

الجمهورية الجزائرية الديمقراطية الشعبية

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

وزارة التعليم العالي والبحث العلمي

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

جامعة سعيدة – د. الطاهر مولاي –

Université Saïda – Dr. Tahar Moulay –
Faculté de Technologie



MEMOIRE

Présenté pour l'obtention du **Diplôme de MASTER en Télécommunications**

Spécialité : Réseaux et Télécommunications

Par : Mlle. Meskine Nadjat
Mlle. Merabti Amira

Détection de la Présence de Drones par les Techniques d'Intelligence Artificielle

Soutenu, le 22 /06/ 2024, devant le jury composé de :

Mr. Bouyeddou Benamar	MCA	Président
Mlle. Ouessai Asmâa	MAB	Rapporteur
Mr. Tami Abdelkader	MCA	Examineur

2023 / 2024

Remerciements

Nous tenons à exprimer notre profonde gratitude à toutes les personnes qui ont contribué à la réalisation de ce projet de fin d'études.

Un merci spécial à notre encadrante de PFE, Dr.Ouessai Asmâa, pour ses précieux conseils, son soutien inébranlable et son expertise.

Nous remercions Dr Bouyeddou Benamar et Dr Tami Abdelkader pour avoir accepté de faire partie de ce jury et pour le temps qu'ils ont consacré à évaluer notre travail.

Nous exprimons notre gratitude au Dr. Boudkhal Abdelhakim, chef du département de Télécommunications, pour avoir accepté notre invitation et donné d'intérêt à notre travail.

Nous sommes également reconnaissants envers l'équipe pédagogique de notre université Dr. Moulay Tahar, ainsi que nos camarades de classe pour leurs esprits de collaboration.

Enfin, nous souhaitons exprimer notre gratitude à nos familles et à nos amis pour leur soutien moral, leur patience et leurs encouragements durant les moments les plus intenses de ce projet.

Ce PFE a été une aventure enrichissante et formatrice, et c'est grâce à l'implication de chacun que nous avons pu mener à bien cette étape importante de notre parcours universitaire.

Dédicace

Je m'incline devant dieu tout puissant qui m'a ouvert la porte du savoir et m'a aidé à la franchir.

Je dédie ce Modest travail :

A mes chers parents et mes grands-parents source d'affectation de courage et d'inspiration qui ont autant sacrifié pour me voir atteindre ce jour.

Une dédicace spéciale à notre encadrante de PFE, Mlle. Ouessai Asmâa

A toute la famille Meskine et Merabti

A

Mes sœurs

Mes frères

Mes nièces

Mes neveux

Tous mes amies

Nadjet

Dédicace

Je m'incline devant dieu tout puissant qui m'a ouvert la porte du savoir et m'a aidé à la franchir.

Je dédie ce Modest travail :

A mes chers parents et mes grands-parents source d'affectation de courage et d'inspiration qui ont autant sacrifié pour me voir atteindre ce jour.

Une dédicace spéciale à notre encadrante de PFE, Mlle. Ouessai Asmâa

A toute la famille Meskine et Merabti

A

Mes sœurs

Mes frères

Mes nièces

Mes neveux

Tous mes amies

Amira

Résumé

Les véhicules aériens sans pilote (UAV ou drones) ont gagné en popularité dans diverses utilisations civiles, commerciales et militaires ; cependant, ils représentent un défi pour la sécurité humaine. Il a été remarqué que les drones peuvent être utilisés à mauvais escient pour des actions nuisibles, causant des dommages à des personnes spécifiques et au grand public. Les drones ont aussi soulevé des problèmes de confidentialité lorsqu'ils sont utilisés pour espionner des personnes dans leurs espaces privés et enregistrer des vidéos. Ce travail propose de fournir une étude sur la détection de la présence de drones en utilisant des données sonores et visuelles, et en comparant des approches d'apprentissage automatique classiques et d'apprentissage profond. Les résultats obtenus sont prometteurs et montrent l'efficacité des modèles proposés notamment dans le cas d'utilisation des données audios.

Mot clé : Drone, MFCC, LPC, ZCR, SVM, KNN, CNN

Abstract

Unmanned Aerial Vehicles (UAVs or drones) have gained popularity in various civil, commercial, and military applications; however, they pose a challenge to human security. It has been noted that drones can be misused for harmful actions, causing damage to specific individuals and the general public. Drones have also raised privacy concerns when used to spy on people in their private spaces and record videos. This work aims to provide a study on the detection of the presence of drones using audio and visual data, and by comparing traditional machine learning approaches and deep learning approaches. The obtained results are promising and demonstrate the effectiveness of the proposed models, especially in the case of using audio data.

Key word: Drone, MFCC, LPC, ZCR, SVM, KNN, CNN

المخلص

اكتسبت المركبات الجوية غير المأهولة (الطائرات بدون طيار) شعبية في مختلف التطبيقات المدنية والتجارية والعسكرية؛ ومع ذلك، فإنها تشكل تحدياً للأمن البشري. وقد لوحظ أنه يمكن استخدام الطائرات بدون طيار في أعمال ضارة، مما يتسبب في أضرار لأفراد معينين وعامة الناس. أثارت الطائرات بدون طيار أيضاً مخاوف تتعلق بالخصوصية عند استخدامها للتجسس على الأشخاص في مساحاتهم الخاصة وتسجيل مقاطع الفيديو. يهدف هذا العمل إلى تقديم دراسة حول الكشف عن وجود الطائرات بدون طيار باستخدام البيانات السمعية والبصرية، ومن خلال مقارنة مناهج التعلم الآلي التقليدية ومناهج التعلم العميق. النتائج التي تم الحصول عليها واثبتت فعالية النماذج المقترحة، خاصة في حالة استخدام البيانات الصوتية.

Sommaire

Introduction générale.....	1
Chapitre 1 : Introduction aux drones et aux systèmes de détection.....	3
1.1. Introduction	3
1.2. Histoire du développement des drones	3
1.3. Définition d'un drone	5
1.4. Les composants principaux du drone	6
1.5. Classification des drones selon le mode de vol	7
1.6. Fonctionnement du drone	8
1.7. Comment ça vole ?.....	8
1.8. Les caractéristiques les plus remarquables des drones	8
1.9. Les domaines d'application et d'utilisation des drone	9
1.10. La détection des drones	10
1.11. Les Systèmes de détection de drones	10
1.11.1. Détection radar	11
1.11.2. Détection des drones par données visuelles	11
1.11.3. Détection par signaux acoustiques	11
1.11.4. Détection des radio fréquences	12
1.12. Les avantages et les inconvénients des systèmes de détection	12
1.13. Conclusion	13
Chapitre 2 : La détection de la présence de drone par les techniques d'apprentissage machine.....	14
2.1 Introduction	14
2.2 Définitions d'un signal audio	14
2.3 Architecture du système de reconnaissance de signal audio de drone	14
2.3.1 Acquisition et traitement du signal audio	14
2.3.2 Extraction de paramètres	15
2.3.2.1 Les coefficients cepstraux de fréquence Mel (MFCC)	15
2.3.2.2 Les coefficients cepstraux Gammatones (GTCC)	17
2.3.2.3 Le codage à prédiction linéaire (LPC)	18
2.3.2.4 Taux de passage par ZCR	19
2.3.3 La classification des signaux audios	20
2.3.3.1 L'intelligence artificielle	20

2.3.3.2 L'apprentissage machine (Machine Learning)	20
2.3.3.2 La classification « KNN » « SVM »	21
2.4 Conclusion	25
Chapitre 3 : La détection de la présence de drone par Apprentissage profond.....	26
3.1 Introduction.....	26
3.2 Définition de l'apprentissage profond (Deep Learning DL)	26
3.3 Les réseaux de neurones.....	27
3.3.1 Définition des réseaux de neurones artificiels	27
3.3.2 Les neurones biologiques.....	27
3.3.3 Les neurones formels.....	27
3.3.3.1 Fonctions d'activation.....	28
3.4 Les trois types de couches d'un réseau de neurones	29
3.5 Architecture de réseau de neurones	29
3.5.1 Les réseaux de neurones non bouclés	29
3.5.2 Les réseaux de neurones bouclés	31
3.6 Les applications des réseaux de neurone	31
3.7 Les avantages et les inconvénients des réseaux de neurones	32
3.8 Procédure de développement d'un réseau de neurones.....	32
3.9 Architecture des réseaux de neurones profond.....	34
3.9.1. Les réseaux de neurones convolutifs (CNNs).....	34
3.9.2. Architecture de réseau neuronal convolutif 1D	Error! Bookmark not defined.
3.9.3 Architecture de réseau neurone convolutif 2D.....	36
3.9.4 Comparaison entre le 1D et le 2D.....	38
3.10 La reconnaissance des signaux audios	39
3.10.1 Etapes de la reconnaissance d'audios	39
3.10.2 Les Application de la reconnaissance d'audios	40
3.11 Reconnaissances d'objets dans une image	40
3.11.1 Classification d'image.....	40
3.11.2 Localisation d'objet.....	40
3.11.3 Fonctionnement de la reconnaissance d'image	40
3.11.3.1 Les pixels d'une image	40
3.11.3.2. Segmentation d'image	40
3.12 L'intérêt de l'utilisation des réseaux de neurones artificiels pour un drone	41

3.13 Conclusion	42
Chapitre 4 : Résultats et Discussions.....	43
4.1. Introduction.....	43
4.2. Reconnaissance de la présence de drone à partir des signaux audios	43
4.2.1. Base de données	43
4.2.2 Environnement de travail.....	44
4.2.3 Langage de programmation	44
4.2.4 Bibliothèques utilisées.....	44
4.2.5. Classification par méthodes classiques (KNN, SVM).....	45
4.2.5.1 Extraction des paramètres à partir des signaux audios	45
a. Visualisation de MFCC	45
b. Visualisation de LPC	46
c. Visualisation de ZCR	47
4.2.5.2 Classification	48
4.2.5.2.1 Les Métriques utilisées	48
4.2.5.3 Résultats de classification.....	49
4.2.6. Classification des signaux audio par Deep Learning (CNN 1D).....	51
4.3. Reconnaissance de la présence de drone à partir d'images	53
4.3.1 Prétraitement des données	53
4.3.2 Modèle CNN utilisé.....	54
4.3.3 Évaluation du modèle	55
4.3.4 Interprétation de la Matrice de Confusion	55
4.3.5 Visualisation des Résultats	55
4.3.5.1 Analyse du Graphique de Précision	55
3.5.2 Analyse du Graphique de Perte.....	56
4.4. Conclusion	57
Conclusion générale	58
Bibliographie et Webographie	60

Liste des figures

Figure 1: l'organisation de notre mémoire	2
Figure 1.1: Hewitt-Sperry Automatic Airplane	3
Figure 1.2: Aéroplane Détable de max Boucher	3
Figure 1.3: Avion Cible Denny qui popularisa le terme de drone	4
Figure 1.4: Unmanned US Navy Grumman F6F – 5K Hellcat drone 1946 © US Navy Naval Aviation News January 1981	4
Figure 1.5: US Navy Culver TD2C-1 target drone © US Navy National Museum of Naval Aviation photo n 2012 075 001	5
Figure 1.6: châssis comporte 4 bras	6
Figure 1.7: (a) Hélices (b) Batterie (c) Contrôleurs ESC	6
Figure 1.8: une carte électronique	6
Figure 1.9: Les drones à voilure fixe	7
Figure 1.10: drone à aile battante	7
Figure 1.11: Système de détection de drones avec Radar	11
Figure 1.12: Schéma synoptique de la méthode de détection acoustique	11
Figure 1.13: Système de détection de drones par Radio Fréquences	12
Figure 2.1: Schéma global du système de reconnaissance de drone par signaux audios	14
Figure 2.2: Prétraitement acoustique du signal audio	14
Figure 2.3: schéma synoptique des étapes d'extraction des coefficients	16
Figure 2.4: Différentes étapes de l'analyse cepstrale	17
Figure 2.5: Schéma synoptique des étapes d'extraction des coefficients GTCC	17
Figure 2.6: Diagramme de Méthode LPC	18
Figure 2.7: ZCR d'un signal de parole de 1s	20
Figure 2.8: Types d'apprentissage automatisé (Machine Learning)	21
Figure 2.9: type de SVM	22
Figure 2.10: Exemple d'un hyperplan séparateur	22
Figure 2.11: Exemple de vecteurs de support	23
Figure 2.12: Classifieur SVM (dans le cas non-linéaire)	23
Figure 2.13: illustration d'un SVM non-linéaire (kernel trick)	24
Figure 3.1: La relation entre l'intelligence artificielle le machine Learning et le deep Learning	26

Figure 3.2: Modèle d'un neurone biologique	27
Figure 3.3: Modèle d'un neurone artificiel	27
Figure 3.4: schéma simple de 3 types de couches d'un réseau de neurones	29
Figure 3.5: les types de réseaux de neurones artificiels	29
Figure 3.6: Schéma d'un réseau de neurones monocouche.....	30
Figure 3.7: Schéma d'un réseau de neurones non bouclé (Perceptron multicouches	30
Figure 3.8: Schéma de réseau de neurones bouclé	31
Figure 3.9: Architecture de réseau neurone convolutif 1D	34
Figure 3.10: Architecture de réseau neurone convolutif 2D	35
Figure 3.11: Fonctionnement d'un réseau neuronal à 2 couches cachées	37
Figure 3.12: exemple de détection d'image et segmentation	41
Figure 3.13: Présentation des taches de vision par ordinateur pour la reconnaissance d'objet	41
Figure 4.1: représentation visuelle de quelques signaux audio de drone et non drone de la base utilisée	43
Figure 4.2: Représentation visuelle des coefficients MFCC pour drone	45
Figure 4.3: Représentation visuelle des coefficients MFCC pour un son d'animale	46
Figure 4.4: Représentation visuelle des coefficients LPC pour audio de drone.....	46
Figure 4.5: Représentation visuelle des coefficients LPC pour son de marteau	47
Figure 4.6: Représentation visuelle des coefficients ZCR pour drone.....	47
Figure 4.7: Représentation visuelle des coefficients ZCR pour un son d'animal	48
Figure 4.8: Matrice de confusion	49
Figure 4.9: Représentation des matrices de confusion pour SVM et KNN	49
Figure 4.10: la précision de KNN et SVM en fonction du nombre de coefficients	50
Figure 4.11: Architecture de modèle CNN 1D proposée	51
Figure 4.12: les résultats de CNN 1D en fonction du Loss et taux de classification pour différents epochs	51
Figure 4.13: la matrice de confusion de model CNN 1D.....	52
Figure 4.14: Comparaison entre les méthodes utilisées pour la détection des drones en utilisant les signaux audios.....	52
Figure 4.15: Images de drone sous différents angles	53
Figure 4.16: Images d'oiseaux sous différents angles.....	54
Figure 4.17: l'architecture de CNN 2D proposé	54
Figure 4.18: La matrice de confusion de model CNN 2D.....	55

Figure 4.19: la courbe d'accuracy de l'entraînement et de la validation du modèle au cours de 100 epochs.....	56
Figure 4.20: la courbe de Loss de l'entraînement et de la validation du modèle au cours de 100 epochs.....	57

Liste des tableaux

Table 1.1: Les avantages et les inconvénients des différents systèmes de détection	12
Table 3.1: Différents fonctions d'activation.....	28
Table 3.2: les avantages et les inconvénients des réseaux de neurones	32
Table 3.3: comparaison entre CNN 1D et CNN 2D	38
Table 4.1: Description de la base audio utilisée	43
Table 4.2: Comparaison des Performances des Méthodes KNN et SVM avec Différentes Caractéristiques Audio	49
Table 4.3: Comparaison des algorithmes de détection de drones en utilisant le MFCC pour les deux classes	50
Table 4.4: Comparaison des algorithmes KNN et SVM pour la classification avec MFCC ..	50
Table 4.5: Résultats de classification de drones par CNN1D.....	52
Table 4.6: La nature des classes de « drones » et « bird ».....	54
Table 4.7: Résultats de classification de drone par CNN2D	55

Liste des Abréviations

CNN: Convolutional Neural Network

DL: Deep Learning

ESC: Electronique speed controller

GPS: Global Positioning System Gammatones

GTCC: Gammatones Cepstral Coefficients

IA : Intelligence Artificielle

KNN: K-Nearst Neighbors

LPC: Linear Predictive Coding

MDS : Signature Micro Doppler

MFCC: Mel Frequency Cepstrum Coefficients

MLP: Multi-Layer Perception

RBF: Radial Basis Function

RCS : Radar Cross Section

RF : Fréquences Radio

RNN: Recurrent Neural Network

SDDD : Système de Défense Des Drones

SVM : Support Vectors Machines

TSF : Télégraphie Son Fil

UAV : Unmanned Air Vehicul

Introduction générale

Introduction générale

Les drones, aussi connus sous le nom de véhicules aériens sans pilote (UAV), sont des appareils volants qui peuvent être pilotés à distance ou fonctionner de manière autonome grâce à des logiciels intégrés. Originellement développés pour des missions militaires où l'envoi d'un pilote humain serait risqué ou impraticable, les drones ont rapidement gagné en popularité et en applicabilité dans de nombreux autres domaines.

De nos jours, la révolution rapide de la technologie des drones provoque une augmentation illustrative de leurs utilisations dans de nombreux domaines tel que la surveillance, la livraison, l'agriculture, etc. Cependant, c'est bien fait peuvent créer des risques potentiels sur la sécurité générale et la vie privée, ils peuvent être utilisé de façon illégale dans des zones interdites pour espionner des individus ou bien pour le transport clandestin de marchandises prohibées. Dans ce contexte, la détection des drones est devenue une préoccupation majeure pour différents secteurs. Les chercheurs de ce domaine sont obligés de développer un système basé soit sur le radar, le signal acoustique, le signal radiofréquence ou le signal visuel qui doit être capable de détecter efficacement la présence d'un drone.

Dans notre travail, Nous utilisons deux méthodes de détection ; la première repose sur la captation acoustique, effectuée à l'aide de microphones ou de capteurs acoustiques, tandis que la deuxième utilise des données visuelles enregistrées par des caméras de surveillance. Pour les données sonores nous avons utilisé des techniques d'extraction des paramètres telles que : MFCC, LPC et ZRC. Les données recueillies sont ensuite classifiées à l'aide du KNN et le SVM.

Pour suivre le rythme d'innovations technologiques et les avancées rapides des techniques de l'intelligence artificielle dans le monde, nous avons utilisé des modèles d'apprentissage profond dans notre projet, spécifiquement les réseaux de neurones CNN « 1D et 2D » pour classer les signaux audios et les images, ces modèles sont capable de détecter efficacement la présence de drone et ainsi de lutter contre leurs utilisations irrégulière dans différente zones et de garantir la sécurité des personne.

L'objectif principal de ce travail est de développer un système capable de détecter la présence de drones avec une grande précision, et de distinguer ainsi différents objets sonores tels que le son d'un drone et celui d'une personne ou d'un animal, ainsi que des images de drones et d'oiseaux.

Organisation du mémoire :

Notre mémoire est se compose de quatre chapitres :

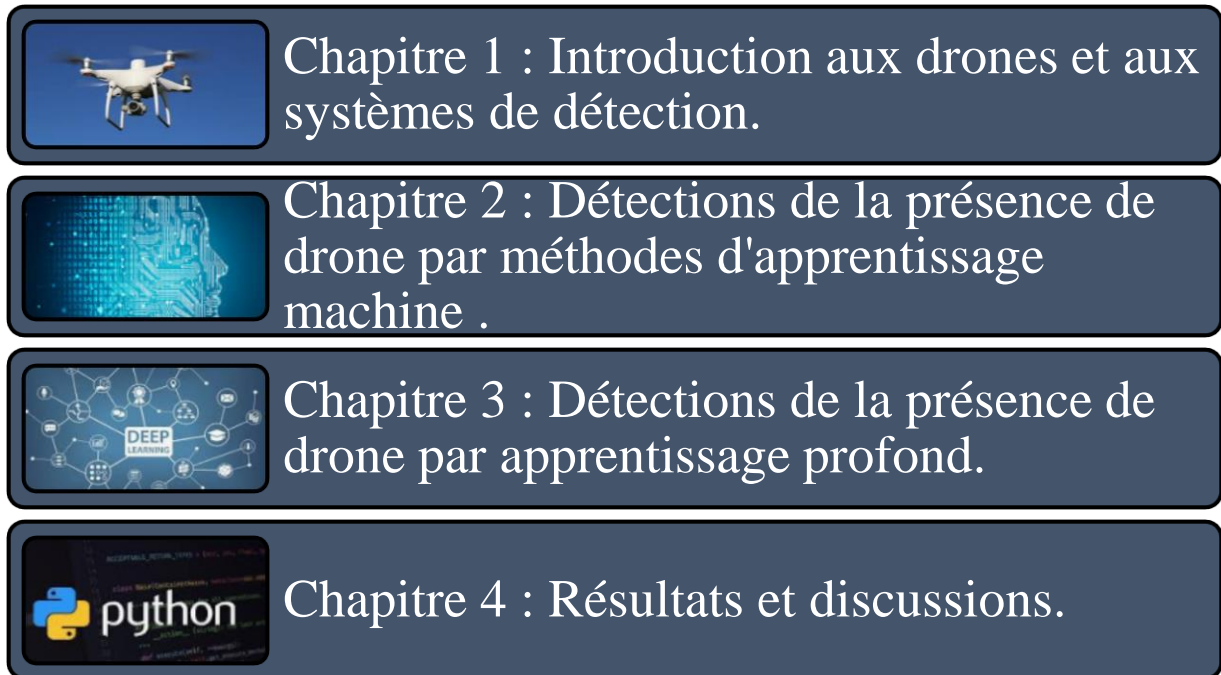


Figure 1 : Organisation de mémoire.

Le chapitre 1 :

Dans le premier chapitre, nous avons présenté de manière détaillée une introduction aux drones ainsi qu'aux divers systèmes de détection disponibles, comme le radar, les signaux visuels, les signaux acoustiques et la radiofréquence, en mettant la lumière sur leurs avantages et leurs limites.

Le chapitre 2 :

Dans le second chapitre nous avons examiné en profondeur la détection de la présence de drones en utilisant des techniques d'apprentissage machine classiques tel que le SVM et le KNN, ainsi que la phase d'extraction des paramètres à partir des signaux audios de drones en utilisant les coefficients MFCC, LPC, ZRC.

Le chapitre 3 :

Le troisième chapitre se concentre sur les réseaux de neurones convolutifs (CNN) notamment les configuration 1D et 2D et leurs applications dans la détection de drones.

Le chapitre 4 :

Le quatrième chapitre présente et discute les résultats obtenus en utilisant les deux types de données (audio et image).

**Chapitre 1 :
Introduction aux drones
et aux systèmes de
détection**

1.1. Introduction :

L'évolution des drones et de leurs systèmes de détection peut être définie comme une progression technologique continue qui englobe le développement et l'élargissement des capacités des véhicules aériens sans pilote (UAVs), ainsi que l'avancement des méthodes et technologies pour identifier, suivre, et parfois neutraliser ces appareils dans le but de garantir la sécurité et la confidentialité. Cette évolution est caractérisée par une diversification des applications des drones, allant des opérations militaires à des usages civils tels que la surveillance, et l'agriculture, parallèlement à l'amélioration des systèmes de détection qui intègrent des technologies avancées comme le radar, visualisation, audio, radio fréquence, et l'intelligence artificielle pour répondre efficacement aux défis posés par l'usage accru des drones.

1.2. Histoire du développement des drones :

L'origine du drone remonte à la Première Guerre Mondiale, durant laquelle il était développé pour les besoins de l'armée. Archibald Low, ingénieur et auteur anglais, développe, dès 1916, l'Aerial Target, un projet d'avion-cible sans pilote, commandé à distance au moyen des ondes de TSF (télégraphie son fil) en Grande Bretagne [1].

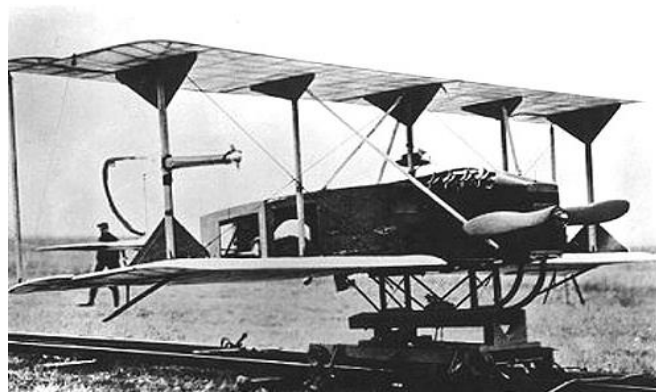


Figure 1.1: Hewitt-Sperry Automatic Airp [1]

Dans le même temps, en 1917 aux Etats Unis, les ingénieurs Elmer Ambrose Sperry, Lawrence Sperry et Peter Cooper imaginent un avion radio-commandé, le Hewitt-Sperry Automatic Airplane (figure 1.1).

Du côté français, le capitaine Max Boucher, grâce aux travaux préalables d'Octave Détable débuté en 1894 [1], réussit à faire voler sans pilote à bord de l'appareil, un avion Voisin, le 2 juillet 1917 sur 1km (figure 1.2).

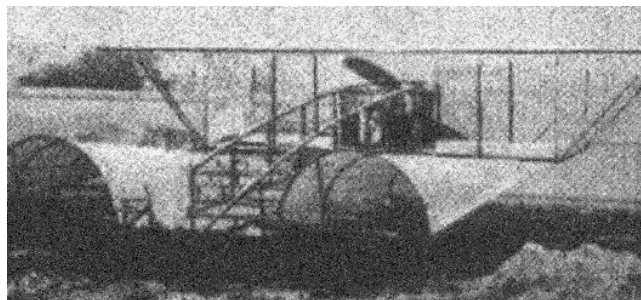


Figure 1.2 : Aéroplane Détable de max Boucher [1]

Georges Clémenceau, alors président de la Commission sénatoriale de l'armée, poussé par cet exploit, lance en 1918 un concours d'avions sans pilotes. En effet, les pilotes de l'armée

coutent cher à l'Etat. La formation est longue et la réduction des pertes humaines est non négligeables à la suite d'une longue guerre meurtrière. Placer le pilote à distance de l'aéronef représente une réduction importante des risques.

Le 14 septembre 1918, Max Boucher, ayant amélioré son système, fait voler un avion Voisin BN3 pendant 51 minutes sur un parcours de 100km. Le 17 avril 1923, accompagné de l'ingénieur Maurice Percheron, il réussit à faire voler un véritable drone piloté via la TSF. L'avion radiocommandé est né ! Mais la guerre est finie, et l'armée s'est désintéressée du projet.

Le nom drone est un mot anglais veut dire faux-bourdon (l'abeille qui bourdonne, son vol est associé à un bruit assez sourd et grave). En effet, en 1935, l'un des constructeurs de version automatisées d'avion, avait baptisé les avions cibles Queen Bee (reine des abeilles) mais leur vol les faisant plus ressembler à de faux bourdons, le mot drone fut choisi. C'est la Navy qui démocratise ce mot en 1941 avec son avion cible Target Drone Denny 1, fabriqué à plusieurs milliers d'exemplaires [1].



Figure 1.3 : Avion Cible Denny qui popularisa le terme de drone [1]

Durant la Seconde Guerre Mondiale, dès 1944, les Allemands développèrent des applications plus belliqueuses avec la mise au point de missiles sol/air radio – commandés.

Toutes ces armes novatrices ne purent être fabriquées en grand nombre, mais les technologies furent utilisées pour le développement de nouvelles armes chez les Alliés après la défaite Nazi.



Figure 1.4: Unmanned US Navy Grumman F6F – 5K Hellcat drone 1946 © US Navy Naval Aviation News January 1981[1].

Il faudra cependant attendre la Guerre Froide pour voir l'essor du drone. Il est développé secrètement par l'armée américaine comme un moyen de supériorité stratégique d'observation, de surveillance et d'intervention militaire dans les lignes ennemies sans encourir les risques humains, qui étaient devenus insupportables. Il permet également le largage des tracts dans le cadre de la guerre psychologique [1].



Figure 1.5: US Navy Culver TD2C-1 target drone © US Navy National Museum of Naval Aviation photo n 2012 075 001 [1].

Les drones sont engagés lors des Guerres du Vietnam et de Corée, puis lors de la guerre du Kipour. Ils font désormais parties des moyens tactiques et stratégiques des champs de bataille. Les missions à risques sont facilitées puisque la crainte de pertes humaines a disparu avec l'essor des drones.

Les drones armés sont développés durant les années 90. Ils peuvent non seulement être utilisés dans des missions de surveillance mais également dans des missions d'attaques au moyen d'armes embarquées. Actuellement, les drones militaires vont de quelques centaines de grammes à plusieurs tonnes en fonction de l'utilisation et de l'autonomie voulues.

Une nouvelle étape est en cours avec le développement des drones de combat. Plusieurs programmes ont été lancés en France, en Europe, aux États-Unis ... mais, pour des raisons de sécurité évidentes, les constructeurs et les États-Majors restent très secrets à ce sujet. D'autant plus que l'emploi des armes totalement automatiques, sans prise de décision humaine, va à l'encontre des lois de la Guerre énoncées par la Convention de Genève.

Historiquement, le secteur de la défense représente un vivier dans le développement de technologies novatrices, qui sont ensuite réutilisées dans le domaine civil. Depuis les années 2000, les drones sont utilisés dans des applications civiles de surveillance d'infrastructures et agricoles, audiovisuelles et cinématographiques.

1.3. Définition d'un drone :

Un drone ou Unmanned Aerial Vehicle (UAV) est un aéronef sans passager ni pilote qui peut voler de façon autonome ou être contrôlé à distance depuis le sol. Le mot « drone » est une extrapolation du terme anglais qui signifie « faux-bourdon ». En français, le terme est employé pour désigner des véhicules aériens, terrestres, de surface ou sous-marins[2].

Il est susceptible d'emporter différentes charges utiles, le rendant capable d'effectuer des tâches spécifiques, pendant une durée de vol qui peut varier en fonction de ses capacités. L'utilisation des drones a d'abord été connue dans les applications militaires, comme la surveillance et la reconnaissance et comme plateforme de désignation de cible ou comme arme. Puis, plusieurs applications civiles sont devenues concurrentes, notamment dans l'observation des phénomènes naturels (Avalanches, volcans...), la pulvérisation des pesticides sur les surfaces agricoles, la surveillance de l'environnement (exemple mesures de la pollution) et des réseaux routiers et la maintenance des infrastructures [3].

1.4. Les composants principaux du drone :

Pour mieux connaître notre objet volant, étudions-le composant par composant. Un drone est composé :

➤ **D'un châssis** : c'est la base du drone. Selon les modèles, le châssis est fait de fibre de carbone, de plastique, d'aluminium, ou d'autres matières plutôt résistantes. Il peut accueillir 3, 4, 6 ou 8 bras pour la fixation des moteurs. Il peut être recouvert d'une coque en plastique pour adopter un look sympa, ou rester « nu » dans le but de peser le moins lourd possible [4].



Figure 1.6 : châssis comporte 4 bras [5]

➤ **D'un système de propulsion** : c'est ce qui va permettre au drone de se déplacer. Il est composé de moteurs, d'hélices, et des contrôleurs de vitesse électronique, appelés ESC (Electronique Speed Controller). Pour animer le tout, les drones de loisir s'appuient sur une alimentation électrique fournie par des batteries. En fonction de ces différents éléments, le temps de vol ainsi que la charge maximale supportée par le drone pourront se différer [4].



(a)



(b)



(c)

Figure 1.7: (a)Hélices (b)Batterie(c) Contrôleurs ESC [4]

➤ **Un contrôleur de vol** : grâce à un microprocesseur et des capteurs et parfois un GPS, il permet de gérer les moteurs pour assurer la stabilité de l'appareil et faire le lien entre les commandes du pilote et le drone [4].



Figure 1.8 : une carte électronique [4]

➤ **La caméra** : Elle permet d'enregistrer le vol pour le visionner plus tard [5].

1.5. Classification des drones selon le mode de vol :

Beaucoup de classifications de drones existes, nous présentons dans ce qui suit une classification selon le mode aérien :

➤ **Les drones à voilure fixe** : ce sont les appareils qui s'appuient sur des ailes pour rester en vol. Ils doivent être en mouvement pour profiter de la portance des ailes. Les avions et ailes volantes sont des appareils à voilure fixe. Ils sont capables de voler longtemps en consommant peu d'énergie, ils peuvent voler vite [6].



Figure 1.9 : Les drones à voilure fixe [12]

Parmi les inconvénients de cette configuration [8] :

- La nécessité d'une piste de décollage.
- Elle n'autorise pas le vol stationnaire.
- Le vol à basse vitesse est impossible.
- Le vol à basse altitude est impossible.

➤ **Les appareils à voilure tournante** : ce sont des appareils qui reposent sur des hélices en rotation à l'horizontale pour rester en vol. Ils n'ont pas besoin d'être en mouvement pour rester en l'air. Ils peuvent rester en vol stationnaire, ils ont besoin d'une surface réduite pour décoller et atterrir, mais ils volent assez lentement.

Parmi leurs inconvénients :

- Consomment plus d'énergie et souffrent donc d'une autonomie réduite [6].
- Complexité en maintenance et réparation [3].
- Moins d'autonomie, de vitesse et d'altitude en vol [3].

➤ **Aile battante** : Le drone est guidé et piloté en battant des ailes. Ils imitent les insectes et peuvent planer à basse vitesse et effectuer des tâches dans des espaces très restreints [8].



Figure 1.10 : drone à aile battante [12].

Parmi leurs inconvénients [3] :

- Sensible aux conditions météorologiques.
- Complexité de réalisation.
- Mémoire trop limitée.
- Ressources d'énergie limitées.

1.6. Fonctionnement du drone :

En fait, il s'agit d'un système multi-hélices spécialement conçu à l'intérieur d'un drone qui rend cet appareil très indépendant et aide également à réduire les risques de pannes. Il est à noter que même si un moteur à l'intérieur de cet appareil cesse de fonctionner, le drone continuera à voler car il obtient le soutien des hélices qui travaillent en groupe. Un drone qui possède un grand nombre de moteurs à l'intérieur peuvent maîtriser leur élévation et peuvent donc porter des charges plus lourdes pendant le vol. Ces hélices reçoivent leur énergie d'une source dédiée et la plupart de ces appareils contiennent des piles amovibles pour pouvoir rester en l'air pour une longue durée. Le temps de vol peut être prolongé grâce à l'utilisation des piles puissantes [7].

1.7. Comment ça vole ?

La commande joue un rôle important dans le mécanisme de vol de l'appareil. Cet accessoire est utilisé pour contrôler tous les mouvements du drone, depuis son lancement, en passant par ses capacités de navigation, jusqu'à son atterrissage. Le marché regorge de diverses commandes de drones aujourd'hui et les développeurs utilisent souvent diverses expériences pour créer des drones avec des fonctionnalités impressionnantes.

La tâche principale d'une commande est d'établir un canal de communication approprié entre l'unité distante et les ondes radio. La plupart des drones fonctionnent généralement sur une gamme de fréquences de 2,4 GHz et plusieurs de ces commandes passent par les réseaux Wi-Fi pour prendre des décisions actives concernant les mouvements. Les fonctionnalités d'un smartphone et d'un drone sont assez identiques ; les deux embarquent en effet un GPS, le Wi-Fi, ainsi que de nombreux capteurs communs. Ces capteurs embarqués aident les drones à rester en l'air pendant longtemps et à prendre de bonnes décisions quant à leur hauteur, leur direction et d'autres mouvements importants.

Le processus d'atterrissage est également contrôlé par un système d'hélice à l'intérieur et les capteurs prennent des décisions concernant sa vitesse, l'altitude et la rotation du moteur, etc. [7]

1.8. Les caractéristiques les plus remarquables des drones :

➤ **Taille et poids** : Les drones sont disponibles dans une large gamme de tailles et de poids, depuis les petits modèles de jouets jusqu'aux gros véhicules utilisés dans les applications industrielles. Cette variété permet d'adapter le drone aux besoins spécifiques de chaque utilisateur.

➤ **Capacité** : Les drones peuvent transporter différents types de marchandises, des caméras et capteurs, des colis et fournitures. La capacité d'emport d'un drone détermine son utilité dans diverses applications, comme la photographie aérienne, la livraison de produits ou la surveillance.

➤ **Temps de vol** : Le temps de vol d'un drone dépend de plusieurs facteurs, tels que la taille de la batterie et l'efficacité du moteur. En général, les drones commerciaux ont une durée de vol limitée, qui peut aller de quelques minutes à plus d'une heure dans les modèles haut de gamme.

➤ **Vitesse et portée** : Les drones peuvent atteindre des vitesses impressionnantes, notamment les modèles conçus pour les courses ou les applications militaires. Quant à la portée,

les drones grand public ont généralement une portée limitée à plusieurs centaines de mètres, tandis que les drones professionnels peuvent opérer à des distances de plusieurs kilomètres.

➤ **Stabilité et contrôle :** Les drones modernes sont équipés de systèmes de stabilisation avancés, tels que des gyroscopes et des accéléromètres, qui leur permettent de maintenir une position stable dans les airs. De plus, les drones peuvent être contrôlés avec précision à l'aide d'une télécommande, d'applications mobiles ou encore de systèmes de navigation GPS.

➤ **Caméras et capteurs :** De nombreux drones sont équipés de caméras et de capteurs qui leurs permettent de capturer des images et des vidéos de haute qualité depuis les airs. Ces capacités rendent les drones idéaux pour des applications telles que la photographie aérienne, la surveillance environnementale ou l'inspection des infrastructures [9].

1.9. Les domaines d'application et d'utilisation des drone :

➤ **Le journalisme :**

- Utilisé pour tourner des documentaires et des films.
- Utilisé pour atteindre des zones sinistrées ou difficilement accessibles.
- Avoir des plans aériens sur de grandes manifestations sans avoir à utiliser un hélicoptère (ce qui est bien plus coûteux).

➤ **La cinématographiques :**

- Grâce aux stabilisateurs, les images capturées par les drones dédiés au cinéma sont nettes.
- Utilisés par de nombreux réalisateurs pour remplacer l'hélicoptère, la grue ou le Steady Cam.
- Utilisés pour réaliser des clips musicaux ou des vidéos YouTube.

➤ **Le tourisme :**

- Utilisée pour réaliser des plans de la ville ou du territoire qu'il souhaite promouvoir.
- Faire des vues aériennes sur les monuments et sites remarquables d'une région pour donner envie aux touristes de se rendre sur place.

➤ **L'agriculture :**

- Surveiller les champs des agriculteurs.
- Fournir des services très pointus par des drones spécialisés dans le domaine : repérage des hétérogénéités du sol, examen de la densité végétale, création de zones sur les parcelles, etc. [10]

➤ **La surveillance des animaux :**

Les drones sont également utilisés depuis longtemps pour mieux connaître les animaux et mieux les protéger, en fait, les drones sont sans danger pour les opérateurs, moins bruyants que les aéronefs habités et capables d'accéder à des zones dangereuses ou éloignées [11].

➤ **Drone solution écologique :**

Les drones sont essentiels pour surveiller la biodiversité, et la pollution, fournissant des données précieuses pour la recherche écologique. Leur capacité à cartographier les paysages et intervenir lors de catastrophes naturelles en fait des outils polyvalents pour la conservation de l'environnement.

➤ **Militaire :**

Les drones militaires offrent une capacité de surveillance et de renseignement stratégique, permettant de collecter des données en temps réel sur les mouvements des troupes ennemies et les activités terroristes. Ils sont utilisés pour des missions de reconnaissance, de ciblage et de frappe précise, réduisant ainsi les risques pour les soldats sur le terrain.

1.10. La détection des drones :

Il existe plusieurs raisons pour lesquelles il est nécessaire de détecter des drones :

- **Sécurité aérienne :**

Les drones non autorisés peuvent constituer une menace pour la sécurité aérienne s'ils interfèrent avec les avions ou s'ils entrent dans des zones restreintes telles que les aéroports, mettant ainsi en danger les avions et les passagers.

- **Protection de la vie privée :**

Les drones équipés de caméras peuvent être utilisés pour espionner des individus ou des infrastructures sensibles (telles que les centrales électriques, les raffineries de pétrole, les installations militaires ou les réseaux de communication), ce qui soulève des préoccupations en matière de vie privée et de sécurité.

- **Sécurité de publics :**

Lors de grands événements publics tels que les manifestations, les rassemblements sportifs ou les festivals, la présence de drones non autorisés peut représenter une menace potentielle pour la sécurité des participants et des spectateurs.

- **Lutte contre les crimes :**

Les drones peuvent être utilisés pour transporter des substances illicites ou effectuer des activités criminelles telles que le trafic de drogue, la contrebande ou l'espionnage industriel, ce qui nécessite une surveillance et une détection efficaces pour l'application des lois.

1.11. Les Systèmes de détection de drones :

C'est un ensemble d'outils qui aident à repérer les drones volants dans le ciel pour s'assurer qu'ils ne se trouvent pas où ils ne devraient pas être comme :

1. Les systèmes Radar.
2. Les systèmes Acoustique.
3. Les systèmes Visuelle.
4. Les systèmes Radio fréquence.

1.11.1. Détection radar :

Les systèmes radar sont largement utilisés dans la vie quotidienne et dans les applications militaires. Ces systèmes radar conventionnels sont ajustés pour détecter des objets relativement grands. Ils ne conviennent pas à la détection de petits drones, surtout lorsque la vitesse du drone est faible, lorsqu'ils volent à basse altitude. La solution radar pour les systèmes de défense des drones (SDDD Système de Défense Des Drones), représente une méthode active d'identification [12].



Figure 1.11: Système de détection de drones avec Radar [12]

1.11.2. Détection des drones par données visuelles :

Malgré son succès traditionnel dans l'identification et le suivi des cibles, le radar reste une technologie hautement professionnelle qui nécessite un personnel qualifié capable d'interpréter les résultats visuels du système radar au moins pour la prise de décision. Cette complexité de la technologie radar et l'évolution rapide de la surveillance informatique ont incité certains chercheurs à envisager la détection des drones à l'aide de données visuelles (image ou vidéo). Les contributions dans ce domaine peuvent être divisées en deux catégories :

- La première catégorie comprend les solutions qui s'appuient sur des fonctionnalités apprises, omettant ainsi l'étape approfondie de l'ingénierie des fonctionnalités (détection visuelle avec caractéristiques apprises).
- La deuxième catégorie dépend des schémas d'apprentissage automatique traditionnels qui devraient alimenter le système avec des fonctionnalités artisanales de bas niveau telles que les bords, les taches et les informations de couleur (caractéristiques faites à la main) [11].

1.11.3. Détection par signaux acoustiques :

Un drone volant produit un bruit de ronflement qui peut être capturé par des capteurs acoustiques (des microphones), et analysé en utilisant différentes méthodes pour identifier l'empreinte digitale audiophone spécifique. Le résultat idéal serait de déterminer le type de drone ou même le drone individuel par son empreinte audio. En général, la détection acoustique des drones repose soit sur des méthodes de corrélation/autocorrélation ou sur la classification d'apprentissage automatique [13].

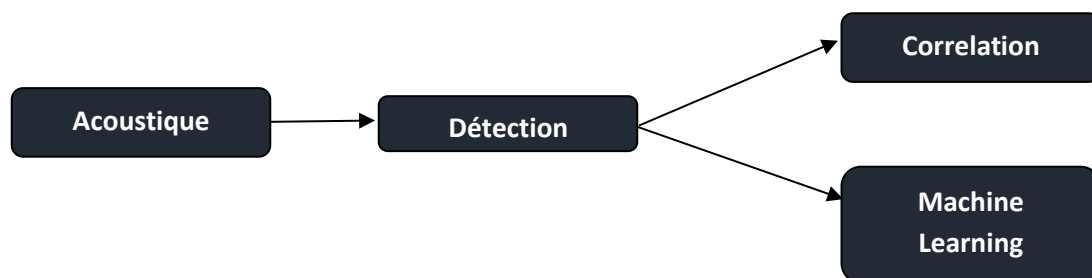


Figure 1.12 : Schéma synoptique de la méthode de détection acoustique. [12]

1.11.4. Détection des radio fréquences :

La détection UAV par fréquences radio (RF) est une technique qui implique l'interception et l'analyse des signaux transmis (Tx, Rx) entre l'UAV et la station au sol. Habituellement, ces signaux consistent en des signaux de contrôle en liaison montante (depuis la station au sol) et en liaison descendante (depuis le drone) signaux de données (données de position et vidéo) [12].



Figure 1.13 : Système de détection de drones par Radio Fréquences [12].

1.12. Les avantages et les inconvénients des systèmes de détection :

Table 1.1 : Les avantages et les inconvénients des différents systèmes de détection.[13]

Le système de détection de drone	Les avantages	Les inconvénients
Radar	<p>Les radars sont résistants au brouillard, aux nuages et à la poussière contrairement à la détection visuelle qui est moins prononcée au bruit.</p> <p>Les radars offrent une résolution plus élevée dans la gamme et permettent de capturer la signature microdoppler (MDS).</p>	<p>Les drones ont de petites sections radar (RCS : radar cross section), ce qui rend la détection plus exigeante. Les ondes ont une perte de chemin plus élevée, ce qui limite la portée de détection des drones. Une petite taille et une faible signature électromagnétique des drones rendant la détection par radar plus difficile.</p>
Acoustique	<p>Il fonctionne dans un environnement à faible visibilité. Faible coût en fonction de la gamme de microphones utilisée.</p>	<p>Sensible au bruit ambiant. L'état du vent affecte les performances de détection. Nécessite une base de données de signatures acoustiques pour différents drones pour la formation et les tests.</p>

Visuelle	<p>Faible coût en fonction des caméras et des capteurs optiques utilisés.</p> <p>La détermination des résultats de détection à l'aide d'écrans par l'homme est plus simple que les autres modalités.</p>	<p>Le niveau de visibilité est affecté par la poussière, le brouillard, les nuages et la journée.</p> <p>Des caméras basées sur un laser et un large champ de vision peuvent être nécessaires.</p>
Radio fréquence	<p>Capteurs RF à faible coût. Longue plage de détection.</p>	<p>Ne convient pas pour détecter les drones volant de manière autonome sans aucun canal de communication.</p>

1.13. Conclusion :

Dans ce chapitre nous avons présenté les composantes principales de drones, leur classification et les systèmes de détection. Opter pour l'acoustique et l'image pour un drone offre de nombreux avantages sans inconvénients majeurs. Cette approche permet une perception complète de l'environnement, améliorant la précision des données. En combinant la détection sonore et visuelle, les drones détectent des éléments échappant à une seule modalité, permettant une utilisation flexible dans divers secteurs.

Dans le chapitre suivant, nous explorerons en détail les différents paramètres audios et leur classification pour la reconnaissance précise des différents types de sons. En combinant l'analyse acoustique avec d'autres modalités telles que l'imagerie, nous pouvons obtenir une perception plus complète de l'environnement. Cette approche offre de nombreux avantages, notamment une meilleure précision des données et une détection plus efficace des éléments échappant à une seule modalité. Ces techniques sont cruciales pour une utilisation flexible des drones dans divers secteurs, de la surveillance à la gestion des catastrophes en passant par la sécurité.

Chapitre 2 : La détection de la présence de drone par les techniques d'apprentissage machine

2.1 Introduction :

L'extraction des paramètres revêt une importance capitale dans le secteur du traitement du signal et de la reconnaissance de formes, notamment à travers des techniques comme les Coefficients Cepstraux de Fréquence de Mel (MFCC), les Coefficients Cepstraux Gammatones (GTCC) et la Prédiction Linéaire (LPC). Les méthodologies qui reposent sur l'apprentissage machine offrent des solutions avancées pour analyser et classifier les signaux, tel que le K-plus Proches Voisins (KNN) et les Machines à Vecteurs de Support (SVM), ces techniques s'avérant particulièrement utiles dans des domaines tels que la surveillance par drone, l'analyse d'images aériennes et la reconnaissance d'objets. Dans ce chapitre nous présentons les méthodes d'extraction des paramètres à partir des signaux audios, ainsi que les méthodes de classification à base d'apprentissage machine utilisées dans notre travail.

2.2 Définitions d'un signal audio :

Un signal audio est une représentation du son, généralement à l'aide d'un niveau de tension électrique variable pour les signaux analogiques, ou d'une série de nombres binaires pour les signaux numériques. Les signaux audios ont des fréquences comprises dans la gamme des fréquences audio, soit environ 20 à 20 000 Hz[14].

2.3 Architecture du système de reconnaissance de signal audio de drone :

La figure suivante présente le schéma générale de reconnaissance des signaux de drones audios en utilisant les méthodes d'apprentissage machine.

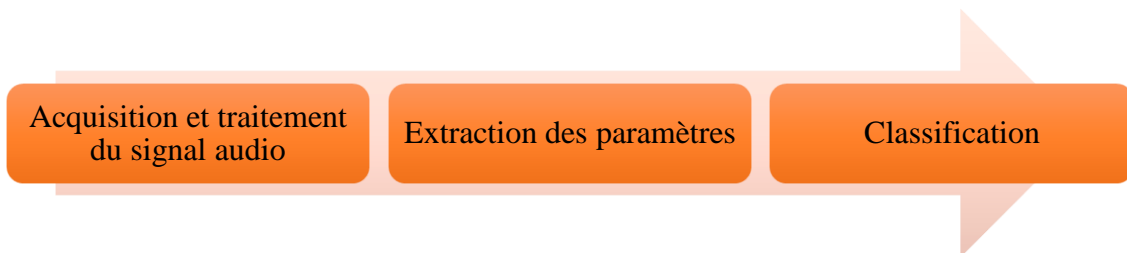


Figure 2.1 : Schéma global du système de reconnaissance de drone par signaux audios

Nous présentons dans ce qui suit chaque bloque séparément.

2.3.1 Acquisition et traitement du signal audio :

La phase d'acquisition et traitement du signal audio est schématisée dans la figure suivante :

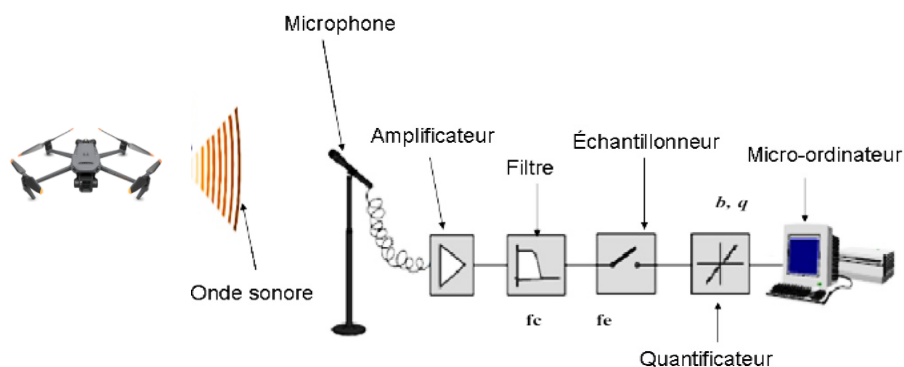


Figure 2.2 : Prétraitement acoustique du signal audio

Le prétraitement acoustique d'un signal audio capturé par un microphone dans un environnement où un drone sonore est présent nécessite plusieurs étapes clés pour garantir la clarté et l'utilité du signal. Initialement, le signal est capturé par un microphone qui convertit les ondes sonores, y compris celles générées par le drone, en signaux électriques. L'amplitude de ces signaux est ensuite ajustée pour optimiser le niveau sonore avant d'être transmis à un filtre. Ce filtre est essentiel pour éliminer les fréquences indésirables et réduire le bruit de fond associé au drone. Après filtration, le signal est envoyé à un échantillonneur, où il est converti en une série de valeurs discrètes pour un traitement numérique. Le quantificateur prend ensuite ces échantillons et les transforme en valeurs numériques précises qui peuvent être traitées par un micro-ordinateur. Ce dernier utilise des algorithmes avancés pour analyser et traiter davantage le signal, permettant ainsi l'extraction de l'information utile tout en minimisant l'impact sonore du drone.

2.3.2 Extraction de paramètres :

L'extraction de paramètres, comme les caractéristiques et les attributs, a pour objectif de convertir le signal acoustique en une série de vecteurs qui représentent de manière précise le signal initial. Ces vecteurs sont ensuite utilisés dans le processus de classification pour identifier la classe du signal parmi celles prises en compte dans le système de classification. On présente par la suite quelques paramètres.

2.3.2.1 Les coefficients cepstraux de fréquence Mel (MFCC) :

Le MFCC est une représentation du spectre de puissance à court terme d'un son, largement utilisée dans le traitement de la parole et de l'audio.

Le processus d'obtention du MFCC implique plusieurs étapes, notamment le cadrage du signal audio en trames courtes, l'application d'une fonction de fenêtre à chaque image, le calcul de la transformée de Fourier discrète (TFD) du signal et enfin la prise du logarithme de l'amplitude de la TFD. Le spectre résultant est ensuite passé à travers un banc de filtres Mel, et les coefficients obtenus après application de la transformée en cosinus discrète (DCT) aux énergies du banc de filtres logarithmiques sont appelés MFCC.

MFCC est une technique d'extraction de fonctionnalités puissante qui capture les caractéristiques essentielles du signal audio, ce qui la rend particulièrement utile pour des tâches telles que la reconnaissance vocale, l'identification du locuteur et la classification audio. Sa capacité à représenter les caractéristiques spectrales des signaux audio de manière compacte et discriminante en a fait une pierre angulaire dans le domaine du traitement du signal audio.

Fonctionnement détaillé :

La MFCC se décompose en 5 phases [15] :

- **Phase 1** : on découpe le signal en plusieurs fenêtres qui se recoupent entre elles. Nous appliquerons la MFCC à chaque fenêtre.
- **Phase 2** : Afin de diminuer la distorsion spectrale nous appliquons une fenêtre de Hamming au signal :

$$W(n) = 0.54 - 0.46 * \cos\left(\frac{2\pi n}{TAILLE - 1}\right) \quad (2.1)$$

n : est une variable qui représente l'index de l'échantillon dans une séquence de données.

Taille : représente la taille totale de la séquence de données pour la fenêtre de Hamming

Par la suite nous multiplions cette fonction par le signal à transformer, nous minimisons ainsi la distorsion spectrale créée par le recoupement.

- **Phase 3** : on applique ensuite la FFT à la fenêtre pour en ressortir la magnitude, on obtient donc le spectre.
- **Phase 4** : On passe à l'échelle de Mel. En effet, après des études sur l'ouïe humaine, il a été montré que l'homme se base sur une échelle fréquentielle spécifique. La formule de transfert est simple :

$$\text{mel}(f) = 2595 * \log_{10} \left(1 + \frac{f}{700} \right) \quad (2.2)$$

Avec f : est la fréquence en Hz.

Pour simuler l'oreille humaine, il faut passer par un Banc Filtre, un filtre pour chaque fréquence que l'on cherche. Ces filtres ont une réponse de bande passante triangulaire. Pour connaître l'intervalle entre chaque filtre, on utilise une constante : Mel-Frequency interval. Nous utilisons 20 filtres.

- **Phase 5** : Pour finir, nous travaillons avec le Cepstre, nous convertissons le spectre logarithmique de Mel en fonction du temps au moyen de la DCT (Discret Cosinus Transform). La formule de cette transformation est simple :

$$y(k) = \sum_{n=1}^N w(n)x(n) \cos\left(\frac{\pi(2n-1)(k-1)}{2N}\right) \text{ avec } k = 1, \dots, N \quad (2.3)$$

Où $w(n) = \sqrt{\frac{1}{N}}$ si $n = 1$ et $w(n) = \sqrt{\frac{2}{N}}$ sinon

N : est un paramètre qui représente la taille totale de la séquence données sur laquelle nous effectuons le calcul.

La figure suivante présente un schéma synoptique des étapes d'extraction des coefficients MFCC.

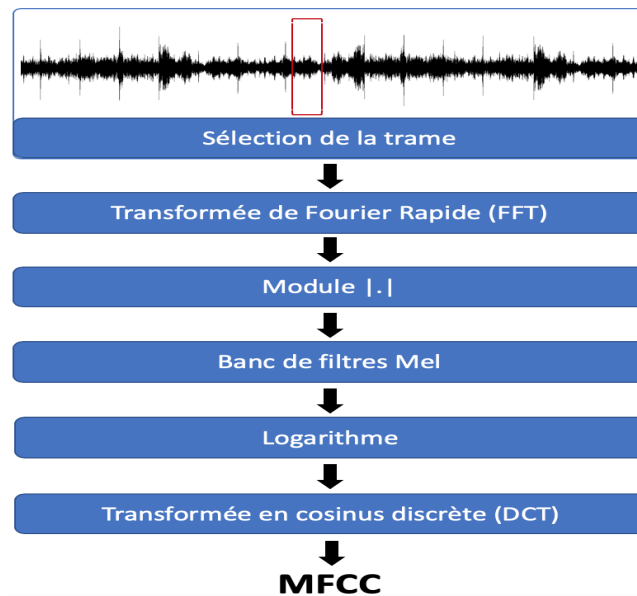


Figure 2.3 : schéma synoptique des étapes d'extraction des coefficients MFCC [16]

Comme cela a été discuté précédemment, l'objectif principal du MFCC est d'imiter le comportement des oreilles humaines. En outre, plutôt que les formes d'onde de la parole elles-mêmes.

Tel qu'il a été déjà mentionné le fenêtrage consiste à fenêtrer chaque trame individuelle de manière à minimiser les discontinuités du signal au début et à la fin de chaque trame. Le concept ici est de minimiser la distorsion spectrale en utilisant la fenêtre pour la réduire à zéro au début et à la fin de chaque trame.

La Transformée de Fourier rapide FFT convertit chaque trame de N échantillons du domaine temporel dans le domaine fréquentiel ; Le résultat obtenu après cette étape est souvent appelé 'Spectre' ou 'Périodogramme du signal'.

Le banc de filtres Mel est appliqué dans le domaine fréquentiel. Ensuite on transforme les données dans l'échelle des Mels-fréquentielle donc vers l'échelle des temps. Le résultat de cette étape sera les MFCC proprement dit.

Généralement le cepstre est obtenu comme indiquer dans la figure suivante.

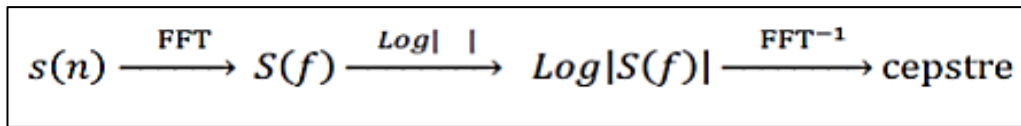


Figure 2.4 : Différentes étapes de l'analyse cepstrale [17]

2.3.2.2 Les coefficients cepstraux Gammatones (GTCC) :

Les GTCC sont des caractéristiques d'inspiration biologique utilisées pour la classification audio non vocale, adaptant les coefficients cepstraux de fréquence Mel (MFCC) largement utilisés et employant des filtres Gammatones avec des bandes passantes rectangulaires équivalentes. Les GTCC sont conçus pour représenter et distinguer avec précision les signaux audios non vocaux, excellant particulièrement dans les basses fréquences [18]. La réponse impulsionnels $g(t)$ du filtre Gammatone est donnée par la fonction suivante [19] :

$$g(t) = kt^{(n-1)}e^{-2\pi Bt} \cos(2\pi f_c t + \varphi) \quad t > 0 \quad (2.4)$$

n : est l'ordre du filtre,

K : est le facteur d'amplitude,

f_c : est la fréquence centrale en Hertz,

φ : est le déphasage,

B : est largeur de la bande.

Les étapes d'extraction des coefficients GTCC sont présentées par la figure suivante.



Figure 2.5 : schéma synoptique des étapes d'extraction des coefficients GTCC [20]

2.3.2.3 Le codage à prédiction linéaire (LPC) :

Le codage prédictif linéaire (LPC) est une méthode principalement utilisée dans le traitement du signal audio et le traitement de la parole pour représenter l'enveloppe spectrale d'un signal numérique de parole sous forme compressée, en utilisant les informations d'un modèle prédictif linéaire.

LPC est la méthode la plus utilisée en codage et en synthèse vocale. Il s'agit d'une technique d'analyse vocale puissante et d'une méthode utile pour coder une parole de bonne qualité à un faible débit.

Un traitement du signal est une activité d'extraction d'une information de signal. Le codage prédictif linéaire (LPC) est une technique d'analyse vocale puissante qui facilite l'extraction des fonctionnalités de bonne qualité. En 1978, LPC était utilisée pour faire une synthèse vocale. La procédure d'obtention des coefficients LPC est illustrée à la figure 6.

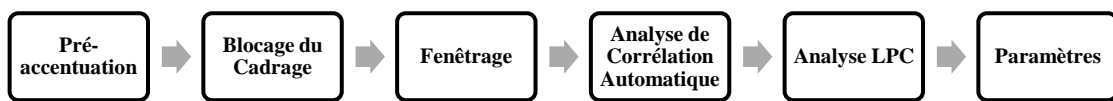


Figure 2.6 : Diagramme de Méthode LPC

- a) **Préaccentuation** : Lors du traitement du signal vocal, un filtre de préaccentuation est nécessaire après le processus d'échantillonnage. Le but du filtrage est d'obtenir une forme spectrale lisse du signal de parole. Un spectral qui a une valeur élevée pour le champ de basse fréquence et diminuée pour une fréquence de champ supérieure à 2000 Hz. Le filtre préemphasis est basé sur la relation d'entrée/sortie dans le domaine temporel est illustré par l'équation (2.5).[21]

$$y(n) = x(n) - ax(n-1) \quad (2.5)$$

a : est une constante du filtre de préemphasis, avec $0,9 < a < 1,0$.

- b) **Blocage du cadre** : segmenté le signal de parole de sorte qu'aucun signal ne soit perdu (suppression).
- c) **Fenêtrage** : Le signal analogique convertit devient un signal numérique lu image par image et chaque cadre est fenêtré. Ce processus de fenêtrage a pour but de minimiser et interrompez le signal du début à la fin de chaque trame. Si la fenêtre est $w(n)$, $0 \leq n \leq N-1$, si N est le nombre total d'échantillons de chaque trame, ainsi le résultat du fenêtrage est un signal :

$$y_1 = x_1(n)w(n), 0 \leq n \leq N - 1 \quad (2.6)$$

Où $w(n)$ utilise la fenêtre de Hamming avec la forme :

$$w(n) = 0.54 - 0.46 \cdot \cos\left(\frac{2\pi n}{N-1}\right), 0 \leq n \leq N - 1 \quad (2.7)$$

- d) **Analyse d'autocorrélation** : L'étape suivante est l'analyse d'autocorrélation vers chaque résultat de trame par fenêtrage $y_1(n)$ (l'équation (2.8)),

$$r_1(m) = \sum_{n=0}^{N-1-m} y_1(n)y(n+m), m = 0, 1, 2, \dots, p \quad (2.8)$$

Où p se situe entre 8 et 16.

- e) **Analyse LPC** : Cette étape convertira chaque trame de l'autocorrélation $p+1$ en compilation de "Paramètre LPC" $a_m = a_m^p$ pour $m = 1, 2, \dots, p$. Cette compilation devient le coefficient LPC ou devient autre transformation LPC. La méthode utilisée pour que le coefficient d'autocorrélation devienne un paramètre de compilation LPC est appelée méthode Durbin d'où la forme est :

$$E^0 = r(0) \quad (2.9)$$

$$k_m = \frac{\{r(m) - \sum_{j=1}^{m-1} r(|m-j|)\}}{E^{m-1}}, \quad 1 \leq m \leq p, \quad (2.10)$$

$$\alpha_m^m = k_m, \quad (2.11)$$

$$\alpha_j^m = \alpha_j^{(m-1)} - k_m \alpha_{m-j}^{(m-1)}, \quad 1 \leq j \leq m-1 \quad (2.12)$$

$$E^m = (1 - k_m^2) E^{(m-1)}, \quad (2.13)$$

Avec $r(0)$ est le résultat d'une autocorrélation, $E(m)$ est une erreur, k_m est un rebond du coefficient, α_j^m est le coefficient de prédiction pour $1 \leq j \leq m$.

- f) **Conversion des paramètres LPC en coefficient LPC** : ceci est obtenu par les équations suivantes:

$$c_m = a_m + \sum_{k=1}^{m-1} \left(\frac{k}{m}\right) c_k a_{m-k}, \quad 1 \leq m \leq p \quad (2.14)$$

$$c_m = \sum_{k=1}^{m-1} \left(\frac{k}{m}\right) c_k a_{m-k}, \quad m > p \quad (2.15)$$

Ce coefficient cepstral est le coefficient de représentation du spectre logarithmique.

2.3.2.4 Taux de passage par ZCR :

ZCR (Zero Crossing Rate) ou le taux de passage par zéro, est une caractéristique fréquemment utilisée pour la classification parole / musique. En effet, les brusques variations du ZCR sont significatives de l'alternance voisée / non voisée [22]. Pour un signal échantillonné, il y a passage par zéro lorsque deux échantillons successifs sont de signes opposés ; le taux de passage par zéro court terme peut être estimé par la formule [23] :

$$Z(n) = \sum_{m=-N/2}^{N/2} |sgn x(n+m) - sgn x(n+m-1)| \cdot w(m) \quad (2.16)$$

Où :

$$w(n) = f(x) = \begin{cases} \frac{1}{N}, & -N/2 \leq n \leq N/2 \\ 0, & \text{ailleurs} \end{cases}$$

Il est clair que les valeurs de $Z(n)$ sont normalement plus élevées pour les sons non voisés que pour les sons voisés (Le voisement est une qualité (ou propriété) de certains sons de la parole. Un son est voisé si sa production s'accompagne d'une vibration des cordes vocales et sinon, il est non voisé. On utilise aussi couramment les termes de sonore et sourd pour désigner cette opposition) [24].

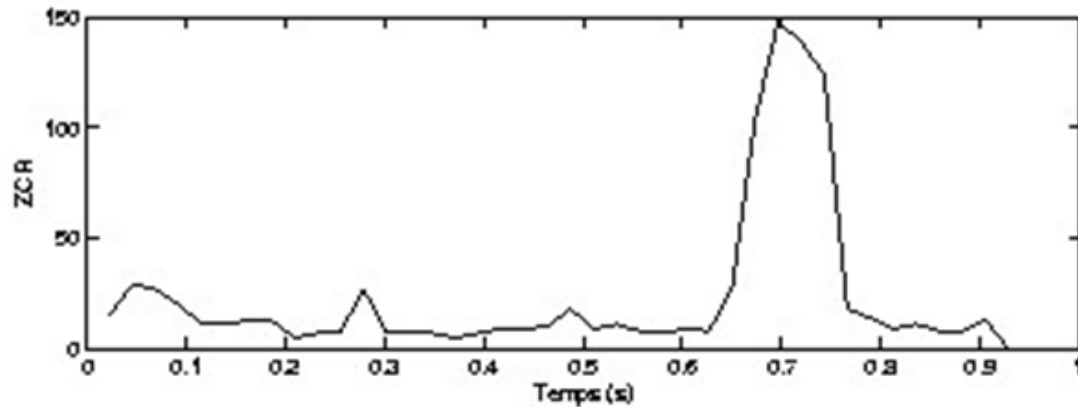


Figure 2.7 : ZCR d'un signal de parole de 1s. [23]

2.3.3 La classification des signaux audios :

2.3.3.1 L'intelligence artificielle :

Le terme IA "intelligence artificielle" a été adopté en 1955 par John McCarthy. En 1956, John McCarthy et ses collègues ont organisé une conférence intitulée « Projet de recherche en intelligence artificielle », qui a mené à l'apprentissage automatique, à l'apprentissage profond, à l'analyse prédictive et, plus récemment, à l'analyse indicative. Et un nouveau domaine d'étude est apparu : la science des données. Aujourd'hui, l'IA est sur une voie express, en raison de trois améliorations sectorielles majeures :

➤ **Les unités de traitement graphique (GPU):** En raison de la demande dans le monde de la vidéo et des jeux.

➤ **Le Big Data:** Le Big Data fait référence aux quantités exponentielles de données, les algorithmes utilisés par l'IA sont principalement façonnés par la grande quantité d'information, l'IA aide à traiter l'information à un rythme élevé et à rendre les données plus accessibles et utilisables.

➤ **Algorithmes:** Les algorithmes évoluent constamment et deviennent plus complexes, ces algorithmes sont utilisés pour compléter des tâches que l'on croyait possibles uniquement en utilisant l'intelligence humaine [25].

2.3.3.2 L'apprentissage machine (Machine Learning) :

C'est l'une des principales disciplines de l'intelligence artificielle. Cela dépend de l'utilisation des données brutes, de leur transformation en connaissances et de leur utilisation automatique pour prendre de meilleures décisions. L'apprentissage automatique permet d'utiliser un modèle qui permet aux machines d'apprendre et de se comporter de la même manière que les humains, d'apprendre indépendamment et d'interagir dans le monde réel. Le processus d'apprentissage repose sur les étapes suivantes :

- Remplir un algorithme avec des données.

- Utiliser ces données pour former un modèle.
- Tester et publier un formulaire.
- Utiliser le formulaire publié pour exécuter une fonction prédictive automatisée [26].

On distingue deux types d'apprentissage machine : l'apprentissage supervisé et l'apprentissage non-supervisé (figure 2.8).

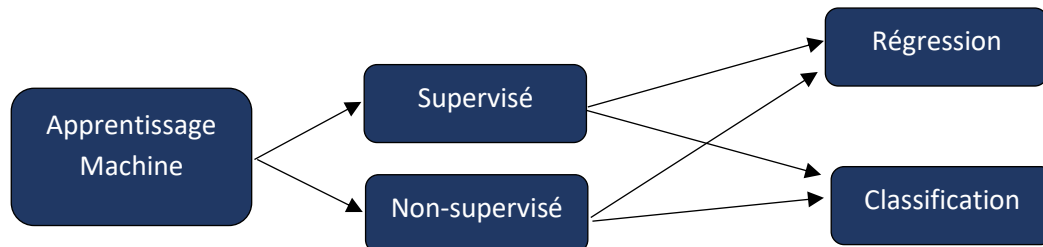


Figure 2.8 : Types d'apprentissage automatisé (Machine Learning) [25]

a. L'apprentissage supervisé :

L'apprentissage supervisé, également appelé apprentissage automatique supervisé, est une sous-catégorie de l'apprentissage automatique et de l'intelligence artificielle. Il se caractérise par l'utilisation de jeux de données étiquetés qui entraînent des algorithmes permettant de classer des données ou de prédire des résultats avec précision. Au fur et à mesure que les données en entrée sont introduites dans le modèle, celui-ci ajuste ses pondérations jusqu'à ce que le modèle soit correctement ajusté [26].

b. L'apprentissage non supervisé :

L'apprentissage non supervisé utilise des données non étiquetées. À partir de ces données, il découvre des modèles qui aident à résoudre les problèmes de regroupement ou d'association. Cette méthode est particulièrement utile lorsque les experts de domaine ne sont pas certains des propriétés communes d'un jeu de données. Les algorithmes de regroupement courants sont les modèles hiérarchiques.

2.3.3.2 La classification « KNN » et « SVM » :

La phase de classification utilise un algorithme pour classer avec précision les données de test dans des catégories spécifiques. Elle reconnaît des entités spécifiques dans l'ensemble de données et tente de tirer des conclusions sur la façon dont ces entités doivent être étiquetées ou définies. Les algorithmes de classification courants sont les machines à vecteurs de support (SVM), les k plus proches voisins (k-NN), qui sont décrits plus en détail ci-dessous [26].

a. Les machines à vecteurs de support (SVM) :

Les SVMs sont une famille d'algorithmes d'apprentissage automatique qui permettent de résoudre des problèmes tant de classification que de régression ou de détection d'anomalie.

Les SVMs ont été développés dans les années 1990. Leur principe est simple : ils ont pour but de séparer les données en classes à l'aide d'une frontière aussi « simple » que possible, de telle façon que la distance entre les différents groupes de données et la frontière qui les sépare soit maximale. Cette distance est aussi appelée « marge ». Ainsi cela revient à

trouver un classificateur qui va séparer les données et maximiser la distance entre ces deux classes. Ce classificateur peut être linéaire ou non linéaire et est appelé hyperplan [24].

a.1) Les types des SVM :

Il existe deux types principaux des SVM :

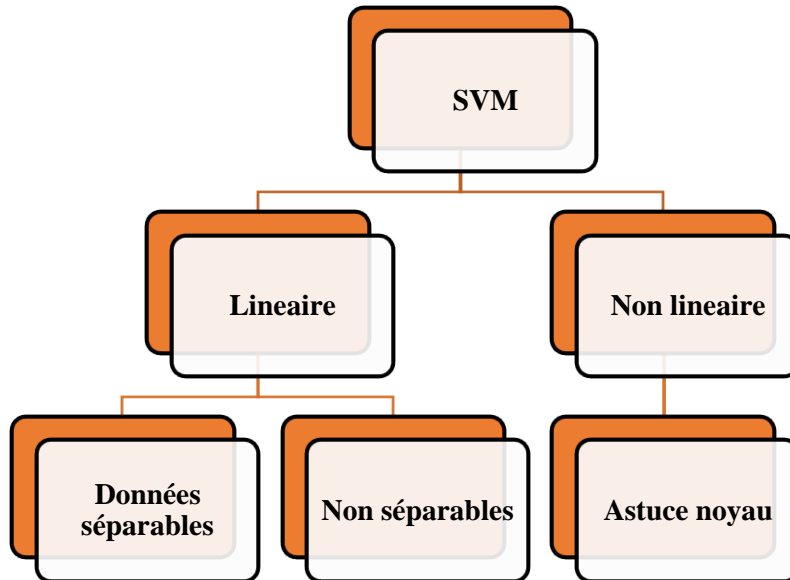


Figure 2.9: type de SVM

➤ linéaire

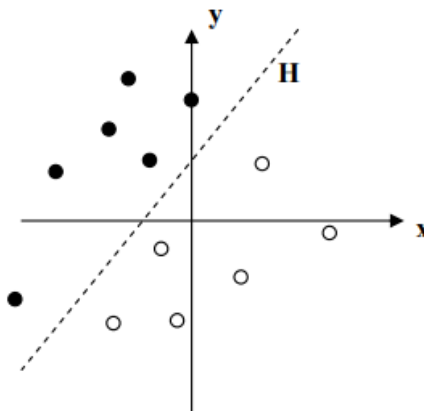


Figure 2.10 : Exemple d'un hyperplan séparateur.[24]

Les points les plus proches, qui seuls sont utilisés pour la détermination de l'hyperplan, sont appelés vecteurs de support.

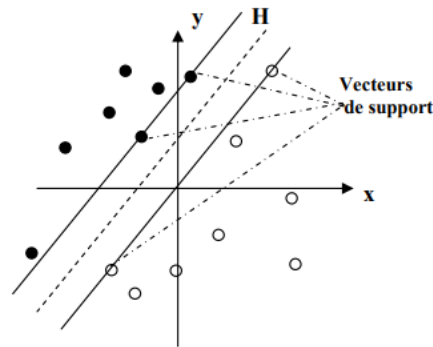


Figure 2.11 : Exemple de vecteurs de support.[24]

Notions de base :

- **Hyperplan** : sépare les deux ensembles de points.
- **Vecteurs de support** : points les plus proches qui déterminent l'hyperplan.
- **Marge** : hyperplan dont la distance minimale aux exemples est maximale.[29]

➤ **non linéairement séparable :**

Précédemment, nous avons décrit le principe des SVM dans le cas où les données sont linéairement séparables. Lorsqu'un problème n'est pas linéairement séparable dans l'espace d'entrée, le SVM linéaire ne peut pas trouver un hyperplan de séparation robuste qui minimise le nombre de points de données mal classés et qui se généralise bien. Pour cela, on utilise l'astuce du noyau, qui permet d'utiliser un classifieur linéaire pour résoudre un problème non linéaire. L'idée est de transformer l'espace de représentation des données d'entrée en un espace de plus grande dimension, dit « espace de redescription », à l'aide d'une fonction ϕ , où un classifieur linéaire peut être utilisé pour obtenir de bonnes performances. La transformation d'espace inverse ϕ^{-1} , permet de retourner à l'espace d'origine après la résolution dans le nouvel espace. En effet, intuitivement, plus grande est la dimension de l'espace de redescription, plus élevée sera la probabilité de pouvoir trouver un hyperplan séparateur entre les exemples.

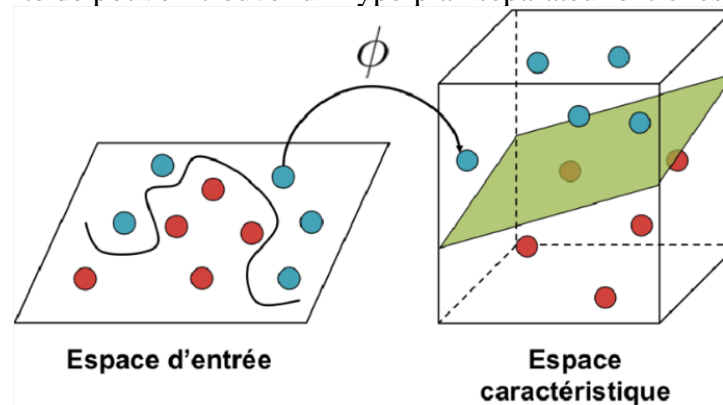


Figure 2.12 : Classifieur SVM (dans le cas non-linéaire)

L'utilisation de l'astuce du noyau rend les SVM plus intéressants et surtout plus riches puisqu'on peut toujours chercher de nouveaux noyaux qui peuvent être mieux adaptés à la tâche qu'on veut accomplir. Les trois noyaux les plus utilisés sont : le noyau linéaire, le noyau polynomial et le noyau gaussien, et le noyau noté RBF (Radial Basis Function).

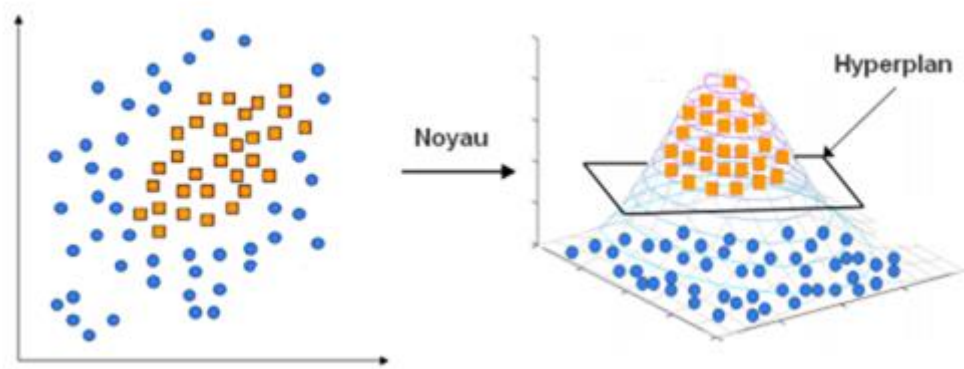


Figure 2.13 : illustration d'un SVM non-linéaire (kernel trick).

Le noyau linéaire : est un simple produit scalaire : $K(x_i, x_j) = \langle x_i, x_j \rangle$

Le noyau polynomial : permet de représenter des frontières de décision par des polynômes de degré d : $K(x_i, x_j) = (a * \langle x_i, x_j \rangle + b)^d$

La dimension de l'espace transformé induit par un noyau polynomial est de l'ordre :

$$\frac{(p+d)!}{p!d!}, \quad (2.17)$$

Où :

p : est la dimension de l'espace de départ.

d : est le degré de noyaux polynomiale.

Le noyau gaussien : a la forme suivante :

$$K(x_i, x_j) = \exp \left(- \frac{\|x_i - x_j\|^2}{2\sigma^2} \right) \quad (2.18)$$

Le paramètre σ permet de régler la largeur de la gaussienne. En prenant un σ grand, la similarité d'un exemple par rapport à ceux qui l'entourent sera assez élevée, alors qu'on prenant un σ tendant vers 0, l'exemple ne sera similaire à aucun autre.

a.2) Avantages des SVM :

- Les exemples de test sont comparés juste avec les supports vecteurs et non pas avec tous les exemples d'apprentissage [30].
- Décision rapide [30].
- Les SVM sont efficaces dans les espaces de grande dimension [24].

a.3) Inconvénients des SVM :

- Elles utilisent des fonctions mathématiques complexes pour la classification [31].

b. Les k plus proches voisins (KNN) :

La méthode des k plus proches voisins (k-Nearest Neighbor, KNN) est une méthode supervisée. L'algorithme KNN est l'un des plus simples de tous les algorithmes d'apprentissage automatique. La méthode KNN suppose que les données se trouvent dans un espace de caractéristiques. Cela signifie que les points de données sont dans un espace métrique. La méthode des k plus proches voisins est utilisée pour la classification et la régression. Dans les deux cas, l'entrée se compose des k données d'entraînement les plus proches dans l'espace de caractéristiques [31].

Pour tester la similarité entre deux vecteurs, une distance est utilisée. Elle permet de mesurer le degré de différence entre deux vecteurs. Il existe plusieurs types de distance parmi lesquels on trouve :[31]

La distance Euclidienne :

$$d(x,y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2.19)$$

Où :x et y sont des vecteurs.

b.1) Avantages des KNN :

- La méthode des k plus proches voisins est efficace si les données sont larges et incomplète.
- Cette méthode est l'une des plus simples de tous les algorithmes d'apprentissage automatique.

b.2) Inconvénients des KNN :

- Le temps de prédiction est très long puisqu'on doit calculer la distance de tous les exemples [31].

2.4 Conclusion :

En conclure que l'efficacité des modèles de machine Learning dans le traitement des signaux audio dépend largement de la qualité de l'extraction des caractéristiques et de la pertinence des méthodes de classification utilisées. MFCC, GTCC, LPC, et ZCR fournissent des informations détaillées sur la structure sonore, tandis que SVM et KNN offrent des mécanismes robustes pour la discrimination et la classification de ces caractéristiques. Cette combinaison est donc fondamentale pour développer des systèmes d'apprentissage automatique plus sophistiqués et performants dans des domaines tels que la reconnaissance de la parole, l'analyse musicale, ou encore la surveillance environnementale et la détection de présence de drones. Nous allons présenter dans le chapitre suivant les notions de l'apprentissage profond explorant ses fondements théoriques, ses architectures courantes et ses applications pratiques dans notre domaine.

Chapitre 3 : La détection de la présence de drone par Apprentissage Profond

3.1 Introduction :

Le Deep Learning (Apprentissage profond), une branche de l'intelligence artificielle, est une approche puissante permettant aux machines d'apprendre à partir de données non structurées ou non étiquetées. Inspiré par le fonctionnement du cerveau humain, le Deep Learning utilise des réseaux neuronaux profonds pour extraire des motifs et des caractéristiques complexes, permettant ainsi aux systèmes informatiques de prendre des décisions autonomes, de reconnaître des formes et de résoudre des problèmes dans divers domaines tels que la vision par ordinateur, le traitement du langage naturel et bien d'autres encore. Cette technologie révolutionnaire a récemment connu des avancées spectaculaires, ouvrant la voie à de nouvelles possibilités et applications dans de nombreux secteurs.

3.2 Définition de l'apprentissage profond (Deep Learning DL) :

Dans l'apprentissage automatique, un programme analyse un ensemble de données afin de tirer des règles qui permettent de tirer des conclusions sur de nouvelles données. L'apprentissage profond est basé sur les « réseaux de neurones artificiels », constitués de milliers de neurones qui effectuent chacune de simples opérations. Les sorties d'une première couche servent comme entrée d'une deuxième couche et ainsi de suite. Les avancées du Deep Learning ont été possible grâce à la grande puissance des ordinateurs et au développement de vaste ensemble de données « big data » [32].

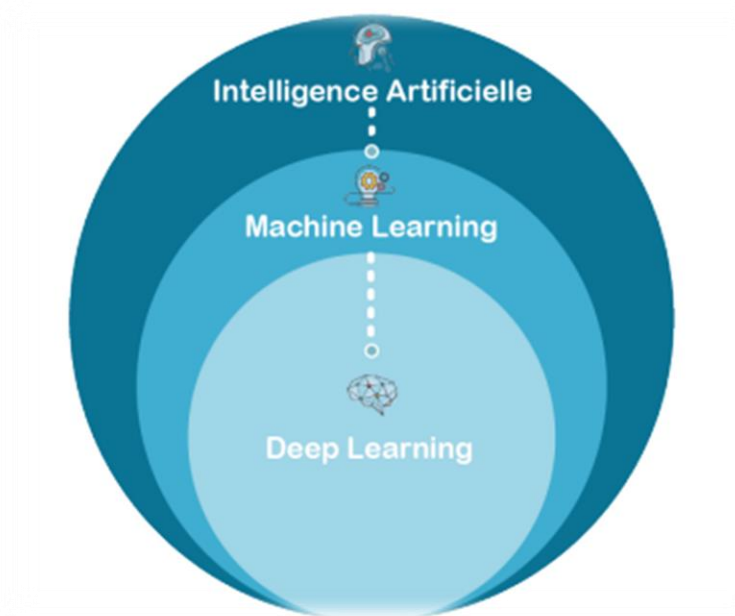


Figure 3.1: La relation entre l'intelligence artificielle le machine Learning et le Deep Learning [32].

3.3 Les réseaux de neurones :

3.3.1 Définition des réseaux de neurones artificiels :

Les réseaux de neurones artificiels sont des réseaux qui connectes des processeurs élémentaires fonctionnant en parallèle. Chaque processeur élémentaire calcule une sortie unique sur la base des informations qu'il reçoit. Toutes structure hiérarchique de réseaux est évidemment un réseau. Les réseaux de neurones artificiels consistent en des modèles plus ou moins inspirés du fonctionnement cérébral de l'être humain en se basant principalement sur le concept de neurones [32].

3.3.2 Les neurones biologiques :

Le cerveau humain est constitué de près de 100 milliards de neurones reliés entre eux, au cœur du neurones on retrouve son noyau, ou « corps cellulaire ». Ces neurones sont connectés à des branches appelées « dendrites » qui transmettent les signaux depuis l'extérieur vers le corps cellulaire ce dernier traite l'information et la transmet via son axone, à la fin de l'axone se trouve les synapses.

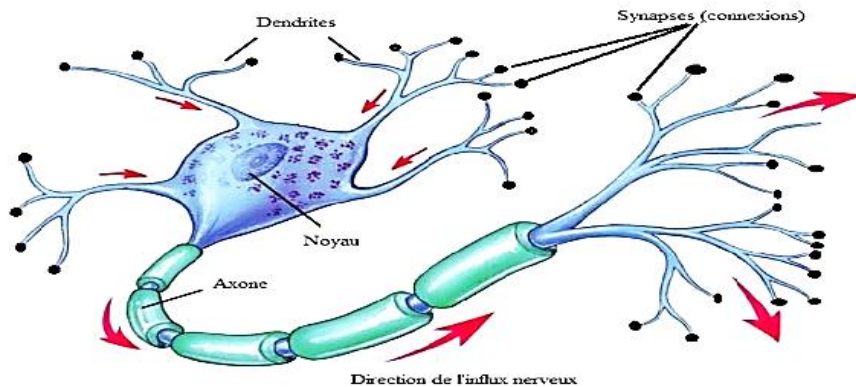


Figure 3.2 : Modèle d'un neurone biologique.

3.3.3 Les neurones formels :

Un neurone formel est une fonction algébrique non linéaire et bornée, dont la valeur dépend des paramètres appelés poids. Les variables utilisées dans cette fonction sont habituellement appelées « entrées » du neurone, et la valeur obtenue à partir de la fonction est appelée sa « sortie ». On peut représenter graphiquement un neurone comme indiqué sur la figure 3.3.

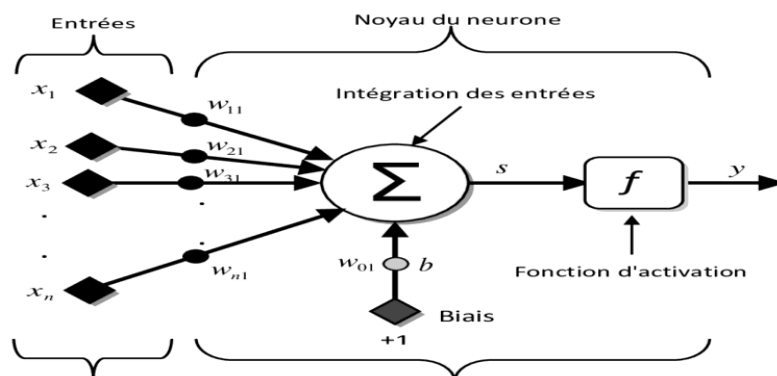


Figure 3.3 : Modèle d'un neurone artificiel

Le modèle du neurone formel est proposé par W.M. Culloch et W. pitts en 1943, on trouve :

Les x_i qui sont les valeurs d'entrées.

Les w_{ij} qui sont les poids synaptiques du neurone j .

Ces poids pondèrent les entrées et peuvent être modifiés par apprentissage.

Biais : permet d'ajouter de la flexibilité au réseau généralement +1 ou -1.

Noyau : intègre toutes les entrées et le biais et calcule la sortie du neurone selon une fonction d'activation.

Un neurone est essentiellement constitué d'un intégrateur qui effectue la somme pondérée de ses entrées. Le résultat de cette somme est ensuite transformé par une fonction de transfert f qui produit la sortie y du neurone. Les n entrées du neurone correspondent au vecteur $= [x_1, x_2, \dots, x_n]^T$, alors que $w = [w_{11}, w_{21}, \dots, w_{n1}]^T$ représente le vecteur des poids du neurone. La sortie de l'intégrateur est donnée par l'équation suivante :

$$S = \sum_{i=1}^n w_i x_i \pm b \quad (3.1)$$

Cette sortie correspond à une somme pondérée des poids et des entrées plus le biais b de neurone. Le résultat S de la somme pondérée s'appelle le niveau d'activation de neurone. Lorsque le niveau d'activation atteint ou dépasse le seuil b alors l'argument de f devient positif (ou nul). Sinon, il est négatif.

Les réseaux de neurones sont caractérisés par l'architecture (l'organisation de neurones), l'apprentissage (méthode de détermination), et par leur fonction d'activation.

3.3.3.1 Fonctions d'activation :

Différentes fonctions d'activation peuvent exister pour activer le neurone dont elles sont énumérées dans le tableau 1. Les fonctions les plus utilisées sont « seuil » (en anglais « hard limit »). « Linéaire » et « sigmoïde ».

Table 3.1 : Différents fonctions d'activation

Nom de la fonction	Relation entrée/sortie	Icone
Seuil	$y = 0$ si $(s < 0)$ $y = 1$ si $(s > 0)$	
Seuil symétrique	$y = -1$ si $(s < 0)$ $y = 1$ si $(s \geq 0)$	
Linéaire	$y = s$	
Linéaire saturée	$y = 0$ si $(s \leq 0)$ $y = s$ si $(0 \leq s \leq 1)$ $y = 1$ si $(s > 1)$	
Linéaire saturée symétrique	$y = -1$ si $(s < -1)$ $y = s$ si $(-1 \leq s \leq 1)$ $y = 1$ si $s > 1$	
Linéaire positive (ReLU)	$y = 0$ si $(s \leq 0)$ $y = s$ si $(s \geq 0)$	
Sigmoïde	$y = \frac{1}{1 + e^{-s}}$	
Tangente hyperbolique (Tanh)	$y = \frac{e^s - e^{-s}}{e^s + e^{-s}}$	

3.4 Les trois types de couches d'un réseau de neurones :

Les combinaisons nœud/connexions sont réparties sur plusieurs couches :

- Couche d'entrée : celle qui reçoit les données.
- Couche cachée : une ou plusieurs couches intermédiaires de traitement des données.
- Couche de sortie : les informations issues du traitement par le réseau de neurones.

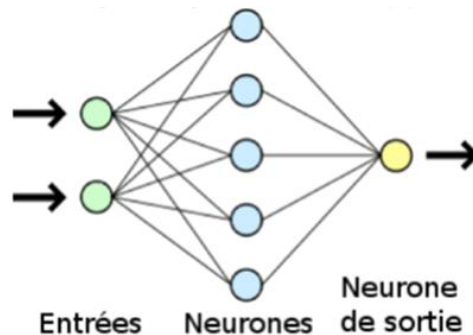


Figure 3.4: schéma simple de 3 types de couches d'un réseau de neurones [33]

3.5 Architecture de réseau de neurones :

Selon la topologie de connexion des neurones, on peut les classer en deux grandes catégories :

- Réseaux non bouclés (statique ou feed forward).
- Réseaux bouclés (dynamique, feed back ou récurrent).

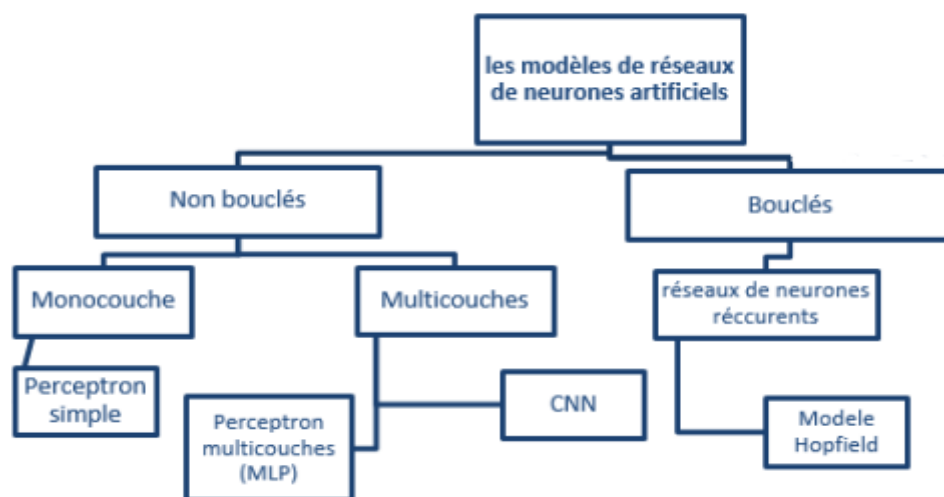


Figure 3.5: les types de réseaux de neurones artificiels » [32]

3.5.1 Les réseaux de neurones non bouclés :

Un réseau de neurones non bouclés appelé aussi famille des réseaux à propagation avant ou feed-forward : est représenté graphiquement par un ensemble de neurones « connectés » entre eux, l'information circulant des entrées vers les sorties « retour en arrière » ; si l'on représente le réseau comme un graphe dont les nœuds sont les neurones et les arêtes les « connexions » entre ceux-ci, le graphe d'un réseau non bouclé est acyclique [32].

A. Réseaux de neurones monocouches :

- **Perception simple :**

Ce réseau est dit simple car il ne se compose que de deux couches : une couche d'entrée et une couche de sortie ce qui implique une seule matrice de poids. L'ensemble des unités de la couche d'entrée sont connectés à celles de la couche de sortie.

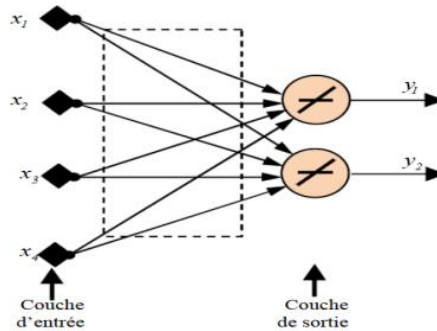


Figure 3.6: Schéma d'un réseau de neurones monocouche.

B. Réseaux de neurones multicouches :

- **La perception multicouche :**

Désigné par le sigle MLP (pour Multi-layer Perception). Le perceptron multicouche se compose d'une couche d'entrée, d'une couche de sortie et d'une ou plusieurs couches cachées. Si le réseau possède n couches, alors il possède $n-1$ matrice de poids (une entre chaque suite de couches).

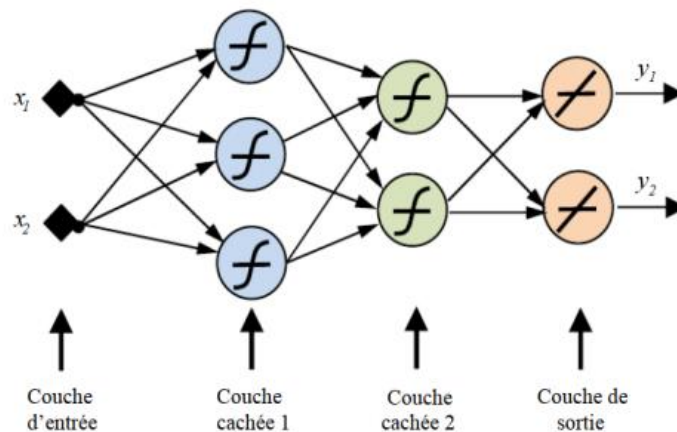


Figure 3.7: Schéma d'un réseau de neurones non bouclé (Perceptron multicouches).

- **Le réseau de neurones à convolution :**

(CNN pour Convolutionnal Neural Network) sont aussi des réseaux feed-forward et qui correspondent, à un empilement de perceptron multicouche. Chacun traitant une portion de l'information globale. Ces réseaux sont surtout utilisés pour la reconnaissance d'image, de vidéos ou encore dans le traitement naturel du langage.[33]

3.5.2 Les réseaux de neurones bouclés :

Contrairement aux réseaux de neurones non bouclés dont le graphe de connexions est acyclique, les réseaux de neurones bouclés peuvent avoir une topologie de connexions quelconque, comprenant notamment des boucles qui ramènent aux entrées la valeur d'une ou plusieurs sorties. Pour qu'un tel système soit causal, il faut évidemment qu'à toute boucle soit associé un retard : un réseau de neurones bouclés est donc un système dynamique, régi par des équations différentielles.

Il s'agit donc de réseaux de neurones avec retour en arrière (feedback network or recurrent network), (figure 3.8).

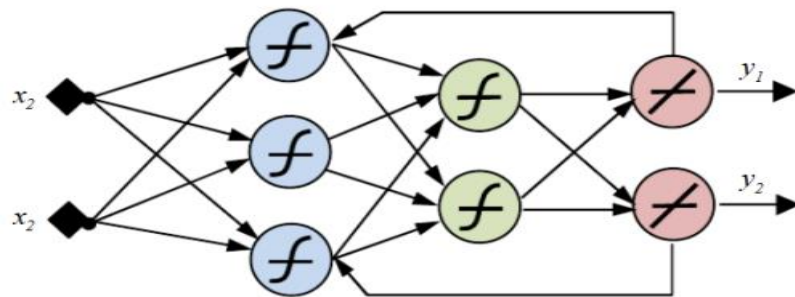


Figure 3.8 : Schéma de réseau de neurones bouclé.

A. Famille des réseaux de neurones récurrents (RNN recurrent network) :

Les RNNs sont des réseaux de neurones qui comportent des cycles dans leurs graphe de connectivité. Ces cycles permettent au réseau d'entretenir une information en l'envoyant à lui-même. Cela change la dynamique du réseau de neurones et l'amène à s'auto-entretenir. Ces modèles étaient souvent reconnus pour le traitement automatique de la parole, et plus généralement de séquence, car leurs caractéristiques leur permettent d'apprendre, de stocker et de prendre en compte l'information contextuelle passée lors de traitement de l'information à l'instant présent.

- **Le modèle de Hopfield :**

C'est un réseau qui se compose d'une seule couche où toutes les unités sont interconnectées. Il s'agit d'une mémoire auto-associative : constituée d'une seule couche qui représente à la fois l'entrée du réseau et sa sortie. Autrement dit, une mémoire auto-associative a autant d'entrées que de sorties. Ce réseau est assimilé à une mémoire adressable par son contenu : les connaissances mémorisées étant distribuées dans le réseau et non localisées à une adresse, il est possible de récupérer l'entièreté d'une donnée, juste en présentant une version dégradée (partielle ou bruitée).

3.6 Les applications des réseaux de neurone :

Voici de nombreux domaine dans lesquelles les réseaux de neurones sont appliqués avec succès [33] :

- Traitement du signal.
- Prédiction.
- Classification.
- Prétraitement des données.
- Reconnaissance de formes.

- Analyse d'image et synthèse vocale.
- Diagnostique et suivi médicale.
- Sécurité.

3.7 Les avantages et les inconvénients des réseaux de neurones :

Le réseau neuronal peut exécuter n'importe quelle fonction non linéaire jusqu'à ce qu'il atteigne un certain degré de fiabilité. Il peut également implémenter les fonctions dynamiques et avoir n'importe quel nombre d'entrées et de sorties.

Nous énumérons les avantages et les inconvénients des réseaux de neurones artificiels dans le tableau ci-dessous :

Table 3.2 : les avantages et les inconvénients des réseaux de neurones.

Les avantages des réseaux de neurones	Les inconvénients des réseaux de neurones
<ul style="list-style-type: none"> -La capacité d'exprimer n'importe quelle fonction, qu'elle soit linéaire ou non, simple ou complexe. -Une facilité d'apprendre à partir d'exemples représentatifs par recto propagation d'erreur. -L'apprentissage (ou la construction de modèles) se fait automatiquement. -Insensible aux données non fiable ou un bruit. -Facile à utiliser, le travail personnel exigeant beaucoup moins qu'une analyse statistique classique. Aucune compétence en mathématiques et calculs statistiques requis. -Moins de mauvais comportements en cas de petite quantité de données. 	<ul style="list-style-type: none"> -Il n'y pas de méthode systématique de définir la topologie optimale du réseau et le nombre de neurones à placer dans la couche cachée. -Le choix des valeurs initiales des poids du réseau et le réglage du pas d'apprentissage jouent un rôle important dans la vitesse de convergence. -Le problème de sur-apprentissage qui ne se généralise pas. -Les connaissances acquises par le réseau neuronal sont encodées par la valeur du poids synaptique, le réseau neuronal est donc une boîte noire et les utilisateurs ne peuvent pas comprendre ces connaissances.

3.8 Procédure de développement d'un réseau de neurones :

Dans les paragraphes suivants, nous expliquerons le développement classique d'un réseau de neurones, que nous décomposerons en sept étapes principales [33] :

➤ Collecte des données :

Le but de cette étape est de collecter des données pour développer et tester des réseaux de neurones. Pour les applications de données réelles, l'objectif est de collecter suffisamment de données pour former une base de données représentative qui peut fonctionner lors de l'utilisation du système neuronal. Les fonctions réalisées par les calculs statistiques, les modèle qu'il fournit, ne sont valable que dans les domaines fournis.

➤ Analyse des données :

En générale, il est préférable d'effectuer une analyse des données pour identifier les caractéristiques discriminantes et détecter ou distinguer les données. Ces caractéristiques sont les entrées du réseau neuronal et doivent généralement être représentées de manière représentative. La détermination des caractéristiques influencera la taille du réseau (et donc le temps de simulation), les performances du système (capacité de séparation, taux de détection) et le temps de développement (temps d'apprentissage).

L'analyse statistique des données peut aider à éliminer les valeurs aberrantes et les données redondantes.

Pour la classification, l'expérimentateur doit déterminer :

- Le nombre de classes appartenant aux données.
- La classe qui appartient à chaque donnée.

➤ **Séparation des bases de données :**

Afin de développer une application à base de réseaux de neurones, il est nécessaire de disposer de deux bases de données : l'une pour l'apprentissage et l'autre pour tester le réseau générer et déterminer ses performances. Pour contrôler la phase d'apprentissage, il est souvent préférable de disposer d'une troisième base de données, appelée « base de validation croisée », qui permet d'arrêter l'apprentissage de manière appropriée pour obtenir de bons résultats de généralisation le critère d'arrêt de l'apprentissage est généralement calculé en fonction de cout, qui caractérise la différence entre la valeur de sortie obtenue et la valeur de référence (la réponse attendue pour chaque exemple donné).

Il n'y a pas de règle pour déterminer quantitativement ce partage, qui est généralement un compromis basé sur la quantité de données disponibles et le temps d'apprentissages, cependant, chaque base de données doit satisfaire aux contraintes de représentativité de chaque donnée et doit généralement refléter la vraie distribution, c'est-à-dire la probabilité d'occurrences de diverses classes.

➤ **Choix d'un réseau de neurones :**

Il existe de nombreux types de réseaux de neurones, chacun ayant ses propres avantages et inconvénients. Le choix du réseau peut dépendre de :

- La tâche à réaliser (classification, association, contrôle de processus, séparation aveugle de sources...).
- La nature de données.

Ce choix dépend également de notre maîtrise ou connaissance de certains réseaux, ou du temps disponible pour tester des architectures plus efficaces.

➤ **Mise en forme des données pour un réseau de neurones :**

Habituellement, la base de données doit être prétraitée pour s'adapter à l'entrée et à la sortie du réseau neuronal. Le prétraitement actuel implique la réalisation d'une normalisation appropriée, qui prend en compte l'amplitude des valeurs acceptées par le réseau.

➤ **Apprentissage du réseau de neurones :**

Tous les modèles des réseaux neurones nécessitent un apprentissage, notons que plusieurs types d'apprentissage peuvent d'adapter au même type de réseau de neurone. Les critères de choix sont généralement la vitesse de convergence ou les performances de généralisation.

Certains algorithmes d'apprentissage sont chargés de déterminer les paramètres architecturaux du réseau neuronal. Si ces techniques ne sont pas utilisées, Les meilleurs paramètres d'architectures peuvent être obtenus en comparant les valeurs de performance de différentes architectures de réseaux neuronaux.

➤ **Validation :**

Après avoir formé le réseau de neurones, il est impératif de le tester dans une base de données différentes de celle utilisée pour l'apprentissage ou la validation croisée. Ce test peut évaluer les performances du système neuronal et détecter le type de données qui ont un problème à l'origine. Si les performances ne sont pas satisfaisantes, il est nécessaire de modifier l'architecture de réseau ou le modifier la base d'apprentissage (caractéristique distinctive ou représentativité des données de chaque classe).

3.9 Architecture des réseaux de neurones profonds :

Il existe un grand nombre d'architecture profondes. Nous allons dans ce qui suit détailler les réseaux de neurones convolutifs (CNNs).

3.9.1. Les réseaux de neurones convolutifs (CNNs) :

Ils sont un type de réseau de neurones spécialisés pour le traitement de données ayant une topologie semblable à une grille. Leurs principes sont le filtrage. Les exemples comprennent des données de type série temporelle sont considérées comme une grille 1D en prenant des échantillons à des intervalles de temps réguliers, les données de type image se représentent en 2D ou 3D, il y a deux dimensions qui correspondent à la largeur et à la hauteur de l'image et une troisième dimension qui correspond à la composante couleur.

Le nom « réseau de neurones convolutifs » indique que le réseau emploie une opération mathématique appelée convolution à la place de la multiplication matricielle dans au moins une de leurs couches. La convolution est une opération linéaire spéciale.

Le rôle du ConvNet est de réduire les dimensions des images sous forme plus facile à traiter, sans perdre les fonctionnalités qui sont essentielles pour obtenir une bonne prédiction.

Il existe diverses architectures de réseaux de neurones convolutifs (CNN) qui ont été cruciales pour la création d'algorithmes alimentant l'intelligence artificielle et continueront de jouer un rôle clé dans son développement futur.

3.9.2. Architecture de réseau neuronal convolutif 1D :

L'architecture d'un réseau de neurones convolutifs (CNN) 1D est adaptée pour traiter des données séquentielles, telles que des séries temporelles ou des signaux audios. Un CNN 1D commence par des couches de convolution qui utilisent des filtres pour extraire des caractéristiques locales le long de la dimension séquentielle de l'entrée. Ces filtres glissent sur la séquence avec une certaine stride et peuvent inclure du padding pour contrôler la taille de la sortie. Les sorties des couches de convolution passent ensuite par des fonctions d'activation, telles que ReLU, qui introduisent des non-linéarités nécessaires pour capturer des relations complexes. Pour réduire la dimensionnalité des cartes de caractéristiques et prévenir le surajustement, des couches de pooling, comme le max pooling ou l'average pooling, sont utilisées. Ces couches réduisent la longueur des séquences tout en conservant les informations les plus importantes. Après plusieurs couches de convolution et de pooling, les cartes de caractéristiques sont souvent aplaties et passées à travers des couches complètement connectées, qui permettent de combiner les caractéristiques extraites en des vecteurs de probabilité pour la classification ou d'autres tâches de prédiction. La fonction de perte, souvent pour les tâches de classification, mesure la différence entre les prédictions du modèle et les étiquettes réelles.

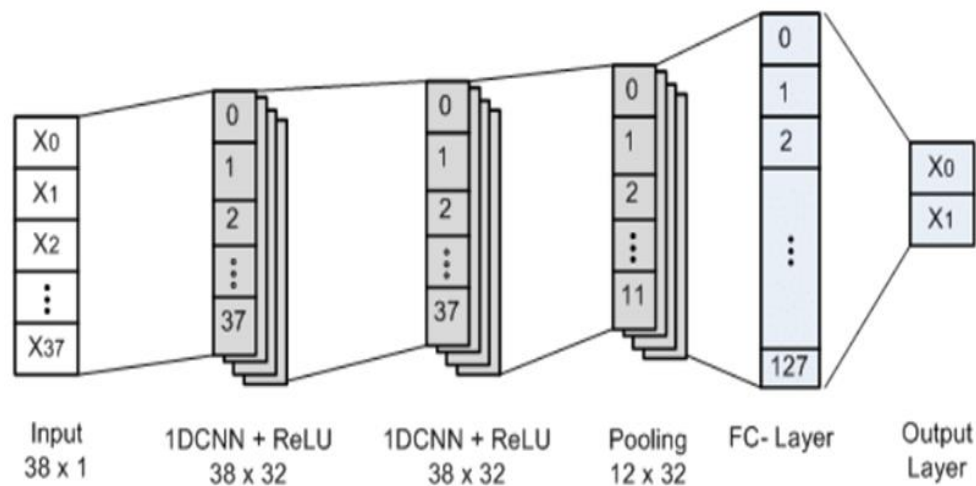
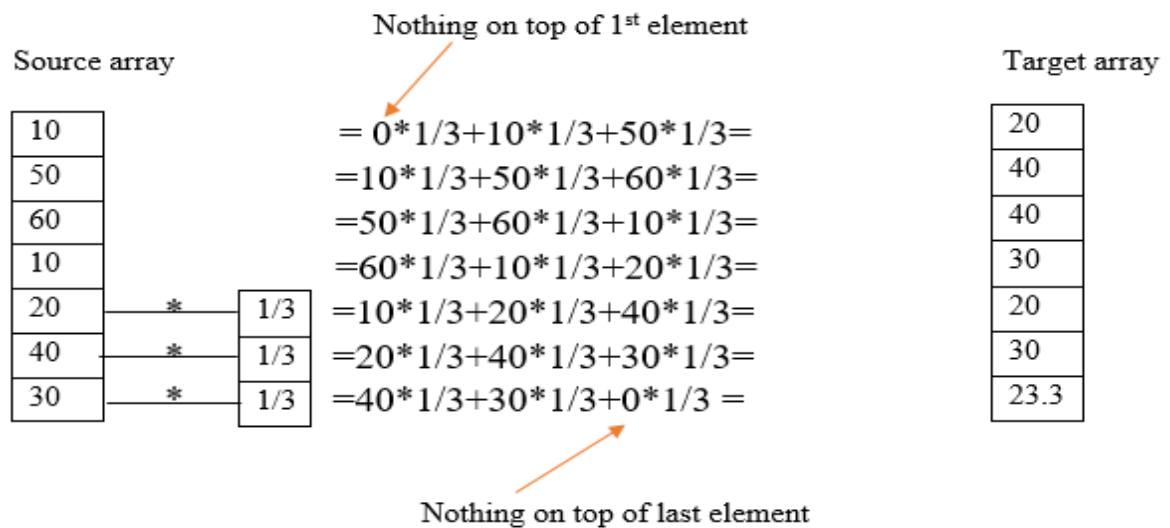


Figure 3.9 : Architecture de réseau neurone convolutif 1D

L'exemple d'architecture de CNN1D présenté dans la figure (3.9) se compose de 2 couches de convolution, 1 couche de pooling, 1 couche entièrement connectée et la couche de sortie.

Exemple :



Les réseaux de neurones à convolution 1D (1D CNN) sont largement utilisés dans diverses applications. Leurs architectures permettent de détecter les motifs locaux et globaux grâce à des filtres convolutifs qui se déplacent le long de la séquence. En utilisant ces filtres, les CNN 1D peuvent extraire des caractéristiques importantes des données, et efficaces pour la classification et la prédiction. Ils sont souvent combinés avec d'autres techniques comme le pooling et le dropout pour améliorer les performances du modèle et éviter le surapprentissage.

3.9.3 Architecture de réseau neurone convolutif 2D :

L'architecture d'un réseau neuronal convolutif 2D (CNN 2D) est conçu pour traiter et reconnaître des images, Elle se compose de plusieurs couches chacune ayant une fonction spécifique.

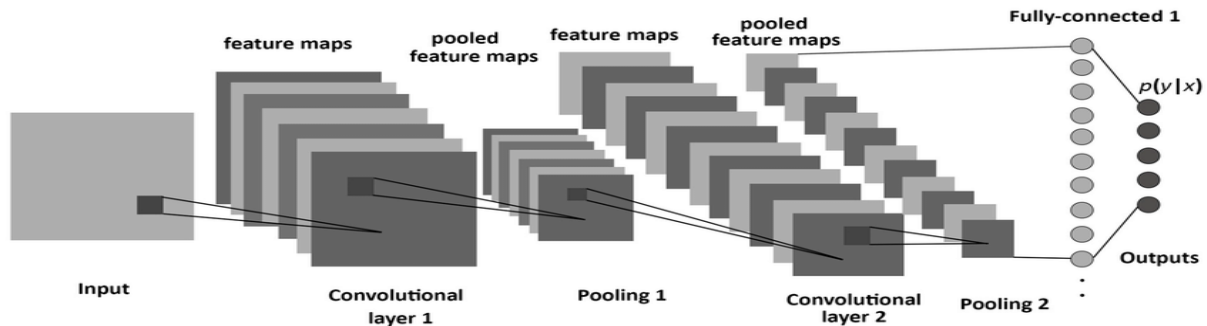


Figure 3.10 : Architecture de réseau neurone convolutif 2D

A. Couche de convolution : la convolution est une opération mathématique utilisée pour traiter et reconnaître les images. Elle agit sur une image de manière similaire à un filtrage. Exemple d'une matrice d'image 5*5 :

1	1	1	0	0
0	1	1	1	0
0	0	1	1	1
0	0	1	1	0
0	1	1	0	0

Matrice « image ».

Exemple d'une matrice utilisée comme filtre.

1	0	1
1	0	1
1	0	1

Matrice « filtre ».

La matrice obtenue après convolution :

4	4	4
3	5	6
3	4	4

Le résultat de convolution

A.1 Le fonctionnement d'un filtre de convolution :

Etape 1 : Tout d'abord, on spécifie la taille de la fenêtre de filtre, située en haut à gauche de l'image.

Etape 2 : Ensuite, cette fenêtre de filtre, représentant la caractéristique recherchée, se déplace progressivement de gauche à droite avec un pas prédéfini, jusqu'à parcourir toute l'image.

Etape 3 : A chaque position de la fenêtre sur l'image, un calcul de convolution est effectué, produisant une carte d'activation ou feature map. Cette carte indique où les caractéristiques sont localisées dans l'image : une valeur plus élevée dans la feature map correspond à une portion de l'image qui ressemble davantage à la caractéristique recherchée.

Pour appliquer un filtre à une image, chaque pixel de l'image est multiplié par la valeur correspondante dans le filtre (par exemple, 1×1 , 1×0 , 1×1 , etc.). Ensuite, ces produits sont additionnés par obtenir une valeur unique, ici '4', qui constitue un pixel de la nouvelle image obtenue après la convolution.

B. Méthode de sous échantillonnage : le Max-pooling :

Le Max-pooling est un processus de discrétisation basé sur des échantillons. Son objectif est de sous-échantillonner une représentation d'entrée (image, matrice de sortie de couche cachée, etc.) en réduisant sa dimension. De plus, son intérêt est qu'il réduit le coût de calcul en réduisant le nombre de paramètres à apprendre et fournit une invariance par translation (si une petite translation ne modifie pas le maximum de chaque région restera le même et donc la nouvelle matrice créée restera indiquée) [35].

Exemple de max-Pooling 4×4 :

12	20	30	15
25	12	2	99
34	70	47	4
113	100	25	12

Matrice « image ».

Le résultat de max-Pooling 2×2 :

25	99
113	47

Matrice « filtre ».

Dans l'exemple ci-dessus, nous avons une matrice « image » de taille 4×4 qui représente notre entrée initiale et une matrice « filtre » de taille 2×2 que nous utiliserons sur notre entrée. Lorsque le filtre balaye différentes zones de l'image, l'opération de max-Pooling sélectionne la valeur maximale de chaque zone balayée pour créer une nouvelle matrice de sortie. Chaque élément de cette matrice résultante correspond au maximum des valeurs dans chaque région balayée.

C. Couche de correction (ReLU) :

Pour améliorer l'efficacité du traitement on intercale entre les couches de traitement une couche qui opère une fonction mathématique sur les signaux de sortie [32].

Elle renvoie 0 si l'entrée $Z < 0$ et Z si $Z > 0$.

Z : représente l'entrée de la fonction d'activation pour unité donnée.

D. Couche entièrement connectée (Fully Connected FC) :

Les neurones de cette couche ont des connexions complètes à toutes les activations de la précédentes dans le but final de classifier l'image d'entrée.

E. Couche de perte (LOSS) :

Elle est normalement la dernière couche du réseau. Elle spécifie comment l'entraînement du réseau pénalise l'écart entre le signalé prévue et réel. Les fonctions les plus utilisées sont : Softmax, la perte par entropie, la perte euclidienne.

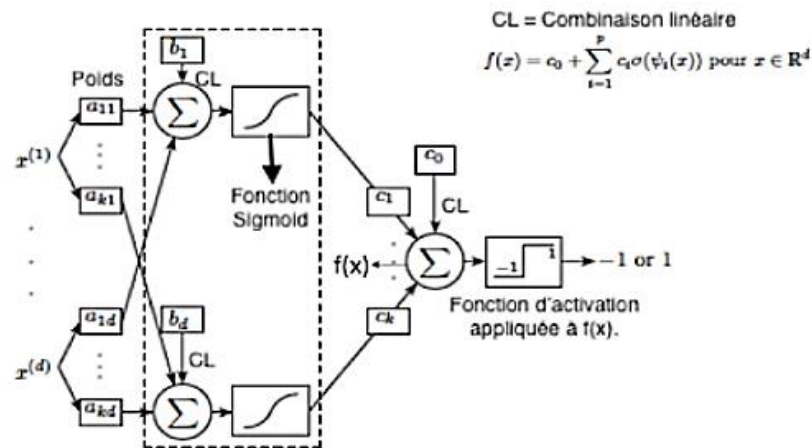


Figure 3.11 : Fonctionnement d'un réseau neuronal à 2 couches cachées [35]

Les réseaux de neurones à convolution 2D (2D CNN) sont largement utilisés pour analyser des images dans des domaines tels que la reconnaissance d'objet et la classification d'images. Ils fonctionnent en appliquant des filtres convolutifs à l'image d'entrée pour extraire des caractéristiques pertinentes. En ajustant ces filtres pendant l'entraînement, le réseau peut apprendre à reconnaître des motifs visuels complexes. En combinant des opérations de pooling pour réduire la taille des caractéristique extraites et des couches entièrement connectées pour la classification, les 2D CNN peuvent produire des prédictions précises sur une grande variété de tâches visuelle.

3.9.4 Comparaison entre le CNN1D et le CNN 2D :

Le tableau suivant représente une comparaison entre le CNN1D et le CNN 2D.

Table 3.3 : comparaison entre CNN 1D et CNN 2D

CNN1D	CNN2D
<ol style="list-style-type: none"> 1. Convient à l'analyse de données séquentielles comme les séries temporelles, le texte et les signaux audio émis par les drones. 2. Ils permettent de détecter la présence de drones en analysant les fréquences et les motifs des signaux radio. 3. Utilise des filtres convolutifs qui se déplacent de long de la séquence unidimensionnelle 4. Idéal pour capturer des motifs locaux et globaux dans les données séquentielles. 5. Architecture souvent plus simple en raison de la nature linéaire des données d'entrée. 	<ol style="list-style-type: none"> 1. Adapté à l'analyse d'images et à la vision par ordinateur telles que des photos ou des vidéos de drones. 2. Ils permettent de détecter et de localiser des drones dans des images en extrayant des caractéristiques de leur texture. 3. Applique des filtres sur des régions 2D de l'image pour capturer les caractéristiques spatiales. 4. Utilise des architectures plus complexes avec des couches de convolution et de Pooling empilées pour capturer des motifs différentes échelles spatiales. 5. Utilisé pour des tâches telles que la classification d'images, la détection d'objet et la segmentation sémantique.

3.10 La reconnaissance des signaux audios :

La reconnaissance audio par les méthodes de l'intelligence artificielle nécessite de convertir les signaux acoustiques en données numériques, puis d'en extraire le sens pour réaliser diverses tâches. Cette technologie nécessite plusieurs étapes essentielles, allant du traitement du signal sonore à sa conversion en un format que les ordinateurs peuvent interpréter.

Des microphones sont utilisés pour capter des sons et comprendre ce que les utilisateurs demandent. Grâce aux avancées en intelligence artificiels, cette technologie est devenue plus populaire. Elle permet aux ordinateurs de parler naturellement avec les gens pour répondre à leurs questions, recueillir des informations ou réaliser différentes tâches. Les progrès réalisés avec ces assistants intelligents ont également ouvert la porte à de nouvelles applications dans les entreprises.

3.10.1 Etapes de la reconnaissance d'audios :

Il existe plusieurs étapes nécessaires au développement d'un programme de reconnaissance audio efficace. Pour bien l'illustrer, prenons l'exemple d'une application de transcription de la narration(speech-to-texte) [36].

1. **La capture du signal sonore :** la collecte des données acoustique se fait à l'aide d'un microphone qui transforme les signaux sonores en impulsions électriques, lesquelles seront par la suite convertie en un signal analogique, qui sera enregistré sous un format numérique (lisible par un ordinateur).
2. **L'utilisation d'un module de traitement du signal :** les données acoustiques collectées sont des données brutes qui nécessitent plusieurs étapes de traitement de signal comme, par exemple, sa représentation suivant un domaine fréquentiel pour en extraire uniquement les informations nécessaires.
3. **L'utilisation d'un modèle de référence :** dans le cas de la reconnaissance vocale, on utilisera une base d'échantillons sonores considérés comme un modèle de référence pour mesurer, à titre d'exemple, la correspondance entre une séquence sonore qui représente un mot (échantillon d'un enregistrement) et une nouvelle séquence introduite dans le système, pour laquelle nous cherchons une correspondance.
4. **L'algorithme de correspondance :** l'algorithme de correspondance servira à mesurer la correspondance entre des séquences sonores introduites en entrée et la séquence sonores préenregistrées (représentant l'échantillon mentionné dans la section précédente) qui représentent les mots connus par le système. Les algorithmes d'apprentissage machine sont d'excellents candidats pour effectuer les correspondances (Matching algorithms). En effet, une fois les données acoustiques numérisées et normalisées, il est possible de faire appel à un modèle apprenant entrainer à reconnaître et à effectuer la correspondance entre les séquences sonores. Les algorithmes d'apprentissage machine modernes sont particulièrement efficaces dans l'exercice d'identification des similarités et des correspondances.
5. **L'utilisation des techniques de traitement du langage naturel :** une fois les séquences sonores transformées en une représentation textuelle, il est possible d'utiliser les techniques du traitement du langage naturel pour extraire de la sémantique à partir du texte, et faire en sorte que le programme de reconnaissance vocale comprenne le sens qui y est exprimé.

3.10.2 Les Application de la reconnaissance d'audios :

L'audio peut être considérée comme une empreinte capable d'identifier quelqu'un d'une manière unique tout comme le font les détecteur rétinien ou les lecteurs d'empreintes digitale. Il existe effectivement des technologies permettant d'identifier une personne par le biais. Elles sont utilisées pour la gestion des accès au sein des structures scientifique et gouvernementales de haute sécurité, et sont généralement accompagnées d'autres outils de reconnaissance comme les lecteurs rétinien ou les mots de passe [36], dans notre travail nous détectons la présence des drones en utilisant des méthodes de reconnaissance d'audio.

3.11 Reconnaissances d'objets dans une image :

La reconnaissance d'objet est une technique d'identification de l'objet présent dans l'image et la vidéo ; c'est l'une des applications les plus importantes de l'apprentissage automatique et l'apprentissage profond. L'objectifs de ce domaine est d'apprendre aux machines à comprendre (reconnaitre) le contenu d'une image comme la font les humains [37].

3.11.1 Classification d'image :

La tâche la plus fondamentale pour les réseaux de neurones convolutifs est la classification d'image qui est une tâche d'apprentissage automatique permettant de déterminer quels objets se trouvent dans une image ou une vidéo. Elle fait référence à la formation de modèles d'apprentissage automatique dans le but de découvrir quelles classes sont présentes tout en représentant la probabilité que l'image appartienne à l'une de plusieurs classes. La classification est utile pour décider si une image contient un objet/une anomalie ou non [37].

3.11.2 Localisation d'objet :

L'algorithme de localisation d'objet localise la position d'un objet dans l'image et le représente avec une boîte englobante. Il prend une image en entrée et affiche l'emplacement de la boîte englobante sous la forme de (position, hauteur et largeur) [37].

3.11.3 Fonctionnement de la reconnaissance d'image :

La reconnaissance d'image fonctionne grâce à des algorithmes qui permettent de détecter des motifs dans les images et de les associer à des objets. Le traitement se fait au niveau des pixels. Les unités fondamentales de l'image numérique.

3.11.3.1 Les pixels d'une image :

Une image numérique est composée de petits carrés appelés pixels, organisés en grille. Dans une image en noir et blanc, chaque pixel est représenté par une valeur numérique allant de 0 (noir) à 255 (blanc) pour indiquer sa luminosité. Pour les images en couleur, chaque pixel contient trois valeurs numériques correspondant aux couleurs rouge (R), vert (G), bleu (B) du système RGB, chaque couleur pouvant varier entre 0 et 255.

3.11.3.2. Segmentation d'image :

La segmentation d'image est une extension supplémentaire de la détection d'objet dans laquelle nous marquons la présence d'un objet à travers des masque pixel par pixel génère pour chaque objet dans l'image. Cette technique est plus granulaire que la génération de boîte englobants car cela peut nous aider à déterminer la forme de chaque objet présent dans l'image. Cette granularité nous aide dans divers domaines tel que le traitement d'image médicales, l'image satellite...

Il existe principalement deux types de la segmentation :

- **Segmentation sémantique** : c'est une technique qui détecte, pour chaque pixel la catégorie d'objet à laquelle il appartient. Tout catégorie objets (étiquette) doivent être connu par un modèle.
- **Segmentation d'instance** : elle est identique à la segmentation sémantique mais en approfondissant un peu, elle identifie, pour chaque pixel, l'instance d'objet à laquelle elle appartient. La principale différence par rapport à la segmentation sémantique c'est qu'elle différencie deux objets avec les mêmes étiquettes [37].

Dans la figure suivante on présente un exemple de détection et de segmentation d'image.

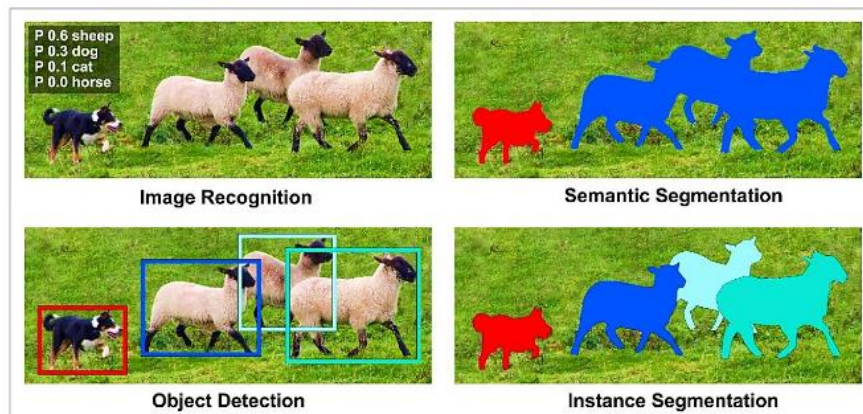


Figure 3.12 : exemple de détection d'image et segmentation [38]

Finalement nous pouvons voir que la reconnaissance d'objet fait référence à une suite de tâches de vision par ordinateur, tel qu'il est illustré par la figure ci-dessous :

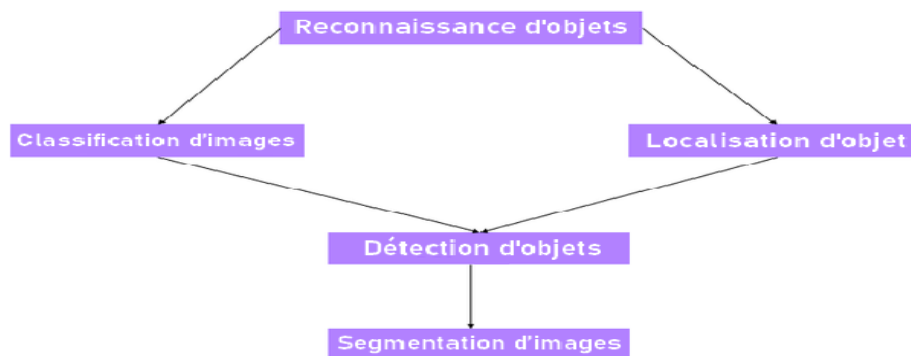


Figure 3.13 : Présentation des tâches de vision par ordinateur pour la reconnaissance d'objet [37].

3.12 L'intérêt de l'utilisation des réseaux de neurones artificiels pour un drone :

L'utilisation de réseau de neurones artificiels pour un drone présente plusieurs avantages significatifs qui transforment ses capacités opérationnelles et ouvrent la porte à de nouvelles applications :

1. **Navigation et vol autonome** : grâce aux réseaux de neurones, un drone apprend à naviguer dans des environnements complexe de manière autonome, en traitant les

données sensorielles en temps réel pour éviter les obstacles, suivre des itinéraires et atterrir de manière sécurisée sans intervention humaine.

2. **Reconnaissance d'objet et de scène :** Les drones équipés de capacités de vision par ordinateur, alimentées par des réseaux de neurones, peuvent identifier et classifier des objets ou des caractéristiques spécifiques dans leur champ de vision.
3. **Traitement des images et des vidéos en temps réel :** Les réseaux de neurones peuvent traiter et analyser les flux d'images et des vidéos capturés par les drones pour fournir des informations et des analyses en temps réel, comme la détection de changements d'environnementaux ou le suivi de la faune.
4. **Amélioration de la sécurité et la fiabilité :** En apprenant des expériences passées et en s'adaptant à de nouvelles conditions, les réseaux de neurones contribuent à améliorer la sécurité et la fiabilité des drones. Cela inclut la capacité de prendre des décisions rapides face à des situations imprévues ou dangereuses.
5. **Optimisation des parcours :** Les réseaux de neurones permettent l'optimisation des trajectoires de vol pour économiser de l'énergie, augmenter l'efficacité des parcours et réduire le temps nécessaire pour accomplir des missions [38].

3.13 Conclusion :

Ce chapitre explore la détection de drones par le biais du Deep Learning, en se concentrant sur l'utilisation de deux méthodes principales : les réseaux de neurones convolutifs (CNN) en 1D et en 2D. En combinant l'analyse d'images et d'audios, ces techniques permettent une détection robuste et précise des drones dans divers environnements. Les CNN 1D sont employés pour l'analyse des signaux audios émis par les drones, capturant leurs caractéristiques temporelles distinctives. Parallèlement, les CNN 2D sont utilisés pour traiter les images des drones capturées par des caméras de surveillance, facilitant leur localisation et leur classification.

Chapitre 4 : Résultats et Discussions

4.1. Introduction :

Dans ce chapitre, nous présentons en détail l'implémentation pratique de nos approches de détection de drones en utilisant des données audios et images, notamment la préparation des données, la construction et l'évaluation des modèles de classification.

4.2. Reconnaissance de la présence de drone à partir des signaux audios :

4.2.1. Base de données :

Les données audios utilisées dans notre travail comprennent des enregistrements de sons de drones ("yes drone") et sans drones ("no drone") [39]. Les chemins vers ces fichiers audios ont été spécifiés, et les caractéristiques pertinentes ont été extraites à l'aide de la bibliothèque Librosa. Les caractéristiques extraites comprennent les coefficients MFCC (Mel-frequency cepstral coefficients), le taux de passage par zéro (ZCR), et les coefficients LPC (Linear predictive coding).

Table 4.1 : Description de la base audio utilisée.

Classe	Nombre d'échantillons	Types d'audios
Drone	400	Original et augmenté
No drone	400	Bruits, urbains, sons humaines, animaux, la pluie, la cloche, le marteau

Au total, la base de données contient 800 fichiers audios, dont 400 échantillons acoustique pour chacune des deux classes. Le tableau ci-dessus détaille la description des échantillons acoustiques sélectionnées dans les classes positives et négatives.

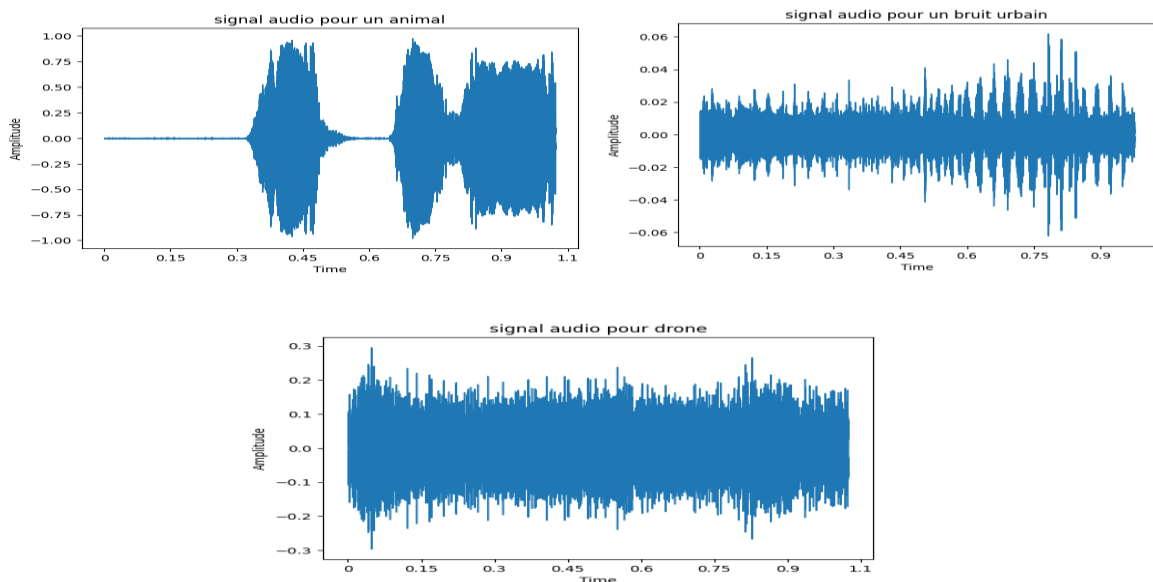


Figure 4.1 : Représentation visuelle de quelques signaux audio de drone et non drone de la base utilisée.

4.2.2 Environnement de travail :

- **Google Colaboratory :**

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.

4.2.3 Langage de programmation :

- **Python :**

Python est un langage de programmation largement utilisé dans les applications Web, le développement de logiciels, la science des données et le machine learning (ML). Les développeurs utilisent Python parce que c'est un langage efficace et facile à apprendre, et qu'il peut s'exécuter sur de nombreuses plateformes différentes.

4.2.4 Bibliothèques utilisées :

- **Numpy :**

NumPy est une bibliothèque pour le langage de programmation Python, destinée à manipuler des matrices ou tableaux multidimensionnels ainsi que des fonctions mathématiques opérant sur ces tableaux.

- **Librosa :**

Librosa est une bibliothèque open-source de traitement du signal audio développée en Python. Elle offre une gamme d'outils et de fonctions permettant d'analyser, de transformer et de visualiser les données audios. Librosa facilite l'extraction de caractéristiques importantes des signaux audio, telles que les coefficients MFCC (Mel-frequency cepstral coefficients), les caractéristiques temporelles et fréquentielles, les enveloppes spectrales, etc. Cette bibliothèque est largement utilisée dans le domaine de l'apprentissage automatique appliqué à l'audio, notamment pour des tâches telles que la classification d'objets sonores, la détection d'événements sonores, la séparation de sources audio, etc. Librosa offre une interface conviviale et des fonctionnalités puissantes, ce qui en fait un outil précieux pour les chercheurs et les développeurs travaillant avec des données audios.

- **Scikit-learn :**

Scikit-Learn est une bibliothèque open source d'apprentissage automatique en Python. Aussi connue sous son nom abrégé Sklearn, elle fournit une vaste gamme d'outils, d'algorithmes et de fonctionnalités permettant de faciliter le développement de modèles de machine learning. Scikit-learn couvre les principaux algorithmes de machine learning généralistes : classification, régression, clustering, etc.

- **Keras :**

Keras est une bibliothèque open source écrite en Python pour l'apprentissage profond. Il s'agit en réalité d'une surcouche à TensorFlow, qui implémente les briques de base (opérations matricielles, couches, fonctions de transfert). TensorFlow gère notamment la possibilité d'exécuter un réseau de neurones sur le processeur (CPU) ou sur l'accélérateur graphique matériel (GPU).

4.2.5. Classification par méthodes classiques (KNN, SVM) :

Nous avons premièrement utilisé des classifieurs classiques, à savoir le K plus proche voisin (K-Nearest Neighbors (KNN)) et les machines à vecteurs de support (Support Vectors Machines (SVM)). Ces modèles ont été entraînés et évalués avec des paramètres spécifiques, et leurs performances ont été comparées à celle du modèle CNN 1D.

4.2.5.1 Extraction des paramètres à partir des signaux audios :

Dans cette section on présente des visualisations pour les caractéristiques extraites à partir des signaux audios afin d'illustrer leurs distributions et leurs pertinences pour la classification des données. Nous avons utilisé le MFCC, ZCR, et LPC.

a. Visualisation de MFCC :

Le spectrogramme que nous avons généré à partir des MFCCs (Mel Frequency Cepstral Coefficients) présenté dans la figure (4.2) offre une représentation visuelle riche et informative des caractéristiques spectrales d'un signal audio enregistré pour un drone (figure 4.2) et un son d'un animal (figure.4.3). Cette visualisation est particulièrement utile pour analyser les propriétés temporelles et fréquentielles du son capturé. Les couleurs dans ce spectrogramme servent à identifier les régions temporelles où les coefficients atteignent des valeurs élevées ou basses, ce qui indique la présence ou l'absence de certaines fréquences.

L'axe des X représente le temps ce qui nous permet de voir comment les caractéristiques du son changent au cours du temps.

L'axe des Y représente les indices des coefficients MFCC, chaque ligne sur l'axe Y correspond à un coefficient MFCC, allant de 1 à 30 coefficients de MFCC.

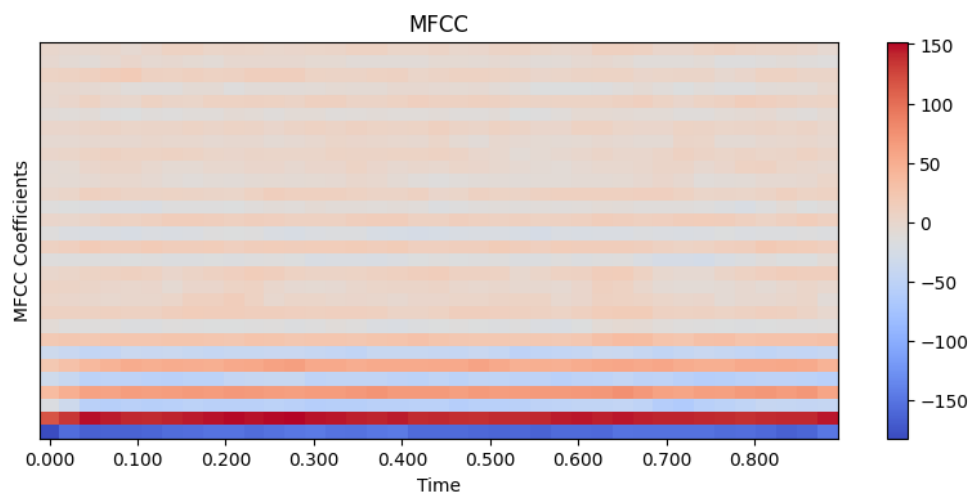


Figure 4.2 : Représentation visuelle des coefficients MFCC pour drone

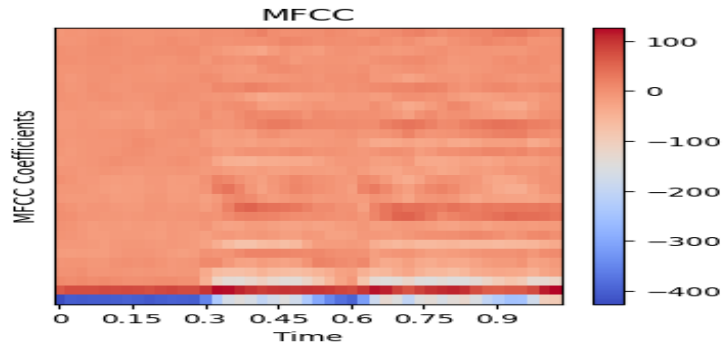


Figure 4.3 : Représentation visuelle des coefficients MFCC pour un son d'animal

L'intensité de la couleur représente l'amplitude ou l'énergie de chaque composante de fréquence à différents intervalles de temps. Ce qui indique non seulement une forte activité de certaines fréquences, en cas de forte énergie, mais aussi signale une capture de son de bonne qualité à cause de la distance courte entre la position de drone et la source d'acquisition du son par exemple.

Dans le cas d'une faible énergie, cela pourrait indiquer un manque de détails ou de nuances dans le son enregistré et pourrait signifier que les fréquences correspondantes ont une amplitude plus basse, ce qui les rend moins perceptibles pour l'oreille à cause du vent, d'autres sources de bruit ou d'autres conditions acoustiques défavorables.

b. Visualisation de LPC :

L'image montre un graphique des coefficients de prédiction linéaire (LPC) d'un signal audio. Les LPC sont une représentation du signal audio dans le domaine temporel qui capture les informations sur la forme d'onde du signal. Voici quelques détails supplémentaires sur les LPC et leur interprétation :

Axes de l'image :

- Axe des x : Représente les indices des coefficients LPC, numérotés de 1 à 9.
- Axe des y : Représente les valeurs des coefficients LPC, numérotés de 1 à 4

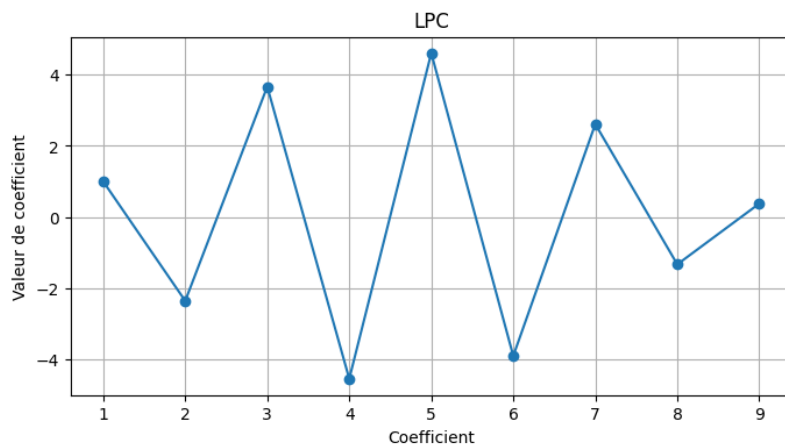


Figure 4.4 : Représentation visuelle des coefficients LPC pour audio de drone

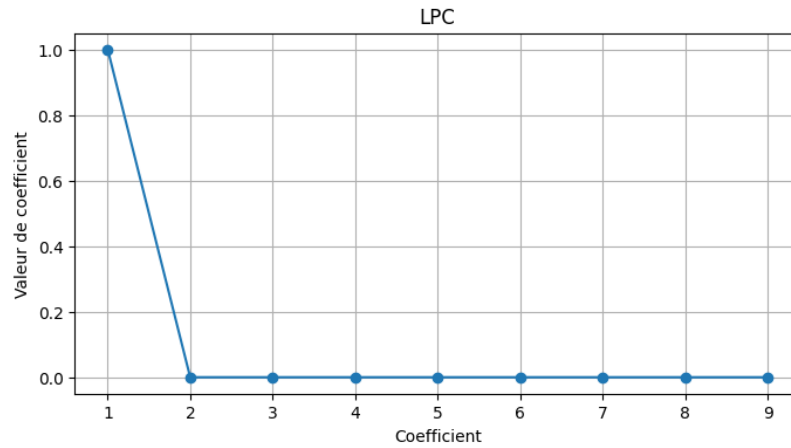


Figure 4.5 : Représentation visuelle des coefficients LPC pour son de marteau

Les pics indiquent une bonne qualité de signal sonore comme indiqué en figure 4.4. Dans la figure 4.5 on remarque que les valeurs des coefficients s'annulent à partir du deuxième coefficient, ce qui correspond à l'enregistrement d'audio.

c. Visualisation de ZCR :

La figure suivante représente une représentation du taux de passage par zéro (ZCR) d'un signal audio. Le ZCR est une mesure de la rapidité avec laquelle le signal audio change de signe. Il est couramment utilisé dans les applications de traitement du signal audio,

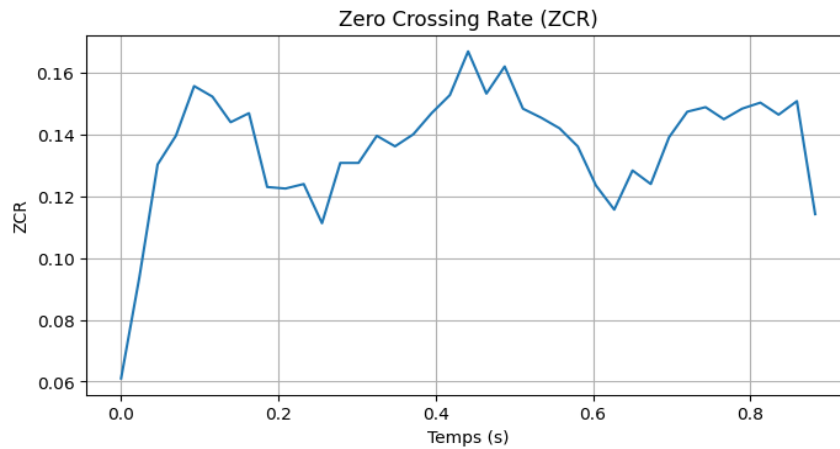


Figure 4.6 : Représentation visuelle des coefficients ZCR pour drone

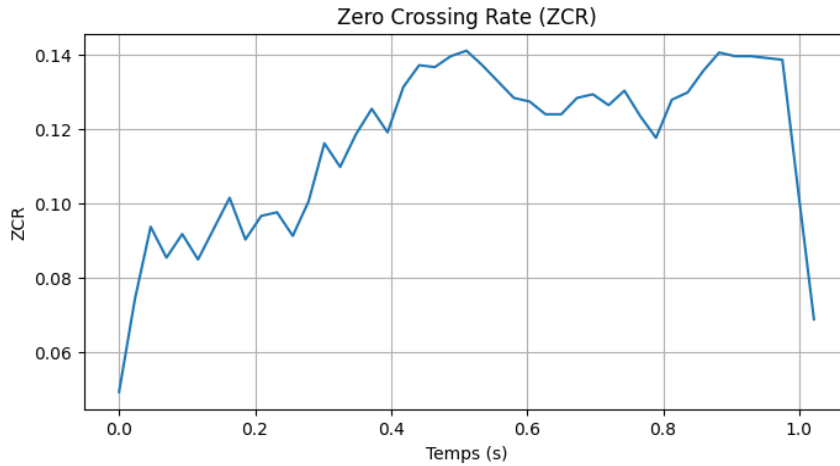


Figure 4.7 : Représentation visuelle des coefficients ZCR pour un son d'animal

Les deux figures (drones et animal) représentent une variation irrégulière du ZRC en fonction du temps, ce qui montre la rapidité avec laquelle le signal change de signe.

4.2.5.2 Classification :

Nous présentons dans ce qui suit les résultats de la classification des signaux sonores par SVM et KNN.

L'ensemble de données contient des signaux provenant de deux classes différentes « yes drone » et « no drone ». Nous avons utilisé les algorithmes SVM et KNN pour classer les signaux dans leurs classes respectives.

80% de la totalité de base est réservé à la phase d'apprentissage et 20% pour le test. Les modèle KNN et SVM sont évalué en utilisant 5 plis de validation croisée.

4.2.5.2.1 Les Métriques utilisées :

Les métriques utilisées dans ce travail pour mesurer les performances des méthodes utiliser pour la détection des drones sont présentées ci-dessous :

$$\text{Précision (accuracy)} = \frac{TP+TN}{TP+TN+FP+FN} \quad (4.1)$$

$$\text{Rappel (Recall)} = \frac{TP}{TP+FN} \quad (4.2)$$

$$\text{Spécificité (precision)} = \frac{TN}{TN+FP} \quad (4.3)$$

$$\text{F1-Score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4.4)$$

D'où les quantités TP, TN, FP, FN représentent respectivement : vrai positif (true positive), (Vrai négatif (True Negative), faux positif (False Positive), Faux négatif (False Negative) telqu'il est présenté dans la matrice de confusion suivante.

		POSITIVE	NEGATIVE
POSITIVE		TP	FN
NEGATIVE		FP	TN

Figure 4.8 : Matrice de confusion

4.2.5.3 Résultats de classification

Le tableau suivant représente les résultats de classification en utilisant le SVM et le KNN et différentes combinaisons de paramètres en fonction de la précision.

Table 4.2 : Comparaison des Performances des Méthodes KNN et SVM avec Différentes Caractéristiques Audio

Paramètre	KNN	SVM
MFCC	0.900	0.96875
ZCR	0.694	0.5
LPC	0.875	0.84375
LPC+ZCR	0.881	0.84375

Nous remarquons que les meilleurs résultats sont obtenus en utilisant le MFCC et le SVM. Par contre le ZCR a donné de très faibles résultats.

Les figures suivantes représentent les matrices de confusion pour les meilleurs résultats obtenus en utilisant le KNN et le SVM.

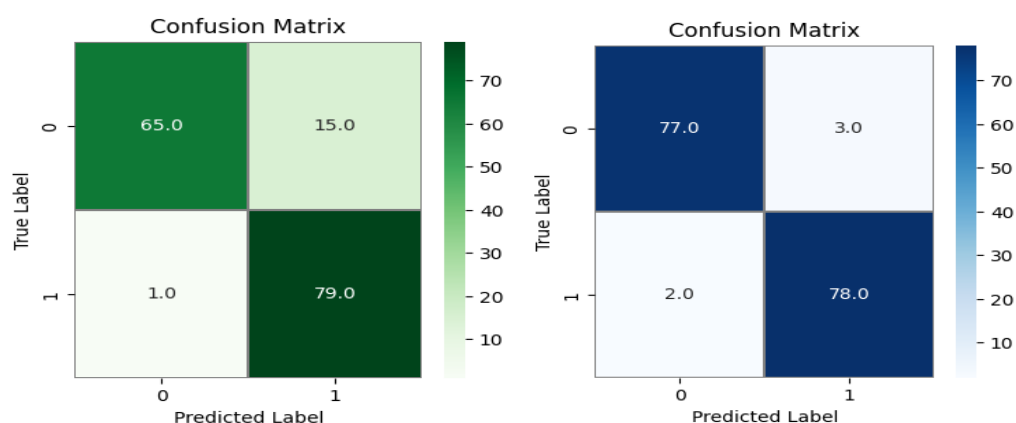


Figure 4.9 : Représentation des matrices de confusion pour SVM et KNN

Les matrices de confusion ci-dessus représentent les performances des deux modèles de classification utilisés pour détecter la présence de drones : le modèle KNN (à gauche en vert) et le modèle SVM (à droite en bleu). Ces matrices permettent d'évaluer comment chaque modèle prédit les classes "drone" (1) et "non-drone" (0) en comparant les prédictions avec les valeurs réelles.

Table 4.3 : Comparaison des algorithmes de détection de drones en utilisant le MFCC pour les deux classes

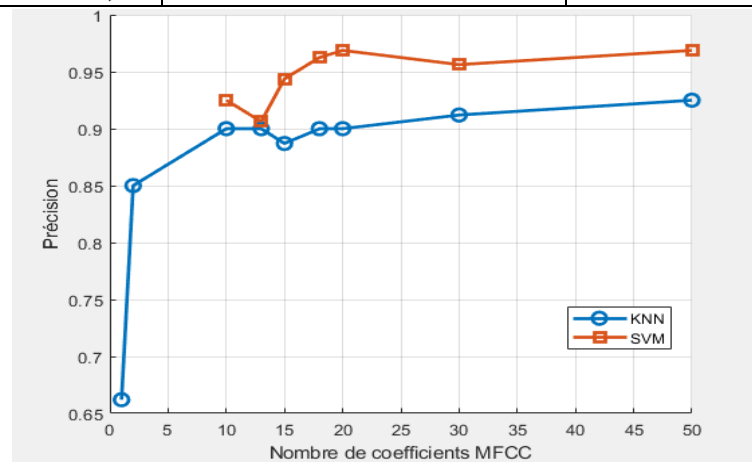
Algorithmes	Precision	F1-score	Recall
KNN (yes drone)	0.89	0.96	0.92
KNN (no drone)	0.96	0.88	0.92
SVM (yes drone)	0.95	0.96	0.96
SVM (no drone)	0.96	0.95	0.96
Mean KNN	0.925	0.885	0.92
Mean SVM	0.955	0.955	0.96

Comme le montre le tableau, l'algorithme SVM a obtenu des performances légèrement meilleures que l'algorithme KNN pour les deux ensembles de données. Cela signifie que l'algorithme SVM était plus efficace pour distinguer correctement les signaux des différentes classes.

Dans le tableau suivant en compare entre la classification par KNN et SVM pour différent nombre de coefficients de MFCC en fonction du taux de classification.

Table 4.4 : Comparaison des algorithmes KNN et SVM pour la classification avec MFCC

Nombre de coefficients	KNN	SVM
MFCC (1 coefficients)	0.662	/
MFCC (2 coefficients)	0.850	/
MFCC (10 coefficients)	0.900	0.925
MFCC (13 coefficients)	0.900	0.90625
MFCC (15 coefficients)	0.887	0.94375
MFCC (18 coefficients)	0.900	0.9625
MFCC (20 coefficients)	0.900	0.96875
MFCC (30 coefficients)	0.912	0.95625
MFCC (50 coefficients)	0.925	0.96875

**Figure 4.10** : la précision de KNN et SVM en fonction du nombre de coefficients.

La figure (4.9) montre que le taux de classification des algorithmes KNN et SVM augmente en fonction du nombre de coefficients MFCC, ceci pourra être interpréter par le nombre de détails élevé qu'on peut capturer à partir un d'un signal audio en utilisant un nombre de coefficients élevé. Le SVM atteint généralement une précision plus élevée que KNN. La précision de KNN plafonne autour de 0,9, tandis que SVM atteint jusqu'à 0.96875.

Ces résultats suggèrent que le SVM est mieux adapté pour cette tâche de classification de drones, bien que le KNN ait également montré de bonnes performances. Les choix entre ces deux modèles peuvent dépendre des exigences spécifiques de l'application, notamment en ce qui concerne l'importance relative de minimiser les faux positifs ou les faux négatifs.

4.2.6. Classification des signaux audio par Deep Learning (CNN 1D) :

Nous avons implémenté un modèle de réseau de neurones Convolutionnel 1D (CNN 1D) pour la classification des données audio. Le modèle CNN 1D proposé (figure 4.10) a été construit avec plusieurs couches de convolution suivie de plusieurs couches de Pooling et des couches entièrement connectées. Le modèle a été compilé avec l'optimiseur Adam et la fonction de perte binaire_crossentropy. Après l'entraînement du modèle sur l'ensemble de données d'apprentissage, son exactitude a été évaluée sur l'ensemble de données de test.

Notre model CNN 1D contient 4 couches cachés avec des différents filtres « 32, 64, 128, 256 ».

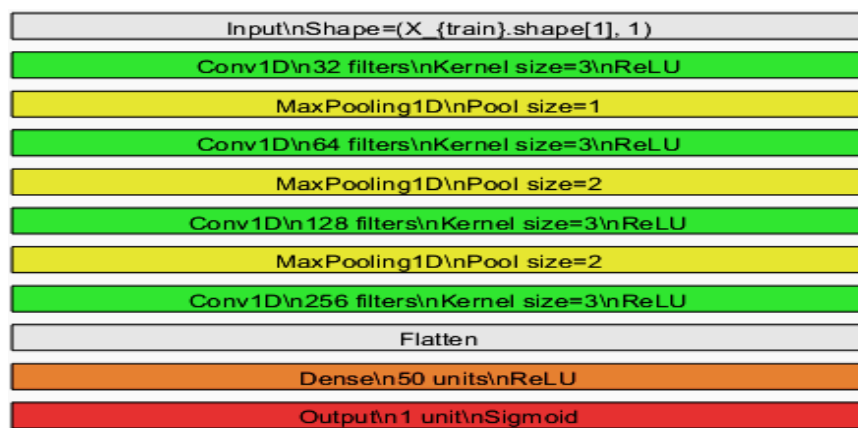


Figure 4.11 : Architecture de modèle CNN 1D proposée

```
Epoch 1/10
20/20 [=====] - 1s 8ms/step - loss: 0.5025 - accuracy: 0.8219
Epoch 2/10
20/20 [=====] - 0s 14ms/step - loss: 0.2245 - accuracy: 0.9172
Epoch 3/10
20/20 [=====] - 0s 16ms/step - loss: 0.1813 - accuracy: 0.9250
Epoch 4/10
20/20 [=====] - 0s 14ms/step - loss: 0.1382 - accuracy: 0.9516
Epoch 5/10
20/20 [=====] - 0s 14ms/step - loss: 0.1029 - accuracy: 0.9594
Epoch 6/10
20/20 [=====] - 0s 14ms/step - loss: 0.1031 - accuracy: 0.9609
Epoch 7/10
20/20 [=====] - 0s 13ms/step - loss: 0.0684 - accuracy: 0.9844
Epoch 8/10
20/20 [=====] - 0s 12ms/step - loss: 0.0873 - accuracy: 0.9719
Epoch 9/10
20/20 [=====] - 0s 12ms/step - loss: 0.0541 - accuracy: 0.9812
Epoch 10/10
20/20 [=====] - 0s 13ms/step - loss: 0.0728 - accuracy: 0.9750
Accuracy of CNN 1D: 0.9937499761581421
```

Figure 4.12 : les résultats de CNN 1D en fonction du Loss et taux de classification pour différents epochs.

Le modèle CNN1D proposé montre une excellente performance avec un taux de classification globale élevée (99.37%). La matrice de confusion correspondante est présentée dans la figure suivante.

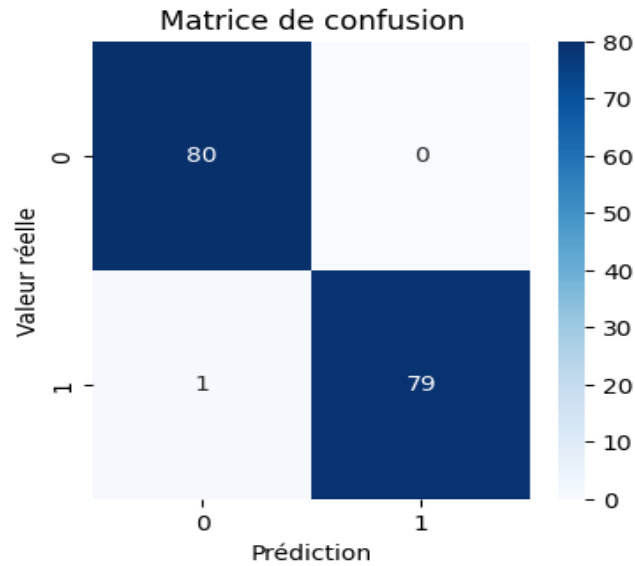


Figure 4.13 : la matrice de confusion de model CNN 1D

Cette matrice présente la prédiction du modèle CNN 1D des classes "drone" (1) et "non-drone" (0) et compare les prédictions avec le cas réelles.

Table 4.5 : Résultats de classification de drones par CNN1D

	Accuracy	Precision	Recall	F1-Score
CNN1D	99.37%	0.98	1	0.98

La figure suivante montre une comparaison entre les résultats des trois méthodes utiliser pour la classification des signaux audio drone et non drones. On remarque que le CNN1D a donné de meilleurs résultats.

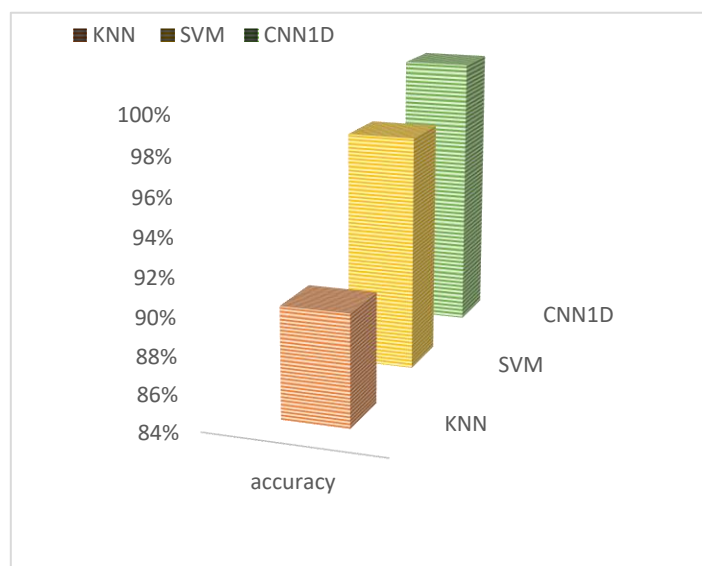


Figure 4.14 : Comparaison entre les méthodes utilisées pour la détection des drones en utilisant les signaux audios

4.3. Reconnaissance de la présence de drone à partir d'images :

Dans cette deuxième partie de notre projet de fin d'études, nous avons mis en œuvre un système de classification visant à distinguer entre les oiseaux et les drones à partir d'images. Pour ce faire, nous avons utilisé TensorFlow et Keras pour développer un modèle de réseau de neurones convolutifs (CNN) capable d'effectuer cette tâche de manière efficace.

4.3.1 Prétraitement des données :

Nous avons commencé par charger les données d'images à partir des répertoires spécifiés sur Google Drive. Les images ont été redimensionnées à une taille fixe de 200x200 pixels pour assurer la cohérence de la taille des entrées pour notre modèle. La base contient un total de 828 images [40]. Dans l'ensemble d'apprentissage (75%) nous avons 621 images, tandis que dans l'ensemble de test (25%) nous avons 207.

Plus précisément, pour l'ensemble d'apprentissage, la classe 0 (oiseaux) compte 295 images, et la classe 1 (drones) en compte 326. Pour l'ensemble de test, la classe 0 (oiseaux) compte 105 images, tandis que la classe 1 (drones) en compte 102.

Les figures suivantes représentent des échantillons d'images de la base utilisée pour les deux classes drone et oiseau dans différents conditions et situations.



Figure 4.15 : Images de drone sous différents angles

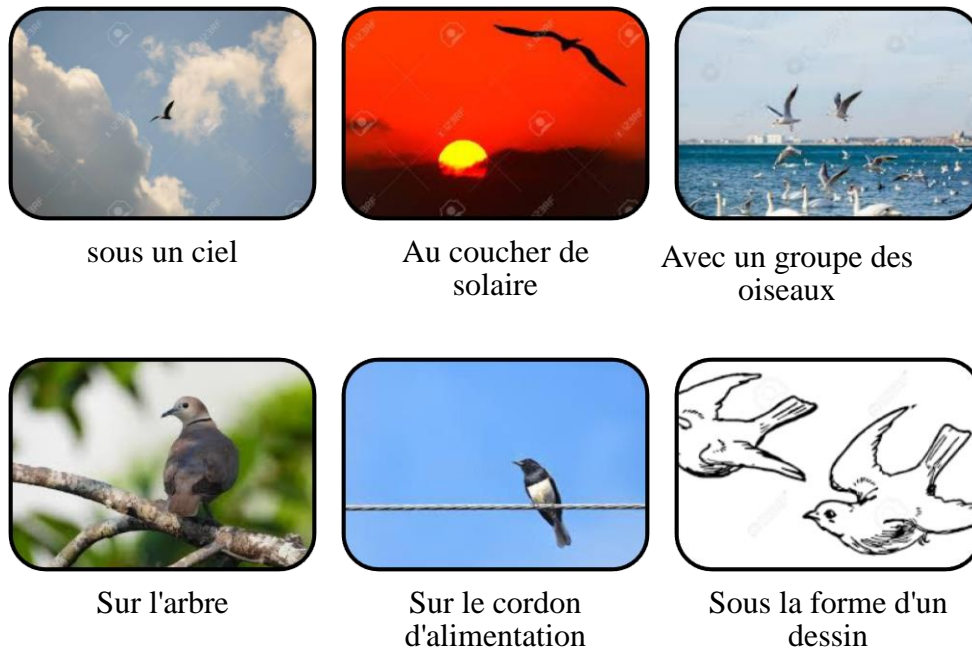


Figure 4.16 : Images d’oiseaux sous différents angles

Table 4.6 : La nature des classes de « drones » et « bird »

Classe	Echantillons
Drone	428
bird	400

Au total, la base de données contient 828 images, dont 400 échantillons d’images pour oiseaux et 428 pour drone. Le tableau ci-dessus détaille la description des échantillons sélectionnées dans les classes positives et négatives.

4.3.2 Modèle CNN utilisé :

Le modèle CNN que nous avons construit comprend plusieurs couches de convolution, suivies de couches de max Pooling pour réduire la dimensionnalité. Nous avons également ajouté des couches entièrement connectées pour la classification finale. Nous avons divisé nos données en ensembles d’entraînement et de test, puis nous avons utilisé la validation croisée k-fold pour évaluer les performances de notre modèle de manière robuste. Cela nous a permis d'estimer la performance du modèle sur de nouvelles données et de vérifier sa généralisation.

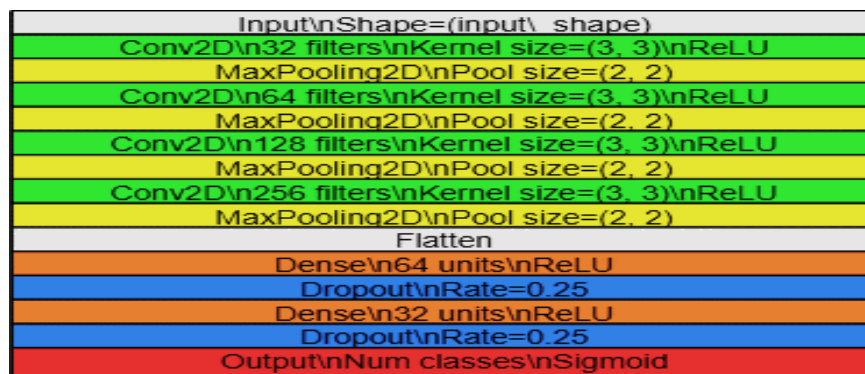


Figure 4.17 : l’architecture de CNN 2D proposé.

4.3.3 Évaluation du modèle :

Nous avons évalué les performances du modèle sur l'ensemble de test en calculant l'exactitude, la précision, le rappel et le score F1. Une matrice de confusion a également été générée pour visualiser les performances de classification par classe.

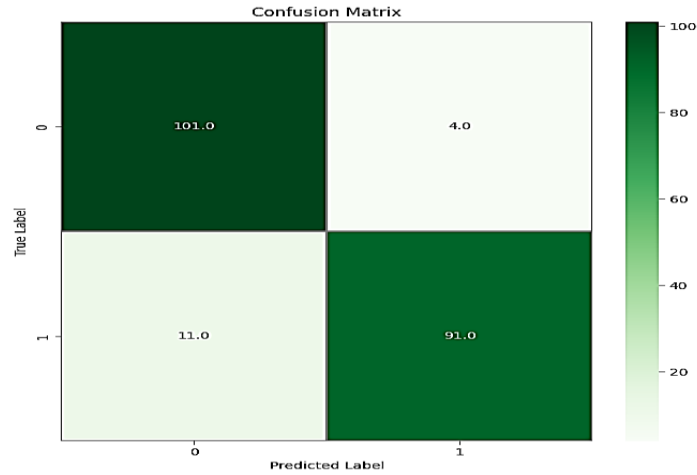


Figure 4.18 : La matrice de confusion de model CNN 2D

4.3.4 Interprétation de la Matrice de Confusion :

En ce qui suit on décrit chaque élément de la matrice de confusion :

- Vrais négatifs (TN) : 101
- Faux positifs (FP) : 4
- Faux négatifs (FN) : 11
- Vrais positifs (TP) : 91

Ces valeurs de la matrice de confusion sont utilisées pour calculer des métriques de performance présentés dans le tableau suivant :

Table 4.7 : Résultats de classification de drone par CNN2D

	Accuracy	Precision	Recall	F1-Score
CNN 2D	92.75%	0.89	0.95	0.91

Les résultats obtenus montrent que le modèle CNN 2D présente une performance globale robuste dans la classification des images de drones.

4.3.5 Visualisation des Résultats :

4.3.5.1 Analyse du Graphique de Précision :

Le graphique suivant illustre l'évolution de la précision du modèle CNN 2D au cours des 100 époques d'apprentissage. Les courbes montrent le taux de classification pour l'ensemble d'entraînement (en bleu) et l'ensemble de validation (en orange).

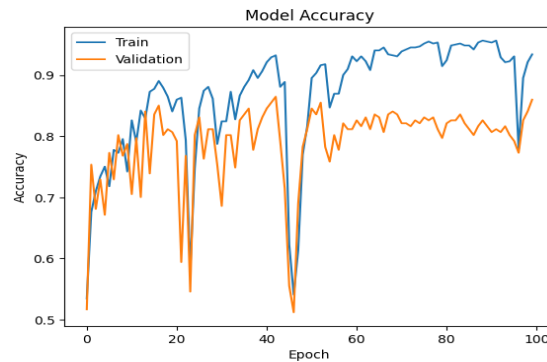


Figure 4.19 : la courbe de précision (accuracy) de l'entraînement et de la validation du modèle au cours de 100 epochs.

a. Courbe de Précision de l'Ensemble d'Entraînement (Train) :

- La courbe bleue montre une tendance générale à la hausse, atteignant des valeurs proches de 0.95 (95%) vers la fin de l'entraînement.
- Cette courbe fluctue initialement, indiquant que le modèle apprend les caractéristiques des données. Les fluctuations diminuent avec le temps, signe que le modèle s'adapte progressivement.
- Les pics et les baisses abruptes peuvent être dus à des ajustements dans les poids du modèle pendant l'apprentissage.

b. Courbe de Précision de l'Ensemble de Validation (Validation) :

- La courbe orange représente la précision du modèle sur des données non vues pendant l'entraînement. Elle atteint environ 0.80 (80%) et montre une stabilisation autour de cette valeur.
- Les fluctuations dans la courbe de validation sont plus importantes que celles de l'ensemble d'entraînement, ce qui peut indiquer des variations dans la capacité du modèle à généraliser sur les nouvelles données.

3.5.2 Analyse du Graphique de Perte :

Le graphique ci-dessous illustre l'évolution de la perte du modèle CNN 2D au cours des 100 époques d'apprentissage. Les courbes montrent la perte pour l'ensemble d'entraînement (en bleu) et l'ensemble de validation (en orange).

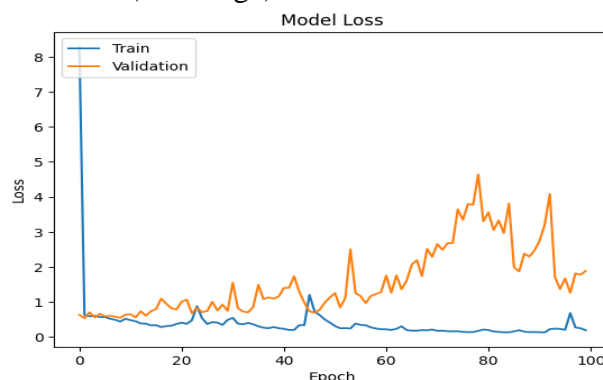


Figure 4.20 : la courbe de Loss de l'entraînement et de la validation du modèle au cours de 100 epochs.

a. Courbe de Perte de l'Ensemble d'Entraînement (Train) :

- La courbe bleue montre une diminution rapide de la perte dès les premières époques, atteignant des valeurs proches de 0.5. Cette diminution rapide initiale indique que le modèle apprend rapidement à minimiser l'erreur sur les données d'entraînement.
- Après cette diminution initiale, la courbe de perte se stabilise avec de légères fluctuations autour de valeurs basses, ce qui suggère que le modèle continue à affiner ses poids tout en gardant une perte basse.

b. Courbe de Perte de l'Ensemble de Validation (Validation) :

- La courbe orange représente la perte du modèle sur des données non vues pendant l'entraînement. Contrairement à la courbe d'entraînement, la courbe de validation montre des fluctuations importantes et une tendance générale à la hausse.
- Les fluctuations et la tendance à l'augmentation de la perte de validation indiquent que le modèle commence à sur ajuster les données d'entraînement et perd sa capacité de généralisation, surtout après les premières époques.

4.4. Conclusion :

Dans ce chapitre nous avons présenté les résultats obtenus pour la détection de la présence de drone en utilisant des algorithmes de classification classiques, basés sur l'extraction des paramètres, et en utilisant le Deep Learning.

Dans un premier temps nous avons utilisé des données sonores contenant des enregistrements de son de drones et des sons divers dont nous avons fait l'extraction de plusieurs paramètres (MFCC, ZCR, LPC). Ces paramètres ont été appliqués aux entrées des classificateurs KNN et SVM, les résultats ont indiqué que le SVM surpasse le KNN. Dans un second temps, nous avons utilisé un CNN1D pour la classification des signaux audio. Les résultats obtenus montrent une très bonne performance du CNN1D pour ce problème de classification.

Dans la deuxième partie du travail nous avons utilisé une base d'images pour détecter des drones. Le CNN a été utilisé pour la classification « drone vs oiseaux » les résultats obtenus sont prometteurs.

Conclusion générale

Conclusion générale

Dans ce travail, nous avons abordé le problème de la détection de drones en utilisant des techniques d'intelligence artificielle. Nous avons implémenté des méthodes d'apprentissage automatique classiques et d'apprentissage profond (Deep Learning), en utilisant une base de données sonores et une base d'images, dans le but d'identifier la présence de drones.

Nous avons commencé par décrire de manière globale les drones et les défis liés à leurs repérages. Ensuite, nous avons mis en œuvre des méthodes d'apprentissage automatique classiques, comme le k plus proches voisins (KNN) et les machines à vecteurs de support (SVM), ces classifieurs ont été alimentés par différents paramètres extraits à partir des signaux audio tel que les coefficients MFCC (Mel-Frequency Cepstral Coefficients), LPC (Linear Predictive Coding) et ZCR (Zero-Crossing Rate). Nous avons obtenu une précision de 0,92 avec KNN et de 0,96 avec SVM.

En ce qui concerne la détection par apprentissage profond ; nous avons utilisé les réseaux de neurones convolutifs. Un modèle proposé de CNN 1D a été appliqué aux données audio et a généré une précision impressionnante de 0.9937. D'un autre côté, concernant les données d'image, un modèle proposé de CNN 2D a produit une précision de 0.92. Cette variabilité est attendue compte tenu de la similitude de ces méthodes et du fait que le traitement d'image est plus complexe.

Ces résultats montrent clairement que, d'une part, le MFCC est une technique robuste par rapport aux LPC et ZCR. D'un autre part, les techniques d'apprentissage automatique et d'apprentissage profond sont efficaces pour la détection des drones. Particulièrement, les méthodes basées sur l'apprentissage profond offrent des solutions robustes et fiables, avec une performance particulièrement élevée pour les données audio. Les résultats sont très prometteurs et indiquent que ces techniques sont très efficaces dans ce domaine.

Bien que ce travail ait démontré l'efficacité des techniques d'intelligence artificielle pour la détection de drones, il existe encore de nombreuses opportunités pour améliorer et étendre ces approches. Les futures recherches dans ces directions pourraient conduire à des systèmes de détection de drones plus avancés, fiables et polyvalents, avec des applications potentiellement révolutionnaires dans divers domaines.

Perspectives :

Dans nos futurs travaux on propose :

- **D'améliorer les modèles de classification d'image** : par utilisation de réseaux de neurones convolutifs plus complexes, et la mise en œuvre de techniques avancées de traitement d'images et d'augmentation des données disponibles.
- **La fusion multimodale** : combinaison des données provenant de sources audio, vidéo et autres capteurs (infrarouges, lidar) pour renforcer la fiabilité et la précision.
- **Optimisation en temps réel** : diminution des délais de traitement et optimisation de l'efficacité dans le calcul.

- **L'adaptabilité aux conditions variables** : tests approfondis et ajustements des modèles pour s'adapter à divers environnements (interférences sonores, conditions météorologiques changeantes, fluctuations lumineuses).
- **Détection et classification des divers types de drones** : cette tâche est essentielle pour les cas où l'utilisation nécessite une distinction entre les différents modèles de drones.
- **Sécurité et contre-mesures** : pour renforcer la protection des systèmes de détection contre les attaques malveillantes, et le développement de stratégies pour neutraliser les drones repérés.

Bibliographie et Webographie

- [1] <https://www.studiofly.fr/les-drones-un-siecle-de-recherches-et-d-elaboration/>.12/03.2024
- [2] <https://www.futura-sciences.com/sciences/definitions/aeronautique-drone-6174>
- [3] M. Salaheddine, B. Yahya, B. Serra, 'Conception et réalisation d'un mini drone'', université Mohamed EL-Ibrahimi – Barj Bou Arreridj, mémoire de master, 2020-2021.
- [4]<https://www.studiosport.fr/guides/drones/les-caracteristiques-et-les-fonctions-des-drones.html>
- [5] http://www.afcadillac.net/_serveurs/drone/composition_dun_drone.html 21/04/2024
- [6] <https://www.studiosport.fr/guides/drones/quels-sont-les-differents-types-de-drones.html>
- [7] <https://filmora.wondershare.fr/drones/what-is-drone-how-does-it-work.html>
- [8] B. Mohammed Reda, B. Youcef, "Etude et implémentation d'une loi de commande pour le pilotage d'un drone quadri-rotor", mémoire de master, université Abou Bakr Belkaid-Tlemcen,2016/2017
- [9] : <https://polaridad.es/fr/caracteristicas-de-los-drones/> 4/03/2024
- [10] A. Boussad, C. Kheirddine, 'Reconnaissance par apprentissage automatique de signatures acoustique de drone' 'mémoire de master, université Saad Blida 1,2022.
- [11] <https://www.clubic.com/drone/dossier-377691-demain-des-drones-pour-sauver-l-environnement-.html>
- [12] B. LALOUANI, H. Younes, "reconnaissance par apprentissage automatique de signature acoustique de drone", mémoire de master, université Saad Dahleb Blida1.2023.
- [13] B. TAHA ET ABDULHADI SHOUFAN, Machine Learning-Based Drone Detection and Classification: State-Of-The-Art in Research, Septembre 23/2019 IEEE Access
- [14] <https://www.wikipédia.org>.
- [15] <http://themarvinproject.free.fr/final/node4.html#SECTION04220000000000000000>.
- [16] https://www.researchgate.net/figure/Les-differentes-etapes-dextractions-des-parametres-MFCC_fig1_339944344
- [17] M. Amel, 'Simulation de la Méthode de Mel pour la Reconnaissance Automatique d'un locuteur'', mémoire master, UNIVERSITE BLIDA 1 ,2018.
- [18] <https://topics-beta.apps.semanticscholar.org/topic/495897993?corpusId=12331561>
- [19] B. Nabila, B. TOUFIQ, N. Benayadi, 'CNN ET LSTM POUR LA CLASSIFICATION TISUR DE LA MALADIER PARKINSON BASEE SUR ON GTCC ET MFCC'', Applied Computeur ,2023.
- [20] F. Rahane " Gammatone Cepstral Coefficient for Speaker Identification "International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering,2013.
- [21] <https://iopscience.iop.org/article/10.1088/1742-6596/1090/1/012046/pdf>
- [22] P. Joulieu 'indexation sonore : parole /musique ', université Paul Sabatier,2001
- [23] H. GamboZeinabou, 'segmentation automatique de la parole en phénomène'', mémoire master, Université SAAD DAHLAB de BLIDA,2018.
- [24][https://dataanalyticspost.com/Lexique/svm/#:~:text=SVM%20\(Support%20Vector%20Machine%20ou,ou%20de%20d%C3%A9tection%20d'anomalie](https://dataanalyticspost.com/Lexique/svm/#:~:text=SVM%20(Support%20Vector%20Machine%20ou,ou%20de%20d%C3%A9tection%20d'anomalie).
- [25] A. AICHA, k, LAID ' Intitulé du sujet Intelligence artificielle dans les archives mémoire master', Université Abdelhamid Ibn Badis Mostaganem, 2022.

- [26] <https://www.ibm.com/fr-fr/topics/supervised-learning#:~:text=L'apprentissage%20supervis%C3%A9%20utilise%20un,apprendre%20au%20fil%20du%20temps>.
- [27] A. Lydia, B. Rosa. Reconnaissance Automatique de la Modulation à base des Séparateurs à Vaste Marge (SVM). Mémoire master. Université Abderrahmane Mira.
- [28] L. Redouane ‘’Application des SVM pour la reconnaissance d’extrasystoles’’. Mémoire Master, université Mouloud Mammeri de Tizi Ouzou.
- [29] D. Sarah. C. Nassima ‘’ Comparaison des différentes approches de classification automatique des documents textuels’’, mémoire master, université de Génie Electrique et D’informatique,2016.
- [30] L. Redouane,’’ Application des SVM pour la reconnaissance d’extrasystoles’’, mémoire master, université de Génie Electrique et D’informatique,2015.
- [31] L. Ali,’’ Comme exigence partielle de la maitrise en mathématiques et informatique appliquées’’, université du Québec, juin 2017.
- [32] M. Soumia Sadiia, K. chaima,’’ Emotion Recognition using deep Learning’’, mémoire de master, Université Moulay Taher Saida ,2020.
- [33] D. Nour EL Houda, H. Souad, ‘’Détection des drones à l’aide des réseaux de neurones artificiels’ ’mémoire de master, Université Saad Dahleb Blida, 2019.
- [34]https://medium.com/@andreimoiceanu_9429/recurrent-neural-networks-vs-1d-convolutional-networks-5ac7b4f68ca9.
- [35] <https://datascientest.com/convolutional-neural-network>
- [36] <https://praedictia.com/page/reconnaissance-vocale/que-peut-faire-la-reconnaissance-vocale.html>.
- [37] L. HASSANE, I. KHAMMAR,’’ Détection automatique d'objets à partir d'un drone’ ’université haut-alsace, mémoire master 1 ,2021.
- [38] <https://blent.ai/blog/a/detection-images-yolo-tensorflow>
- [39] <https://www.kaggle.com/datasets/harshwalia/birds-vs-drone-dataset>
- [40] <https://github.com/saraalemadi/DroneAudioDataset>