SE-ORNet: Self-Ensembling Orientation-aware Network for Unsupervised Point Cloud Shape Correspondence

Jiacheng Deng[1]* Chuxin Wang[1]* Jiahao Lu[1] Jianfeng He[1] Tianzhu Zhang[1,3] Jiyang Yu[2] Zhe Zhang[3]

[1]University of Science and Technology of China [2]China Academy of Space Technology [3]Deep Space Exploration Lab

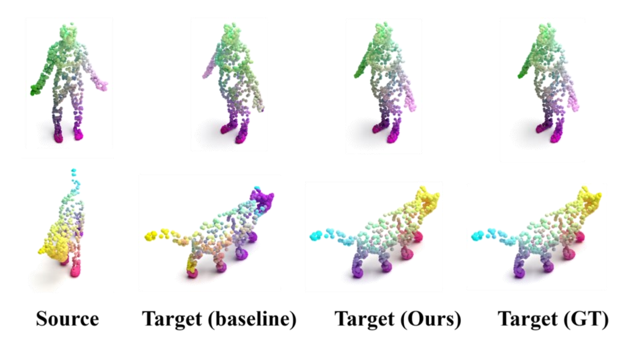
JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

Point Cloud Shape Correspondence

Input: a Pair of Point Clouds

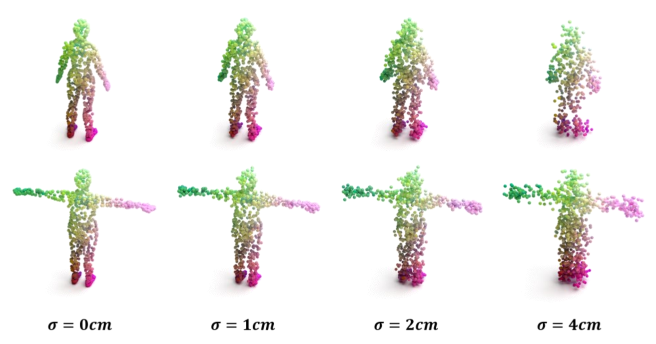
Output: Point-to-point Correspondence

Challenges and Contributions



1. Mismatching issue of symmetrical parts in point clouds with different body orientations

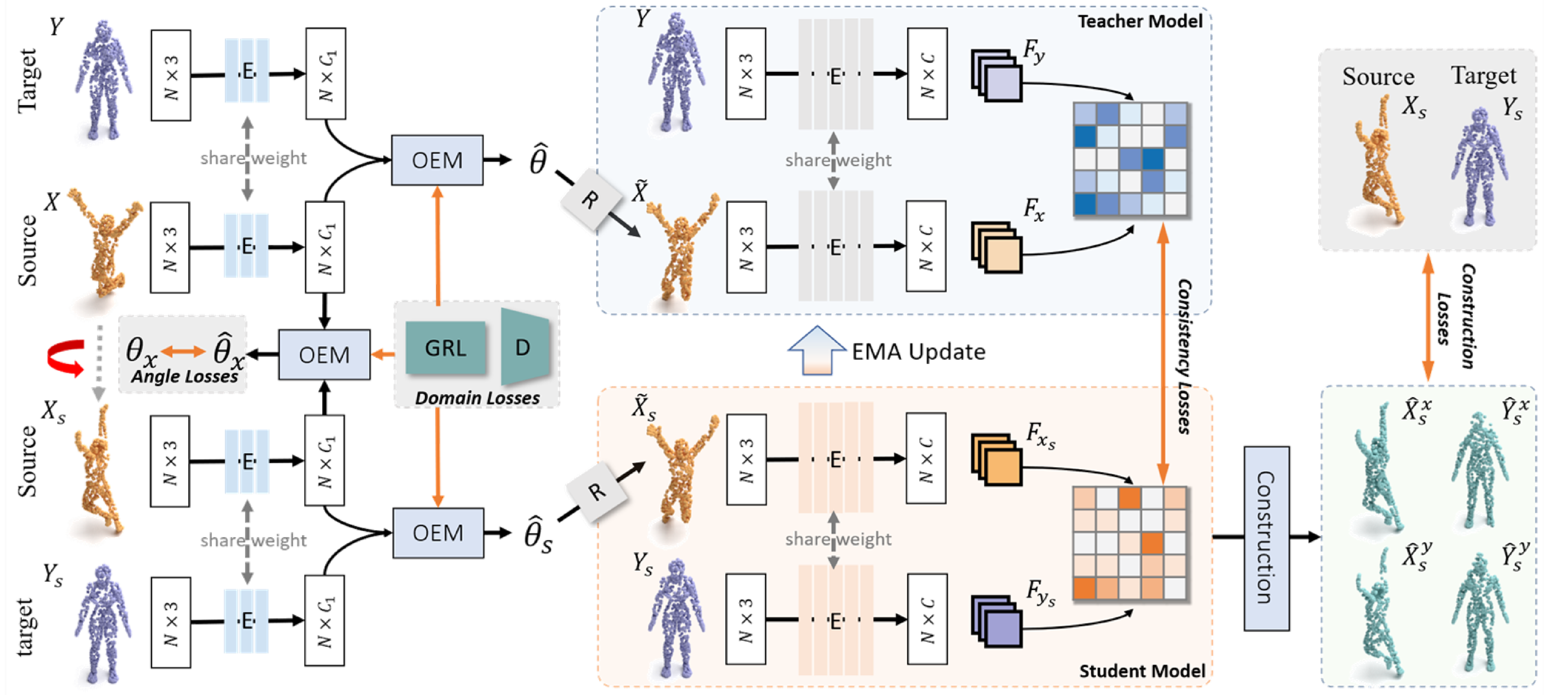
We design a lightweight **Orientation Estimation Module** that accurately aligns the orientations of point cloud pairs to achieve correct matching results of symmetrical parts.



2. Point cloud noise perturbs the spatial coordinates of point cloud and interferes with local structure modeling

We integrate orientation modeling and consistent point cloud representation learning with the disturbance of noise into **a unified self-ensembling framework**.

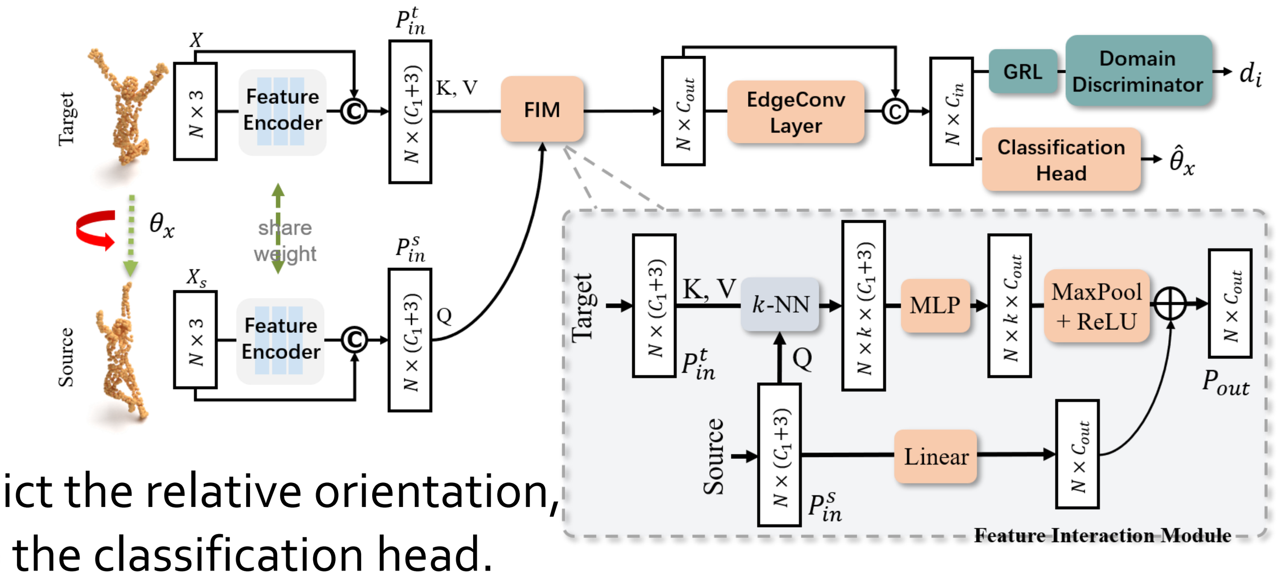
Our Framework



Our Approach

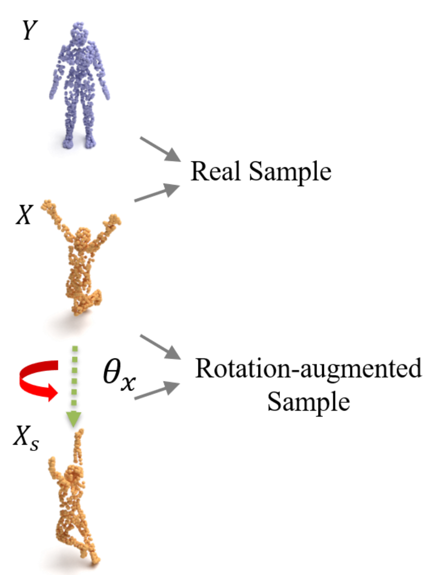
(a) Orientation Estimation Module

- Input & Process: The encoded features P_{in}^s, P_{in}^t are fed into the Feature Interaction Module for feature fusion. The fused features are enhanced by a single-layer edgeconv.
- Classification Head: To predict the relative orientation, the features $\widehat{P_{in}}$ are fed into the classification head.
- Domain Adaptation: We also input $\widehat{P_{in}}$ into the discriminator to determine whether the pair comes from the same shape or from different shapes.



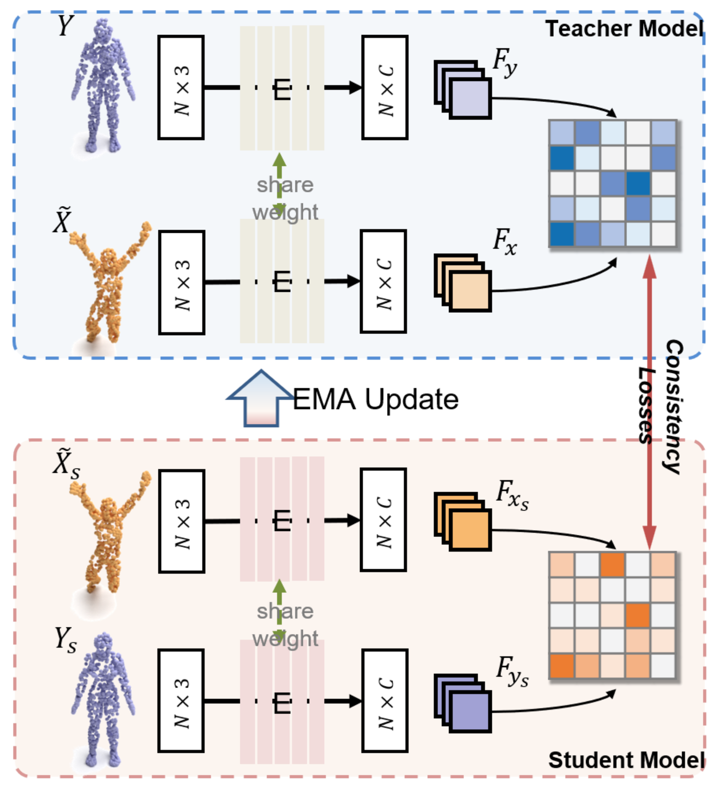
(b) Adversarial Domain Adaptation

- OEM learning: Due to the absence of relative rotation angle of the real samples, we utilize the relative rotation angle of the rotation-augmented samples to guide the OEM learning.
- To eliminate the domain gap: we use a discriminator to identify whether the input features of the classification head are from real samples or not.



(c) Self-Ensembling Framework

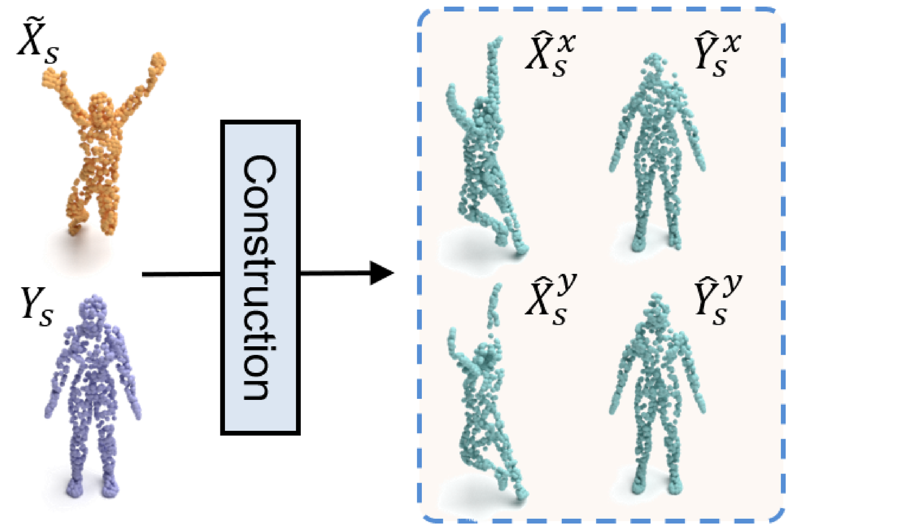
- Stochastic Transform: We apply stochastic transformations that include rotation and Gaussian noise on the point clouds for the student network formulated.
- Teacher & Student Models: Our approach follows the Mean Teacher paradigm and inputs the aligned point cloud pairs into the student and teacher models, respectively.
- Soft label: We take the output of the teacher model as soft labels and design two consistency losses to maintain feature consistency.



(d) Total Loss

- Consistence loss:
  - Computed with a smooth L1 loss
  - The consistency of the cross-similarities
  - The consistency of the self-similarities
- Relative Rotation Angle Loss
  - Computed with a cross entropy loss
- Adversarial Domain Loss:
  - Computed with a Focal Loss
- Construction Loss
  - Following DPC, we use the cross-construction operation to construct the target shape by using the feature similarity.
- Regularization Term
  - Correspond close points in the source to close points in the target.

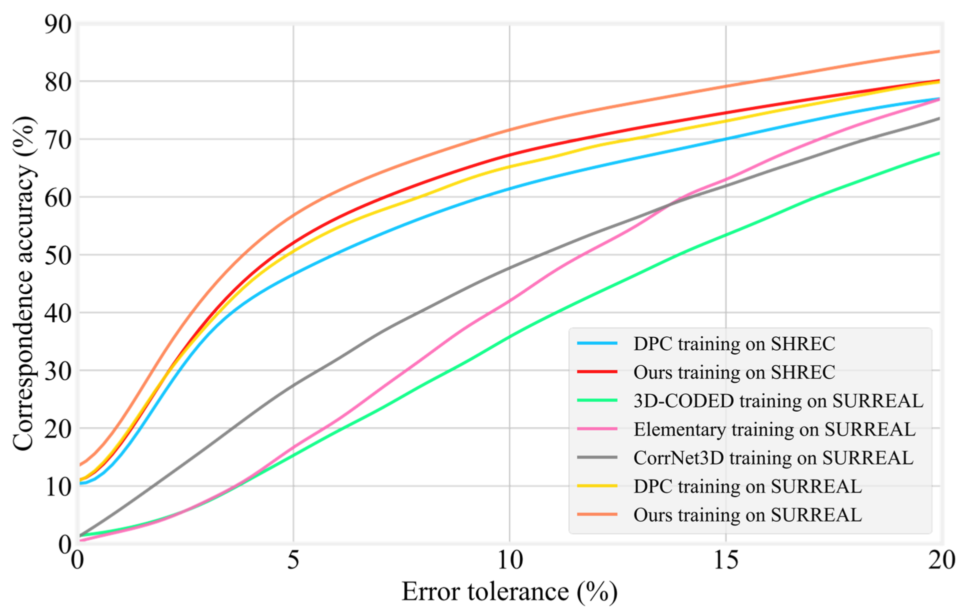
$$\mathcal{L}_{total} = \lambda_1 \mathcal{L}_{ccs} + \lambda_2 \mathcal{L}_{css} + \lambda_3 \mathcal{L}_{angle} + \lambda_4 \mathcal{L}_{domain} + \mathcal{L}_{cons} + \lambda_5 \mathcal{L}_{norm}$$



Experiments

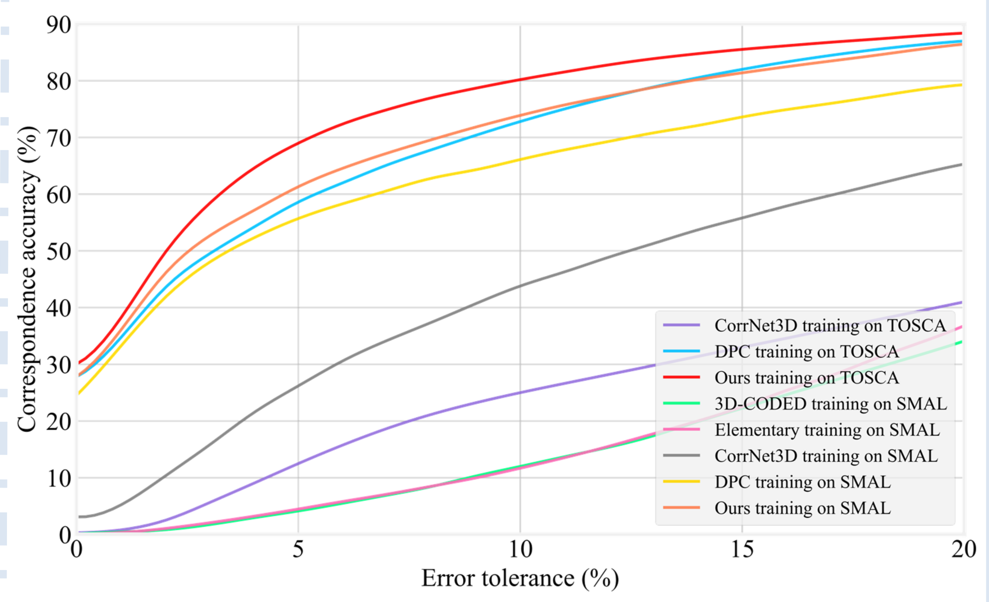
(a) On Human Benchmarks

Method	Input	SHREC acc ↑	SHREC err ↓	SURREAL acc ↑	SURREAL err ↓
Diff-FMaps [16]	Point	/	/	4.0%	7.1
3D-CODED [8]	Point	/	/	2.1%	8.1
Elementary [4]	Point	/	/	2.3%	7.6
CorrNet3D [32]	Point	0.4%	33.8	6.0%	6.9
DPC [12]	Point	15.3%	5.6	17.7%	6.1
Ours	Point	**17.5%**	**5.1**	**21.5%**	**4.6**

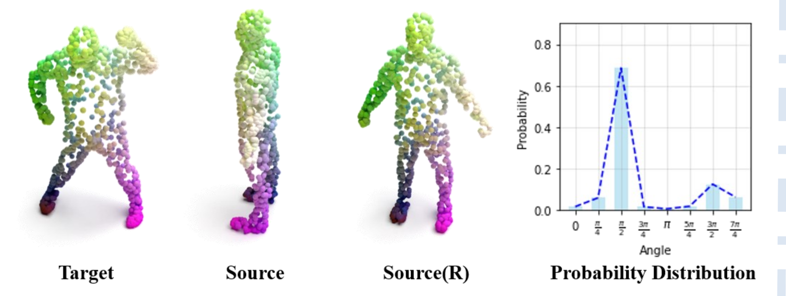


(b) On Animal Benchmarks

Method	TOSCA acc ↑	TOSCA err ↓	SMAL acc ↑	SMAL err ↓
3D-CODED [8]	/	/	0.5%	19.2
Elementary [4]	/	/	0.5%	13.7
CorrNet3D [32]	0.3%	32.7	5.3%	9.8
DPC [12]	34.7%	2.8	33.2%	5.8
Ours	**38.2%**	**2.7**	**36.4%**	**3.9**



(c) Effect of the OEM



(d) On Real Scanned Owlii Dataset