Transferable joint attribute-identity deep learning for unsupervised person re-identification

Jingya Wang, Xiatian Zhu, Shaogang Gong, Wei Li

Published in: *Conference on Computer Vision and Pattern Recognition (CVPR) 2018*

Centre for Intelligent Sensing
Queen Mary University of London

CIS centre for intelligent sensing

Queen Mary University of London

Transferable joint attribute-identity deep learning for unsupervised person re-identification

- **Research Problem**
- Methodology
- Experiments

Transferable joint attribute-identity deep learning for unsupervised person re-identification

Person re-identification (re-id)

Transferable joint attribute-identity deep learning for unsupervised person re-identification

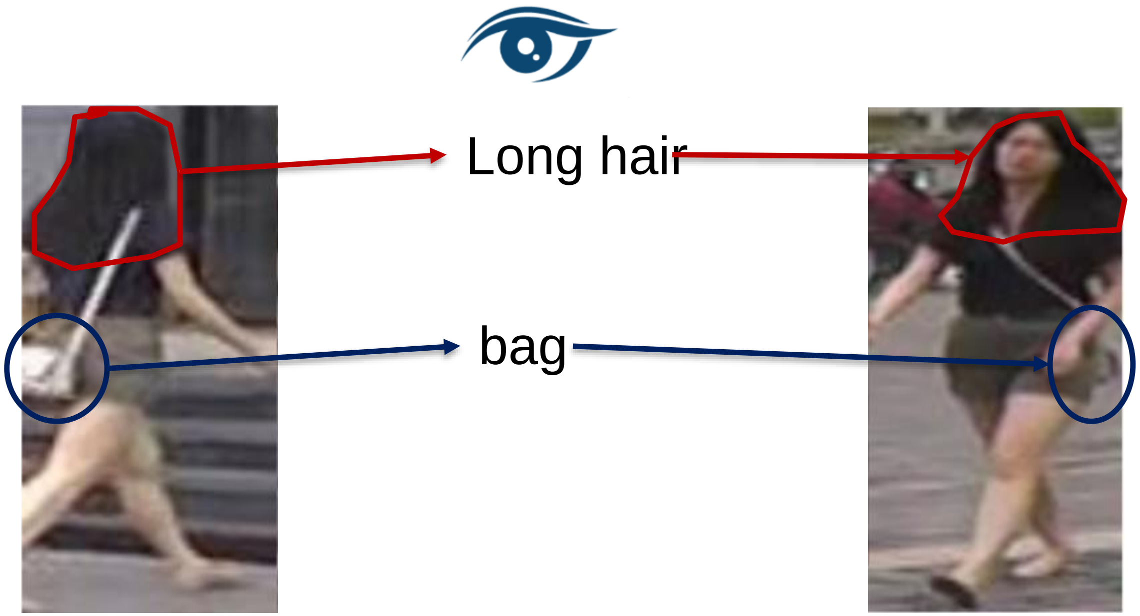
Person re-identification (re-id) aims at matching people across non-overlapping camera views distributed at distinct locations.

Camera A

Camera B

Transferable joint attribute-identity deep learning for unsupervised person re-identification

- How do human brain match person?



Transferable joint attribute-identity deep learning for unsupervised person re-identification

- **Supervised learning:**
  - Metric learning
  - Deep learning

Limitation: need a large number of manually labelled matching pairs for each pair of camera views, poor scalability in practical re-id deployments, expensive to collect

- **Unsupervised Transfer Learning: (Our Focus)**

lack the necessary knowledge on how visual appearance of identical objects changes cross-views due to different view angles, background and illumination -> weaker re-id performances

Transferable joint attribute-identity deep learning for unsupervised person re-identification

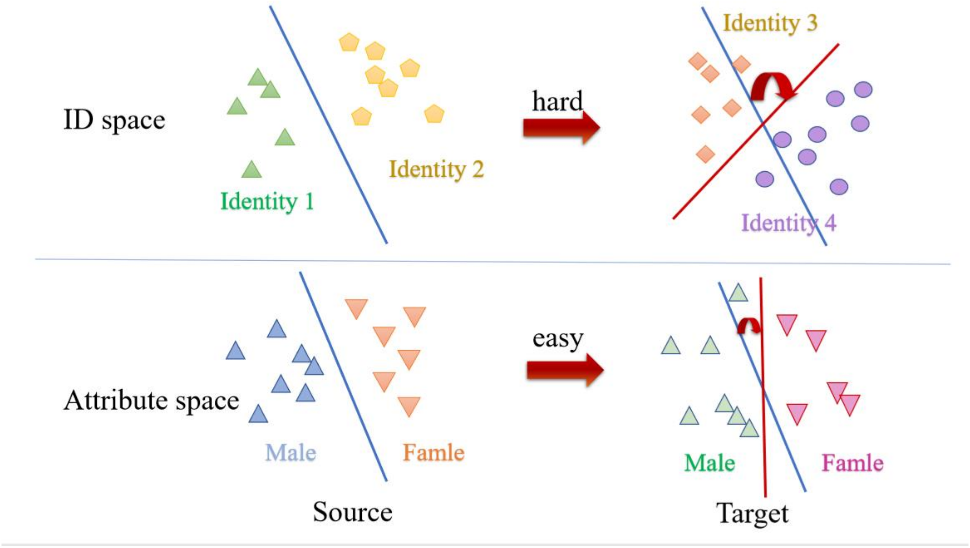
- Research Problem
- **Methodology**
- Experiments

Transferable joint attribute-identity deep learning for unsupervised person re-identification

Challenges:

- Source and target domains have unknown camera viewing conditions
- The identity/class between source and target domains are non-overlapping therefore presents a more challenging open-set recognition problem

-> Transferring knowledge of the source domain to target domain **in attribute space**

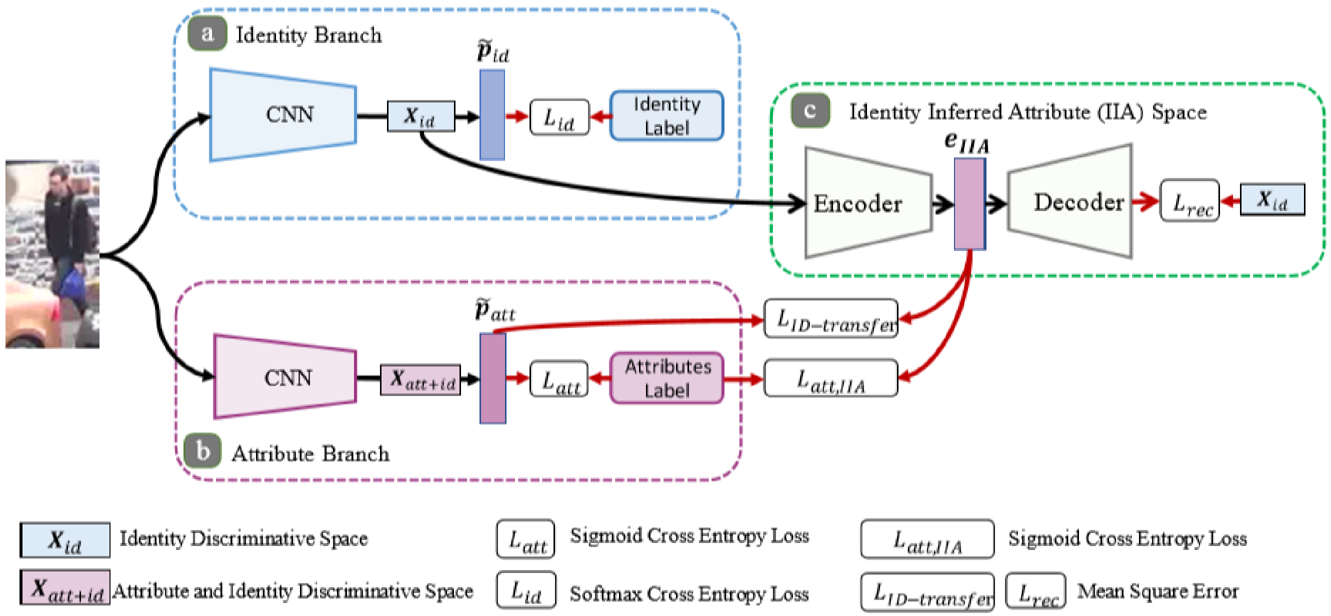


Transferable joint attribute-identity deep learning for unsupervised person re-identification

Challenges:

- The joint exploitation of attribute and identity labels gives rise to the heterogeneous problem

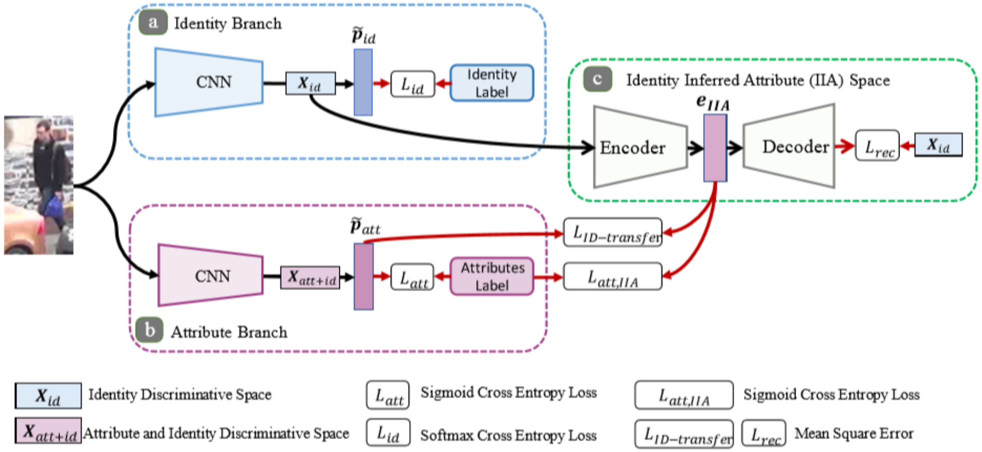
-> **smoothly transferring** the global identity information into the local attribute feature representation space



CIS centre for intelligent sensing

Queen Mary University of London

Transferable joint attribute-identity deep learning for unsupervised person re-identification



(a) Identity Branch

$$L_{\text{id}} = -\frac{1}{n_{\text{bs}}} \sum_{i=1}^{n_{\text{bs}}} \log \left(p_{\text{id}}(\boldsymbol{I}_i^s, y_i^s) \right)$$

(b) Attribute Branch

$$L_{\text{att}} = -\frac{1}{n_{\text{bs}}} \sum_{i=1}^{n_{\text{bs}}} \sum_{j=1}^{m} \Big(a_{ij} \log \left(p_{\text{att}}(\boldsymbol{I}_i, j) \right) + (1 - a_{i,j}) \log \left(1 - p_{\text{att}}(\boldsymbol{I}_i, j) \right) \Big)$$

$$L_{\text{att-total}} = L_{\text{att}} + \lambda_2 \sum_{i=1}^{n_{\text{bs}}} L_{\text{ID-transfer},i}$$

(c) Identity Inferred Attribute Space

$$L_{\text{rec}} = \|\boldsymbol{x}_{\text{id}} - f_{\text{IIA}}(\boldsymbol{x}_{\text{id}})\|^2$$

$$L_{\text{ID-transfer}} = \|\boldsymbol{e}_{\text{IIA}} - \tilde{\boldsymbol{p}}_{\text{att}}\|^2$$

$$L_{\text{attr, IIA}} = -\frac{1}{n_{\text{bs}}} \sum_{i=1}^{n_{\text{bs}}} \sum_{j=1}^{m} \Big(a_{i,j} \log \left(p_{\text{IIA}}(\boldsymbol{I}_i, j) \right) + (1 - a_{i,j}) \log \left(1 - p_{\text{IIA}}(\boldsymbol{I}_i, j) \right) \Big)$$

$$L_{\text{IIA}} = L_{\text{attr, IIA}} + \lambda_1 L_{\text{rec}} + \lambda_2 L_{\text{ID-transfer}}$$

Transferable joint attribute-identity deep learning for unsupervised person re-identification

- Research Problem
- Methodology
- **Experiments**

Transferable joint attribute-identity deep learning for unsupervised person re-identification

1. Datasets:

- **Market-1501**:contains 32,668 images of 1,501 pedestrians, each of which was captured by at most six cameras at a university campus.
  (27 classes of attributes)
- **DukeMTMC-ReID:** contains 2 ~ 426 images per person captured by 8 non-overlapping camera views.
  (23 classes of attributes)
- **VIPeR** : contains 632 identities each with two images captured from two camera views with low resolution.
- **PRID**: consists of person images from two camera views: View A captures 385 people, whilst View B contains 749 people. Only 200 people appear in both views.

(a) Market-1501

(b) DukeMCMT-ReID


(c) VIPER


(d) PRID

Transferable joint attribute-identity deep learning for unsupervised person re-identification

Dataset	VIPeR	PRID	Market-1501		DukeMCMT	
Metric (%)	R1	R1	R1	mAP	R1	mAP
SDALF[9]	19.9	16.3	-	-	-	-
DLLR [18]	29.6	21.1	-	-	-	-
CPS [6]	22.0	-	-	-	-	-
GL [17]	33.5	25.0	-	-	-	-
GTS [46]	25.2	-	-	-	-	-
SDC[55]	25.8	-	-	-	-	-
ISR [31]	27.0	17.0	40.3	14.3	-	-
Dic[19]	29.9	-	50.2	22.7	-	-
RKSL[48]	25.8	-	34.0	11.0	-	-
SAE[25]	20.7	-	42.4	16.2	-	-
AML[38]	23.1	-	44.7	18.4	-	-
UsNCA [38]	24.3	-	45.2	18.9	-	-
CAMEL [53]	30.9	-	**54.5**	**26.3**	-	-
PUL [8]	-	-	44.7	20.1	**30.4**	**16.4**
kLFDA_N [52]	15.9	9.1	-	-	-	-
SADA+kLFDA [52]	15.2	8.7	-	-	-	-
AdaRSVM [33]	10.9	4.9	-	-	-	-
UDML [36]	31.5	24.2	-	-	-	-
SSDAL [43]	**37.9**	20.1	39.4	19.6	-	-
TJ-AIDL$^{\text{Duke}}$	35.1	**34.8**	**58.2**	**26.5**	N/A	N/A
TJ-AIDL$^{\text{Market}}$	**38.5**	**26.8**	N/A	N/A	**44.3**	**23.0**

Conclusion

- ✓ Novel heterogeneous **multi-task joint deep learning** framework for **unsupervised person re-id**
- ✓ **Progressive knowledge fusion** for smoothly transferring the global identity information into the local attribute feature representation space
- ✓ Introduce an **attribute consistency scheme** for cross **domain adaptation**

Thank you