



جامعة العربي التبسي - تبسة
Université Larbi Tébessi - Tébessa

République Algérienne Démocratique et Populaire
Ministère de l'enseignement supérieur et de la
recherche scientifique

Université Larbi Tébessi - Tébessa

Faculté des Sciences Exactes et des Sciences de la Nature et de la Vie
Département : Mathématiques et Informatique



كلية العلوم الدقيقة وعلوم الطبيعة والحياة
FACULTÉ DES SCIENCES EXACTES
ET DES SCIENCES DE LA NATURE ET DE LA VIE

Mémoire de fin d'étude
Pour l'obtention du diplôme de *MASTER*
Domaine : Mathématiques et Informatique
Filière : Informatique
Option : Systèmes multi média

Thème

identification de scripteurs en utilisant l'apprentissage en profondeur

Présenté Par :

Maatoub Soumaia

Devant le jury :

Mr.. <i>A zaggari</i>	<i>MAA</i>	Université Larbi Tébessa	Président
Mr. <i>A zamaar</i>	<i>MAB</i>	Université Larbi Tébessa	Examineur
Mr. <i>abdejalili gattal</i>	<i>MCA</i>	Université Larbi Tébessa	Encadreur
Mr. <i>faiçel abbas</i>	<i>MCB</i>	Université Larbi Tébessa	Co-Encadreur

Date de soutenance : juin 2019

Résumé

Les recherches antérieures sur l'identification de scripteurs ont été basées sur l'identification de caractéristiques spécifiques aux scripteurs conçues par des experts. Cependant, au cours de la dernière décennie, des méthodes d'apprentissage en profondeur ont été appliquées avec succès pour apprendre automatiquement les caractéristiques à partir de données. Nous proposons ici une méthode d'apprentissage en profondeur pour l'identification de scripteurs indépendante du texte qui ne nécessitant pas d'identification préalable des caractéristiques. Un réseau de neurones convolutifs (CNN) est initialement formé pour extraire les caractéristiques locales, qui représentent les caractéristiques de l'écriture manuscrites dans les images des documents manuscrites entières et leurs sous-régions. Les images des documents manuscrites à partir base de données d'apprentissage sont utilisées pour trainer le CNN et agréger l'extraction des caractéristiques locales des images pour former des caractéristiques globales. Le processus d'échantillonnage est répété à chaque époque de formation, ce qui équivaut à un grand nombre de modèles de formation en cours de préparation pour former le CNN à l'identification de scripteurs. Nous avons mené des expériences sur une partie de la base de données ICFHR2018, notre méthode a atteint une précision de 68.5 % pour classer 80document de 40 écrivains, Globalement, nous avons obtenu de meilleures performances que les meilleurs résultats publiés précédemment, ce qui démontre également l'efficacité de notre méthode l'identification de scripteurs à partir d'un document manuscrit

Les mots clés : l'identification de scripteurs, réseau de neurones convolutionnelle , base de données ICFHR2018.

Abstract

Previous research on writer identification has been based on identifying writer-specific features designed by experts. However, in the last decade, deep learning methods have been successfully applied to automatically learn characteristics from data. We propose here a deep-learning method for identifying the writer that does not require prior identification of features. A convolutional neural network (CNN) is initially formed to extract local characteristics, which represent the characteristics of individual writing in images of whole handwritten documents and their sub-regions. The images of the handwritten documents from the training database are used to train the CNN and aggregate the extraction of the local characteristics of the images to form global characteristics. The sampling process is repeated at each training epoch, which represent a large number of training models being prepared to train the CNN for writer identification. We conducted experiments on a part of the ICFHR2018 database, our method reached an accuracy of 68.5% to classify 80 documents from 40 writers, Overall, we have performed better than the best results published previously, which also demonstrates the effectiveness of our method writer identification from a handwritten document.

Keywords: writer identification , convolutional neural network (CNN) , ICFHR2018 database.

ملخص

استند البحث السابق في تحديد هوية الكاتب إلى تحديد الميزات الخاصة بالكاتب التي صممها الخبراء. ومع ذلك ، في العقد الماضي ، تم تطبيق أساليب التعلم العميق بنجاح لتعلم خصائص البيانات تلقائيًا. نقترح هنا طريقة للتعلم العميق لتحديد الكاتب الذي لا يتطلب تحديدًا مسبقًا للميزات. يتم تكوين شبكة عصبية تلافيفية (CNN) في البداية لاستخراج الخصائص المحلية ، والتي تمثل خصائص الكتابة الفردية في صور من وثائق مكتوبة بخط اليد بالكامل ومناطقها الفرعية. تُستخدم صور المستندات المكتوبة بخط اليد من قاعدة بيانات التدريب لتدريب شبكة CNN وتجميع استخراج الخصائص المحلية للصور لتشكيل خصائص عالمية. تتكرر عملية أخذ العينات في كل فترة تدريب ، والتي تمثل عددًا كبيرًا من نماذج التدريب التي يتم إعدادها لتدريب شبكة CNN للتعرف على الكاتب. لقد أجرينا تجارب على جزء من قاعدة بيانات ICFHR2018 ، وصلت دقتنا إلى 68.5 ٪ لتصنيف 80 وثيقة من 40 كاتبًا ، وعموماً ، لقد حققنا أداء أفضل من أفضل النتائج المنشورة سابقاً ، مما يدل أيضاً على فعالية تحديد طريقة الكاتب لدينا من وثيقة مكتوبة بخط اليد.

الكلمات المفتاحية: تحديد هوية الكاتب ، الشبكة العصبية التلافيفية (CNN) ، قاعدة بيانات ICFHR2018

Remerciements

En guise de reconnaissance, je tiens à témoigner mes sincères remerciements à toutes les personnes qui ont contribué de près ou de loin au bon déroulement de mon stage de fin d'étude et à l'élaboration de ce modeste travail.

Mes sincères gratitude à M.abdejalil gattal et M.faiçel abbas pour la qualité de leurs enseignements, leur intérêt et leurs conseils incontestables qu'il porte à tous les étudiants. Dans l'impossibilité de citer tous les noms, nos sincères remerciements vont à tous ceux et celles, qui de près ou de loin, ont permis par leurs conseils et leurs compétences la réalisation de ce mémoire

Enfin, je n'oserais oublier de remercier tout le corps professoral de , pour le travail énorme qu'il effectue pour nous créer les conditions les plus favorables pour le déroulement de nos études.

Didicace

Je dédie ce travail qui n'aura jamais pu voir le jour sans les soutiens indéfectibles et sans limite de mes chers parents qui ne cessent de me donner avec amour le nécessaire pour que je puisse arriver à ce que je suis aujourd'hui. Que dieux vous protège et que la réussite soit toujours à ma portée pour que je puisse vous combler de bonheur.

Je dédie aussi ce travail à :

- ✚ Mes grands-parents.***
- ✚ Mes frères, mes sœurs .***
- ✚ Mon future mari et sa famille .***
- ✚ Mes oncles, mes tantes et leur famille.***
- ✚ Tous mes cousins et cousines.***
- ✚ Tous mes amis, mes collègues et tous ceux qui m'estiment.***

Table des matières

I- Introduction général	1
II- Chapitre I: introduction à identification de document manuscrits	3
II.1- introduction	3
II.2 - Différents aspects du système d'identification.....	3
II.2.1- système d'identification en-ligne et hors-ligne.....	3
II.2.1.1- système d'identification En-Ligne.....	3
II.2.1.2- système d'identification hors-ligne.....	3
II.3- Processus d'un système d'identification des écritures manuscrits.....	4
II.3.1- Phase acquisition.....	5
II.3.2 -Phase de prétraitement.....	5
II.3.2.1- La binarisation.....	5
II.3.2.2- Squelettisation	6
II.3.2.3- Elimination de bruit	7
II. 3.2.4- Suppression de bordure	7
II.3.2.5- Lissage	8
II.3.3- Phase d'extraction des caractéristiques.....	8
II. 3.3.1- Les types de caractéristiques.....	8
II.3.3.1.1- Caractéristiques structurelles.....	8
II.3.3.1.2- Caractéristiques statistiques.....	8
II.3.3.1.3- Caractéristiques globales	8
II.3.3.1.4- Caractéristiques topologiques or métriques.....	8
II.3.3.1.5- Caractéristiques locales	8
II.3.4- Phase de classification.....	9
II.3.4.1 Classification non supervise.....	9
II.3.4.2 classification supervise	9
II.4-conclusion.....	10
III- Chapitre 2: L'apprentissage en profondeur (deep learning)	11

III.1- Introduction.....	11
III.2- Réseaux de neurons artificiels.....	11
III.2.1- Perceptron.....	11
III.2.2- Perceptron Multi Couches.....	11
III.3- Apprentissage en profondeur.....	12
III.4- Architecture de L'apprentissage en profondeur	12
III.4.1- CNN (convolutional neural network).....	13
III.4.1.1- définition.....	13
III.4.1.2- Architecture de Réseaux de neurons convolutifs.....	14
III.4.2- Recurrent Neural Network.....	15
III.4.3- LSTM.....	16
III.4.4- Hopfield.....	16
III.5- Conclusion.....	16
IV- Chapitre 03: Etat de l'art.....	17
IV.1- introduction.....	17
IV.2- Les travaux uulteriers	17
IV.2.1- Travaux de Durou et al.....	17
IV.2.2- Travaux Keglevic et al.....	17
IV.2.3- Travaux Nguyen et al.....	18
IV.2.4- Travaux Abdul Hassan et al.....	18
IV.2.5- Travaux Siddiqi et al.....	19
IV.2.6- Travaux de Sheng He et al.....	19
IV.2.7 - Travaux de Hannad et al.....	19
IV.2.8 - Travaux de Tang et al.....	19
IV.2.9 - Travaux de Khalifa et al.....	20
IV. 2.10 - Travaux de Djeddi et al.....	20
IV.3- Comparaison des travaux.....	21
IV.4- Compétitions sur d'identification de rédacteur.....	22
IV.4.1- ICFHR2016 Competition on Multi-script Writer Demographics Classification Using "QUWI" Database.....	22
IV.4.1.1 - Les résultats de compétitions ICFHR 2016	24

IV. 4.2- ICFHR 2018 Competition on Multi-script Writer Identification	25
IV.4.2.1- Le résultat de la competition ICFHR 2018.....	26
IV.5- conclusion.....	29
V- Chapitre 04: rasultats expérimentaux	30
V.1- Introduction.....	30
V.2- Presentation des outils de développement.....	30
V.2.1- Matériel.....	30
V.2.2- Google colab.....	30
V.2.3- Pythone.....	30
V.3- La base de données.....	30
V.4- La phase de prétraitement	31
V.4.1- Normalization des images	31
V.5- Résultat et décision	35
V.6- Conclusion.....	35
VI- Conclusion général.....	37
VII- Annex.....	38
Référence	

Liste des figures

Figure 1: Exemple d'une écriture hors-ligne.....	4
Figure 2: Les étapes de processus d'identification des écritures manuscrits.....	4
Figure 3: Image en niveau de Gris.....	5
Figure 4: Image binaire.....	6
Figure 5: image Squelettisation.....	6
Figure 6: de application de technique de suppression de bordure.....	7
Figure 7: perceptron multi couches.....	12
Figure 8: le principe de la technique de Max pooling d'image.....	14
Figure 9: Les réseaux de neurone convolutif.....	15
Figure 10: Échantillons manuscrits arabe et anglaise de la base de données QUWI.....	24
Figure11: document chinois.....	25
Figure12: document anglais	25
Figure 13: exemple image de test originale	32
Figure 14: : extraction des mots connexe en utilisant la méthode détection de contours	32
Figure 15: extraction des mots connexe en utilisant la méthode détection de contours	33
Figure 16 : image normalisé (100*100).....	33
Figure17: structure de CNN	34
Figure 18: création de la base.....	38
Figure 19: importer les modules utilisés.....	38
Figure 20: afficher l'image existe un BDD.....	39
Figure 21: déclaration les variable.....	39
Figure 22: structure de CNN.....	40
Figure 23: description de CNN.....	40
Figure 24: loss fonction.....	41
Figure 25: train on minibatches.....	42

Liste des tableaux

Tableau 1: principaux travaux de recherche dans le domaine d'identification de document manuscrits.....	21
Tableau 2 : résultat de compétition ICFHR 2016 pour d'identification de rédacteur.....	24
Tableau 3: taux d'identification et classement de système participants aux tache 1 et 2 (BASE DE DONNÉES CERUG).....	27
Tableau 4: taux d'identification et classement de système participants aux tache 3 et 4 (BASE DE DONNÉES LAMIS-MSHD).....	27
TABLE 5: taux d'identification et classement de système participants aux tache 5 et 6 (BASE DE DONNÉES WDAD).....	28
Tableau 6 : grades globaux de système participants aux six taches	28
Tableau 7: description de la base de données	31
Tableaux 8: les taux d'identification.....	35

Introduction générale

I. Introduction générale

Le problème d'identification de scripteur d'un document manuscrit (modernes ou anciens, ou paléographie) est un domaine de recherche actif au cours des dernières années et a des applications potentielles, telles que la reconnaissance de l'écriture manuscrite, l'identification automatique de scripteur, l'examen de documents et l'analyse de documents.

Dans ce contexte, un grand nombre de scripts ont été examinés dans la littérature, y compris des textes manuscrits à script unique (arabe, latin, chinois, ...). Plus récemment, la recherche sur l'identification de scripteur a été étendue à un environnement multi-script. Dans ce travail, nous concentrons le problème de l'attribution d'un scripteur à partir d'un ensemble de documents en comparant le document de requête avec des documents d'attribution connus (modèles de référence).

Ensuite, nous utilisons dans ce travail des manuscrits modernes, et des documents manuscrits à manuscrits multi-scripts.

L'apprentissage en profondeur est une partie des algorithmes d'apprentissage automatique et donc classé dans une section plus large de l'intelligence artificielle. L'apprentissage en profondeur de diverses architectures concerne les réseaux de neurones profonds, les réseaux de neurones récurrents, les réseaux de croyances profondes et ceux-ci ont été appliqués à divers domaines du monde de l'informatique tels que la reconnaissance vocale, la traduction automatique, le traitement du langage naturel, le filtrage des réseaux sociaux, la bioinformatique et la conception de médicaments. Le réseau de neurones convolutifs est l'un des formes les plus importants d'apprentissage en profondeur. Réseau de neurones convolutifs CNN traite plusieurs couches d'un réseau de neurones simple et constitue donc l'algorithme le plus important pour classer des images ou des documents manuscrits. Les réseaux neuronaux en profondeur constituent le domaine émergent qui est plus efficace avec les GPU (unité de traitement graphique) car il nécessite une énorme quantité de données à traiter et consomme moins de temps de calcul.

Le travail présenté dans cette thèse est organisé comme suit:

Dans le chapitre 1, nous présentons les étapes constituant le système d'identification de l'écriture manuscrite. Et on montrera les différentes caractéristiques sur les écritures manuscrites et on terminera avec le processus de ce système.

Dans le chapitre 2, consacré aux différentes de l'apprentissage en profondeur en mentionnant quelques types de l'apprentissage en profondeur.

Introduction générale

Dans Le chapitre 3, consacré à la description de l'état de l'art dans le domaine de l'identification d'écrivain à partir d'un document manuscrit.

Dans Le chapitre 4, on présente les résultats obtenus pour la validation du système.

I1.1-Introduction

La classification de l'écriture manuscrite est une tâche assez difficile. Cette difficulté réside dans le fait que le système doit être capable d'identifier des écritures manuscrites dans son ensemble, sans aucune segmentation préalable de l'image dans ces éléments. Afin d'obtenir une transcription précise de celles-ci, différentes méthodes et techniques du domaine, d'identification de documents manuscrits sont nécessaires. Parmi ces domaines, de l'écriture manuscrite est relativement peu exploré avec seulement quelques contributions significatives.

Dans ce chapitre, on va parler de différents aspects d'un système d'identification des écritures manuscrites, et puis on montrera les différentes caractéristiques sur les écritures manuscrites et on terminera avec le processus de ce système.

II. 2-Différents aspects du système d'identification

Pour l'identification d'un scripteur à partir son écriture, le système existant s'appuient sur deux différents modes, en ligne et hors ligne chaque mode ayant ses propres outils, concepts et méthodologies, On va décrire brièvement ces deux aspects

II.2.1- système d'identification en-ligne et hors-ligne

Ayant chacun ses outils propres d'acquisition et ses algorithmes correspondants d'identification.

II. 2.1.1- système d'identification En-Ligne

Ce mode s'appelle aussi dynamique, système d'identification en-ligne est effectuée en temps réel, c'est-à-dire elle est effectuée pendant le traçage de caractère.

Les moyens de saisie en ligne sont nombreux où la tablette graphique avec un stylo électronique et l'écran tactile sont couramment utilisés.

II.2.1.2- système d'identification hors-ligne

L'écriture hors ligne est obtenue par la saisie d'un texte déjà existant, obtenue par un scanner ou une caméra. Dans ce cas, on dispose d'une image binaire ou en niveaux de gris, ayant perdu toute information temporelle sur l'ordre des points, dans notre travail on a concentré sur ce mode. Voir figure 1.

قد يبتدأ الزلزال كنتيجة لأنشطة البراكين أو نتيجة لوجود إزلاقات في طبقات الأرض. و
تؤدي الزلازل إلى تشقق الأرض ونضوب الينابيع وأظهور الينابيع الجديدة أو
حدوث أمواج عالية لها آثارها التخريرية للمباني والمواصلات والمنشآت.

Figure 1: Exemple d'une écriture hors-ligne

II.3- Processus d'un système d'identification des écritures manuscrites

Les systèmes d'identification des caractères manuscrits sont généralement basés sur les étapes principales suivantes : Acquisition, prétraitements, extraction des caractéristiques, et une phase de classification. Le schéma suivant explique ce processus. Voir figure 02.

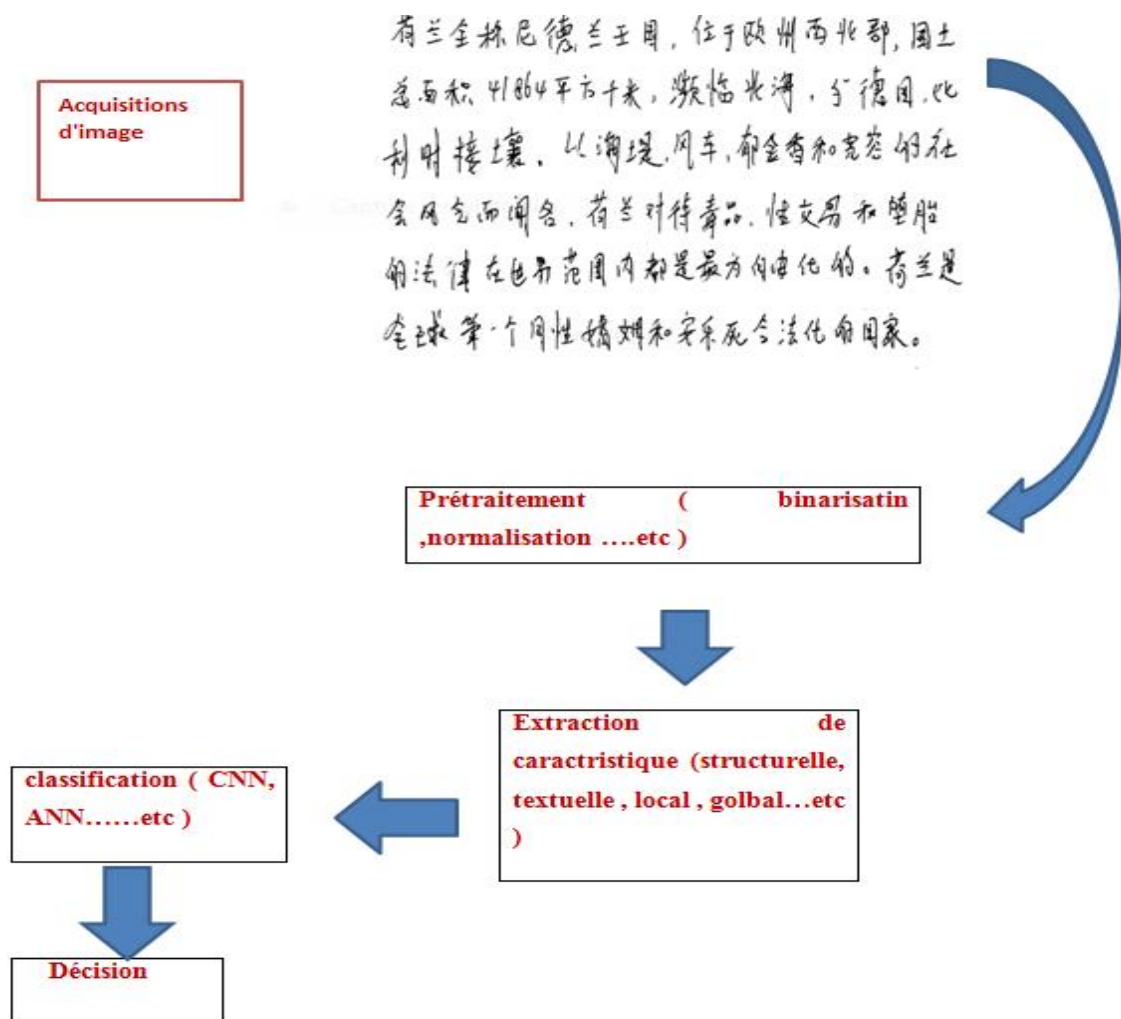


Figure 2 : Les étapes de processus d'identification des écritures manuscrites.

II.3.1- Phase d'acquisition

L'acquisition permettant la conversion du document papier sous la forme d'une image numérique. Cette étape est importante car elle se préoccupe de la préparation des documents à saisir, du choix et du paramétrage du matériel de saisie (scanner), ainsi que du format de stockage des images. L'image sera probablement en niveau de gris ou en couleur.

II.3.2- Phase de prétraitement

Dans l'étape de prétraitement, D'abord la sortie est convertie en forme binaire. Cela peut contenir un peu de clarté ou de bruit. Le bruit peut avoir un effet sur l'exactitude de l'étape de la classification. Opérations de prétraitements généralement utilisées, citons : la binarisation, la squelettisation, le lissage.

II.3.2.1- La binarisation : c'est le passage d'une image en couleur ou définie par plusieurs niveaux de gris en image bitonale (composée de deux valeurs 0 et 1) qui permet une classification entre le fond (image du support papier en blanc) et la forme (traits des gravures et des caractères en noir) .



Figure 3: Image en niveau de gris.



Figure 4 : Image binaire.

II.3.2.2- Squelettisation : est une classe d'algorithmes utilisée en analyse de formes. Elle consiste à réduire une forme en un ensemble de courbes, appelées squelettes, centrées dans la forme d'origine. La squelettisation est un outil d'analyse de forme non-scalaire, qui conserve les propriétés topologiques de la forme d'origine ainsi que les propriétés géométriques, selon la méthode employée.

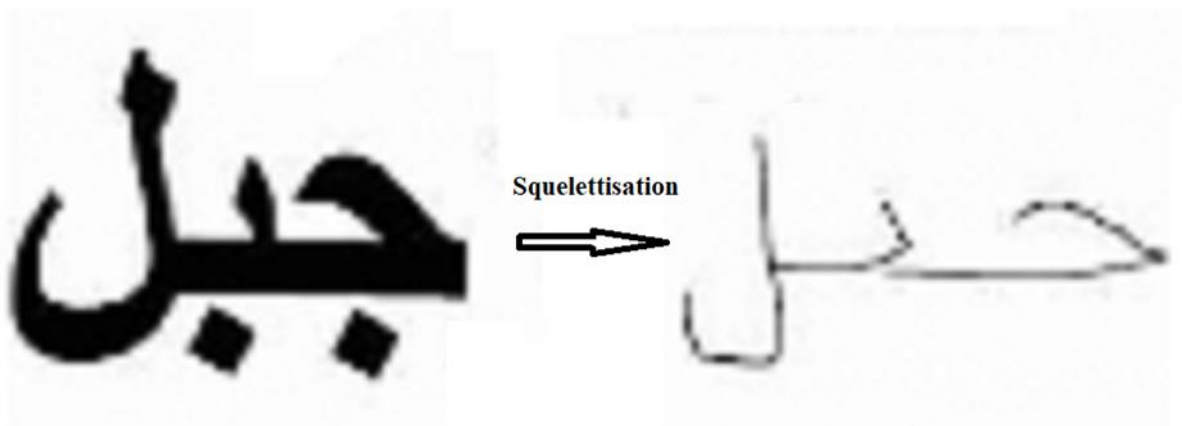


Figure 5 : image squelettisation.

II.3.2.3- Elimination de bruit : Il existe diverses techniques de nettoyage, de lissage, d'aiguisage de caractère. Pour éliminer le bruit. La supériorité des documents manuscrits dépend de nombreux facteurs différents, comme la ligne déconnectée, les trous, les boucles débordant, le vieillissement des papiers, la couleur de l'encre et bien d'autres. Certaines techniques de filtrage et d'opération de dilatation et d'érosion peuvent enlever des bruits dans une certaine mesure. Il est obligatoire de se débarrasser de ces imperfections. La pratique de réduction du bruit disponible peut être classée en deux ensembles majeurs : le filtrage et les opérations morphologiques. Différents filtres spatiaux peuvent être conçus pour cette utilisation comme le filtrage passe-bas, qui est utilisé pour supprimer un bruit à haute fréquence spatiale d'une image numérique. L'idée fondamentale est de convoluer un masque prédéfini avec l'image pour assigner une valeur à un pixel en tant que but des valeurs de gris de ses pixels voisins. Les filtres peuvent être mesurés pour le lissage, le seuillage et l'arrière-plan coloré.

II.3.2.4- Suppression de bordure : Lorsqu'une page d'un livre est numérisée, le texte d'une page adjacente peut également être capturé dans l'image de la page en cours. Ces régions indésirables sont appelées "régions de texte bruyantes (noisy text regions)".chaque fois qu'une page numérisée ne couvre pas complètement la taille de l'image selon la configuration du scanner, il y aura généralement des bordures noires dans l'image. Ces régions indésirables sont appelées " bordures noires bruyantes (noisy black borders) ". La figure montre des bordures noires bruitées ainsi que des régions de texte bruyantes [01].

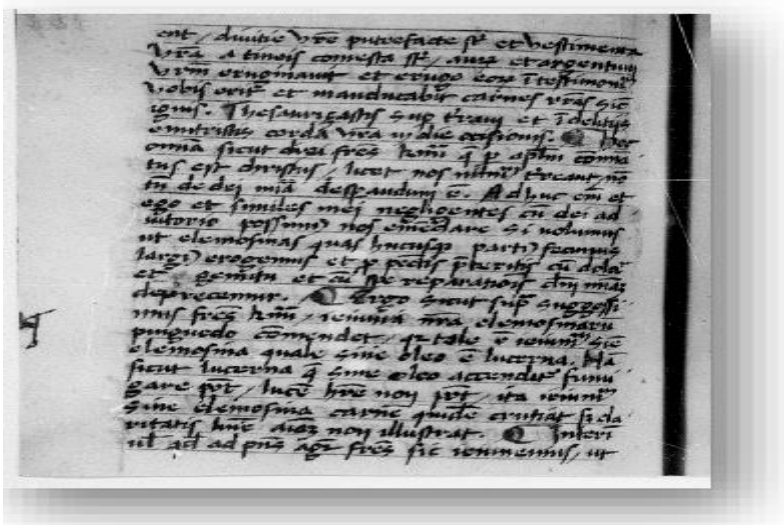


Figure 6: Application de technique de suppression de bordure.

II.3.2.5- Lissage : L'image des caractères peut être entachée de bruits dus aux artefacts de l'acquisition et à la qualité du document, conduisant soit à une absence de points ou à une surcharge de points. Les techniques de lissage permettent de résoudre ces problèmes par des opérations locales qu'on appelle opérations de nettoyage. L'opération de nettoyage permet de supprimer les petites tâches et les excroissances de la forme. Pour le bouchage il s'agit d'égaliser les contours et de boucher les trous internes à la forme du caractère en lui ajoutant des points noirs.

II. 3.4- Phase d'extraction des caractéristiques

L'extraction des caractéristiques peut être définie comme extraire les informations les plus représentatives des données brutes, ce qui minimise la variabilité du modèle de classe interne en l'améliorant avec la variabilité du modèle de classe. Généralement l'ensemble des caractéristiques est extrait pour chaque classe qui permet de la distinguer d'autres classes.

II.3.4.1- Les types de caractéristiques

II.3.4.1.1- Caractéristiques structurelles: Les caractéristiques structurelles décrivent une forme en termes de sa topologie et sa géométrie en donnant ses propriétés globales et locales. Parmi ces caractéristiques on peut citer.

II.3.4.1.2- Caractéristiques statistiques : Les primitives statistiques portant des informations concernant la distribution des pixels dans l'image du caractère, Les primitives statistiques décrivent une forme en termes d'un ensemble de mesures extraites à partir de cette forme [02].

II.3.4.1.3- Caractéristiques globales : Il s'agit des primitives globales basées sur une transformation globale de l'image. La caractéristique d'une primitive globale est de dépendre de la totalité des pixels d'une image, ces primitives sont donc dérivées de la distribution des pixels [03].

II.3.4.1.4- Caractéristiques topologiques ou métriques : Dans Primitives topologiques ou métriques basé essentiellement sur des densités de pixels. Il s'agit d'effectuer une mesure sur l'échantillon au moyen d'une métrique.

II.3.4.1.5- Caractéristiques locales : Ces caractéristiques sont calculées localement, sur des sous-images en utilisant des petites fenêtres positionnées sur l'image, considérées comme

des parties de l'image de texte manuscrit. Par exemple en utilisant un algorithme de positionnement de fenêtre adaptatif [04].

II.3.5- Phase de classification

Un certain nombre de chercheurs qui ont développé différents algorithmes ont largement exploré le domaine de la reconnaissance des formes. Le modèle de conception de tous ces algorithmes repose sur la classification, pour la prise de décision permettant de classer autant que possible un objet en fonction de certains critères d'optimisation, en fonction de l'appartenance à la classe (passage de l'espace cryptographique au domaine décisionnel). Le type de méthode de classification se divise généralement en deux catégories: mode supervisé et statut non supervisé.

a. Classification non supervisé :

Apprendre sans supervision nécessite que le système soit doté d'un mécanisme automatique basé sur les règles d'assemblage exactes pour trouver des classes de référence avec une assistance minimale. Dans ce cas, les échantillons sont présentés en grand nombre par l'utilisateur sans référence à sa classe, il s'agit également d'un algorithme qui découvre la structure des données cachées sous une forme ou une autre et forme des groupes d'individus dont les caractéristiques sont communes.

b. classification supervisé :

Cette technique est basée sur le tri des commentaires en affectant chaque note à une catégorie (sous supervision où la sortie correcte doit être fournie à l'avance). L'expert doit d'abord distinguer les exemples. Le processus d'apprentissage se déroule sous deux phases de supervision. Lors de la première étape, il s'agit de définir un modèle de données nommées. La deuxième étape est la prédiction de la nouvelle carte de données, avec la connaissance du modèle déjà appris. Parfois, il est préférable de lier une partie des données non pas à une classe, mais avec la possibilité d'appartenir à chacune des classes prédéterminées (on parle alors d'apprentissage supervisé).

Il existe plusieurs algorithmes d'apprentissage supervisé, cette section présente quelques-uns des plus connus parmi eux, il s'agit de k plus proche voisin (KNN), les Réseaux de Neurones(ANN), Naive Bayes(NB) et SVM.

II.4-conclusion

Dans ce chapitre, nous avons présenté les différents concepts généraux d'un système de classification d'images de documents manuscrits. Le module d'extraction de caractéristiques et la classification constituent le principal objectif de notre travail. Aussi, nous présentons dans le chapitre suivant, les différentes informations sur apprentissage en profondeur.

III.1- Introduction

L'apprentissage en profondeur est un type d'intelligence artificielle dérivé du machine learning où des différentes techniques ont été développées pour créer des algorithmes capables d'apprendre et de s'améliorer de manière autonome. Parmi ces techniques, on compte les réseaux de neurones artificiels. C'est sur ces algorithmes que reposent l'apprentissage en profondeur, Ce chapitre est consacré les différentes notions de l'apprentissage en profondeur en mentionnant quelque types de l'apprentissage en profondeur.

III.2- Réseaux de neurones artificiels

Les réseaux de neurones artificiels sont des réseaux fortement connectés de processeurs élémentaires fonctionnant en parallèle. Chaque processeur élémentaire calcule une sortie unique sur la base des informations qu'il reçoit. Toute structure hiérarchique de réseaux est évidemment un réseau.

III.2.1- Perceptron

Perceptron est un algorithme d'apprentissage d'un classificateur binaire appelé fonction de seuil, est un classificateur linéaire. En outre, il est utilisé dans l'apprentissage supervisé. Il est utile de classer les données d'entrée données (un vecteur à valeur réelle) en une valeur de sortie, qui fait partie de l'apprentissage en profondeur.

III.2.2- Perceptron Multi Couches

Perceptron Multi Couches(PMC) est une généralisation du modèle de perceptron simple lorsque les variables explicatives sont finies, développé en utilisant des variables comme variables indépendantes, qui une combinaison de séparateurs linéaires permet de produire un séparateur global non linéaire et l'augmentation du nombre de couches et du nombre de neurones accroît le pouvoir de séparation, à la même structure d'un perceptron à une seule couche avec une ou plusieurs couches cachées. L'algorithme de rétro propagation comprend deux phases: la phase avant où les activations sont propagées de la couche d'entrée à la couche de sortie et la phase arrière où l'erreur entre la valeur réelle observée et la valeur nominale demandée dans la couche de sortie est propagée dans l'ordre modifier les poids et les valeurs de biais [05].

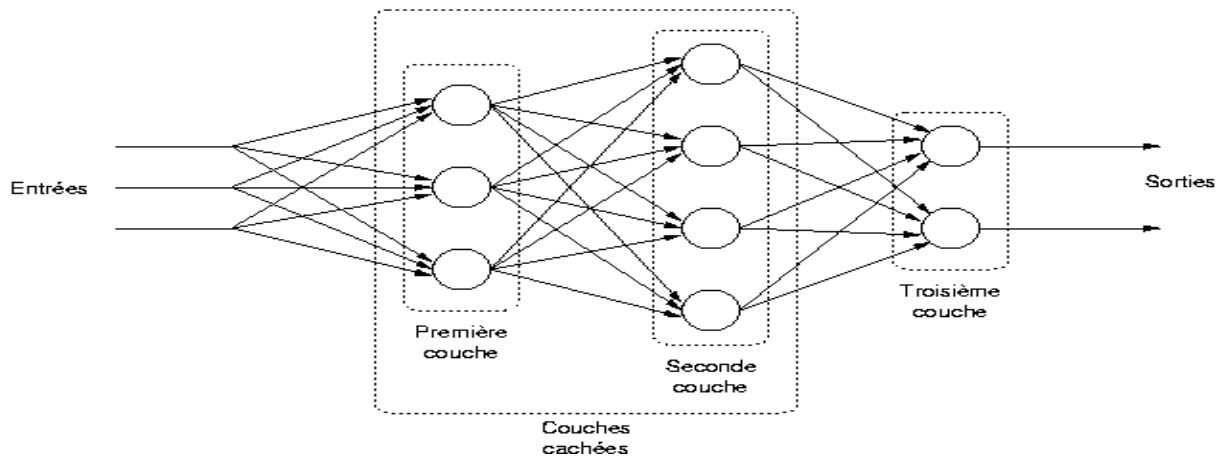


Figure 7 : perceptron multi couche [06].

III.3- Apprentissage en profondeur

L'apprentissage en profondeur est type de l'apprentissage automatique basée sur un ensemble d'algorithmes qui tente de modéliser des données à un haut niveau d'abstraction, en utilisant des traitements multi-couches. Les éléments élémentaires de l'apprentissage en profondeur sont les réseaux de neurones, qui sont combinés pour former les réseaux de neurones profonds.

Ces techniques ont permis des progrès significatifs dans les domaines du traitement du son et des images, notamment la reconnaissance faciale, la reconnaissance vocale, la vision par ordinateur, le traitement automatique du langage, la classification de texte.

Il existe plusieurs types d'architectures apprentissage en profondeur:

- ✓ Les perceptrons multicouches, les plus anciens et les plus simples
- ✓ Les réseaux de neurone convolutif (CNN), particulièrement adaptés au traitement d'images.
- ✓ Les réseaux de neurones récurrents, utilisés pour des données séquentielles telles que du texte ou séries chronologiques

III.4- L'architecture de L'apprentissage en profondeur

Généralement, il existe plusieurs architectures de l'apprentissage comme démontre au-dessous

III.4.1- Réseaux de neurone convolutif (CNN)

Un réseau à convolutifs est un type de réseau neuronal très utilisé pour la classification d'image ou la reconnaissance visuelle. Le but principal d'un réseau à convolution est d'extraire des caractéristiques d'une image donnée.

III.4.1.1- définition

Les réseaux de neurones convolutifs sont un type particulier de réseaux de neurones, inspirés biologiquement du système du cortex visuel humain, qui consistent en un très grand nombre de nœuds interconnectés, ou neurones. L'architecture d'un CNN est conçue pour tirer parti de la structure 2D d'une image d'entrée. Ceci est réalisé avec des connexions locales et des poids liés suivis de certaines formes de regroupement qui aboutissent à des caractéristiques invariantes. Les nœuds du réseau sont organisés en plusieurs couches empilées, chacune effectuant une opération simple sur l'entrée.

L'ensemble d'opérations dans un CNN comprend généralement une convolution, une normalisation d'intensité, une activation et un seuillage non linéaires et une mise en commun locale. En minimisant une fonction de coût à la sortie de la dernière couche, les poids du réseau sont réglés de manière à pouvoir capturer des modèles dans les données d'entrée et à extraire des caractéristiques distinctives.

Dans un CNN, les caractéristiques sont apprises en utilisant la rétropropagation [07] couplée à une méthode d'optimisation telle que la descente de gradient (stochastique)[8] et l'utilisation de grands ensembles de données d'apprentissage . Les couches moins profondes des réseaux apprennent généralement les caractéristiques visuelles de bas niveau telles que les edges, les formes simples et le contraste de couleurs, tandis que les couches plus profondes combinent ces caractéristiques pour identifier des formes visuels complexes. Enfin, les couches

entièrement connectées (fullyconnected) associées à une couche softmax sont couramment utilisés pour générer une étiquette de classe de sortie en minimisant la fonction d'activation

le filtre : Alignant les caractéristique avec chaque sous-partie de l'image ce filtre va être déplacé par pas successifs sur l'ensemble de image pour créer une map indiquant où la caractéristique a été trouvée dans l'image donnée (couleur, contour, luminosité, netteté, etc.....)

Le couche convolution : cette couche permet d'extraire les informations et les caractéristiques globales et locales en utilisant la convolution par filtre.

Le Pooling : est un procédé important dans un réseau à convolution. En extrayant les valeurs importantes des pixels, il permet de réduire une image tout en conservant les caractéristiques pertinentes. La méthode la plus utilisée est le "Max Pooling". Elle consiste à réduire l'image

Chapitre 2:L'apprentissage en profondeur (deep learning).

en conservant les valeurs les plus grandes des pixels. Pour ce faire, on a une tuile qui se déplace (comme un filtre) sur image. À chaque position de la tuile, on extrait la valeur la plus haute et ne retient que celle là. Cela produit une nouvelle image avec uniquement les valeurs remarquable de l'image

L'image ci dessous montre un exemple de Pooling. La tuile a, ici, des dimensions de 2x2.

L'image de 6x6 pixel de depart et reduit en une image de 3x3 pixel.

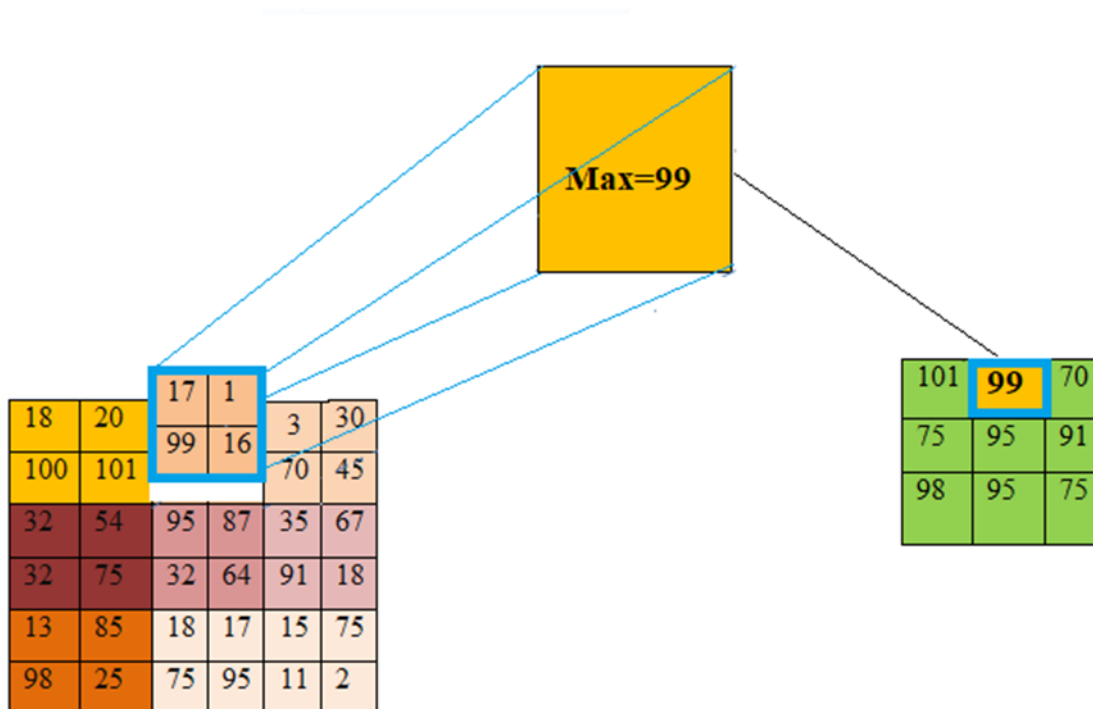


Figure 8: le principe de la technique de Max pooling d'image

(Pas de filter=2 et la tailles=2x2)

III.4.1.2- Architecture de Réseaux de neurones convolutifs

Les réseaux de neurone convolutif sont actuellement les modèles les plus efficaces pour la classification des images, ils ont deux parties distinctes.

En entrée, une image est fournie sous la forme d'une matrice de pixels. Il a 2 dimensions pour une image en niveaux de gris, la couleur est représentée par une troisième dimension, de profondeur 3 pour représenter les couleurs fondamentales [rouge, vert, bleu].

La première partie d'un CNN est la partie convoluté réelle, il fonctionne comme le model d'extraction de caractristique d'images, une image passe par une succession de filtres ou de noyaux de convolutifs, créant de nouvelles images appelées cartes de convolution. Certains filtres intermédiaires réduisent la résolution de l'image d'un maximum local.

Finalement, les cartes de convolution sont posées à plat et concaténées en un vecteur caractéristique appelé code CNN. Ce code CNN à la sortie de la partie convolutive est ensuite connecté à l'entrée d'une deuxième partie, constituée de couches entièrement connectées .Le rôle de cette partie est de combiner les caractéristiques du code CNN pour classifier l'image.

La sortie est une dernière couche avec un neurone par catégorie. Les valeurs numériques obtenues sont généralement normalisées entre 0 et 1, de la somme 1, pour produire une distribution de probabilité sur les catégories.

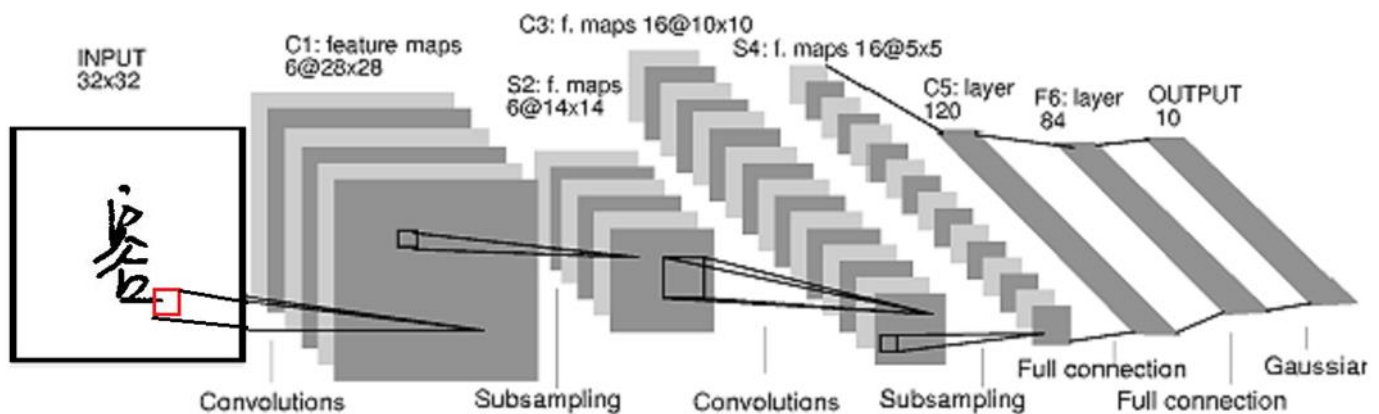


Figure 9: Les réseaux de neurone convolutif (Image Originale publier en [09]).

III.4.2- Recurrent Neural Network

Un réseau de neurones récurrent ressemble est un réseau hiérarchique où il n'y a vraiment aucun aspect temporel dans la séquence d'entrée mais où l'entrée doit être traitée de manière hiérarchique sous forme d'arborescence, il montre comment apprendre l'arbre d'analyse d'une phrase en prenant de manière récursive le résultat de l'opération effectuée sur une plus petite partie du texte, l'architecture de réseau neuronal récurrent est composée d'une matrice à poids partagés et d'une structure arborescente binaire qui permet au réseau récurrent.

Chapitre 2:L'apprentissage en profondeur (deep learning).

Il est utile comme analyseur de phrases et de scènes. Les réseaux de neurones récurrents utilisent une variation de la rétro-propagation[10].

III.4.3- LSTM (long short_term memory)

Les LSTM représentent un modèle qui possède une représentation interne et explicite du temps. Les LSTM ont été proposés initialement par Hochreiter et Schmidhuber [1997]. L'inspiration à leur origine était le souhait de trouver un nouveau modèle permettant de pallier les faiblesses des réseaux de neurones récurrents traditionnels. Les auteurs ont proposé un nouveau système avec une unité complexe, comparable à un automate, dotée d'une circuiterie interne reliant et exploitant des éléments simples, à savoir des neurones artificiels utilisant la rétro propagation pour leur apprentissage. Le modèle LSTM se distingue de ses prédécesseurs par une nouveauté : les cellules mémoires. Ces unités spéciales sont utilisées, en plus d'unités standards, au niveau de la couche cachée d'un réseau de neurones, afin de permettre le maintien d'information en mémoire pour de longues périodes de temps. Ce sont ces cellules mémoires qui permettent de garder une représentation explicite du passé [11].

III.4.5- réseaux de hopfield

Le réseau de neurones d'Hopfield est un modèle de réseau de neurones récurrents à temps discret dont la matrice des connexions est symétrique et nulle sur la diagonale et où la dynamique est asynchrone (un seul neurone est mis à jour à chaque unité de temps). Il a été découvert par le physicien John Hopfield en 1982. Sa découverte a permis de relancer l'intérêt dans les réseaux de neurones qui s'était essoufflé durant les années 1970 à la suite d'un article de Marvin Minsky et Seymour Papert.

Un réseau de Hopfield est une mémoire adressable par son contenu : une forme mémorisée est retrouvée par une stabilisation du réseau, s'il a été stimulé par une partie adéquate de cette forme.

III.5- Conclusion

Dans ce chapitre on a présenté les notions importantes qui sont en relation avec l'apprentissage profondeur (définition, Architectures....etc.). Aussi qu'une vision générale sur l'apprentissage profondeur, toute on donnant en détail la méthode choisie dans notre travail de recherche qui est le CNN, et ANN, Le prochain chapitre, traite l'état de l'art.

IV.1- introduction

Ces dernières années, de nombreuses méthodes ont été créées, développées ou améliorées pour la classification automatique des documents manuscrits. Ce chapitre est consacré à la description de l'état de l'art dans le domaine de l'identification du type d'écrivain à partir d'un document manuscrit. Dans les premières sections, nous présentons les principaux travaux de recherche dans ce domaine en les classant par ordre chronologique. Applicable à la classification automatique de documents manuscrite. Nous concluons chaque section en comparant le champ, tel que les critères de comparaison, la base de données utilisée.

IV.2- Les travaux ultérieurs

Dans cette section, nous allons présenter les travaux l'identification du type d'écrivain en fonction de l'approche utilisée.

IV. 2.1- Travaux de Durou et al :

Dans les travaux de Durou et al [12] propose de combiner les deux approches l'approches statistiques et basée sur des modèles en utilisant les caractéristique Oriented Basic Image features (OBIF) et le concept de codebook de graphèmes. Pour réduire la haute dimensionnalité résultante du vecteur de caractéristiques, une analyse en composantes principales(ACP) a été utilisée. Pour évaluer l'efficacité de la méthode proposée, une analyse de performance, utilisant la base de données IAM qui contient des documents manuscrits anglaise et la base de données ICFHR 2012 pour des manuscrits arabe, a été réalisée. Les résultats obtenus ont atteint une précision de 96%, démontrant ainsi sa supériorité par rapport à des techniques similaires.

IV. 2.2- Travaux Keglevic et al :

Keglevic et al. 2016 [13] présente une méthode de Trouver et d'identification de rédacteur de document utilisant un descripteur de caractéristique appris par un réseau de neurones de convolution(CNN). Au lieu d'utiliser un réseau pour la classification, ils proposent l'utilisation d'un réseau de triplets qui apprend une mesure de similarité pour les patches d'image. Les patches de l'écriture manuscrite sont extraits et mappés où cette mesure de similarité est définie par la distance L 2. Le réseau de triplets est formé en maximisant la distance interclasse, tout en minimisant la distance intraclasse dans cette intégration. Les patches d'image sont codés à l'aide du descripteur de caractéristique appris. En appliquant le codage de vecteur de descripteurs

agrégés localement (Vector of Locally Aggregated) à ces caractéristiques, nous générons un vecteur de caractéristique pour chaque image de document. Une évaluation détaillée des paramètres est fournie, qui montre que cette méthode permet d'obtenir une précision moyenne de 86,1% sur la base de données d'identification d'écrivain ICDAR 2013,.

IV.2.3- Travaux Nguyen et al :

Nguyen et al 2018 [14] proposent une méthode base sur réseau de neurones convolutifs (CNN) qui initialement formé pour extraire les caractéristiques locales, qui représentent les caractéristiques de l'écriture individuelle dans les images entières et leurs sous-régions. Les petites images prise de manière aléatoire à partir de la base d'apprentissage sont utilisées pour former le CNN et agréger les caractéristiques locales extraites des images des n-uplets pour former des caractéristiques globales. Le processus d'échantillonnage aléatoire est répété à chaque époque d'apprentissage, ce qui correspond à un grand nombre de modèles pour former le CNN à l'identification. Ils effectués des expériences sur la base de données JEITA-HP de modèles de manuscrits japonais hors ligne. Avec 200 caractères, cette méthode a atteint une précision de 99,97% pour classer 100 écrivains. D'autres expériences faite sur les bases de données Firemaker et IAM de texte anglais hors ligne. La méthode a permis d'obtenir une précision de 91,81% lors de la classification de 900 rédacteurs. Dans l'ensemble, cette approche obtenue des meilleurs résultats que résultats publiés précédemment, sur la base de caractéristiques manuelles et d'algorithmes de classification, ce qui démontre L'efficacité de méthode pour le manuscrit anglais également.

IV.2.4- Travaux Abdul Hassan et al :

Abdul Hassan et al 2019 [15] proposent une méthode d'identification manuscrite d'un mot arabe sans segmentation en sous-lettres basée sur l'extraction de caractéristiques, permettant d'accélérer la transformation robuste des entités (SURF) et la classification KNN (voisin le plus proche) afin d'améliorer la précision de l'identification de l'écrivain.

Une fois l'extraction de caractéristiques pouvant être groupée par l'algorithme Kmeans afin de normaliser le nombre de caractéristiques, l'extraction de caractéristiques et la mise en grappes de caractéristiques appelée pour rassembler Bag of Word (BOW) permettent de convertir un nombre arbitraire de caractéristiques d'image en unités de longueur uniforme de vecteur , la méthode proposée expérimentée à l'aide de la base de données (IFN / ENIT). Le résultat de l'expérience est (96,666) le taux de reconnaissance.

IV.2.5- Travaux Siddiqi et al :

Siddiqi et al. 2014 [16]. Proposé des caractéristiques d'écriture suggérées telles que l'inclinaison, la flexion et la lecture ont été suggérées en calculant les caractéristiques locales et globales afin de distinguer les écritures masculine et féminine, et les caractéristiques d'inclinaison et de flexion étaient également extraites des informations fournies par l'auteur local concernant l'orientation et la forme. Pour la classification, ces propriétés sont utilisées dans trois catégories: SVM, ANN et CNN ont une évaluation d'environ 68% basée sur les données QUWI.

IV.2.6- Travaux de Hannad et al.

Hannad et al. 2015 [17] la méthode proposée présente une nouvelle approche pour l'identification des écrivains indépendants du texte hors ligne à l'aide d'images manuscrites en arabe. Exploitant l'idée que des fragments graphiques de l'écriture manuscrite caractérisent l'écrivain, ils proposent une approche locale basée sur l'analyse de la texture de petits fragments d'écriture, chaque fragment étant représenté par son histogramme LBP (Local Binary Pattern). La méthode proposée tire parti de l'efficacité de la LBP en tant que descripteur de texture et du pouvoir discriminant élevé des fragments manuscrits pour améliorer les performances d'identification de l'écrivain. La technique proposée évaluée sur une base de données IFN/ENIT de 130 auteurs permet d'obtenir des taux d'identification prometteurs avec un temps d'exécution réduit.

IV.2.7 - Travaux de Tang et al.

Tang et al 2013. [18] ont proposé l'utilisation de document manuscrits car les méthodes basées sur les textures nécessitent un nombre important de texte manuscrits pour obtenir des fonctionnalistes stables. De plus, lors de l'utilisation d'une méthode basées sur les caractéristiques de texture sur ensemble du document, les régions non textuelles génèrent des caractéristiques instable causes par les régions vides d'arrière-plan. la distance chi-square utiliser pour la mesure des similarités entre les fragments de référence et de test, Ces caractéristiques sont utilisées pour former les classifies (K-Means) et (k-Centers), la performance du système proposé a été évaluée sur deux bases de données, IAM et IFN/ENIT.

IV.2.8 - Travaux de Khalifa et al:

Khalifa et al. 2015 [19] ont proposé une méthode d'identification d'écrivain utilisant un modèle d'extension codebook et un ensemble de codebook dans lesquels une analyse discriminante du noyau utilisant une régression spectral (SR-KDA) a été déployée en tant que technique de réduction de la dimensionnalités afin d'éviter tout problème de sur adaptation. Deux jeux de données ont été utilisés en évaluation avec un seul codebook et en utilisant plusieurs tailles de codebook .En outre, les auteurs concluent que la méthode proposée est compétente par rapport aux méthodes existentes. Les bases de données IAM (anglais) et ICFHR-2012 (arabe) ont été utilisées pour l'évaluation,

IV.2.9 - Travaux de Djeddi et al:

Djeddi et al. 2014 [20] a proposé un système d'identification basé sur l'écriture manuscrite pour les documents manuscrits en arabes. La méthode proposée comprend deux étapes principales: premièrement, le système traite chaque image manuscrite et extrait des caractéristiques telles que la direction de bord, la charnière de bord et la longueur. Ensuite, ces fonctionnalités sont transmises à une SVM Multi-class en vue de leur classification. La méthode a été formée et testée sur une vaste base de données des documents arabes rédigés par 1 000 écrivains différents. Les auteurs ont indiqué que le meilleur résultat avait été obtenu en combinant des caractéristiques de longueur de trait et de la charnière atteignant un taux de classification de 84,10%.

Ces caractéristiques sont utilisées pour former et classifier (SVM), la performance du système proposé a été évaluée sur la base de données KHATT.

IV.2.10 - Travaux de Sheng He et al:

Sheng He et al 2019 [21] Présentent une méthode d'apprentissage adaptatif en profondeur pour l'identification du rédacteur, basée sur des images d'un seul mot utilisant un apprentissage multitâche. Une tâche auxiliaire est ajoutée au processus de d'apprentissage pour renforcer l'émergence de caractéristiques. la méthode proposée transfère les avantages des caractéristiques d'apprentissage acquises d'un réseau de neurones convolutionnels (CNN) d'une tâche auxiliaire telle que la reconnaissance explicite de contenu à la tâche principale d'identification d'un scripteur au cours d'une procédure unique. ils proposent une nouvelle couche convolutionnelle adaptative pour exploiter les caractéristiques profondes apprises. Un réseau neuronal multitâche avec une ou plusieurs couches convolutives adaptatives est apprenant de bout

en bout, afin d'exploiter des caractéristiques génériques robustes pour une tâche principale spécifique, à savoir l'identification de l'écrivain. Trois tâches auxiliaires, correspondant à trois attributs explicites d'images de mots manuscrites (contenu lexical, longueur de mot et attributs de caractère), sont évaluées. Les résultats expérimentaux de deux ensembles de données CVL et IAM montrent que la méthode d'apprentissage adaptatif approfondie proposée peut améliorer les performances d'identification de rédacteur basée sur des images d'un mot, par rapport aux approches non adaptatives et linéaires adaptatives.

IV.3- Comparaison des travaux:

Le tableau au-dessus, illustre la base de données utilisé, les caractéristiques et les performances pour chaque travail :

Travaux	BDD	Caractéristiques	Classification
Durou et al [12]	IAM (Anglais)	OBIF + codebook	KNN+ SVM
	ICFHR 2012(Arabic		
Keglevic et al. 2016 [13]	ICDAR 2013	Vector of Locally Aggregated	CNN
Nguyen et al 2018 [14]	JEITA-HP(japonais)	Local et global	CNN
	Firemaker et IAM (anglais)		
Abdul Hassan et al 2019 [15]	IFN / ENIT	BOW	KNN + k_means
Sidiqqi & al 2016 [16]	QUWI (arabe)	Locale et global	CNN+ANN
	QUWI (anglais)		
Sheng He et Al 219 [17]	CVL+IAM	caractéristiques générique et profonds	CNN
Hannad & al 2015 [18]	INF/ENIT (arab)	LBP+LPQ+LTP	SVM+KNN
Tang & al 2013 [19]	INF/ENIT + IAM	Texteur +fragmentation	K_means+ K_centers

Khalifa & al 2015 [20]	ICFHR 2012 (Arabe)	SIFT+OBI	SVM
	IAM (anglais)	SIFT+OBI	
Djeddi & al 2014 [21]	KHATT.	direction de bord + charnière de bord et la longueur +	SVM

Tableau 1: principaux travaux de recherche dans le domaine d'identification de document manuscrits.

IV.4- Compétitions sur d'identification de scripteur :

L'intérêt des chercheurs à l'identification de scripteur et aux méthodes d'évaluation des systèmes de classification a évolué massivement les dernières années. En raison de l'importance du domaine, les chercheurs de différentes équipes ont organisé plusieurs compétitions dans le cadre des conférences spécialisées telles que ICFHR (International Conference on Frontiers in Handwriting Recognition) et visait à rendre compte et à comparer objectivement les techniques les plus récentes sur le problème l'identification de rédacteur dans les mêmes environnements expérimentaux parmi ces compétitions on cite ICFHR2016 et ICFHR 2018 (objectif de notre expérimentation).

IV.4.1 ICFHR2016 Competition on Multi-script Writer Demographics Classification Using "QUWI" Database

La série d'évaluations de la compétition a été réalisée sur un sous-ensemble de la base de données QUWI. La base de données comprend 4 068 échantillons d'écriture fournis par un total de 1 017 personnes différentes. Chaque auteur a contribué d'écrire quatre pages, deux en arabe et deux en anglais. Une page de chaque auteur comprend le même texte pour tous les auteurs en arabe (anglais), tandis que les autres pages contiennent des textes arbitraires en arabe et en anglais. la tâche utiliser pour d'identification du rédacteur (1A - 1D) est décrite comme suite :

Des échantillons d'écriture de 800 écrivains de la base de données QUWI ont été utilisés. Pour chaque sous-tâche, 400 échantillons d'écrivains en anglais (arabe) ont été fournis en tant qu'ensemble d'apprentissage, 100 échantillons d'écriture utilisée pour la validation.

Quatre systèmes ont été soumis à cette compétition [22] :

Méthode de laboratoire MISC: La méthode se repose principalement sur une approche locale basée sur l'analyse de la texture de petits fragments d'écriture. Chaque fragment d'écriture extrait est représenté par son histogramme de Quantification de Phase Locale (LPQ) et est utilisé comme caractéristique dans l'étape d'identification d'écrivain.

Méthodes de Nuremberg: le système soumis repose sur le calcul des moments de Zernike (Zernike moments) aux contours de l'image de documents manuscrits. Ces descripteurs locaux sont ensuite codés dans des vecteurs de descripteurs agrégés localement (VLAD) et tous les VLAD sont normalisés. Pour la tâche d'identification d'écrivain, les descripteurs extraits sont utilisés pour former les SVM.

Méthode CVC: La méthode est basée sur une variante des modèles binaires locaux développés spécifiquement pour le contenu textuel appelé SRS-LBP. Les SRS-LBP sont calculés pour plusieurs rayons et leurs histogrammes sont concaténés dans un vecteur de caractéristiques. Pour l'identification du scripteur, une analyse en composantes principales (ACP) est utilisée pour la réduction de la dimensionnalité et l'extraction des caractéristiques pertinentes pour la caractérisation du scripteur.

Méthode CSMC: le système d'identification du scripteur soumis exploite un algorithme de correspondance basé sur le fait que les probabilités de classe conditionnelles peuvent être approximées de manière approximative par la distance euclidienne à l'entité la plus proche appartenant à la classe correcte ; sur la base de ce fait, une image à classe au lieu d'une distance image à image est calculée. Seules les classes situées à l'intérieur d'un certain voisinage du descripteur de requête dans l'espace des caractéristiques sont considérées. La méthode soumise prend en compte la particularité des modèles écrits à la main en ajoutant des contraintes pour empêcher la correspondance de points-clés non pertinents en termes d'échelle et de rotation. L'implémentation de la méthode est basée sur des recherches au plus proche voisin, ce qui peut être effectué efficacement à l'aide de la mise en œuvre d'arbres K-D fournie par la bibliothèque FLANN pour les recherches au plus proche voisin. L'algorithme SIFT est utilisé pour détecter et décrire les points clés. Cette méthode a été développée pour être testée uniquement dans un environnement dépendant du script.

The International Organization for Migration (IOM). Said there are more than 200 million migrants around the world today. Europe hosted the largest number of immigrants, with 70.6 million people in 2005, the latest year for which figures are available.

الزلازل هو ظاهرة طبيعية عبارة عن اهتزاز أرضي سريع يعود إلى تكسر الصخور وازاحتها بسبب تراكم الجهود داخلية نتيجة لمؤثرات جيولوجية ينتج عنها تحرك الصفائح الأرضية. وتكونه أجمالا "آثاره كارثية على أي منطقة تتعرض له.

الزلازل هو ظاهرة طبيعية عبارة عن اهتزاز أرضي سريع يعود إلى تكسر الصخور وازاحتها بسبب تراكم الجهود داخلية نتيجة لمؤثرات جيولوجية ينتج عنها تحرك الصفائح الأرضية. وتكونه أجمالا آثار كارثية على أي منطقة تتعرض له.

Figure10: Échantillons manuscrits arabe et anglaise de la base de données QUWI

IV.4.1.1- Les résultats de compétitions ICFHR 2016 :

Au total, La tâche d'identification de scripteur a reçu quatre soumissions. MISC, Nuremberg, CVC et CSMC,

Méthode	Taux de classification			
	1A	1B	1C	1D
MISC	74.33	69.00	03.33	03.67
Nuremberg	89.33	84.67	10.33	18.67
CVC	80.67	80.33	04.07	07.33
CSMC	90.33	87.67	-----	-----

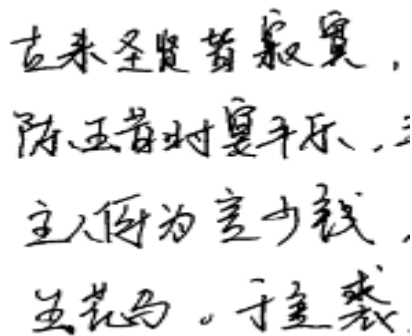
Tableau 2 : résultat de compétition ICFHR 2016 pour d'identification de rédacteur [22]

IV.4.2- ICFHR 2018 Compétition on Multi-script Writer Identification :

Continuation des compétitions similaires précédemment organisés utilisant des textes manuscrits multi-scripts, la compétition ICFHR 2018 [23] visait à explorer le problème d'identification d'écrivain traditionnel dans un scénario plus complexe d'environnement multi-scripts où la formation et le test document scripteur proviennent de scripts différents. Trois bases de données différentes contenant des documents manuscrits en arabe, français, anglais, chinois et farsi ont été utilisées pour les six tâches du concours. Les résultats obtenus indiquent que, bien que élevé les taux d'identification sont rapportés dans la littérature par les systèmes d'identification classiques des auteurs, Les base des données utilisé dans cette compétition sont, LAMIS-MSHD et WDAD et CERUG.

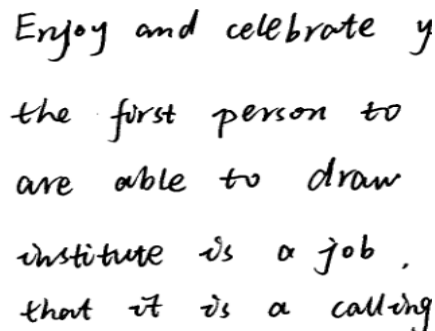
Dans notre travail nous avons étudié une seul tâche, pour la Classification de système d'identification dans la même expérience de la compétition ICFHR 2018 d'identification.

- ❖ **Tâche 1:** document manuscrite chinois en apprentissage et document manuscrite anglais dans la base de test.



古来圣贤皆寂寞，
陈王昔时宴平乐，
主人何为言少钱，
五花马，千金裘。

Figure 11: document chinois



Enjoy and celebrate y
the first person to
are able to draw
institute is a job,
that it is a calling

figure 12: document anglais

- ❖ **Un exemple sur bases de données de la compétition icfhr 2018**

Les bases des données utilisées dans cette compétition sont, LAMIS-MSHD et WDAD et CERUG.

La base de données LAMIS-MSHD:

Comprend 1200 données manuscrites. Images de texte écrites par 100 écrivains différents. Chaque écrivain produit 12 documents manuscrits, les 6 premières pages contiennent Arabe alors que les 6 pages restantes contiennent du texte français. Le texte sur chacune des 12 pages est différent et tous les rédacteurs sont copiés le même texte.

La base de données CERUG:

Contient des documents manuscrits recueillies auprès de 105 sujets chinois, Chaque personne devait écrire quatre pages A4 différentes. En page 1, les participants étaient demandés de copier un texte de deux paragraphes en chinois. Sur la page 2, les sujets ont décrit certains sujets qu'ils aimaient dans leurs propres mots en chinois. La page 3 contient le texte anglais copié de deux paragraphes. Cette page est divisée en deux sous-pages et chaque sous-page contient un paragraphe. Au total, il y a quatre des échantillons manuscrits de chaque écrivain, deux en chinois et deux en anglais. Tous les documents ont été numérisés à 300 dpi, 8 images en niveaux de gris bits / pixel.

La base de données WDAD:

Comprend 800 textes manuscrits images écrites par 200 écrivains différents. Toutes les formes en les jeux de données sont remplies par des rédacteurs représentant un âge différent groupes, degré de justesse, niveaux d'éducation et sexes. Chaque l'auteur a produit 4 documents manuscrits, les 2 premières pages contiennent du farsi tandis que les 2 pages restantes contiennent de l'anglais texte.

IV.4.2.1- Le résultat de la compétition ICFHR 2018:

Trois équipes différentes dont L'Algérie, l'Allemagne et le Japon ont été soumis à cette compétition, LIMPAF-I, LIMPAF-II, Tokyo et Nuremberg

- Les meilleurs taux d'identification de tous les systèmes soumis sont très faibles une fois comparés à l'état de cœur systèmes d'identification d'écrivain principalement évalués en utilisant cadre expérimental traditionnel. Le plus haut réalisé Le taux d'identification est de 57,50% pour la tâche 1 (document chinois, en formation et en anglais dans le test).
- Pour toutes les tâches, les systèmes basés sur les caractéristiques texturales (LIMPAF-1 et LIMPAF-2) sont plus performants que les systèmes compter sur l'extraction automatique de caractéristique à l'aide de CNN. Il est également intéressant de noter que dans la plupart des cas, pour un données, les taux d'identification sont également sensibles une fois que le la formation et les tests sont inversés. Comparer les performances de différents systèmes à travers les trois bases de données, on peut observer que l'identification la plus basse les taux sont rapportés dans l'ensemble de données WDAD. Un raison possible pourrait être la quantité relativement limitée de texte par page dans cet ensemble de données par rapport à LAMISMSHD et ensembles de données CERUG.

Chapitre 03: Etat de l'art

Système	Tâche 1				Tâche 2				AR
	Top-1	Top-2	Top-5	Top-10	Top-1	Top-2	Top-5	Top-10	
LIMPAF-1	42.50(2)	53.75(2)	72.50(2)	83.75(2)	56.25(1)	70.00(1)	81.25(1)	90.00(1)	12
LIMPAF-2	57.50(1)	67.50(1)	80.00(1)	86.25(1)	46.25(2)	55.00(2)	71.25(2)	80.00(2)	12
Tokyo	23.75(4)	42.50(4)	60.00(4)	68.75(4)	16.25(4)	28.75(4)	46.25(4)	75.50(4)	32
Nuremberg	32.50(3)	46.25(3)	66.25(3)	82.50(3)	27.50(3)	40.00(3)	61.25(3)	80.00(3)	23

Tableau 3: taux d'identification et classement de système participants aux tâche 1 et 2
(BASE DE DONNÉES CERUG)

Système	Tâche 3				Tâche 4				AR
	Top-1	Top-2	Top-5	Top-10	Top-1	Top-2	Top-5	Top-10	
LIMPAF-1	40.83(1)	52.92(1)	67.92(2)	83.33(1)	42.08(1)	51.67(1)	73.83(1)	85.00(1)	9
LIMPAF-2	37.50(2)	50.00(2)	68.75(1)	80.83(2)	40.00(2)	47.50(2)	68.75(2)	79.58(2)	15
Tokyo	30.00(3)	40.42(3)	56.67(3)	71.25(3)	17.08(4)	29.17(4)	51.25(3)	60.83(4)	27
Nuremberg	19.58(4)	24.17(4)	36.67(4)	55.42(4)	31.25(3)	36.67(3)	46.67(4)	63.75(3)	29

Tableau 4: taux d'identification et classement de système participants aux tâche 3 et 4(BASE DE DONNÉES LAMIS-MSHD)

Chapitre 03: Etat de l'art

Système	Tâche 5				Tâche 6				AR
	Top-1	Top-2	Top-5	Top-10	Top-1	Top-2	Top-5	Top-10	
LIMPAF-1	29.37(1)	38.75(1)	58.12(2)	70.62(1)	28.75(1)	38.12(1)	59.37(1)	68.75(1)	8
LIMPAF-2	26.25(2)	36.87(2)	49.37(1)	64.37(2)	24.37(2)	33.75(2)	47.50(2)	64.37(2)	16
Tokyo	9.37(4)	16.87(4)	31.25(4)	50.62(4)	6.87(4)	17.50(4)	33.12(4)	47.50(4)	32
Nuremberg	20.62(3)	28.12(3)	45.00(3)	59.37(3)	17.50(3)	21.87(3)	38.75(3)	51.87(3)	24

TABLE 5: taux d'identification et classement de système participants aux tâches 5 et 6 (BASE DE DONNÉES WDAD)

Système	Tâche 1 et 2	Tâche 3 et 4	Tâche 5 et 6	Overall AR (rank)
LIMPAF-1	12	9	8	29(1)
LIMPAF-2	12	15	16	43(2)
Tokyo	32	27	32	91(4)
Nuremberg	23	29	24	76(3)

Tableau 6 : grades globaux de système participants aux six tâches.

- **Distribution de base de données CERUG pour la tâche 1:**

Les tâches 1 est constitué des documents manuscrites anglais et chinois en appelant la Base de données CERUG. La base d'apprentissage comprend 80 doucement manuscrits en chinois et 80 document en anglais de 40 auteurs différents, tandis que les données de validation contiennent 50 documents d'écriture en chinois et en anglais de 25 auteurs différents. L'ensemble de test comprend 160 images manuscrites non étiquetées, 80 en chinois et 80 en anglais sur un total de 40 écrivains différents.

IV.5- conclusion

Dans ce chapitre nous avons présenté un aperçu sur état de l'art dans le domaine de l'identification l'écriture manuscrite et les différentes compétitions organisées dans ce domaine. Nous concentrons dans sa première section sur la présentation des principaux travaux de recherche dans ce domaine. Ensuite, nous déterminons les compétitions dans ce domaine et les résultats obtenus dans chaque compétition. On basant sur la compétition ICFHR 2018 qui est l'objectif de notre expérimentation dans le chapitre suivant.

V.1- Introduction

Dans ce chapitre, on va présenter l'expérimentation que nous avons effectuées afin d'évaluer l'efficacité de notre architecture de CNN proposé pour identifier le scripteur d'un document manuscrite considérée. L'évaluation est menée sur la tache 1 de la base de données ICFHR2018.

V.2- Présentation des outils de développement :

V.2.1- Matériel:

Le matériel utilisé est PC personnel HP I3 avec un 4GB capacité mémoire, et un processeur Intel® coré™ 2 Duo CPU 2 GHz, avec Windows 7 professionnel, service pack 1 64 bit type système.

V.2.2- Google colab :

Google Colab est un service en nuage gratuit qui prend désormais en charge les GPU gratuits Colab travaille sur votre propre disque Google Drive, Vous pouvez:

- améliorez des compétences de codage en langage de programmation Python.
- développer des applications d'apprentissage en profondeur en utilisant des bibliothèques populaires telles que Keras, TensorFlow, PyTorch et OpenCV.
- La fonctionnalité la plus importante qui distingue Colab des autres services cloud gratuits est la suivante: Colab fournit un processeur graphique et est totalement gratuit [24].

V.2.3- Python :

Python est un langage de programmation puissant et facile à apprendre. Il possède des structures de données de haut niveau et permet une approche simple mais efficace de la programmation orientée objet. De par sa syntaxe élégante, son typage dynamique et interprété, Python est un langage idéal pour les scripts et le développement rapide d'applications dans de nombreux domaines et sur la plupart des plates-formes. Par exemple apprentissage en profondeur [25].

V.3- La base de données

Dans ce travaille, nous avons effectués les expérimentations sur une partie de la base de données de la tache 1 CERUG de la compétition ICFHR2018 [14] . Cette partie comprend 160 de documents manuscrits, 80 de documents manuscrits ont été fournis pour l'apprentissage, et 80 en test. L'aspect le plus intéressant est l'utilisation d'une base de données avec des documents

Chapitre 04: Résultats expérimentaux

manuscrits du de différents scripts. La répartition de système d'identification et les échantillons de système d'identification qui utilisés pour l'apprentissage et le test avec d'autre information est présenté dans le tableau 7.

Base de données	Protocole de la base de données
Script	Chinois et anglais
Nombre de scripteur	40 * 2
La base l'apprentissage	80 documents manuscrits chinois
La base de test	80 documents manuscrits anglais

Tableau 7: description de la base de données

V.4- La phase de prétraitement :

Le prétraitement des entrées en général ou la normalisation des entrées est une phase très important dans les réseaux neurones convectifs (CNN). L'idée est que nous souhaitons standardiser autant que possible les entrées du réseau, afin que ;

- a) L'apprentissage soit plus stable (en réduisant la variabilité entre les données de formation)
- b) Le réseau généralisera mieux les données originales (car la normalisation réduit la variabilité entre les données d'apprentissage et de test)

V.4.1- Normalisation des images

La normalisation des données est une étape importante qui garantit que chaque entrée (pixel, dans ce cas) a une distribution de données similaire, Cela accélère la convergence lors de phase d'apprentissage du réseau.

Pour les ensembles d'image (apprentissage et test) nous avons décidé de normaliser le document manuscrite à l'image à 100x100 on suit les étapes suivant :

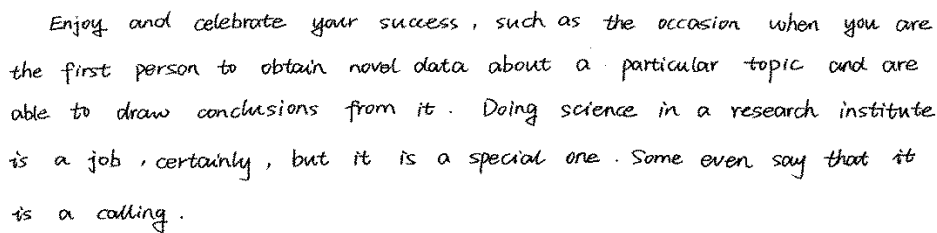
- ✓ 1ere étape : suppression le bordeur : on prend une image redressé sans avoir par 5 pixels de chaque cote.
- ✓ 2ime Etape : extraction de mots connexe depuis l'image original par l'utilisateur de la technique de détection de contour en utilisant la fonction de Matlab BWLABEL , la

Chapitre 04: Résultats expérimentaux

sélection de ces composant s'appuient sur 8 direction de code Freeman après la délimitation de l'image originale voire le figure 13.

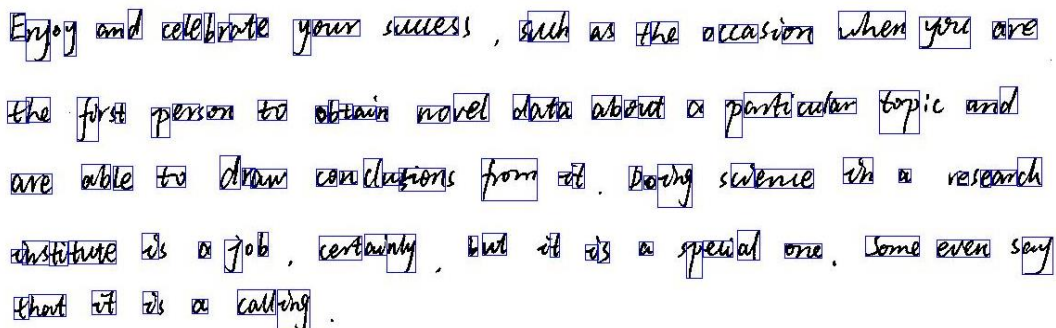
- ✓ Le dernière étape: consacre de concatiner les éléments connexes dans un image de dimension 10x10.

Cette image utilise de CNN comme entré, voire figure 14.



Enjoy and celebrate your success , such as the occasion when you are the first person to obtain novel data about a particular topic and are able to draw conclusions from it . Doing science in a research institute is a job , certainly , but it is a special one . Some even say that it is a calling .

Figure 13: exemple image de test originale



Enjoy and celebrate your success , such as the occasion when you are the first person to obtain novel data about a particular topic and are able to draw conclusions from it . Doing science in a research institute is a job , certainly , but it is a special one . Some even say that it is a calling .

Figure 14: extraction des mots connexe en utilisant la méthode détection de contours

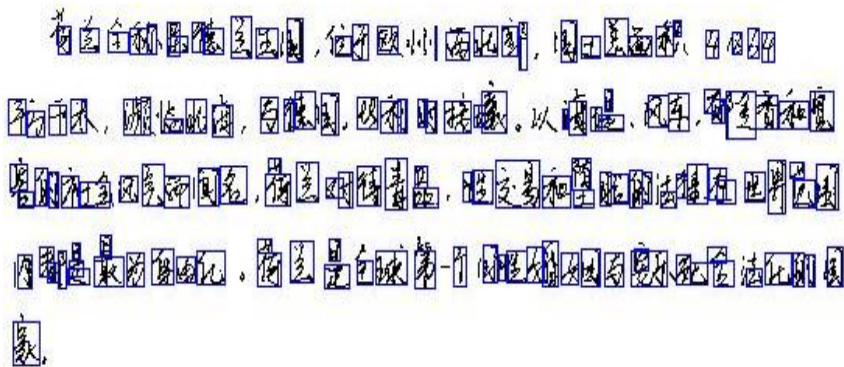


Figure 15: extraction des mots connexe en utilisant la méthode détection de contours

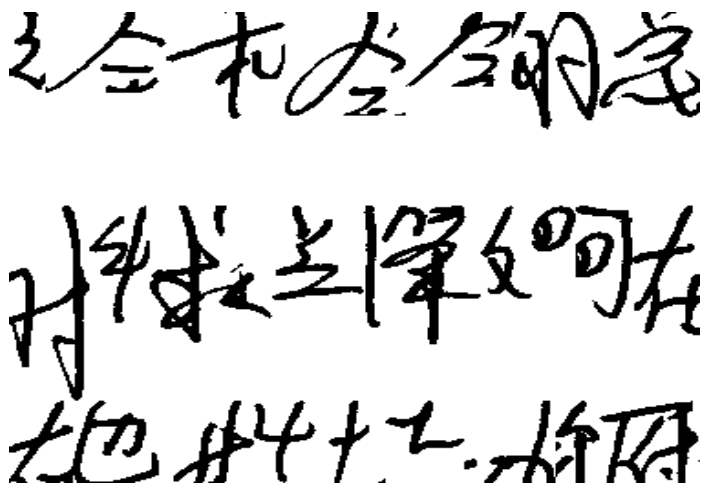


Figure 16: image normalisé (100*100)

❖ La structure CNN :

L'étape de prétraitement est appliquée à la fois dans toutes les images de tache 1 de la base de données ICFHR 2018. Le but de cette phase est de rendre le document d'identification standard pour la classification CNN.

Le technique propose de la normalisation est permet d'augmenter le nombre de échantillons par doucement de la base d'apprentissage à 7 images par doucement c-à-dire 14 image par scripteur, et pour le test en gardent seul image normalise de tailles 100x100.

❖ La structure

- 1- Input (100x100 image)-Output (40 écrivain).

Chapitre 04: Résultats expérimentaux

- 2- Convolution layer: convolution 1, convolution 2.
- 3- Sub sampling layer: pooling 1, pooling 2.
- 4- Full connections: FC.

Dans ce travail nous allons utiliser l'architecture suivante: couches de convolution + couches max pooling+ couche convolution + couche max pooling suivies de couches de full connexion.

Input : Image de tailles 100x100 de texte.

Convolution 1: dans la première couche de convolution, nous créons 7 feature map de taille 5x5. Stride est 1 pour les couches de convolution.

MaxPooling2D 1: Nous utilisons une fenêtre 2x2 car elle est la plus courante, et pas=2.

Convolution 2 : la deuxième couche est une convolution nous créons 14 feature map de taille 5x5. Stride est 1 pour les couches de convolution.

MaxPooling2D 2 : crée une couche pooling 2, nous utilisons une fenêtre 2x2 car elle est la plus courante. Par défaut, et le stride est 1.

Connexion complète: après les couches de convolution + pooling, est un résultat de structure complète

Output : c'est un résultat de cette méthode.

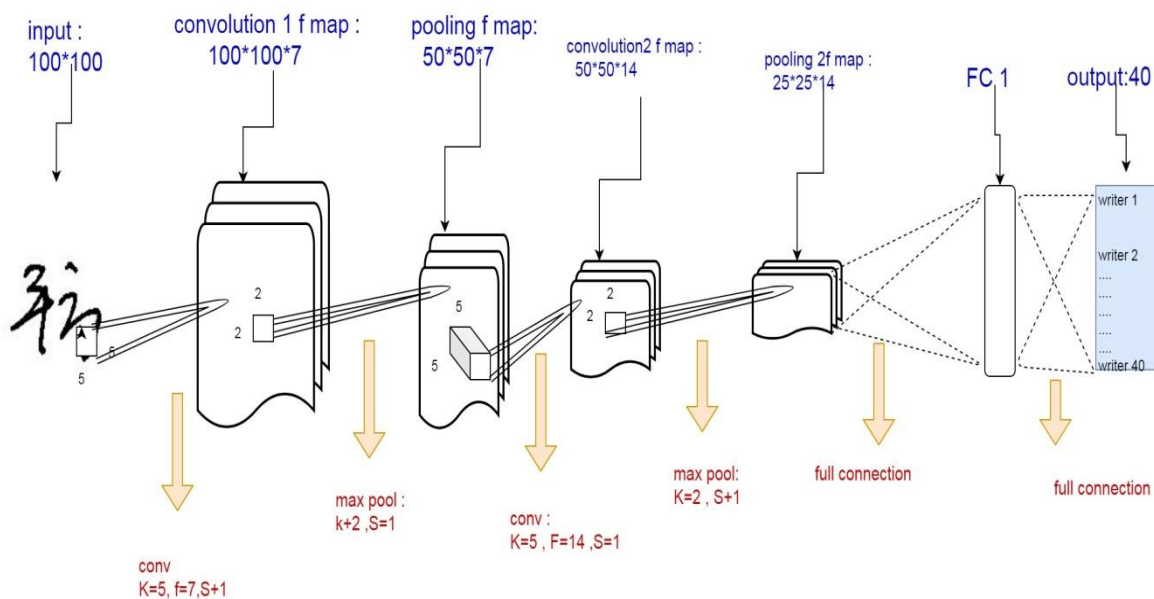


Figure 17: structure de CNN

V.5- Résultat et décision

- 1- dans l'expérimentation nous fixons le paramètre de CNN comme le nombre d'épochè à 100, size batch à 7, nombre de filtre de la premier couche de convolution est 7 et la deuxième couche à 14.
- 2- Après 100 épochè, nous avons atteint un taux de 68,75% sur notre base de donnes de la tache 1.
- 3- Le résultat de système propose avec les quatre systèmes sur la base de tache 1, sont résumés dans les tableaux 6,

Systeme	Top-1
LIMPAF-1	42.50
LIMPAF-2	57.50
Tokyo	23.75
Nuremberg	32.50
Notre système	68.75

Tableaux 8: les taux d'identification

On a noté dans le tableau aux dessus, que notre approche proposé attent meilleur taux de identification par rapport les autres méthodes.

V.6- Conclusion

Dans ce chapitre, nous avons proposé une méthode pour extraire les caractéristiques et les différents paramètres des systèmes d'identification de la base de données ICFHR 2018, après avoir achevé notre conception nous avons donné les outils nécessaires pour la réalisation de notre travail. Nous avons présenté aussi l'environnement de développement. Nous avons aussi utilisé le

Chapitre 04: Résultats expérimentaux

classifieurs CNN pour classer ces mêmes documents. Globalement, nous avons obtenu de meilleures performances que les meilleurs résultats publiés précédemment (ICFHR 2018), ce qui démontre également l'efficacité de notre méthode d'identification de rédacteur d'un document manuscrite.

Conclusion générale

VI- Conclusion général

Malgré la croissance des documents numériques au cours de la dernière années, l'écriture manuscrite et les documents manuscrits ont conservé leur importance. L'écriture manuscrite représente une tâche complexe qui met en jeu différents traitements cognitifs par opposition à la dactylographie. Le fait de dessiner des caractères à la main améliorés considérablement la reconnaissance ultérieure, mais conduit également à un meilleur développement intellectuel. L'analyse de l'écriture présente un sujet intéressant pour les psychologues, les paléographes, les examinateurs de documents et les analystes légistes.

Cependant, au cours de la dernière décennie, l'apprentissage en profondeur est en train de devenir un sous-ensemble très populaire de l'apprentissage automatique en raison de ses performances élevées sur de nombreux types de données. Un excellent moyen d'utiliser l'apprentissage en profondeur pour classer les images consiste à créer un réseau de neurones convolutifs (CNN), ces méthodes d'apprentissage en profondeur ont été appliquées avec succès pour apprendre automatiquement les caractéristiques à partir de données. Nous proposons dans ce mémoire une méthode d'apprentissage en profondeur pour l'identification d'écrivain qui ne nécessitant pas d'identification préalable des caractéristiques. Un réseau de neurones convolutif (CNN) est initialement formé pour extraire les caractéristiques locales, qui représentent les caractéristiques de l'écriture individuelle dans les images de caractères entières et leurs sous-régions

Nous avons vu comment définir les paramètres de base des données de documents manuscrits pour notre modèle et avons suivi les étapes d'apprentissage de base requises. Les étapes de prétraitement devraient nous aider à utiliser un modèle de réseau neuronal très efficace.

La méthode proposée est testée sur une base de données de la tâche 1, de les compétitions ICFHR2018. Les résultats expérimentaux démontrent que la méthode proposée obtient la meilleure performance par rapport aux états de l'art. Ces résultats encourageants une suite d'amélioration sur le modèle de CNN et les méthodes nécessaires pour augmenter le nombre d'échantillons de la base d'apprentissage utilisé.

VII- Annexe

❖ Téléchargé la base

```
from google.colab import drive
drive.mount('/content/gdrive')
import scipy.io as scio
input_data = scio.loadmat('/content/gdrive/My Drive/Colab Notebooks/1005.mat')
```

↳ Drive already mounted at /content/gdrive; to attempt to forcibly remount, call drive.mount("/content/gdrive", force_remount=True).

Figure 18: création de la base

Où est Les bases de données et téléchargé sont situées et folder un converties en fichier matlab

❖ Importe les méthodes

```
[ ] import tensorflow as tf
import numpy as np
import time
```

Figure 19: importer les modules utilisés

Importons les modules requis ici. Nous allons discuter ces modules pendant les utilisons dans les segments de code.

❖ Afficher les images en vecteur

```
[ ] # print input_data
# from tensorflow.examples.tutorials.100S import input_data
#100S = input_data.read_data_sets('100S_data', one_hot=True)

sess = tf.InteractiveSession()
# test
x_train = input_data['x_train'].astype('float32')/255
y_train = input_data['y_train'].astype('float32')
x_test = input_data['x_test'].astype('float32')/255
y_test = input_data['y_test'].astype('float32')
print(x_train[0])
print(np.shape(x_train[0]))
```

↳ [0.00392157 0.00392157 0.00392157 ... 0.00392157 0.00392157 0.00392157]
(10000,)

Figure 20: afficher l'image existe un BDD

Il est également possible de dire qu'Interactive Session prend en charge moins de frappe, car cela permet d'exécuter des variables sans avoir à se référer constamment à l'objet de session.

Par conséquent, avant de l'introduire dans le réseau, vous devez convertir son type en float32 et redimensionner les valeurs de pixel

❖ Déclaration de variable

```
[ ] # declaration
def weight_variable(shape):
    initial = tf.truncated_normal(shape, stddev=0.1, seed=1)
    return tf.Variable(initial)

def bias_variable(shape):
    initial = tf.constant(0.1, shape=shape)
    return tf.Variable(initial)

def conv2d(x, w):
    return tf.nn.conv2d(x, w, strides=[1, 1, 1, 1], padding='SAME')

def max_pool_2x2(x):
    return tf.nn.max_pool(x, ksize=[1, 2, 2, 1], strides=[1, 2, 2, 1], padding='SAME')

# input&output
xs = tf.placeholder(tf.float32, [None, 100*100])
ys = tf.placeholder(tf.float32, [None, 40])
keep_prob = tf.placeholder(tf.float32)
x_image = tf.reshape(xs, [-1, 100, 100, 1])
```

Figure 21 : déclaration les variable

Initialisation des poids et des biais

Nous devons maintenant définir une fonction de perte pour optimiser nos poids et nos biais, avec les logos pour le label prédit et correct. et définissez les poids et les biais à l'aide de la fonction, initialisez avec une petite valeur positive aléatoire, Fonctions de base de construction de couches de regroupement et de convolution nous définissons un espace réservé pour nos variables d'entrée et de sortie et pour toutes les variables que nous souhaitons suivre.

❖ Structure CNN

```
# CNN structure
Featuremaps1=16;
Featuremaps2=32;
sizefmfinal=25;
W_conv1 = weight_variable([5, 5, 1, Featuremaps1]) # Feature maps =7
b_conv1 = bias_variable([Featuremaps1])
h_conv1 = tf.nn.relu(conv2d(x_image, W_conv1) + b_conv1)

h_pool1 = max_pool_2x2(h_conv1)

W_conv2 = weight_variable([5, 5, Featuremaps1, Featuremaps2]) # Feature maps=14
b_conv2 = bias_variable([Featuremaps2])
h_conv2 = tf.nn.relu(conv2d(h_pool1, W_conv2) + b_conv2)

h_pool2 = max_pool_2x2(h_conv2)
h_pool2_flat = tf.reshape(h_pool2, [-1, 25*25*Featuremaps2])

W_fc1 = weight_variable([25*25*Featuremaps2, 40])
b_fc1 = bias_variable([40])
y_conv = tf.nn.relu(tf.matmul(h_pool2_flat, W_fc1) + b_fc1)
```

Figure 22: structure de CNN.

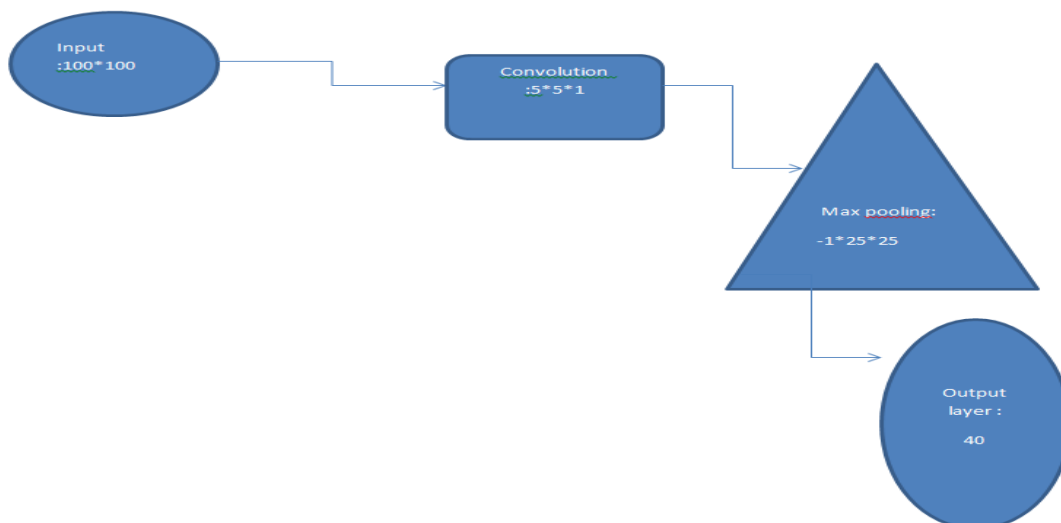


Figure 23: description de CNN

Annexes

Nous pouvons simplement empiler des calques en ajoutant le calque souhaité, un par un. C'est exactement ce que nous allons faire ici: nous allons d'abord ajouter une première couche convolutionnelle avec `Conv2D()`. Notions que nous utilisons cette fonction car nous travaillons avec des images, ensuite, nous allons ajouter la couche max-pooling avec `MaxPooling2D()`, etc. La dernière couche est une couche dense dotée d'une fonction d'activation softmax à 40 unités, nécessaire pour résoudre ce problème de classification multi-classes.

❖ Loss fonction

```
# loss function
cross_entropy = tf.nn.l2_loss(ys - y_conv)
train_step = tf.train.GradientDescentOptimizer(0.0005).minimize(cross_entropy)
# testà)
correct_prediction = tf.equal(tf.argmax(y_conv, 1), tf.argmax(ys, 1))
accuracy = tf.reduce_mean(tf.cast(correct_prediction, tf.float32))
sess.run(tf.initialize_all_variables())
```

Figure 24: loss fonction.

Nous devons maintenant définir une fonction de perte pour optimiser nos poids et nos biais, et nous allons utiliser l'entropie croisée softmax avec des logis pour le label prédit et correct. Nous allons également ajouter des régularisations de niveau 1 à notre réseau.

❖ Train on minibatches:

```
[7] # train on minibatches
batch_size=7
nTraindb=560
imgsize=100
numbatches=(560//7)
num_epochs=100
kk=np.random.RandomState(seed=1).permutation(numbatches)*batch_size
# Loop over number of epochs
for epoch in range(num_epochs):
    start_time = time.time()

    for i in range(numbatches):
        if i % 100 == 0:
            train_accuracy = accuracy.eval(feed_dict={xs: x_train[kk[i]:kk[i]+batch_size].reshape(batch_size, imgsize*imgsize), ys: y_train[kk[i]:kk[i]+batch_size]})
            #print("step %d, training accuracy %g" % (i, train_accuracy))
            sess.run(train_step, feed_dict={xs: x_train[kk[i]:kk[i]+batch_size].reshape(batch_size, imgsize*imgsize), ys: y_train[kk[i]:kk[i]+batch_size].reshape(batch_size, imgsize*imgsize)})
            print(sess.run(accuracy, feed_dict={xs: x_test, ys: y_test}))
            end_time = time.time()
            print("Epoch "+str(epoch+1)+" completed : Time usage "+str(int(end_time-start_time))+" seconds")
            print (sess.run(accuracy, feed_dict={xs: x_test, ys: y_test}))
        else:
            print(0.6875)

Epoch 100 completed : Time usage 0 seconds
0.6875
```

Figure 25: train on minimarchés

Référence bibliographique

- [01] D.X. Le and G.R. Thoma, "Automated Borders Detection and Adaptive Segmentation for Binary Document Images", International Conference on Pattern Recognition, p. III: 737-741, 1996.
- [02] ZARAA Abderrazak FARES Souhail, Détermination du sexe d'un scripteur a partir d'un document manuscrit, MEMOIRE DE MASTER, Université de Tébessa , Mai 2017, Algerian
- [03] MENASRIA Abdelaali ZEMOULI Saddam , Reconnaissance hors ligne des chiffres manuscrite isolé (Base de donnée M.N.I.S.T) , MEMOIRE DE MASTER , Université de Tébessa , Algerian
- [04] Imran SIDDIQI, Florence CLOPPET, Nicole VINCENT "Contour Based Features for the Classification of Ancient Manuscripts," presented at the 14th Conference of the International Graphonolics Society, (IGS), Dijon, 2009
- [05] Pierre Buysens, Fusion de différents modes de capture pour la reconnaissance du visage appliquée aux e_transactions doctorat de l'universite de Caen, France , Janvier 2011
- [06] Sylvain TERTOIS, RÉDUCTION DES EFFETS DES NON-LINÉARITÉS DANS UNE MODULATION MULTIPORTEUSE A L'AIDE DE RÉSEAUX DE NEURONES, Équipe d'accueil : ETSN, Supélec, Campus de Rennes École Doctorale : MATISSE Composante Universitaire : S.P.M., SOUTENUE LE 12 décembre 2003 devant la commission d'Examen
- [07] Y. Bengio, "Learning Deep Architectures for AI," Foundations and Trends in Machine Learning, vol. 2, no. 1, pp. 1– 127, January 2009.
- [08] L. Bottou, "Large-scale machine learning with stochastic gradient descent," Proceedings of the International Conference on Computational Statistics, pp. 177–186, August 2010, Paris, France
- [09] Y. Le Cun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. Proceedings of IEEE, 86(11) :2278–2324, November 1998.
- [10] Adam Gibson, Josh Patterson, "Deep Learning", O'Reilly Media, Inc, August 2017, ISBN: 9781491924570
- [11] Ikram Chraïbi Kaadoud, Apprentissage de séquences et extraction de règles de réseaux récurrents : application au traçage de schémas techniques, L'UNIVERSITÉ DE BORDEAUX ÉCOLE DOCTORALE DE MATHÉMATIQUES ET D'INFORMATIQUE, Mars 2018
- [12] Durou A, Aref I, Al-Maadeed S, Bouridane A, Benkhelifa E "Writer identification approach based on bag of words with OBI features" (2019) Information Processing and Management, 56 (2) , pp. 354-366.
- [13] M. Keglevic, S. Fiel and R. Sablatnig, "Learning Features for Writer Retrieval and Identification using Triplet CNNs," 2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR), Niagara Falls, NY, 2018, pp. 211-216. doi: 10.1109/ICFHR-2018.2018.00045
- [14] H.T. Nguyen et al., Text-independent writer identification using convolutional neural network, Pattern Recognition Letters (2018), <https://doi.org/10.1016/j.patrec.2018.07.022>

- [15] Alia Karim Abdul Hassan, Bashar Saadoon Mahdi ,Asmaa Abdullah Mohammed,“ Writer Identification Based on Arabic Handwriting Recognition by using Speed Up Robust Feature and K- Nearest Neighbor Classification», Journal of University of Babylon, Pure and Applied Sciences, Vol. (27), No. (1): 2019
- [16] Hannad Y, Siddiqi I, Kettani MEYE. Arabic writer identification using local binary patterns (LBP) of handwritten fragments. Cham: Springer International Publishing978-3-319-19390-8; 2015. p. 237–44.
- [17] Hannad Y., Siddiqi I., Kettani M.E.Y.E. (2015) Arabic Writer Identification Using Local Binary Patterns (LBP) of Handwritten Fragments. In: Paredes R., Cardoso J., Pardo X. (eds) Pattern Recognition and Image Analysis. IbPRIA 2015. Lecture Notes in Computer Science, vol 9117. Springer, Cham
- [18] Tang, Youbao & Wu, Xiangqian & Bu, Wei. (2013). Offline text-independent writer identification using stroke fragment and contour based features. Proceedings - 2013 International Conference on Biometrics, ICB 2013. 1-6. 10.1109/ICB.2013.6612988
- [19] Khalifa, E., Al-maadeed, S., Tahir, M., & Bouridane, A. (2015). Off-line writer identification using an ensemble of grapheme codebook features. Pattern Recognition Letters, 59, 18–25.
- [20] Djeddi, C., Meslati, L., Siddiqi, I., Ennaji, A., El Abed, H., & Gattal, A. (2014). Evaluation of texture features for offline arabic writer identification. Proceeding of the 11thIAPR international workshop on document analysis systems (pp. 106–110).
- [21] Sheng He and Lambert Schomaker ; Deep adaptive learning for writer identification based on single handwritten word images ; Pattern Recognition ; page(64—74) ; 2019
- [22] Chawki Djeddi and Somaya Al Madeed andImran Siddiqi and Abdeljalil Gattal and Abdellatif Ennaji and Haikal El Abed ; ICFHR2016 Competition on Multi-script Writer Demographics Classification Using "QUWI" Database; 15th International Conference on Frontiers in Handwriting Recognition, ICFHR 2016,Shenzhen,china,October, 23-26,2016;page(602-606);2016
- [23] Djeddi, Chawki & Al-ma'adeed, Somaya & Siddiqi, Imran & Gattal, Abdeljalil & Sheng, He & Akbari, Y. (2018). ICFHR 2018 Competition on Multi-Script Writer Identification. 10.1109/ICFHR-2018.2018.00094.
- [24] <https://www.kdnuggets.com/2018/02/google-colab-free-gpu-tutorial-tensorflow-keras-pytorch.html>
- [25] <https://docs.python.org/fr/3/tutorial/index.html>