

Rapport du mini-projet

Module : Bases de Données Multimedia

*Sujet: Reconnaissances des chiffres en
utilisant des algorithmes de machine
learning*



Réalisé par :

Yasser Kouddan

Fatim Ezzahera Belkoudia

Encadré par :

Mr.El Ayachi Rachid

Année universitaire : 2022 / 2023

Sommaire

Chapitre 1 : Introduction.....	3
Qu'est-ce qu'une image?.....	3
Qu'est-ce que le traitement d'image?.....	3
Qu'est-ce que l'intelligence artificielle?.....	3
Qu'est-ce que le machine learning?.....	4
Les types du machine learning.....	5
Les étapes du projet:.....	5
Présentation des outils.....	5
I. Python.....	5
II. OpenCV.....	6
III. NumPy.....	6
IV. Matplotlib.....	6
V. Pandas.....	7
VI. Spyder.....	7
Chapitre 2 : Prétraitement.....	8
En quoi consiste le prétraitement pour la reconnaissance d'images?.....	8
Chapitre 3 : Extraction.....	10
Les moments de Hu.....	10
Chapitre 4 : Classification.....	12
En quoi consiste la classification pour la reconnaissance d'images?.....	12
En quoi consiste l'algorithme KNN ?.....	15
Chapitre 5 : Résultats.....	16
Résultats de l'algorithme KNN codé.....	16
Résultats de l'algorithme KNN prédéfinie.....	16
Résultats de l'algorithme RN codé.....	16
Résultats de l'algorithme RN prédéfinie.....	16
Chapitre 6 : Conclusion.....	17

Chapitre 1 : Introduction

Qu'est-ce qu'une image?

Une image est un ensemble de données numériques qui représente une scène visuelle ou une photographie. Elle est généralement composée de pixels, qui sont des points individuels de couleur et d'intensité lumineuse. Les images peuvent être en noir et blanc ou en couleur, et peuvent être stockées dans différents formats, tels que JPEG, PNG, GIF, TIFF, etc. Les images peuvent être créées à l'aide d'appareils photo numériques, de scanners ou de logiciels de dessin, et peuvent être modifiées et traitées à l'aide de logiciels d'édition d'images.

Qu'est-ce que le traitement d'image?

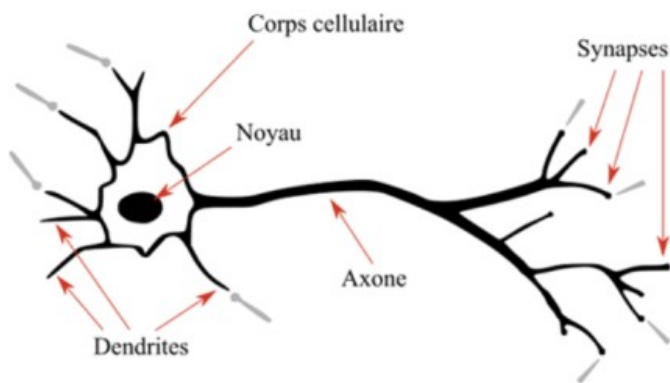
Le traitement d'images est un domaine très vaste qui a connu, et qui connaît encore, un développement important depuis quelques dizaines d'années. On désigne par traitement d'images numériques l'ensemble des techniques permettant de modifier une image numérique afin d'améliorer ou d'en extraire des informations. De ce fait, le traitement d'images est l'ensemble des méthodes et techniques opérant sur celles-ci, dans le but de rendre cette opération possible, plus simple, plus efficace et plus agréable, d'améliorer l'aspect visuel de l'image et d'en extraire des informations jugées pertinentes.

Qu'est-ce que l'intelligence artificielle?

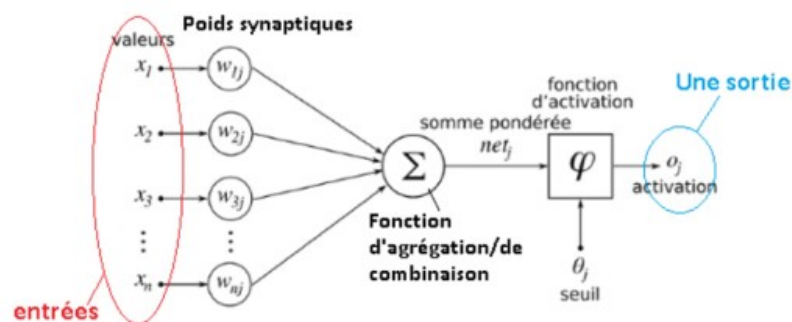
L'intelligence artificielle (IA) est un domaine de l'informatique qui vise à créer des systèmes capables de réaliser des tâches qui nécessitent normalement une intelligence humaine, telles que la reconnaissance de la parole ou d'images, la compréhension du langage naturel, la prise de décision et l'apprentissage automatique. L'IA utilise des techniques mathématiques, statistiques et computationnelles pour simuler et améliorer les capacités cognitives de l'être humain, telles que la perception, la mémoire, la résolution de problèmes et la compréhension de la complexité. Les systèmes d'IA peuvent être basés sur des réseaux de neurones, des arbres de décision, des algorithmes génétiques, des systèmes experts et d'autres méthodes. L'IA est un domaine en constante évolution qui a un impact croissant sur de nombreux aspects de la vie moderne, allant de la robotique à la santé en passant par les finances et les transports.

Qu'est-ce que sont Les réseaux neuronaux artificiels?

Les réseaux neuronaux artificiels (RNA) sont des systèmes informatiques inspirés de la structure et du fonctionnement des réseaux de neurones biologiques dans le cerveau humain, utilisés pour résoudre des tâches complexes comme la reconnaissance d'images ou la reconnaissance de la parole. Un RNA est composé de nœuds appelés neurones artificiels connectés entre eux via des arcs, il utilise des calculs mathématiques pour traiter les informations numériques en entrée et résoudre des tâches. En somme, un RNA est une simulation informatique d'un réseau neuronal biologique.



NEURONE BIOLOGIQUE



NEURONE ARTIFICIEL

Comparaison entre un neurone biologique et artificiel

Qu'est-ce que le machine learning?

Le Machine Learning (apprentissage automatique) est un sous-domaine de l'intelligence artificielle qui vise à créer des systèmes informatiques capables d'apprendre à partir de données et d'améliorer leur performance sans être explicitement programmés. Il s'agit d'un processus automatisé d'apprentissage qui permet aux ordinateurs de découvrir des modèles cachés dans les données, de les utiliser pour effectuer des tâches spécifiques et de s'améliorer avec le temps.

Les types du machine learning

Il existe différents types d'apprentissage automatique, notamment l'apprentissage supervisé, où les données d'entraînement comprennent des étiquettes ou des résultats attendus, et l'apprentissage non supervisé, où les données d'entraînement ne comprennent pas d'étiquettes. Il existe également des méthodes d'apprentissage renforcé où l'apprentissage est guidé par des récompenses.

Les étapes du projet:

Pour créer un algorithme de reconnaissance de chiffres utilisant l'extraction, on va suivre les étapes suivantes:

- **Prétraitement des données:** On obtient un ensemble de données d'images de chiffres et effectuez un prétraitement approprié pour préparer les images pour l'analyse. Cela peut inclure une normalisation de la taille de l'image, un redimensionnement, un filtrage, etc.
- **Extraction:** On applique les algorithmes d'extraction des moments de Hu et Tchebychev sur chaque image pour obtenir un ensemble de caractéristiques numériques représentant chaque image.
- **Classification:** On utilise des données d'entraînement (80%) pour entraîner des modèles de classification KNN et RN en utilisant les caractéristiques extraites des moments de Tchebychev.
- **Évaluation et Test:** Utilisez l'ensemble de données de test (20%) pour évaluer les performances des modèles KNN et RN en utilisant des métriques appropriées, telles que la précision et le rappel.

Présentation des outils

I. Python

Python est un langage de programmation open-source, facile à utiliser et flexible, utilisé pour diverses tâches comme la programmation web, la science des données, l'apprentissage automatique et la création de scripts. Il a une syntaxe lisible et une grande communauté de développeurs et bibliothèques étendent

ses fonctionnalités. Il est utilisé dans de nombreux domaines tels que la finance, la météorologie, les jeux vidéo, l'IA, la recherche et le développement de logiciels.



II. OpenCV

OpenCV est une bibliothèque libre, initialement développée par Intel, spécialisée dans le traitement d'images en temps réel. La société de robotique Willow Garage, puis la société ItSeez se sont succédé au support de cette bibliothèque.



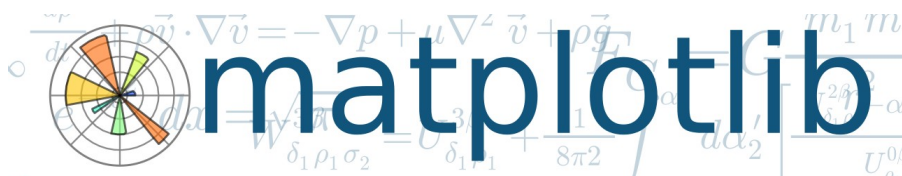
III. NumPy

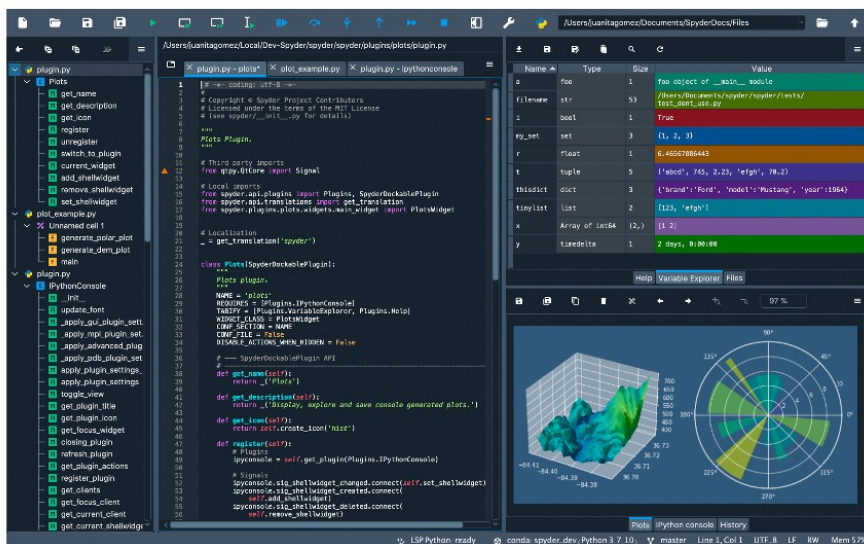
NumPy est une bibliothèque pour langage de programmation Python, destinée à manipuler des matrices ou tableaux multidimensionnels ainsi que des fonctions mathématiques opérant sur ces tableaux.



IV. Matplotlib

Matplotlib est une bibliothèque du langage de programmation Python destinée à tracer et visualiser des données sous forme de graphiques. Elle peut être combinée avec les bibliothèques python de calcul scientifique NumPy et SciPy.





Chapitre 2 : Prétraitement

En quoi consiste le prétraitement pour la reconnaissance d'images?

Le prétraitement en reconnaissance d'images consiste à préparer les images avant de les utiliser pour l'apprentissage automatique ou la reconnaissance. Cela peut inclure des étapes telles que la conversion en niveaux de gris, le redimensionnement, la suppression du bruit, la normalisation, la segmentation, etc. Le but est de maximiser les informations utiles de l'image et de minimiser les perturbations pour améliorer les performances de reconnaissance.

Les étapes du prétraitement

Transformation aux niveaux de gris

La transformation d'une image en niveaux de gris consiste à remplacer les valeurs de couleur RGB de chaque pixel par une seule valeur de niveau de gris, généralement obtenue en utilisant une formule mathématique qui pondère les composantes de couleur en fonction de leur sensibilité relative à l'œil humain. Cette opération permet de simplifier l'image, de réduire la taille des données et de rendre l'analyse des données plus facile, mais elle peut également entraîner une perte d'informations de couleur.

Binarisation

La binarisation est une technique de traitement d'image qui consiste à convertir une image en noir et blanc en utilisant un seuil de luminosité défini. Tous les pixels de l'image qui ont une luminosité supérieure au seuil sont remplacés par blanc et tous les pixels qui ont une luminosité inférieure au seuil sont remplacés par noir. Cette technique permet de simplifier l'image et de mettre en évidence les zones d'intérêts, mais elle peut également entraîner une perte d'informations et un bruit sur l'image.

Redimensionner l'image

Le redimensionnement d'une image consiste à changer la taille de cette image en modifiant le nombre de pixels qui la composent. Cela peut être fait en augmentant ou en réduisant la résolution de l'image, en modifiant les proportions de l'image ou en utilisant des techniques d'interpolation pour ajouter ou supprimer des pixels de manière à conserver le contenu de l'image. Le redimensionnement d'une image peut être utilisé pour adapter les images aux besoins d'un projet ou d'une application spécifique, mais il peut également entraîner une perte de qualité de l'image.

Inverser l'image

L'inversion d'une image est une technique qui consiste à remplacer chaque pixel par son inverse colorimétrique pour obtenir une image négative. Cela peut être fait en inversant les valeurs de chaque composante de couleur ou en utilisant une formule mathématique. Il peut être utilisé pour mettre en évidence certaines caractéristiques de l'image ou pour créer des effets visuels spéciaux, mais il peut également entraîner une perte d'informations et un bruit sur l'image.



Exemple de prétraitement

Chapitre 3 : Extraction

Les moments de Hu

Les moments de Hu sont une série de 7 invariants de forme définis pour une image binarisée ou une région d'intérêt d'une image. Ces invariants de forme sont mathématiquement définis à partir des moments d'ordre supérieur de l'image, et ils peuvent être utilisés pour caractériser la forme, la taille, l'orientation et la symétrie d'une région d'intérêt. Les moments de Hu sont souvent utilisés dans la reconnaissance d'images pour identifier et classer des objets dans une image, car ils sont insensibles à des transformations comme la rotation, la translation et l'échelle. Les moments d'image standard sont implémentés dans OpenCV via la fonction **cv2.moments**. Les Hu Moments sont disponibles via la fonction **cv2.HuMoments**.

Le type de moment le plus simple que nous pouvons définir est donné ci-dessous

$$M = \sum_x \sum_y I(x, y)$$

Ces moments sont souvent appelés **moments bruts**, pour un moments plus complexes on écrit

$$M_{ij} = \sum_x \sum_y x^i y^j I(x, y)$$

Les moments centraux:

$$\mu_{ij} = \sum_x \sum_y (x - \bar{x})^i (y - \bar{y})^j I(x, y)$$

avec

$$\bar{x} = \frac{M_{10}}{M_{00}}$$
$$\bar{y} = \frac{M_{01}}{M_{00}}$$

Les 7 moments sont calculés à l'aide des formules suivantes :

$$h_0 = \eta_{20} + \eta_{02}$$

$$h_1 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$$

$$h_2 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$$

$$h_3 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$$

$$h_4 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$

$$h_5 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})]$$

$$h_6 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$

Les moments de Tchebychev

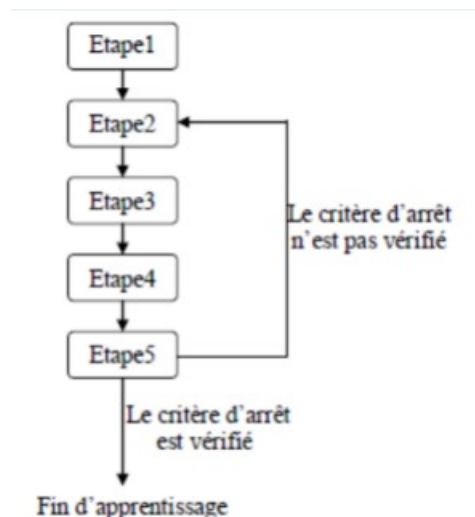
Les moments de Tchebychev sont une autre série d'invariants de forme utilisés pour caractériser des objets dans une image. Ils sont également appelés moments de Tchebychev-Vandermonde ou moments d'intégrales de Tchebychev. Ces moments sont définis comme des intégrales sur une région d'intérêt d'une image, où la fonction de pondération est un polynôme de Tchebychev. Les moments de Tchebychev sont souvent utilisés pour la reconnaissance des formes, car ils sont insensibles à certaines transformations géométriques, comme la rotation, la translation et la symétrie. Cependant, ils sont moins utilisés que les moments de Hu car ils nécessitent un calcul plus complexe.

Chapitre 4 : Classification

En quoi consiste la classification pour la reconnaissance d'images?

L'étape de classification en reconnaissance d'images consiste à utiliser un modèle d'apprentissage automatique pour assigner une classe ou une étiquette à une image ou une région d'intérêt d'une image. Ce modèle est généralement entraîné sur des données d'entraînement étiquetées, où chaque image est associée à une étiquette ou une classe spécifique. Une fois entraîné, le modèle peut être utilisé pour classer des images non étiquetées en utilisant des caractéristiques extraites de ces images.

Les étapes d'apprentissage d'un réseaux de neurones



Etape 1 : Initialisation des poids des connexions

Initialisation des poids a des petites valeurs et des seuils du réseau.

Etape 2 : Propagation des entrées

Les entrées Z , sont présentées a la couche d'entrée : $X_i = E_i$, La propagation vers la couche cachée se fait a l'aide de la formule suivante :

$$Y_j = f\left(\sum_{i=1}^n X_i V_{ji} + X_0\right) = \frac{1}{1 + \exp\left(-\left(\sum_{i=1}^n X_i V_{ji} + X_0\right)\right)}$$

Ensuite de la couche cachée vers la couche de sortie, on adopte:

$$Z_k = f\left(\sum_{j=1}^m Y_j W_{kj} + Y_0\right) = \frac{1}{1 + \exp\left(-\left(\sum_{j=1}^m Y_j W_{kj} + Y_0\right)\right)}$$

X0 et Y0 sont des seuils, f est la fonction d'activation (fonction sigmoïde)

$$f(a) = \frac{1}{1 + \exp(-a)}$$

avec

$$a = \sum_{i=1}^n X_i V_{ji} + X_0 \quad (\text{Le cas de la couche cachée})$$

Et

$$a = \sum_{j=1}^m Y_j W_{kj} + Y_0 \quad (\text{Le cas de la couche de sortie})$$

Etape 3 : Rétropropagation de l'erreur

Au niveau de la couche de sortie, l'erreur entre la sortie désirée S_k , et la sortie réelle Z_k , est calculée par:

$$E_k = Z_k(1 - Z_k)(S_k - Z_k)$$

L'erreur calculée est propagée sur la couche cachée en utilisant la formule suivante:

$$F_j = Y_j(1 - Y_j) \sum_{k=1}^r W_{kj} \cdot E_k$$

Etape 4 : Correction des poids des connexions

On corrige les poids de connexions entre la couche d'entrée et la couche cachée par:

$$\Delta V_{ji} = \eta X_i F_j \quad \text{et} \quad \Delta Y_0 = \eta F_j$$

Puis, on modifie les connexions entre la couche cachée et la couche de sortie par :

$$\Delta W_{kj} = \eta Y_j E_k \quad \text{et} \quad \Delta X_0 = \eta E_k$$

Avec η une constante positive appelée taux d'apprentissage (un paramètre à déterminer empiriquement).

Etape 5 : Boucle

Boucler a l'étape 2 jusqu'a l'obtention d'un critère d'arrêt a définir (seuil d'erreur, nombre d'itérations).

Après l'apprentissage et l'exécution d'OCR, on utilise la distance euclidienne pour identifier les caractères de Tifinaghe :

$$d(t_k, o) = \left(\sum_{i=1}^r (t_{ki} - o_i)^2 \right)^{1/2}$$

Avec T_k la sortie désirée et o la sortie du réseau.

En quoi consiste l'algorithme KNN ?

KNN (K-Nearest Neighbors) est un algorithme de classification qui fonctionne en comparant une image à classifier avec les images de la base de données d'entraînement qui sont les plus proches d'elle selon un certain critère de distance. Les k images les plus proches (déterminées par un hyperparamètre k) sont utilisées pour déterminer la classe de l'image à classifier. La classification se fait en attribuant à l'image la classe qui apparaît le plus souvent parmi les k plus proches voisins.

En utilisant des algorithmes de KNN pour la reconnaissance d'images, il est possible de classifier automatiquement les images en différentes catégories telles que les animaux, les véhicules, les objets ou les personnes. Les algorithmes KNN sont souvent utilisés en combinaison avec des algorithmes de traitement d'images pour améliorer la précision de la reconnaissance et pour éviter les erreurs de classification. En général, KNN est considéré comme un algorithme rapide et simple pour la reconnaissance d'images, ce qui en fait un choix populaire pour de nombreuses applications de reconnaissance de la vision.

Chapitre 5 : Résultats

Résultats de l'algorithme KNN codé

	HU	Tchebychev
Taux de classification	90.48%	13.73%
Taux d'erreur	9.52%	86.27%

Résultats de l'algorithme KNN prédéfinie

	HU	Tchebychev
Taux de classification	46.48%	53.81%
Taux d'erreur	53.52%	46.19%

Résultats de l'algorithme RN codé

	HU	Tchebychev
Taux de classification	10%	11.82%
Taux d'erreur	90%	88.18%

Résultats de l'algorithme RN prédéfinie

	HU	Tchebychev
Taux de classification	6%	8.5%
Taux d'erreur	94%	91.5%

Chapitre 6 : Conclusion

Pour conclure, nous pouvons dire que le traitement d'images est un domaine complexe mais passionnant de l'informatique qui utilise des techniques mathématiques pour analyser, modifier et améliorer les images numériques. Les concepts tels que la transformation en niveaux de gris, la binarisation et le redimensionnement d'images, ainsi que l'inversion d'images, en sont des exemples pertinents. Les algorithmes d'apprentissage automatique, les réseaux de neurones, le KNN et les moments de Hu et Tchebychev, jouent également un rôle crucial dans le traitement d'images en fournissant des moyens de classifier et de caractériser les images de manière plus efficace et plus précise. Ces techniques sont souvent utilisées pour améliorer la qualité de l'image, en extraire des informations ou pour créer des effets visuels spéciaux. Le développement continu de ces techniques est crucial pour l'évolution de nombreuses applications telles que la reconnaissance de la vision, la médecine, les sciences et les médias.