Person Re-identification

Introduction and Future Trends

Shengcai Liao

Institute of Automation
Chinese Academy of Sciences

ICPR 2018 Tutorial • Beijing

CONTENT

01

PART ONE

Introduction

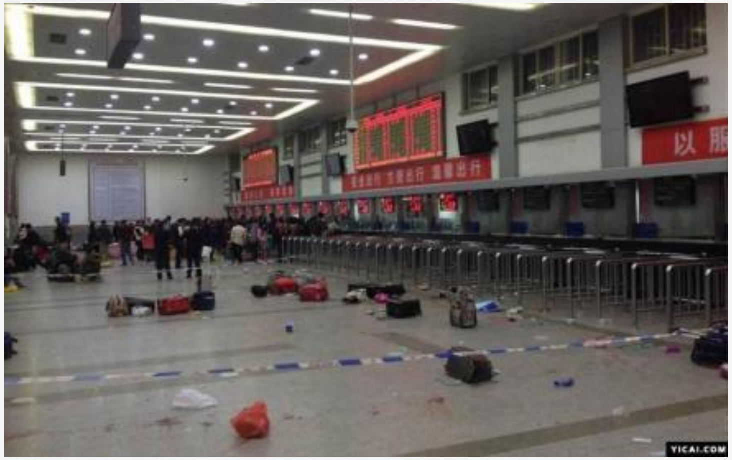
Background

- Security concerns

2011 riot in London

2013 Boston Marathon bombings

2012 “8.10” serial killer Zhou Kehua



2014 “3.1” Kunming terror attack

Background

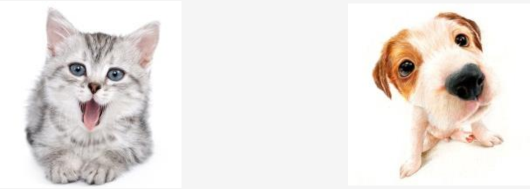
- Surveillance cameras everywhere
- However,
  - Mostly, searching suspects still requires large amount of labors
  - Automatic algorithms are still poor
  - But the real demand is increasing

Background

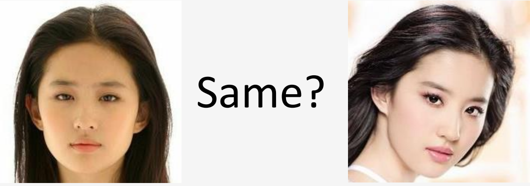
Search suspects in a large amount of videos

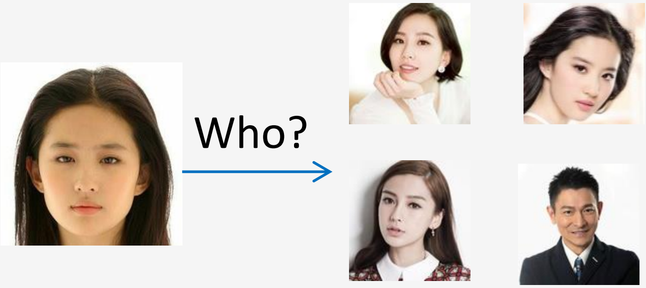
Concepts

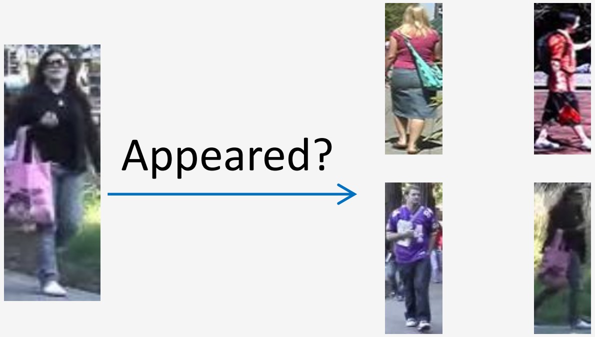
- **Classification:** classes fixed
- **Verification:** pairwise
- **Identification:** gallery IDs known
- **Re-identification :** gallery IDs unknown



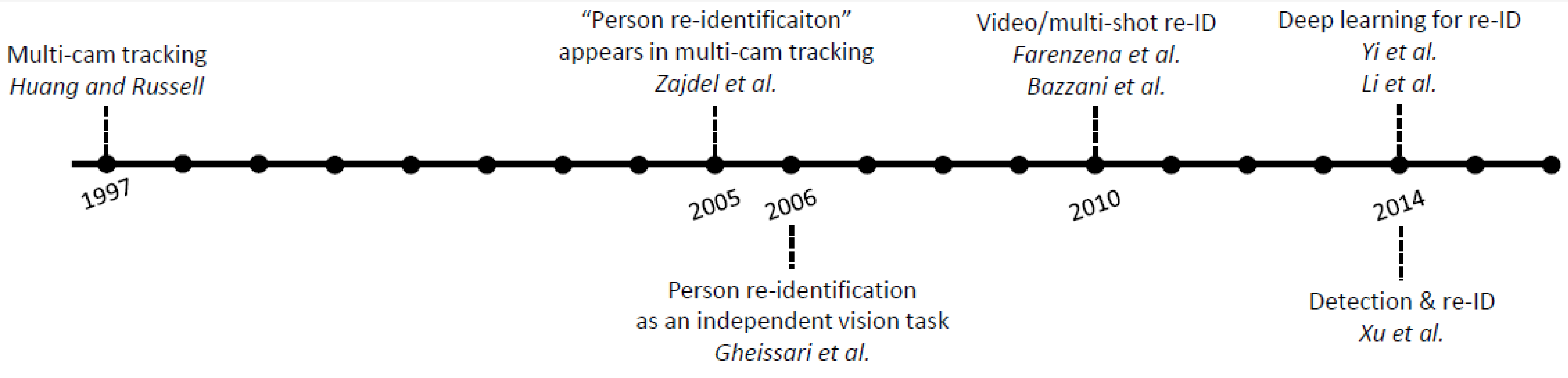
Cat Dog







History



From Zheng et al. 2016.

Difference with Multi-camera Tracking

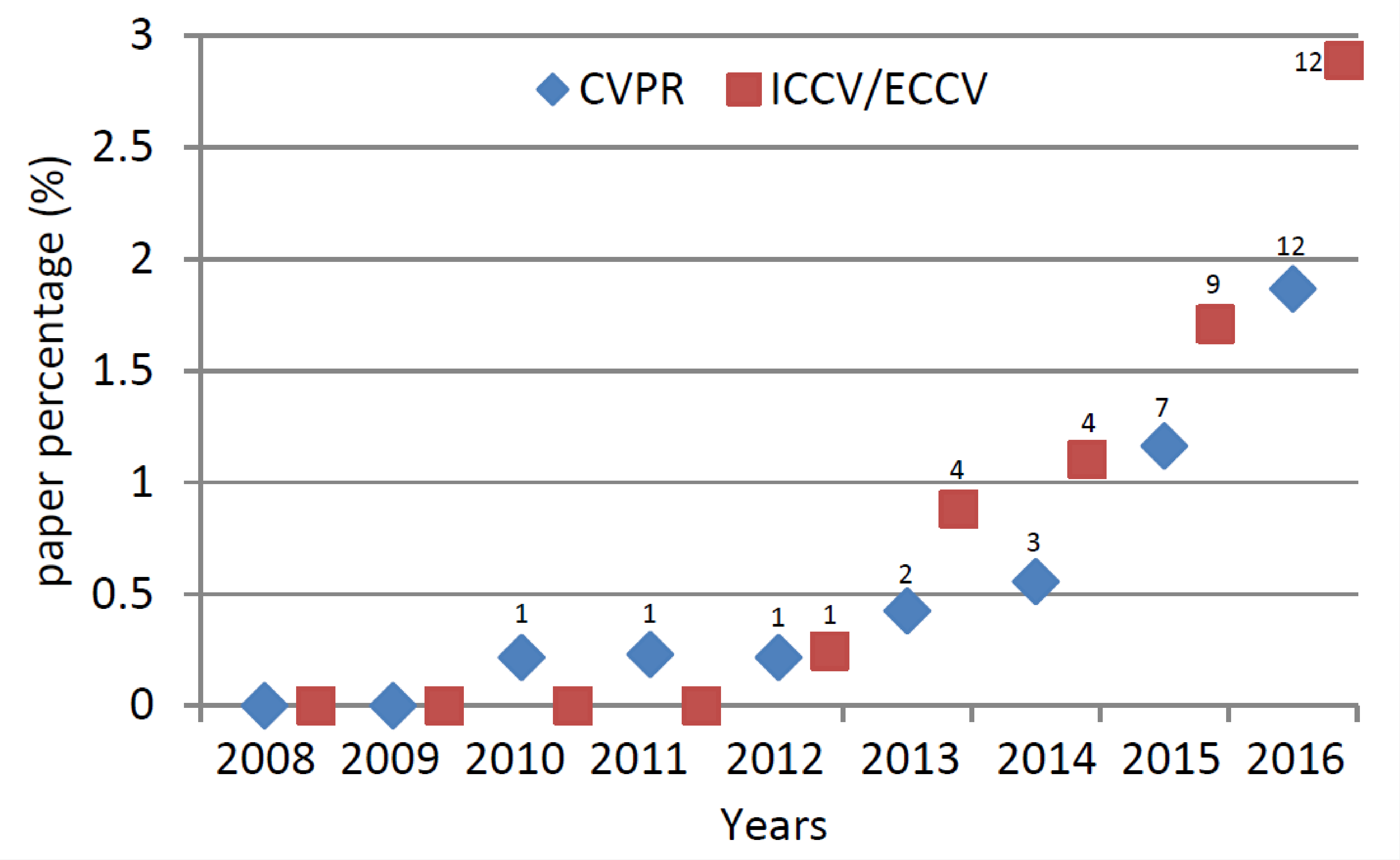
- Multi-camera tracking
  - Usually online
  - Need to track all persons in all cameras
  - In a local area
  - In a short duration
- Person Re-identification
  - Usually offline, for retrieval
  - Re-identify one specific person
  - Across broad areas
  - With a possible long time

Multi vs. multi

One vs. multi

**Oriented from multi-camera tracking,
but is a particular independent task now.**

Popularity



From Zheng et al. 2016.

Pipeline

Preprocess

- Pedestrian detection
- Single-camera Tracking

Representation

- Hand-crafted features
- Feature learning

Matching

- Traditional Distances
- Metric learning
- Re-ranking

Challenges

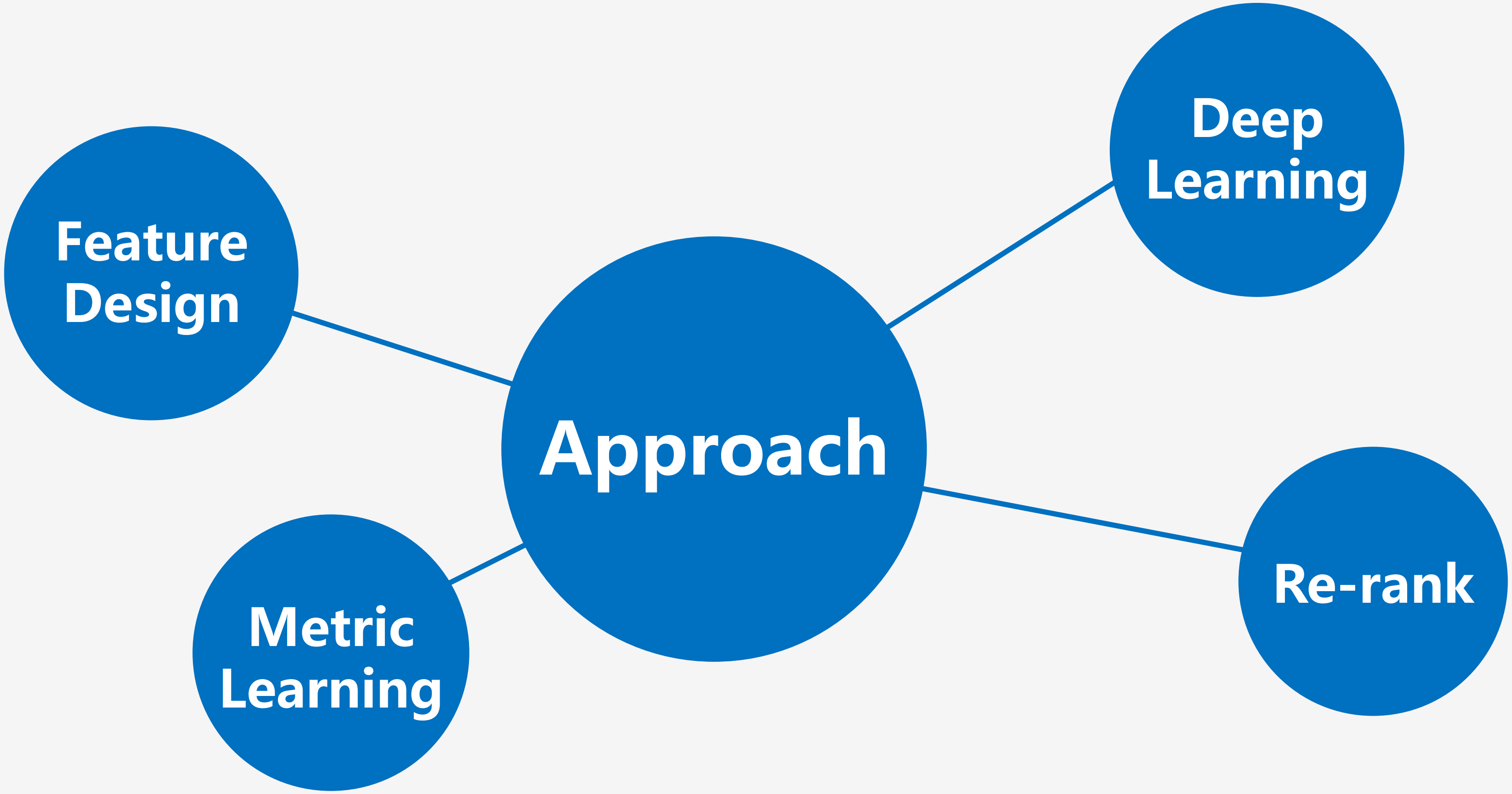
- Viewpoint changes
- Pose changes
- Illumination variations
- Occlusions
- Low resolutions
- Limited labeled data
- Generalization ability

02

PART TWO

Approach

Approach



Main research directions in person re-identification

Feature Design

Color
RGB, HSV, YCbCr, Lab, Color names

Texture
Gabor, LBP, SILTP, Schmid, BiCov

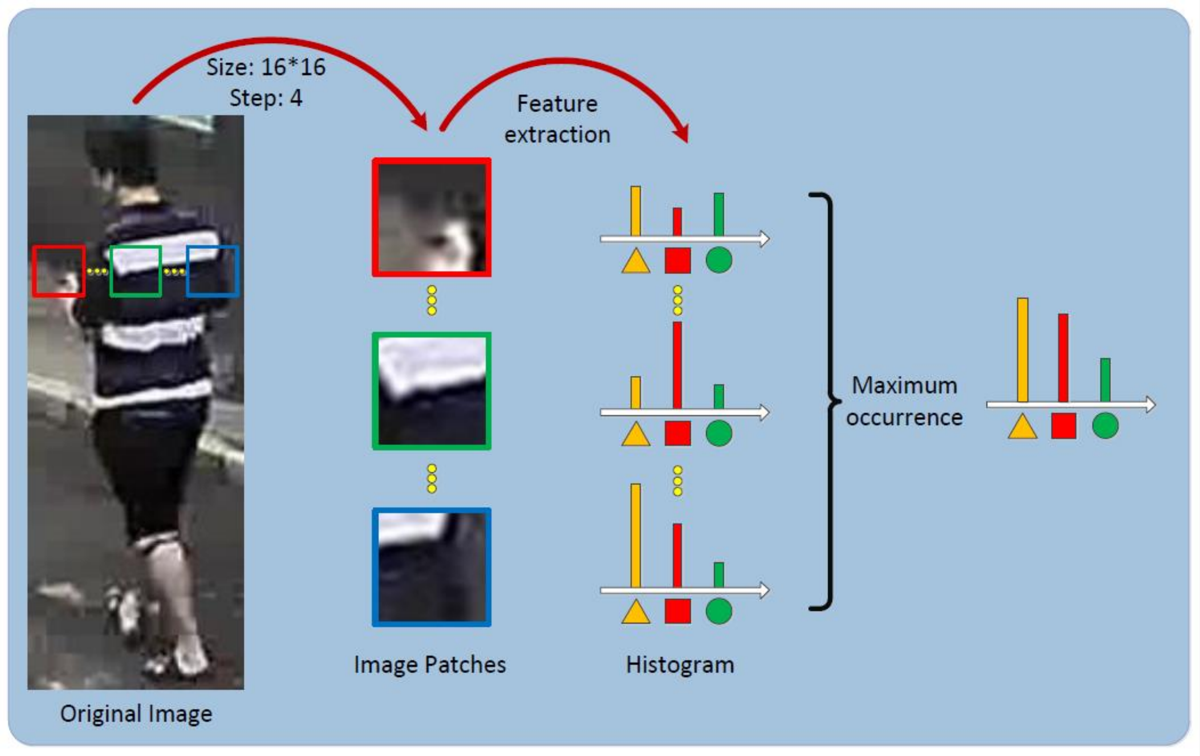
Hybrid
ELF, LOMO, GOG

Structure
Pictorial, SDALF, Saliency

Attribute
Age, gender, bag

Feature Design

- Typical feature: LOMO
  - Viewpoint changes: local maximal occurence
  - Illumination variations: retinex and SILTP



S. Liao et al., "Person Re-identification by Local Maximal Occurrence Representation and Metric Learning," In CVPR 2015.

Metric Learning

Traditional Methods

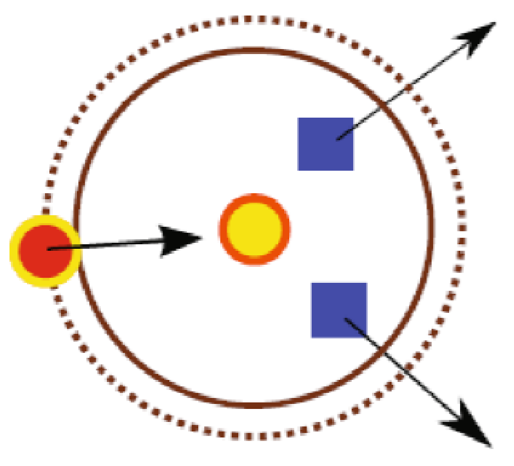
ITML, LMNN, LDML

Optimization Methods

PRDC, MLAPG

Fast Methods

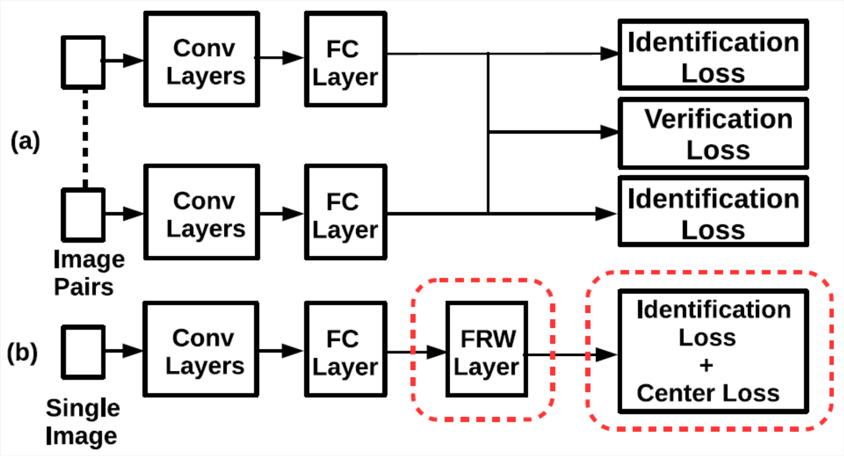
KISSME, XQDA, LSSL

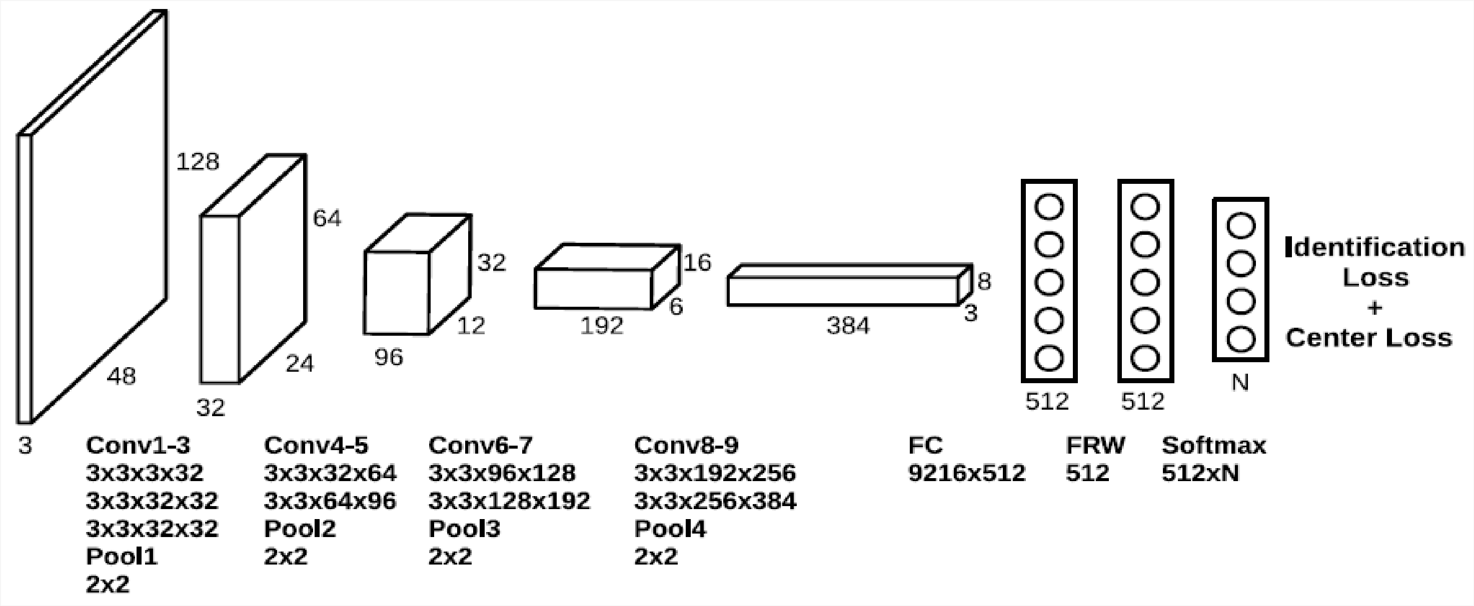


$$D^2_{\mathbf{M}}(\mathbf{x},\mathbf{z}) = \|\mathbf{x}-\mathbf{z}\|^2_{\mathbf{M}} = (\mathbf{x}-\mathbf{z})^T\mathbf{M}(\mathbf{x}-\mathbf{z})$$

Deep Learning

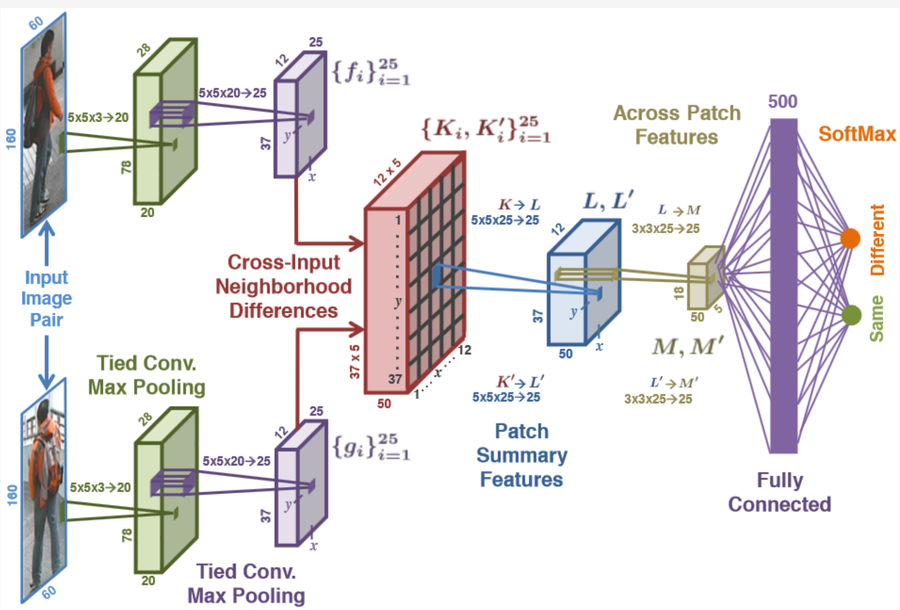
- Deep metric learning
  - Cosine similarity
  - Contrastive loss
  - Triplet loss
  - Center loss

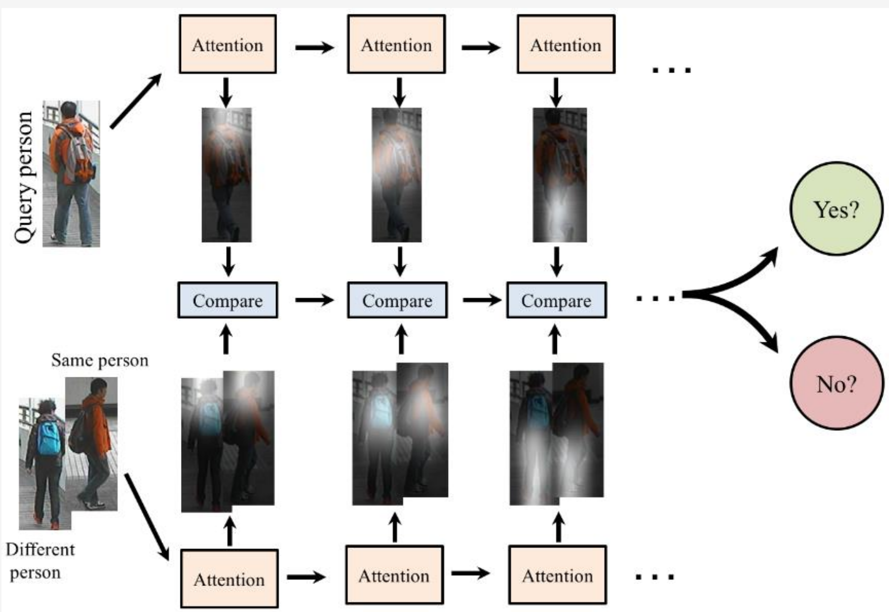




Deep Learning

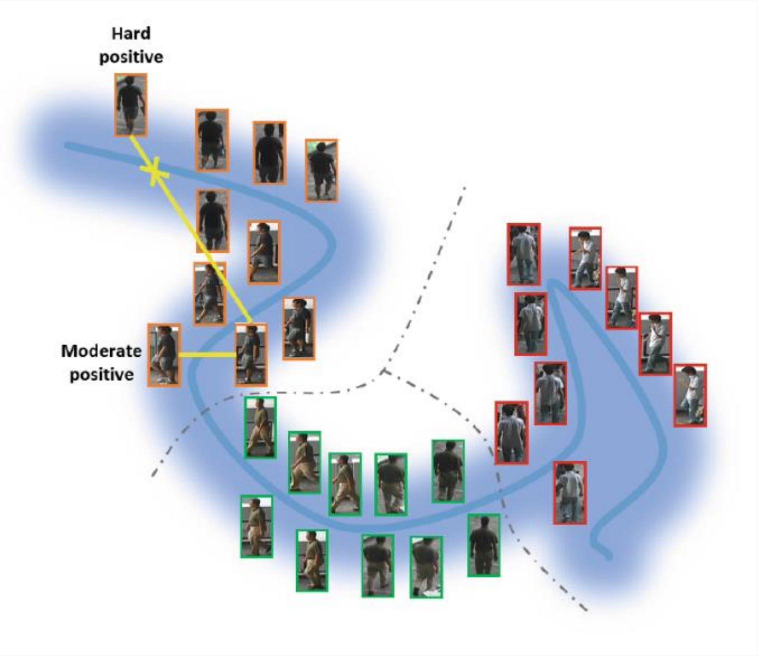
- Deep structures
  - Siamese CNN
  - Cross-input neighborhood, patch summary
  - Gating CNN
  - Contextual LSTM
  - Attention network

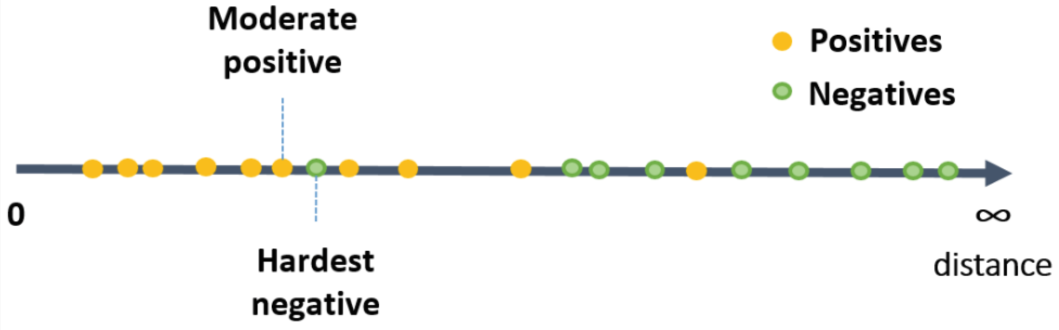




Deep Learning

- Sample mining
  - Hard negative mining
  - Moderate positive sample mining

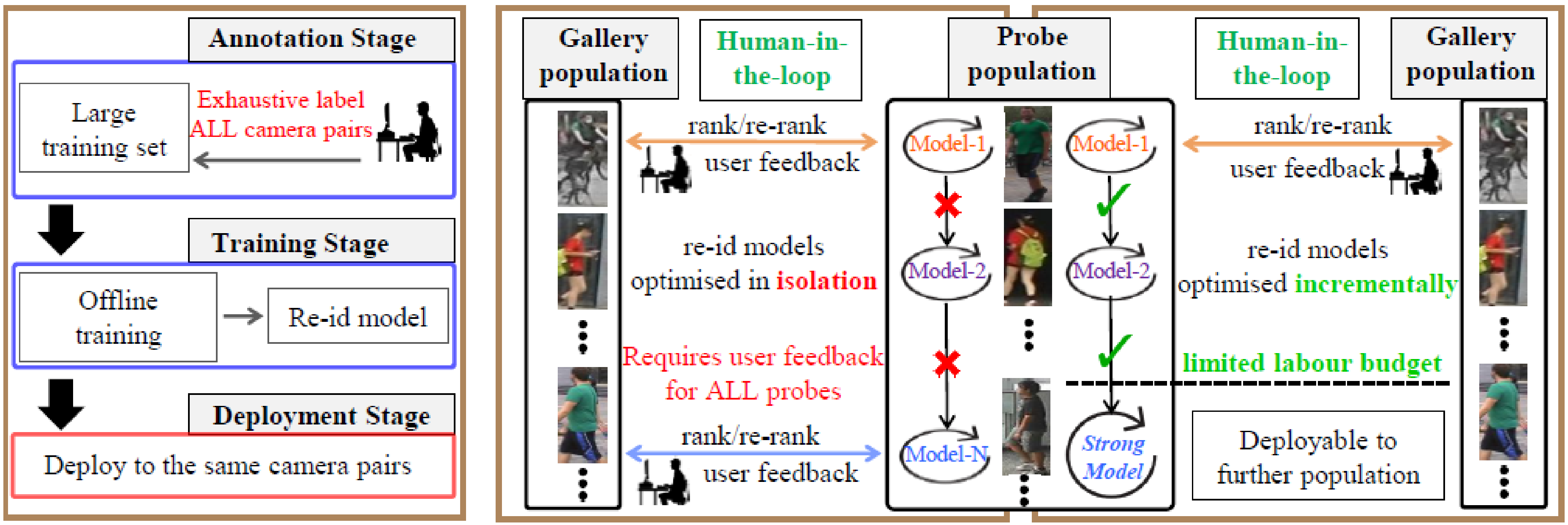




H. Shi et al., "Embedding Deep Metric for Person Re-identi cation: A Study Against Large Variations," In ECCV 2016.

Re-ranking

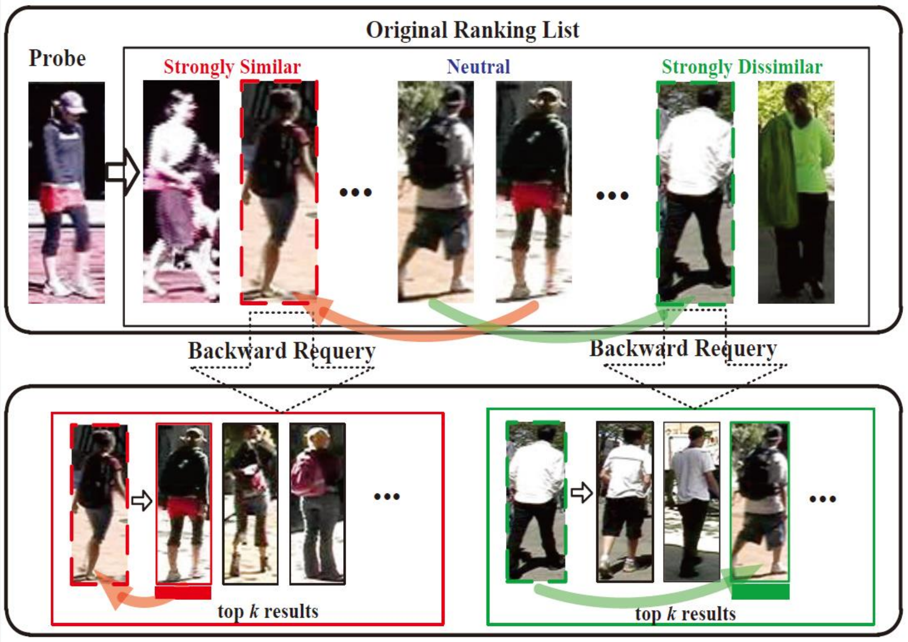
- User feedback based methods (human in the loop)
  - POP
  - HVIL



(a) Train-once-and-deploy re-id models (b) POP: Post rank optimisation [15] (c) HVIL: Human Verification Incremental Learning

Re-ranking

- Context based methods
  - DCIA
  - Bidirectional ranking
  - DSAR



Rank →	1	5	10	25	50
Euc. Dist.+ DCIA	16.29	33.38	47.46	58.86	72.78
DDC [10]	19	-	52	69	80
KISSME+SB [2]	19.3	50.7	63.3	78.2	90.6
KISSME+CCRR [17]	22	49	69	87	95
RIRO [37] (1 Iteration)	28	30	34	51	64
PRRS [4]	33.29	-	78.35	-	97.53
KISSME+ DCIA	38.87	67.96	82.01	93.62	98.36
IRT [1] (1 Iteration)	43	45	46	53	61
LADF+ DCIA	44.67	71.54	83.56	93.82	98.52
POP [23] (1 Iteration)	59.05	60.95	63.10	72.20	-
KCCA+ DCIA	**63.92**	**78.48**	**87.50**	**96.36**	**99.05**

DCIA on VIPeR

Garcia et al., "Person Re-Identification Ranking Optimization by Discriminant Context Information Analysis," In ICCV 2015.

03

PART THREE

Evaluation and Benchmark

Evaluation

- Closed-set scenario
  - Probe:
    - query images to be re-identified
  - Gallery:
    - a set of images from surveillance videos to re-identify probe images
  - Performance measure:
    - Cumulative Matching Characteristic (CMC) curves
    - mAP: mean average precision

mAP is from image retrieval. CMC is more practical for person re-id, because one correct retrieval is already enough for forensic search.

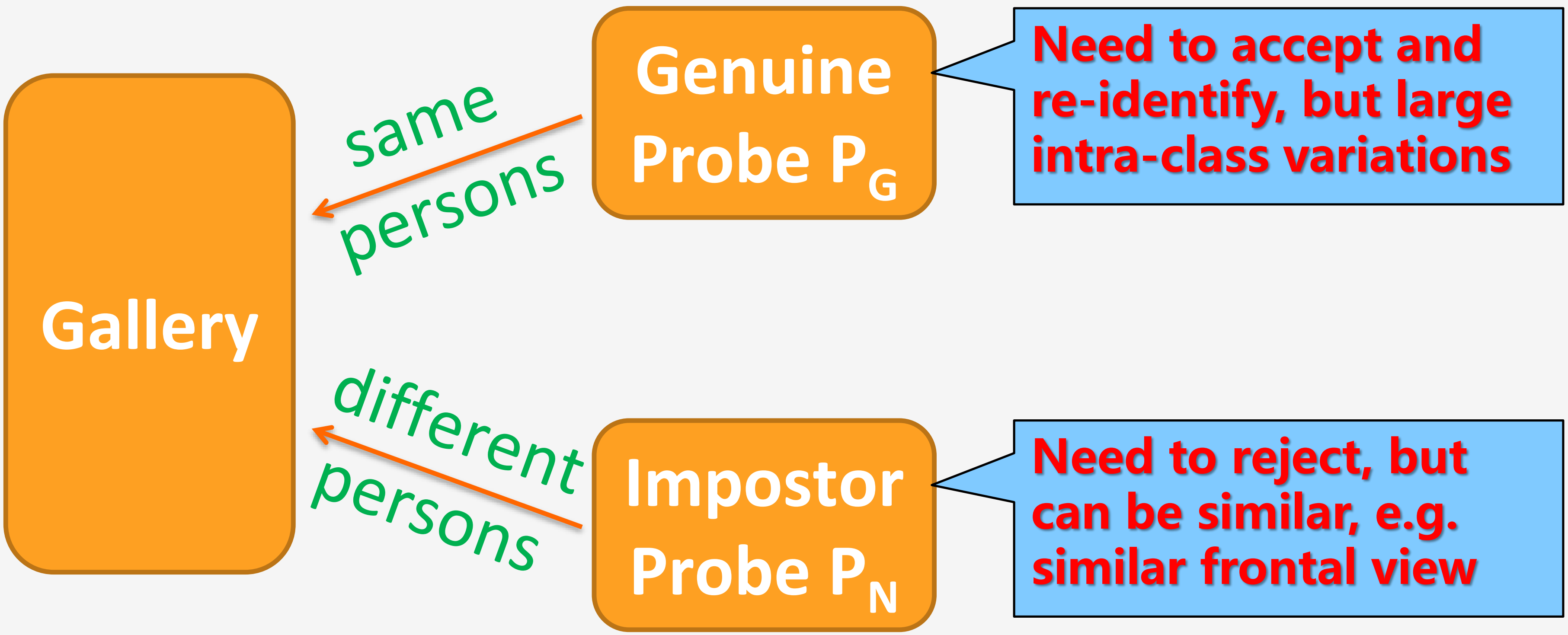
Constraint: each probe image must have the same person appearing in the gallery

Evaluation

- Open-set scenario

Open-set Person Re-identification

- Task: determine the same person of the probe in the gallery, or reject the probe
- Two subsets of probes



Open-set Person Re-identification

- Performance measures:
  - Detection and Identification Rate (DIR): percentage of images in P_G that are correctly accepted and re-identified
  - False Accept Rate (FAR): percentage of images in P_N that are falsely accepted

Closed-set Benchmark Datasets

Dataset	#Cameras	#Persons	#Images	#Views
VIPeR	2	632	1,264	2
ETHZ	1	146	8,555	1
i-LIDS	5	119	476	2
QMUL GRID	8	250	1,275	2
PRID2011	2	200	1,134	2
CUHK01	2	971	3,884	2
CUHK02	5 pairs	1,816	7,264	2
CUHK03	6	1,360	13,164	2
CAMPUS-Human	3	74	1,889	3
Market-1501	6	1,501	32,668	-
MARS	6	1,261	1,191,003	-
DUKE	8	1,404	36,411	-

Open-set Benchmark Datasets

Dataset	#Cameras	#Persons	#Images	#Views
Open-world	6	28	4,096	-
OPeRID	6	200	7,413	5

Closed-set Benchmark Results

Benchmark on DukeMTMC-reID

Methods	Rank@1	mAP
BoW+kissme	25.13%	12.17%
LOMO+XQDA	30.75%	17.04%
PSE	79.8%	62.0%
ATWL(2-stream)	79.80%	63.40%
Mid-level Representation	80.43%	63.88%
HA-CNN	80.5%	63.8%
Deep-Person	80.90%	64.80%
MLFN	81.2%	62.8%
DuATM (Dense-121)	81.82%	64.58%
PCB	83.3%	69.2%
Part-aligned （Inception V1, OpenPose)	84.4%	69.3%
GP-reID	85.2%	72.8%
SPreID (Res-152)	85.95%	73.34%

Open-set Benchmark Results

- On OPeRID

	FAR=1%		FAR=10%	
	Rank=1	Rank=10	Rank=1	Rank=10
IDENTITY	0.84	0.91	7.36	9.21
MAHAL [13]	1.89	1.99	10.50	11.97
KISSME [13]	1.82	1.92	9.99	11.46
LMNN [29]	0.41	0.41	3.97	4.58
ITML [6]	1.18	1.21	8.39	9.27
LADF [19]	1.53	1.74	9.11	10.82
RRDA	**3.99**	**4.35**	**14.51**	**16.72**

Very poor!

S. Liao et al., "Open-set Person Re-identification," In arXiv 2014.

04

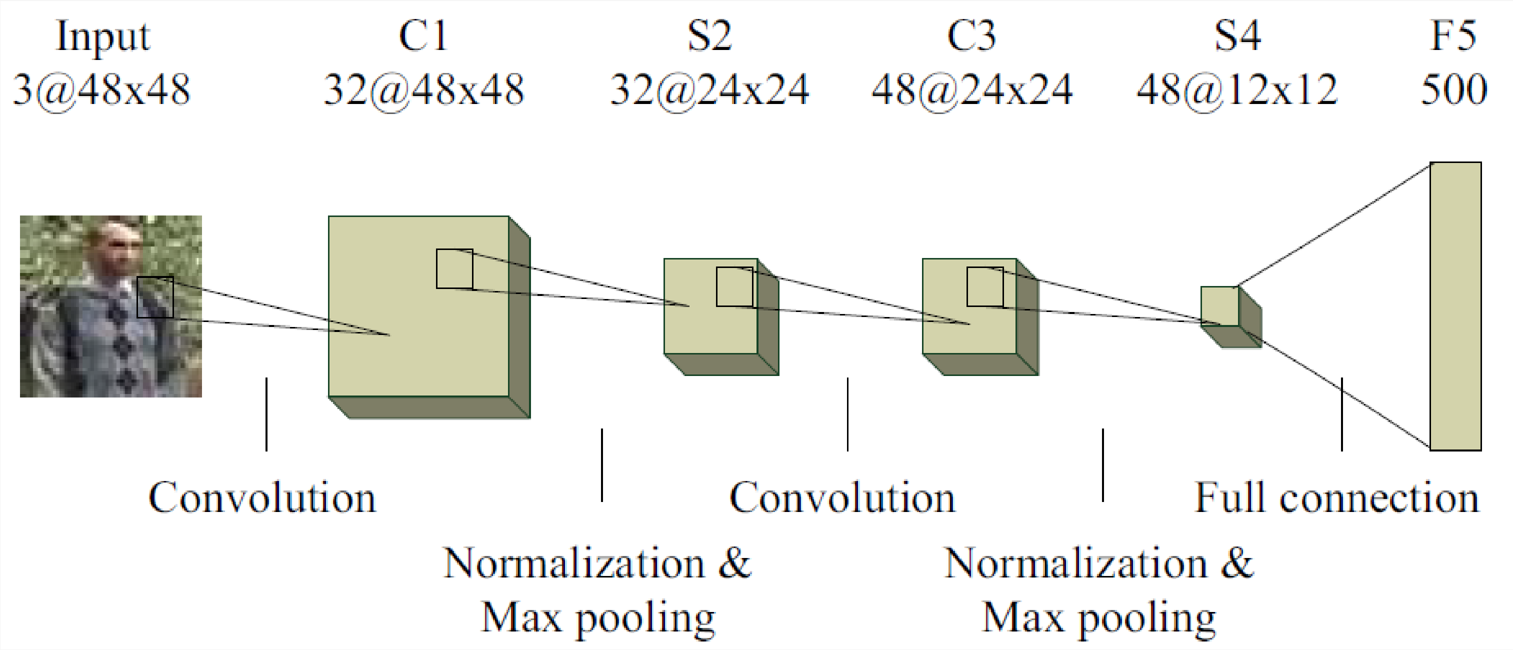
PART FOUR

Future Directions

Future Directions

1

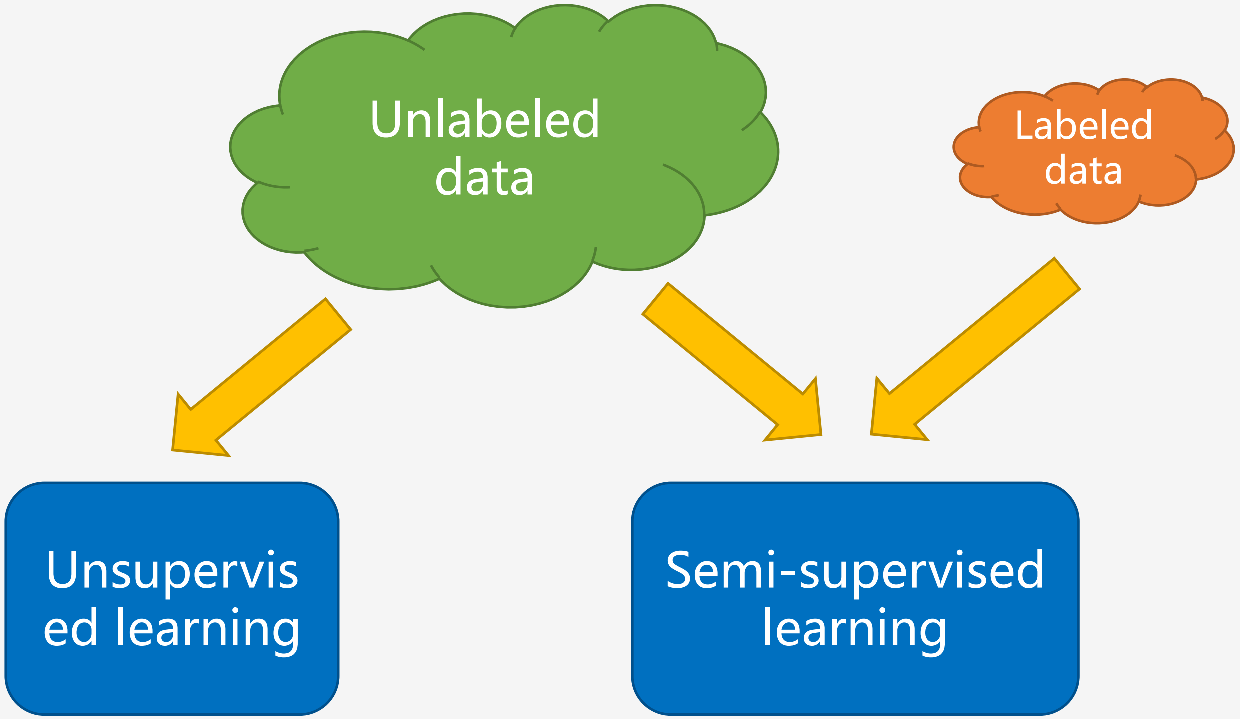
With the help of large datasets, deep learning methods have achieved much better performance, and are becoming more and more important for person re-identification.



Future Directions

2

Due to limited labeled data and large diversity in practical scenarios, semi-supervised learning or unsupervised learning will be potentially useful for practical applications in exploring large amount of unlabeled data.



Future Directions

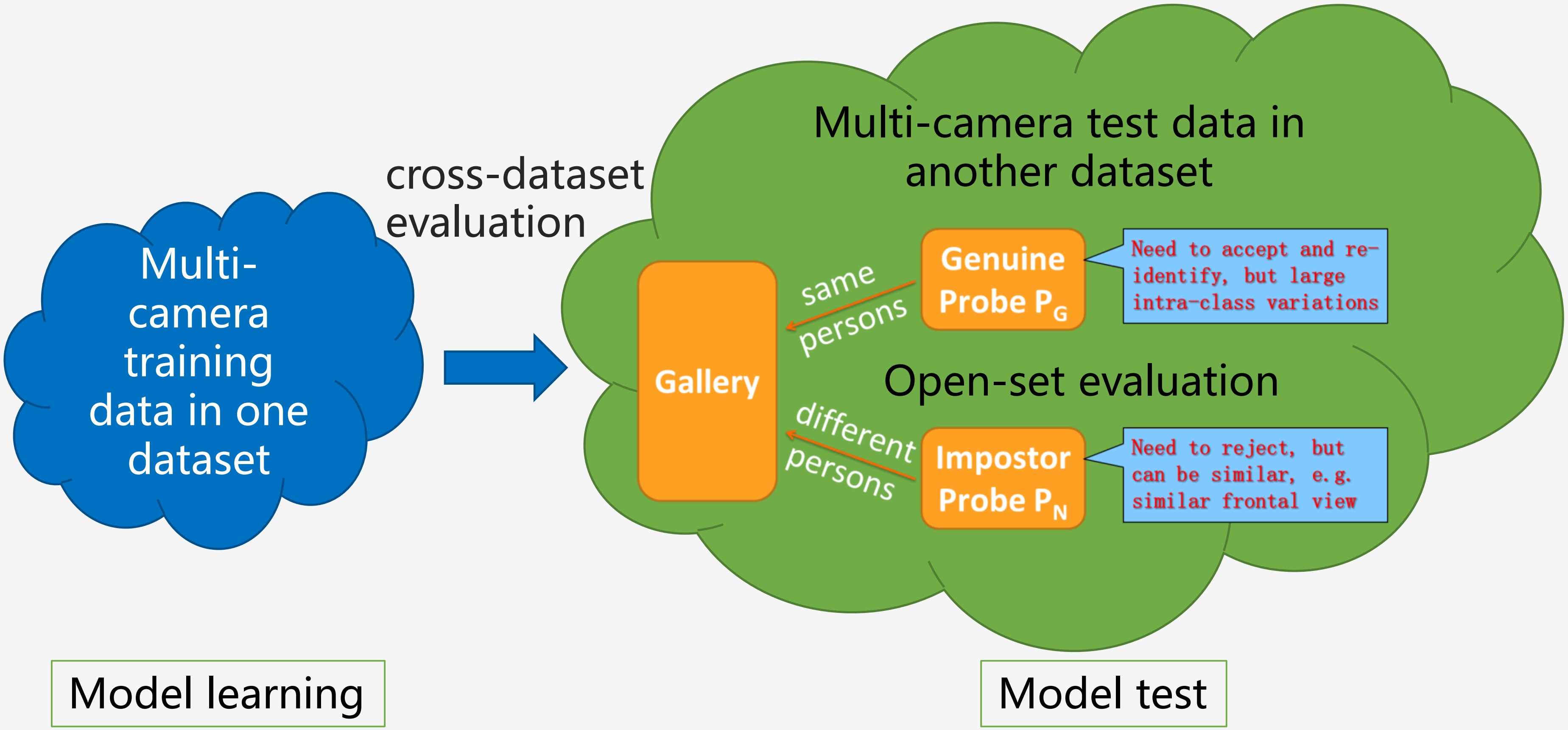
3

Performance of cross-dataset evaluation is still poor. Unsupervised transfer learning and Re-ranking methods may be very useful in improving the performance.

Future Directions

4

For evaluation, open-set person re-identification and cross-dataset evaluation will be preferred in evaluating practical performance.



Thanks!

Shengcai Liao
Institute of Automation
Chinese Academy of Sciences

http://www.cbsr.ia.ac.cn/users/scliao/