CSE 591: Person Re-Identification using Deep Learning(Final)

Jitesh Kamble
jkamble@asu.edu
ASU ID: 1211223244

Vishal Naik
vnaik1@asu.edu
ASU ID: 1211173623

***Abstract*—Area of visual computing has evolved rapidly due to advances in deep learning. Convolutional Neural Network has made it easy to deal with computer vision problems. Person Re-Identification aims to match an individual across images. Change in lighting conditions, orientation, background, and viewing angle make this process a challenging task. In this project we plan to design a model which accepts pair of images with same dimensions and outputs a binary value indicating whether the images correspond to the same individual. We use Convolutional layer and max-pooling kernel to extract the features from both the images and perform cross-input neighborhood differences. We plan to experiment with different filter parameters to evaluate the performance of the model. We also aim to investigate the model's performance on the dataset by varying the architecture.**

1. Introduction

Person Re-identifications deals with spotting individuals across images captured by different cameras from different vantage points, or images captured by a camera at different times. One of the major applications of Person Re-Identification is within the domain of Security, especially surveillance. Person Re-Identification is particularly challenging accounting to the variance in image's illumination and background. At times, these variances lead to images of different people looking quite alike, and images of the same individual looking unalike. Because the source of the images are surveillance cameras pointing at completely different directions, conventional tracking of individual's movement isn't always possible. At times, the variation between the two camera pictures is so much that the features present in one image are completely absent from the other image. Owing to the lack of high quality images, using face-detection or similar techniques on the images isn't always possible [2]. Figure 1 illustrates a few such challenging image pairs.

Figure 1. Image pairs in the first row (outlined in green) are matching, whereas, image pairs in the second row (outlined in red) are dissimilar.

Re-identification models typically take two same sized images, and output either a similarity score, or a binary value indicating whether the images correspond to the same individual. We aim to establish a model which outputs a binary value based on the similarity scores for a pair of images. We plan to build on top of "Improved Deep Learning Architecture for Person Re-Identification" [1]

by experimenting with different distance metrics like Manhattan, Euclidian and observe the behavior of the model. We are going to ensemble already existing neighborhood difference layer with Manhattan and Euclidian distance layer. This way we will be able to arrive at more accurate difference. We also plan to observe the performance of the model by changing the depth of the model. We will use ViPeR [8] and CUHK01 [7] datasets having about 600 and 900 identities each for training, validation, and testing. If we are able to secure computing resources, we intend to use medium sized CUHK02 dataset [7], and large sized CUHK03 dataset [7] for training, validation, and testing of our model.

2. Related Work

Existing research has used different feature extraction methods, like [3] which focused on establishing a probability based color distribution to generate a feature representation. Other research work has focused on establishing similarity metrics. Li et. al. established a distance metric by learning a decision function combined with locally adaptive thresholding rule [4]. F. Xiong et. al. used a ranking ensemble voting scheme, in conjunction with different sizes of sets of histogram-based features and linear kernels based metrics [5].

With respect to the use of deep learning in person re-identification, Yi et. al. designed a model with three independent convolutional neural network with max pooling [6]. Ahmed et. al designed two-layer convolutional network with maxpooling to extract the features, followed by a cross-input difference layer to compute the similarity indices between the image pairs [1]. The model predicted the binary label by passing the similarity indices through a convolutional and a fully-connected dot-product layer [1].

3. Timeline

TABLE 1. TENTATIVE PROJECT TIMELINE

Task	Date	Contributor
Data Analysis	02/24	Jitesh
Literature Survey	02/24	Vishal
Design	03/10	Vishal, Jitesh
Initial Implementation	03/24	Vishal, Jitesh
Experiment	04/14	Vishal, Jitesh
Update Model	04/20	Vishal, Jitesh
Benchmark	04/21	Vishal
Report	04/25	Jitesh

References

[1] E. Ahmed, M. Jones, and T. K. Marks, An improved deep learning architecture for person re-identification, in Proc. IEEE Conf. Comp. Vis. Patt. Recogn., 2015.

[2] Vezzani, R., Baltieri, D., Cucchiara, R.: People reidentification in surveillance and forensics: A survey. ACM Comput. Surv. 46(2), 29:129:37 (2013)

[3] Y. Yang, J. Yang, J. Yan, S. Liao, D. Yi, and S. Li. Salient color names for person re-identification. In ECCV, 2014.

[4] Z. Li, S. Chang, F. Liang, T. Huang, L. Cao, and J. Smith. Learning locally-adaptive decision functions for person verification. In CVPR, 2013.

[5] F. Xiong, M. Gou, O. Camps, and M. Sznaier. Person re-identification using kernel-based metric learning methods. In ECCV, 2014.

[6] D. Yi, Z. Lei, and S. Z. Li. Deep metric learning for practical person re-identification. ICPR, 2014.

[7] http://www.ee.cuhk.edu.hk/ xgwang/CUHK_identification.html

[8] https://vision.soe.ucsc.edu/node/178

[9] Z. Zhang, Y. Chen, and V. Saligrama. A novel visual word co-occurrence model for person re-identification. In ECCV Workshop on Visual Surveillance and Re-identification, 2014.

[10] C. Liu, S. Gong, C. C. Loy, and X. Lin. Person re-identification: What features are important? In A. Fusiello, V. Murino, and R. Cucchiara, editors, ECCV Workshops (1), volume 7583 of Lecture Notes in Computer Science, pages 391401. Springer, 2012.