Object Detection in Panoramic Images

Guangwei Zhou, Yixuan Huang, Prasad Vagdargi, Luis Madera

- Project description:

The objective of this project is to perform object detection based on panoramic images, that is, given a panorama, detect and identify the objects that appear in it. Panoramic images have a wide range of applications in many fields with their ability to perceive all-round information, especially for autonomous driving with road panoramas. However, object recognition in such images can be challenging due to the distortion introduced in the panoramic images. We aim to approach the problem by using deep learning methods that have achieved remarkable results in the field of image classification and object detection.

The dataset we want to use is called Pano-RSOD. Pano-RSOD contains panoramas images with vehicles, pedestrians, traffic signs and guiding arrows. The objects of Pano-RSOD are labelled by bounding boxes in the images. Different from traditional object detection datasets, Pano-RSOD contains more objects in a panoramic image, and the high-resolution images have 360-degree environmental perception, more annotations, more small objects and diverse road scenes.

A research paper regarding this topic is "Pano-RSOD: A Dataset and Benchmark for Panoramic Road Scene Object Detection". https://www.mdpi.com/2079-9292/8/3/329/htm. Another dataset which can be used is the KITTI 3D object detection dataset (http://www.cvlibs.net/datasets/kitti/eval_object.php?obj_benchmark=3d)

One potential way to tackle this problem is to directly apply state-of-the-art object detection CNNs on the panoramic images, including Faster R-CNN (2-stage) or YOLO v4 (1-stage). However, the network's ability to overcome optical distortion is to be tested. The other method is to unstitch the panoramic image and perform object detection with CNNs on the component images, and map the detections back to the original panoramic image.

- Team member assignments (who does what)

1. Guangwei: Preprocess and extract features from panorama
2. Yixuan: Build Faster R-CNN for object detection
3. Prasad: Validation and testing
4. Luis: Unstitch panoramic image and build YOLO

- Backup Plan: Since the access to the dataset mentioned is dependent on the authors, we came up with a backup project. Using a large scale dataset of faces with people wearing masks (correctly and incorrectly, in different ways) we aim to identify which people are wearing a mask, and of that, what percentage of people are wearing it correctly. This can be later extended to tracking that person in surveillance video using person detection and motion tracking, for identifying the potential spread of disease caused due to that person. For this, we will be using the MaskedFace-Net Dataset (https://arxiv.org/abs/2008.08016), developed on Flickr-Faces dataset followed by testing person detection and tracking on surveillance video.