Relation Preserving Triplet Mining for Stabilizing the Triplet Loss in Re-identification Systems: Supplementary Material

Adhiraj Ghosh[1,2], Kuruparan Shanmugalingam[1,3], and Wen-Yan Lin[1]

[1]Singapore Management University, [2]University of Tübingen, [3]University of New South Wales

Overview

In this document, we provide additional results and technical details. Firstly, Section 1 provides a detailed description of the network architecture. Section 2 analyses the presence of outliers in reID datasets and how they degrade feature learning. Section 3 provides qualitative analysis of the triplet mining strategy used by RPTM in semi-hard positive sampling. Section 4 demonstrates the role played by reranking in improving evaluation results. Finally, we show more visual comparisons of our model against current state-of-the-art results in Section 5.

1. Architecture Details and Hyperparameters

A major focus of our research was to create a universal model that performs well across benchmarks, for both vehicle re-id and person re-id datasets. As mentioned in Section 5.2 of the main paper,we observed that RPTM performs well on the ResNet backbone with Squeeze-Excitation [2] and an Instance Batch Normalisation (IBN) appendage. Table 1 provides the details of the hyperparameters of the universal network that generated state-of-the-art results in vehicle re-id and comparable results in person re-id.

	RPTM (ResNet101 Baseline)
Input Size(Veri-776)	240×240
Input Size(VehicleID)	240×240
Input Size(DukeMTMC)	300×150
Train Batch Size	24
Test Batch Size	100
Workers	8
Optimizer	SGD
Momentum	0.9
Weight Decay	$5e^{-4}$
Learning Rate	0.005
Scheduler	MultiStepLR
Decay Factor	0.1
Margin (triplet loss)	0.3
λ_{tri}	2
λ_{xent}	0.5
Stride	1
Droprate	0.2
Pooling	Average
Pre-trained	ImageNet
Feature Dimension	2048

Table 1: Hyperparameters of RPTM (ResNet101 Baseline) used for implementing RPTM on reID benchmark datasets.

2. Outlier Analysis

Figure 1 displays several cases of outliers in reID datasets. Outliers prevent images from mapping the correct semantic meaning of the IDs the images belong to, making outliers extremely problematic for reID. One of the most common outliers in reID datasets is images with occlusions that tend to mask the object being studied, wholly or partially. In person reID, bounding boxes sometimes capture two persons and the focus is put on the wrong object. In vehicle reID, especially for large datasets like VehicleID, a major issue is the presence of the same model across IDs. This outlier issue is further exacerbated when these models are of the same colour as well. All these cases make it difficult to train CNNs to learn accurate features. Hence, RPTM is proposed to exclude such outlier cases and learns additional features during positive selection for triplet mining. By cleaning up the reID process as described, RPTM can train reID models robust to outliers.

3. Triplet Mining and Triplet Loss

The triplet mining method proposed in Section 4.2 of the main paper is quite effective is selecting a suitable semi-hard positive sample for an anchor image during training. The triplet mining strategy used estimates threshold τ which reflects the bare minimum GMS matches $RPTM_{min}$, the non-zero average value of GMS matches

Figure 1: Outliers serve as a hindrance to proper training of reID models. Standard models are unable to focus on fine-grained details and resolve outlier cases during training. Some examples of outliers in the DukeMTMC(top two rows) and VehicleID(bottom row) datasets. The first row shows the effect of occlusions. The second row shows an overlap between object tracklets of two persons. The last row shows the distribution of the same vehicle model with the same colour across several IDs.

$RPTM_{mean}$ or the maximum GMS matches $RPTM_{max}$, from which we select $RPTM_{mean}$ as the threshold for positive sample selection. Figure 2 reflects the above strategy across all three benchmark datasets used in this paper.

Table 2 describes the role of triplet loss. We have established how important positive mining is to the triplet loss cost function but it is prudent to evaluate the role triplet loss plays in reID. To that end, we manipulate the co-efficient of the triplet loss function λ_{tri}. Set at 2 for normal experiments, we change the value of λ_{tri} to 0.5 and 1 and train the RPTM model on Veri-776 and DukeMTMC. Since cross-entropy loss is significantly larger than triplet loss, λ_{xent} is set at 0.5 throughout our experiments.

4. Re-Ranking

Here, we test the variation of re-ranked mAP for the Veri-776 and DukeMTMC dataset. Implementing the process used by [4], Figure 3 involves manipulating and setting the values of three coefficients, k_1, k_2 and η, which represent co-efficients(k_1, k_2) and penalty factors(η) used to revise the ranking list calculated from standard evaluation and re-calculate the pairwise distance for the new ranking list. Figure 3a shows mAP results with the variation of $k1$ keeping $k2$ and η at 15 and 0.2 for Veri-776 and at 10 and 0.2 for DukeMTMC. Figure 3b shows the impact of $k2$ on mAP, with $k1$ and η set at 60 and 0.2 for Veri-776 and 20 and 0.2 for DukeMTMC. Finally, the impact of η is studied in figure 3c, with $k1$ and $k2$ fixed at 60 and 15 for Veri-776 and at 20 and 10 for DukeMTMC respectively.

	mAP	r = 1	r = 5
Duke(λ_{tri}=1)	80.6	86.8	94.1
Duke(λ_{tri}=0.5)	77.2	83.8	92.6
Veri(λ_{tri}=1)	81.8	94.7	96.9
Veri(λ_{tri}=0.5)	79.9	93.4	95.8

Table 2: Evaluation results with the λ co-efficient for triplet-cost set at 0.5. We observe a significant drop in mAP and top-k rank across both datasets.

5. More Experimental Results

In this section, we provide more visual comparisons with state-of-the-art methods, based on the top-k gallery matches for a given query sample, where we set the value of k to 20. Correct ID matches are denoted in green boxes whereas wrong matches are enclosed in red boxes. We compare our vehicle reID results on Veri-776 (Figure 4) with DMT [1] and our person reID results on DukeMTMC(Figure 5) with st-ReID [3]. The proposed algorithm generates strong pose-aware results, especially in the top-10 matches.

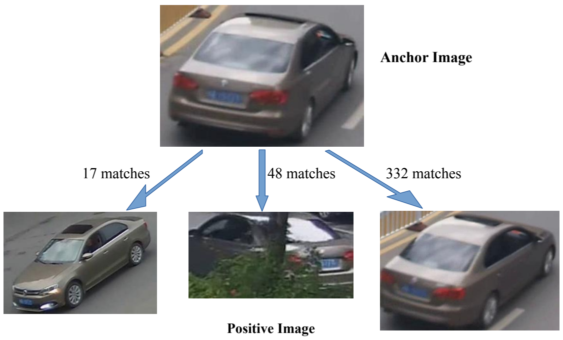
References

[1] Shuting He, Hao Luo, Weihua Chen, Miao Zhang, Yuqi Zhang, Fan Wang, Hao Li, and Wei Jiang. Multi-domain learning and identity mining for vehicle re-identification. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pages 582–583, 2020. 2, 4

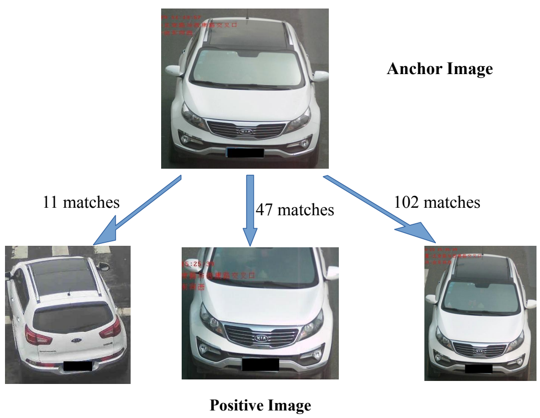
[2] Jie Hu, Li Shen, and Gang Sun. Squeeze-and-excitation networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7132–7141, 2018. 1

[3] Guangcong Wang, Jianhuang Lai, Peigen Huang, and Xiaohua Xie. Spatial-temporal person re-identification. In *Proceedings of the AAAI conference on artificial intelligence*, volume 33, pages 8933–8940, 2019. 2, 5

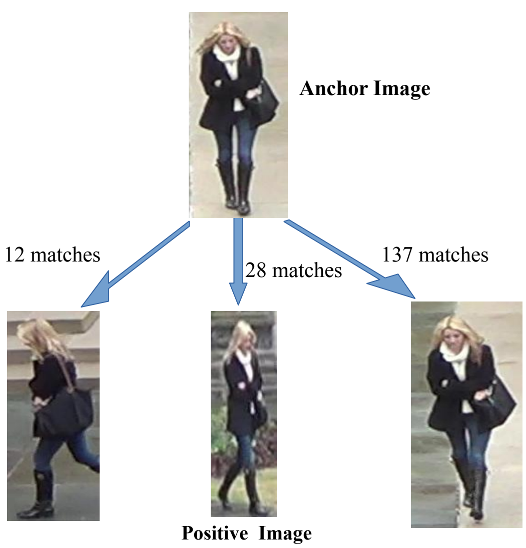
[4] Zhun Zhong, Liang Zheng, Donglin Cao, and Shaozi Li. Re-ranking person re-identification with k-reciprocal encoding. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1318–1327, 2017. 2



(a) Anchor-Positive Pair Selection by RPTM: Veri-776

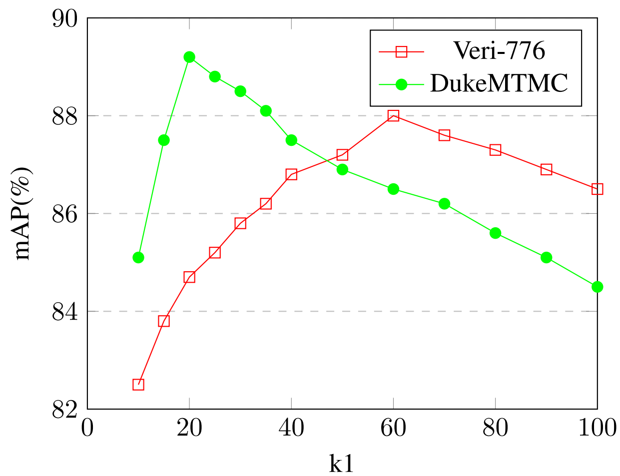


(b) Anchor-Positive Pair Selection by RPTM: VehicleID

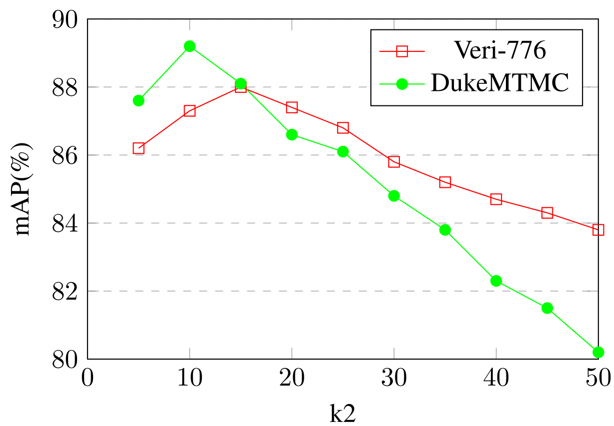


(c) Anchor-Positive Pair Selection by RPTM: DukeMTMC

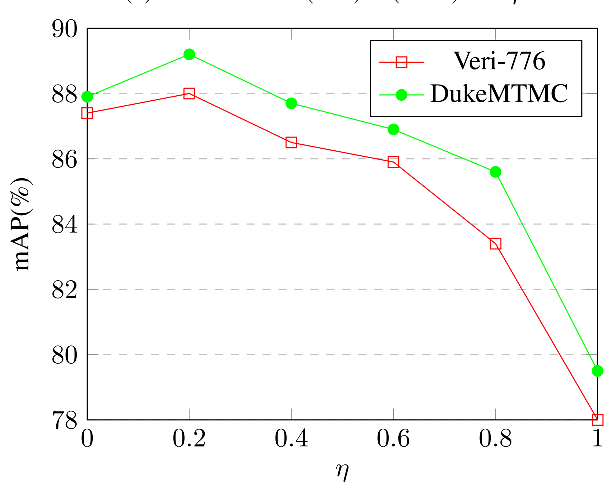
Figure 2: Visual representation of the Anchor-Positive Pair Selection methodology for Veri-776, VehicleID and DukeMTMC by the proposed RPTM algorithm, as explained in Section 4.2 of the main paper.



(a) We fix $k2$ at 15(Veri)/10(Duke) and η at 0.2.



(b) We fix $k1$ at 60(Veri)/20(Duke) and η at 0.2.



(c) We fix $k1$ at 60(Veri)/20(Duke) and $k2$ at 15(Veri)/10(Duke).

Figure 3: Insight into the variation of mAP results with the manipulation of re-ranking parameters $k1, k2$ and η for Veri-776(red) and DukeMTMC(green). Optimal values of parameters $k1$ and $k2$ vary with datasets, while the best mAP results are seen when η is set to 0.2.

Figure 4: Example results of two query images from the Veri-776 dataset. We compare the top-20 gallery retrieval results between our proposed RPTM model and current state-of-the-art for Veri-776, DMT [1]. RPTM shows relatively better results throughout the retrieval task, but more notably in the top-11 to top-20 retrievals.

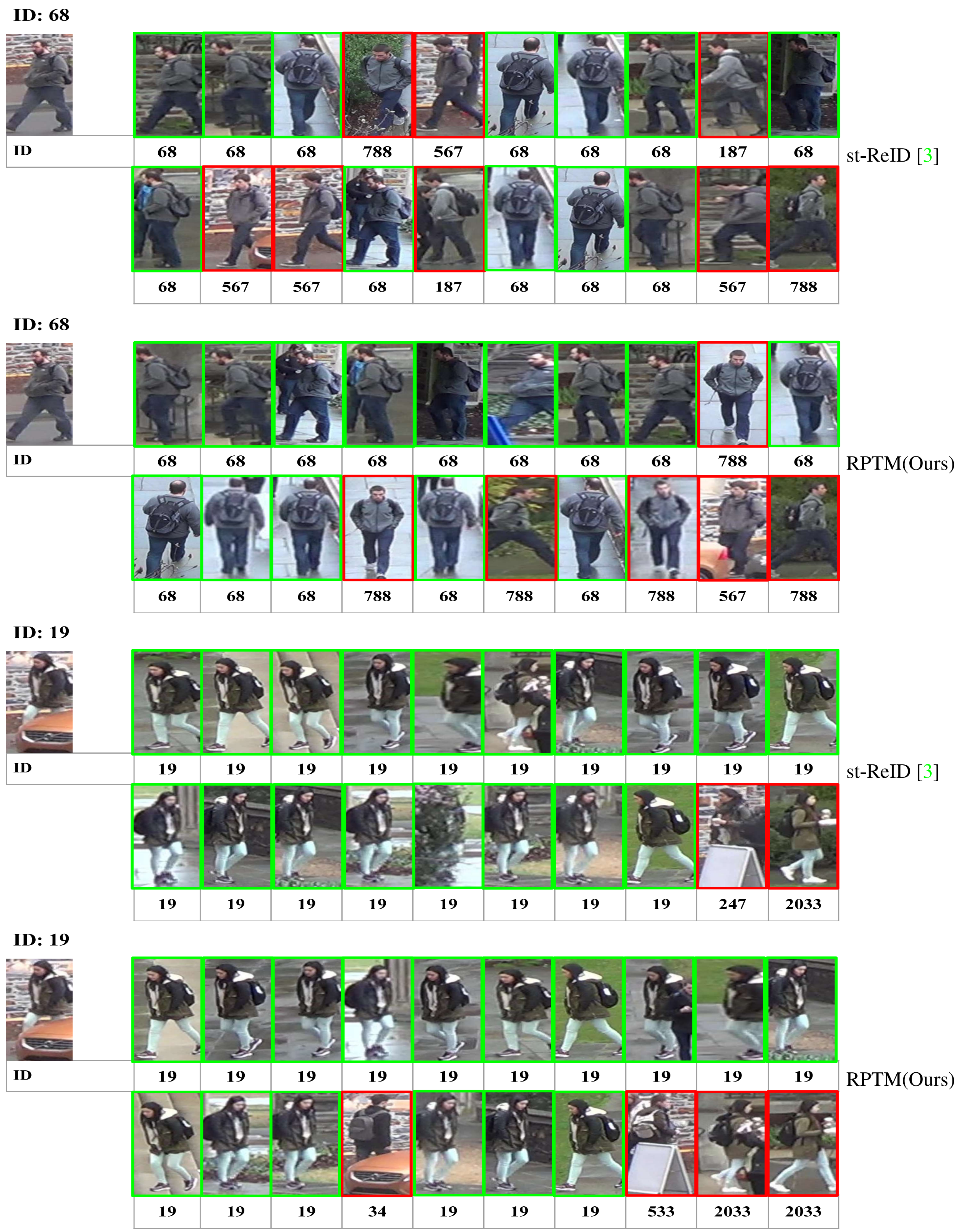


Figure 5: Top-20 retrieval results for two query images taken from DukeMTMC. We compare our RPTM model with the current state-of-the-art, st-ReID [3]. Despite st-ReID showing higher values of mAP, rank-1 and rank-5 results, and RPTM being fine-tuned for vehicle reID, RPTM shows equivalent retrieval results compared to st-ReID.