Institut für Photogrammetrie
und GeoInformation

Leibniz Universität Hannover

Fakultät für Bauingenieurwesen
und Geodäsie

Institut für Photogrammetrie
und GeoInformation
Prof. Dr.-Ing. habil. Christian Heipke

Hubert Kanyamahanga, M.Sc.
Tel. +49 511 762 19388
Fax +49 511 762 2483
E-Mail: kanyamahanga@ipi.uni-hannover.de

10. Dezember 2024

Deep Learning-based Multiscale Fusion of Satellite Images Time Series and Aerial Images for Land Cover Classification

Proposal for a Master thesis topic (EN)

Recent developments in remote sensing have significantly encouraged the use of multi-sensor data for various applications such as land cover mapping. Multiple sensors can be used to acquire data with complementary information about the same observed region. For example, aerial imagery can deliver textural information at very high resolution, but usually with high revisit times, so that there is no information about the changing appearance of objects during the vegetation cycle. On the other hand, satellite systems, such as Sentinel-2, have short revisit times, so that the resultant images can capture temporal changes, but usually at a coarser spatial resolution, e.g. with a ground sampling distance (GSD) of 10 m or more. This causes problems in detecting smaller objects. Thus, integrating aerial and satellite images is of interest for land cover classification, the task of assigning a class label representing the physical material of the Earth surface to each pixel in the image.

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) to predict, at a pixel-level, the land cover of a given input image at the GSD of aerial images (see Fig. 1). A baseline model **[1]** based on a U-Net architecture is provided to fuse Sentinel-2 SITS with aerial images. 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. The aerial images are processed by a U-Net to produce pixel-wise class predictions; in order to fuse the two modalities, the SITS features are added to the encoder features of all levels in the skip connections.

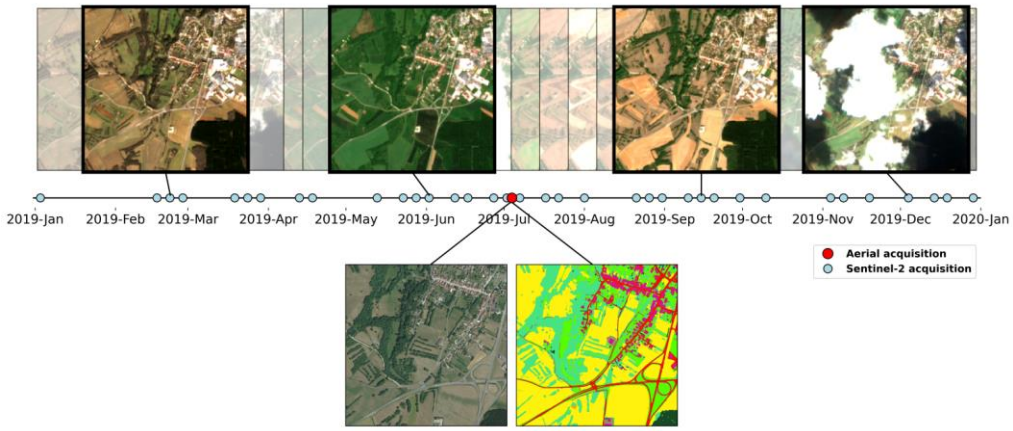

Figure 1: Example of mono-temporal aerial acquisition and annotations over an area (*bottom*) and Sentinel-2 time series acquisitions of the corresponding area (*top*) with four acquisitions samples.

Potential methodological modifications to the baseline could be, for instance to investigate Transformer-based models to extract spatial and temporal information from aerial and satellite images, respectively, and fusing them with an attention-based module to leverage the complementary information contained in both sensors. Another aspect to consider would be to investigate existing self-supervised learning approaches, with the advantage of using pre-trained models on large datasets. Beyond that, another issue to be addressed could be to improve the lower-resolution of the SITS data; perhaps, super-resolution approaches **[4]** such as diffusion models can be used to generate higher-resolution SITS images that align more closely with aerial images. The developed approach will be compared to existing approaches from the literature **[1, 2, 3]** which also combine multiscale data for land cover classification. The code for the baseline and the dataset is available on github (https://github.com/IGNF/FLAIR-2).

References

1. Garioud, A., Wit, A. D., Poupée, M., Valette, M., Giordano, S., Wattrelos, B., 2023. FLAIR #2: textural and temporal information for semantic segmentation from multi-source optical imagery. ArXiv, abs/2305.14467.

2. Kanyamahanga, H., Rottensteiner, F., 2024. Land Cover classification based on Multiscale aerial and satellite time series images. Proceedings, 44th Annual Scientific and Technical Conference of the DGPF in Remagen, 32, 223–235.

3. Heidarianbaei, M., Kanyamahanga, H., and Dorozynski, M., 2024. Temporal ViT-U-Net Tandem Model: Enhancing Multi-Sensor Land Cover Classification Through Transformer-Based Utilization of Satellite Image Time Series, ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., X-3-2024, 169–177, https://doi.org/10.5194/isprs-annals-X-3-2024-169-2024.

4. Okabayashi, A., Audebert, N., Donike, S., Pelletier, C., 2024. Cross-sensor super-resolution of irregularly sampled Sentinel-2 time series. IEEE International Conference on Computer Vision and Pattern Recognition (ICVPR), 502–511.

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