One-Shot Learning for Face Recognition

Data Science Institute
COLUMBIA UNIVERSITY

Pranav Gopal, Vikram Singh, Aniket Shahane, Jannik Wiedenhaupt, Tianxiang Li
JPMC Mentors: Zhen Zeng, Kassiani Papasotiriou
Faculty: Sining Chen, Adam Kelleher

Data Science Capstone Project
with JPMorgan Chase & Co.

Problem Statement and Introduction

One unique challenge in the banking industry is face verification. Specifically, there is typically only one image of each customer's face in the database. It is critical for security to robustly verify a person's identity despite varying conditions like expressions, lighting, accessories, and hairstyles. Thus, we researched approaches and developed our own model architecture to tackle one-shot face recognition.

Methods

We used a FaceNet-based architecture to create distinct embeddings for each person which is shown in Figure 1. With the triplet loss function, the model learns to best distinguish the embeddings of different classes (here: people) without a classifier.

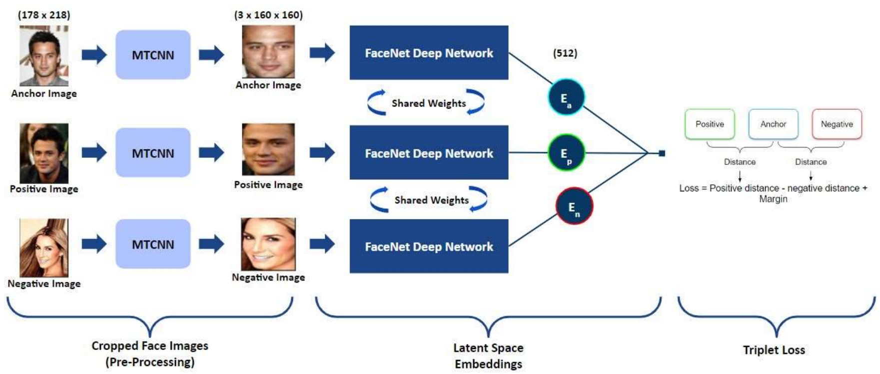

Figure 1. Face Recognition Model Training

We extracted weights from a pre-trained FaceNet model and fine-tuned on the CelebA dataset. However, we held out 1,000 people to test generalizability to unknown faces (see Figure 2). We then used the HFGI & IDInvert GANs to generate additional semantically augmented images for each person e.g., with a different pose (see Figure 3). Finally, we created embeddings for our one-shot vault set and tested our model's accuracy on the unseen test set.

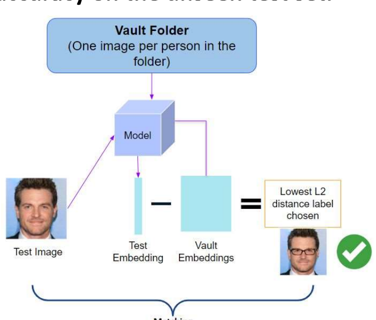

Figure 2. Matching Test

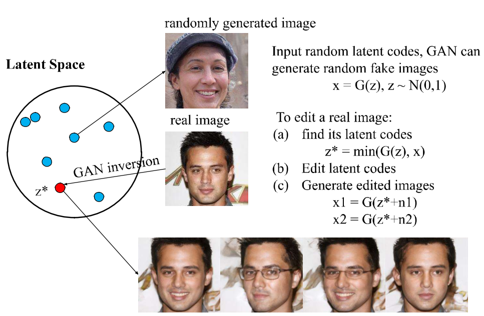

Figure 3. GAN Augmentation

Results

We conducted five experiments to measure the success of our model architecture. We tested FaceNet pre-trained on VGG-Face-2 (PT), FaceNet with images pre-processed by MTCNN (PT + MTCNN), Feature Extraction (FE) with and without MTCNN and finally FE with MTCNN and the GAN-augmented images. The results can be seen in Figures 4 & 5. While our image recognition pipeline is promising and we address the challenges of one-shot face recognition, our model falls short of PT + MTCNN. Due to the small image dataset and the insufficient generalization of the GANs, the model struggles to learn the underlying features.

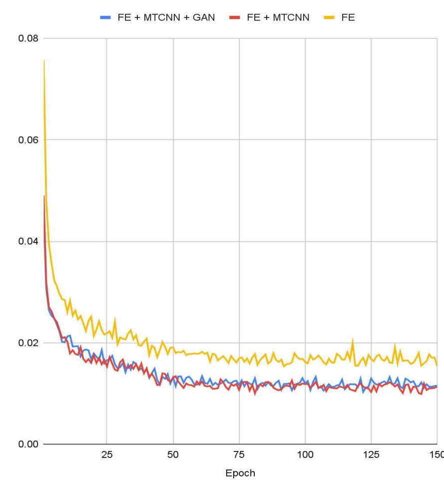

Figure 4. Training Loss - Feature Extraction Experiments

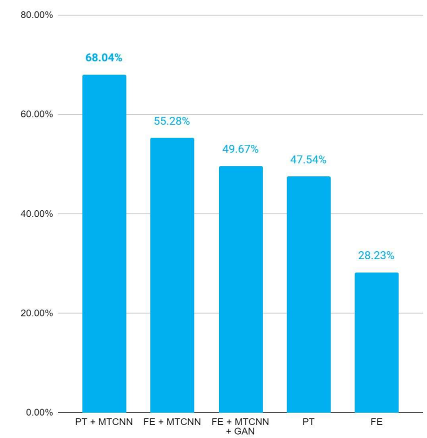

Figure 5. Test Accuracies

Conclusion:

One-shot learning is a quickly advancing research field trying to reduce the reliance on large datasets. However, as our experiments show, even with a proven pre-trained model, complex loss functions, and synthetic images, a model that is purely trained on a large dataset like VGG-Face-2 outperforms fine-tuning on smaller datasets. For future research, we suggest increasing training time and adding more complex semantic augmentation.

Acknowledgments

Special thanks to Zhen Zeng and Kassiani Papasotiriou for their support, guidance, and encouragement throughout this Capstone Project.

References

Schroff, Florian, Dmitry Kalenichenko, and James Philbin. "Facenet: A unified embedding for face recognition and clustering." Proceedings of the IEEE conference on computer vision and pattern recognition. 2015.