# Recent Advances in Person Re-identification for Real-world Scenarios

조윤기 (yoonki@kaist.ac.kr)

SGVR Lab, KAIST

# Contents

# Introduction

*Person re-identification and its challenges*

# Person Re-identification (Person Re-ID)

- Person re-ID aims to ***retrieve a person corresponding to a given query*** across disjoint camera views or different time stamps.
- Applications: Surveillance system, Finding a missing person, etc.

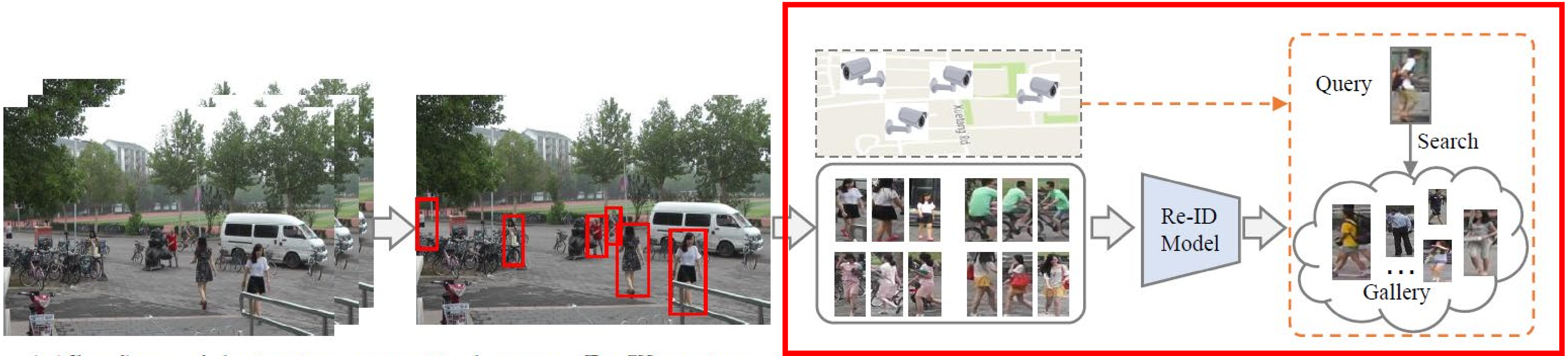



Fig. 1: The flow of designing a practical person Re-ID system, including five main steps: 1) *Raw Data Collection*, (2) *Bounding Box Generation*, 3) *Training Data Annotation*, 4) *Model Training* and 5) *Pedestrian Retrieval*.

SGVR Lab KAIST

Ye et al. Deep Learning for Person Re-identification: A Survey and Outlook. In TPAMI 2021.

# Person Re-identification (Person Re-ID)

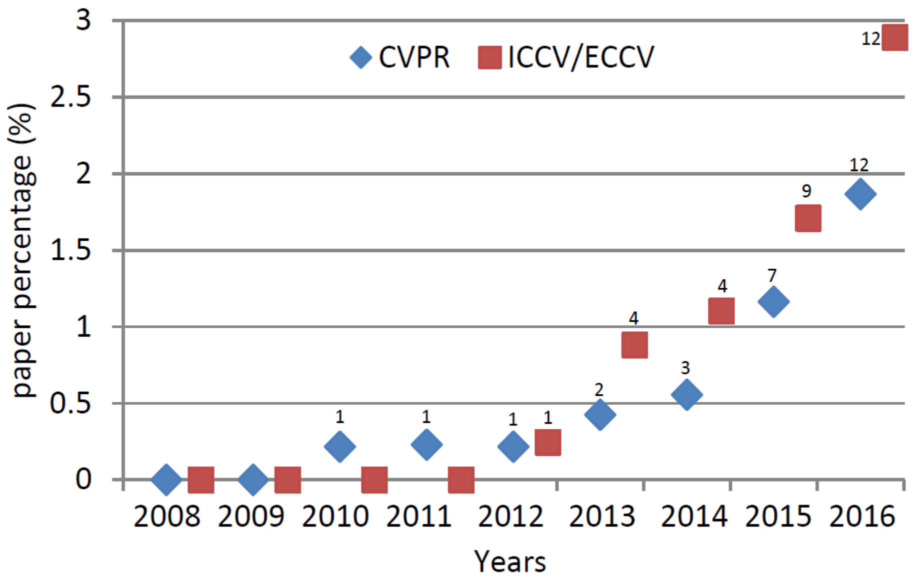



Fig. Percentage of person re-ID papers on top conferences over the years. Numbers above the markers indicate the number of re-ID papers.

| Conference | Link | #Total | Person Re-ID |
|---|---|---|---|
| ICCV2021 | click | 34 | 24 |
| CVPR2021 | click | 32 | 25 |
| ECCV2020 | Click | 30 | 23 |
| CVPR2020 | Click | 34 | 24 |
| ICCV2019 | Click | 39 | 33 |
| CVPR2019 | Click | 29 | 21 |
| ECCV2018 | Click | 19 | 15 |
| CVPR2018 | Click | 31 | 30 |
| ICCV2017 | Click | 16 | 14 |
| CVPR2017 | Click | 16 | 14 |



Zheng et al. Person Re-identification: Past, Present and Future. In arXiv 2016.
https://github.com/bismex/Awesome-person-re-identification.

# Datasets for Person Re-ID

- Evaluation metric
  - mean Average Precision (mAP)
  - Cumulative Matching Characteristics (CMC)

| | *Image datasets* | | | | | | |
|---|---|---|---|---|---|---|---|
| Dataset | Time | #ID | #image | #cam. | Label | Res. | Eval. |
| VIPeR | 2007 | 632 | 1,264 | 2 | hand | fixed | CMC |
| iLIDS | 2009 | 119 | 476 | 2 | hand | vary | CMC |
| GRID | 2009 | 250 | 1,275 | 8 | hand | vary | CMC |
| PRID2011 | 2011 | 200 | 1,134 | 2 | hand | fixed | CMC |
| CUHK01 | 2012 | 971 | 3,884 | 2 | hand | fixed | CMC |
| CUHK02 | 2013 | 1,816 | 7,264 | 10 | hand | fixed | CMC |
| CUHK03 | 2014 | 1,467 | 13,164 | 2 | both | vary | CMC |
| Market-1501 | 2015 | 1,501 | 32,668 | 6 | both | fixed | C&M |
| DukeMTMC | 2017 | 1,404 | 36,411 | 8 | both | fixed | C&M |
| Airport | 2017 | 9,651 | 39,902 | 6 | auto | fixed | C&M |
| MSMT17 | 2018 | 4,101 | 126,441 | 15 | auto | vary | C&M |
| | *Video datasets* | | | | | | |
| Dataset | time | #ID | #track(#bbox) | #cam. | label | Res. | Eval |
| PRID-2011 | 2011 | 200 | 400 (40k) | 2 | hand | fixed | CMC |
| iLIDS-VID | 2014 | 300 | 600 (44k) | 2 | hand | vary | CMC |
| MARS | 2016 | 1261 | 20,715 (1M) | 6 | auto | fixed | C&M |
| Duke-Video | 2018 | 1,812 | 4,832 (-) | 8 | auto | fixed | C&M |
| Duke-Tracklet | 2018 | 1,788 | 12,647 (-) | 8 | auto | C&M | |
| LPW | 2018 | 2,731 | 7,694(590K) | 4 | auto | fixed | C&M |
| LS-VID | 2019 | 3,772 | 14,943 (3M) | 15 | auto | fixed | C&M |

Ye et al. Deep Learning for Person Re-identification: A Survey and Outlook. In TPAMI 2021.

# Challenges in Person Re-ID

- Challenges by different camera views and time stamps.
  - Variance of viewpoints, illumination, pose, etc.
  - Occlusions.
  - Low resolutions.

- Large intra-variation & Small inter-variation

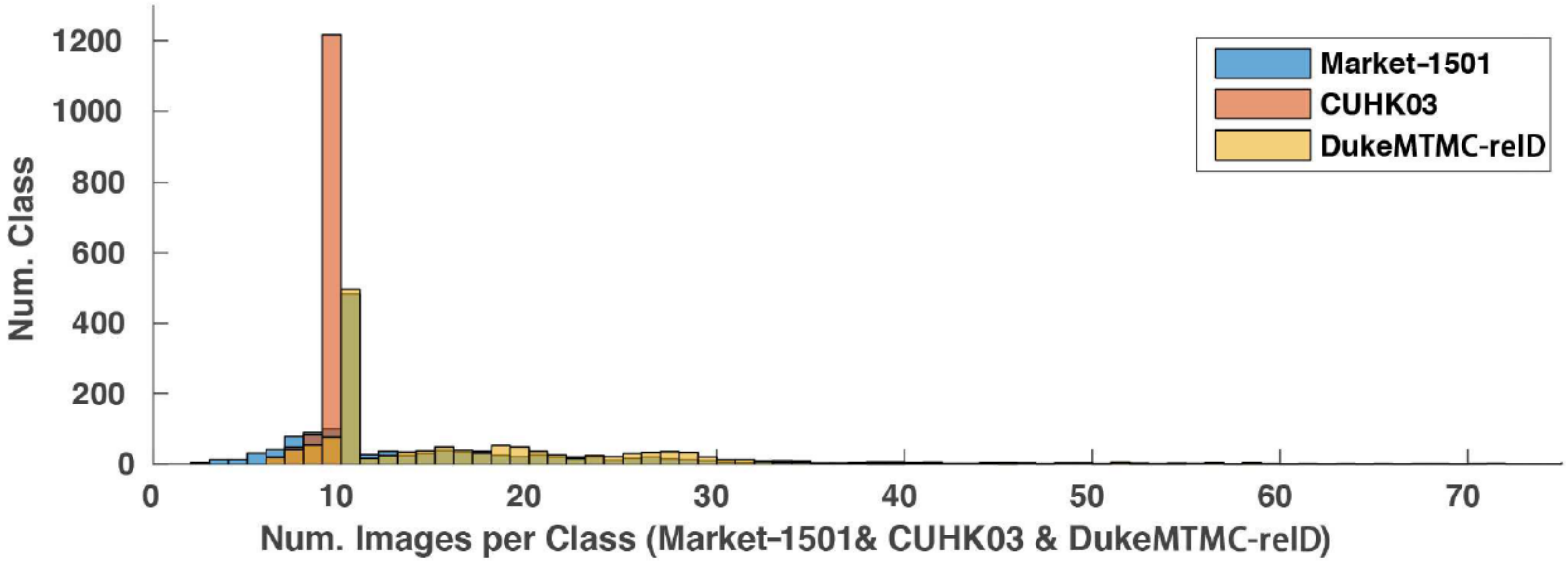


True match False match

SGVR Lab KAIST

# Challenges in Person Re-ID

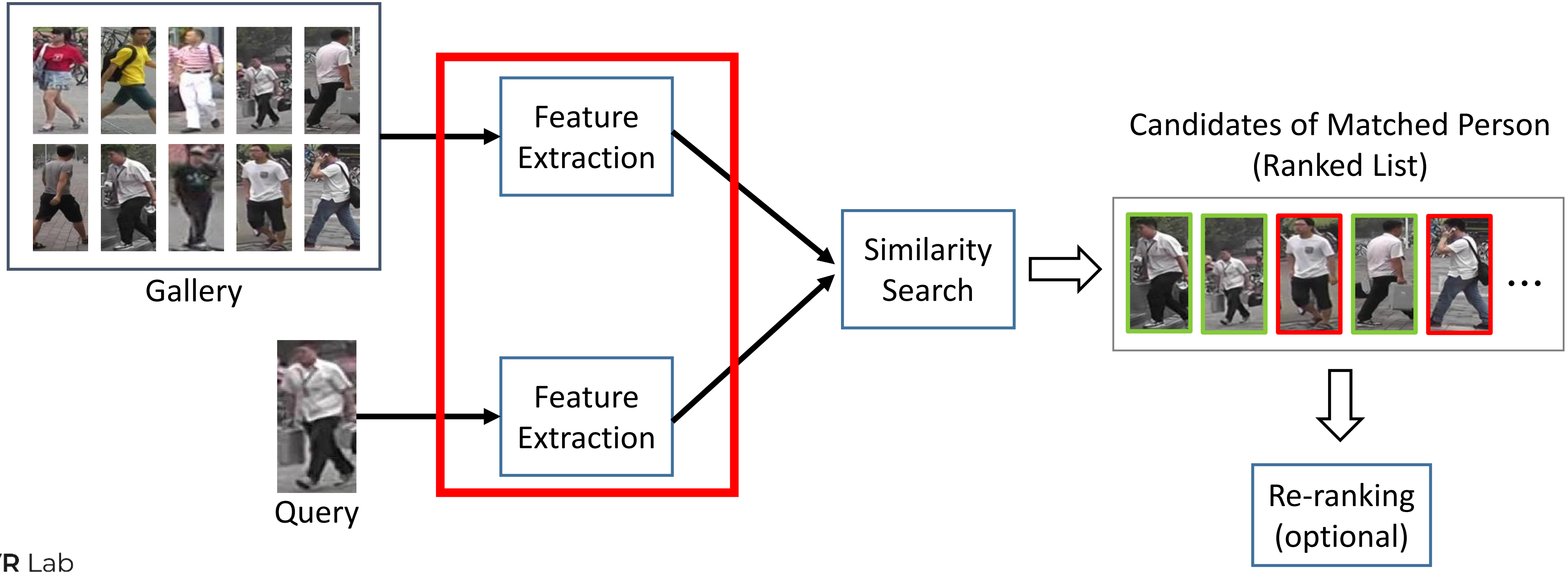
- Long-tail problem.
  - In person re-ID, all datasets suffer from the insufficient training set.
  - Insufficient training sets can yield overfitting and unstable convergences.
  - MNIST 10 class/ 5000 per class, CIFAR 100 class/500 per class.




Ro et al. Backbone Can Not be Trained at Once: Rolling Back to Pre-trained Network for Person Re-Identification. In AAAI 2019.
Zheng et al. Unlabeled Samples Generated by GAN Improve the Person Re-identification Baseline in vitro. In ICCV 2017.

# General Protocol of Person Re-ID

- Person re-identification pipeline.



# Feature Representation Learning for Person Re-ID

- Most studies focus on ***learning discriminative representation for person retrieval.***
- Recently, deep neural networks (DNN) have provided powerful descriptors.

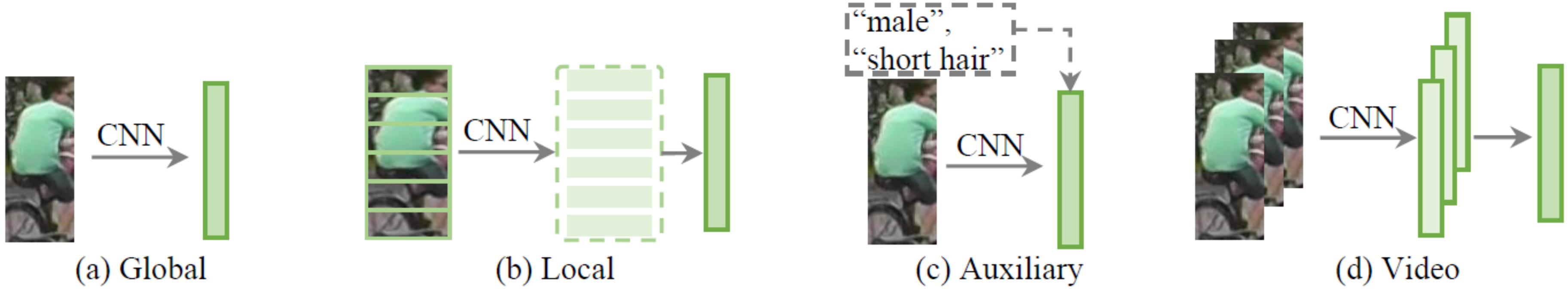



Fig. 2: Four different feature learning strategies. a) Global Feature, learning a global representation for each person image in § 2.1.1; b) Local Feature, learning part-aggregated local features in § 2.1.2; c) Auxiliary Feature, learning the feature representation using auxiliary information, *e.g.*, attributes [62], [63] in § 2.1.3 and d) Video Feature , learning the video representation using multiple image frames and temporal information [64], [65] in § 2.1.4.

Ye et al. Deep Learning for Person Re-identification: A Survey and Outlook. In TPAMI 2021.

# Supervised Person Re-identification

*Recent techniques for supervised approaches*

# Problem Setting

- *Goal:* ***Learn discriminative features*** for person retrieval with given labels.

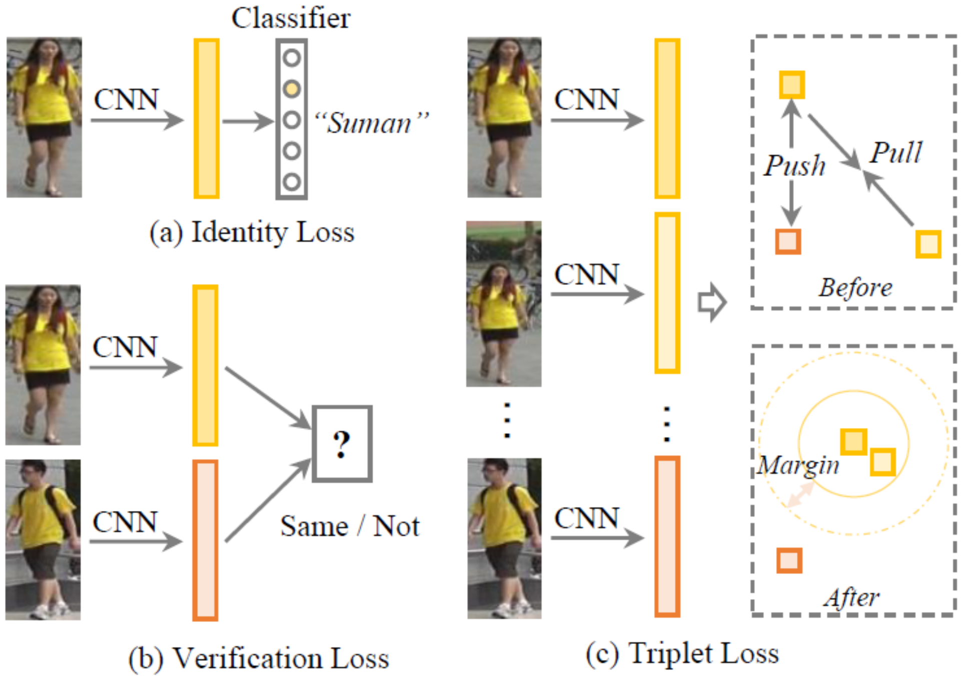
- *Protocol:* Training on target domain w/ labels → Testing on target domain.

- *Challenges:* Large intra-variation & Small inter-variation.
  Occlusions and misalignments.

SGVR Lab
KAIST

# Widely-used Loss Functions

- Identification loss ← classification problem.
- Verification loss ← binary classification problem.
- Triplet loss ← metric learning problem.

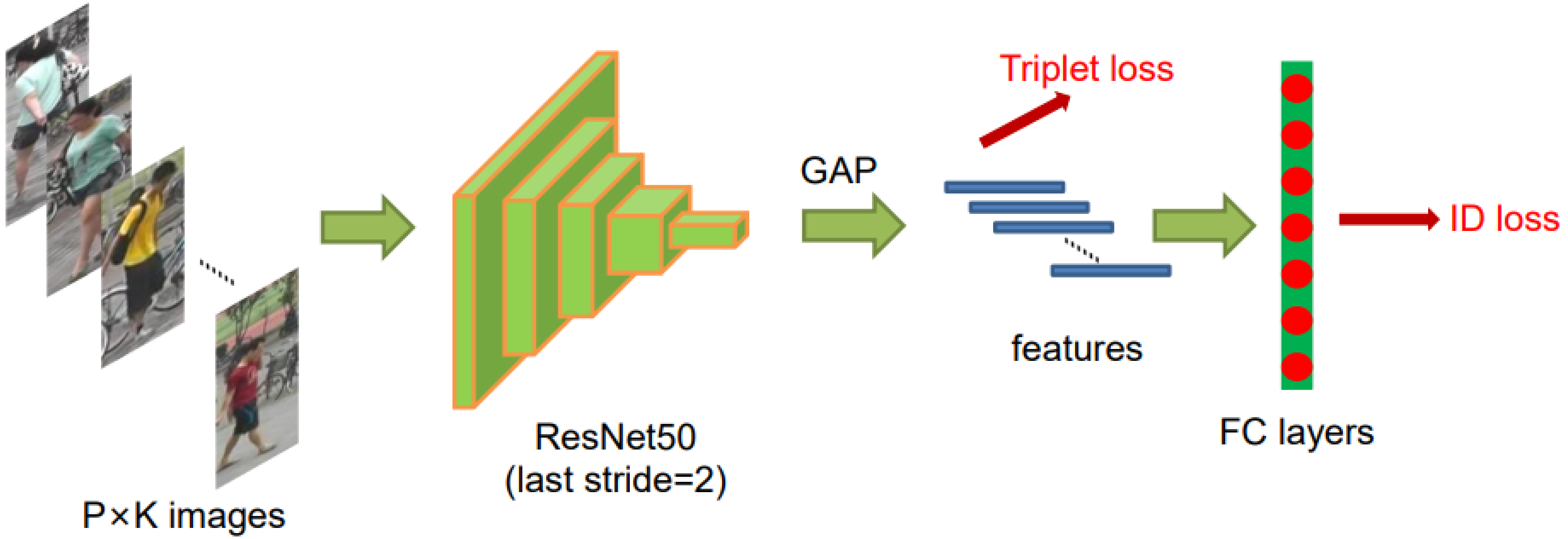



Ye et al. Deep Learning for Person Re-identification: A Survey and Outlook. In TPAMI 2021.

SGVR Lab KAIST

# Standard Approaches

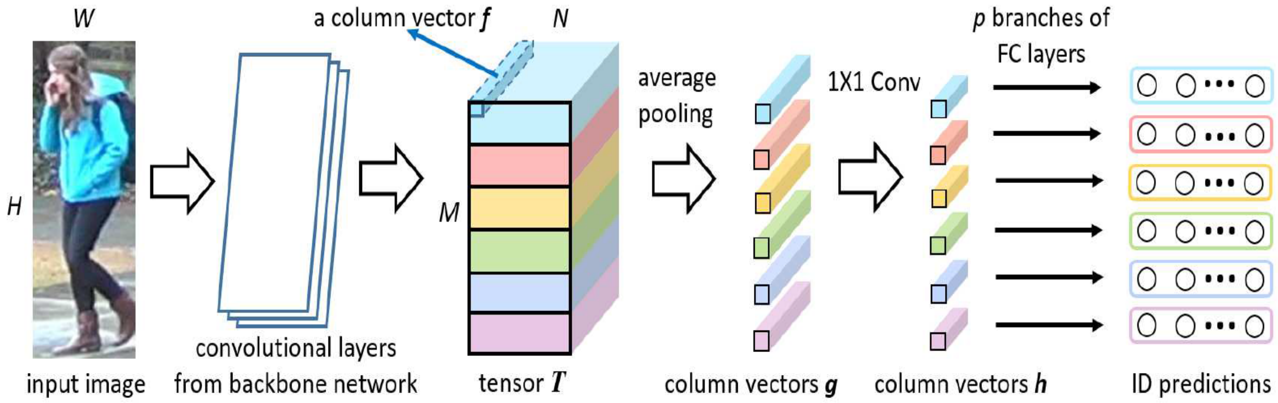
- Recent approaches ***utilize both identification and triplet loss.***
- ***ResNet-50*** is the standard architecture in recent studies.



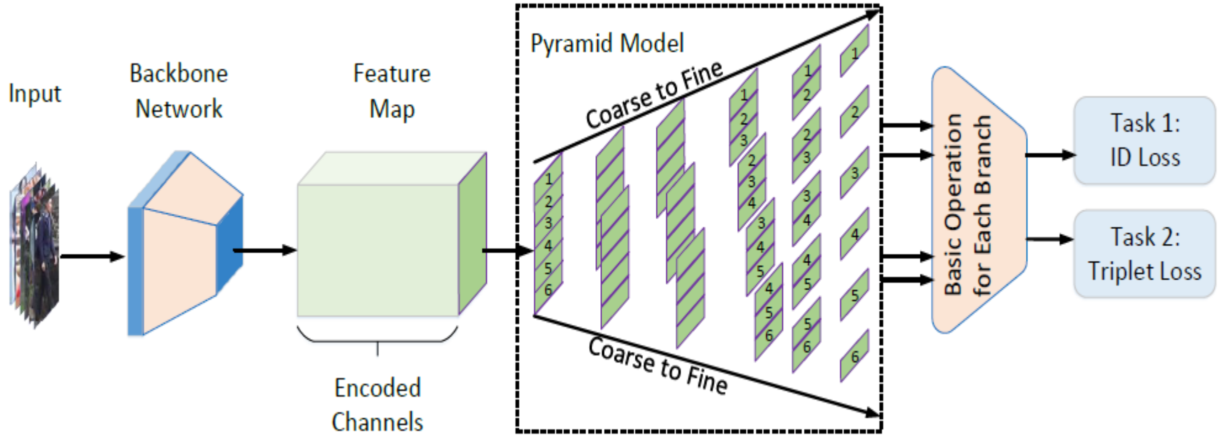

Luo et al. A Strong Baseline and Batch Normalization Neck for Deep Person Re-identification. In Transactions on Multimedia 2020.

# Part-based Approaches

- Key idea: ***Learn part(local) features to search a person*** via part-wise matchings.
- Divide a feature map uniformly and extract part features for training.




Part-based Convolutional Baseline (ECCV 18)



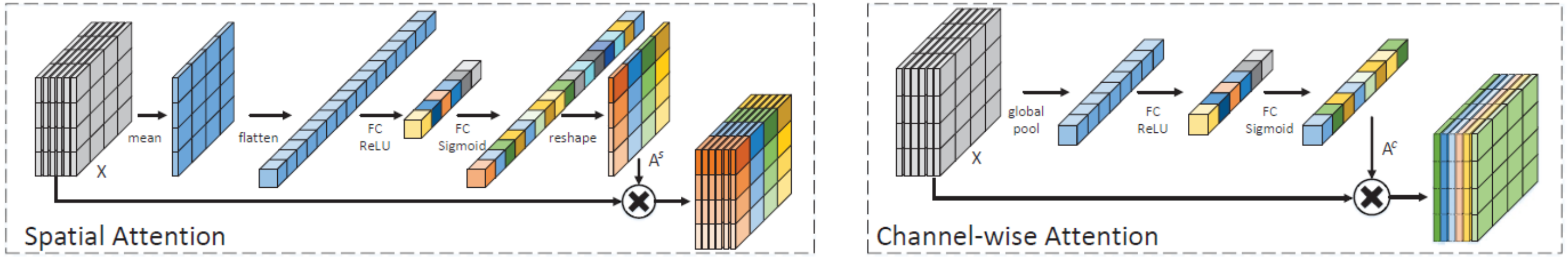

Pyramidal Model (CVPR 19)

Sun et al. Beyond Part Models: Person Retrieval with Refined Part Pooling (and a strong convolutional baseline). In ECCV 2018.
Zheng et al. Pyramidal Person Re-IDentification via Multi-Loss Dynamic Training. In CVPR 2019.

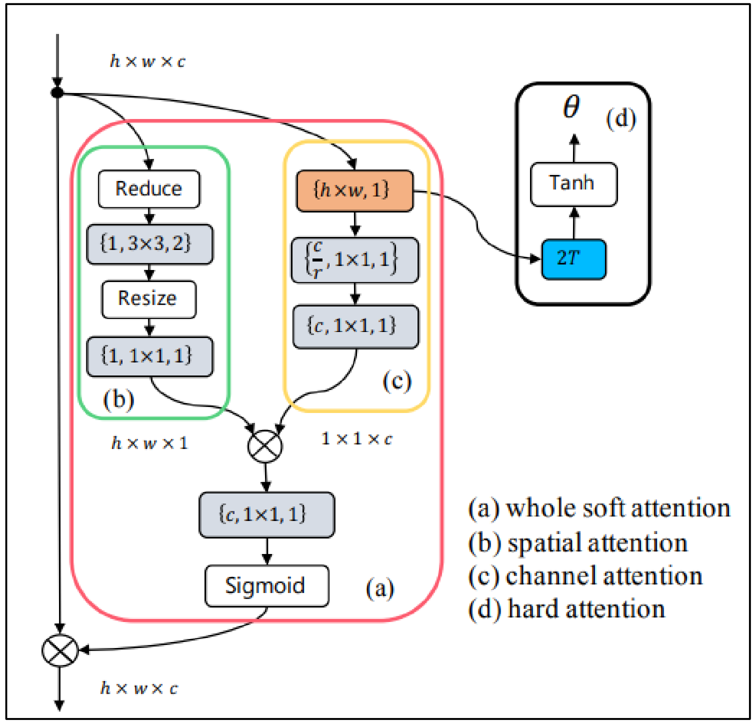
# Attention-based Approaches

- Key idea: Learn features ***robust to occlusions and misalignments via an attention***.
- Focus discriminative parts using a attention mechanism.

*Examples of ID 166 in Market-1501*




Self-Critical Attention Learning (ICCV 19)

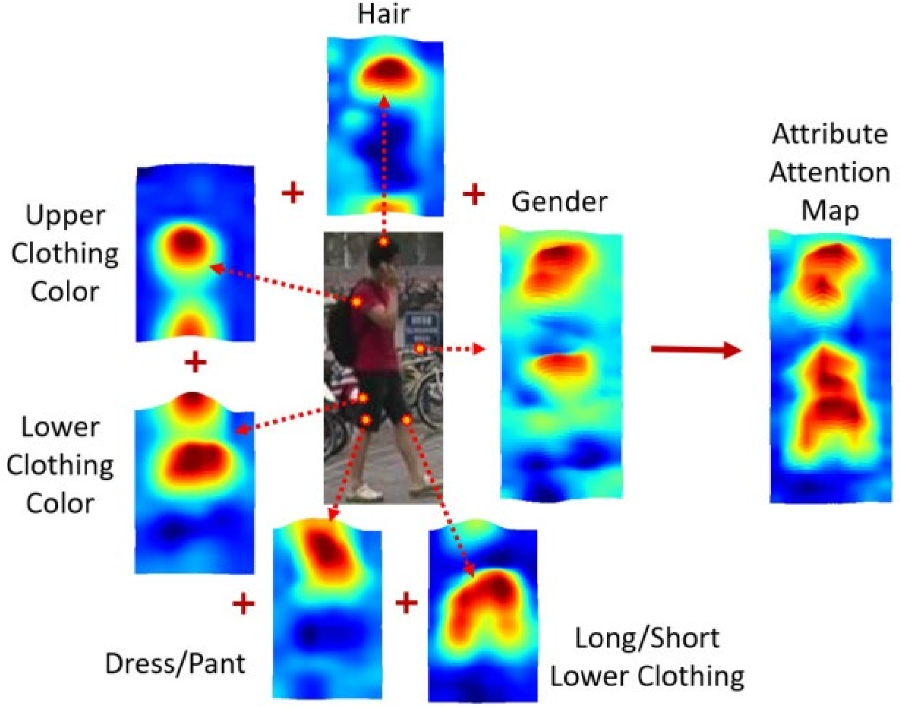



Harmonious Attention (CVPR 18)

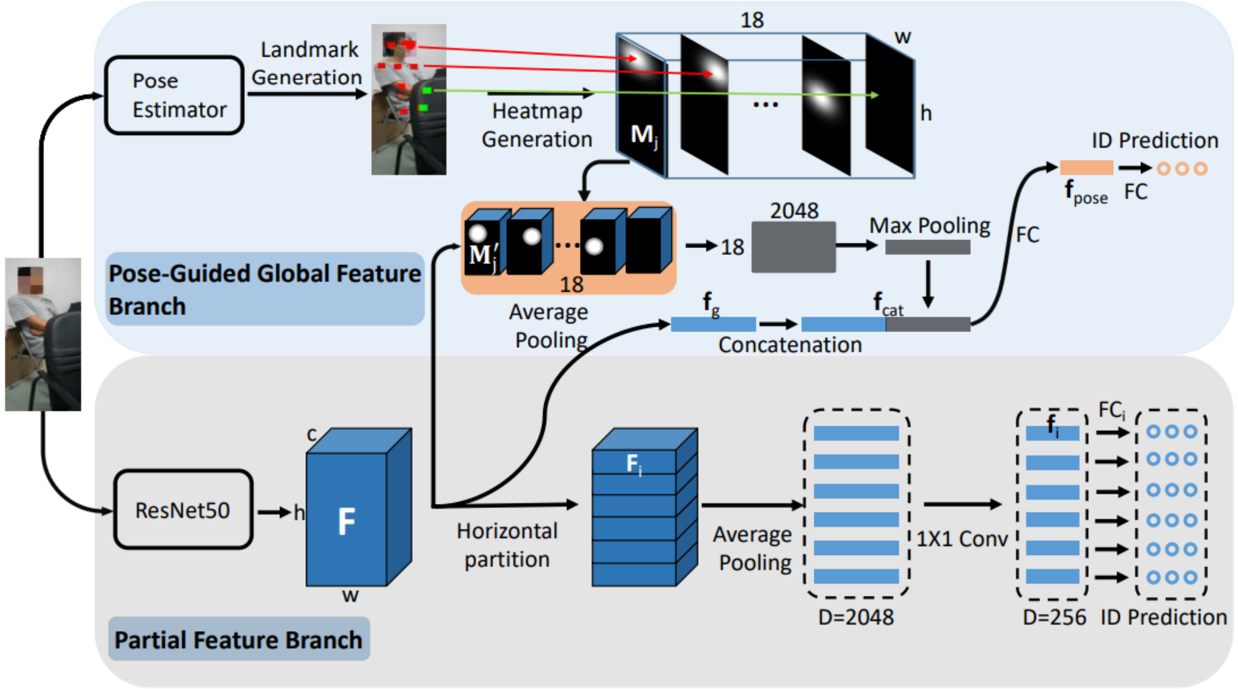
Chen et al. Self-Critical Attention Learning for Person Re-Identification. In ICCV 2019.
Li et al. Harmonious Attention Network for Person Re-Identification. In CVPR 2018.

# Approaches using Auxiliary Information

- Key idea: Learn discriminative features ***with additional clues***.
- Attribute labels, camera labels, pose estimator, human parsing, …




Attribute Attention Map (CVPR 19)



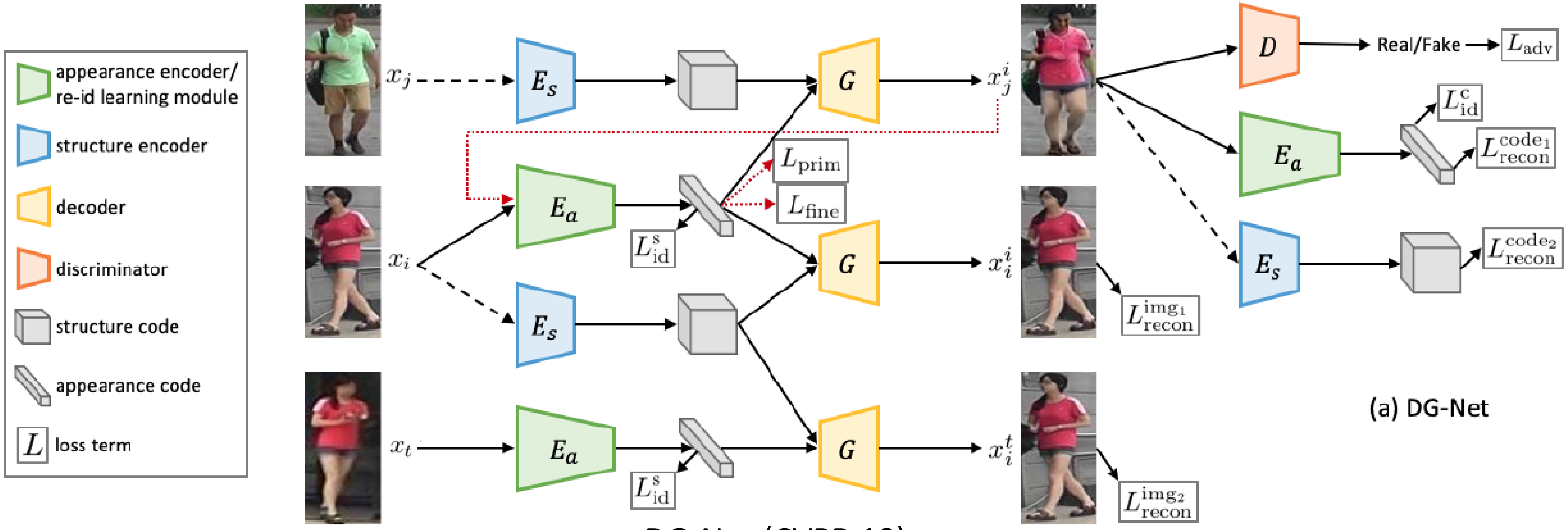

Pose-Guided Feature Alignment (ICCV 19)


Tay et al. AANet: Attribute Attention Network for Person Re-Identification. In CVPR 2019.
Miao et al. Pose-Guided Feature Alignment for Occluded Person Re-Identification. In ICCV 2019.


SGVR Lab KAIST

# Feature Disentanglement-based Approaches

- Key idea: Disentangle features into ***ID-relevant and ID-irrelevant*** features.
- Gave an additional supervision for feature disentanglements.

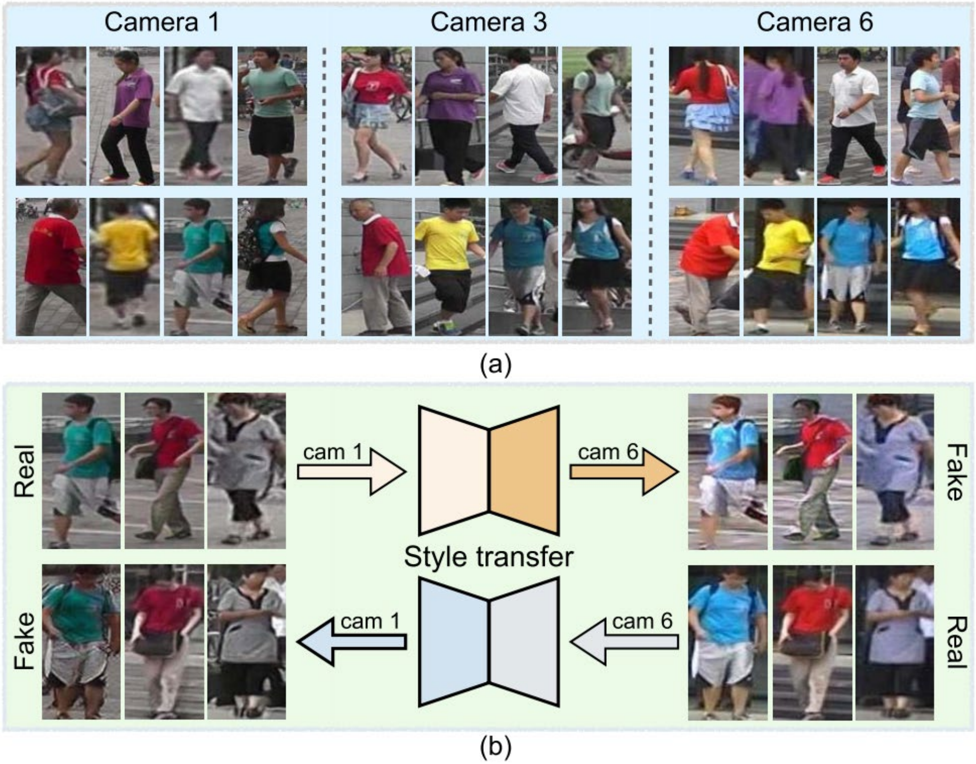



DG-Net (CVPR 19)

Zheng et al. Joint Discriminative and Generative Learning for Person Re-identification. In CVPR 2019.

# Data Augmentation for Person Re-ID

- Key idea: Augment input images for being robust to occlusions, camera bias, …
- Use input image distortion, style transfer, GAN, …




CamStyle (CVPR 18)


Random Erasing (AAAI 20)

Zhong et al. Camera Style Adaptation for Person Re-identification. In CVPR 2018.
Zhong et al. Random Erasing Data Augmentation. In AAAI 2020.

SGVR Lab KAIST

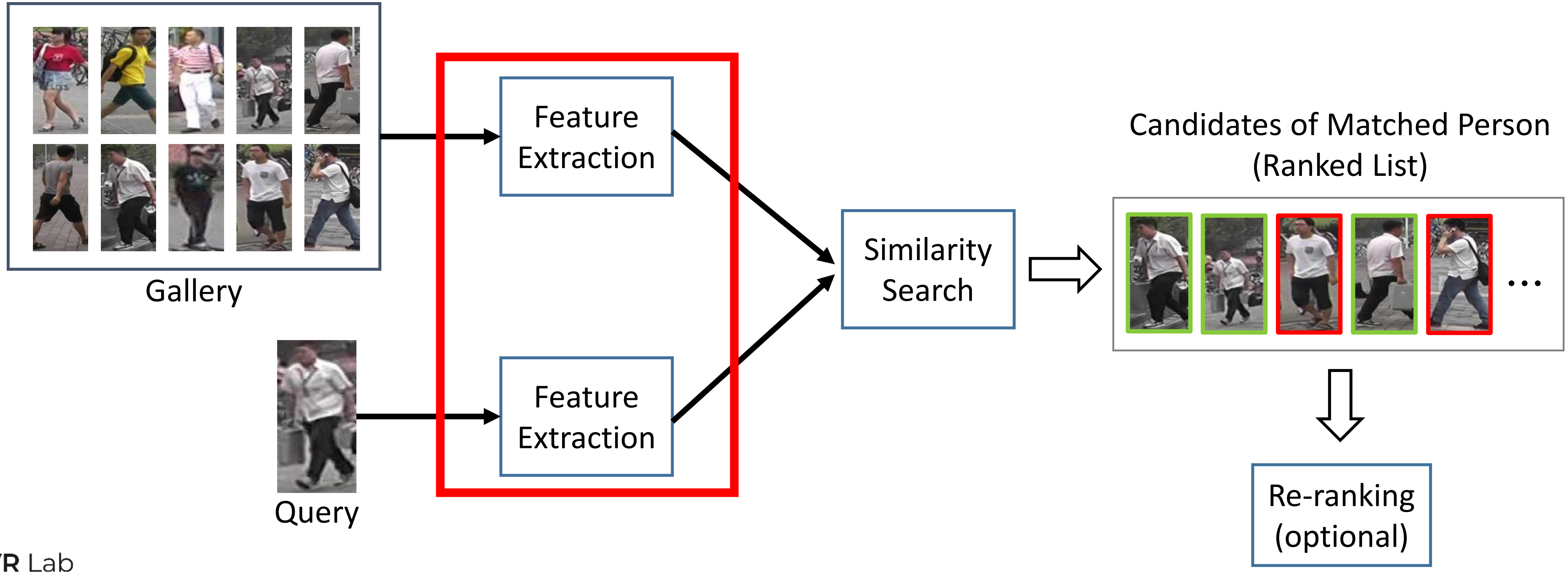
# Unsupervised Person Re-identification

*Recent techniques for unsupervised approaches*

SGVR Lab
KAIST

# General Protocol of Person Re-ID

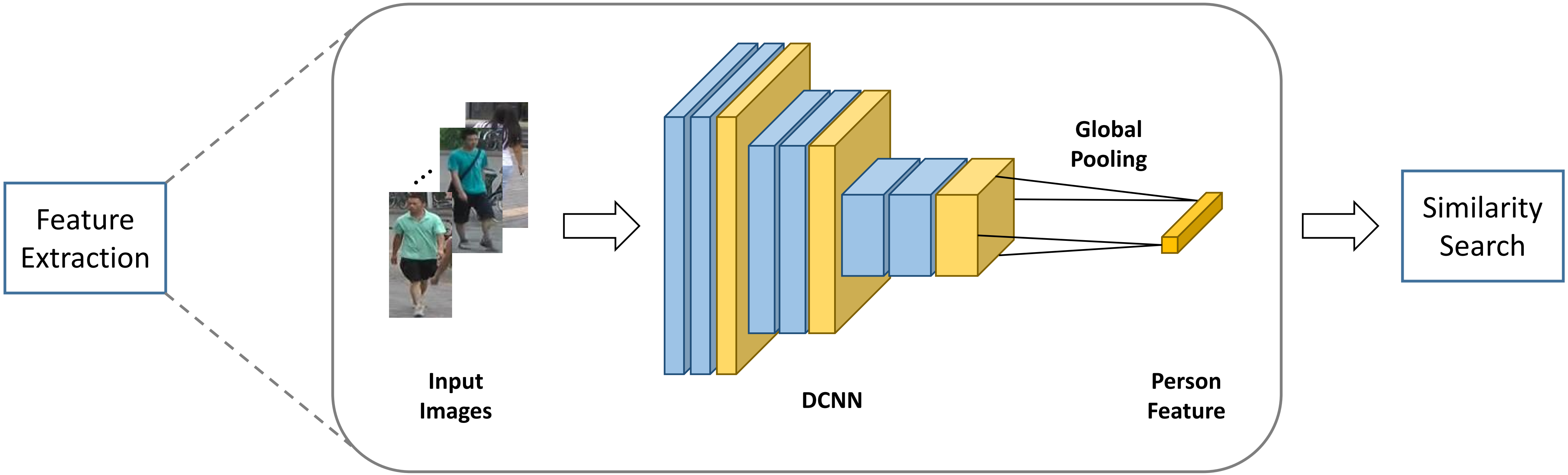
- Person re-identification pipeline.




SGVR Lab KAIST

# General Protocol of Person Re-ID

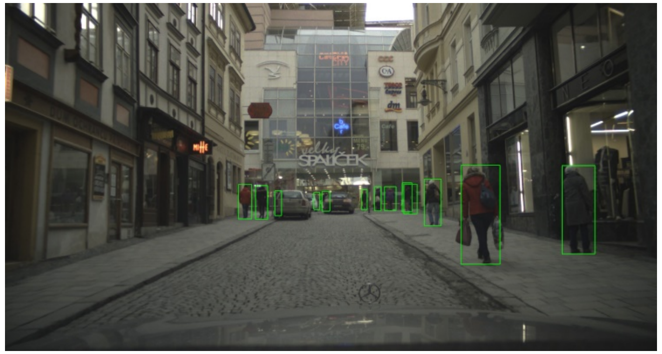
- Deep convolutional neural network (DCNN) brings impressive improvements in p erson re-ID fields.



# Problems in DCNN

- Require many **training data with labels.**
- Challenges in identity annotation.
  - illumination changes.
  - Low resolution.
  - **Occlusions.**

**Camera view**



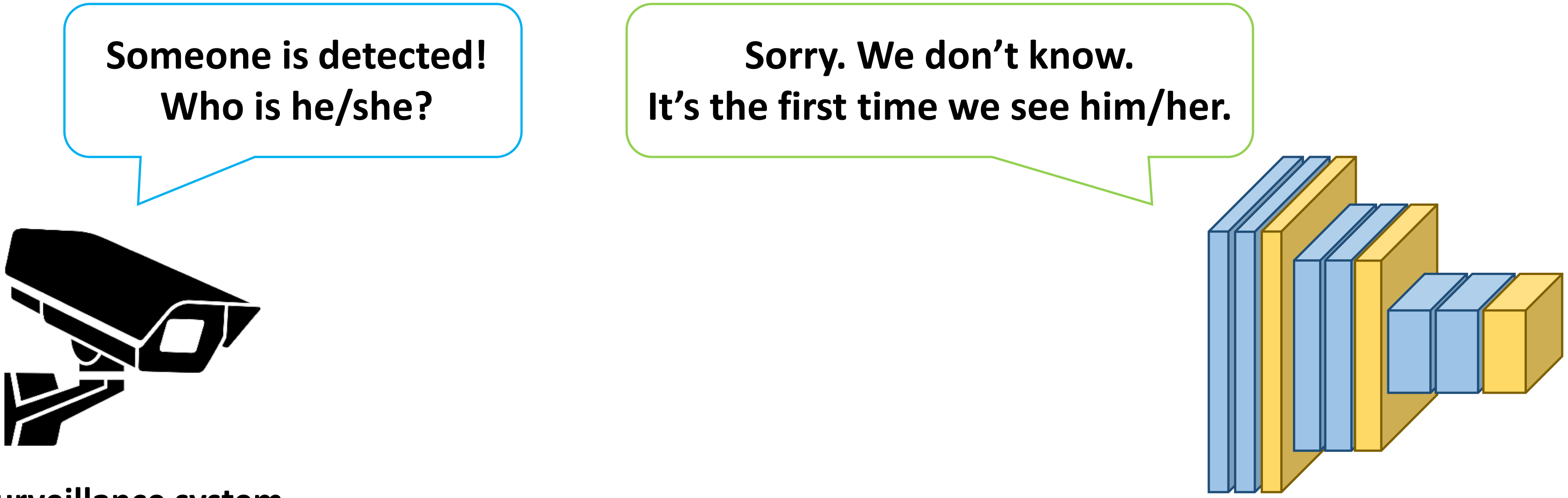
**Person Detection**


**Identity Annotation**

SGVR Lab
KAIST

# Problems in DCNN

- The real-world scenario of person re-ID is an open set problem.
- New people (= new class) will appear from the camera views.




Surveillance system

Person re-ID system

SGVR Lab KAIST

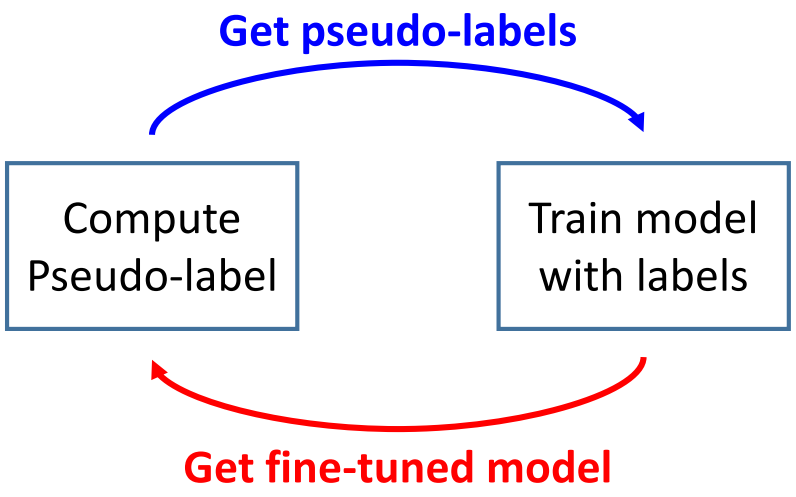
# Problem Setting

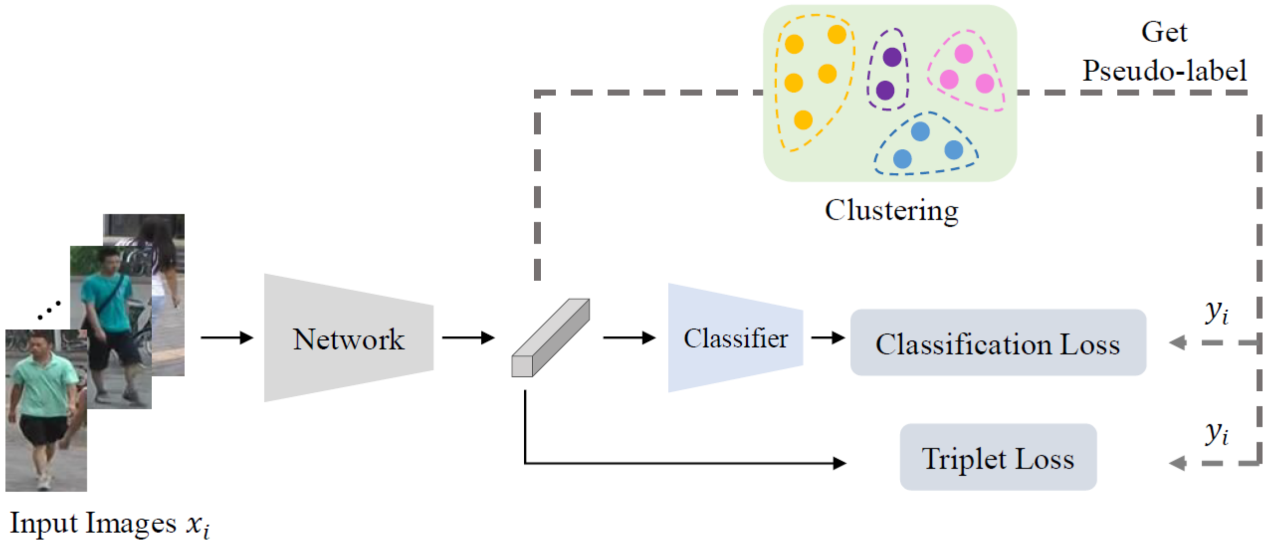
- *Goal:* Learn discriminative features for person retrieval ***without ID labels.***
- *Protocol:* Training on target domain w/o labels → Testing on target domain.
- *Challenges:* Poor pseudo-supervision from unlabeled data.

SGVR Lab KAIST

# Pseudo Label-based Approaches

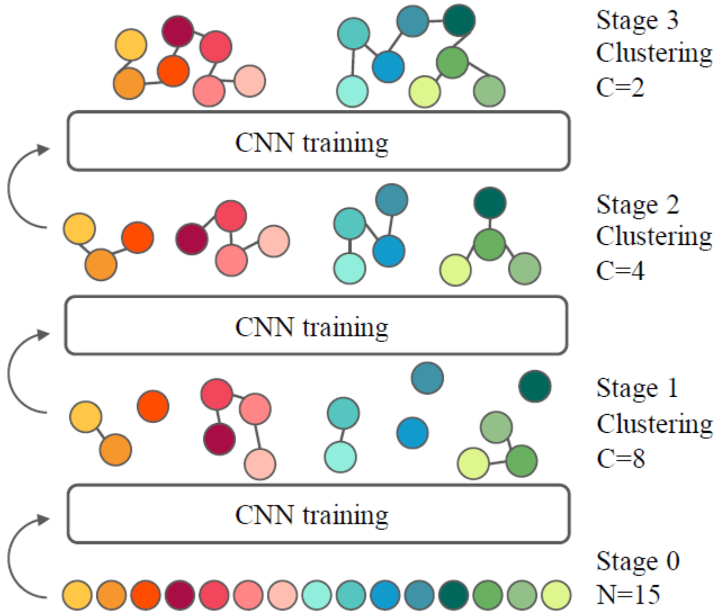
- Most recent studies ***utilize pseudo-labels to train a re-ID model.***
  - K-nearest neighbor search.
  - Clustering.
- Clustering-based approaches dominate unsupervised task.



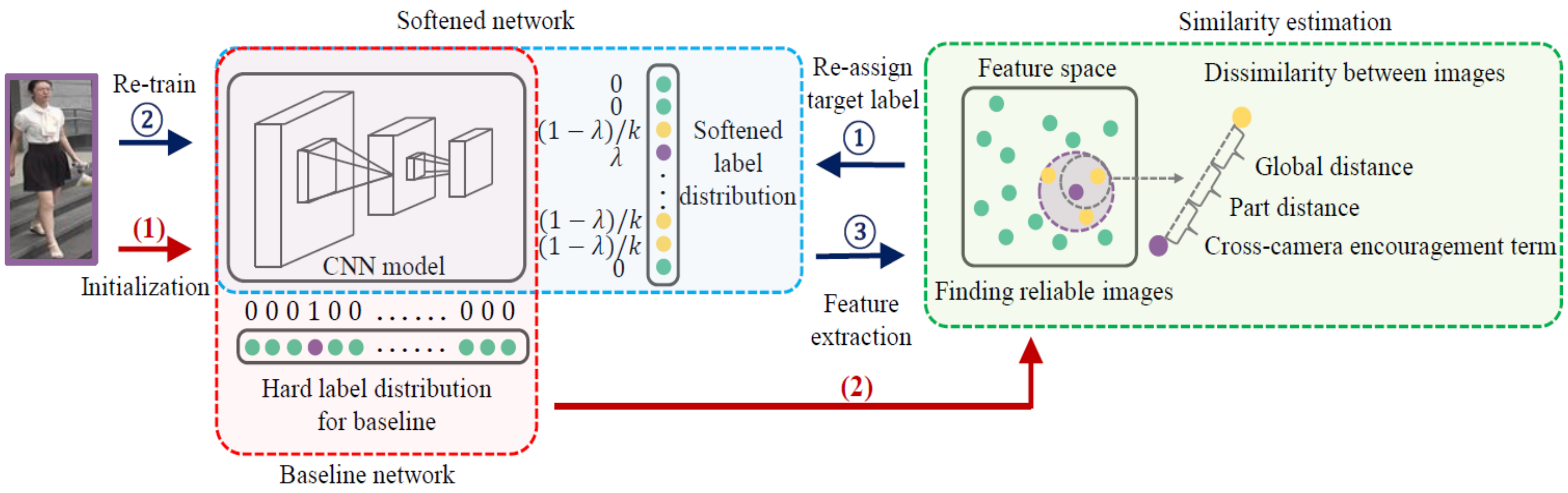



# Pseudo Label-based Approaches

- In early studies of this field focus on ***how to obtain pseudo-labels.***
- Nowadays, most methods utilize ***DBSCAN*** clustering with ***re-ranked distances***.




Bottom-Up Clustering (AAAI 19)



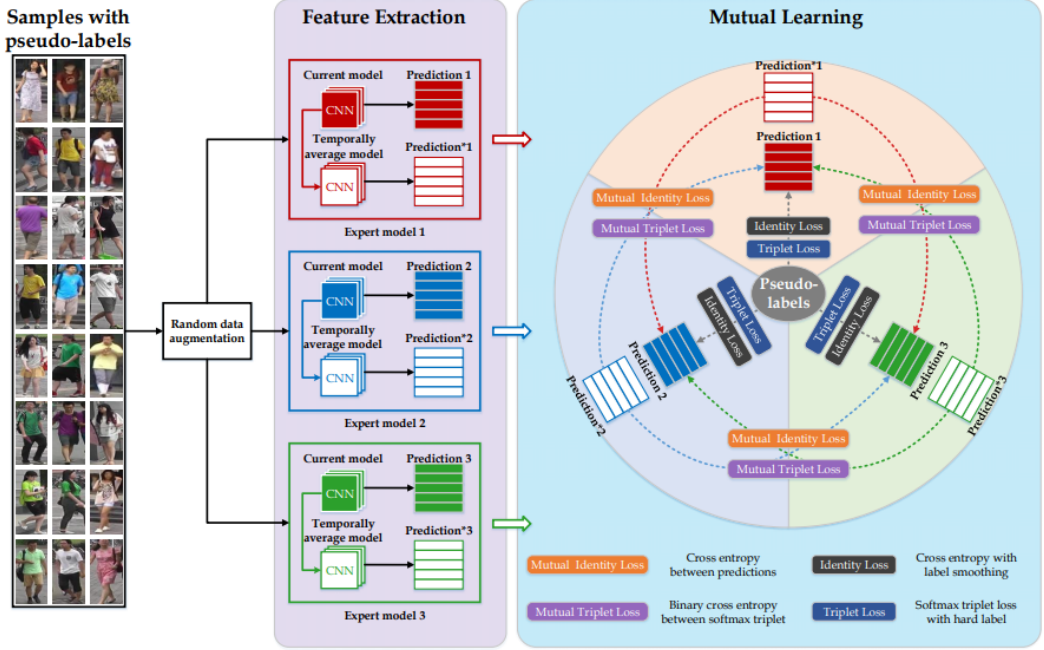

Softened Similarity Learning (CVPR 20)

Lin et al. A Bottom-Up Clustering Approach to Unsupervised Person Re-Identification. In AAAI 2019.
Lin et al. Unsupervised Person Re-identification via Softened Similarity Learning. In CVPR 2020.

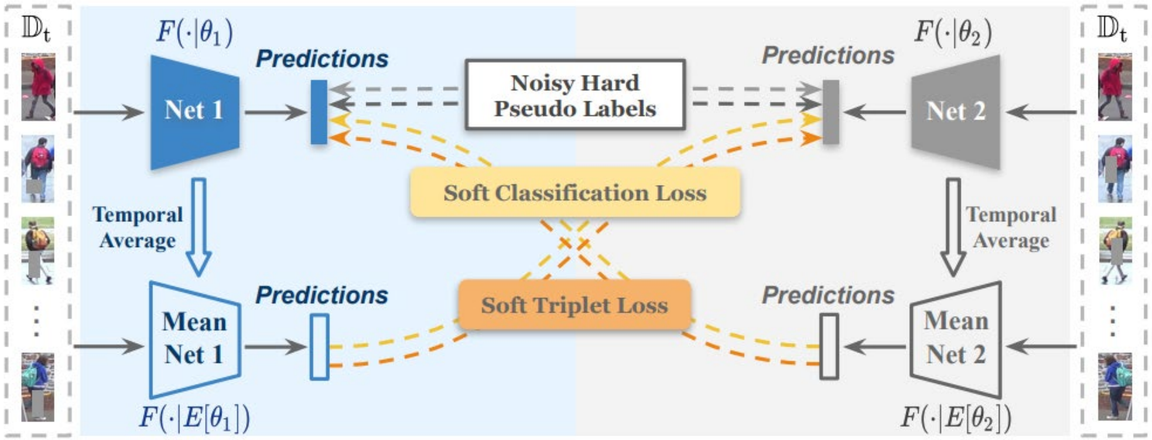
SGVR Lab KAIST

# Pseudo Label Refinement

- Key idea: Re-ID performance ∝ Quality of pseudo-labels.
- There are inevitable noised in pseudo-labels (noisy label problem), and some studies utilize predictions of an auxiliary network to refine labels.




MEB-Net (ECCV 20)



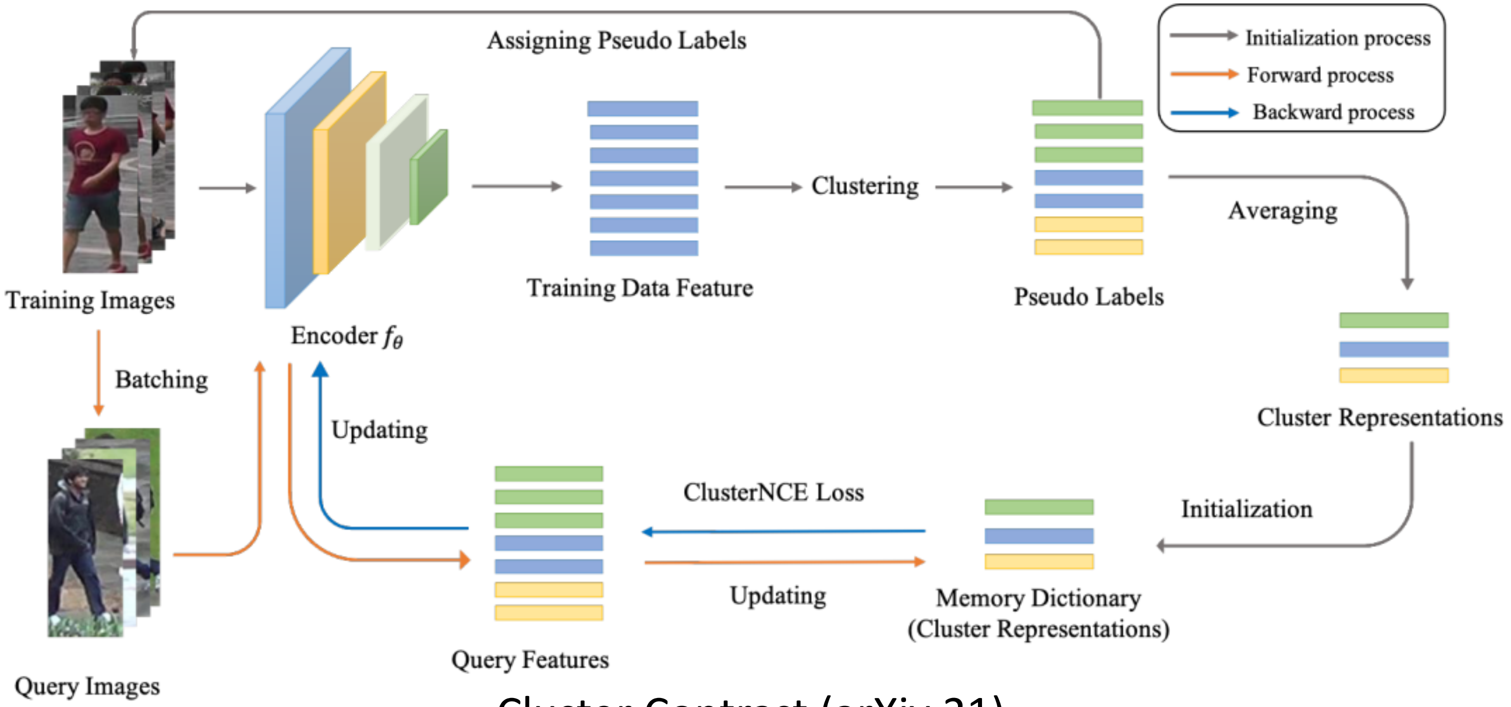

MMT (ICLR 20)

Zhai et al. Multiple Expert Brainstorming for Domain Adaptive Person Re-identification. In ECCV 2020.
Ge et al. Mutual Mean-Teaching: Pseudo Label Refinery for Unsupervised Domain Adaptation on Person Re-identification. In ICLR 2020.

SGVR Lab KAIST

# Cluster-based Contrastive Learning

- Key idea: Utilize clustering results for contrastive learning which is demonstrated its effectiveness in various unsupervised (self-supervised) tasks.
- Apply a contrastive learning in cluster-level.



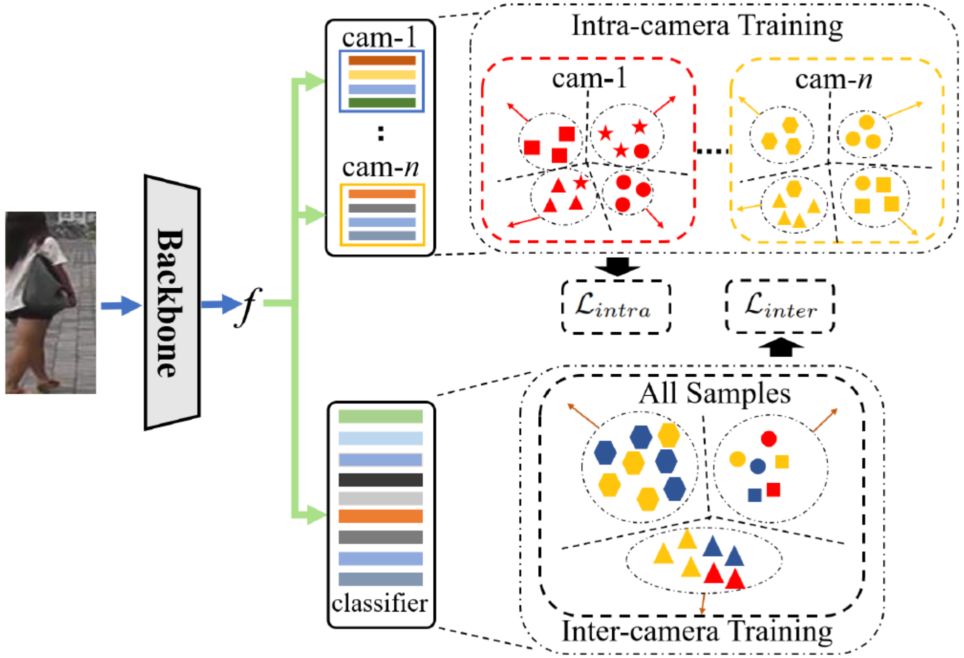

Cluster Contrast (arXiv 21)

Dai et al. Cluster Contrast for Unsupervised Person Re-Identification. arXiv preprint arXiv:2103.11568, 2021.

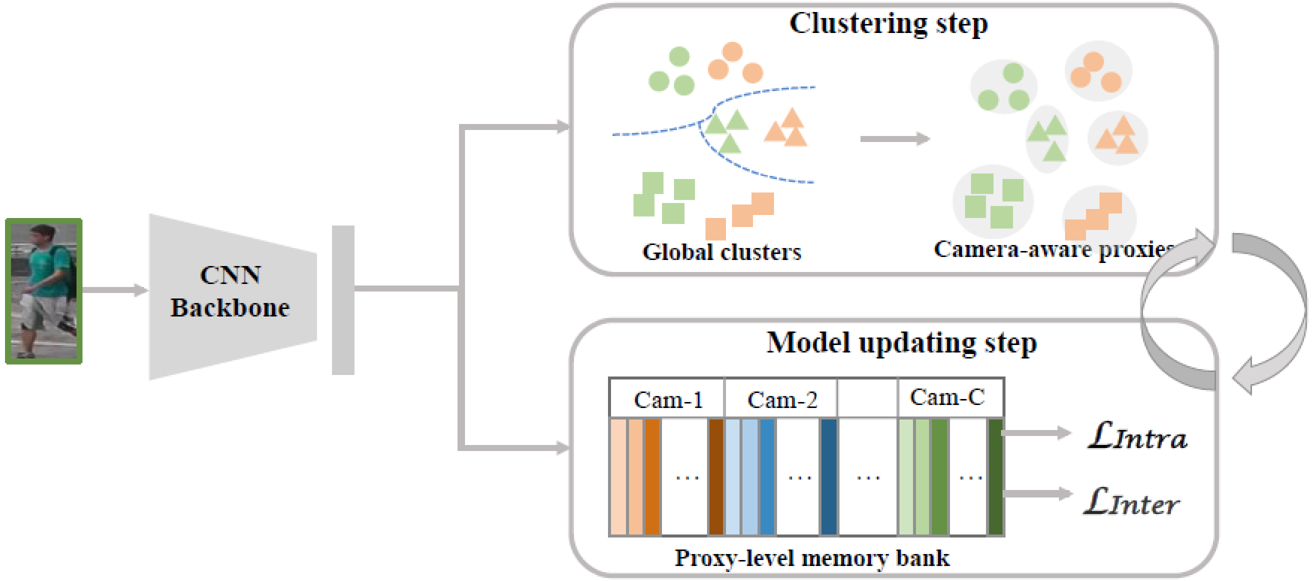
SGVR Lab KAIST

# Camera-based Approaches

- Key idea: Camera labels are easy to obtain and can be used for re-ID training.
- Apply intra-camera level training and inter-camera level training.
- Directly reduce the intra-class variance by different camera views.




Intra-Inter Camera Similarity (CVPR 21)




Camera-aware Proxy (AAAI 21)

SGVR Lab KAIST


Xuan and Zhang. Intra-Inter Camera Similarity for Unsupervised Person Re-Identification. In CVPR 2021.
Wang et al. Camera-aware Proxies for Unsupervised Person Re-Identification. In AAAI 2021.

# Generalizable Person Re-identification

*Recent techniques for domain generalizable approaches*

SGVR Lab
KAIST

# Domain Generalization (DG) Problem

- Trained re-ID models show degraded performance in unseen domains.
- Practically, it is not easy to train a re-ID model according to each environment (domain).
  - High cost (e.g., GPU, engineering, …) is required.
  - Data collection problem w/ privacy issues.

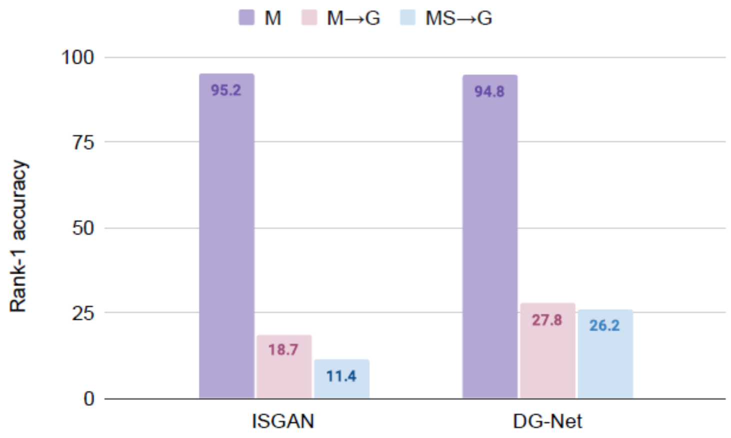



Figure 1. Performance of two representative models. M: Train and test on Market1501. M→G: Trained on Market1501 and tested on GRID . MS→G: Trained on multi-source datasets and tested on GRID.

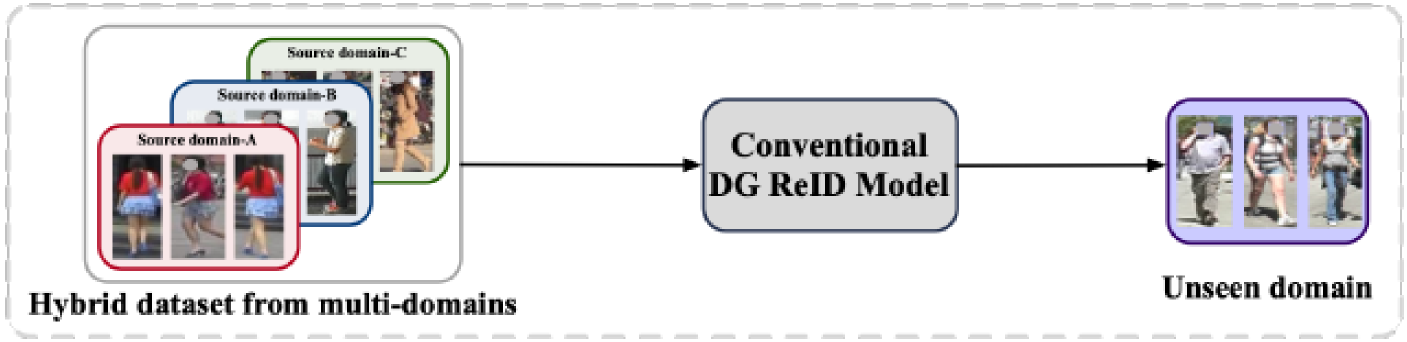
Zhang et al. Learning Domain Invariant Representations for Generalizable Person Re-Identification. arXiv preprint arXiv:2103.15890.

SGVR Lab KAIST

# Generalizable Person Re-identification

- In contrast to most DG studies focusing on close-set scenarios, ***generalizable person re-ID focuses on the open-set problem***.

(a) Digits (b) VLCS (c) PACS




Zhou et al. Domain Generalization in Vision: A Survey. arXiv preprint arXiv:2103.02503.
Wang et al. Camera-aware Proxies for Unsupervised Person Re-Identification. In AAAI 2021.

# Generalizable Person Re-identification

- Generalizable person re-ID aims to solve ***"Domain Generalization (DG)" for person retrieval.***

- What makes it challenging?
  - Each dataset (domain) has different characteristics.
  - Season(Weather), Viewpoint, etc.


Zheng et al. Joint Discriminative and Generative Learning for Person Re-identification. In CVPR 2019.

# Problem Setting

- *Goal:* Learn ***representations with good generalization capability on unseen target domains***.
- *Protocol:* Training on source domain with labels → Testing on unseen target domains.
- *Challenges:* Domain gap between source and target domains.

# Style Normalization-based Approaches

- CNNs are strongly biased to style (texture) of images.

(a) Texture image
81.4% **Indian elephant**
10.3% indri
8.2% black swan

(b) Content image
71.1% **tabby cat**
17.3% grey fox
3.3% Siamese cat

(c) Texture-shape cue conflict
63.9% **Indian elephant**
26.4% indri
9.6% black swan

- Style normalization-based approaches claim that “Style differences between the domains make a domain gap”.

Geirhos et al. ImageNET-Trained CNNS Are Biased Towards Texture; Increasing Shape Bias Improves Accuracy And Robustness. In ICLR 2019.

# Style Normalization-based Approaches

- Key idea: Reduce style variations via a feature normalization.
- Style Normalization can be achieved by Instance Normalization (IN).
- Some works train a style-invariant network by using both BN and IN.

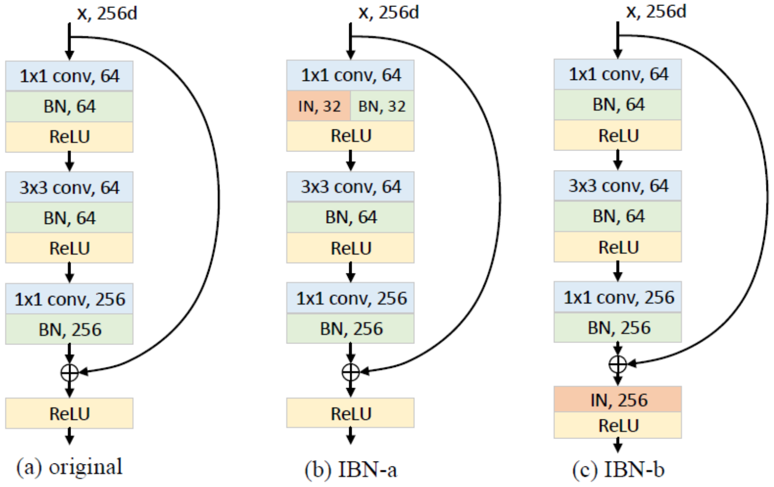
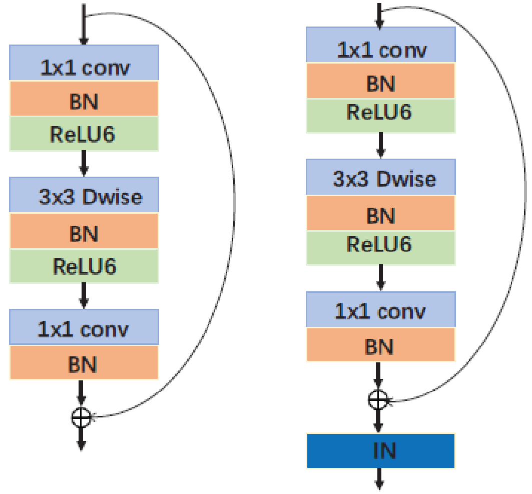



**Fig. 3.** Instance-batch normalization (IBN) block.

IBN-Net (ECCV 18)




(a) Original bottleneck

(b) bottleneck-IN

DualNorm (BMVC 19)

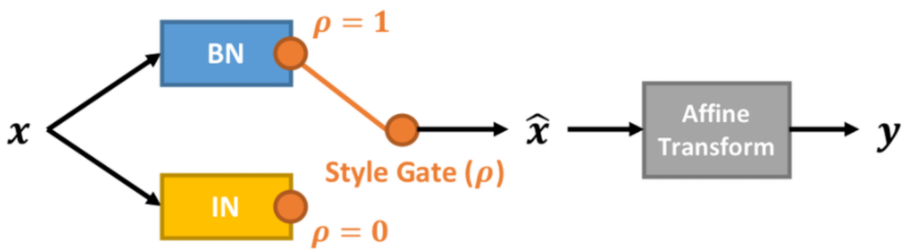


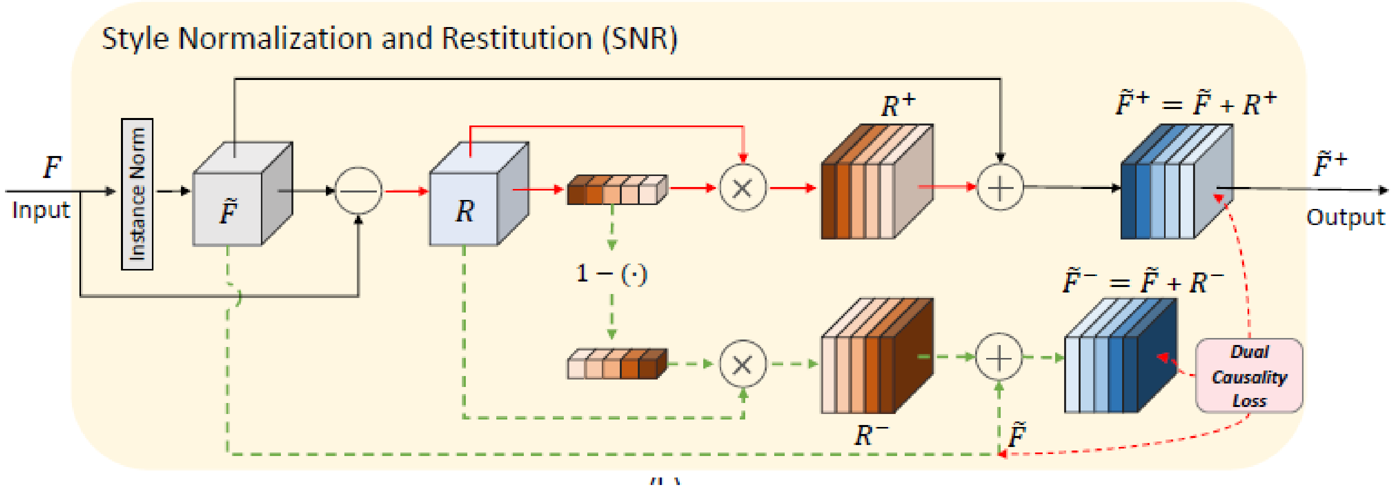

Diagram of Batch-Instance Normalization. The style gate determines the proper normalization method based on the importance of the style information.

BIN (NeurIPS 18)


Pan et al. Two at Once: Enhancing Learning and Generalization Capacities via IBN-Net. In ECCV 2018.
Jia et al. Frustratingly Easy Person Re-Identification: Generalizing Person Re-ID in Practice. In BMVC 2019.
Nam et al. Batch-Instance Normalization for Adaptively Style-Invariant Neural Networks. In NeurIPS 2018.

# Style Normalization-based Approaches

- Style normalization also can remove the discriminative information.
- SNR (CVPR 21) disentangles identity-relevant and –irrelevant features.




Style Normalization and Restitution (CVPR 20)

Jin et al. Style Normalization and Restitution for Generalizable Person Re-identification. In CVPR 2020.
Choi et al. Meta Batch-Instance Normalization for Generalizable Person Re-Identification. In CVPR 2021.

# Meta Learning-based Approaches

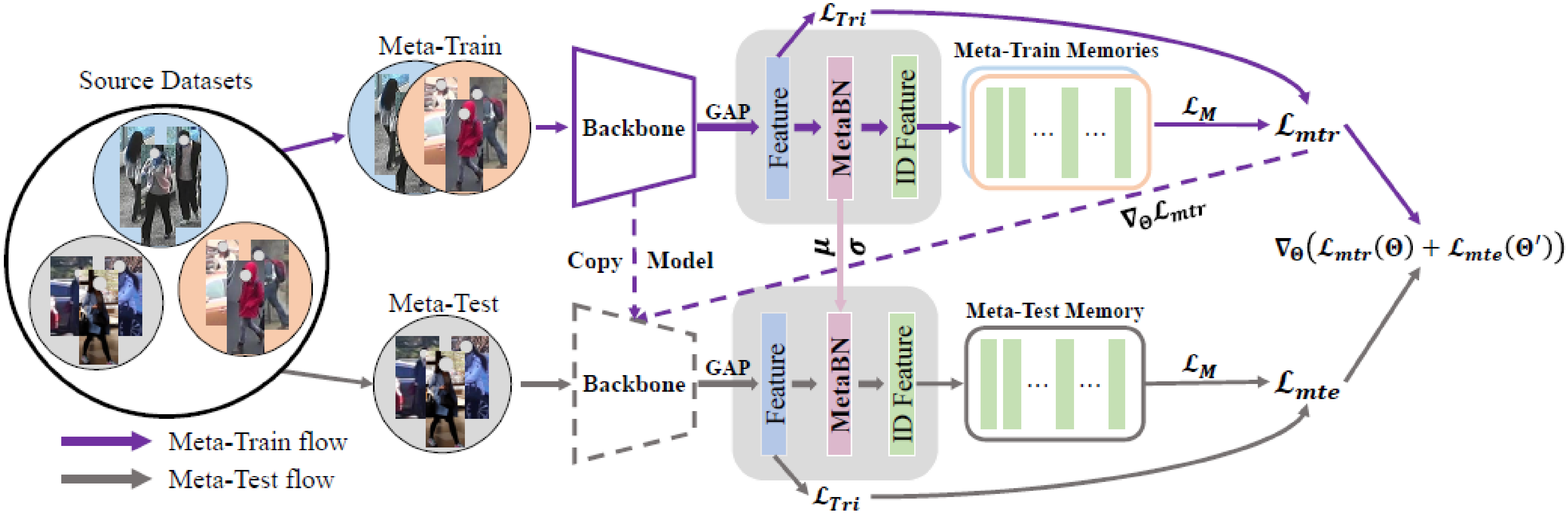
- Meta learning approaches adapt the concept of “learning to learn” to simulate the train-test process of domain generalization scenarios.
  - Make the training process like a domain generalization task.

- It divides given datasets into a meta-train domain and a meta-test domain.

- Meta learning process can be divided into “meta-train” and “meta-test”.
  - Meta-train = conventional re-ID training.
  - Meta-test = domain generalization training.

# Meta Learning-based Approaches

- Compute the loss on a meta-test domain with the trained model in a meta-train domain.

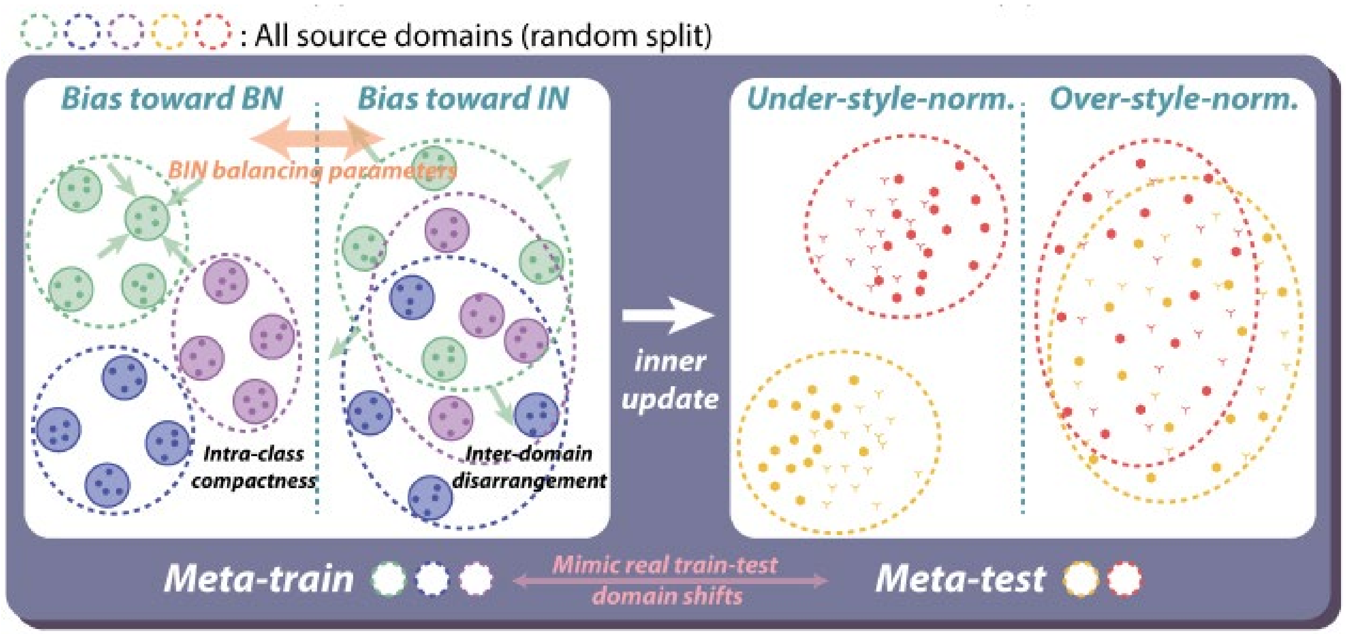



Memory-based Multi-Source Meta- Learning (CVPR 21)

Zhao et al. Learning to Generalize Unseen Domains via Memory-based Multi-Source Meta-Learning for Person Re-Identification. In CVPR 2021.

# Hybrid Approaches

- MetaBIN (CVPR 21) trains a style normalization module by meta learning pipeline.
  - Train BIN (Batch-instance normalization) parameters in a meta learning manner.




MetaBIN (CVPR 21)

Choi et al. Meta Batch-Instance Normalization for Generalizable Person Re-Identification. In CVPR 2021.

SGVR Lab KAIST

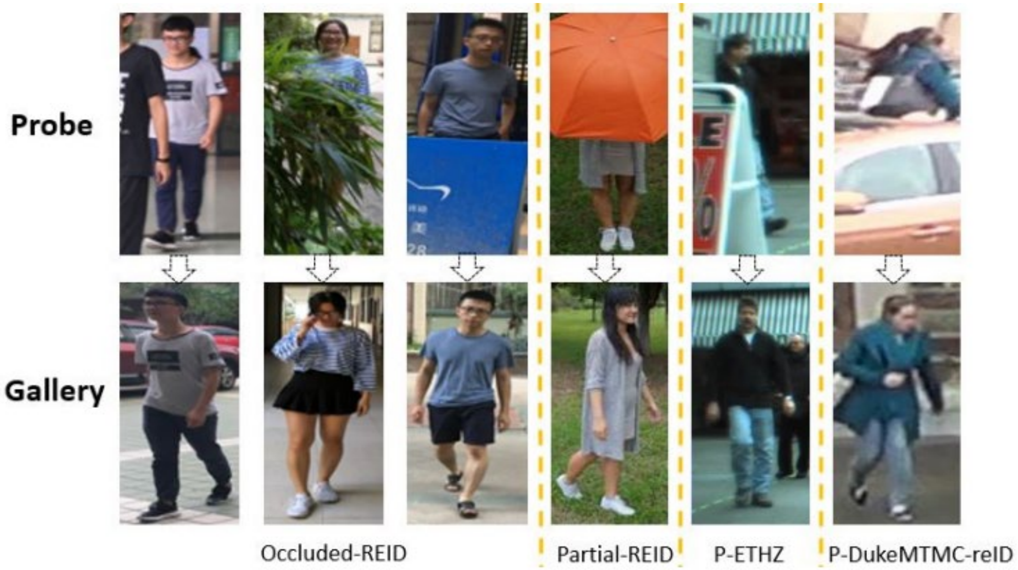
# Conclusion

- Person re-ID aims to retrieve a person corresponding to a given query across disjoint camera views or different time stamps.
  - Large intra-variation & Small inter-variation, Open-set, Data collection, …

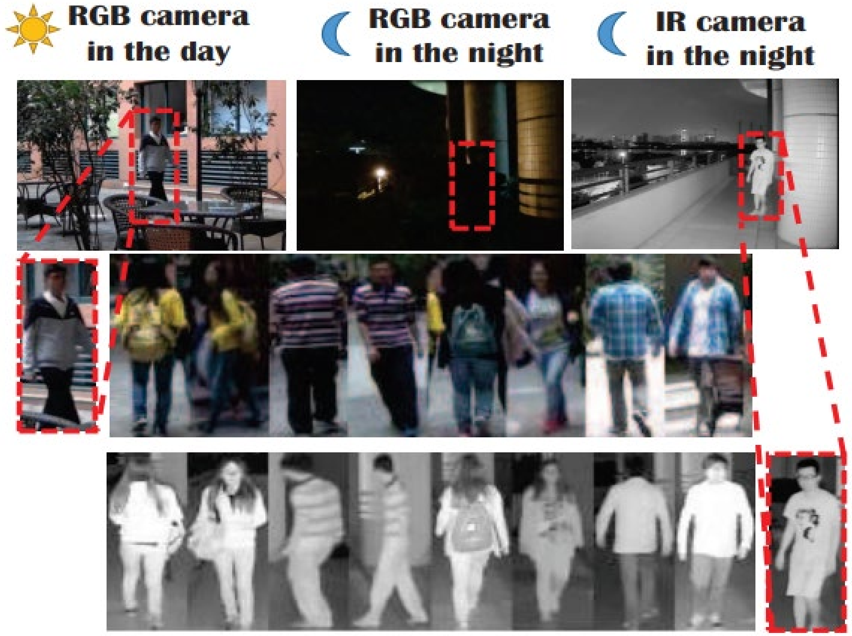
- Supervised approaches.
  - Part feature learning, Attention mechanism, Feature disentanglement, …
- Unsupervised approaches.
  - Pseudo-label refinement, Cluster-level contrastive learning, …
- Domain generalizable approaches.
  - Style normalization, Meta learning, …

# Conclusion

- Other research topics for practical usage of person re-ID techniques.
  - Continual (incremental) learning.
  - Occluded person re-ID.
  - Cross-modality person re-ID.




Occluded person re-ID




Cross-modality person re-ID

Zhuo et al. Occluded Person Re-identification. In ICME 2018.
Wu et al. RGB-Infrared Cross-Modality Person Re-Identification. In ICCV 2017.

# Q&A

**Thank You!**

SGVR Lab KAIST