

PROPOSITION DE STAGE EN COURS D'ETUDES

Référence : **DEMR-2025-11**
(à rappeler dans toute correspondance)

Lieu : ONERA/Palaiseau

Département/Dir./Serv. : DEMR/SEM

Tél. : 01 80 38 72 23

Responsable(s) du stage : Nathan Letheule,
Solène Debuysere

Email : nathan.letheule@onera.fr
solene.debuysere@onera.fr

DESCRIPTION DU STAGE

Thématique(s) : Imagerie radar, IA générative, télédétection

Type de stage : Fin d'études bac+5 Master 2 Bac+2 à bac+4 Autres

Intitulé : Repousser les limites de l'imagerie radar : Création d'un dataset d'entraînement à très haute résolution pour l'IA à partir d'images satellites

Sujet : Le département DEMR de l'Onera est reconnu comme un expert de premier plan en acquisition et traitement de données radar, en particulier pour les images à très haute résolution. Au sein de ce département, l'unité SEM a été pionnière dans le développement de méthodes d'intelligence artificielle générative, adaptées aux spécificités des images radar dites SAR ([4], [5]), un type d'imagerie aussi particulier que rare.

Les images SAR posent de nombreux défis techniques : elles présentent des statistiques et des dynamiques spécifiques, sont souvent acquises en format complexe, parfois en multipolarisation, et leur taille est très importante (jusqu'à plusieurs dizaines de milliers de pixels de côté). Ces particularités exigent des adaptations rigoureuses pour que les images puissent être exploitées dans des bases de données d'apprentissage conventionnelles, adaptées aux architectures de réseaux de neurones classiques.

Pour ce faire, il est nécessaire de transformer ces données en des extraits plus petits « vignettes », de format standard (par exemple, 512x512 pixels) et d'harmoniser les résolutions, tout en préservant la nature complexe des données et les données auxiliaires associées. Nous avons déjà amorcé le développement de ces algorithmes de préparation de données sur des images aéroportées. Le stagiaire devra dans un premier temps s'approprier ces algorithmes et les approfondir. Dans un second temps, son objectif sera de les adapter et appliquer à une base de données d'images satellites haute résolution en source ouverte, les données UMBRA [1].

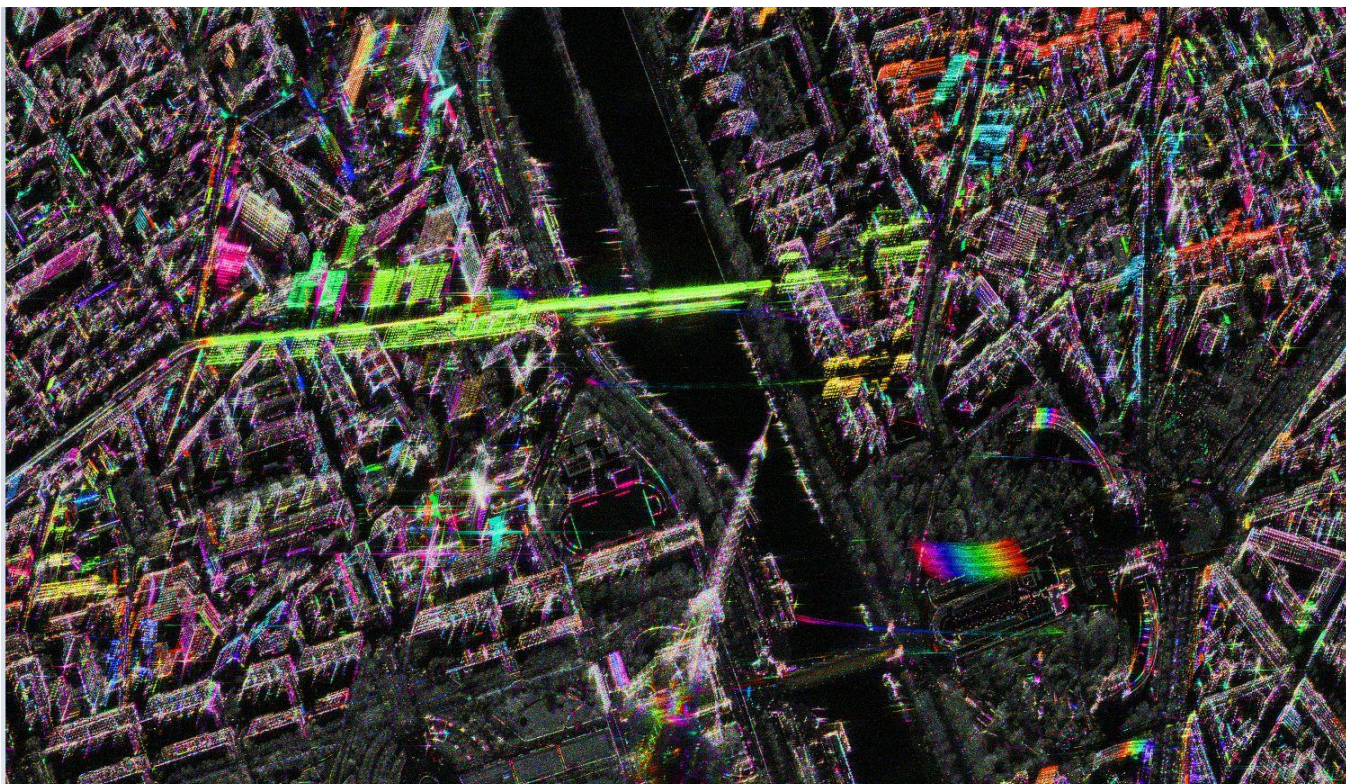
L'un des défis de ce projet sera d'intégrer la fonctionnalité de récupération d'images optiques à l'échelle mondiale. Cette tâche nécessitera de créer ou de perfectionner un module qui utilise des coordonnées géographiques (latitude et longitude) pour obtenir les images correspondantes. Une possibilité envisagée serait d'automatiser cette récupération en utilisant l'API de Google Maps, permettant ainsi d'accéder efficacement aux images optiques en fonction des coordonnées spécifiées.

La mission se poursuivra par la création d'un dataset structuré et adapté, qui sera utilisé pour entraîner des modèles d'IA [2], [3] en imagerie radar. Le ou la stagiaire aura la responsabilité d'organiser et de documenter le dataset de manière à ce qu'il soit facile à exploiter par les équipes de recherche et développement.

Enfin, la dernière phase du stage consistera à tester l'efficacité du dataset créé. Des entraînements de modèles d'IA seront réalisés afin de valider la qualité et l'utilité des données préparées. Le ou la stagiaire pourra également proposer des améliorations ou des optimisations en fonction des résultats obtenus.

Ce stage est une opportunité unique d'acquérir des compétences techniques en traitement d'images radar, en adaptation de code et en manipulation de données géospatiales. Il permettra également de renforcer ses compétences en programmation Python, en traitement des données massives, ainsi qu'en intelligence artificielle générative.

Une dimension importante de ce projet est son caractère open source. À l'issue du stage, le code développé sera publié en open source, permettant ainsi à la communauté scientifique de l'exploiter pour leurs recherches ou applications. Le ou la stagiaire aura donc la responsabilité de documenter soigneusement l'ensemble des codes et des processus, en s'assurant que les outils et méthodes développés soient clairs, bien structurés, et facilement réutilisables par d'autres chercheurs ou ingénieurs.



Composition colorée faite à partir d'un mode UMBRA "dwell" et de l'algorithme REACTIV (la couleur code les angles d'acquisitions)

[1] Umbra. (s.d.). Open data. Consulté le 12 novembre 2024, sur <https://umbra.space/open-data/>

[2] Rombach, R., Blattmann, A., Lorenz, D., Esser, P., and Ommer, B. (2022a). Highresolution image synthesis with latent diffusion models. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pages 10684– 10695.

[3] Fluximage. (s.d.). Consulté le 12 novembre 2024, sur <https://fluximage.org/>

[4] Nicolas Trouvé, Nathan Lethéule, Olivier Lévêque, Ilias Rami, and Elise Colin (Avril 2024) SAR image synthesis using text-conditioned pre-trained generative AI models. EUSAR 2024

[5] Synthesizing SAR Images with Generative AI: Expanding to LargeScale Imagery Solène Debuysere, Nicolas Trouve, Nathan Lethéule, Elise Colin and Olivier Lévêque, RADAR 2024

Est-il possible d'envisager un travail en binôme ? Non

Méthodes à mettre en oeuvre :

- | | |
|---|---|
| <input type="checkbox"/> Recherche théorique | <input checked="" type="checkbox"/> Travail de synthèse |
| <input checked="" type="checkbox"/> Recherche appliquée | <input checked="" type="checkbox"/> Travail de documentation |
| <input type="checkbox"/> Recherche expérimentale | <input checked="" type="checkbox"/> Participation à une réalisation |

Possibilité de prolongation en thèse : Oui

Durée du stage : Minimum : 5 mois Maximum : 6 mois

Période souhaitée : Février à Septembre 2025

PROFIL DU STAGIAIRE

Connaissances et niveau requis : Programmation Python, Mathématiques appliquées, Traitement d'images	Ecoles ou établissements souhaités : M2 ou 3 ^{ème} année d'Ecole d'Ingénieur
---	--