CS803 Project Proposal

Minchul Kim, Evan Masters, Redwan Sony, Roberto Muntaner Whitley

October 2021

1 Problem Definition

Face recognition is a field of AI which benefited from the advances in deep learning. The task is considered to be well-studied when the face images are taken in a controlled setting. The task becomes a more challenging task when the the face images are taken from extreme or wild settings such as suveillance or drone footage. In this project, we would like to explore the effect of state of the art augmentations in the training data such as GAN, adversarial examples, or 3D mesh projection. Introducing these augmented dataset into the training data could benefit the face recognition performance.

2 Objective

2.1 Face Recognition Pipeline

Deep learning is widely used for conducting face recogntion (FR). As opposed to conventional classification where training and testing classes are the same, FR requires learning discriminative features that can be generalized even when the identity in the test set was not present in the training dataset. For this method, metric learning methods have been used to increase the euclidean distance among identities and decrease the euclidean distance within the identities [1, 2]. Recently works on modifying softmax loss to accomodate margin in the decision boundary have shown to work well in FR [3, 4, 5]. Margin based softmax increases the inter-class distance while trying to minimize the intra-class distance in the feature space. During test time, the learned representations are compared using cosine similarity. We will test 1:1 recognition performance on well-known datasets such as LFW[6], CFP-FP[7], CPLFW[8] AgeDB[9] and CALFW[10].

2.2 Augmetations

2.2.1 GAN based

General Adversarial Nets (GANs) utilize deep learning architecture to produce synthetic data that has a probability distribution as close to a dataset of in-

terest as possible. Using a framework that estimates generative models for an adversial process, two seperate training models can be trained simultaneously. The first is a generative model that produces a data distribution and the second is a discriminitve model that estimates the probability that a sample came from the training dataset rather than the generative model [11]. This architecture has achieved tremendous success in recent years, producing high-fidelity virtual images that can be hard to distinguish from real ones. Within the field of FR, synthesizing face images of virtual people with varying expressions, pose and, lighting have become a topic of great interest. Given a collection of real face images, the generative model is trained to create realistic face images from random noise, consisting of multiple variables that all follow the normal distribtuion [12].Using this framework, we will create synthetic, unique images that can be used for face recognition. Generating a range of images, we will be able to asses the effectiveness of this data augmentation technique, and potentially increase the accuracy by introducing more challenging and diverse data.

2.2.2 Adversarial Example based

One of the important aim of the adversarial learningis to make a understandable high level feature representation so that it is possible to tune up certain features to make it robust against certain attacks. With that view in mind, an image will be represented in its high level latent feature space and then some of the features of that will be changed specifically to alter the image in such a way that it adds the robustness to the original image to prevent adversarial attacks on them.

2.2.3 3D mesh based

For each training image, we will use 3D mesh detection to create a mesh for the face. The facial image will then be projected onto the created mesh to create a 3D version of the face. This face will then be projected to a 2D image with a different view as the source image. These additional views will be then be combined with the original dataset to create a new dataset. The new dataset will be used in the Face Recognition Pipeline, and compared to the results. We will experiment with various combinations of views to study the impact on the results from the pipeline.

3 Group Member Work Distribution

Minchul Kim - Face Recognition Pipeline
Roberto Muntaner Whitley - GAN augmentation
Redwan Sony - Adversarial Example augmentation
Evan Masters - 3D mesh augmentation

References

[1] Florian Schroff, Dmitry Kalenichenko, and James Philbin. Facenet: A unified embedding for face recognition and clustering. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 815–823, 2015.

[2] Yair Movshovitz-Attias, Alexander Toshev, Thomas K Leung, Sergey Ioffe, and Saurabh Singh. No fuss distance metric learning using proxies. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 360–368, 2017.

[3] Weiyang Liu, Yandong Wen, Zhiding Yu, Ming Li, Bhiksha Raj, and Le Song. Sphereface: Deep hypersphere embedding for face recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 212–220, 2017.

[4] Hao Wang, Yitong Wang, Zheng Zhou, Xing Ji, Dihong Gong, Jingchao Zhou, Zhifeng Li, and Wei Liu. Cosface: Large margin cosine loss for deep face recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 5265–5274, 2018.

[5] Jiankang Deng, Jia Guo, Niannan Xue, and Stefanos Zafeiriou. Arcface: Additive angular margin loss for deep face recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4690–4699, 2019.

[6] Gary B Huang, Marwan Mattar, Tamara Berg, and Eric Learned-Miller. Labeled faces in the wild: A database forstudying face recognition in unconstrained environments. In *Workshop on faces in'Real-Life'Images: detection, alignment, and recognition*, 2008.

[7] Soumyadip Sengupta, Jun-Cheng Chen, Carlos Castillo, Vishal M Patel, Rama Chellappa, and David W Jacobs. Frontal to profile face verification in the wild. In *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 1–9. IEEE, 2016.

[8] Tianyue Zheng and Weihong Deng. Cross-pose lfw: A database for studying cross-pose face recognition in unconstrained environments. *Beijing University of Posts and Telecommunications, Tech. Rep*, 5:7, 2018.

[9] Stylianos Moschoglou, Athanasios Papaioannou, Christos Sagonas, Jiankang Deng, Irene Kotsia, and Stefanos Zafeiriou. Agedb: the first manually collected, in-the-wild age database. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 51–59, 2017.

[10] Tianyue Zheng, Weihong Deng, and Jiani Hu. Cross-age lfw: A database for studying cross-age face recognition in unconstrained environments. *arXiv preprint arXiv:1708.08197*, 2017.

[11] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial networks, 2014.

[12] Yu Deng, Jiaolong Yang, Dong Chen, Fang Wen, and Xin Tong. Disentangled and controllable face image generation via 3d imitative-contrastive learning, 2020.