SESUG Paper 237-2019

Face Recognition using SAS® Viya®: Guess who the person is!

Pratyush Dash, Oklahoma State University; Anvesh Reddy Minukuri, Comcast Corporation

ABSTRACT

Humans can take a look at an image and instantly recognize what the object is in the image, identify the person in the image or the location of the photo. The human cognitive system is fast and reliable, allowing people to perform very complex tasks like driving or operating a machine with little conscious thought. Performing these tasks for a computer would be very tough. Using fast, accurate and reliable algorithms could make computers to drive cars with sensors, enable them to recognize humans, operate different machines and even perform surgeries. The digital universe is expected to reach 44 zetabytes by 2020 because of the growth of Internet of Things (IoT). This shows us about the massive opportunity we have in terms of digital content analytics. Facial recognition and classification algorithms like deep learning and neural networks can extract information from photos or videos and classify them almost instantaneously after it is posted online. There are many other applications of such algorithms like in security screening, medical image processing, and insurance claims.

It is very challenging to perform this task as it requires extensive data preparation and lot of levels are needed to classify an image. In this project we have used the machine learning capabilities available in SAS® Viya® for image processing. In this paper we have used a dataset which consists of images of different people. The data was then divided into training and validation sets. For training 9 out of 10 images for each person were taken into consideration. The rest one image was used for validation. Several models will be designed using deep learning techniques like deep fully-connected neural networks (DNN), convolutional neural networks (CNN), and recurrent neural networks (RNN). The entire work has been done by using SAS DLPy which is by pulling in the Jupyter notebook into SAS® Viya®.

This work proposes to analyze celebrity images and classify them as different people using the deep learning techniques and measuring accuracy. Through this project the objective of automatically detecting who the celebrity is achieved and it can be further used to segregate them into different folders.

INTRODUCTION

What is Face Recognition?

Face Recognition is one of the many wonders that AI research has brought forward to the world. It is a kind of computer vision technology that detects and identifies people's faces within digital images. It involves identifying the face in the image as belonging to a particular person X and not another person Y. This is very easy for a human brain to process and interpret, but computers would need instructions. The images might contain several other objects in the background like a car, house, animals and so on.

How do computers see images?

The smallest element of an image is called a pixel. It is just like a dot in the image. An image constitutes of multiple pixels arranged in rows and columns. A computer does not understand these dots. It only understands numbers. So we need to convert all images into numbers so that the computer can understand it. Every pixel is represented by three numbers corresponding to the amount of red, green and blue present in the pixel. For grayscale images, each pixel is a single number representing the intensity of the light. If each pixel is a number, then an image for a computer to understand can be represented as a matrix of numbers.

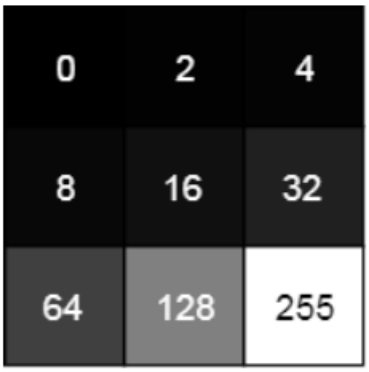

Figure 1. An example 3*3 image with pixel values and colors

What are Features?

Features are information of an image that are relevant to identifying a face. They can be combined into complex features to detect edges, corners, eyes of the face. Using deep learning techniques in SAS Viya these features can be used to create, train and tune deep fully-connected neural networks (DNN), convolution neural networks (CNN), and recurrent neural networks (RNN) for the applications of face detection and face recognition.

This paper provides a high-level overview of neural networks and walks through a facial recognition model using SAS Viya and Python.

NEURAL NETWORKS

Neural networks are models that were initially designed to mimic how neurons in the brain function. They consist of components including an input layer, hidden layer(s) and an output layer. Each layer is comprised of nodes. Figure 2 shows a diagram of the core components of a neural network with a single hidden layer.

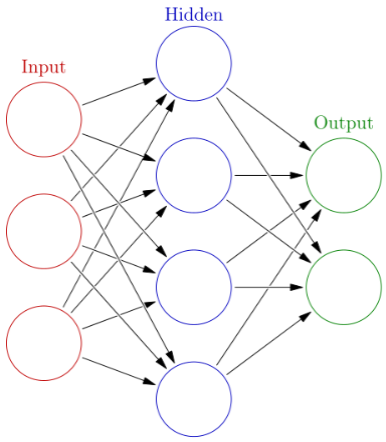

Figure 2. Neural Network Diagram

Neural networks are commonly named after their number of layers. The number of layers in a neural network can impact the accuracy of the model. A deep neural network is one that has many hidden layers between the data input and output layers. Hence, deep learning is a term used to describe larger and more complicated networks that are used for extracting valuable information from unstructured data such as video or images.

As neural networks deepen, and hidden layers are added, more complex features can be detected within the data, further improving the output predictions of the model. One advantage of neural networks compared to other predictive models is that they can model complex, non-linear relationships for better classification. In the context of facial recognition, shallower sections of a network may be able to detect lines whereas deeper layers can detect more complex features such as parts of faces.

MODEL BUILDING

For this paper a Neural network model using SAS VIYA was developed for facial recognition. We have also used YOLO (you only look once) algorithm to compare results between two algorithms. There are ten different images of each of 40 different subjects. All the images were taken against a dark homogenous background with the subjects in an upright, frontal position. The images were taken at different times, varying the lighting and facial expressions. Python was used to interface with CAS through the SAS SWAT package.

Generally, data modeling tasks require a lot of data cleaning and transformations to bring the input variables into proper formatting for modeling. But for face recognition images need to be labelled and resized properly for modeling. With use of machine learning algorithms, we can then go forward to detect and learn patterns from the data. The following sections would cover details of data preprocessing and model building for our face recognition model. For both the models some of the steps were same and just the model building algorithm was different. The data preparation steps were same for both the models.

DATA PREPARATION

Figure 3. Data Preparation steps

The image files were read in PGM format for modeling in SAS VIYA and they were taken and converted to jpg for YOLO. Then they were resized into 100*100 pixels. Each image is divided into 25 blocks and each block is 16*16 pixels in size. Each block was then converted into 1*256 vector and all the blocks were then concatenated in an image into a 1*6400 vector. The resized images were then uploaded into the CAS session using SAS SWAT package. The data was partitioned in such that 9 out of 10 images of a person was used for training the model and the remaining image was used for validating the model. Below is an image which shows how the images look after data preparation step is complete.

100*100 Pixels

Figure 4. Image resized to 100*100

RESULT COMPARISON OF MODELS

Neural network model was individually trained on all the blocks for all the images in order to predict the person ID for both the models. The model used the concept of majority voting from all of the blocks for a given person ID in order to make final decisions to predict the person ID. In case of YOLO we got the probability prediction scores. YOLO uses a single stage detector and is supposed to be a faster algorithm. It uses regression to identify and assess the probability score of images.

The model performs training of the images based on the 9 images of each person. IT also performs scoring based on the remaining images and the predicted person ID is printed alongside the features list in the output data set. For both models the same procedure was used to train and validate the images

To next check how efficient our model was in predicting the person ID's from the features we calculated the misclassification rate. And to check the efficiency of YOLO we passed a test image to check if the model is able to correctly identify the person.

Figure 5. YOLO test image results

Description	Value
Number of observations read	360
Number of observations used	360
Misclassification rate(%)	10%

Table 1. Misclassification rate from Neural network model on SAS VIYA

From the results we see that SAS VIYA model was able to correctly identify for about 90% data. It means that our model was accurate during 90% of the time. And it could not correctly identify for the rest 10% of the time. This shows us that our model was predicting quite well and is performing well. But the high accuracy could be because of the data we used for building our model. All the images had similar backgrounds, the same resolution and had just one person in them. This helped to train the model better for telling people apart. The challenges that we could face when there are multiple people in an image could be in able to distinctly detect all the faces present in the image, cropping each person's face, adjusting for each individual facial alignment and accounting for different lighting condition. In such a case it could be possible to develop many deep learning models which individually could target one specific problem and provide input to the next model.

For the results we got from the YOLO Model, we see that it is incorrectly assigning a high probability score to the test image and is not performing that well. The accuracy of YOLO is a problem, but the advantage it has is its high speed processing. If more number of training images had been used to train the YOLO model for a specific person it could have led to better results. But in this case since the number of training images we used for a specific person is less the accuracy for this algorithm is not so good. But the accuracy that we got from the neural network model on SAS VIYA is very good. This helps us to compare results from both algorithms and help us conclude.

CONCLUSION

Even though YOLO was believed to be a better and faster model, but due to less number of images for training the dataset, we find that the neural network model built on SAS Viya turned out to be our champion model. Using these results, we can build facial recognition models on SAS Viya which can have practical use in virtually every kind of industry. They can be used to improve some workflows and be able to solve new business problems in the industry. It can help in the insurance industry by helping to curb false insurance claims, it could also be used in the medical industry for imaging. This technology can improve a lot of processes at airports and security in appliances like phones or laptops.

Using this technology, we could also build as part of future work an API program which would help segregate different people and help us cluster them into different folders and store images of each person in their respective folders. Developing models that can achieve these duties needs state-of-the-art analytics platforms that can process large quantities of information in reasonable time. SAS deep learning enables us to use Cloud analytics services for in-memory processing. Furthermore, we can integrate with open source pre-trained models which could lead to rapid creation, training, and customization of deep learning models.

REFERENCES

Huetter, J. (2016, October 11). “Tractable photo-analyzing AI to assess vehicle damage for Mitchell.” Accessed February 24, 2018. Retrieved from http://www.repairerdrivennews.com/2016/10/11/tractableceo-says-ai-will-assess-vehicle-damage-for-mitchell-company-specializes-in-ai-photo-analysis/

Perret, E. (2017, January 29). “Here’s How Many Digital Photos Will Be Taken in 2017.” Accessed February 24, 2018. Retrieved from https://mylio.com/true-stories/tech-today/how-many-digital-photos-willbe-taken-2017-repost

Leign Ann Herhold and Ivan Gomez, Zencos Consulting LLC, “Image processing: Seeing the world through the eyes of SAS Viya” Retrieved from

https://www.sas.com/content/dam/SAS/support/en/sas-global-forum-proceedings/2018/2759-2018.pdf

Swastik Somani. “Face recognition using Tensorflow Pre-Trained model & OpenCV” Retrieved from

https://medium.com/@somaniswastik/face-recognition-using-tensorflow-pre-trained-model-opencv-91184efa4aaf

Adrian Rosebrok. “Face recognition with OpenCV, Python and deep learning” Retrieved from

https://www.pyimagesearch.com/2018/06/18/face-recognition-with-opencv-python-and-deep-learning/

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Pratyush Dash
Oklahoma State University
pratyush.dash@okstate.edu

Anvesh Reddy Minukuri
Comcast Corporation
anvcse2007@gmail.com