



République Algérienne Démocratique et Populaire  
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique  
Université de Mohamed El Bachir El Ibrahimi  
Faculté des Mathématiques et Informatique  
Département d'Informatique

**Mémoire du projet de fin d'études**  
En vue l'obtention du diplôme de  
Master en Informatique  
**Spécialité : Ingénierie de l'Informatique Décisionnelle**

---

**THEME :**  
**VALIDATION D'UNE RÉCITATION CORANIQUE**  
**EN UTILISANT UN CNN**

---

**Réalisé par :**

- ❖ Chouitah Marwa
- ❖ Benslimane Aya

**Supervisé par :**

Mme. BADAOUI Atika

**Devant le jury composée de :**

**Président : Zouache Djaafar**

**Examinatrice : Lamri Manal**

**Année universitaire : 2023/2024**

## *Remerciements*

*Au terme de cette étude, nous tenons à présenter nos sincères remerciements au bon dieu de nous avoir accordé la connaissance de la science et de nous avoir aidé à réaliser ce travail.*

*Nous tenons à remercier en premier lieu notre encadrante de mémoire Mme. BADAUD Atika pour sa disponibilité, ses conseils, sa gentillesse et toute l'aide qu'elle nous a rapportée.*

*Un grand remerciement aux membres du jury chacun par son propre nom pour l'honneur et l'intérêt qu'ils nous ont accordé en acceptant d'examiner et d'évaluer notre mémoire.*

*Nous présentons également notre gratitude à tous les professeurs du département d'informatique.*

*Ainsi, nous remercions pour leur soutien tant moral, spirituel et matériel, nos parents, tous nos frères, sœurs, amis et compagnons qui nous ont aidés, conseillé et encouragé.*

*Enfin, à tous ceux que nous n'avons pas pu citer, auxquels on réitère nos sincères remerciements, à vous tous, Merci !*

## *Dédicace*

*C'est avec un grand plaisir que je dédie cet événement marquant de ma vie :*

*À mon très cher père **Derradji***

*Tu as toujours été pour moi un exemple du père respectueux, honnête.*

*Grâce à toi papa j'ai appris le sens du travail et de responsabilité.*

*Ce modeste travail est le fruit de tous les sacrifices que tu as déployés pour mon éducation et ma formation. J'implore le tout-puissant pour qu'il t'accorde une bonne santé et une vie longue et heureuse*

*À ma très chère mère **Salima***

*Quoi que je fasse ou que je dise, je ne saurai point te remercier comme il se doit. Ton affection me couvre, ta bienveillance me guide et ta présence à mes côtés a toujours été ma source de force pour affronter les différents et les obstacles*

*À mon unique frère **Abd elbasset** et ma cher sœur **Takwa** pour leur patiente durant mes moments difficiles ainsi que pour leur amour.*

*À ma grand-mère, mes oncles et mes tantes en particulier ma tante **Soria** pour leur amour, et leur tendresse.*

*À une personne spéciale qui a été toujours là pour moi, mon fiancé **Belkorchia Belkacem** qui n'a pas cessé de m'encourager et soutenir tout au long de mes études et qui*

*est toujours à mes côtés.*

*À ma belle-famille pour son soutien et son amour tout au long de cette aventure.*

*À mes belles-sœurs **imen** et **salsabile** je tiens à vous exprimer ma profonde gratitude,*

*À mon amie **Hanane** et mon binôme **Aya**, Je ne peux pas trouver les mots justes et sincères pour vous exprimer mon affection et mes pensées, vous êtes pour moi des sœurs et des amies sur qui je peux compter.*

*Enfin, je tiens à remercier tous ceux qui m'ont soutenu et qui m'ont connu sans exception.*

*Manna*

## *Dédicace*

*Je dédie ce mémoire à ma chère mère **Salima**, pour son amour, sa patience et*

*Ses sacrifices innombrables qui m'ont permis de poursuivre mes rêves.*

*À mon cher père **Toumi** pour ses encouragements constants, sa sagesse et son soutien inébranlable*

*À mes chers frères **Fayçal** et **Marmar** pour leur camaraderie et leurs encouragements constants*

*À ma belle-sœur **Achwek** pour sa gentillesse et son soutien.*

*À mes chers neveux **Amir** et **Tadj Eddine** pour leur joie et leur inspiration.*

*À mes chers **Oncles** et mes chères **Tantes** pour leur bienveillance et leurs encouragements.*

*À ma chère grand-mère **Fatma** pour son amour et sa sagesse qui m'ont toujours guidé.*

*À mes chers **cousins** et mes chères **cousines** pour les moments de partage et de soutien.*

*À mes chères **amies** pour leur présence réconfortante et leurs encouragements inestimables.*

*À mes chers **collègues** de travail **crèche Baraa** et **Taim les petits génies** pour leur soutien*

*Professionnel et leur amitié précieuse.*

*À la famille de ma collègue **Marna** pour leur gentillesse et leur soutien durant ce parcours.*

*À toute ma famille élargie pour leur soutien inconditionnel et leur présence réconfortante.*

*Merci à vous tous pour votre présence et votre contribution inestimable à cette réussite.*

*Aya*

## Résumé

Le Coran est considéré comme la principale source de la loi islamique, mais beaucoup trouvent difficile de le mémoriser et de le réciter correctement. Les modèles d'apprentissage profond ont émergé comme une solution efficace pour faciliter ce processus. Dans ce mémoire, nous introduisons une nouvelle approche utilisant un réseau de neurones convolutifs (CNN) pour valider la récitation du Coran. Le modèle a été entraîné et testé à l'aide de la base de données "quran-reciters" et s'est avéré efficace avec un taux de précision de 95% sur les données de test. Cette méthode promet d'aider les pratiquants à vérifier et à améliorer leur récitation du Coran.

**Mots clés:** L'apprentissage en profondeur, CNN, Quran, Verset, Spectrogramme, Convolution, Réseau de neurones, MFCC.

## Abstract

The Quran is considered the primary source of Islamic law, but many find it difficult to memorize and recite it accurately. Deep learning models have emerged as an effective means to facilitate this process. In this work, we introduced a novel deep learning approach using a CNN model to identify the numbers of verses and surah from audio recordings of Quranic recitation, thus facilitating memorization. We trained and tested the model using the "quran-reciters" database, which proved effective with an accuracy rate of 95% on the test data.

**Keywords:** Deep Learning, CNN, Quran, Verset, Spectrogramme, Convolution, Réseau de neurones, MFCC.

## ملخص

القرآن هو المصدر الرئيسي للشريعة الإسلامية، ولكن الكثير يجدون صعوبة في حفظه وتلاوته بدقة. ظهرت النماذج العميقة للتعليم كوسيلة فعالة لتسهيل هذه العملية. في هذا العمل، قدمنا نهجاً عميقاً جديداً باستخدام نموذج CNN لتحديد أرقام الآيات والسور من التسجيلات الصوتية لتلاوة القرآن، مما يسهل عملية الحفظ. قمنا بتدريب واختبار النموذج باستخدام قاعدة بيانات "quran-reciters"، والتي أثبتت فعاليتها بدقة تبلغ 95% على بيانات الاختبار.

**الكلمات الرئيسية:** التعلم العميق، CNN، القرآن، الآية، السورة، الطيف، التلاوة، التحوير التكراري للمعاملات، الشبكة العصبية CNN

## Table Des Matières

Introduction Générale .....	1
1.1 Introduction.....	3
1.2 Définition de l'intelligence Artificielle .....	3
1.3 L'historique d'IA.....	4
1.4 Les fondements de l'intelligence artificielle .....	4
1.4.1 Agent intelligent.....	5
1.4.2 L'intelligence Artificielle Distribuée (IAD).....	5
1.5 Les réseaux de neurones artificiels .....	5
1.6 Deep Learning.....	6
1.7 La relation entre Intelligence Artificielle, Deep Learning, Machine Learning.....	6
1.8 Les différents type des modèles Deep Learning .....	8
1.8.1 Les réseaux de neurones récurrents (RNN) .....	8
1.8.2 Les réseaux de neurones convolutifs (CNN) .....	8
1.8.3 Les réseaux antagonistes génératifs (GAN).....	9
1.8.4 Les réseaux LSTM.....	9
1.8.5 Les transformer .....	10
1.9 Les techniques d'apprentissage .....	10
1.9.1 L'apprentissage supervisé .....	10
1.9.2 L'apprentissage non supervisé .....	10
1.9.3 L'apprentissage semi supervisé .....	11
1.10 Les algorithmes d'apprentissage supervisé .....	11
Les arbres de décision .....	11
La régression linéaire .....	12
Les machines à vecteurs de support.....	12
1.11 Les algorithmes d'apprentissage non supervisé .....	13
Le clustering K-means .....	13
1.12 La reconnaissance automatique de la parole.....	13
1.13 Le processus de la reconnaissance vocale .....	13
1.13.1 Acquisition du signal vocal (Signal de parole) .....	14
1.13.2 Prétraitement du signal .....	14
1.13.3 La modélisation acoustique.....	14
1.13.4 Décodage phonétique.....	15
1.14 Conclusion .....	15

Chapitre 2 : La science de la récitation du Coran .....	16
2.1 Introduction.....	16
2.2 Définition du Coran .....	16
2.3 Définition du Tajwid .....	16
2.4 Sections du Tajwid et Les jugement de chaque elle.....	16
2.5 Les niveaux de récitation .....	17
2.6 La faute dans la lecture (Al-lahn) .....	17
2.7 L'arrêt.....	17
2.7.1 L'arrêt Contraint.....	18
2.7.2 L'arrêt évaluative .....	18
2.7.3 L'arrêt attente .....	18
2.7.4 L'arrêt optionnelle.....	18
2.8 Le début .....	19
2.8.1 Le véritable début .....	19
2.8.2 Le début additionnel .....	19
2.8.3 Le début absolu .....	20
2.8.4Le début approprié .....	20
2.9 Dessin coranique الرسم القرآني.....	20
2.9.1 Définition du dessin coranique .....	20
2.9.2 Avantages du dessin coranique .....	20
2.10 Conclusion .....	22
Chapitre 3: L'apport du Deep Learning dans la reconnaissance vocale de la récitation coranique.....	23
3.1 Introduction.....	23
3.2 Les différentes approches de Deep Learning pour la reconnaissance vocale .....	23
3.2.1 Réseaux de Neurones Récurrents (RNN) .....	23
3.2.2 Réseaux de Neurones Convolutifs (CNN).....	24
3.2.3 Modèles de langage neuronal (Les transformers).....	25
3.3 La reconnaissance vocale de la récitation coranique .....	25
3.4 présentation des solutions envisagées .....	26
3.4.1 Objectifs du système .....	26
3.5 Solution 1 : Convertir la récitation audio en texte .....	27
3.6 Solution 2 : Comparer la récitation avec une référence audio .....	27
3.7 Choix de la solution .....	28
1.7 Conclusion .....	28

Chapitre 4 : La présentation de l'architecture du système et les résultats .....	29
4.1 Introduction.....	29
4.2 La description modulaire du système .....	29
4.3 Collecte des données.....	30
4.4 Prétraitement des données audio.....	30
4.4.1 Connecter le Google Drive à Google Colab .....	30
4.4.2 Conversion des fichiers audio en spectrogrammes.....	31
4.4.3 Division des données en ensembles d'entraînement, de validation et de test .....	31
4.4.4 Création d'un fichier de métadonnées .....	31
4.4.5 Conversion des images en tenseurs NumPy .....	32
4.5 Définition de l'architecture du réseau neuronal .....	32
4.5.1 Principe d'architecture d'un CNN .....	32
4.5.2 Couches de réseaux neuronaux convolutifs.....	33
4.6 Modèle Implementation .....	38
4.6.1 Architecture du modèle .....	39
4.7 Les outils et environnement de travail .....	40
4.7.1 Le Langage de Programmation Python.....	40
4.7.2 Environnement de l'exécution .....	40
4.7.3 Les bibliothèques utilisées .....	40
4.8 Présentation des méthodes .....	41
4.8.1 L'évolution du signal .....	41
4.8.2 Les métriques d'évaluations .....	42
4.8.3 La Fonction de Perte (Loss Function).....	44
4.9 L'entraînement .....	44
4.9.1 Courbe de Perte.....	45
4.9.2 Courbe de Précision .....	45
4.10 Les résultats des métriques d'évaluation .....	46
4.11 Démonstration de l'interface.....	47
4.11.1 Architecture Technique .....	47
4.11.2 Composants Clés.....	47
4.12.3 Flux de Travail .....	47
4.13.4 Intégration du Modèle.....	48
.....	48
.....	49
4.14 Discussion .....	50



4.15 Conclusion .....	51
Conclusion Générale.....	52
Bibliographie.....	53

## Liste Des Abréviations

<b>CNN</b>	: Convolutional Neural Networks
<b>CAN</b>	: Conversion Analogique Numérique
<b>CONV</b>	: Couche de Convolution
<b>DL</b>	: Deep Learning
<b>DNN</b>	: Deep Neural Network
<b>GAN</b>	: Generative Adversarial Networks
<b>GRU</b>	: Gated Recurrent Units
<b>IA</b>	: Intelligence Artificielle
<b>IAD</b>	: Intelligence Artificielle Distribuée
<b>KNN</b>	: K-Nearest Neighbors
<b>LSTM</b>	: Long-Short Term Memory
<b>ML</b>	: Machine Learning
<b>MLP</b>	: Multi-Layer Perceptron
<b>MFCC</b>	: Mel Frequency Cepstrum Coefficients
<b>RNN</b>	: Recurrent Neural Network
<b>RAP</b>	: Reconnaissance Automatique de la Parole
<b>SVM</b>	: Support Vector Machine
<b>SMA</b>	: Système Multi-Agents
<b>Relu</b>	: Rectified Linear Unit

## Liste Des Figures

**Figure 1** : Le processus de la reconnaissance vocale

**Figure 2** : La conversion d'un signal analogique

**Figure 3** : Classification du système de reconnaissance vocale

**Figure 4** : La description modulaire du système

**Figure 5** : Exemple d'architecture d'un CNN à 2 couches

**Figure 6** : Convolution 3D

**Figure 7** : Exemple d'application de Pooling

**Figure 8** : Une Couche entièrement connectée

**Figure 9** : La fonction Softmax

**Figure 10** : Comparaison de taux d'erreur de classification avec et sans dropout

**Figure 11** : Architecture du modèle

**Figure 12** : les différentes bibliothèques utilisées dans notre implémentation

**Figure 13** : Exemple d'un spectrogramme normalisé

**Figure 14** : La table de matrice de confusion

**Figure 15** : Valeurs Loss et Val-Loss du modèle

**Figure 16** : Précision de la formation et précision de la validation

**Figure 17** : Interface de l'application

**Figure 18** : Interface sélection de sourah et verset

**Figure 19** : Interface de test (correct)

**Figure 20** : Interface de test (incorrect)

## Liste Des Tableaux

**Tableau 1** : Comparaison entre Intelligence Artificielle, l'apprentissage automatique et l'apprentissage profond.

**Tableau 2** : Fonction d'activation utilisée dans les réseaux neuronaux synthétiques

**Tableau 3** : Les résultats de notre méthodologie



### Introduction Générale

Le Saint Coran, ouvrage sacré vénéré par les musulmans du monde entier, occupe une place centrale dans leur vie spirituelle et culturelle. Toutefois, mémoriser et réciter avec exactitude ses versets représente un défi considérable, surtout pour les débutants. Dans cette optique, l'intelligence artificielle, et spécifiquement l'apprentissage profond, émerge comme un acteur majeur.

L'assistance apportée aux musulmans, en particulier aux apprentis, pour mémoriser et réciter le Coran correctement est un sujet d'une grande envergure. Cette initiative contribue non seulement à préserver l'intégrité du texte sacré, mais également à faciliter la compréhension et la diffusion de son message divin. En concevant un système basé sur l'apprentissage profond capable de validation de récitation, nous mettons à disposition des apprenants un outil précieux pour leur entraînement et leur perfectionnement. En plus, ce travail illustre le potentiel de l'intelligence artificielle dans la sauvegarde du patrimoine culturel et religieux.

Cela pose plusieurs questions : Comment l'apprentissage profond peut-il contribuer à améliorer la mémorisation et la récitation précise du Coran malgré sa complexité linguistique? Quelle est la manière de concevoir un système qui puisse vérifier de manière précise et fiable une récitation Coranique? Pour quelle raison est-il essentiel de mettre en place de tels outils afin de simplifier l'apprentissage du Coran?

Selon une récente étude, près de 10 millions de musulmans à travers le monde ont mémorisé l'intégralité du Coran. Néanmoins, même les récitants les plus expérimentés peuvent commettre des erreurs de récitation, en raison de la complexité du texte et des défis associés à sa mémorisation. C'est dans ce contexte que l'intelligence artificielle peut intervenir pour aider les fidèles à parfaire leur récitation du texte sacré.

En combinant des techniques avancées d'apprentissage profond avec une connaissance approfondie du Coran et de ses règles de récitation, ce projet ambitionne de développer un outil novateur. Cet outil permettra aux apprenants de tous niveaux de progresser dans la mémorisation et la récitation du texte sacré, tout en préservant sa richesse et sa beauté intrinsèques.

Notre mémoire est ainsi organisé :



Un premier chapitre «Introduction à l'intelligence artificielle», nous faisons un bref tour d'horizon sur l'intelligence artificielle de manière générale, et un aperçu sur La reconnaissance automatique de la parole et exposons sa définition ainsi que ses processus.

Dans le deuxième chapitre «La science de la récitation du Coran», Nous nous sommes fait une étude dans des règles des Tajwid et du dessin coranique.

En troisième chapitre : «L'apport du deep learning dans la reconnaissance vocale de la récitation coranique ».

En fin, la quatrième chapitre : «La présentation de l'architecture du système et les résultats ».



# Chapitre 1 : Introduction à l'intelligence artificielle

## 1.1 Introduction

L'intelligence artificielle (IA) est un domaine interdisciplinaire visant à créer des systèmes capables d'effectuer des tâches nécessitant normalement l'intelligence humaine. Depuis les années 1950, l'IA a évolué de manière spectaculaire, transformant notre quotidien et nos interactions avec le monde. Un élément central de l'IA est la capacité d'apprentissage des machines. Contrairement aux programmes traditionnels qui suivent des instructions préétablies, les systèmes d'IA peuvent apprendre à partir des données, identifier des schémas et prendre des décisions de manière autonome. Cette adaptabilité les distingue des technologies informatiques conventionnelles.

Les applications de l'IA sont vastes et variées, allant des assistants virtuels sur nos smartphones à la conduite autonome des véhicules, en passant par la prédiction économique et la découverte de nouveaux médicaments. L'IA est un moteur majeur de l'innovation dans de nombreux secteurs, offrant des opportunités inédites tout en soulevant des questions éthiques et sociales cruciales.

Un domaine spécifique de l'IA qui a suscité beaucoup d'intérêt est la reconnaissance automatique de la parole. La communication vocale, étant un moyen naturel et efficace d'interaction humaine, a poussé les chercheurs à développer des systèmes capables de comprendre et de répondre à la parole humaine. Cette technologie vise à simplifier la vie quotidienne en concevant des machines capables de remplacer l'homme dans diverses tâches complexes, telles que la reconnaissance vocale.

## 1.2 Définition de l'intelligence Artificielle

Depuis ses débuts en tant que domaine d'étude au milieu du XXe siècle, l'intelligence artificielle (IA) représente une frontière constamment repoussée. [1]

L'intelligence Artificielle est un domaine de l'informatique qui vise à créer des systèmes capables de réaliser des tâches nécessitant normalement une intelligence humaine. Ces systèmes apprennent à partir de données, s'adaptent à de nouvelles situations et résolvent des problèmes de manière autonome. L'IA comprend diverses techniques telles que l'apprentissage automatique, les réseaux neuronaux, le traitement du langage naturel et la



vision par ordinateur. Son objectif est de développer des systèmes intelligents capables de prendre des décisions autonomes dans des environnements complexes. [2]

### 1.3 L'histoire d'IA

L'histoire de l'intelligence artificielle (IA) remonte aux travaux de pionniers comme Kurt Gödel, Alonso Church et Alan Turing dans les années 1930. Gödel a formulé des théorèmes fondamentaux sur la logique mathématique, révélant à la fois la complétude et l'incomplétude de certains systèmes logiques. Alan Turing a quant à lui démontré l'indécidabilité du problème de l'arrêt, montrant qu'il n'existe pas de programme capable de déterminer si un autre programme s'exécute indéfiniment.

Dans les années 1940 et 1950, les bases théoriques de l'IA ont été établies, avec des contributions telles que le modèle de neurones formels de McCulloch et Pitts, les premières réflexions sur l'intelligence artificielle de Turing, et les premiers pas dans l'apprentissage automatique avec le programme de jeu de dames d'Arthur Samuel.

Le véritable essor de l'IA a eu lieu dans les années 1950 et 1960, avec des événements marquants tels que la conférence de Dartmouth en 1956, qui a introduit pour la première fois le terme "intelligence artificielle", ainsi que le développement de programmes symboliques comme le Logician Theorist de Newell et Simon.

Dans les décennies suivantes, l'IA a connu des avancées significatives, notamment avec l'émergence des réseaux neuronaux dans les années 1980, l'utilisation de la logique probabiliste dans les années 1990, et les percées dans l'apprentissage profond et l'IA créative au 21<sup>e</sup> siècle.

Donc, l'histoire de l'IA est marquée par une succession de découvertes et d'avancées théoriques et technologiques, alimentant un domaine en constante évolution qui continue de repousser les limites de ce que les machines peuvent accomplir. [3]

### 1.4 Les fondements de l'intelligence artificielle

Les fondements de l'IA reposent sur plusieurs disciplines, notamment les mathématiques, la logique, la philosophie, la psychologie cognitive et les neurosciences. L'objectif est de développer des algorithmes et des modèles qui permettent aux machines d'apprendre à partir de données, de raisonner, de percevoir leur environnement et d'interagir de manière autonome avec celui-ci. Les principales approches de l'IA incluent l'apprentissage automatique, qui repose sur l'utilisation de données pour entraîner des modèles prédictifs, et





l'intelligence artificielle symbolique, qui utilise des règles logiques pour manipuler des symboles et résoudre des problèmes. L'IA est omniprésente dans de nombreux domaines, de la recherche scientifique à la médecine, en passant par les applications industrielles et les systèmes autonomes. Son développement soulève également des questions éthiques et sociales importantes, notamment en ce qui concerne la confidentialité des données, la discrimination algorithmique et l'impact sur l'emploi. [2]

### 1.4.1 Agent intelligent

Un agent intelligent représente un programme informatique capable de fonctionner de manière autonome. Il interagit avec son environnement en percevant des événements et en agissant en fonction d'objectifs spécifiques définis par sa programmation. [4] Par exemple un système de contrôle de la circulation urbaine. Ces systèmes utilisent des capteurs de circulation, des caméras de surveillance, des données GPS et d'autres sources pour surveiller et gérer le flux de circulation dans une ville. L'agent intelligent peut percevoir les conditions actuelles de la circulation, telles que la densité du trafic, les accidents, les travaux routiers, etc...[5]

### 1.4.2 L'intelligence Artificielle Distribuée (IAD)

L'intelligence artificielle distribuée (IAD) représente une branche de l'intelligence artificielle axée sur la création de systèmes décentralisés, typiquement composés de multiples agents, qui peuvent collaborer et se synchroniser entre eux. Son objectif principal est d'explorer les techniques permettant à ces agents autonomes d'interagir de manière efficace et de partager la charge de travail de manière équitable. Inspirée par les structures complexes observées dans certaines sociétés d'insectes telles que les fourmis, l'IAD cherche à résoudre des problèmes de coordination, notamment dans des domaines tels que la gestion du trafic aérien ou la navigation des véhicules autonomes, où les agents doivent apprendre à éviter les collisions tout en respectant des contraintes de temps et de trajectoire. [6]

## 1.5 Les réseaux de neurones artificiels

Les réseaux de neurones artificiels, s'inspirant du fonctionnement des neurones du cerveau humain. Ils sont composés de neurones artificiels interconnectés, organisés en couches, qui transmettent des signaux entre eux.[7] Les neurones, qui sont les unités de base des réseaux de neurones, reçoivent une entrée, appliquent une fonction d'activation et



transmettent une sortie aux neurones de la couche suivante. Les couches sont composées d'une couche d'entrée, d'une ou plusieurs couches cachées qui effectuent des transformations non linéaires sur les données d'entrée, et d'une couche de sortie.

Les neurones sont connectés entre les couches par des poids synaptiques, qui déterminent l'importance de chaque connexion dans le calcul de la sortie. Le processus de propagation avant transmet les données d'entrée à travers le réseau pour obtenir une sortie, tandis que la rétro-propagation de l'erreur est l'algorithme d'apprentissage qui ajuste les poids du réseau en propageant l'erreur de la sortie vers les couches précédentes, afin de minimiser l'erreur globale. Les réseaux de neurones peuvent être entraînés de manière supervisée avec des données étiquetées ou de manière non supervisée sans étiquettes. [8]

### 1.6 Deep Learning

Le Deep Learning (DL) est une méthode d'apprentissage automatique qui utilise des structures multicouches pour traiter l'information de manière complexe. Il se concentre sur la création de représentations hiérarchiques des données, en identifiant des caractéristiques de plus haut niveau à partir de celles de plus bas niveau. Ces caractéristiques sont apprises à partir des données elles-mêmes, ce qui permet au système de reconnaître des motifs et de classer les informations de manière automatique. Le point principal est que le Deep Learning utilise des réseaux neuronaux artificiels pour apprendre des niveaux de représentation des données, facilitant ainsi la compréhension et le traitement des informations. [9]

### 1.7 La relation entre Intelligence Artificielle, Deep Learning, Machine Learning

Dans le paysage technologique actuel, les termes intelligence artificielle (IA), apprentissage automatique (ML) et apprentissage profond (DL) sont souvent utilisés de manière interchangeable, mais ils représentent en fait des concepts distincts au sein d'une hiérarchie.

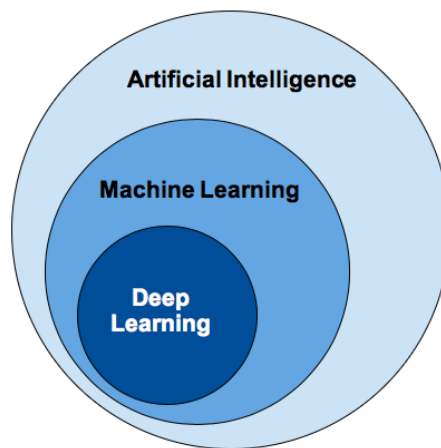
Dans cette hiérarchie, l'intelligence artificielle englobe l'idée d'infuser des comportements et une intelligence similaires à ceux des humains dans des machines ou des systèmes. L'apprentissage automatique, quant à lui, est une méthodologie permettant aux systèmes d'apprendre à partir de données ou d'expériences, automatisant ainsi la création de modèles analytiques.



## CHAPITRE 1: Introduction à l'intelligence artificielle

Au sein de l'apprentissage automatique, l'apprentissage profond occupe une position particulière. Il utilise des réseaux neuronaux à plusieurs couches pour traiter les données, en intégrant plusieurs niveaux de traitement pour construire des modèles basés sur les données. Le terme "profond" dans l'apprentissage profond met l'accent sur cette utilisation de multiples niveaux de traitement.

L'intelligence artificielle représente le concept global d'infusion d'intelligence dans les machines, tandis que l'apprentissage automatique est une technique pour automatiser la création de modèles à partir de données. L'apprentissage profond se situe au sein de l'apprentissage automatique, utilisant des réseaux neuronaux à plusieurs couches pour créer des modèles basés sur les données, en mettant en avant l'aspect "profond" de traitement en plusieurs niveaux. Cette hiérarchie clarifie les rôles distincts de l'IA, de l'apprentissage automatique et de l'apprentissage profond dans le domaine des systèmes et des logiciels intelligents. [10]



**Figure 1** : la relation entre AI, ML, DL [10]

Voici une comparaison entre Intelligence Artificielle, l'apprentissage automatique et l'apprentissage profond est résumé dans le tableau 1 :

<b>Intelligence Artificielle (IA)</b>	<b>apprentissage automatique (ML)</b>	<b>apprentissage profond (DL)</b>
IA simule l'intelligence humaine pour effectuer des tâches et prendre des décisions.	ML est un sous-ensemble d'AI qui utilise des algorithmes pour apprendre des modèles à partir de données.	DL est un sous-ensemble du ML qui utilise des réseaux de neurones artificiels pour des tâches complexes.
IA peut nécessiter ou non de grands ensembles de	Le ML s'appuie fortement sur des données étiquetées pour	DL nécessite de nombreuses données étiquetées et



données ; il peut utiliser des règles prédéfinies.	la formation et la réalisation de prédictions.	fonctionne exceptionnellement avec de grands ensembles de données.
IA peut être basé sur des règles, nécessitant une programmation et une intervention humaines.	ML automatise l'apprentissage à partir des données et nécessite moins d'intervention manuelle.	DL automatise l'extraction de fonctionnalités, réduisant ainsi le besoin d'ingénierie manuelle.
Tous les algorithmes peuvent être simples ou complexes, selon l'application.	ML utilise divers algorithmes tels que des arbres de décision, SVM et des forêts aléatoires.	DL s'appuie sur des réseaux de neurones profonds, qui peuvent comporter de nombreuses couches cachées pour un apprentissage complexe.
IA peut gérer diverses tâches, des plus simples aux plus complexes, dans tous les domaines.	ML se spécialise dans les tâches basées sur les données telles que la classification, la régression, etc.	DL excelle dans les tâches complexes telles que la reconnaissance d'images, le traitement du langage naturel, etc.

**Tableau 1** : Comparaison entre Intelligence Artificielle, l'apprentissage automatique et l'apprentissage profond. [11]

## 1.8 Les différents type des modèles Deep Learning

### 1.8.1 Les réseaux de neurones récurrents (RNN)

Les réseaux de neurones récurrents (RNN) sont un type de réseau de neurones artificiels (RNA) conçus pour traiter des données séquentielles, telles que le texte, la parole ou les signaux temporels. Contrairement aux réseaux de neurones artificiels classiques, qui traitent chaque entrée indépendamment, les RNN peuvent prendre en compte le contexte des données d'entrée, ce qui leur permet de capturer des dépendances temporelles et d'effectuer des tâches plus complexes.

Les RNN sont composés de neurones artificiels interconnectés, organisés en couches. Chaque neurone reçoit une entrée, la traite et produit une sortie. La particularité des RNN est que la sortie d'un neurone peut être réinjectée comme entrée dans d'autres neurones, ce qui permet au réseau de "se souvenir" des informations des entrées précédentes. [10]

### 1.8.2 Les réseaux de neurones convolutifs (CNN)

Les réseaux neuronaux convolutifs (CNN) ont été introduits par Yann Lecun en 1988 pour le traitement d'image et la détection d'objets, se distinguant par une topologie en grille. Extrêmement efficaces, ils sont utilisés dans divers domaines tels que la reconnaissance d'images et de vidéos, la conduite autonome, et bien d'autres. En machine learning, un CNN



est un type de réseau feed-forward inspiré par des processus biologiques, comprenant cinq opérations clés : convolution, ReLU, pooling, couche entièrement connectée, et couche de perte.

Le CNN, est une architecture de deep learning qui apprend directement de l'entrée sans nécessiter d'extraction manuelle de caractéristiques. Il améliore la conception des réseaux de neurones traditionnels, en réduisant la complexité du modèle et en optimisant les paramètres à chaque couche. Les CNN sont spécifiquement conçus pour traiter des formes en 2D et sont largement utilisés dans des domaines tels que la reconnaissance visuelle, l'analyse d'images médicales, la segmentation d'images, et le traitement du langage naturel. Leur capacité à découvrir automatiquement des caractéristiques sans intervention humaine les rend plus puissants que les réseaux traditionnels. Diverses variantes existent, adaptées à des applications spécifiques en fonction de leurs capacités d'apprentissage. [10]

### 1.8.3 Les réseaux antagonistes génératifs (GAN)

Les réseaux antagonistes génératifs (GAN) sont des algorithmes de deep learning permettant de créer de nouvelles données ressemblant à celles sur lesquelles ils ont été formés. Composés d'un générateur et d'un discriminateur, les GAN sont utilisés dans divers domaines tels que la synthèse de mouvement respiratoire en imagerie tomodensitométrie 4D, la reconstruction super-résolution et la segmentation en imagerie par résonance magnétique (IRM). Conçus par Ian Goodfellow, les GAN impliquent deux réseaux neuronaux en compétition : un générateur qui crée des données réalistes et un discriminateur qui cherche à les distinguer des données réelles. Les GAN sont déployés pour des tâches d'apprentissage non supervisé, mais sont également utilisés dans des contextes semi-supervisés et de renforcement. Leur champ d'application couvre la santé, l'analyse d'images, l'augmentation de données, la génération vidéo et vocale, ainsi que des domaines émergents comme le cyber sécurité et le contrôle du trafic. En résumé, les GAN représentent une solution complète pour étendre indépendamment les données et résoudre divers problèmes nécessitant une approche générative. [10]

### 1.8.4 Les réseaux LSTM

Les LSTM, ou Réseaux Neuronaux à Long Terme et Court Terme, sont des variantes de RNN qui excellent dans l'apprentissage de dépendances à long terme. Ils surmontent les



limitations des RNN traditionnels en utilisant des unités spéciales et en introduisant des portes pour contrôler le flux d'information. Cette capacité en fait un choix prisé dans des applications telles que la composition musicale, la reconnaissance, et le développement de médicaments. En particulier, l'architecture LSTM est reconnue pour résoudre le défi de la disparition des gradients, ce qui en fait l'une des formes de RNN les plus réussies. [10]

### 1.8.5 Les transformer

Le Transformer est un modèle de deep learning qui a été largement utilisé dans différents domaines, tels que le traitement du langage naturel (NLP), la vision par ordinateur (CV) et le traitement de la parole. Le Transformer, initialement destiné à la traduction automatique, s'est révélé être un outil puissant grâce à ses versions améliorées et aux modèles pré-entraînés qui s'en sont inspirés. On l'a réussi à utiliser dans des domaines tels que la reconnaissance d'images, le traitement audio et même dans des domaines non conventionnels tels que la chimie et les sciences de la vie. Cette adoption généralisée témoigne de la capacité du Transformer à offrir des performances de pointe dans diverses tâches et à être une architecture de choix pour de nombreuses applications de deep learning. [12]

## 1.9 Les techniques d'apprentissage

### 1.9.1 L'apprentissage supervisé

L'apprentissage supervisé est une méthode qui utilise des données étiquetées, où un système apprend à partir d'un ensemble d'entrées et de leurs sorties correspondantes. Dans les approches de deep learning supervisées, le système ajuste itérativement ses paramètres pour mieux approximer les sorties souhaitées. Les principales techniques incluent les réseaux de neurones profonds (DNN), les réseaux de neurones convolutifs (CNN), et les réseaux de neurones récurrents (RNN), tels que les mémoires à court et long terme (LSTM) et les unités récurrentes fermées (GRU). [13]

### 1.9.2 L'apprentissage non supervisé

L'apprentissage non supervisé est une méthode où le système apprend sans la présence d'étiquettes de données. Dans ce processus, le réseau apprend à représenter les données internes ou à identifier des caractéristiques importantes pour découvrir des relations ou une structure inconnue dans les données d'entrée. Les approches courantes incluent le regroupement, la réduction de la dimensionnalité et les techniques génératives. Les techniques



populaires de deep learning pour cela comprennent les encodeurs automatiques (AE), les machines de Boltzmann restreintes (RBM) et les réseaux adversaires génératifs (GAN). De plus, les réseaux de neurones récurrents (RNN) tels que le LSTM sont également utilisés dans de nombreux domaines pour l'apprentissage non supervisé. [13]

### 1.9.3 L'apprentissage semi supervisé

L'apprentissage semi-supervisé est une méthode qui repose sur des ensembles de données partiellement étiquetés, parfois également appelés apprentissage par renforcement. Cette approche utilise des techniques telles que les réseaux adversaires génératifs (GAN), les réseaux de neurones récurrents (RNN) comme LSTM et GRU, ainsi que l'apprentissage par renforcement profond (DRL). Elle permet d'exploiter à la fois les données étiquetées et non étiquetées pour entraîner un modèle. [13]

## 1.10 Les algorithmes d'apprentissage supervisé

### Les arbres de décision

La technique des arbres de décision repose sur la notion simple de classifier un objet en effectuant une série de tests sur ses attributs. Ces tests sont organisés de manière à ce que la réponse à l'un d'eux détermine le test suivant à appliquer à cet objet. Cette approche de classification est couramment utilisée dans les sciences naturelles, où l'espace de représentation est défini par l'observation des caractéristiques anatomiques pertinentes de la plante, telles que les étamines, la corolle, le calicule, les bractées, etc.

Chaque attribut est évalué pour déterminer quel test effectuer ensuite, dans le but de classer la plante dans l'une des classes prédéfinies. Cette méthode nécessite la connaissance et la mesure d'une variété d'attributs pour chaque plante répertoriée, ce qui définit la taille de l'espace de représentation.

Le processus de classification est organisé comme un arbre, où chaque nœud représente un test sur un ou plusieurs attributs, et chaque feuille correspond à une classe spécifique. La classification se déroule en partant de la racine de l'arbre et en suivant récursivement les tests jusqu'à atteindre une feuille, ce qui définit ainsi l'espèce de la plante. L'objectif principal est d'apprendre à construire de tels arbres de décision à partir d'exemples, ce qui permet de généraliser la méthode de classification pour de nouvelles observations. [14]



### La régression linéaire

La régression linéaire est un algorithme de machine Learning qui vise à prédire une valeur scalaire ( $y$ ) à partir d'un vecteur d'entrée ( $x$ ) en utilisant une fonction linéaire. Cette fonction linéaire est déterminée par un ensemble de paramètres ( $w$ ) qui sont ajustés afin de minimiser l'erreur de prédiction. Les paramètres contrôlent l'impact de chaque caractéristique d'entrée sur la prédiction finale : un poids positif augmente la prédiction en augmentant la valeur de la caractéristique correspondante, tandis qu'un poids négatif diminue la prédiction. L'objectif est de trouver les meilleurs poids ( $w$ ) en minimisant l'erreur quadratique moyenne (MSE) entre les prédictions et les valeurs réelles sur un ensemble de données d'entraînement. [15]

### Les machines à vecteurs de support

Les machines à vecteurs de support ou Support Vector Machine (SVM) sont une méthode influente d'apprentissage supervisé. Elles utilisent une fonction linéaire pour prendre des décisions de classification grâce à un concept appelé le "kernel trick", elles peuvent être étendues pour effectuer des classifications non linéaires. Le "kernel trick" permet de calculer des produits scalaires dans des espaces de caractéristiques de dimension supérieure sans réellement calculer ces transformations. Cela permet aux SVM d'apprendre des modèles non linéaires tout en conservant des propriétés mathématiques favorables. Les SVM sont capables de généraliser et de classer de nouvelles données en utilisant des exemples de formation qui sont les plus significatifs, appelés vecteurs de support. Bien que puissantes, les SVM peuvent être coûteuses en termes de temps de calcul, en particulier avec de grands ensembles de données, et elles peuvent être moins efficaces pour généraliser dans des situations où les données sont hautement non linéaires. [14]

Les machines à vecteurs de support (SVM) peuvent ne pas être adaptées pour traiter des données complexes et fortement interconnectées, comme les images haute résolution. Leur coût computationnel élevé et leur difficulté à définir des frontières de décision claires dans des espaces de grande dimension peuvent poser problème. Dans de tels cas, les approches d'apprentissage profond, comme les réseaux de neurones convolutifs, sont souvent préférées. Ces derniers sont plus efficaces pour extraire des caractéristiques complexes et capturer des relations non linéaires.





### 1.11 Les algorithmes d'apprentissage non supervisé

#### Le clustering K-means

Le clustering K-means est un algorithme de représentation simple qui divise un ensemble d'entraînement en  $k$  clusters différents, regroupant les exemples similaires. Chaque exemple est assigné à un cluster représenté par un vecteur one-hot de dimension  $k$ , où une seule entrée est égale à 1 pour le cluster auquel l'exemple appartient, et toutes les autres entrées sont nulles.

Cependant, le problème de clustering est intrinsèquement mal posé car il n'existe pas de critère unique pour mesurer la qualité du regroupement des données. Les différentes méthodes de clustering peuvent produire des regroupements différents, chacun reflétant différentes propriétés des données. Une représentation distribuée peut être préférable à une représentation one-hot car elle permet de capturer plusieurs attributs pour chaque exemple, offrant ainsi une manière plus riche de mesurer la similarité entre les objets. [15]

### 1.12 La reconnaissance automatique de la parole

La reconnaissance automatique de la parole (RAP) est une technologie qui permet à un ordinateur de comprendre et d'interpréter le langage parlé, facilitant ainsi les interactions homme-machine via des commandes vocales, des dictées et des réponses vocales interactives. Elle offre également des avantages significatifs en matière d'accessibilité, permettant aux personnes handicapées de communiquer plus facilement, ainsi que des possibilités d'apprentissage des langues étrangères et d'assistance personnelle. [16]

### 1.13 Le processus de la reconnaissance vocale

La reconnaissance automatique de la parole consiste à transcrire en temps réel une série de mots prononcés en un texte interprétable par la machine, de manière entièrement automatisée.

Ce processus implique plusieurs étapes, comme illustré dans la Figure 2.

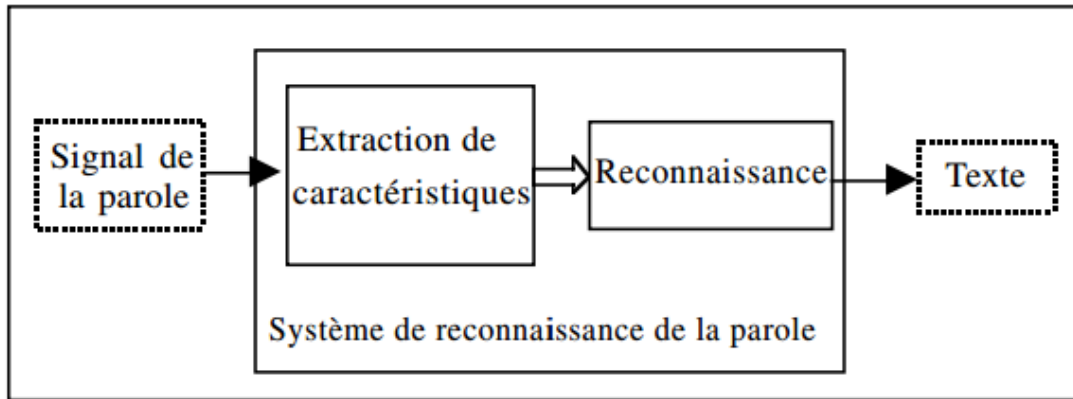


Figure 2: Le processus de la reconnaissance vocale [17]

### 1.13.1 Acquisition du signal vocal (Signal de parole)

Le processus commence par l'acquisition du signal vocal à l'aide d'un microphone ou d'un autre dispositif d'enregistrement audio. Ce signal est une représentation analogique de la parole humaine.

### 1.13.2 Prétraitement du signal

Avant d'être analysé, le signal vocal brut subit un prétraitement visant à supprimer le bruit de fond, à ajuster le volume pour une uniformité et éventuellement à réduire sa complexité dimensionnelle.[18]

L'objectif principal de cette phase est de convertir le signal analogique recueilli précédemment en un signal numérique (conversion analogique-numérique, ou CAN), afin qu'il puisse être interprété par la machine.

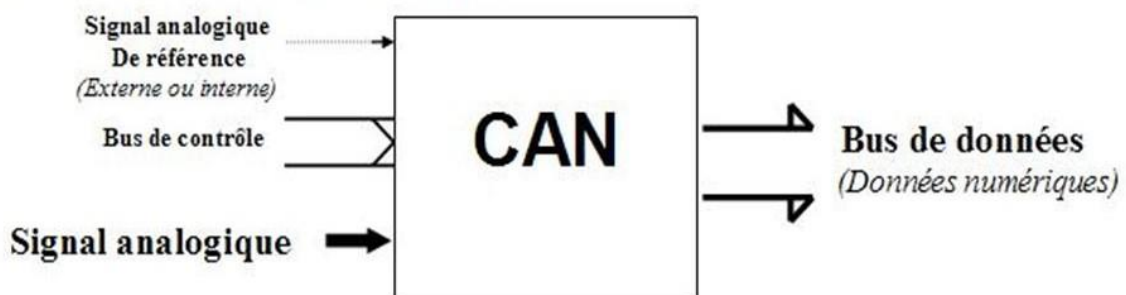


Figure 3 : La conversion d'un signal analogique [19]

### 1.13.3 La modélisation acoustique

La modélisation acoustique dans la reconnaissance vocale commence par la capture et le nettoyage du signal audio, suivi par l'extraction de caractéristiques à travers des techniques



comme les Coefficients Cepstraux de Fréquence Mel (MFCC) ou spectrogramme. Ces caractéristiques sont utilisées pour entraîner des modèles acoustiques, souvent basés sur des réseaux de neurones récurrents (RNN) ou convolutifs (CNN), qui alignent les séquences audio avec les phonèmes. Ensuite, un modèle linguistique fournit un contexte pour améliorer la précision en prédisant la probabilité des séquences de mots. Le décodeur combine ces informations pour générer la transcription la plus probable du signal audio, et un post-traitement valide la récitation pour affiner le résultat final.

### 1.13.4 Décodage phonétique

Dans le processus de décodage phonétique, le modèle acoustique est appliqué pour traduire le signal vocal en une série probable de phonèmes. Cette étape fait souvent appel à des techniques d'analyse telles que les modèles de Markov cachés ou les réseaux de neurones convolutifs.[20]

## 1.14 Conclusion

Ce chapitre a fourni une introduction exhaustive à l'intelligence artificielle, explorant ses concepts fondamentaux, son historique, et les bases des réseaux de neurones et du Deep Learning, tout en clarifiant les relations entre le DL, le ML et l'IA. De plus, nous avons examiné la reconnaissance automatique de la parole (RAP), en détaillant son processus, de l'acquisition du signal vocal au décodage phonétique. Cette technologie révolutionnaire améliore les interactions entre les individus et les machines, offrant des avantages significatifs en termes d'accessibilité et d'assistance personnelle, et ouvrant la voie à de nouvelles opportunités pour l'avenir technologique.



### Chapitre 2 : La science de la récitation du Coran

#### 2.1 Introduction

Dans ce chapitre, nous aborderons un élément fondamental de la pratique religieuse islamique la récitation du Coran et les règles qui la régissent, communément appelées Tajwid. La récitation correcte du Coran revêt une importance primordiale dans la tradition islamique, tant du point de vue spirituel que culturel. Ce chapitre se concentrera sur l'importance de la récitation coranique, ainsi que sur les principaux aspects du Tajwid, offrant ainsi un aperçu approfondi de cette pratique essentielle pour les musulmans du monde entier.

#### 2.2 Définition du Coran

Le Coran est la parole d'Allah révélée à Son prophète Muhammad, (ﷺ). Il confirme de manière miraculeuse la mission de Muhammad en tant que prophète choisi par Allah.

Malgré les répétitions qui peuvent se trouver dans le Coran, ses merveilles infinies continueraient d'émerveiller les lecteurs et les auditeurs. Parmi les versets qui indiquent sa vertu et encouragent à le lire, citons : La parole de Dieu Tout-Puissant : ( إِنَّ الَّذِينَ يَتْلُونَ كِتَابَ اللَّهِ ) (Fatir : 29). [21]

#### 2.3 Définition du Tajwid

Le Tajweed, un art de lecture psalmodiée du Coran qui vise à prononcer chaque lettre avec précision, tout en préservant ses caractéristiques distinctives. Des imams dignes de confiance l'ont documenté, posé ses principes et fondé ses décisions sur les méthodes de lecture correctes transmises par le Prophète (que Dieu le bénisse et lui accorde la paix), ainsi que par ses compagnons et disciples. Le but de la maîtrise de cette spécialité est d'assurer la prononciation correcte du Coran et d'éviter toute erreur linguistique lors de la récitation. [22]

#### 2.4 Sections du Tajwid et Les jugement de chaque elle

La science du Tajwid est divisée en deux Aspects : une Aspect pratique et une Aspect théorique.

**Aspect pratique**, c'est-à-dire qu'il s'agit du Tajweed lu par la langue lors de la récitation du Coran.

**Jugement:** La pratique du Tajwid lors de la récitation du Coran est **obligatoire (wājib)**. [23]



**Aspect théorique**, ce sont ces règles du Tajweed placées dans les livres. Cette section est l'application des règles du Tajweed dans la récitation, qui est l'enregistrement de ces règles lues. [23]

**Jugement** : L'apprentissage des règles du Tajwid écrites est **recommandé (mandoub)**, mais pas obligatoire.

### 2.5 Les niveaux de récitation

**La Méditation (Tahqiq)** : C'est similaire au Tartil, mais avec un niveau de réflexion plus profond, et c'est une étape utilisée dans l'enseignement.

**La vigilance (Hadur)** : implique une récitation rapide tout en respectant strictement les règles du Tajwid.

**La modulation (Tadwir)** : correspond à une récitation se situant entre la méditation et la vigilance, tout en respectant rigoureusement les règles du Tajwid. [24]

### 2.6 La faute dans la lecture (Al-lahn)

**Al-lahn, اللحن** est l'erreur dans la lecture du Saint Quran. On parle d'Al-lahn lorsque l'on dévie de la justesse dans la récitation.

Il y a 2 types d'erreurs (lahn) de la récitation :

**1. Lahn Jalî** : Il s'agit d'une erreur qui se produit sur les mots, perturbant ainsi la prononciation correcte, et modifiant la structure du mot, comme le remplacement de la lettre "ت" par "د" (« تَنْبُغُهَا الرَّادِفَةُ ») (Sourate An-Nazi'at, verset 7), ou de la lettre "ح" par "ع" (« فَاصْفَحْ عَنْهُمْ ») (Sourate Az-Zukhruf, verset 89). Ou le changement de mouvement vocalique, comme le passage d'un damma à une kasra (« أَنْعَمْتَ عَلَيْهِمْ ») (Sourate Al-Fatiha, verset 7). [25]

**2. Lahn Khafi** : C'est une erreur qui se produit sur les mots, violant ainsi les règles du Tajwid, comme la suppression d'une ghunnah ou d'une prolongation, et ainsi de suite.

Son jugement : Il est désapprouvé, voire interdit selon certains avis. Il est appelé "khafi" (caché) en raison de la spécialisation des experts de ce domaine dans sa connaissance. [24]

### 2.7 L'arrêt

L'arrêt est une pause effectuée à la fin ou au milieu des versets du Coran, permettant à la lecture de reprendre sa respiration. Il se fait en arrêtant la lecture à la fin d'un mot, en maintenant la dernière lettre si c'est une voyelle mobile, et en reprenant la lecture sans interrompre le flux. Cette pause ne nécessite pas de réciter la formule de protection, même si



elle est prolongée. Il est important de ne pas interrompre un mot ou une liaison graphique lors de cette pause, pour maintenir la continuité de la lecture. [25]

Les différentes catégories sont définies en fonction de la situation du 'waqf' (L'arrêt) :

Contraint, Évaluatif, Attendant, Optionnel.

### 2.7.1 L'arrêt Contraint

L'arrêt Contraint, ou «الوقف الاضطراري» est une pause qui survient pour le lecteur en raison de circonstances telles qu'une difficulté respiratoire, un étouffement, une toux ou un oubli, et il n'y a aucune faute commise par le lecteur car la cause de cette pause est indépendante de sa volonté. Dans ce cas, le lecteur est autorisé à faire une pause sur n'importe quel mot, même si le sens n'est pas complet, tel que faire une pause après une condition sans sa réponse, ou après une conjonction sans sa relation. Cependant, il est nécessaire de reprendre la lecture par le mot sur lequel on s'est arrêté si cela est approprié pour reprendre la lecture. [25]

### 2.7.2 L'arrêt évaluative

L'arrêt évaluative, ou "الوقف التقييمي" intervient lorsqu'un enseignant pose une question à un élève ou le guide dans son apprentissage. L'objectif est d'inciter l'élève à faire une pause sur un mot afin d'identifier certains éléments linguistiques tels que ce qui est isolé ou relié, ce qui est constant ou omis. Cela permet à l'enseignant d'enseigner à l'élève comment effectuer des pauses appropriées dans sa lecture. Cette technique n'est généralement pas utilisée dans la lecture courante, mais plutôt comme un outil pédagogique pour améliorer la compréhension et l'analyse linguistique. [25]

### 2.7.3 L'arrêt attente

L'arrêt en attente, ou "الوقف الانتظاري", est une pratique qui consiste à faire une pause sur chaque mot afin de le relier à d'autres en utilisant différentes méthodes de récitation. Elle permet de choisir la récitation la plus appropriée parmi plusieurs réceptions fréquentes, en comparant notamment les variantes de prononciation ou de vocalisation.

### 2.7.4 L'arrêt optionnelle

L'arrêt optionnelle, se produit lorsque le lecteur décide de faire une pause de sa propre initiative, sans aucune contrainte externe, cela signifie que le lecteur a la liberté de choisir



quand et où faire une pause dans sa lecture, sans qu'il y ait des directives spécifiques ou de règle imposée.

Les catégories de la pause optionnelle :

- ❖ **L'arrêt complète** se produit lorsqu'on fait une pause après un discours qui est en lui-même complet, sans connexion sémantique ou syntaxique avec ce qui suit, c'est-à-dire, sans lien de sens ou de structure grammaticale avec le contenu qui vient ensuite.
- ❖ **L'arrêt adéquate** : il s'agit de faire une pause sur un discours complet en soi qui est lié à ce qui suit du point de vue du sans nécessairement être lié du point de vue de la formulation verbale.

**Exemple** [النمل: ٣٤] « إِنَّ الْمُلُوكَ إِذَا دَخَلُوا قَرْيَةً أَفْسَدُوهَا وَجَعَلُوا أَعْرَبةَ أَهْلِهَا أَذْلةً وَكَذَلِكَ يَفْعَلُونَ » قال تعالى

- ❖ **L'arrêt convenable** : il s'agit de faire une pause après un discours complet en discours complet en soi, cependant entre celui-ci et ce qui suit, il existe une connexion à la fois sémantique et verbale.

**Exemple** [البقرة: ٢٥٥] « اللَّهُ لَا إِلَهَ إِلَّا هُوَ الْحَيُّ الْقَيُّومُ لَا تَأْخُذُهُ سِنَّةٌ وَلَا نَوْمٌ » قوله تعالى

- ❖ **L'arrêt inappropriée** : il s'agit d'une pause sur un discours qui ne transmet pas un sens correct en raison de son étroite connexion avec ce qui suit, à la fois verbalement et sémantiquement. Cependant, la pause sur ce discours donne un sens incomplet, erroné ou invalide, qui n'est pas acceptable. [25]

**Exemple** : La pause sur "اهْدِنَا" dans le verset "اهْدِنَا الصِّرَاطَ الْمُسْتَقِيمَ" du Coran.

## 2.8 Le début

Commencer la lecture après une pause ou un arrêt

Les différents types :

### 2.8.1 Le véritable début

Il survient après une pause dans la lecture précédente, marquant ainsi la fin de celle-ci ou la transition vers d'autres activités autre que la lecture. Lorsque l'on revient à la lecture, le début à ce moment-là est considéré comme un véritable commencement. [24]

### 2.8.2 Le début additionnel

C'est ce qui intervient après une pause sur le dernier mot du Coran pendant un court laps de temps, pendant lequel on respire généralement avec l'intention de reprendre la lecture plutôt que de s'en détourner. Il existe également trois types de cela.



### 2.8.3 Le début absolu

C'est le commencement avec un discours complet en lui-même, sans connexion verbale ou sémantique avec ce qui précède. Ainsi, le début de chaque sourate du Coran est considéré comme un début absolu, de même que le début de chaque récit coranique et chaque segment qui n'a pas de lien verbal ou sémantique avec ce qui précède.

**Exemple :** « [البقرة: ٧، «خَتَمَ اللَّهُ عَلَى قُلُوبِهِمْ وَعَلَى سَمْعِهِمْ» ]

### 2.8.4 Le début approprié

C'est le commencement avec des paroles qui sont liées, à la fois sur le plan du sens et de la grammaire, à ce qui précède. Cette forme de commencement est valide uniquement au début des versets et nécessite d'être un commencement additionnel après une pause. Il n'est pas valide de le considérer comme un début véritable, même s'il s'agit de la tête d'un verset, après une pause. [22]

**Exemple :** قوله تعالى: ﴿إِلَى فِرْعَوْنَ وَمَلَأْنَاهُ. بعد قوله تعالى: ﴿ثُمَّ أَرْسَلْنَا مُوسَى وَأَخَاهُ هَارُونَ بِآيَاتِنَا وَسُلْطَانٍ مُّبِينٍ﴾: (المؤمنون: ٤٥-٤٦).

## 2.9 Dessin coranique الرسم القرآني

### 2.9.1 Définition du dessin coranique

Dessiner le Coran, ou comme certains érudits l'appellent le dessin d'Uthmani, c'est écrire quelques mots du Saint Coran, dans un dessin autre que le dessin orthographique connu.

Elle a été approuvée par le Prophète - que la paix et la bénédiction de Dieu soient sur lui - et a été acceptée par les nobles compagnons. Elle a été adoptée par Uthman ibn Affan - qu'Allah soit satisfait de lui - dans le mus'haf (coran) qu'il a fait écrire par le comité de Zaid ibn Thabit. Ce mus'haf a été nommé le mus'haf d'Uthman, et il a été transmis par la nation avec un consensus continu jusqu'à nos jours. [26]

### 2.9.2 Avantages du dessin coranique

#### ❖ Indiquer l'origine du mouvement

La lettre "ya" a été ajoutée dans la parole de Dieu Tout-Puissant : « وَلَقَدْ جَاءَكَ مِنْ نَبِيِّ الْمُرْسَلِينَ » : [Al-An'am : 34] pour indiquer le kasrah, et la lettre "wa" a été ajoutée dans la parole du Tout-





Puissant : « سَأُورِيكُمْ دَارَ الْفَسِقِينَ » [Al-A'raf : 1 45] pour indiquer le dhamma, car le Coran a été rédigé sans hamza, ponctuation et voyelles. [27]

### ❖ Indication de l'origine de la lettre

Écrire "صلاة" avec un "و" au lieu d'un "ا" dans : "الصَّلَاةُ" [Al-Baqarah: 3] indique que l'origine de la lettre est le "و" dans le mot "صلوة" et écrire "ا" et "ياء" indique qu'ils sont parmi les lettres "ياء" qui, de par leur nature, inclinent vers leur inclinaison, comme dans : "وَالضُّحَىٰ" [Ad-Duha: 1].

### ❖ Prise en compte des différentes lectures du dessin

Le mot "عباد" dans la parole de Dieu Tout-Puissant : "وَجَعَلُوا الْمَلَائِكَةَ الَّذِينَ هُمْ عِبُدُ الرَّحْمَنِ ابْنًا" [AZ-Zukhruf: 19] a été lu comme "عباد" et "عبد" et le dessin est possible pour les deux lectures.[27]

### ❖ Se référant à une autre lecture

Le mot "بينة" apparaît dans dix-neuf endroits. Dans tous ces cas, il est écrit avec un "ت" attaché, sauf dans un seul endroit où il est écrit avec un "ت" ouvert, dans le verset : "أَمْ آتَيْنَاهُمْ" [Fatir: 40]. Cela indique une autre lecture. Certains récitants lisent ce verset au pluriel et d'autres au singulier, donc il est écrit avec un "ت" ouvert pour permettre les deux lectures.

### ❖ Indiquer un sens caché précis

- L'ajout de la lettre "ا" après le verbe dans le mot "وَيَعْفُوا" dans la parole divine : "وَيَعْفُوا عَنْ" [Fatir: 34] indique l'abondance et la continuité du pardon de Dieu. L'augmentation de la construction renforce le sens.

- L'écriture de "أيد" avec deux "ياء" dans la parole divine: "وَالسَّمَاءَ بَنَيْنَاهَا بِأَيْدٍ وَإِنَّا لَمُوسِعُونَ" [Adh-Dhariyat: 47] est pour souligner la grandeur de Dieu.

- L'ajout d'un "همزة" sur la "واو" et l'ajout de la lettre "ا" après cela dans le mot "تَقْتُلُوا" dans la parole divine: "فَالأَوْ تَاللَّهِ تَقْتُلُوا تَذَكَّرُ يُوسُفُ" [Yusuf: 85] indique la fréquence de la mention de Joseph par Jacob, que la paix soit sur eux.

- L'ajout de la lettre "ا" dans le mot "وَجَاءَ" dans la parole divine : "وَجَاءَ يَوْمَئِذٍ بِجَهَنَّمَ" [Al-Fajr: 23] est pour l'accentuation, l'exagération et la menace.

- La suppression de la lettre "ا" après la "واو" du pluriel dans le mot "وَجَاءُوا" dans la parole divine : "وَجَاءُوا بِسِحْرِ عَظِيمٍ" [Al-A'raf: 116] indique que leur arrivée est dominée par le mensonge et la tromperie. [27]

### ❖ Fournir des significations différentes d'une manière qui n'est pas cachée

La séparation de "أم" de "من" dans la parole divine : "أَمْ مَنْ يَكُونُ عَلَيْهِمْ وَكَيْلًا" [An-Nisa: 109] indique que "أم" est détaché, ce qui signifie "plutôt", tandis que "من" est lié.



La liaison de "أم" à "من" dans la parole divine : "أَمَّنْ يَمْشِي سَوِيًّا عَلَى صِرَاطٍ مُسْتَقِيمٍ" [Al-Mulk: 22] indique que "أم" est relié à "من". [28]

### ❖ Encouragez les gens à recevoir le Saint Coran de sources fiables

-Le dessin uthmanique différait du dessin standard par certains mots. Par exemple, au début de 29 sourates du Saint Coran, il y a des lettres disjointes composées de 14 lettres regroupées en : "طرق سمعك النصيحة" ; Trois d'entre elles commençaient par une lettre : "ن، ق، ص", neuf commençaient par deux lettres : "طس، حم، يس، طه", treize commençaient par trois lettres telles que : "الم، الر، طسم", et deux commençaient par quatre lettres : "المص، المر", et deux commençaient par cinq lettres : "كهيعص، حم عسق".

-Ces lettres sont lues par leurs noms, et non par leur prononciation. Cinq de ces lettres regroupées dans : "حي طهر" se prononcent comme deux lettres comme ceci : "ها، ها، را", et les huit lettres restantes regroupées dans : "نقص عسلكم" se prononcent comme trois lettres comme ceci : "نون، قاف، صاد، عين، سين، لام، كاف، ميم". Si la question avait été laissée au format standard, les gens ne se seraient pas précipités pour recevoir le Saint Coran de sources fiables, et "آلم" [Al-Baqarah : 1] aurait été récité "آلف لام ميم". [28]

## 2.10 Conclusion

Ce chapitre a souligné l'importance de la récitation correcte du Coran, guidée par les règles de Tajwid, pour préserver l'intégrité et la beauté du texte sacré. En expliquant la définition du Coran et du Tajwid, ainsi que les différents niveaux de récitation et types d'erreurs, nous avons démontré combien il est crucial de maîtriser ces aspects pour une récitation correcte. Le rasm al-mushaf, ou dessin coranique, joue également un rôle vital en facilitant une lecture précise et respectueuse. Ces pratiques sont essentielles pour maintenir la tradition et la spiritualité islamiques.



## Chapitre 3: L'apport du Deep Learning dans la reconnaissance vocale de la récitation coranique

### 3.1 Introduction

Le domaine de la reconnaissance vocale de la récitation coranique a connu une évolution fulgurante ces dernières années grâce à l'avènement du Deep Learning. Les nouvelles approches basées sur les réseaux de neurones profonds ont surpassé les méthodes traditionnelles, permettant une analyse précise des variations subtiles de la récitation, l'identification des différents styles et la reconnaissance des voix individuelles. Dans le cadre de la lecture et de l'analyse du Coran, où la précision est primordiale, deux principales méthodes de validation de la récitation sont utilisées : la comparaison vocale et la comparaison textuelle

### 3.2 Les différentes approches de Deep Learning pour la reconnaissance vocale

#### 3.2.1 Réseaux de Neurones Récurents (RNN)

Les RNN sont utilisés pour modéliser des séquences temporelles comme le signal audio. Ils prennent en entrée la séquence d'échantillons audio et produisent une sortie qui peut être la transcription du texte prononcé. [29]

Les RNN présentent plusieurs avantages pour la reconnaissance vocale

- **Modélisation séquentielle** : Les RNN sont spécialement conçus pour traiter des données séquentielles, ce qui en fait un choix naturel pour la reconnaissance vocale où les signaux audio sont des séquences temporelles.
- **Meilleure précision** : Les RNN peuvent prendre en compte le contexte temporel de la parole, ce qui leur permet d'obtenir une meilleure précision que les approches traditionnelles.
- **Robustesse au bruit** : Les RNN peuvent être entraînés à être robustes au bruit et aux variations de l'environnement.
- **Adaptabilité aux variations temporelles** : Les RNN peuvent s'adapter à des séquences de longueurs variables, ce qui est essentiel pour la reconnaissance vocale où les mots peuvent être prononcés à des vitesses différentes et avec différentes durées.

Cependant, les RNN ont des inconvénients :



- **Temps de calcul élevé** : Le traitement séquentiel des RNN peut être lent, en particulier pour des tâches de reconnaissance en temps réel.
- **Complexité** : Les RNN sont plus complexes à entraîner et à déployer que les approches traditionnelles.
- **Oubli à long terme** : Bien que les RNN soient capables de modéliser les dépendances à long terme, ils peuvent avoir du mal à retenir des informations importantes sur de longues séquences, ce qui peut entraîner une perte d'information contextuelle.

### 3.2.2 Réseaux de Neurones Convolutifs (CNN)

Bien que principalement utilisés pour le traitement d'images, les CNN peuvent également être employés pour extraire des caractéristiques pertinentes à partir de représentations audio, telles que les spectrogrammes qui sont ensuite utilisées pour la reconnaissance vocale.

Les avantages qu'offrent les CNN dans ce domaine sont nombreux :

- **Meilleure précision** : Les CNN peuvent extraire des caractéristiques plus discriminantes du signal vocal que les approches traditionnelles, ce qui est particulièrement utile dans la reconnaissance vocale où les caractéristiques pertinentes peuvent être complexes et difficiles à définir manuellement. [29]
- **Traitement parallèle efficace** : Contrairement aux RNN, les CNN permettent un traitement hautement parallélisable, accélérant considérablement l'entraînement et l'inférence.
- **Efficacité dans des applications réelles** : Les CNN sont utilisés par des entreprises telles que Google pour leur système de reconnaissance vocale, comme "Hey Google!", démontrant ainsi leur efficacité dans des applications réelles.

Cependant, malgré leur potentiel et leur efficacité, l'application des réseaux de neurones convolutifs n'est pas sans défis :

- **Prétraitement des données** : Les CNN nécessitent un prétraitement des données plus important que les RNN.
- **Besoin en grandes quantités de données** : L'entraînement des CNN nécessite de grandes quantités de données annotées, ce qui peut être coûteux



### 3.2.3 Modèles de langage neuronal (Les transformers)

Les modèles de langage neuronal, comme les Transformers, permettent de modéliser les relations statistiques entre les mots d'une langue. En intégrant ces modèles dans la reconnaissance vocale, on peut améliorer la compréhension du contexte et la prise en compte des variations grammaticales et syntaxiques. [29]

#### Les avantages des Transformers pour la reconnaissance vocale sont multiples

- **Meilleure gestion du contexte:** Les Transformers capturent efficacement les dépendances à longue distance dans les séquences grâce à leur mécanisme d'attention, ce qui leur permet de mieux comprendre le contexte de la parole.
- **Parallélisation efficace:** les Transformers permettent un calcul massivement parallèle ce qui peut conduire à des temps de traitement plus rapides.
- **Reconnaissance multilingue:** Un seul modèle Transformer peut être entraîné sur plusieurs langues simultanément. [30]

#### Inconvénients des Transformers

- **Besoins en données élevés:** Les Transformers nécessitent généralement d'énormes quantités de données pour être entraînés efficacement.
- **Complexité:** L'architecture des Transformers est plus complexe que celle des RNN et des CNN, ce qui peut les rendre plus difficiles à mettre en œuvre et à déployer.
- **Sensibilité aux variations de bruit :** Les transformers peuvent être sensibles aux variations de bruit dans les signaux audio, ce qui peut entraîner une dégradation de la performance dans des environnements acoustiques bruyants. [30]

### 3.3 La reconnaissance vocale de la récitation coranique

La reconnaissance vocale de récitation coranique a fait l'objet de plusieurs études en particulier avec l'avènement de l'intelligence artificielle. Elle implique la capacité de reconnaître et d'évaluer la récitation du Coran en fonction des règles de tajwid et des différentes lectures autorisées.

Un aperçu complet est présenté dans [31]. Les auteurs ont fait une étude récente concernant la reconnaissance vocale du Saint Coran (HQSR), soulignant les difficultés et les progrès réalisés dans ce domaine. Diverses approches de traitement du signal sont examinées dans les études, telles que l'utilisation de MFCC et de réseaux de neurones profonds pour extraire et reconnaître les versets du Coran. Les réseaux LSTM, HMM et SVM sont des modèles d'apprentissage couramment employés afin d'améliorer la précision de la



reconnaissance. Les défis qui demeurent sont l'exigence d'un ensemble de données complet et la prise en considération des particularités de la récitation.

Dans leur contribution [32], les auteurs ont proposé un modèle de reconnaissance de la récitation du Saint Coran basé sur un modèle d'apprentissage de Machine à Vecteurs de Support (SVM). La recherche s'est concentrée sur la reconnaissance des différents styles de récitation autorisés ("Qira'ah" en arabe) et aider les étudiants du Saint Coran à apprendre la récitation correcte basée sur les règles de l'"Ahkam Al-Tajweed". Le modèle utilise l'extraction des coefficients fréquence Mel (MFCC) et l'apprentissage par SVM et atteint une précision de reconnaissance de 96%.

Un peu proche de ce contexte, [33] propose de reconnaître le récitant du coran. Les chercheurs ont employé deux approches de traitement de l'audio reposant sur les coefficients cestraux de fréquence Mel (MFCC) et la hauteur tonale afin de détecter les versets coraniques, et dans la deuxième approche, l'audio est traité comme des images via spectrogram.

Pour la classification, ils ont utilisé des modèles d'apprentissage automatique tels que Naïve Bayes, J48 et Random Forest. Les travaux réalisés ont atteint une précision de reconnaissance de 88% pour les caractéristiques audio et de 81% pour les caractéristiques basées sur des images de spectrogramme.

Dans une autre optique, les travaux de [34] visent à aider les étudiants à corriger leur prononciation (hijaiyah) des versets coraniques. Les auteurs ont choisi d'utiliser MFCC pour l'extraction des caractéristiques, et les Modèles Markov Caché (HMM), les Réseaux de Neurones Artificiels (ANN) pour les modèles d'apprentissage employés.

### **3.4présentation des solutions envisagées**

#### **3.4.1 Objectifs du système**

L'objectif est de concevoir un système efficace pour la validation des récitations du Coran. Qu'il se base sur l'analyse directe des enregistrements vocaux, l'approche vise à garantir l'exactitude et l'authenticité de la récitation coranique. En validant et en confirmant la justesse des récitations, le but est de permettre aux récitants de renforcer leur confiance et d'améliorer leur récitation conformément aux normes établies de précision et de dévotion.



### 3.5 Solution 1 : Convertir la récitation audio en texte

#### Description de la solution

- La récitation du Coran est convertie en texte à l'aide de techniques de reconnaissance vocale.
- Ce texte est ensuite comparé à une version correcte du texte coranique.

#### Avantage

- Facilité de traitement : le texte est généralement plus facile à traiter que le signal audio, ce qui peut rendre cette solution plus efficace en termes de temps de traitement et de ressources.

#### Inconvénients

- Cette solution ne permet pas de corriger les erreurs liées aux règles de Tajwid,
- Le texte reproduit peut comporter des erreurs ce qui contrevient à la sanctité et l'intégrité du saint coran.

### 3.6 Solution 2 : Comparer la récitation avec une référence audio

#### Description de la solution

- Ces données vocales sont comparées à une référence de récitation correcte en utilisant des techniques de traitement du signal audio.
- Les caractéristiques audio telles que la fréquence, l'intensité, le rythme, etc., peuvent être analysées pour détecter les écarts par rapport à la récitation correcte.

#### Avantage

- Précision vocale : Ce scénario peut être plus efficace pour valider et détecter les erreurs liées à la prononciation ou à l'intonation.
- Amélioration de la qualité de la récitation en identifiant les zones où des corrections sont nécessaires, ce qui peut aider les récitants à perfectionner leur technique.
- Comparer avec une référence audio authentique garantit que la récitation suit les règles de Tajwid
- Cette méthode de validation est conforme à ce qui a été rapporté par le prophète, qui enseignait verbalement le Coran à ses compagnons.

#### Inconvénients

- Des erreurs techniques, telles que des problèmes de synchronisation ou de qualité audio, peuvent entraîner des résultats incorrects lors de la comparaison.



- En lecture coranique les styles de réceptions sont très variés, ce qui rend difficile de définir une référence unique.
- Difficulté d'obtenir un Feedback personnalisé.

### 3.7 Choix de la solution

Le choix du deuxième scénario, qui consiste à transformer la source en vocal et à comparer vocal-vocal pour validation de la récitation du Coran, est judicieux pour plusieurs raisons.

Ce scénario permet de préserver l'authenticité de la récitation en analysant directement les enregistrements vocaux, ce qui est essentiel pour maintenir l'intégrité du texte sacré. En plus, c'était la méthode avec laquelle le prophète a message reçu le Coran de Gabriel et l'enseigna à ses compagnons

En se basant sur l'analyse des caractéristiques audio telles que la fréquence, l'intensité et le rythme, ce scénario est capable de validation des récitation subtiles qui pourraient échapper à d'autres méthodes de comparaison.

Ce scénario offre une grande sensibilité aux variations régionales et individuelles dans la récitation du Coran, ce qui le rend adapté à une grande diversité de pratiquants de la religion islamique.

### 1.7 Conclusion

Ce chapitre nous a permis de découvrir en détail les diverses méthodes de Deep Learning employées dans le domaine de la reconnaissance vocale pour la récitation coranique. En analysant deux scénarios de conception d'un système de validation de la récitation du Coran, nous avons examiné les caractéristiques et avantages de chaque approche. Quel que soit le choix, l'objectif demeure de concevoir un système efficace pour soutenir et améliorer la pratique de la récitation du Coran, contribuant ainsi à préserver l'authenticité et la dévotion dans la religion islamique.





## Chapitre 4 : La présentation de l'architecture du système et les résultats

### 4.1 Introduction

Le projet vise à développer un système d'assistance basé sur l'intelligence artificielle pour faciliter la mémorisation du Coran, une tâche exigeante pour de nombreux musulmans. Utilisant des réseaux de neurones convolutifs et des données audio du Coran, ce système valide les récitations des étudiants pour les aider dans leur apprentissage. Ce chapitre détaille la conception, la mise en œuvre et l'évaluation des performances du système. Nous examinerons les différentes étapes de nos tests, les métriques d'évaluation, les résultats obtenus, ainsi que les défis rencontrés et les leçons apprises, afin de tirer des conclusions sur l'impact et le potentiel de cette approche dans le domaine de la mémorisation du Coran.

### 4.2 La description modulaire du système

Le système de reconnaissance vocale des versets coraniques basé sur un modèle CNN est composé de plusieurs modules. Le module de prétraitement transforme les données audio brutes en spectrogrammes. Le modèle CNN traite ces spectrogrammes pour classifier les versets en utilisant plusieurs blocs convolutifs, des couches entièrement connectées et une couche de sortie Softmax. Le module de post-traitement convertit les probabilités en étiquettes de versets et applique des techniques pour améliorer la précision. L'interface utilisateur permet de sélectionner une sourate, d'enregistrer un verset et d'afficher les résultats de la reconnaissance vocale.

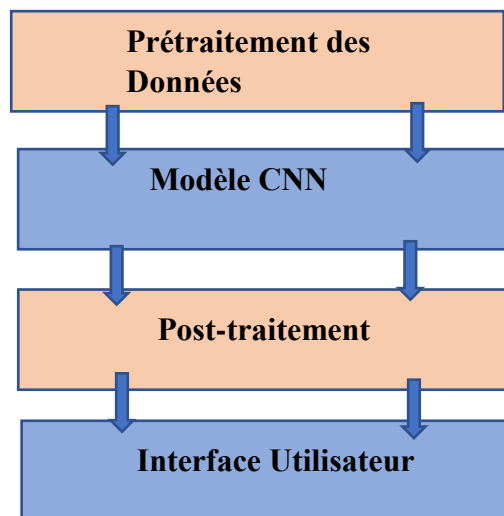


Figure 4 : La description modulaire du système



### 4.3 Collecte des données

Pour constituer la base de données audio utilisée dans cette étude, les enregistrements ont été obtenus à partir du jeu de données "Quran Reciters" disponible sur Kaggle (<https://www.kaggle.com/datasets/omartariq612/quran-reciters>). Cette ressource contient les récitations complètes du Coran par 8 récitants différents. Après avoir téléchargé ce jeu de données, les fichiers audio ont été importés sur Google Drive afin de faciliter leur utilisation avec Google Colab. Afin de simplifier le traitement, les noms des 8 récitants ont été remplacés par des appellations génériques "Récitation 1", "Récitation 2", etc. De plus, bien que le jeu de données initial couvre les 114 sourates du Coran, seules 11 sourates ont été conservées pour cette étude spécifique.

Cette sélection réduite de données facilite la gestion des fichiers audio tout en fournissant un échantillon représentatif pour l'analyse et l'entraînement des modèles d'intelligence artificielle.

Les fichiers audio (.mp3) ne contiennent pas des récitations complètes de sourates, mais sont plutôt segmentés de manière à représenter chaque verset (āyah) individuellement.

Les enregistrements audio contiennent diverses traditions de récitation du Coran, connues sous le nom de "narrations". Chaque narration représente une manière spécifique de réciter le texte coranique, transmise au fil des générations selon des chaînes de transmission orales distinctes. Parmi les différentes narrations présentes dans ce jeu de données, on trouve notamment celle de Warsh, qui est l'une des narrations faisant autorité dans la récitation du Coran.

### 4.4 Prétraitement des données audio

#### 4.4.1 Connecter le Google Drive à Google Colab

La première étape consiste à établir une connexion entre Google Drive et Google Colab. Cela permettra d'accéder directement aux fichiers audio et autres données requises pour effectuer le prétraitement des données.



#### **4.4.2 Conversion des fichiers audio en spectrogrammes**

Pour transformer nos fichiers audio au format MP3 en données exploitables par un modèle d'apprentissage profond, nous les convertissons en images de spectrogrammes. Un spectrogramme représente visuellement la distribution de l'énergie d'un signal audio dans les domaines temps et fréquence. Nous parcourons l'ensemble des fichiers MP3 présents dans le répertoire "audio". Pour chaque fichier, nous chargeons d'abord le signal audio à l'aide d'une bibliothèque de traitement du signal (librosa). Ensuite, nous calculons le spectrogramme correspondant en appliquant une transformée mathématique sur le signal. Nous créons alors une figure affichant ce spectrogramme sous forme d'image. Cette image est sauvegardée au format PNG dans le répertoire "images". Enfin, nous redimensionnons l'image à une taille fixe de 224x224 pixels avec la bibliothèque (resize) pour uniformiser les données d'entrée. À l'issue de ce processus, nous disposons d'un ensemble d'images de spectrogrammes converties à partir des fichiers audio, prêtes à être utilisées pour l'entraînement du modèle.

#### **4.4.3 Division des données en ensembles d'entraînement, de validation et de test**

La division des données en ensembles d'entraînement, de validation et de test est une étape cruciale dans la préparation des données pour l'entraînement d'un modèle d'apprentissage profond. Après avoir récupéré tous les chemins des images de spectrogrammes et les avoir mélangés aléatoirement, nous divisons ces chemins en trois ensembles distincts : l'ensemble d'entraînement (60% des données), l'ensemble de validation (20% des données) et l'ensemble de test (20% des données). Cette répartition garantit que le modèle sera entraîné sur un large échantillon de données, tout en conservant des ensembles de validation et de test représentatifs pour évaluer ses performances de manière fiable.

#### **4.4.4 Création d'un fichier de métadonnées**

Cette étape consiste à créer un fichier CSV contenant les métadonnées associées à chaque fichier audio et son image de spectrogramme correspondante. Nous parcourons d'abord tous les fichiers audio MP3 présents dans le répertoire "audio\_uniques". Pour chaque fichier audio, nous récupérons son chemin complet ainsi que le chemin de l'image de spectrogramme correspondante présente dans le répertoire "images\_uniques". Nous extrayons également le numéro de verset à partir du nom du fichier audio.



Il est nécessaire de créer un fichier de métadonnées contenant les associations entre les fichiers audio, les images de spectrogrammes et les numéros de versets pour plusieurs raisons importantes tel que l'Organisation des données, Récupération des labels, Chargement des données.

#### 4.4.5 Conversion des images en tenseurs NumPy

Nous chargeons les métadonnées du fichier CSV, ouvrons les images de spectrogrammes, les convertissons en tenseurs NumPy et normalisons les valeurs de pixels entre 0 et 1. Cette normalisation facilite l'entraînement du modèle. Finalement, nous constituons deux listes distinctes : l'une contenant l'ensemble des tenseurs NumPy normalisés représentant les images, et l'autre contenant les labels de versets correspondants. Ces deux listes, représentant respectivement les données d'entrée et les labels, sont désormais prêtes à être utilisées pour l'entraînement du modèle d'apprentissage profond.

#### 4.5 Définition de l'architecture du réseau neuronal

Il existe une grande variété d'architectures de réseaux de neurones profonds, la plupart étant des variantes dérivées de quelques architectures de base originales. Parmi cette diversité, notre choix s'est porté sur les réseaux de neurones convolutifs (CNN). Ces derniers sont particulièrement adaptés pour le traitement de données telles que des images ou des signaux audio, grâce à leur capacité à extraire automatiquement des caractéristiques pertinentes à différents niveaux de représentation.

##### 4.5.1 Principe d'architecture d'un CNN

Les réseaux de neurones convolutifs, ou CNN (Convolutional Neural Networks), représentent actuellement les modèles les plus performants pour la tâche de classification d'images. Leur architecture se compose de deux parties distinctes. En entrée, une image est fournie sous forme d'une matrice de pixels, ayant deux dimensions pour les images en niveaux de gris, et trois dimensions (rouge, vert, bleu) pour les images couleur.

L'architecture principal d'un réseau de neurone à convolution se décompose souvent de 5 types de couche : une couche des entrées (input Layer), des couches à convolution (Convolutionnal Layer) chacune suivie par une couche de polarisation (Pooling Layer), une couches FC (Fully Connected Layer) et enfin une couche des sorties (output Layer). **Figure 5**

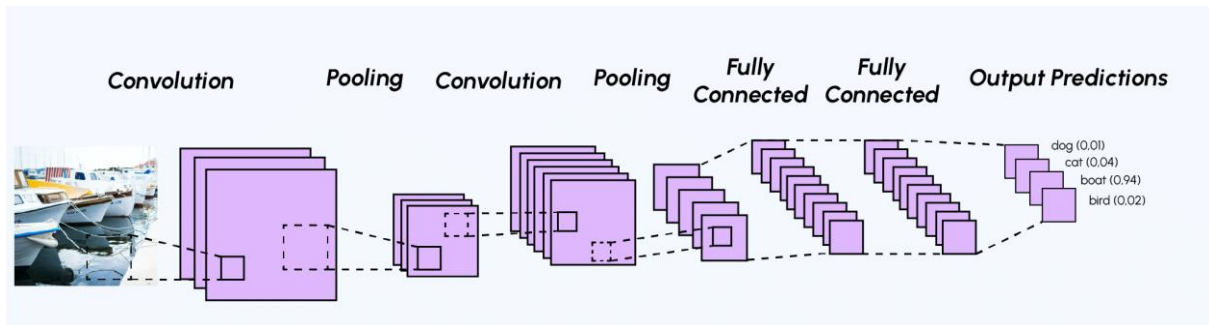


Figure 5 : Exemple d'architecture d'un CNN à 2 couches [35]

#### 4.5.2 Couches de réseaux neuronaux convolutifs

Dans les CNN, il existe plusieurs couches, notamment :

##### a) Couche d'entrée

Dans les CNN, l'entrée comprend des données décrivant une image qui est initialement Représentée sous la forme d'une matrice tridimensionnelle. Par la suite, ces données sont remodelées en une seule colonne formant une représentation vectorielle.

##### b) Couche convolutionnelle (CONV)

Les couches CONV, également appelées ainsi d'après leur opération de convolution, sont des composants essentiels des CNN. Elles sont responsables de l'extraction automatique des caractéristiques pertinentes des données telles que les images et les signaux audio. En utilisant des noyaux ou des filtres apprenables, elles appliquent une opération de convolution à l'entrée, permettant ainsi de détecter des motifs spécifiques tels que les bords et les textures. Ces motifs sont capturés dans des cartes de caractéristiques, et les couches CONV successives les combinent pour détecter des caractéristiques de plus en plus complexes. Au cours de l'entraînement, les poids des noyaux sont ajustés pour optimiser l'extraction de caractéristiques discriminantes pour la tâche visée, ce qui confère aux CNN leur capacité d'apprentissage hiérarchique des caractéristiques. Ce processus de convolution permet d'extraire des modèles locaux, renforçant ainsi la capacité du réseau à discerner les caractéristiques critiques au sein des données, comme le montre la **Figure 6**. [36]

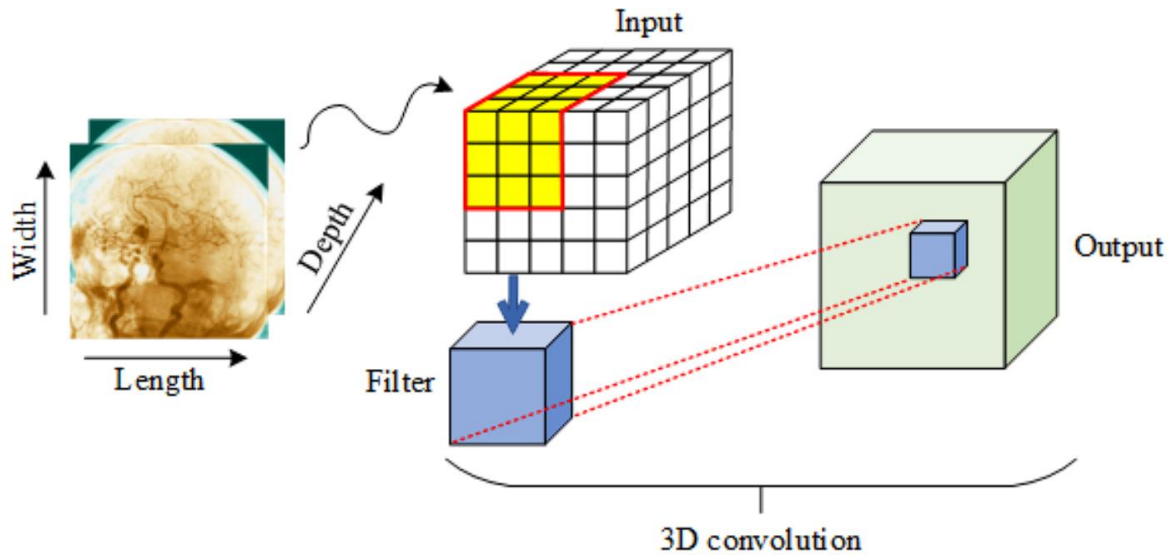


Figure 6 : Convolution 3D [36]

### c) Couche de pooling (POOL)

Ce type de couche est souvent placé entre deux couches de convolution : il prend en compte de nombreuses cartes de caractéristiques et applique l'opération de pooling sur chacune d'elles. L'opération de pooling réduit la taille des photos tout en préservant leurs caractéristiques cruciales.

Pour ce faire, l'image est divisée en cellules régulières, et la valeur maximale est conservée à l'intérieur de chaque cellule. En pratique, on utilise souvent de petites cellules carrées pour éviter de perdre trop d'informations. Les choix les plus courants sont des cellules voisines de taille  $2 \times 2$  pixels qui ne se chevauchent pas, ou des cellules de taille  $3 \times 3$  pixels qui ne sont séparées que de 2 pixels (et qui se chevauchent donc).

Nous obtenons le même nombre de cartes de caractéristiques qu'auparavant, mais elles sont beaucoup plus petites. La couche de regroupement réduit le nombre de paramètres et de calculs dans le réseau. Cela améliore l'efficacité du réseau et évite le sur-apprentissage.[37]

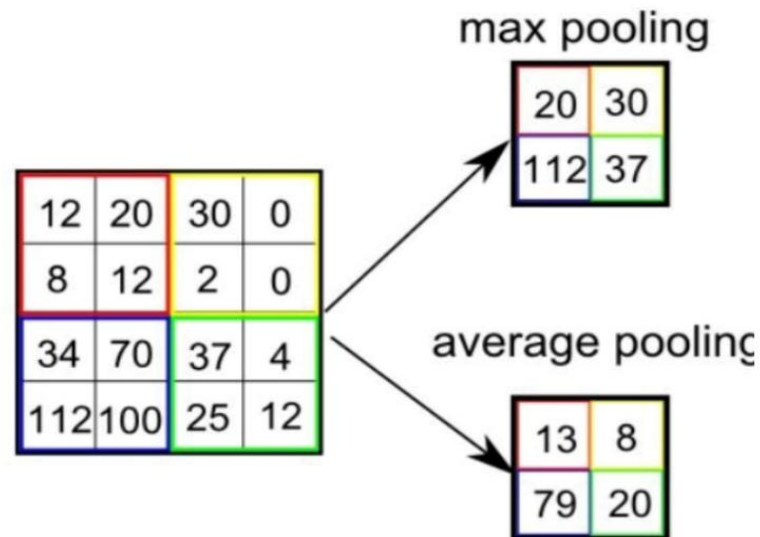


Figure 7 : exemple d'application de Pooling [38]

#### d) Couche entièrement connectée (Fully Connected)

Cette couche englobe les poids, les biais et les neurones facilitant les connexions entre les neurones d'une couche à ceux d'une autre couche. Elle joue un rôle crucial dans les tâches de classification d'images en combinant les caractéristiques extraites des couches précédentes pour classer les images en différentes catégories.

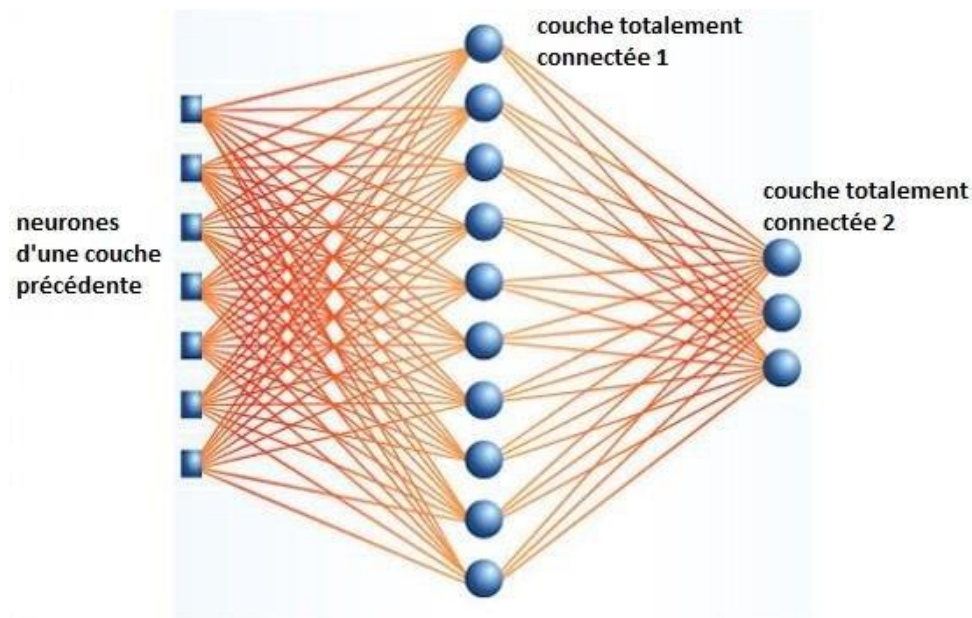


Figure 8: une Couche entièrement connectée [39]



### e) Fonction d'activation

Dans un réseau neuronal, chaque neurone utilise une fonction d'activation pour traiter les données entrantes. De plus, une valeur seuil est attribuée à chaque neurone, représentant l'entrée cumulative minimale requise pour activer le neurone. Si l'entrée cumulative tombe en dessous de ce seuil, le neurone reste inactif (sortie 0), tandis que le dépassement du seuil entraîne l'activation du neurone (sortie 1)

Diverses fonctions d'activation aux caractéristiques distinctes sont utilisées dans les réseaux neuronaux pour atteindre des objectifs spécifiques et les qualités de sortie souhaitées. Voici quelques fonctions d'activation couramment utilisées :

Step Function	Sigmoid Function	ReLU Function	Linear function
<p>Cette fonction identifie la sortie du neurone entre 0 et 1</p>	<p>Il s'agit d'une fonction non linéaire et sa sortie est limitée entre 0 et 1 et est l'une des fonctions les plus importantes utilisées dans les réseaux de neurones artificiels pour un calcul facile.[40]</p>	<p>est une fonction d'activation très populaire, surtout dans les couches convolutionnelles et fully-connected. Elle applique l'opération <math>\max(0, x)</math> aux entrées, remplaçant les valeurs négatives par zéro.[41]</p>	<p>La fonction dans laquelle l'image de sortie est similaire à l'image d'entrée et fournit des notes multiples et illimitées.[42]</p>

**Tableau 2 :** Fonction d'activation utilisée dans les réseaux neuronaux synthétiques





## f) Couche SoftMax/Logistique

Positionnée à l'extrémité de la couche entièrement connectée, cette couche est essentielle dans les tâches de classification. La fonction logistique est utilisée pour la classification binaire tandis que **SoftMax** est utilisée pour la classification multi-classes en attribuant des probabilités à chaque sortie de classe.[43]

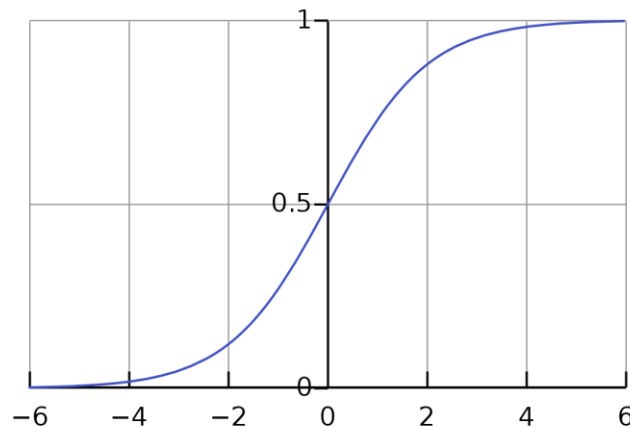


Figure 9 : La fonction Softmax [43]

## g) Couche Dropout

La régularisation est un ensemble de techniques utilisées lors de l'entraînement des réseaux de neurones, et en particulier des réseaux de neurones convolutifs (CNN), afin de réduire les risques de sur-apprentissage (overfitting).

En plus d'obtenir plus de données, certaines techniques sont utilisées pour lutter contre le sur-apprentissage, voici les plus courantes : Dropout, Data Augmentation, L1/L2 Régularisation...

**La Couche Dropout** Consiste à désactiver aléatoirement un certain pourcentage de neurones durant l'entraînement. Cela empêche les neurones de trop se spécialiser et les force à extraire des caractéristiques robustes.

La **Figure 10** présente une étude comparant les performances de deux modèles entraînés sur la même tâche. Le premier modèle a été entraîné sans aucune technique de régularisation, tandis que le second modèle a intégré la méthode du dropout lors de son entraînement. [44]

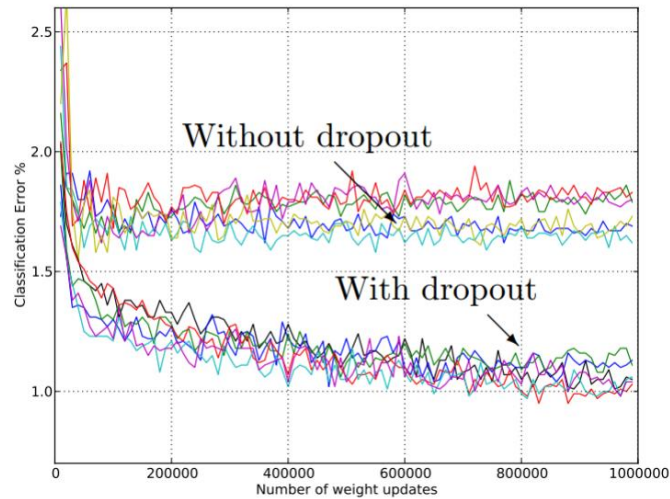


Figure 10 : Comparaison de taux d'erreur de classification avec et sans dropout [44]

### h) Couche de perte (LOSS)

La couche de perte (loss layer) dans un réseau de neurones convolutif sert à mesurer l'erreur commise par le modèle sur les prédictions par rapport aux véritables étiquettes des données. Cette erreur, appelée fonction de perte (loss function), est utilisée pour guider l'entraînement du modèle.

### i) Couche de sortie

La couche de sortie comprend les étiquettes généralement représentées sous forme codé.

## 4.6Modèle Implementation

Pour ce modèle CNN, nous avons proposé l'architecture illustrée dans la **Figure 11**. Après la récupération des entrées sous forme d'images de spectrogrammes, ces dernières vont être entraînées sur un réseau de neurones convolutionnels.

Chaque couche de convolution est suivie d'une couche de sous-échantillonnage (MaxPooling2D), qui permet de réduire les dimensions des cartes de caractéristiques tout en conservant les informations les plus importantes, ce qui permet de réduire le nombre de paramètres et le coût computationnel. Après ces trois couches de convolution et MaxPooling2D.

Nous aplatissons ensuite les cartes de caractéristiques avec une couche de flattening. Ce vecteur aplati est passé à une première couche entièrement connectée avec activation ReLU, suivie d'une couche Dropout pour minimiser le sur-apprentissage. Enfin, nous avons une



deuxième couche entièrement connectée utilisant la fonction d'activation Softmax pour attribuer une probabilité à chaque classe et prédire la classe avec la probabilité la plus élevée comme sortie du modèle.

#### 4.6.1 Architecture du modèle

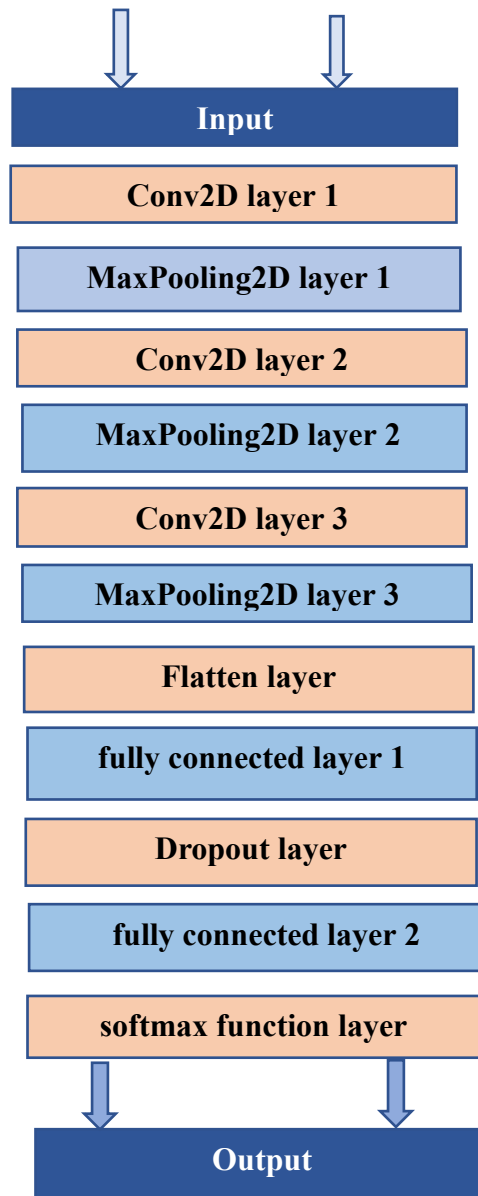


Figure 11 : Architecture du modèle



## 4.7 Les outils et environnement de travail

Dans cette partie nous allons présenter les outils ainsi que les bibliothèques utilisées afin de réaliser notre système.

### 4.7.1 Le Langage de Programmation Python

Python est un langage de programmation puissant, simple d'utilisation et polyvalent. Il se distingue par sa syntaxe claire, son typage dynamique et son mode d'exécution interprétée qui facilitent le développement rapide. Python propose des structures de données de haut niveau et une approche pragmatique de la programmation orientée objet. Malgré sa simplicité, il offre d'excellentes performances. Python trouve des applications dans de nombreux domaines comme le traitement de données, le web ou l'IA, et sur la plupart des plateformes grâce à sa portabilité. Son écriture rationnelle, sa flexibilité et son vaste écosystème de bibliothèques expliquent son succès croissant auprès des développeurs. [45]

### 4.7.2 Environnement de l'exécution

**Google Colaboratory** : est un service Cloud basé sur Jupyter Notebooks, permet de développer des applications en Deep Learning en Python, il offre un processeur GPU gratuit, 12 Go de RAM et plus de 100 Go de stockage. Pour l'accès dans ce service il nous suffit simplement d'avoir qu'un compte Google. [46]

### 4.7.3 Les bibliothèques utilisées

Pour nos projets de deep learning, nous avons principalement eu recours aux bibliothèques suivantes :

- **Librosa**: Est une bibliothèque utilisée pour analyser le comportement de l'audio. Il aide à charger des fichiers audio, à extraire les caractéristiques de la musique et à visualiser les données audio. Avec l'aide de librosa, travailler avec l'audio en Python est devenu simple.
- **TensorFlow**: Développée par Google, c'est l'une des bibliothèques que nous avons le plus utilisées. Elle nous a offert un large éventail d'outils pour construire, entraîner et déployer des réseaux de neurones, ainsi que des APIs dédiées à des tâches spécifiques.



- **Keras:** Nous avons souvent utilisé Keras comme interface conviviale par-dessus TensorFlow ou d'autres backends. Sa simplicité pour construire des réseaux de neurones à partir de blocs de haut niveau nous a été très utile.
- **Scikit-Learn :** Bien que plus généraliste, ses outils pour le prétraitement et les tâches annexes au deep learning nous ont permis de compléter notre workflow.

Nous allons besoin de plusieurs bibliothèques dans notre projet (**Figure 12**)

```
import os
import librosa
import librosa.display
import numpy as np
import pandas as pd
from PIL import Image
import matplotlib.pyplot as plt
import tensorflow as tf
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Conv2D, MaxPooling2D, Flatten, Dense, Dropout
from sklearn.preprocessing import LabelEncoder
from tensorflow.keras.utils import to_categorical
from tensorflow.keras.optimizers import Adam
from PIL import Image
import random
from sklearn.model_selection import train_test_split
```

**Figure 12 :** les différentes bibliothèques utilisées dans notre implémentation

## 4.8 Présentation des méthodes

### 4.8.1 L'évolution du signal

Afin de représenter les informations d'un signal numérique, on a choisi la méthode du spectrogramme. Le spectrogramme utilisé est normalisé en 3D qui nous affiche les informations Temps-Fréquence. Voir Figure 13

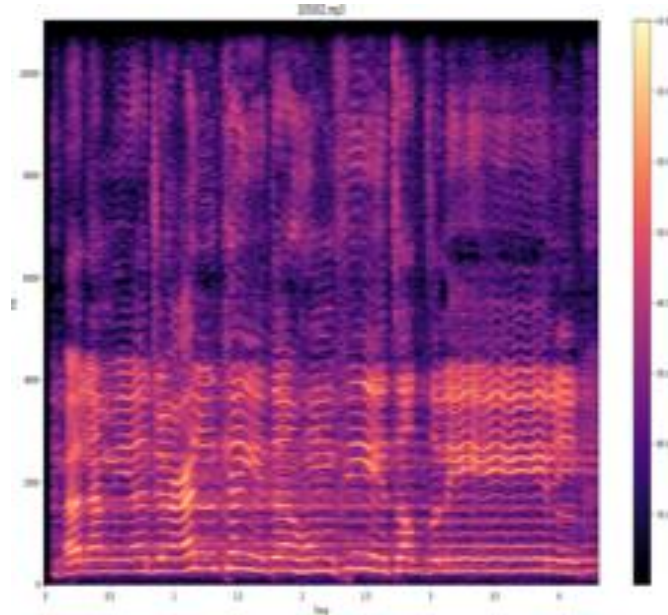


Figure 13 : Exemple d'un spectrogramme normalisé

#### 4.8.2 Les métriques d'évaluations

Dans cette section, nous définirons les principaux éléments que nous utiliserons pour évaluer les modèles de classification.

##### a) Matrice de confusion

La matrice contient plusieurs confusion chaque cellule de la matrice de confusion représente un facteur d'évaluation.

**True Positive (TP):** est un résultat où le modèle prédit qu'une observation fait partie d'une classe et que l'observation fait effectivement partie de cette classe.

**True Negative (TN) :** un résultat où le modèle prétend que l'observation ne fait pas partie d'une classe et qu'elle n'est en réalité pas membre de cette classe.

**False Positive (FP) :** est un résultat lorsque le modèle prétend qu'une observation fait partie d'une classe alors qu'elle n'en fait pas partie.

**False Negative (FN) :** est un résultat se produit lorsque le modèle prétend que l'observation ne fait pas partie d'une classe alors qu'elle en fait partie. **Figure 14**



		True/Actual Class	
		Positive (P)	Negative (N)
Predicted Class	True (T)	True Positive (TP)	False Positive (FP)
	False (F)	False Negative (FN)	True Negative (TN)
		P=TP+FN	N=FP+TN

Figure 14 : La table de matrice de confusion [47]

### b) Precision

Est une mesure de l'exactitude des prédictions positives faites par un modèle, calculé comme le nombre de prédictions positives correctes (TP) divisé par le nombre total de prédictions positives (TP + FP).[47]

$$\text{Précision} = \frac{TP}{TP+FP} \quad (1)$$

### c) Accuracy

Le ratio de prédictions correctes sur le nombre total d'instances évaluées est mesuré par la métrique de précision.[48]

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

### d) Recall

On le nomme aussi sensibilité pour évaluer la proportion de modèles positifs correctement classés.[49]

$$\text{Recall} = \frac{TP}{TP+FN} \quad (3)$$

### e) F1-score:

La moyenne harmonique entre les valeurs de rappel et de précision est représentée par cette métrique.[49]



$$\text{F1-score} = \frac{2 * \text{precision} * \text{recall}}{\text{precision} + \text{recall}} \quad (4)$$

### 4.8.3 La Fonction de Perte (Loss Function)

La fonction de perte `categorical_crossentropy` est couramment utilisée dans les réseaux de neurones pour classifier plusieurs classes. Elle analyse la différence entre les probabilités de distribution prédites par le modèle et les réelles distributions des étiquettes de classe.

$$L = - \sum_{i=1}^N \sum_{j=1}^C y_{ij} \log(\hat{y}_{ij})$$

Grâce à l'utilisation de fonctions telles que `softmax`, elle évalue la concordance des prédictions du modèle avec les étiquettes réelles. L'apprentissage supervisé vise à réduire cette perte en utilisant des algorithmes d'optimisation tels que la descente de gradient. La métrique quantitative de la `categorical_crossentropy` joue un rôle essentiel dans l'ajustement des paramètres du modèle et dans l'amélioration de sa capacité à classifier efficacement les données multi classe.

#### La formule mathématique

- N est le nombre d'échantillons dans le lot.
- C'est le nombre de classes.
- $y^{ij}$  est la probabilité prédite par le modèle que l'échantillon i appartienne à la classe j.
- $Y_{ij}$  est l'indicateur binaire (0 ou 1) si l'échantillon i appartient à la classe j.

### 4.9 L'entraînement

Dans le cadre de la formation du modèle, les fichiers audio de récitations sont convertis en spectrogrammes, qui sont ensuite enregistrés sous forme d'images de 224x224. La répartition de ces images se fait en ensembles d'entraînement (60%), de validation (20%) et de test (20%). Chaque fichier audio est associé à son image correspondante et à son label grâce à un fichier de métadonnées.

Ensuite, on définit et compile un modèle de réseau de neurones convolutif (CNN) en utilisant l'optimiseur 'adam' et la fonction de perte 'categorical\_crossentropy'. La formation du modèle se déroule sur 50 époques, avec des rappels pour arrêter précocement et réduire le taux d'apprentissage, assurant ainsi une convergence optimale sans surutilisation.

Afin d'évaluer la résistance du modèle, on utilise une validation croisée K-Fold à 5 plis. Dans cette méthode, on subdivise les données d'entraînement en 5 sous-ensembles (ou "plis"). Le modèle subit cinq entraînements, avec un pli différent chaque fois utilisé comme





ensemble de validation et les quatre autres comme ensemble d'entraînement. Pour obtenir une estimation plus précise de la performance du modèle, il est nécessaire de le tester sur diverses parties des données.

Une fois que le modèle a été exécuté, les résultats obtenus sont les suivants : une accuracy de 0,95 et une perte de 0.15 sur l'ensemble de précision. Le modèle a obtenu une validation accuracy de 0.93 sur l'ensemble de validation et une perte de 0.20. Le modèle final est modifié sur l'ensemble d'entraînement et approuvé sur l'ensemble de validation, avec une évaluation finale effectuée sur l'ensemble de test indépendant, garantissant ainsi la généralisation du modèle sur des données non observées.

Afin d'approfondir notre compréhension de l'évolution des performances du modèle pendant l'entraînement, nous avons établi les graphiques de la perte et de la précision pour l'ensemble d'entraînement et l'ensemble de validation sur les 50 époques.

#### 4.9.1 Courbe de Perte

La courbe de perte montre comment la perte (erreur) diminue au fil des époques, tant pour l'ensemble d'entraînement que pour l'ensemble de validation. Une diminution régulière de la perte indique que le modèle apprend correctement, tandis qu'une divergence entre les pertes d'entraînement et de validation peut indiquer un sur-apprentissage. **Figure 14**

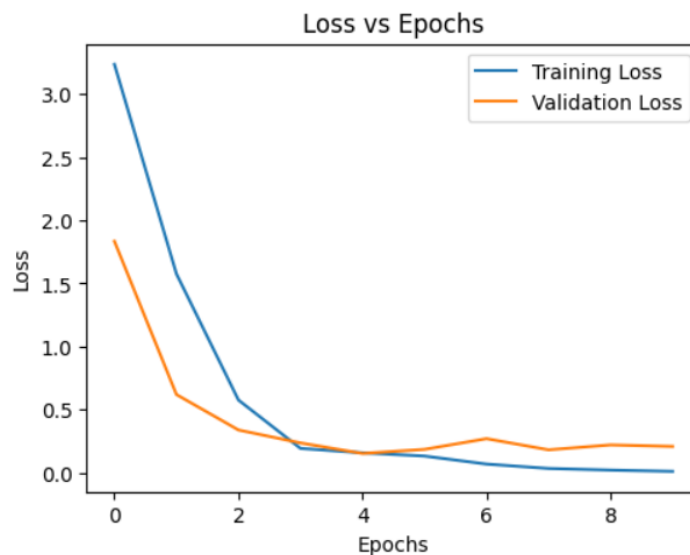


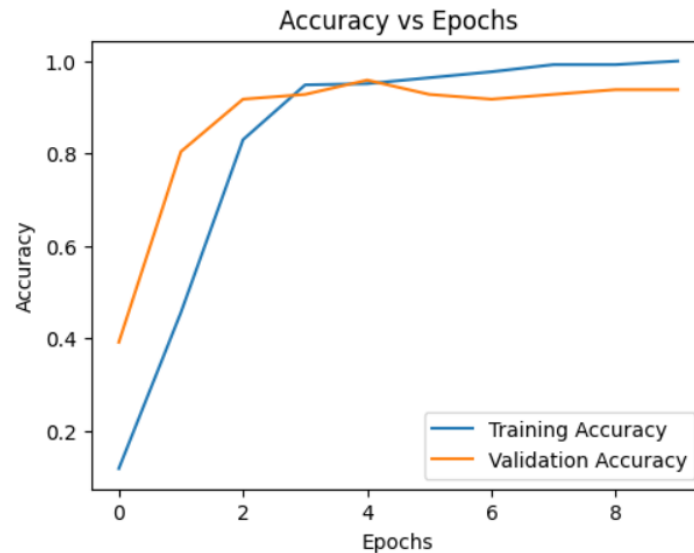
Figure 15 : Valeurs Loss et Val-Loss du modèle

#### 4.9.2 Courbe de Précision

La courbe de précision illustre l'amélioration de la précision du modèle au fil du temps. Une précision croissante pour l'ensemble d'entraînement montre que le modèle devient



de plus en plus performant pour prédire correctement les classes des données d'entraînement. De même, une précision de validation croissante indique une bonne généralisation du modèle sur des données non vues. **Figure 15**



**Figure 16:** Précision de la formation et précision de la validation

Grâce à ces graphiques, il est possible de visualiser de manière claire les tendances d'entraînement et de validation, ce qui offre des informations précieuses sur la performance et la stabilité du modèle tout au long de l'apprentissage.

#### 4.10 Les résultats des métriques d'évaluation

Suite à la formation réussie du modèle proposé, celui-ci a été appliqué à l'ensemble de test, dans lequel ses performances ont été évaluées à l'aide de mesures telles que la précision, le rappel et le score F1 pour chaque classe. Nous avons obtenu une valeur moyenne de **95 %**

Le tableau présente les résultats

Classe	Précision	Accuracy	Recall	F1-score
Verset	0.96	0.95	0.93	0.97

**Tableau 3 :** Les résultats de notre méthodologie



## 4.11 Démonstration de l'interface

### 4.11.1 Architecture Technique

Flutter a été choisi comme framework principal pour son approche "write once, run anywhere", permettant de cibler iOS et Android avec une seule base de code. Cette décision a permis de se concentrer sur la fonctionnalité plutôt que sur l'adaptation à chaque plateforme.

### 4.11.2 Composants Clés

**1. Interface Utilisateur (Flutter) :** Flutter a été utilisé pour créer une interface élégante et intuitive. Des widgets comme `DropDownButtonFormField` pour la sélection de la Sourate ont rendu le développement plus rapide. La fonction Hot Reload de Flutter a été inestimable, permettant d'itérer rapidement sur le design.

**2. Capture Audio (record) :** Ce package a aidé à gérer l'enregistrement audio de manière transparente sur iOS et Android. Une attention particulière a été portée aux permissions du microphone pour assurer une expérience utilisateur fluide.

**3. Lecture Audio (audioplayers) :** Après l'enregistrement, `audioplayers` permet une relecture instantanée, offrant un retour auditif immédiat aux utilisateurs.

**4. Cœur IA (tflite) :** Le package `tflite` intègre le modèle TensorFlow Lite pré-entraîné dans l'application. Optimisé pour mobile, il offre une inférence rapide sur l'appareil, assurant confidentialité et réactivité.

**5. Traitement Audio (FlutterFFmpeg) :** `FFmpeg`, intégré via `flutter_ffmpeg`, convertit l'enregistrement audio en spectrogramme. Cette transformation visuelle est cruciale pour que l'IA "voie" la récitation.

**6. Manipulation d'Images (image) :** Le package `image` prétraite le spectrogramme, le redimensionnant à 224x224 pixels pour correspondre à l'entrée du modèle d'IA.

**7. Gestion des Fichiers (path\_provider) :** `path_provider` identifie les emplacements sûrs pour stocker les fichiers temporaires, renforçant la fiabilité et la sécurité.

### 4.12.3 Flux de Travail

**1. Enregistrement Audio :** Capture d'un audio de haute qualité.



2. **Conversion** : Transformation au format MP3 pour la cohérence.

3. **Génération du Spectrogramme** : Utilisation de FFmpeg pour créer une représentation visuelle.

4. **Prétraitement** : Redimensionnement à 224x224 pixels.

#### 4.13.4 Intégration du Modèle

- Au démarrage, l'application charge le modèle TensorFlow Lite.



Figure 17 : Interface de l'application

- L'utilisateur sélectionne une Sourate et un Ayah, puis enregistre sa récitation



Figure 18 : Interface sélection de sourah et verset

- L'application rejoue l'audio et le convertit en spectrogramme.
- Le modèle analyse l'image et prédit la Sourate récitée.
- Le résultat est affiché.

Dans le cas où l'enregistrement audio correspond au verset et sourah :

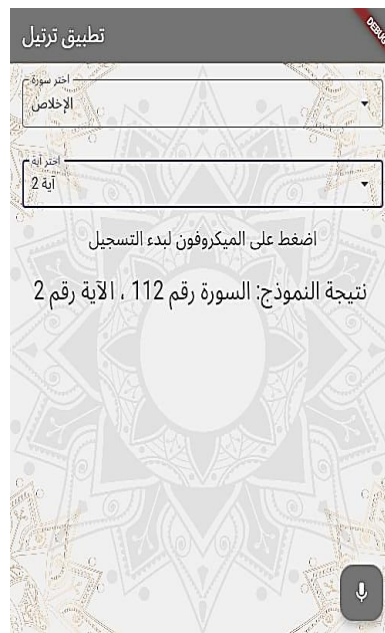


Figure 19 : Interface de test (correct)



Dans le cas où l'enregistrement audio ne correspond pas au verset et sourah :



Figure 20 : Interface de test (incorrect)

#### 4.14 Discussion

L'évaluation du modèle de validation du Coran basé sur l'apprentissage profond a donné des résultats très prometteurs, démontrant ainsi l'efficacité de cette approche innovante. Le modèle, avec une accuracy de 95 % sur l'ensemble des tests, montre une grande capacité à distinguer et classer les différentes parties du Coran. L'extrême accuracy de cette mémoire témoigne de sa capacité à valider de manière fiable les séquences audio et à les relier de manière adéquate aux versets respectifs du Coran.

En analysant attentivement les indicateurs tels que la précision, le rappel et le score F1 et l'accuracy pour chaque catégorie, on observe une performance équilibrée et harmonisée. Pour la classe "Verset", par exemple, le modèle présente une accuracy de 0,95, une précision de 0,96, un rappel de 0,93 et un score F1 de 0,97, ce qui démontre une capacité accrue à identifier et à classer précisément les versets du Coran.



## 4.15 Conclusion

En conclusion, ce chapitre souligne l'importance cruciale de la collecte et du prétraitement des données pour le développement d'un modèle d'intelligence artificielle. En utilisant des enregistrements audio du Coran, nous les avons convertis en spectrogrammes et répartis en ensembles d'entraînement, de validation et de test. De plus, l'architecture du réseau de neurones convolutifs a été définie, mettant en évidence le rôle essentiel des différentes couches et fonctions d'activation.



### Conclusion Générale

Le projet a démontré le potentiel des technologies de l'intelligence artificielle, en particulier le Deep Learning, pour améliorer la validation et la mémorisation de récitation du Coran. En combinant des concepts fondamentaux de l'IA avec des techniques avancées de reconnaissance vocale, nous avons développé un système capable de valider la récitation coranique. Cette initiative offre un outil précieux pour les apprenants du Coran, facilitant un apprentissage précis et respectueux des normes religieuses.

Nous avons utilisé un modèle de réseau de neurones convolutifs (CNN) pour extraire les caractéristiques des spectrogrammes de récitations coraniques. Ce modèle a été enrichi avec des couches supplémentaires de convolution, de normalisation et de pooling pour améliorer les capacités de classification. Les résultats obtenus montrent une accuracy élevée, atteignant une valeur moyenne d'intersection sur union de 95 %

Le projet a impliqué une collecte rigoureuse des données, leur prétraitement et la mise en place d'une architecture de réseau de neurones convolutifs adaptée. Les résultats obtenus montrent une grande précision, démontrant l'efficacité de notre approche. Ce système non seulement préserve l'authenticité de la récitation coranique, mais valorise également le patrimoine culturel et religieux en intégrant les avancées technologiques modernes.

Enfin, le projet montre comment l'intelligence artificielle peut être utilisée pour répondre à des besoins spécifiques et sensibles, tout en respectant les traditions et les pratiques culturelles. Il ouvre la voie à de futures innovations et applications, soulignant l'importance de l'IA dans l'amélioration et la facilitation de diverses activités humaines, y compris celles de nature religieuse et éducative.





## Bibliographie

- [1] C. Villani *et al.*, *Donner un sens à l'intelligence artificielle: pour une stratégie nationale et européenne*. Conseil national du numérique, 2018.
- [2] S. Russell et P. Norvig, *Intelligence artificielle: Avec plus de 500 exercices*. Pearson Education France, 2010.
- [3] W. Ertel, *Introduction to Artificial Intelligence*. Springer, 2018.
- [4] M. Sarazin, *Intelligence artificielle en santé*. ISTE Group, 2023.
- [5] A. Doniec, « Prise en compte des comportements anticipatifs dans la coordination multi-agent : application à la simulation de trafic en carrefour », phdthesis, Université de Valenciennes, 2006. Consulté le: 16 février 2024. [En ligne]. Disponible sur: <https://hal.science/tel-01992947>
- [6] F. Iafrate, *Intelligence artificielle et Big Data: Naissance d'une nouvelle intelligence*. ISTE Group, 2018.
- [7] Mouelhi-Chibani, Wiem., « Apprentissage autonome de réseaux de neurones pour le pilotage en temps réel des systèmes de production basé sur l'optimisation via simulation . ».
- [8] C. Gershenson, « Artificial Neural Networks for Beginners », sept. 2003.
- [9] L. Deng et D. Yu, « Deep Learning: Methods and Applications », mai 2014, Consulté le: 22 mars 2024. [En ligne]. Disponible sur: <https://www.microsoft.com/en-us/research/publication/deep-learning-methods-and-applications/>
- [10] I. Sarker, « Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions », *SN Comput. Sci.*, vol. 2, août 2021, doi: 10.1007/s42979-021-00815-1.
- [11] S. Bansal, « "The Tech Trio: Artificial Intelligence vs Machine Learning vs Deep Learning" », Medium. Consulté le: 9 mai 2024. [En ligne]. Disponible sur: <https://medium.com/@seepika.mittal/artificial-intelligence-vs-machine-learning-vs-data-learning-7f1ddd26598b>
- [12] T. Lin, Y. Wang, X. Liu, et X. Qiu, « A survey of transformers », *AI Open*, vol. 3, p. 111-132, janv. 2022, doi: 10.1016/j.aiopen.2022.10.001.
- [13] Md. Z. Alom *et al.*, « The History Began from AlexNet: A Comprehensive Survey on Deep Learning Approaches », mars 2018.
- [14] Cornuéjols Antoine, *Apprentissage artificiel / Antoine Cornuéjols, Laurent Miclet ; [préface de Jean-Paul Haton]*, 2e édition. in Algorithmes. Paris: Eyrolles, 2010.
- [15] I. Goodfellow, Y. Bengio, et A. Courville, *Deep Learning*. MIT Press, 2016.
- [16] H. Satori, H. Hiyassat, M. Harti, et N. Chenfour, « Investigation Arabic Speech Recognition Using CMU Sphinx System », *Int Arab J Inf Technol*, vol. 6, p. 186-190, avr. 2009.
- [17] « Application à la reconnaissance vocale et de caractère », [intelligenceartificielle.com](http://intelligenceartificielle.com). Consulté le: 22 mars 2024. [En ligne]. Disponible sur: <http://intelligence-artificielle-robotique.weebly.com/application-agrave-la-reconnaissance-vocale-et-de-caractegravere.html>
- [18] J.-P. Haton, C. Cerisara, D. Fohr, Y. Laprie, et K. Smaïli, *Reconnaissance automatique de la parole: Du signal à son interprétation*. Dunod, 2006.
- [19] « Signaux et Systèmes pour l'audiovisuel et la transmission ». Consulté le: 22 mars 2024. [En ligne]. Disponible sur: <https://public.iutenligne.net/telecommunications/berthet/module-signaux-systeme/AnalogiqueNumerique/index.html>



- [20] A. Waibel, T. Hanazawa, G. Hinton, K. Shikano, et K. J. Lang, « Phoneme recognition using time-delay neural networks », *Acoust. Speech Signal Process. IEEE Trans. On*, vol. 37, p. 328-339, avr. 1989, doi: 10.1109/29.21701.
- [21] *مذكرة في أحكام التجويد / جامع الكتب الإسلامية*. Consulté le: 27 février 2024. [En ligne]. Disponible sur: <https://ketabonline.com/ar/books/25819>
- [22] *الموجز المفيد في أحكام التجويد / جامع الكتب الإسلامية*. Consulté le: 27 février 2024. [En ligne]. Disponible sur: <https://ketabonline.com/ar/books/97724>
- [23] *أحكام التجويد في سؤال وجواب / جامع الكتب الإسلامية*. Consulté le: 27 février 2024. [En ligne]. Disponible sur: <https://ketabonline.com/ar/books/92668>
- [24] *الخلاصة من أحكام التجويد / جامع الكتب الإسلامية*. Consulté le: 29 février 2024. [En ligne]. Disponible sur: <https://ketabonline.com/ar/books/69967>
- [25] *الميزان في أحكام تجويد القرآن / جامع الكتب الإسلامية*. Consulté le: 29 février 2024. [En ligne]. Disponible sur: <https://ketabonline.com/ar/books/4658>
- [26] م. ع. الباز, *الحذف البلاغي في الرسم القرآني*. دار اللؤلؤة للنشر والتوزيع, 2021.
- [27] Al Manhal, 2010. ع. ا. ا. مسلم, *سر المقطوع والموصول والتاءات التي بسطت في الرسم القرآني*.
- [28] « مجلة البحوث والدراسات القرآنية » مزايا الرسم العثماني وفوائده. (للدكتور: طه عابدين طه) Consulté le: 1 mars 2024. [En ligne]. Disponible sur: <https://jqrs.qurancomplex.gov.sa/?p=317>
- [29] I. Papastratis, « Speech Recognition: a review of the different deep learning approaches », AI Summer. Consulté le: 12 mai 2024. [En ligne]. Disponible sur: <https://theaisummer.com/speech-recognition/>
- [30] Kalelia.23, « Les Transformers : Voyage au coeur du Deep Learning ! », Valoway. Consulté le: 12 mai 2024. [En ligne]. Disponible sur: <https://www.valoway.com/les-transformers-un-voyage-au-coeur-du-deep-learning/>
- [31] « Paper\_97-Speech\_Recognition\_Models\_for\_Holy\_Quran\_Recitation.pdf », Google Docs. Consulté le: 30 mars 2024. [En ligne]. Disponible sur: [https://drive.google.com/file/d/1GK09xc6err0Yn01ssDNeUXmdT0FnDWD\\_/view?usp=sharing&usp=embed\\_facebook](https://drive.google.com/file/d/1GK09xc6err0Yn01ssDNeUXmdT0FnDWD_/view?usp=sharing&usp=embed_facebook)
- [32] « EFFICIENT HOLY QURAN RECITATION\_JJCIT-2020-06-069\_17\_9\_2020\_Updated.pdf », Google Docs. Consulté le: 30 mars 2024. [En ligne]. Disponible sur: [https://drive.google.com/file/d/16ejIQ23zXjNyEm6bgo--aeb7sTs627yx/view?usp=sharing&usp=embed\\_facebook](https://drive.google.com/file/d/16ejIQ23zXjNyEm6bgo--aeb7sTs627yx/view?usp=sharing&usp=embed_facebook)
- [33] « ASTESJ\_040621.pdf », Google Docs. Consulté le: 30 mars 2024. [En ligne]. Disponible sur: [https://drive.google.com/file/d/1kj3FDChqy0VXdnPZ99i-C4rjK1FWQwwy/view?usp=sharing&usp=embed\\_facebook](https://drive.google.com/file/d/1kj3FDChqy0VXdnPZ99i-C4rjK1FWQwwy/view?usp=sharing&usp=embed_facebook)
- [34] « Quranic\_Verses.pdf », Google Docs. Consulté le: 30 mars 2024. [En ligne]. Disponible sur: [https://drive.google.com/file/d/1AqcpOZqs7IoMvsJDpM5is-fxBCh6jEtz/view?usp=sharing&usp=embed\\_facebook](https://drive.google.com/file/d/1AqcpOZqs7IoMvsJDpM5is-fxBCh6jEtz/view?usp=sharing&usp=embed_facebook)
- [35] J. Robert, « Convolutional Neural Network : Tout ce qu'il y a à savoir », Formation Data Science | DataScientest.com. Consulté le: 16 mai 2024. [En ligne]. Disponible sur: <https://datascientest.com/convolutional-neural-network>
- [36] N. Klingler, « Convolution Operations: an In-Depth 2024 Guide », viso.ai. Consulté le: 1 juin 2024. [En ligne]. Disponible sur: <https://viso.ai/deep-learning/convolution-operations/>
- [37] « Découvrez les différentes couches d'un CNN », OpenClassrooms. Consulté le: 1 juin 2024. [En ligne]. Disponible sur: <https://openclassrooms.com/fr/courses/4470531-classez-et-segmentez-des-donnees-visuelles/5083336-decouvrez-les-differentes-couches-dun-cnn>
- [38] Jorgecardete, « Convolutional Neural Networks: A Comprehensive Guide », The Deep Hub. Consulté le: 16 mai 2024. [En ligne]. Disponible sur: <https://>



- //medium.com/thedeephub/convolutional-neural-networks-a-comprehensive-guide-5cc0b5eae175
- [39] R. Ayachi, Y. Said, et M. Atri, « STUDY AND IMPLEMENTATION OF DEEP LEARNING ALGORITHMS FOR OBJECT DETECTION », 2018.
- [40] « Activation Functions », GeeksforGeeks. Consulté le: 1 juin 2024. [En ligne]. Disponible sur: <https://www.geeksforgeeks.org/activation-functions/>
- [41] B. Maurice, « Fonction d'activation », Deeply Learning. Consulté le: 16 mai 2024. [En ligne]. Disponible sur: <https://deeplylearning.fr/cours-theoriques-deep-learning/fonction-dactivation/>
- [42] O. Qasim et I. Mohammed, « An Analytical Mathematical Study of Artificial Neural Network Algorithms in the Suitability of a Model for Medical Diagnosis », *AL-Rafidain J. Comput. Sci. Math.*, vol. 10, p. 183-194, mars 2013, doi: 10.33899/csmj.2013.163451.
- [43] J. Dominguez, « Answer to “What is the difference between softmax or sigmoid activation for binary classification?” », Stack Overflow. Consulté le: 16 mai 2024. [En ligne]. Disponible sur: <https://stackoverflow.com/a/72771942>
- [44] « Dropout Layer | Artificial Intelligence ». Consulté le: 17 mai 2024. [En ligne]. Disponible sur: [https://leonardoaraujosantos.gitbook.io/artificial-intelligence/machine\\_learning/deep\\_learning/dropout\\_layer](https://leonardoaraujosantos.gitbook.io/artificial-intelligence/machine_learning/deep_learning/dropout_layer)
- [45] « Le tutoriel Python », Python documentation. Consulté le: 18 mai 2024. [En ligne]. Disponible sur: <https://docs.python.org/3/tutorial/index.html>
- [46] S. Prud'homme, « Initiation au Deep Learning avec Google Colab », Moov AI. Consulté le: 18 mai 2024. [En ligne]. Disponible sur: <https://moov.ai/fr/blog/deep-learning-avec-google-colab>
- [47] A. Tharwat, « Classification assessment methods », *Appl. Comput. Inform.*, vol. 17, n° 1, p. 168-192, janv. 2020, doi: 10.1016/j.aci.2018.08.003.
- [48] « (PDF) A Review on Evaluation Metrics for Data Classification Evaluations ». Consulté le: 2 juin 2024. [En ligne]. Disponible sur: [https://www.researchgate.net/publication/275224157\\_A\\_Review\\_on\\_Evaluation\\_Metrics\\_for\\_Data\\_Classification\\_Evaluations](https://www.researchgate.net/publication/275224157_A_Review_on_Evaluation_Metrics_for_Data_Classification_Evaluations)
- [49] « Recall, Precision, F1 Score - Explication Simple Métrique en ML ». Consulté le: 2 juin 2024. [En ligne]. Disponible sur: <https://inside-machinelearning.com/recall-precision-f1-score/>