ANALYSIS OF FACE RECOGNITION USING DOCFACE+ SELFIE MATCHING

DR. V.SUBEDHA, A.ABARNA, AMREEN FATHIMA M.R, R.ARTHI

COMPUTER SCIENCE DEPARTMENT, PANIMALAR INSTITUTE OF TECHNOLOGY CHENNAI

---***---

Abstract-Numerous activities in our day -to-day life require us to verify who we are by showing our ID documents containing face images, such as passports and driver licenses and other documents to human operators. However, this process is labor intensive, unreliable and very slow. As such, an automated system for matching ID document photos to face images live (selfies) in real time and with high accuracy is required. The gradient-based optimization methods converge slowly when many classes have few samples, a characteristic of existing ID-selfie datasets. To meet the requirement, the project proposes a method, called Dynamic Weight Imprinting (DWI), to update the classifier weights, which allows more generalizable representations and faster convergence. This project proposes DocFace+ to meet the objective. Next, a pair of sibling networks with partially shared parameters are trained to learn the unified face representation with domain-specific parameters. Cross-validation on an ID selfie dataset shows that publicly available general face matcher.

Index Terms:

I. INTRODUCTION

Identification is the process of classifying an input image into a large number of identity classes, while verification is the process of classifying a pair of images as belonging to the same identity or not(binary classification) [6]. The value of both pattern recognition research and practical applications, face recognition has attracted a large attention over many years, and so the performance of face recognition algorithms has increased significantly. Still the large visual variations of faces has high challenges for these tasks in real world applications [3]. Machine learning has shown great success in a variety of tasks with large amounts of labeled data in imageclassification, machine translation, and speech modelling, computer vision, natural language processing and speech recognition. However, deep networks seem less useful when the goal is to learn a new concept on the fly, from a few or even a single example as in one shot learning. A BlockChain is a list of records (blocks).It is linked using cryptography. It is a new and growing technique in recent technology. Each list of records contains a cryptographic hash of the previous block, a timestamp, and transaction data. A BlockChain is a decentralized, distributed, and oftentimes public and a A BlockChain database is managed autonomously using a peer-to-peer network and a distributed time stamping server. They are protected by mass collaboration sponsored by collective self-interests where security is marginal. Thus, BlockChain is a decentralized system, it is more secure, faster and cost effective ,has no downtime, user empower and can record historical and current records in one place [2]. We are proposing a certificate system based on blockchain to overcome the problem. Datas are stored in different nodes, and anyone who wants to modify a particular internal data must request the nodes and the other nodes will modify it simultaneously. The process is highly reliable. This is a decentralized application and designed a certificate system based on Ethereum Blockchain. This technique is incorruptible, encrypted, and trackable and permits data synchronization. By using the features of BlockChain methodology, the system improves the efficiency operations at each stage. The system saves on paper, cuts management costs, prevents document forgery, and provides accurate and reliable information on digital certificates and compare user live face with verified document face.

Fig:1 Selfie Matching

Fig 2: Comparison of both DocFace+ Models and Fine-Tuned

II. RELATED WORK

A. FACE VERIFICATION

Face verification is to decide whether two images containing faces represent the same person or two different people, and thus is important for access control or re-identification tasks. Face verification using deep learning techniques achieved a series of breakthroughs in recent years there are mainly two types of methods according to their loss functions. One type uses metric learning loss functions, such as triplet loss and contrastive loss[1] .The other type uses softmax loss and treats the problem as a classification task, but also constrains the intra-class variance to get better generalization for comparing face features. Some works also combine both kinds of loss functions. Deep Identification features are embedded into the conventional face verification pipeline off ace alignment, feature extraction, and face verification. The recently proposed SDM algorithm is used to detect 21 facial landmarks. Then the images of the faces are globally aligned by similarity transform according to the detected landmarks. Face verification involves cropping 400 face patches, which vary in horizontal flipping, color, scales and positions according to the globally aligned faces and the position of the facial landmarks [15]. Accordingly, 400 DeepID2 feature vectors are extracted by a total of 200 deep ConvNets, each of which is trained to extract two 160-dimensional DeepID2 feature vectors on one particular face patch and its horizontally flipped counterpart of each face. To reduce the redundancy among the large number of DeepID2 features and make our system practical, we use the forward-backward greedy algorithm to select a small number of effective and complementary DeepID2 feature vectors which saves most of the feature extraction time during test. Deep identification feature vectors are extracted and are concatenated to a 4000-dimensionalDeepID2feature vector. Joint Bayesian has been used successfully to model the joint probability of two faces being the same or different persons, The Joint Bayesian model for face verification based on the extracted DeepID2 features [1].

Fig 3: Matching Effect: Matching Live Image with Photo Proof

B. FEATURE REPRESENTATION

For feature representation, the face recognition utilized three kinds of features, namely a learning based descriptor LE, the hand-crafted feature LBP and a high-dimensional LBP feature, denoted by HighDimLBP [2]. These three kinds of features have been extracted from the LFW database. The LBP features were extracted by firstly dividing the face image into10× 10 non-overlapping sub-windows, then computing the 59-dimensional uniform LBP histogram on each of the sub-window, and finally concatenating all histograms. This kind of feature has 5,900 dimensions[5]. The LE descriptor is based on random projection tree learning to encode local image structures into discrete codes, which are considered to be uniformly distributed Patch histogram is further applied with the learned encodings, resulting in 20,735 feature dimensions. The HighDimLBP feature is based on recently developed accurate dense facial landmark detection and 27 facial landmarks are detected, and the face image is aligned according to the detected facial landmarks. Then, 5 scales of local patches are sampled around each facial landmark, and each patch is further divided into 4 4 cells. A 59 -dimensional uniform LBP histograms are exacted in each cell. Finally, all LBP histograms are concatenated, resulting in a 127,440-dimensional HighDimLBP feature [7]

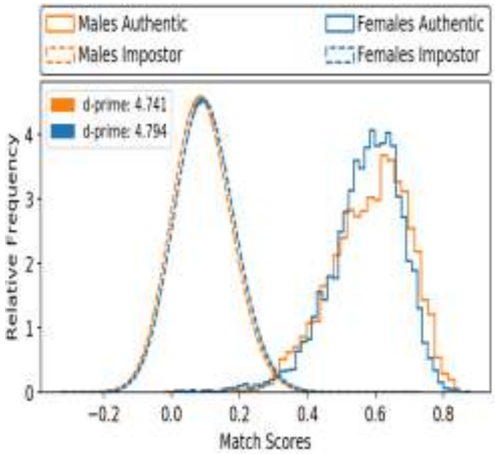

Fig 4: Poster Matching

C. AUTOMATIC FILTERING BY CLASSIfiCATION

The aim of this automatic filtering is to remove outlier faces for each identity automatically. This is achieved by learning a classifier to identify the different faces, and removing possible erroneous faces below a classification score[10]. To this end, 1-vs-rest classifiers are trained to discriminate between the 9244 subjects. Specifically, faces from the top 150 retrieved images of each identity are used as positives, and the top 100 of all other identities are used as negative for training. The face descriptor features are obtained from the VGGFace model. Then, the scores (between 0 and 1) from the trained model is used to sort images for each subject from least likely to most likely [16]. By manually checking through images from a random 500 subjects, we choose a threshold of 0.5 and remove any faces below that threshold value [8]

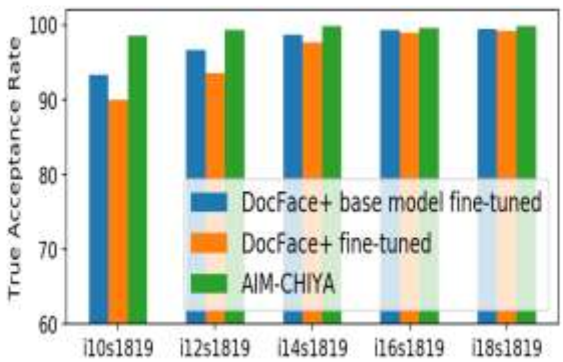

Fig 5: Acceptance Rate

D.FINAL AUTOMATIC AND MANUAL FILTERING

At this point, two types of error may still remain first, one classes contain a mixture of faces of more than one person, or they overlap with another class in the dataset; and second, some classes still have outliers (i.e. images that do not belong to the person) some classes contain a mixture of faces of more than one person, or they overlap with another class in the dataset [9]. This stage addresses these two types of errors with a mix of automated and manual algorithms.

III .METRIC LEARNING

Metric learning tries to learn semantic distance measures and embeddings such that similar samples are nearer and different samples are further apart from each other on a manifold. With the help of neural networks'enormous ability of representation learning, deep metric learning can do even better than the traditional methods. Recently, more complicated loss functions were proposed to get better local embedding structures [14].

Fig 6: Metric Learning

IV. FUTURE ENHANCEMENT

Facial recognition technology once seemed like something but it is increasingly being incorporated into our everyday lives and also in future development. Both private and public sector organizations are also incorporating facial recognition into services and product to create substantial benefits [4] for consumers. Many other new users are contemplated by the commercial sector, but they are still in development as the accuracy and usability of the technology progresses which is largely driven by increased demand by government agencies for their security systems, and the global market for facial recognition hardware and technology is forecasted. Further it will provide us the various advantages such as Transparency, Data Security, Privacy and Design and Integrity and Access. The process of Face Comparison is done faster and leads to better generalization performance The BlockChain technique provides more security to the documents and reduces forgery of the information. The process also avoids Overfitting [5].

V. EXPERIMENTAL SETTING

The experiments are conducted using Tensor flow library.10 when training the base model with original AM-Softmax on MS-Celeb-1M, Experiments use a batch size of 258 and keep training for 280K steps. The experiment start with a learning rate of 0.1 point, which is decreased to 0.01, 0.001 after 160K and 240K steps, respectively. When fine-tuning on the Private ID-selfie data, we use a batch size of 248 and train the sibling networks for 4, 000 steps. We start with a learning rate lower of 0.1 decrease the learning rate to 0.001 after 3, 200 steps. For both training stages, the feature networks are optimized by a Stochastic Gradient Descent (SGD) optimizer with a momentum of 0.9 and a weight decay of 0.0005. All the images are aligned via similarity transformation based on landmarks detected by MTCNN, and are resized to 96 ?112. We set margin parameters m as 5.0 in both stages. All the training and testing experiments are run on a single Nvidia Geforce GTX 1080Ti GPU with 11GB memory[9]. The inference speed of our model on this GPU is 3ms per image. By utilizing the MS-Celeb-1M dataset and the AM-Softmax loss function in Equation (1), our base model achieves 99.67% accuracy on the standard verification protocol of LFW and a Verification Rate (VR) of 99.60% at False Accept Rate (FAR) of 0.1% on the BLUFR protocol. Similar to the protocol of LFW, we define two views of the Private ID-Selfie dataset for the following experiments. In the development view, we tune the hyper-parameters including learning schedule, optimizer, m and α using 80% random identities for training and 20% for validation [12]. In the evaluation view, the dataset is equally split into 5 partitions for cross-validation. In each fold, one split is used for testing and other is used for training, In particular 873, 10, 718 and entites used for training and testing, respectively, in each fold. All the following analysis experiments are conducted on the evaluation view. Cosine similarity is used as score for all experiments [11].

VI. CONCLUSION

A new system, named DocFace+, for matching ID document photos to selfies is proposed. This technique of transfer learning is used where a base model for unconstrained face comparison is fine-tuned on a private ID-selfie dataset. A pair of sibling networks with highly shared-level modules are used as domain-specific parameters. Based on our observation of the weight-shift problem of classification-based embedding learning loss functions on shallow datasets, we propose an alternative optimization method, called dynamic weight imprinting (DWI) and a variant of AM-Softmax, DIAM-Softmax. Previous experiments prove that the proposed system not only helps the loss converge much faster but also leads to better generalization performance. A comparison with static weight imprinting methods confirms that DWI is capable of capturing the global distribution of embeddings accurately. Futher, the BlockChain methodology protects the document information and is the secured and fastest technique till now.

VII REFERENCES

[1] Y. Sun, Y. Chen, X. Wang, and X. Tang, "Deep learning face representation by joint identification-verification," in Proc. NIPS, 2014, pp. 1988–1996.

[2] S. Liao, Z. Lei, D. Yi, and S. Z. Li, "A benchmark study of large-scale unconstrained face recognition," in Proc. IJCB, 2014, pp. 1–8.

[3] L. Bertinetto, J. F. Henriques, J. Valmadre, P. H. S.Torr, and A. Vedaldi, "Learning feed-forward one-shot learners," in Proc. NIPS, 2016, pp. 523–531.

[4] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint face detection and alignment using multitask cascaded convolutional networks," IEEE Signal Process. Lett., vol. 23, no. 10, pp. 1499–1503, Oct. 2016.

[5] S. Ravi and H. Larochelle, "Optimization as a model for few-shot learning," in Proc. ICLR, 2017, pp. 1–11.

[6] Y. Movshovitz-Attias, A. Toshev, T. K. Leung, S. Ioffe, and S. Singh, "No fuss distance metric learning using proxies," 2017, pp. 360–368.

[7] F. Wang, X. Xiang, J. Cheng, and A. L. Yuille, "NormFace: L2 hypersphere embedding for face verification," in Proc. ACM MM, 2017

[8] Q. Cao, L. Shen, W. Xie, O. M. Parkhi, and A.Zisserman, "VGGFace2: A dataset for recognising faces across pose and age," in Proc. IEEE Conf. Autom. Face Gesture Recognit., 2018, pp. 1–8.

[9] H. Qi, M. Brown, and D. G. Lowe, "Low-shot learning with imprinted weights," in Proc. CVPR, 2018, pp. 5822–5830.

[10] H. Wang et al., "Cosface: Large margin cosine loss for deep face recognition," in Proc. CVPR, 2018, pp. 5265–5274.

[11] DocFace+ID Document to Selfie Matching Yichun Shi, and Anil K. Jain, Life Fellow, IEEE

[12] X. Wu, L. Song, R. He, and T. Tan, "Coupled deep learning for heterogeneous face recognition," in Proc. AAAI, 2018.

[13] G. Koch, R. Zemel, and R. Salakhutdinov, "Siamese neural networks for one-shot image recognition," in Proc. ICML Workshop Deep Learn., 2015, pp. 1–8.

[14] O. Vinyals et al., "Matching networks for one shot learning," in Proc. NIPS, 2016, pp. 3637–3645.

[15] J. Snell, K. Swersky, and R. S. Zemel, "Prototypical networks for fewshot learning," in Proc. NIPS, 2017, pp. 1–9.

[16] S. Chopra, R. Hadsell, and Y. Lecun, "Learning a similarity metric discriminatively, with application to face verification," in Proc. CVPR, 2005, pp. 539–546.