Assignment: Gesture Recognition

Team: Samrat Sengupta, Deepshikha Ranjan, Dinesh Veluri, Govinda Padaki

Date: October 28, 2019

Due to size limitation, we have uploaded a model called “video\_classifier\_resnet50\_rnn\_rgb\_model” in Google Drive. The link to access the same is <https://drive.google.com/open?id=1vWj0iz5PuqgvNhdibIpzjW5d-F5gkYe3>

Alternatively you can also find the model in following github repository <https://github.com/SamratSengupta/imageNet>

We sincerely request you to take this file when reviewing our RESNET50 -RNNModel.

# Experiment Number: 1

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | **Epochs** | **Batch Size** | **Channels** | **Image Processing** | **Optimizer** | **Time Taken** | **Accuracy** |
| Conv3D | 20 | 10 | 3 | Crop: Centre Resize: 120x120 | Adam |  | 0.77 |
| Experiment Description: | | | | | | | |
| At first, we tried with 3 layers of conv3d with maxpool. Then, increased addl layer. On increasing further layers or Increasing epoch size, the accuracy was not improving much. | | | | | | | |

# Experiment Number: 2

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | **Epochs** | **Batch Size** | **Channels** | **Image Processing** | **Optimizer** | **Time Taken** | **Accuracy** |
| Conv3D | 20 | 10 | 3 | Crop: Centre Resize: 120x120 | Adam |  | 0.81 |
| Experiment Description: | | | | | | | |
| Changed filter sizes in 3rd and 4th layers to (1,3,3) from (3,3,3). | | | | | | | |

# Experiment Number: 3

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | **Epochs** | **Batch Size** | **Channels** | **Image Processing** | **Optimizer** | **Time Taken** | **Accuracy** |
| Conv3D | 20 | 10 | 1 | Crop: Centre Resize: 120x120 | Adam |  | 0.80 |
| Experiment Description: | | | | | | | |
| Keeping the model architecture same changed to ***gray scale image*** keeping cropping and resizing same. | | | | | | | |

# Experiment Number: 4

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | **Epochs** | **Batch Size** | **Channels** | **Image Processing** | **Optimizer** | **Time Taken** | **Accuracy** |
| Conv3D | 20 | 10 | 3 | Crop: Centre Resize: 120x120 | SGD |  | 0.79 |
| Experiment Description: | | | | | | | |
| Changed the optimizer, accuracy decreased a bit not overall much change. | | | | | | | |

# Experiment Number: 5

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | **Epochs** | **Batch Size** | **Channels** | **Image Processing** | **Optimizer** | **Time Taken** | **Accuracy** |
| Conv3D | 20 | 10 | 1 | Crop: Centre Resize: 120x120 | SGD |  | 0.78 |
| Experiment Description: | | | | | | | |
| Changed the image processing into Grey Scale. | | | | | | | |

# Experiment Number: 6

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | **Epochs** | **Batch Size** | **Channels** | **Image Processing** | **Optimizer** | **Time Taken** | **Accuracy** |
| Resnet50-GRU | 10 | 10 | 3 | Crop: Centre Resize: 197x197 | Adam |  | 0.82 |
| Experiment Description: | | | | | | | |
| Used pretrained imagenet weights. For resnet50. Retained weights of 140 layers and trained the model for same. Added globalAveragepooling2D. Used gru and dense layer to complete the model. | | | | | | | |

# Experiment Number: 7

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | **Epochs** | **Batch Size** | **Channels** | **Image Processing** | **Optimizer** | **Time Taken** | **Accuracy** |
| Resnet50-GRU | 15 | 10 | 3 | Crop: Centre Resize: 197x197 | Adam |  | **0.94** (Best so far) |
| Experiment Description: | | | | | | | |
| Increased epochs to 15. | | | | | | | |

# Experiment Number: 8

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | **Epochs** | **Batch Size** | **Channels** | **Image Processing** | **Optimizer** | **Time Taken** | **Accuracy** |
| Conv2D-GRU | 20 | 10 | 3 | Crop: Centre Resize: 120x120 | Adam |  | 0.60 |
| Experiment Description: | | | | | | | |
| A timedistributed conv2d model followed by gru built from scratch was not giving satisfactory accuracy. Hence resorted to pre-trained resnet with GRU approach. | | | | | | | |

# Experiment Number: 9

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | **Epochs** | **Batch Size** | **Channels** | **Image Processing** | **Optimizer** | **Time Taken** | **Accuracy** |
| Conv2D-GRU | 20 | 10 | 3 | Crop: Centre Resize: 120x120 | SGD |  | 0.49 |
| Experiment Description: | | | | | | | |
| Changing optimizer to SGD | | | | | | | |

# Highlights of Code:

## Cropping Images

The following code creates a 120x120 image.

def crop\_image(image):

# Cropping non symmetric frames

if image.shape[0] != image.shape[1]:

image=image[0:120,20:140]

return image

## Resizing of Image

The following code resizes the image. We use OpenCV library to resize the image.

def resize\_image(image,nb\_rows,nb\_cols):

resized = cv2.resize(image, (nb\_rows,nb\_cols), interpolation = cv2.INTER\_AREA)

return resized

## Hyperparameters:

These are the hyperparameters that we used in the final model.

# Parameters initialization

nb\_rows = 120

nb\_cols = 120

nb\_frames = 30

nb\_channel = 3

batch\_size=10

nb\_labels = 5

## Generators

We use two different types of Generators. The b2g\_image\_generator is used to *yield* images in Grayscale. The rgb\_image\_generator is used to yield images in 3 channels RGB.

## Models

We have tried with 5 different models and a combination of optimizers. We have used ***Batch Normalization*** in all the models. Our best result was from RESNet50.

1. build\_conv3d\_model
2. build\_conv2d\_rnn\_model
3. build\_vgg16\_rnn\_model
4. build\_vgg19\_rnn\_model
5. build\_resnet50\_rnn\_model