

# Rapport sur la Classification d'Images

Notre projet est de faire du Deep Learning sur 3 bases de donnée.

## 1. Classification Dogs & Cats

### Objectif

L'objectif de ce projet est de classer des images de chiens et de chats en utilisant un réseau de neurones convolutifs (CNN). Ce type de modèle est particulièrement adapté pour la reconnaissance d'images en raison de sa capacité à extraire des caractéristiques visuelles importantes.

### Méthodologie

Le prétraitement des images a été une étape essentielle pour assurer un bon entraînement du modèle. Les images ont été normalisées en mettant à l'échelle leurs valeurs de pixels entre 0 et 1. Une augmentation des données a également été appliquée afin de rendre le modèle plus robuste aux variations dans l'apparence des chiens et des chats.

L'architecture du modèle repose sur plusieurs couches de convolution, qui permettent d'extraire les caractéristiques des images, suivies de couches de pooling pour réduire la dimensionnalité et améliorer la généralisation du modèle. Enfin, des couches denses sont utilisées pour la classification finale, avec une activation sigmoïde permettant de distinguer les chiens des chats.

L'entraînement du modèle a été réalisé en utilisant l'optimiseur Adam, qui est efficace pour ce type de problème, et la fonction de perte binaire cross-entropy. Après plusieurs époques d'entraînement, une évaluation du modèle a été effectuée sur un ensemble de test afin de mesurer sa capacité à généraliser ses prédictions à de nouvelles images.

## Résultats

Le modèle a obtenu une précision supérieure à 80 % sur l'ensemble de validation, indiquant une bonne capacité à différencier les chiens et les chats. Une analyse des erreurs a permis d'identifier certains cas où le modèle confond les deux catégories, notamment lorsque les images sont floues ou que l'angle de prise de vue est atypique. Une amélioration du modèle pourrait être envisagée en utilisant des architectures plus complexes comme les réseaux pré-entraînés.

## 2. Classification Intel Images

### Objectif

Ce projet vise à classer des images de scènes naturelles en différentes catégories, telles que les forêts, les bâtiments, les glaciers, les montagnes et les rues urbaines. La classification d'images de paysages est un défi intéressant, car certaines classes peuvent présenter des similarités visuelles qui rendent la distinction plus difficile.

### Méthodologie

Le prétraitement des données a consisté en un redimensionnement des images pour les normaliser à une taille commune et en une mise à l'échelle des pixels. Ces transformations ont permis de s'assurer que le modèle puisse traiter les images de manière homogène.

L'architecture du modèle repose sur un réseau de neurones convolutifs comportant plusieurs couches de convolution et de pooling, permettant d'extraire progressivement les caractéristiques des images. La dernière couche applique une activation softmax afin de classer les images en plusieurs catégories distinctes. L'entraînement du modèle a été réalisé en utilisant la fonction de perte categorical cross-entropy, particulièrement adaptée aux problèmes de classification multi-classes.

L'évaluation du modèle a été effectuée en mesurant sa précision globale et en générant une matrice de confusion pour observer les erreurs commises entre les différentes classes.

## Résultats

Le modèle a réussi à classifier les images avec une précision globale satisfaisante, bien que certaines confusions aient été observées entre certaines classes visuellement proches, comme les montagnes et les glaciers. Une analyse des erreurs a montré que l'intégration de données supplémentaires et l'utilisation de modèles plus avancés, comme les réseaux de neurones pré-entraînés, pourraient améliorer les performances.

## 3. Classification Digits

### Objectif

Ce projet porte sur la reconnaissance de chiffres manuscrits à partir de la base de données MNIST. Il s'agit d'un problème classique en apprentissage automatique et en vision par ordinateur, utilisé pour tester l'efficacité des réseaux de neurones convolutifs.

### Méthodologie

Les images ont été normalisées afin de s'assurer que les valeurs des pixels soient comprises entre 0 et 1. Cette normalisation est une étape clé pour améliorer l'apprentissage du modèle et éviter les disparités dans les données d'entrée.

L'architecture du modèle est basée sur un réseau de neurones convolutifs comportant plusieurs couches de convolution et de pooling. Ces couches permettent d'extraire les caractéristiques essentielles des chiffres manuscrits, en capturant leurs formes et leurs courbes distinctives. Une couche finale dense, avec une activation softmax, permet d'attribuer chaque image à l'une des dix catégories correspondant aux chiffres de 0 à 9.

L'entraînement a été réalisé en utilisant l'optimiseur Adam et la fonction de perte categorical cross-entropy. Après plusieurs époques d'entraînement, une évaluation du modèle a été réalisée afin de mesurer sa performance sur un ensemble de validation.

## Résultats

Le modèle a atteint une précision proche de 99 % sur l'ensemble de validation, démontrant ainsi son efficacité pour la reconnaissance des chiffres manuscrits. L'analyse des erreurs a révélé que les confusions les plus fréquentes concernaient des chiffres aux formes similaires, comme le 3 et le 8 ou le 4 et le 9. Une amélioration pourrait être apportée en utilisant des architectures plus complexes ou en appliquant des techniques de régularisation supplémentaires pour limiter le sur-apprentissage.

## Conclusion

Ces trois projets illustrent l'efficacité des réseaux de neurones convolutifs pour la classification d'images. Chaque modèle a été conçu en fonction des spécificités des données traitées et a démontré des performances intéressantes. La classification des chiens et des chats a montré que même un problème simple peut présenter des défis liés à la diversité des images et aux variations des poses. La classification des paysages naturels a révélé l'importance d'un bon prétraitement des images et d'une architecture de réseau adaptée pour différencier des classes visuellement proches. Enfin, la reconnaissance des chiffres manuscrits a permis d'atteindre une très haute précision, démontrant ainsi la puissance des CNN pour ce type de tâche.

Des améliorations potentielles pourraient être envisagées, comme l'intégration de modèles pré-entraînés, l'augmentation des données ou l'utilisation de techniques d'optimisation avancées. Ces éléments pourraient permettre d'améliorer encore davantage les performances des modèles et d'étendre leur application à d'autres types de classification d'images.