Aligning Building Cadastral Footprints to Aerial Images by a Deep Learning Multi-Resolution Approach

Nicolas Girard (nicolas.girard@inria.fr)[1] Yuliya Tarabalka[1] Guillaume Charpiat[2]
[1] TITANE team, INRIA, Université Côte d'Azur. [2] TAU team, INRIA Saclay, LRI, Université Paris-Sud.

Introduction

- **Problem**: groundtruth data of building rooftops often do not align with the buildings in images
- Due to **different angles of capture** which make the rooftops move (even on orthorectified images because the Digital Terrain Model is not precise and does not include buildings)
- or **human error** when annotating the buildings
- or **lack of precision** of the groundtruth data

Method

Objectives:

- Primary:
  1. Output: **displacement map** that aligns the building polygons to the image (for example building from [1])
  2. Loss: mean squared error of the predicted displacement vectors for each pixel
- Secondary:
  1. Output: a **segmentation of the buildings** from the optical image
  2. Loss: cross entropy of the predicted class for each pixel
  3. The segmentation loss helps to train the network as the model has to learn where building are in order to predict the displacement map
- The Convolutionnal Neural Network is **iteratively applied at different resolutions** to reduce the range of displacement to deal with

How:

- Deep Learning method that builds on [2]
- Use of a modified U-Net [3] to have 2 image inputs and 2 image outputs

Conclusion

- Very effective at aligning buildings over aerial images and generalizes very well
- The segmentation helps training the model and also detects new buildings to update the map
- Another application of this method could be building height estimation

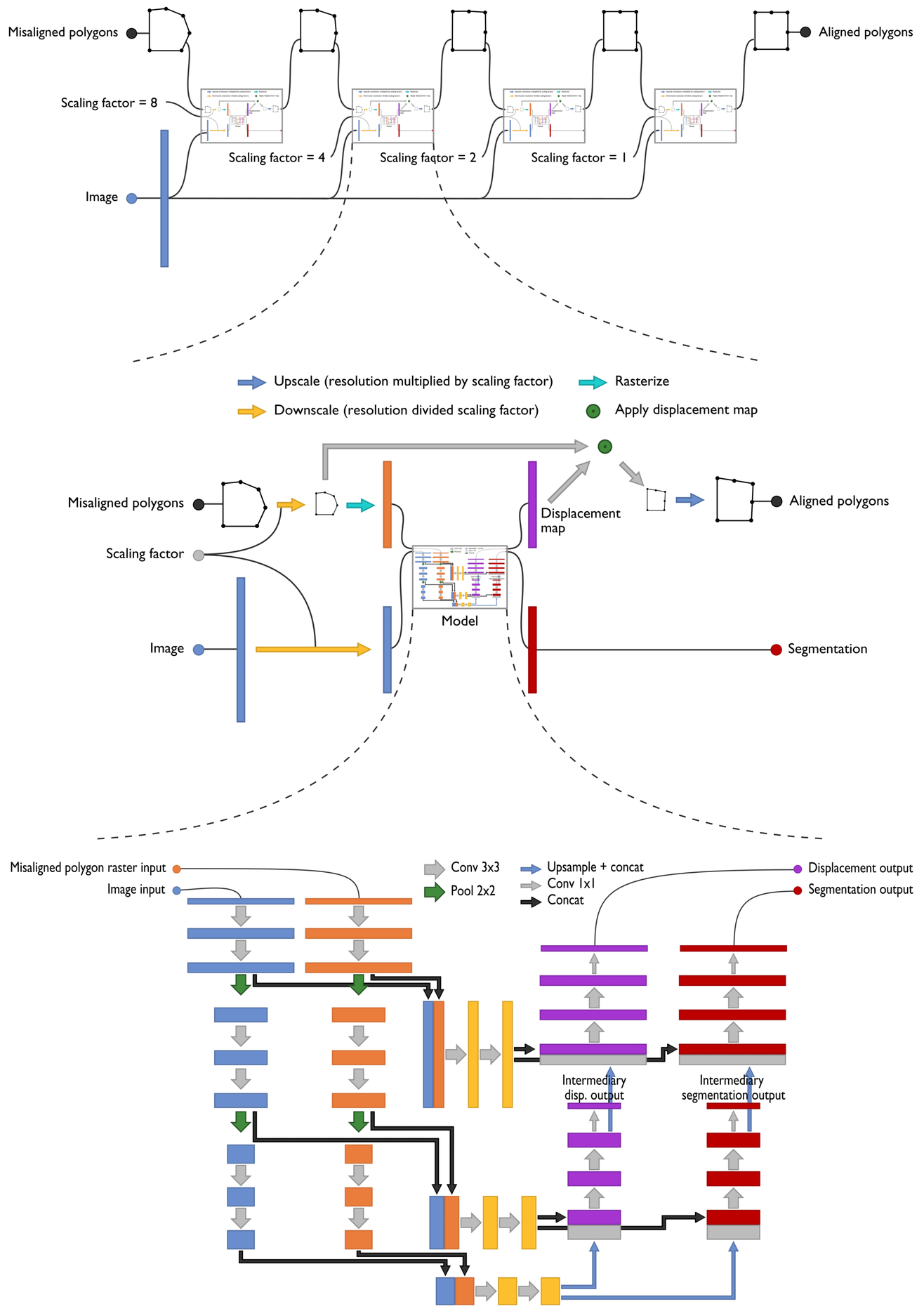
References

[1] OpenStreetMap contributors. Planet dump retrieved from https://planet.osm.org , 2017.

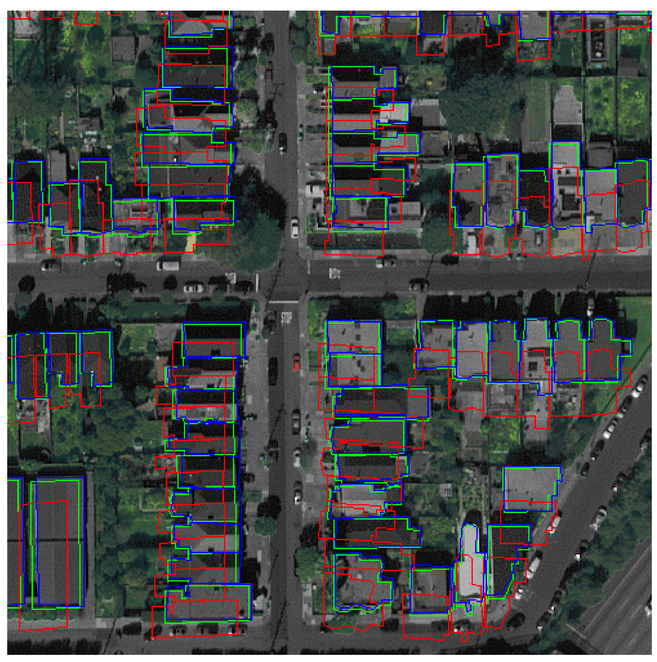
[2] A. Zampieri et al. Coarse to fine non-rigid registration: a chain of scale-specific neural networks for multimodal image alignment with application to remote sensing. *arXiv:1802.09816*, 2018.

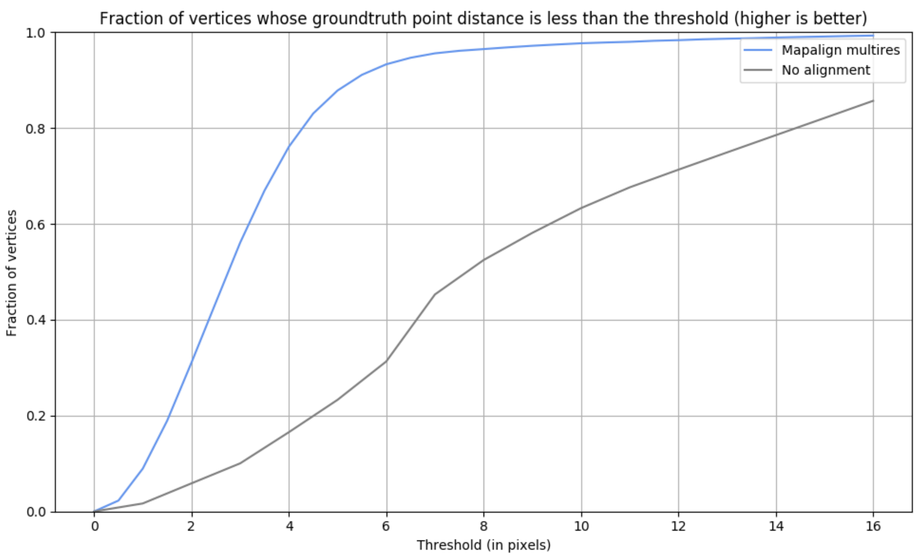
[3] O. Ronneberger et al. U-net: Convolutional networks for biomedical image segmentation. *arXiv:1505.04597*, 2015.

Complete pipeline



Results





- Green buildings: ground truth; red: misaligned [input]; blue: aligned [our output]
- **Accuracy measure**: Euclidean distance in pixels between ground truth vertices and aligned vertices
- Accumulated distribution of distances is plotted