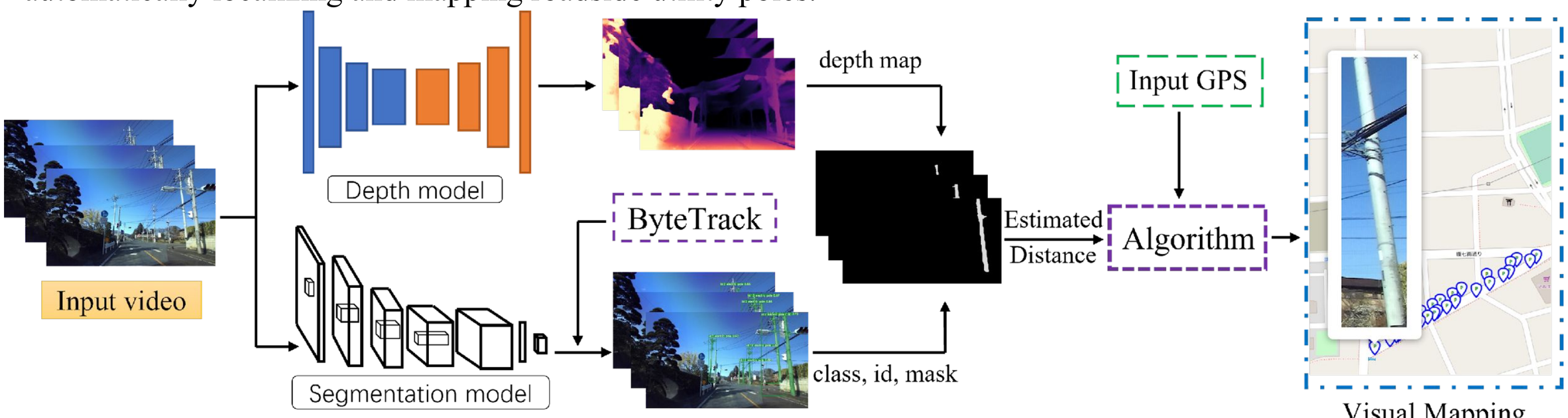
東京大学　関本研究室 / Sekimoto Lab. IIS, the University of Tokyo.

Are Utility Poles within Sight? Unleashing the Power of Vision for Effortless Roadside Management

Zhehui Yang[1], Deeksha Arya[1], Chenbo Zhao[1], Hiroya Maeda[2], Yoshihide Sekimoto[1]

Background

In 2021, the Japanese Road Bureau released documents indicating that the Japanese government is planning to promote the removal of utility poles to improve the disaster-prevention capability of roads in Japan. In this research, we present an approach to support digitalization of road-side objects specifically focusing on utility poles, which are vital infrastructure components for power distribution, telecommunications, and other services. The proposed approach aims to streamline and improve the management of utility poles, by offering a Computer Vision-based method for their automatic localization and mapping. To this aim, we have developed a novel deep learning-based vision-method for automatically localizing and mapping roadside utility poles.



System overview

Method

Sample training images

Our vision-based system consists of three modules: YOLOv8 performs utility pole detection and segmentation, generating the bounding box and semantic mask of detected objects. Distance estimation is performed by MonoViT, the current state-of-the-art self-supervised depth estimation network. Finally, for visual mapping, we developed an algorithm to visualize and reconstruct the GPS location map of utility poles.

Results

The test results show that our approach achieves a high detection recall score of 82.7% on our custom dataset and a mean error rate of 6.31% for locating utility poles. The proposed method opens new horizon applications in the field of Intelligent Transportation Systems (ITS) and autonomous driving research. Since the input data consists of consecutive images and the depth prediction network usually has stable and good perception performance on large objects or close range, we only need to ensure that the instance segmentation model has stable performance and sufficient accuracy within our designated range. Therefore, as shown in the right image , the predicted bounding boxes of the YOLOv8 model are more than our ground truth labels. This is because we only labeled the utility poles which close to the camera in the images. In other words, we only require the model have a sufficient recall rate on custom dataset.

Ground truth and Detection results

Dataset	Precision	Recall	F1-score
Mapillary	**79.5%**	35.0%	48.6%
ZED	26.4%	**82.7%**	40.0%

1.Sekimoto Lab. @ IIS Human Centered Urban Informatics, the University of Tokyo
2. UrbanX Technologies, Inc., Japan