Project Proposal

”Improving Point Cloud Shape Analysis of a Point Cloud Transformer using Curve Aggregation”

1 Abstract

Local feature aggregation is an operation that assembles point features of a given key point set, computes the position encodings of the subject point and the neighboring points, and passes the results into relevant transformation and aggregation modules in furtherance of local feature extraction. Even though these operations are feasible for depicting relative local patterns, they are inept with regard to long-range point relations. To that extent, the aggregation strategy introduced by Xiang et al. [5] proposes a new long-range feature aggregation method, namely curve aggregation, for point clouds shape analysis. Initiative of our project is to implement the curve aggregation method upon the Point Cloud Transformer (PCT) of Guo et al. [2], replacing the local neighbor embedding strategy.

2 Requirements

2.1 Overview

The permutation-invariant PCT achieves very high performance on classification and segmentation tasks, in addition to its improved semantic feature learning capability. Nonetheless, competency of PCT has not been tested on very large datasets, attesting the feature learning capabilities through large amounts of training data. To that extent, long-range feature aggregation method utilizes the intra-channel feature variety of relative encodings, which then projects the difference into a higher dimension. Our plan is to use this more sophisticated feature aggregation method of CurveNet with the transformer based point cloud feature encoding structure of PCT to improve the given results.

2.2 Methodology

For the project, we will first implement the curve aggregation method for feature points into Point Cloud Transformer, replacing the local neighbor embedding strategy. Followingly, we will train the resulting transformer on different datasets, examining conclusive advantages and disadvantages in tandem.

2.3 Dataset

Apart from the baseline datasets that are used in PCT (which are ModelNet40 [6] for point cloud classification, ShapeNet [4] for object segmentation, and Stanford 3D Indoor Dataset [1]), we will be exploring various other point cloud datasets (e.g. Partnet [3]) to evaluate and verify our findings.

3 Team

Cakir, Cavit - 03751930
Hiebl, Christian - 03737941
Nalcaci, Atilla - 03736102
Seiler, Fabian - 03756784

References

[1] Iro Armeni, Ozan Sener, Amir R Zamir, Helen Jiang, Ioannis Brilakis, Martin Fischer, and Silvio Savarese. 3d semantic parsing of large-scale indoor spaces. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1534–1543, 2016.

[2] Meng-Hao Guo, Jun-Xiong Cai, Zheng-Ning Liu, Tai-Jiang Mu, Ralph R Martin, and Shi-Min Hu. Pct: Point cloud transformer. *Computational Visual Media*, 7(2):187–199, 2021.

[3] Kaichun Mo, Shilin Zhu, Angel Chang, Li Yi, Subarna Tripathi, Leonidas Guibas, and Hao Su. Partnet: A large-scale benchmark for fine-grained and hierarchical part-level 3d object understanding. pages 909–918, 06 2019.

[4] Zhirong Wu, Shuran Song, Aditya Khosla, Fisher Yu, Linguang Zhang, Xiaoou Tang, and Jianxiong Xiao. 3d shapenets: A deep representation for volumetric shapes. pages 1912–1920, 06 2015.

[5] Tiange Xiang, Chaoyi Zhang, Yang Song, Jianhui Yu, and Weidong Cai. Walk in the cloud: Learning curves for point clouds shape analysis. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 915–924, 2021.

[6] Li Yi, Vladimir G. Kim, Duygu Ceylan, I-Chao Shen, Mengyan Yan, Hao Su, Cewu Lu, Qixing Huang, Alla Sheffer, and Leonidas Guibas. A scalable active framework for region annotation in 3d shape collections. *SIGGRAPH Asia*, 2016.