3D Point Cloud Data Classification

Background

3D point cloud data is a set of data points represented by its three coordinates (x, y, z) in space (Qi et al. 2017b), some also include colour information (R, G, B). One of the most well-known representatives is Light Detection and Ranging (Lidar) data (Reutebuch et al., 2005), which is collected by illuminating pulsed laser light at a target area and calculating the distance via the difference of return time that laser back to sensor.

The most different part of point cloud data from images is that it is disorderly, so common deep learning structures cannot be applied.

Motivation

Though deep learning is very successful in the field of image and video, there are some limitations on these kinds of data. For example, RGB-video cannot detect the range and the distance of the objects. Furthermore, videos will be influenced by some condition such as bright sunlight, pitch darkness, rain, which Lidar would not; hence, Lidar can perform equally well day and night.

Thanks to the advantages of Lidar, it can be used on autonomous vehicles, agriculture, river survey etc. It can have many usages and be able to improve the life of human if point cloud data can be handled properly.

Data Introduction

The data is point cloud of ModelNet40 made up of 40 objects categories, provided by Princeton University. It contains 9840 training data and 2468 for testing, each made up by 2048 points with 3D coordinates (x, y, z), that is, each object with dimension (2048, 3).

Apart from using original data, this paper adds some random noise data, structured data that extract from other objects to imitate the situation that may encounter in real life, and also tests the robustness to missing value by evaluating model with data only about one eighth points remain.

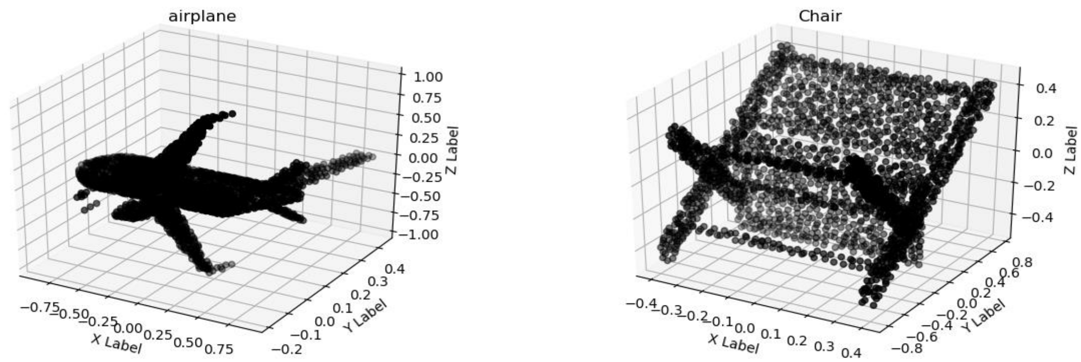

Figure 1 Visualization of ModelNet40 Data

Methods

This report uses the current best-performing 3D point cloud data deep learning structure, based on PointCNN, and introduces the concepts of Inception Model and The All Convolutional Net to reduce the required training parameters and improve the accuracy of the model in classification.

Then use the original data, noise-added data, and data with missing values to train and test, and compare the amount of training parameters required and accuracy with other structures (PointNet model) to measure each model.

Conclusion

From the comparison of the following data, it can be clearly seen that PointCNN after introducing the concepts of Inception Model and The All Convolutional Net performs better than other structures on all types of data (Number of parameters decrease, accuracy increases).

Through comparing and improving the performance of the models, this paper expect that the analysis of 3D point cloud data can be more mature and accurate, and be applied to various tasks.

	Original Data	15% Random Noise Data	15% Structured Noise Data	Missing Point (256 points)
PointNet	88.61	57.05	39.22	70.26
PointCNN	92.03	89.59	87.03	87.12

Comparison between PointNet and PointCNN

	Structure 1	Structure 2	Structure 3	Structure 4	Structure 5
Number of Parameters	2,261,340	549,084	599,340	505,764	560,244
Accuracy	91.69	91.25	91.53	91.05	92.03
Avg. Class Accuracy	87.42	87.03	87.47	86.95	87.84

Comparison between Different PointCNN Structures

Structure 1: Convolution

Structure 2: With 1x1 Convolution

Structure 3: Depthwise Separable Convolution

Structure 4: Inception Model

Structure 5: Inception Model (Replace Max-pool with Convolution)