MITSUBISHI-A project

Title: Exploration of Useful Geometric Structures for Object Recognition using Point Clouds

Industrial Partner: Mitsubishi Electric Corp. Advanced Technology R&D Center.

Mitsubishi Electric (Mitsubishi Electric Global Website) founded in 1921, is an electrical and electronic equipment manufacturer, developing products and solutions in widely diverse fields, including home appliances, industrial equipment, and space technologies. The Advanced Technology R&D Center was established to support the business of Mitsubishi Electric Group through the development of a broad scope of projects covering both basic and new advanced technologies. The main research themes include power electronics, mechatronics, satellite communications, next generation key devices, system solutions for electric power, transportation, factory automation, and automobiles.

Industry Mentor:

Akinobu SASADA, Mitsubishi Electric Corp.

Masashi YAMAZAKI, Ph.D., Mitsubishi Electric Corp.

Background:

In recent years, demand has increased for object detection using deep learning with limited point cloud data annotated manually. Deep learning is a method that uses input data to create a model that outputs predictions such as class and bounding box. The model of deep learning is created by iteration to modify its parameter to minimize the error between prediction obtained using the model and the actual value for the prediction. The actual value for the prediction is called ground truth. In general, many input data with ground truth are required to accomplish accurate prediction.

Remarkable progress has occurred in products equipped with object detection systems using point clouds such as self-driving cars and autonomous robots. The image available at this link [1] is useful for you to understand how a point cloud is used for a self-driving car. To enable these products to be operated more safely and to extend their service, more accurate object detection using point clouds is desired.

When a point cloud is used as the input of a detector, it is possible to detect objects, and to obtain information about three-dimensional (3D) positions. A point cloud is a dataset consisting of points represented by three coordinates, x, y, and z, which differs from image data because it includes depth information. Therefore, point cloud-based object detection enables us to know the locations of the detected objects.

Manual preparation of data used for deep learning is time consuming and costly. Deep learning has achieved success in various fields including object detection by achieving a high accuracy rate in prediction. However, an important shortcoming when using point cloud data for object detection is the need to add ground truth to each dataset. This process is called annotation. Additionally, it is also necessary to obtain features of data of various kinds for improvement of the accuracy rate. Therefore, using human effort for preparing data for deep learning entails high costs, which poses great difficulties for the manufacturing of products.

We explore a method to enable accurate object detection using deep learning even if we cannot gather a sufficient amount of manually annotated point cloud data. The importance of object detection using point clouds is becoming increasingly significant, especially in fields related to self-driving and autonomous robots. When using point cloud as input data of object detection, we can use the information of a detected object's 3D locations. However, there is a challenge to gather manually annotated data in object detection using deep learning with point clouds, just as there is for other tasks using deep learning. Therefore, a need exists for accurate object detection using deep learning with point clouds that require less annotation effort.

Project and Expectations:

For this project, we expect participants to develop an implementation that enables accurate object detection, even with limited point clouds, using artificially generated geometric structures. Then we expect participants to evaluate it. In [2], it has been demonstrated that features which are useful for object detection and object classification can be extracted from real objects using artificially generated fractal patterns. Panel (a) of Figure 1 in [2] provides an overview of paper [2], where fine-tuning is used. We expect the following three things from you.

- To devise a geometric structure specific to each real object of interest to extract their features successfully.
- To create an algorithm to generate figures with a geometric structure devised by you.
- To implement detection and identification of real objects using figures created by your own algorithm. You can use existing programs for object detection and object classification such as [3] and [4].

Required:

- Knowledge of calculus and linear algebra at the level of first and second-year undergraduate students.
- Elementary geometry

Optional:

- Knowledge about deep learning at the level of [5], [6] and [7] (You can start these tutorials if you have a Google account.)
- Knowledge of point cloud as in [8]
- Knowledge of geometry at a graduate school level.

References:

[1] Y. Muramatsu, "Autoware – the world's leading open-source software project for autonomous driving," [Online]. Available: https://github.com/autowarefoundation/autoware. [Accessed 25 1 2024].

[2] R. Yamada, H. Kataoka, N. Chiba, Y. Domae and T. Ogata, "Point Cloud Pre-training with Natural 3D Structures," IEEE, 2022.

[3] "OpenMMLab's next-generation platform for general 3D object detection," [Online]. Available: https://github.com/open-mmlab/mmdetection3d.

[4] C. R. Qi, O. Litany, K. He and L. Guibas, "Deep Hough Voting for 3D Object Detection in Point Clouds," [Online]. Available: https://github.com/facebookresearch/votenet.

[5] "TorchVision Object Detection Finetuning Tutorial," [Online]. Available: https://colab.research.google.com/github/pytorch/tutorials/blob/gh-pages/_downloads/4a542c9f39bedbfe7de5061767181d36/torchvision_tutorial.ipynb.

[6] PyTorch, "Torchvision Object Detection Finetuning Tutorial," [Online]. Available: https://pytorch.org/tutorials/intermediate/torchvision_tutorial.html.

[7] S. Chilamkurthy, "Transfer Learning for Computer Vision Tutorial," [Online]. Available: https://colab.research.google.com/github/pytorch/tutorials/blob/gh-pages/_downloads/74249e7f9f1f398f57ccd094a4f3021b/transfer_learning_tutorial.ipynb.

[8] C. R. Qi, H. Su, K. Mo and L. J. Guibas, "Pointnet: Deep learning on point sets for 3d classification and segmentation," IEEE, 2017.