

Supplementary Materials: Robust Pseudo-label Learning with Neighbor Relation for Unsupervised Visible-Infrared Person Re-Identification

I Datasets

SYSU-MM01 stands as a large-scale, publicly available benchmark tailored for the VI-ReID task, boasting a diverse collection of 491 identities captured across four RGB cameras and two IR cameras, spanning both indoor and outdoor environments. Within this dataset, a total of 22,258 RGB images and 11,909 IR images, portraying 395 distinct identities, have been meticulously curated for training purposes. During the inference phase, the query set encompasses 3,803 IR images, representative of 96 unique identities, while the gallery set comprises 301 randomly selected RGB images. In contrast, the RegDB dataset, captured by a single RGB camera and a single IR camera, features 4,120 RGB images and 4,120 IR images, each depicting 412 distinct identities. To elaborate further, the dataset is partitioned into two disjoint sets: one designated for training and the other for testing.

II Implementation Details

The Adam optimizer is adopted to train the network with a weight decay of $5e-4$. The initial learning rate is set to $3.5e-4$, which decays to $1/10$ of its previous value every 20 epochs. In the first 50 epochs, we employ the DCL [4] framework to alternately offline pseudo-labels generation and online representation learning. The proposed framework is trained in the last 50 epochs. Additionally,

we store multiple proxies for each cluster to provide complementary representation at each stage when constructing the memory following [3, 6]. The parameter κ for κ -reciprocal nearest neighbors is set to 30 following [5] and K is fixed to 20. The hyper-parameter λ is set to 25 following [2]. Following ADCA [4], the momentum value μ is set to 0.1 and the temperature τ is 0.5. The margin hyper-parameter γ and the kernel bandwidth σ are both set to 1.0 following [1]. The trade-off hyper-parameter α is set to 0.5, β_1 and β_2 is set to 0.5 and 10.0, respectively.

References

- [1] Sungyeon Kim, Dongwon Kim, Minsu Cho, and Suha Kwak. 2021. Embedding transfer with label relaxation for improved metric learning. In *CVPR*. 3967–3976.
- [2] Jiangming Wang, Zhizhong Zhang, Mingang Chen, Yi Zhang, Cong Wang, Bin Sheng, Yanyun Qu, and Yuan Xie. 2022. Optimal Transport for Label-Efficient Visible-Infrared Person Re-Identification. In *ECCV*. 93–109.
- [3] Yuhang Wu, Tengting Huang, Haotian Yao, Chi Zhang, Yuanjie Shao, Chuchu Han, Changxin Gao, and Nong Sang. 2022. Multi-Centroid Representation Network for Domain Adaptive Person Re-ID. In *AAAI*. 2750–2758.
- [4] Bin Yang, Mang Ye, Jun Chen, and Zesen Wu. 2022. Augmented Dual-Contrastive Aggregation Learning for Unsupervised Visible-Infrared Person Re-Identification. In *ACM MM*. 2843–2851.
- [5] Zhun Zhong, Liang Zheng, Donglin Cao, and Shaozi Li. 2017. Re-ranking person re-identification with k-reciprocal encoding. In *CVPR*. 1318–1327.
- [6] Chang Zou, Zeqi Chen, Zhichao Cui, Yuehu Liu, and Chi Zhang. 2023. Discrepant and Multi-Instance Proxies for Unsupervised Person Re-Identification. In *ICCV*. 11058–11068.