

Convulaitonal Neural Network for Point Cloud Classification

EEC 289Q - Final Project

Ahmed Mahmoud

Muhammad Awad

ABSTRACT TO DO

ACM Reference Format:

Ahmed Mahmoud and Muhammad Awad. 2018. Convulaitonal Neural Network for Point Cloud Classification: EEC 289Q - Final Project. In *Proceedings of ACM Conference (Conference'17)*. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/nnnnnnnn.nnnnnnnn>

1 INTRODUCTION

1.1 Problem Statement:

In this project we are trying to classify 3D point cloud. The data set is sets of *models*. Each model is comprised of a set of n data points. A data points is a tuple of 3D coordinates (x, y, z) of the point in the space. Figure on the left shows few examples of the data set we have. Each model is associated with a label (for training) shown under the model.

The data set consists (ModelNet40 [5]) of 9840 models for training and 2468 models for testing. Our task is to create a CNN model that can train such data set and achieve good accuracy within the test models.

1.2 Challenges:

At first glance the classification looks similar to classifying images which has been investigated thoroughly in the past couple of years. However, we a careful look at the problem and the type of input we have, we can realize the following set of challenges that will influence our model

Irregularity: While standard deep neural network models take as input data with regular structure (images of a fixed sizes), point clouds are irregular. This means we have no expectation of the input size (the number of points per models). The reason behind this is that these models either come from scan operations or hand-crafted which can not be constrained to give models of fixed size.

Unordered: Point positions are continuously distributed in the space and any permutation of their ordering does not change the spatial distribution. The model is read from a file which could store the point in a certain order that is different than other similar models. Thus, associating the feature with the point order could be misleading.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

Conference'17, July 2017, Washington, DC, USA

© 2018 Association for Computing Machinery.

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM...\$15.00

<https://doi.org/10.1145/nnnnnnnn.nnnnnnnn>

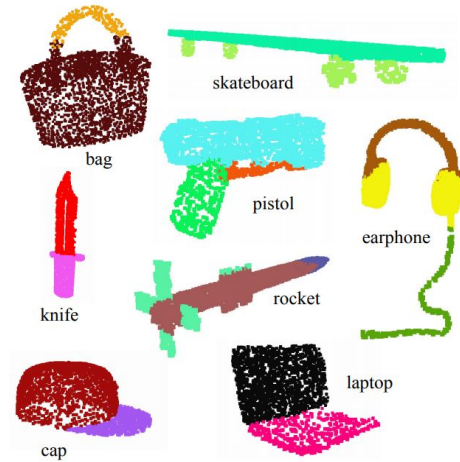


Figure 1: Example Models (color here are added for visualization only).

Invariance: The learned features should be invariant to translation and rotation. Our model should be able to accurately similar models even if they are rotated or shifted. Similar case exists in image where a certain feature could be located anywhere in the image. However, with point cloud in 3D space the problem is augmented as there is a third degree of freedom for translation and rotation.

1.3 Previous Work:

Previous work on point cloud classification using CNN has three main directions:

View-based Methods: These methods represents a single model as a collection of 2D views to which standard CNNs used in image classification can be employed directly. View-based appraoches are good match for applications where the input comes from 3D sensor and represented as a range image in which case a single view can be used [4].

Volumetric Methods: Voxelization is another straightforward way to convert irregular data set to a regular 3D grid over which standard CNN can be applied [3]. Voxelization produces a sparsely-occupied grids and induce quantization artifacts. Cleverer space partition like k -d tree can be used to mitigate these problem [2].

Direction Methods: These methods try to learn from the raw data points without voxelization or alternating the views. It should be expected that these methods to have greater accuracy since we do not manipulate the feature being learned. However, proper model design is challenging due to the challenges been mention earlier. The pioneer work on tackling point cloud learning from the raw

data set is PointNet [1]. PointNet applies a symmetric function to the 3D coordinates in a manner invariant to permutation where every point is treated individually. However, PointNet employs a complex and computationally expensive spatial transformer network to learn the 3D alignment of the points. This is due to the fact that the features being learned are global features and no local features are utilized.

2 XYZ CNN:

2.1 Model

2.2 Results:

REFERENCES

- [1] Kaichun Mo, Leonidas J. Guibas, Charles R. Qi, Hao Su. 2017. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. In *Conference on Computer Vision and Pattern Recognition (CVPR)*.
- [2] Roman Klokov and Victor Lempitsky. 2017. Escape from cells: Deep kd-networks for the recognition of 3d point cloud models. In *2017 IEEE International Conference on Computer Vision (ICCV)*. IEEE, 863–872.
- [3] Daniel Maturana and Sebastian Scherer. 2015. Voxnet: A 3d convolutional neural network for real-time object recognition. In *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*. IEEE, 922–928.
- [4] Lingyu Wei, Qixing Huang, Duygu Ceylan, Etienne Vouga, and Hao Li. 2016. Dense human body correspondences using convolutional networks. In *Computer Vision and Pattern Recognition (CVPR), 2016 IEEE Conference on*. IEEE, 1544–1553.
- [5] Zhirong Wu, Shuran Song, Aditya Khosla, Fisher Yu, Linguang Zhang, Xiaoou Tang, and Jianxiong Xiao. 2015. 3d shapenets: A deep representation for volumetric shapes. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 1912–1920.