sur le Deep learning il est nécessaire de définir les grands axes où se déroule notre projet.[1]

#### **3.1.Intelligence artificielle**

L'intelligence artificielle est une discipline de l'informatique qui a pour but de créer des machines intelligentes, en "opposition" avec l'intelligence naturelle des êtres vivants. Le terme a beaucoup évolué au fil du temps, il englobe dorénavant toutes les idées visant à permettre à une machine de pouvoir émuler les capacités cognitives de l'Homme, et de les surpasser. Ce terme "d'intelligence artificielle" voit le jour en 1956 après les nombreux travaux débutés post Seconde Guerre Mondiale et constitue l'un des plus récents champs d'études parmi les sciences et l'ingénierie. Cela fait suite aux nombreuses inventions au fil des siècles permettant à des machines de calculer que l'hypothèse qu'elles puissent penser et agir par elles-mêmes virent le jour également.[2]

L'intelligence artificielle (IA) est un processus d'imitation de l'intelligence humaine qui repose sur la création et l'application d'algorithmes exécutés dans un environnement informatique dynamique. Son but est de permettre à des ordinateurs de penser et d'agir comme des êtres humains. Tout simplement on peut dire que l'intelligence artificielle c'est une machine intelligente qui simule la pensée ainsi le comportement de l'être humain.

Pour y parvenir, trois composants sont nécessaires :

- Des systèmes informatiques
- Des données avec des systèmes de gestion
- Des algorithmes d'IA avancés (code)

#### **3.2.Machine Learning**

Machine Learning : Appelé aussi apprentissage automatique. C'est une méthode qui utilise l'Intelligence Artificielle en se fondant sur des approches statistiques pour

donner aux machines la capacité d'apprendre à partir de données provenant du big data, c'est à dire de s'améliorer en ayant la capacité à résoudre des tâches sont être explicitement programmées.

L'apprentissage automatique, également appelé apprentissage machine ou apprentissage artificiel et en anglais machine learning, est une forme d'intelligence artificielle (IA) qui permet à un système d'apprendre à partir des données et non à l'aide d'une programmation explicite. Cependant, l'apprentissage automatique n'est pas un processus simple. Au fur et à mesure que les algorithmes ingèrent les données de formation, il devient possible de créer des modèles plus précis basés sur ces données. Un modèle de machine learning est le résultat généré lorsque vous entraînez votre algorithme d'apprentissage automatique avec des données. Après la formation, lorsque vous fournissez des données en entrée à un modèle, vous recevez un résultat en sortie. Par exemple, un algorithme prédictif crée un modèle prédictif. Ensuite, lorsque vous fournissez des données au modèle prédictif, vous recevez une prévision qui est déterminée par les données qui ont servi à former le modèle.

### **3.3.Deep Learning**

Le Deep Learning ou apprentissage profond est un type d'intelligence artificielle dérivé de la machine learning (apprentissage automatique) où la machine est capable d'apprendre par elle-même, contrairement à la programmation où elle se contente d'exécuter à la lettre des règles prédéterminées.

L'apprentissage profond s'appuie sur un réseau de neurones artificiels s'inspirant du cerveau humain. Ce réseau est composé de dizaines voire de centaines de « couches » de neurones, chacune recevant et interprétant les informations de la couche précédente. Le système apprendra par exemple à reconnaître les lettres avant de s'attaquer aux mots dans un texte, ou détermine s'il y a un visage sur une photo avant de découvrir de quelle personne il s'agit.

Ces trois domaines peuvent être vus comme étant imbriqués mais ils ne sont pas au même niveau comme il indique cette figure :

### **3.4 La Computer Vision**

La vision par ordinateur est un domaine de l'intelligence artificielle (IA) qui permet aux ordinateurs et aux systèmes de dériver des informations significatives à partir d'images numériques, de vidéos et d'autres entrées visuelles, et de prendre des mesures ou de faire des recommandations sur la base de ces informations. Si l'intelligence artificielle permet



aux ordinateurs de penser, la vision par ordinateur leur permet de voir, d'observer et de comprendre.

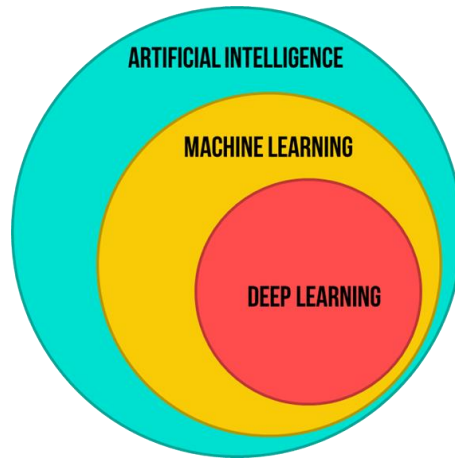
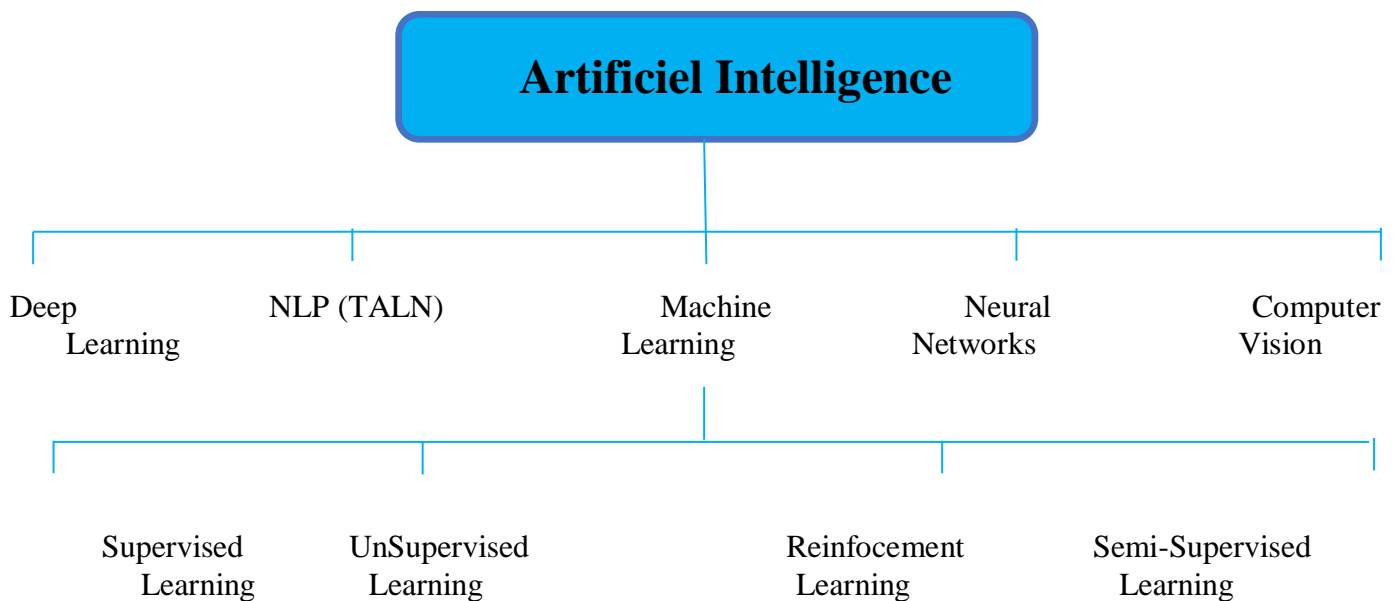


Figure 1. 1: les différences entre IA, ML et DL[4]



## 4. Le Comportement

### 4.1.Définition

On peut définir le comportement comme étant la manière d'agir ou la façon de se comporter au quotidien. C'est donc un terme très général qui ne dispose pas de définition «

scientifique ». « Les troubles du comportement sont liés à des circonstances sociales ou culturelles et attestent d'un désordre psychique. Ils peuvent se traduire par de l'agitation, de l'agressivité, des troubles obsessionnels compulsifs (TOC), des désordres alimentaires (anorexie, boulimie...), de l'hyperactivité, de l'addiction (alcool, tabac, autres drogues, jeu, travail, sport, écrans...) ou des phobies ». [3]

Pour être diagnostiquée comme telle, chacune de ces anomalies doit entraîner une altération cliniquement significative du fonctionnement social, scolaire ou professionnel. Ces troubles peuvent se manifester à n'importe quel moment de la vie, de l'enfance à l'âge adulte.

## **4.2. Les différents types de troubles du comportement**

### **4.2.1. Les troubles du comportement sport**

Dans les esprits, pratiquer une activité sportive est gage de bonne santé. Mais quand la pratique s'intensifie et que la course à la performance vire à l'obsession, certains peuvent basculer dans des comportements compulsifs associés à des troubles du comportement alimentaire (TCA).



*Figure 1. 2: Les troubles du comportement [5]*

### **4.2.2. Troubles de l'étude**

Récemment, nombreux ont relayé l'étude d'une équipe de chercheurs français de l'Université de Rennes qui pensaient avoir trouvé une cause anatomique potentielle à la dyslexie. Malheureusement, il s'avère que cette dernière n'est que parcellaire. En effet, divers spécialistes des troubles de l'apprentissage ont mis en exergue les manquements et le côté simpliste de cette étude. En effet, la dyslexie se manifesterait sous diverses formes et aurait des causes multiples. Par ailleurs, le panel des personnes observées dans l'étude ne serait pas suffisamment représentatif pour tirer des conclusions scientifiquement valides.

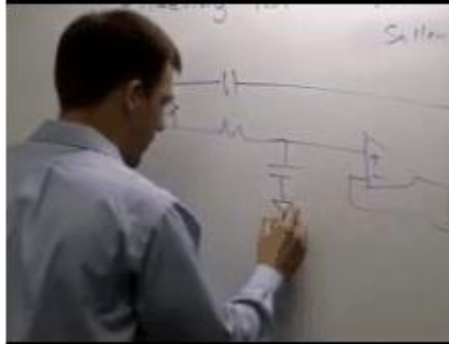


Figure 1. 3: Les troubles de l'humeur [6]

#### **4.3. Les causes des troubles du comportement**

Les troubles du comportement pour certains (troubles bipolaires, etc.) peuvent être liés à une forte prédisposition familiale qui entraîne une mauvaise humeur et une incapacité à réguler ses émotions. Elle peut également résulter d'un traumatisme émotionnel (rupture, exposition à la violence, difficultés financières), d'un traumatisme crânien ou être le symptôme d'une autre maladie comme une maladie fébrile (paludisme, septicémie), la maladie d'Alzheimer ou une tumeur au cerveau.[3]

### **5. la détection d'anomalies**

La détection d'anomalies est une technique utilisée pour identifier des modèles inhabituels qui ne correspondent pas au comportement attendu des données, et est un processus délibéré pour déterminer ce qui est normal et ce qui est anormal à partir des échantillons.

Ils sont souvent appliqués à des données non étiquetées, connues sous le nom de détection d'anomalies non supervisées.

#### **5.1 Etude de quelques approches pour la détection d'anomalies**

Les algorithmes d'apprentissage automatique peuvent être utilisés pour détecter les anomalies des citoyens plus efficacement et plus précisément que les méthodes manuelles. Certains des algorithmes les plus couramment utilisés pour la détection d'anomalies sont :

### 5.1.1 K-Means Clustering

K-Means Clustering est un algorithme d'apprentissage automatique non supervisé qui regroupe des points de données en groupes distincts. Il est utilisé pour détecter les anomalies en séparant les points de données qui ne rentrent pas dans le même cluster.

L'algorithme de clustering k-means effectue principalement deux tâches :

- Détermine la meilleure valeur pour K points centraux ou centroïdes par un processus itératif.
- Attribue chaque point de données à son centre k le plus proche. Les points de données proches du centre k particulier créent un cluster.

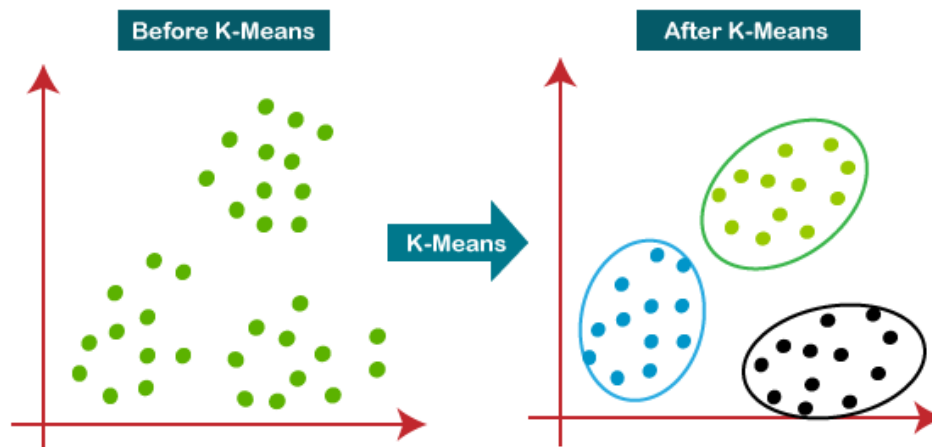


Figure 1. 4: Means Clustering[7]

### 5.1.2 Forêt d'isolement

les forêts d'isolement sont un algorithme d'apprentissage automatique non supervisé qui fonctionne en isolant des points de données individuels et en formant un ensemble d'arbres de décision. L'algorithme est capable de détecter les anomalies en fonction du temps nécessaire pour séparer les points de données.

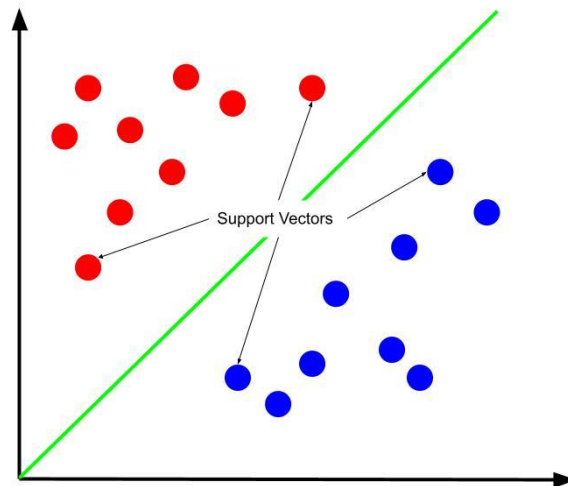
### 5.1.3 Machines vecteurs de support (SVM)

Machines à vecteurs de support : les machines à vecteurs de support sont des algorithmes d'apprentissage automatique supervisés qui utilisent une fonction mathématique pour classer les points de données. Ils sont utilisés pour détecter les anomalies en identifiant les points de données éloignés de l'ensemble de données global.

Les machines à vecteur support (Support Vector Machines) (**Figure 1.5**) sont une méthode de machine learning très utilisée dans les classifications . Leur but est d'apprendre à placer une délimitation entre deux classes. La frontière choisie doit être aussi lointaine

que possible des premiers éléments de chaque côté, ceci afin de ne pas biaiser la classification et d'optimiser la capacité de généralisation. De façon générale, un SVM aura pour tâche de trouver un hyperplan qui sépare les catégories. Dès lors, un tel classificateur déterminera la classe du nouvel objet selon sa position dans l'hyperplan par rapport à la frontière. En effet, les hyperplans sont placés de sorte que les catégories sont séparées par un « fossé ». Les nouveaux objets sont placés sur le plan et la classification de l'intent s'effectuera en fonction du côté du fossé vers lequel « tombe » ou « penche » l'objet.

Ce système peut générer ce type de frontières :



*Figure 1. 5: SVM [8]*

#### **5.1.4 Réseaux de neurones**

Les réseaux de neurones sont un type d'algorithme d'apprentissage automatique supervisé qui utilise plusieurs couches de nœuds pour traiter les données. Ils sont utilisés pour détecter les anomalies en identifiant des modèles et des corrélations qui ne correspondent pas au comportement attendu.

Ces algorithmes d'apprentissage automatique peuvent être utilisés en combinaison pour détecter les anomalies avec plus de précision et d'efficacité. À mesure que le monde numérique continue d'évoluer, ces algorithmes deviendront de plus en plus importants pour détecter.

#### **5.1.5 Convolutional neural network (CNN)**

Un CNN demande moins de prétraitement que d'autres méthodes. Cette technique est très utilisée pour la reconnaissance et la classification d'images. Comparativement aux RNN, un CNN est plus rapide car, lorsque le RNN calcule de façon séquentielle les mots

d'une phrase, le CNN le fait simultanément. Il est donc possible d'obtenir des résultats similaires en abaissant les contraintes calculatoires, ce qui en fait un modèle intéressant pour le TAL. Il semble l'être particulièrement pour les tâches de classification.

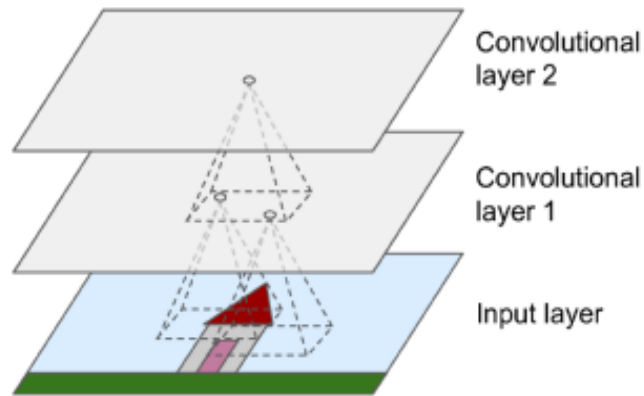


Figure 1. 6: CNN [9]

## 6. Utilisation de la détection d'anomalies basée sur l'apprentissage automatique

Cette partie se concentre sur la partie de la littérature qui utilise l'apprentissage automatique pour détecter les anomalies

### 6.1. Jeux de données utilisés

Avant d'appliquer une technique d'apprentissage automatique à un problème, il est nécessaire de disposer d'un jeu de données pour entraîner et tester le modèle. Cependant, il n'existe pas de consensus autour d'un jeu de données pour les anomalies que les chercheurs peuvent utiliser pour évaluer et comparer leurs travaux. En conséquence, chaque article utilise son propre jeu de données. Cependant, les auteurs suivent généralement une approche similaire pour la construction de leur jeu de données qui est composée de quatre étapes :

- ❖ le choix des anomalies à inclure dans le jeu de données
- ❖ la collecte des données
- ❖ l'extraction d'attributs
- ❖ l'étiquetage.

### 6.2. Méthodologies appliquées

Parmi les travaux examinés, différentes techniques ont été utilisées par les auteurs pour chaque étape du processus d'entraînement de leurs modèles d'apprentissage automatique. Ces choix méthodologiques sont présentés ci-après, selon quatre critères : le prétraitement appliqué aux données, les méthodes de sélection d'attributs, les algorithmes d'apprentissage automatique utilisés et le type de classification effectué.

### **6.2.1. Pré-traitements**

Au lieu d'alimenter directement leur algorithme d'apprentissage automatique avec leur jeu de données, certains articles appliquent des techniques de prétraitement afin d'améliorer les performances de leur modèle. Un pré-traitement qui est couramment utilisé lors dans le domaine de l'apprentissage automatique est la normalisation des données.

### **6.2.2. Sélection d'attributs**

La sélection d'attributs est une technique couramment utilisée en apprentissage automatique qui consiste à sélectionner un sous-ensemble d'attributs à partir de l'ensemble des attributs initiaux. Elle permet de filtrer les attributs redondants et non pertinents.

### **6.2.3. Algorithmes d'apprentissage automatique**

Les machines à vecteurs de support ou Support Vector Machine (SVM) est une classe d'algorithmes d'apprentissage automatique largement utilisé et réputé pour ses bonnes performances sur des tâches de classification variée. L'objectif d'un classificateur SVM est de trouver le meilleur hyperplan qui divise l'ensemble de données en fonction des classes à prédire.

### **6.2.4. Type de classification**

Tous les articles présentés utilisent une approche de classification binaire où chaque enregistrement issu du jeu de données est classé comme "1" si une anomalie est prédite et "0" dans le cas contraire. Cependant, certains articles vont plus loin en appliquant une approche multi-classes.

## **7. Conclusion**

Dans ce chapitre nous avons défini l' Etude technologique et la notion de Comportement et la détection anomalies. On a présenté les notions importantes qui sont en relation avec la Comportement humaine.

Le chapitre suivant sera consacré à la présentation des étapes de création des système de détection et nous allons bien détailler avec les réseaux de neurones convolutifs.

# Chapitre2 :

## **Les réseaux de neurones convolutifs**



## 1. Introduction

Dans ce chapitre nous allons focaliser sur un des algorithmes les plus performants du Deep Learning, les Convolutional Neural Network ou CNN : Réseaux de neurones convolutifs en français.

## 2. Neurone formel

De manière très simplifiée, un neurone biologique est une cellule caractérisée par : les points de connexion avec les autres neurones, fibres nerveuses ou musculaires des dendrites ou entrées des neurones les axones, ou sorties du neurone vers d'autres neurones ou fibres musculaires le noyau qui active les sorties en fonction des stimulations en entrée. [10]

Par analogie, le neurone formel est un modèle qui se caractérise par un état interne  $\in S$ , des signaux d'entrée  $x_1, \dots, x_p$  et une fonction d'activation.

$$S = h(x_1, \dots, x_p) = g(\alpha_0 + \sum_{j=1}^p \alpha_j x_j) = g(\alpha_0 + \alpha'x). \quad (1)$$

La fonction d'activation opère une transformation d'une combinaison affine des signaux d'entrée,  $\alpha_0$ , terme constant, étant appelé le biais du neurone.

Les différents types de neurones se distinguent par la nature  $g$  de leur fonction d'activation. Les principaux types sont :

- ✓ Linéaire  $g$  est la fonction identité,
- ✓ Seuil  $g(x) = 1[0, +\infty[(x)$
- ✓ Sigmoides  $g(x) = 1/(1 + e^{-x})$
- ✓ ReLU  $g(x) = \max(0, x)$  (rectified linear unit),
- ✓ Soft max  $g(x)_j = \frac{e^{x_j}}{\sum_{k=1}^K e^{x_k}}$  pour tout  $k \in \{1 \dots K\}$
- ✓ Radiale  $g(x) = \sqrt{1/2\pi} e^{-x^2/2}$ ,
- ✓ Stochastique  $g(x) = 1$  avec la probabilité  $1/(1 + e^{-x/H})$ , 0 sinon ( $H$  intervient comme une température dans un algorithme de recuit simulé).

Les modèles linéaires, sigmoïdes, ReLU, softmax sont bien adaptés à l'apprentissage d'algorithmes qui incluent la rétro propagation de gradient car leur fonction d'activation est différentiable, ils sont les plus utilisés.

Le modèle à seuil est sans doute plus conforme à la réalité biologique mais pose des problèmes d'apprentissage. Enfin le modèle stochastique est utilisé pour des problèmes d'optimisation globale de fonctions perturbées ou encore pour les analogies avec les systèmes de particules.

### 3. L'architecture des réseaux de neurones convolutifs (CNN)

Les réseaux de neurones convolutifs sont de loin le modèle le plus efficace pour classer les images. [11]

Il y a plusieurs couches différentes dans CNN comme le montre la figure 8 :

- Couche d'entrée (Input layer).
- Couche de convolution (Convo layer : Convolution + ReLU).
- Couche de Pooling.
- Couche entièrement connectée (Couche Fully connected).
- Couche Softmax/logistique.
- Couche de sortie (Output layer).

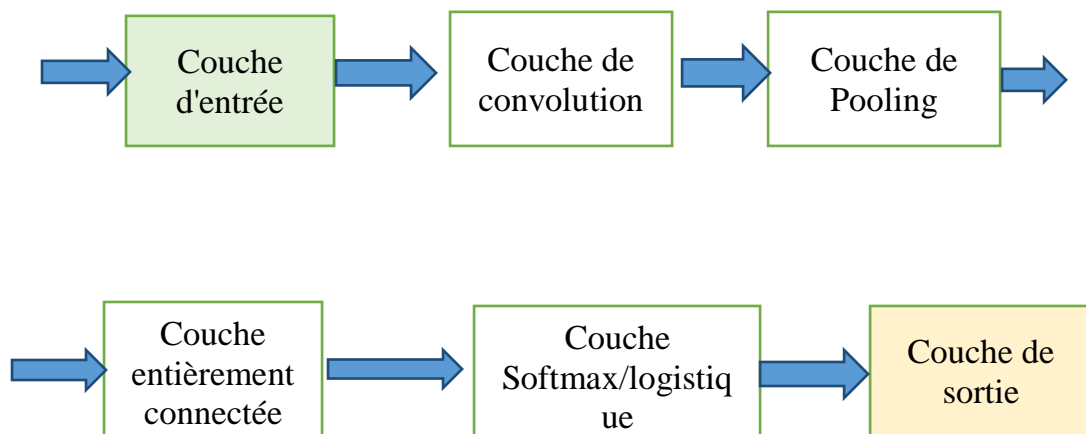


Figure 2. 1: Les couches de CNN [12]

Désignés par l'acronyme CNN, ils se composent de deux parties distinctes. En entrée, une image est fournie sous la forme d'une matrice de pixels. Elle a deux dimensions pour une image en niveaux de gris. La couleur est représentée par une troisième dimension, de profondeur 3 pour représenter les couleurs fondamentales.

La première partie de CNN est la partie convolutive elle-même (**figure 2.2**). Elle fonctionne comme un extracteur de caractéristiques des images. Une image est passée à travers une succession de couches, ou noyaux de convolution, créant de nouvelles images appelées cartes de convolutions. Figure représentant les deux parties d'un CNN

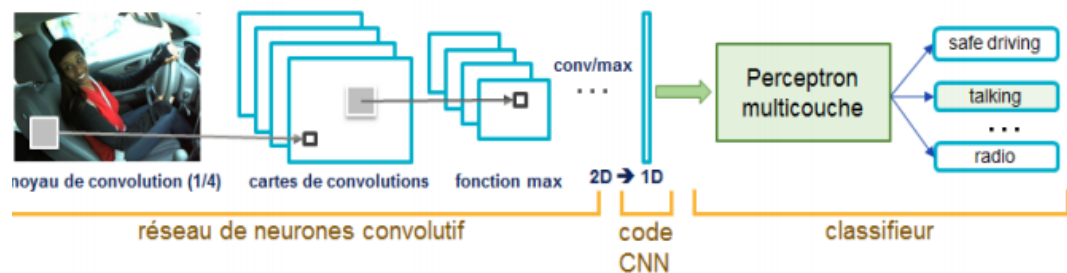


Figure 2. 2: Figure représentant les deux parties d'un CNN [13]

### 3.1. Couche entièrement connecté (Fully Connected)

Une couche entièrement connectée implique des poids, des biais et des neurones. Il connecte les neurones d'une couche aux neurones d'une autre couche. Il est utilisé pour classer les images entre différentes catégories par formation.

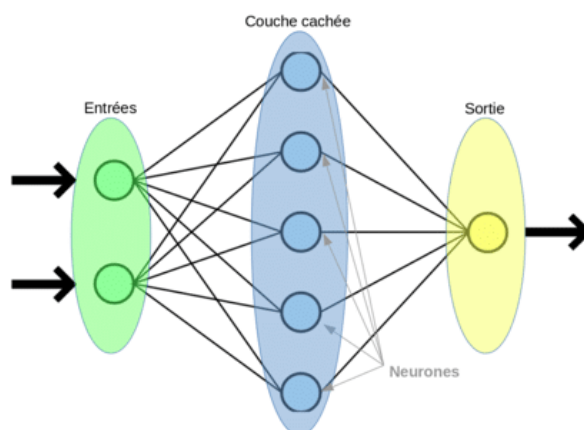


Figure 2. 3: Principe de la couche entièrement connectée

### 3.2. Couches convolutives

Les couches convolutives forment le cœur du réseau convolutif. Ces couches sont constituées d'un réseau allongé de neurones, qui ont un petit champ réceptif étendu à travers toute la profondeur du volume d'entrée. Ainsi, la couche convolutionnelle est juste une convolution d'image de la couche précédente, où les poids spécifient le filtre de convolution.

### 3.3. Couches de pooling

Après chaque couche convolutive, il peut y avoir une couche de pooling. La couche de pooling sous échantillonne leur entrée. Il y a plusieurs façons de faire cette mise en commun, comme prendre la moyenne ou le maximum, ou une combinaison linéaire apprise des neurones dans le bloc. Par exemple, la (Figure 2.4) montre max pooling sur une fenêtre  $2 \times 2$ .

Pooling: réduire la pile d'images

- Choisissez une taille de fenêtres (généralement 2 ou 3).
- Choisissez un pas (généralement 2).
- Parcourez votre fenêtre à travers vos images filtrées.
- De chaque fenêtre, prenez la valeur maximale.

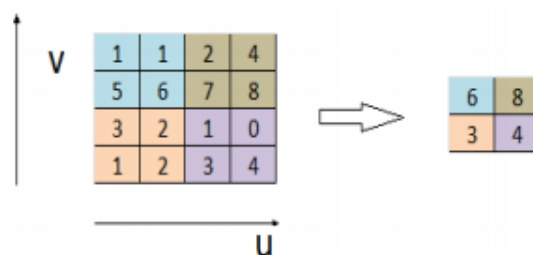


Figure 2. 4: Max pooling

### 3.4. Couches totalement connectées

Enfin, après plusieurs couches de torsion et d'assemblage, Pensée de haut niveau dans un réseau de neurones. Cela se fait par des coutures totalement connectées. Dans les réseaux de neurones convolutifs, chaque couche agit comme un filtre de détection pour la

présence de caractéristiques spécifiques ou de motifs présents dans les données d'origine. Les premières couches d'un convolutif détectent des caractéristiques qui peuvent être reconnues et interprétées relativement facilement. Les couches ultérieures détectent de plus en plus des caractéristiques plus abstraites. La dernière couche du réseau convolutif est capable de faire une classification ultra-spécifique en combinant toutes les caractéristiques spécifiques détectées par les couches précédentes dans les données d'entrée.

## 4. Avantage du CNNs

Le principal avantage des réseaux convolutifs est l'utilisation du poids unique associé aux signaux que tous les neurones entrent dans le même noyau convolutif. Cette méthode réduit l'empreinte mémoire, améliore les performances et permet la stabilité du traitement.

## 5. L'application des réseaux neuronaux convolutifs

Les réseaux neuronaux convolutifs (CNNs) ont de larges applications, notamment :

- La reconnaissance d'image et vidéo.
- Les systèmes de recommandation.
- Le traitement du langage naturel.

L'efficacité des CNN dans la reconnaissance de l'image (Voir Figure 2.16) est l'une des principales raisons pour lesquelles le monde reconnaît le pouvoir de l'apprentissage profond.



Figure 2. 5: Reconnaissance d'une image avec CNN [17]

## 6. Principe général de CNN

Dans les réseaux de neurones convolutifs (Figure 2.5), on retrouve la même base que pour les réseaux de neurones artificiels vue précédemment. La différence notable est due à la présence de couches convolutives et de max-pooling. Celles-ci ont permis grâce à la puissance des ordinateurs actuels de faire apprendre des réseaux très grands et très profonds. Effectivement, sans ce mécanisme convolutif il y aurait eu beaucoup trop de paramètres dans des architectures de telle ampleur

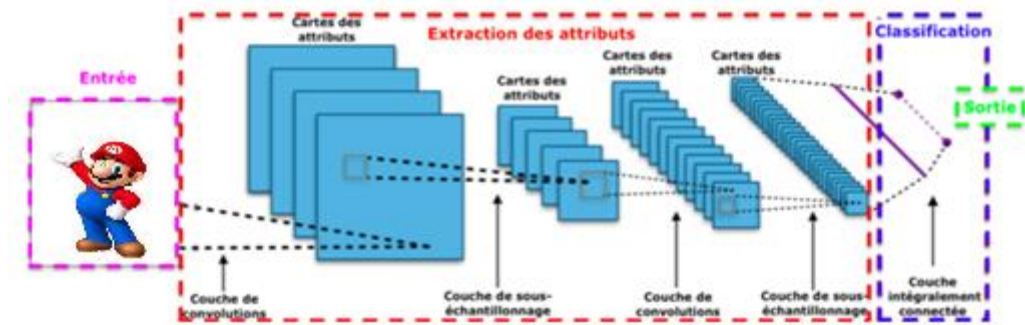


Figure 2. 6: Fonctionnement d'un CNN

Le principe de cette architecture est de séparer le traitement des données de type différent dans des branches différentes du réseau et de fusionner les branches avant d'effectuer la classification. Dans notre figure, la branche représentée dans le cadre bleu traite les images avec les canaux bleu, rouge et vert, et la branche dans le cadre rouge traite l'image de profondeur.

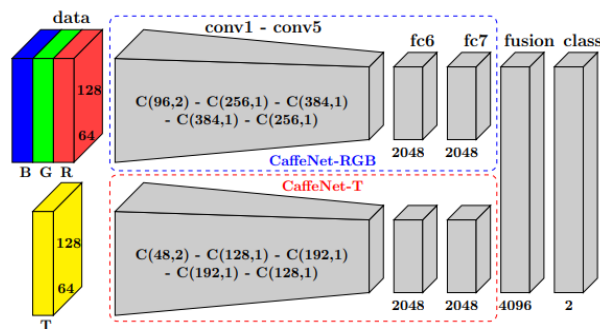


Figure 2. 7 : Schéma de l'architecture de permettant de faire de la fusion de données multi-sources dans un CNN [14]

## 7. Différents types de filtres à convolutions

Les principaux filtres de convolution passe-bas sont le filtre moyennneur et le filtre gaussien. Des filtres non linéaires et qui ne sont donc pas définis par un masque de convolution. On verra dans cette section le filtre médian et le filtre bilatéral qui sont des filtres non linéaires.

## **7.1. Filtres linéaires**

Les possibilités offertes par le système de détection anormaux doivent rester simples en évitant notamment des chemins qui se “croisent”. le travail de système cadrée sera plus qualitative et aboutira à un résultat plus cohérent.

Il faut pour cela, dès le début de la conception du système établir les cas d'utilisation et les liens entre eux. C'est essentiel afin de clarifier les possibilités du système que ce soit pour les concepteurs ou l'utilisateur final.

De plus, les conversations, du bot, sont rarement linéaires ou fluides. Souvent, il y a des interruptions, des ambiguïtés ou encore des malentendus. Il est donc important de tester le bot avant la mise en service.

## **7.2. filtres non linéaires**

Des filtres non linéaires et qui ne sont donc pas définis par un masque de convolution. On verra dans cette section le filtre médian et le filtre bilatéral qui sont des filtres non linéaires. Vous verrez, dans la prochaine partie, les filtres morphologiques qui sont très utiles pour le débruitage des images, et qui sont aussi des filtres non linéaires.

# **8. Apprentissage des réseaux de neurones convolutifs**

## **8.1. Données d'entraînement**

Chaque modèle de machine learning (apprentissage automatique) nécessite une base de données (data set) pour donner lieu à un entraînement. Et plus ces données d'entraînement sont volumineuses et représentatives, plus les probabilités de classer correctement les messages s'améliorent. En général, la base de données est organisée autour des intents et donne à chaque exemple d'intent des exemples des phrases le contenant.

A ce modèle on trouve que l'intention est identique, mais la façon de la formuler peut se diffère et varier d'un point de vue lexical, orthographique, syntaxique, etc. Il est donc primordial de faire la création(ou de trouver) d'un jeu de données qui sont correspondant à ces intents et suffisamment vaste et grande pour donner à ce modèle une bonne robustesse face aux variations de ces volumineux données.

Donc la phase la plus important lorsque nous sommes en train de réaliser ce projet est que nous avons faire la création de ces données manuellement, c'est le défi alors ! , donc grâce à la qualité de ces données, il est possible de bien classer ces intentions.

Et pour avoir un bon système de détection il faut savoir un bon mécanisme et une classification bien étudiée .Il consiste de classer un vidéos ou bien . La bonne détection decette intention dépend du résultat de la classification des intents.

Nous utilisons un dropout de 0,5 dans les deux couches entièrement connectées pour éviter le sur apprentissage. L'apprentissage des paramètres du réseau est réalisé sur 500 cycles ou itérations.

## **8.2. Fonctionnement de détection**

Grâce au Machine Learning et plus précisément le Deep Learning, nous avons établi une base de données de référence associée au système, Lorsque le système détecte que limage, il déclenche un algorithme de classification, dans notre cas, nous utilisons souvent les CNNs : Réseaux de neurones à convolution.

Un premier traitement sur le contenu de l'image à travers l'algorithme de classification est effectué ; ceci va permettre de détecter dans l'image Comportement anormaux des citoyens. Grâce aux réseaux de neurones, nous allons extraire un vecteur représentatif de l'image, ce vecteur sera comparé avec la base de données de référence du système pour afficher le résultat

## **9. Apprentissage automatique**

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 d'émotions [15].

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 (les données), d'en assimiler la nature afin de pouvoir appliquer ce qu'ils ont ainsi appris aux cas futurs.

Il existe plusieurs méthodes en apprentissage automatiques que ce soit pour la régression ou la classification. Pour bien choisir une méthode il faut comprendre les fondements de ces méthodes existantes et de ce qui permet de les distinguer afin de déterminer les modèles qui traiteraient au mieux un problème particulier.



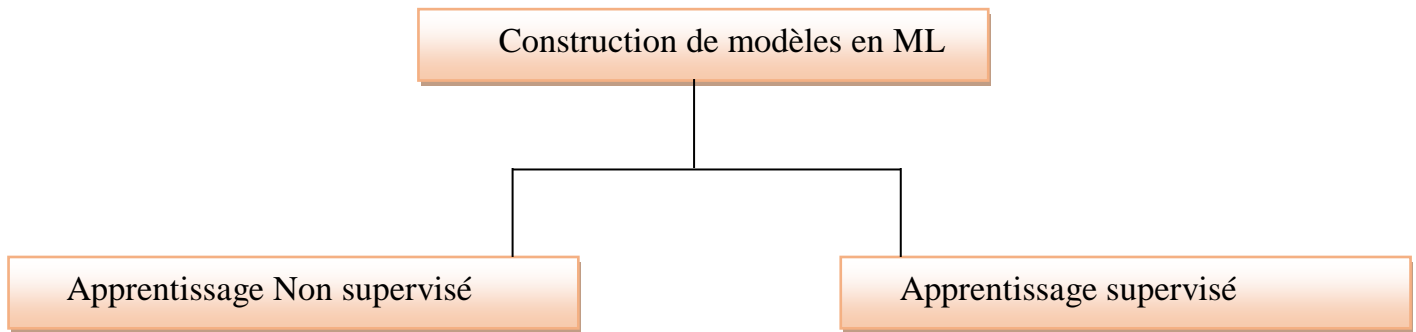


Figure 2. 8: Schéma des différentes techniques issues de l'apprentissage automatique

### 9.1.1. L'apprentissage supervisé

Si les classes sont prédéterminées et les exemples connus, le système apprend à classer selon un modèle de classement ; on parle alors d'apprentissage supervisé.

L'apprentissage supervisé consiste à utiliser un ensemble de données pour prédire des événements futurs statistiquement probables, c'est-à-dire qu'il forme un modèle de prédiction à partir des événements déjà prédits auparavant. Les modèles de ML peuvent être utilisés dans des applications de « prédiction » ou de « classification ». On distingue deux types de problèmes d'apprentissage supervisé :

### 9.1.2. Apprentissage non supervisé

Dans l'apprentissage non supervisé il n'y a pas de valeurs de sortie, il s'agit de trouver des structures cachées à partir d'un ensemble de données qui doivent être regroupé d'où le terme « clustering ». Le but de ce type d'apprentissage est de séparer les données en groupes ou en catégories.

## 10. Conclusion

Dans ce chapitre, nous avons présenté l'architecture des réseaux de neurones convolutifs leur principe et les Différents types de filtres à convolutions, par la suite nous avons expliqué Apprentissage des réseaux de neurones convolutifs et dans le chapitre suivant intitulé « Les réseaux de neurones convolutifs » on va présenter qu'est-ce qu'un réseau de neurones convolutifs.

# **Chapitre3 :**

## **Validation**

## 1. Introduction

Le but de ce chapitre est de présenter notre système réalisé, alors nous commençons par les différents outils du Deep Learning utilisés. Et par la suite, nous présenterons le résultat de modèle proposé.

## 2. Les outils de développement

### 2.1 . L'environnement logiciel

#### 2.1.1 Python

Python est un langage de programmation interprété, orienté objet et de haut niveau, connu pour sa syntaxe claire et lisible. Il offre une grande variété de bibliothèques et de frameworks, ce qui en fait un choix populaire pour le développement d'applications web, scientifiques, et bien d'autres. Python favorise la lisibilité du code grâce à l'indentation obligatoire et dispose d'une vaste communauté de développeurs qui contribuent à son écosystème.



*Figure3.1 : Python*

#### 2.1.2 Google Colab

Google Colab ou Colaboratory est un service cloud, offert par Google (gratuit), basé sur Jupyter Notebook et destiné à la formation et à la recherche dans l'apprentissage automatique. Cette plateforme permet d'entraîner des modèles de Machine Learning directement dans le cloud. Sans donc avoir besoin d'installer quoi que ce soit sur notre ordinateur à l'exception d'un navigateur.

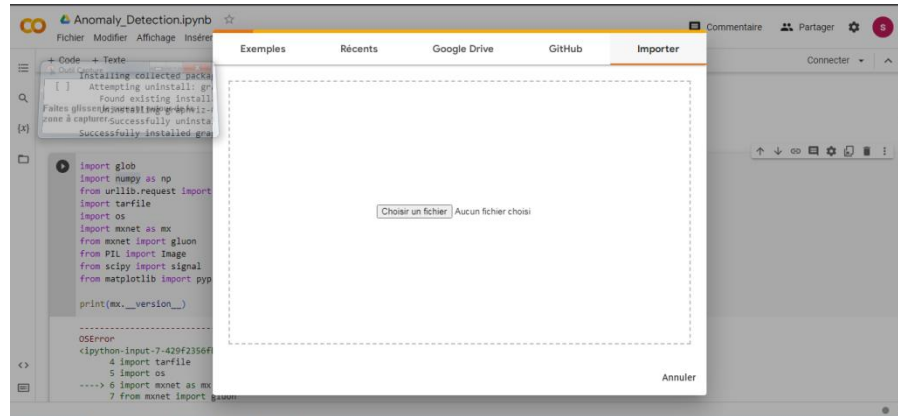


Figure3. 2: google Colab

## 2.2 Les différentes bibliothèques

### 2.2.1 Numpy

Numpy est la bibliothèque de base pour l'informatique scientifique en Python. Il contient des tableaux multidimensionnels et des structures de données matricielles. Elle peut-être utilisée pour effectuer un certain nombre d'opérations mathématiques sur des tableaux tels que des routines trigonométriques, statistiques et algébriques. Par conséquent, la bibliothèque contient un grand nombre de fonctions mathématiques, algébriques et de transformation.

Pour réaliser notre application, nous avons besoin d'installer NumPy comme outil pour stocker et traiter de grandes matrices, ce qui est beaucoup plus efficace que la propre structure de liste imbriquée de Python. On dit que NumPy transforme Python en un MatLabsystem gratuit et plus puissant.



Figure3. 3: bibliothèque Nump

### 2.2.2 TensorFlow

Tensorflow Figure 3.4 est une plateforme (Framework) open source de machine Learning créée par Google, permettant de développer et d'exécuter des applications de Machine Learning et de Deep Learning. Il s'agit d'une boîte à outils permettant de résoudre facilement des problèmes mathématiques extrêmement complexes. il permet aux chercheurs de développer des architectures d'apprentissage expérimentales et de les transformer en logiciels.

Cette bibliothèque open-source est utilisée pour concevoir et former un modèle de deep learning comme CNN ou RNN ou ANN. Aussi, il permet aux développeurs de créer des graphiques de dataflow (dataflow graphs).

Tensorflow permet notamment d'entraîner et d'exécuter des réseaux de neurones pour la classification de chiffres écrits à la main, la reconnaissance d'image, les plongements de mots, les réseaux de neurones récurrents, les modèles séquence-to-séquence pour la traduction automatique, ou encore le traitement naturel du langage. Les applications TensorFlow peuvent s'exécuter sur CPU, GPU ou TPU.

TensorFlow fonctionne en 3 phases :

- Prétraitement des données.
- Construction du modèle.
- Apprentissage et test du modèle



*Figure3. 4: Logo de tensorflow*

### **2.2.3 OpenCV**

OpenCV est une bibliothèque libre, initialement développée par Intel, spécialisée dans le traitement d'images en temps réel. La société de robotique Willow Garage, puis la société ItSeez se sont succédé au support de cette bibliothèque.



Figure 3. 5: bibliothèque OpenCV

## 2.3 L'environnement matériel

Le Deep Learning est un domaine avec des exigences en calculs intenses et la disponibilité des ressources (surtout en GPU) dédiés à cette tâche vont fondamentalement influencer sur l'expérience de l'utilisateur car sans ses ressources, il faudra trop de temps pour apprendre de ses erreurs ce qui peut être décourageant.

Les expérimentations ont tous été effectuées sur une machine qui offre des performances acceptables dont voici les caractéristiques :

*Tableau 1: Architecture de machine*

CPU	Intel Core i5-4440 (3.1 GHz)
GPU	MSI GEFORCE GTX 770 TWIN FROZR GAMING OC 2GB
RAM	8GB

## 3. Base d'apprentissage

1. Actuellement, le plus grand ensemble de données sur les actions humaines.
2. Se compose de 101 classes d'action réparties sur 13 000 clips et 27 heures de données vidéo Figure (3.7 & 3.8).
3. Se compose de vidéos réalistes téléchargées par les utilisateurs, la plupart d'entre elles provenant de Youtube au hasard.

4. Actuellement, l'ensemble de données d'actions le plus difficile en raison de son grand nombre de classes, d'un grand nombre de clips et de la nature non contrainte de ces clips
5. Vise à encourager d'autres recherches sur la reconnaissance d'action en apprenant et en explorant de nouvelles catégories d'action réalistes

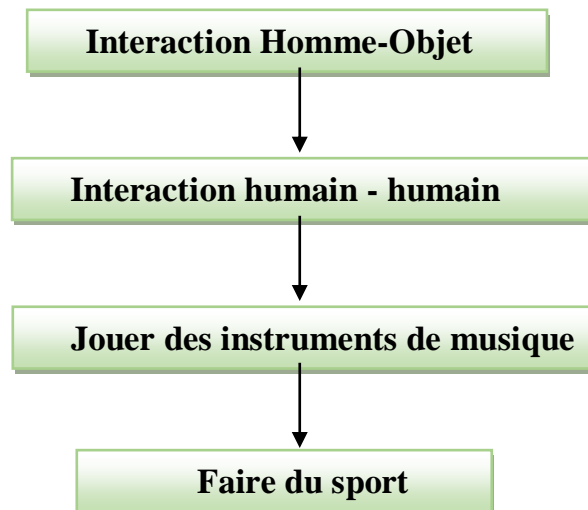


Figure3. 6: Catégories de clips vidéo dans l'ensemble de données UCF101



Figure3. 7: - QUELQUES ILLUSTRATIONS

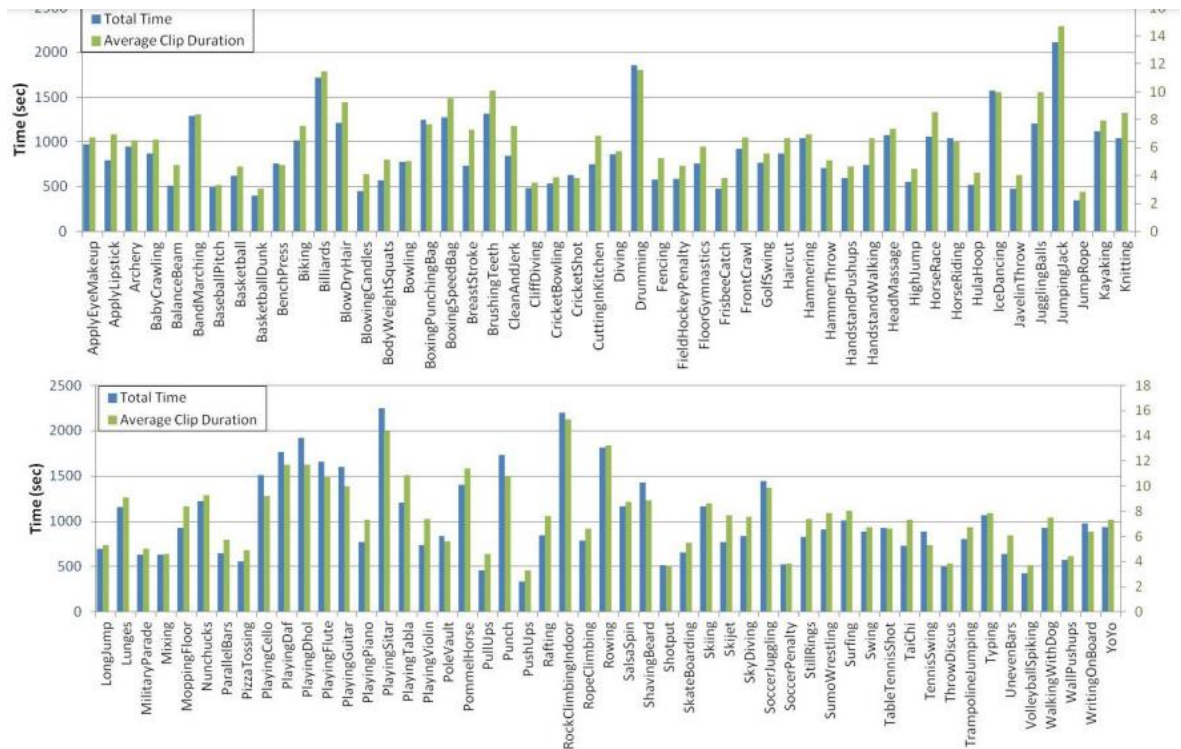


Figure3. 8: Catégories vs durée de la vidéo dans chaque catégorie

Certains des clips vidéo présents dans la liste de fichiers vidéo UCF101 comme représente la figure ci-dessous

## Index of /THUMOS14/UCF101/UCF101













 [ICO]	Name	Last modified	Size	Description
 [PARENTDIR]	<a href="#">Parent Directory</a>		-	
 [VID]	<a href="#">v_ApplyEyeMakeup_g01_c01.avi</a>	2012-10-01 08:15	288K	
 [VID]	<a href="#">v_ApplyEyeMakeup_g01_c02.avi</a>	2012-10-01 08:15	205K	
 [VID]	<a href="#">v_ApplyEyeMakeup_g01_c03.avi</a>	2012-10-01 08:15	408K	
 [VID]	<a href="#">v_ApplyEyeMakeup_g01_c04.avi</a>	2012-10-01 08:15	400K	
 [VID]	<a href="#">v_ApplyEyeMakeup_g01_c05.avi</a>	2012-10-01 08:15	461K	
 [VID]	<a href="#">v_ApplyEyeMakeup_g01_c06.avi</a>	2012-10-01 08:15	213K	
 [VID]	<a href="#">v_ApplyEyeMakeup_g02_c01.avi</a>	2012-10-01 08:21	249K	
 [VID]	<a href="#">v_ApplyEyeMakeup_g02_c02.avi</a>	2012-10-01 08:21	243K	
 [VID]	<a href="#">v_ApplyEyeMakeup_g02_c03.avi</a>	2012-10-01 08:21	366K	
 [VID]	<a href="#">v_ApplyEyeMakeup_g02_c04.avi</a>	2012-10-01 08:22	252K	

Figure3. 9: liste de fichiers vidéo UCF101

Nous avons créé un fichier texte nommé étiquettes dans ce fichier, nous avons créé des étiquettes qui se trouvent dans le jeu de données Kinetics la figure ci-dessous montre le fichier étiquettes :



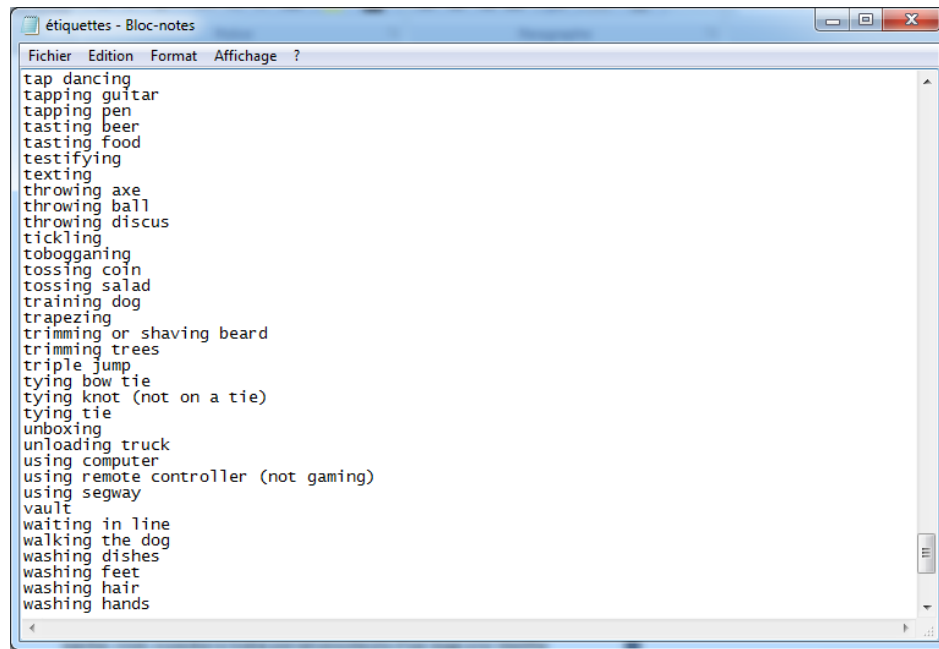


Figure3. 10: fichier étiquettes

#### 4. Model proposé

La reconnaissance des actions humaines est un problème très intéressant où l'objectif principal est d'analyser une vidéo et de reconnaître les actions effectuées par la personne dans la vidéo. Certaines actions simples comme se tenir debout peuvent être identifiées en utilisant une seule image, mais des actions plus complexes telles que marcher, courir, se pencher ou tomber peuvent nécessiter plus d'une image pour identifier correctement l'action. Le but de notre modèle est de prendre une entrée vidéo et d'identifier l'action (Fig 3.7) qui est effectuée par la personne dans la vidéo.



*Figure3. 11: QUELQUES ILLUSTRATIONS*

#### **4.1. Reconnaissance et détection d'actions humaines**

De nos jours, les industries manufacturières font appel à la fois à des humains et à des robots d'une grande complexité pour améliorer le rendement et la qualité de leurs produits [16]. Les capacités de perception et de décision des robots évoluant progressivement dans des environnements plus peuplés d'humains et notamment vers des situations nécessitent une collaboration humain-robot dans les environnements de travail industriels. Par conséquent, outre le fait que les robots doivent fonctionner de manière sûre et compétente, ils doivent également être capables d'interagir et de communiquer avec les humains.

Afin d'aider l'humain à atteindre un ensemble d'objectifs, il appartient au robot d'estimer son intention et d'agir en conséquence. Une personne peut communiquer son intention soit délibérément par une communication explicite, soit implicitement par des actions. Les manières de communiquer l'intention telles qu'elles sont élaborées dans [17], sont présentées dans le Tableau 2.

*Tableau 2: Principaux moyens de communiquer les intentions*

<b>Communication d'intention</b>	
<b>Discours</b>	Informations explicites
	Émotions
<b>Geste</b>	Tête/yeux
	Gestes communicatif
<b>Action</b>	Gestes de manipulation
	Exécution proactive des tâches
<b>Signal haptique</b>	Force/Couple
	Angles/Orientation
<b>Signal physiologique</b>	Approbation

La reconnaissance d'actions a attiré l'attention des chercheurs ces dernières années. Les méthodes actuelles sont principalement conçues pour la reconnaissance d'actions segmentées, c'est-à-dire que le type d'action est reconnu une fois que la séquence d'action entière est observée. Cette méthode est également appelée reconnaissance d'action hors ligne. Néanmoins, il serait souhaitable que la reconnaissance d'action se fasse pendant l'exécution de l'action ce qui est nécessaire pour des applications en temps réel. Les travaux récents s'intéressent plutôt à ce type de reconnaissance d'actions appelé aussi reconnaissance en ligne vu leurs larges champs d'applications dans plusieurs domaines comme la surveillance vidéo, les interactions homme-machine, etc.

Les mouvements humains sont classés en trois types différents : les gestes, les actions et les activités

#### **4.1.1. Gestes**

Les gestes sont les éléments élémentaires décrivant le mouvement significatif d'une personne, et sont perceptibles visuellement par les humains et qui sont facilement annotés. Ce type de mouvement n'implique qu'une partie du corps humain, comme le bras, la tête ou la jambe, etc. En général, les gestes sont très brefs et ne durent que quelques secondes. De plus, ils sont réalisés sans l'utilisation d'un objet quelconque. Exemples de gestes : agiter un bras, lever un pied, lever un bras, etc.

#### **4.1.2. Actions**

Les actions sont définies comme une séquence de gestes organisée temporellement. Ainsi, une telle variété de mouvements peut inclure le mouvement de plusieurs parties du

corps, contrairement aux gestes. La durée des actions est évidemment plus importante que la durée du geste et peut durer jusqu'à une minute. Les actions peuvent impliquer un objet, mais celui-ci doit être présent du début à la fin de l'action. Exemples d'actions : sauter, marcher, courir, nager, balancer une balle de golf, lancer un objet, etc.

#### **4.1.3. Activités**

Les activités représentent le plus haut niveau de mouvement et sont composées d'une séquence d'actions. Ils impliquent des interactions avec des objets, ce qui les rend plus complexes à comprendre et à reconnaître. De plus, la connaissance du contexte est parfois nécessaire pour évaluer correctement l'activité. Leur durée est relativement importante par rapport aux actions et aux gestes (de l'ordre de quelques minutes). Exemples d'activités : parler au téléphone, manger une pomme, jouer à des jeux vidéo, etc.

### **4.2. Étapes de création du modèle proposé**

Nous avons présenté avec des organigrammes le fonctionnement du modèle proposé comme représente la figure ci-dessous

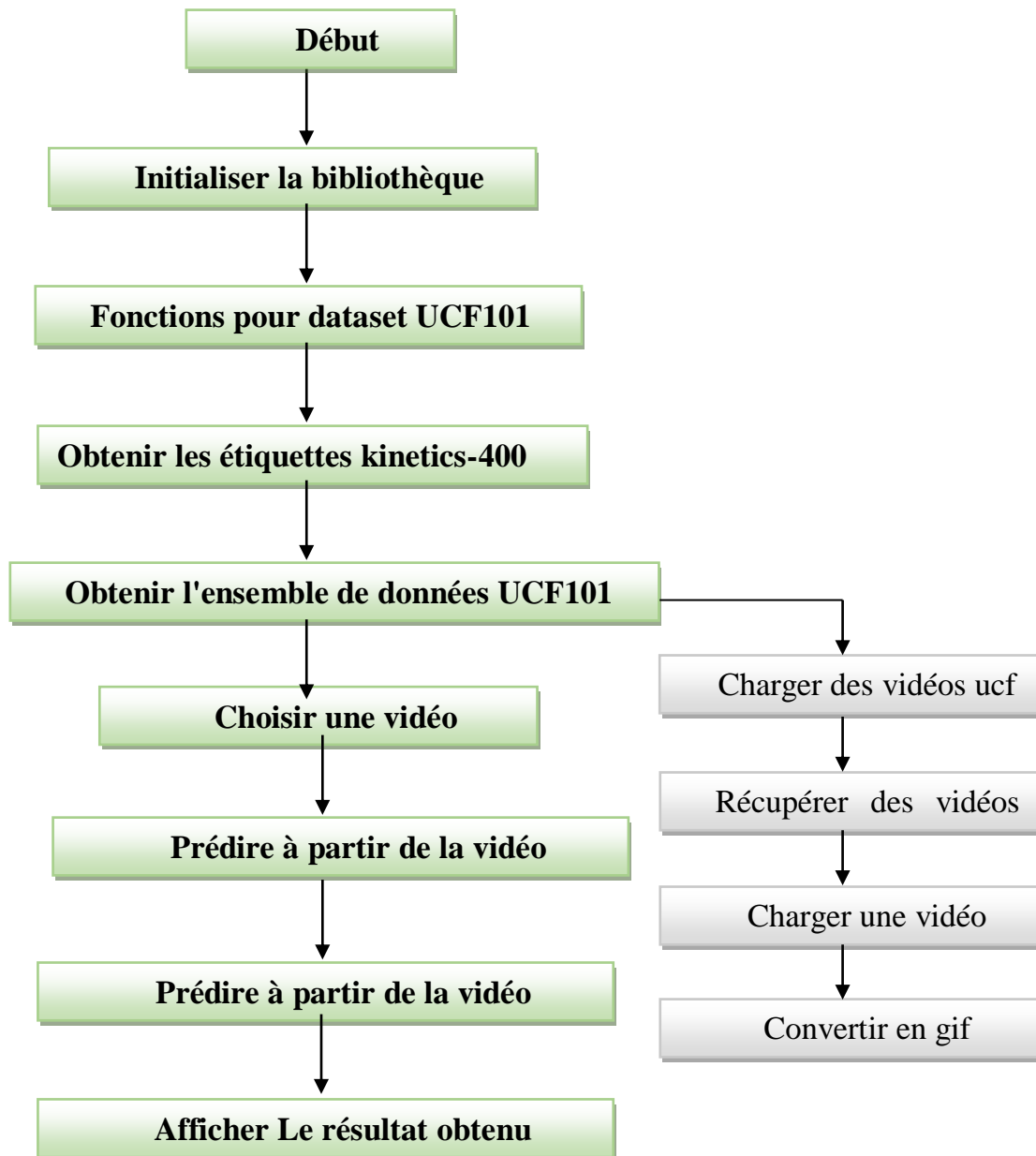


Figure3. 12: organigrammes le fonctionnement du Model

## 5. Résultats obtenus et discussion

Dans cette partie nous mettons en œuvre les différents résultats. Dans les images suivantes le programme fait les prédictions à l'aide du modèle préformé utilisé. Après le calcul des prédictions : le système affiche la résultat.

J'ai utilisé plusieurs vidéos pour tester les différentes anomalies comme montre les figures ci-dessous :

## 5.1. Teste1 :

Sur la photo (Fig. 3.13), un joueur de football tire au but, ce qui signifie que le résultat obtenu est correct. Le comportement peut être considéré comme normal.

Dans notre travail, nous visons toujours à obtenir un excellent résultat de détection d'action, D'après les résultats obtenus (Fig. 3.13), nous remarquons que notre solution obtient une précision plus de 98,90 %.

Top 5 actions:  
shooting goal (soccer): 98.90%  
playing tennis : 0.44%  
hurling (sport) : 0.27%  
catching or throwing frisbee: 0.11%  
playing kickball : 0.06%



Figure3. 13: Test1 de validation

## 5.2. Teste2 :

Sur la photo (Fig. 3.14), l'enseignant écrit au tableau, ce qui signifie que le résultat obtenu est correct. Le comportement peut être considéré comme normal, D'après les résultats obtenus (Fig. 3.14), nous remarquons que notre solution obtient une précision plus de 99,99 %.

Top 5 actions:  
writing : 99.99%  
shaking hands : 0.01%  
spray painting : 0.00%  
brush painting : 0.00%  
drawing : 0.00%

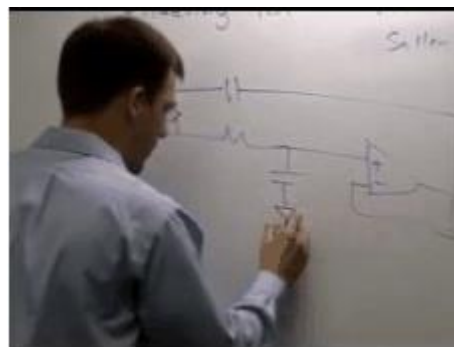


Figure3. 14: Test2 de validation

### 5.3. Teste3 :

D'après les résultats obtenus (Fig. 3.15), nous remarquons que notre solution obtient une précision plus de 86,20 %.

```
Top 5 actions:
stretching arm      : 86.20%
exercising with an exercise ball: 4.27%
exercising arm      : 3.36%
stretching leg      : 1.10%
applying cream      : 1.03%
```



*Figure3. 15: Test3 de validation*

## 6. Conclusion

Dans ce chapitre nous avons décrit l'environnement de travail : L'éditeur de code, le langage les différentes bibliothèques choisies pour traiter les images. En deuxième partie, nous avons établi la reconnaissance d'anomalies.

# *Conclusion Générale & perspectives*

Le travail que nous avons présenté s'inscrit dans le cadre de l'apprentissage automatique et s'intéresse essentiellement au problème de la classification supervisée. Notre objectif était de détecter des comportements anormaux des citoyens en effectuant une classification de ces dernières en deux catégories : normales et anomalies.

Au début de ce travail, nous avons conçu un modèle de détection Deep Learning pour détecter une base de données "UCF101" composée de 101 classes d'action réparties sur 13 000 clips et 27 heures. Par la suite, nous avons exploité son architecture pour l'appliquer à notre modèle de détection de l'automatique des comportements en ajustant quelques paramètres.

Comme perspectives, ce travail pourra être exploité pour plusieurs divers applications, telles que, la détection des maladies de la peau et le cancer etc., et même pour la classification des images satellitaires des terrains, etc. On peut aussi travailler sur les comportements anormaux à travers certains comportements comme la violence et le vol, etc.

En fin, Ce travail a été très bénéfique pour nous car il nous a permis d'apprendre beaucoup de nouvelles notions que nous n'avons pas apprises durant notre cursus de formation, telles que les réseaux de neurones artificiels, l'apprentissage profond avec le Deep-Learning, la détection automatique de vidéos et des images avec les techniques de l'intelligence artificielle (computer vision, TALN...), la programmation avec le logiciel Python, etc. et nous espérons qu'il aidera les futurs étudiants et chercheurs désirant travailler sur ce sujet.



## *Références/bibliographiques*

- [1]. <https://blog.smart-tribune.com/fr/>
- [2]. Adrian PETRICEVIC L'intelligence artificielle : une solution à la prise de décision quotidienne ?
- [3]. [https://www.passeportsante.net/fr/Maux/Problemes/Fiche.aspx?doc=trouble-comportement-causes-symptomes-traitements#:~:text=Ils%20peuvent%20se%20traduire%20par,%E2%80%A6\)%20ou%20des%20phobies%20%C2%BB.](https://www.passeportsante.net/fr/Maux/Problemes/Fiche.aspx?doc=trouble-comportement-causes-symptomes-traitements#:~:text=Ils%20peuvent%20se%20traduire%20par,%E2%80%A6)%20ou%20des%20phobies%20%C2%BB.)
- [4]. <https://kobia.fr/quest-ce-que-le-machine-learning/>
- [5]. <https://img.passeportsante.net/1200x675/2019-07-15/i90006-troubles-alimentaires.webp>
- [6]. <https://trustmyscience.com/wp-content/uploads/2021/04/chercheurs-identifient-mecanisme-responsable-trouble-bipolaire-couv-1.jpeg>
- [7]. <https://www.javatpoint.com/k-means-clustering-algorithm-in-machine-learning>
- [8]. <https://geekflare.com/fr/support-vector-machine/>
- [9]. Géron, A. (2017). Hands-on machine learning with Scikit-Learn and TensorFlow : concepts, tools, and techniques to build intelligent systems. " O'Reilly Media, Inc."
- [10]. <https://www.math.univ-toulouse.fr/~besse/Wikistat/pdf/st-m-app-rn.pdf>
- [11]. <https://blog.octo.com/classification-dimages-les-reseaux-de-neurones-convolutifs-en-toute-simplicite/>
- [12]. <https://towardsdatascience.com/covolutional-neural-networkcb0883dd6529?fbclid=IwAR0UWoPkFYTEAqFitcR4fuuQUcNBvxV8ig1oJGZ3EhbRypu8Qf9pk9C Xdy4> Consulté le 28/05/2021

- [13]. Classification d'images : les réseaux de neurones convolutifs en toute simplicité. url : <https://blog.octo.com/classification-dimages-les-reseaux-de-neurones-convolutifs-en-toute-simplicité/>.
- [14]. Jörg WAGNER, Volker FISCHER, Michael HERMAN et Sven BEHNKE : Multispectral pedestrian detection using deep fusion convolutional neural networks. In European Symposium on Artificial Neural Networks (ESANN), Bruges, Belgium, April 2016.
- [15]. P. Vincent, « Modèles à noyaux à structure locale », Thèse de Phd en informatique, Université de Montréal, 2003.
- [16]. El Zaatari, Shirine, Mohamed Marei, W. Li, et Zahid Usman. «Cobot programming for collaborative industrial tasks: An overview.» Robotics and Autonomous Systems 116 (6 2019): 162-180.
- [17]. Bauer, Andrea Maria, Dirk Wollherr, et Martin Buss. «Human-Robot Collaboration: a Survey.» Int. J. Humanoid Robotics 5 (2008): 47-66.