**Title**

Deep learning-based Image similarity estimation for geo-localization of Historical Aerial imagery

**Author**

[Liu, Yushan](https://repository.tudelft.nl/islandora/search/author%3A%22Liu%2C%20Yushan%22?collection=education) (TU Delft Civil Engineering & Geosciences)

**Date**

2024-02-26

**Abstract**

Historical aerial imagery serves as a valuable data source for observing Antarctica, facilitating an extended temporal scale of observation and enabling comparisons to deepen understanding of glacier dynamics. However, many historical aerial datasets, including the Antarctica Single Frames dataset utilized in this study, lack geo-referencing and orientation metadata essential for spatial analysis. One method of geo-referencing these historical images involves image matching to establish Ground Control Points (GCPs). This study focuses on the prerequisite for image matching: ensuring alignment between unreferenced historical images and already geo-referenced images in terms of scene and approximate resolution, a process termed 'geo-localization' herein.

Geo-localization is achieved by comparing the historical image with positions within a predefined geo-referenced Area of Interest (AoI). Two predefined remote sensing datasets are used: Sentinel-2 and Quantarctica Rock Outcrop Mask, from which AoIs are generated. Positions within the AoI exhibiting the highest similarity to the historical image are likely to correspond to the same ground area, thus providing the location of the historical imagery.

This similarity assessment employs two Siamese Networks: SigNet and ResNet-50. SigNet, originally designed for signature verification tasks, consists of four convolutional layers. In contrast, ResNet-50, initially developed for image classification purposes, is characterized by its deep architecture comprising approximately 50 convolutional layers, as suggested by its name. In this study, these two models are initially pre-trained on cross-domain datasets and subsequently adaptively trained with task-specific datasets created in this study. The adaptive training datasets comprise triplets of similar and dissimilar images pre-processed using methods devised in this study. An evaluation methodology based on confidence level is developed to assess the model and workflow performance, which is then applied to 51 test historical image samples.

Overall, the results indicate that the ResNet-50 based network outperforms SigNet, achieving a 95.5% average confidence level. However, the method does not meet the initial expectation of directly providing the location of the historical image within the AoI. Instead, it identifies potential locations. Nevertheless, this outcome is valuable as it streamlines the search process for subsequent image matching steps. For instance, a 95.5% average confidence level for the ResNet-50 based network correlates with an approximate 95.5% reduction in processing time for geo-referencing when integrated with image matching in subsequent steps.