Learning Deep Context-aware Features over Body and Latent Parts for Person Re-identification

Dangwei Li, Xiaotang Chen, Zhang Zhang, Kaiqi Huang
{dangwei.li, xtchen, zhang, kqhuang}@nlpr.ia.ac.cn

Contributions
1) An efficient multi-scale context-aware network for fine-grained pedestrian feature learning.
2) A latent part localization network to atomically localize pedestrian parts without explicitly human parts supervision.
3) An efficient pedestrian representation by fusing global full body and local body parts for pedestrian retrieval.

Object function

1) Prior constraints of body part's center

\[ L_{cen} = \frac{1}{2} \max(0,(x_s - C_x)^2 + (y_s - C_y)^2 - \alpha) \]

2) Prior constraints on value range of scale

\[ L_{pos} = \max(0,\beta - s_y) + \max(0,\beta - s_y) \]

3) Prior constraints of cropped body parts

\[ L_{vis} = \frac{1}{2} \max(0,|s_x + t_x|^2 - \gamma) + \frac{1}{2} \max(0,|s_y + t_y|^2 - \gamma) \]

4) Prior constraints on localization

\[ L_{loc} = L_{cen} + \lambda L_{pos} + \lambda L_{vis} \]

5) Identification loss

\[ L_{cls} = -\sum_{i=1}^{N} \log \frac{e^{W^T x_i + b_y}}{\sum_{j=1}^{N} e^{W^T x_i + b_j}} \]

6) Final object function

\[ L = L_{cls} + M_{loc} \]

Experimental results

- Market-1501
- CUHK03-labeled
- CUHK03-detected

Challenges of person re-identification

- Background
- Pose
- Occlusion
- Light
- Scale
- Viewpoint

Our method

- MSCAN
- Spatial Transformer Networks
- FC_loc
- FC_part
- FC_body

Definition of person re-identification

- CAM A
- Gallery set from CAM B
- CAM A
- Ranking list from CAM B

Visualization of query image and its corresponding retrieval results (body, parts, fusion)

Visualization of original image, rigid parts, and learned latent body parts

Name: Dangwei Li (党伟)
Tel: +86 15600616646
Homepage: http://dangwei.li.github.io