

A multi-scale body-part mask guided attention network for person re-identification

May 8 2019, by Ingrid Fadelli



Image 1: We can pay more attention to human body, upper body and bottom body. Credit: Cai, Wang & Cheng.

Person re-identification entails the automated identification of the same person in multiple images from different cameras and with different backgrounds, angles or positions. Despite recent advances in the field of artificial intelligence (AI), person re-identification remains a highly challenging task, particularly due to the many variations in a person's pose, as well as other differences associated with lighting, occlusion, misalignment and background clutter.

Researchers at the Suning R&D Center in the U.S. have recently developed a new technique for person re-identification based on a multi-scale body-part mask guided attention network (MMGA). Their paper, pre-published on arXiv, will be presented during the 2019 CVPR Workshop spotlight presentation in June.

"Person re-identification is becoming a more and more important task due to its wide range of potential applications, such as [criminal investigation](#), [public security](#) and image retrieval," Honglong Cai, one of the researchers who carried out the study, told *TechXplore*. "However, it remains a challenging task, due to occlusion, misalignment, variation of poses and background clutter. In our recent study, our team tried to develop a method to overcome these challenges."

Instead of focusing on entire images, Cai and his colleagues developed a model for person re-identification that only pays attention to the person of interest, ignoring the background. Taking this idea one step further, their model analyses different body parts of the person in a given image.

"To implement our idea, we creatively proposed a multi-scale body-part mask guided attention network," Cai said. "We apply body masks to guide the training of our model so that it can pay more attention to the [human body](#) in the image. Our model contains two parts: a feature extractor and an attention module."



The top-5 retrieve results for query images are all correct. Credit: Cai, Wang & Cheng.

The feature extractor component of the model devised by Cai and his colleagues can extract discriminative features of people's bodies from images. The model's attention module, on the other hand, guides the MMGA network, highlighting areas of the image (i.e. pixels) that it should pay closer attention to.

The researchers used body masks to guide the training of their model's attention module, as this allows it to discern human bodies from background information. In addition, they split body masks into upper body and bottom body masks, so that the attention module can learn to distinguish between upper and lower parts of a person's body.

"Differently from most current person re-identification methods, which split images into fixed slides, our model can tell exactly where the upper body and lower body are," Cai explained. "Moreover, body masks are only used in the training phase, and we don't require body masks in the inference phase, which makes our model very efficient in practical applications."

To evaluate their model, Cai and his colleagues carried out a series of experiments testing its performance on two datasets, namely the Market-1501 and DukeMTMC-reID datasets. They found that their model can reduce the negative effects of variations in a person's pose, misalignment and background clutter, outperforming state-of-the-art re-identification methods.

The findings gathered by the researchers suggest that attention mechanisms can significantly improve the accuracy of person re-identification networks. Moreover, their study introduced a mask guide attention training method that can further improve this accuracy.

"In our recent work, upper body masks and lower body masks are used to guide the training of the attention module," Cai said. "In the future, we would like to try dividing body masks into finer details such as head, hand, arm, leg, etc., as this could further improve the accuracy of person re-identification."

More information: Honglong Cai, Zhiguan Wang, Jinxing Cheng.
Multi-scale body-part mask guided attention for person re-identification.
arXiv:1904.11041 [cs.CV]. arxiv.org/abs/1904.11041

© 2019 Science X Network

Citation: A multi-scale body-part mask guided attention network for person re-identification

(2019, May 8) retrieved 9 May 2024 from <https://techxplore.com/news/2019-05-multi-scale-body-part-mask-attention-network.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.