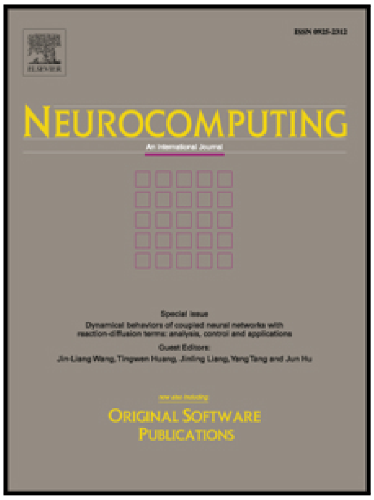
Communicated by Prof. M. Bianchini

Accepted Manuscript



DeepDiff: Learning deep difference features on human body parts for person re-identification

Yan Huang, Hao Sheng, Yanwei Zheng, Zhang Xiong

PII: S0925-2312(17)30346-6
DOI: 10.1016/j.neucom.2017.02.055
Reference: NEUCOM 18129

To appear in: *Neurocomputing*

Received date: 14 June 2016
Revised date: 30 December 2016
Accepted date: 17 February 2017

Please cite this article as: Yan Huang, Hao Sheng, Yanwei Zheng, Zhang Xiong, DeepDiff: Learning deep difference features on human body parts for person re-identification, *Neurocomputing* (2017), doi: 10.1016/j.neucom.2017.02.055

DeepDiff: Learning deep difference features on human body parts for person re-identification

Yan Huang[a,b], Hao Sheng[a,b,*], Yanwei Zheng[a], Zhang Xiong[a]

[a] State Key Laboratory of Software Development Environment, School of Computer Science and Engineering, Beihang University, Beijing, P.R. China

[b] Shenzhen Key Laboratory of Data Vitalization, Research Institute in Shenzhen, Beihang University, Shenzhen, P.R. China

Abstract

Person re-identification is an important part of smart surveillance systems which match people across different scenarios. The most challenging aspect of person re-identification is to design robust features and similarity models that reduce the impact of viewpoints, lightings, background clutters, camera settings, occlusions, and pedestrian poses under different scenarios. DeepDiff is a learning method proposed in this paper which uses deep neural networks to identify the different features of various human body parts, and then evaluate the similarities between those corresponding parts. Given those two corresponding parts, we propose three subnets that show deep difference using original data, feature maps, and spatial variations. We focus on a part-based method which introduces a pyramid partition architecture through different partition granularities on human images. In order to determine whether those two parts belong to the same identity, we utilize two combinations of our three subnets. Lastly, we present a part-based integration validation method so as to achieve better performance. The DeepDiff's effectiveness is validated on three public datasets, including: CUHK03 (labeled, detected), VIPeR and CUHK01 (100 and 486 identities settings). The experimental study shows that the proposed DeepDiff model has promising potential.

Keywords: Person Re-identification, Neural Network, Body Parts, Difference Features, Smart Surveillance

1. Introduction

The demand to analyze pedestrians using intelligent surveillance systems is growing alongside the rapid development of computer vision technology. Between person detection, tracking, and re-identification, the latter underpins crucial applications such as long-term multi-camera tracking and forensic searches [1]. Person re-identification gained a lot of interest in the recent years [2, 3, 4, 5, 6, 7], because it addresses regarding matching people across disjointed camera views in a multi-camera system, or identifying the same identity from different time periods within the same camera view. It can also be formulated to match persons between two sets (”probe set” and ”gallery set” [8]), which contain individuals captured in different scenarios (different views or time periods); the person re-identification algorithm calculates the similarities between any two persons respectively taken from these two sets.

Different scenarios provide complex variations (e.g. viewpoints, lightings, background clutters, camera settings, occlusions and pedestrian poses (see Fig.1)). As such, person re-identification is challenging since the biological characteristics (e.g. gait, face and physical appearance) are hardly used on low resolution images. To design person re-identification system, two fundamental modules are necessary: robust feature extraction and similarity measure. The former can reduce the impact of variations under different scenarios, which is the foundation of similarity measure. Unlike biological characteristics, the features used in person re-identification range from color his-

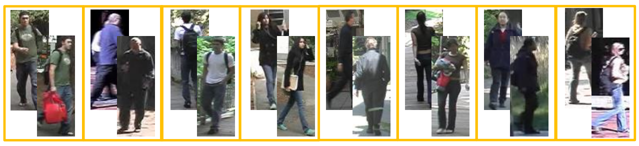


Figure 1: Typical examples of person images captured from different camera views. Each box contains two images with the same identity. Huge variations are illustrated on each person.

togram [9], texture character [10], body structure representation [11] or the combination of multiple features [3]. Given these handcrafted features, similarity measure algorithms are used to identify the distance between person images with the same identity (D_{same}) or different identities (D_{diff}) under certain conditions, such as $D_{same} > D_{diff}$. Although this proves effective, it depends on the quality of the selected features which requires a great deal of domain knowledge and experience [12].

The combined handcrafted features and similarity measure algorithms try to find the differences amongst different images so as to identify whether or not two images belong to the same person. However, this kind of combination is ineffective in sharing these differences between two modules, because they are not integrally optimized and the information on those differences cannot be propagated to each other. To deal with this problem, the deep neural network is the most promising for the overall design because it establishes automatic interaction between the two modules [5], and has made a large contribution