

# APPRENTISSAGE INCRÉMENTAL POUR LE PASSAGE A L'ÉCHELLE DES ALGORITHMES D'INTERPRÉTATION DES IMAGES DE TÉLÉDÉTECTION

Marie-Ange BOUM<sup>1,2</sup>, Dawa DERKSEN<sup>2</sup>, Pierre FOURNIER<sup>1</sup>, Stéphane HERBIN<sup>1</sup>  
<sup>1</sup>ONERA/DTIS, <sup>2</sup>CNES

## INTRODUCTION

Les images de télédétection en expansion offrent de nouvelles opportunités pour la cartographie et le suivi à grande échelle de l'utilisation des sols.

### Objectif Global

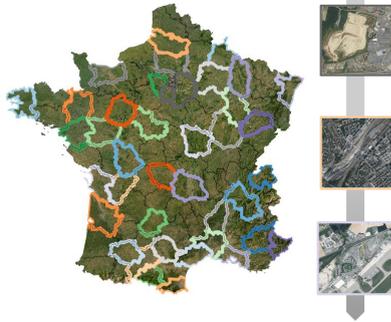
Concevoir un système de classification à grande échelle (ex. France), optimisé en ressources, et capable de s'étendre progressivement à de nouvelles zones géographiques.

### Défis :

1. Adaptation de domaines successives.
2. Atténuation de l'oubli catastrophique.

## FLAIR#1 (French Landcover Aerial ImageRy) [1]

- Images aériennes THR (0,2 m/pixel)
- Variabilité géographique plus ou moins importantes (paysages et climats).
- Domaines spatiaux définis par les départements de la France Métropolitaine.

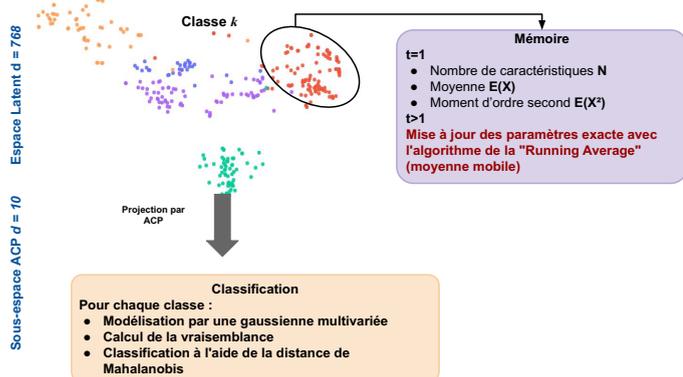


## CLASSIFICATION

### Domain Incremental Probabilistic PCA [2]

Extraction de caractéristiques à l'aide d'un modèle de fondation :

- L'encodeur est fixe : il génère des caractéristiques invariantes.
- Le classificateur linéaire est remplacé par un classificateur génératif afin d'atténuer l'oubli.



## CLASSIFICATION : Expériences

Pre-training Data	SL	SSL	FT	DIPPGA			ER			SUP.
				ACC	BWT	DEGRAD.	ACC	BWT	DEGRAD.	
ImageNet21K	x			91,7	-2	-4,8	90,9	-6,9	-5,6	96,9
ImageNet1K		DINO		91,6	-3,2	-4,1	91	-6,9	-4,8	95,8
SA-1B		MAE		90	-4	-4,1	90,1	-5,6	-3,9	94
MillionAID		MAE		69,3	-24,5	-20,3	74,3	-9,5	-11,4	85,7
RSSCN7	x		x	81,3	-11,6	-7,8	81,7	-8,2	-7,4	89,1

Utilisation des modèles de fondation (FM) pour la télédétection :

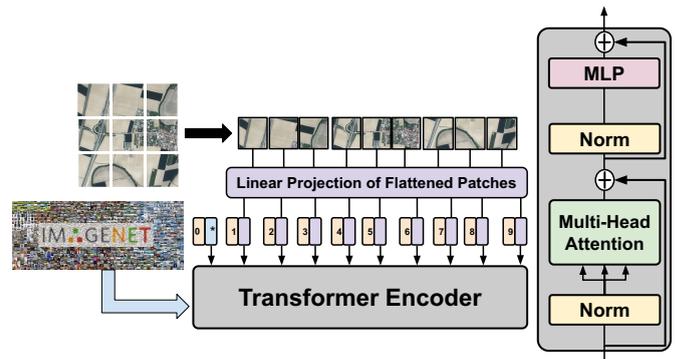
- Les FM pré-entraînés sur des images naturelles sont efficaces pour extraire des caractéristiques expressives.
- L'ajout d'un classificateur génératif réduit l'oubli, à condition que les caractéristiques soient pertinentes.
- Cette méthode surpasse les baselines (ER).

### Limites :

- Les caractéristiques extraites dépendent des données de pré-entraînement.
- Les RSFM sont confrontés au surajustement et à la faible diversité des datasets.

## LES MODÈLES DE FONDATION

- **Pré-entraînement** sur des datasets diversifiés et vastes.
- **Généralisation** : performants sur diverses tâches en aval.
- **Accessibilité** : modèles open-source.

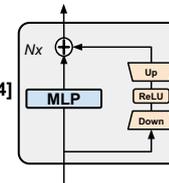


Comment exploiter efficacement les modèles de fondation pour l'apprentissage incrémental en télédétection ? [3]

## SEGMENTATION

### Atténuer l'oubli catastrophique par PEFT

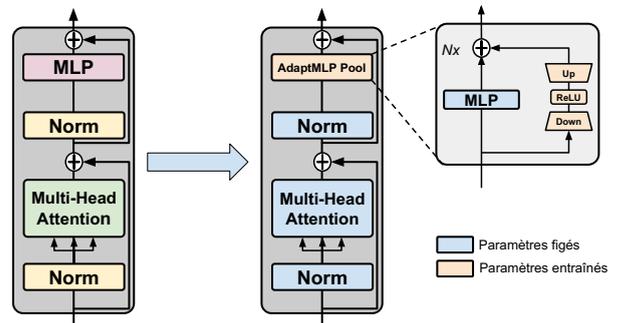
Les **Parameter Efficient Fine Tuning (PEFT)** permettent d'ajuster efficacement un modèle pré-entraîné à de nouvelles tâches ou domaines en ne modifiant que certaines couches tout en gardant la majorité des paramètres fixes.



Choix de PEFT : AdaptMLP [4]

### Approche proposée :

L'idée de base est d'apprendre une série de modules d'adapter, domaine par domaine, à l'aide d'un transformer pré-entraîné.



## 2 stratégies d'apprentissages à tester :

- **Dépendante du passé** : Les nouveaux modules AdaptMLP sont influencés par les paramètres des modules déjà appris.
- **Indépendante du passé** : Chaque nouveau module est appris de manière indépendante des modules précédemment appris.

**Objectif final** : Créer un modèle capable de généraliser sur l'ensemble des domaines spatiaux tout en minimisant l'oubli catastrophique.

## RÉFÉRENCES

1. Garioud, Anatol, et al. "Flair# 1: semantic segmentation and domain adaptation dataset." *arXiv preprint arXiv:2211.12979* (2022).
2. Boum, Marie-Ange, et al. "Continual Learning in Remote Sensing: Leveraging Foundation Models and Generative Classifiers to Mitigate Forgetting." *IGARSS 2024-2024 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2024.
3. Dosovitskiy, Alexey. "An image is worth 16x16 words: Transformers for image recognition at scale." *arXiv preprint arXiv:2010.11929* (2020).
4. Chen, Shoufa, et al. "Adaptformer: Adapting vision transformers for scalable visual recognition." *Advances in Neural Information Processing Systems* 35 (2022): 16664-16678.