

UNIVERSITÉ NATIONALE DU VIETNAM À HANOÏ

Institut de la Francophonie pour l'Innovation



Reconnaissance de Formes

Thème : **Reconnaissance de caractères manuscrits**

Option : Systèmes Intelligents et Multimédia (SIM)

Promotion : 22

Rédigé par :

KINDA Zakaria

ZONGO Sylvain

Enseignant :

Dr Ho Huong Vinh

Année académique : 2018 - 2019

Table des matières

1	Introduction	1
2	Analyse du sujet	2
2.1	Problématique	2
2.2	Objectifs	2
3	Etat de l'art	2
3.1	La méthode proposée par M. Pechwitz et V. Maergner [24]	2
3.2	La méthode proposée par P.Dreuw et al [25]	3
3.3	La méthode proposée par A. Benouareth et al [26]	3
3.4	La méthode proposée par H. Nemmour et Y. Chibani [32]	3
4	Comparaison entre les travaux ultérieurs	3
5	Solution proposée	4
5.1	Phase 1 : Recherche de données	4
5.2	Phase 2 : Algorithme d'apprentissage	5
5.2.1	Algorithme utilisé : K-NN	5
5.2.2	Principe de kNN	5
5.3	Phase 3 : L'application de l'algorithme à notre modèle de données	6
6	Implémentation de la solution proposée	6
7	Résultats de l'implémentation	7
7.1	Test 1	7
7.2	Test 2	10
8	Outils et technologies	11
9	Evaluation	12
9.0.1	Matrice de confusion	12
10	Conclusion	14

Table des figures

1	Données d'entraînement	4
2	Données de test	5
3	Architecture complète du processus	6
4	Architecture de la solution proposée	7
5	Image en entré pour le test	8
6	Résultat du test	8
7	Image d'entrée	9
8	Résultat de sortie	9
9	Image d'entrée	10
10	Image d'entrée	10
11	Image d'entrée	11
12	Image de sortie	11

1 Introduction

En matière d'évolution technologique, les dix dernières années ont permis d'affiner la précision et la pertinence de la capture et de la reconnaissance de caractères imprimés au point que le marché soit considéré aujourd'hui comme totalement mature, voire presque parfait. "Nous ne cherchons aujourd'hui plus à affiner le taux de reconnaissance, confirme Venceslas Cartier ; c'est plutôt sur les fonctions annexes que nous nous concentrons". De son côté, Abbyy se positionne davantage sur l'ICR, dont le marché semble principalement axé sur les champs manuscrits de formulaires à reconnaître. "Des améliorations peuvent encore être apportées dans ce domaine, explique Grégory Laborderie, et ce même si le niveau de reconnaissance est déjà élevé dans le cas de champs dits contraints". Comme disent plusieurs auteurs notre travail sera centré sur la reconnaissance des caractères manuscrits.

2 Analyse du sujet

Après analyse profonde du sujet “ Reconnaissance des caractères manuscrits”, s’inscrit dans le domaine de reconnaissance de langues. Toute personne qui s’intéresse à l’apprentissage des langues peut être directement concernée par ce sujet. Il est appliqué dans le domaine reconnaissance de caractères écrits à la main vue que pour un caractère donné chaque personne peut avoir sa manière de l’écrire.

2.1 Problématique

La distinction des caractères manuscrits et des caractères électroniques est très importante dans tous les milieux. Il existe des outils ou des méthodes permettant la reconnaissance des caractères manuscrits dans un document ou dans une image. Cependant ces outils ou méthodes utilisés pour la reconnaissance des caractères manuscrits donnent des résultats très peu satisfaisants. Ainsi dit dit la distinction et la reconnaissance des caractères manuscrits des caractères électroniques demeure une sérieuse problématique à résoudre.

2.2 Objectifs

L’analyse automatique de documents concerne essentiellement la rétro-conversion d’une image d’un document en un format compatible avec des données sémantiques. Il existe deux types d’écriture le manuscrit et l’imprimé. En effet, la régularité de l’imprimé permet d’utiliser des techniques beaucoup plus fiables et beaucoup plus directes et rapides que celles pour le manuscrit dont la complexité et la variabilité sont très importantes. Nous allons présenter les concepts de bases et les méthodes de reconnaissance en fonction du mode d’acquisition de l’écriture et nous allons nous intéresser plutôt sur les méthodes de reconnaissance d’écriture manuscrite hors ligne.

3 Etat de l’art

De nombreux articles ont été écrits concernant la reconnaissance des caractères manuscrits. Reconnaissance de l’écriture manuscrite avec des réseaux récurrents présentée et soutenue publiquement par Luc Mioulet.

3.1 La méthode proposée par M. Pechwitz et V. Maergner [24]

Ce travail présente un système de reconnaissance basé sur HMM de dimension semi-continue. Pour chaque image de mot binarisé, les paramètres de normalisation ont été estimés. Les expériences ont été réalisées sur les quatre (4) ensembles distincts (a, b, c, d) de la base IFN / ENIT qui contient 26.459 mots arabes manuscrites, l’ensemble (a, b, c) est utilisée pour l’apprentissage et l’ensemble d est utilisé pour le test. Le système est obtenu un taux de reconnaissance d’environ 89%.

3.2 La méthode proposée par P.Dreuw et al [25]

Ce travail consiste à présenter un système basé sur HMM pour la reconnaissance hors ligne de l'écriture arabe qui modélise explicitement les espaces blancs entre les caractères et les morceaux des mots arabes (PAW). Les modèles des espaces blancs proposés pour l'écriture arabe pourrait améliorer la performance de système et surpasse le meilleur taux d'erreur rapportés. Une inspection visuelle des modèles de l'apprentissage a montré le besoin d'une modélisation précise et d'une adaptation des longueurs de caractères. Les expérimentations sont menées sur la base de données l'IFN / ENIT. Cette base de données est divisée en quatre dossiers de l'apprentissage avec un dossier supplémentaire pour le test. La version utilisée contient un total de 32.492 mots arabes écrits à la main par plus de 1000 écrivains, et a une taille de vocabulaire de 937 noms de ville tunisien. Ce système atteint un taux de reconnaissance de 92.86%.

3.3 La méthode proposée par A. Benouareth et al [26]

Ce travail décrit un système de reconnaissance des mots arabes hors-ligne sans contrainte basée sur l'approche sans segmentation (segmentation-free) et les modèles de Markov cachés semi-continus (SCHMMs) avec durée d'état explicite. Pour effectuer un modèle d'apprentissage de mot et de lettre et la reconnaissance plus efficace, les auteurs proposent une nouvelle version de l'algorithme Viterbi prenant en compte la modélisation de durée d'état explicite. Trois distributions (le Gamma, Gauss et Poisson) pour la modélisation de durée d'état explicite ont été utilisées et les comparaisons entre eux ont été rapportées. Pour augmenter le taux de reconnaissance de mot, le système proposé utilise la fenêtre glissante basée sur l'histogramme de projection verticale du mot et extrait un nouvel ensemble de caractéristiques statistiques et structurales. Les résultats obtenus sont très promoteurs et réalisent un de taux de 90.20% (dans le top 1) sur l'ensemble de 26.459 mots arabes de la base de donnée IFN / ENIT.

3.4 La méthode proposée par H. Nemmour et Y. Chibani [32]

Dans ce travail, un système de reconnaissance des mots arabes manuscrits est proposé. Spécifiquement, par l'utilisation d'un système de reconnaissance immunitaire artificiel pour la reconnaissance de script arabe dans un vocabulaire moyen. Les caractéristiques appropriées sont obtenu on utilisant transformé de Ridgelets. Les expériences sont effectuées sur des échantillons extraits de la base de données IFN/ENIT qui contient des noms des villes tunisiennes. L'évaluation des performances est effectuée comparativement à SVM. L'expérimentation est effectuée sur vingt-quatre mots (24) mots correspondant à vingt-quatre mots (24) classes de la base de donnée IFN/ENIT en réalisant un taux de reconnaissance de 79.8%.

4 Comparaison entre les travaux ultérieurs

Le domaine de reconnaissance des mots arabes manuscrits est un vaste domaine qui contient un grand nombre de méthodes de classification qui sont plus ou moins bien adaptés à la reconnaissance de l'écriture. Cependant, cela n'a pas permis de mettre en évidence la supériorité

incontestable et le choix d'une méthode de classification par rapport à d'autres. La comparaison entre les autres travaux de la littérature doit être effectuée et pour que cette comparaison ait un sens, elle doit se faire avec des systèmes exploitant la même base de mots.

5 Solution proposée

Après étude des différents travaux effectués sur le thème “ Reconnaissance des caractères manuscrits” nous optons de travailler sur la “ Reconnaissance des caractères manuscrits en anglais”. Ainsi pour notre solution proposée nous avons plusieurs étapes à suivre. Celles-ci peuvent s'étaler comme suit :

5.1 Phase 1 : Recherche de données

Nos données ont été récupérées sur le site <http://yaroslavvb.com/upload/notMNIST/>. ensuite ces données ont été partagées en deux parties qui sont les données d'apprentissage et les données de test.

- Les données d'entraînement représentent 75% de toutes les données

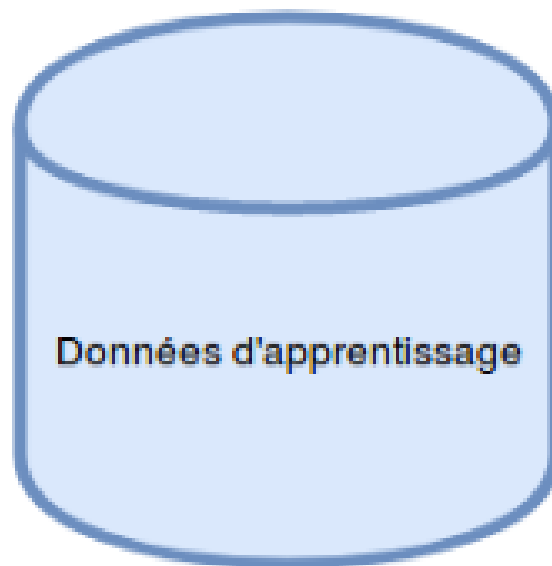


FIGURE 1 – *Données d'entraînement*

- Les données de test qui elles représentent les 25% des données

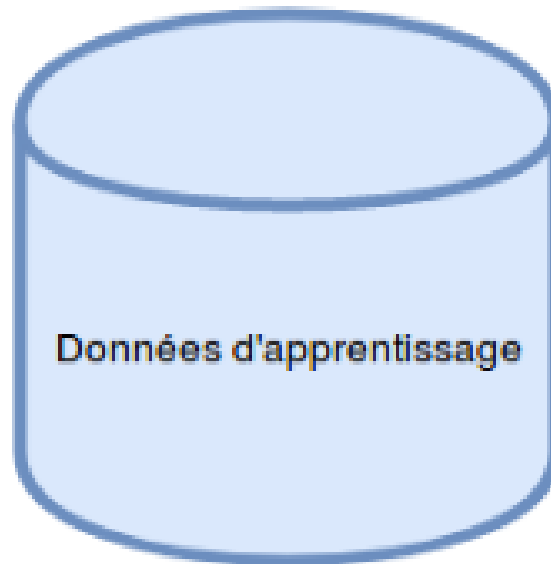


FIGURE 2 – *Données de test*

5.2 Phase 2 : Algorithme d'apprentissage

5.2.1 Algorithme utilisé : K-NN

l'algorithme K-NN (K-nearest neighbors) est une méthode d'apprentissage supervisé. Il peut être utilisé aussi bien pour la régression que pour la classification. L'algorithme du plus proche voisin fait parti des algorithmes les plus utilisés dans le Machine Learning. Pour son implémentation il est donc important de connaître ses principes.

5.2.2 Principe de kNN

Nous pouvons illustrer le principe de KNN comme suite .

- On possède un ensemble de données de départ (données d'apprentissage) ayant chacun un label ;
- On dispose d'un ensemble de données que l'on cherche à classer dans une des catégories ;
- Pour l'élément de données à classer, on calcule "la distance" entre celui-ci et chaque point des données d'apprentissage ;
- On classe les distances par ordre croissant ;
- On choisit les k premiers éléments par ordre de distances ;
- On choisit donc le label dominant parmi les k éléments, ce qui représente la catégorie de l'élément du jeu de données.

5.3 Phase 3 : L'application de l'algorithme à notre modèle de données

Dans cette partie nous faisons une conception complète du processus de reconnaissance de forme. D'abord une extraction des caractéristiques de la base données ensuite nous faisons l'apprentissage à partir de ce modèle et enfin nous évaluons le système sur la base des tests.

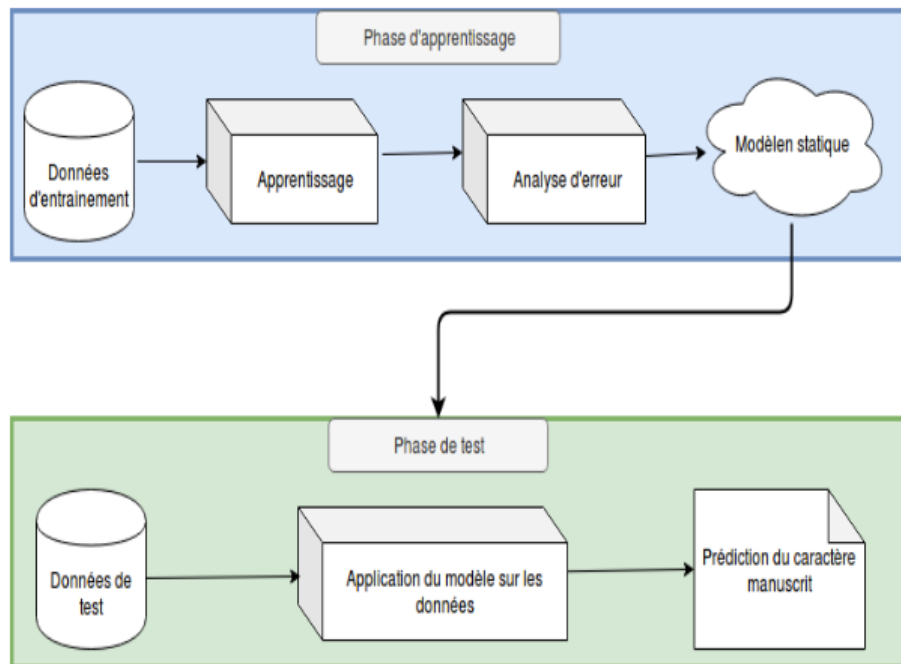


FIGURE 3 – Architecture complète du processus

6 Implémentation de la solution proposée

Après la phase théorique de notre projet qui a consisté à effectuer une analyse de sujet, l'état de l'art (études des différents travaux connexe) et proposer une solution, nous allons passer à l'implémentation de notre solution proposée. La phase d'implémentation consiste donc à développer une application nous permettant de distinguer un caractère manuscrit d'un caractère électronique. A cet effet en pratique lorsque nous avons une image sur laquelle nous voulons récupérer des caractères manuscrits, il faut d'abord les extraire. Et dans notre cas nous avons utilisé le OCR(Optical Character Recognition) pour l'extraction des caractères. Le schéma ci-dessous présente l'architecture d'implémentation de la solution proposée.

Modèle optique : reconnaît les formes

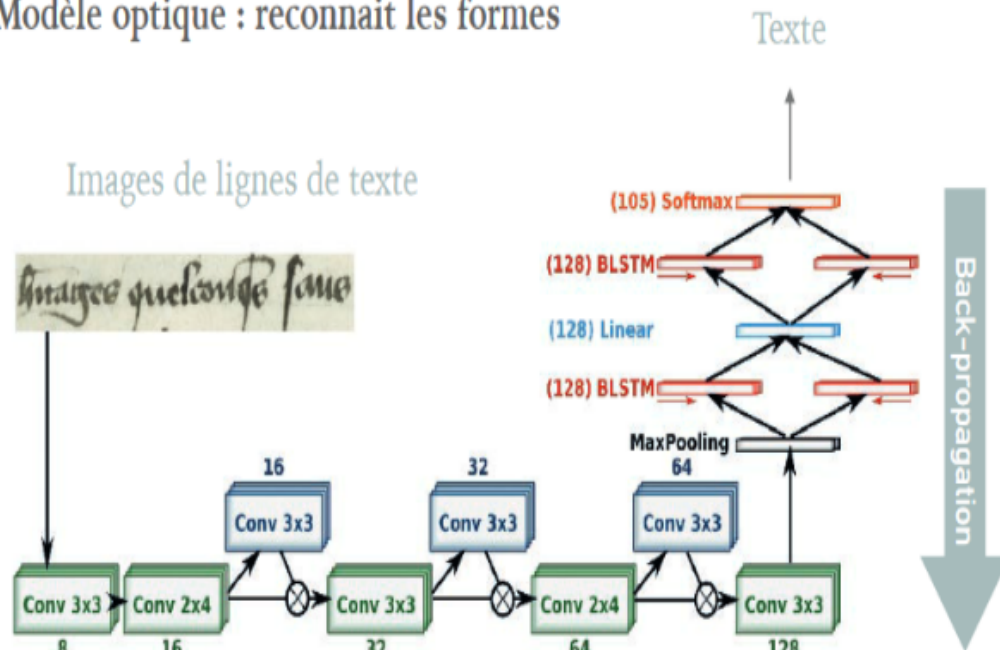


FIGURE 4 – Architecture de la solution proposée

7 Résultats de l'implémentation

L'implémentation de notre solution proposée étant effective, nous obtenons cependant plusieurs résultats. Les résultats obtenus ont été appliqués sur notre base de test, ainsi que d'autres types d'images afin de connaître la performance de notre modèle. A cet effet nous obtenons des résultats différents en fonction de la qualité de l'image, donc des écritures. Tout d'abord, nous donnons en entrée de notre programme une un fichier image déjà pré-traité afin de faire l'extraction des différents caractères.

7.1 Test 1

Pour ce premier test nous avons des écritures manuscrites sur les images que nous chercherons à reconnaître.



FIGURE 5 – Image en entrée pour le test



FIGURE 6 – Résultat du test

Le résultat obtenu par cette image présente certaines erreurs. Nous remarquons à cet effet que le programme distingue difficilement certains caractères manuscrits due au format de l'écriture, cependant il arrive à bien distinguer les caractères électroniques.

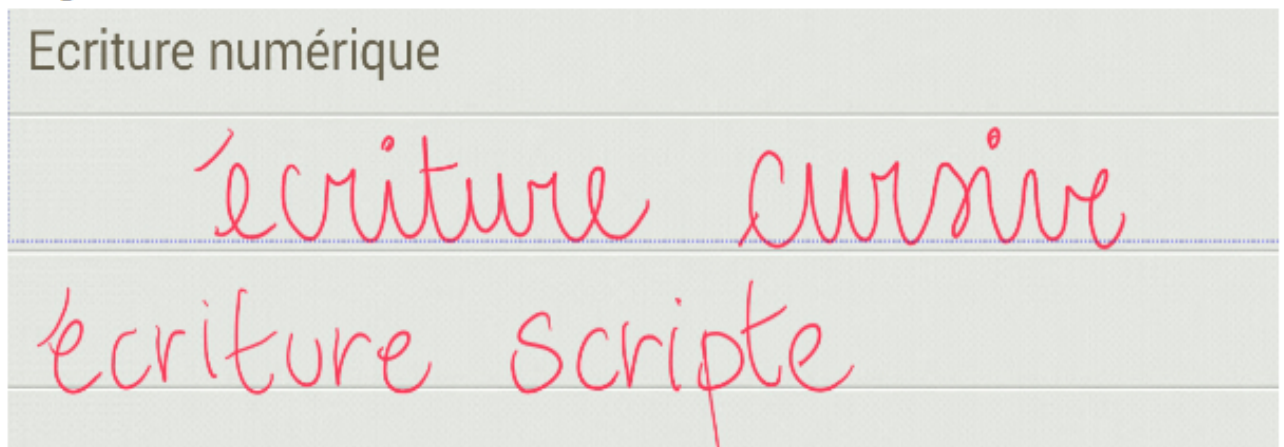


FIGURE 7 – Image d'entrée



FIGURE 8 – Résultat de sortie

Dans cette image nous constatons que l'extraction des informations ne donne pas de meilleur résultat. Le format des caractères manuscrits joue un rôle important dans la reconnaissance des caractères. Par contre pour les caractères électronique la reconnaissance s'effectue plus facilement.



FIGURE 9 – Image d'entrée

```
(base) onezongoforall@onezongoforall-CX61-2QF:~/Desktop/rform$ python d.py
0 Use Simple Storage

- - Data Lake

- Keep mules cl all your raw data in Slmple, cheap storage
[53, HFDS, file system)

- Data Restarz: Be able [a back up and vesmre vow
databases easily

2 a 0 - "My Own Database": Data Marts On Demand

~ Create parameterized varlations of your plocessthat

allow you In assemble data for experimentation
i developmznt, and production
```

FIGURE 10 – Image d'entrée

Pour les images à caractères électroniques uniquement la distinctions des différents caractères est plus visible, cependant les lettres et chiffres sont souvent confondu. Ces constat sont illustrés dans la figure ci-dessus (figure 10).

7.2 Test 2

Pour ce deuxième test nous avons des images sur lesquelles il n'y a que des chiffres de 0 à 9. Nous cherchons ici à reconnaître ces chiffres à l'aide de notre système.

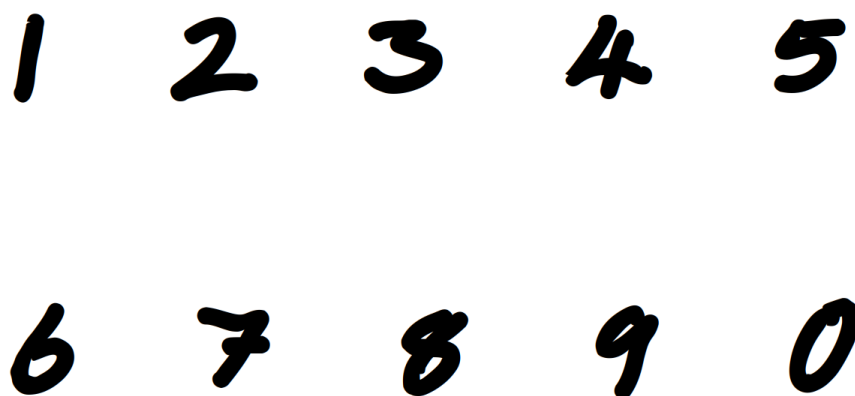


FIGURE 11 – Image d'entrée

Pour ce test l'image d'entrée est constituée uniquement de chiffres et en sortie. Notre système reconnaît tous les chiffres présents sur l'image entrée.

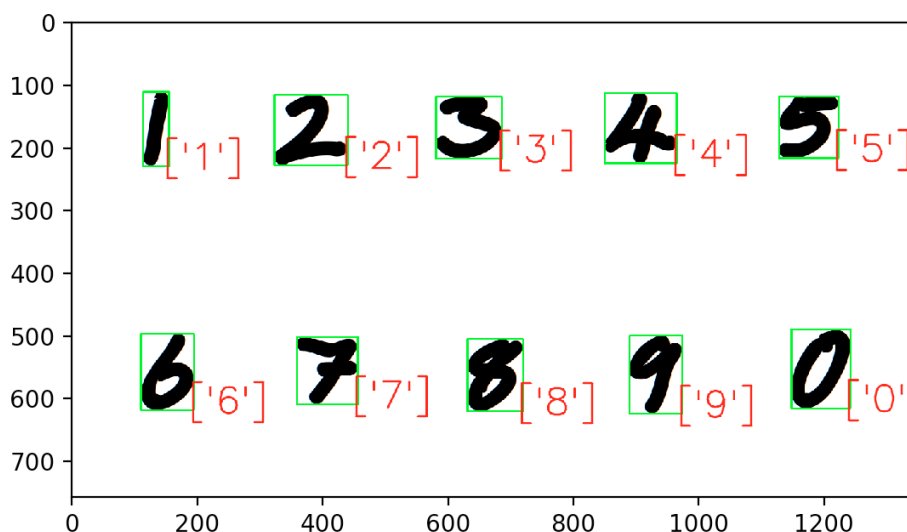


FIGURE 12 – Image de sortie

8 Outils et technologies

Pour mener à bien notre projet, nous avons utilisé plusieurs outils et technologie.

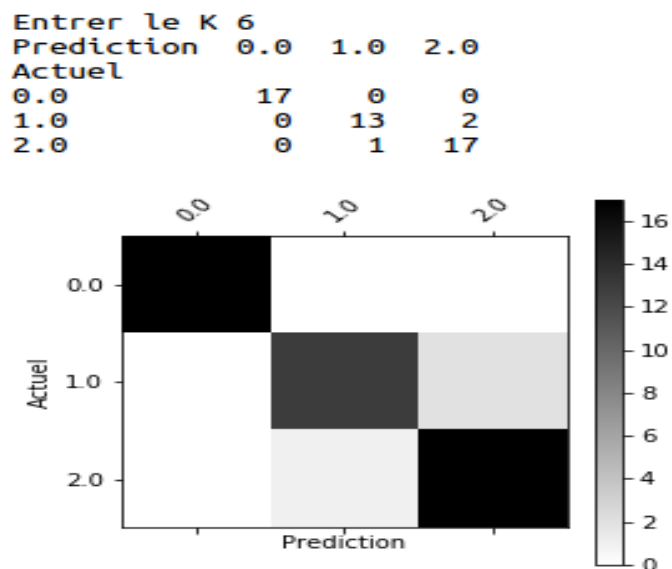
- Outils : Comme outils nous avons utilisés Draw IO pour la modélisation, et spyder pour l'outil de programmation
- Technologie : Nous avons utilisé Python comme langage de programmation de notre projet.

9 Evaluation

Pour évaluer notre système nous avons utilisé la matrice de confusion pour illustrer la précision de la prédiction qu'un caractère pris en entrée est bien prédit ou mal.

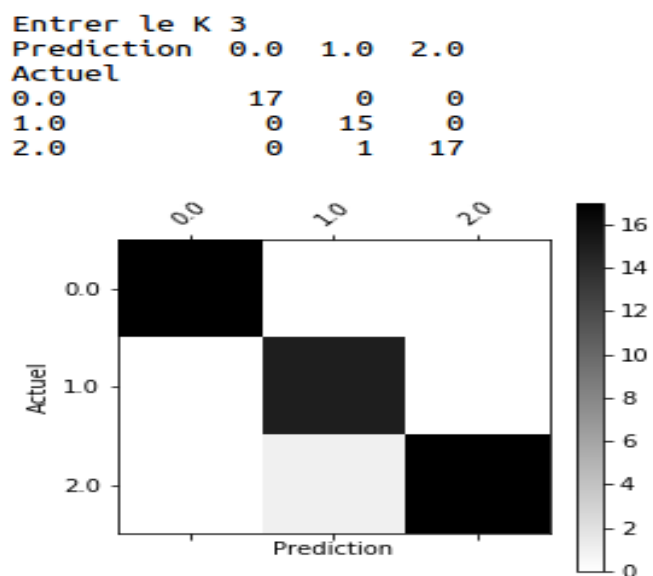
9.0.1 Matrice de confusion

- Classification avec K=6



Avec K=6 nous remarquons que 2 éléments ont été mal classés dans la classe 2.0 et pour bien illustrer cette erreur nous l'avons encadré en rouge.

- Classification avec K=3



Avec $K=3$ nous remarquons que l'erreur précédente a été corrigée et pour bien illustrer cette correction nous l'avons encadrée en rouge. Les 2 éléments qui étaient mal classés dans la classe 2.0 sont maintenant bien classés.

Dans la matrice de confusion les éléments de la diagonale représentent les éléments bien prédits par le système.

10 Conclusion

Dans ce projet il s'agissait pour nous de mettre en place un programme de reconnaissance de caractères manuscrits. Pour la mise en place de ce projet, nous avons procédé par une analyse du sujet afin de comprendre le contexte la problématique ainsi que les objectifs à atteindre du projet. Par la suite nous avons l'état de l'art, c'est à dire une étude des travaux connexe en faisant une comparaison sur ces différentes techniques et outils utilisés pour la mise en place effective du projet. Dans la solution proposée, nous avons utilisés l'algorithme KNN (k plus proche voisin) qui est l'un des algorithme les plus utilisé dans le domaine d'apprentissage supervisé. L'implémentation de la solution proposée a été faite en utilisant le langage de programmation python. Avec le KNN, nous avons extrait les caractéristiques de l'image en calculant la distance euclidienne de de l'objet par rapport aux autres objets afin de trouver la distance la plus petite. Pour la phase d'implémentation, nous avons à notre disposition un dataset que nous avons divisé en données d'apprentissage (75%) et en données de test (25%). Pour la phase de test nous avons utilisé d'autre types de données afin de connaître la performance de notre modèle. Nous avons donc utilisé la matrice de confusion pour déterminer la performance de notre modèle et nous obtenons un taux de 95%. En somme nous pouvons dire que l'étude de cet projet nous a été bénéfique dans la connaissance des différentes méthodes existant dans le domaine de machine learning, donc la reconnaissance de forme. A travers nos différentes études nous avons puis réaliser ce projet, cependant il reste toujours à améliorer pour plus de performance.

Comme perspective, nous pouvons dire qu'il sera important de distinguer la couleur des différent caractères manuscrits existant dans l'image

Références

<https://github.com/ZongAFRIC/Reonnaissance-caracteres-manuscrits>