

Course Final Project

CSE 402

Group 9

Joel Nataren, Kushi Vora, Kofi Debrah, Ajuisiwon Azantilow, Sydney

Gruczelak, Jared Singh Sekhon

Introduction

The final course project involves using a pretrained model with a dataset of 17 celebrities with 100 face pictures each. The aim is to use a model to classify a series of test images in order to study the accuracy of the alexnet model from the torchvision module, along with other metrics such as precision, recall and f-score to identify how robust and useful the model is.

Each team member is responsible for a specific part of the project. Joel Nataren was tasked with separating the data into readable sets, Jared was tasked with preprocessing and reading in the data, separating it into the training and testing data to be passed into the model. Azantilow and Kofi were tasked with constructing/using the models with the datasets. Finally, Sydney and Kushi were tasked with analysis of the output from the model and translating that into the previously stated metrics, as well as constructing a confusion matrix.

Data Separation

Before going fully into the preprocessing, images for each celebrity were separated into 10 images per folder to be able to use the LOOCV method for training and testing the model. There were a total of 17 celebrities in each of the classes. Ten folders were created, each containing 10 images of each of the 17 classes. A json/ dictionary was used to store the name of the image to match with the name of the class it belonged to.

Data Preprocessing and LOOCV

In the preliminary parts of the code, images were compiled into 10 folders of 17 celebrities each, with 100 face images per person. To ease image processing and to keep labels, a `preprocessedImage` class was created with image, label, filename and tensor attributes. A

preprocessedImage object was then made for each image, and a method preprocess() was called to read in and convert images into the tensor attribute to be passed into the model later on. Each folder of preprocessedImages was then added to a master list of folders.

To carry out LOOCV (Leave One Out Cross Validation) the code loops through the master folder list 10 times, each time appending the folder at the current iteration number (i) in the master list to a master test folder list, adding every other folder for that iteration to a temporary training list, finally adding that training list to a master training list at the end of the iteration. This results in a master test list of 10 folders, and a master train list of 10 lists of 9 folders. With this, LOOCV is achieved by running the model 10 times with the train and test sets changing each time.

Face recognition model

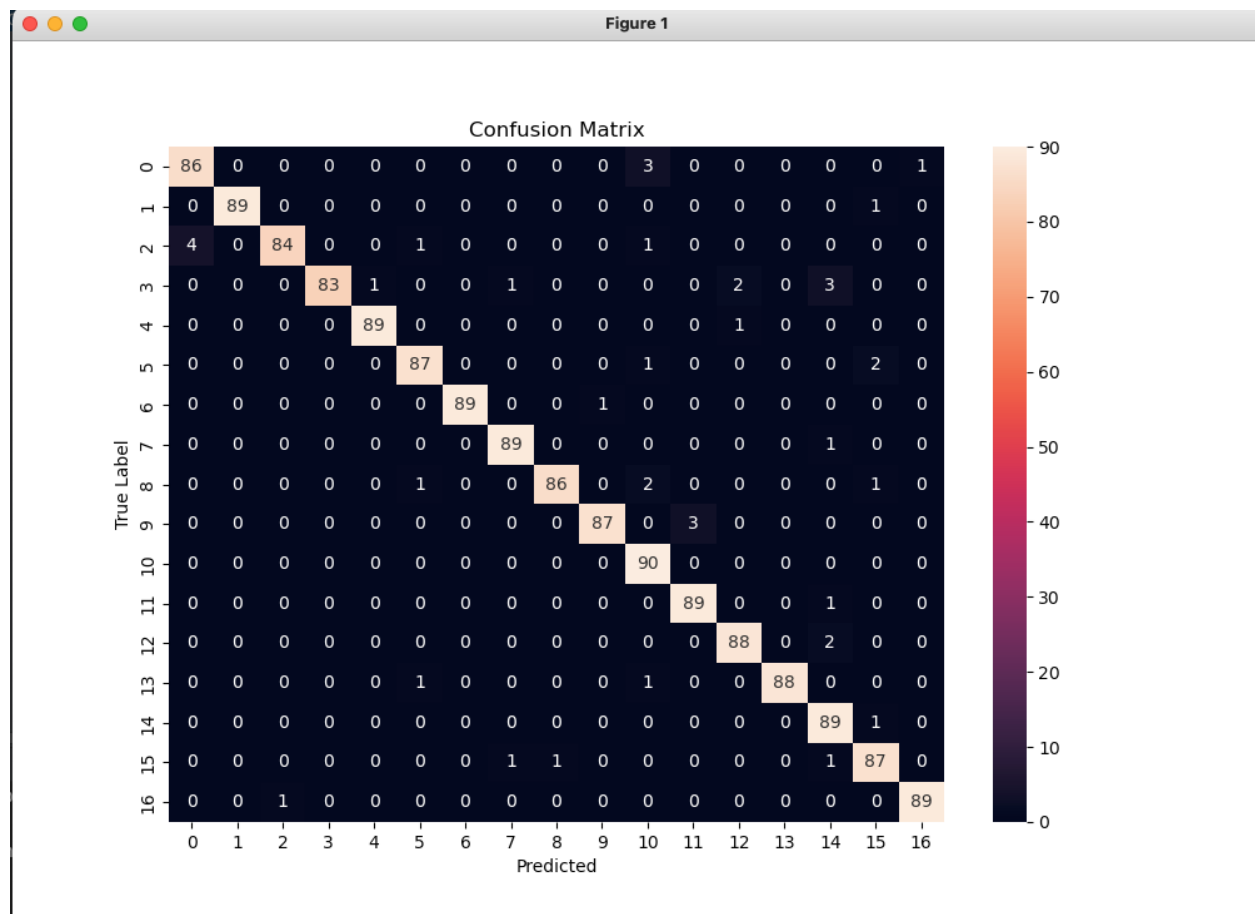
The AlexNet model was used to train the face images and used to make the predictions on the testing data. The weights of the model was initialized with the default AlexNet model at the beginning of the training. Cross-entropy loss function was used as the loss function to fine tune the weights of the model to ensure a better accuracy in prediction . An initial learning rate of 0.001 was used to train the model. To ensure that the types were consistent with the preset AlexNet types, the image labels which were initially strings (names of celebrities) were converted to integers which were subsequently converted into tensors. Also, to ensure that the model ran faster, the model was moved to the GPU if the GPU was available.

To improve the model's initial performance of 5%, the model's initial learning rate of 0.001 was updated to 0.0001, a reduction to ensure that the model is adjusting the weights slowly and not changing it greatly for every sample. The accuracy of the model increased to approximately 99% after the learning rate of the model was updated.

Accuracy

As mentioned before, the final accuracy obtained using the LOOCV scheme was from 95-99%.

Confusion matrix:



Given the results of our confusion matrix, we can draw conclusions about which faces may get confused with others. The face corresponding to label 0 is Sandra Bullock's. Her face can be confused with those corresponding to labels 2 and 10, which are Hugh Jackman and Megan Fox. Denzel Washington's face can be confused with Leonardo Dicaprio's face; their faces are labeled as 9 and 11. Jennifer Lawrence's face has the label 3, which can be confused

with Tom Cruise's face with label 14. Overall, the matrix shows that the model does not result in a significant amount of confusion and its outcomes are fairly accurate.