International Journal of Computer Vision (2023) 131:772–795
https://doi.org/10.1007/s11263-022-01731-4

A Closer Look at Few-Shot 3D Point Cloud Classification

Chuangguan Ye[1] · Hongyuan Zhu[2] · Bo Zhang[3] · Tao Chen[1]

Received: 31 May 2022 / Accepted: 30 November 2022 / Published online: 15 December 2022

Abstract

In recent years, research on few-shot learning (FSL) has been fast-growing in the 2D image domain due to the less requirement for labeled training data and greater generalization for novel classes. However, its application in 3D point cloud data is relatively under-explored. Not only need to distinguish unseen classes as in the 2D domain, 3D FSL is more challenging in terms of irregular structures, subtle inter-class differences, and high intra-class variances when trained on a low number of data. Moreover, different architectures and learning algorithms make it difficult to study the effectiveness of existing 2D FSL algorithms when migrating to the 3D domain. In this work, for the first time, we perform systematic and extensive investigations of directly applying recent 2D FSL works to 3D point cloud related backbone networks and thus suggest a strong learning baseline for few-shot 3D point cloud classification. Furthermore, we propose a new network, Point-cloud Correlation Interaction (PCIA), with three novel plug-and-play components called Salient-Part Fusion (SPF) module, Self-Channel Interaction Plus (SCI+) module, and Cross-Instance Fusion Plus (CIF+) module to obtain more representative embeddings and improve the feature distinction. These modules can be inserted into most FSL algorithms with minor changes and significantly improve the performance. Experimental results on three benchmark datasets, ModelNet40-FS, ShapeNet70-FS, and ScanObjectNN-FS, demonstrate that our method achieves state-of-the-art performance for the 3D FSL task.

Keywords Machine learning · Few-shot learning · Meta learning · Point cloud classification

1 Introduction

Deep learning models have achieved promising performance on various computer vision (CV) and natural language processing (NLP) tasks, with the advent of powerful computing resources and large-scale annotated datasets. The 3D point cloud understanding studies with deep learning techniques are also fast-growing (Qi et al., 2017a, b; Li et al., 2018; Liu et al., 2019b; Zhang et al., 2022; Cosmo et al., 2022). Unlike traditional point cloud recognition algorithms (Johnson and Hebert, 1999; Zhong, 2009; Rusu et al., 2009) that extract shape features based on the hand-crafted operators, deep-learning-based methods can get more informative descriptors for point cloud instances from shape projections (Yu et al., 2018; Qi et al., 2016; Feng et al., 2018) or raw points (Qi et al., 2017a, b; Liu et al., 2019b; Li et al., 2018; Wang et al., 2019b) with deep networks, which achieve better performance.

However, deep-learning-based methods have two crucial issues for point cloud classification. Firstly, the deep-learning technique usually requires large-scale labeled data to train, but it is cumbersome and costly to annotate extensive point cloud data. Secondly, the models trained on base classes often fail to generalize to novel or unseen classes. To overcome the data annotation problem, data-augmentation techniques (Tanner and Wong, 1987), semi-supervised learning (Chapelle et al., 2009), or other learning paradigms (Liao

Communicated by Ming-Hsuan Yang.

This work was completed under the partial supervision of Dr. Hongyuan Zhu at A*STAR. Chuangguan Ye and Hongyuan Zhu are with equal contribution.

✉ Tao Chen
eetchen@fudan.edu.cn

Chuangguan Ye
cgye19@fudan.edu.cn

Hongyuan Zhu
hongyuanzhu.cn@gmail.com

Bo Zhang
zhangbo@pjlab.org.cn

1 School of Information Science and Technology, Fudan University, Shanghai, China

2 Institute for Infocomm Research, A*STAR, Singapore, Singapore

3 Shanghai AI Laboratory, Shanghai, China