On Mitigating Hard Clusters for Face Clustering

Yingjie Chen[*,1], Huasong Zhong[*,2], Chong Chen[†,2], Chen Shen[2], Jianqiang Huang[2], Tao Wang[1], Yun Liang[1], Qianru Sun[3]

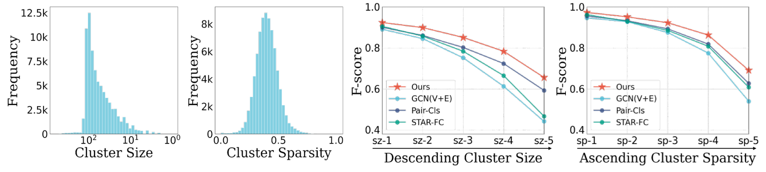
[1]Peking University, [2]DAMO Academy, Alibaba, [3]Singapore Management University

ALIBABA DAMO ACADEMY

TEL AVIV 2022

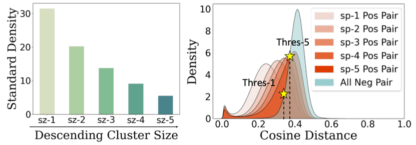
Motivation

- Face clustering is challenging due to that 1) recognizing person identities is a fine-grained task; 2) the number of identities is always large; 3) the derived face clusters are often of high variations in both size and sparsity, and small or sparse clusters – we call **hard clusters** – are hard to identify.



(a) (b) (c) (d)

- We find that both density and distance are highly influenced by the size and sparsity of latent clusters in face data.
- For example, 1) smaller clusters tend to have lower density, so they could be misclassified as big ones by DPC.
- 2) to identify positive pairs, higher-sparsity clusters prefer a higher distance threshold, so it is hard to determine a uniform threshold for DPC.



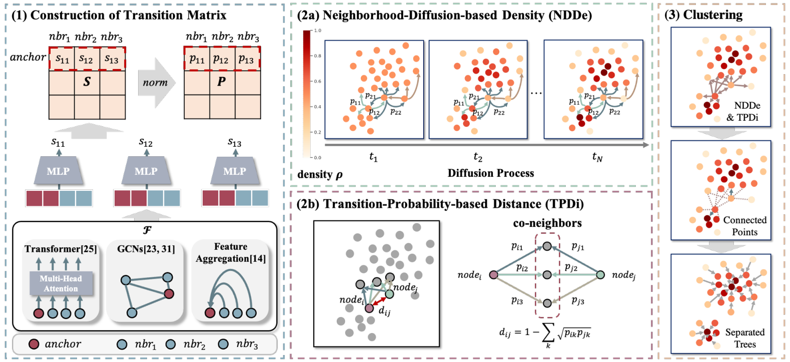
(a) (b)

Contributions

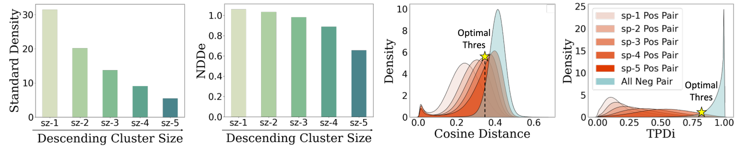
- 1) We inspect face clustering problem and find existing methods failed to identify hard clusters---yielding significantly low recall for small or sparse clusters.
- 2) To mitigate the issue of small clusters, we introduce NDDe based on the diffusion of neighborhood densities.
- 3) To mitigate the issue of sparse clusters, we propose the relative distance TPDi that can facilitate a uniform sparsity in different clusters.
- In experiments, we evaluate NDDe and TPDi on large-scale benchmarks and incorporate them into multiple baselines to show their efficiency.

Methodology

- **Overview**



- **Size-invariant density NDDe & sparsity-aware distance TPDi**



(a) (b) (c) (d)

Experimental Results

- **Results on MS1M & DeepFashion**

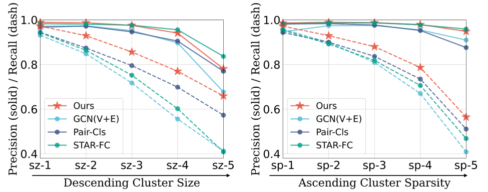
Method	#Clusters	F_P	F_B	Time
K-means [18]	3991	32.86	53.77	573s
HAC [24]	17410	22.54	48.7	112s
DBSCAN [6]	14350	25.07	53.23	2.2s
ARO [20]	10504	26.03	53.01	6.7s
CDP [33]	6622	28.28	57.83	1.3s
L-GCN [29]	10137	28.85	58.91	23.3s
LTC [32]	9246	29.14	59.11	13.1s
GCN(V+E) [31]	6079	38.47	60.06	18.5s
Pair-Cls [14]	6018	37.67	62.17	0.6s
STAR-FC [23]	-	37.07	60.60	-
Ada-NETS [28]	-	39.30	61.05	-
Ours	8484	40.91 +1.61	63.61 +1.44	4.2s

#Images	584K		1.74M		2.89M		4.05M		5.21M	
Method / Metrics	F_P	F_B	F_P	F_B	F_P	F_B	F_P	F_B	F_P	F_B
K-means [18]	79.21	81.23	73.04	75.20	69.83	72.34	67.90	70.57	66.47	69.42
HAC [24]	70.63	70.46	54.40	69.53	11.08	68.62	1.40	67.69	0.37	66.96
DBSCAN [6]	67.93	67.17	63.41	66.53	52.50	66.26	45.24	44.87	44.94	44.74
ARO [20]	13.60	17.00	8.78	12.42	7.30	10.96	6.86	10.50	6.35	10.01
CDP [33]	75.02	78.70	70.75	75.82	69.51	74.58	68.62	73.62	68.06	72.92
L-GCN [29]	78.68	84.37	75.83	81.61	74.29	80.11	73.70	79.33	72.99	78.60
LTC [32]	85.66	85.52	82.41	83.01	80.32	81.10	78.98	79.84	77.87	78.86
GCN(V+E) [31]	87.93	86.09	84.04	82.84	82.10	81.24	80.45	80.09	79.30	79.25
Clusformer [19]	88.20	87.17	84.60	84.05	82.79	82.30	81.03	80.51	79.91	79.95
Pair-Cls [14]	90.67	89.54	86.91	86.25	85.06	84.55	83.51	83.49	82.41	82.40
STAR-FC [23]	91.97	90.21	88.28	86.26	86.17	84.13	84.70	82.63	83.46	81.47
Ada-NETS [28]	92.79	91.40	89.33	87.98	87.50	86.03	85.40	84.48	83.99	83.28
GCN(V+E)++	90.72 +2.79	89.28 +3.19	86.06 +2.02	84.36 +1.52	85.97 +3.87	84.24 +3.00	84.76 +4.31	83.10 +3.01	83.69 +4.39	82.26 +3.01
Pair-Cls++	91.70 +1.03	89.94 +0.40	88.17 +1.26	86.50 +0.25	86.49 +1.43	84.76 +0.21	85.25 +1.74	83.50 +0.01	83.74 +1.33	82.61 +0.21
STAR-FC++	92.35 +0.38	90.50 +0.29	89.03 +0.75	86.94 +0.68	86.70 +0.53	85.16 +1.03	85.38 +0.68	83.93 +1.30	83.94 +0.48	82.95 +1.48
Ours	**93.22**	**92.18**	**90.51**	**89.43**	**89.09**	**88.00**	**87.93**	**86.92**	**86.94**	**86.06**

- **Ablation Study**

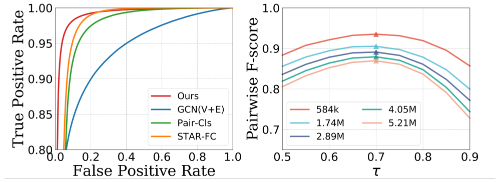
	NDDe	TPDi	584K		1.74M		2.89M		4.05M		5.21M	
			F_P	F_B	F_P	F_B	F_P	F_B	F_P	F_B	F_P	F_B
M_1			53.03	56.75	47.80	53.84	45.07	52.41	43.29	51.56	41.49	50.76
M_2		✓	61.07	59.81	59.29	58.26	58.66	57.40	58.37	57.00	57.88	56.48
M_3	✓		82.98	80.33	78.79	77.87	76.32	76.42	74.08	75.28	72.47	74.39
M_4	✓	✓	**93.22**	**92.18**	**90.51**	**89.43**	**89.09**	**88.00**	**87.93**	**86.92**	**86.94**	**86.06**

- **Results on Hard Clusters**

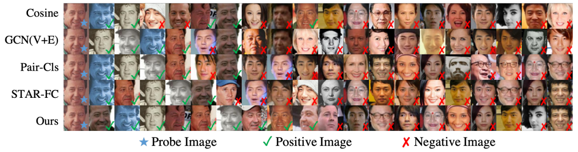


(a) (b)

- **The Superiority of TPDi**



(a) (b)



- **Face Recognition**
- To further show the potential of our method in scaling up face recognition systems, we use our method to generate pseudo-labels for unlabeled face images to train face recognition models.