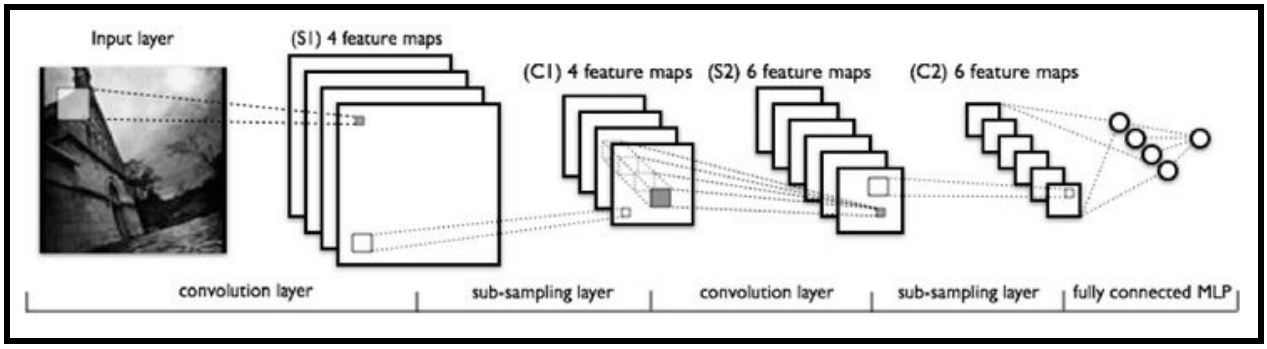
Facial Recognition Proposal

Group 1 : Jacob Dineen, Jewell Anderson, Stephanie Chavers, Vincent Toney

IST 718 - Advanced Information Analytics

The Foundations of Facial Recognition

Computer vision is a subfield of artificial intelligence, which 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 both previous and present experiences. Computers lack characteristics that make them human such as their ability to explicitly see, at least as we are known to understand it. Images must be transformed into data that reflects the discrete representation of image properties. Data is modeled using convolutional neural networks to discern patterns between training data and testing data. CNNs, which we'll get into during our research, are the ideal deep learning solution for image recognition due to their ability to make compute manageable. This is accomplished, in part, with their ability to bypass input mapping to each neuron, instead focusing only on a small subsection of the previous layer. In short, each neuron works as part of a whole, and together they produce vision.



https://www.kdnuggets.com

Roles

Researcher(s) : Jacob Dineen, Jewell Anderson, Stephanie Chavers, Vincent M Toney
Asset Gathering : Jewell Anderson, Vincent M Toney
Analyst : Jacob Dineen, Jewell Anderson
Modelers : Stephanie Chavers, Vincent M Toney

Obstacle

Can we utilize dlib's facial recognition library to detect all the faces present within a live classroom setting, or through a recorded live session, and successfully classify them based on supplied training data?

Goal

After compiling and transforming our training data we will be able to utilize a script that we developed to use the OpenCV library. This script will read images in realtime from a webcam (at

a slowed down frame rate), or take a stored MP4 file as an input and reframe the output with our wrapped functions aimed at identifying and classifying our fellow classmates. Upon completion, we will be able to feed in frames in real time, or upload a previously recorded file, and a box will be drawn around their faces and correctly classify them based on our database (array) of already recognized people, hopefully at an accuracy of 90%.

Approach

We will 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). The data will be produced from a compilation of pictures from our fellow classmates, and also procured through the scraping of respective social media pages for various static image types : jpgs, bmp and other. The image detection segment of this project will utilize known patterns of human faces, pre-trained in a Convolutional Neural Network (CNN) that we will run against our training data. There are a number of issues that we have to account for when performing a Computer Vision (CV) task on human faces. We have to review the image quality, transform the image into an array, while updating the original image from RGB(red, green blue) to B&W(black and white). Rotation of a framed face also plays a role in detection and classification so we have to understand how to detect images from different angles - There is an algorithm that accounts for this in the form of landmark estimation and it can identify a face even if directionality is a concern. The next step deals with image encoding. This includes machine learning and subsequent classification. Based on what the written program is receiving in real time, the frames are encoded and then the pre-trained neural net is used to generate measurements that match a trained image's encoding with that of the testing image.

Concerns

There are some initial development concerns regarding this project. Dlib's packages appear to only work on Macs/Linux, which we are troubleshooting to find workarounds such as virtual machines or operating system mirrors. Additionally, the quality of video from live classroom varies from student webcams. Also, inadequate quality of the live feed itself (due to network strain) may negatively impact our Convolutional Neural Network's (CNN) 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). In initial tests we have successfully built a detector and classifier with an uploaded mp4 file.

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/davisking/dlib-models

https://github.com/ageitgey/face_recognition

https://www.kdnuggets.com/2017/08/convolutional-neural-networks-image-recognition.html