International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 04 Issue: 04 | April -2020 ISSN: 2582-3930

DISTANCE METRIC LEARNING BASED PERSON REIDENTIFICATION.

M.Shiny Rose[1], **N.Neslin**[2]

[1]*Assistant Professor/ECE, Bethlahem Institute Of Engineering,Karungal,Tamilnadu,India.*
[2]*II year ME Communication Systems, Bethlahem Institute Of Engineering,Karungal,Tamilnadu,India*

---***---

Abstract - Person re-identification is a method matching individuals with data set. The prevailing methods face many problems such as (1) Pair wise constraints be affected from slow convergence and poor local optima, (2) Hard negative positive samples which also used in the training of network but this is not attained the enough level. These problems sting out by all positive pair is allowed to compare with negative pair in a mini batch. Then each positive pair is adaptively assigned a weight to modulate its contribution. The new positive pair initiates the algorithm to fine tuning the more hard positive samples. Due to this, some loss function is generated, this loss function reduced by adding a global loss term, to avoid the variances of positive and negative pair distances. By this method person images can be nonlinearly mapped into a low dimensional embedding space where similar samples are kept closer and unlike samples are pushed to more remotely. This method has achieved better performance and is becoming more and more importance for person Re- identification.

Key Words: *Person re-identification, Negative and positive samples, Distance metric learning.*

1.INTRODUCTION

In recent years, person reidentification (re-ID) [1]–[4]has attracted attracted increasing attention in the computer vision community for its critical role in security surveillance applications. It aims to recognize the person of interest across multiple nonoverlapping camera views. Given a probe person image (query), the task is to rank all the person images in the gallery set by the similarity between the query and candidate images and return the most relevant images as retrieval results. To tackle this problem, massive efforts [5]–[8] have been made over the last decade. However, it remains a challenging problem since a person's appearance usually undergoes dramatic variations across camera views due to the changes in view angle, body pose, illumination, and background clutter.

Traditional methods mainly consist of two parts: feature extraction and metric learning. The first part focuses on designing robust hand-crafted features [6]–[10]. The second part [5], [6], [7] aims to learn a suitable distance/ similarity function.

Existing deep CNN-based re-ID approaches can roughly be classified into two groups. The first group [26]–[30], [42] considers each sample independently using an identification loss, which directly casts person re-ID as a multiclass recognition task and usually learns a nonlinear mapping from an input person image to its person identity using a cross-entropy loss.

Deep Metric Learning

Distance metric learning [8], [9] aims to learn a mapping from the input data space to a low-dimensional embedding space, where similar instances are kept closer, while dissimilar instances are pushed farther apart. In recent years, with the remarkable development of deep learning techniques], deep metric learning has shown promising results on multiple computer vision tasks, e.g., face recognition, image retrieval, and fine-grained image recognition. The major difference with standard metric learning is that deep metric learning optimizes feature and metric jointly in an endto- end manner. Existing works can be roughly classified into the following three groups.

The first group of deep metric learning methods trains Siamese networks with a contrastive loss [5]–[7], in which paired data are fed into neural networks. They usually minimize intraclass distance and penalize interclass distance for being smaller than a data-independent threshold. The contrastive loss usually results in the absolute distance.

The second group of methods aims to learn deep embeddings using the triplet loss [3], [4], [6], which takes triplets as an input. Each triplet consists of three samples (anchor, positive, and negative), where the former two samples share the same class label and the third one is from a different class.

The third group of methods focuses on improving the performance by exploiting more negative samples for each update [5], [6] or exploiting the global structure of embedding space [7], [9]. Song *et al.* [7] proposed a lifted structured embedding loss by lifting the vector of pairwise distances within the minibatch to the dense matrix of pairwise distances.

Person Reidentification

This paper focuses on tackling person re-ID with the proposed deep metric learning scheme. We roughly categorize most existing works into two types: traditional methods and deep model-based methods. In this section, we only briefly introduce some representative works. Traditional methods usually cope with two subproblems, i.e., feature learning and metric learning, separately. Some works focus on designing hand-crafted features [6]–[8], [10] that are expected to be robust to complex variations in human appearances from different camera views, e.g., local maximal occurrence feature [7] and Gaussian of Gaussian [6]. Some works focus on learning an optimal distance/similarity

II. METHODOLOGY

This paper addresses the problem of person re-ID by training a deep convolutional network discriminatively. We aim to find a suitable distance function between two person images $\mathbf{x}_i$ and $\mathbf{x}_j$, which is expected to be small if $\mathbf{x}_i$ and $\mathbf{x}_j$ are from the same class or large if they are from different classes. In this paper, it is defined as a squared Euclidean distance between deep embeddings of person images:

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 04 Issue: 04 | April -2020 ISSN: 2582-3930

$d2(\mathbf{x}i, \mathbf{x}j) = \_f\theta(\mathbf{x}i) - f\theta(\mathbf{x}j)\_22$, where $f\theta(\cdot)$ is a nonlinear feature mapping parameterized by the network parameters θ (weight matrices, bias vectors, and so on). Therefore, the key step of this problem .

The basic idea is to learn a nonlinear mapping from person images to discriminative embeddings based on a deep CNN. For each iteration, in the training stage, we feed the network with an identity-balanced minibatch generated by an online random sampling strategy and generate triplets online. The output of the last fully connected layer is _2-normalized and passed into the loss layer. The network parameters are updated by backpropagation supervised by a hardness-aware structural metric learning objective. In the testing stage, given a probe image, we compute the Euclidean distances between the probe and gallery images using the learned deep embeddings. Our network architecture is simpler yet effective.

The lifted structured loss and N-pair loss treat all positive pairs equally. We argue that hard positive pairs contribute more to the training of network than easy positive pairs. For this reason, we improve the lifted structured loss by adaptively assigning larger weights to hard positive pairs. Our loss function also includes a global loss term that minimizes the variances of positive/negative pair distances. It can improve the generalization capability of the network. By this way, the learning performance can be further improved. The problem of person re-ID by training a deep convolutional network discriminatively. We aim to find a suitable distance function between two person images $\mathbf{x}i$ and $\mathbf{x}j$, which is expected to be small if $\mathbf{x}i$ and $\mathbf{x}j$ are from the same class or large if they are from different classes

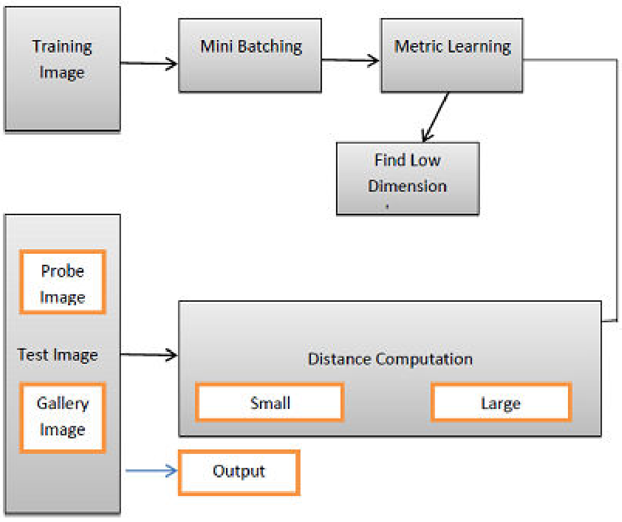

Fig 1.Block Diagram of Proposed System

1.Mini Batch

A lifted structured embedding loss by lifting the vector of pairwise distances within the minibatch to the dense matrix of pairwise distance

2. Metric learning

To learn a mapping from the input data space to a low-dimensional embedding space, where similar instances are kept closer, while dissimilar instances are pushed farther apart. The major difference with standard metric learning is that deep metric learning optimizes feature and metric jointly in an end to end manner. which takes triplets as an input. Each triplet consists of three samples (anchor, positive, and negative), where the former two samples share the same class label and the third one is from a different class. Triplet loss encourages the network to find an embedding space where the anchor sample is closer to the positive sample than the negative sample. The triplet loss results in a relative distance that has been shown to be better than an absolute distance in most tasks. For the triplet-loss-based methods, it is crucial to mine hard samples, e.g., the semi-hard or hardest negative samples, to select triplets violating the triplet constraints for fast convergence. a lifted structured embedding loss by lifting the vector of pairwise distances within the minibatch to the dense matrix of pairwise distances.AN-pair loss that optimizes the log probability of identification loss directly, which is a special case of the lifted structured loss. Then, a structured prediction framework is applied to ensure that the score of the ground-truth clustering assignment is higher than the score of any other clustering assignment, which exploits the global structure of embedding space and results in an impressive performance.

3.Distance Computation

A distance and penalize interclass distance for being smaller than a data-independent threshold. The contrastive loss usually results in the absolute distance. The triplet loss results in a relative distance that has been shown to be better than an absolute distance in most tasks.

A relative distance learning problem by maximizing the probability that relevant samples have a smaller distance than the irrelevant ones. A simple and effective metric learning method by computing the difference between the intraclass and the interclass covariance matrix. As an improvement, a cross-view quadratic discriminates analysis method by learning a more discriminative distance metric and a low-dimensional subspace simultaneously. To take paired images with a binary label $(\mathbf{x}i, \mathbf{x}j, yij)$ as an input to formulate a contrastive loss based on pairwise constraints. It aims to minimize the positive pair distance and penalize the negative pair distance that is smaller than a margin α.

4.Algorithm of CNN Classification

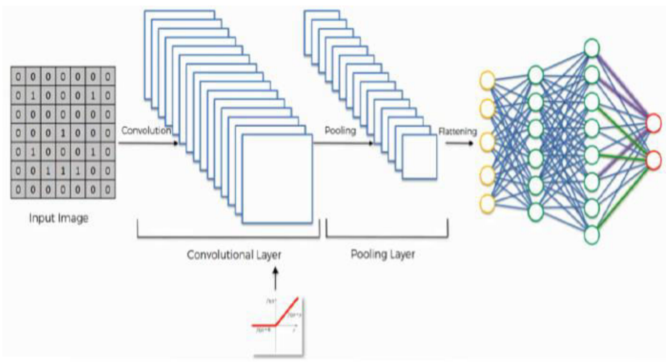

Fig 2. Convolution network Algorithm Diagram

We start off with an input image.We apply filters or feature maps to the image, which gives us a convolutional layer. We then break up the linearity of that image using the rectifier function. The image becomes ready for the pooling step, the purpose of which is providing our convolutional neural network with the faculty of "spatial invariance" which you'll see explained in more detail in the pooling tutorial.

After we're done with pooling, we end up with a pooled feature map. We then flatten our pooled feature map before inserting into an artificial neural network.

Step1 .a Convolutional operation

The first building block in our plan of attack is onvolution operation. In this step, we will touch on feature detectors, which basically serve as the neural network's filters. We will also discuss feature maps, learning the parameters of such maps, how patterns are detected, the layers of detection, and how the findings are mapped out.

Step 1.b ReLU Layer

The second part of this step will involve the Rectified Linear Unit or ReLU. We will cover ReLU layers and explore how linearity functions in the context of Convolutional Neural Networks.

Step 2: Pooling

In this part, we'll cover pooling and will get to understand exactly how it generally works. Our nexus here, however, will be a specific type of pooling; max pooling. We'll cover various approaches, though, including mean (or sum) pooling. This part will end with a demonstration made using a visual interactive tool that will definitely sort the whole concept out for you.

Step 3: Flattening

This will be a brief breakdown of the flattening process and how we move from pooled to flattened layers when working with Convolutional Neural Networks.

The discipline of distance metric learning, a branch of machine learning that aims to learn distances from the data. Distance metric learning can be useful to improve similarity learning algorithms, and also has applications in dimensionality reduction. We describe the distance metric learning problem and analyze its main mathematical foundations. We discuss some of the most popular distance metric learning techniques used in classification, showing their goals and the required information to understand and use them. To ensure that this be a metric—satisfying non-negativity and the triangle inequality. we require that A be positive semi-definite, A>0 gives Euclidean distance.

III. RESULTS AND DISCUSSION

In this paper, we adopt a new data set split for CUHK03, where767 identities are used for training and the remaining 700 identities are employed for testing. The new protocol is more challenging, since the number of training images becomes much less. In the matching process, we calculate the similarities between each query and all the gallery images and return the rank list according to the similarities. All the experiments are under the single query setting. The performance is evaluated by the cumulative matching characteristics (CMC) curve that is an estimation of the expectation of finding the correct match in the top-K matches and the mean average precision (mAP) score. A new testing protocol for the evaluation on the CUHK03 data set.

Minibatch Training means that the gradient is calculated across the entire batch before updating weights. Minibatch having more number of person images.

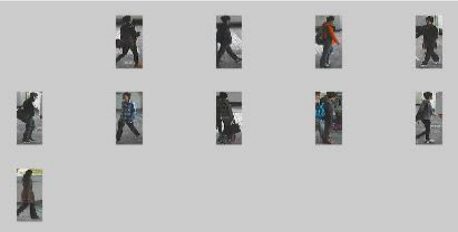

Fig -3. Minibatch Training Image

Hard positive pair is the input image. Input image is select from the minibatch image set. Entire the file location and select the image.

Fig -4. Input Image

Grayscale image is one which the value of each pixel is a single sample representating only an amount of light that is it carries only an intensity information.

Fig -5. Gray Scale Image

Filtering is a technique used for modifying or enhancing an image or part of an image by altering shades and colors of the pixels.Filter is used to increase brightness and contrast

Fig - 6Filter Image

Histogram equalization is a computer image processing technique uesd to improve contrast in images. Histogram intensity value ranges from 0 to L-1

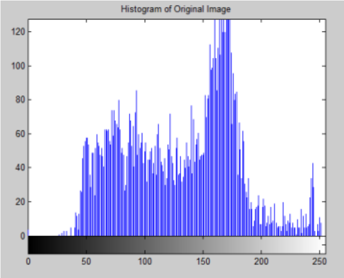

Fig - 7. Histogram Equaliation

Feature extraction is a process of dimensionality reduction by which an initial set of raw data is reduced to more manageable group for processing.

Fig 8. Feature Extraction

The obtained output image is the pairwise image.

Fig 9. Pairwise Image

4. CONCLUSION

An effective person re-ID framework by discriminatively learning a nonlinear deep feature mapping from person images to low-dimensional embeddings, where similar samples are mapped closer to each other, while dissimilar samples are pushed farther apart. The proposed approach jointly learns feature representation and distancemetric in an end-to-end manner. The main contribution of this paper is that we develop a hardness-aware structural metric learning objective where each positive pair is allowed to be compared with all the corresponding negative pairs within minibatch and each positive pair is assigned a hardness-aware weight to adaptively modulate its contribution. Moreover, we incorporate a global loss term that penalizes large variance of positive/negative pair distances into the proposed objective function, which improves the generalization capability of the network effectively. Extensive experimental evaluations on three large-scale data sets have demonstrated the effectivenessof our approach.

ACKNOWLEDGEMENT

I would like to thank God Almighty for his blessings, strength and support to complete this work successfully. I am grateful to all of those with whom I have had the pleasure to work during this and other related projects.

REFERENCES

[1] D. Chen, Z. Yuan, B. Chen, and N. Zheng, "Similarity learning withspatial constraints for person re-identification," in *Proc. IEEE Conf.Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 1268–1277.

[2] M. Geng, Y. Wang, T. Xiang, and Y. Tian. (2016). "Deep transfer learning for person re-identification." [Online]. Available:https://arxiv.org/abs/1611.05244

[3] I. Kviatkovsky, A. Adam, and E. Rivlin, "Deep view aware metric learning for person reidentification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 7,pp. 1622–1634, Jul. 2013.

[4] S. Liao, Y. Hu, X. Zhu, and S. Z. Li, "Relaxed pairwise learned metric for Person re-identification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 2197–2206.

[5] T. Matsukawa, T. Okabe, E. Suzuki, and Y. Sato, "People reidentification in surveillance and forensics;A survey," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 1363–1372

[6] B. Shi, X. Bai, W. Liu, and J. Wang, "Face alignment with deep regression," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 1,pp. 183–194, Jan. 2016.

[7] R. R. Varior, G. Wang, J. Lu, and T. Liu, "Comparative study of several feature extraction methods for person reidentification," *IEEE Trans. Image Process.*, vol. 25, no. 7, pp. 3395–3410, Jul. 2016.

[8] L.Xibi Zhao, G.Nan Wang, "Beyond pairwise matching person reidentification via higher order relevance learning,"*IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1089–1102, Jun. 2017.

[9] T. Xiao, H. Li, W. Ouyang, and X. Wang, "Learning deep feature representations with domain guided dropout for person re-identification,"in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016,pp. 1249–1258.

[10] Y. Yang, J. Yang, J. Yan, S. Liao, D. Yi, and S. Z. Li, "Deeply part aligned representations for person re-identification," in *Proc. Eur. Conf. Comput. Vis.*,2014, pp. 536–551.