Enhancing Land Cover Classification with Diffusion-based Super-Resolution of Satellite Image Time Series and Aerial Imagery (EN)

Nienburger Straße 1, 30167 Hannover
Fakultät für Bauingenieurwesen und Geodäsie

Institut für Photogrammetrie und GeoInformation

Prof. Dr.-Ing. habil. Christian Heipke

Hubert Kanyamahanga

Tel. +49 511 762 2488
Fax +49 511 762 2483
E-Mail:
kanyamahanga@ipi.uni-hannover.de

26.11.2025

Accurate land cover classification requires both spatial details and temporal information of remote sensing data. While publicly available satellite image time series (SITS) offer short revisit times, they suffer from limited spatial resolution. In contrast, aerial imagery provides fine-grained spatial details, but their temporal coverage is limited. Thus, combining data from those sensors is of interest as their properties are complementary w.r.t. the problem domain. However, the large gap in spatial resolution between these two sensors makes their integration challenging.

The objective of this master thesis is to develop a method to jointly use aerial images and multi-temporal information from co-registered satellite image time series (SITS) **(Figure 1)** to predict, at a pixel-level, the land cover of a given input image at the ground sampling distance (GSD) of aerial images. A baseline model **[1] UPerFuse** is provided to fuse Sentinel-2 SITS with aerial imagery. This baseline consists of the U-TAE model, a modified U-Net with a Temporal self-Attention Encoder (TAE), to extract temporal information from a SITS. A Swin Transformer model is used to process aerial imagery. The features from both modalities are provided to a UPerNet decoder to produce pixel-wise class predictions at the GSD of aerial imagery.

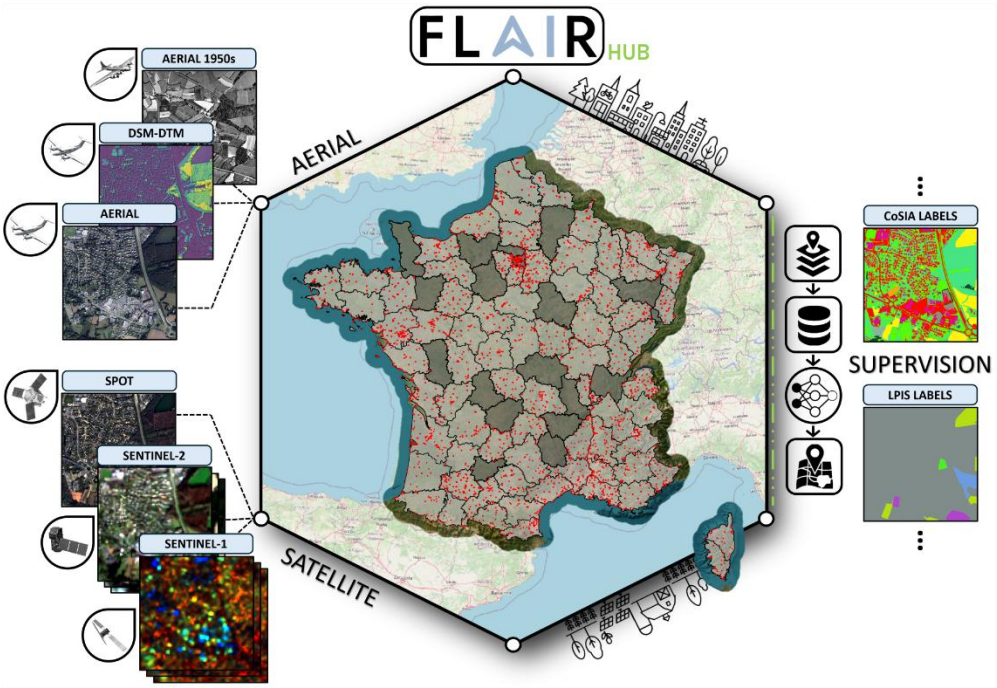

Figure 1. French Land cover from Aerospace ImageRy (FLAIR): Aerial and Satellite images.

Potential methodological modifications to the baseline could include, for instance, exploiting diffusion-based super-resolution approaches to generate intermediate higher-resolution SITS images that align more closely with aerial images before fusion. This can be achieved by adapting existing diffusion models **[2]** to generate SITS imagery constrained on higher-resolution SPOT-6 images. The code for the baseline and the dataset is available on github (https://github.com/IGNF/FLAIR-HUB). This thesis will be supervised by Hubert Kanyamahanga.

References

1. Garioud, Anatol, et al. "FLAIR-HUB: Large-scale Multimodal Dataset for Land Cover and Crop Mapping." *arXiv preprint arXiv:2506.07080* (2025).
2. Donike, Simon, et al. "Trustworthy super-resolution of multispectral sentinel-2 imagery with latent diffusion." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* (2025).

Besucheradresse:
Nienburger Straße 1
30167 Hannover
www.ipi.uni-hannover.de