

Supplementary Material for Partial Person Re-identification with Part-Part Correspondence Learning

Tianyu He¹, Xu Shen¹, Jianqiang Huang¹, Zhibo Chen², and Xian-Sheng Hua¹

¹DAMO Academy, Alibaba Group

²University of Science and Technology of China

timhe.hty@alibaba-inc.com

1. Summary

The supplementary material are organized as follows:

- Section 2: We introduce the implementation details for all experiments.
- Section 3: We visualize some examples in the testing set.

2. Implementation Details

In this section, we give details of experiments for the reproducibility.

2.1. Experiments on Person Re-Identification

We pre-train the *gated layout rectifier* (GLRec) in a self-supervised manner as illustrated in the paper. The model is pre-trained using SGD with initial learning rate 5×10^{-2} to train the models for the first 60 epochs. Then we linearly decay the learning rate to 5×10^{-5} in the next 100 epochs. The batch size is set to 128.

For the backbone and the *corresponding region locator* (CRLoc) training, we use the Adam optimizer for model optimization. We first train the backbone network for 120 epochs with initial learning rate as 5×10^{-4} . The learning rate is decayed by a factor of 0.1 for every 40 epochs. The batch size is 64, constructed from 16 identities. After that, the GLRec, the backbone network and the CRLoc are jointly trained with 40 epochs in the assistance of Adam optimizer. The initial learning rate is also set to 5×10^{-4} . Our models are implemented on PyTorch.

2.2. Experiments on Face Recognition

Similar to person re-identification, the GLRec is pre-trained with the same hyper-parameters. While for the joint training, We set batch size as 200. We exploit the SGD optimizer with initial learning rate 1×10^{-1} , which is decayed by a factor of 0.1 for every 5 epochs. The backbone network is able to achieve 98.45% verification accuracy on

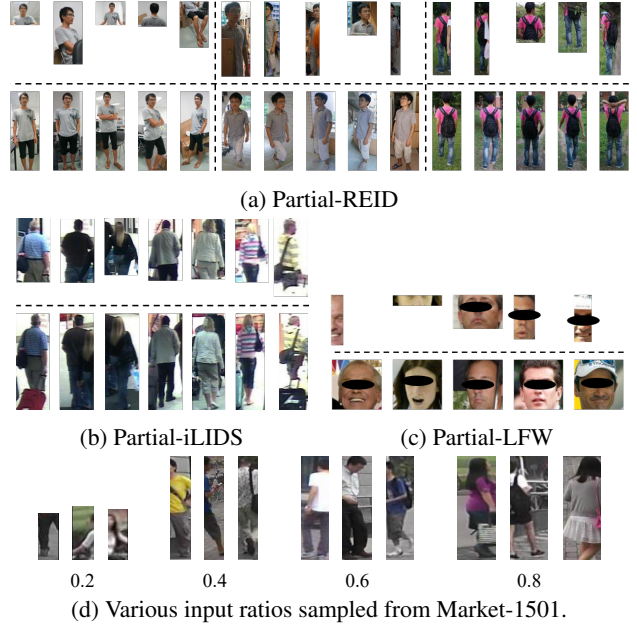


Figure 1: Visualization of the datasets. The top row is the partial images. The middle row is the holistic images for each group. The bottom row is the examples for various input ratios.

LFW dataset when trained on CASIA-Webface [2], which is comparable with the original VGGFace [1] model. The models are implemented with PyTorch.

3. Examples of the Datasets

In this section, we visualize some examples in the testing set, as shown in Figure 1. We can observe that the partial images in Partial-REID typically consist of nearly half of the holistic one (*e.g.*, the left body part), while in Partial-iLIDS, the partial images are typically the upper bodies. For Partial-LFW, the partial inputs are randomly sampled from the holistic images with a small ratio, establishing challenges for re-identifying partial query image from the gallery set.

References

- [1] Omkar M Parkhi, Andrea Vedaldi, and Andrew Zisserman. Deep face recognition. 2015. [1](#)
- [2] Dong Yi, Zhen Lei, Shengcai Liao, and Stan Z Li. Learning face representation from scratch. *arXiv preprint arXiv:1411.7923*, 2014. [1](#)