Supplementary Material: Augmented and Softened Matching for Unsupervised Visible-Infrared Person Re-Identification

Zhiqi Pang[1] Chunyu Wang[1*] Lingling Zhao[1] Junjie Wang[2]
[1] Harbin Institute of Technology, China
[2] Nanjing Medical University, China
zqpang98@gmail.com, chunyu@hit.edu.cn, zhaoll@hit.edu.cn, junjiehit@163.com

1. Algorithm

Algorithm 1 Augmented and softened matching

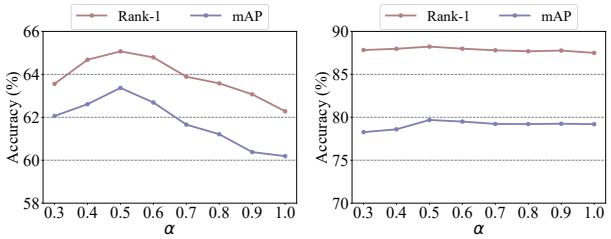
Input: Unlabeled data $\{x_i^v\}_{i=1}^{N^v}$ and $\{x_i^r\}_{i=1}^{N^r}$, encoder parameterized by θ_E, training $epochs$ and $iters$.

for $i = 1$ to $epochs$ **do**
  Extract features of $\{x_i^v\}_{i=1}^{N^v}$ and $\{x_i^r\}_{i=1}^{N^r}$ with encoder
  Calculate cluster centroids based on Eq. 1 and Eq. 2 respectively
  for $j = 1$ to $iters$ **do**
    Optimize θ_E to minimize the loss in Eq. 5
  end for
end for
Save the intra-modality pseudo-labels
for $i = 1$ to $epochs$ **do**
  Apply channel augmentation to $\{x_i^v\}_{i=1}^{N^v}$ to obtain $\{x_i^a\}_{i=1}^{N^a}$
  Extract features of $\{x_i^v\}_{i=1}^{N^v}$, $\{x_i^a\}_{i=1}^{N^a}$ and $\{x_i^r\}_{i=1}^{N^r}$ with encoder
  Perform cross-modality augmented matching
  Perform soft-labels momentum update
  for $j = 1$ to $iters$ **do**
    Optimize θ_E to minimize the loss in Eq. 18
  end for
end for
return Trained θ_E.

2. Sensitivity Analysis

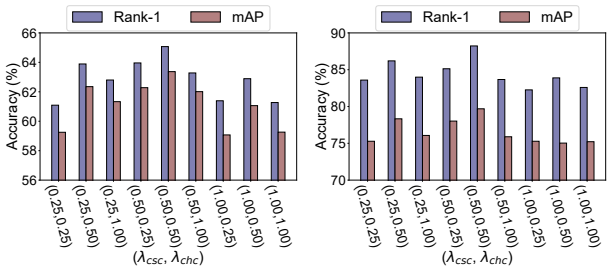
2.1. Update rate α

α controls the update rate of the inter-modality soft pseudo-labels. When $\alpha = 1.0$, the soft-labels momentum update becomes ineffective, and the soft-labels are converted into hard-labels. In this case, the model only relies on the matching results of the current epoch for optimization. As shown in Fig. 1, when $\alpha = 0.5$, the model achieves the best performance on both datasets. When $\alpha = 1.0$, the model's performance is poorer, which also validates the effectiveness of the soft-labels momentum update.



(a) SYSU-MM01(All Search) (b) RegDB(Visible to Thermal)

Figure 1. The effect of α on model performance.



(a) SYSU-MM01(All Search) (b) RegDB(Visible to Thermal)

Figure 2. The effect of λ_{csc} and λ_{chc} on model performance.

2.2. λ_{csc} and λ_{chc}

As shown in Fig. 2, we observe that when $\lambda_{csc} = 0.5$ and $\lambda_{chc} = 0.5$, the model achieves optimal performance on both datasets. When λ_{csc} and λ_{chc} are set to 0.25 or 1, the model performs poorly. This is because small weights limit the contribution of L_{csc} and L_{chc} to the model's optimization, while large weights may introduce excessive incorrect inter-modality optimization directions. Since the reliability of intra-modality pseudo-labels is generally higher

*Corresponding author: Chunyu Wang

than that of inter-modality pseudo-labels, the optimization should still primarily rely on the more reliable supervision signals within each modality.