BRNO FACULTY OF INFORMATION TECHNOLOGY
UNIVERSITY OF TECHNOLOGY

CrossLocate

cphoto.fit.vutbr.cz/crosslocate

Cross-modal Large-scale Visual Geo-localization in Natural Environments using Rendered Modalities

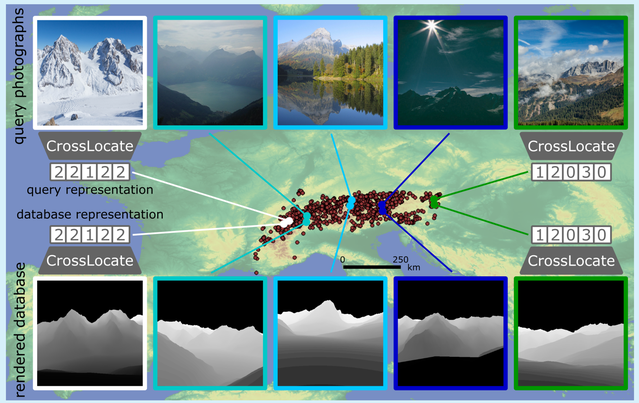
Jan Tomešek itomesek@fit.vutbr.cz
Martin Čadík cadik@fit.vutbr.cz
Jan Brejcha ibrejcha@fit.vutbr.cz

1. Large-scale Localization in Nature

We solve the challenging task of visual geo-localization of photographs captured in nature, i.e. we determine the position and approximate orientation of ground-level photographs. We demonstrate our approach on the area of the **Alps mountain range** spanning approx. 250 000 km^2.

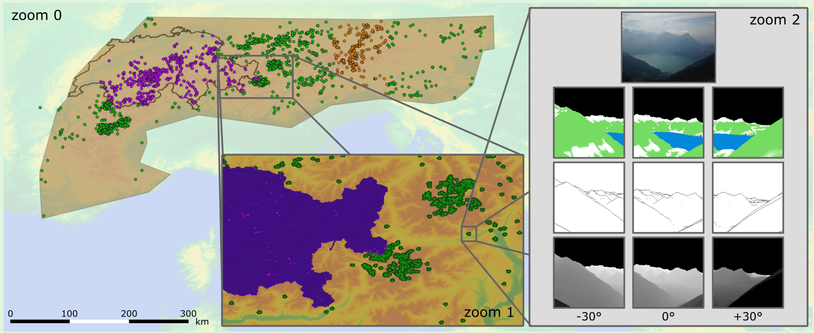
2. Cross-modal Image Retrieval

Photographs are encoded into **deep representations** and matched with deep representations extracted from rendered database imagery. This is achieved with only a single deep model with only a single branch.



3. Databases of Rendered Image Modalities

We introduce databases of various rendered image modalities and pair these databases with datasets of query photographs. This results into two new and **unique datasets**:
Sparse dataset - for fast and simple experimentation and task exploration
Uniform dataset - for real-world localization across vast areas and millions of images

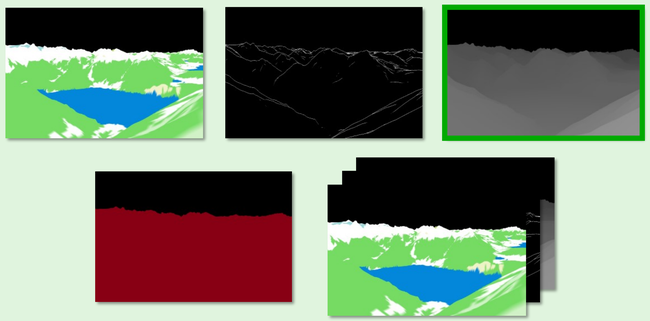


4. Efficient Architecture

We propose a simple, yet efficient **deep convolutional architecture** for our method. Among other crucial modifications, we use maximum pooling as an aggregation technique, which proved crucial in our task, compared to other aggregations, which failed.

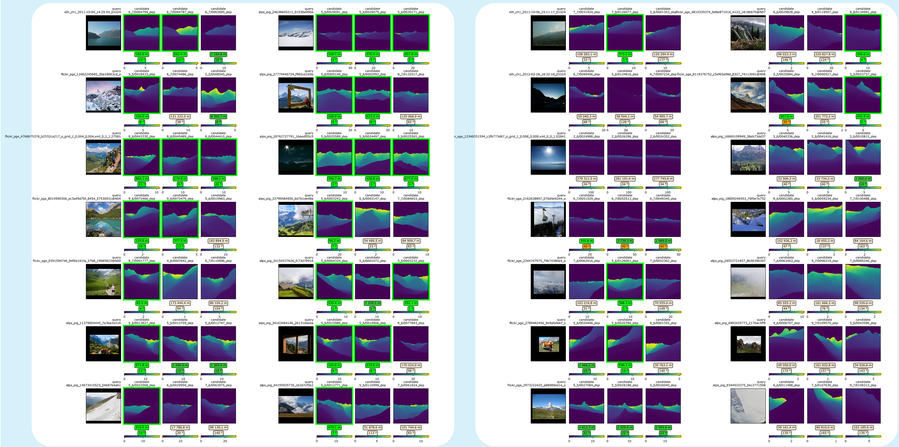
5. Depth Information for Localization

We evaluate the **localization potential** of different database modalities. Besides semantic segmentations, silhouette maps and depth maps, we measure the performance for previously used horizon lines and for the combination of the modalities.
We show that **depth information** is the most important one by a significant margin.



6. Qualitative Evaluation with Examples

We provide qualitative evaluation of our CrossLocate approach by presenting examples of both successful and unsuccessful localizations.
CrossLocate can deal with **severe occlusions** and **challenging lighting conditions**.



7. Experiments and Comparisons

Our work provides the evaluation of different **architectures**, underlines the importance of individual **components** and explains the benefits of known **field of view** information.

Our CrossLocate localization approach **significantly outperforms** state-of-the-art localization methods utilizing retrieval.