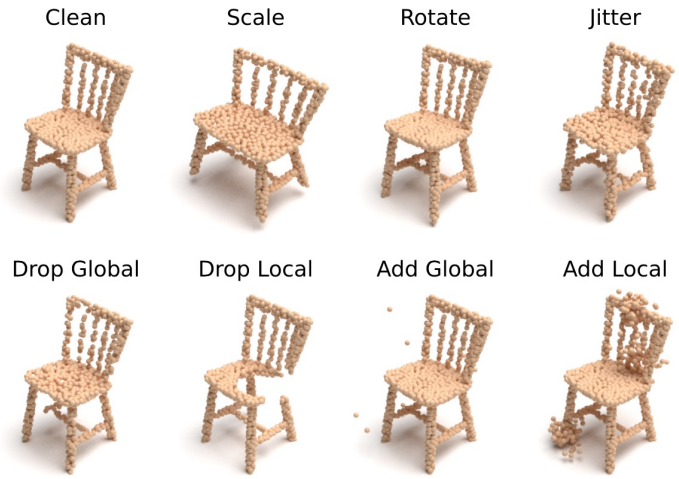
Benchmarking and Analyzing Point Cloud Classification under Corruptions

Jiawei Ren, Liang Pan, Ziwei Liu

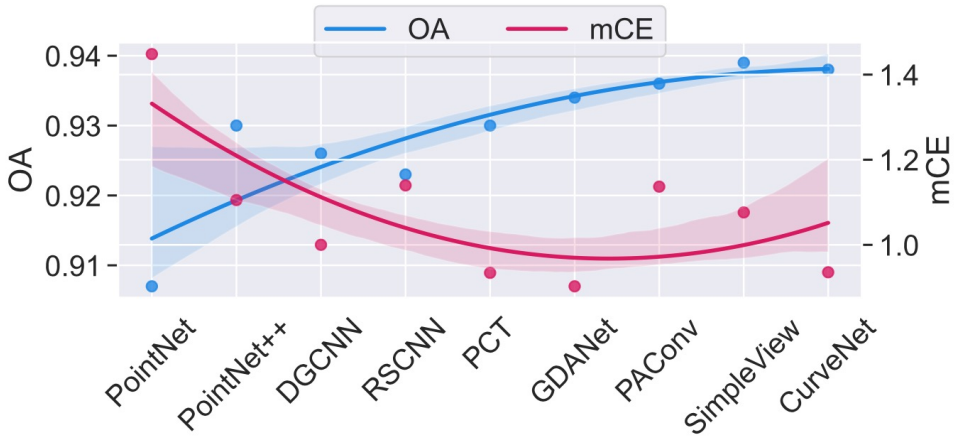
NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE | S-LAB FOR ADVANCED INTELLIGENCE

TL;DR

- We present a test suite *ModelNet-C* for benchmarking Point Cloud classification under corruptions.
- Our benchmark result suggests that point cloud classifiers are at the risk of getting less robust. We discuss effective designs to improve the robustness.



ModelNet-C



Benchmark Result

Overview

- Introduction
  - **Why is robustness crucial in point cloud classification?**
  - How has the robustness been evaluated?
- ModelNet-C
  - What should a comprehensive corruption set include?
  - How to design the corrupted test set?
  - How to measure the robustness?
- Benchmark Results
  - Are point cloud classifier getting more robust?
  - What makes a robust point cloud classifier?
- Conclusion: A call for robustness in point cloud classification

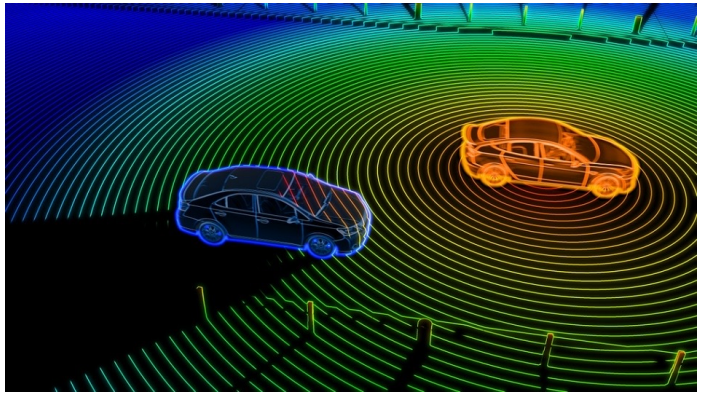
Why is robustness crucial in point cloud classification?

- Point clouds are used in **safety-critical** applications but often suffer from severe **OOD corruptions**.

Clean

Real-world

Corruptions are severe and OOD
e.g., occlusions, sensory noise



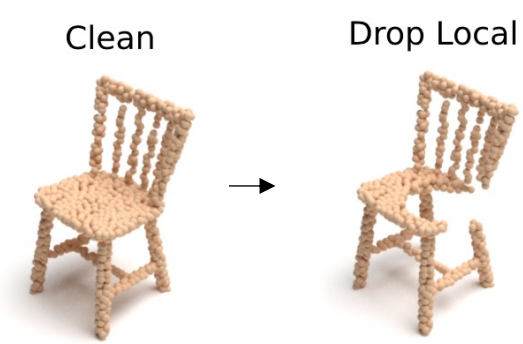
Applications are safety-critical
e.g., autonomous driving

Overview

- Introduction
  - Why is robustness crucial in point cloud classification?
  - **How has the robustness been evaluated?**
- ModelNet-C
  - What should a comprehensive corruption set include?
  - How to design the corrupted test set?
  - How to measure the robustness?
- Benchmark Results
  - Are point cloud classifier getting more robust?
  - What makes a robust point cloud classifier?
- Conclusion: A call for robustness in point cloud classification

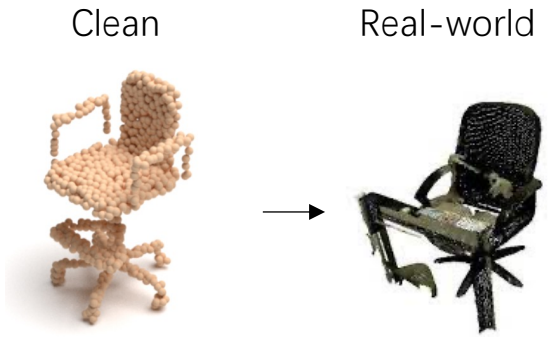
How has the robustness been evaluated?

- Prior works use different protocols:



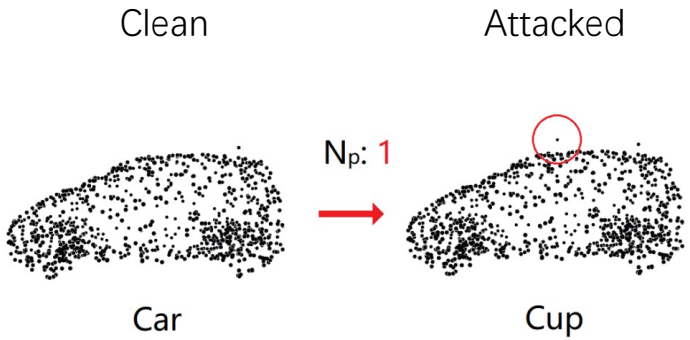
e.g., local point dropping

Protocol-1
Robustness to selected corruptions



e.g., ModelNet40 -> ScanObjectNN

Protocol-2
Robustness to the sim-to-real gap

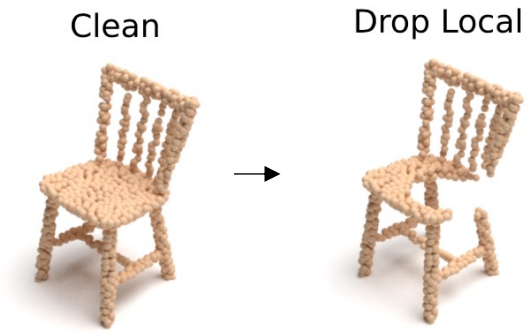


e.g., one-point attack

Protocol-3
Robustness to adversarial attacks

How has the robustness been evaluated?

- Prior works use different protocols:



e.g., local point dropping

Protocol-1
Robustness to selected corruptions

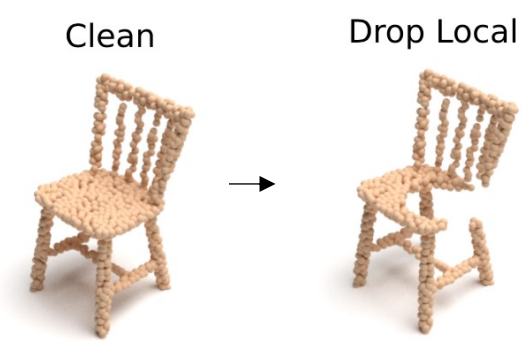
[X] Not comprehensive
[X] Difficult to compare across methods

	Jitter	Drop Global	Drop Local	Add Global	Add Local	Scale	Rotate
PointNet (Qi et al., 2017b)	✓	✓		✓			
ECC (Simonovsky & Komodakis, 2017)	✓	✓					
PointNet++ (Qi et al., 2017a)		✓					
DGCNN (Wang et al., 2019)		✓					
RSCNN (Liu et al., 2019)		✓					✓
PointASNL (Yan et al., 2020)		✓		✓			
Orderly Disorder (Ghahremani et al., 2020)	✓						
PointAugment (Li et al., 2020)	✓	✓				✓	✓
PointMixup (Chen et al., 2020)	✓	✓				✓	✓
PAConv (Xu et al., 2021a)	✓					✓	✓
OcCo (Wang et al., 2021)		✓					
Triangle-Net (Xiao & Wachs, 2021)	✓	✓				✓	✓
Curve-Net (Xiang et al., 2021)	✓	✓					
RSMix (Lee et al., 2021)	✓	✓				✓	✓
PointWolf (Kim et al., 2021)	✓	✓	✓		✓		
GDANet (Xu et al., 2021b)		✓					✓

Prior works evaluate point cloud classification robustness on different sets of corruptions, hence their evaluations can be partial and incomparable.

How has the robustness been evaluated?

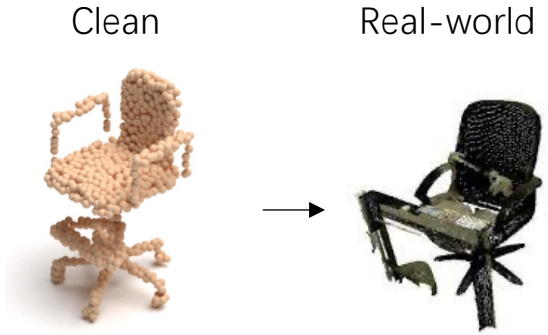
- Prior works use different protocols:



e.g., local point dropping

Protocol-1
Robustness to selected corruptions

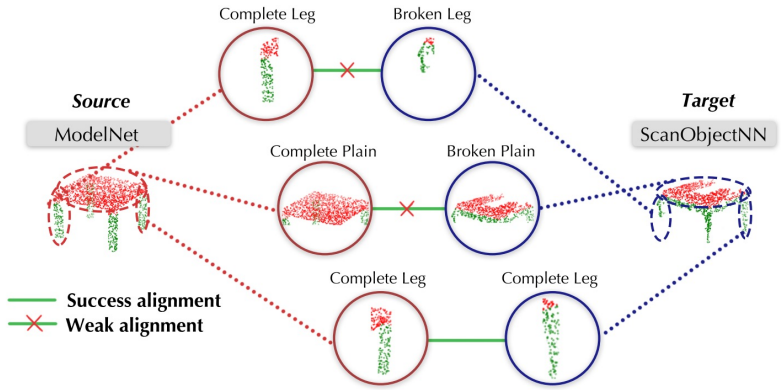
[X] Not comprehensive
[X] Difficult to compare across methods



e.g., ModelNet40 -> ScanObjectNN

Protocol-2
Robustness to the sim-to-real gap

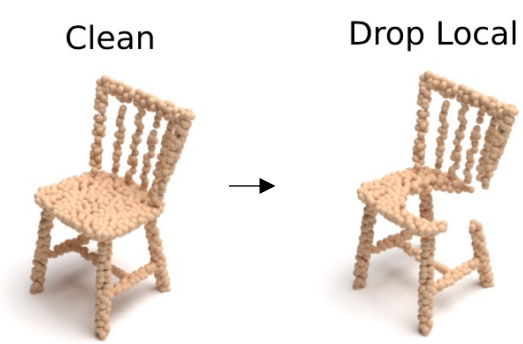
[X] Not allowing fine-grained analysis
[X] Coupled with domain gap



Corruptions are composite and hard to analyze

How has the robustness been evaluated?

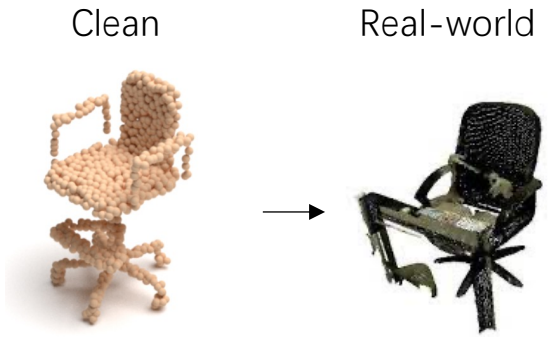
- Prior works use different protocols:



e.g., local point dropping

Protocol-1
Robustness to selected corruptions

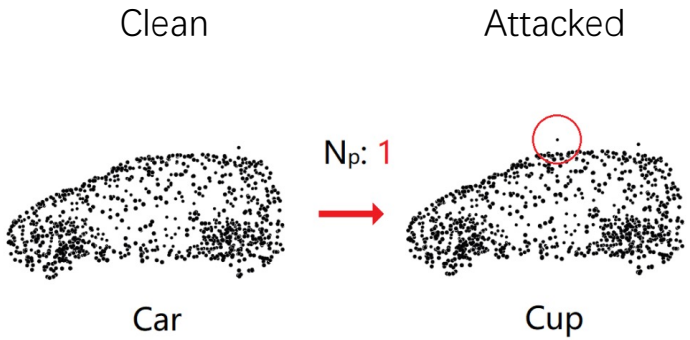
[X] Not comprehensive
[X] Difficult to compare across methods



e.g., ModelNet40 -> ScanObjectNN

Protocol-2
Robustness to the sim-to-real gap

[X] Not allowing fine-grained analysis
[X] Coupled with domain gap



e.g., one-point attack

Protocol-3
Robustness to adversarial attacks

[X] Can not reflect robustness to corruptions in the natural world.

Introduction

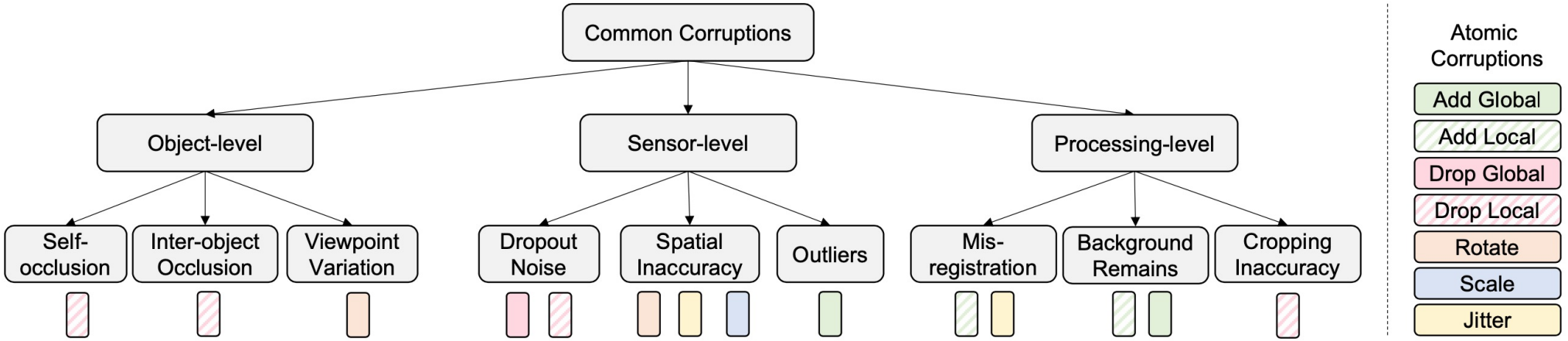
- Summary:
- Point clouds are used in **safety-critical** applications but often suffer from severe **OOD corruptions**.
- Despite various ways to evaluate a point cloud classifier's robustness, there lacks a **standard**, **comprehensive** benchmark.
- => We present *ModelNet-C*, a full corruption test suite, to close this gap.

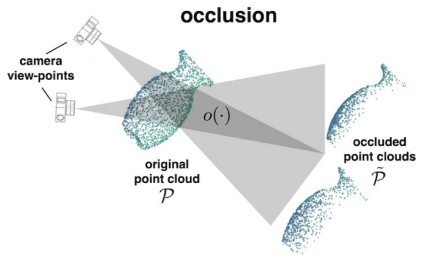
Overview

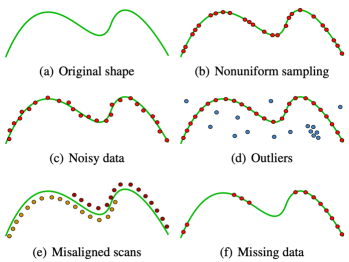
- Introduction
  - Why is robustness crucial in point cloud classification?
  - How has the robustness been evaluated?
- ModelNet-C
  - **What should a comprehensive corruption set include?**
  - How to design the corrupted test set?
  - How to measure the robustness?
- Benchmark Results
  - Are point cloud classifier getting more robust?
  - What makes a robust point cloud classifier?
- Conclusion: A call for robustness in point cloud classification

What should a comprehensive corruption set include?

- **Corruptions Taxonomy**: We break down common corruptions into detailed corruption sources, and further simplify them into a combination of atomic corruptions.







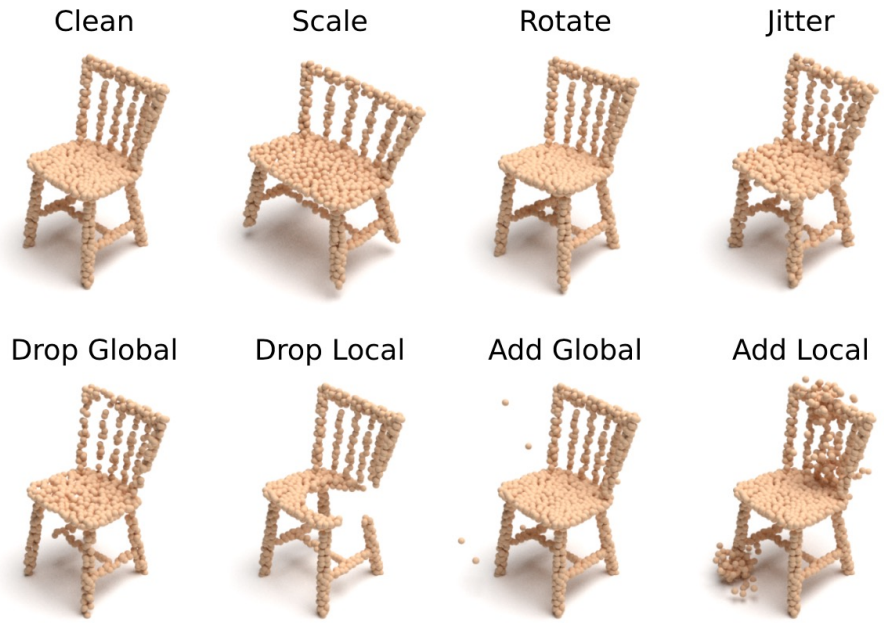
(a) OBJ_BG (b) PB_T25 (c) PB_T25_R

Overview

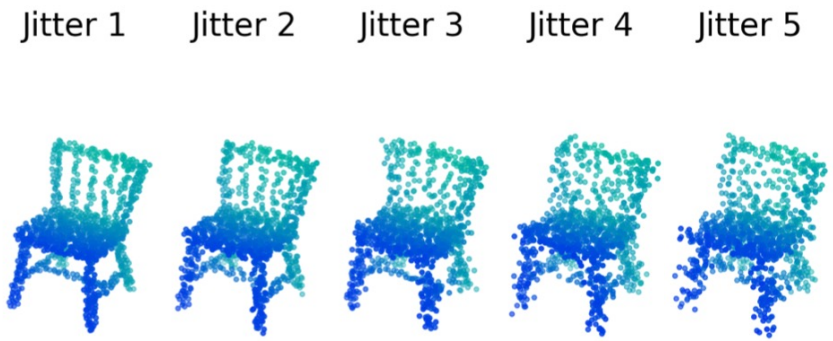
- Introduction
  - Why is robustness crucial in point cloud classification?
  - How has the robustness been evaluated?
- ModelNet-C
  - What should a comprehensive corruption set include?
  - **How to design the corrupted test set?**
  - How to measure the robustness?
- Benchmark Results
  - Are point cloud classifier getting more robust?
  - What makes a robust point cloud classifier?
- Conclusion: A call for robustness in point cloud classification

How to design the corrupted test set?

ModelNet-C: ModelNet40 is one of the most used benchmarks. We corrupt the ModelNet40 testset using the atomic corruptions with varying severities.



Atomic Corruptions



Different Severities

Overview

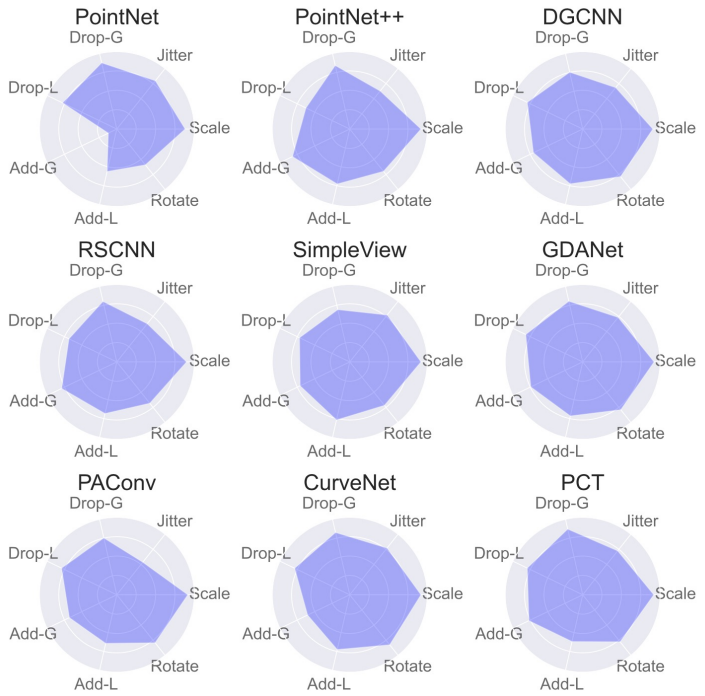
- Introduction
  - Why is robustness crucial in point cloud classification?
  - How has the robustness been evaluated?
- ModelNet-C
  - What should a comprehensive corruption set include?
  - How to design the corrupted test set?
  - **How to measure the robustness?**
- Benchmark Results
  - Are point cloud classifier getting more robust?
  - What makes a robust point cloud classifier?
- Conclusion: A call for robustness in point cloud classification

How to measure the robustness?

Evaluation Metrics: Inspired by the ImageNet-C, we use mean CE (mCE), as the primary metric. Compared to the commonly used Overall Accuracy (OA), mCE shows average performance under all types of corruptions.

$$\mathrm{CE}_i = \frac{\sum_{l=1}^{5}(1-\mathrm{OA}_{i,l})}{\sum_{l=1}^{5}(1-\mathrm{OA}_{i,l}^{\mathrm{DGCNN}})},$$

$$\mathrm{mCE} = \frac{1}{N}\sum_{i=1}^{N}\mathrm{CE}_i$$

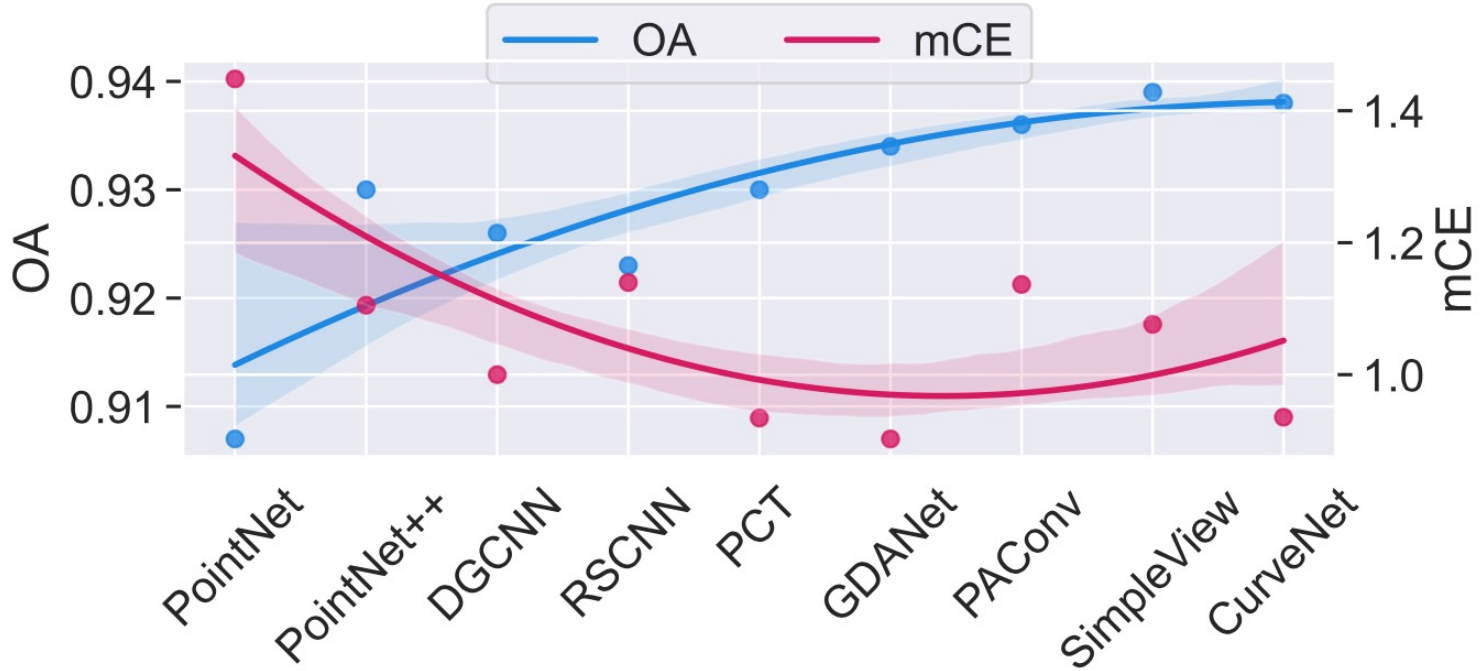


Overview

- Introduction
  - Why is robustness crucial in point cloud learning?
  - How has the robustness been evaluated?
- ModelNet-C
  - What should a comprehensive corruption set include?
  - How to design the corrupted test set?
  - How to measure the robustness?
- Benchmark Results
  - **Are point cloud classifiers getting more robust?**
  - What makes a robust point cloud classifier?
- Conclusion: A call for robustness in point cloud classification

Are point cloud classifier getting more robust?

- **No.** Although the accuracy on ModelNet40 gradually saturates, the robustness is at the risk of getting worse, due to the lack of a standard test suite.

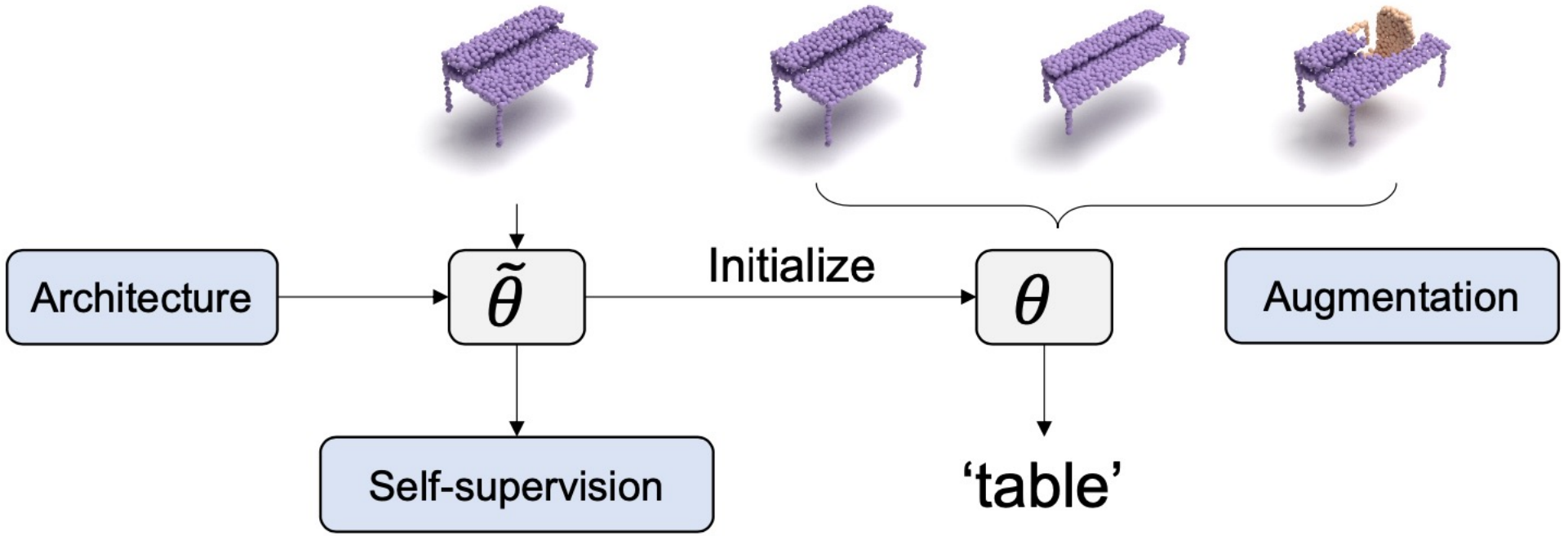


Overview

- Introduction
  - Why is robustness crucial in point cloud classification?
  - How has the robustness been evaluated?
- ModelNet-C
  - What should a comprehensive corruption set include?
  - How to design the corrupted test set?
  - How to measure the robustness?
- Benchmark Results
  - Are point cloud classifier getting more robust?
  - **What makes a robust point cloud classifier?**
- Conclusion: A call for robustness in point cloud classification

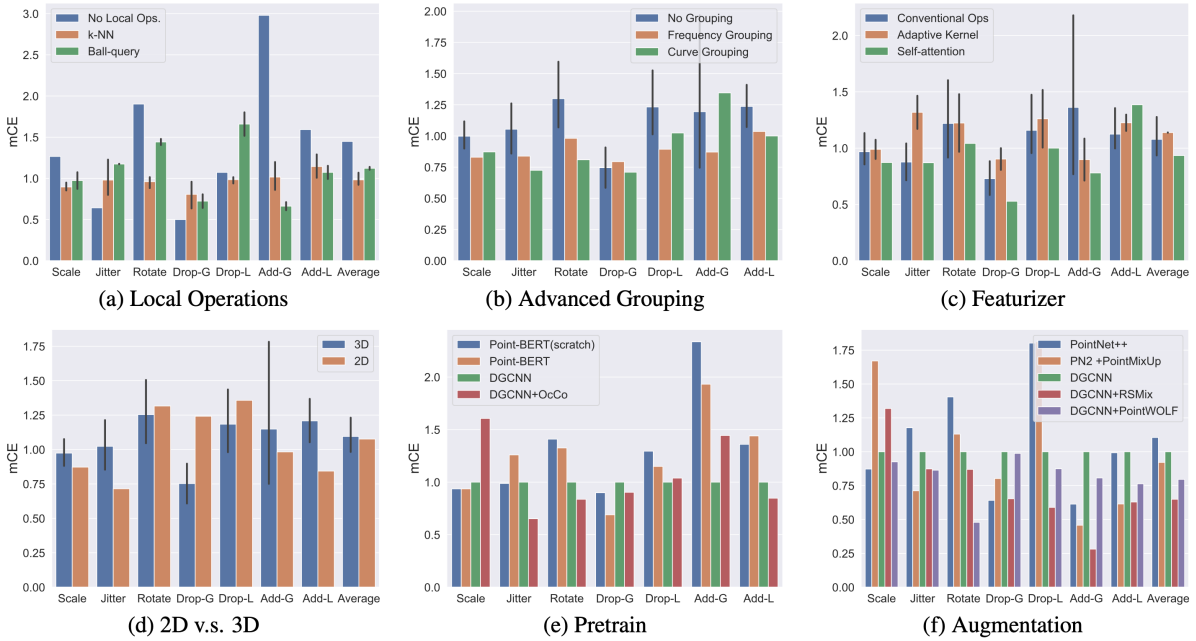
What makes a robust point cloud classifier?

- **Three main components**: 1) architecture design, 2) self- supervised pretraining 3) augmentation methods.



What makes a robust point cloud classifier?

- We conduct a comprehensive analysis and observe:
  - Proper architecture designs can improve robustness, e.g., advanced grouping and self-attention.
  - Pretrain signals can be transferred, benefiting robustness under specific corruptions.
  - Mixing and deformation augmentations can bring significant improvements to model robustness.

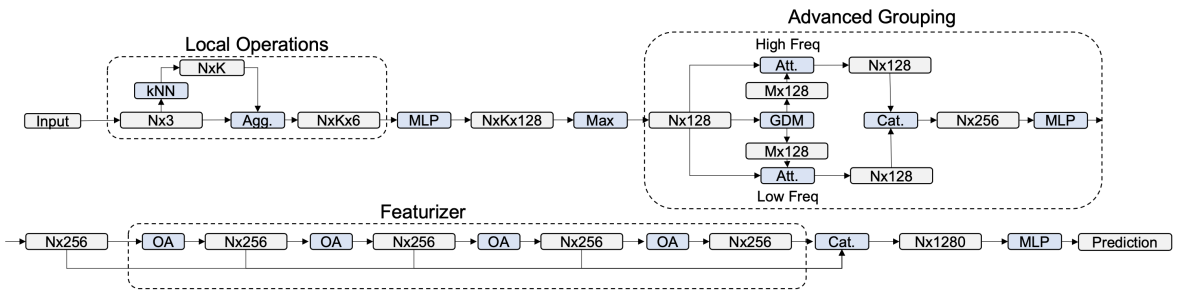


(a) Local Operations (b) Advanced Grouping (c) Featurizer

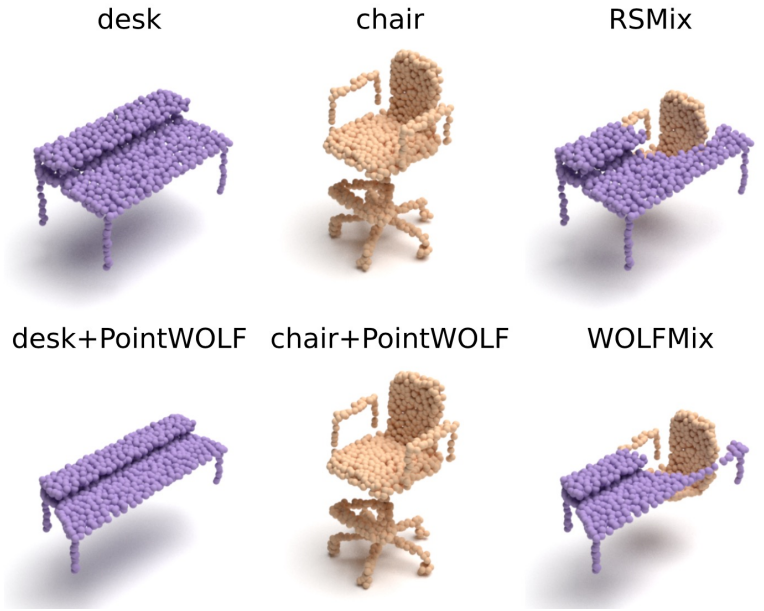
(d) 2D v.s. 3D (e) Pretrain (f) Augmentation

What makes a robust point cloud classifier?

- For verification, we propose a new architecture and a new augmentation technique strictly following our empirical findings.
- They *outperform* existing methods.



Our proposed architecture *RPC*



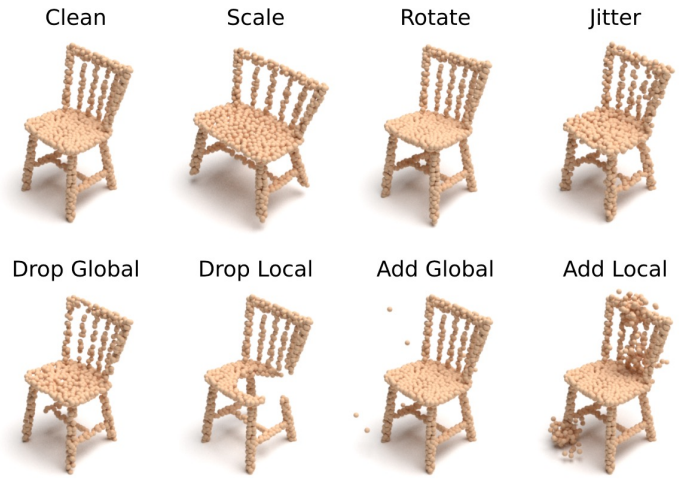
Our proposed augmentation *WolfMix*

Overview

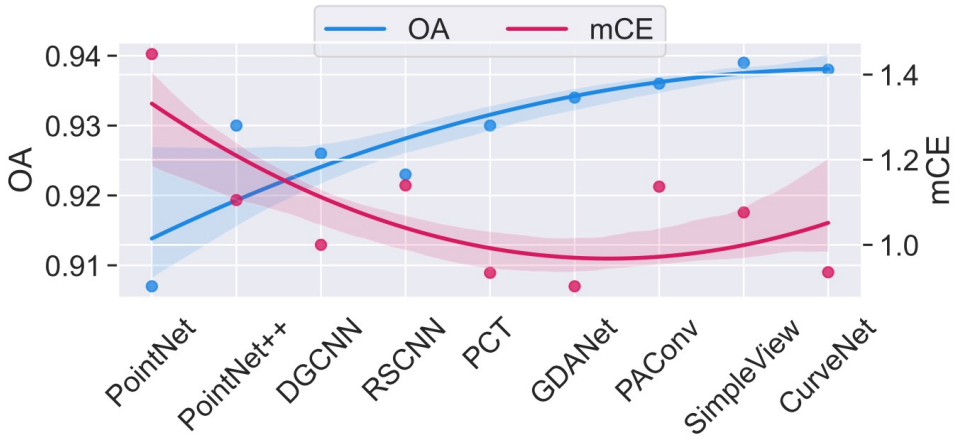
- Introduction
  - Why is robustness crucial in point cloud classification?
  - How has the robustness been evaluated?
- ModelNet-C
  - What should a comprehensive corruption set include?
  - How to design the corrupted test set?
  - How to measure the robustness?
- Benchmark Results
  - Are point cloud classifier getting more robust?
  - What makes a robust point cloud classifier?
- **Conclusion: A call for robustness in point cloud classification**

Conclusion

- The SoTA methods for point cloud classification on clean data are becoming **less robust** to random real-world corruptions.
- We highly encourage future research to **focus on classification robustness** so as to benefit real applications.



ModelNet-C



Benchmark Result

Thank you for listening!

- Code: https://github.com/jiawei-ren/ModelNet-C