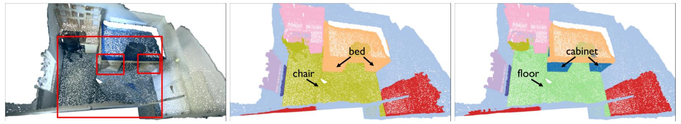
Learning with Noisy Labels for Robust Point Cloud Segmentation

Shuquan Ye[1] Dongdong Chen[2] Songfang Han[3] Jing Liao[1]

[1]City University of Hong Kong, [2]Microsoft Cloud AI, [3]University of California

2021 ICCV OCTOBER 11-17 VIRTUAL

Background and Introduction



Input Scenes | Real-World Noisy GT Labels | Predictions of PNAL

- Clean 3D data labels are difficult to obtain because of massive point number and complex annotation process. Even the commonly used **3D scene dataset** ScanNetV2 suffers from **noisy label** problem.
- Most noisy-robust works focus on image classification. And they are **not applicable or suboptimal** to apply on point segmentation.
- In this work, we take the lead in solving this issue by proposing a novel Point Noise-Adaptive Learning (PNAL) framework.

Problem Description

- Formally, we denote point cloud data as $\mathbf{X} \in \mathcal{R}^{N\times C}$ of N points with C features of coordinates and RGB values possibly, and its **noisy** semantic label as Y, and M as the class number. Our target is to train a model $\mathbf{f}_\theta(\mathbf{X})$ robust to the label noise in the training set. Based on our observation and its official annotation pipeline of popular real-world dataset, **label noise** is mainly at **instance level**.

Motivation

- To correct the **instance-level** label noise, we create a **cluster-based** noise correction method, since instance label may not be available.
- While noise rate is **unknown**, **variant** and **possibly heavy** for real-world noisy dataset, we design a novel **noise-rate blind** framework.
- To take **neighbor** point correlations into consideration and generate the best possible label, we propose a **voting** strategy.

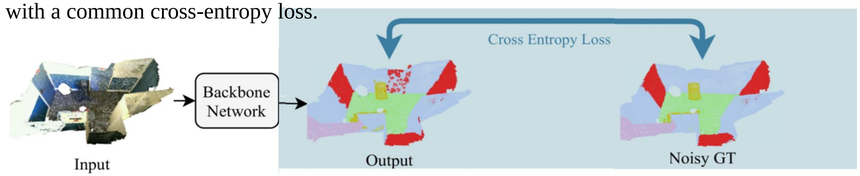
Highlight

- We are the **first** to investigate noisy label problem on point cloud data which has a wide and urgent need for 3D applications where the volume of data is growing drastically.
- A novel and effective noise-rate blind **framework** PNAL is proposed with point-level confidence selection and cluster-level label correction with voting mechanism.
- We **refine** the validation set of **ScanNetV2** with more accurate labels to facilitate point segmentation and noisy label learning.

PNAL Framework

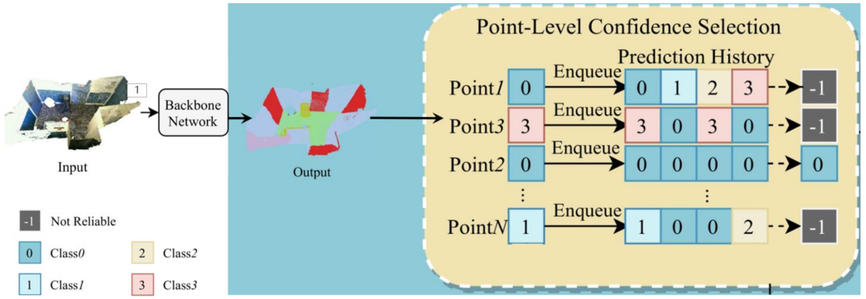
Warm-up Stage

- According to the study of memorization effects, in the presence of noisy labels, DNNs are prone to learn clean, easy samples first. Therefore, in the warm-up stage, we train the network with a common cross-entropy loss.

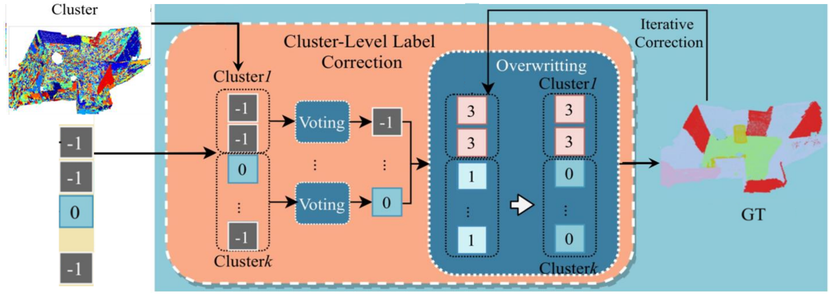


Noise Cleaning Stage

- **Point-Level Confidence Selection** aims to select **reliable samples** from each mini-batch and obtain the **confidential labels** for these samples that can be corrected with high probability. In detail, a sample with consistent label prediction history is regarded as the reliable sample, and its most frequently predicted label is the reliable label.

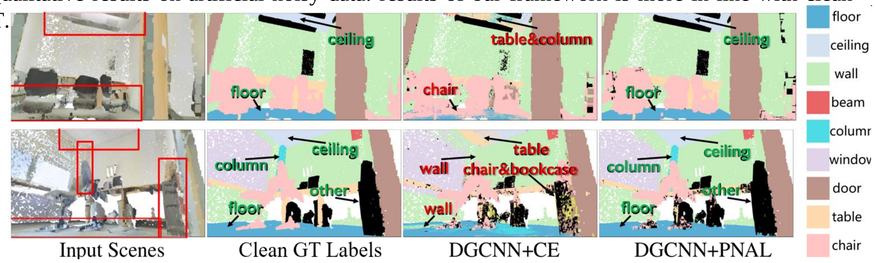


- **Cluster-Level Label Correction** aims to correct the noise at the cluster level, considering local relationship between point labels. Cluster-wisely, we count the occurrences of reliable labels for each class. Then we get a winner label by **voting** from the top reliable labels, and overwrite the labels within this cluster with it and iteratively correct the **training set**.



Results

- Qualitative results on artificial noisy data: results of our framework is more in line with clean GT.

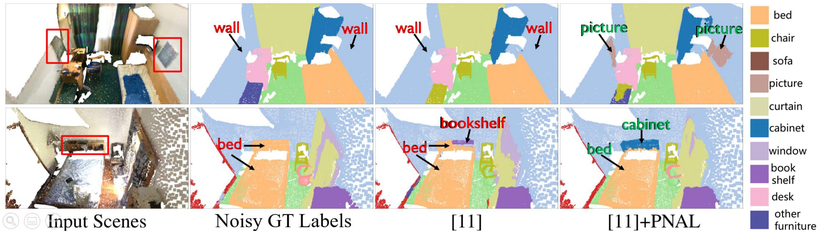


Input Scenes | Clean GT Labels | DGCNN+CE | DGCNN+PNAL

- Quantitative results on artificial noisy data: our framework produces much better results than all baselines. Even on 60% noise, our results is comparable to trained completely clean dataset.

Methods	0%	Symmetric Noise (τ)				Asymmetric Noise
		20%	40%	60%	80%	$\tau_{pair}=40\%,\tau=60\%$
DGCNN[31]+CE	**0.8692**	0.7506	0.6732	0.6390	0.5060	0.5634
DGCNN[31]+SCE[30]	0.7768	0.7524	0.7230	0.6509	0.5705	0.7084
DGCNN[31]+GCE[35]	0.7067	0.7003	0.6997	0.6967	0.6880	0.6614
DGCNN[31]+SELFIE[26]*	0.8673	0.8158	0.7914	0.7725	0.7163	0.7500
DGCNN[31]+PNAL	0.8686	**0.8569**	**0.8378**	**0.8236**	**0.7651**	**0.7968**
PointNet2[22]+CE	**0.8898**	0.7008	0.6796	0.5850	0.5204	0.5648
PointNet2[22]+PNAL	0.8852	**0.8385**	**0.8271**	**0.8067**	**0.7708**	**0.8202**

- Qualitative results on real-world noisy data: results of our framework is more reasonable than GT labels given by ScanNetV2 validation set.



Input Scenes | Noisy GT Labels | [11] | [11]+PNAL

- Quantitative results on real-world noisy data: our framework achieves significant performance gain.

Methods	our re-labeled ScanNetV2	
	mIoU	OA
SparseConvNet[11]	0.7103	0.8807
SparseConvNet[11]+PNAL	0.7416	0.9211

- Visualization of the label correction process during training. The correction process spreads from large areas to small objects and to whole training set as the training proceeds.

Input Scenes | Noisy Labels | 5 | 15 | 29 | GT Labels

Code: https://github.com/pleaseconnectwifi/PNAL

☑ Computationally efficient implementation

Dataset: https://shuquanye.com/PNAL_website/

☑ A re-labeled clean ScanNetV2 validation