

# 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



فوجنا



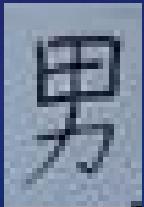
Grilled



ফার্মিচার



学校区域



男



뉴턴하고



283

```
36 word_36.png,Latin,Shoe
37 word_37.png,Latin,Shoe
38 word_38.png,Latin,District
39 word_39.png,Latin,The
40 word_40.png,Latin,Dubai
41 word_41.png,Latin,Fountain
42 word_42.png,Latin,The
43 word_43.png,Latin,Waterfall
44 word_44.png,Arabic,عراش
45 word_45.png,Arabic,تجهيزات
46 word_46.png,Arabic,قطعات
47 word_47.png,Latin,BAHGAT
```

## Snippet of Ground Truth - Training Dataset

```
20 word_20.png,latin
21 word_21.png,korean
22 word_22.png,latin
23 word_23.png,korean
24 word_24.png,korean
25 word_25.png,latin
26 word_26.png,latin
27 word_27.png,korean
28 word_28.png,korean
29 word_29.png,korean
30 word_30.png,latin
```

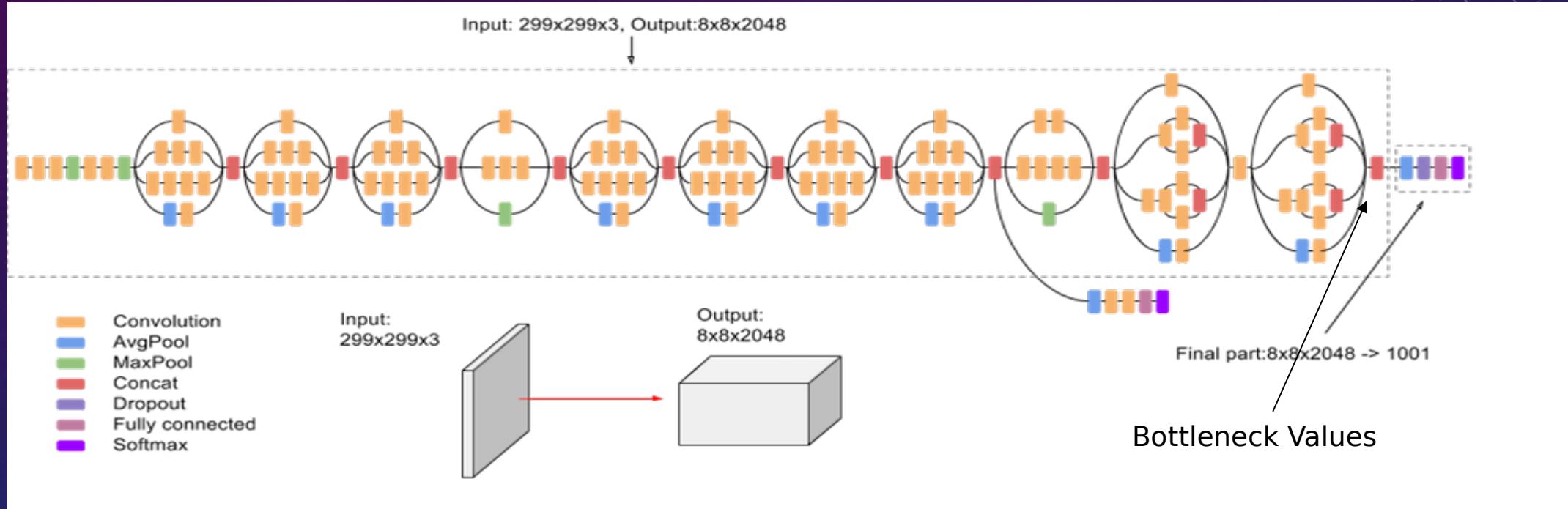
## Snippet of Test Dataset Result

# 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

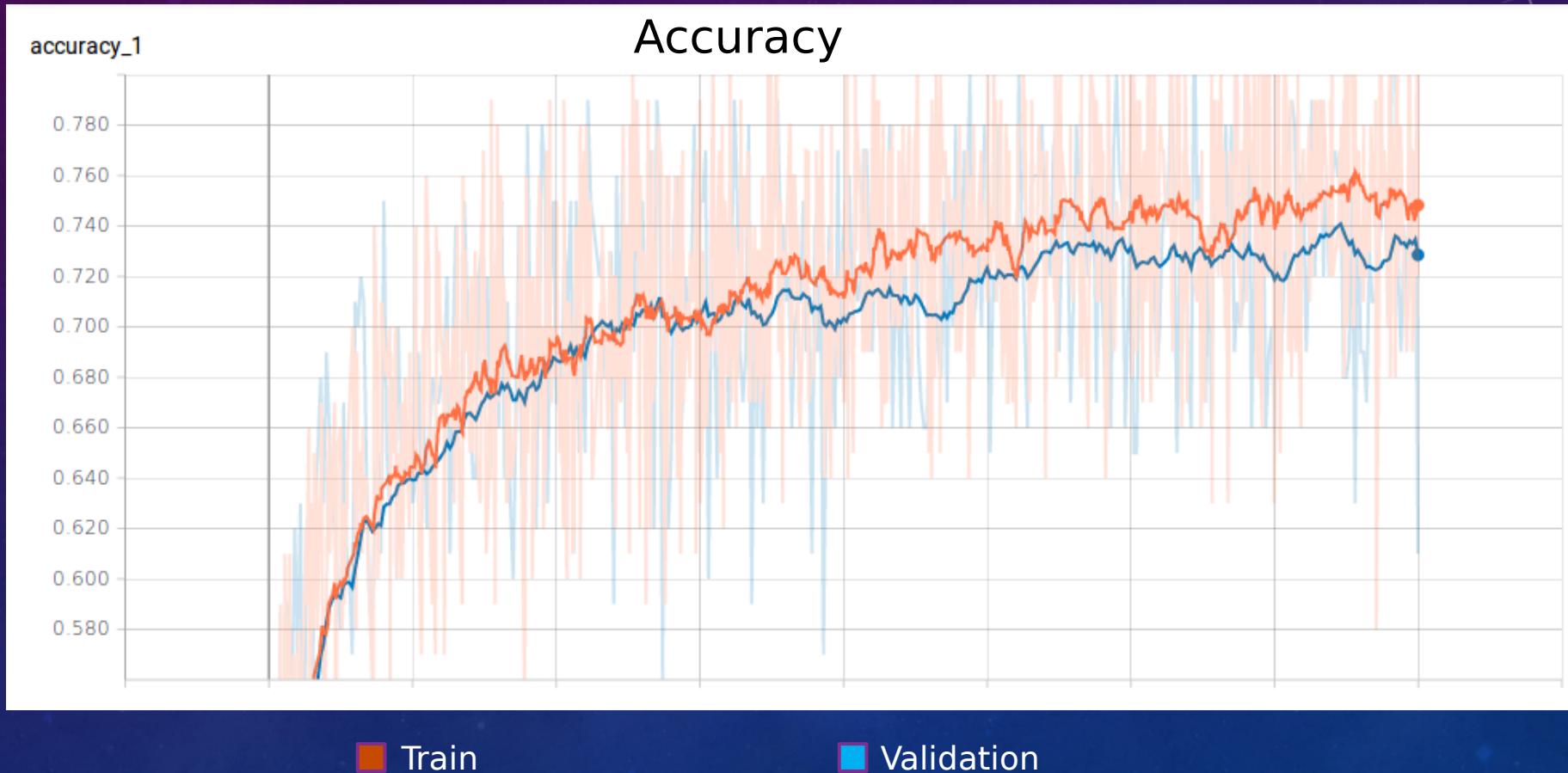
For this task we used Inception V3 architecture trained on ImageNet



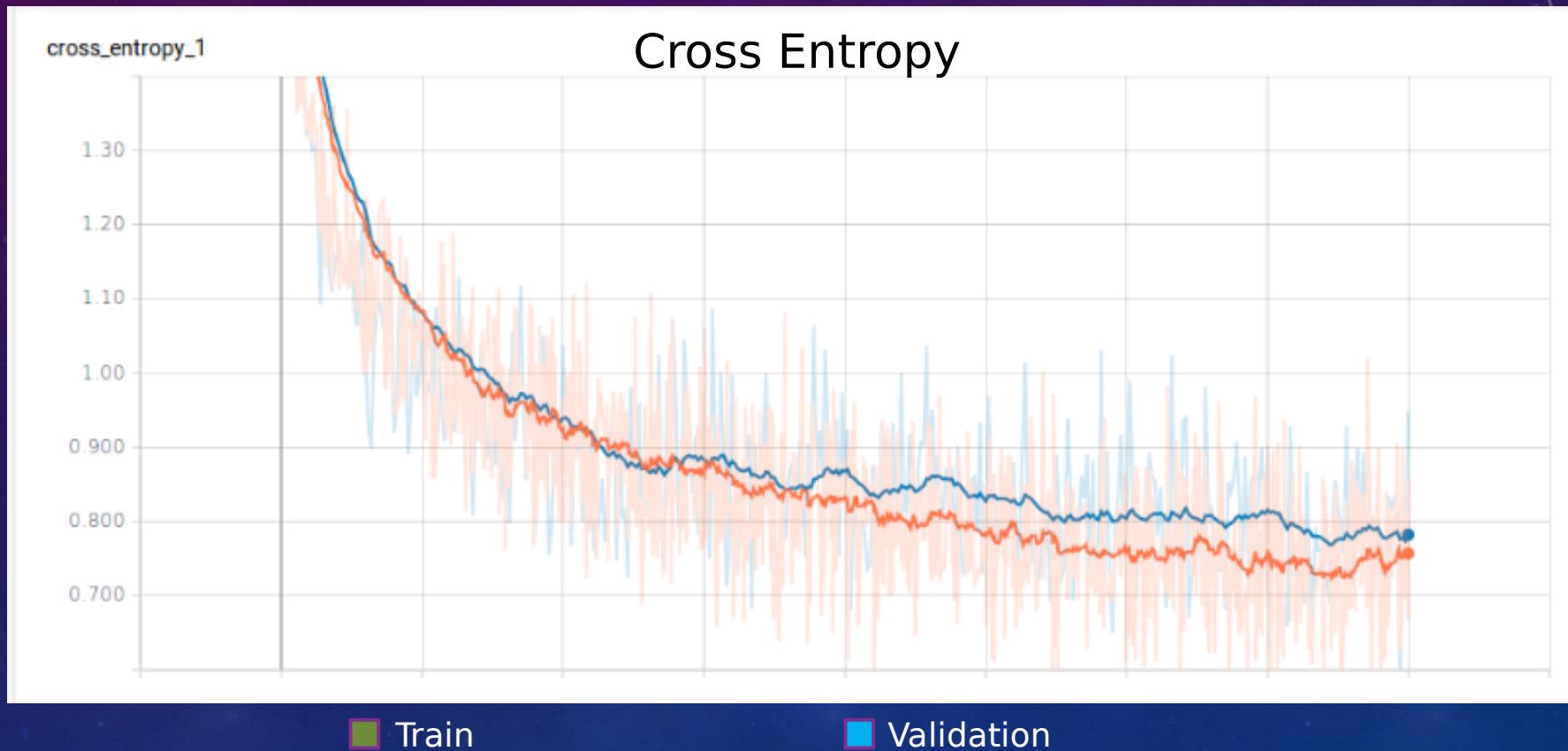
# RESULTS

- 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 !