

Aerial image geolocalization from recognition and matching of roads and intersections

Dragos Costea^{1,2}
onorabil@gmail.com

Marius Leordeanu^{1,2}
leordeanu@gmail.com

¹ University "Politehnica" of Bucharest
Bucharest, Romania

² Autonomous systems
Bucharest, Romania

We propose a complete pipeline for aerial image geolocalization based on roads and intersections. The main steps are road detection, followed by intersection detection and intersection matching with publicly available road vectors from the OpenStreetMap project (OSM)¹. An initial localization is proposed, further improved by geometric alignment of roads.

The main goal of this project is to provide drones a reliable localization system in the areas we expect most autonomous deployments will be made - that is, urban and suburban areas.

For intersection matching, we introduce a novel dataset consisting of images centered on intersections from two cities (one for training and the other one for testing), totalling 7204 600x600px images. A 4096-element descriptor is generated for each intersection using the surrounding detected roads and a neural network trained for intersection detection, in a way that is similar to [2]. We further fine-tune the network in a Siamese-like fashion in order to improve matching performance.

After an initial set of corresponding intersections is returned, we pick the best one by geometrical alignment of road maps for intersection candidates using shape context[1] and RANSAC. Further road enhancements for OSM are possible once the location has been determined.

We notice that most errors (around or above 90% of them) are below 2.5 meters, that is below 3 pixels for the image resolution available in our experiments. We believe that our results demonstrate high level of localization accuracy for our system, which could be very effective in most cases when the GPS signal is lost, for both day and nighttime.

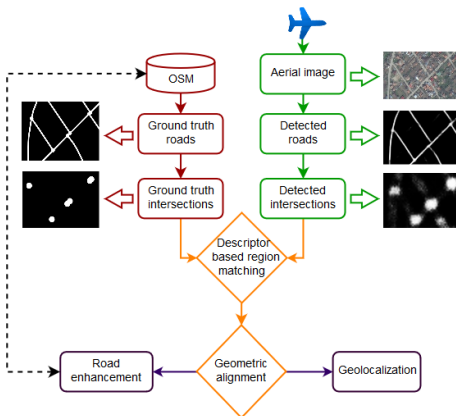


Figure 1: Framework overview

A key insight of our approach is the observation that intersections tend to have a unique road pattern surrounding them and thus can play a key role in localization, by reducing this difficult task to a sparse feature matching problem followed by a local refined road map alignment.

For pixel-wise road detection we used a dual-stream local-global CNN model proposed in [3]. Alexnet was used for intersection detection based on the detected road pattern.

- [1] Serge Belongie, Jitendra Malik, and Jan Puzicha. Shape context: A new descriptor for shape matching and object recognition. In *NIPS*, volume 2, page 3, 2000.
- [2] Kevin Lin, Hwei-Fang Yang, Jen-Hao Hsiao, and Chu-Song Chen. Deep learning of binary hash codes for fast image retrieval. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 27–35, 2015.
- [3] Alina Marcu and Marius Leordeanu. Dual local-global contextual pathways for recognition in aerial imagery. *arXiv preprint arXiv:1605.05462*, 2016.

¹<https://www.openstreetmap.org/>