Combinative Matching for Geometric Shape Assembly

Nahyuk Lee [1*] Juhong Min [1,2*] Junhong Lee [1] Chunghyun Park [1] Minsu Cho [1,3]

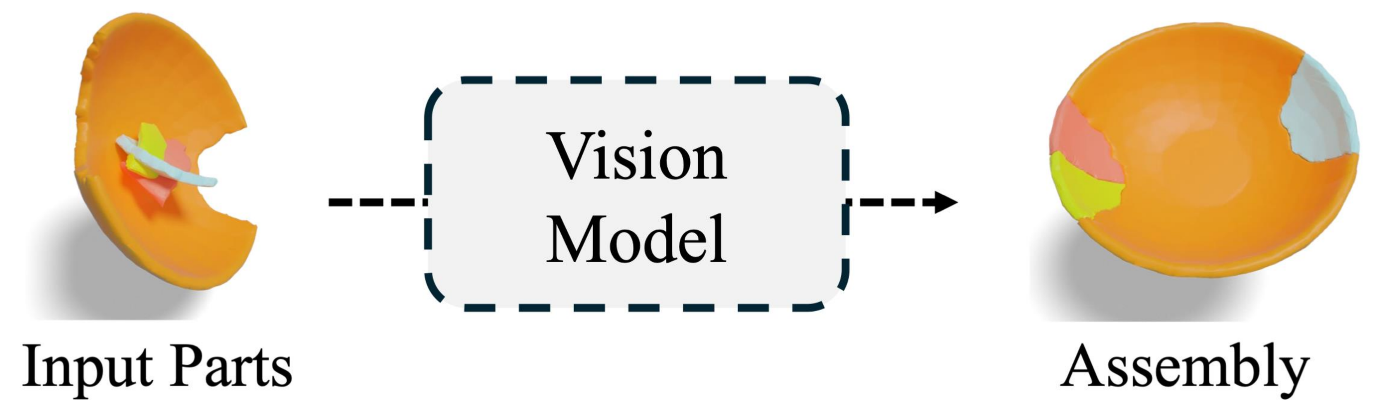
[1] Pohang University of Science and Technology (POSTECH) [2] Samsung Research America [3] RLWRLD

Project page

Paper

ICCV HONOLULU HAWAII OCT 19-23, 2025

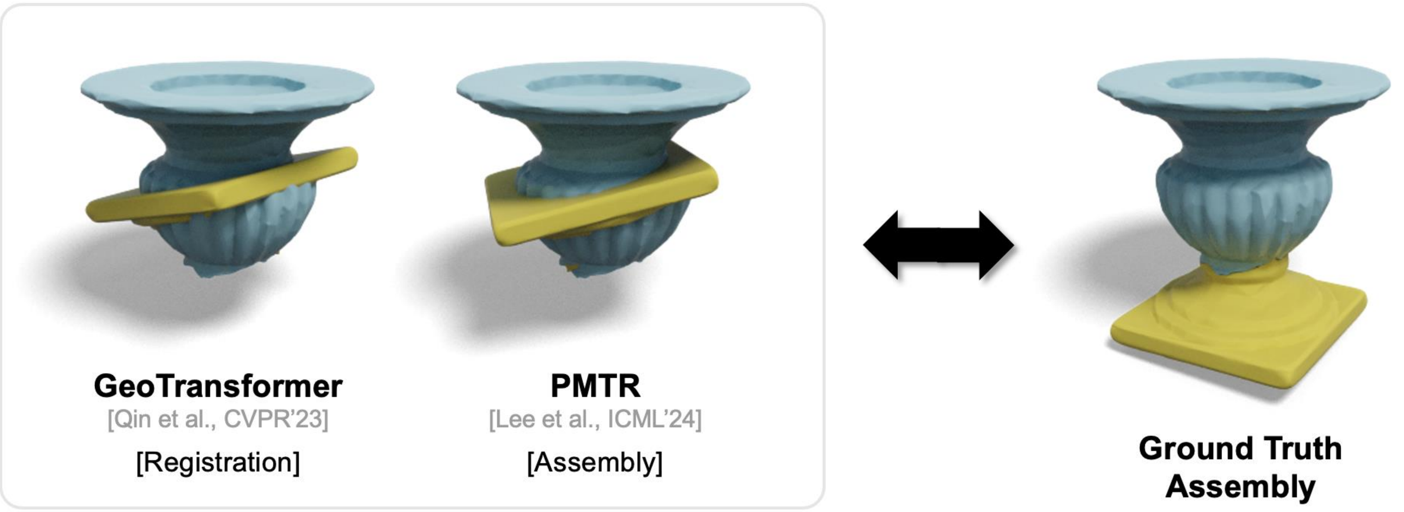
Geometric Shape Assembly?



Given a set of fractured parts, our goal is to predict the **6-DoF pose in SE(3) for each part** to restore the underlying object

Motivation and Overview

Traditional shape matching methods rely on ***equative matching strategy***, which assumes that mating parts resemble each other at their surfaces.



➡ Often falls short in geometric shape assembly, where parts are **not merely visually similar**, but are **structurally complementary**.

Then, what do we miss? 🤔

➡ ***Complementary properties:*** *Surfaces must not only look alike but also* ***fit by occupying opposite volumes****.*

Main Contributions

- **Combinative Matching:** a new shape-matching methodology to combine interlocking parts for shape assembly.
- **Combinative Matching Network:** a framework utilizing combinative matching, achieving SoTA on Breaking Bad.

Learning to interlock parts: Combinative Matching

We propose ***Combinative Matching***, explicitly modeling both **identical surface shapes** and **opposite volume occupancy**, enabling robust combination of interlocking parts.

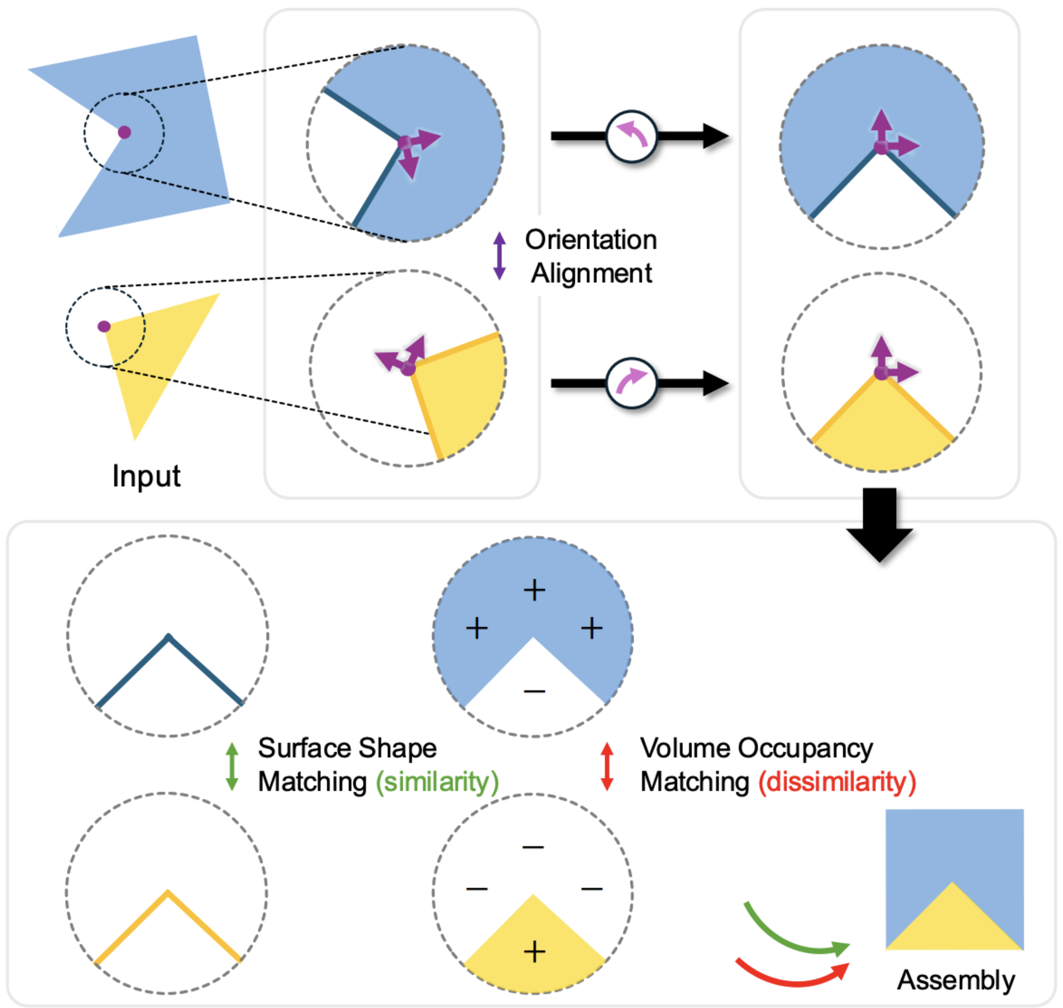
$$\mathcal{L}_{\mathrm{d}} = \frac{1}{|\mathcal{C}|} \sum_{(i,j)\in\mathcal{C}} \left\| (\mathbf{F}_{\mathrm{d}}^{\mathrm{P}})_i \,\mathbf{R}^{\mathrm{P}} - (\mathbf{F}_{\mathrm{d}}^{\mathrm{Q}})_j \,\mathbf{R}^{\mathrm{Q}} \right\|_F$$

$$\mathcal{L}_{\mathrm{s}} = \mathop{\mathbb{E}}_{i\sim\mathcal{I}} \left[\log\left(1 + \sum_{j\in\mathcal{E}_{\mathrm{p}}(i)} e^{\alpha_{ij}(d_{ij}^{\mathrm{p}}-\Delta_{\mathrm{p}})} \cdot \sum_{k\in\mathcal{E}_{\mathrm{n}}(i)} e^{\beta_{ik}(\Delta_{\mathrm{n}}-d_{ik}^{\mathrm{n}})}\right)\right]$$

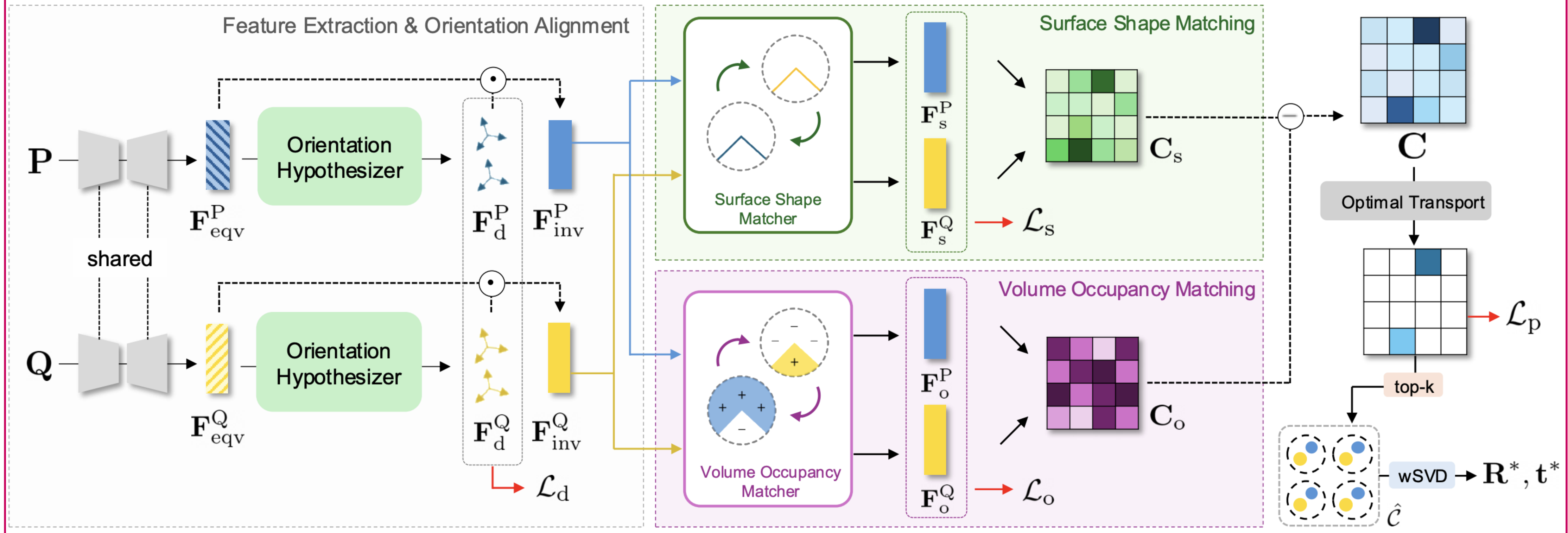
$$\alpha_{ij} = \gamma[d_{ij}^{\mathrm{p}} - \Delta_{\mathrm{p}}]_+, \quad \beta_{ik} = \gamma[\Delta_{\mathrm{n}} - d_{ik}^{\mathrm{n}}]_+ \qquad d_{ij}^{\mathrm{p}} = \|\hat{\mathbf{F}}_{\mathrm{s},i}^{\mathrm{P}} - \hat{\mathbf{F}}_{\mathrm{s},j}^{\mathrm{Q}}\|_2$$

$$\mathcal{L}_{\mathrm{o}} = \mathop{\mathbb{E}}_{i\sim\mathcal{I}} \left[\log\left(1 + \sum_{j\in\mathcal{E}_{\mathrm{p}}(i)} e^{\alpha_{ij}(s_{ij}^{\mathrm{p}}-\Delta_{\mathrm{p}})} \cdot \sum_{k\in\mathcal{E}_{\mathrm{n}}(i)} e^{\beta_{ik}(\Delta_{\mathrm{n}}-s_{ik}^{\mathrm{n}})}\right)\right]$$

$$\alpha_{ij} = \gamma[s_{ij}^{\mathrm{p}} - \Delta_{\mathrm{p}}]_+, \quad \beta_{ik} = \gamma[\Delta_{\mathrm{n}} - s_{ik}^{\mathrm{n}}]_+ \qquad s_{ij}^{\mathrm{p}} \approx \cos\left(\mathbf{F}_{\mathrm{o},i}^{\mathrm{P}}, \mathbf{F}_{\mathrm{o},j}^{\mathrm{Q}}\right)$$



Combinative Matching Network (CMNet)



Step 1. Orientation Alignment
- Encode rotation-equivariant features (w/ VN-Layers)
- Predict per-point orientations
- Derive **rotation-invariant embeddings** via dot-product

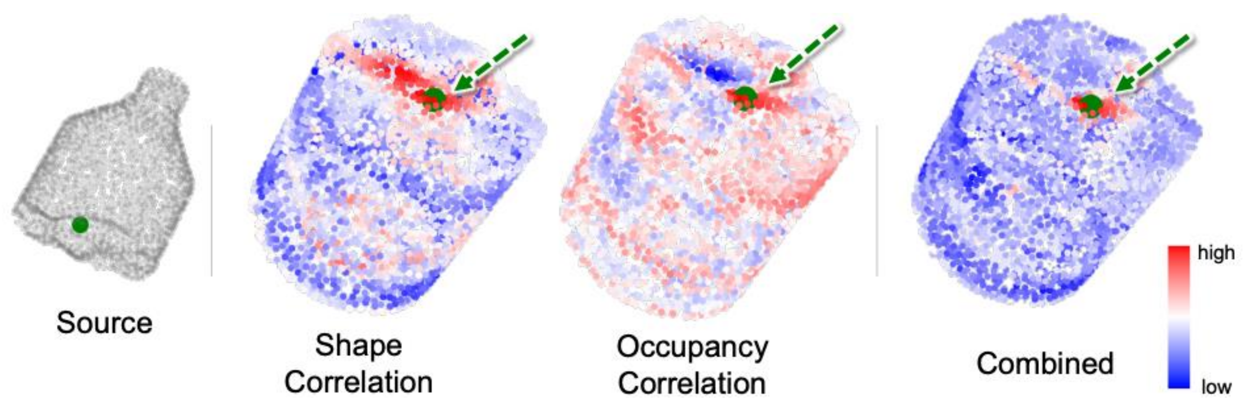
Step 2. Combinative Matching
- **Dual matching branches:** shape & occupancy matching
- Learn to align both identical surface shape and opposite volume occupancy

Step 3. Transformation Estimation
- Combine **shape / occupancy correlations** → Optimal Transport
- Estimate SE(3) pose via wSVD
- Further utilized as pairwise matcher for multi-part assembly

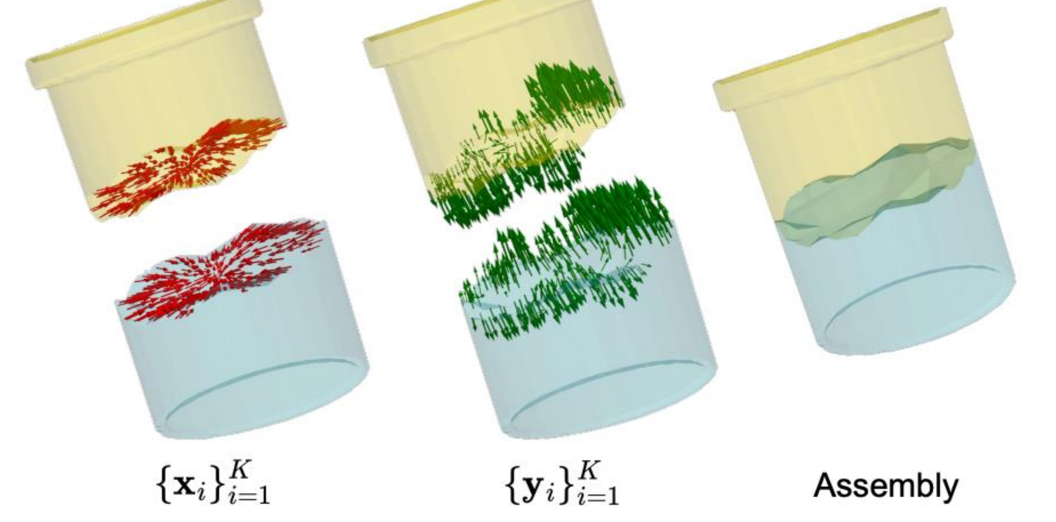
Experiments and Analysis

Learned Descriptor Analysis

(left: correlation distribution, right: orientation visualization)



By combining shape and occupancy information, the local ambiguity and match confidence uncertainty are resolved.



Learned orientations capture complementarity between parts ***without any explicit supervision***.

Ablation Studies on Combinative Matching

Occupancy Affinity	Orientation Loss ($\mathcal{L}_d$)	CRD ↓ (10^{-2})	CD ↓ (10^{-3})	RMSE(R) ↓ (°)	RMSE(T) ↓ (10^{-2})
L2 dist	✗	0.42	0.31	14.88	4.31
cosine	✗	0.31	0.21	14.58	4.44
L2 dist	✓	0.38	0.30	13.29	3.81
cosine	✓	**0.28**	**0.17**	**12.88**	**3.78**

(a) Ablation study on combinative matching.

Equivariant Embedding	Shape Matching	Occupancy Matching	CRD ↓ (10^{-2})	CD ↓ (10^{-3})	RMSE(R) ↓ (°)	RMSE(T) ↓ (10^{-2})
✗	✓	✓	0.74	0.53	38.74	11.88
✓	✗	✓	0.38	0.28	13.17	3.86
✓	✓	✗	0.35	0.25	14.01	4.24
✓	✓	✓	**0.28**	**0.17**	**12.88**	**3.78**

(b) Ablation study on model components.

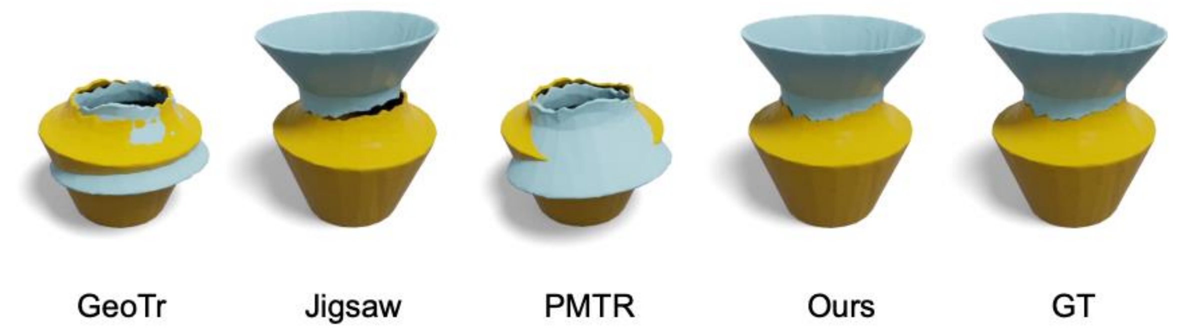
Learning **complementary volume occupancy** under **orientation alignment** is crucial for accurate geometric shape assembly.

Incorporating **joint shape and occupancy matching** under **SO(3)-equivariance** significantly enhances the assembly accuracy and robustness.

Improving SOTA for Geometric Shape Assembly

(top: pairwise assembly, bottom: multi-part assembly)

Method	CRD ↓ (10^{-2})	CD ↓ (10^{-3})	RMSE(R) ↓ (°)	RMSE(T) ↓ (10^{-2})
everyday				
Wu et al. [44]	20.65	11.66	84.58	22.90
GeoTransformer [32]	0.61	0.51	22.81	7.28
Jigsaw [22]	5.48	1.34	38.73	**2.73**
PMTR [14]	<u>0.39</u>	<u>0.25</u>	<u>17.14</u>	5.53
CMNet (Ours)	**0.28**	**0.17**	**12.88**	<u>3.78</u>



Method	CRD ↓ (10^{-2})	CD ↓ (10^{-3})	RMSE(R) ↓ (°)	RMSE(T) ↓ (10^{-2})	PA_{CRD} ↑ (%)	PA_{CD} ↑ (%)
everyday						
Wu et al. [48]	28.18	19.70	54.98	15.59	35.66	36.28
Jigsaw [24]	14.13	11.82	41.12	11.74	52.48	60.26
PMTR [16]	<u>6.51</u>	<u>5.56</u>	<u>31.57</u>	<u>9.95</u>	<u>66.95</u>	<u>70.56</u>
CMNet (Ours)	**5.18**	**3.65**	**27.11**	**8.13**	**73.88**	**77.88**