Supplementary Material:

SIRA-PCR: Sim-to-Real Adaptation for 3D Point Cloud Registration

Suyi Chen[1,4*] Hao Xu[2*] Ru Li[3] Guanghui Liu[1†] Chi-Wing Fu[2] Shuaicheng Liu[1,4†]

[1]University of Electronic Science and Technology of China
[2]The Chinese University of Hong Kong [3]Harbin Institute of Technology [4]Megvii Technology

*Equal contribution.
†Corresponding authors.

This supplementary material document contains the following two parts.

Part 1 presents details of the evaluation metrics, additional evaluation results, and detailed network architecture of SIRA.

Part 2 presents more visualization and qualitative results.

Part 1.1: Evaluation Metrics in Experiments

Following the existing works, we present the definitions of the evaluation metrics we employed.

(i) **Registration Recall (RR)** is the fraction of successfully registered point cloud pairs. A point cloud pair is said to be successfully registered when its transformation error is lower than threshold τ_1. In addition, the transformation error is defined as the root mean square error of the ground-truth correspondences $\mathcal{C}^*$, to which the estimated transformation $\mathbf{T}_{est}(\cdot)$ has applied:

$$\mathrm{RMSE} = \sqrt{\frac{1}{|\mathcal{C}^*|} \sum_{(\mathbf{p}_x^*, \mathbf{q}_y^*) \in \mathcal{C}^*} \|\mathbf{T}_{est}(\mathbf{p}_x^*) - \mathbf{q}_y^*\|_2^2}, \tag{1}$$

$$\mathrm{RR} = \frac{1}{M} \sum_{i=1}^{M} [\mathrm{RMSE}_i < \tau_1], \tag{2}$$

where $\mathbf{p}_x$ and $\mathbf{q}_y$ denote the x-th point in source $\mathbf{P}$ and y-th point in target $\mathbf{Q}$, respectively; $[\cdot]$ is the Iversion bracket; and M is the number of all point cloud pairs.

(ii) **Inlier Ratio (IR)** is the fraction of inlier correspondences among all hypothesized correspondences $\mathcal{C}$. A correspondence is regarded as an inlier if the distance between the two points is lower than a certain threshold τ_2 under the ground-truth transformation $\mathbf{T}_{gt}(\cdot)$:

$$\mathrm{IR} = \frac{1}{|\mathcal{C}|} \sum_{(\mathbf{p}_x, \mathbf{q}_y) \in \mathcal{C}} [\|\mathbf{T}_{gt}(\mathbf{p}_x) - \mathbf{q}_y\|_2 < \tau_2]. \tag{3}$$

(iii) **Feature Matching Recall (FMR)** is the fraction of point cloud pairs whose IR $>$ threshold τ_3:

$$\mathrm{FMR} = \frac{1}{M} \sum_{i=1}^{M} [\mathrm{IR}_i > \tau_3]. \tag{4}$$

(iv) **Relative Rotation Error (RRE)** is the geodesic distance in degrees between the estimated and ground-truth rotation matrices $\mathbf{R}_{est}$ and $\mathbf{R}_{gt}$:

$$\mathrm{RRE} = \arccos\left(\frac{\mathrm{trace}(\mathbf{R}_{est}^T \cdot \mathbf{R}_{gt} - 1)}{2}\right). \tag{5}$$

(v) **Relative Translation Error (RTE)** is the Euclidean distance between estimated and ground-truth translation vectors $\mathbf{t}_{est}$ and $\mathbf{t}_{gt}$:

$$\mathrm{RTE} = \|\mathbf{t}_{est} - \mathbf{t}_{gt}\|_2. \tag{6}$$

(vi) **Transformation Recall (TR)** is the fraction of successfully registered point cloud pairs, similar to RR but using a different definition of successful registration:

$$\mathrm{TR} = \frac{1}{M} \sum_{i=1}^{M} [\mathrm{RRE}_i < \tau_4 \text{ and } \mathrm{RTE}_i < \tau_5]. \tag{7}$$

TR is also called RR in [1, 11], while we follow [13] to use TR to avoid ambiguity.

Following [2, 7, 11, 15], we set $\tau_1 = 0.2\mathrm{m}$, $\tau_2 = 0.1\mathrm{m}$, $\tau_3 = 0.05$, $\tau_4 = 15°$, and $\tau_5 = 0.3\mathrm{m}$ for evaluation on the 3DMatch and 3DLoMatch benchmarks. Following [13], we set $\tau_1 = 0.5\mathrm{m}$, $\tau_2 = 0.2\mathrm{m}$, and $\tau_3 = 0.05$ for evaluation on the ETH benchmark.

Model	3DMatch									3DLoMatch								
	Kitchen	Home_1	Home_2	Hotel_1	Hotel_2	Hotel_3	Study	Lab	Mean	Kitchen	Home_1	Home_2	Hotel_1	Hotel_2	Hotel_3	Study	Lab	Mean
Registration Recall (%) ↑																		
PerfectMatch [5]	90.6	90.6	65.4	89.6	82.1	80.8	68.4	60.0	78.4	51.4	25.9	44.1	41.1	30.7	36.6	14.0	20.3	33.0
FCGF [4]	98.0	94.3	68.6	96.7	91.0	84.6	76.1	71.1	85.1	60.8	42.2	53.6	53.1	38.0	26.8	16.1	30.4	40.1
D3Feat [2]	96.0	86.8	67.3	90.7	88.5	80.8	78.2	64.4	81.6	49.7	37.2	47.3	47.8	36.5	31.7	15.7	31.9	37.2
PREDATOR [7]	97.6	97.2	74.8	**98.9**	**96.2**	88.5	85.9	73.3	89.0	71.5	58.2	60.8	77.5	64.2	61.0	45.8	39.1	59.8
CoFiNet [16]	96.4	**99.1**	73.6	95.6	91.0	84.6	89.7	84.4	89.3	76.7	66.7	64.0	81.3	65.0	63.4	53.4	69.6	67.5
GeoTrans [11]	98.9	97.2	81.1	**98.9**	89.7	88.5	88.9	88.9	91.5	85.9	73.5	**72.5**	89.5	**73.2**	66.7	55.3	**75.7**	74.0
RegTR* [15]	97.8	90.6	75.5	97.8	94.9	**100.0**	88.5	**91.1**	92.0	66.0	58.2	64.9	72.7	61.3	70.7	53.4	71.0	64.8
Our SIRA-PCR†	97.1	96.2	77.4	97.3	**96.2**	92.3	89.3	86.7	91.5	80.2	65.4	66.2	83.8	61.6	64.3	53.2	64.3	67.4
Our SIRA-PCR‡	**99.1**	98.1	**83.0**	98.4	94.9	96.2	**94.4**	88.9	**94.1**	**89.7**	**78.8**	**72.5**	**91.0**	71.7	**73.8**	**59.9**	**75.7**	**76.6**
Relative Rotation Error (°) ↓																		
PerfectMatch [5]	1.926	1.843	2.324	2.041	1.952	2.908	2.296	2.301	2.199	3.020	3.898	3.427	3.196	3.217	3.328	4.325	3.814	3.528
FCGF [4]	1.767	1.849	2.210	1.867	1.667	2.417	2.024	1.792	1.949	2.904	3.229	3.277	2.768	2.801	2.822	3.372	4.006	3.147
D3Feat [2]	2.016	2.029	2.425	1.990	1.967	2.400	2.346	2.115	2.161	3.226	3.492	3.373	3.330	3.165	2.972	3.708	3.619	3.361
PREDATOR [7]	1.861	1.806	2.473	2.045	1.600	2.458	2.067	1.926	2.029	3.079	2.637	3.220	2.694	2.907	3.390	3.046	3.412	3.048
CoFiNet [16]	1.910	1.835	2.316	1.767	1.753	1.639	2.527	2.345	2.011	3.213	3.119	3.711	2.842	2.897	3.194	4.126	3.138	3.280
GeoTrans [11]	1.797	1.353	1.797	1.528	1.328	1.571	1.952	1.678	1.625	2.356	2.305	2.541	2.455	2.490	2.504	3.010	2.716	2.547
RegTR* [15]	1.729	1.356	1.781	1.660	**1.297**	1.807	**1.578**	**1.352**	1.570	**3.350**	2.458	3.231	2.725	2.435	2.992	3.136	2.434	2.845
Our SIRA-PCR†	**1.659**	1.435	1.848	1.478	1.475	1.446	1.970	1.562	1.609	2.366	2.149	**2.518**	2.434	2.455	2.670	**2.885**	2.314	2.474
Our SIRA-PCR‡	1.687	**1.320**	**1.731**	**1.409**	1.350	**1.443**	1.846	1.530	**1.539**	2.452	**2.027**	2.896	**2.276**	**2.391**	**2.357**	2.836	**1.870**	**2.388**
Relative Translation Error (m) ↓																		
PerfectMatch [5]	0.059	0.070	0.079	0.065	0.074	0.062	0.093	0.065	0.071	0.082	0.098	0.096	0.101	0.080	0.089	0.158	0.120	0.103
FCGF [4]	0.053	0.056	0.071	0.062	0.061	0.055	0.082	0.090	0.066	0.084	0.097	0.076	0.101	0.084	0.077	0.144	0.140	0.100
D3Feat [2]	0.055	0.065	0.080	0.064	0.078	0.049	0.083	0.064	0.067	0.088	0.101	0.086	0.099	0.092	0.075	0.146	0.135	0.103
PREDATOR [7]	0.048	0.055	0.070	0.073	0.060	0.065	0.080	0.063	0.064	0.081	0.080	0.084	0.099	0.096	0.077	0.101	0.130	0.093
CoFiNet [16]	0.047	0.059	0.063	0.063	0.058	0.044	0.087	0.075	0.062	0.080	0.078	0.078	0.099	0.086	0.077	0.131	0.123	0.094
GeoTrans [11]	0.042	0.046	0.059	0.055	0.046	0.050	0.073	0.053	0.053	**0.062**	0.070	0.071	0.080	0.075	**0.049**	0.107	0.083	0.074
RegTR* [15]	0.040	**0.041**	0.058	0.057	**0.042**	**0.039**	**0.053**	0.057	**0.048**	0.080	**0.064**	0.077	0.093	**0.073**	0.060	**0.094**	0.079	0.077
Our SIRA-PCR†	0.041	0.049	**0.056**	0.050	0.048	0.049	0.073	**0.047**	0.052	0.064	0.070	**0.062**	0.080	0.075	0.064	0.114	**0.063**	0.074
Our SIRA-PCR‡	**0.039**	0.048	0.058	**0.049**	0.046	0.047	0.071	0.052	0.051	**0.062**	0.065	0.069	**0.076**	**0.073**	0.063	0.104	0.067	**0.072**

Table 1. Scene-wise registration results on the 3DMatch and the 3DLoMatch. †: the model is totally trained on the synthetic dataset. ‡: the model is fine-tuned on the 3DMatch. *: the results produced using the released model. For better comparison, the best and second-best results are marked in **bold** and underlined.

Part 1.2: More Experimental Results

In this section, we present additional experiments that evaluate our method. Note that we present these additional experiments because different existing methods use different evaluation metrics/approaches.

Following [7, 11], we evaluate the scene-wise registration on the 3DMatch and 3DLoMatch datasets. Tab. 1 reports the results. Since extremely large errors generated from failure cases can easily dominate the results, median RRE and RTE for each scene are reported only on the successfully-registered pairs. As RegTR [15] did not report such scene-wise results, we produce them using the released model. Our method achieves *the best performance* (see the "mean" columns) for *almost all cases* (two datasets and three metrics).

For a fair comparison with Lepard [9], we also report the results using the metrics following them. Here, RR is averaged on all scan pairs, different from the one averaged over scenes in Tab. 1. As Tab. 2 shows, our method achieves *the best performance on all metrics for both datasets*. Specifically, it outperforms Lepard by 2.5/3.3 pp on 3DMatch/3DLoMatch with lower RRE and RTE, showing the robustness and effectiveness of our method.

Following OIF-PCR [14], we report the mean median RRE and RTE obtained by RANSAC for the successfully-registered scan pairs. As Tab. 3 shows, although methods with higher RR tend to generate higher RRE and RTE since more challenging cases are taken into consideration, we can see that our method always has better improvements in RRE and RTE regardless of RR (no matter higher/lower than the competitors); see again Tab. 3. The robustness shown in this experiment is mainly attributed to the absolutely accurate annotations generated by our approach.

Model	3DMatch			3DLoMatch		
	RRE ↓	RTE ↓	RR ↑	RRE ↓	RTE ↓	RR ↑
PREDATOR [7]	2.72	0.078	91.8	4.44	0.116	62,4
Lepard [9]	2.48	0.072	93.5	4.10	0.108	69.0
PREDATOR [7] + ICP	2.06	0.062	92.3	3.46	0.098	65.2
Lepard [9] + ICP	1.96	0.060	93.9	3.17	0.089	71.3
GeoTrans [11]	1.84	0.061	94.1	2.86	0.086	76.1
Our SIRA-PCR†	1.85	0.060	92.7	2.84	0.087	70.2
Our SIRA-PCR‡	**1.80**	**0.059**	**95.4**	**2.74**	**0.084**	**79.4**

Table 2. Registration results on 3DMatch and 3DLoMatch following the metrics used in [9]. †: the model is only trained on the synthetic dataset. ‡: the model is fine-tuned on 3DMatch.

Model	3DMatch										3DLoMatch									
	RRE ↓					RTE ↓					RRE ↓					RTE ↓				
	5000	2500	1000	500	250	5000	2500	1000	500	250	5000	2500	1000	500	250	5000	2500	1000	500	250
GeoTrans [11]	1.871	1.924	1.929	1.959	2.047	0.065	0.067	0.066	0.066	0.068	2.954	3.007	3.129	3.089	3.187	0.090	0.091	0.093	0.093	0.093
OIF-PCR [14]	1.859	1.895	1.940	1.981	2.023	0.064	0.064	0.067	0.070	0.068	3.040	3.026	3.117	3.073	3.203	0.092	0.092	0.092	0.093	0.095
Our SIRA-PCR†	**1.787**	1.894	1.943	2.048	**1.983**	**0.060**	0.061	**0.059**	**0.061**	**0.061**	**2.735**	2.817	2.989	2.972	**2.941**	**0.085**	0.087	0.087	**0.084**	**0.084**
Our SIRA-PCR‡	1.862	**1.767**	**1.891**	**1.907**	2.012	0.063	**0.056**	0.060	**0.061**	**0.061**	2.846	**2.813**	**2.873**	**2.940**	3.110	**0.085**	**0.083**	**0.083**	0.086	0.087

Table 3. Supplementary registration results obtained by RANSAC on 3DMatch and 3DLoMatch. †: the model is only trained on the synthetic dataset. ‡: the model is fine-tuned on 3DMatch.

Following [1, 8, 11, 13], we evaluate our method using RRE, RTE, and TR. As Tab. 4 shows, the difference between RR and TR lies in the definition of successful registration. RR uses RMSE to measure whether a scan pair is finely registered, whereas TR uses RRE and RTE. As Tab. 4 shows, even totally trained only on synthetic data, our method can still achieve comparable performance as the closest competitor GeoTrans [11] on 3DMatch. After being further fine-tuned on the training set of 3DMatch, our method shows the best performance. Particularly for TR, our method surpasses the previous best by 1.5/3.3 pp on 3DMatch/3DLoMatch.

Model	3DMatch			3DLoMatch		
	RRE ↓	RTE ↓	TR ↑	RRE ↓	RTE ↓	TR ↑
FCGF+FGR [17]	2.82	8.36	78.6	5.28	12.98	20.0
FCGF+DGR [3]	2.40	7.48	91.3	4.17	10.82	43.8
FCGF+DHVR [8]	2.25	7.08	91.9	4.14	12.56	54.4
FCGF+PointDSC [1]	2.06	6.55	93.3	3.87	10.39	56.1
GeoTrans [11]	1.98	5.69	95.0	2.98	8.55	77.5
RoReg [13]	**1.84**	6.28	95.5	3.09	9.30	72.0
Our SIRA-PCR†	1.98	5.79	94.7	2.98	8.75	71.5
Our SIRA-PCR‡	1.91	**5.66**	**97.0**	**2.88**	**8.48**	**80.8**

Table 4. Registration results on 3DMatch and 3DLoMatch following the metrics used in [1, 8, 11, 13]. †: the model is only trained on the synthetic dataset. ‡: the model is fine-tuned on 3DMatch.

In [6], an unsupervised domain adaptation approach, named UDGE, has been proposed to apply cubic cropping on 3DMatch to generate pairs. However, as shown in Tab. 5, UDGE shows limited gains compared with the baseline when it is applied to domain adaptation from our synthetic indoor dataset FlyingShapes to the real-world indoor dataset 3DMatch. Since generated point cloud pairs have the same coordinates in overlap regions, it may cause over-fitting of the feature extractor during training, leading to performance degradation. The results also show the necessity of the sim-to-real adaptation (SIRA) in our pipeline.

In Tab. 6, we present supplementary ablation results on FMR and IR metrics. Our SIRA improves performance on all metrics consistently, showing its effectiveness.

Model	3DMatch RRE ↓	3DMatch RTE ↓	3DMatch RR ↑	3DMatch FMR ↑	3DMatch IR ↑	3DLoMatch RRE ↓	3DLoMatch RTE ↓	3DLoMatch RR ↑	3DLoMatch FMR ↑	3DLoMatch IR ↑
(a) baseline	1.672	0.051	87.5	95.5	59.5	2.521	**0.074**	63.8	77.8	30.6
(b) UDGE [6]	1.710	**0.050**	87.1	94.4	65.0	2.666	**0.074**	55.5	71.8	31.5
(c) SIRA†	**1.609**	0.052	**91.5**	**97.7**	**65.5**	**2.474**	**0.074**	**67.4**	**81.7**	**35.8**

Table 5. Comparison with another UDA method in [6].

	Model	3DMatch FMR ↑	3DMatch IR ↑	3DMatch RR ↑	3DLoMatch FMR ↑	3DLoMatch IR ↑	3DLoMatch RR ↑
FlyingShapes	(a) Structured3D	94.1	49.8	85.5	65.9	18.9	48.6
	(b) 3DFront	87.9	42.7	74.8	48.7	12.3	33.5
	(c) + ShapeNet	93.8	57.0	85.6	69.4	25.2	55.1
	(d) + Delete planes	**95.6**	**59.7**	87.0	72.5	28.5	57.4
	(e) + Structured3D	95.5	59.5	**87.5**	**77.8**	**30.6**	**63.8**
SIRA	(a) baseline	95.5	59.5	87.5	77.8	30.6	63.8
	(b) uniform noise	96.3	62.5	88.6	77.2	30.7	62.7
	(c) Gaussian noise	97.1	63.9	89.3	78.6	32.6	63.8
	(d) TSDF	95.3	60.1	87.9	78.3	31.0	63.9
	(e) w/o ARM	97.3	65.4	90.8	**81.7**	**36.2**	65.3
	(f) w/o multi-scale	97.5	65.3	91.2	80.6	35.8	66.6
	(g) SIRA†	**97.7**	**65.5**	**91.5**	**81.7**	35.8	**67.4**

Table 6. Supplementary results of ablation studies.

Part 1.3: Network Architecture Details

Fig. 1 depicts the detailed network architecture of SIRA. The generator of SIRA is an encoder-decoder structure based on the KPConv backbone [12]. Group normalization and LeakyReLU are applied after each KPConv layer. To better capture the point pattern distributions, the first KPConv layer is applied to the point clouds with 30,000 points, which are randomly sampled rather than voxel downsampled from the original dense point clouds. The l-th strided convolution is applied to a voxel-downsampled point cloud with voxel size $0.025 \cdot 2^{l}$ m. The upsampling module in the generator is performed by searching the corresponding feature of the closest point from the previous layer. ARMs are applied to each upsampling step to provide adaptive point positions. The discriminator is a multi-scale architecture. Encoders based on the same simplified PointNet structure [10], *i.e.*, a shared MLP followed by max-pooling, first extracts the features of multi-scale local patches with 5/10/20 points. The features are then concatenated together, being fed into the subsequent 5-layer shared MLP for classification. For more details, please see our code

Part 1.4: Details in evaluation on ETH

There are many differences, such as the density and the scale, between the point clouds from indoor 3DMatch and outdoor ETH. Thus, it is improper for models totally trained on indoor datasets to be directly evaluated on outdoor benchmarks. For this reason, following [13], we scale the point cloud from ETH to 1/10 of its original size before inference. Note that the point cloud will be rescaled to its original size when computing metrics for a fair comparison.

Part 2: More Visualization and Qualitative Results

For better observation, we provide some visualization results in the video in our supplementary material. Layouts of 10 selected indoor scenes in FlyingShapes are displayed, including depth and segmentation images from different viewpoints. We also provide more qualitative results on 3DMatch and 3DLoMatch benchmarks in our video. The results from GeoTrans [11] and RegTR [15] are shown for comparison. More qualitative results on the ETH benchmark are presented in the video as well.

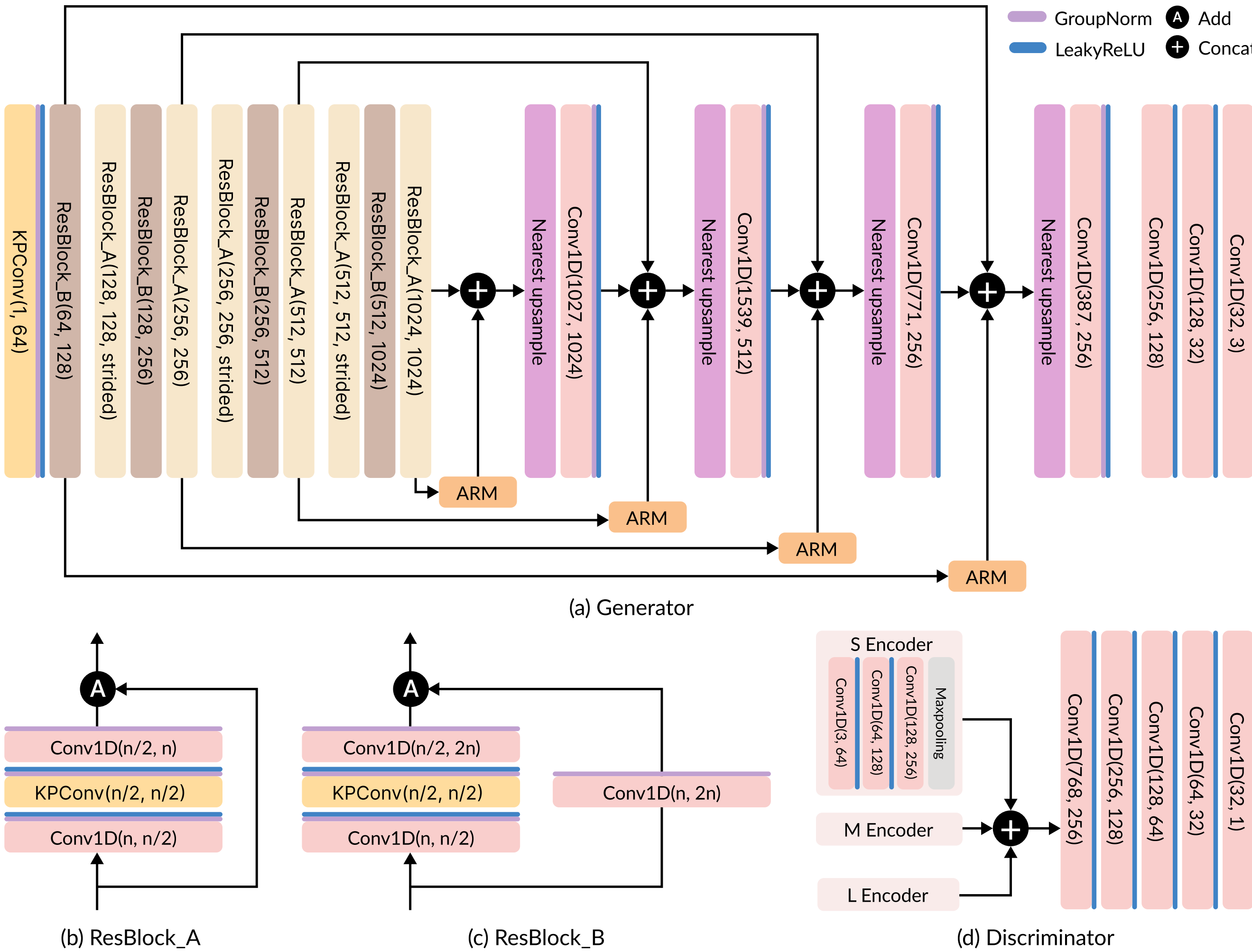

Figure 1. Architecture details of SIRA.

References

[1] Xuyang Bai, Zixin Luo, Lei Zhou, Hongkai Chen, Lei Li, Zeyu Hu, Hongbo Fu, and Chiew-Lan Tai. Pointdsc: Robust point cloud registration using deep spatial consistency. In *Proc. CVPR*, pages 15859–15869, 2021. 3, 5

[2] Xuyang Bai, Zixin Luo, Lei Zhou, Hongbo Fu, Long Quan, and Chiew-Lan Tai. D3feat: Joint learning of dense detection and description of 3d local features. In *Proc. CVPR*, pages 6359–6367, 2020. 3, 4

[3] Christopher Choy, Wei Dong, and Vladlen Koltun. Deep global registration. In *Proc. CVPR*, pages 2514–2523, 2020. 5

[4] Christopher Choy, Jaesik Park, and Vladlen Koltun. Fully convolutional geometric features. In *Proc. ICCV*, pages 8958–8966, 2019. 4

[5] Zan Gojcic, Caifa Zhou, Jan D Wegner, and Andreas Wieser. The perfect match: 3d point cloud matching with smoothed densities. In *Proc. CVPR*, pages 5545–5554, 2019. 4

[6] Sofiane Horache, Jean-Emmanuel Deschaud, and François Goulette. 3d point cloud registration with multi-scale architecture and unsupervised transfer learning. In *2021 International Conference on 3D Vision (3DV)*, pages 1351–1361. IEEE, 2021. 5, 6

[7] Shengyu Huang, Zan Gojcic, Mikhail Usvyatsov, Andreas Wieser, and Konrad Schindler. Predator: Registration of 3d point clouds with low overlap. *arXiv:2011.13005*, 2020. 3, 4, 5

[8] Junha Lee, Seungwook Kim, Minsu Cho, and Jaesik Park. Deep hough voting for robust global registration. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 15994–16003, 2021. 5

[9] Yang Li and Tatsuya Harada. Lepard: Learning partial point cloud matching in rigid and deformable scenes. *Proc. CVPR*, 2022. 4, 5

[10] Charles R. Qi, Hao Su, Kaichun Mo, and Leonidas J Guibas. Pointnet: Deep learning on point sets for 3d classification and segmentation. In *Proc. CVPR*, pages 652–660, 2017. 6

[11] Zheng Qin, Hao Yu, Changjian Wang, Yulan Guo, Yuxing Peng, and Kai Xu. Geometric transformer for fast and robust point cloud registration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 11143–11152, June 2022. 3, 4, 5, 6

[12] Hugues Thomas, Charles R Qi, Jean-Emmanuel Deschaud, Beatriz Marcotegui, François Goulette, and Leonidas J Guibas. Kpconv: Flexible and deformable convolution for point clouds. In *Proc. ICCV*, pages 6411–6420, 2019. 6

[13] Haiping Wang, Yuan Liu, Qingyong Hu, Bing Wang, Jianguo Chen, Zhen Dong, Yulan Guo, Wenping Wang, and Bisheng Yang. Roreg: Pairwise point cloud registration with oriented descriptors and local rotations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023. 3, 5, 6

[14] Fan Yang, Lin Guo, Zhi Chen, and Wenbing Tao. One-inlier is first: Towards efficient position encoding for point cloud registration. In *Advances in Neural Information Processing Systems*. 4, 5

[15] Zi Jian Yew and Gim Hee Lee. Regtr: End-to-end point cloud correspondences with transformers. In *Proc. CVPR*, pages 6677–6686, 2022. 3, 4, 6

[16] Hao Yu, Fu Li, Mahdi Saleh, Benjamin Busam, and Slobodan Ilic. Cofinet: Reliable coarse-to-fine correspondences for robust pointcloud registration. *Proc. NeurIPS*, 34:23872–23884, 2021. 4

[17] Qian-Yi Zhou, Jaesik Park, and Vladlen Koltun. Fast global registration. In *Proc. ECCV*, pages 766–782, 2016. 5