Queen Mary University of London
Computer Vision Group

Deep Reinforcement Learning Attention Selection for Person Re-Identification

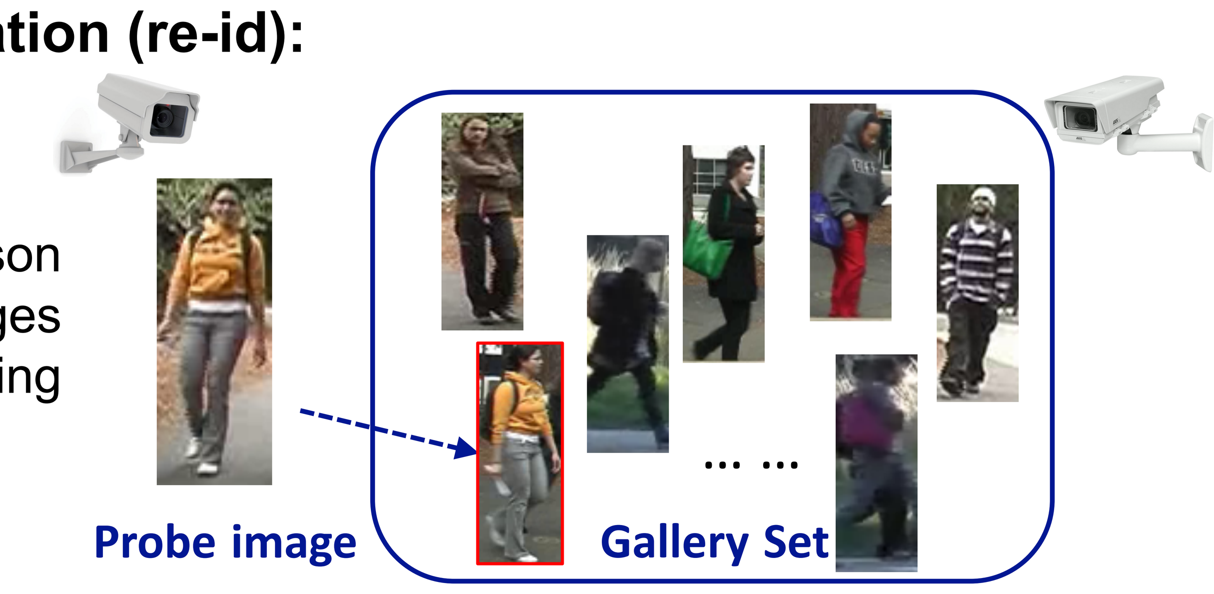
Xu Lan[1] Hanxiao Wang[2] Shaogang Gong[1] Xiatian Zhu[1]
x.lan@qmul.ac.uk hxw@bu.edu s.gong@qmul.ac.uk xiatian.zhu@qmul.ac.uk
[1]Queen Mary University of London, London, UK [2]Boston University, Boston, USA

BMVC 2017 LONDON
28th British Machine Vision Conference

1. Introduction

Person Re-Identification (re-id):

Task: Matching person identity in person images across non-overlapping camera views.

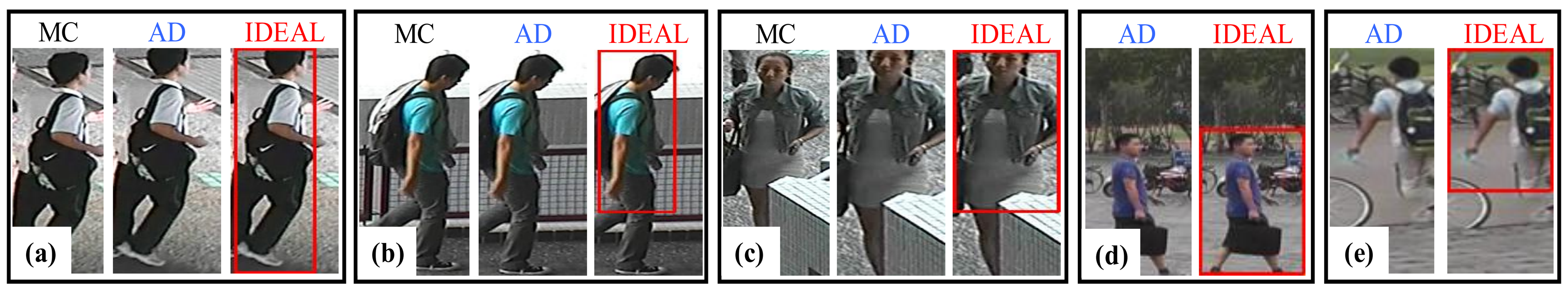


Limitations of existing methods: Assuming accurately labelled person bounding boxes by manually cropping (MC). However, in practice person bounding boxes must be automatically detected (AD) for scalability.

Motivation:

- Automatically detection person suffering from the misalignment (Fig. 1 a,d,e) and occlusion problems (Fig. 1 c)

Figure1: Comparisons of bounding boxes by MC, AD, and our model IDEAL



- Re-id performance drop on AD, compared to MC (**8%** rank-1 drop CUHK03)

Contributions:

- A novel **I**dentity **D**iscriminativ**E** **A**ttention reinforcement **L**earning (**IDEAL**) model for re-id attention selection.
- IDEAL model is trained by pairwise re-id constraints without the need for accurate object bounding box annotations, more scalable to large size data.

2. Methodology

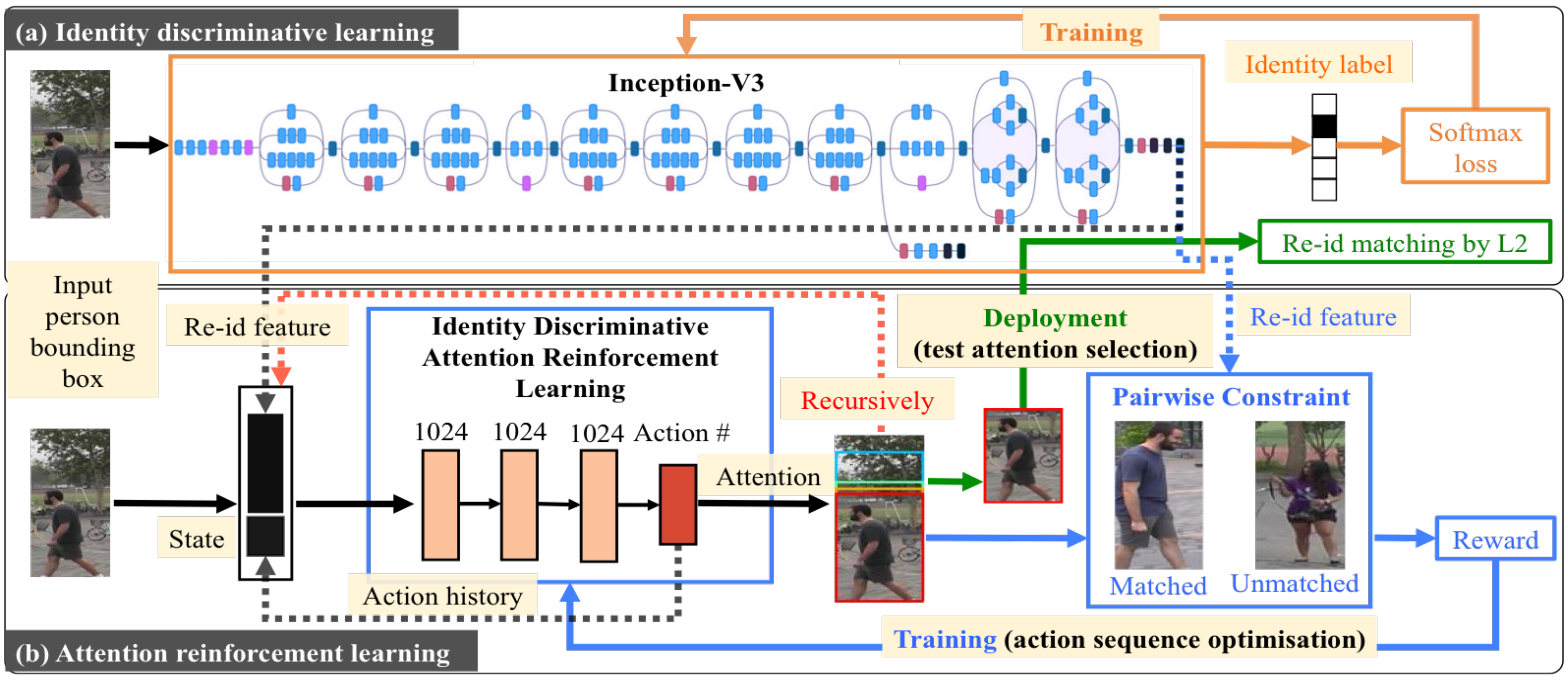
Reinforcement learning re-id attention sequence: Specific Markov Decision Process for re-id attention selection in auto-detected bounding boxes.

- **Environment**: Input person bounding box image.
- **Actions**: Each action defined by changes in location and size of input image.

(a) Input image (b) Attending actions (Each red dotted box represents the attention window after the action)

- **State**: Define by the current attention window feature and an action history vector.
- **Reward**: Directly relating to the re-id matching criterion.

3. Model Framework and Reward Design



IDEAL model has two subnetworks:

1. A multi-class discrimination network trained by a set of auto-detected person bounding boxes (Fig. (a)).
2. A re-identification attention network by reinforcement learning recursively selecting a salient sub-region (Fig. (b)).

Notations:

I_t: Current attention window
I_t^+: Same identity, different camera
I_t^-: Different identity, same camera
I_t^a: Attention window after action a
x_t, x_t^+ x_t^- and x_t^a: the feature for I_t ,I_t^+, I_t^-, I_t^a

- **Reward by Relative Comparison:**

$$R_t = R_{rc}(\mathbf{s}_t, a) = \left(f_{\text{match}}(\mathbf{x}_t^a, \mathbf{x}_t^-) - f_{\text{match}}(\mathbf{x}_t^a, \mathbf{x}_t^+)\right) - \left(f_{\text{match}}(\mathbf{x}_t, \mathbf{x}_t^-) - f_{\text{match}}(\mathbf{x}_t, \mathbf{x}_t^+)\right)$$

- **Reward by Absolute Comparison:**

$$R_t = R_{ac}(\mathbf{s}_t, a) = \left(f_{\text{match}}(\mathbf{x}_t, \mathbf{x}_t^+)\right) - \left(f_{\text{match}}(\mathbf{x}_t^a, \mathbf{x}_t^+)\right)$$

- **Reward by Ranking:**

$$R_t = R_r(\mathbf{s}_t, a) = \begin{cases} +1, & \text{if } \text{Rank}(\mathbf{x}_t^+|\mathbf{x}_t) > \text{Rank}(\mathbf{x}_t^+|\mathbf{x}_t^a) \\ -1, & \text{otherwise} \end{cases}$$

4. Experiments

- **Comparisons to the State-of-the-Arts re-id performance**

Dataset	CUHK03(AD) [23]				Market-1501(AD) [64]			
					Single Query		Multi-Query	
Metric (%)	R1	R5	R10	R20	R1	mAP	R1	mAP
ITML[10]	5.1	17.7	28.3	-	-	-	-	-
LMNN[55]	6.3	18.7	29.0	-	-	-	-	-
KISSME[21]	11.7	33.3	48.0	-	40.5	19.0	-	-
MFA[58]	-	-	-	-	45.7	18.2	-	-
kLFDA[58]	-	-	-	-	51.4	24.4	52.7	27.4
BoW[64]	23.0	42.4	52.4	64.2	34.4	14.1	42.6	19.5
XQDA[25]	46.3	78.9	83.5	93.2	43.8	22.2	54.1	28.4
MLAPG[24]	51.2	83.6	92.1	**96.9**	-	-	-	-
L_1-Lap [20]	30.4	-	-	-	-	-	-	-
NFST[59]	53.7	83.1	93.0	94.8	55.4	29.9	68.0	41.9
LSSCDL[60]	51.2	80.8	89.6	-	-	-	-	-
SCSP[6]	-	-	-	-	51.9	26.3	-	-
TMA[32]	-	-	-	-	47.9	22.3	-	-
HL[46]	-	-	-	-	59.5	-	-	-
HER[51]	60.8	87.0	**95.2**	**97.7**	-	-	-	-
FPNN[23]	19.9	-	-	-	-	-	-	-
DCNN+[2]	44.9	76.0	83.5	93.2	-	-	-	-
EDM[43]	52.0	-	-	-	-	-	-	-
SICI[49]	52.1	84.9	92.4	-	-	-	-	-
SSDAL[44]	-	-	-	-	39.4	19.6	49.0	25.8
S-LSTM [48]	57.3	80.1	88.3	-	-	-	61.6	35.3
eSDC[61]	7.7	21.9	35.0	50.0	33.5	13.5	-	-
CAN[26]	63.1	82.9	88.2	93.3	48.2	24.4	-	-
GS-CNN[47]	**68.1**	**88.1**	**94.6**	-	**65.8**	**39.5**	**76.0**	**48.4**
IDEAL	**71.0**	**89.8**	93.0	95.9	**86.7**	**67.5**	**91.3**	**76.2**

- **IDEAL attention selection visualisation:** (a) Two examples of action sequence for attention selection; (b) Two examples of IDEAL attention selection for re-id; (c) Seven examples of IDEAL attention selection; (d) A failure case in reducing distraction when the original auto-detected (AD) bounding box contains two people; (e) Four examples of IDEAL selection on significantly poor auto-detected bounding boxes.

- **Evaluations on Different Attention Selection Strategy**

Dataset	CUHK03 [23]				Market-1501 [64]			
Metric (%)	R1	R5	R10	R20	R1(SQ)	mAP(SQ)	R1(MQ)	mAP(MQ)
eSDC [61]	7.7	21.9	35.0	50.0	33.5	13.5	-	-
CAN [26]	63.1	82.9	88.2	93.3	48.2	24.4	-	-
GS-CNN [47]	68.1	88.1	**94.6**	-	65.8	39.5	76.0	48.4
No Attention	64.9	84.5	92.6	95.7	84.5	64.8	89.4	72.5
Random Attention	54.1	79.2	85.9	90.4	80.3	54.6	85.1	66.7
Centre Attention (95%)	66.1	86.7	91.1	94.9	84.1	64.2	88.6	69.4
Centre Attention (90%)	64.1	85.3	90.3	93.5	82.7	60.3	87.5	65.3
Centre Attention (80%)	51.9	76.0	83.0	89.0	74.7	48.5	83.4	57.6
Centre Attention (70%)	35.2	62.3	73.2	81.7	63.8	39.0	72.3	43.5
Centre Attention (50%)	16.7	38.8	49.5	62.5	39.9	18.5	46.3	23.9
IDEAL(Ranking)	70.3	89.1	92.7	95.4	86.2	66.3	90.8	74.3
IDEAL(Absolute Comparison)	69.1	88.4	92.1	95.0	85.3	65.5	87.5	72.3
IDEAL(Relative Comparison)	**71.0**	**89.8**	93.0	**95.9**	**86.7**	**67.5**	**91.3**	**76.2**

5. Conclusion

- **Problem**: Attention learning for improving re-id in auto-detected person images.
- **Method:** Explore reinforcement learning for sequential attention learning.
- **Result**: Our auto-generated attention achieves similar re-id performance as manually labelled.