Supplementary Material for "Dual-Refinement: Joint Label and Feature Refinement for Unsupervised Domain Adaptive Person Re-Identification"

Anonymous Author(s)
Submission Id: 679

A OVERALL TRAINING PROCEDURE

Algorithm 1: Alternative Training Procedure of Our Dual-Refinement Method

Input: Labeled source dataset D_s; Unlabeled target dataset D with N images; Feature encoder F pretrained on ImageNet; Identity classifier ϕ; Instant memory bank V; Maximum training epoch max_epoch; Maximum training iteration max_iter

Output: Optimized feature encoder F for target domain;

Pretain feature encoder F on the labeled source dataset D_s with classification loss and triplet loss;
for $epoch = 1$ *to* max_epoch **do**
  // *Off-line pseudo label generation and refinement*
  Extract features of the target dataset D using F and calculate Jaccard distance $d_J(i, j)$ by Eq. (1) (2);
  Perform DBSCAN clustering on D with d_J and assign the coarse pseudo labels $\widetilde{D} = \{(x_i, \widetilde{y}_i)|_{i=1}^{N}\}$;
  Perform fine clustering and assign refined pseudo labels by Eq. (6) (7) to obtain the refined target dataset $\widehat{D} = \{(x_i, \widehat{y}_i)|_{i=1}^{N}\}$;
  // *On-line feature learning and refinement*
  for $iter = 1$ *to* max_iter **do**
    Sample $(x_i, \widetilde{y}_i, \widehat{y}_1)$ from $\widetilde{D} \cup \widehat{D}$;
    Update the feature encoder F, classifier ϕ and instant memory bank V by computing the gradients of the overall loss (Eq. (12)) with back-propagation;
  end
end
return feature encoder F;

We demonstrate the details of the alternative training procedure for our Dual-Refinement method in Algorithm 1. Our method can be easily incorporated into the general clustering-based UDA framework [4]. In detail, we perform the DBSCAN clustering [1] by the implementation in scikit-learn [3].

B COMPUTATIONAL COST COMPARISONS

As shown in Table 1, we compare our Dual-Refinement method with the baseline and the state-of-the-art method MMT [2] in the computional cost. The experiments are conducted when Market1501 [5] → DukeMTMC-ReID [6]. MMT uses two networks to train with each other, which is not memory efficient. Compared with MMT, our Dual-Refinement can achieve higher performance by costing less training time and GPU memory. Compared with the baseline method, our Dual-Refinement only introduces little extra GPU memory cost (about 908 MB) and little extra time cost (about 0.36 hours) because of the proposed instant memory bank. However, our Dual-Refinement outperforms the baseline method's rank-1 accuracy (R1) by a large margin. Based on the above analyses, our proposed Dual-Refinement is superior not only in the performance but also in the computational cost.

Table 1: Computional cost comparisions.

Method	Market1501 → DukeMTMC-ReID		
	R1 (%)	Time (hours)	GPU Memory (MB)
Baseline	72.5	3.17	8692
Dual-Refinement	82.1	3.53	9600
MMT	78.0	11.45	15068

REFERENCES

[1] Martin Ester, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu. 1996. A density-based algorithm for discovering clusters in large spatial databases with noise.. In *Kdd*, Vol. 96. 226–231.
[2] Yixiao Ge, Dapeng Chen, and Hongsheng Li. 2020. Mutual Mean-Teaching: Pseudo Label Refinery for Unsupervised Domain Adaptation on Person Re-identification. *arXiv preprint arXiv:2001.01526* (2020).
[3] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. 2011. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 12 (2011), 2825–2830.
[4] Liangchen Song, Cheng Wang, Lefei Zhang, Bo Du, Qian Zhang, Chang Huang, and Xinggang Wang. 2020. Unsupervised domain adaptive re-identification: Theory and practice. *Pattern Recognition* (2020), 107173.
[5] Liang Zheng, Liyue Shen, Lu Tian, Shengjin Wang, Jingdong Wang, and Qi Tian. 2015. Scalable person re-identification: A benchmark. In *Proceedings of the IEEE international conference on computer vision*. 1116–1124.
[6] Zhedong Zheng, Liang Zheng, and Yi Yang. 2017. Unlabeled samples generated by gan improve the person re-identification baseline in vitro. In *Proceedings of the IEEE International Conference on Computer Vision*. 3754–3762.