Resolution based Feature Distillation for Cross Resolution Person Re-Identification

Presenter: Asad Munir

Authors : Asad Munir, Chengjin Lyu, Bart Goossens, Wilfried Philips, Christian Micheloni

Affiliations: University of Udine (Machine Learning and Perception Lab)

TELIN-IPI, Ghent University - imec

Grant No. 765866

Person Re-Identification

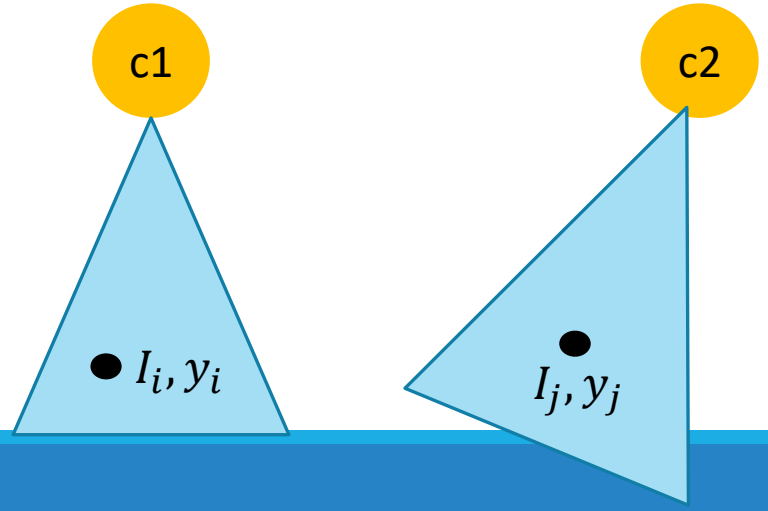
Given a dataset of N persons

$$D_{Tr} = \{I_k, y_k\}_{k=1}^{N}$$

Where I_k and y_k are the person image and ID of the k^{th} person in the dataset.

For a given pair of person images $\{I_i, I_j\}$, the task of re-identification is to compute weather

$$y_i = y_j \text{ or } y_i \neq y_j$$



Categories

- Person Re-identification.
- Cross dataset Person Re-identification.
- Unsupervised Person Re-Identification.
- Cross Modalities Person Re-identification.
- Cross Resolution Person Re-Identification.

Contributions

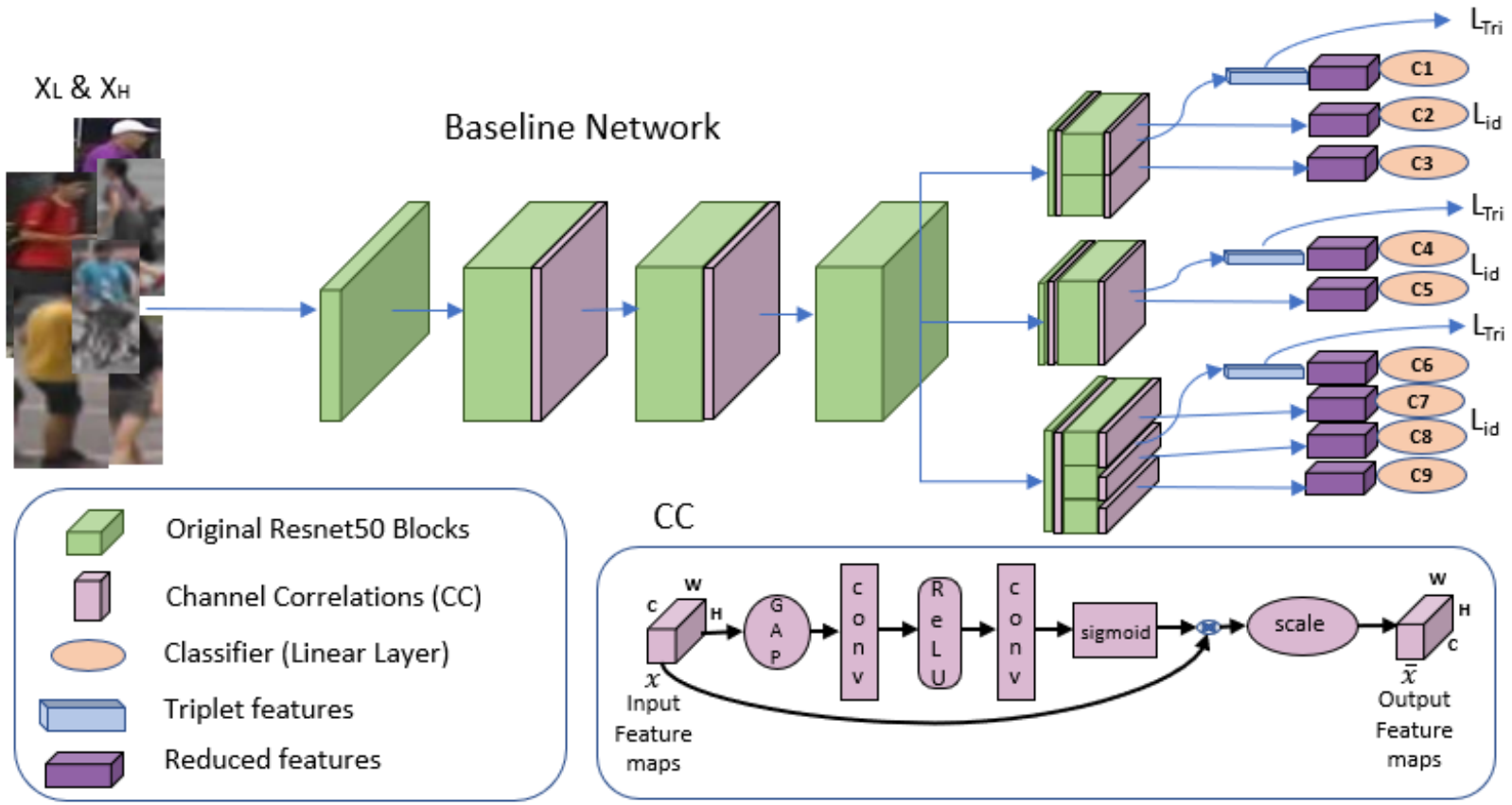
- Cross Resolution Person Re-Identification.
  - Improve the baseline network with the addition of multiple channels correlation modules to learn better representations from degraded data.
  - Propose a resolution based feature distillation with feature and resolution baselines to match the features of different resolution to perform the cross resolution re-id.
  - Adopt a new and more realistic scenario assuming that the gallery images are also collected in LR form. Thus, the LR query image is matched with multiple resolutions (HR and LR) instead of a single resolution (HR).

Proposed Method

- Learning Mechanism from multiple resolutions degraded data
  - MGN based baseline with modifications.
  - Channels Correlation (Channel attention modules)
  - Different placement from the original channel attention networks.

- Resolution based features distillation (RFD) mechanism.
  - Train the first baseline with original labels while the second by using resolutions as labels.
  - Pseudo labeling for the dataset with real degradations.
  - ID and Triplet losses

Proposed Network



Proposed RFD

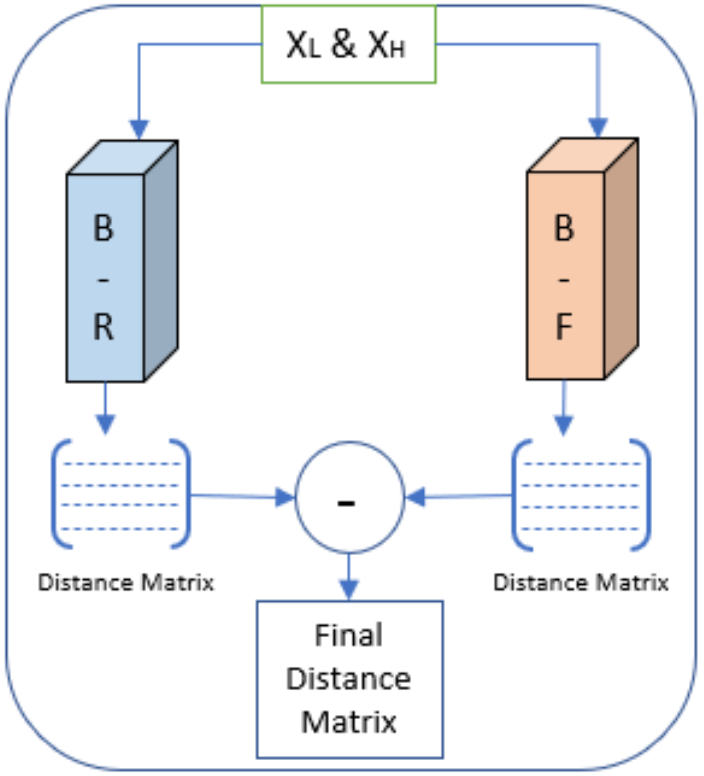
- Distance matrix calculated as cosine similarity

$$D_r(x, y) = \frac{f(x).f(y)}{\|f(x)\|\|f(y)\|}$$

- Final Distance matrix is

$$D(x, y) = D_f(x, y) - \alpha D_r(x, y)$$

- Pseudo Labeling for resolutions
  - Down sampling scales used as labels to train.
  - We divide these number of pixels into 5 patches of equal length.



Experiments

Datasets:

- **MLR-Market**
  - The dataset consists of images from 6 cameras
  - 12936 images in training set with 751 identities
  - 19281 images in testing set with 750 identities (3368 queries)
- **MLR-Duke**
  - The dataset consists of images from 8 cameras
  - 16522 images in training set with 702 identities
  - 17661 images in testing set with 702 identities (2228 queries)
- **CAVIAR**
  - A real multi resolution dataset and is composed of 1220 images of 72 persons taken from 2 different cameras. We discard 22 identities those appear in one camera and split the dataset into two non-overlapping halves.

Results

- Comparisons of the proposed method with the state-of-the-art re-id methods on MLR-Market dataset. The best and second best results are highlighted in red and blue respectively.
- The results on baseline used and other proposed baselines are shown below the dashed line.

Methods	Reference	MLR-Market		
		R1	*R5*	*R10*
SING [13]	AAAI18	74.4	87.8	91.6
CSR-GAN [27]	IJCAI18	76.4	88.5	91.9
CamStyle [32]	CVPR18	74.5	88.6	93.0
FD-GAN [8]	NeurIPS18	79.6	91.6	93.5
RIPR [19]	IJCAI19	66.9	84.7	-
CRGAN [17]	ICCV19	83.7	92.7	95.8
MSA [1]	IEEE Access20	68.3	85.7	-
INTACT [6]	CVPR20	88.1	95.0	96.9
PRI [9]	ECCV20	88.1	94.2	96.5
baseline	-	84.1	92.3	95.9
baseline B-F	-	85.5	94.1	96.0
(B-F+RFD) Proposed	-	86.9	95.6	97.4

Results

- Results and comparisons of the proposed method with the state-of-the-art re-id methods on MLR-Duke dataset. The best and second best results are highlighted in red and blue respectively.
- The results on baseline used and other proposed baselines are shown below the dashed line.

Methods	Reference	MLR-Duke		
		R1	*R5*	*R10*
SING [13]	AAAI18	65.2	80.1	84.8
CSR-GAN [27]	IJCAI18	67.6	81.4	85.1
CamStyle [32]	CVPR18	64.0	78.1	84.4
FD-GAN [8]	NeurIPS18	67.5	82.0	85.3
CRGAN [17]	ICCV19	75.6	86.7	89.6
MSA [1]	IEEE Access20	79.1	90.0	-
INTACT [6]	CVPR20	81.2	90.1	92.8
PRI [9]	ECCV20	82.1	91.1	92.8
baseline	-	81.2	90.1	91.9
baseline B-F	-	82.0	90.8	92.7
(B-F+RFD) Proposed	-	82.9	92.0	94.0

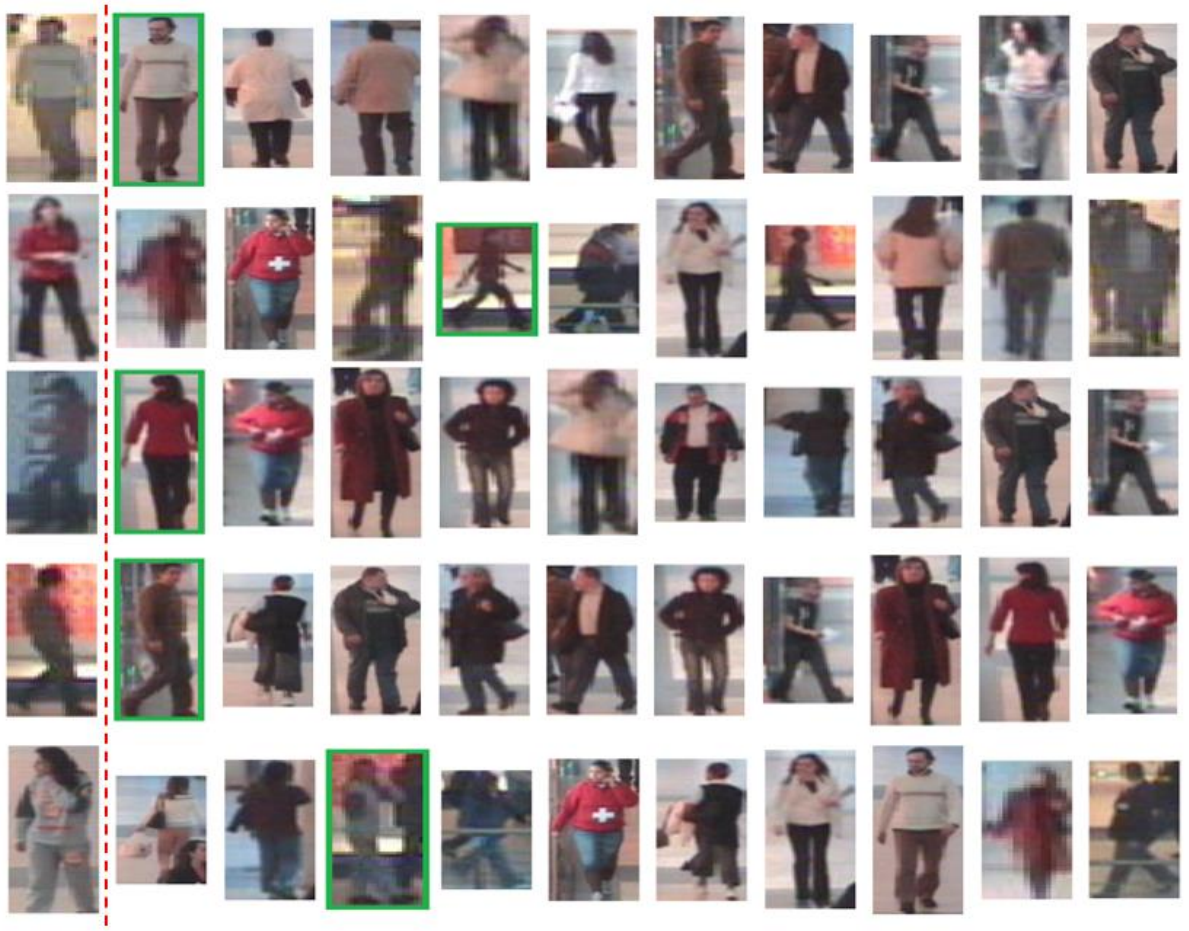
Results

- Comparisons of the proposed method with the state-of-the-art re-id methods on the real world multi resolution CAVIAR dataset. The best and second best results are highlighted in red and blue respectively.
- The results on baseline used and other proposed baselines are shown below the dashed line.

Methods	Reference	CAVIAR		
		R1	*R5*	*R10*
SING [13]	AAAI18	33.5	72.7	89.0
CSR-GAN [27]	IJCAI18	34.7	72.5	87.4
CamStyle [32]	CVPR18	32.1	72.3	85.9
FD-GAN [8]	NeurIPS18	33.5	71.4	86.5
RIPR [19]	IJCAI19	36.4	72.0	-
CRGAN [17]	ICCV19	42.8	76.2	91.5
INTACT [6]	CVPR20	44.0	81.8	93.9
PRI [9]	ECCV20	45.2	84.1	94.6
baseline	-	42.1	87.6	92.9
baseline B-F	-	44.3	88.4	94.2
(B-F+RFD) Proposed	-	47.6	89.2	96.0

Visual Results

- Visual results extracted by the proposed method. Query image with its first ten matches in the gallery for CAVIAR dataset.
- Red dashed line separating the probe images from the ranked gallery images.
- We did not resize the images to show multiple resolutions in the ranking list.



Ablation Study

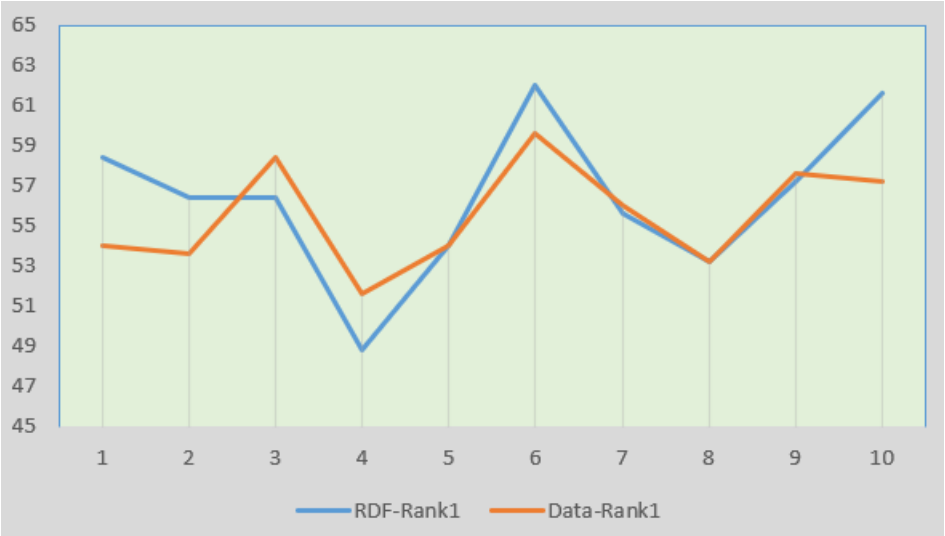
Components	MLR-Market		MLR-Duke	
	R1	*R5*	*R1*	*R5*
Bicubic	83.1	96.0	82.4	92.0
Bilinear	82.9	95.9	82.2	92.1
Single-Reso	85.5	94.1	82.0	90.8
Multi-Reso	84.2	93.8	81.5	90.0
Multi-Reso+RFD	85.6	94.4	83.0	91.8

Table 4. Ablation study of the proposed method on two synthetic datasets MLR-Market and MLR-Duke.

Components	CAVIAR		
	R1	*R5*	*R10*
Single-Reso	44.3	88.4	94.2
Multi-Reso	55.5	90.6	91.5
multi-Reso+RFD	56.4	90.6	92.2

Table 5. Results of the proposed method for single and multi resolution gallery.

The effect of proposed pseudo labeling on CAVIAR dataset. Random splits of the dataset are shown on horizontal axis with their performance on vertical axis.



Conclusion

- We proposed a resolution based feature distillation (RFD) approach for cross resolution person re-identification.
- We improved a baseline by means of channels correlation to solve the general low resolution person re-id problem.
- We proposed a resolution based feature distillation technique which filters out the resolution dependent features to compute the final distance matrix for matching.
- A pseudo labelling technique for computing the resolution label is also introduced to train the RFD.

This work was supported by EU H2020 MSCA through Project ACHIEVE-ITN (Grant No 765866)

THANK YOU

Questions?

asad.munir@uniud.it