World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:16, No:09, 2022

AS-Geo: Arbitrary-Sized Image Geolocalization with Learnable Geometric Enhancement Resizer

Authors : Huayuan Lu, Chunfang Yang, Ma Zhu, Baojun Qi, Yaqiong Qiao, Jiangqian Xu

Abstract : Image geolocalization has great application prospects in fields such as autonomous driving and virtual/augmented reality. In practical application scenarios, the size of the image to be located is not fixed; it is impractical to train different networks for all possible sizes. When its size does not match the size of the input of the descriptor extraction model, existing image geolocalization methods usually directly scale or crop the image in some common ways. This will result in the loss of some information important to the geolocalization task, thus affecting the performance of the image geolocalization method. For example, excessive down-sampling can lead to blurred building contour, and inappropriate cropping can lead to the loss of key semantic elements, resulting in incorrect geolocation results. To address this problem, this paper designs a learnable image resizer and proposes an arbitrary-sized image geolocation method. (1) The designed learnable image resizer employs the self-attention mechanism to enhance the geometric features of the resized image. Firstly, it applies bilinear interpolation to the input image and its feature maps to obtain the initial resized image and the resized feature maps. Then, SKNet (selective kernel net) is used to approximate the best receptive field, thus keeping the geometric shapes as the original image. And SENet (squeeze and extraction net) is used to automatically select the feature maps with strong contour information, enhancing the geometric features. Finally, the enhanced geometric features are fused with the initial resized image, to obtain the final resized images. (2) The proposed image geolocalization method embeds the above image resizer as a fronting layer of the descriptor extraction network. It not only enables the network to be compatible with arbitrary-sized input images but also enhances the geometric features that are crucial to the image geolocalization task. Moreover, the triplet attention mechanism is added after the first convolutional layer of the backbone network to optimize the utilization of geometric elements extracted by the first convolutional layer. Finally, the local features extracted by the backbone network are aggregated to form image descriptors for image geolocalization. The proposed method was evaluated on several mainstream datasets, such as Pittsburgh30K, Tokyo24/7, and Places365. The results show that the proposed method has excellent size compatibility and compares favorably to recently mainstream geolocalization methods.

Keywords : image geolocalization, self-attention mechanism, image resizer, geometric feature
Conference Title : ICCV 2022 : International Conference on Computer Vision
Conference Location : Dubai, United Arab Emirates
Conference Dates : September 27-28, 2022