MULTI-SCRIPT TEXT IDENTIFICATION

CSN 383 MACHINE LEARNING – COURSE PROJECT

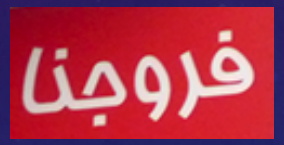
Under The Supervision Of Dr. Partha Partim Roy

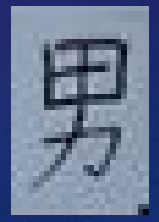
OBJECTIVE

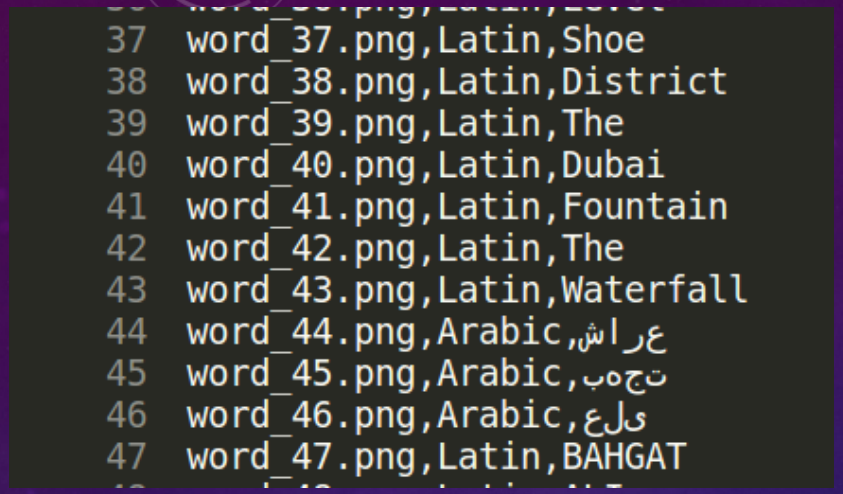
Train a Model that can detect and predict which language a word belongs to when an image containing the word is given as input

DATASET DESCRIPTION

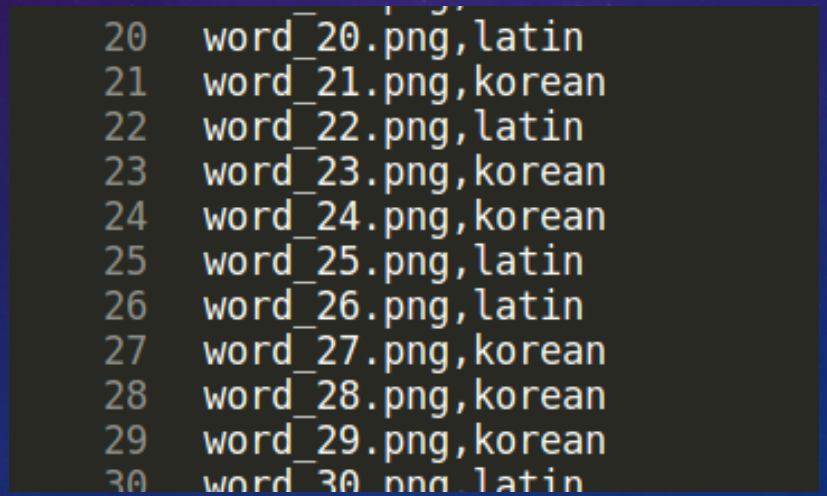
- Dataset from ICDAR2017 Competition on Multi-Lingual Scene Text Detection and Script Identification Task 2 is used for this project
- Training Set of about 63,000 images and validation set of about 16,000 images
- Two text files each describing the ground truth of each image in training set and validation set
- There are total 7 class labels that is 6 different languages that a word from the dataset may belong to and a symbol class to which mathematical symbols belong







Snippet of Ground Truth – Training Dataset



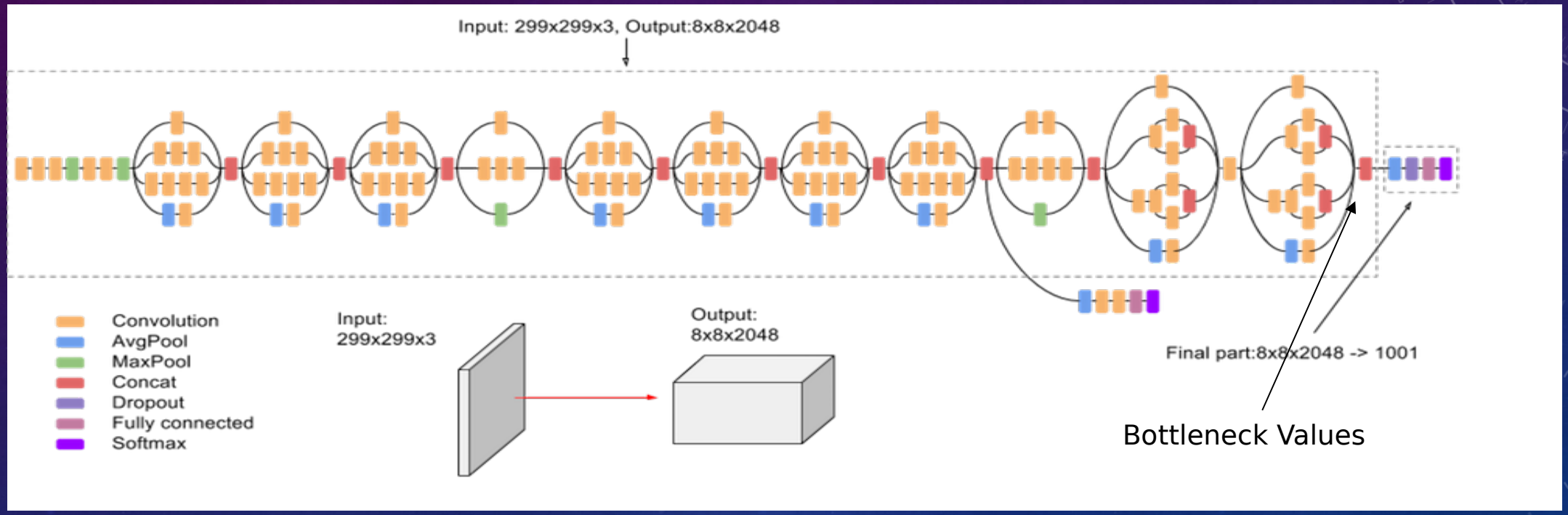
Snippet of Test Dataset Result

PROCEDURE

- Sort all the images into seven folders based on their languages
- Input the directory containing these seven folders to the retrain python script
- Image classification model will be trained on these seven labels and weights are saved in a .pb file
- The first phase analyzes all the images on disk and calculates and caches the bottleneck values for each of the images
- Once the bottlenecks are complete, the actual training of the top layer of the network begins
- There are about 4000 training steps and 10 images are chosen in random from the training dataset and the weights of the final layer are tuned based on the these predictions

ARCHITECTURE

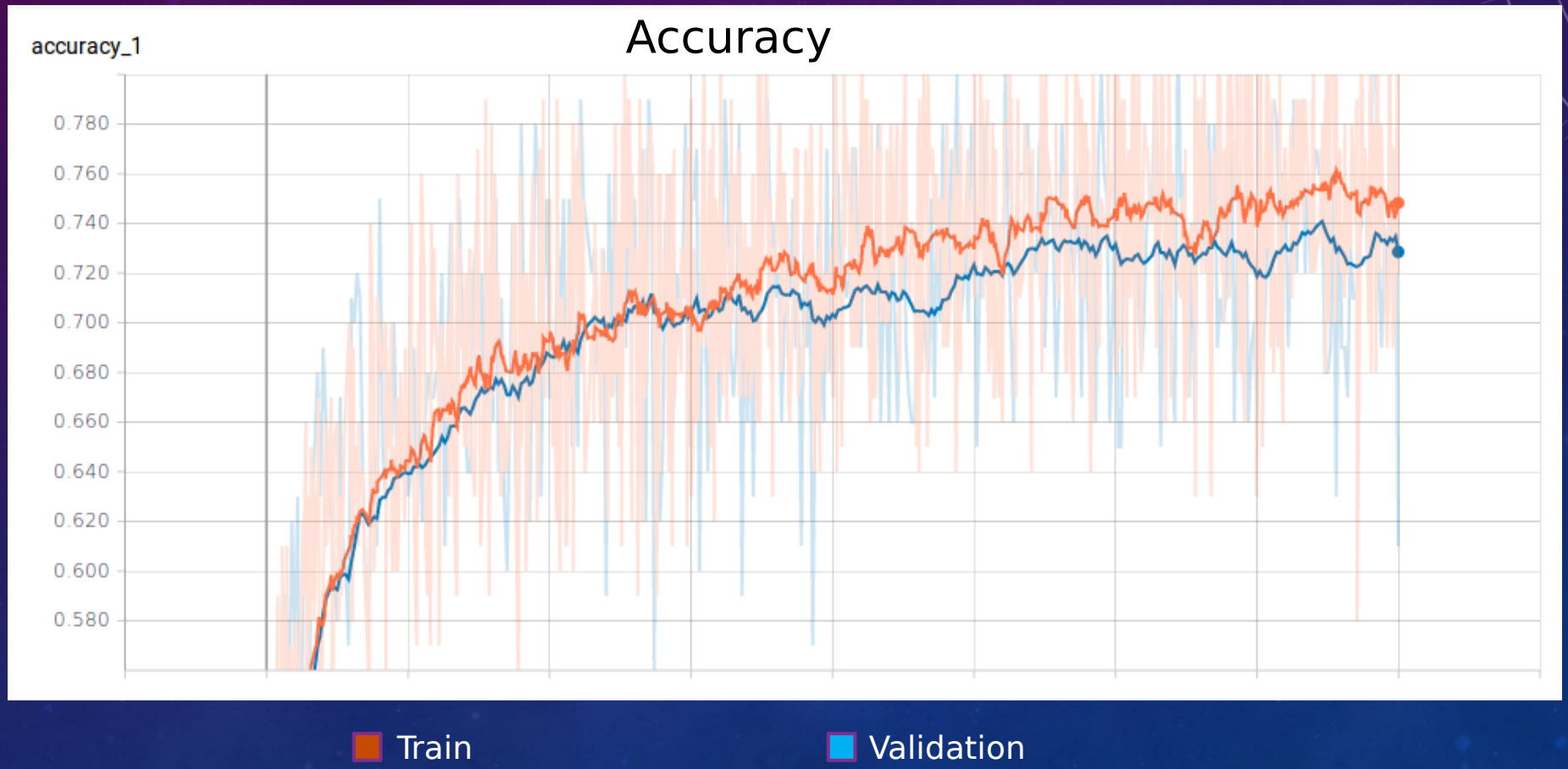
For this task we used Inception V3 architecture trained on ImageNet



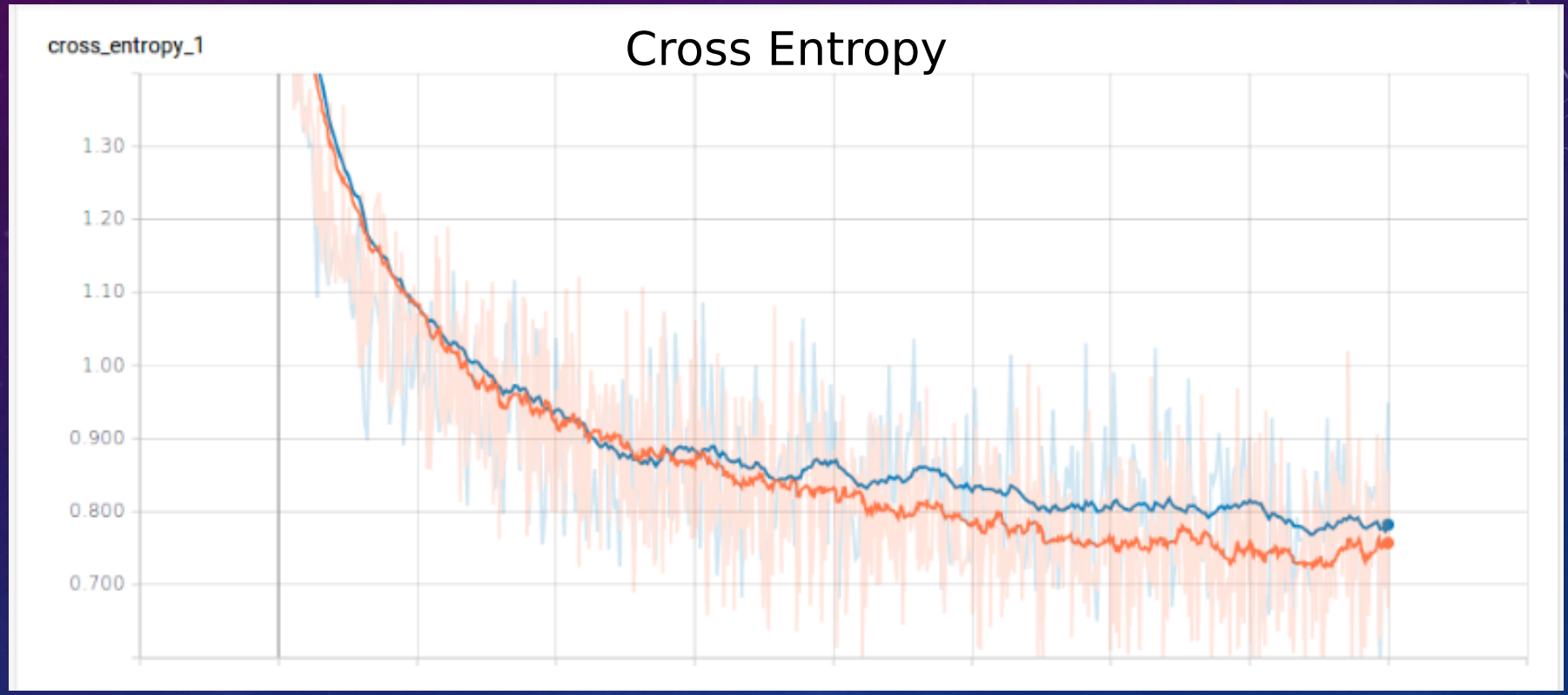
RESULTS

- The retrained model is tested on 16,000 images and about 11,000 of the images were correctly predicted
- The accuracy of the model is about 70%
- Accuracy can be improved by training the weights of previous layers using these images

RESULTS CONTINUED



RESULTS CONTINUED



THANK YOU !