Supervised Deep Learning
For Real-Time Face Recognition

Amir Sotoodeh AASOTOODEH@CPP.EDU
California Polytechnic University: Pomona
3801 W Temple Ave
Pomona, CA 91768, USA

Abstract

Facial recognition is a prominent topic of discussion as an application to a variety of systems pertaining to security measures. In terms of machine learning, face recognition has always been an abstract and seemingly complex task with respect to computer vision given that facial landmarks must first be accurately examined and observed.

Although there have been many recent advancements in facial recognition, the greatest challenge to overcome has been the means of finding an efficient and quick ways to compute facial similarities or matches. A typical deep feed-forward neural network can be quite computationally expensive in its forward and backward propagation of data, requiring costly hardware to keep up with its capabilities. Hence, the task of real-time facial recognition for security purposes is a difficult task to complete with respects to limited hardware and efficiency.

This research paper will look into various techniques to allow for facial recognition via hardware such as webcams and/or security cameras in real time.

Keywords: Deep Learning, Supervised Learning, Convolutional Neural Network

1. Introduction/Background

There are many uses for facial recognition; namely, in video surveillance and aiding in search of specific criminals from available cameras in public. Additionally, it can also be applied as a safety/security measure for high authority officials to enter restricted premises. As such, it is an extremely active field in research by many data scientists and ML enthusiasts.

Given that there is a surplus of data from the proliferation of social media in recent years, powerful machine learning and statistical models could be built to an alarmingly high accuracy rate. This technology can even be seen in Apple's or Facebook's photo face recognition feature when searching for a name; it is commonly asked to tag a specific individual after detection of a face in a photo.

This goal of this project is to research the potential usage of machine learning models such as deep convolutional neural networks to solve the task of facial recognition (potentially) in real-time video feed. Additionally, research will be conducted to obtain a greater or more efficient performance of this task by using state of the art statistical models while training on large public/open source dataset.

2. Review of Literature

The several works discussed in this project will be based on the papers within the references section. According to [3], In 2014, DeepFace had acquired the state-of-the-art accuracy on the famous LFW benchmark, and had later been dramatically improved to 99.80% in the next three years. When referring to Google's FaceNet, it became apparent that larger datasets contribute largely to the accuracy of such models; additionally, no one other than Google had access to the 260 million images that were used to obtain its high accuracy. As such, this inspired varying architectures that required less data. In this project I will refer to and utilize several different deep learning architectures mentioned in [1][2][3], namely, variants of Inception (v1 and v2).

3. Objectives

Upon completion of this project, I hope to have successfully created a deep learning model ready for deployment on a system of my choice (via Docker). With this in mind, the model should ideally accurately recognize my face and any additional specified faces with greater than 90% accuracy. Additionally, the model should be relatively compact and produce an acceptable frame rate when run in real-time. I chose to work on this topic because I found the topic to have many use cases in the industry as well as personal use. Furthermore, this project will be great practice for gaining more experience in the growing field of Machine Learning and Artificial Intelligence.

4. Tentative Technical Approach

To begin, I will begin researching state of the art models with high accuracy with respect to its performance. I will prioritize models in which have faster performance and smaller size relative to its accuracy; this will allow for a model that can perform adequately in a real-time environment. Once a model or several models have been chosen, I will research any tweaks to the model's architecture as well as any impactful preprocessing steps that may contribute to the overall performance/efficiency of the models. Additionally, it is important to consider the dataset that will be used. Based on the papers referenced, it seems that the LFW (labeled faces in the wild) public dataset is ideal for this task given that they are rotated and cropped in the proper aspect. Lastly, I will attempt to tweak the model's hyper-parameters and/or apply regularization/ dropout to obtain a greater accuracy.

4. Timeline/Plan

Week 1-2: Research architecture of choice and highest performing techniques used.
Week 3-7: Data pre-processing & build DNN
Week 8-9: Model testing/tweaking
Week 9-11: Deployment and Integration
Week 12: Complete research paper and finalize project.

4. Deliverables

Python Code: High quality python code containing pre-processing functions, deep learning models via tensorflow, and web-server code.

Github Page: Github page for my peers and fellow computer scientists containing instructions for running the code as well as an in-depth explanation of the methods/architectures used.

[Tentative] *Web-server:* Deployment of code to my personal portfolio AWS instance

4. Resources

Given the complicated and expensive nature of Deep Neural Networks, the use of a GPU will be necessary to train the models that will be used. Additionally, software dependencies will be required:

CPU: Intel i7 7700k
GPU [High Performance]: MSI Geforce GTX 1060 6GB
Python 2.7+
Tensorflow
Numpy

Acknowledgements

Thank you to Dr. Hao Ji for the inspiration and opportunity to enter to exciting and amazing field of Machine Learning.

References

[1] Florian, S., Kalenichenko, D., Philbin, J., (March 2015) *FaceNet: A Unified Embedding for Face Recognition and Clustering.* Obtained From: https://arxiv.org/pdf/1503.03832.pdf

[2] Kulkarni, H., (May 2018) *Deep Learning for Facial Recognition*. Obtained From: https://www.researchgate.net/publication/325071878_Deep_Learning_for_Facial_Recognition

[3] Wang, M., Deng, W., (Sep 2018) *Deep Face Recognition: A Survey.* Obtained From: https://arxiv.org/pdf/1804.06655.pdf