Unsupervised Visible-Infrared Person Re-identification under Unpaired Settings

Haoyu Yao[1,2,3,4*] Bin Yang[1,2,3*] Wenke Huang[1,2,3] Bo Du[1,2,3] Mang Ye[1,2,3†]
[1] School of Computer Science, Wuhan University, China
[2] National Engineering Research Center for Multimedia Software, Wuhan University, China
[3] Hubei Key Laboratory of Multimedia and Network Communication Engineering, Wuhan University, China
[4] School of Remote Sensing and Information Engineering, Wuhan University, China

https://github.com/USL-VI-ReID/MCL

Evaluation of hyper-parameter η. We explore the influence of hyper-parameter η as shown in Fig. 1. The η determines the balance of L_{SL} and L_{DL} in static and dynamic learning. When $\eta = 0$, the dynamice learning is excluded. When $\eta = 1.0$, the method achieves a balance in static and dynamic learning.

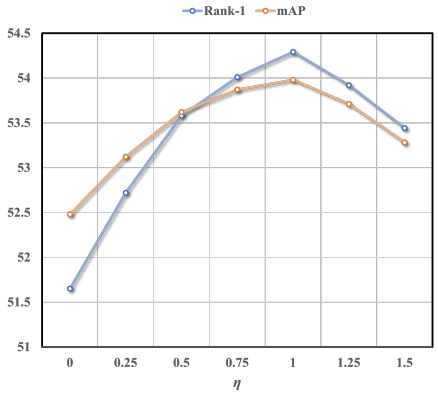

Figure 1. Evaluation of hyper-parameter η. The results are based on all search mode of unpaired SYSU-MM01 dataset, where α is set to 0.5. Rank-1 accuracy (%) and mAP (%) are reported.

Details of unpaired replacement process. As illustrated in Fig. 2, suppose there are 4 identities (*i.e.*, 1, 2, 3, 4) in both the visible and infrared training sets. If we set $\alpha = 0.5$, two identities (*e.g.*, 2 and 3) are randomly selected as overlapping identities. The visible training set then consists of 1^*, 2, 3, and $4^\dagger$, while the infrared set remains 1, 2, 3, 4. Here, 1^* and $4^\dagger$ are different identities sampled from external datasets (e.g., 1^* from Market-1501 and $4^\dagger$ from MSMT17), replacing the original identities 1 and 4 in the visible modality.

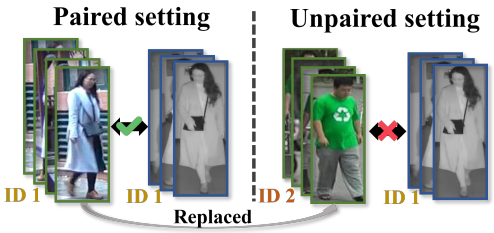

Figure 2. Illustration of the paired and unpaired settings. In the unpaired settings, the visible images of selected identities are replaced with external images, preserving image count per identity.

*Equal contribution.
†Corresponding author.