Supplementary Material for
Local Connectivity-Based Density Estimation for Face Clustering

Junho Shin, Hyo-Jun Lee, Hyunseop Kim, Jong-Hyeon Baek, Daehyun Kim, Yeong Jun Koh*
Chungnam National University
{wnsgh8190, gywns6287, hyunseop95, whdgusdl97, seven776484}@gmail.com, yjkoh@cnu.ac.kr

S-1. Supplemental of Ablation Study

In the main paper, we provided partial ablation study results due to the page limit. In this supplementary material, we additionally provide ablation study results in Tables S-1∼ S-6. Tables S-1 and S-2 provide ablation study results on F_{512} and F_{1024} in IJB-B, which are omitted in Tables 4 and 6 in the main paper. Tables S-1 and S-2 show that the proposed local connectivity-based method and PCENet, respectively, are effective for clustering. Tables S-3∼ S-5 supplement Table 5 in the main paper, which analyzes the proposed edge selection strategies. Tables S-3∼ S-5 show ablation study results of edge selection strategies on remaining datasets not in Table 5. They demonstrate that the proposed edge selection method significantly improves recall scores by increasing positive pairs (P.P) while reducing negative pairs (N.P).

Table S-1. Ablation study on IJB-B for the local connectivity.

Datasets	IJB-B F_{512}		IJB-B F_{1024}	
Methods/ Metrics	F_P	F_B	F_P	F_B
Similarity (Sim.)	92.50	83.96	92.24	83.98
Local Connectivity (LC)	92.70	85.04	91.17	84.33
Sim. + LC	**93.05**	**85.09**	**92.66**	**85.15**

Table S-2. Comparison of PCENet with LCENet on IJB-B.

Datasets	IJB-B F_{512}		IJB-B F_{1024}	
Methods/ Metrics	F_P	F_B	F_P	F_B
LCENet	89.09	80.31	85.21	79.64
PCENet	**93.05**	**85.09**	**92.66**	**85.15**

Table S-3. Ablation study on MS-Celeb-1M (1.74M, 2.89M) according to edge selection strategies.

	Sim.	LC	$\tilde{\mathcal{E}}_d$	$\tilde{\mathcal{E}}_s$	MS-Celeb-1M (1.74M) BCubed Precision	BCubed Recall	F_B	P.P.	N.P.	$\frac{\text{P.P.}}{\text{P.P.+N.P.}}$	MS-Celeb-1M (2.89M) BCubed Precision	BCubed Recall	F_B	P.P.	N.P.	$\frac{\text{P.P.}}{\text{P.P.+N.P.}}$
S_1	✓		✓		96.86	81.36	88.44	1,626,026	105,452	93.91%	96.49	78.64	86.66	2,677,760	198,549	93.10%
S_2	✓	✓	✓		95.65	85.75	90.43	1,643,394	86,820	94.98%	95.16	83.39	88.88	2,711,091	163,001	94.33%
S_3	✓	✓	✓	✓	95.63	86.39	90.78	4,464,245	94,835	97.92%	95.13	84.11	89.28	7,333,456	189,103	97.49%

Table S-4. Ablation study on MS-Celeb-1M (4.05M, 5.21M) according to edge selection strategies.

	Sim.	LC	$\tilde{\mathcal{E}}_d$	$\tilde{\mathcal{E}}_s$	MS-Celeb-1M (4.05M) BCubed Precision	BCubed Recall	F_B	P.P.	N.P.	$\frac{\text{P.P.}}{\text{P.P.+N.P.}}$	MS-Celeb-1M (5.21M) BCubed Precision	BCubed Recall	F_B	P.P.	N.P.	$\frac{\text{P.P.}}{\text{P.P.+N.P.}}$
S_1	✓		✓		96.13	76.66	85.30	3,721,871	302,085	92.49%	95.81	75.20	84.26	4,768,518	413,872	92.01%
S_2	✓	✓	✓		94.76	81.65	87.72	3,775,715	248,040	93.84%	94.44	80.35	86.83	4,837,244	340,691	93.42%
S_3	✓	✓	✓	✓	94.72	82.43	88.15	10,152,648	296,686	97.16%	94.40	81.16	87.28	12,916,446	419,347	96.86%

Table S-5. Ablation study on IJB-B (F_{512}, F_{1024}) according to edge selection strategies.

	Sim.	LC	$\tilde{\mathcal{E}}_d$	$\tilde{\mathcal{E}}_s$	IJB-B (F_{512}) BCubed Precision	BCubed Recall	F_B	P.P.	N.P.	$\frac{\text{P.P.}}{\text{P.P.+N.P.}}$	IJB-B (F_{1024}) BCubed Precision	BCubed Recall	F_B	P.P.	N.P.	$\frac{\text{P.P.}}{\text{P.P.+N.P.}}$
S_1	✓		✓		96.63	74.34	84.03	15,527	2,615	85.59%	96.24	72.71	82.84	30,928	5,107	85.83%
S_2	✓	✓	✓		96.37	74.25	83.88	15,662	2,481	86.33%	95.99	74.66	83.99	31,217	4,816	86.63%
S_3	✓	✓	✓	✓	96.37	76.18	85.09	144,985	2,982	97.98%	95.99	75.51	85.15	257,186	6,835	97.41%

Table S-6. Ablation study on DeepFashion according to edge selection strategies.

	Sim.	LC	$\tilde{\mathcal{E}}_d$	$\tilde{\mathcal{E}}_s$	DeepFashion BCubed Precision	BCubed Recall	F_B	P.P.	N.P.	$\frac{\text{P.P.}}{\text{P.P.+N.P.}}$
S_1	✓		✓		83.12	52.21	64.13	17,451	8,458	67.35%
S_2	✓	✓	✓		83.48	52.16	64.20	17,688	7,743	69.55%
S_3	✓	✓	✓	✓	82.75	52.93	64.56	22,325	7,985	73.66%

S-2. How performance depends on K

The proposed method has hyperparameter K to construct KNN graph. Tables S-7∼ S-9 list clustering performance according to K on MS-Celeb-1M, IJB-B, and DeepFashion respectively. We observe that stable performance is achieved when K is large enough. These results indicate that the proposed method does not depend on K.

For MS-Celeb-1M, we pick $K = 80$ as done in most existing face clustering methods. In contrast, diverse hyperparameters are used in L-GCN [24] ($K = 200$), DANet [6] ($K = 256$), and Pairwise [12] ($K = 40$). Therefore, we select $K = 120$, which produces the best performance on IJB-B dataset. For DeepFashion, we use $K = 8$ to maintain the consistent model of eight multi-heads in LCENet and PCENet regardless of datasets, even if $K = 10$ produces the best result.

Table S-7. Clustering results on MS-Celeb-1M according to K. The results of the proposed method ($K = 80$) are boldfaced.

Datasets	MS-Celeb-1M 584K		1.74M		2.89M		4.05M		5.21M	
K/ Metrics	F_P	F_B	F_P	F_B	F_P	F_B	F_P	F_B	F_P	F_B
$K = 64$	94.33	93.25	91.21	90.40	89.35	88.76	87.77	87.55	86.21	86.60
$K = 72$	93.97	93.07	91.07	90.49	89.08	88.99	87.44	87.89	86.08	87.03
$K = 80$	**94.64**	**93.36**	**91.90**	**90.78**	**90.27**	**89.28**	**88.69**	**88.15**	**87.35**	**87.28**
$K = 88$	94.43	93.29	91.76	90.59	89.98	89.05	88.52	87.82	87.19	86.87
$K = 96$	93.97	93.06	90.77	90.31	88.74	88.75	86.74	87.54	84.98	86.59

Table S-8. Clustering results on IJB-B according to K. The results of the proposed method ($K = 120$) are boldfaced.

Datasets	IJB-B F_{512}		F_{1024}		F_{1845}	
K/ Metrics	F_P	F_B	F_P	F_B	F_P	F_B
$K = 80$	92.71	84.88	91.42	84.60	88.80	83.94
$K = 104$	92.71	84.91	90.79	84.54	85.06	83.99
$K = 112$	92.65	85.06	90.98	84.70	84.83	84.04
$K = 120$	**93.05**	**85.09**	**92.66**	**85.15**	**90.78**	**84.81**
$K = 128$	93.09	85.10	92.40	84.91	87.53	84.46

Table S-9. Clustering results on DeepFashion according to K. The results of the proposed method ($K = 8$) are boldfaced.

Datasets	DeepFahsion	
K/ Metrics	F_P	F_B
$K = 5$	39.79	64.47
$K = 8$	**41.76**	**64.56**
$K = 10$	42.30	65.08
$K = 16$	42.84	63.85

S-3. The connecting threshold τ

The proposed method computes the connecting threshold τ to construct $\tilde{\mathcal{E}}_s$ by averaging similarities between each node and its 3-nearest neighbors for each dataset as in Table S-10. In Tables S-11 and S-12, we compare clustering performance according to the number of nearest neighbors to compute the connecting threshold. We observe that stable performance is obtained regardless of the number of nearest neighbors, which indicates that the proposed method does not depend on τ. We pick 3-nearest neighbors, which provide reliable performance.

Table S-10. τ according to datasets.

Datasets	MS-Celeb-1M					IJB-B			DeepFashion
	584K	1.74M	2.89M	4.05M	5.21M	F_{512}	F_{1024}	F_{1845}	-
τ (3NN)	0.8339	0.8347	0.8358	0.8363	0.8367	0.7498	0.7464	0.7550	0.8840

Table S-11. Clustering results on MS-Celeb-1M according to the number of nearest neighbors to compute τ.

Datasets	MS-Celeb-1M 584K		1.74M		2.89M		4.05M		5.21M	
Neighbors/ Metrics	F_P	F_B	F_P	F_B	F_P	F_B	F_P	F_B	F_P	F_B
1-neighbors	94.47	93.26	91.84	90.69	90.21	89.18	88.76	88.04	87.31	87.16
2-neighbors	94.52	93.31	91.88	90.74	90.26	89.24	88.70	88.11	87.31	87.23
3-neighbors	94.64	93.36	91.90	90.78	90.27	89.28	88.69	88.15	87.35	87.28
4-neighbors	94.71	93.39	91.95	90.80	90.28	89.32	88.67	88.18	87.21	87.32
5-neighbors	94.71	93.40	91.95	90.82	90.31	89.35	88.66	88.22	87.17	87.35

Table S-12. Clustering results on IJB-B and DeepFashion according to the number of nearest neighbors to compute τ.

Datasets	IJB-B F_{512}		F_{1024}		F_{1845}		DeepFashion -	
Neighbors/ Metrics	F_P	F_B	F_P	F_B	F_P	F_B	F_P	F_B
1-neighbors	93.05	85.09	92.66	85.15	81.01	83.64	41.59	64.45
2-neighbors	93.05	85.09	92.66	85.15	84.88	84.19	41.54	64.51
3-neighbors	93.05	85.09	92.66	85.15	90.78	84.81	41.76	64.56
4-neighbors	93.05	85.09	92.66	85.15	90.78	84.81	41.59	64.61
5-neighbors	93.05	85.09	92.66	85.15	90.78	84.81	41.54	64.61

S-4. Additional experiments

Ratios of true positive connection: Table S-13 shows ratios of true positive connection according to the procedures: KNN graph, LCENet for $\tilde{\mathcal{E}}_d$ PCENet, and Graph edge selection for $\tilde{\mathcal{E}}_s$.

Table S-13. True positive connection ratio according to the procedures.

	MS-Celeb-1M (584K) P.P.	N.P.	$\frac{\text{P.P.}}{\text{P.P.+N.P.}}$
KNN graph	32,930,669	13,790,371	70.48%
LCENet for $\tilde{\mathcal{E}}_d$	558,304	22,032	96.20%
LCENet for $\tilde{\mathcal{E}}_d$ + PCENet	554,271	6,383	98.86%
LCENet for $\tilde{\mathcal{E}}_d$ + PCENet + $\tilde{\mathcal{E}}_s$	1,525,634	7,198	99.53%

Clustering with PCENet: There are two options for clustering with PCENet only. The first is S_1 in Table 5 in main paper. The second is to estimate all local connectivity KN pairs for each node using PCENet. We refer this option to S_4 as in Table S-14.

Table S-14. Clustering with PCENet only.

	MS-Celeb-1M (584K)		IJB-B (F_{1845})		DeepFashion	
	F_P	F_B	F_P	F_B	F_P	F_B
S_4	94.12	93.10	90.44	84.82	41.23	64.94

Binary classification: Table S-15 shows binary classification scores of PCENet. It provides reliable F_1 scores on MS-Celeb-1M and IJB-B, while providing relatively low scores on DeepFashion.

Table S-15. Binary classification performance.

	MS-Celeb-1M (584K)			IJB-B (F_{1845})			DeepFashion		
	Precision	Recall	F_1	Precision	Recall	F_1	Precision	Recall	F_1
PCENet	98.86	99.28	99.07	97.37	96.70	97.03	86.18	87.99	87.08

BCubed scores for PCENet: As in Table S-16, PCENet reduces the BCubed recall, but it extremely improves the BCubed precision, resulting in the good F scores.

Table S-16. Clustering performance according to PCENet.

	MS-Celeb-1M (584K)			IJB-B (F_{1845})			DeepFashion		
	Pre	Rec	F_B	Pre	Rec	F_B	Pre	Rec	F_B
w/o PCENet	45.33	92.58	60.86	17.66	81.01	29.00	29.42	66.78	40.84
PCENet	96.65	90.28	93.36	95.78	76.10	84.81	82.75	52.93	64.56