

# **Livrable 1 Ingénieur ML/DL :**

## **Différentes méthodes pour de la classification du sol**

Notre objectif en tant qu'ingénieur DL/ML est de trouver un modèle et de l'optimiser pour de la segmentation. Les images que nous avons sont multispectrales et conséquente en taille (7000x7000 pixels). Nous avons de la végétation, des habitations, de l'eau, du désert, des montagnes, etc. Nous avons donc un problème avec plusieurs classes, et elles ne sont pas équilibrées. Notre modèle doit donc supporter des images multispectrales, et il doit être capable de faire une segmentation dense et fine.

Nous avons choisi d'étudier un U-Net, un Random Forest, et un k-NN.

### **A.1. U-Net**

Le réseau U-Net est une architecture de réseau neuronal convolutif principalement utilisée pour des tâches de segmentation d'images. Cette architecture est particulièrement appréciée pour sa capacité à capturer à la fois des informations globales et des détails fins, ce qui la rend performante dans les tâches nécessitant une délimitation précise.

Le réseau U-Net prend une image multispectrale comme entrée, où chaque pixel appartient à une certaine classe (par exemple, végétation, eau). L'objectif est de prédire quelle classe appartient à chaque pixel (c'est une tâche de segmentation sémantique).

Grâce à sa structure en "U" avec une phase de contraction (encodeur) et une phase d'expansion (décodeur), U-Net combine les caractéristiques d'image à grande et petite échelle, ce qui est essentiel pour la délimitation fine des objets dans des images complexes telles que celles de la télédétection.

La sortie attendue est une carte de segmentation de taille  $H \times L$ , où chaque pixel contient un indice correspondant à la classe à laquelle il appartient (végétation, sol, eau, etc.). Chaque valeur dans la carte de sortie représente une probabilité d'appartenance à une certaine classe.

### **3. Avantages du réseau U-Net**

L'architecture U-Net est capable de produire de bons résultats même avec un nombre limité d'images annotées grâce à son efficacité dans l'utilisation des informations spatiales.

U-Net excelle dans la détection d'objets de petite taille ou dans la délimitation d'objets ayant des frontières complexes, ce qui est crucial pour la segmentation des types de sol dans des images de télédétection.

#### **4. Limites de la méthode**

Dans le cas d'images multispectrales, les variations spectrales dues à des conditions atmosphériques ou des différences de capteurs peuvent influencer la précision du modèle. De plus si les étiquettes sont imprécises ou bruitées, cela peut affecter les résultats de segmentation.

##### **A.2 Random Forest**

Un Random forest est un modèle de machine learning qui combine plusieurs arbres de décisions, pour obtenir en sortie une meilleure précision qu'avec un seul arbre.

Pour faire de la segmentation, un RF peut prendre en entrée des images satellites multispectrales. Chaque pixel est donc représenté par un vecteur à 7 valeurs, dans notre cas, et le RF va prédire une classe pour chaque pixel. Aucun autre paramètre ne va être pris en compte, contrairement au U-net. La sortie doit être une image labélisée, bien souvent à la main.

Utiliser un RF pour de la segmentation présente plusieurs avantages et inconvénients. Comme son fonctionnement repose sur un ensemble d'arbres de décision, il n'est pas très sensible au surapprentissage. Il est également capable de donner de bons résultats, même si les données sont bruitées, ou si elles sont incomplètes. Ce modèle peut également être utile lorsque les classes sont déséquilibrées.

Cependant, pour traiter des données volumineuses, la mémoire d'un RF devient conséquente, car il faut stocker un certain nombre d'arbres. Ce modèle ne prend pas en compte les dépendances spatiales, ce qui peut être problématique pour de la segmentation d'images satellites. Lorsque le problème de segmentation devient trop important, le RF a besoin de beaucoup d'arbres et d'une grande profondeur sur chacun d'eux, ce qui rend le modèle très lourd.

##### **A.3 k-Nearest-Neighbors**

La méthode des k-Nearest-Neighbors peut permettre de faire de la segmentation, mais ce n'est que rarement retenu, ce n'est pas la méthode la plus efficace.

Cet algorithme définit une classe pour un pixel en fonction des classes attribuées aux pixels voisins. La distance entre les pixels est calculée selon leurs caractéristiques. De plus cet algorithme nécessite un premier prétraitement, il ne peut pas recevoir une image satellite en entrée, il faut donc traiter toutes nos images avant de les envoyer au modèle. Ce traitement nous force à avoir des données très lourdes, pour une image de 1000x1000 pixels nous devons créer un tableau de 1 000 000 de lignes et les colonnes selon le nombre de spectres.

Cet algorithme offre plusieurs avantages, comme sa simplicité de compréhension et d'implémentation, et son apprentissage qui ne nécessite pas une durée très longue. Cependant, pour des jeux de données lourds, il lui faudra beaucoup de temps et d'énergie. Ces performances peuvent aussi être altérées si les images ont trop de spectres.

#### **A.4 Supervisé vs Non supervisé**

Travailler avec des modèles non supervisés a une utilité pour les approches exploratoires, ils peuvent mettre en avant des patterns que l'œil humain n'avait pas vu auparavant. Cependant, il y aura un manque de précision, et il faudra ensuite interpréter les différentes classes, qui peuvent ne pas avoir de signification pour notre travail.

Les modèles supervisés ont besoin de recevoir des données labélisées pour pouvoir être efficaces, mais ils seront normalement meilleurs, plus précis sur les classes et leurs frontières. Les résultats n'ont pas besoin d'être interprétés. Cependant, pour obtenir de bonnes performances, il faut que la labélisation soit de bonne qualité, ce qui peut vite demander beaucoup de temps.

#### **A.5 Fonction de perte**

Le choix de la fonction de perte est aussi important. Selon nos besoins, nous n'allons pas choisir la même. Dans notre cas, nous avons des classes déséquilibrées, et nous voulons avoir une segmentation fine et dense. Nous avons donc plusieurs fonctions de perte que nous pouvons choisir.

- **Weighted Cross entropy** : Cette fonction est utile lorsque les classes sont déséquilibrées, en donnant plus de poids aux classes minoritaires. Les erreurs sur ces classes sont plus coûteuses que sur celles surreprésentées.
- **Cross-Entropy Loss** : C'est la fonction de perte standard pour les tâches de classification, ici adaptée pour la segmentation, où elle calcule la perte pixel par pixel entre les prédictions et les vérités terrain.

- **Dice Loss** : Utilisée pour des tâches où il y a un fort déséquilibre entre les classes, la Dice Loss se base sur le Dice Coefficient et pénalise les erreurs sur les classes rares. Son calcul se base sur les différences des prédictions et des vérités terrains.
- **Focal Loss** : En combinaison avec la Cross-Entropy Loss, la Focal Loss permet de mieux gérer les déséquilibres de classes, en accord plus de poids aux exemples mal classés.

Les fonctions de pertes adaptées au Random forest fonctionne sur le même principe, donner plus de poids aux classes minoritaires, en pénalisant plus fortement leurs erreurs.

Nous avons pour le RD la fonction Weighted Gini Impurity, ou encore la Balanced Class Weight.

Pour le k-NN, il n'existe pas de fonctions de pertes à proprement parlé, mais nous pouvons essayer d'ajuster les pondérations des voisins avec Weighted Distance.

## A.6 Métrique

Le choix des métriques nous ait aussi imposé avec les mêmes conditions que pour les fonctions de perte.

- **IoU (Intersection over Union)** : La métrique IoU est l'une des plus utilisées pour la segmentation. Elle mesure l'intersection entre les pixels prédits comme appartenant à une classe spécifique et les pixels de la classe réelle.

$$\text{IoU} = \frac{\text{Aire de l'intersection}}{\text{Aire de l'union}}$$

Un IoU élevé indique que la segmentation prédite correspond bien à la vérité terrain.

- **Dice Coefficient** (ou Dice Similarity Coefficient) : Très similaire à IoU, le Dice Coefficient est souvent utilisé dans les tâches de segmentation pour mesurer la similarité entre les prédictions et les vérités terrain.

$$\text{Dice} = \frac{2 \times |\text{Prédictions} \cap \text{Vérité terrain}|}{|\text{Prédictions}| + |\text{Vérité terrain}|}$$

Le Dice est particulièrement pertinent pour des classes déséquilibrées.

- **Précision pixel-wise** : Cette métrique mesure simplement le pourcentage de pixels correctement classés. Cependant, elle peut être trompeuse dans le cas de classes

déséquilibrées (où une classe dominante pourrait masquer des erreurs sur les classes rares).

### **A.7 Données d'entrées et de sorties**

Les réseaux U-net et les Random Forest peuvent prendre en entrée des images satellites multispectrales, mais ce n'est pas le cas du k-NN. Il nécessite un prétraitement pour convertir l'image en tableaux avec une ligne par pixel, et le nombre de colonne est défini par le nombre de dimensions de l'image.

Les données de sorties doivent être des images labelisées, et cette labélisation se fera à la main. Plusieurs techniques sont possibles. Nous pouvons faire sélectionner un seul pixel par classe, ou alors tracer une ligne de pixel, ou encore faire un polygone représentant des pixels d'un même classe.