New Trends of Person re-ID System

Zheng Wang

National Institute of Informatics, Japan

Outline

- General Person Re-identification
- Person Re-identification New Trends
    - Cross-modality
    - Long-term
    - Group
- Discussion

Person Re-identification

XXX Case in Nanjing, China

Search for XXX

1500 Police, One Month

329 shots

Person Re-identification

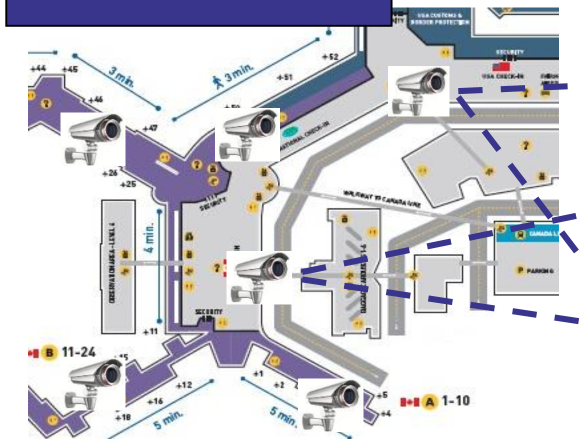
XXX Case in Nanjing, China

Search for XXX

1500 Police, One Month

329 shots

Person Re-identification



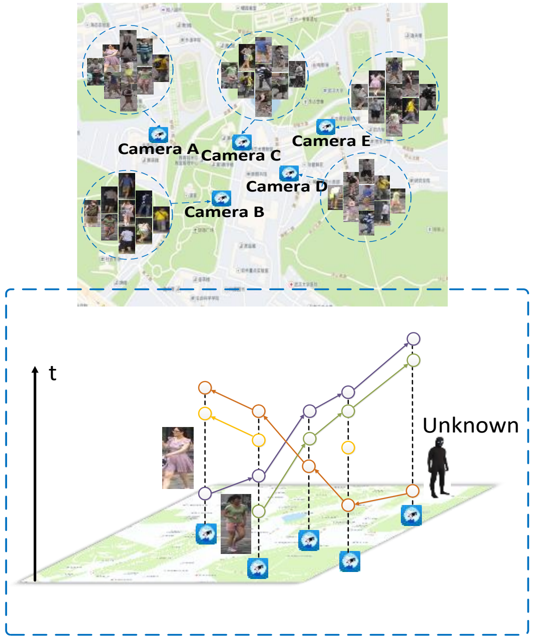
Same person?

Camera a

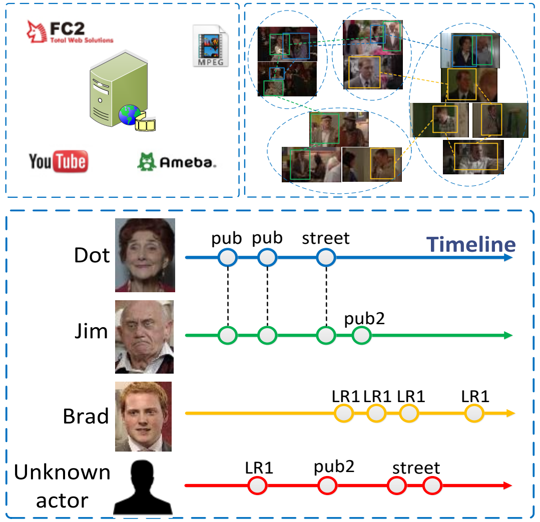
Camera b

Person Re-identification

- Image Retrieval / Instance Search
  - Target: Certain Person

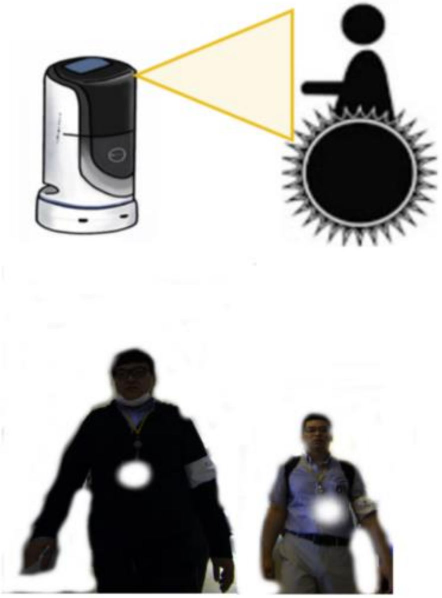


Surveillance Camera [1]



TV Media [2]

First Person Vision [3]



Robot Vision

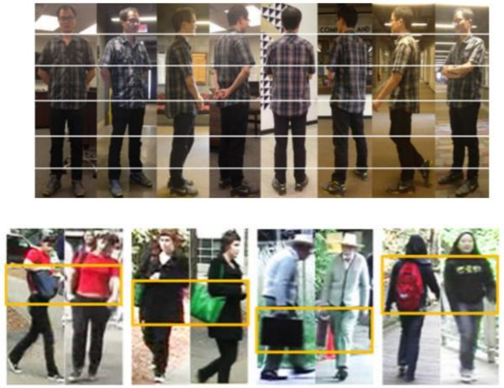
[1] Wang, et al., Incremental Re-identification by Cross-Direction and Cross-Ranking Adaption, TMM, 2019
[2] Fischer, et al., Person re-identification in tv series using robust face recognition and user feedback, MTAP, 2011
[3] Chakraborty, et al., Person re-identification using multiple first-person-views on wearable devices, WACV, 2016

General Person Re-identification

- Challenges
  - Appearance changes / No Clothes change
  - Intra-Modality

Occlusion [4]

Illumination [5]



Viewpoint [6]

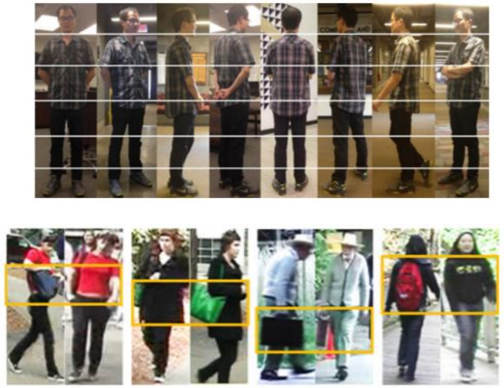
[4] Luo, et al., STNReID: Deep Convolutional Networks with Pairwise Spatial Transformer Networks for Partial Person Re-identification, TMM, 2020
[5] Zeng, et al., Illumination-Adaptive Person Re-identification, TMM, 2020
[6] Wu, et al., Viewpoint Invariant Human Re-Identification in Camera Networks Using Pose Priors and Subject-Discriminative Features, TPAMI, 2014

General Person Re-identification

- Challenges
  - Appearance changes / No Clothes change
  - Intra-Modality

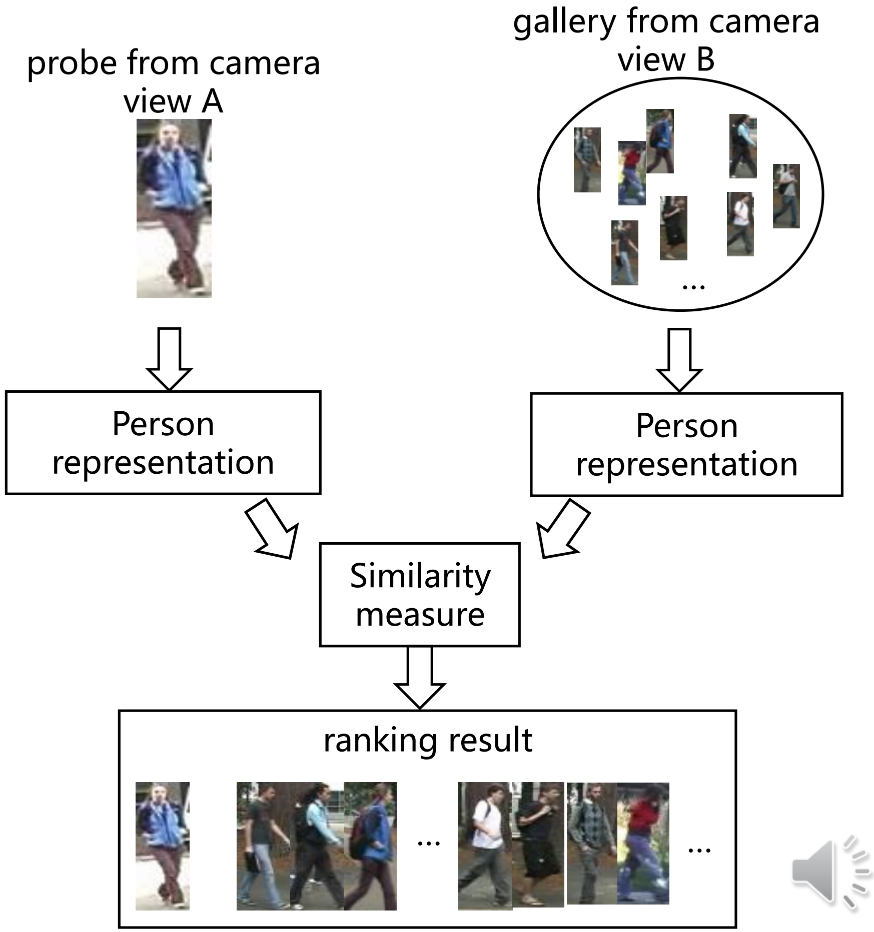
Occlusion [4]

Illumination [5]



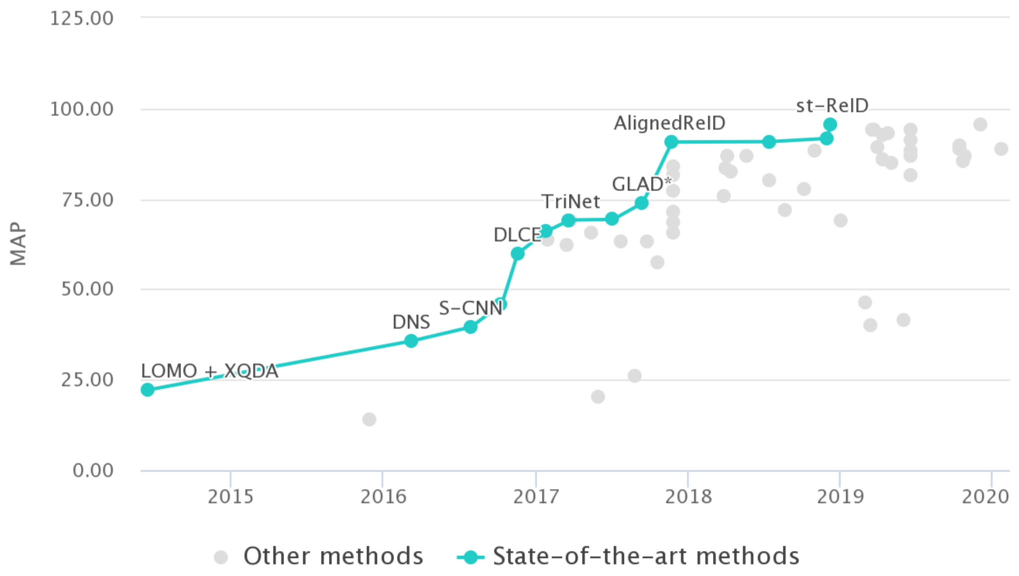
Viewpoint [6]

Diagram

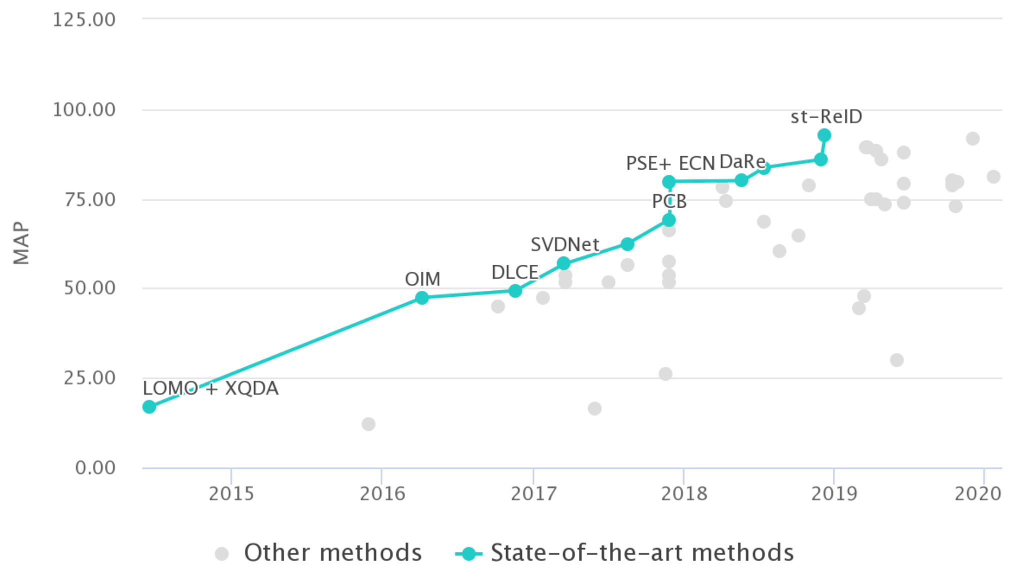


[4] Luo, et al., STNReID: Deep Convolutional Networks with Pairwise Spatial Transformer Networks for Partial Person Re-identification, TMM, 2020
[5] Zeng, et al., Illumination-Adaptive Person Re-identification, TMM, 2020
[6] Wu, et al., Viewpoint Invariant Human Re-Identification in Camera Networks Using Pose Priors and Subject-Discriminative Features, TPAMI, 2014

General Person Re-identification



Market-1501 [7]



DukeMTMC-reID [8]

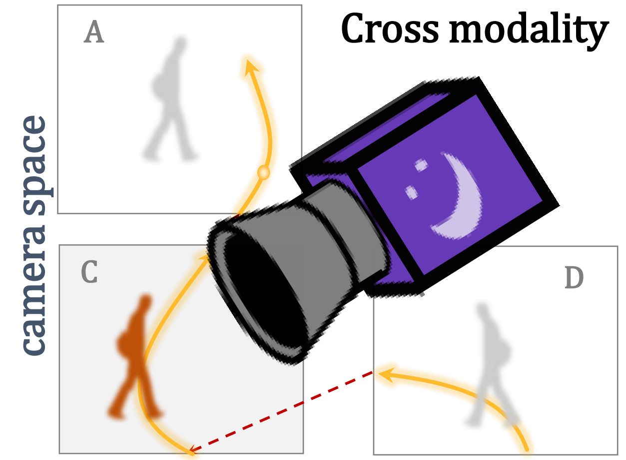
- Rank-1 accuracy surpass the human performance [9]
  - Intra-modality discrepancy has been well addressed
  - daytime, visible spectrum, sufficient details

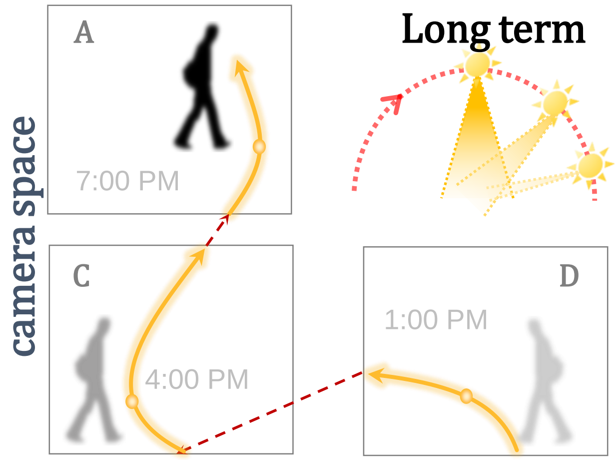
[7] https://paperswithcode.com/sota/person-re-identification-on-market-1501
[8] https://paperswithcode.com/sota/person-re-identification-on-dukemtmc-reid
[9] Zhang, et al., AlignedReID: Surpassing Human-Level Performance in Person Re-Identification, arXiv, 2018

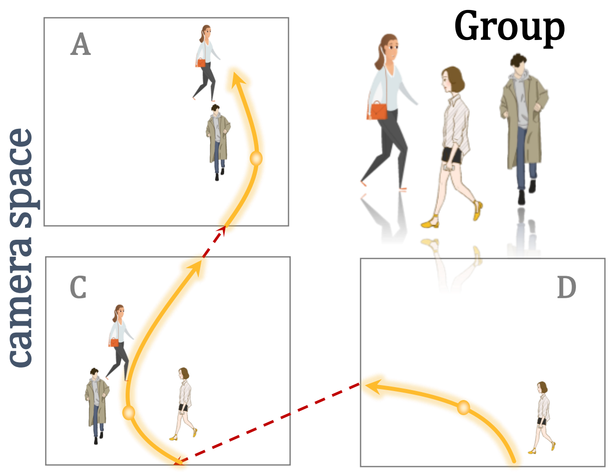
Outline

- General Person Re-identification
- Person Re-identification New Trends
  - Cross-modality
  - Long-term
  - Group
- Discussion

Person Re-identification New Trends







Cross-modality

different camera specifications and settings
(low- vs. high resolution data)

different sensory devices
(infrared vs. visible light devices)

reproduction of human memory and direct recording by a camera
(sketch/text description vs. digital images)

Cross-modality

	Main Focus	Feature
[10]	Gait sequences	a special and different focus
[11]	Appearance	a multi-dimensional overview
[12]	Appearance	a systematic evaluation with different features and metrics
[13]	Appearance	a limited summary of current efforts or problems present in different modalities

[10] Nambiar, et al., Gait-based person re-identification: A survey. ACM Computing Surveys, 2019
[11] Vezzani, et al., People reidentification in surveillance and forensics: A survey. ACM Computing Surveys, 2013.
[12] Gou, et al. A systematic evaluation and benchmark for person re-identification: Features, metrics, and datasets. TPAMI, 2018
[13] Leng, et al., A survey of open-world person re-identification. TCSVT, 2019.

Cross-modality

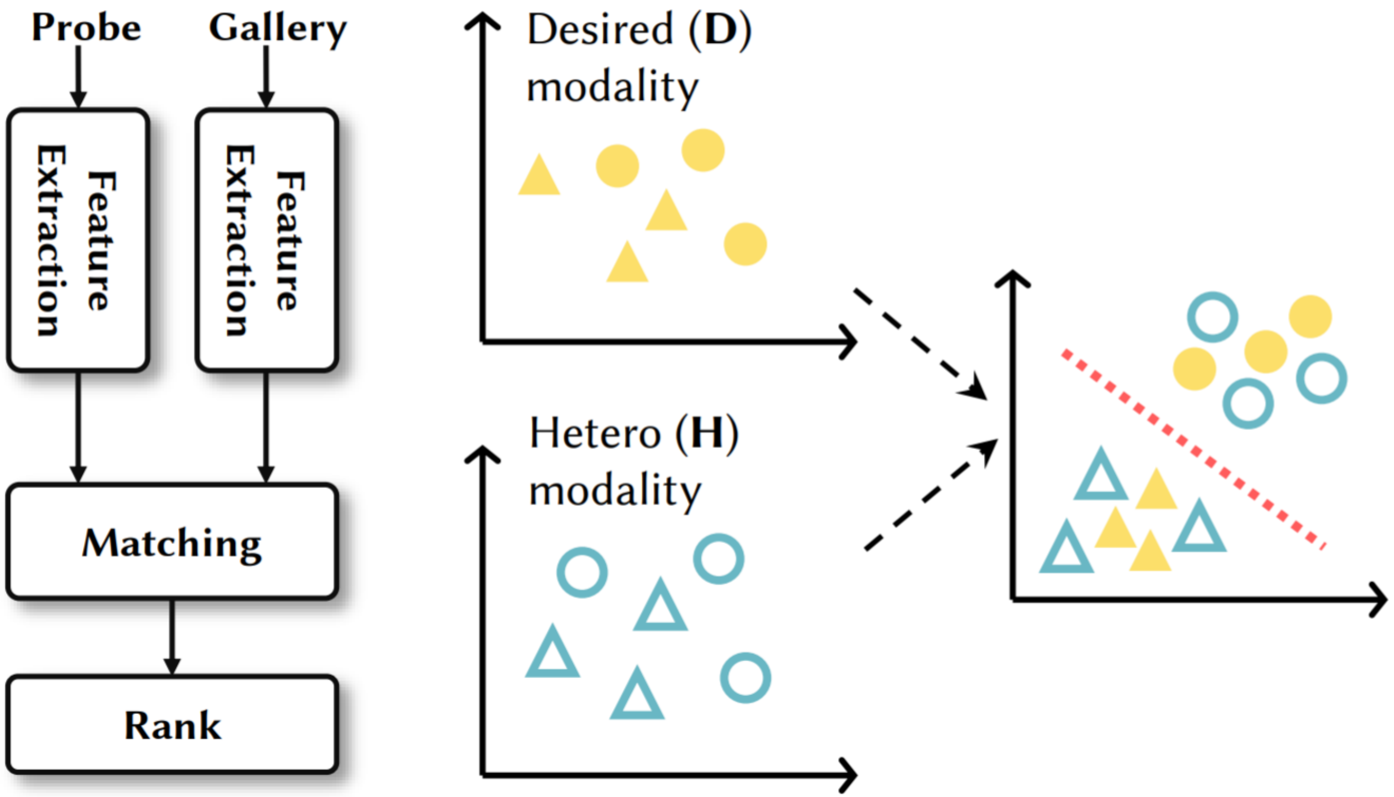
	Main Focus	Feature
[10]	Gait sequences	a special and different focus
[11]	Appearance	a multi-dimensional overview
[12]	Appearance	a systematic evaluation with different features and metrics
[13]	Appearance	a limited summary of current efforts or problems present in different modalities

	Homo-ReID	**Hetero-ReID**
Media Type	Desired image	+ LR / IR / Sketch / Text
Participant	Machine	Machine (+Human)
Main Challenge	Intra-modality	Intra- + Inter-modality
#Publications	>1000	<100
Performance	96.1	42.50 / 28.90 / 34.00 / 53.14

There is also a big performance gap between Homo-ReID and Hetero-ReID.

[10] Nambiar, et al., Gait-based person re-identification: A survey. ACM Computing Surveys, 2019
[11] Vezzani, et al., People reidentification in surveillance and forensics: A survey. ACM Computing Surveys, 2013.
[12] Gou, et al. A systematic evaluation and benchmark for person re-identification: Features, metrics, and datasets. TPAMI, 2018
[13] Leng, et al., A survey of open-world person re-identification. TCSVT, 2019.

The Diagram


Probe
Gallery
Feature Extraction
Feature Extraction
Matching
Rank
Desired (**D**) modality
Hetero (**H**) modality

Datasets

No.	Dataset	Appl.	Type	#Cam.	#ID	#Sam.
	Market-1501 [Zheng *et al.*, 2015a]	Desired	Real	6	1,501	32,668
	MSMT17 [Wei *et al.*, 2018]	Desired	Real	15	4,101	126,441
1	CAVIAR [Cheng *et al.*, 2011]	LR	Real	2	50	1,000
2	LR-VIPeR [Li *et al.*, 2015]	LR	Simulated	2	632	1,264
3	LR-3DPES [Li *et al.*, 2015]	LR	Simulated	8	192	1,011
4	LR-i-LIDS [Jing *et al.*, 2015]	LR	Simulated	2	119	238
5	LR-PRID [Jing *et al.*, 2015]	LR	Simulated	2	100	200
6	SALR-VIPeR [Wang *et al.*, 2016b]	LR	Simulated	2	632	1,264
7	SALR-PRID [Wang *et al.*, 2016b]	LR	Simulated	2	450	900
8	MLR-VIPeR [Jiao *et al.*, 2018]	LR	Simulated	2	632	1,264
9	MLR-SYSU [Jiao *et al.*, 2018]	LR	Simulated	2	502	3,012
10	MLR-CUHK03 [Jiao *et al.*, 2018]	LR	Simulated	2	1,467	14,000
11	SYSU-MM01 [Wu *et al.*, 2017b]	IR	Real	6	491	38,271
12	RegDB [Nguyen *et al.*, 2017]	IR	Real	2	412	8,240
13	PKU-Sketch [Pang *et al.*, 2018]	Sketch	Real	2	200	400
14	CUHK-PEDES [Li *et al.*, 2017b]	Text	Real	–	13,003	80,412

Market-1501

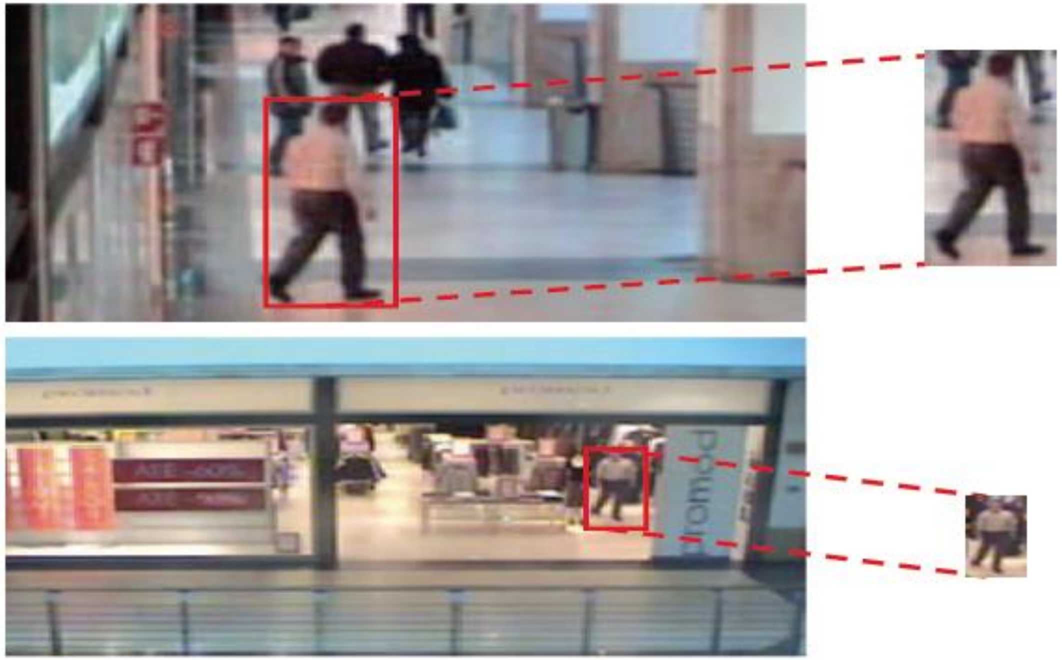
MLR-VIPeR

SYSU-MM01

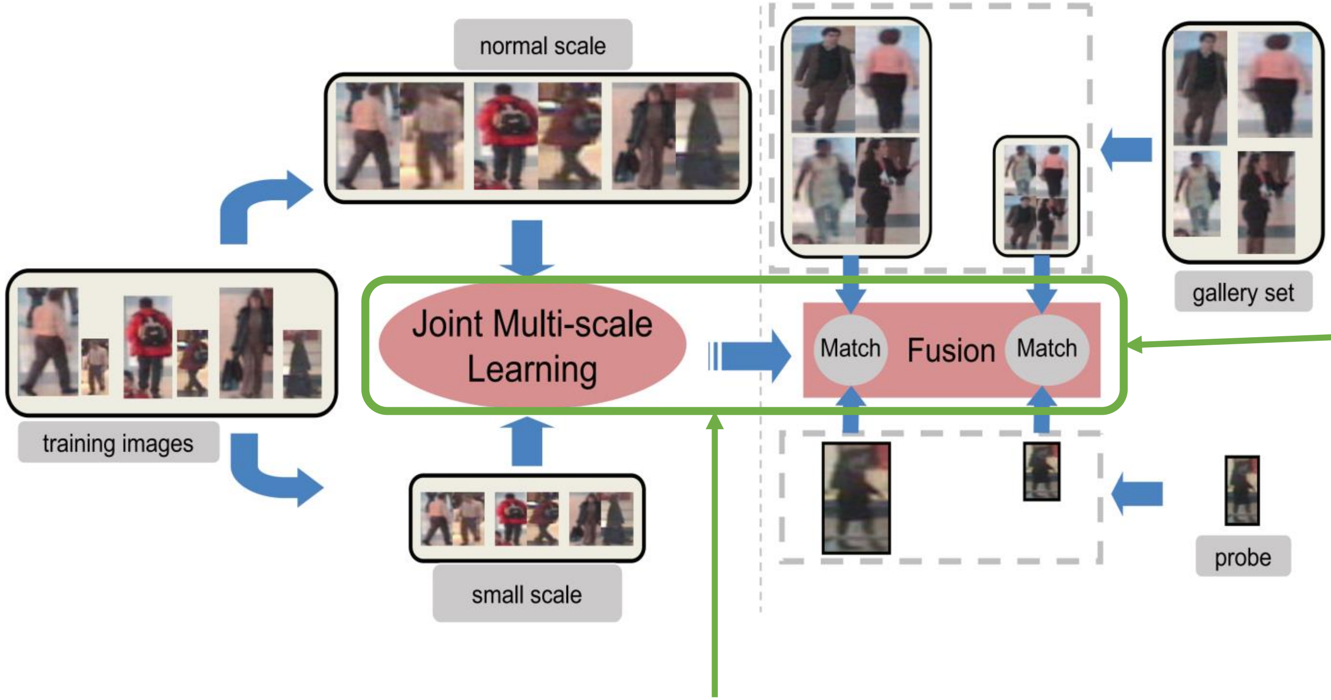
CUHK-PEDES

PKU-Sketch

LR-HR



LR-1-JUDEA [14]



- Contributions
  - The first work focusing on the LR ReID
  - a multi-scale discriminant distance metric learning model
  - Existing ReID models have a clear performance drop at the LR task, but the proposed method does not.

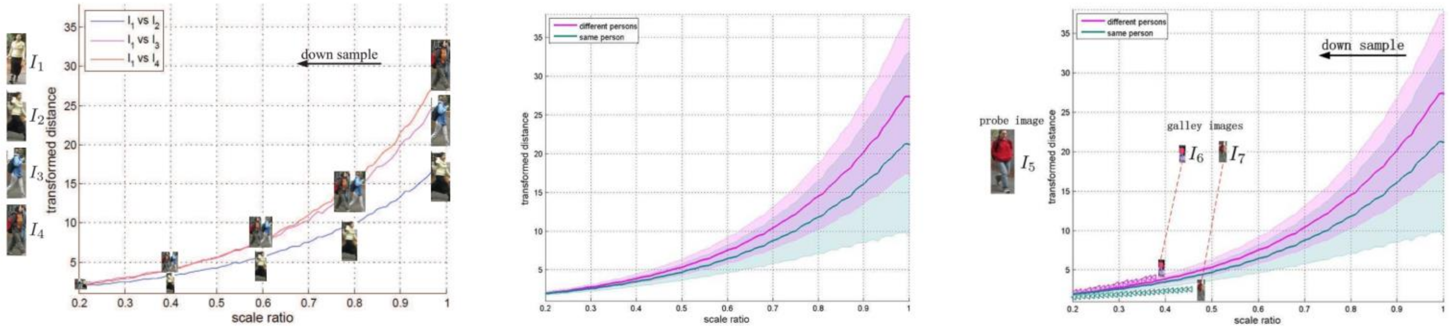
Heterogeneous class mean discrepancy (HCMD)

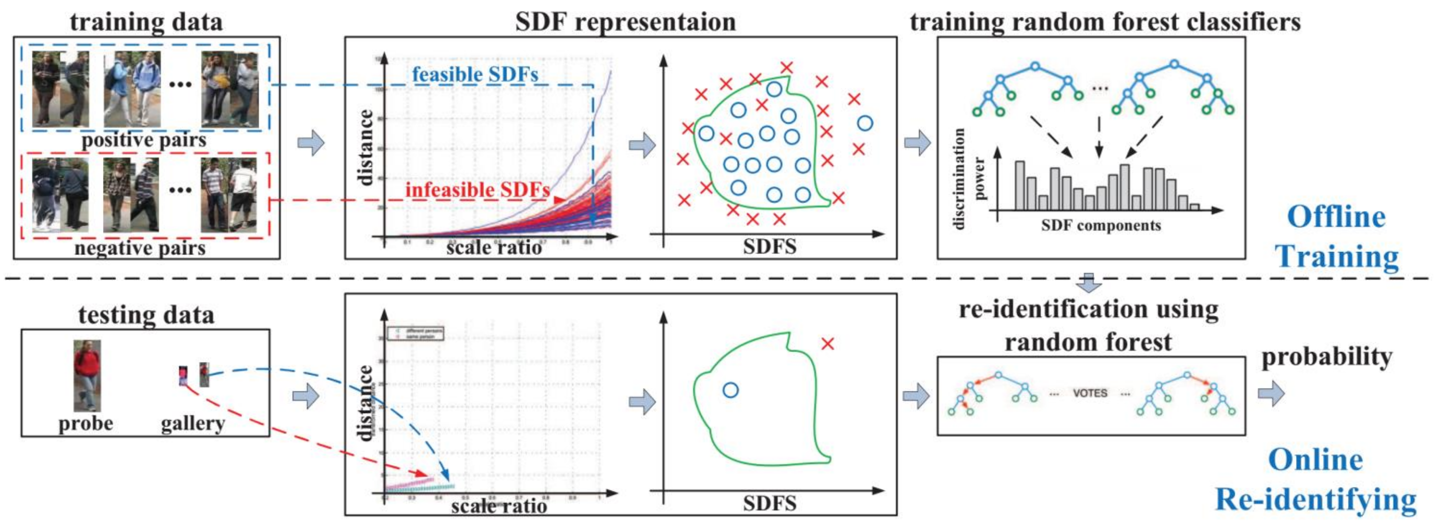
$$\min_{\boldsymbol{W}_h,\boldsymbol{W}_s} \mathrm{HCMD}(\boldsymbol{W}_h,\boldsymbol{W}_s) = \frac{1}{C}\sum_{i=1}^{C} ||\boldsymbol{W}_h^T\boldsymbol{u}_i^h - \boldsymbol{W}_s^T\boldsymbol{u}_i^s||_2^2$$

[14] Li, et al., Multi-scale Learning for Low-resolution Person Re-identification, ICCV, 2015

Metric Learning

LR-2-SDF [15]

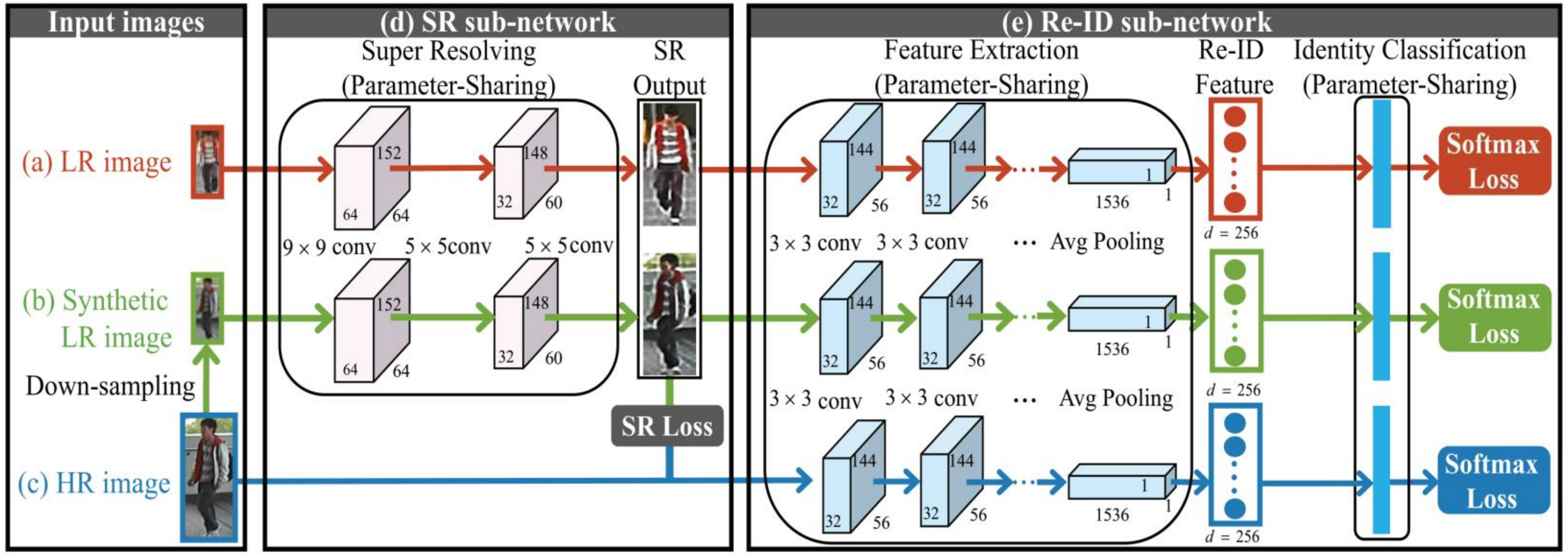




- Contributions
  - a new issue - Scale-adaptive Low-resolution Person Re-identification
  - the discriminating power of the feasible and infeasible SDFs respectively generated by positive and negative image pairs

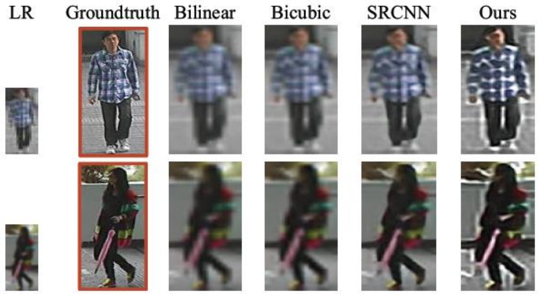
[15] Wang, et al., Scale-Adaptive Low-Resolution Person Re-Identification via Learning a Discriminating Surface, IJCAI, 2016

LR-3-SING [16]



- Contributions
  - image SR and ReID techniques in a novel unified formulation
  - a joint loss function on optimsing a hybrid CNN architecture
  - a multi-resolution adaptive fusion mechanism by aggregating a set of anchor SING CNN models

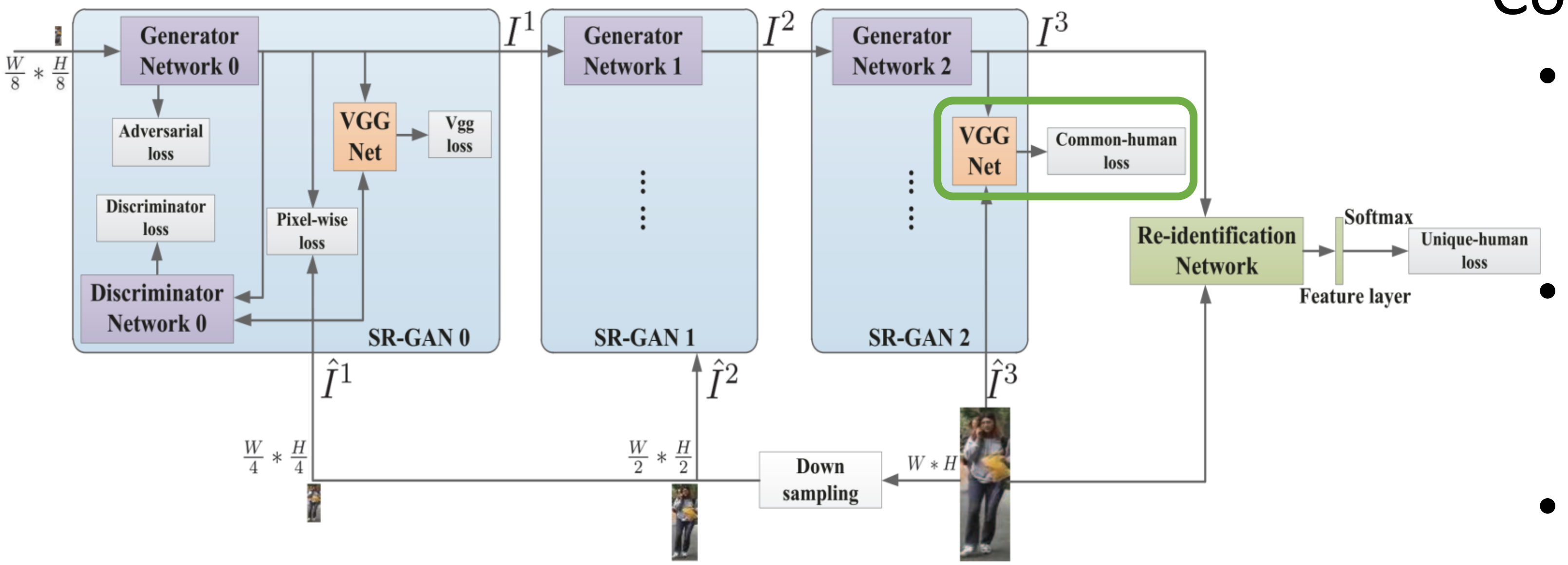
$$L\left(\{(\boldsymbol{x}_i^l, \boldsymbol{x}_i^h, y_i^l, y_i^h)\}_{i=1}^N\right) =$$
$$L_{reid}\left(\{(\boldsymbol{x}_i^l, \boldsymbol{x}_i^h, y_i^l, y_i^h)\}_{i=1}^N\right) + \alpha L_{sr}\left(\{\boldsymbol{x}_i^h\}_{i=1}^N\right)$$



[16] Jiao, et al., Deep Low-Resolution Person Re-Identification, AAAI, 2018

Modality Unification

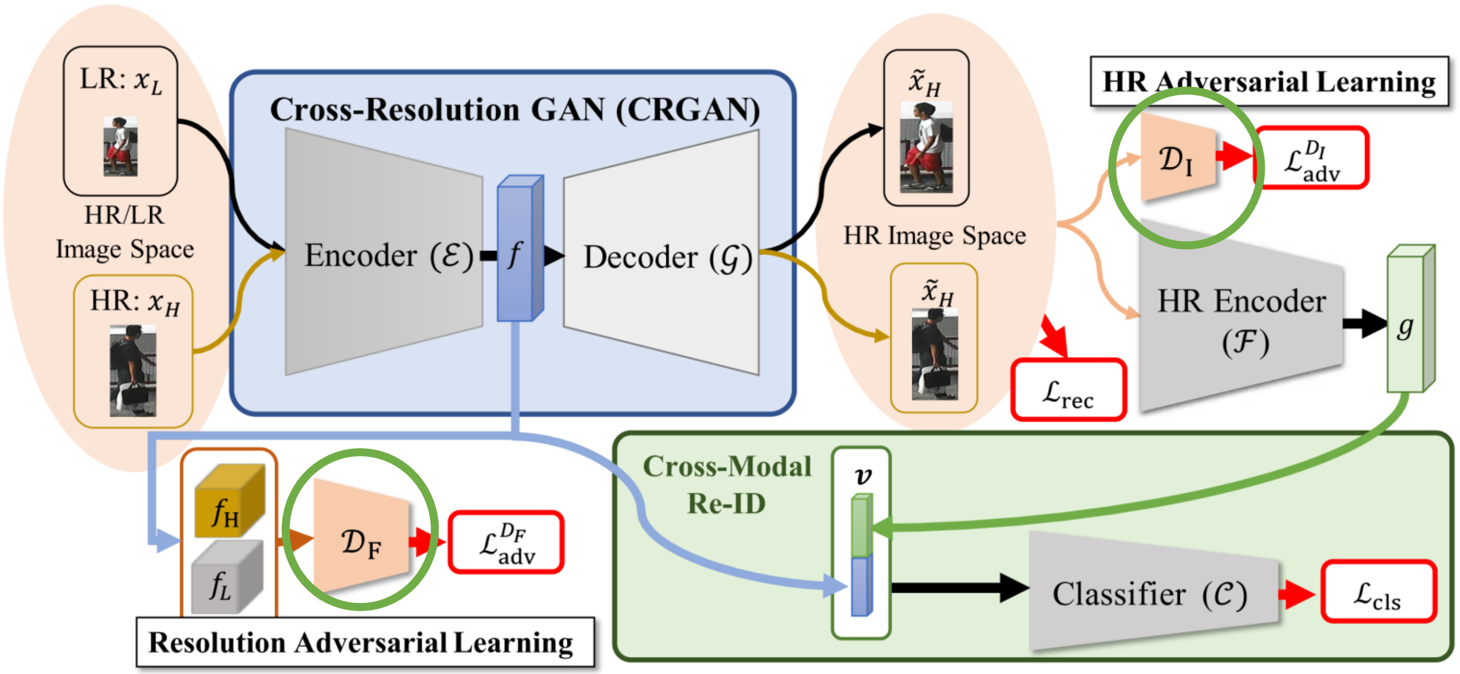
LR-4-CSR-GAN [17]



- Contributions
  - cascade multiple SRGANs in series, capable of super-resolving LR images with multi-scale upscaling
  - the integration compatibility between scale-adaptive super-resolution and re-identification
  - a common-human loss to make the super-resolved image look more like human

[17] Wang, et al., Cascaded SR-GAN for Scale-Adaptive Low Resolution Person Re-identification, IJCAI, 2018

LR-5-CAD [18]



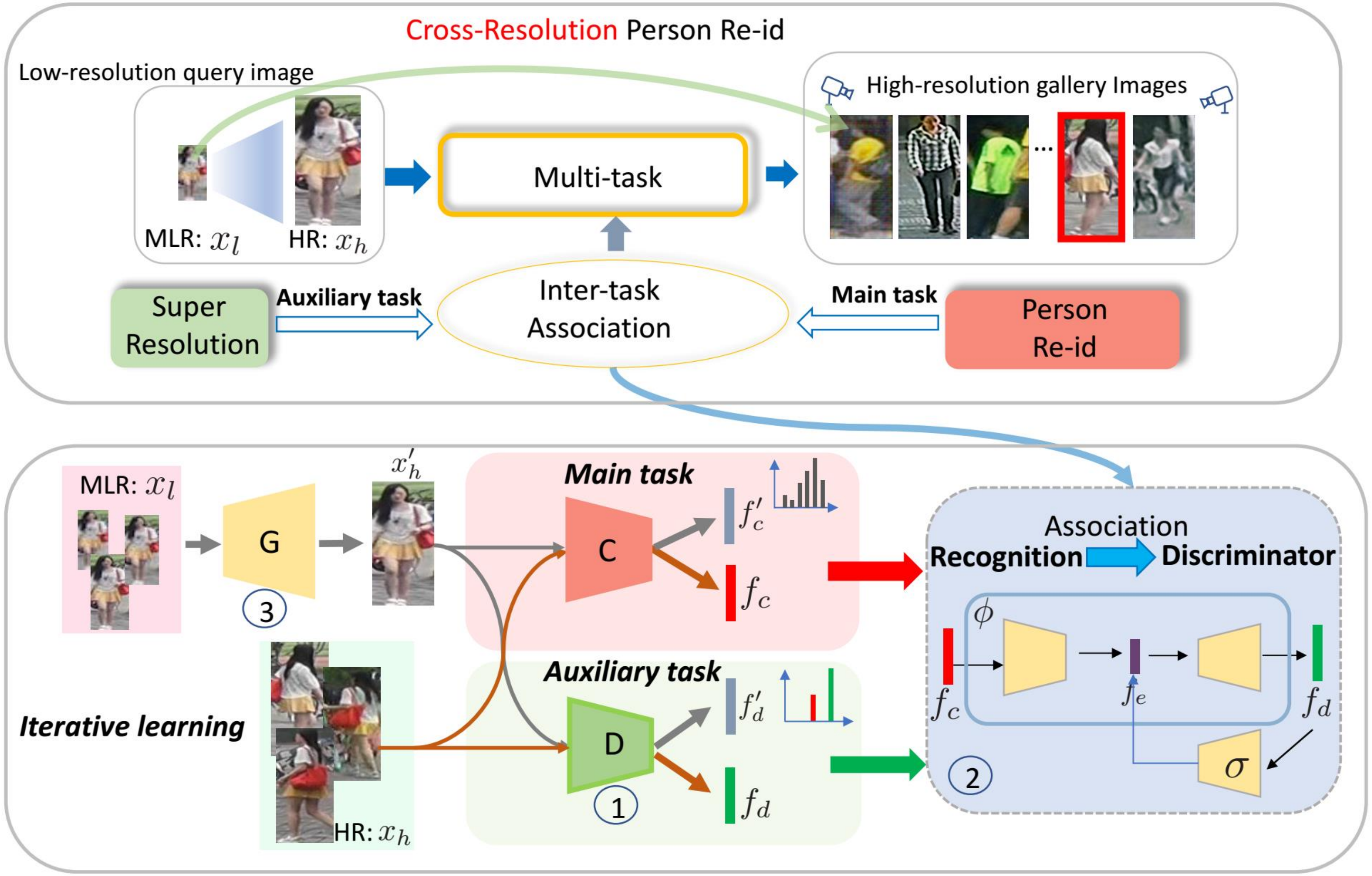
Contributions

- advances adversarial learning strategies
- learns resolution-invariant representations while recovering the missing details in LR input images

Modality Unification + Representation Learning

[18] Li, et al., Recover and Identify: A Generative Dual Model for Cross-Resolution Person Re-Identification, ICCV, 2019

LR-6-INTACT [19]

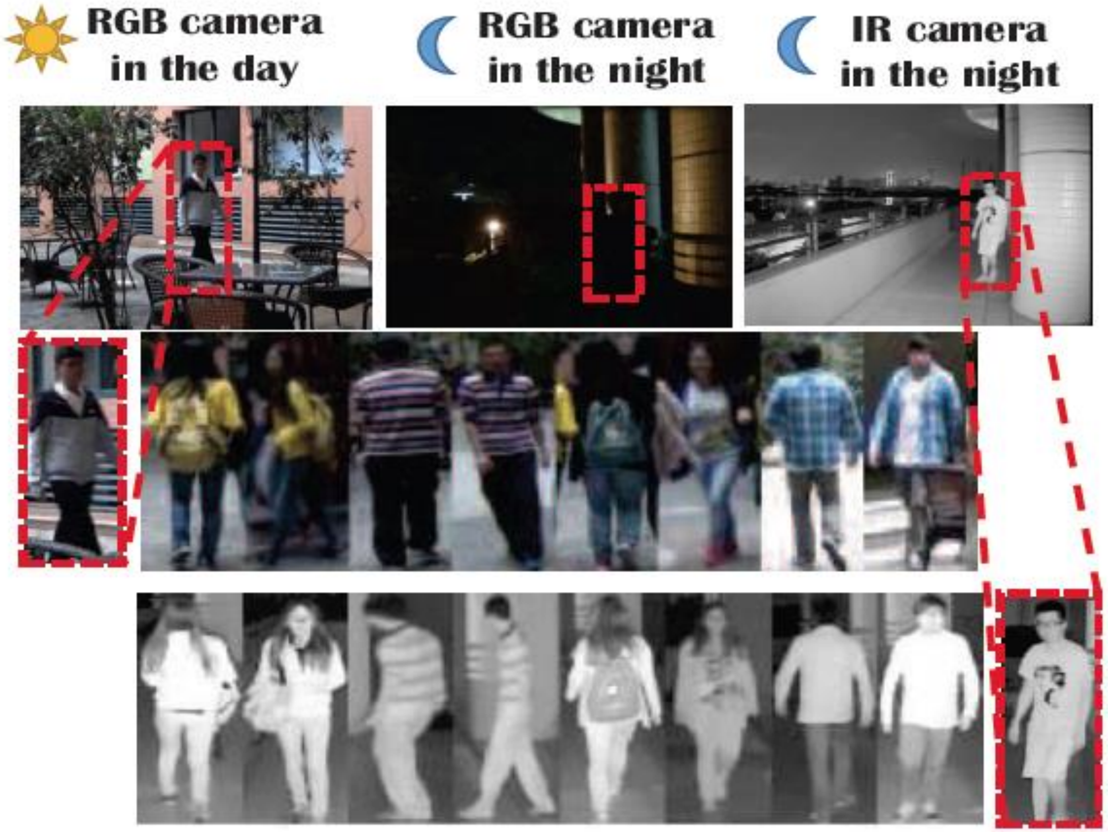


- Contributions
  - an idea of leveraging the association between image SR and person re-id tasks
  - a regularisation method implements the proposed inter-task association

Modality Unification + Representation Learning

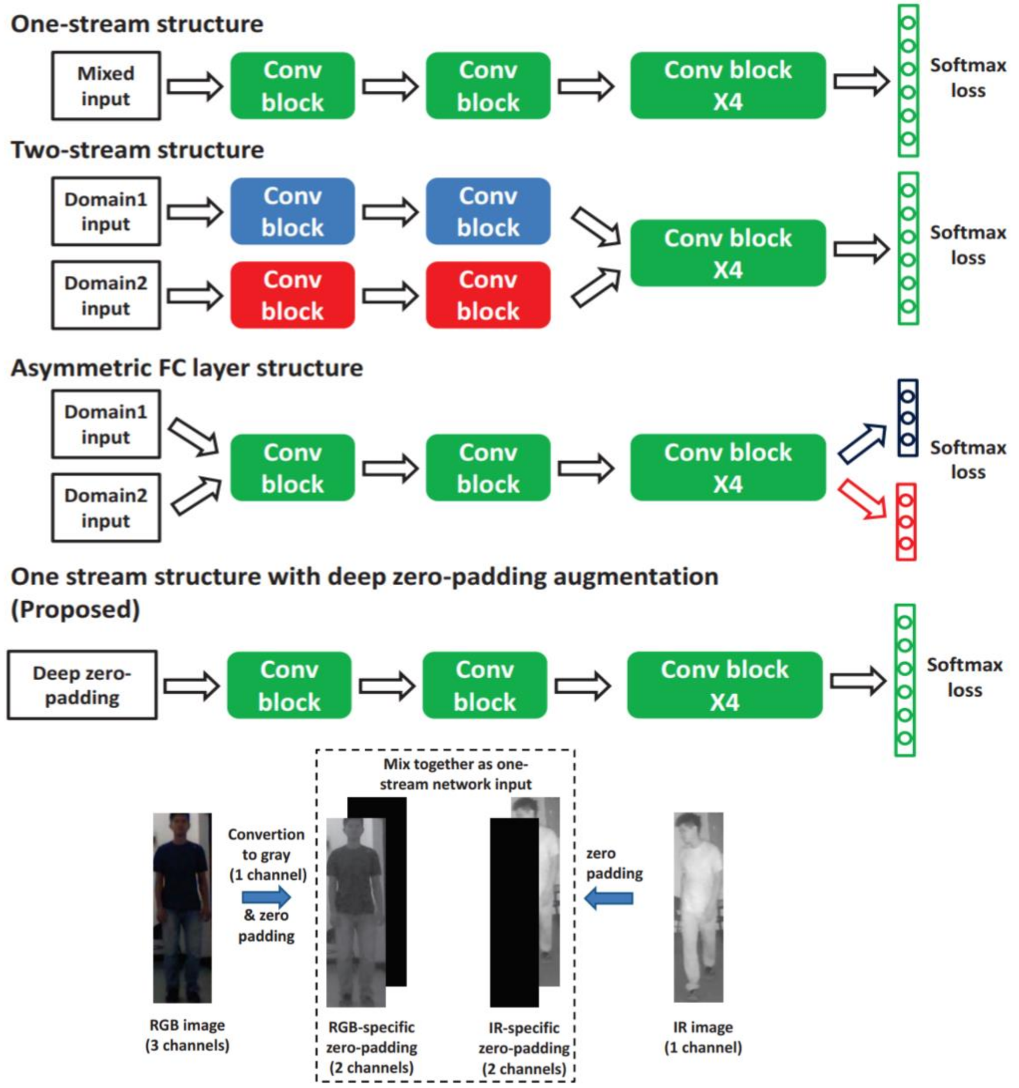
[19] Cheng, et al., Inter-Task Association Critic for Cross-Resolution Person Re-Identification, CVPR, 2020

IR-RGB



RGB camera
in the day
RGB camera
in the night
IR camera
in the night

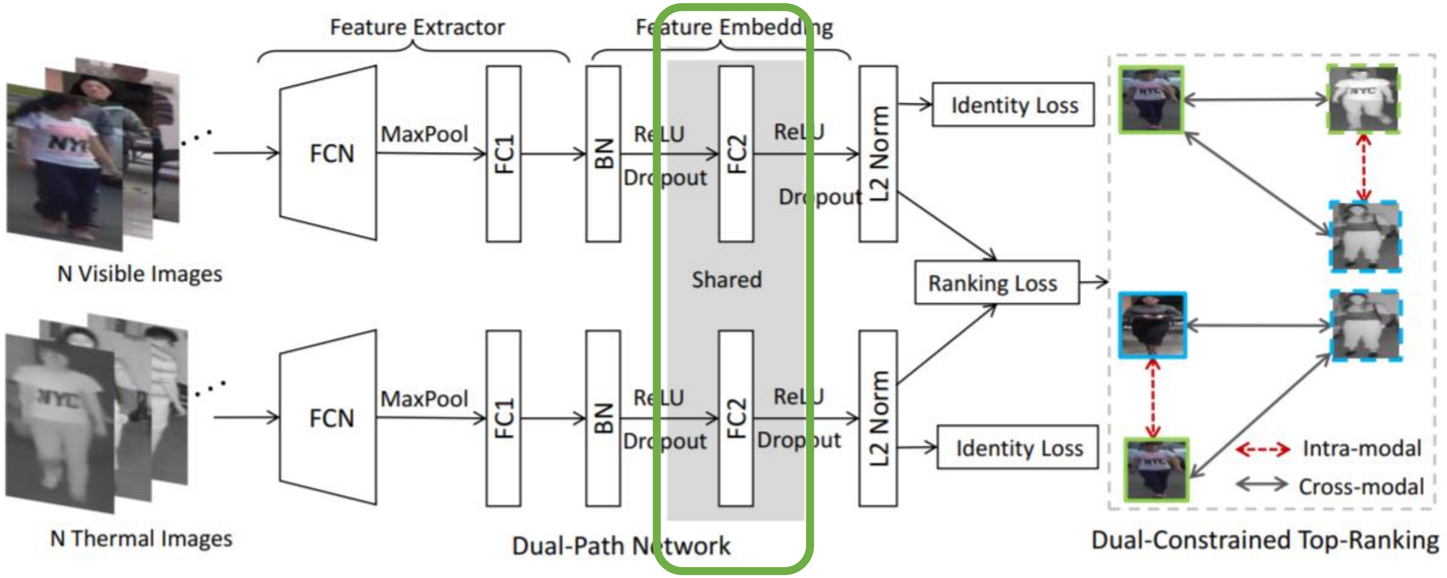
IR-1-Zero-padding [20]



- Contributions
  - study the RGB-IR Re-ID for the first time and raise a standard benchmark
  - analyse three different network structures (one-stream, two-stream and asymmetric FC layer)
  - deep zero-padding

[20] Wu, et al., RGB-Infrared Cross-Modality Person Re-Identification, ICCV, 2017

IR-2-BDTR [21]



$$\mathcal{L}_{cross} = \sum_{\forall y_i = y_j} \max[\rho_1 + D(x_i, z_j) - \min_{\forall y_i \neq y_k} D(x_i, z_k), 0] + \sum_{\forall y_i = y_j} \max[\rho_1 + D(z_i, x_j) - \min_{\forall y_i \neq y_k} D(z_i, x_k), 0]$$

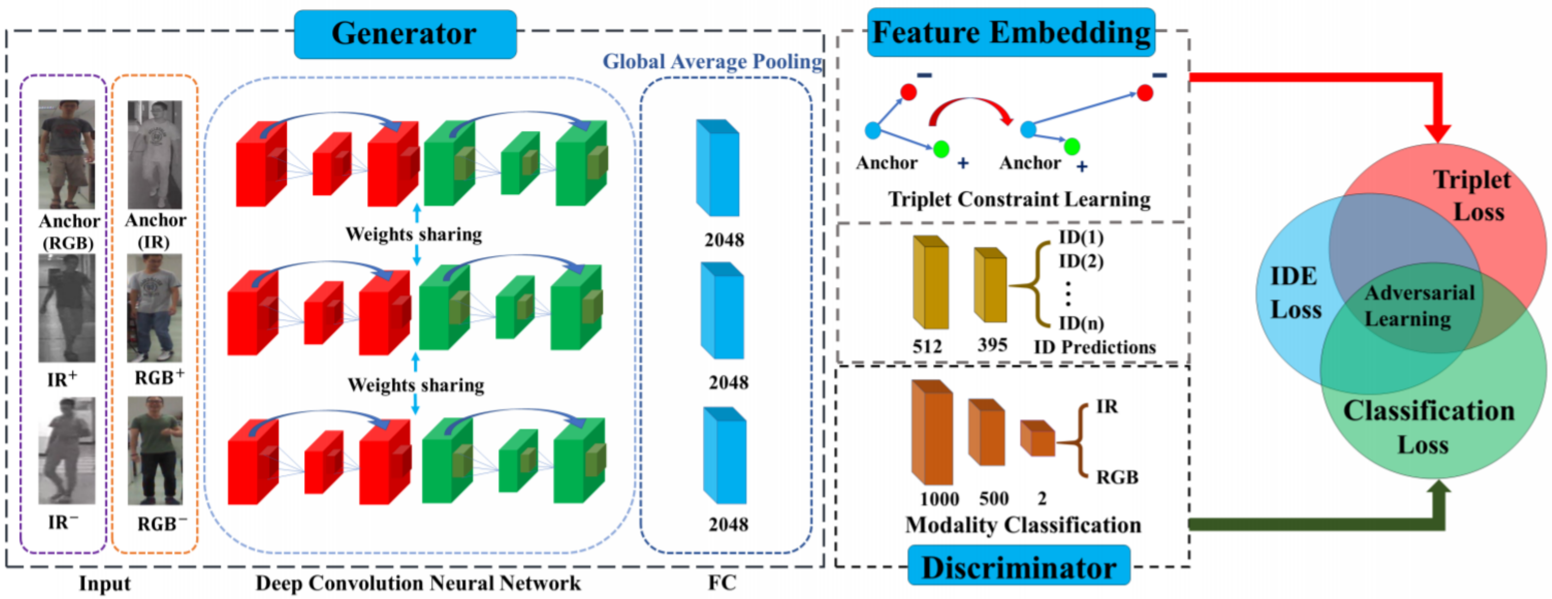
$$\mathcal{L}_{intra} = \sum \max[\rho_2 - D(z_j, z_k), 0] + \sum \max[\rho_2 - D(x_j, x_k), 0]$$

- Contributions
  - an end-to-end dual-path feature and metric learning framework
  - a bi-directional dual-constrained top-ranking loss to simultaneously consider the cross-modality and intra-modality variations

[21] Ye, et al., Visible thermal person re-identification via dual constrained top-ranking, IJCAI, 2018

Representation Learning

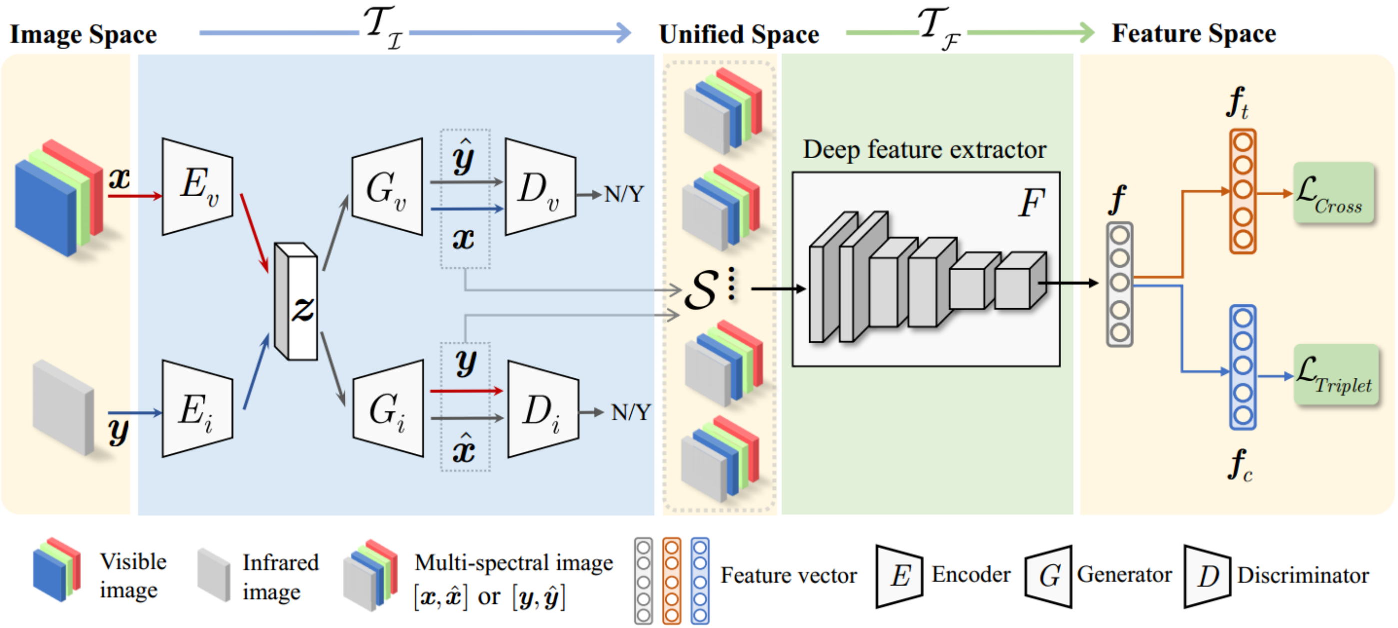
IR-3-cmGAN [22]

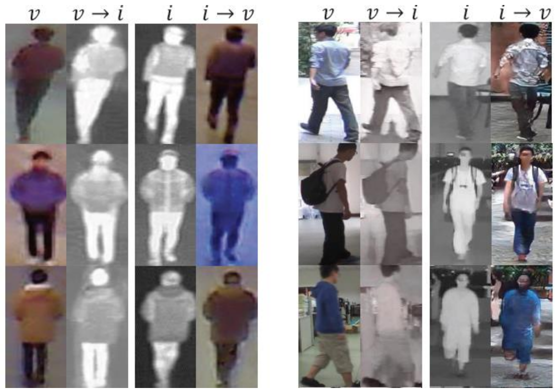


- Contributions
  - a loss function for cross-modality generative adversarial network
  - identification loss and cross-modality triplet loss together for generator
  - a modality classifier as discriminator

[22] Dai, et al., Cross-Modality Person Re-Identification with Generative Adversarial Training, IJCAI, 2018

IR-4-D^{2}RL [23]

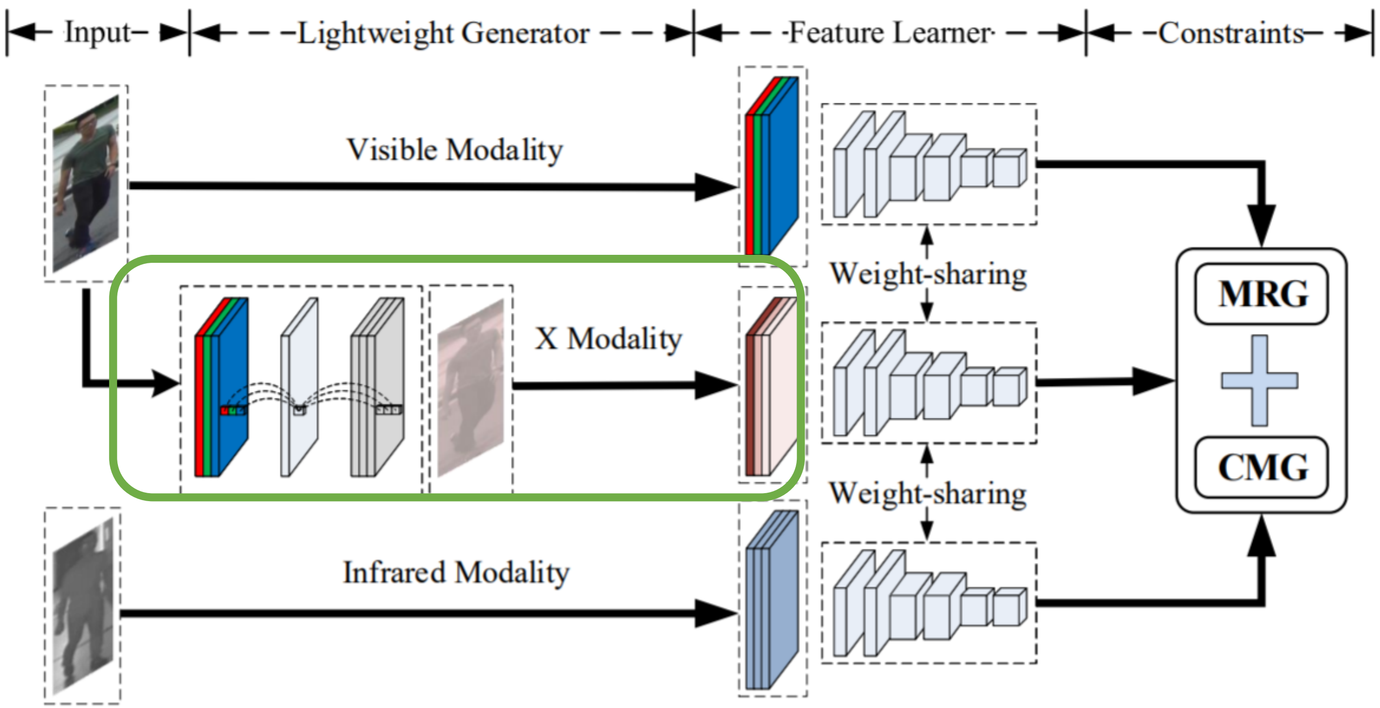




- Contributions
  - A dual-level discrepancy reduction learning scheme. the first to decompose the mixed modality and appearance discrepancies.
  - An end-to-end scheme enforces these two sub-networks benefit each other.

[23] Wang, et al., Learning to Reduce Dual-level Discrepancy for Infrared-Visible Person Re-identification, CVPR, 2019

IR-5-XIV [24]



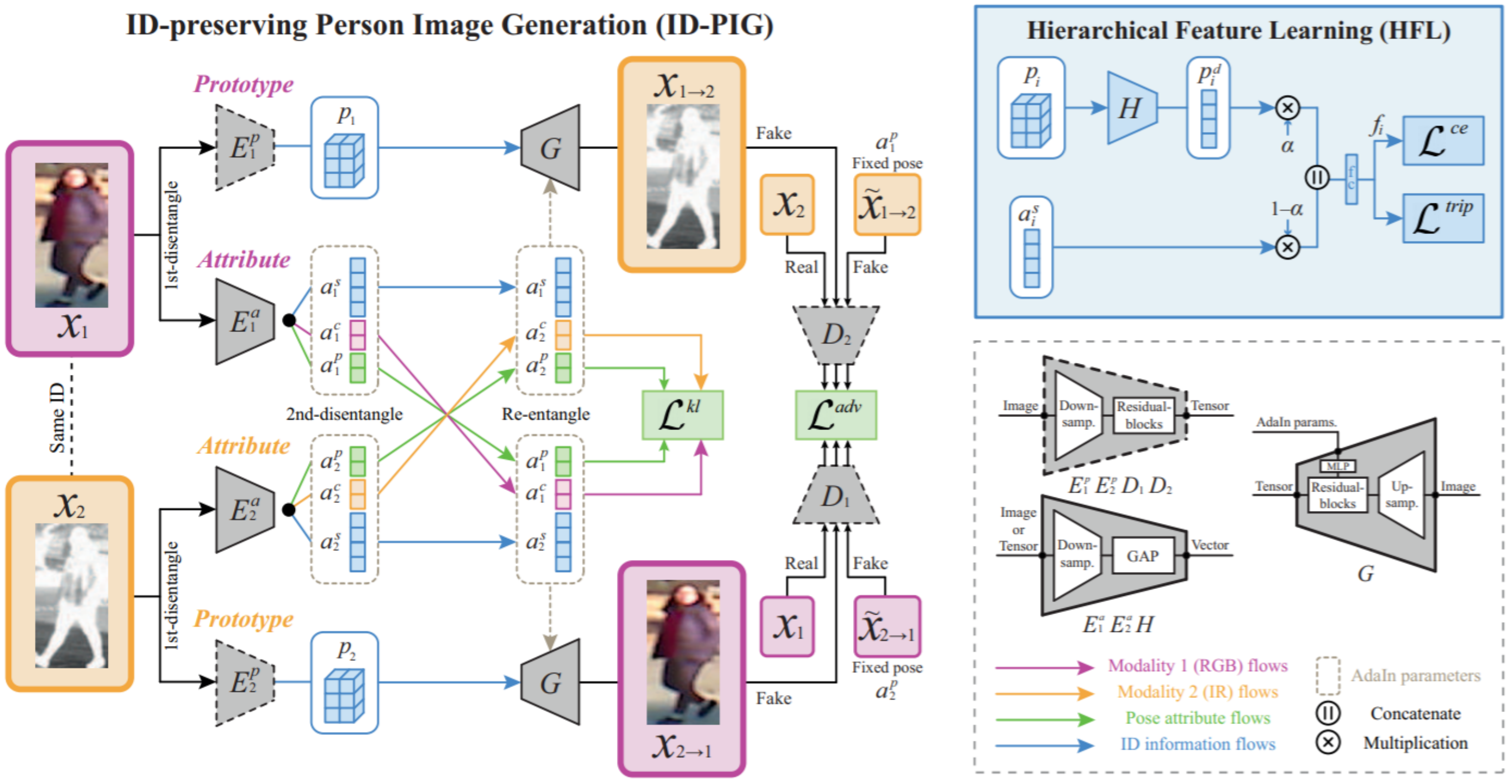
cross modality gap (CMG) and the modality respective gap (MRG)

- Contributions
  - an adjoint and auxiliary X modality.
  - an extra lightweight network to generate the X modality through self-supervised learning
  - a modality gap constraint to direct the learning and knowledge communication across modalities

[24] Li, et al., Infrared-Visible Cross-Modal Person Re-Identification with an X Modality, AAAI, 2020

Representation Learning

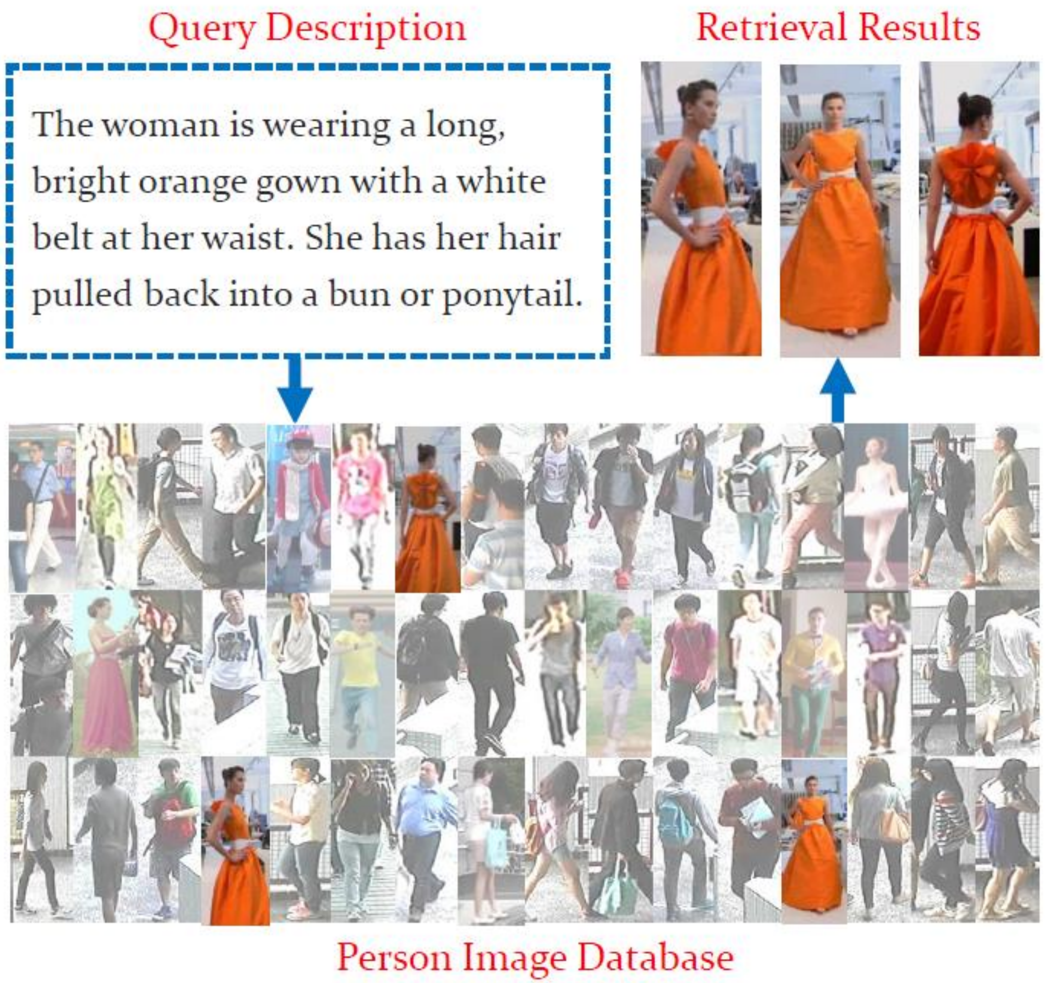
IR-6-Hi-CMD [25]



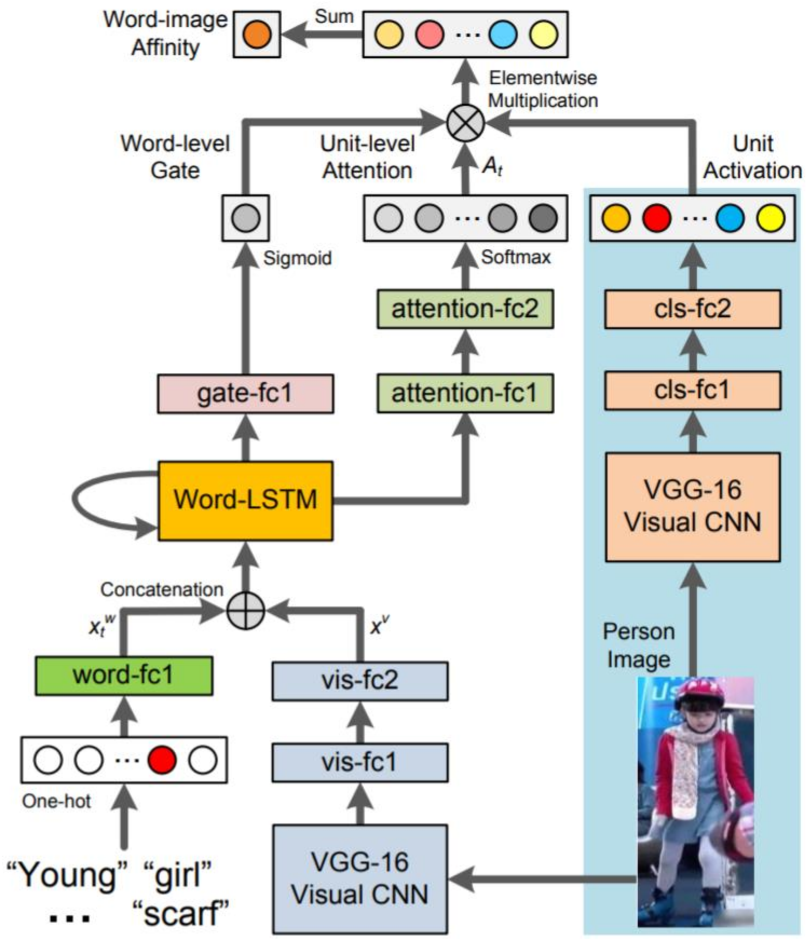
- Contributions
  - A Hierarchical Cross-Modality Disentanglement (Hi-CMD) method extracts pose- and illumination-invariant features for cross-modality matching.
  - The proposed ID-preserving Person Image Generation (ID-PIG) network changes the pose and illumination attributes while maintaining the identity characteristic of a specific person.

[25] Choi, et al., Hi-CMD: Hierarchical Cross-Modality Disentanglement for Visible-Infrared Person Re-Identification, CVPR, 2020

Text-Image


Query Description
The woman is wearing a long, bright orange gown with a white belt at her waist. She has her hair pulled back into a bun or ponytail.
Retrieval Results
Person Image Database

Text-1-GNA-RNN [26]

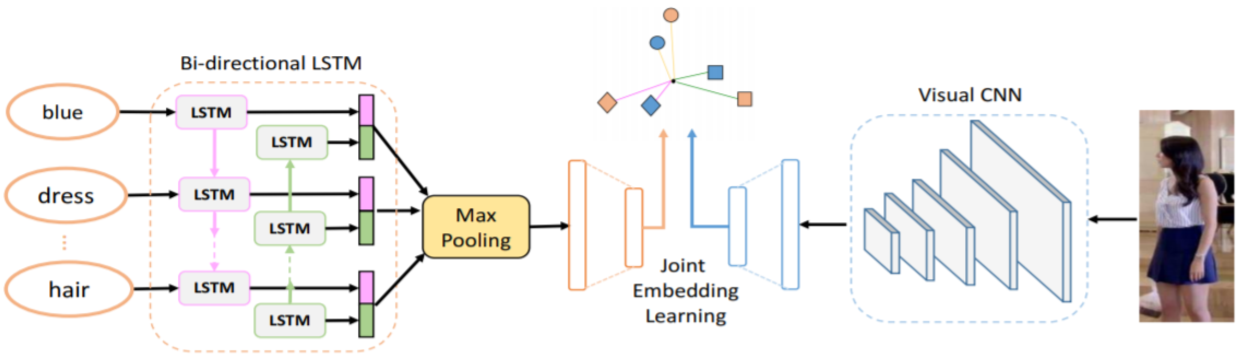


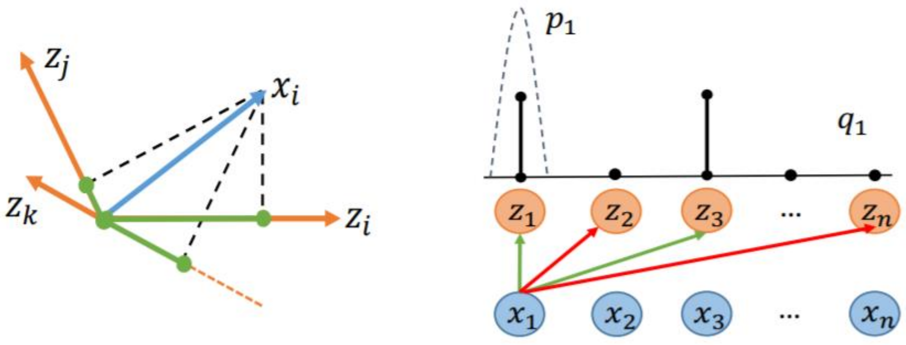
- Contributions
  - study the problem of searching persons with natural language
  - a novel Recurrent Neural Network with Gated Neural Attention (GNA-RNN) for person search

[26] Li, et al., Person search with natural language description, CVPR, 2017

Affinity/Metric Learning

Text-2-CMPM+CMPC [27]





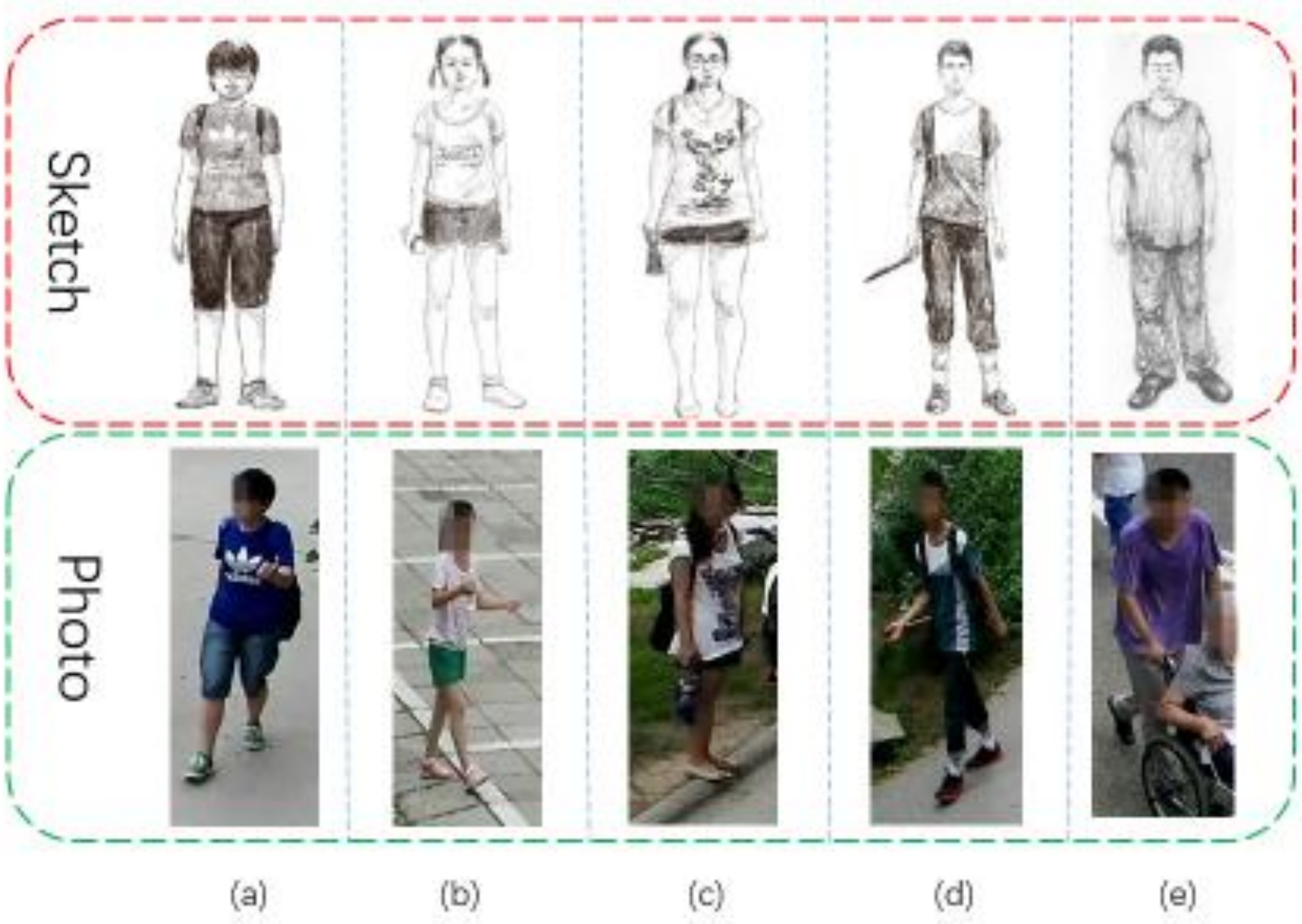
(a) cross-modal projection (b) cross-modal matching

- Contributions
  - a cross-modal projection matching (CMPM) loss attempts to minimize the KL divergence between projection compatibility distributions and the normalized matching distributions
  - a cross-modal projection classification (CMPC) loss attempts to classify the vector projection of the features from one modality onto the matched features from another modality

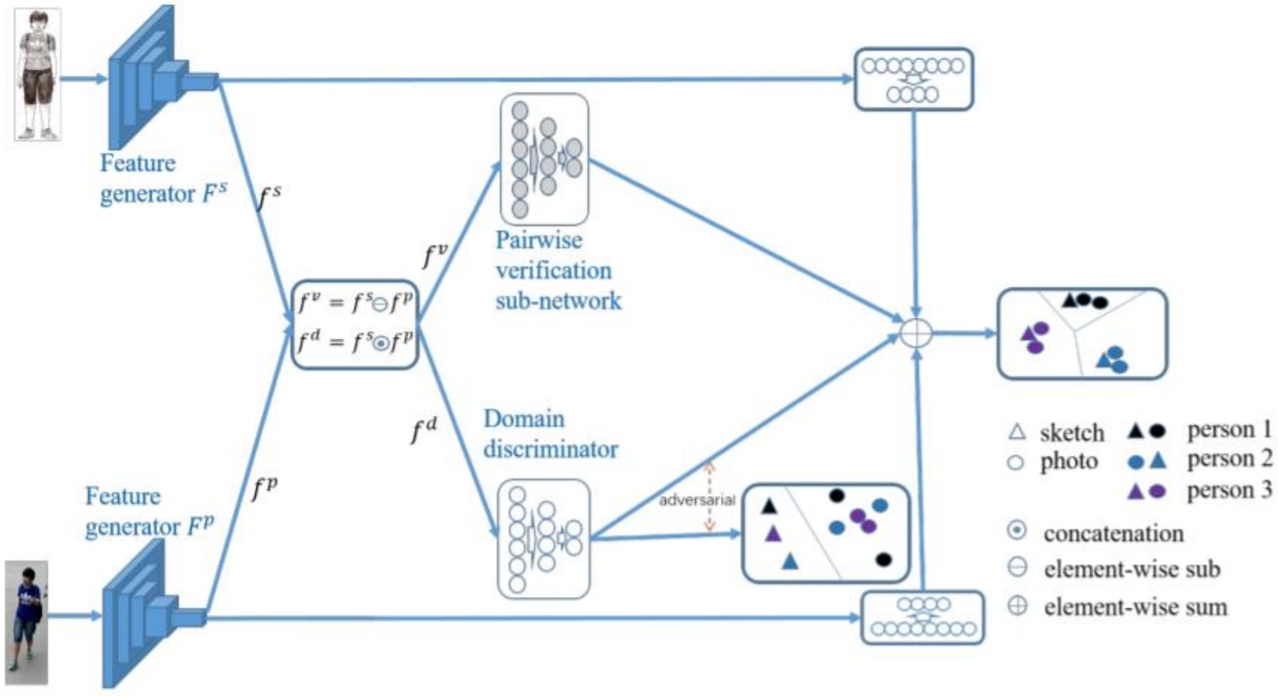
[27] Zhang, et al., Deep cross modal projection learning for image-text matching, ECCV, 2018

Representation Learning

Sketch-Photo



Sketch-1-CDAFL [28]



- Contributions
  - A deep adversarial learning architecture to jointly learn identity features and domain-invariable features
  - filtering low-level features and remaining high-level semantic features.
  - A sketch Re-ID dataset containing 200 persons, in which each person has one sketch and two photos

[28] Pang, et al., Cross-domain adversarial feature learning for sketch re-identification, ACM MM, 2018

Our Sketch-Photo Focus

Large-scale
Multi-style
Semi-professional

From the perspective of application scenario

- Most of the methods selected a deep learning framework.

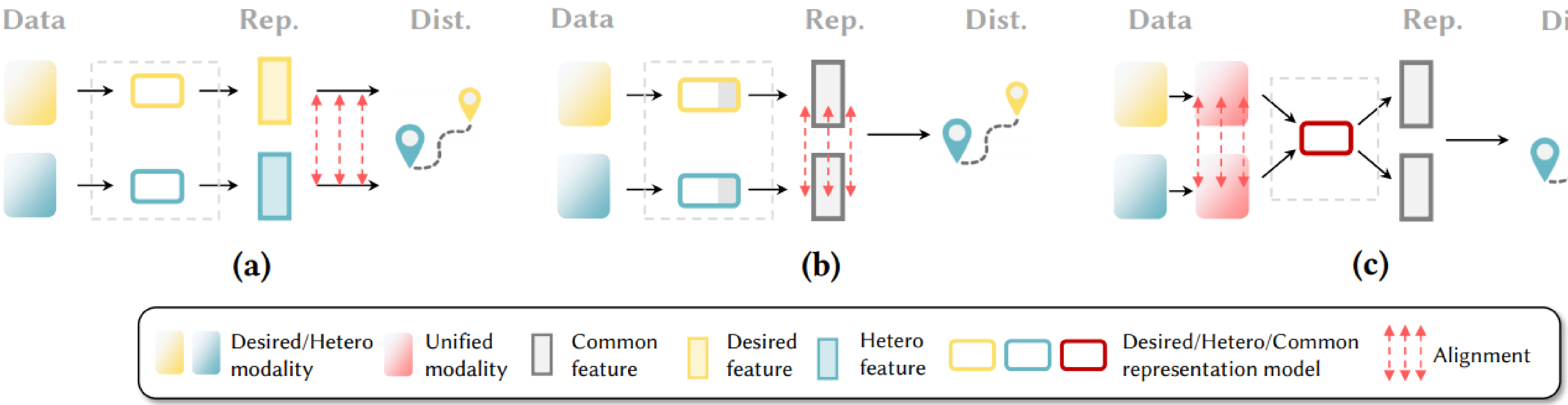
From the perspective of application scenario

- Most of the methods selected a deep learning framework.
- Different methods have different focuses.

From the perspective of application scenario

- Most of the methods selected a deep learning framework.
- Different methods have different focuses.
- The existing researches in each application scenario still have many limitations.

From the Perspective of Learning Pipeline



Method	Method	Strategy	Focus	CMC-1	CMC-5	CMC-10	CMC-20	mAP
LR MLR-VIPeR	JUDEA	Multi-scale Metrics	ML	26.0	55.1	69.2		
	SLD2L	Dictionary Learning	RL	20.3	44.0	62.0		
	SDF	Resolution-Distance Variation	RL	9.3	38.1	52.4		
	SING	Super Resolution	MU	33.5	57.0	66.6		
	CSR-GAN	**Cascaded SR and ReID**	**MU**	**37.2**	**62.3**	**71.6**		
	FFSR+RIFE	**Foreground Focus SR**	**MU**	**41.6**	**64.9**	**--**		
	CAD	**Adversarial Learning**	**MU**	**43.1**	**68.2**	**77.5**		
	INTACT	**Inter-task Association**	**MU**	**46.2**	**73.1**	**81.6**		
IR SYSU-MM01	Zero-padding	One-stream and Zero-padding	RL	14.80	--	54.12	71.33	15.95
	HCML	Feature & Metric Learning	ML	14.32	--	53.16	69.17	16.16
	BCTR	End-to-End	RL	17.01	--	55.43	71.96	19.66
	cmGAN	Adversarial Learning	RL	26.97	--	67.51	80.56	27.80
	D2RL	**Dual-level Reduction**	**MU**	**28.90**	**--**	**70.60**	**82.40**	**29.20**
	XIV	X Modality	RL	49.92	--	89.79	95.96	50.73
	Hi-CMD	Disentanglement	MU	34.94	--	77.58	--	35.94
	cm-SSFT	Affinity Modeling	RL	61.6	--	89.2	93.9	63.2
Text CUHK-PEDES	GNA-RNN	Affinity Learning	ML	19.05	--	53.64		
	CNN-LSTM	Two-Stage Matching	ML	25.94	--	60.48		
	CMPM+CMPC	Cross-modal Projection	RL	49.37	--	79.27		
	GDA+LRA	**Local and Global Association**	**RL**	**43.58**	**66.93**	**76.26**		
Sketch PKU-Sketch	**CDAFL**	**Adversarial Learning**	**RL**	**34.0**	**56.3**	**72.5**	**84.7**	

Multi-task Learning

Modality Unification

Adversarial Learning

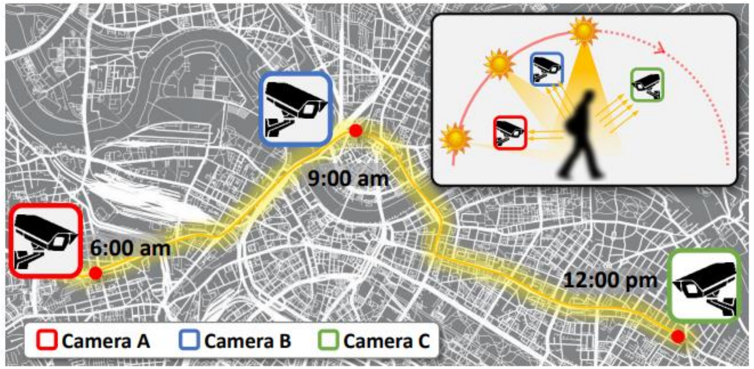
Focus on Person Details

ML: Metric Learning
RL: Representation Learning
MU: Modality Unification

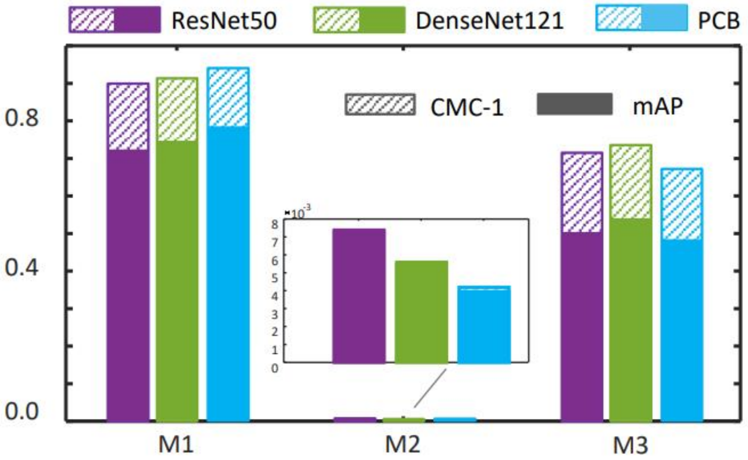
Outline

- General Person Re-identification
- Person Re-identification New Trends
    - Cross-modality
    - Long-term
    - Group
- Discussion

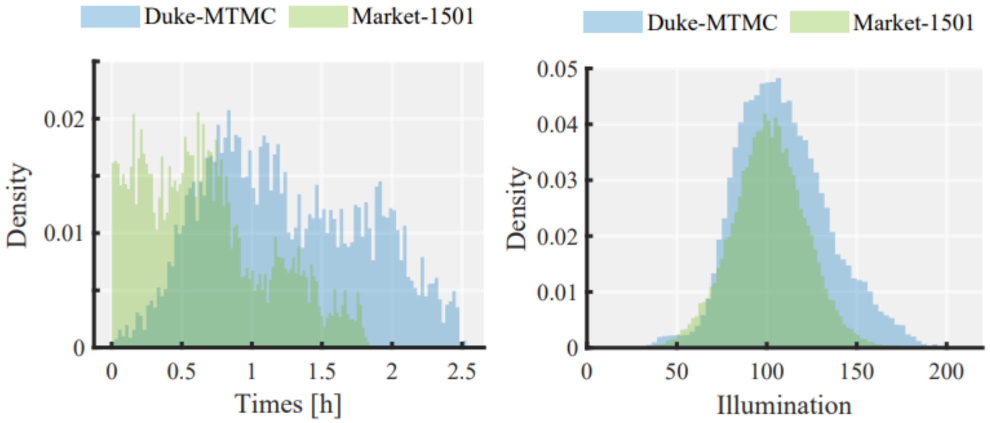
Challenge1-Illumination

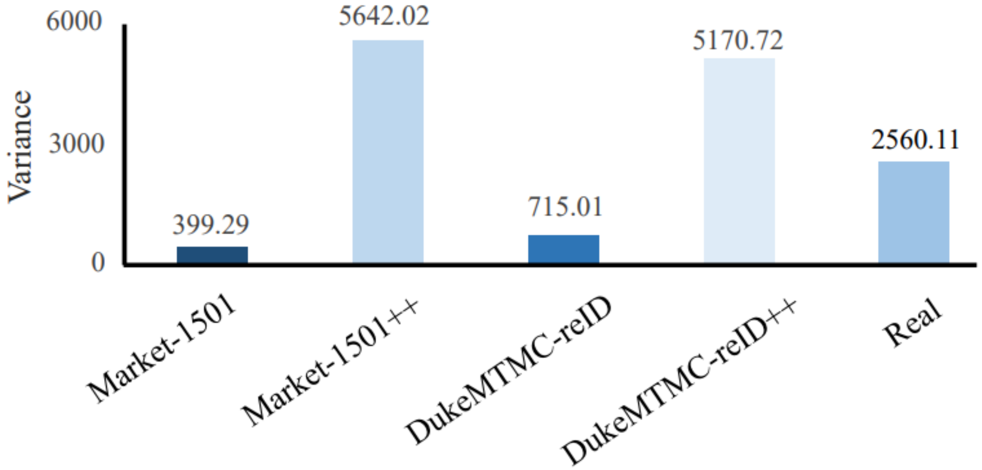


(a)



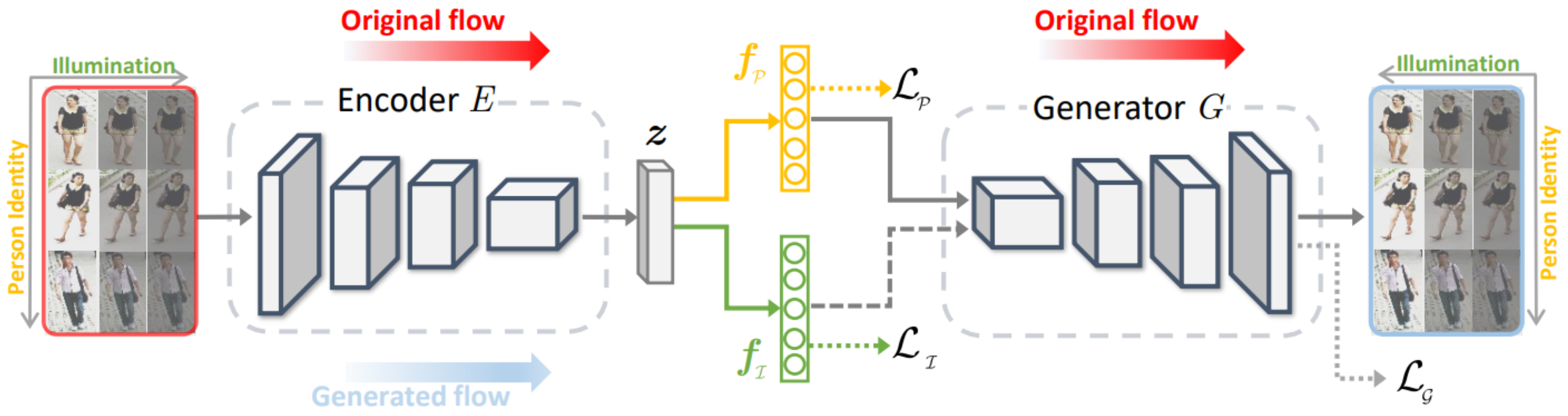
(b)





[29] Z Zeng, Z Wang, **Z Wang***, Y Zheng, YY Chuang, S Satoh, Illumination-Adaptive Person Re-identification, ***IEEE T-MM***, 2020

Illumination Adaptive



$$\mathcal{L}_{\mathcal{P}} = \lambda_1 \mathcal{L}_{\mathcal{P}}^T + \lambda_2 \mathcal{L}_{\mathcal{P}}^S$$

$$\mathcal{L}_{\mathcal{P}}^T = \sum_{\boldsymbol{f}_{\mathcal{P}}^a, \boldsymbol{f}_{\mathcal{P}}^p, \boldsymbol{f}_{\mathcal{P}}^n \in \mathcal{B}} [\mathcal{D}(\boldsymbol{f}_{\mathcal{P}}^a, \boldsymbol{f}_{\mathcal{P}}^p) - \mathcal{D}(\boldsymbol{f}_{\mathcal{P}}^a, \boldsymbol{f}_{\mathcal{P}}^n) + \xi]_+$$

$$\mathcal{L}_{\mathcal{P}}^S = -\frac{1}{N} \sum_{j=1}^{N} \log \hat{\boldsymbol{y}}_{\mathcal{P}}^j$$

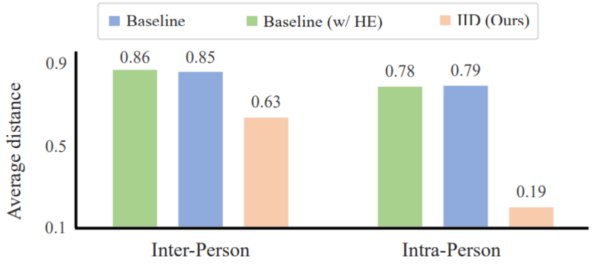
$$\mathcal{L}_{\mathcal{I}} = \frac{1}{N} \sum_{j=1}^{N} \left\| \hat{c}_{\mathcal{I}}^j - (\mathbf{W}_{\mathcal{I}} \boldsymbol{f}_{\mathcal{I}}^j + \mathbf{b}_{\mathcal{I}}) \right\|_2^2$$

$$\hat{c}_{\mathcal{I}} = c_{\mathcal{I}} + \epsilon, \quad \text{with} \quad \epsilon \in \mathcal{N}(0, 1)$$

Illumination Adaptive

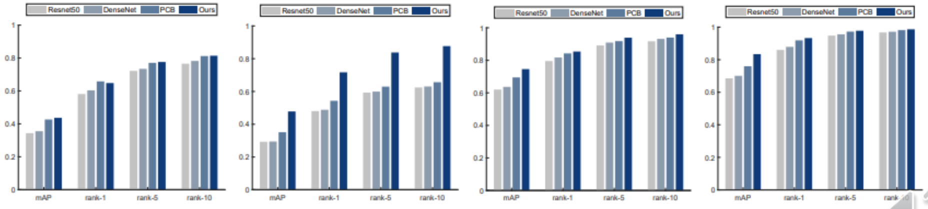
(a) Market-1501++ (b) DukeMTMC-reID++

Method	Components				Market-1501++	
	$\mathcal{L}_\mathcal{P}^T$	$\mathcal{L}_\mathcal{P}^S$	$\mathcal{L}_\mathcal{I}$	$\mathcal{L}_\mathcal{G}$	CMC-1	mAP
Baseline	✓	✓	×	×	66.18	47.71
IID (no G)	✓	✓	✓	×	71.54	55.17
IID (no triplet for id)	×	✓	✓	✓	64.14	45.87
IID (no softmax for id)	✓	×	✓	✓	65.21	46.53
IID (no illum.)	✓	✓	×	✓	70.79	54.57
IID	✓	✓	✓	✓	73.37	56.22



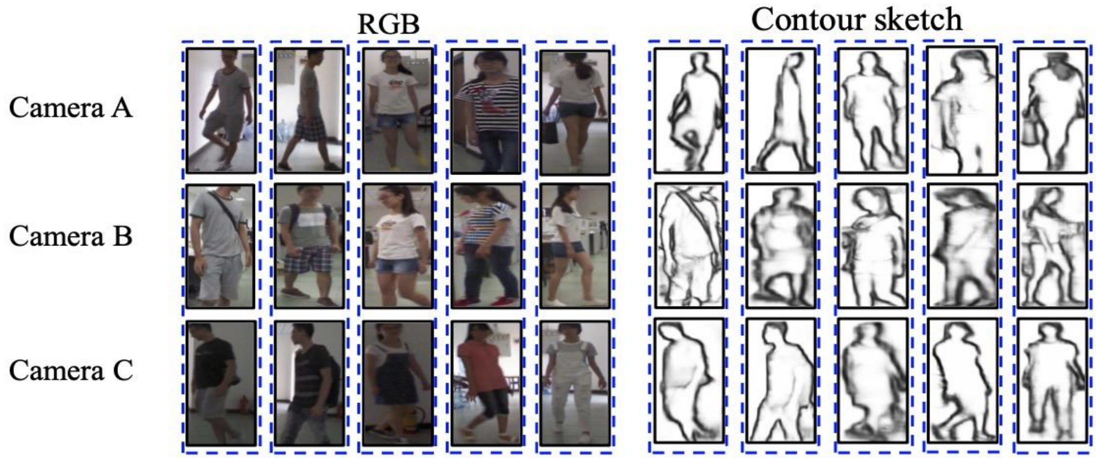
Method	Market-1501++				DukeMTMC-reID++			
	CMC-1	CMC-5	CMC-10	mAP	CMC-1	CMC-5	CMC-10	mAP
DenseNet121 [45]	0.74	2.29	3.53	0.73	1.21	2.74	4.13	0.80
DenseNet121 w/ Train	70.60	85.36	89.66	49.79	64.45	77.82	82.45	45.12
PCB [46]	0.56	1.69	2.91	0.54	0.72	2.15	3.23	0.49
PCB w/ Train	72.55	85.22	90.08	53.11	65.98	77.93	82.21	45.15
ResNet50 [47]	0.42	1.16	2.05	0.39	0.54	1.97	3.14	0.50
ResNet50 w/ Train (Baseline)	66.18	81.97	87.02	47.71	62.07	75.54	88.08	42.63
IID	**73.37**	**86.55**	**91.01**	**56.22**	**68.11**	**79.75**	**91.27**	**49.20**
Improvement over baseline	7.19↑	4.58↑	3.99↑	8.51↑	6.04↑	4.21↑	3.19↑	6.57↑

(a) Foreground Duke (b) Foreground Market (c) Patch Duke (d) Patch Market



Challenge2-Cloth Changing

The same identity with different clothes



PRCC Dataset

- 33698 images from 221 identities
- each person in Cameras A and B is wearing the same clothes
- for Camera C, the person wears different clothes

https://www.isee-ai.cn/~yangqize/clothing.html

LTCC Dataset

- one cloth-change set where 91 persons appearing with 416 different sets of outfits in 14,783 images
- one cloth-consistent subset containing the remaining 61 identities with 2,336 images without outfit changes

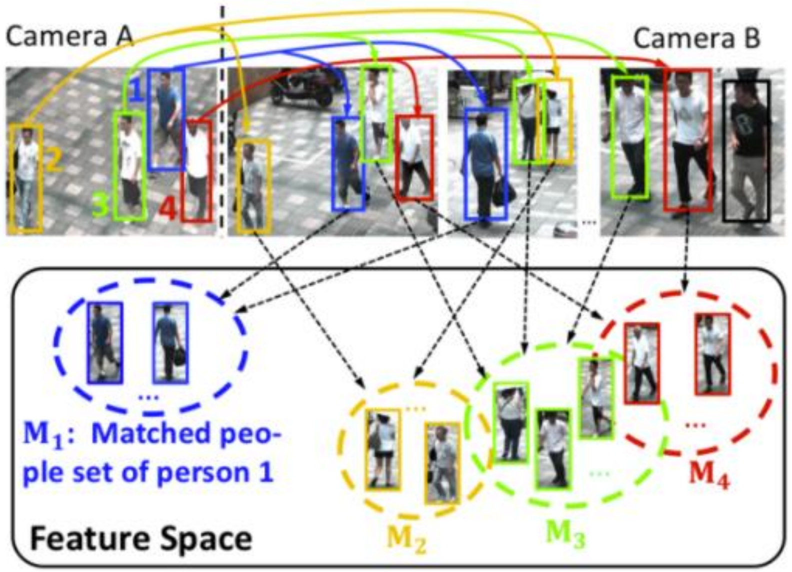
https://naiq.github.io/LTCC_Perosn_ReID.html

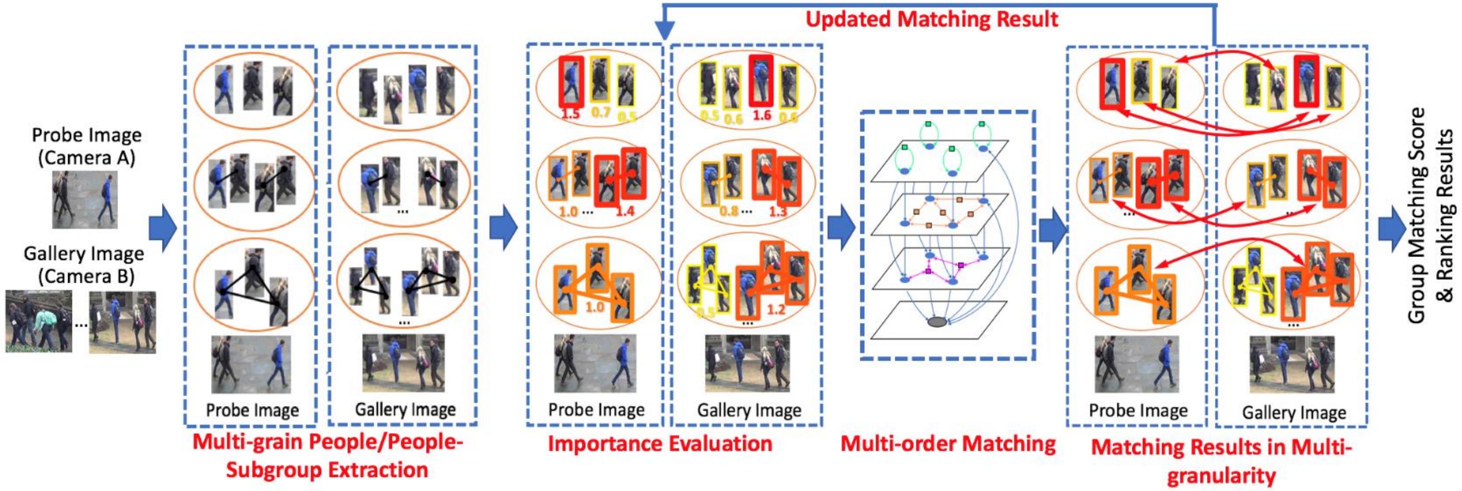
[30] Qize Yang, Ancong Wu, Wei-Shi Zheng. "Person Re-identification by Contour Sketch under Moderate Clothing Change." IEEE TPMI, 2020
[31] Xuelin Qian, Wenxuan Wang, et al., Long-Term Cloth-Changing Person Re-identification, ECCV, 2020

Outline

- General Person Re-identification
- Person Re-identification New Trends
  - Cross-modality
  - Long-term
  - Group
- Discussion

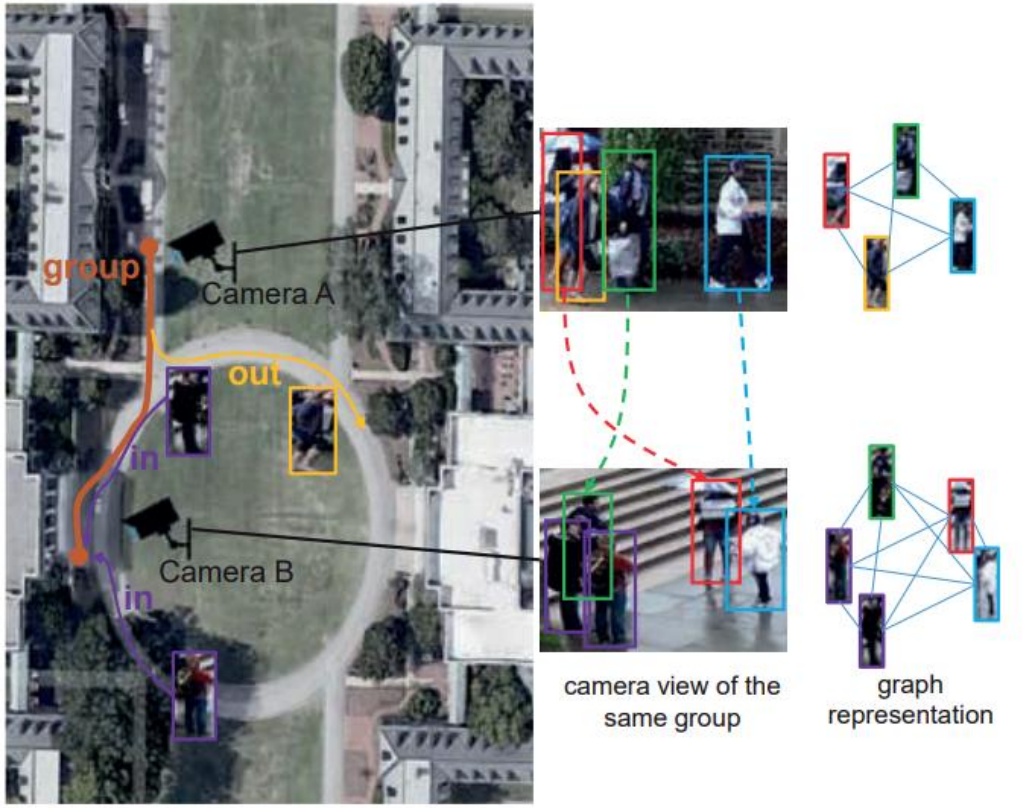
Group: Challenge





[32] Group Re-Identification: Leveraging and Integrating Multi-Grain Information, ACM MM, 2018

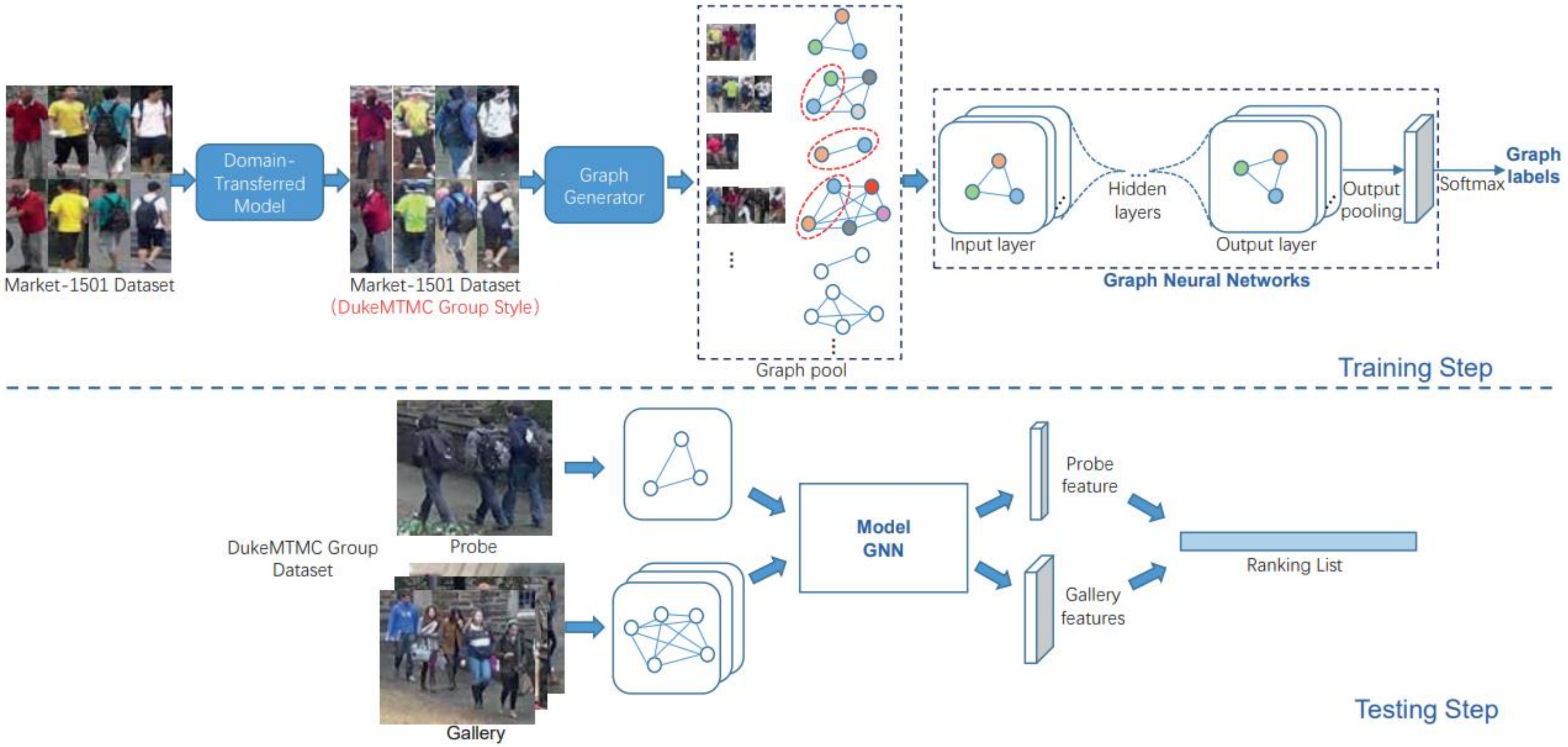
DoT-GNN

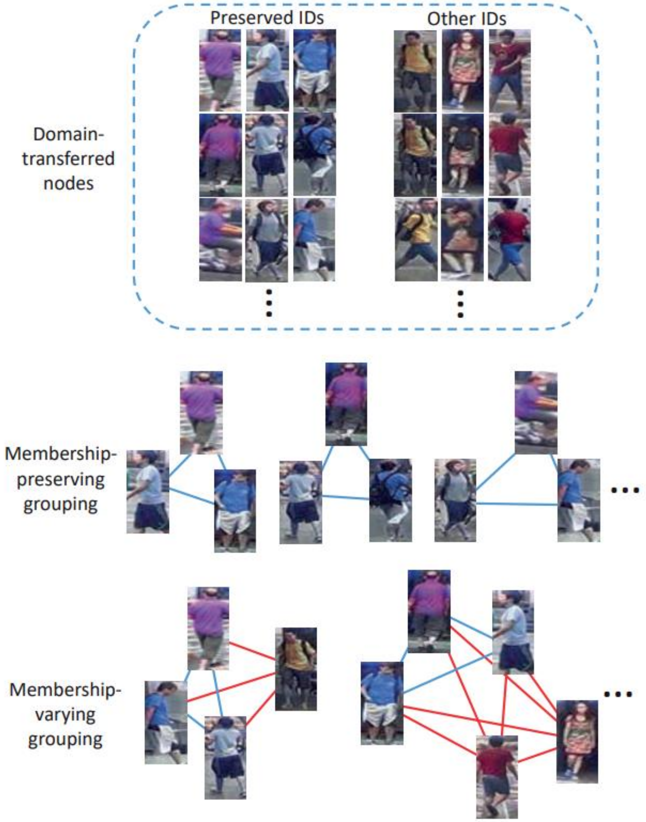


Challenge	ReID	G-ReID	Strategy
Training Set	Abundant	Insufficient	
Appearance	✓	✓	Node generating (transfer)
Layout	×	✓	Membership-preserving grouping
Membership	×	✓	Membership-varying grouping

[33] Z Huang, **Z Wang***, W Hu, CW Lin, S Satoh, DoT-GNN: Domain-Transferred Graph Neural Network for Group Re-identification, ***ACM MM***, 2019

DoT-GNN





DoT-GNN: Domain-Transferred Graph Neural Network

DoT-GNN

Variant	Settings				DukeMTMC Group			
	Tr.	*S1*	*S2*	*GNN*	CMC-1	CMC-5	CMC-10	CMC-20
①	✓	×	×	×	35.2	46.6	46.6	47.7
②	✓	✓	×	✓	44.3	72.2	78.4	86.4
③	✓	×	✓	✓	35.2	62.5	78.7	86.4
④	×	✓	✓	✓	44.3	67.0	76.1	85.2
⑤	✓	✓	✓	✓	**53.4**	**72.7**	**80.7**	**88.6**

Method	DukeMTMC Group				Road Group			
	CMC-1	CMC-5	CMC-10	CMC-20	CMC-1	CMC-5	CMC-10	CMC-20
CRRRO-BRO [34]	9.9	26.1	40.2	64.9	17.8	34.6	48.1	62.2
Covariance [4]	21.3	43.6	60.4	78.2	38.0	61.0	73.1	82.5
PREF [12]	22.3	44.3	58.5	74.4	43.0	68.7	77.9	85.2
BSC+CM [38]	23.1	44.3	56.4	70.4	58.6	80.6	87.4	92.1
MGR [28]	47.4	68.1	77.3	87.4	72.3	**90.6**	**94.1**	97.5
Resnet50 + Feature Fusion	31.8	56.8	73.9	80.7	38.3	58.0	67.9	77.8
DoT + Feature Fusion	40.9	69.3	77.3	83.0	43.2	65.4	70.4	76.5
DoT + Distance Fusion	35.2	46.6	46.6	47.7	9.9	9.9	55.6	65.4
DoT + GNN	**53.4**	**72.7**	**80.7**	**88.6**	**74.1**	90.1	92.6	**98.8**

Outline

- General Person Re-identification
- Person Re-identification New Trends
  - Cross-modality
  - Long-term
  - Group
- Discussion

Conclusion and Future Directions

- Dataset Construction [5]

[5] Zeng, et al., Illumination-Adaptive Person Re-identification, TMM, 2020
[34] Yang,, et al., Mining on heterogeneous manifolds for zeroshot cross-modal image retrieval. AAAI, 2020.
[35] Mirjalili , et al., Soft biometric privacy: Retaining biometric utility of face images while perturbing gender, IJCB, 2017.

Conclusion and Future Directions

- Dataset Construction [5]
- Taking Advantages of general ReID Datasets and Methods [34]

[5] Zeng, et al., Illumination-Adaptive Person Re-identification, TMM, 2020
[34] Yang,, et al., Mining on heterogeneous manifolds for zeroshot cross-modal image retrieval. AAAI, 2020.
[35] Mirjalili , et al., Soft biometric privacy: Retaining biometric utility of face images while perturbing gender, IJCB, 2017.

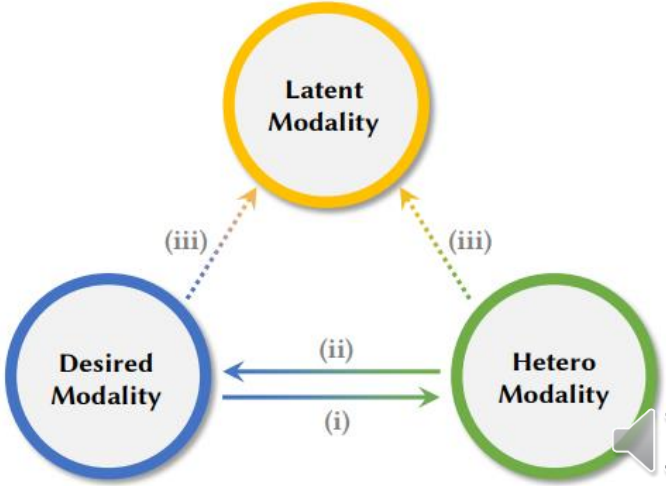
Conclusion and Future Directions

- Dataset Construction [5]
- Taking Advantages of general ReID Datasets and Methods [34]
- Human Interaction and Crowd-sourcing

[5] Zeng, et al., Illumination-Adaptive Person Re-identification, TMM, 2020
[34] Yang,, et al., Mining on heterogeneous manifolds for zeroshot cross-modal image retrieval. AAAI, 2020.
[35] Mirjalili , et al., Soft biometric privacy: Retaining biometric utility of face images while perturbing gender, IJCB, 2017.

Conclusion and Future Directions

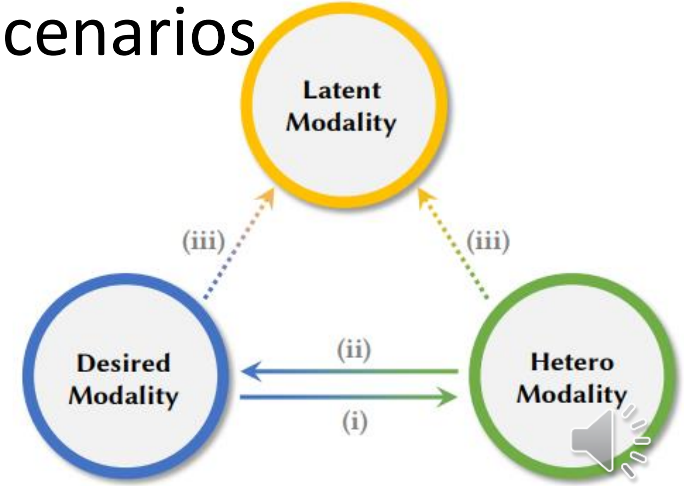
- Dataset Construction [5]
- Taking Advantages of general ReID Datasets and Methods [34]
- Human Interaction and Crowd-sourcing
- Investigation on Unifying the Modality



[5] Zeng, et al., Illumination-Adaptive Person Re-identification, TMM, 2020
[34] Yang,, et al., Mining on heterogeneous manifolds for zeroshot cross-modal image retrieval. AAAI, 2020.
[35] Mirjalili , et al., Soft biometric privacy: Retaining biometric utility of face images while perturbing gender, IJCB, 2017.

Conclusion and Future Directions

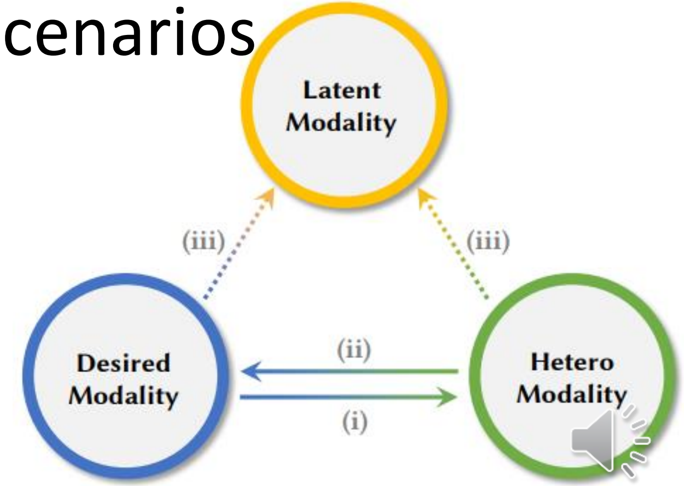
- Dataset Construction [5]
- Taking Advantages of general ReID Datasets and Methods [34]
- Human Interaction and Crowd-sourcing
- Investigation on Unifying the Modality
- Integrating Multiple cross-modality ReID Application Scenarios



[5] Zeng, et al., Illumination-Adaptive Person Re-identification, TMM, 2020
[34] Yang,, et al., Mining on heterogeneous manifolds for zeroshot cross-modal image retrieval. AAAI, 2020.
[35] Mirjalili , et al., Soft biometric privacy: Retaining biometric utility of face images while perturbing gender, IJCB, 2017.

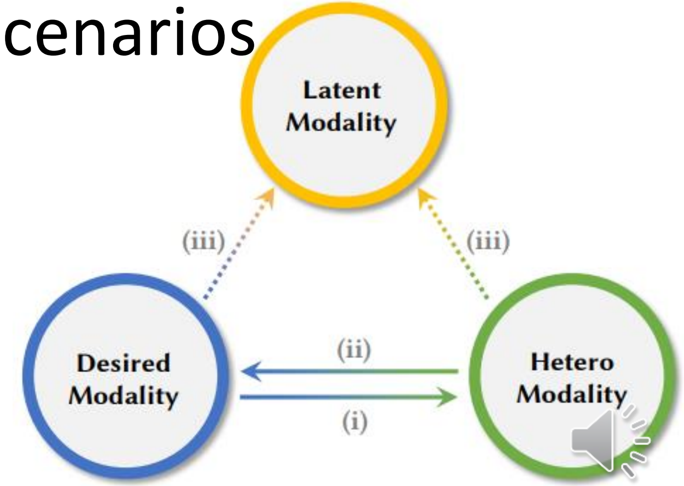
Conclusion and Future Directions

- Dataset Construction [5]
- Taking Advantages of general ReID Datasets and Methods [34]
- Human Interaction and Crowd-sourcing
- Investigation on Unifying the Modality
- Integrating Multiple cross-modality ReID Application Scenarios
- Considering the Privacy Issue [35]



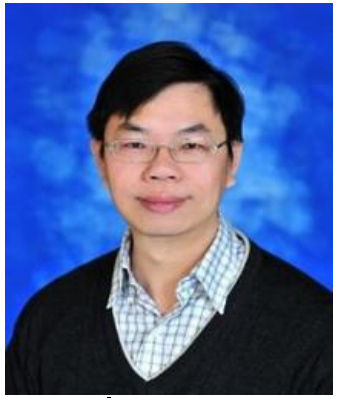
Conclusion and Future Directions

- Dataset Construction [5]
- Taking Advantages of general ReID Datasets and Methods [34]
- Human Interaction and Crowd-sourcing
- Investigation on Unifying the Modality
- Integrating Multiple cross-modality ReID Application Scenarios
- Considering the Privacy Issue [35]
- Common Model for General ReID and New Trends



Collaborators

Shin'ichi
NII/UTokyo



Chia-wen
NTHU

Wenjun
Microsoft

Mang
IIAI

Yinqiang
NII

Yang
KyotoU

Zhixiang
NTU

Fan
UTokyo

Zelong
Utokyo

Ziling
NTHU

Kajal
IIIT Delhi

Thanks!