

**Master Intelligence Artificielle et  
Analyse des Données**



# RECONNAISSANCE DE L'ÉCRITURE MANUSCRITE

**PRÉSENTÉ À** Brahim Elhamdi

**ENCADRÉ PAR :** Pr. Ali BEKRI

# SOMMAIRE

TABLE DES FIGURES.....	3
Introduction générale.....	4
La reconnaissance d'écriture manuscrite.....	5
I.    Revue de la littérature.....	5
1.    Historique de l'OCR :.....	5
II.   Techniques et algorithmes :.....	5
1.    Réseaux de Neurones Convolutifs (CNN) :.....	6
2.    Réseaux de Neurones Récurrents (RNN) :.....	6
3.    CTC (Connectionist Temporal Classification) :.....	6
4.    Approches basées sur l'apprentissage profond :.....	6
III.  Travaux existants.....	6
1.    Tesseract OCR :.....	7
2.    Deep Learning for Text Recognition :.....	7
3.    Projet MADCAT (Multilingual Automatic Document Classification Analysis and Translation) :.....	7
IV.   Types de reconnaissance d'écriture.....	7
1.    Reconnaissance hors ligne.....	8
2.    Reconnaissance en ligne.....	8
V.    Approches de reconnaissance.....	9
1.    Approche globale.....	9
2.    Approche analytique.....	12
VI.   Les différents styles d'écriture.....	13
1.    Écriture Cursive.....	13
2.    Écriture Script :.....	13
3.    Écriture Imprimée.....	14
4.    Styles Culturels et Linguistiques.....	15

5. Influence des Outils d'Écriture .....	16
VII. Applications de l'OCR .....	16
1. Archivage et Préservation des Documents.....	16
2. Gestion de Documents.....	17
3. Accessibilité.....	17
Traitement et Analyse des Images.....	19
I. Prétraitement de l'Image .....	19
1. Binarisation :.....	19
2. Filtrage de Bruit :.....	20
3. Redressement et Rotation : .....	20
4. Segmentation des Lignes et des Mots : .....	21
II. Extraction de Caractéristiques.....	21
III. Classification.....	22
Méthodologie.....	23
I. Jeu de données utilisée :.....	23
1. Pour la langue Anglais :.....	23
2. Pour la langue Arabe : .....	24
II. Modèle :.....	25
1. Modèle CNN pour la Reconnaissance de Texte en Anglais.....	25
2. Modèle CNN pour la Reconnaissance de Texte en Arabe : .....	26
III. Outils et Technologies.....	28
1. Langages de Programmation .....	28
2. Bibliothèques et Frameworks.....	29
3. Environnements de Développement.....	31
IV. Interface graphique pour mon OCR :.....	32
Conclusion.....	33
Bibliographie .....	34

## TABLE DES FIGURES

Figure 1: Reconnaissance d'écriture en ligne .....	8
Figure 2 :IAM Dataset.....	10
Figure 3 : KHATT Database .....	11
Figure 4 : Approche analytique .....	12
Figure 5:Écriture Cursive.....	13
Figure 6 : Écriture Script .....	14
Figure 7 : Écriture Imprimée .....	15
Figure 8:Schéma de traitement et analyse des images.....	19
Figure 9 : binarisation .....	20
Figure 10:Filtrage de Bruit .....	20
Figure 11:Redressement et Rotation.....	21
Figure 12:Segmentation .....	21
Figure 13:Nombre d'occurrences de chaque lettre dans le jeu de données .....	23
Figure 14:Diagramme circulaire .....	24
Figure 15:distribution des classes des lettres.....	25
Figure 16: Loss d'entraînement et de test pour anglais. ....	26
Figure 17: Accuracy d'entraînement et de test pour anglais.....	26
Figure 18:Accuracy d'entraînement et de validation pour Arabe .....	27
Figure 19:matrice de confusion.....	28
Figure 20:Keras.....	29
Figure 21:TensorFlow.....	29
Figure 22:NumPY.....	30
Figure 23:OpenCV.....	30
Figure 24: Matplotlib .....	31
Figure 25:Interface graphique.....	32

# INTRODUCTION GENERALE

La reconnaissance de l'écriture manuscrite (en anglais, handwritten text recognition ou HTR) est un domaine de traitement informatique qui vise à traduire un texte écrit à la main en un texte codé numériquement. Cette technologie est cruciale dans diverses applications, allant de la numérisation des documents historiques à la facilitation de la prise de notes numériques. L'OCR (Optical Character Recognition) est une branche spécifique de cette technologie, focalisée sur la reconnaissance des caractères imprimés ou manuscrits à partir d'images scannées ou photographiées.

L'objectif principal de l'OCR est de permettre aux ordinateurs de lire et de comprendre le texte écrit, ce qui ouvre la voie à l'automatisation des processus de traitement des documents, à la recherche textuelle avancée, et à l'analyse de grandes quantités de données textuelles. L'OCR pour les langues comme l'arabe présente des défis uniques en raison de la nature cursive de l'écriture et de la complexité des caractères.

Dans ce projet, nous nous concentrons sur la reconnaissance de texte arabe à partir d'images. La reconnaissance du texte arabe est particulièrement complexe en raison de la diversité des formes des lettres, de la nécessité de gérer les ligatures et de l'importance du contexte pour la compréhension correcte des caractères individuels.

L'objectif de ce projet est de développer un système OCR capable de reconnaître avec précision les caractères arabes et anglais manuscrits à partir d'images, en utilisant des techniques avancées de traitement d'image et d'apprentissage automatique. Ce rapport détaille les méthodes employées, les défis rencontrés, les solutions mises en œuvre, et les résultats obtenus, avec une attention particulière à l'amélioration de la précision du texte reconnu à travers diverses techniques de post-traitement.

En suivant une approche systématique qui inclut la revue de la littérature existante, la conception et l'implémentation du modèle, et l'évaluation des performances, ce projet vise à contribuer à l'avancement des technologies de reconnaissance de l'écriture manuscrite pour la langue arabe, facilitant ainsi l'accès et l'utilisation des documents manuscrits dans un format numérique.

# LA RECONNAISSANCE D'ECRITURE MANUSCRITE

## I. Revue de la littérature

### 1. Historique de l'OCR :

La reconnaissance optique de caractères (OCR) a une histoire riche qui remonte aux premières tentatives de traduction de textes imprimés en formats lisibles par machine. Les premières techniques d'OCR ont été développées dans les années 1920 et 1930, où des dispositifs mécaniques étaient utilisés pour lire des caractères imprimés spécifiques. Ces premiers systèmes étaient rudimentaires et limités à des caractères très simplifiés.

Dans les années 1950 et 1960, avec l'avènement des ordinateurs électroniques, l'OCR a fait des progrès significatifs. L'un des pionniers dans ce domaine, Ray Kurzweil, a développé un système OCR capable de reconnaître du texte en différentes polices d'écriture. Les années 1970 et 1980 ont vu l'introduction de techniques basées sur des algorithmes plus sophistiqués, permettant une reconnaissance plus précise et la possibilité de traiter des polices variées.

Avec l'émergence de l'intelligence artificielle et du machine learning dans les années 1990 et 2000, l'OCR a bénéficié de nouvelles méthodes basées sur les réseaux de neurones artificiels. Ces techniques ont considérablement amélioré la précision et la robustesse des systèmes OCR. Aujourd'hui, les méthodes modernes utilisent des réseaux de neurones convolutifs (CNN) et des réseaux de neurones récurrents (RNN), offrant des performances remarquables, même pour des textes manuscrits complexes et des langues cursives comme l'arabe.

## II. Techniques et algorithmes :

Les techniques modernes d'OCR se basent principalement sur des algorithmes avancés de traitement d'image et d'apprentissage automatique. Parmi les plus couramment utilisés, on trouve :

## 1. Réseaux de Neurones Convolutifs (CNN) :

Les CNN sont utilisés pour extraire des caractéristiques pertinentes des images de texte. Ils sont particulièrement efficaces pour reconnaître des motifs complexes et sont la base de nombreux systèmes OCR modernes.

## 2. Réseaux de Neurones Récurents (RNN) :

Les RNN, et en particulier les variantes comme les LSTM (Long Short-Term Memory), sont utilisés pour traiter des séquences de données. Dans le contexte de l'OCR, ils permettent de modéliser la séquence des caractères dans un texte.

## 3. CTC (Connectionist Temporal Classification) :

La CTC est une technique utilisée pour entraîner des modèles de séquence à séquence sans alignement explicite des entrées et des sorties. Cela est particulièrement utile dans l'OCR pour les langues où les caractères sont liés de manière complexe, comme l'arabe.

## 4. Approches basées sur l'apprentissage profond :

Les architectures hybrides combinant CNN et RNN, souvent intégrées dans des frameworks comme TensorFlow et PyTorch, sont utilisées pour développer des modèles OCR robustes et précis.

# III. Travaux existants

Plusieurs études et projets ont contribué de manière significative au développement de l'OCR pour les textes manuscrits et les langues cursives. Quelques exemples notables incluent :

## 1. Tesseract OCR :

Initialement développé par HP et maintenant maintenu par Google, Tesseract est un moteur OCR open-source largement utilisé. Bien qu'il ait commencé par traiter principalement des textes imprimés, des versions récentes incluent des capacités pour le texte manuscrit et le texte arabe.

## 2. Deep Learning for Text Recognition :

De nombreuses études académiques ont exploré l'utilisation des CNN et RNN pour la reconnaissance de texte manuscrit. Par exemple, les travaux utilisant les architectures CRNN (Convolutional Recurrent Neural Networks) montrent des performances impressionnantes pour les textes manuscrits dans diverses langues.

## 3. Projet MADCAT (Multilingual Automatic Document Classification Analysis and Translation) :

Ce projet, financé par DARPA, vise à développer des technologies pour la traduction automatique de documents multilingues. Une partie essentielle de ce projet concerne la reconnaissance du texte manuscrit en langues comme l'arabe.

Malgré ces avancées, plusieurs défis demeurent. La reconnaissance de l'écriture manuscrite, en particulier pour les langues cursives comme l'arabe, présente des difficultés liées à la variabilité des écritures, la complexité des ligatures et l'importance du contexte. Les recherches futures doivent se concentrer sur l'amélioration de la robustesse des modèles et sur le traitement efficace des variations inter-écrivains.

# IV. Types de reconnaissance d'écriture

La reconnaissance de l'écriture manuscrite est un processus informatique qui vise à convertir un texte écrit en une forme numérique. Il existe deux principales catégories de reconnaissance :



## 1. Reconnaissance hors ligne

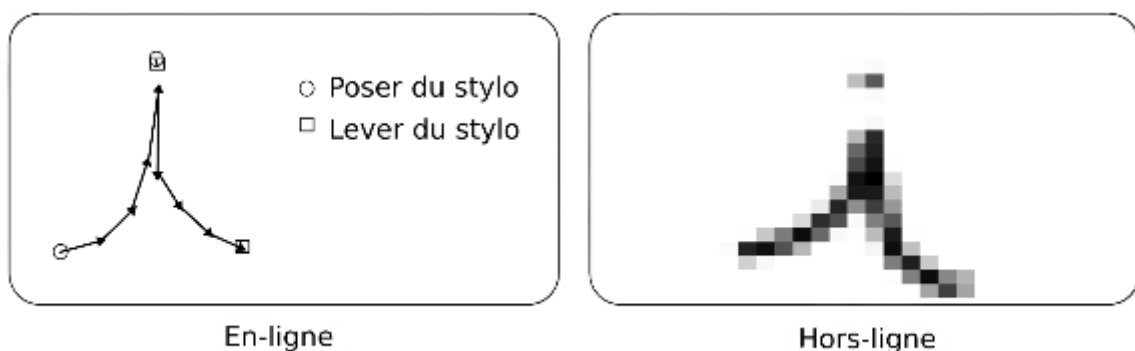
La reconnaissance hors ligne concerne la reconnaissance statique d'images, où le texte manuscrit est reconnu à partir de documents déjà écrits. L'image du texte manuscrit est capturée à l'aide d'un scanner ou d'un appareil photo, produisant une représentation discrète de l'image composée d'un ensemble de pixels. Dans ce contexte, l'écriture apparaît comme un signal spatial bidimensionnel numérisé.

Cette méthode est plus difficile en raison de l'absence d'informations temporelles sur la manière dont les lettres et les mots ont été formés. Les systèmes doivent donc se baser uniquement sur l'image statique pour extraire les caractéristiques des caractères et les reconnaître. Des techniques de traitement d'image et d'apprentissage automatique sont souvent utilisées pour segmenter les caractères, extraire les caractéristiques et effectuer la reconnaissance.

## 2. Reconnaissance en ligne

La reconnaissance en ligne, quant à elle, vise à reconnaître l'écriture au fur et à mesure de son tracé. Cela implique la reconnaissance en temps réel du texte saisi à l'aide d'un stylo et d'une tablette de numérisation. Les informations collectées comprennent une séquence ordonnée de points, définis par leurs coordonnées, échantillonnés à une fréquence fixe.

L'écriture en ligne se présente sous la forme d'une paire de signaux temporels numérisés. Cette méthode bénéficie des informations temporelles sur l'ordre et la dynamique du tracé des caractères, ce qui peut améliorer la précision de la reconnaissance. Par exemple, la direction, la vitesse et la pression du stylo peuvent fournir des indices supplémentaires pour différencier les caractères similaires.



**Figure 1: Reconnaissance d'écriture en ligne**

## V. Approches de reconnaissance

En reconnaissance hors-ligne des mots manuscrits, on distingue deux approches principales :

### 1. Approche globale

L'approche globale consiste à considérer l'image du mot manuscrit comme une entité indivisible, sans la segmenter en parties plus petites. Cette méthode traite le mot entier comme une seule unité de reconnaissance, utilisant des techniques de correspondance de modèles pour identifier le mot le plus probable dans un vocabulaire donné.

Cette approche est particulièrement utile lorsque le vocabulaire est limité, car elle permet de réduire la liste des mots candidats en fonction de la similarité globale de l'image du mot. Cependant, elle présente plusieurs inconvénients :

- Manque de détails :

En ne segmentant pas le mot, cette méthode peut manquer d'informations cruciales nécessaires pour discriminer entre des mots visuellement similaires.

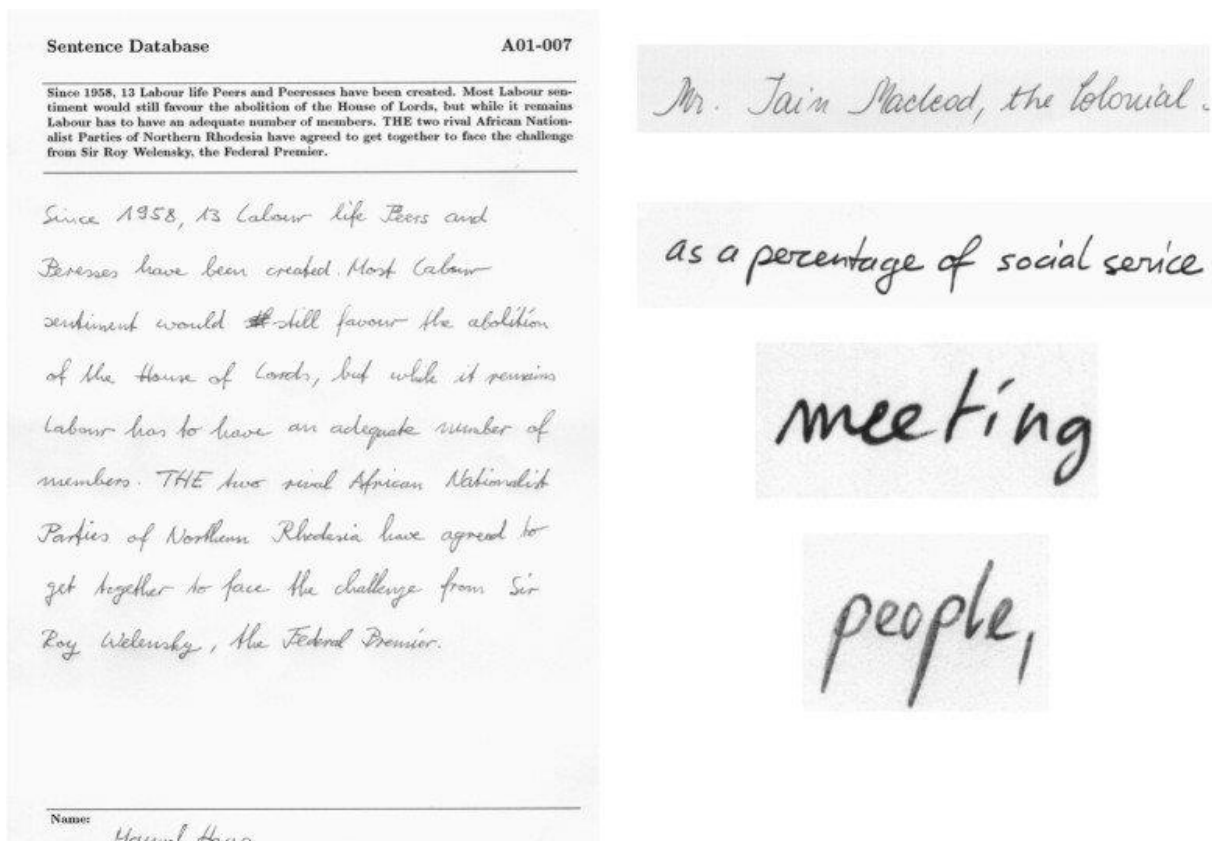
- Exigence en échantillons :

L'apprentissage des modèles globaux nécessite un grand nombre d'échantillons de chaque mot possible pour atteindre une reconnaissance précise, ce qui peut être impraticable dans des applications à grand vocabulaire.

- Jeux de Données de Référence

IAM Handwriting Database :

La IAM Handwriting Database est une base de données couramment utilisée pour l'évaluation de la reconnaissance de l'écriture manuscrite en anglais. Elle contient des images de texte manuscrit écrit par de nombreux auteurs différents, offrant ainsi une grande variété de styles d'écriture. La base de données comprend des images de mots individuels ainsi que des lignes de texte, permettant de tester à la fois la reconnaissance de caractères isolés et de mots entiers.



**Figure 2 :IAM Dataset**

Les avantages de la IAM Handwriting Database incluent :

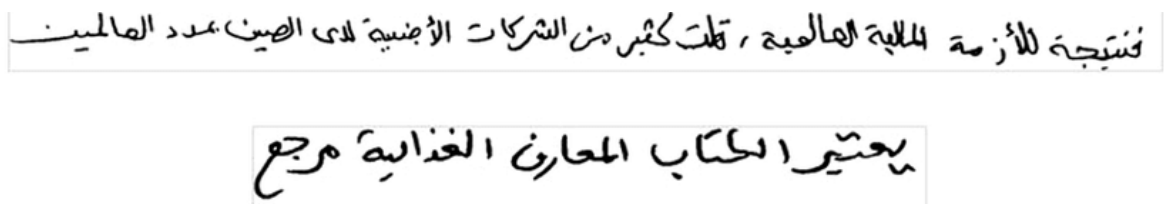
**Variabilité des Écritures :** La diversité des styles d'écriture manuscrite permet de tester la robustesse des algorithmes face à des variations inter-écrivains.

**Standardisation :** En étant largement adoptée, elle permet de comparer les performances des différentes méthodes de reconnaissance de manière cohérente.

**Annotations :** Les images sont annotées avec le texte correspondant, facilitant l'évaluation de la précision des systèmes OCR.

## KHATT Database

La KHATT Database (Kuwaiti Handwritten Text Database) est utilisée pour la reconnaissance de l'écriture manuscrite arabe. Elle contient une vaste collection d'images de texte manuscrit en arabe, écrites par divers auteurs. La base de données inclut des mots et des phrases, offrant un contexte riche pour tester les algorithmes de reconnaissance de texte arabe manuscrit.



**Figure 3 : KHATT Database**

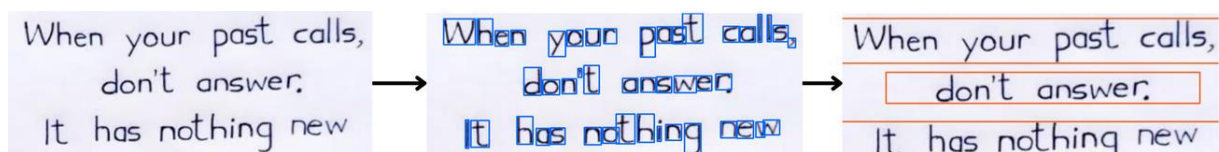
Les avantages de la KHATT Database incluent :

- Complexité de l'Arabe Manuscrit : La base de données prend en compte la complexité des ligatures et des formes contextuelles des lettres arabes, offrant un défi significatif pour les algorithmes OCR.
- Variabilité des Écritures : Comme avec la IAM Handwriting Database, la diversité des styles d'écriture permet de tester la robustesse des algorithmes.
- Standardisation : Utilisée par de nombreux chercheurs, elle permet une comparaison cohérente des performances des systèmes OCR pour l'écriture arabe.

Ces jeux de données de référence sont essentiels pour le développement et l'évaluation des systèmes OCR utilisant l'approche globale. Ils fournissent des benchmarks fiables permettant de mesurer la précision et la robustesse des algorithmes, contribuant ainsi à l'amélioration continue des technologies de reconnaissance de l'écriture manuscrite.

## 2. Approche analytique

L'approche analytique en reconnaissance hors-ligne des mots manuscrits repose sur la segmentation du mot en graphèmes, des unités inférieures aux lettres. Cette méthode permet de décomposer le mot en segments plus petits, souvent correspondant à des caractères ou des groupes de caractères, et de reconstruire le mot en combinant ces graphèmes.



**Figure 4 : Approche analytique**

Les avantages de l'approche analytique incluent :

- **Flexibilité :** En traitant des segments plus petits, cette méthode permet une reconnaissance plus précise, même pour des vocabulaires étendus.
- **Adaptabilité :** Elle est particulièrement adaptée aux applications nécessitant la reconnaissance d'un grand nombre de mots différents, car elle peut gérer les variations dans l'écriture des caractères individuels.
- **Réduction de la complexité :** En segmentant le mot en graphèmes, l'approche analytique peut réduire la complexité de la reconnaissance en permettant l'utilisation de modèles plus simples et plus spécifiques pour chaque segment.

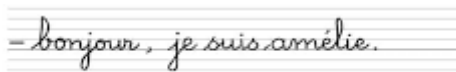
Contrairement à l'approche globale, l'approche analytique offre une meilleure précision et est plus adaptée à des vocabulaires larges et variés, ce qui élargit les possibilités d'applications de la reconnaissance d'écriture manuscrite.

## VI. Les différents styles d'écriture

La reconnaissance de l'écriture manuscrite doit faire face à la grande variabilité des styles d'écriture, ce qui constitue un défi majeur. Chaque personne a un style d'écriture unique, influencé par divers facteurs tels que la culture, l'éducation, et les préférences personnelles. En outre, les styles d'écriture peuvent varier en fonction du contexte et de l'outil d'écriture utilisé. Dans cette section, nous examinerons les principaux styles d'écriture manuscrite et leur impact sur la reconnaissance de texte.

### 1. Écriture Cursive

L'écriture cursive est un style d'écriture où les lettres sont reliées entre elles par des traits continus, créant un flux ininterrompu de mots. Ce style est souvent enseigné aux enfants dès le début de leur éducation formelle.



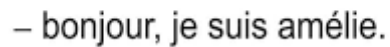
**Figure 5: Écriture Cursive**

Caractéristiques : Les lettres sont souvent inclinées et les connexions entre les lettres peuvent varier considérablement.

Défis pour l'OCR : La nature connectée des lettres peut rendre difficile la segmentation en caractères individuels. Les variations dans les connexions et les ligatures ajoutent une couche de complexité supplémentaire pour les algorithmes de reconnaissance.

### 2. Écriture Script :

L'écriture script est une forme d'écriture où les lettres sont formées individuellement, sans connexions continues entre elles. Ce style est souvent utilisé pour une écriture plus lisible et formelle.



– bonjour, je suis amélie.

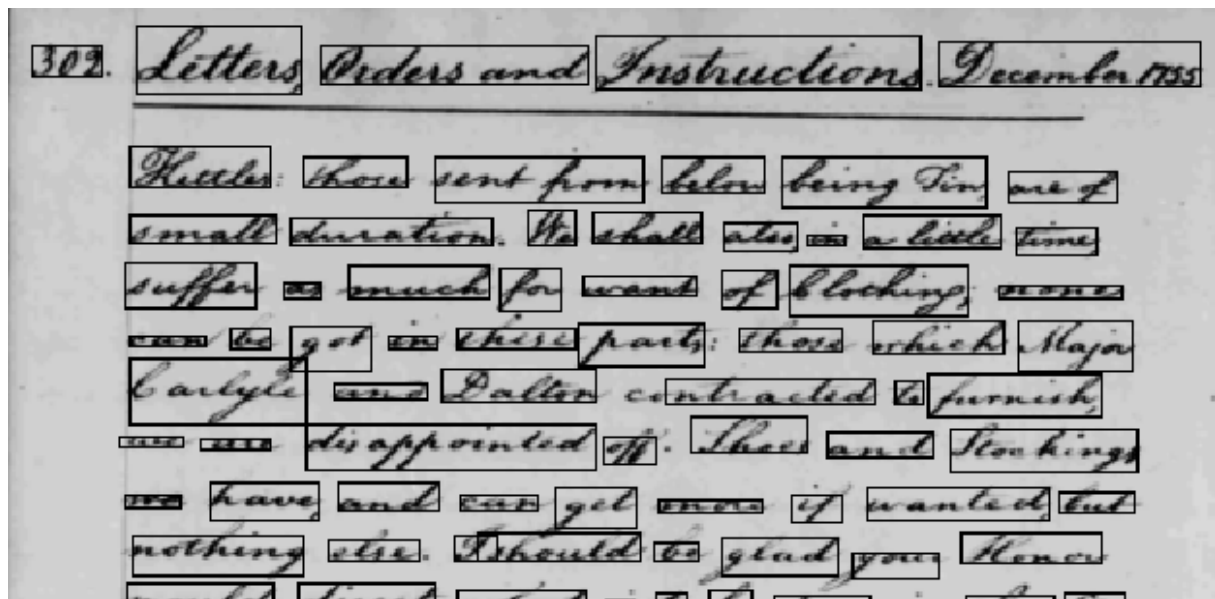
**Figure 6 : Écriture Script**

Caractéristiques : Chaque lettre est distincte, ce qui peut faciliter la reconnaissance de caractères.

Défis pour l'OCR : Bien que plus facile à segmenter que l'écriture cursive, les variations dans la formation des lettres et la présence de bruit (comme des traits de stylo irréguliers) peuvent encore poser des défis pour la précision de la reconnaissance.

### 3. Écriture Imprimée

L'écriture imprimée, ou "print handwriting", est un style où les lettres ressemblent à celles utilisées dans les textes imprimés. Ce style est souvent utilisé pour des formulaires ou des documents officiels.



**Figure 7 : Écriture Imprimée**

Caractéristiques : Lettres distinctes et souvent sans inclinaison, ressemblant aux caractères d'imprimerie.

Défis pour l'OCR : Bien que ce style soit généralement le plus facile à reconnaître, les variations personnelles et les imperfections d'écriture peuvent encore affecter la précision de la reconnaissance.

#### 4. Styles Culturels et Linguistiques

Les styles d'écriture peuvent également varier considérablement d'une culture à l'autre, influençant la reconnaissance de l'écriture manuscrite dans différentes langues. Par exemple :

Écriture Arabe : L'écriture arabe se caractérise par des lettres qui changent de forme en fonction de leur position dans le mot (initiale, médiane, finale, isolée). Les ligatures et les variations de forme ajoutent une complexité unique à la reconnaissance de l'écriture arabe.

Écriture Chinoise et Japonaise: Les caractères chinois et japonais sont souvent complexes, composés de nombreux traits individuels. La reconnaissance de ces



caractères nécessite des techniques spécifiques pour gérer la complexité et la variabilité des traits.

## 5. Influence des Outils d'Écriture

Le type d'outil d'écriture utilisé peut également affecter le style d'écriture et, par conséquent, la reconnaissance de l'écriture manuscrite. Par exemple :

- Stylo à bille vs. Stylo plume : Les variations dans la pression et le débit d'encre peuvent affecter l'apparence des lettres.
- Crayon vs. Stylet numérique : Les outils numériques peuvent offrir des informations supplémentaires, telles que la pression et l'angle du stylet, qui peuvent être utilisées pour améliorer la reconnaissance.

## VII. Applications de l'OCR

La reconnaissance optique de caractères (OCR) a une multitude d'applications dans divers domaines. Ces applications exploitent la capacité des systèmes OCR à convertir des images de texte manuscrit ou imprimé en données numériques. Voici une description plus détaillée de certaines des principales applications de l'OCR :

### 1. Archivage et Préservation des Documents

L'archivage et la préservation des documents historiques et culturels sont des domaines clés où l'OCR joue un rôle essentiel.

- Numérisation de Manuscrits Historiques : Les bibliothèques et les archives utilisent l'OCR pour numériser des manuscrits anciens, des lettres, et des

journaux. Cela permet de préserver ces documents et de les rendre accessibles au public via des bases de données numériques.

- **Accessibilité et Recherche** : En transformant les documents physiques en formats numériques, l'OCR permet des recherches textuelles rapides, facilitant ainsi l'accès et l'étude de documents historiques par les chercheurs et le grand public.
- **Restitution de Textes Endommagés** : Les techniques avancées de l'OCR peuvent aider à reconstruire des textes partiellement endommagés ou illisibles, en complétant les parties manquantes grâce à des modèles prédictifs.

## 2. Gestion de Documents

L'OCR est largement utilisé dans les environnements professionnels pour la gestion et l'organisation des documents.

- **Automatisation des Entrées de Données** : Les entreprises utilisent l'OCR pour automatiser l'extraction de données à partir de factures, de formulaires, de contrats, et d'autres documents administratifs. Cela réduit les erreurs humaines et accélère les processus de traitement des documents.
- **Archivage Électronique** : Les documents papier sont numérisés et convertis en formats électroniques, facilitant leur stockage, leur recherche, et leur récupération.
- **Indexation de Documents** : Les systèmes OCR peuvent indexer les documents numérisés, permettant une recherche efficace par mots-clés dans de grandes bases de données documentaires.

## 3. Accessibilité

Les technologies OCR jouent un rôle crucial dans l'amélioration de l'accessibilité pour les personnes ayant des déficiences visuelles.

- **Lecteurs d'Écran** : Les logiciels OCR peuvent convertir du texte imprimé en texte numérique, qui peut ensuite être lu à haute voix par des lecteurs

d'écran, aidant ainsi les personnes aveugles ou malvoyantes à accéder aux informations écrites.

- Applications Mobiles : Des applications mobiles équipées de la technologie OCR permettent aux utilisateurs de prendre des photos de textes imprimés et de les convertir instantanément en formats lisibles ou audibles.
- Sous-titrage Automatique : Les systèmes OCR peuvent être utilisés pour générer des sous-titres à partir de textes affichés dans des vidéos, rendant les contenus vidéo accessibles à un public plus large.

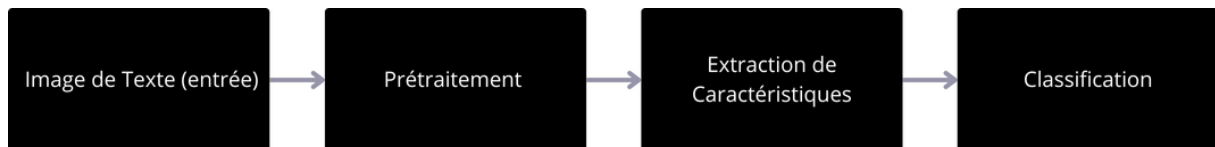
#### 4. Automatisation de la Donnée

L'OCR est utilisé pour automatiser l'extraction de données dans diverses industries, rendant les processus plus efficaces et moins sujets aux erreurs.

- Traitement des Chèques : Les banques utilisent l'OCR pour lire et traiter automatiquement les chèques, en extrayant les informations nécessaires comme les montants, les numéros de compte, et les signatures.
- Reconnaissance des Formulaires Médicaux : Les hôpitaux et les cliniques utilisent l'OCR pour numériser les formulaires médicaux, facilitant l'intégration des données patient dans les systèmes de gestion des informations de santé.
- Lecture des Plaques d'Immatriculation : Les systèmes OCR sont utilisés dans les applications de surveillance du trafic pour lire les plaques d'immatriculation des véhicules, facilitant ainsi le contrôle de la circulation et l'application des lois.

# TRAITEMENT ET ANALYSE DES IMAGES

Ce chapitre couvre les diverses techniques et processus utilisés pour préparer les images de texte afin de les rendre prêtes pour la reconnaissance de caractères, ainsi que les méthodes de classification et de post-traitement.



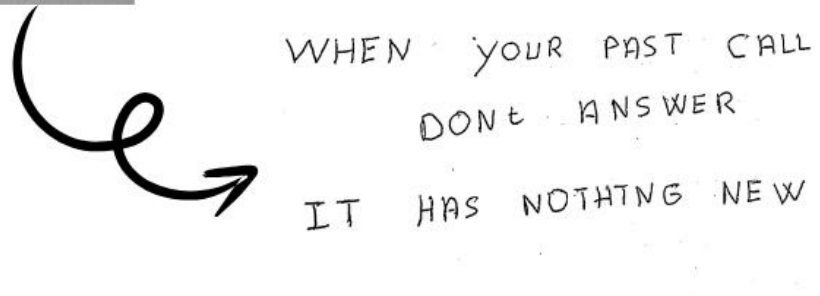
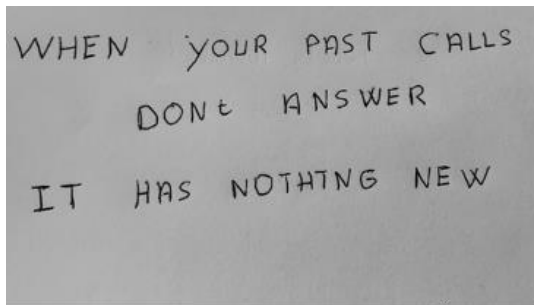
**Figure 8:Schéma de traitement et analyse des images**

## I. Prétraitement de l'Image

Le prétraitement est une étape essentielle qui prépare les images pour la reconnaissance des caractères. Il inclut plusieurs sous-étapes :

### 1. Binarisation :

Ce processus convertit une image en niveaux de gris en une image binaire (noir et blanc). Cela simplifie l'image en éliminant les nuances de gris, ce qui facilite la détection des contours et des formes des caractères. Les méthodes courantes incluent la binarisation d'Otsu et la binarisation adaptative.



**Figure 9 : binarisation**

## 2. Filtrage de Bruit :

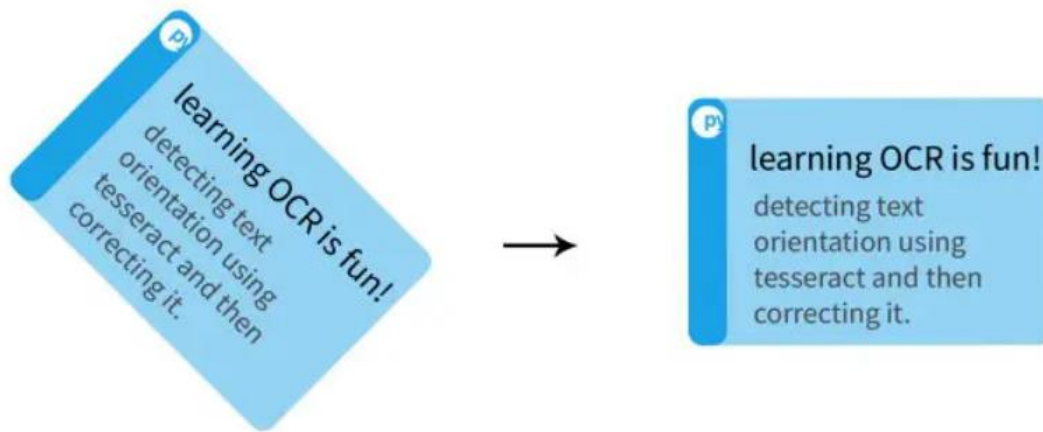
Les images scannées peuvent contenir du bruit (taches, distorsions, etc.) qui peut perturber la reconnaissance des caractères. Des techniques de filtrage, comme le filtrage médian ou gaussien, sont utilisées pour réduire ce bruit tout en préservant les contours des caractères.



**Figure 10: Filtrage de Bruit**

## 3. Redressement et Rotation :

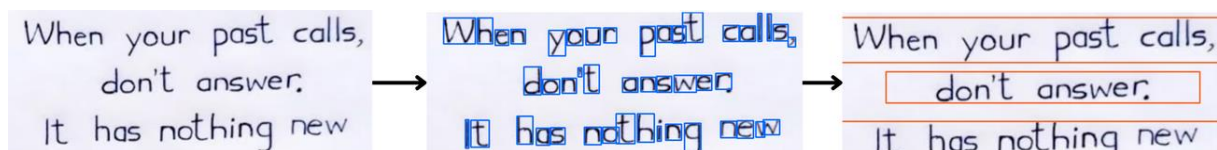
Les images de texte peuvent être inclinées en raison d'un mauvais positionnement lors de la numérisation. Des algorithmes de détection et de correction de l'angle d'inclinaison sont appliqués pour aligner correctement le texte.



**Figure 11: Redressement et Rotation**

#### 4. Segmentation des Lignes et des Mots :

Cette étape divise l'image en lignes de texte et en mots individuels, facilitant l'analyse ultérieure. Les méthodes incluent la projection horizontale pour segmenter les lignes et la projection verticale pour segmenter les mots.



**Figure 12: Segmentation**

## II. Extraction de Caractéristiques

L'extraction de caractéristiques est une étape où les attributs distinctifs des caractères sont identifiés et extraits. Ces caractéristiques servent de base pour la classification et la reconnaissance des caractères.

- **Contours et Formes** : Les contours des caractères sont détectés en utilisant des techniques comme la détection de contours de Canny. Les formes géométriques des caractères sont analysées pour identifier les traits distinctifs.
- **Descripteurs de Texture** : Les descripteurs comme les Histogrammes de Gradients Orientés (HOG) capturent les gradients et les orientations dans les

images, fournissant des informations précieuses sur la texture et les motifs des caractères.

- Transformée de Fourier et de Wavelet : Ces techniques sont utilisées pour extraire les caractéristiques de fréquence et d'échelle des caractères, respectivement. Elles sont particulièrement utiles pour la reconnaissance des caractères manuscrits et des polices variées.

### III. Classification

La classification est l'étape où les caractéristiques extraites sont utilisées pour identifier les caractères individuels. Divers algorithmes de classification sont utilisés en fonction des besoins et des complexités des données.

- Support Vector Machines (SVM) : Les SVM sont efficaces pour les petits ensembles de données et les tâches de classification linéaire. Ils sont souvent utilisés dans les premières étapes de la reconnaissance OCR.
- Réseaux de Neurones Convolutifs (CNN) : Les CNN sont largement utilisés pour la reconnaissance d'images complexes. Ils sont capables d'apprendre automatiquement les caractéristiques pertinentes à partir des données d'entraînement, offrant une haute précision dans la reconnaissance des caractères manuscrits et imprimés.
- Réseaux de Neurones Récurrents (RNN) : Les RNN, et en particulier les LSTM (Long Short-Term Memory), sont utilisés pour la reconnaissance de séquences de caractères, comme dans les textes manuscrits cursifs où les lettres sont connectées.

# METHODOLOGIE

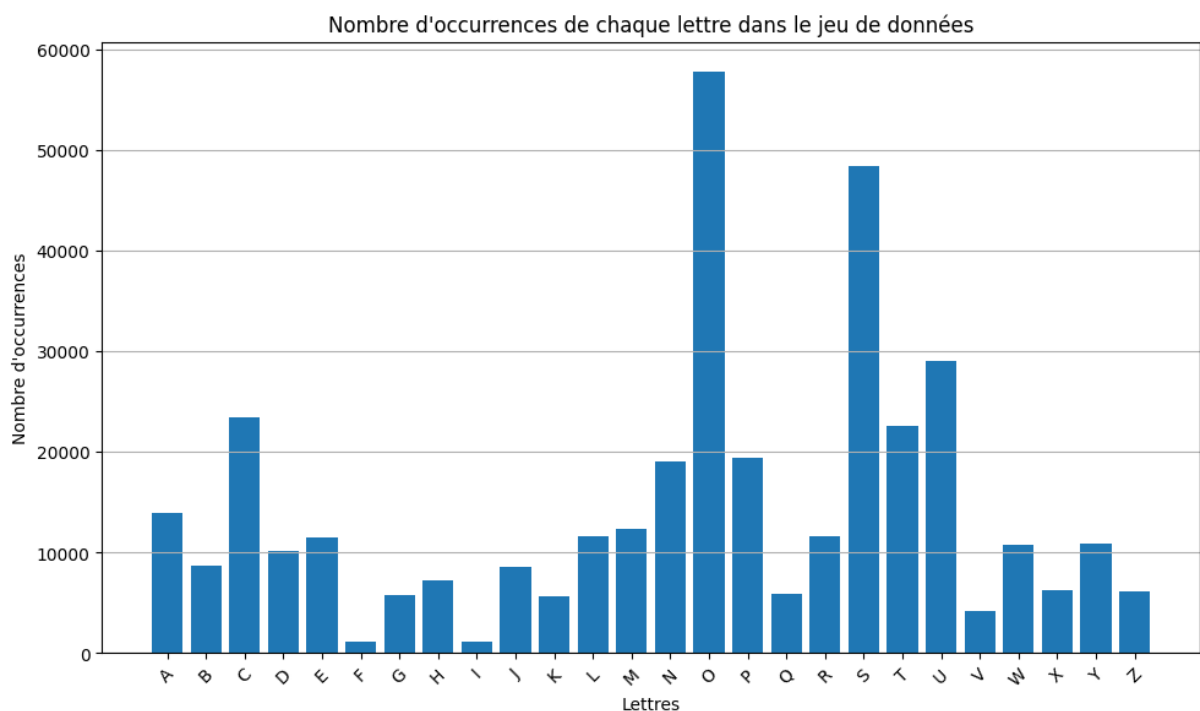
## I. Jeu de données utilisée :

### 1. Pour la langue Anglais :

Le jeu de données contient des images manuscrites des lettres de A à Z. Chaque image est stockée en niveaux de gris et représente une lettre centrée. Le dataset est constitué de 26 classes, correspondant à chaque lettre de l'alphabet anglais.

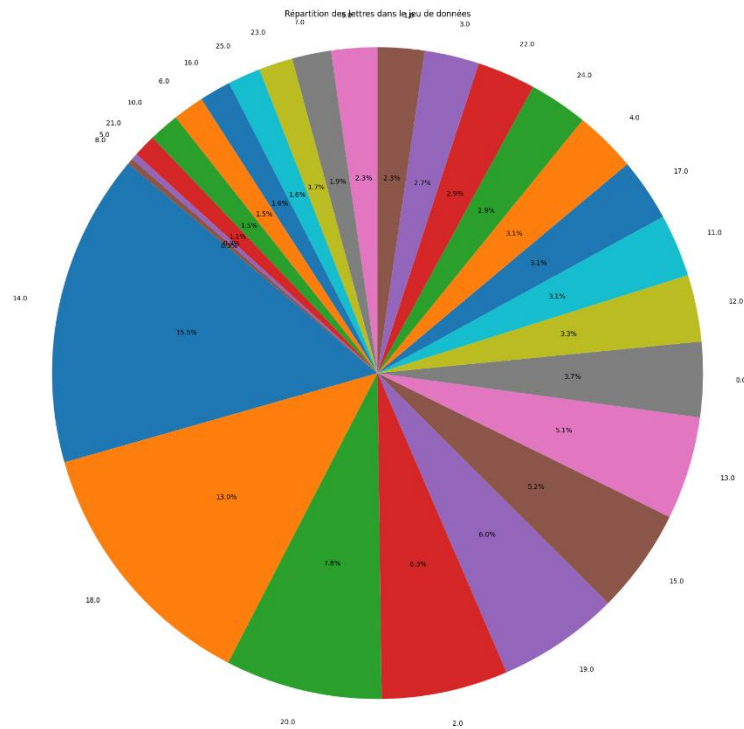
La distribution des classes dans le jeu de données montre le pourcentage d'occurrence de chaque lettre, ce qui permet d'analyser la fréquence de chaque caractère manuscrit. Les données sont préparées de manière à ce que chaque lettre soit facilement identifiable et utilisable pour des tâches de reconnaissance de caractères.

Visualiser la distribution des classes des lettres :



**Figure 13: Nombre d'occurrences de chaque lettre dans le jeu de données**



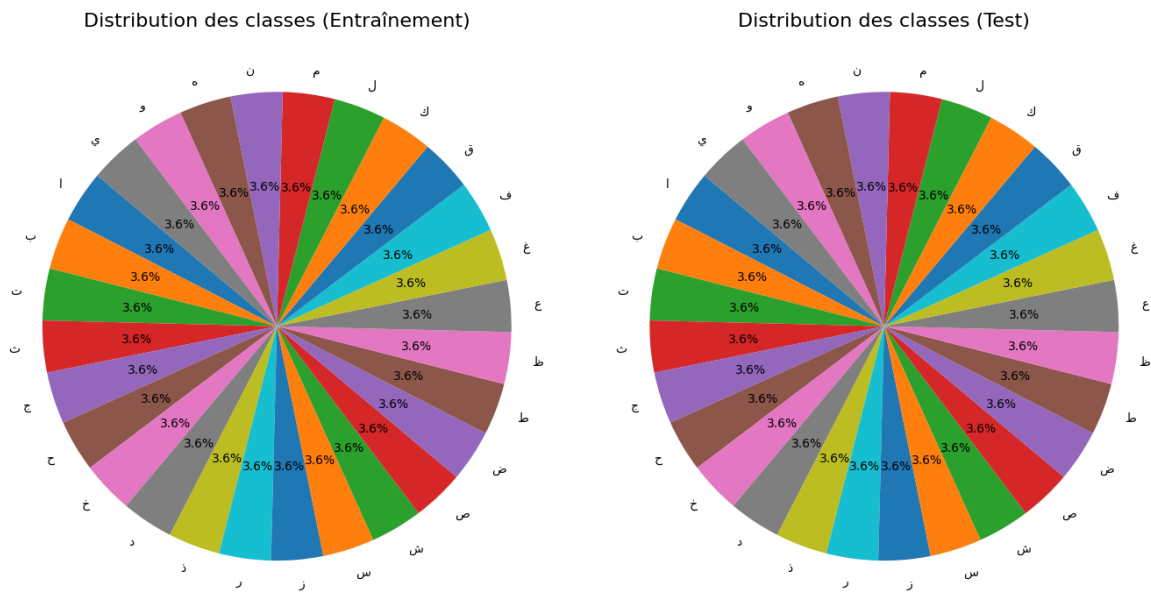


**Figure 14:Diagramme circulaire**

## 2. Pour la langue Arabe :

Le jeu de données contient des lettres arabes écrites à la main, réparties en 28 classes (ا - ي). Le dataset comprend un total de 16,799 images, avec 13,439 images dans l'ensemble d'entraînement et 3,360 images dans l'ensemble de test. Chaque image est de taille 32x32 pixels et est stockée en niveaux de gris. Les lettres sont centrées dans chaque image pour assurer une représentation uniforme et claire de chaque caractère.

Visualiser la distribution des classes des lettres :



**Figure 15: distribution des classes des lettres**

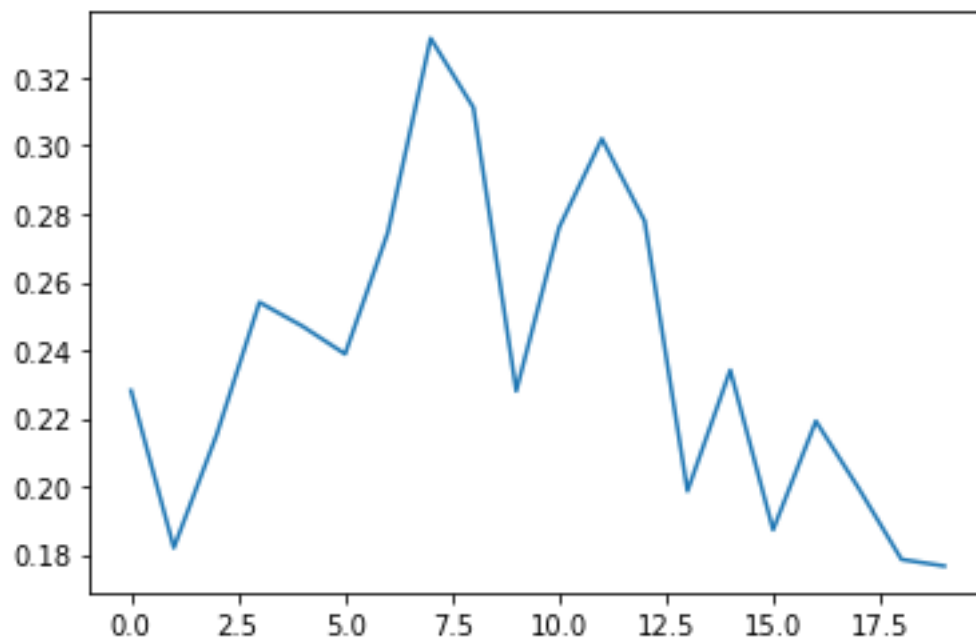
## II. Modèle :

### 1. Modèle CNN pour la Reconnaissance de Texte en Anglais

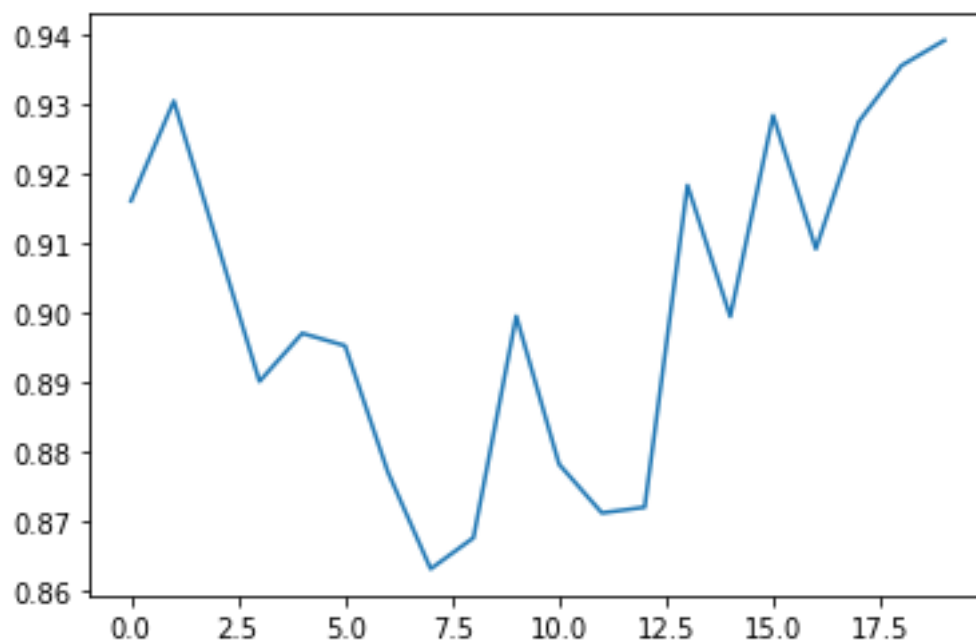
Pour la reconnaissance de caractères manuscrits en anglais, un modèle de réseau de neurones convolutifs (CNN) est souvent utilisé en raison de sa capacité à apprendre automatiquement les caractéristiques spatiales des images.

Structure du Modèle :

Le modèle est composé de plusieurs couches de convolution et de pooling, suivies de couches entièrement connectées (dense). Chaque couche a un rôle spécifique dans l'extraction des caractéristiques et la classification des caractères manuscrits.



**Figure 16: Loss d'entraînement et de test pour anglais.**



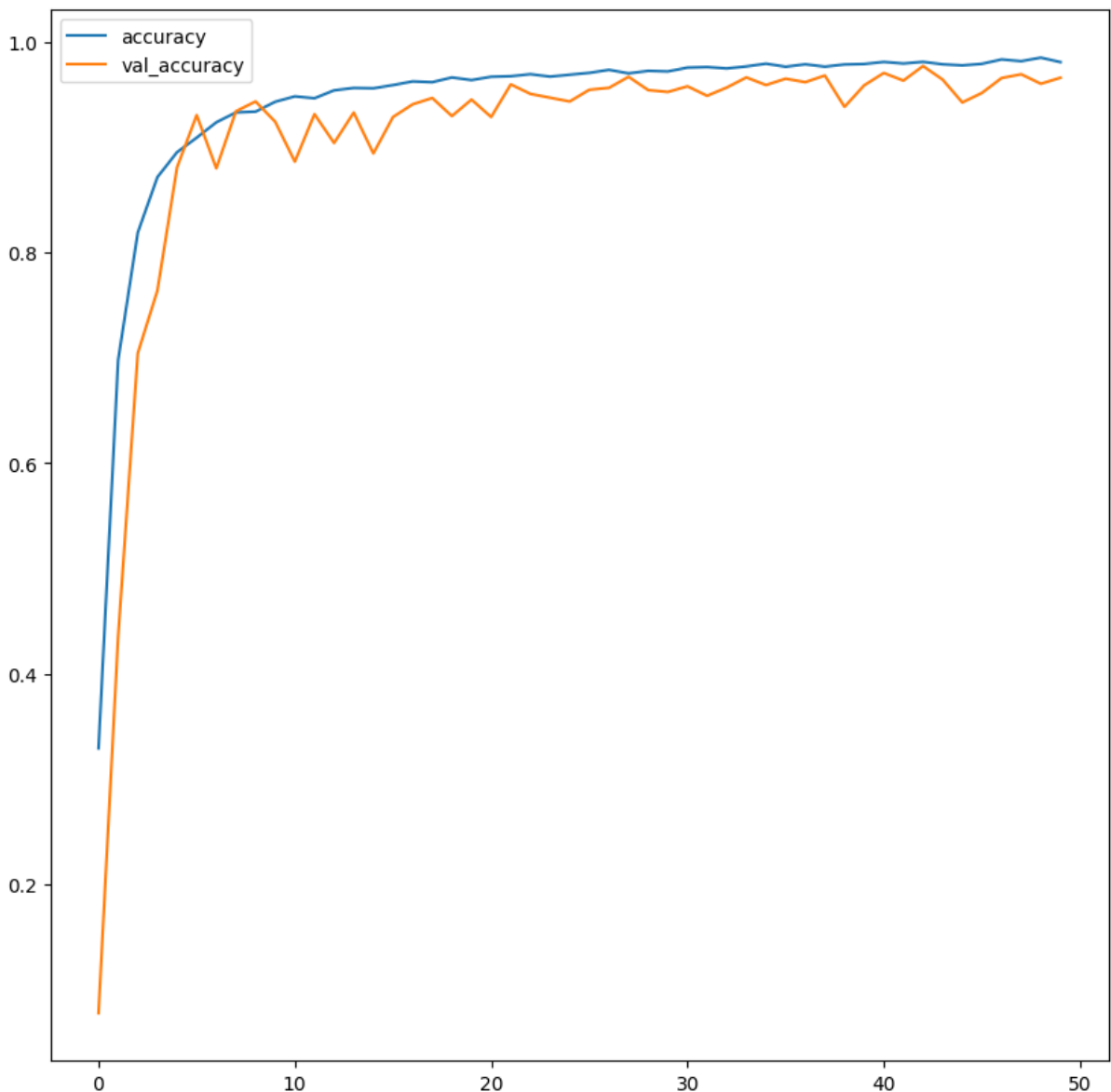
**Figure 17: Accuracy d'entraînement et de test pour anglais.**

## 2. Modèle CNN pour la Reconnaissance de Texte en Arabe :

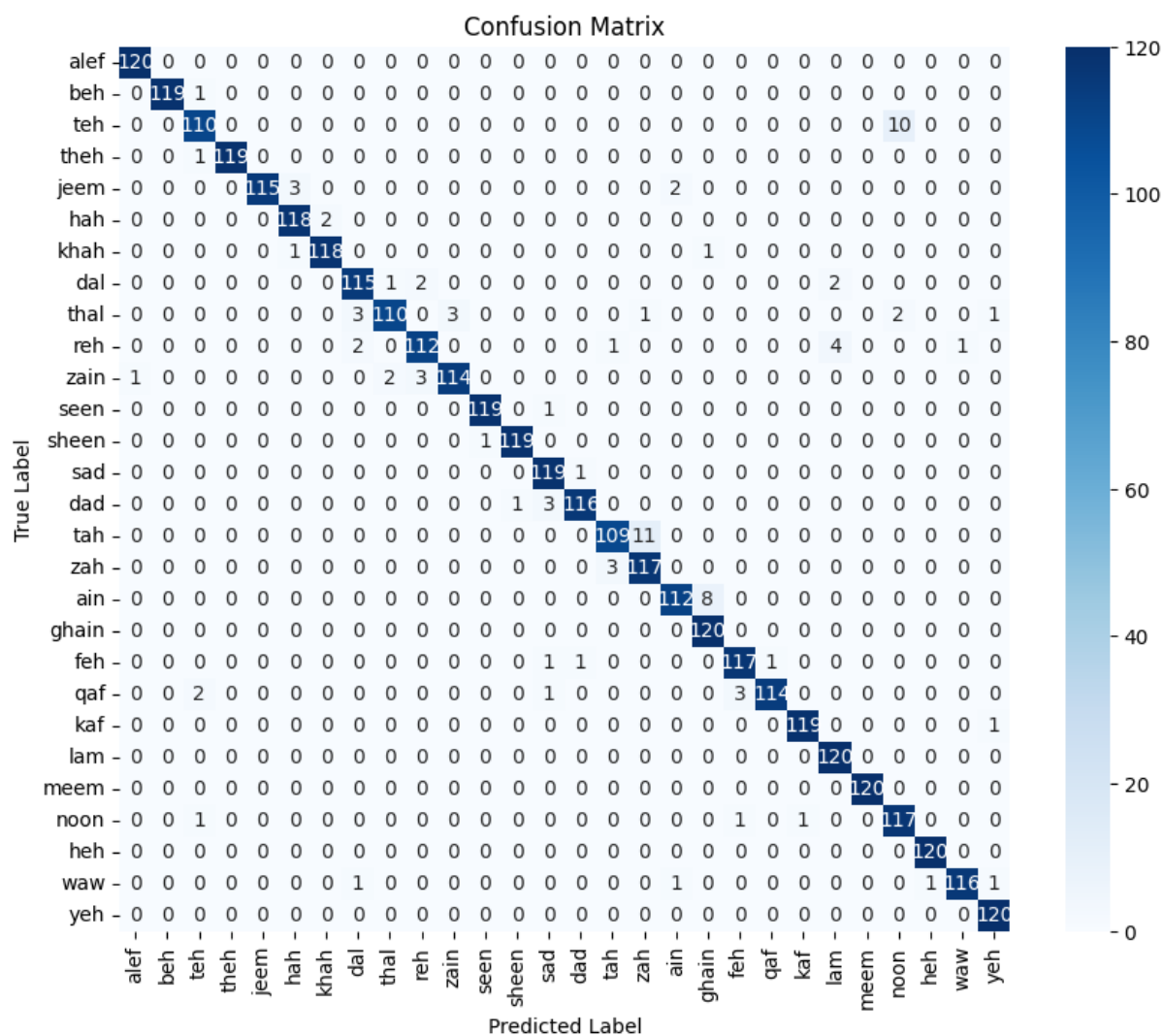
Pour la reconnaissance de caractères manuscrits en arabe, un modèle de réseau de neurones convolutifs (CNN) est également utilisé. Ce modèle est conçu pour traiter des images de 32x32 pixels en niveaux de gris (1 canal).

Structure du Modèle :

Le modèle est composé de plusieurs couches de convolution et de pooling, suivies de couches entièrement connectées (dense) avec une régularisation par Dropout pour éviter le surapprentissage. Il utilise également des couches de normalisation par lot (BatchNormalization) pour améliorer la stabilité et la vitesse d'apprentissage.



**Figure 18:Accuracy d'entraînement et de validation pour Arabe**



**Figure 19:matrice de confusion**

### III. Outils et Technologies

Pour le développement et la mise en œuvre de systèmes de reconnaissance de l'écriture manuscrite (HTR), plusieurs outils et technologies sont essentiels. Voici un aperçu des principales technologies utilisées dans ce projet :

#### 1. Langages de Programmation

Python : Principal langage utilisé pour le développement de modèles d'apprentissage automatique et d'apprentissage profond. Python est préféré pour sa simplicité et son vaste écosystème de bibliothèques scientifiques.

## 2. Bibliothèques et Frameworks

- **Keras** : Une API de haut niveau pour les réseaux de neurones, qui permet de créer et d'entraîner des modèles d'apprentissage profond de manière simple et efficace. Keras fonctionne sur des backend comme TensorFlow.



Figure 20:Keras

- **TensorFlow** : Une bibliothèque open source développée par Google pour le calcul numérique et l'apprentissage automatique. TensorFlow est utilisé pour construire et entraîner des modèles d'apprentissage profond.



Figure 21:TensorFlow

- **NumPy** : Une bibliothèque fondamentale pour le calcul scientifique en Python. NumPy est utilisé pour la manipulation et le traitement des données, notamment pour les opérations sur les tableaux et les matrices.



Figure 22:NumPY

- **OpenCV** : Une bibliothèque open source pour la vision par ordinateur. OpenCV est utilisée pour le prétraitement des images, y compris la binarisation, le redimensionnement et la rotation.



Figure 23:OpenCV

- **Matplotlib** : Une bibliothèque de visualisation en Python utilisée pour créer des graphiques et des visualisations des données, ce qui est utile pour l'analyse et la compréhension des performances des modèles.



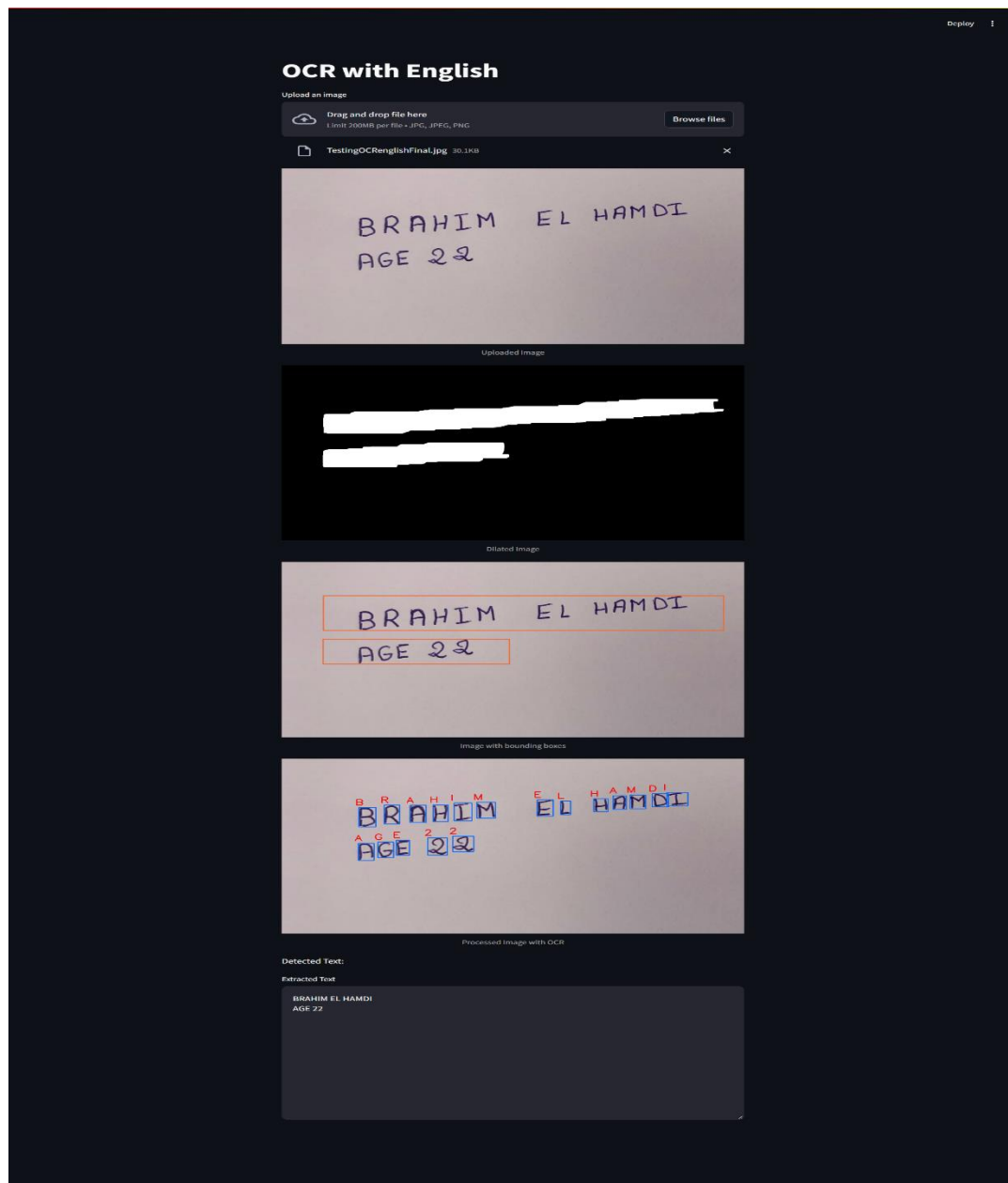
**Figure 24: Matplotlib**

### 3. Environnements de Développement

- **Jupyter Notebook** : Un environnement interactif qui permet de créer et de partager des documents contenant du code, des visualisations et du texte narratif. Jupyter Notebook est particulièrement utile pour le prototypage et l'analyse exploratoire des données.



#### IV. Interface graphique pour mon OCR :



**Figure 25:Interface graphique**

## CONCLUSION

Ce projet de reconnaissance de l'écriture manuscrite (HTR) a exploré efficacement les techniques modernes pour la conversion précise de texte manuscrit en format numérique, en se concentrant principalement sur les langues anglaise et arabe. À travers l'utilisation de modèles de réseaux de neurones convolutifs (CNN), nous avons réussi à construire des systèmes capables de traiter des images de texte manuscrit avec une bonne précision.

Nous avons examiné deux approches principales de reconnaissance : hors ligne et en ligne, chacune présentant des défis uniques liés à la capture et à la représentation des données. L'évaluation des performances s'est appuyée sur des bases de données standardisées comme IAM pour l'anglais et KHATT pour l'arabe, offrant ainsi une comparaison significative des résultats obtenus.

L'infrastructure technologique, incluant des outils comme Keras, TensorFlow, et OpenCV, a joué un rôle crucial dans le développement et l'entraînement des modèles. Ces outils ont permis une manipulation efficace des données, une construction optimale des réseaux neuronaux, et une évaluation rigoureuse des performances.

En conclusion, ce projet illustre le potentiel prometteur des modèles de HTR basés sur CNN pour des applications diverses, de la numérisation de documents à l'amélioration de l'accessibilité numérique. Les prochaines étapes pourraient inclure l'amélioration continue des modèles avec des techniques avancées d'apprentissage profond et l'expansion vers d'autres langues et domaines d'application.

## BIBLIOGRAPHIE

- <https://www.datacamp.com/tutorial/introduction-to-convolutional-neural-networks-cnns>
- <https://www.coursera.org/learn/neural-networks-deep-learning>
- <https://www.sciencedirect.com/topics/computer-science/handwriting-recognition#:~:text=First%2C%20the%20handwriting%20to%20be,using%20lexicons%20or%20spelling%20checkers.>