Relation Preserving Triplet Mining for Stabilising the Triplet Loss in Re-identification Systems

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

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

WACV WAIKOLOA HAWAII JAN 3-7 • 2023

Goal

- We present **Relation Preserving Triplet Mining (RPTM)**, a feature-matching guided scheme ensuring that triplets respect natural subgroupings within an object ID.
- We use this triplet mining mechanism to establish a pose-aware, well-conditioned triplet loss by implicitly enforcing view consistency.
- It allows us to keep the training pipeline simple, as a standard SGD to optimise a cost function, and is also well-conditioned enough to permit the use of constant training parameters across datasets.

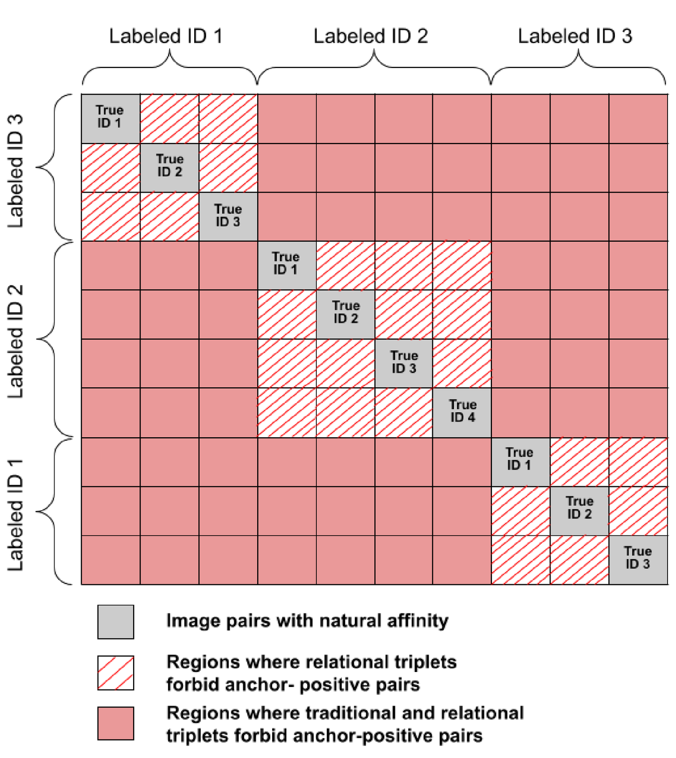
Motivation

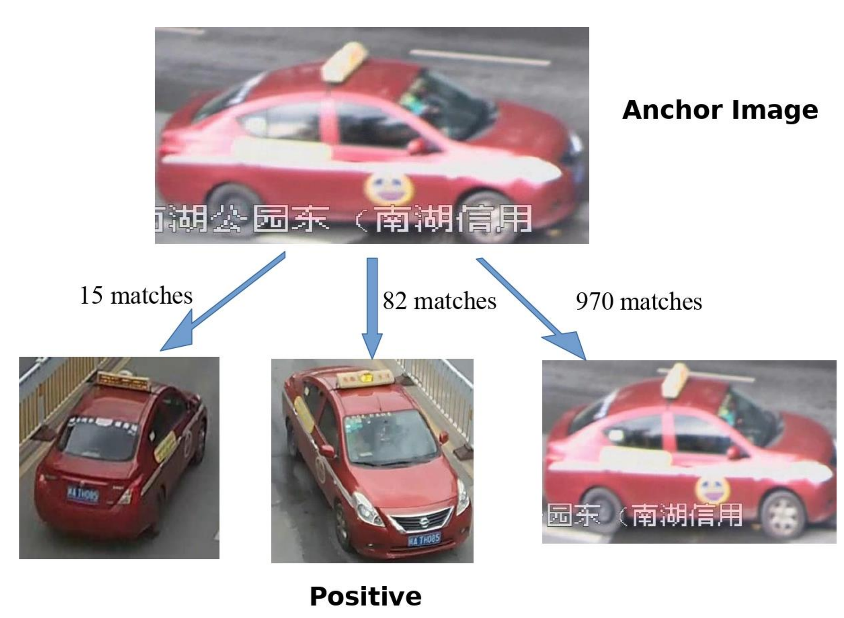
- Large intra-class variation
- Leads to bad triplet definitions
- Causes non-optimal and longer training scenarios
- Triplet Mining fails at pose awareness

+ Exploit internal groupings for selecting implicitly pose-aware anchor-positive maps
+ Better triplet mining scheme
+ Improved generalisation for reID

Relational Triplets

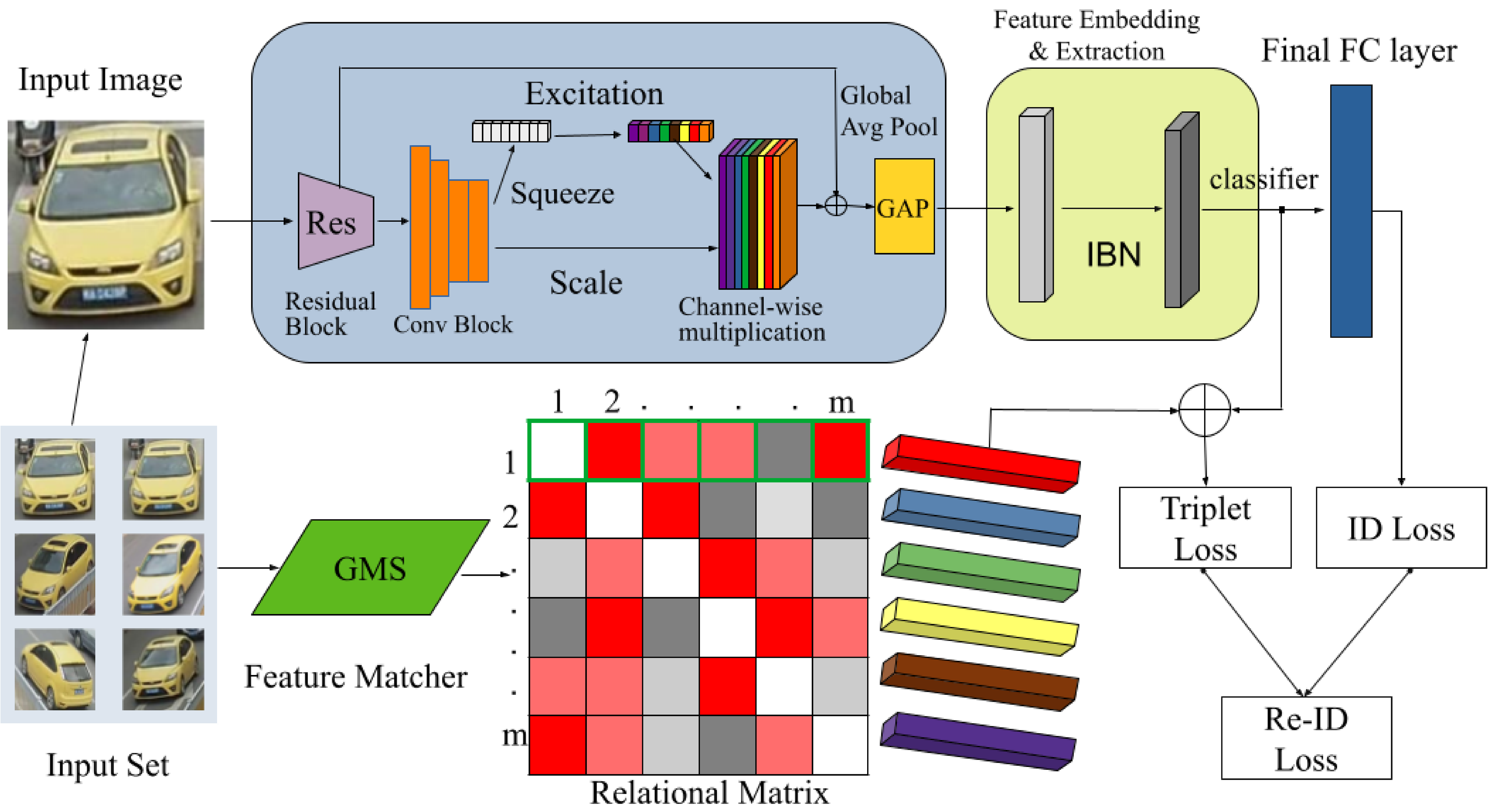
- Relational triplets change the triplet definition from one based on human assigned IDs to naturally occurring groups.
- The set of subsets is denoted by N = {S_m} and $\mathcal{S} = \bigcup_{\mathcal{S}_m \in \mathcal{N}} \mathcal{S}_m$. Whether two instances share a natural subset, we use the relational indicator $C(\mathrm{x}_i, \mathrm{x}_j) = \begin{cases} 1, & \text{if } \mathrm{x}_i, \mathrm{x}_j \text{ share a subset in } \mathcal{N}, \\ 0, & \text{otherwise.} \end{cases}$. Anchor-positive pairs share a common subset, negative does not.





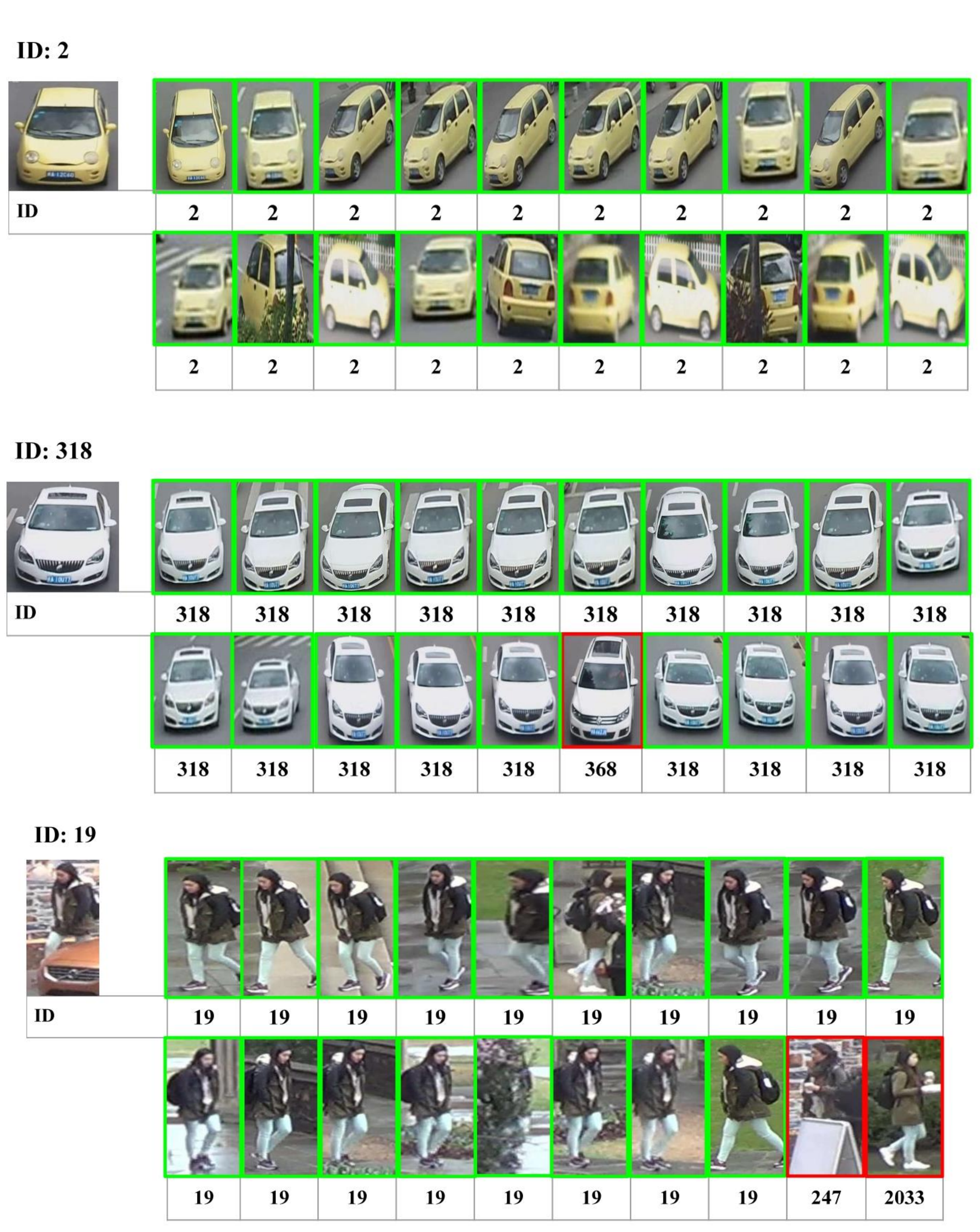
- **Class-adaptive Thresholding System:** positive images are chosen using a threshold τ which is the mean number of non-zero matching results.
- **RPTM:** semi-hard positive mining, anchor-positive pairs satisfy the relational indicator AND positive differs significantly from the anchor.
- **Implicitly Enforced View Consistency:** RPTM cleans up the triplet mining process with a triplet filtration step and prevents erroneous local minimas.
- Models with larger parameters can optimise just as fast as smaller networks.

Architecture



1. **Relational Matrix:** class-based image pair correspondences
   - **1.a. GMS[1]:** Grid-based Motion Statistics for a quick and effective feature matching strategy
2. **ResNet:** generate discriminative feature embeddings
3. **Squeeze-Excitation:** improve channel interdependencies at no extra computational cost
4. **IBN:** diverse features impact performance in proportion
5. **Re-ID Loss:** weighted sum of triplet loss and cross-entropy loss

Results: Qualitative

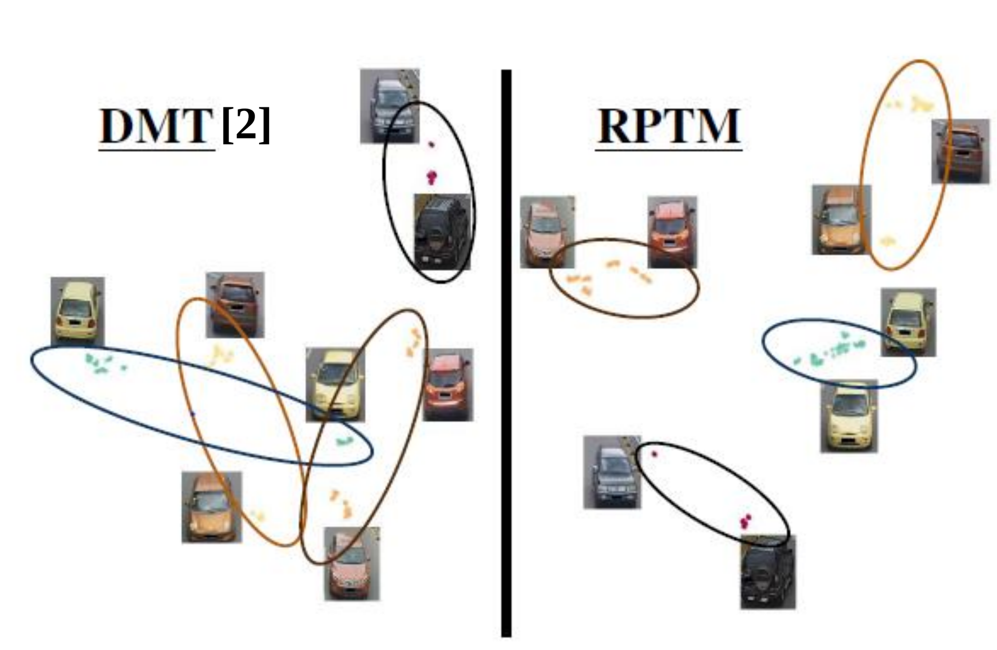


Quantitative

Dataset	mAP	r = 1	r = 5
Veri-776	**88.00**	**97.30**	**98.40**
VehicleID(800)	**84.80**	**95.50**	**97.40**
VehicleID(1600)	**81.20**	**93.30**	**96.50**
VehicleID(2400)	**80.50**	**92.90**	**96.30**
DukeMTMC	89.20	**93.50**	**96.10**

Table 1: Quantitative results of RPTM on the Veri-776, VehicleID and DukeMTMC datasets. We compare (bold being SoTA) with works using a ResNet101 backbone and re-ranking(Veri and Duke).

UMAP Visualisation



References

[1] Jia Wang Bian, et al. "Gms: Grid-based motion statistics for fast, ultra-robust feature correspondence.", CVPR 2017.

[2] Shuting He, et al. "Multi-domain learning and identity mining for vehicle re-identification.", CVPRW 2020.

Code & Contact

Code is available open-source

Contact:
Adhiraj Ghosh:
adhiraj.ghosh@student.uni-tuebingen.de