

Livrable 1 :

Plateforme de Prédiction Géographique Basée sur l'IA pour l'Oasis de Tafilalet à partir d'Imagerie Satellitaire

(Partie ML/DL)

I. Introduction :

Ce projet vise à développer une plateforme intelligente pour prédire l'évolution géographique de l'oasis de Tafilalet grâce à l'analyse d'images satellites. Ce livrable 1 présente une enquête sur les méthodes de classification des images satellites, essentielles pour déterminer l'occupation du sol. Ces approches permettront d'éclairer les transformations spatiales et de soutenir la gestion durable des ressources naturelles de cet environnement unique.

II. Méthodes de Classification :

1. **Deep learning** : couramment utilisé pour la classification d'images satellitaires grâce à sa capacité à traiter de grandes quantités de données et à identifier des caractéristiques complexes.
 - a) **Convolutional Neural Networks (CNN)** : Ils permettent d'extraire automatiquement les caractéristiques des images, sans avoir besoin de faire d'ingénierie des features à la main.
 - b) **U-Net** : U-Net est particulièrement adapté pour la segmentation d'images satellitaires. Il permet de classer chaque pixel d'une image, ce qui est crucial pour déterminer l'occupation du sol sur des cartes géographiques.
 - c) **SegNet** : est connu pour son efficacité en termes de mémoire, ce qui est utile pour les images satellitaires haute résolution
2. **Machine learning** : Les méthodes de machine learning classiques restent utiles, notamment dans les environnements où la puissance de calcul ou les volumes de données sont limités.
 - a) **Support Vector Machines (SVM)** : Utilisé pour la classification binaire et multiclass, SVM est efficace pour les images satellitaires, surtout lorsqu'il est combiné avec des méthodes de réduction de la dimensionnalité comme PCA.

- b) **Random Forest (Forêts Aléatoires)** : Un algorithme basé sur les arbres décisionnels, très populaire pour la classification des images satellitaires. Il est robuste face au bruit et aux variations dans les données.,
 - c) **K-Nearest Neighbors (K-NN)** : C'est une méthode simple mais souvent efficace pour la classification de petites zones d'images. Elle est facile à implémenter et peut être utile dans des contextes moins complexes.
3. **Méthodes Statistiques Simplifiées** : Les approches statistiques classiques peuvent également être employées pour la classification des images satellitaires dans des environnements avec peu de données ou des calculs limités.
- a) **K-means clustering** : Permet de regrouper des pixels similaires en différentes classes sans avoir à leur attribuer d'étiquettes. Cela peut être utile pour détecter des changements géographiques non identifiés.

III. Supervisé vs Non Supervisé :

1. Approches Supervisées :

Les méthodes supervisées nécessitent des données labellisées, c'est-à-dire que chaque image d'entraînement doit avoir une étiquette correspondant à la classe d'occupation du sol.

-Exemple de techniques :

CNN, U-Net, SVM : Ces algorithmes sont efficaces lorsque des données labellisées sont disponibles. Par exemple, une image satellite de l'oasis de Tafilalet pourrait être classée en plusieurs catégories : végétation, zones désertiques, eau, zones urbanisées.

-Avantages : Haute précision lorsqu'il existe un grand volume de données labellisées.

-Limites : Les données labellisées sont coûteuses à obtenir, surtout pour les images satellitaires de haute résolution.

2. Approches Non Supervisées

Les méthodes non supervisées ne nécessitent pas de données labellisées. Elles peuvent être utilisées pour détecter des structures dans les données sans connaissance préalable des classes.

-Exemple de techniques :

K-means clustering : Permet de regrouper des pixels similaires en différentes classes sans avoir à leur attribuer d'étiquettes. Cela peut être utile pour détecter des changements géographiques non identifiés.

Autoencoders : Un réseau neuronal non supervisé qui apprend à représenter les données sous une forme compressée. Cela peut être utile pour des tâches de réduction de dimensionnalité ou de prétraitement.

-Avantages : Ne nécessite pas de données labellisées et peut découvrir des structures ou des changements non anticipés.

-Limites : La précision est souvent inférieure à celle des méthodes supervisées. Les classes détectées peuvent ne pas correspondre à des catégories significatives pour les experts.

3. Entrées et Sorties

Entrées pour Deep Learning : Les données d'entrée sont généralement des images satellitaires haute résolution sous forme de matrices de pixels. Ces images peuvent inclure plusieurs bandes spectrales (par exemple, visible, infrarouge, etc.).

Sorties pour Deep Learning : Selon l'architecture, les sorties peuvent être :

- Des catégories d'occupation du sol (pour la classification d'image).
- Des masques binaires ou multiclassés (pour la segmentation d'image), où chaque pixel est classifié.

Entrées pour Machine Learning traditionnel : Les entrées incluent souvent des caractéristiques extraites des images, comme des statistiques sur les couleurs, textures, ou indices de végétation comme l'NDVI (Normalized Difference Vegetation Index).

Sorties pour Machine Learning traditionnel : Les sorties incluent généralement des étiquettes de classes d'occupation du sol pour chaque pixel ou segment d'image.

Méthode	Avantages	Limites
CNN, U-Net (Deep Learning)	Hautement performantes sur des données complexes et volumineuses, extraction automatique des caractéristiques.	Besoin d'une grande quantité de données et d'une puissance de calcul élevée.
SVM, Random Forest (Machine Learning)	Rapides à entraîner sur des petites données, efficaces même avec moins de calculs.	Moins performantes sur des images complexes et des données à forte dimensionnalité.
K-means, Autoencoders (Non supervisé)	Ne nécessitent pas de données labellisées, capables de détecter des modèles cachés.	Moins précises, difficultés à interpréter les clusters sans connaissance approfondie.