Jacob Dineen, Jewell Anderson, Stephanie Chavers, Vincent Toney

IST 718 – Advanced Info Analytics

Group 1

Project Checkpoint 1 – Proposal

Computer vision is a field, sometimes said to be under the umbrella of artificial intelligence, that focuses on a machine’s ability to process and analyze images. It is becoming increasingly popular with the imminent arrival of autonomous vehicles, and the subsequent need to detect, recognize and learn from previous experience. Because computers lack characteristics that make them human, namely sight, in this instance, images must be transformed into data that is reflective of discrete representations of image properties, and that data is modeled, mainly using convolutional neural networks, to discern patterns between training data and testing data.

Our question as it stands; Can we utilize dlib’s facial recognition distribution to be able to detect all the faces present within a live classroom setting, and be able to classify those faces (matching them to the correct name)? This problem is not pertinent to any of our real-world work experience, at least as we currently know, but it will provide us with the framework to perform deep learning on image classification problems in the future, should the need arise.

To begin, we perform extensive research and begin structuring the design of our project being mindful of the encoded functions that are available to us (See citations below). Our data, which we will need to compile from snippets of fellow classmates, or scrape from their respective social media pages, will be static jpgs. The image detection segment of this project will utilize known patterns of human faces, pretrained in a CNN, which we can run against our training data. There are a number of issues that we have to account for when performing a CV task on human faces, such as image quality, transforming the original image from RGB to B&W and understanding of rotational impact on the ability to detect from different angles. Luckily, there is a an algorithm that accounts for this in the form of landmark estimation, which will be able to identify a face even if directionality is a concern. The next step deals with image encoding, machine learning and subsequent classification. Based on what the written program is receiving in real time, the frames are encoded and the pretrained neural net is used to generate measurements which, for lack of a better word, match a trained images’ encodings with that of the testing image.

Our hope is that, after compiling and transforming our training data we will be able to utilize/script a program that uses OpenCV, reads images in real-time from a webcam (at a slowed down frame rate), draws boxes around their faces and correctly classifies them based on our database (array) of already recognized people, hopefully at an accuracy of 90%

Some initial concerns regarding this project: Dlib’s packages appear to only work on macs/linux. Troubleshooting commenced to find workarounds. Quality of video from live classroom varies from students’ webcams. Inadequate quality of live feed may negatively impact our NN’s ability to classify encodings. If we are unable to get a prototype of this project running, we may revert to building an application that uses responses from a text pad (PYQT) to classify digits in real-time (Using the MNIST Digit Recognizer from Kaggle, and probably CNN for modeling).

Sources:

<https://medium.com/@ageitgey/machine-learning-is-fun-part-4-modern-face-recognition-with-deep-learning-c3cffc121d78>

<https://www.lynda.com/Python-tutorials/Application-template-matching/601786/660496-4.html>

[https://github.com/ageitgey/face\_recognition](https://github.com/ageitgey/face_recognition%20)