"Re-ID BUFF": An Enhanced Similarity Measurement Based on Genetic Programming for Person Re-identification

Yiming Li
mf1933052@smail.nju.edu.cn
State Key Laboratory for Novel Software Technology
Department of Computer Science and Technology
Nanjing University
Nanjing, Jiangsu, China

Lin Shang
shanglin@nju.edu.cn
State Key Laboratory for Novel Software Technology
Department of Computer Science and Technology
Nanjing University
Nanjing, Jiangsu, China

ABSTRACT

Person re-identification (re-ID) is a fundamental link to ensure successful cross-camera tracking. Its overall process is divided into two stages, feature extraction and matching. In the matching stage, the simple Euclid measurement has limitations in score due to the lack of consideration of the vector directions. The majority of recent contributions focus on the use of neighborhood information of the query image, which makes the calculation process more complex. In addition, these similarity measurements are predefined and therefore evaluated. In this paper, we propose a re-ID similarity measurement based on Genetic Programming (GP). The similarity measurement formula is automatically evolved and constructed through the optimization using specific function set, terminal set and fitness function targeting re-ID. For the training process of GP, we propose the feature triplet-dataset like the triplet-loss based on the existing re-ID datasets. We conduct sufficient experiments on three benchmark datasets, comparing with features of different quality and various measurements. The proposed method has a better effect than the Euclid measurement and achieves a general improvement on the mAP and rank-1 of combination with other metrics. Therefore, our method can be used as a gain "buff" to enhance the score of other methods on the original basis.

CCS CONCEPTS

• **Computing methodologies → Computer Vision Task**; **Matching**; **Genetic Programming**; *Discrete space search.*

KEYWORDS

Person re-identification, Genetic programming

ACM Reference Format:
Yiming Li and Lin Shang. 2021. "Re-ID BUFF": An Enhanced Similarity Measurement Based on Genetic Programming for Person Re-identification. In *2021 Genetic and Evolutionary Computation Conference Companion (GECCO '21 Companion), July 10–14, 2021, Lille, France.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3449726.3459432

GECCO '21 Companion, July 10–14, 2021, Lille, France

ACM ISBN 978-1-4503-8351-6/21/07...$15.00.
https://doi.org/10.1145/3449726.3459432

1 INTRODUCTION

Given a query image containing a person (e.g., pedestrian, child, suspect, etc.), person re-identification (re-ID) aims to find the same pedestrian at different locations across the cross-domain cameras. Re-ID has become a hot topic in the field of computer vision with the increasing demand for security. Despite of the recent massive research on re-ID due to the neural network development, especially the CNNs (Convolutional Neural Networks) [1, 2], it is still a challenging mission along with large variations on persons like pose, occlusion, background clutter and detection failure.

The overall framework of re-ID is composed of two modules, namely feature extraction and feature matching. Due to the effectiveness of CNNs, related studies on feature extraction module have sprung up and achieved a huge improvement in effect. However, limited attention has been paid to the feature matching module in the re-ID community. Because of the limitation of the noise factors mentioned above and the simple Euclid distance metric, pedestrian matching errors has frequently appeared in the initial ranking list. Using a re-ranking method with better matching accuracy is preferred yet complicated. In this paper, we also focus on the re-ranking link in feature matching module.

Contributions have exploited the similarity relationships between top-ranked images (such as k-nearest neighbors and k-recipro-cal nearest neighbors) in the original ranking list [3, 4]. We observe that initial distances such as Euclid distance or Cosine distance often contain many errors in high-order pictures due to their respective limitations, and some distance formulas using k-nearest neighbors are often onerous and time-consuming, therefore we consider to find a simple and feasible formula by automatically searching. For this reason, searching algorithms like genetic programming (GP) may be useful for finding suitable solutions.

We propose a measurement formula for re-ID's re-ranking based on GP automatic searching. Our approach is composed of three steps. Firstly, we introduce the concept of feature triplet-dataset to help GP train to get a best performing GP-tree, and then use this GP-tree to calculate the new similarity distance for the features of the query and gallery in the re-ID. Finally, we conduct sufficient experiments on three benchmark datasets by defining the new similarity distance as a "BUFF" to other measures. Results demonstrate that our method is able to find an effective measurement method, which not only is better than the original Euclid measurement, but also can be used as a plug-in to improve the effect of most other measurement methods.

2 PROPOSED METHOD

2.1 Overall Algorithm

Intended to search for a credible and effective distance calculation formula, the overall framework of the proposed "Re-ID BUFF" algorithm should be divided into three stages. The first phase is how to deal with the dataset problem, because the dataset given by traditional re-ID cannot be directly applied to GP training. For dealing with this problem, we propose the feature triplet-dataset to form a new dataset. The form of the original dataset is like this, $[image1, id1]$,$[image2, id2]$, ..., $[imageK,\ idK]$. It has a total of K person images, and each picture has a ID. While the data in new dataset is described as $\{Same[1], Diff[1], id1\}$, $\{Same[2], Diff[2], id2\}$, ..., $\{Same[H], Diff[H], idH\}$. Its members are divided into groups, with three ones in each group. H is the total number of pedestrian categories in the original dataset. $Imagex$ and $imagey$ with the same ID, idi, constitute $Same[i]$ while $imagex$ and $imagez$ with the different IDs make up $Diff[i]$. In fact, these "$image$" exactly represent the deep features of the $image$ extracted by CNNs and the specific details are introduced later. The second stage is aimed at the GP training task. In this stage, the training effect is achieved by setting specific terminal set, function set and fitness function, and the details are explained in the following subsections. The last stage is to apply the training results of GP to the deep features of different quality of re-ID which are published or obtained by our reproduced CNNs in order to ensure the fairness.

2.2 Terminal Set, Function Set and Fitness Function

2.2.1 Terminal Set.

The $var\_featureA$ represents the variance of the N-dimensional feature vector extracted by CNNs from a pedestrian picture A, usually 2048 dimensions. Similarly, the $var\_featureB$ is the variance of another pedestrian B while there is no restriction on the identities of A and B, they can be of the same ID or different IDs. The var_diff represents the variance of the difference between the N-dimensional feature vectors of the two pictures A and B. In the same way, the definitions of these three variables, i.e., $mean\_featureA$, $mean\_featureB$ and $mean\_diff$ are the same as before, but they are the average of the feature vectors. We use the variance and average of the feature vectors instead of the individual value of the feature vector as inputs, because they are more likely to contain the global information of this feature than the individual value. In addition, we also set a variable with the range of -1 to 1.

2.2.2 Function Set.

We have adopted nine alternative functions in total: six arithmetic operators $\{Add, Sub, Mul, ProtectedDiv, ProtectedSqrt, exp\}$, two trigonometric operators $\{cos, sin\}$ and one logical operator $\{max\}$. Since GP adopts the form of tree structure in this paper, we stipulate that the number of inputs of function nodes is 1 or 2. Among the arithmetic operators, we add an extremely decimal to the denominator before performing the division for the case where the denominator is 0 and we achieve the square root of a negative number by multiplying the sign of the value. Other functions are the same as usual mathematical calculations. In order to ensure the diversity of functions, we add a logical operator and what the max node realizes is the function of finding a larger value between two inputs.

2.2.3 Fitness Function.

We utilize a single objective fitness function and it is defined as Eq.1.

$$\begin{aligned} fitness &= \{\frac{\sum_{i=1}^{H} Dist\_Same[i] - \sum_{i=1}^{H} Dist\_Diff[i]}{H}\} \\ weights &= [-1] \end{aligned} \tag{1}$$

Our goal is that fitness function designed can obtain the formula which achieves the following effects: 1. The calculated distance between the pictures of similar pedestrians becomes closer; 2. The calculated distance between the pictures of heterogeneous pedestrians gets farther. Consequently, we propose the fitness functions Eq.1, where H represents the number of candidates in the set, $Dist\_Same[i]$ and $Dist\_Diff[i]$ respectively represents the similar distance and heterogeneous distance of category i through GP, and $weights$ denotes the weight parameter of fitness function, i.e., a value of -1 represents that the minimum value will be selected during the genetic process and a value of 1 has the opposite meaning.

3 CONCLUSION

In this paper, we step further on the re-ranking link in person re-identification (re-ID). We first propose a re-ID measurement method based on genetic programming (GP). In order to adapt to re-ID, we proposed specific function set, terminal set and fitness function. The performance of the proposed method for four kinds of features on three datatsets demonstrate that our method firstly is more effective for re-ranking than the original Euclid distance, and secondly can generally have good gains for other methods regardless of the quality of the features. In addition, our method is not sensitive to parameters, which further confirms the rationality of the GP's settings proposed by our method. In the future, we would like to focus on better genetic programming settings and consider the integration of multiple genetic programming.

ACKNOWLEDGEMENT

This work is supported by the National Natural Science Foundation of China (No.61672276, No.51975294).

REFERENCES

[1] Honglong Cai, Zhiguan Wang, and Jinxing Cheng. 2019. Multi-Scale Body-Part Mask Guided Attention for Person Re-Identification. In *CVPR*. Computer Vision Foundation / IEEE, 1555–1564.

[2] Yifan Sun, Liang Zheng, Yi Yang, Qi Tian, and Shengjin Wang. 2018. Beyond Part Models: Person Retrieval with Refined Part Pooling (and A Strong Convolutional Baseline). In *Computer Vision - ECCV 2018 - 15th European Conference, Munich, Germany, September 8-14, 2018, Proceedings, Part IV (Lecture Notes in Computer Science, Vol. 11208)*. Springer, 501–518.

[3] Mang Ye, Chao Liang, Yi Yu, Zheng Wang, Qingming Leng, Chunxia Xiao, Jun Chen, and Ruimin Hu. 2016. Person Reidentification via Ranking Aggregation of Similarity Pulling and Dissimilarity Pushing. *IEEE Trans. Multim.* 18, 12 (2016), 2553–2566.

[4] Zhun Zhong, Liang Zheng, Donglin Cao, and Shaozi Li. 2017. Re-ranking Person Re-identification with k-Reciprocal Encoding. In *2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, July 21-26, 2017*. IEEE Computer Society, 3652–3661.