# Supplementary Material of CVPR 2018 Review Paper: A Deep CNN-Based Framework For Enhanced Aerial Imagery Registration With Applications to UAV Geolocalization

Paper ID 3121

## 1. Supplementary Material: Geolocalization Dataset

As explained in the paper, our dataset is split into two sets: 1) Geolocalization Dataset. 2) Training Dataset. Geolocalization Dataset is composed of Aerial & Satellite Imagery to act as our $\boldsymbol{R}$ (Reference Map), and $\boldsymbol{S}$ (UAV Image Sequence). In Figure 2, samples of our $\boldsymbol{R}$ are shown which are used in the Geolocalization and Training Dataset. Figure 1 shows the approach of using similar Google Satellite imagery cloning ISPRS' tile extents. In continuation, Figure 3 shows another cloned ISPRS tile with an ISPRS label, and other areas nearby Potsdam with their OSM label. A sample of $\boldsymbol{S}$ from Potsdam and Famagusta is also shown in Figure 4.

Figura 1: A comparison between and ISPRS Potsdam (left) and a Google Satellite image (right).

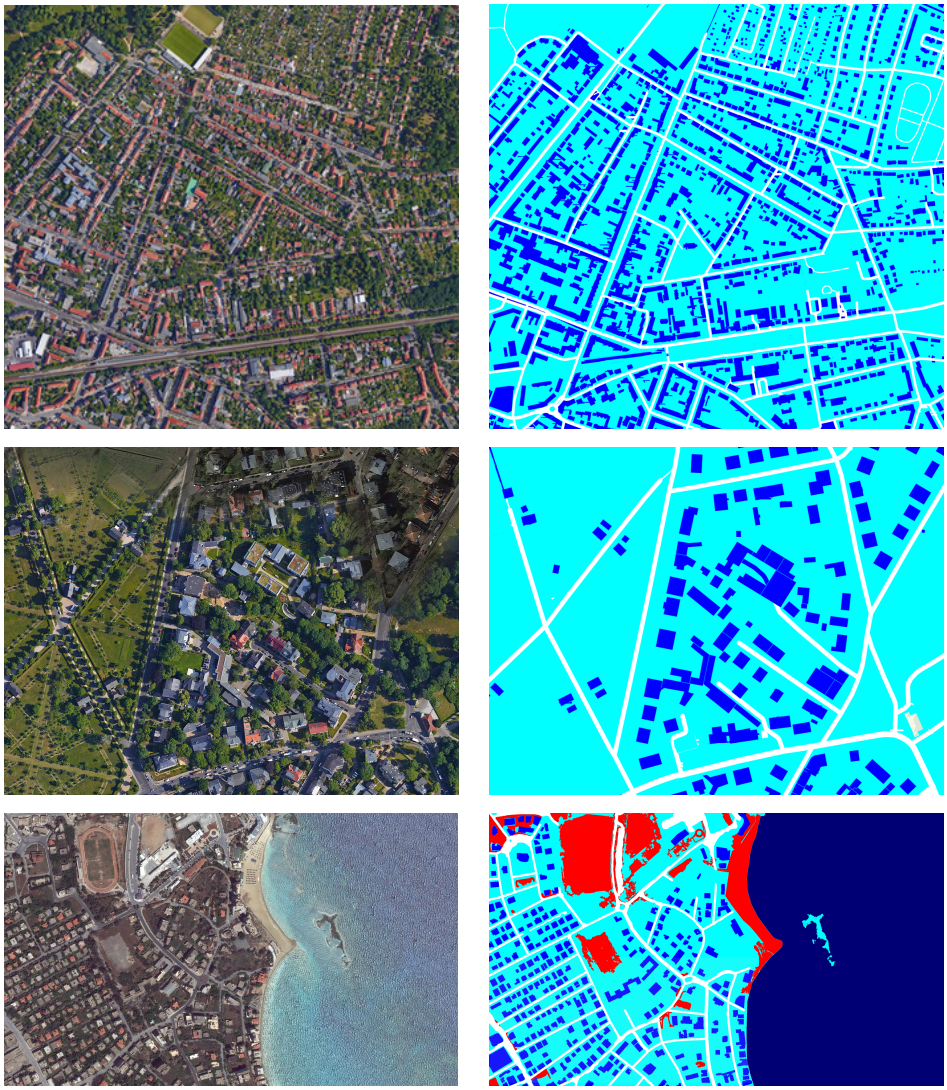

Figura 2: Top: $\boldsymbol{R}$ (Reference Map) of Potsdam and it's label used for registration. Middle: A sample of Potsdam's area used for training with it's label. Bottom: A sample of Famagusta's area used for training with it's label.

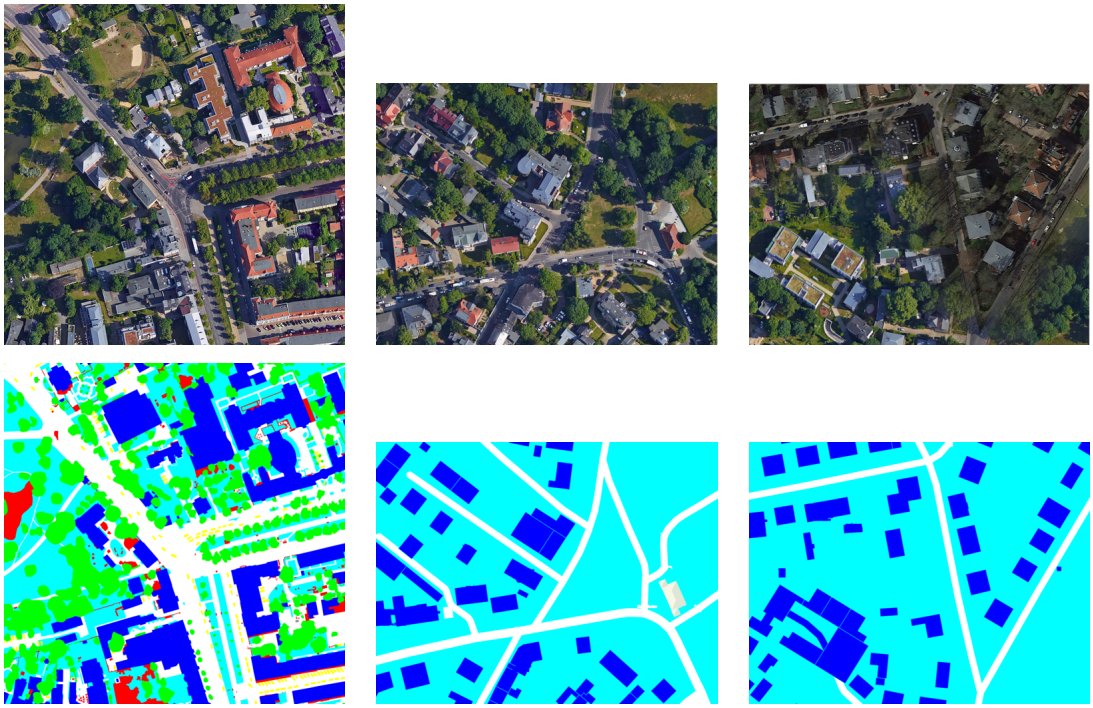

Figura 3: A sample of the method used to extend the dataset. On the left is a corresponding Google Satellite image copying ISPRS's tile, and using its ground truth. In the middle and on the right, tiles produced from other areas in Potsdam using OSM as the ground truth.

Figura 4: Sample of Potsdam (rows 1-4) & Famagusta's (rows 5-6) $\boldsymbol{S}$ (UAV Sequence Imagery).

## 2. Supplementary Material: Evaluation by distance Equation

Equations 1a and 1b calculate the distance between 2 GPS coordinates $(lat_1, lon_1)$ and $(lat_2, lon_2)$. This equation is used in our evaluation to find out the displacement between our predicted path, and the ground truth as shown in Figure 5.

$$2 \times 6371000 \times \arcsin\sqrt{x} = y \tag{1}$$

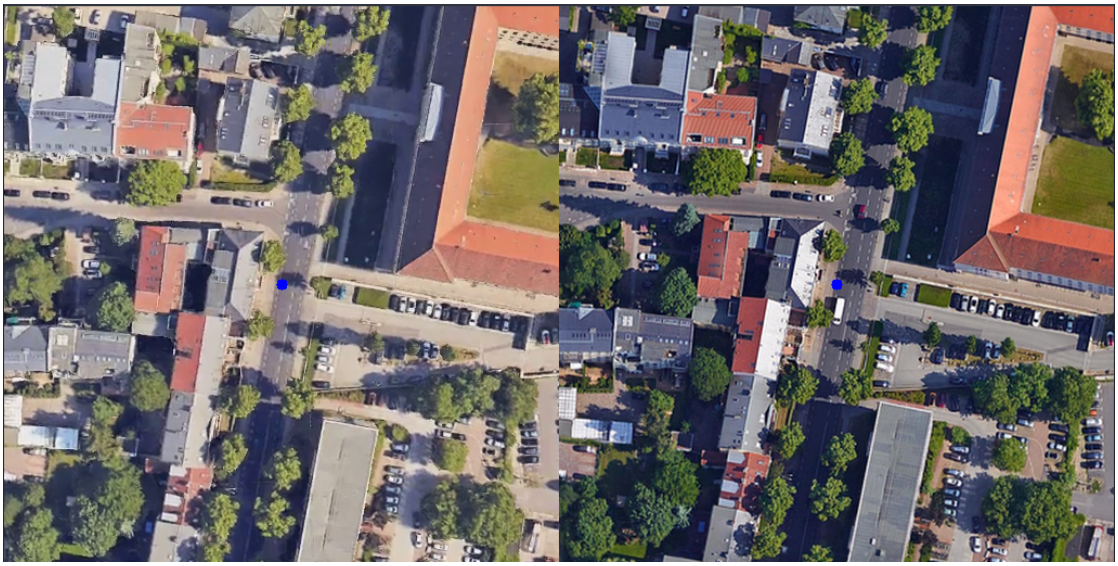

Figura 5: On the left is the $\boldsymbol{S_i}$, and on the right is the estimated $\boldsymbol{r}$. The actual UAV GPS coordinate is {52.404758, 13.053362}, while the estimated GPS coordinate is {52.404737, 13.053338}

.

## 3. Supplementary Material: Calibration

An example of the initial Calibration step is shown in Figure 6 as presented in Section 3.2 in the paper.

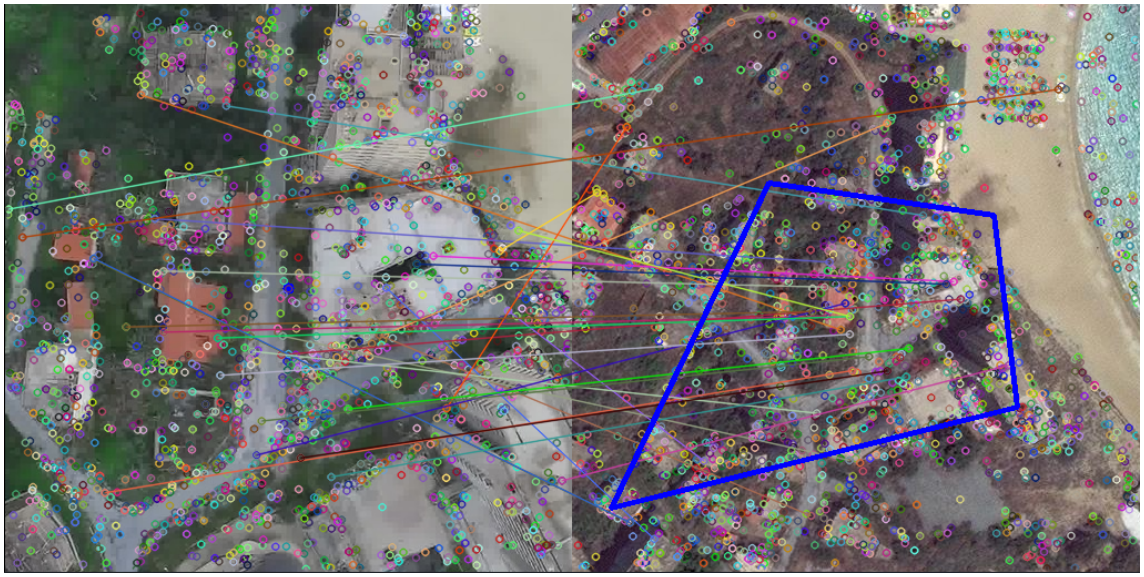

Figura 6: This figure demonstrates the registration using SIFT for calibration. The circles represent feature keypoints extracted, and the lines represent the matching of them. The blue polygon represents the boundaries of the $\boldsymbol{S}_{(i)}$ (left) in $\boldsymbol{r}$ (right) after the transformations. As shown, $\boldsymbol{r}$ is in a higher scale than $\boldsymbol{S}_{(i)}$ to be a search space for $\boldsymbol{S}_{(i)}$.

## 4. Supplementary Material: Semantic Segmentation Results

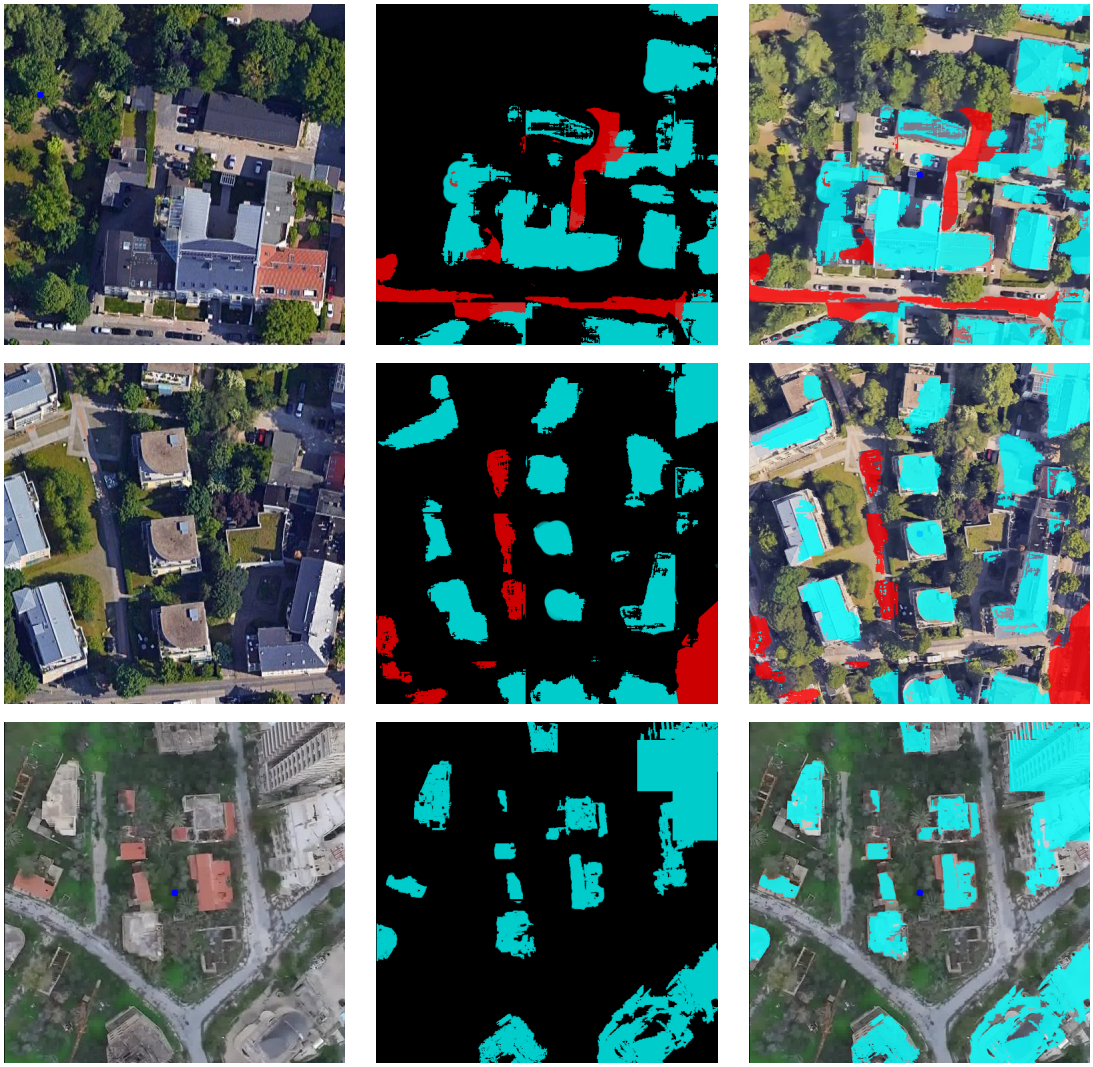

Figura 7: A sample of the predictions made using Experiment 2. Showing left column the target image, center the prediction, and right combined. These first 2 rows are $\boldsymbol{r}$ images in Potsdam and the last row is a $\boldsymbol{S}_{(i)}$ in Famagusta. Buildings are labeled cyan, and roads are labelled red.