Multidimensional Semantic Disentanglement Network for Clothes-Changing Person Re-Identification

Yongkang Ding (NUAA)  Anqi Wang (NUAA)  Liyan Zhang (NUAA)

Motivation

- Performance degradation of ReID system brought by changing clothes
- Completely filtering out clothing limits the system’s ability to identify people who haven't changed their clothes

Contribution

1. **MSD-NET**
   Multidimensional semantic disentanglement network is designed which diminishes the dependency on clothing attributes.

2. **Semantic segmentation**
   Employ semantic segmentation for pedestrian feature disentanglement and erasing clothing features.

3. **Spatial attention mechanism**
   Employ spatial attention mechanism on semantically segmented features and integrates the learned weights with global features.

4. **Multi-dimensional experiments**
   Multi-dimensional experiments on two mainstream datasets for CC-ReID.

Components

(a) RGB global feature steam (RGB) : use the pretrained ResNet50 as Backbone to extract basic visual features.
(b) Semantic spatial attention steam (Sem-spat) : enhance perception of specific pedestrian parts (head, torso, etc.)
(c) Clothing attribute filtering stream (Attr) : use SCHP to extract decoupled features and filter out clothing-related attributes.

Loss Function

- Dual-stage optimization strategy

Results

Extensive experiments demonstrates MSD-Net’s validity with an average increase of 4.8%

<table>
<thead>
<tr>
<th>Methods</th>
<th>PRCC-CC Top-1</th>
<th>mAP</th>
<th>VC-Clothes-CC Top-1</th>
<th>mAP</th>
</tr>
</thead>
<tbody>
<tr>
<td>TransReID</td>
<td>47.1</td>
<td>49.3</td>
<td>70.0</td>
<td>71.8</td>
</tr>
<tr>
<td>CAL</td>
<td>55.2</td>
<td>55.8</td>
<td>81.4</td>
<td>81.7</td>
</tr>
<tr>
<td>AFL</td>
<td>57.4</td>
<td>56.5</td>
<td>82.5</td>
<td>83.0</td>
</tr>
<tr>
<td>MSD-Net(Ours)</td>
<td>65.3</td>
<td>60.7</td>
<td>84.1</td>
<td>83.4</td>
</tr>
</tbody>
</table>

Visualization