Dispersion based Clustering for Unsupervised Person Re-identification

Guodong Ding[1*], Salman Khan[2], and Zhenmin Tang[1]
[1]Nanjing University of Science and Technology, [2]Australian National University
* Corresponding author, guodong.ding@njust.edu.cn

Problem Definition and Contribution

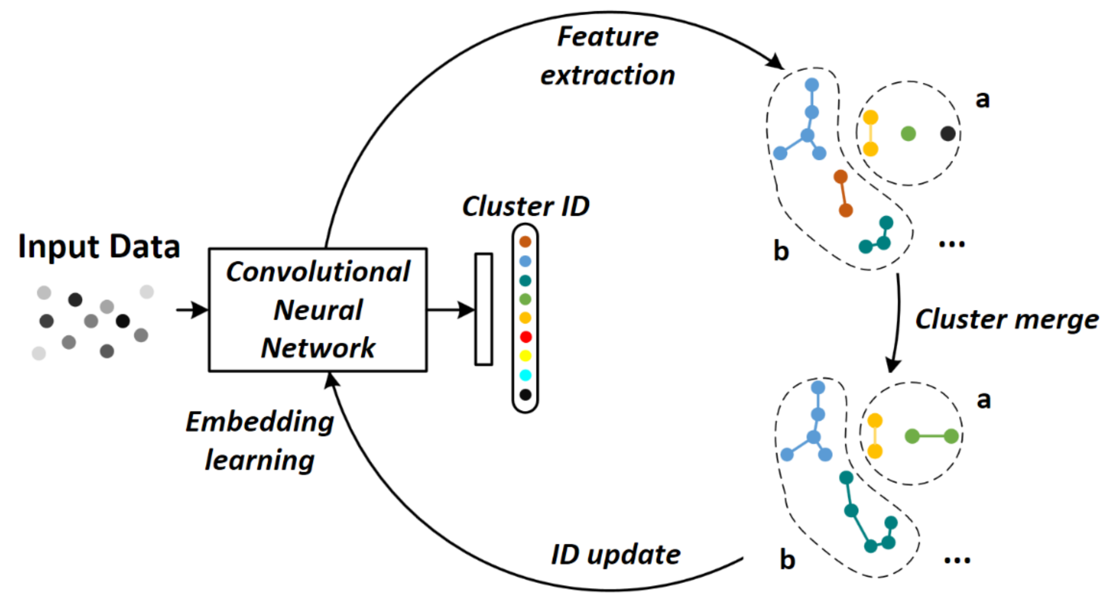
Goal: To perform unsupervised person re-identification with a robust criterion in an agglomerative clustering framework.

Contributions: An unsupervised deep learning framework that

- proposes a dispersion based clustering criterion which considers both within cluster compactness and between clusters separation.
- has two major advantages manifested by the criterion, *i.e.*, automatic prioritization of isolated data points for merging and prevention of poor clustering.
- demonstrates superior performances over the SOTA methods on both image-based and video-based person re-ID datasets.

Approach

Learning Framework: Starting from sample specificity learning, *i.e.*, each sample as a class, alternatively train feature extractor and update clustering results for subsequent training.



Embedding Learning

Repelled loss: Given an input image x, a target class y , we compute its feature representation v, we calculate its class distribution as:

$$p(y|x,V) = \frac{exp(V_y^T v/\tau)}{\Sigma_{j=1}^{N} exp\left(V_j^T y/\tau\right)}, \quad (1)$$

where τ is a temperature parameter that controls the softness of probability distribution over classes, V is a lookup table (LUT) containing the centroid feature of each class. After which a cross entropy (CE) loss is imposed to align $p(y|x,V)$ with the clustering resulting labels.

Advantages: this form jointly considers inter-class and intra-class variances.

Merging Criterion (DBC)

Dispersions: Given a cluster $\mathcal{C}$ scattered in feature space, we define its dispersion $d(\mathcal{C})$ as:

$$d(\mathcal{C}) = \frac{1}{n}\Sigma_{i,j\in\mathcal{C}}\, dist(\mathcal{C}_i, \mathcal{C}_j), \quad (2)$$

where n is the cardinality of $\mathcal{C}$. As such, dispersion between clusters is written as:

$$d(\mathcal{C}_a, \mathcal{C}_b) = \frac{1}{n_a n_b}\Sigma_{i\in\mathcal{C}_a, j\in\mathcal{C}_b}\, dist(\mathcal{C}_{a_i}, \mathcal{C}_{b_j}), \quad (3)$$

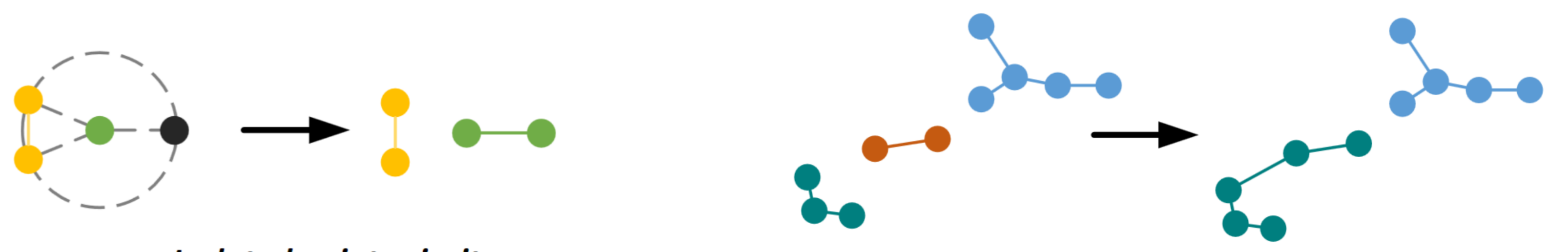
Thus, the **dispersion based merging criterion** is a combine of both above dispersion terms as follows:

$$\mathcal{D}_{ab} = d_{ab} + \lambda(d_a + d_b) \quad (4)$$

where λ is a trade off parameter between two components.

Discussion

On one hand: Proposed criterion (Eq. (4)) brings two major advantages to the merging process, *i.e.*, isolated point priority and poor clustering prevention.



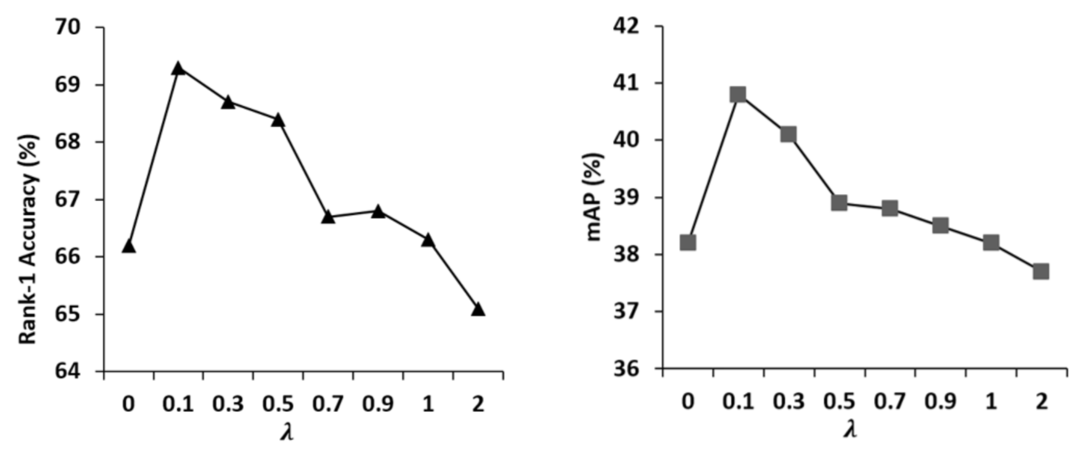
a. Isolated point priority — *b. Poor clustering prevention*

- Isolated point has zero within cluster dispersion, thus can be prioritized in first few merging stages.
- Poor cluster has large within cluster dispersion which can defer itself from merging.

On the other hand: Proposed criterion (Eq. (4)) ensures the forming of compact and well-separated cluster results, serves the same purpose as the repelled loss, eventually speeds up learning in a reciprocal way.

Experiments & Results

Tradeoff Parameter

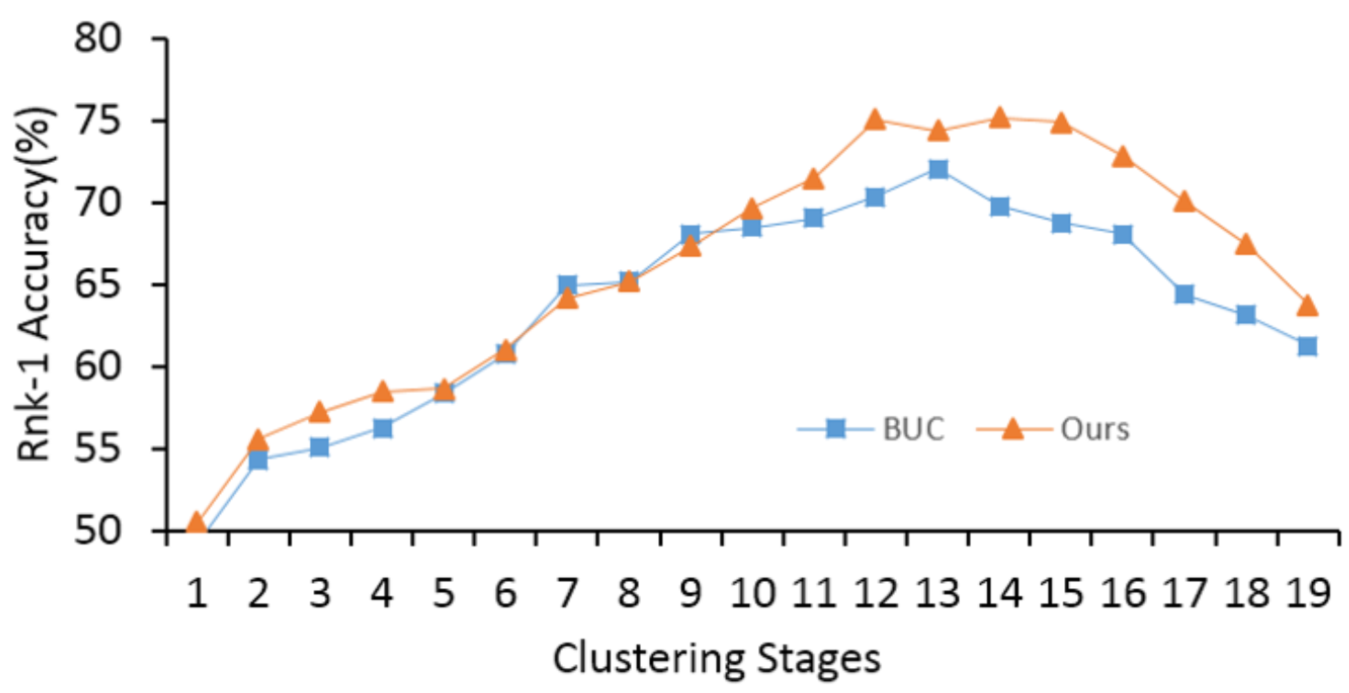


Ablation Study

Methods	Market-1501		DukeMTMC-reID		MARS		DukeMTMC-VideoReID	
	rank-1	mAP	rank-1	mAP	rank-1	mAP	rank-1	mAP
BUC⁻[15]	62.9	33.8	41.3	22.5	55.5	31.9	60.7	50.8
BUC [15]	66.2	38.3	47.4	27.5	61.1	38.0	69.2	61.9
DBC⁻	66.2	38.7	48.2	27.5	59.8	37.2	71.8	63.2
DBC	**69.2**	**41.3**	**51.5**	**30.0**	**64.3**	**43.8**	**75.2**	**66.1**

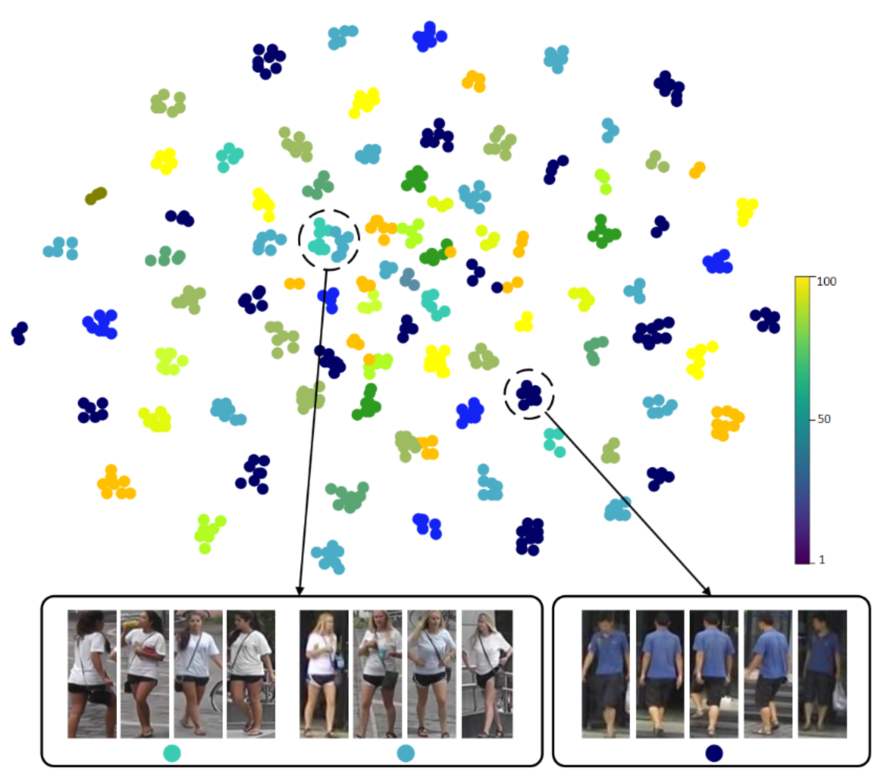
⁻ denotes removal of regularization term, *i.e.*, *cardinality* in BUC and *intra-cluster* term (Eq. (3)) in BUC.

Learning Speed & Robustness**



- DBC has faster learning speed over BUC on DukeMTMC-VideoReID dataset and enjoys a better performance robustness against varying target cluster numbers.

Qualitative Analysis**



T-SNE visualization of the clustering results on a reduced Market-1501 subset.

- In most cases, samples from the same identity are group together (see collections of same colored points).
- Some incorrect merging of identities are also present. For example, bottom left box where two ladies with similar appearance (white t-shirts and dark sports wear) are clustered together.

Image-based Datasets

Methods	Labels	Market-1501				DukeMTMC-reID			
		rank-1	rank-5	rank-10	mAP	rank-1	rank-5	rank-10	mAP
BOW[42]	**None**	35.8	52.4	60.3	14.8	17.1	28.8	34.9	8.3
OIM[36]	**None**	38.0	58.0	66.3	14.0	24.5	38.8	46.0	11.3
UMDL[22]	Transfer	34.5	52.6	59.6	12.4	18.5	31.4	37.6	7.3
PUL[5]	Transfer	44.7	59.1	65.6	20.1	30.4	46.4	50.7	16.4
EUG[35]	OneEx	49.8	66.4	72.7	22.5	45.2	59.2	63.4	24.5
SPGAN[3]	Transfer	58.1	76.0	82.7	26.7	46.9	62.6	**68.5**	26.4
TJ-AIDL[33]	Transfer	58.2	-	-	26.5	44.3	-	-	23.0
BUC[15]	**None**	**66.2**	**79.6**	**84.5**	**38.3**	**47.4**	**62.6**	68.4	**27.5**
DBC	**None**	**69.2**	**83.0**	**87.8**	**41.3**	**51.5**	**64.6**	**70.1**	**30.0**

- "Transfer“: External dataset with annotations used.
- "OneEx“: One labeled example per person is used.
- "Camera ": Camera view information is used.
- "None“: No extra information used.

Video-based Datasets

Methods	Labels	MARS				DukeMTMC-VideoReID			
		rank-1	rank-5	rank-10	mAP	rank-1	rank-5	rank-10	mAP
OIM[42]	**None**	33.7	48.1	54.8	13.5	51.1	70.5	76.2	43.8
DGM+IDE[37]	OneEx	36.8	54.0	-	16.8	42.3	57.9	69.3	33.6
Stepwise[17]	OneEx	41.2	55.5	-	19.6	56.2	70.3	79.2	46.7
RACE[38]	OneEx	43.2	57.1	62.1	24.5	-	-	-	-
DAL[1]	Camera	49.3	65.9	72.2	23.0	-	-	-	-
BUC[15]	**None**	61.1	**75.1**	**80.0**	38.0	69.2	81.1	**85.8**	61.9
EUG[35]	OneEx	**62.6**	74.9	-	**42.4**	**72.7**	**84.1**	-	**63.2**
DBC	**None**	**64.3**	**79.2**	**85.1**	**43.8**	**75.2**	**87.0**	**90.2**	**66.1**

** Details of additional information in this poster can be found via our journal version “Towards better Validity: Dispersion based Clustering for unsupervised Person Re-identification” at https://arxiv.org/pdf/1906.01308.pdf