Person Re-IDentification

Final presentation
Third milestone

UNIVERSITÀ DI TRENTO

Bonomi Andrea - Ismail Khouloud - Laiti Francesco
Lobba Davide - Turri Evelyn

Trends and Applications of Computer Vision
Academic Year 2022/2023

Introduction

What is Person Re-ID?

(a) Pedestrian Detection

(b) Person Re-identification

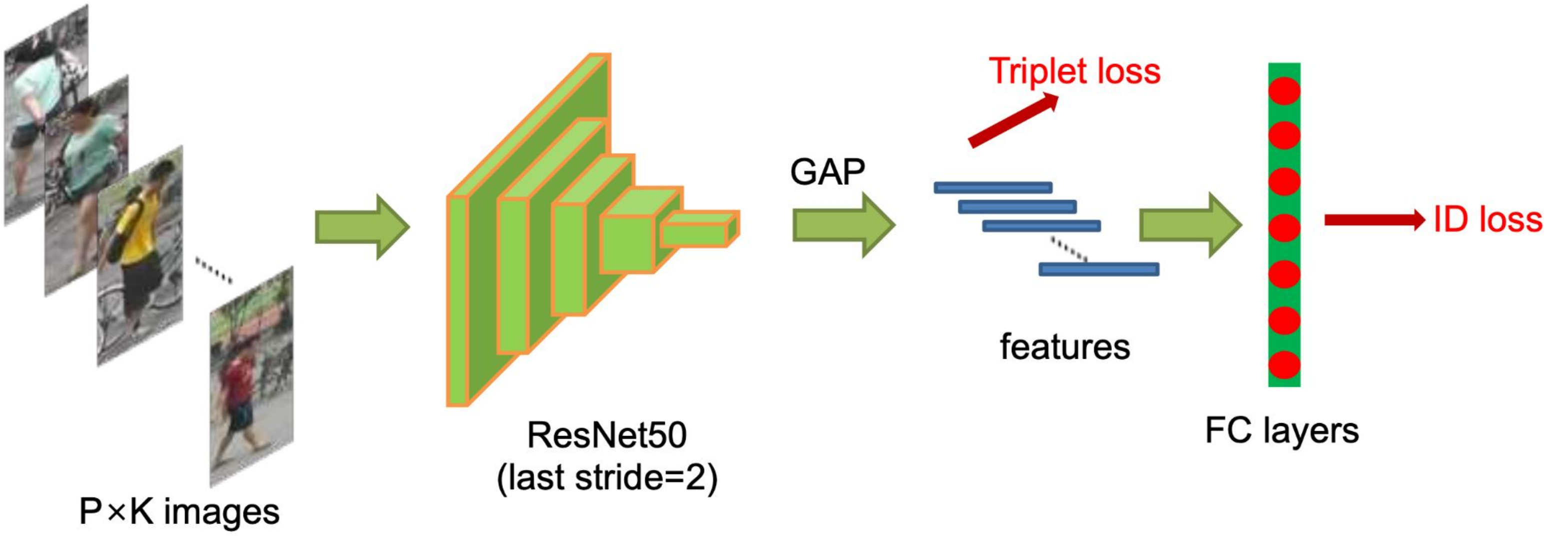
Why Person Re-ID is relevant

Applications

- Video surveillance for crimes
- Help in identify and locate missing persons
- Build smart cities
- Monitor the movement of people in crowded areas to prevent accidents or other incidents

Theoretical Background

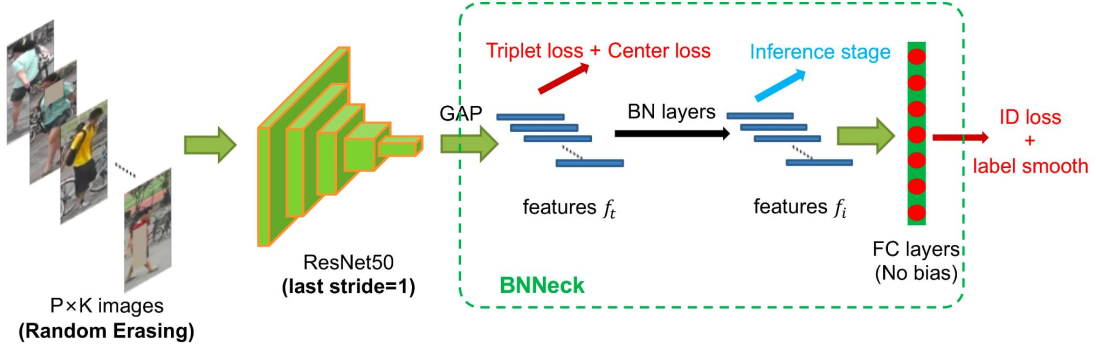
Deep Person Re-ID



github.com/Cysu/open-reid

Bag of Tricks

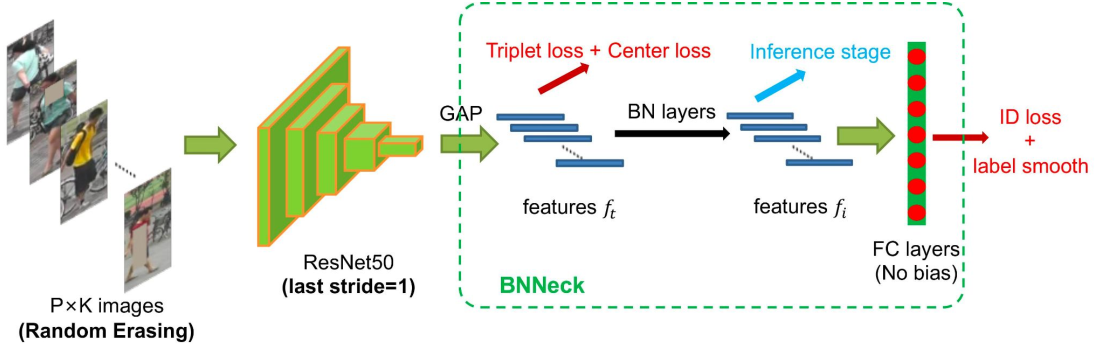
- Important baseline for Person Re-ID
- Old baseline + 6 tricks
- Main tricks : BNNeck and Center Loss



github.com/michuanhaohao/reid-strong-baseline

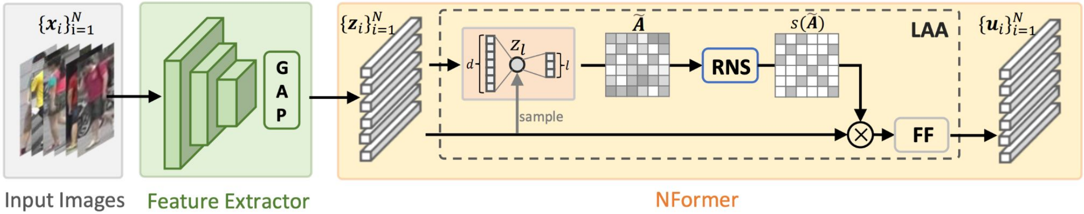
Bag of Tricks | Tricks

- Warmup Learning Rate
- Random Erasing Augmentation (REA)
- Label Smoothing
- Last Stride
- BNNeck
- Center Loss



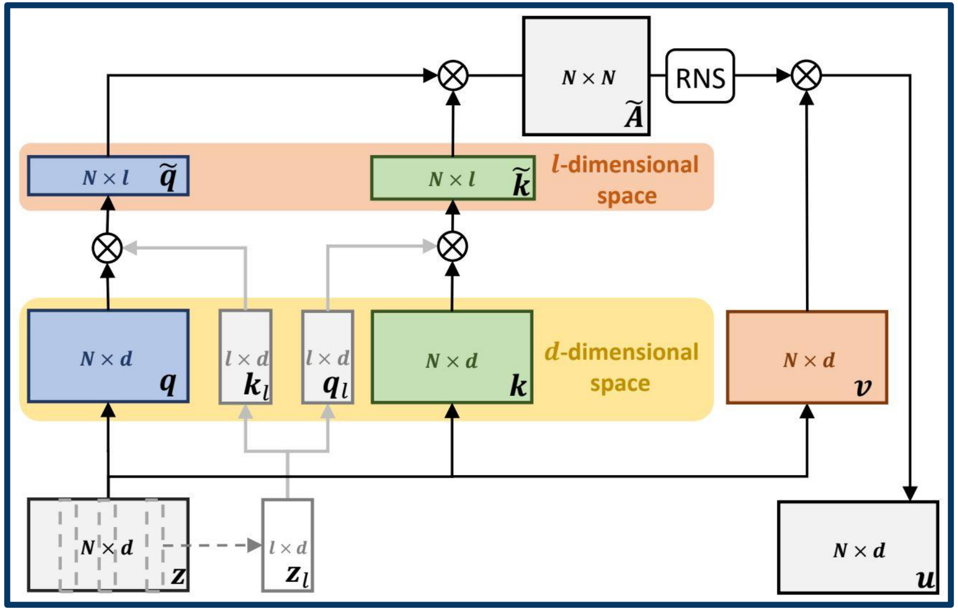
NFormer

- Extraction of combined features
- A new softmax that takes into account only the closest neighbors
- Improve the performances in case of occlusions or problems as change color clothes

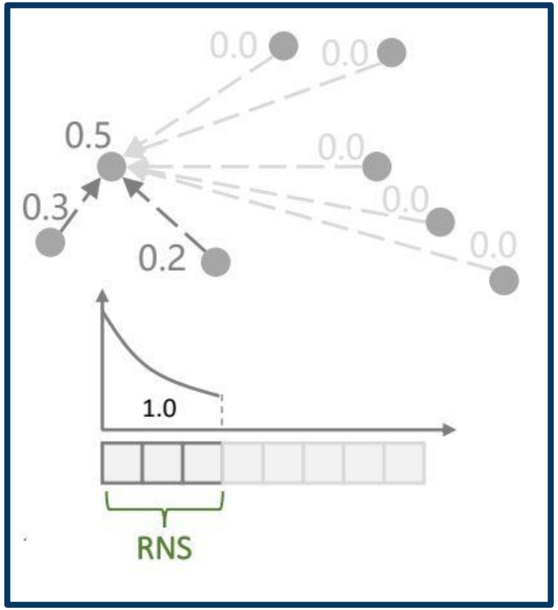


github.com/haochenheheda/NFormer

NFormer | LAA & RNS



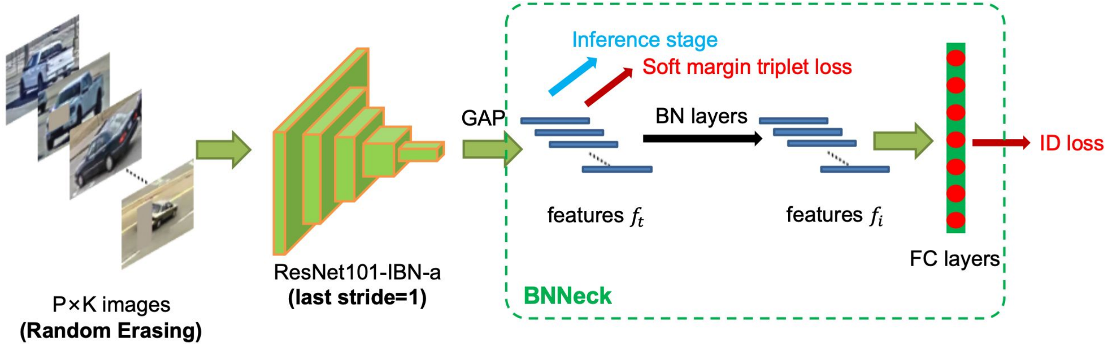
Landmark Agent Attention



Reciprocal Neighbor Softmax

Vehicle Re-ID

- Bag of Tricks architecture with some changes
- Changes : ResNet101, Inference stage moved before and loss moved before BN layers



github.com/heshuting555/AICITY2020_DMT_VehicleReID.git

Results

Results | Person Re-ID with masked faces

Without mask

With mask

Results | Person Re-ID with masked faces

Without masked query

mAP	Rank-1	Rank-5	Rank-10
84.4 %	93.6 %	98.0 %	98.8 %

With masked query

mAP	Rank-1	Rank-5	Rank-10
84.4 %	93.6 %	98.0 %	98.8 %

Results | Person Re-ID with masked faces

Without masked query	mAP	Rank-1	Rank-5	Rank-10
	84.4 %	93.6 %	98.0 %	98.8 %

=

With masked query	mAP	Rank-1	Rank-5	Rank-10
	84.4 %	93.6 %	98.0 %	98.8 %

Results | NFormer Vehicles Re-ID

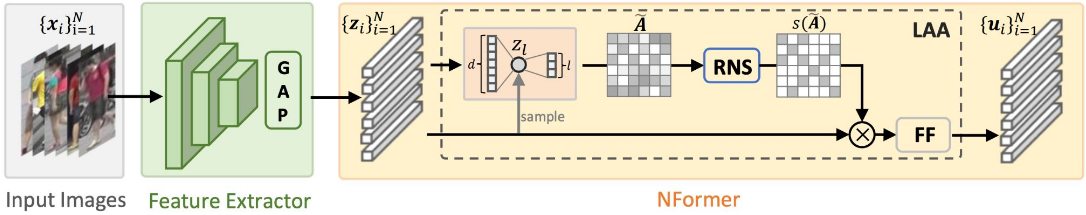
Person Re-ID		Vehicles Re-ID
BoT-BS	→	Multi-Domain Learning and Identity Mining for Vehicle Re-Identification
NFormer	→	?

Lack of vehicle datasets

- Datasets on Vehicle-ReId are not many
- A lot of datasets needs an agreement even if license plates are masked

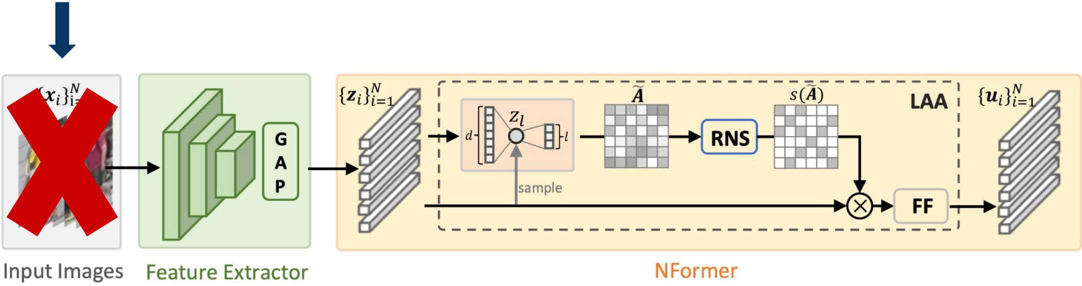
Vehicle ReID	
Dataset	**Public Availability**
VeRi-776	Not Available
VehicleReId	Not Available
VRIC	Available

Results | NFormer Vehicles Re-ID without changes

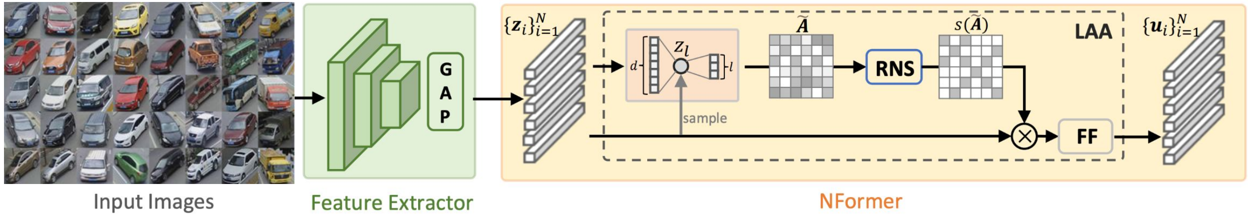


Results | NFormer Vehicles Re-ID without changes

Veri-776



Results | NFormer Vehicles Re-ID without changes



mAP	Rank-1	Rank-5	Rank-10
76.5 %	95.1 %	97.3 %	98.3 %

Results | NFormer Vehicles Re-ID with changes

Description	mAP	Rank-1	Rank-5	Rank-10
ResNet-50 + Adam + S Triplet L. w/ Center	76.5 %	95.1 %	97.3 %	98.3 %
ResNet-101 + Adam + S Triplet L. w/ Center	76.1 %	94.5 %	97.6 %	98.4 %
ResNet-101 + Adam + Soft-Margin Triplet Loss	71.3 %	93.4 %	97.1 %	98.5 %
ResNet-101 + SGD* + Soft-Margin Triplet Loss	69.4 %	93.7 %	97.2 %	97.9 %
ResNet-50 + Adam + Soft-Margin Triplet Loss	73.1 %	93.9 %	97.7 %	98.7 %

NFormer without changes

* Learning rate adjusted accordingly

Results | NFormer Vehicles Re-ID with changes

Description	mAP	Rank-1	Rank-5	Rank-10
ResNet-50 + Adam + S Triplet L. w/ Center	76.5 %	95.1 %	97.3 %	98.3 %
ResNet-101 + Adam + S Triplet L. w/ Center	76.1 %	94.5 %	97.6 %	98.4 %
ResNet-101 + Adam + Soft-Margin Triplet Loss	71.3 %	93.4 %	97.1 %	98.5 %
ResNet-101 + SGD* + Soft-Margin Triplet Loss	69.4 %	93.7 %	97.2 %	97.9 %
ResNet-50 + Adam + Soft-Margin Triplet Loss	73.1 %	93.9 %	97.7 %	98.7 %

ResNet-101 does not improve

○ NFormer without changes
* Learning rate adjusted accordingly

Results | NFormer Vehicles Re-ID with changes

Description	mAP	Rank-1	Rank-5	Rank-10
ResNet-50 + Adam + S Triplet L. w/ Center	76.5 %	95.1 %	97.3 %	98.3 %
ResNet-101 + Adam + S Triplet L. w/ Center	76.1 %	94.5 %	97.6 %	98.4 %
ResNet-101 + Adam + Soft-Margin Triplet Loss	71.3 %	93.4 %	97.1 %	98.5 %
ResNet-101 + SGD* + Soft-Margin Triplet Loss	69.4 %	93.7 %	97.2 %	97.9 %
ResNet-50 + Adam + Soft-Margin Triplet Loss	73.1 %	93.9 %	97.7 %	98.7 %

SGD vs Adam

NFormer without changes

* Learning rate adjusted accordingly

Results | NFormer Vehicles Re-ID with changes

Description	mAP	Rank-1	Rank-5	Rank-10
ResNet-50 + Adam + S Triplet L. w/ Center	76.5 %	95.1 %	97.3 %	98.3 %
ResNet-101 + Adam + S Triplet L. w/ Center	76.1 %	94.5 %	97.6 %	98.4 %
ResNet-101 + Adam + Soft-Margin Triplet Loss	71.3 %	93.4 %	97.1 %	98.5 %
ResNet-101 + SGD* + Soft-Margin Triplet Loss	69.4 %	93.7 %	97.2 %	97.9 %
ResNet-50 + Adam + Soft-Margin Triplet Loss	73.1 %	93.9 %	97.7 %	98.7 %

No improvements with
Soft-Margin Triplet Loss

NFormer without changes
* Learning rate adjusted accordingly

Results | NFormer Vehicles Re-ID with changes

Description	mAP	Rank-1	Rank-5	Rank-10
ResNet-50 + Adam + S Triplet L. w/ Center	76.5 %	95.1 %	97.3 %	98.3 %
ResNet-101 + Adam + S Triplet L. w/ Center	76.1 %	94.5 %	97.6 %	98.4 %
ResNet-101 + Adam + Soft-Margin Triplet Loss	71.3 %	93.4 %	97.1 %	98.5 %
ResNet-101 + SGD* + Soft-Margin Triplet Loss	69.4 %	93.7 %	97.2 %	97.9 %
ResNet-50 + Adam + Soft-Margin Triplet Loss	73.1 %	93.9 %	97.7 %	98.7 %

No improvements with Soft-Margin Triplet Loss

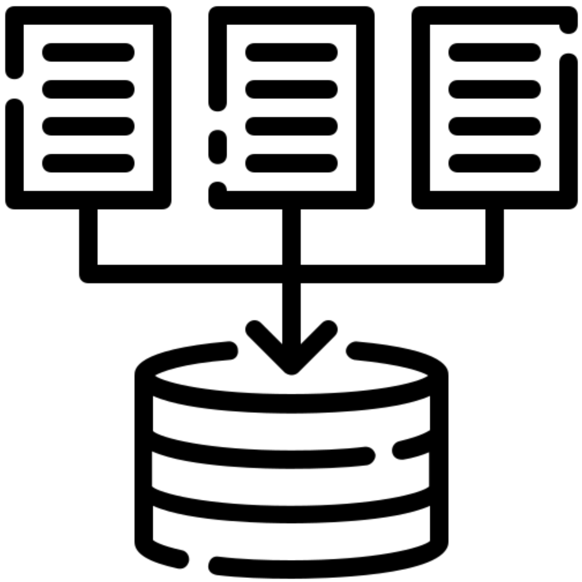
NFormer without changes

* Learning rate adjusted accordingly

Drawbacks

Ethical problems

- Lack of consent
- Misuse of personal data
- Unfair targeting
- Loss of anonymity

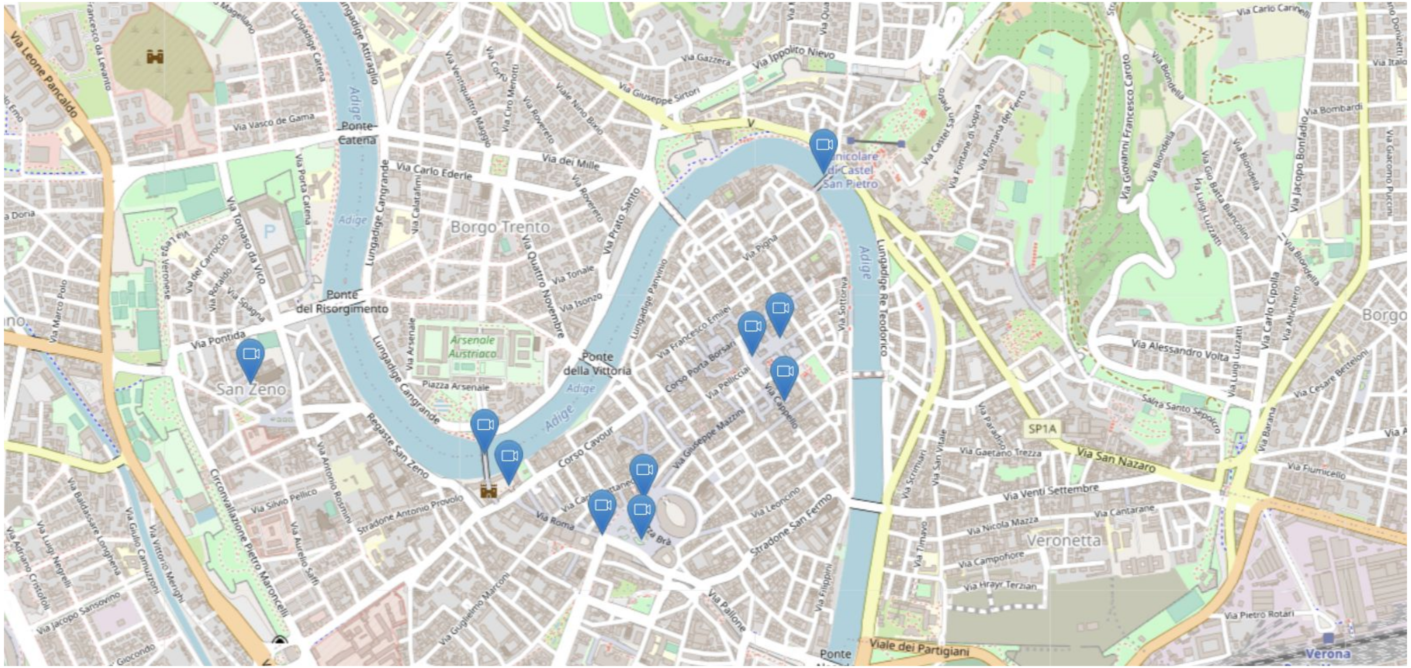


Mass surveillance

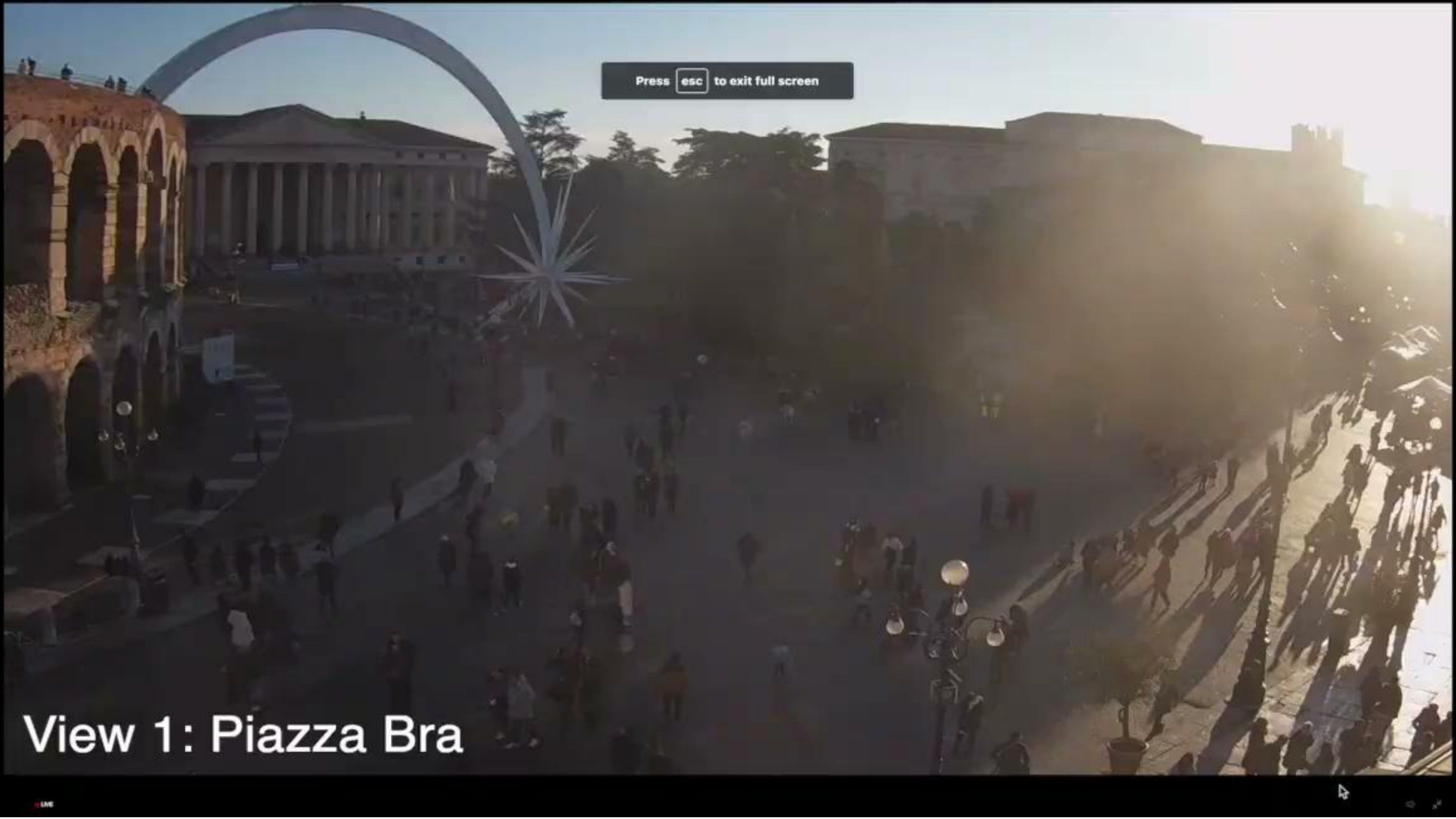
Stalking

End-to-end person Re-ID system from live cameras available to the public

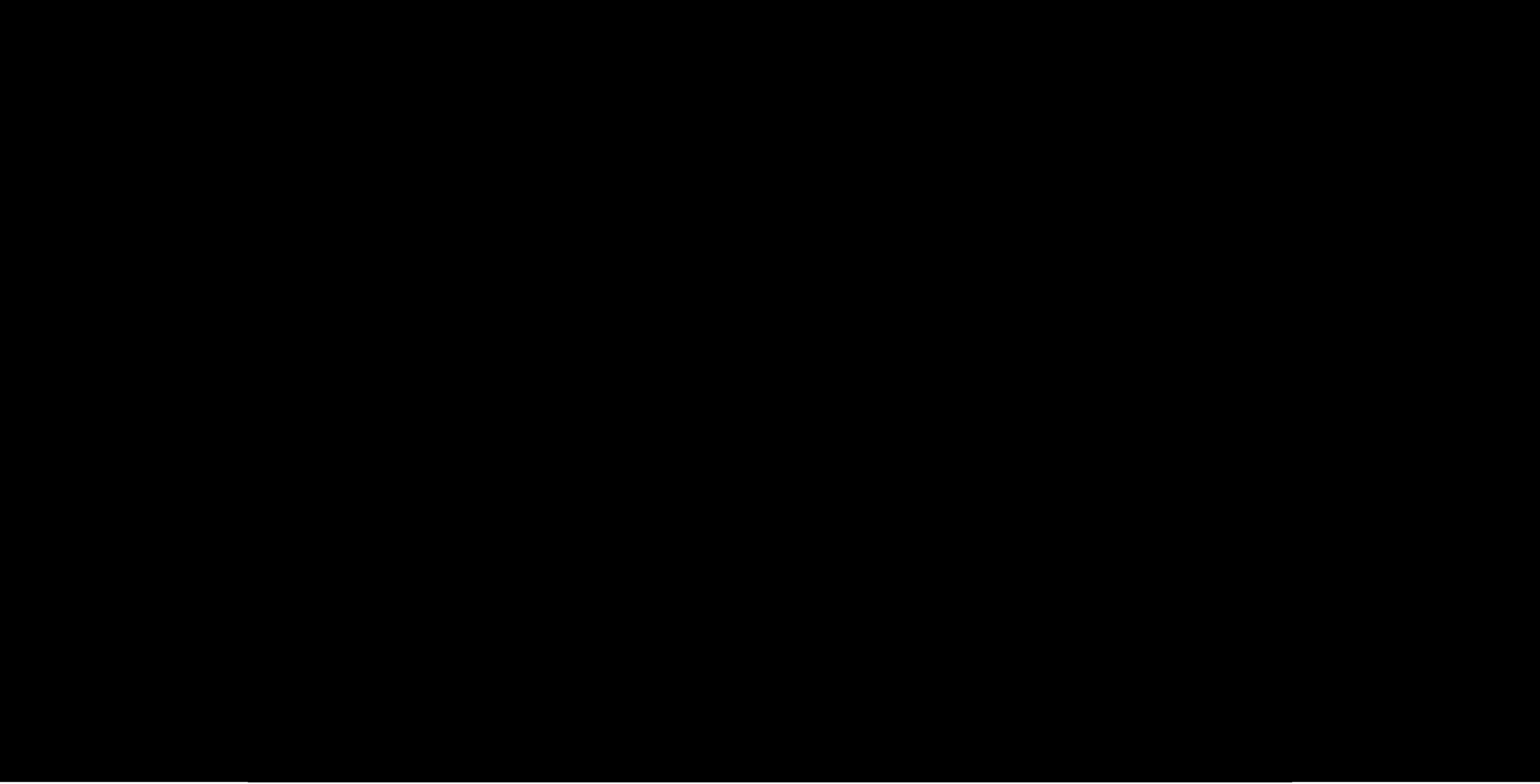
Live webcam



Camera positions in Verona



Different live view of Verona from webcam.comune.verona.it/



4K resolution. Live view of Times Square, New York, from youtube.com

Visual comparison with Market dataset

Visual comparison with Market dataset

Market dataset

Verona dataset
created on the fly

Times Square dataset
created on the fly

Outlook

Outlook | Open Issues

- Poorly generalized global model
- Scalability model issues
- Multiple heterogeneous modalities
- Occlusions
- Clothes changing
- Human Annotation Minimization

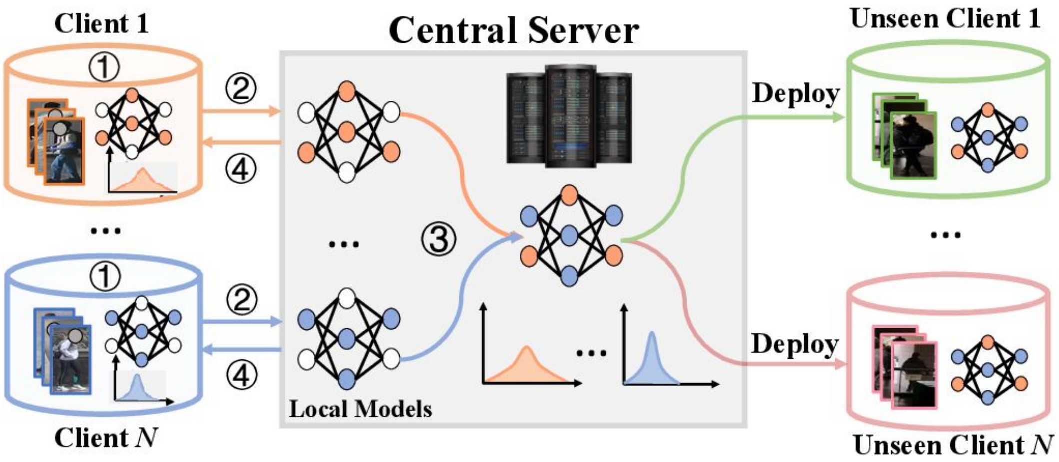
Federated average learning

Problem

Privacy Issues
Real-world applications’ limitation

Solution

Training local models individually and averaging them to a global model, for deploying in unseen target domains



1- Local training
2- Client-to-server updating
3- Server-side aggregation
4- Redistributing

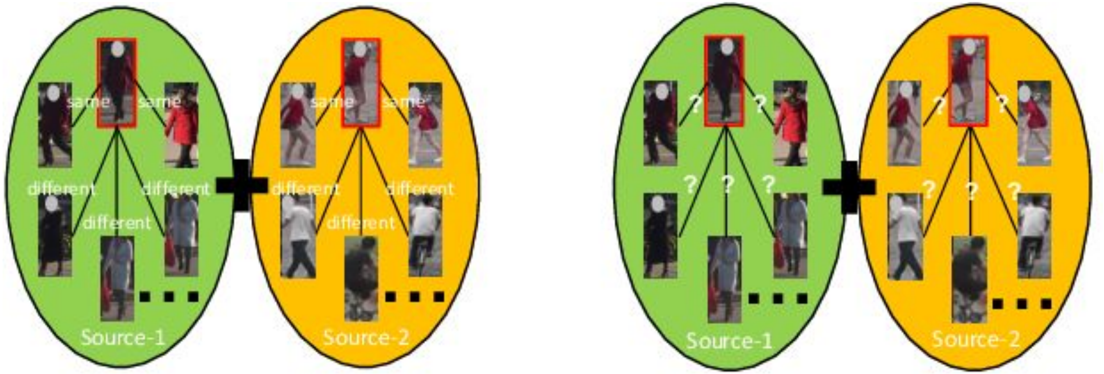
Domain Generalization

Problem

Local models overfit local data
Poorly-generalized global model

Solution

Optimizing re-ID models with several source domains and locally fine-tuning or directly deploying the obtained model to target domain



(a) Typical setting (Supervised)   (b) Our setting (Unsupervised)

Source Domain   Target Domain

Disjoint label space in the training and testing set

Domain Generalization

Other Problems

- Model Adaptation to New Domain/Camera
- Model Updating with Newly Arriving Data

Solutions

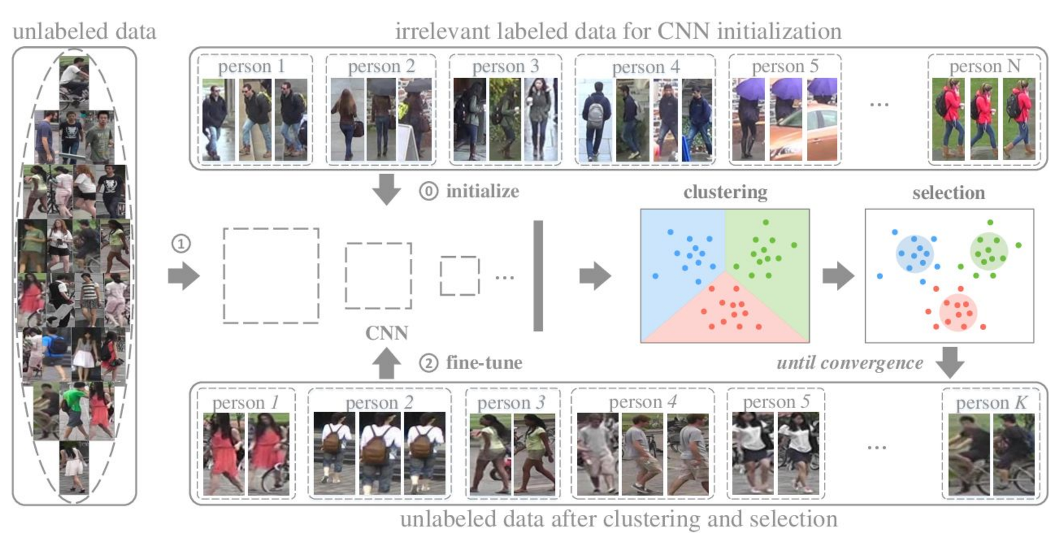
- Domain Adaptation (DA)
- Domain and Feature Hallucinating (DFH)

Domain Adaptation

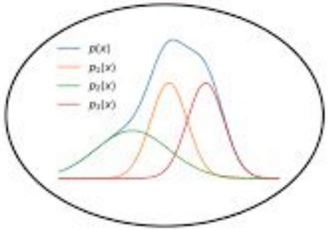
Transferring pre-trained deep representations to unseen domains

★ Initialize the model

1. Extract features for unlabeled data
2. Clustering and assign pseudo-labels based on centers
3. Select reliable pseudo-labeled samples
4. Fine-tune the model with pseudo-labeled data
5. Repeat 1-4 until convergence.



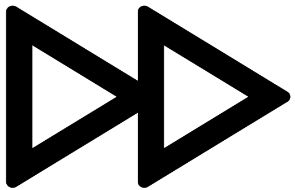
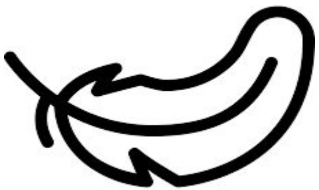
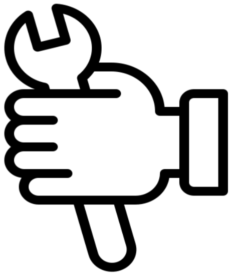
Domain and Feature Hallucinating

	FedAvg	DFH
Local Training		Transform local features to another domain's distribution through scaling and shifting
Server-side Aggregation		Aggregation of local models with weighted average

Outlook | Open Issues

- Poorly generalized global model
- Scalability model issues
- Multiple heterogeneous modalities
- Occlusions
- Clothes changing
- Human Annotation Minimization

Efficient models

Fast Re-ID	Lightweight model	Resource Aware Re-ID
		

Efficient models

Fast Re-ID | Lightweight model | Resource Aware Re-ID

- Hashing
- Transform high-dim to compact-dim
- In study: mix of long and short binary codes

Efficient models

Fast Re-ID

Lightweight model

Resource Aware Re-ID

- Change the network
- Model distillation (Teacher-student network)

Efficient models

Fast Re-ID

Lightweight model

Resource Aware Re-ID

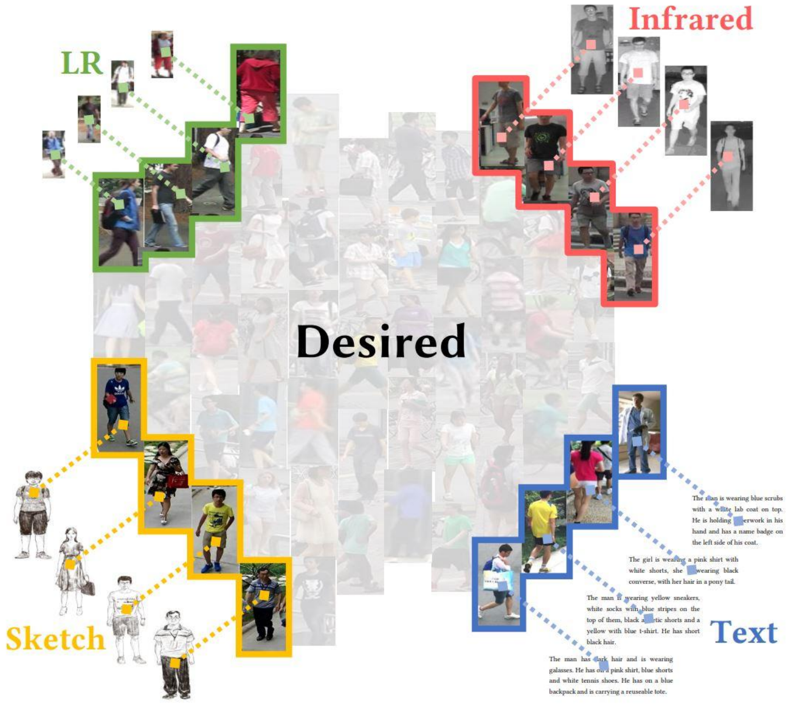
Adjust model to the hardware in use

Outlook | Open Issues

- Poorly generalized global model
- Scalability model issues
- **Multiple heterogeneous modalities**
- Occlusions
- Clothes changing
- Human Annotation Minimization

Multiple heterogeneous modalities

- Visual and text modalities
- Multiple visual modalities
- Visual and audio modalities



Wang, Z., Wang, Z., Zheng, Y., Wu, Y., Zeng, W., & Satoh, S. (2019). Beyond Intra-modality: A Survey of Heterogeneous Person Re-identification. *arXiv*. https://doi.org/10.48550/arXiv.1905.10048

Multiple heterogeneous modalities

- Multi stream neural network
- Feature-level fusion
- Decision-level fusion

Demonstration of RGB, NI (Near Infrared) and TI (Thermal Infrared) multi-modality person re-identification.

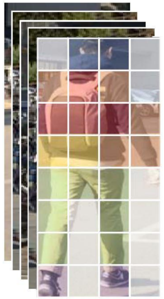
Zheng, A., Wang, Z., Chen, Z., Li, C., & Tang, J. (2021). Robust Multi-Modality Person Re-identification. *Proceedings of the AAAI Conference on Artificial Intelligence*, *35*(4), 3529-3537. https://doi.org/10.1609/aaai.v35i4.16467

Text-to-Image Person Re-ID (TIRe-ID)

Query

A young man with short black hair is wearing a blue hooded jacket over a white t-shirt. He is also wearing beige fitted pants and grey sneakers with a white design and soles, He is carrying a grey backpack with a black patch.

Dataset



CFine

CLIP-driven Fine-grained information excavation framework

S. Yan, N. Dong, L. Zhang, and J. Tang, "Clip-driven fine-grained text-image person re-identification," 2022. [Online]. Available: https://arxiv.org/abs/2210.10276

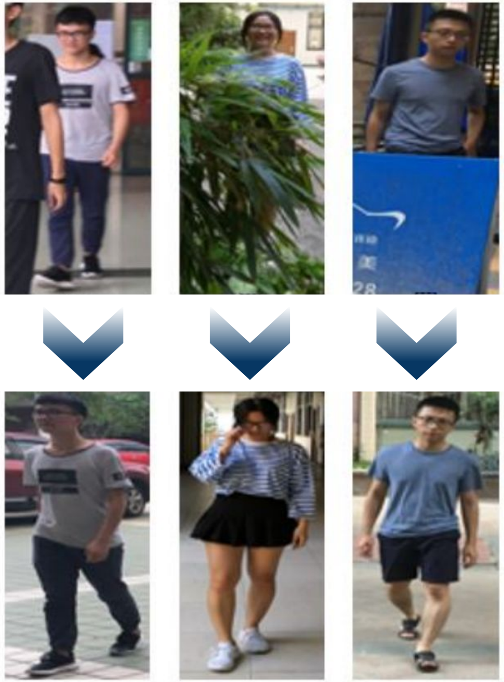
Outlook | Open Issues

- Poorly generalized global model
- Scalability model issues
- Multiple heterogeneous modalities
- Occlusions
- Clothes changing
- Human Annotation Minimization

Occlusions

Probe

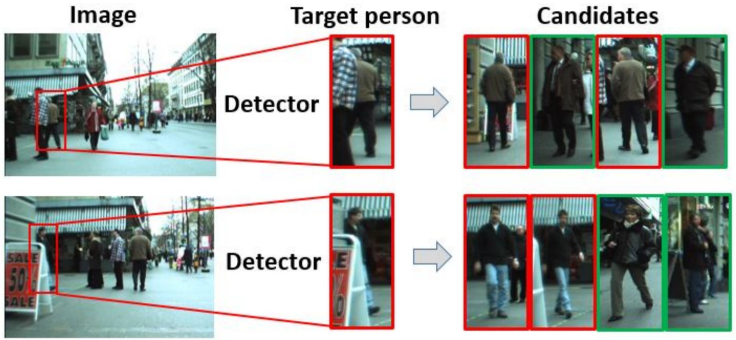
Gallery

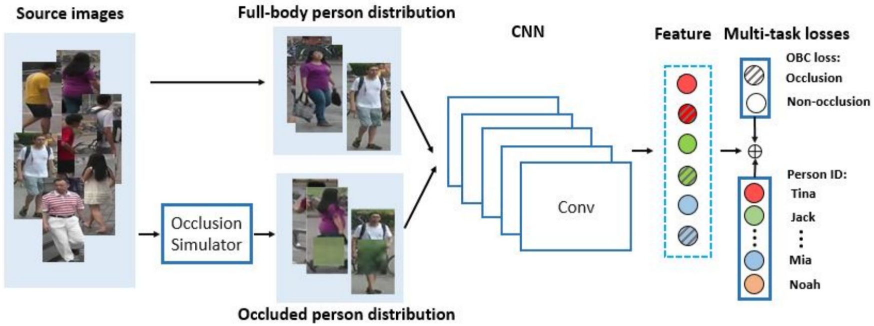


Zhuo, J., Chen, Z., Lai, J., & Wang, G. (2018). Occluded Person Re-identification. *arXiv*. https://doi.org/10.48550/arXiv.1804.02792
Peng, Y., Hou, S., Cao, C., Liu, X., Huang, Y., & He, Z. (2022). Deep Learning-based Occluded Person Re-identification: A Survey. *arXiv*. https://doi.org/10.48550/arXiv.2207.14452

Occlusions

- Data augmentation
- Smart features
- Multiple views
- Attention mechanisms





Outlook | Open Issues

- Poorly generalized global model
- Scalability model issues
- Multiple heterogeneous modalities
- Occlusions
- **Clothes changing**
- Human Annotation Minimization

Clothes changing

Qian, X., Wang, W., Zhang, L., Zhu, F., Fu, Y., Xiang, T., Jiang, Y., & Xue, X. (2020). Long-Term Cloth-Changing Person Re-identification. *arXiv*. https://doi.org/10.48550/arXiv.2005.12633

Clothes changing

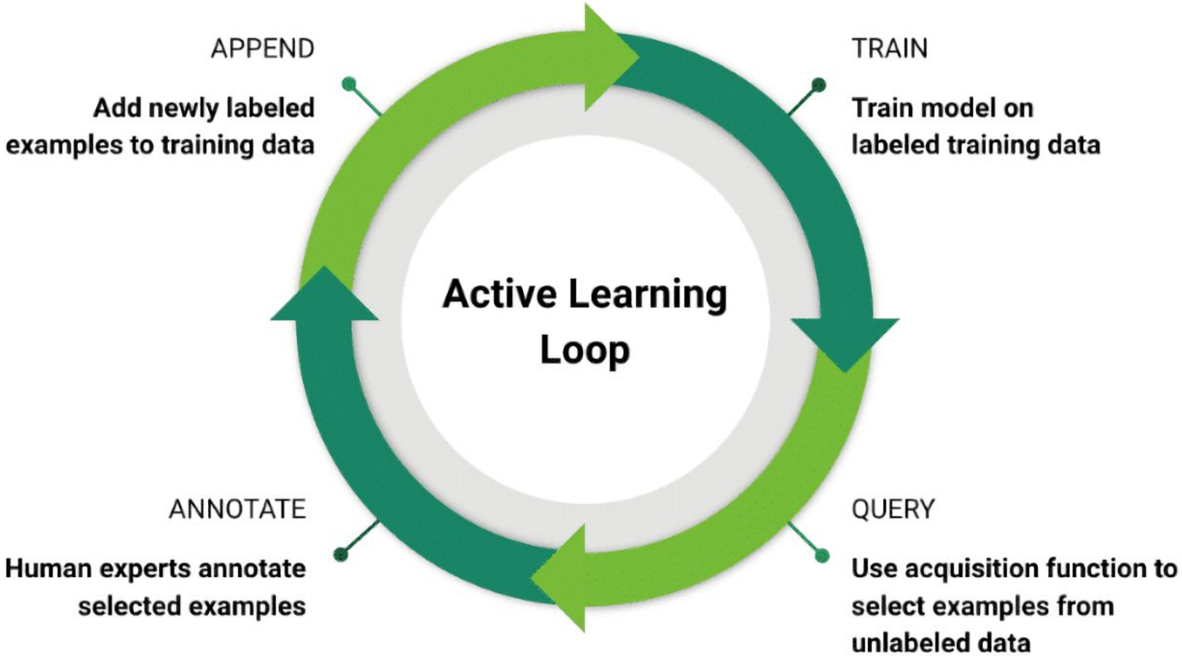
- Part-based approach
- Human pose estimation
- Attributes-based approach

Qian, X., Wang, W., Zhang, L., Zhu, F., Fu, Y., Xiang, T., Jiang, Y., & Xue, X. (2020). Long-Term Cloth-Changing Person Re-identification. *arXiv*. https://doi.org/10.48550/arXiv.2005.12633

Outlook | Open Issues

- Poorly generalized global model
- Scalability model issues
- Multiple heterogeneous modalities
- Occlusions
- Clothes changing
- Human Annotation Minimization

Active learning | human-in-the-loop



Learning from synthetic data

PersonX
dataset

VehicleX
dataset

Sun, X., & Zheng, L. (2018). Dissecting Person Re-identification from the Viewpoint of Viewpoint. arXiv. https://doi.org/10.48550/arXiv.1812.02162
Yao, Y., Zheng, L., Yang, X., Naphade, M., & Gedeon, T. (2019). Simulating Content Consistent Vehicle Datasets with Attribute Descent. arXiv. https://doi.org/10.48550/arXiv.1912.08855

References

1. He, T., Zhang, Z., Zhang, H., Zhang, Z., Xie, J., & Li, M. (2018). Bag of Tricks for Image Classification with Convolutional Neural Networks. arXiv. https://doi.org/10.48550/arXiv.1812.01187
2. Wang, H., Shen, J., Liu, Y., Gao, Y., & Gavves, E. (2022). NFormer: Robust Person Re-identification with Neighbor Transformer. arXiv. https://doi.org/10.48550/arXiv.2204.09331
3. He, S., Luo, H., Chen, W., Zhang, M., Zhang, Y., Wang, F., Li, H., & Jiang, W. (2020). Multi-Domain Learning and Identity Mining for Vehicle Re-Identification. arXiv. https://doi.org/10.48550/arXiv.2004.10547
4. Zheng, L., Yang, Y., & Hauptmann, A. G. (2016). Person Re-identification: Past, Present and Future. arXiv. https://doi.org/10.48550/arXiv.1610.02984
5. Wang, Z., Wang, Z., Zheng, Y., Wu, Y., Zeng, W., & Satoh, S. (2019). Beyond Intra-modality: A Survey of Heterogeneous Person Re-identification. arXiv. https://doi.org/10.48550/arXiv.1905.10048
6. Zheng, A., Wang, Z., Chen, Z., Li, C., & Tang, J. (2021). Robust Multi-Modality Person Re-identification. Proceedings of the AAAI Conference on Artificial Intelligence, 35(4), 3529-3537. https://doi.org/10.1609/aaai.v35i4.16467
7. Sun, X., & Zheng, L. (2018). Dissecting Person Re-identification from the Viewpoint of Viewpoint. arXiv. https://doi.org/10.48550/arXiv.1812.02162
8. Yao, Y., Zheng, L., Yang, X., Naphade, M., & Gedeon, T. (2019). Simulating Content Consistent Vehicle Datasets with Attribute Descent. arXiv. https://doi.org/10.48550/arXiv.1912.08855
9. “Open-reID.” [Online]. Available: https://github.com/Cysu/open-reid
10. “Pytorch reid.” [Online]. Available: https://github.com/layumi/Person_reID_baseline_pytorch
11. “Nformer.” [Online]. Available: https://github.com/haochenheheda/NFormer
12. “Bag of tricks and a strong reid baseline.” [Online]. Available: https://github.com/michuanhaohao/reid-strong-baseline