PERSON RE-IDENTIFICATION USING TEXTURE DRIVEN DEEP LEARNING

[1]BONTHA KANCHANA, [2]B.SUNEETHA, [3]PAMIDI MALATHI

[1]152W1D3802 GVR&S College of Engineering and Technology
[2]Asst.Prof. Dept. Of ECE GVR&S College of Engineering and Technology
[3]132W1D3803 Dept.of ECE.., GVR&S College of Engineering and Technology
E-mail: [1]bonthakanchana.405@gmail.com., [2]suneetha.bobbillapati@gmail.com., [3]malathipamidi42@gmail.com

Abstract- Person re-identification largely involves input patterns or attributes. Recognizing a person across non-overlapping camera views, with different pose, illumination, and camera characteristics. To propose to tackle this problem by training a deep convolutional network to represent a person's appearance as a low-dimensional feature vector that is invariant to common appearance variations encountered in the re-identification problem. Specifically, a Siamese-network architecture is used to train a feature extraction network using pairs of similar and dissimilar images.

Keywords- person re-identification, deep learning, neural networks, feature embedding

I. INTRODUCTION

Person re-identification models the problem of tracking a person as they move through a non-overlapping camera network. Reliable person re-identification is vital for multiple camera tracking in realistic conditions, where there is little control over the image acquisition process, where camera coverage may be sparse, and where the subject may be noncooperative. The core of this problem is to decide whether pedestrian detections from several non-overlapping cameras, acquired at different times, were all caused by the same person. In the most general case, the individual cameras will have different hardware, will have non-overlapping fields of view, will capture the person from different angles, with different pose, and with differing illumination. Due to the large number of uncontrolled sources of variation, as well as generally poor image quality, this task remains very challenging.

II. IMAGE DATA HIDING

This paper differs from other re-identification methods based on deep networks [61], [37] by inclusion of multitask learning to improve re-identification performance and prevent over-fitting to the training dataset. In particular, the multi-task network will be trained to perform verification, identification, clothing attribute labelling, and pose labelling. This aims to encourage the learned feature representation to better generalize to unseen data such as other re-identification datasets. We show that by using multi-task learning, the reidentification performance of a simple convolutional network can match that of the more complex network, such as [61], given the same training and testing datasets.

There are three types of redundancy can be identified:
• *Spatial Redundancy* i.e. correlation between neighboring pixel values.
• *Spectral Redundancy* i.e. correlation between different color planes or spectral bands.

III. INVARIANT FEATURE EXTRACTION

Invariant feature extraction Ideally, the features used for re-identification should be invariant under common image transformations, while having a high degree of inter-person variation, and a low degree of intra-personal variation. Colour is commonly used as it exhibits a degree of pose invariance [53]. However, colour features, especially RGB or HSV, tend not to be illumination invariant [56] or invariant to different camera set-ups. Attempts have therefore been made to use physical illumination models, such as the Retinex model [30], to understand how colour features are affected by illumination to improve their invariance. Brightness transfer functions (BTF) can be used to transform colour features as a person moves between a pair of cameras

A. Deep Convolutional Neural Networks

Deep Convolutional Neural Networks There has been renewed interest in using neural networks for computer vision, sparked by the significant performance improvements over previously state of the art methods recently achieved using deep convolutional networks (CNNs) [28]. An application of neural networks that is particularly suited to person re-identification is that of learning embeddings, which involves mapping images into a low dimensional feature space, while preserving semantic relationships between the images.

IV. DISTANCE METRIC LEARNING FOR PERSON RE-IDENTIFICATION

The Euclidean distance between the feature representation of each image is used for training the network to perform veri- fication. For a training image pair (x1, x2), the cost function, V, is dependent

on whether the images are from the same, or different people. We will first introduce the cost functions for both cases, then we will show how these cost functions can be combined.

Therefore, in the same person case, the cost increases as the Euclidean distance between the feature representations increases, and when the feature representations are identical, the cost is zero.

V. PROPOSED METHOD

In practice, each layer of a convolutional network learns several small filters, which are convolved with the layer's input i.e., the previous layer's activation maps, to produce a new set of activation maps. Note that the filters in the first convolutional layer are connected to the colour channels of the input image. The activation maps are typically passed through a non-linear activation function, such as hyperbolic tangent, before further processing. Finally, a pooling operation, such as max-pooling.

Pre-process: The pixels in the range of and are processed excluding the first 16 pixels in the bottom row. A location map is generated and the image histogram is calculated with counting the first 3 LSB bits of pixels in the bottom row.

The extraction and recovery process: 1) On the recipient side, additional messages can be extracted if the receiver has the key KEMB. The marked encrypted image is separated to the Square set, the Triangle set and the Circle set again. With the embedding key, the recipient permutes pixels in each set independently, and divides the permuted sets into segments, each of which contains Li (i=1,2,3) pixels. Collect the bits of three LSB-layers in each segment and reconstruct the groups. The recovery operations are carried out by processing pixels except the excluded ones. The process of extraction and recovery is repeated until all of the split peak are restored and the data embedded with them are extracted.

VI. IMPLEMENTATION RESULTS

To improve re-identification performance, data augmentation was used during testing, as in [23]. To calculate the similarity-score between a given testing image and a given gallery image, features were extracted from 10 samples (the four corner crops, the centre crop, and their horizontal flips) of the testing image and gallery image. The similarity-score was then calculated as the mean Euclidean distance between the features of all one hundred test/gallery image pairs, where a smaller distance indicates two images are more similar

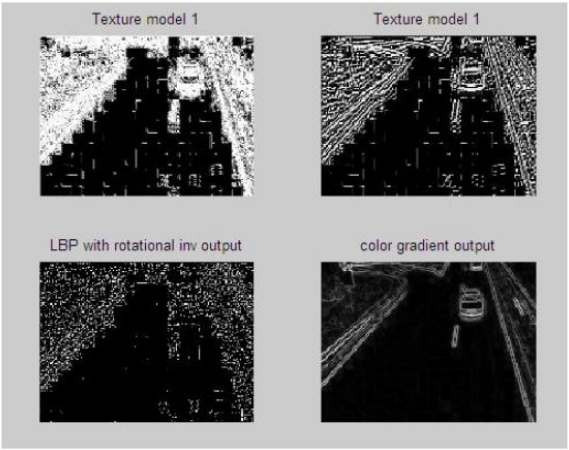
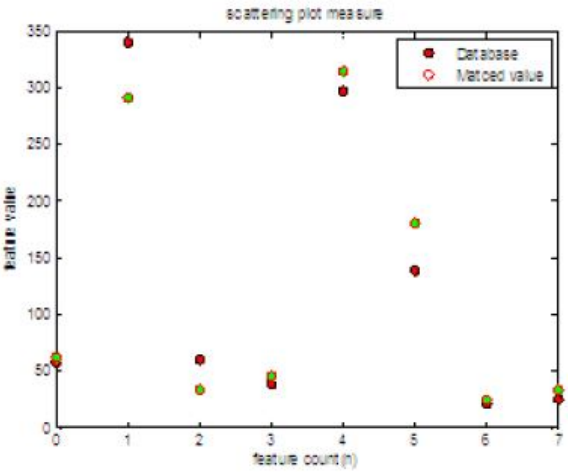


Fig 2 uniform and texture invariant model output



VII. DISTORTION ANALYSIS

The large boost in performance could be attributed to the fact that each person has a unique identity, meaning the identification task involves predicting from a diverse set of labels, forcing the network to learn the subtle differences between the appearances of individuals, rather than simply performing verification by comparing image pairs. A further improvement, achieving the best performance, occurs when the verification, identification, and attribute labelling tasks are used together. Use of attribute labelling together with verification produces a large improvement in performance compared with verification used alone.

CONCLUSION

Attribute and Pose Classification: While the primary aim of using the auxiliary attribute classification tasks is to help improve re-identification performance, it is interesting to observe the network's classification accuracy for these tasks, as they may have applications independent of person re-identification. In this work, we propose an approach for person re−identification by joint SIR and CIR learning. Since

SIR is efficient in matching, while CIR is effective in modeling the relationship between probe and gallery images, we fuse their losses together to utilize the advantages of both these representations.

REFERENCES

[1] E. Ahmed, M. Jones, and T. K. Marks. An improved deep learning architecture for person re-identification. In Proceedings of the IEEE Conference on CVPR, pages 3908–3916, 2015.

[2] L. An, M. Kafai, S. Yang, and B. Bhanu. Reference-based person reidentification. In Advanced Video and Signal Based Surveillance, pages 244–249, Aug 2013.

[3] S. Bak, G. Charpiat, E. Corvee, F. Bremond, and M. Thonnat. Learning to match appearances by correlations in a covariance metric space. In ECCV, pages 806–820. 2012.

[4] D. Baltieri, R. Vezzani, and R. Cucchiara. 3dpes: 3d people dataset for surveillance and forensics. In International ACM Workshop on Multimedia access to 3D Human Objects, pages 59–64, Scottsdale, Arizona, USA, Nov. 2011.

[5] A. Bar-Hillel, T. Hertz, N. Shental, and D. Weinshall. Learning a mahalanobis metric from equivalence constraints. JMLR, 6(6):937–965, 2005.

[6] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ra− manan. Object detection with discriminatively trained part− based models. IEEE Trans. Pattern Anal. Mach. Intell., 32(9):1627–1645, 2010.

[7] N. Gheissari, T. B. Sebastian, and R. Hartley. Person reiden− tification using spatiotemporal appearance. In CVPR, 2006.

[8] S. Gong, M. Cristani, S. Yan, and C. C. Loy, editors. Person Re-Identification. Springer, 2014. [9] D. Gray and H. Tao. Viewpoint invariant pedestrian recogni− tion with an ensemble of localized features. In ECCV, 2008.

[9] M. Guillaumin, J. Verbeek, and C. Schmid. Is that you? metric learning approaches for face identification. In ICCV, 2009.

[10] M. Hirzer, P. M. Roth, M. K¨ostinger, and H. Bischof. Re− laxed pairwise learned metric for person re−identification. In ECCV, 2012.

★★★