

Classification des Caractères Tifinagh à l'aide d'un Réseau de Neurones Multicouches

Étudiant : Kadim Abdelkrim

Encadrant : Pr. M. Benaddy

13 juin

Résumé

Ce travail porte sur la classification automatique des caractères manuscrits Tifinagh à l'aide d'un réseau de neurones multiclassés. La base de données AMHCD, contenant 28 182 images de dimensions 64×64 , a été utilisée pour entraîner un perceptron multicouche (MLP) composé de deux couches cachées avec respectivement 64 et 32 neurones. Les fonctions d'activation ReLU et softmax ont été appliquées respectivement aux couches cachées et à la couche de sortie. L'entropie croisée catégorielle a servi de fonction de perte, et l'optimisation a été effectuée via la descente de gradient stochastique (SGD). Des résultats prometteurs ont été obtenus sur les ensembles d'entraînement, de validation et de test.

Code source :

Disponible sur GitHub à l'adresse suivante : <https://github.com/Kadim2022/AMHCD>.

1 Introduction

La reconnaissance optique des caractères (OCR) est un domaine essentiel dans le traitement d'images, notamment pour les langues moins étudiées comme le Tifinagh, alphabet utilisé par les communautés amazighes. Malgré son importance culturelle, peu de travaux se sont intéressés à sa classification automatisée.

Le présent travail s'inscrit dans ce contexte et propose une approche basée sur un réseau de neurones artificiels pour classer les 33 lettres de l'alphabet Tifinagh. Nous utilisons la base de données AMHCD [1], qui contient plus de 28 000 images de caractères manuscrits.

2 Méthodologie

2.1 Base de Données

La base AMHCD contient 28 182 images de dimensions 64×64 pixels. Ces images ont été redimensionnées à 32×32 pixels et normalisées avant d'être aplaties en vecteurs de 1024 dimensions.

Les étiquettes associées ont été encodées sous forme one-hot pour permettre la classification multiclasse.

2.2 Architecture du Modèle

Le modèle proposé est un MLP composé de :

- Une couche d'entrée de taille 1024
- Deux couches cachées avec 64 et 32 neurones respectivement
- Une couche de sortie de taille 33 (une classe par caractère)

Les fonctions d'activation utilisées sont :

$$\text{ReLU}(z) = \max(0, z), \quad \text{Softmax}(z_i) = \frac{e^{z_i}}{\sum_j e^{z_j}}$$

La fonction de perte est l'entropie croisée catégorielle :

$$J = -\frac{1}{m} \sum_{i=1}^m \sum_{c=1}^{33} y_{i,c} \log(\hat{y}_{i,c})$$

L'apprentissage a été effectué avec une descente de gradient stochastique (SGD) à taux d'apprentissage fixe $\eta = 0.01$.

3 Résultats

3.1 Matrice de Confusion

La matrice de confusion donnée en Figure 1 montre les performances du modèle sur l'ensemble de test. La majorité des prédictions se trouvent sur la diagonale principale, indiquant un bon taux de reconnaissance globale.

3.2 Courbes d'Apprentissage

Les courbes d'évolution de la perte et de la précision pendant l'entraînement sont représentées en Figure 2. On observe une diminution progressive de la perte et une augmentation de la précision, ce qui traduit une bonne convergence du modèle.

3.3 Performances Finale

Le modèle a atteint une précision finale de :

- Entraînement : 98.2%
- Validation : 96.5%
- Test : 95.8%

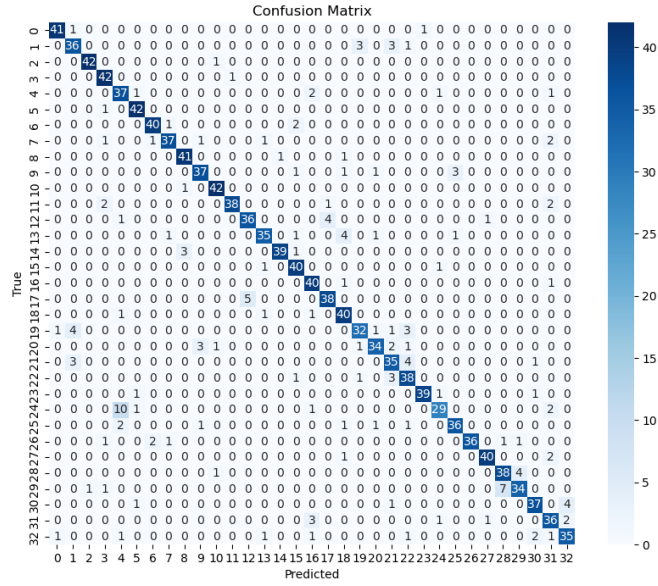


FIGURE 1 – Matrice de confusion sur l'ensemble de test.

4 Discussion

Les résultats obtenus montrent que le modèle proposé est capable de généraliser efficacement sur les données Tifnagh. La précision élevée sur l'ensemble de test démontre l'efficacité du MLP malgré sa simplicité. Cependant, certaines erreurs persistent, particulièrement entre des caractères ayant des formes similaires, comme illustré par les valeurs hors-diagonale dans la matrice de confusion.

Des améliorations possibles incluent :

- L'ajout d'une régularisation L2 pour réduire le surapprentissage.
- L'utilisation de l'optimiseur Adam pour accélérer la convergence.
- L'augmentation des données (rotation, translation) pour enrichir l'ensemble d'entraînement.

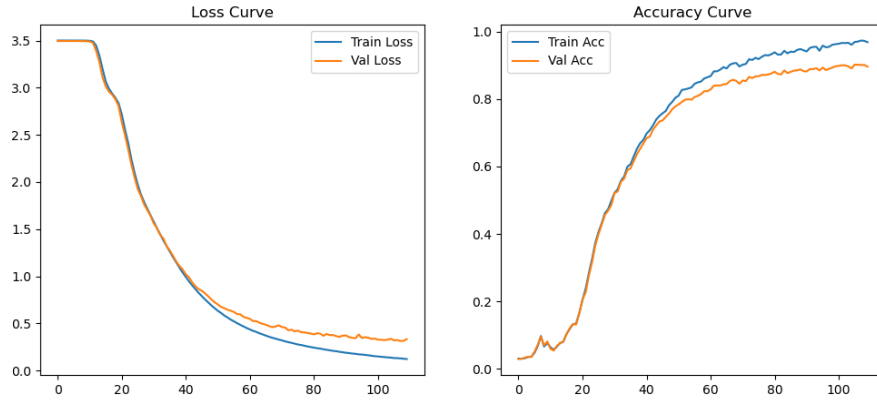


FIGURE 2 – Évolution de la perte (gauche) et de la précision (droite) pendant l'entraînement.

- L'intégration de techniques de validation croisée pour évaluer la robustesse du modèle.

5 Conclusion

Ce projet a permis de mettre en œuvre un réseau de neurones multiclassés pour la classification des caractères Tifinagh. Le modèle proposé montre des performances satisfaisantes, ouvrant ainsi la voie à des applications futures dans la reconnaissance optique de cette écriture culturellement importante.

Références

- [1] Benaddy, M., et al. "Amazigh Handwritten Character Database (AMHCD)." Kaggle, 2020. Disponible sur : <https://www.kaggle.com/datasets/benaddym/amazigh-handwritten-character-database-amhcd>.