Unsupervised Person Re-identification via Mining Label Homogeneity

1st Qing Tang, 2nd Kang-Hyun Jo
Department of Electrical, Electronic and Computer Engineering
University of Ulsan
Ulsan, Korea
tangqing@islab.ulsan.ac.kr; acejo@ulsan.ac.kr

***Abstract*—This paper studies the fully unsupervised person reidentification (Re-ID) problem which does not need any labeled data to train the re-ID network. To make fully unsupervised training possible, the model needs to generate pseudo labels. Unlike human-labeled annotation, the generated pseudo labels contain noisy labels. One of the best fully unsupervised methods till now formulated person Re-ID as a multi-label classification task. This approach only considered similarities between images and ignore relations among the generated labels. In this paper, we follow the multi-label classification training strategy but utilize the generated labels as auxiliary labels to further refine the noisy labels. The proposed approach seeks auxiliary labels by investigating two underlying homogeneity in generated labels, i.e. Symmetric Homogeneity (S) and Neighbor Homogeneity (N). The network is optimized under the joint supervision of the main pseudo label and two auxiliary labels in our proposed mutual label learning training strategy. The experimental results confirm the effectiveness of our discovered label homogeneities and the proposed mutual label learning on two mainstream datasets, Market-1501 and DukeMTMC-reID.**

***Index Terms*—Person re-identification, fully unsupervised learning, pseudo label prediction**

I. Introduction

The person re-identification (re-ID) system aims to retrieve images that contain the same person. Thanks to the development of Convolutional Neural Network (CNN), supervised person re-ID [1]–[3] achieved significant progress by training with substantial labeled data. However, it is expensive and time-consuming to annotate identities across multiple cameras. Moreover, the performance of a re-ID model trained on an existing large-scale dataset will significantly decline when it is directly used on real-world systems because of domain gaps. Therefore, some recent works focus on the unsupervised person re-ID methods, which can be trained without using any human-annotated labels.

To make unsupervised training possible, the model needs to self-generate pseudo labels by clustering algorithm [4]–[6] or exemplar memory-based algorithm [7], [8]. Unlike human-annotated labels used in supervised learning, the self-generated pseudo labels contain noisy labels that substantially hinder the model's capability because the network learns to extract discriminative features based on these pseudo labels. Therefore, unsupervised person re-ID performance still significantly falls behind the supervised person re-ID.

Consequently, the key to improving the unsupervised person re-ID model performance is to generate high-quality pseudo labels which can represent the unlabeled data domain distribution. Several studies [5]–[8], [11]–[16] utilized unsupervised domain adaption (UDA) which transferred knowledge from a labeled source dataset to an unlabeled target dataset by transfer learning, thereby learning better image representation. The key of UDA-based methods is to reduce the domain gap between the source and target dataset [11], [12], [21]. ECN [8] discovered that the intra-domain information in the target dataset is also important; thus, ECN introduced an feature memory to enforce learning over the whole target dataset. Although the UDA-based methods have achieved some improvements, UDA-based methods still need a labeled source dataset. Inspired by the exemplar memory in ECN [8], a fully unsupervised method MMCL [7] formulated unsupervised person Re-ID as a multi-label classification task and achieved good re-ID performance. MMCL did not require any manually labeled dataset and learn re-identification information only from unlabeled images, thus it is easily deployed to a new scenario. ECN [8] and MMCL [7] addressed the person re-ID problem only consider the similarities among all images to predict pseudo labels, but ignore the underlying relations among generated pseudo labels. Based on MMCL [7], Tang et al. [9] leveraged neighbors information as additional reference information to ease the impact of noisy pseudo labels. Subsequently, Multiple pseudo Labels Joint Training (MLJT) [10] improved the quality pseduo labels by predicting multiple pseudo labels for each image.

In fact, after the networks can roughly capture the training data distribution and predict pseudo labels, the predicted pseudo labels also can be served as auxiliary labels. In this study, we mine the relations among generated pseudo labels and use them as additional supervisions for an unlabeled image to further refine noisy labels. We construct an Pseudo Label Memory (PLM) for storing the up-to-date generated pseudo labels $\hat{Y} = \{y_1, y_2, ..., y_n\}$ of all images in $X = \{x_1, x_2, ..., x_n\}$. With the help of PLM, we inquire two auxiliary labels based on two discovered underlying relations, i.e, Symmetric Homogeneity (S) and Neighbor Homogeneity (N), thereby refining the main label. Figure 1 shows our re-ID framework SNNet which optimizes the network $\mathcal{F}(\cdot)$ under the joint supervisions of main pseudo label y_i and two

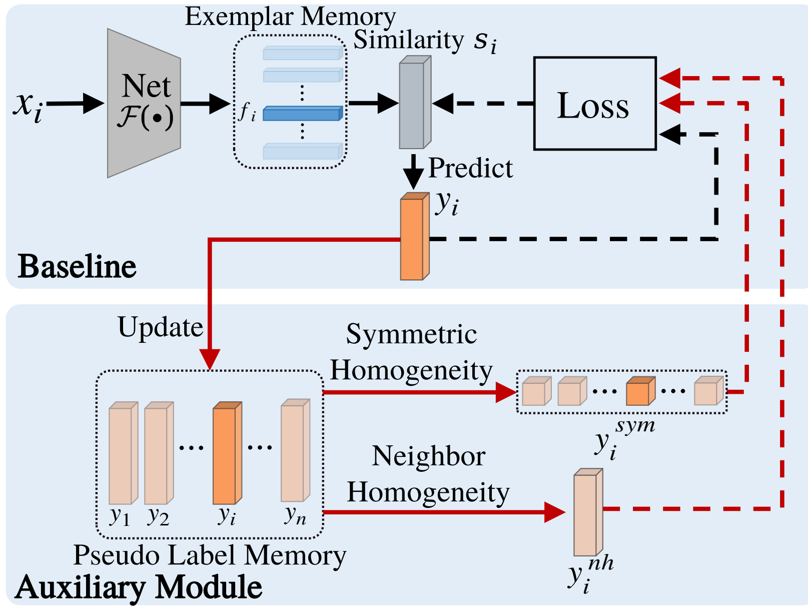

Fig. 1. The framework of our proposed SNNet. The black line indicates the baseline network. Apart from the predicted main pseudo label y_i, we proposed an auxiliary module to seek auxiliary labels y_i^{sym} and y_i^{nh} as additional supervisions to optimize the network.

auxiliary pseudo labels y_i^{sym} and y_i^{nh}. Different from training several student networks collaboratively and mutually in deep mutual learning [17], our proposed SNNet framework trains one single network via labels of different samples, thereby training different samples mutually.

Based on the above aspects, we propose a fully unsupervised person re-ID training strategy. The contributions could be summarized as three-fold. (1) We discovered two underlying label homogeneity constraints in the exemplar memory-based person re-ID method. (2) To make the network learn two label homogeneities possible, we design an unsupervised mutual label learning framework, which optimizes the network under the joint supervisions of main pseudo labels and auxiliary labels. (3) The proposed fully unsupervised person re-ID framework SNNet shows exceptionally strong performances.

II. Proposed Method

Our proposed re-ID framework is shown in Figure 1. The model is mainly divided into two parts: the baseline network and the proposed auxiliary module. The baseline network is the general exemplar memory-based re-ID network [7]. Given a unlabeled dataset $X = \{x_1, x_2, ..., x_n\}$, the baseline network computes the similarity score s_i to seek main pseudo label y_i for each image $x_i \in X$. Unlike the human-annotated label, generated main pseudo label y_i contain the noisy. To mitigate the effects of noisy pseudo labels, we propose an auxiliary module that seeks two auxiliary pseudo labels y_i^{sym} and y_i^{nh} to optimize the neural network with y_i jointly.

A. The Baseline Network

Given a set of unlabeled person images $\{x_1, x_2, ..., x_n\} \in X$, the d-dimensional feature f_i of x_i are extracted by back-

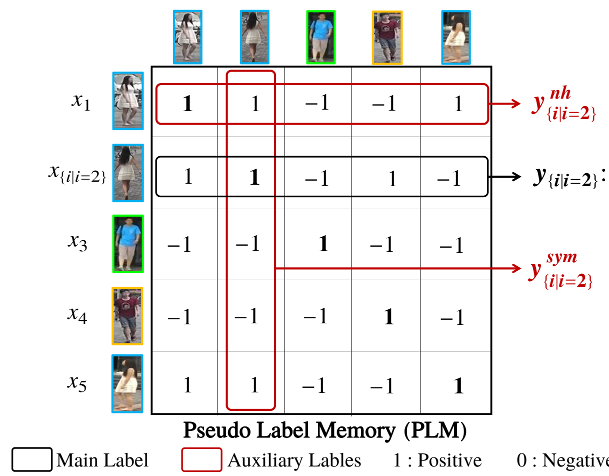

Fig. 2. The illustration of pseudo label memory (PLM). $n = 5$ is assumed in this figure. Ideally, $y_i = y_i^{sym}$ because of symmetric homogeneity constrain, and $y_i = y_i^{nh}$ because of neighbor homogeneity constrain.

bone network $\mathcal{F}(\cdot)$ to form the exemplar memory $\mathcal{M}$. n is the number of images in X. The size of $\mathcal{M}$ is $n \times d$. Using $\mathcal{M}$, the cosine similarity between x_i and the other image $x_j \in X$ can be computed as,

$$s_i[j] = f_i \times f_j^{\top}, \qquad j = 1, ..., n. \tag{1}$$

where s_i is an n-dimensional vector, range in $[-1, 1]$. Based on s_i, the Memory-based Positive Label Prediction (MPLP) [7] is used to predict the main pseudo multi-class label y_i for x_i.

In each training iteration, y_i is predicted and used to fine-tune the model until convergence. The Memory-based Multi-label Classification Loss (MMCL) [7] is used in baseline network to directly regress similarity score s_i to y_i as follows:

$$\mathcal{L}_{baseline} = \|s_i - y_i\|^2 \tag{2}$$

B. Auxiliary Module

The re-ID model trained only with y_i is usually sensitive to the noise in y_i, which hinders the feature learning in $\mathcal{F}(\cdot)$. To mitigate the effects of noisy pseudo labels, we proposed the auxiliary module, which seeks two auxiliary labels by investigating two underlying constraints in PLM, i.e., symmetric homogeneity and neighbor homogeneity. Ideally, the generated main label and two auxiliary labels should be the same because of two homogeneity constraints. Based on them, we design a mutual label learning training strategy to enforce one single network mutual training with different labels.

1) Pseudo Label Memory (PLM): To mine the generated labels relation on the whole dataset, the Pseudo Label Memory (PLM) is first constructed in this paper for storing the up-to-date main pseudo labels $\hat{Y} = \{y_1, y_2, ..., y_n\}$ of all images in X. The PLM is notated as $\mathcal{P}$, contains n slots, in which each slot storing a n-dimensional pseudo label y_i. Therefore, the size of PLM is $n \times n$. It is noteworthy that our proposed PLM is different with exemplar memory $\mathcal{M}$ in [7], [8], $\mathcal{M}$ is used to store the up-to-date features of all images in X. An illustration of PLM is shown in Figure 2 which assumes $n = 5$.

To store and inquire up-to-date labels for all images, PLM is updated using generated y_i in every training iteration through,

$$\mathcal{P}[i] \leftarrow y_i \tag{3}$$

2) Symmetric Homogeneity: Ideally, two images x_i and x_j should be mutual positive samples or mutual negative samples for each other. However, because the network is updated in each iteration based on the input batch size, the x_i and x_j might not be predicted as mutual positive or negative samples if they are fed-forward into the network in a different iteration. To effectively avoid the asymmetric error amplification, we seek auxiliary symmetric label y_i^{sym} of image x_i to enforce symmetric constraint into the network via our proposed mutual label learning strategy.

An illustration of the symmetric homogeneity of labels is shown in Figure 2. The main pseudo label $y_{\{i|i=2\}}$ of $x_{\{i|i=2\}}$ is save in PLM. The auxiliary symmetric label of y_i is inquired from PLM as follows,

$$y_i^{sym} = \mathcal{P}[:, i] \tag{4}$$

If there is no noisy pseudo labels, $\mathcal{P}$ is a symmetric matrices; in other words, $y^i = y_i^{sym}$, ideally. We utilize the symmetric constraint of generated pseudo label y_i in $\mathcal{P}$ to mitigate the effects of noisy pseudo labels by also regressing similarity score s_i to y_i^{sym} as follows:

$$\mathcal{L}_{sym} = \left\| s_i - y_i^{sym\top} \right\|^2 \tag{5}$$

Algorithm 1: SNNet Algorithm

Input: x_i: Input image in unlabeled dataset X

Initialize weighting factors λ^{sym} and λ^{nh}
Initialize $\mathcal{P}$ by identity matrix with size $n \times n$
for *epoch in* $[1, num\_epochs]$ **do**
  for *iter in* $[1, num\_iter]$ **do**
    1. Predict y_i for input image x_i
    2. Inquire y_i^{sym} from $\mathcal{P}$ based on symmetric homogeneity
    3. Inquire y_i^{nh} from $\mathcal{P}$ based on neighbor homogeneity
    4. Update $\mathcal{P}$ using predicted pseudo labels y_i
    5. Joint update $\mathcal{F}(\cdot)$ by the loss function Eq. 8

3) Neighbor Homogeneity: For an image $x_i \in X$, there may have another image x_e in X share same multi-class label with x_i. We aim to seek the x_e and exploit its main pseudo label y_e as an auxiliary label of x_i to further mitigate the effects of noise in y_i. To achieve this objective, we find the nearest neighbors of x_i according to the similarity between x_i and other images in X. The image share the highest similarity with x_i is the nearest neighbor of x_i. Then, we define the indexes of the nearest neighbor as e. Ideally, the predicted labels of x_i and its nearest neighbor x_e are same. Based on this constraint, we utilize the main label of x_e as the auxiliary neighbor label y_i^{nh} of image x_i to enforce neighbor homogeneity constraint into the network via our proposed mutual label learning strategy. The y_i^{nh} is inquired from PLM as follows,

$$y_i^{nh} = \mathcal{P}[e] \tag{6}$$

An illustration of auxiliary neighbor label y_i^{nh} is shown in Figure 2. Apart from the y_i and y_i^{sym}, We further mitigate the effects of noisy pseudo labels by regressing similarity score s_i to y_i^{nh} as follows:

$$\mathcal{L}_{nh} = \left\| s_i - y_i^{nh} \right\|^2 \tag{7}$$

C. Overall Loss

Different from training several student networks collaboratively and mutually in deep mutual learning [17], our proposed mutual label learning trains one single network $\mathcal{F}(\cdot)$ via the main pseudo label y_i and two auxiliary labels in a collaborative training manner.

The overall loss is the summation of $\mathcal{L}_{baseline}$, $\mathcal{L}_{sym}$, and $\mathcal{L}_{nh}$, which combines Eq. 2, Eq. 5, and Eq. 7 and is formulated as,

$$\mathcal{L} = \frac{1}{\eta}(\lambda^{base}\mathcal{L}_{baseline} + \lambda^{sym}\mathcal{L}_{sym} + \lambda^{nh}\mathcal{L}_{nh}) \tag{8}$$

where λ^{base}, λ^{sym} and λ^{nh} are the weighting parameters of $\mathcal{L}_{baseline}$, $\mathcal{L}_{sym}$ and $\mathcal{L}_{nh}$. η is the normalized parameter, $\eta = (\lambda^{base} + \lambda^{nh} + \lambda^{sym})$ to keep the scale of the gradient of loss = 1.0. Our model update steps are summarized in Algorithm 1.

TABLE I. Comparison with other fully unsupervised person re-ID methods on Market-1501 and DukeMTMC-ReID Dataset. "*": Baseline method, reproduced by us based on the authors' code. "↑": Results that outperforms baseline. Results that surpass all methods are **bold**.

Method	Reference	Market		Duke	
		R-1	mAP	R-1	mAP
CAMEL [18]	ICCV17	54.5	26.3	-	-
DECAMEL [19]	TPAMI18	60.2	32.4	-	-
BUC [4]	AAAI19	66.2	38.3	47.4	27.5
DBC [20]	BMVC19	69.2	41.3	51.5	30.3
SSL [22]	CVPR20	71.7	37.8	52.5	28.6
MMCL* (Resnet-50) [7]	CVPR20	79.1	44.1	63.6	39.0
MMCL* (OSNet) [7]	CVPR20	80.3	45.0	66.3	41.9
SNNet (Resnet-50)	**Ours**	80.2↑	48.3↑	66.1↑	41.5↑
SNNet (OSNet)	**Ours**	**85.1**↑	**58.3**↑	**69.2**↑	**46.7**↑

TABLE II. Methods comparison when tested on Market-1501 and DukeMTMC-reID. **Baseline**: Baseline model trained with main pseudo labels. **S**: Auxiliary symmetric labels y_i^{sym}. **N**: Auxiliary neighbor label y_i^{nh}.

Method	Market-1501				DukeMTMC-reID			
	R-1	R-5	R-10	mAP	R-1	R-5	R-10	mAP
Baseline	80.3	89.9	93.0	45.0	66.3	77.7	81.1	41.9
Baseline + S	83.8	91.3	93.8	49.0	67.5	78.5	81.8	44.1
Baseline + S + N	**85.1**	**92.6**	**94.3**	**58.3**	**69.2**	**79.6**	**83.2**	**46.7**

III. Experiments

A. Datasets and Evaluation Metrics

Market-1501 (Market) [23] has six cameras and 32,668 person images of 1,501 identities. DukeMTMC-reID (Duke) [24], [25] has eight cameras and 36,411 person images of 1,404 identities in total. Two evaluation metrics are used to measure model performance. The first one is the Cumulative Matching Characteristic (CMC) curves. The CMC represents the probability of top-k ranked gallery samples containing the query identity. The CMCs (%) of rank-1 (R-1), rank-5 (R-5), and rank-10 (R-10) are reported in this paper. The second evaluation metric is the Mean Average Precision (mAP) (%).

B. Implementation Details

The experiments are performed using one NVIDIA GeForce Titan 1080Ti GPU with 11 GB of memory. The experiments are implemented on PyTorch. The Resnet-50 [26] and OSNet [1] is adopted as the backbone network. Following the previous works [7], [8], we remove the subsequent layers after the pooling-5 layer and add a batch normalization layer. The backbone network is pre-trained on ImageNet [27]. During training, the initial learning rate is 0.1. The learning rate is divided by ten after 30 epochs. The training batch size is 128 with ResNet-50, and 64 with OSNet backbones because of the limited memory of the GPU. The network is trained in an end-to-end fashion by the Stochastic Gradient Descent (SGD). The weighting parameters $\lambda^{sym} = 1.0$, $\lambda^{sym} = 0.8$ and $\lambda^{nh} = 0.6$ are set for achieving the best performance.

C. Comparison with Other Methods

As shown in Table I, we compare our proposed SNNet with other fully unsupervised methods with ResNet-50 [26] and OSNet [1] backbones. We do not compare our method with the UDA-based methods, because UDA-based methods still required a labeled source dataset which is not the fully unsupervised method. The weighting parameters $\lambda^{sym} = 1.0$, $\lambda^{sym} = 0.8$ and $\lambda^{nh} = 0.6$ are set in Table I for achieving the best performance.

MMCL* [7] is the baseline approach in this letter which is reproduced by us based on the authors' code. Compared to the baseline, the proposed SNNet improves the model performance with ResNet-50 and OSNet on two datasets consistently. More specifically, on DukeMTMC-reID, 2.7% Rank-1 and 2.9% mAP improvements with ResNet-50 backbone, and 3.1% Rank-1 and 5.2% mAP improvements with OSNet backbone are observed. The results indicate the importance of our proposed two label homogeneity constraints. Moreover, the proposed SNNet achieves the best performance among the compared methods with ResNet-50 and OSNet on Market-1501 and DukeMTMC-reID. The superior performance with different backbones indicates the robustness and effectiveness of our designed mutual label learning structure.

D. Ablation Studies

In this section, we evaluate each components in SNNet by conducting ablation studies on Market-1501 and DukeMTMC-reID with OSNet [1] backbones.

1) Effectiveness of Mutual Label Learning: we conduct ablation studies to investigate the effectiveness of the proposed mutual label learning in Table II. The weighting parameters $\lambda^{sym} = 1.0$, $\lambda^{sym} = 0.8$ and $\lambda^{nh} = 0.6$ are set in Table II for achieving the best performance. First, we report the result of the baseline network as "Baseline" in Table II, which not incorporates auxiliary labels into the training. Our proposed "Baseline + S" and "Baseline + S + N" consistently improve the results over "baseline", e.g., from baseline 80.3% to 85.1% in rank-1 and 45.0% to 58.3% on Market-1501, and 66.3% to 69.2% in rank-1 and 41.9% to 46.7% on DukeMTMC-reID. The results demonstrate the proposed mutual label learning

TABLE III. Evaluation with different values of λ^{sym}.

λ^{sym}	Market-1501				DukeMTMC-reID			
	R-1	R-5	R-10	mAP	R-1	R-5	R-10	mAP
0.0 (baseline)	80.3	89.9	93.0	45.0	66.3	77.0	81.5	41.9
0.2	81.9	90.9	93.4	47.6	66.9	77.8	82.1	42.6
0.4	82.2	91.1	93.3	49.6	67.1	77.5	80.9	43.3
0.6	82.9	90.6	93.2	49.3	**67.5**	78.2	**82.2**	43.5
0.8	**83.8**	**91.3**	**93.8**	49.0	**67.5**	**78.5**	81.8	**44.1**
1.0	83.5	91.2	93.2	**50.2**	66.1	77.5	81.8	42.8

TABLE IV. Evaluation with different values of λ^{nh}.

λ^{nh}	Market-1501				DukeMTMC-reID			
	R-1	R-5	R-10	mAP	R-1	R-5	R-10	mAP
0.0 (baseline)	80.3	89.9	93.0	45.0	66.3	77.0	81.5	41.9
0.2	**84.8**	**91.7**	**94.1**	**55.7**	**69.2**	**79.8**	83.0	46.1
0.4	82.8	90.6	92.8	55.2	68.6	78.5	82.2	45.7
0.6	80.1	88.7	91.9	51.7	69.1	79.7	81.8	46.2
0.8	80.3	89.3	91.6	51.4	68.8	79.5	82.5	46.0
1.0	78.5	88.2	91.1	50.5	69.1	**79.8**	**83.3**	**46.7**

with auxiliary labels is an effective way to improve the re-ID model performance, and the necessity of high-quality pseudo labels in the unsupervised person re-ID task.

2) Effectiveness of Symmetric Homogeneity: In Table III, we compare different values of λ^{sym} by keeping $\lambda^{nh} = 0.0$ in Eq. 8 for analyzing the effect of symmetric homogeneity to baseline network. When $\lambda^{sym} = 0.0$, the re-ID model is only trained with main pseudo labels y_i. Comparing $\lambda^{sym} = 1.0$ to $\lambda^{sym} = 0.0$, we observe 3.2% Rank-1 and 5.2% mAP improves on Market-1501. Comparing $\lambda^{sym} = 0.8$ to $\lambda^{sym} = 0.0$, we observe 1.2% Rank-1 and 2.2% mAP improves on DukeMTMC-reID. The improvements demonstrate the effectiveness of our proposed symmetric homogeneity. We observe that higher λ^{sym} value achieves better result on Market-1501. Different to Market-1501, we achieve best results on DukeMTMC-reID when $\lambda^{sym} = 0.8$ rather than $\lambda^{sym} = 1.0$ because of over-fitting.

3) Effectiveness of Neighbor Homogeneity: Table IV investigates the effect of neighbor homogeneity in our method by varying λ^{nh} form 0.0 to 1.0 on both Market-1501 and DukeMTMC-reID datasets. We keep $\lambda^{sym} = 0.0$ in Eq. 8 for analyzing the effect of neighbor homogeneity to baseline network. Using auxiliary neighbor labels significantly boosts the performance on Market-1501 and DukeMTMC-reID, e.g., from baseline 80.3% to 84.8% in rank-1 and 45.0% to 55.7% on Market-1501. With the increasing of λ^{nh}, we observe the model is very easy to over-fitting. However, Comparing $\lambda^{nh} = 1.0$ to $\lambda^{nh} = 0.0$, we still observe improvements on both Market-1501 and DukeMTMC-reID. The significant improvements demonstrate the necessity of our proposed neighbor homogeneity constraint.

IV. Conclusion

This paper introduces an exemplar memory-based fully unsupervised method for person re-ID task via mining two underlying label homogeneities, symmetric homogeneity and neighbor homogeneity. We design a mutual label learning framework to enforce two label homogeneities constraints into the network training by optimizing the network under the joint supervision of the main pseudo label and two auxiliary labels in every training iteration. The experiment results on Market-1501 and DukeMTMC-reID demonstrate the effectiveness of our approach.

Acknowledgment

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the government (MSIT). (No.2020R1A2C200897212)

References

[1] K. Zhou, Y. Yang, A. Cavallaro, and T. Xiang. Omni-scale feature learning for person re-identification. In 2019 IEEE/CVF International Conference on Computer Vision (ICCV), pages 3701-3711, 2019.

[2] Q. Zhou, B. Zhong, X. Lan, G. Sun, Y. Zhang, B. Zhang, and R. Ji. Fine-grained spatial alignment model for person re-identification with focal triplet loss. IEEE Transactions on Image Processing, 29:7578-7589, 2020.

[3] H. Huang, W. Yang, J. Lin, G. Huang, J. Xu, G. Wang, X. Chen, and K. Huang. Improve person re-identification with part awareness learning. IEEE Transactions on Image Processing, 29:7468-7481, 2020.

[4] Yutian Lin, Xuanyi Dong, Liang Zheng, Yan Yan, and Yi Yang. A bottom-up clustering approach to unsupervised person re-identification. Proceedings of the AAAI Conference on Artificial Intelligence, 33:8738-8745, Jul. 2019.

[5] Y. Fu, Y. Wei, G. Wang, Y. Zhou, H. Shi, U. Uiuc, and T. Huang. Self-similarity grouping: A simple unsupervised cross domain adaptation approach for person reidentification. In 2019 IEEE/CVF International Conference on Computer Vision (ICCV), pages 6111-6120, 2019.

[6] X. Zhang, J. Cao, C. Shen, and M. You. Self-training with progressive augmentation for unsupervised cross-domain person re-identification. In 2019 IEEE/CVF International Conference on Computer Vision (ICCV), pages 8221-8230, 2019.

[7] D. Wang and S. Zhang. Unsupervised person reidentification via multi-label classification. In 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pages 10978-10987, 2020.

[8] Z. Zhong, L. Zheng, Z. Luo, S. Li, and Y. Yang. Invariance matters: Exemplar memory for domain adaptive person re-identification. In 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pages 598-607, 2019.

[9] Q. Tang and K. -H. Jo, "Unsupervised Person Re-Identification Via Nearest Neighbor Collaborative Training Strategy," 2021 IEEE International Conference on Image Processing (ICIP), 2021, pp. 1139-1143.

[10] Q. Tang, G. Cao and K. -H. Jo, ”Fully Unsupervised Person Re-Identification via Multiple Pseudo Labels Joint Training,” in IEEE Access, vol. 9, pp. 165120-165131, 2021.
[11] L. Wei, S. Zhang, W. Gao, and Q. Tian. Person transfer gan to bridge domain gap for person re-identification. In 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 79-88, 2018.
[12] Zhun Zhong, Liang Zheng, Shaozi Li, and Yi Yang. Generalizing a person retrieval model hetero- and homogeneously. In Proceedings of the European Conference on Computer Vision (ECCV), September 2018.
[13] L. Qi, L. Wang, J. Huo, L. Zhou, Y. Shi, and Y. Gao. A novel unsupervised camera-aware domain adaptation framework for personre-identification. In 2019 IEEE/CVF International Conference on Computer Vision (ICCV), pages 8079-8088, 2019.
[14] K. Jiang, T. Zhang, Y. Zhang, F. Wu, and Y. Rui. Selfsupervised agent learning for unsupervised cross-domain person re-identification. IEEE Transactions on Image Processing, 29:8549-8560, 2020.
[15] Y. Li, C. Lin, Y. Lin, and Y. F. Wang. Cross-dataset person re-identification via unsupervised pose disentanglement and adaptation. In 2019 IEEE/CVF International Conference on Computer Vision (ICCV), pages 7918-7928, 2019.
[16] H. Yu, W. Zheng, A. Wu, X. Guo, S. Gong, and J. Lai. Unsupervised person re-identification by soft multilabel learning. In 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pages 2143-2152, 2019.
[17] Y. Zhang, T. Xiang, T. M. Hospedales, and H. Lu. Deep mutual learning. In 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 4320-4328, 2018.
[18] H. Yu, A. Wu, and W. Zheng. Cross-view asymmetric metric learning for unsupervised person re-identification. In 2017 IEEE International Conference on Computer Vision (ICCV), pages 994-1002, 2017.
[19] H. Yu, A. Wu, and W. Zheng. Unsupervised person reidentification by deep asymmetric metric embedding. IEEE Transactions on Pattern Analysis and Machine Intelligence, 42(4):956-973, 2020.
[20] Q. Yin G. Ding, S. Khan and Z. Tang. Dispersion based clustering for unsupervised person re-identification. In 2019 BMVC, 2019.
[21] K. Zeng, M. Ning, Y. Wang, and Y. Guo. Hierarchical clustering with hard-batch triplet loss for person reidentification. In 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pages 13654-13662, 2020.
[22] Y. Wu C. Yan Y. Lin, L. Xie and Q. Tian. Unsupervised person re-identification via softened similarity learning. In 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), page 3390-3399, 2017.
normalization and restitution for generalizable person reidentification. In 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), page 31403149, 2020.
[23] L. Zheng, L. Shen, L. Tian, S. Wang, J. Wang, and Q. Tian. Scalable person re-identification: A benchmark. In 2015 IEEE International Conference on Computer Vision (ICCV), pages 1116-1124, 2015.
[24] Z. Zheng, L. Zheng, and Y. Yang. Unlabeled samples generated by gan improve the person re-identification baseline in vitro. In 2017 IEEE International Conference on Computer Vision (ICCV), pages 3774-3782, 2017.
[25] Roger Zou Rita Cucchiara Ergys Ristani, Francesco Solera and Carlo Tomasi. Performance measures and a data set for multi-target, multi-camera tracking. In ECCV, 2016.
[26] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 770-778, 2016.
[27] J. Deng, W. Dong, R. Socher, L. Li, Kai Li, and Li FeiFei. Imagenet: A large-scale hierarchical image database. In 2009 IEEE Conference on Computer Vision and Pattern Recognition, pages 248-255, 2009.