

Dual-Branch Network Fused With Two-Level Attention Mechanism for Clothes-Changing Person Re-Identification

Yong Lu, Minzu University of China, China*

Ming Zhe Jin, Minzu University of China, China

ABSTRACT

Clothes-changing person re-identification is a hot topic in the current academic circles. Most of the current methods assume that the clothes of a person will not change in a short period of time, but they are not applicable when people change clothes. Based on this situation, this paper proposes a dual-branch network for clothes-changing person re-identification that integrates a two-level attention mechanism and captures and aggregates fine-grained person semantic information in channels and spaces through a two-level attention mechanism and suppresses the sensitivity of the network to clothing features by training the clothing classification branch. The method does not use auxiliary means such as human skeletons, and the complexity of the model is greatly reduced compared with most methods. This paper conducts experiments on the popular clothes-changing person re-identification dataset PRCC and a very large-scale cross-spatial-temporal dataset (LaST). The experimental results show that the method in this paper is more advanced than the existing methods.

KEYWORDS

Channels, Clothes Features, Clothes-Changing Person Re-Identification, Complexity, Dual-Branch Network, Fine-Grained Person Semantic Information, Spaces, Two-Level Attention Mechanism

INTRODUCTION

Person re-identification technology, a key technology within intelligent surveillance systems, is regarded as an image retrieval problem. Person re-identification technology is a necessary technology for intelligent surveillance systems in public places for instances like locating criminals. It can also be applied to intelligent security, epidemiological investigations, and intelligent transportation. Through all-weather monitoring, the technology can prevent the occurrence of crimes like theft and robbery, locate lost persons, and assist intelligent transportation systems in completing the automatic dispatching of people, vehicles, and roads.

When monitoring large amounts of data, traditional manual processing methods are inefficient and costly. The person re-identification technology can improve such problems by quickly locating

DOI: 10.4018/IJWSR.322021

*Corresponding Author

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

and tracking the target. This saves labor costs, improves the accuracy of detection, and has a high application value in intelligent monitoring systems.

Person re-identification aims to search for a targeted person via surveillance videos at different locations and times. Due to factors like the limitations of technology, most of the current research on person re-identification assume that the target's clothes are unchanged (Huang et al., 2018; Jin et al., 2022; Li et al., 2018). Thus, it uses the color, texture, and other features of the clothes as discriminant conditions. However, the problem of changing clothes is unavoidable when re-identifying a person over an extended time. There is also the problem of changing clothes in some short-term scenarios. For example, suspects usually change clothes to avoid identification and tracking. The original method will no longer be applicable in the clothes-changing scenario because people may be wrongly matched if wearing similar clothes. To address the issue, this article studies problems related to clothes-changing person re-identification.

To avoid the interference of clothes, some clothes-changing re-identification methods attach modal inputs along with the input image (Chao et al., 2019; Chen et al., 2021; Qian et al., 2020; Shu et al., 2021; Yang et al., 2019). These include three-dimensional (3D) shapes, bones, and contour (Chao et al., 2019; Chen et al., 2021; Qian et al., 2020). However, these methods often require additional models to capture multimodal information. This, in turn, increases the complexity of the model. In fact, original images contain rich clothing-independent information, which is largely underutilized.

This article aims to better mine information unrelated to clothes in the image. Thus, it adds a two-level attention module to the model, acting on the features extracted by the backbone network in space and channel, respectively. Then, it obtains a multi-scale fine-grained attention map. The module can more effectively capture the semantic information of persons in the channel and space, as well as eliminate the influence of irrelevant background as it focuses on features related to an individual. In view of the influence of the clothes feature, this article sets up a clothes classification branch. It also suppresses the sensitivity of the model to clothes features by training this branch. Experiments on popular datasets show that the proposed method is competitive (Shu et al., 2021; Yang et al., 2019).

The contributions of this article can be summarized as follows. First, the article uses IBN-Net as the backbone network. It uses the instance normalization (IN) layer to eliminate individual contrast and extracts features, such as person texture and outline, through the batch normalization (BN) layer. It uses a two-level attention module to capture and aggregate more fine-grained features in space and channels. Experiments show that this kind of two-level attention module enables the network to learn complementary global and local features. This is more suitable for clothes-changing person re-identification. Second, to be more suitable for the clothes-changing person re-identification, this article adds a clothes classification branch to suppress the sensitivity of the model to clothes features. Third, this article conducts extensive experiments on commonly used data sets and large-scale cross-temporal dataset. All have achieved strong results.

RELATED WORK

Person re-identification was first recognized as an independent task in the field of computer vision at the 2016 CVPR conference. The task of clothes-changing (long-term) person re-identification was first proposed in 2019. At present, many researchers have devoted themselves to the research of person re-identification algorithm technology without changing clothes. They have achieved remarkable results. However, there are few studies on clothes-changing person re-identification. The results obtained are not ideal.

Non-Clothes-Changing Person Re-Identification

Person re-identification has developed rapidly in recent years. The research on person re-identification is divided into traditional methods and deep learning methods. Traditional methods include two steps of feature extraction and similarity measurement (Liao et al., 2015; Weinberger & Saul, 2008). The

12 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/article/dual-branch-network-fused-with-two-level-attention-mechanism-for-clothes-changing-person-re-identification/322021

Related Content

Result Refinement in Web Services Retrieval Based on Multiple Instances Learning

Yanzhen Zou, Lu Zhang, Yan Li, Bing Xie and Han Mei (2008). *International Journal of Web Services Research* (pp. 77-93).

www.irma-international.org/article/result-refinement-web-services-retrieval/3120

User Interaction Within Online Innovation Communities: A Social Network Analysis

Jiali Chen, Yiyi Li, Mengzhen Feng and Xinru Zhang (2023). *International Journal of Web Services Research* (pp. 1-19).

www.irma-international.org/article/user-interaction-within-online-innovation-communities/330988

Service Discovery and Composition Based on Contracts and Choreographic Descriptions

Mario Bravetti and Gianluigi Zavattaro (2013). *Adaptive Web Services for Modular and Reusable Software Development: Tactics and Solutions* (pp. 60-88).

www.irma-international.org/chapter/service-discovery-composition-based-contracts/69470

Linked Data: Connecting Spatial Data Infrastructures and Volunteered Geographic Information

Carlos Granell, Sven Schade and Gobe Hobona (2011). *Geospatial Web Services: Advances in Information Interoperability* (pp. 189-226).

www.irma-international.org/chapter/linked-data-connecting-spatial-data/51488

Fully Automated Web Services Discovery and Composition Through Concept Covering and Concept Abduction

Azzurra Ragone, Tommaso Di Noia, Eugenio Di Sciascio, Francesco M. Donini, Simona Colucci and Francesco Colasuonno (2007). *International Journal of Web Services Research* (pp. 85-112).

www.irma-international.org/article/fully-automated-web-services-discovery/3106