Point Cloud Augmentation with Weighted Local Transformation

Sihyeon Kim[1*], Sanghyeok Lee[1*], Dasol Hwang[1], Jaewon Lee[1], Seong Jae Hwang[2], Hyunwoo J. Kim[1†]

Korea University[1], University of Pittsburgh[2]

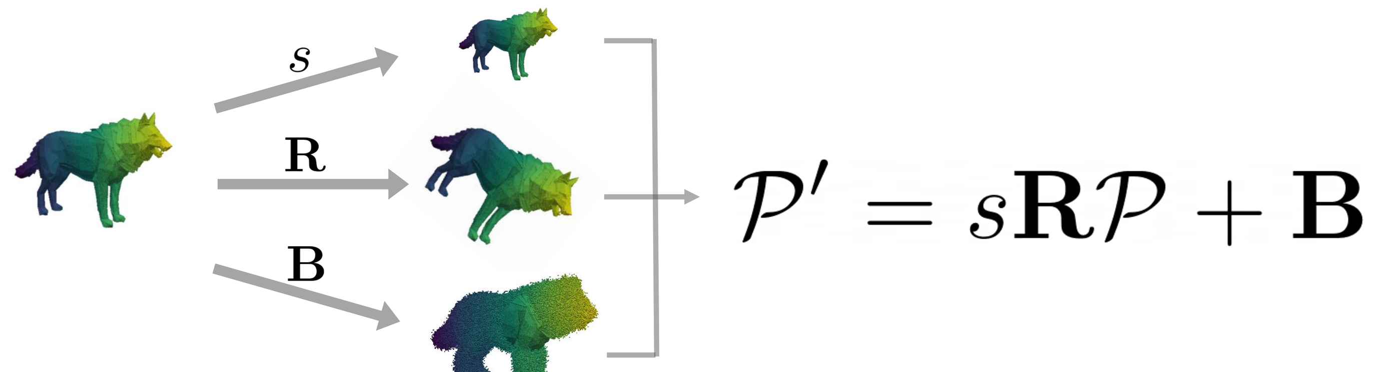
2021 ICCV OCTOBER 11-17 VIRTUAL

Motivation

Only limited data exist for 3D Deep Learning

Data Augmentation is ESSENTIAL!

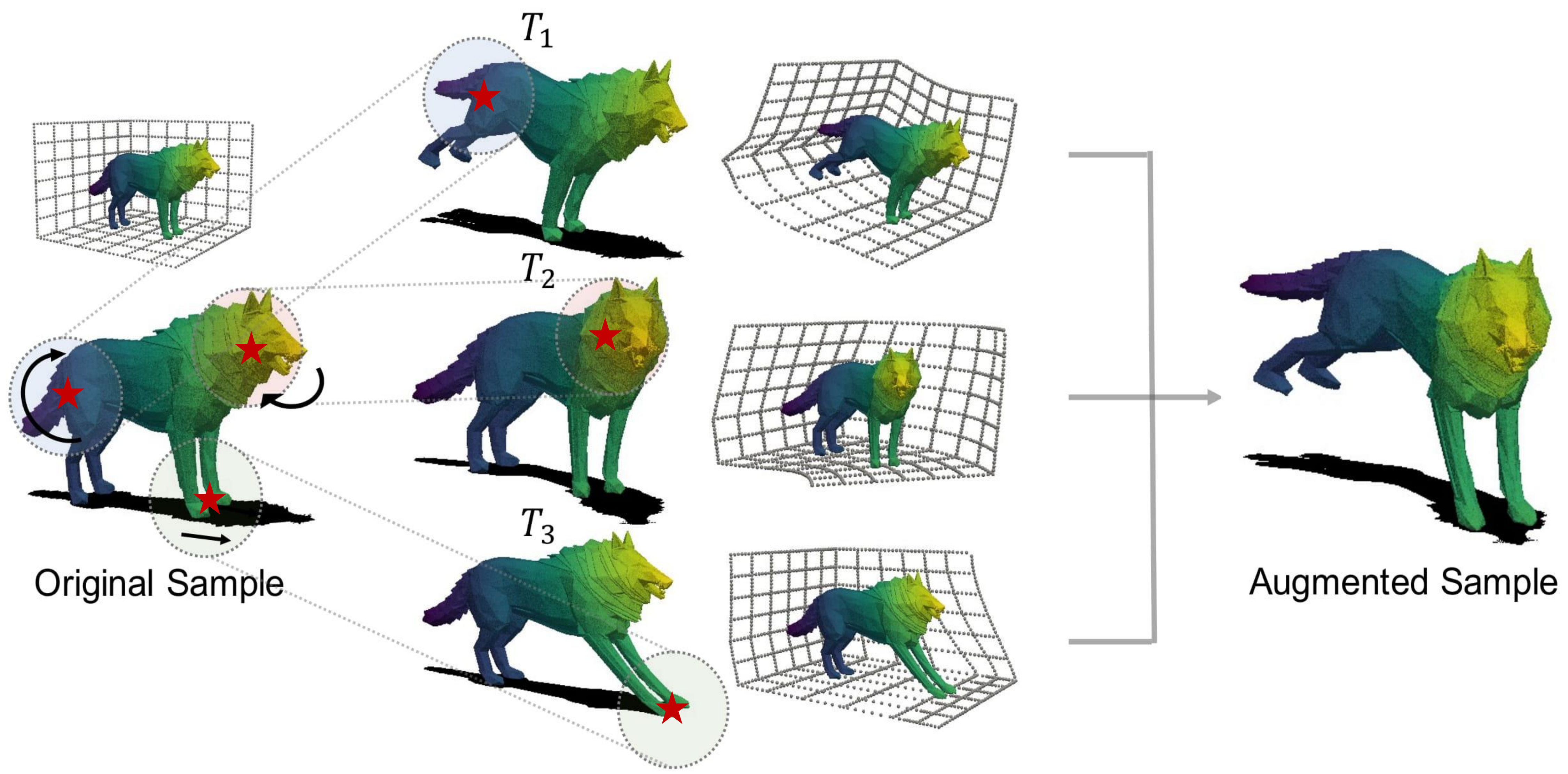
- Previous works are limited to *Global* augmentations.
- For example, Conventional Data Augmentation (CDA) generates the augmented $\mathcal{P}'$ with a global scaling factor s, a rotation matrix $\mathbf{R}$, and a translation matrix $\mathbf{B}$ including translation and point-wise noise.



$$\mathcal{P}' = s\mathbf{R}\mathcal{P} + \mathbf{B}$$

- PointWOLF explicitly considers the **local transformation**, which is crucial for point cloud data.

PointWOLF



- Given an anchor point $\mathbf{p}_j^{\mathcal{A}}$ (★), the **local transformation** for an input point $\mathbf{p}_i$ can be written as:

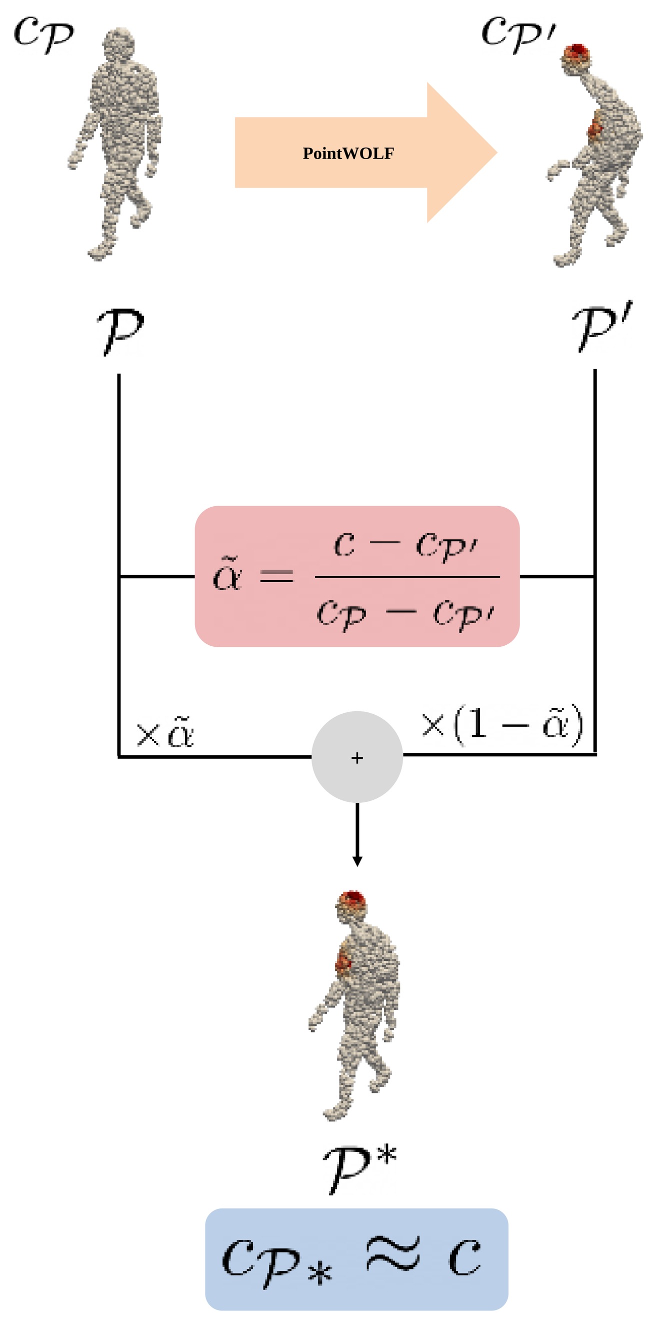
$$\mathbf{p}_i^j = \mathbf{S}_j\mathbf{R}_j\left(\mathbf{p}_i - \mathbf{p}_j^{\mathcal{A}}\right) + \mathbf{b}_j + \mathbf{p}_j^{\mathcal{A}}$$

  where $\mathbf{R}_j$, $\mathbf{S}_j$ and $\mathbf{b}_j$ are rotation matrix, scaling matrix and translation vector specifically correspond to $\mathbf{p}_j^{\mathcal{A}}$.
- By employing **kernel regression** to smoothly interpolate the local transformations, **smooth** deformations $\hat{T}(\cdot)$ generates diverse and realistic augmented samples.

$$\hat{T}(\mathbf{p}_i) = \frac{\sum_{j=1}^{M} K_h(\mathbf{p}_i, \mathbf{p}_j^{\mathcal{A}})T_j}{\sum_{k=1}^{M} K_h(\mathbf{p}_i, \mathbf{p}_k^{\mathcal{A}})}$$

  $K_h(\cdot,\cdot)$ is a kernel function and T_j is the local transformation centered at $\mathbf{p}_j^{\mathcal{A}}$.
- Below are examples with ModelNet40

AugTune: Effective DA Tuning Method



- An effective scheme to adaptively adjust the strength of DA with a single hyperparameter λ.
- Defines a target confidence score c for the final augmented $\mathcal{P}^*$ as below:

$$c = \max(c_{\mathcal{P}'}, (1-\lambda)c_{\mathcal{P}})$$

- To generate $\mathcal{P}^*$ with the target confidence c, AugTune uses **linear interpolation** between the original $\mathcal{P}$ and the augmented $\mathcal{P}'$ with α^*, which is the solution to

$$\alpha^* = \underset{\alpha}{\operatorname{argmin}} \parallel c - f(\alpha\mathcal{P} + (1-\alpha)\mathcal{P}') \parallel^2$$

- Due to computational expense, we approximate α^* by $\tilde{\alpha}$ as below:

$$\alpha c_{\mathcal{P}} + (1-\alpha)c_{\mathcal{P}'} = c$$

Experiments

Shape Classification

Table 1. **Overall accuracy on ModelNet40.**

Dataset	Model	CDA	CDA (w/o R)	PointAugment [4]	PointMixup [5]	**PointWOLF**
MN40	PointNet	89.2	89.7	90.8	89.9	**91.1**
	PointNet++	91.3	92.5	92.4	92.7	**93.2**
	DGCNN	91.7	92.7	92.9	93.1	**93.2**
ReducedMN40	PointNet	81.9	82.7	84.1	83.4	**85.7**
	PointNet++	85.9	87.8	87.0	88.6	**88.7**
	DGCNN	87.5	88.8	88.3	89.0	**89.3**

Table 2. **Overall accuracy on ScanObjectNN.**

Dataset	Model	CDA	PointAugment [4]	PointMixup [5]	**PointWOLF**
OBJ_ONLY	PointNet	76.1	74.4	-	**78.7**
	PointNet++	86.6	85.4	88.5	**89.7**
	DGCNN	85.7	83.1	-	**88.8**
PB_T50_RS	PointNet	64.0	57.0	-	**67.1**
	PointNet++	79.4	77.9	80.6	**84.1**
	DGCNN	77.3	76.8	-	**81.6**

Part Segmentation

Table 3. **Overall mean IoU ($mIoU$) on ShapeNetPart.**

Method	air plane	bag	cap	car	chair	ear phone	guitar	knife	lamp	labtop	motor bike	mug	pistol	rocket	skate board	table	mIoU
PointNet	81.8	**74.7**	**80.2**	71.9	89.6	71.5	90.3	84.9	79.5	95.2	**65.2**	91.1	**81.1**	55.1	72.8	82.2	83.5
+PointWOLF	**82.5**	73.3	78.8	**73.2**	89.6	**72.2**	**91.2**	**86.2**	**79.7**	95.2	64.6	**92.5**	80.2	**56.6**	**73.1**	82.2	**83.8**
PointNet++	81.9	83.4	86.4	**78.6**	90.5	64.7	**91.4**	83.1	83.4	95.1	**69.6**	**94.7**	**82.8**	56.9	**76.0**	**82.3**	84.8
+PointWOLF	**82.0**	**83.9**	**87.3**	77.6	**90.6**	**78.4**	91.1	**87.6**	**84.7**	**95.2**	62.0	94.5	81.3	**62.5**	75.7	83.2	**85.2**
DGCNN	82.2	**75.1**	81.3	**78.2**	90.6	73.6	90.8	87.8	84.4	95.6	**57.8**	92.8	**80.6**	51.5	73.9	82.8	84.8
+PointWOLF	**82.9**	73.3	**83.5**	76.7	**90.8**	**76.7**	**91.4**	**89.2**	**85.2**	**95.8**	53.7	**94.0**	80.1	**54.9**	**74.3**	**83.4**	**85.2**

Robustness to Global and Local Corruption

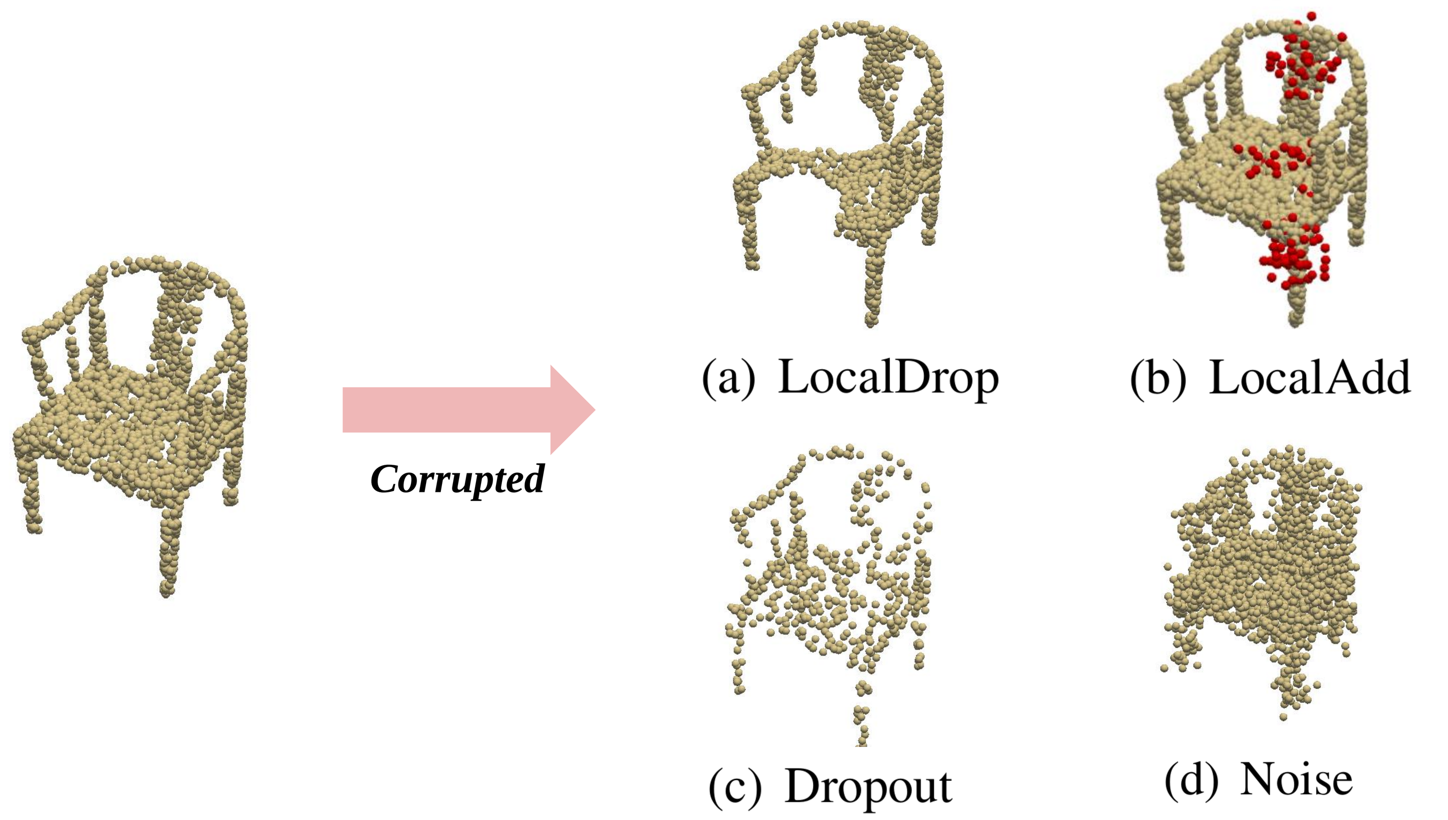


Table 7. **Robustness to Corruption.**

Corruption		CDA	PointWOLF
LocalDrop	$\mathcal{C}$=3	67.0	68.8 (1.8↑)
	$\mathcal{C}$=5	63.2	66.5 (3.3↑)
	$\mathcal{C}$=7	52.8	60.0 (7.2↑)
LocalAdd	$\mathcal{C}$=3	73.9	77.2 (3.3↑)
	$\mathcal{C}$=5	63.5	69.4 (5.9↑)
	$\mathcal{C}$=7	51.5	64.6 (13.1↑)
Dropout	r=0.25	91.2	92.2 (1.0↑)
	r=0.5	84.0	90.4 (6.4↑)
	r=0.75	29.5	60.8 (31.3↑)
Noise	σ=0.01	91.5	93.0 (1.5↑)
	σ=0.03	78.8	87.6 (8.8↑)
	σ=0.05	22.9	45.1 (22.2↑)

Conclusion

- 3D Data Augmentation produces smoothly varying non-rigid deformations by locally weighted transformations.
- Diverse and realistic augmented samples bring consistent improvements over existing augmentations.
- Validates the importance of understanding the local structure and improves the robustness against corruptions.
- Suggests an efficient scheme to adaptively adjust the strength of augmentation during training with a single hyperparameter.

Codes are available at https://github.com/mlvlab/PointWOLF