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

➢ How do human brain match person?

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

(a) Identity Branch
\[ L_{id} = -\frac{1}{n_{bs}} \sum_{i=1}^{n_{bs}} \log \left( p_{id}(I_i^s, y_i^s) \right) \]

(b) Attribute Branch
\[ L_{att} = -\frac{1}{n_{bs}} \sum_{i=1}^{n_{bs}} \sum_{j=1}^{m} \left( a_{i,j} \log \left( p_{att}(I_i, j) \right) + (1 - a_{i,j}) \log \left( 1 - p_{att}(I_i, j) \right) \right) \]
\[ L_{att\text{-total}} = L_{att} + \lambda_2 \sum_{i=1}^{n_{bs}} L_{ID\text{-transfer},i} \]

(c) Identity Inferred Attribute Space
\[ L_{rec} = \| x_{id} - f_{IIA}(x_{id}) \|^2 \]
\[ L_{ID\text{-transfer}} = \| e_{IIA} - \tilde{p}_{att} \|^2 \]
\[ L_{attr, II A} = -\frac{1}{n_{bs}} \sum_{i=1}^{n_{bs}} \sum_{j=1}^{m} \left( a_{i,j} \log \left( p_{IIA}(I_i, j) \right) + (1 - a_{i,j}) \log \left( 1 - p_{IIA}(I_i, j) \right) \right) \]
\[ L_{II A} = L_{attr, II A} + \lambda_1 L_{rec} + \lambda_2 L_{ID\text{-transfer}} \]
Transferable joint attribute-identity deep learning for unsupervised person re-identification

- Research Problem
- Methodology
- Experiments
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.
Transferable joint attribute-identity deep learning for unsupervised person re-identification

<table>
<thead>
<tr>
<th>Dataset</th>
<th>VIPeR R1</th>
<th>PRID R1</th>
<th>Market-1501 R1, mAP</th>
<th>DukeMCMT R1, mAP</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDALF [9]</td>
<td>19.9</td>
<td>16.3</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>DLLR [18]</td>
<td>29.6</td>
<td>21.1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CPS [6]</td>
<td>22.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>GL [17]</td>
<td>33.5</td>
<td>25.0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>GTS [46]</td>
<td>25.2</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SDC[55]</td>
<td>25.8</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ISR [31]</td>
<td>27.0</td>
<td>17.0</td>
<td>40.3, 14.3</td>
<td>-</td>
</tr>
<tr>
<td>Dic [19]</td>
<td>29.9</td>
<td>-</td>
<td>50.2, 22.7</td>
<td>-</td>
</tr>
<tr>
<td>RKSL [48]</td>
<td>25.8</td>
<td>-</td>
<td>34.0, 11.0</td>
<td>-</td>
</tr>
<tr>
<td>SAE [25]</td>
<td>20.7</td>
<td>-</td>
<td>42.4, 16.2</td>
<td>-</td>
</tr>
<tr>
<td>AML [38]</td>
<td>23.1</td>
<td>-</td>
<td>44.7, 18.4</td>
<td>-</td>
</tr>
<tr>
<td>UsNCA [38]</td>
<td>24.3</td>
<td>-</td>
<td>45.2, 18.9</td>
<td>-</td>
</tr>
<tr>
<td>CAMEL [53]</td>
<td>30.9</td>
<td>-</td>
<td>54.5, 26.3</td>
<td>-</td>
</tr>
<tr>
<td>PUL [8]</td>
<td>-</td>
<td>-</td>
<td>44.7, 20.1</td>
<td>30.4, 16.4</td>
</tr>
<tr>
<td>kLFDA_N [52]</td>
<td>15.9</td>
<td>9.1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SADA+kLFDA [52]</td>
<td>15.2</td>
<td>8.7</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>AdaRSVM [33]</td>
<td>10.9</td>
<td>4.9</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>UDML [36]</td>
<td>31.5</td>
<td>24.2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SSDAL [43]</td>
<td>37.9</td>
<td>20.1</td>
<td>39.4, 19.6</td>
<td>-</td>
</tr>
<tr>
<td><strong>TJ-AIDL</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Duke</strong></td>
<td>35.1</td>
<td>34.8</td>
<td>58.2, 26.5</td>
<td>N/A, N/A</td>
</tr>
<tr>
<td><strong>Market</strong></td>
<td>38.5</td>
<td>26.8</td>
<td>N/A</td>
<td>N/A, 44.3, 23.0</td>
</tr>
</tbody>
</table>
Transferable joint attribute-identity deep learning for unsupervised person re-identification

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

Thank you