Negative Samples are at Large: Leveraging Hard-distance Elastic Loss for Re-identification (Supplementary Material)

Hyungtae Lee[1] Sungmin Eum[1,2] Heesung Kwon[1]

[1]DEVCOM Army Research Laboratory [2]Booz Allen Hamilton

A Details of Datasets Used for Evaluation

In this section, we provide details of three re-ID datasets used for evaluation.

VeRi-776 [2] was developed for vehicle re-ID. This dataset consists of 49,357 images of 776 vehicles taken by 20 different cameras. 37,775 images with 576 IDs are used for training and the remaining images are used for testing. In the test set, 1,678 images are selected for the query set.

Market-1501 [6] was developed for person re-ID. It contains 32,668 images of 1,501 IDs. Images of each ID are taken by at most six cameras. 12,936 images and 19,732 images were used for training and testing, respectively. 3,368 images are selected as the query.

VeRi-Wild [3] is a large-scale vehicle ID dataset which contains 416,314 images of 40,671 IDs from 174 cameras. 138,517 images of 10,000 IDs are used as a test set and evaluations are carried out with three subsets with 3,000 (`small`), 5,000 (`medium`), and 10,000 (`large`) IDs. The remaining images are used for training.

B Implementation: N-pair Loss and Ranked List Loss

In the main manuscript, we chose the N-pair [4] loss and the Ranked List loss [5] as the two representative losses that can leverage multiple negative samples. The original N-pair loss is developed for multi-class tasks and employs hard-negative class mining, which selects a pre-defined number of hard samples from each class except the class where the query belongs to. To be used as a tripletwise loss, however, we selected hard negative samples regardless of their classes, as negative representative samples. When evaluating the N-pair loss within the two re-ID frameworks (baseline and MoReID), we defined the number of hard negative samples to be 15 which was found to be optimal for both frameworks. As the N-pair loss does not allow multiple positives to be used as representative samples, only the hardest positive sample is leveraged.

Ranked List loss considers two sets of multiple hard examples to represent the positive and negative sets, respectively. Hard example mining is controlled by two parameters[1], α and β where the positives satisfying $d_{pq} > \alpha - \beta$ and the

[1] Although the margin parameter is denoted as m in [5], we use β instead in this paper to avoid confusion with the momentum parameter.

negatives satisfying $d_{nq} < \alpha$ are chosen as hard examples. When Ranked List loss is used with either baseline or MoReID, the optimal parameters were found to be 1.2 for α and 0.4 for β after an exhaustive search.

C Ablation Study: Positive Samples Are Also at Large?

To validate our decision to only leverage the "negative samples at large" (i.e., excluding the past samples with the same ID as the query from training), we test out a "positive samples are also at large" scenario where all excluded samples are pulled back in as positive samples. The comparison is shown below:

w/ past positive samples	m		
	0.995	0.997	0.999
✓	80.6	80.7	81.3
✗ (Ours)	82.4	83.3	83.0

As training the forward-only encoder progresses faster (i.e., m getting smaller), the degree of inconsistency of past samples with respect to a given query becomes larger. To observe the effect of the degree of inconsistency, experiments were conducted with various m. When past positive samples are included for training, using a larger m to maintain consistency was crucial to yield a better accuracy. Nevertheless, using past positive samples expressed in the maximally consistent feature space (i.e., $m = 0.999$ case) performed worse than the case where these past samples are excluded.

D Ablation Study: Similarity Measure.

As the proposed HE loss is defined using the similarity between the query sample and the key sample, we seek to find the most appropriate similarity measure by comparing the two methods: negative cosine similarity and the Euclidean distance. The similarity measures are defined as follows:

- cosine: $d(x, y) = -\left(x/||x||_2\right)^T \left(y/||y||_2\right)$
- euclidean: $d(x, y) = ||x - y||_2$

The table below shows mAP of models trained with the HE loss using different similarity measures:

w/ Triplet loss	w/ HE loss	
	cosine	euclidean
79.4	82.9	**83.3**

Regardless of the similarity measure, our model using the HE loss outperforms the Triplet loss. Although subtle, using the Euclidean distance yields the better than negative cosine similarity. Accordingly, we have used the Euclidean distance for the HE loss throughout all experiments.

E Ablation Study: Backbone Components

As studied and suggested by [1], we used a ResNet-50 architecture with new components as the backbone for optimal accuracy. The notable components are auto-augment pre-processing (AutoAug), instance batch normalization (IBN), and non-local modules (non-local). To verify that using the proposed HE loss is effective regardless of the influence of these components, we performed a set of "Triplet loss vs. HE loss" experiments where several variations of ResNet-50 were tested including the vanilla version as shown below:

AutoAug	IBN	non-local	w/ Triplet	w/ HE (gain)
✗	✗	✗	77.7	**80.2** (+2.5)
✗	✓	✓	79.4	**83.2** (+3.8)
✓	✗	✓	78.6	**82.1** (+3.5)
✓	✓	✗	78.7	**82.8** (+4.1)
✓	✓	✓	79.7	**83.3** (+3.6)

Using the proposed HE loss consistently yields better accuracy than using the Triplet loss regardless of which components are being used.

References

1. He, L., Liao, X., Liu, W., Liu, X., Cheng, P., Mei, T.: FastReID: A pytorch toolbox for general instance re-identification. arXiv:2006.02631 (2020), https://arxiv.org/abs/2006.02631
2. Liu, X., Liu, W., Mei, T., Ma, H.: A deep learning-based approach to progressive vehicle re-identification for urban surveillance. In: ECCV (2016)
3. Lou, Y., Bai, Y., Liu, J., Wang, S., Duan, L.Y.: VERI-Wild: A large dataset and a new method for vehicle re-identification in the wild. In: CVPR (2019)
4. Sohn, K.: Improved deep metric learning with multi-class n-pair loss objective. In: NeurIPS (2016)
5. Wang, X., Hua, Y., Kodirov, E., Robertson, N.M.: Ranked list loss for deep metric learning. In: CVPR (2019)
6. Zheng, L., Shen, L., Tian, L., Wang, S., Wang, J., Tian, Q.: Scalable person re-identification: A benchmark. In: ICCV (2015)