IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-7, Issue-6, June 2019 E-ISSN: 2347-2693

Person Re-identification with feature Aggregation

Madhavi Dayaram Bhamare[1*], Nilesh R. Wankhade[2]

[1,2]Dept. of Computer Engineering, Late G N Sapkal College of Engineering, Nashik, India

Corresponding Author: bhamaremadhavi@gmail.com, Tel.: +91-86981-65207

DOI: https://doi.org/10.26438/ijcse/v7i6.466469 | Available online at: www.ijcseonline.org

Accepted: 12/Jun/2019, Published: 30/Jun/2019

Abstract -Person Re-identification (re-ID) is a critical problem in video analytics applications such as security and surveillance. Although many approaches have been proposed, it remains a challenging problem since persons appearance usually undergoes dramatic changes across camera views due to changes in view angle, body pose and background clutter. Person re-id aims to retrieve a person of interest across spatially disjoint cameras. The system focuses on tackling the person re-ID problem with the proposed metric learning scheme. There is a discriminant metric learning strategy for this testing issue. Most existing metric learning algorithms, it takes both original data and auxiliary data during training which is motivated by the new machine learning paradigm - Learning Using Privileged Information. This system is based on features aggregation. Image dataset is load and the basic operation is performing that is to convert those load images into gray scale. And also create the HOG (Histogram of oriented gradient) descriptor, in this features extraction task completed based on EHD (Edge of histogram descriptor), CLD (Color Layout descriptor), and SCD (Scale Color descriptor). The system aggregates all Features and Generate Train metric. After that an unknown image is load which is comes through gray scale process and HOG descriptor. Classify that images and identify the correct image. Such system is used in many sectors for security purpose.

Keywords— Person Re-identification, Metric Learning, Feature Aggregation, HOG descriptor

I. INTRODUCTION

Person re-id aims to retrieve a person of interest across spatially disjoint cameras. It can be seen as a image retrieval problem [1].To compare a person of interest from a probe camera view to a gallery of candidates captured from a camera that does not overlap with the probe one. If a true match to the probe exists in the gallery, it should have a high matching score, compared to incorrect candidates. Generally speaking, person re-ID involves two sub-problems: feature representation and metric learning. An effective feature representation [3], [4] is critical for person re-ID, which should be robust to complex variations in human appearances from different camera views. The general strategy is to concatenate multiple low-level visual features into a long feature vector.

Most existing metric learning methods are limited in that they compare the distance between a pair of similar/dissimilar instances with a global threshold. Such global threshold based pairwise constraints may suffer from sub-optimal learning performance when coping with some real-world tasks with Complex interclass and intra-class variations, e.g., person re-ID. In our setting, each training instance is represented with two forms of features: one is from the original space and the other is from the privileged space. There are two distance metrics by minimizing the empirical loss penalizing the difference between the distance in the original space and the distance in the privileged space. During training, the distance in the privileged space functions as a local decision threshold to guide the metric learning in the original space like a teacher. The finally learned metric from the original space is used to compute the distance between a probe image and a gallery image during testing. Moreover, we extend the proposed algorithm from the single-view setting to a multiview setting which is able to explore the complementation of multiple feature representations. In the multi-view setting, we simultaneously learn multiple distance metrics from different original feature. A distance metric by minimizing the distance between similar instances while keeping that between dissimilar instances larger than a predefined threshold. Presented a metric learning algorithm by collapsing all examples in the same class to a single point and pushing examples in other classes infinitely far away. Developed a method for learning a distance metric from relative comparison. Presented an information-theoretic metric learning approach, which formulates the problem as that of minimizing the differential relative entropy between two multivariate Gaussians under pairwise constraints on the distance function.

Rest of the paper is organized as follows, Section I contains the introduction of Person Re-identification, Section II contain the related work of Metric Learning and Re-identification, Section III contain the some measures of the topic, Section IV contain the architecture and essential steps of Re-identification, section V explain the Re-identification methodology with flow chart, Section VI describes results and discussion Person Re-identification, Section VII contain the recommendation of Re-identification and Section VIII concludes research work with future directions.

II. RELATED WORK

A. Metric Learning

Many algorithms have been developed to learn distance metric. In this subsection, some classical or related distance metric learning works. Xing et al. [13] proposed to learn a distance metric by minimizing the distance between similar instances while keeping that between dissimilar instances larger than a predefined threshold. Globerson et al. [14] presented a metric learning algorithm by collapsing all examples in the same class to a single point and pushing examples in other classes infinitely far away. Schultz et al. [15] developed a method for learning a distance metric from relative comparison. Davis et al. [16] presented an information-theoretic metric learning approach, which formulates the problem as that of minimizing the differential relative entropy between two multivariate Gaussians under pairwise constraints on the distance function.

Traditional pairwise constrained methods only use the original data while training process. It normally takes the global threshold based decision function, which is too rough to get a reasonable metric. System designs a locally adaptive decision rule by exploiting additional knowledge. A logistic discriminant metric learning scheme is presented that utilize auxiliary information to build a locally adaptive decision rule during training process, this method as the LDML+ for simplicity [1]. LDML+ method only considers a single original feature and another is Extend LDML+ from the single-view setting one to a multi-view setting to use the complementation of multiple original features.

The existing system calculates and compares the distance between two training instances with a global threshold and it decides whether both are similar or not. In this, such global threshold is replaced with the squared distance of the privileged space. There are some conflicts in previous methods because there is no clear identification of images. Tackling the video is most difficult to those systems and tough to identify proper image.

B. HOG Descriptor

The histogram of oriented gradients (HOG) is a feature descriptor used in computer vision and image processing for the purpose of object detection. The technique counts occurrences of gradient orientation in localized portions of an image. This method is similar to that of edge orientation histograms, scale-invariant feature transform descriptors, and shape contexts, but differs in that it is computed on a dense grid of uniformly spaced cells and uses overlapping local contrast normalization for improved accuracy. The essential thought behind the histogram of oriented gradients descriptor is that local object appearance and shape within an image can be described by the distribution of intensity gradients or edge directions. The image is divided into small connected regions called cells, and for the pixels within each cell, a histogram of gradient directions is compiled. The descriptor is the concatenation of these histograms. For improved accuracy, the local histograms can be contrast-normalized by calculating a measure of the intensity across a larger region of the image, called a block, and then using this value to normalize all cells within the block. This normalization results in better invariance to changes in illumination and shadowing.

C. Feature Extraction

In machine learning, feature extraction starts from an initial set of measured data and builds derived values (features) intended to be informative and non-redundant, facilitating the subsequent learning and generalization steps, and in some cases leading to better human interpretations. Feature extraction is related to dimensionality reduction. When the input data to an algorithm is too large to be processed and it is suspected to be redundant, then it can be transformed into a reduced set of features. Determining a subset of the initial features is called feature selection. The selected features are expected to contain the relevant information from the input data, so that the desired task can be performed by using this reduced representation instead of the complete initial data. Feature extraction involves reducing the amount of resources required to describe a large set of data. When performing analysis of complex data one of the major problems stems from the number of variables involved. Analysis with a large number of variables generally requires a large amount of memory and computation power; also it may cause a classification algorithm to over-fit to training samples and generalize poorly to new samples. Feature extraction is a general term for methods of constructing combinations of the variables to get around these problems while still describing the data with sufficient accuracy. Many machine learning practitioners believe that properly optimized feature extraction is the key to effective model construction.

III. METHODOLOGY

The system first extract features from the given image and after that performs feature aggregation that is collects features from given image. Image dataset is load and the basic operation is performing that is to convert those load images into gray scale. And also create the HOG (Histogram of oriented gradient) descriptor is a Feature descriptor used in image processing. Purpose of the HOG Descriptor is to detect Object. In this features extraction task completed based on EHD (Edge of histogram descriptor), CLD (Color Layout descriptor), and SCD (Scale Color descriptor). The system aggregates all Features and Generate Train metric. After that an unknown image is load which is comes through gray scale process and HOG descriptor. Classify that images and identify the correct image. Such system is used in many sectors for security purpose.

A. Architecture

Following figure shows the architecture of proposed methodology.

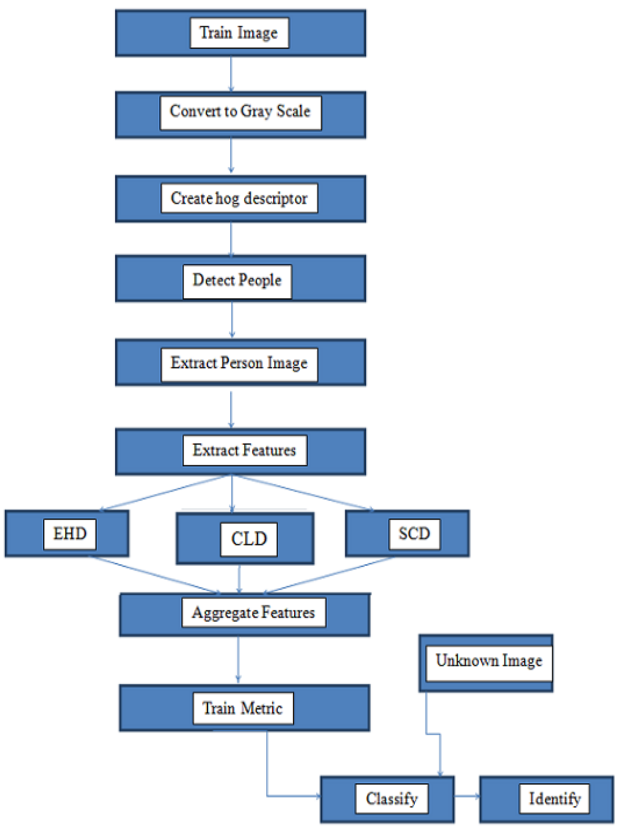

Figure 1. Architecture of proposed methodology.

IV. RESULTS AND DISCUSSION

Basically a gray scale image is simply one in which the only colours are shades of gray. The reason for differentiate such image from any other sort of colour image is that less information needs to be provided for each pixel. A gray colour is one in which the red, green and blue components all have equal intensity in RGB space and so it is only necessary to specify a single intensity value for each pixel.

The grayscale intensity is stored as an 8-bit integer giving 256 possible different shades of gray from black to white. If the significantly levels are evenly spaced then the differences. Between successive gray levels is significantly better than the gray level resolving power of the human eye.
After the gray scale conversion system create the HOG descriptor and performs the feature extraction and feature aggregation and identify correct image.
.

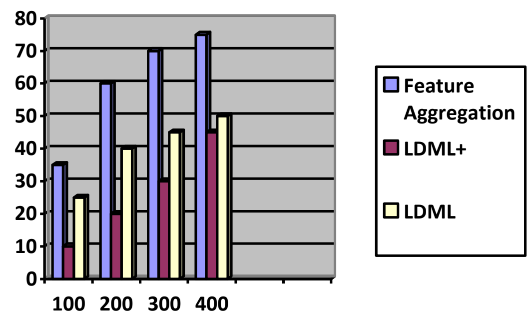

Figure 2. Performance comparison of Feature Aggregation with LDML+ and LDML at varying PCA dimensions.

Above figure represents a graph which shows the performance comparison of feature aggregation with LDML+ and LDML at varying PCA dimension. In this the X-axis denotes PCA dimension and Y- axis denotes Rank. Here the feature aggregation performs better than the other.

V. CONCLUSION AND FUTURE SCOPE

Person re-identification is a challenging task because of complex varieties in human appearances from different camera views. By using this system we can process on video such as capturing the correct image of person from the given video clip. There are so many benefits of this system in image processing. Problems like Queue counting are easily solve by this system and it also used for security purpose. An input video should be in .avi format; this one is a limitation of this system. In future work, various types of video formats will be valid input video to the system.

REFERENCES

[1] Xun Yang, Meng Wang, Dacheng Tao, *"Person Re-identification with Metric Learning using Privileged Information"*, IEEE transactions on Image processing, **2017**.
[2] L. Zheng, Y. Yang, and A. G. Hauptmann, *"Person re-identification: Past, present and future"*, arXiv preprint arXiv: **1610**.**02984**, **2016**.
[3] T. Matsukawa, T. Okabe, E. Suzuki, and Y. Sato, *"Hierarchical Gaussian descriptor for person re-identification"*, in IEEE Conference on Computer Vision and Pattern Recognition, pp. **13631372**,**2016**.
[4] S. Liao, Y. Hu, X. Zhu, and S. Z. Li, *"Person re-identification by local maximal occurrence representation and metric learning"*, in

IEEE conference on Computer Vision and Pattern Recognition, pp. **21972206**, **2015**.
[5] A. Mignon and F. Jurie, Pcca, *"A new approach for distance learning from sparse pairwise constraints",* in IEEE Conference on Computer Vision and Pattern Recognition, pp. **26662672**, **2012**.
[6] S. Liao and S. Z. Li, *"Efficient psd constrained asymmetric metric learning for person re-identification",* in IEEE International Conference on Computer vision, pp. **36853693**, **2015**.
[7] L. Ma, X. Yang, and D. Tao, *"Person re-identification over camera networks using multi-task distance metric learning",* IEEE Transactions on Image Processing.
[8] L. Zhang, T. Xiang, and S. Gong, *"Learning a discriminative null space for person re-identification",* in IEEE Conference on Computer Vision and Pattern Recognition, **2016**.
[9] M. Kstinger, M. Hirzer, P. Wohlhart, P. M. Roth, and H. Bischof, *"Large scale metric learning from equivalence constraints"*, in IEEE Conference on Computer Vision and Pattern Recognition, pp. **22882295**, **2012**.
[10] S. Paisitkriangkrai, C. Shen, and A. van den Hengel, *"Learning to rank in person re-identification with metric ensembles",* in IEEE Conference on Computer Vision and Pattern Recognition, pp. **18461855**, **2015**.
[11] Z. Li, S. Chang, F. Liang, T. S. Huang, L. Cao, and J. R. Smith, *"Learning locally-adaptive decision functions for person verification*", in IEEE Conference on Computer Vision and Pattern Recognition, pp. **36103617**, **2013**.
[12] Xiong, M. Gou, O. Camps, and M. Sznaier, *"Person re-identification using kernel-based metric learning methods",* in European Conference on Computer Vision, pp. **116**,**2014**.
[13] E. P. Xing, M. I. Jordan, S. Russell, and A. Y. Ng, *"Distance metric learning with application to clustering with side-information"*, in NIPS, pp. **505–512**, **2002**.
[14] A. Globerson and S. T. Roweis, *"Metric learning by collapsing classes"*, in NIPS, pp. **451–458**, **2005**.
[15] M. Schultz and T. Joachims, *"Learning a distance metric from relative comparisons,"* in NIPS, **2004**.
[16] J. V. Davis, B. Kulis, P. Jain, S. Sra, and I. S. Dhillon, *"Information theoretic metric learning,"* in ICML, pp. **209–216**, **2007**.
[17] C. Jose and F. Fleuret, *"Scalable metric learning via weighted Approximate rank component analysis"*, in European Conference on Computer Vision. Springer, pp. **875–890**, 2016.
[18] H. Liu, J. Feng, M. Qi, J. Jiang, and S. Yan, *"End-to-end Comparative attention networks for person re-identification"*, IEEE Transactions on Image Processing, **2017**.

Authors Profile

Miss Madhavi D. Bhamare pursed Diploma in Computer technology from Mumbai University, Bachelor of Engineering in computer engineering from Late G. N. Sapkal College Of Engineering, Savitribai Phule Pune University. She is currently pursuing Master of Engineering from Savitribai Phule Pune University.

Prof. (Dr.) N. R. Wankhade pursed Bachelor of Engineering and Master of Engineering degree in Computer Sc. and Engineering. Also pursed Ph.D in Computer Sc. and Engineering from Amravati University. He is currently working as Head and Professor of Department of computer engineering in Late G. N. Sapkal college of Engineering, Nashik in Savitribai Phule Pune university. He is member of LISTE, IEEE, and CSI.