Gesture Recognition Write-Up

Google Drive Link:

https://drive.google.com/drive/folders/1T7zi3eRZnFIV7LoinbtOkmFs3DCS_tNU?usp=sharing

Model Path: model_init_2022-05-2521_49_00.765099\model-00030-1.48737-0.34161-1.84481-0.14474.h5

Model: https://drive.google.com/file/d/1T5VMk-o6CmA_G618CtYBzAS-AQPEBDaz/view?usp=sharing

GitHub Repo:

https://github.com/98abhilash/Gesture-Recognition

Problem Statement

Imagine you are working as a data scientist at a home electronics company which manufactures state of the art smart televisions. You want to develop a cool feature in the smart-TV that can recognise five different gestures performed by the user which will help users control the TV without using a remote

The gestures are continuously monitored by the webcam mounted on the TV. Each gesture corresponds to a specific command:

- Thumbs up: Increase the volume
- Thumbs down: Decrease the volume
- Left swipe: 'Jump' backwards 10 seconds
- Right swipe: 'Jump' forward 10 seconds
- Stop: Pause the movie

Each video is a sequence of 30 frames (or images).

Understanding the Dataset

The training data consists of a few hundred videos categorised into one of the five classes. Each video (typically 2-3 seconds long) is divided into a sequence of 30 frames(images). These videos have been recorded by various people performing one of the five gestures in front of a webcam - similar to what the smart TV will use.

The data is in a zip file. The zip file contains a 'train' and a 'val' folder with two CSV files for the two folders. These folders are in turn divided into subfolders where each subfolder represents a video of a particular gesture. Each subfolder, i.e. a video, contains 30 frames (or images).

Approach

Experiment 1: Type Conv3d

Frames Considered: 18 Frames

Batch Size = 16

Activation Functions: Relu and Softmax

Kernel Size = 3,3,3

Number of Trainable Parameters: 12,322,949

Image Size = 100,100

Epochs = 30

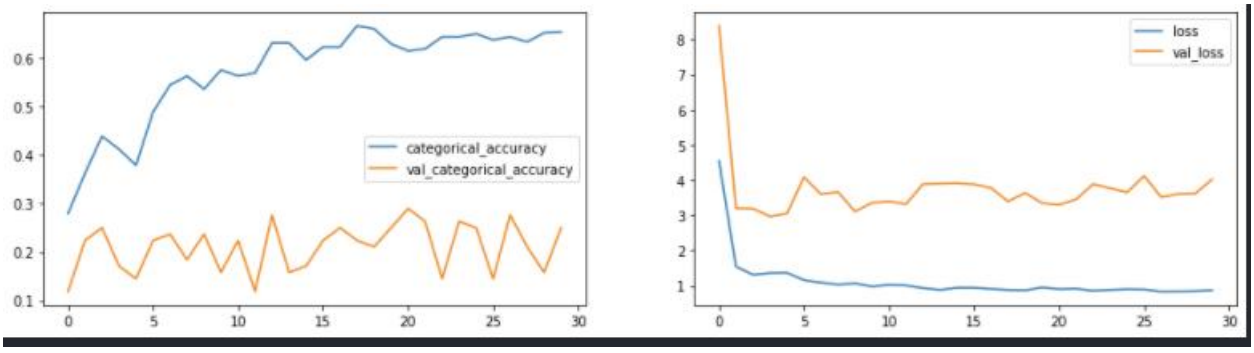
Optimizer = SGD(lr=0.001, decay=1e-6, momentum=0.9, nesterov=True)

Model Path = model_init_2022-05-2423_40_30.698988\model-00030-0.86914-0.65424-4.02549-0.25000.h5

Model Summary:

```
Output exceeds the size limit. Open the full output data in a text editor
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv3d (Conv3D)              (None, 18, 100, 100, 64)  5248
_________________________________________________________________
batch_normalization (BatchNo (None, 18, 100, 100, 64)  256
_________________________________________________________________
activation (Activation)      (None, 18, 100, 100, 64)  0
_________________________________________________________________
max_pooling3d (MaxPooling3D) (None, 9, 50, 100, 64)    0
_________________________________________________________________
conv3d_1 (Conv3D)            (None, 9, 50, 100, 128)   221312
_________________________________________________________________
batch_normalization_1 (Batch (None, 9, 50, 100, 128)   512
_________________________________________________________________
activation_1 (Activation)    (None, 9, 50, 100, 128)   0
_________________________________________________________________
max_pooling3d_1 (MaxPooling3 (None, 4, 25, 50, 128)    0
_________________________________________________________________
conv3d_2 (Conv3D)            (None, 4, 25, 50, 256)    884992
_________________________________________________________________
batch_normalization_2 (Batch (None, 4, 25, 50, 256)    1024
_________________________________________________________________
activation_2 (Activation)    (None, 4, 25, 50, 256)    0
...
Trainable params: 12,322,949
Non-trainable params: 1,408
```

Accuracy and Loss:



Experiment 2: Type Conv3d

Frames Considered: 15 Frames (alternate)

Batch Size = 16

Activation Functions: Relu and Softmax

Kernel Size = 3,3,3

Number of Trainable Parameters: 40,437,125

Image Size = 100,100

Epochs = 30

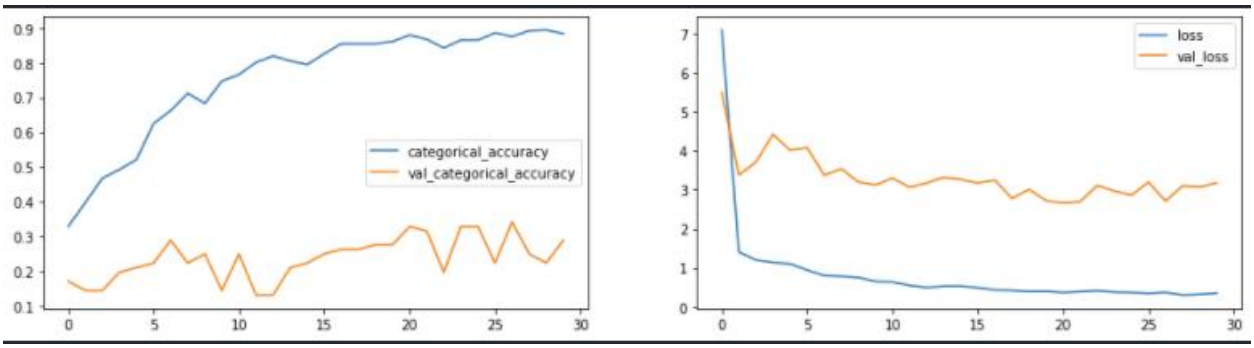
Optimizer = SGD(lr=0.001, decay=1e-6, momentum=0.9, nesterov=True)

Epoch 00030: saving model to model_init_2022-05-2500_25_39.273273\model-00030-0.35095-0.88406-3.17772-0.28947.h5

Model Summary:

```
Output exceeds the size limit. Open the full output data in a text ed
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv3d (Conv3D)              (None, 15, 100, 100, 64)  5248
_________________________________________________________________
batch_normalization (BatchNo (None, 15, 100, 100, 64)  256
_________________________________________________________________
activation (Activation)      (None, 15, 100, 100, 64)  0
_________________________________________________________________
max_pooling3d (MaxPooling3D) (None, 7, 50, 100, 64)    0
_________________________________________________________________
conv3d_1 (Conv3D)            (None, 7, 50, 100, 128)   221312
_________________________________________________________________
batch_normalization_1 (Batch (None, 7, 50, 100, 128)   512
_________________________________________________________________
activation_1 (Activation)    (None, 7, 50, 100, 128)   0
_________________________________________________________________
max_pooling3d_1 (MaxPooling3 (None, 3, 25, 50, 128)    0
_________________________________________________________________
conv3d_2 (Conv3D)            (None, 3, 25, 50, 256)    884992
_________________________________________________________________
batch_normalization_2 (Batch (None, 3, 25, 50, 256)    1024
_________________________________________________________________
activation_2 (Activation)    (None, 3, 25, 50, 256)    0
...
Trainable params: 40,437,125
Non-trainable params: 896
_________________________________________________________________
None
```

Accuracy and Loss:



Experiment 3: Type Conv3d

Frames Considered: 15 Frames (alternate)

Batch Size = 16

Activation Functions: Relu and Softmax

Kernel Size = 3,3,3

Number of Trainable Parameters: 40,437,125

Image Size = 100,100

Epochs = 30

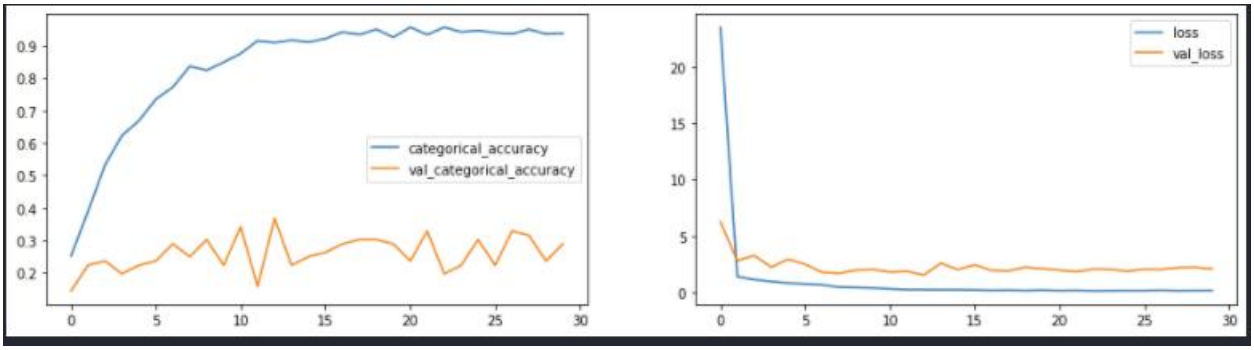
Optimizer = Adam(lr=0.001, decay=1e-6)

Epoch 00030: saving model to model_init_2022-05-2500_51_37.032489\model-00030-0.17731-0.93789-2.08135-0.28947.h5

Model Summary:

```
Output exceeds the size limit. Open the full output data in a text edit
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv3d (Conv3D)              (None, 15, 100, 100, 64)  5248
_________________________________________________________________
batch_normalization (BatchNo (None, 15, 100, 100, 64)  256
_________________________________________________________________
activation (Activation)      (None, 15, 100, 100, 64)  0
_________________________________________________________________
max_pooling3d (MaxPooling3D) (None, 7, 50, 100, 64)    0
_________________________________________________________________
conv3d_1 (Conv3D)            (None, 7, 50, 100, 128)   221312
_________________________________________________________________
batch_normalization_1 (Batch (None, 7, 50, 100, 128)   512
_________________________________________________________________
activation_1 (Activation)    (None, 7, 50, 100, 128)   0
_________________________________________________________________
max_pooling3d_1 (MaxPooling3 (None, 3, 25, 50, 128)    0
_________________________________________________________________
conv3d_2 (Conv3D)            (None, 3, 25, 50, 256)    884992
_________________________________________________________________
batch_normalization_2 (Batch (None, 3, 25, 50, 256)    1024
_________________________________________________________________
activation_2 (Activation)    (None, 3, 25, 50, 256)    0
...
Trainable params: 40,437,125
Non-trainable params: 896
_________________________________________________________________
None
```

Accuracy and Loss:



Experiment 4: Type Conv3d

Frames Considered: 30 Frames

Batch Size = 8

Activation Functions: Relu and Softmax

Kernel Size = 3,3,3

Number of Trainable Parameters: 12,322,949

Image Size = 100,100

Epochs = 30

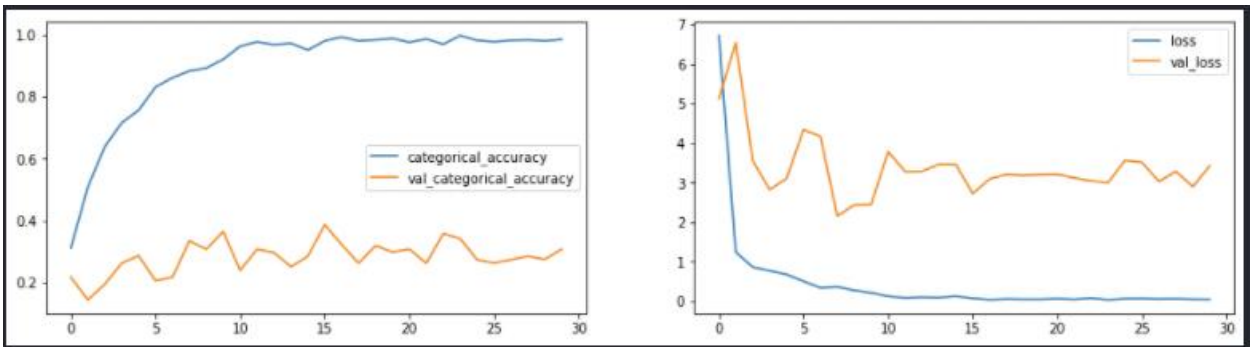
Optimizer = Adam(lr=0.001, decay=1e-6)

Epoch 00030: saving model to model_init_2022-05-2502_08_27.982955\model-00030-0.04855-0.98585-3.42767-0.30682.h5

Model Summary:

```
Output exceeds the size limit. Open the full output data in a text edit
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv3d (Conv3D)              (None, 30, 100, 100, 64)  5248
_________________________________________________________________
batch_normalization (BatchNo (None, 30, 100, 100, 64)  256
_________________________________________________________________
activation (Activation)      (None, 30, 100, 100, 64)  0
_________________________________________________________________
max_pooling3d (MaxPooling3D) (None, 15, 50, 100, 64)   0
_________________________________________________________________
conv3d_1 (Conv3D)            (None, 15, 50, 100, 128)  221312
_________________________________________________________________
batch_normalization_1 (Batch (None, 15, 50, 100, 128)  512
_________________________________________________________________
activation_1 (Activation)    (None, 15, 50, 100, 128)  0
_________________________________________________________________
max_pooling3d_1 (MaxPooling3 (None, 7, 25, 50, 128)    0
_________________________________________________________________
conv3d_2 (Conv3D)            (None, 7, 25, 50, 256)    884992
_________________________________________________________________
batch_normalization_2 (Batch (None, 7, 25, 50, 256)    1024
_________________________________________________________________
activation_2 (Activation)    (None, 7, 25, 50, 256)    0
...
Trainable params: 12,322,949
Non-trainable params: 1,408
_________________________________________________________________
None
```

Accuracy and Loss:



Experiment 5: Type Conv3d

Frames Considered: 30 Frames

Batch Size = 8

Activation Functions: Relu and Softmax

Kernel Size = 3,3,3

Number of Trainable Parameters: 12,322,949

Image Size = 100,100

Epochs = 30

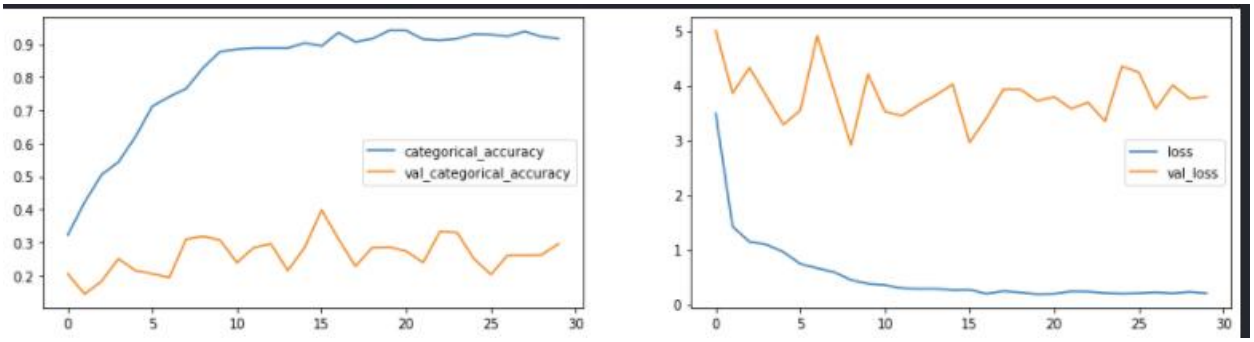
```
Optimiser = .SGD(lr=0.001, decay=1e-6, momentum=0.9, nesterov=True)
```

Epoch 00030: saving model to model_init_2022-05-2503_10_26.268891\model-00030-0.20177-0.91667-3.79952-0.29545.h5

Model Summary:

```
Output exceeds the size limit. Open the full output data in a text ed
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv3d (Conv3D)              (None, 30, 100, 100, 64)  5248
_________________________________________________________________
batch_normalization (BatchNo (None, 30, 100, 100, 64)  256
_________________________________________________________________
activation (Activation)      (None, 30, 100, 100, 64)  0
_________________________________________________________________
max_pooling3d (MaxPooling3D) (None, 15, 50, 100, 64)   0
_________________________________________________________________
conv3d_1 (Conv3D)            (None, 15, 50, 100, 128)  221312
_________________________________________________________________
batch_normalization_1 (Batch (None, 15, 50, 100, 128)  512
_________________________________________________________________
activation_1 (Activation)    (None, 15, 50, 100, 128)  0
_________________________________________________________________
max_pooling3d_1 (MaxPooling3 (None, 7, 25, 50, 128)    0
_________________________________________________________________
conv3d_2 (Conv3D)            (None, 7, 25, 50, 256)    884992
_________________________________________________________________
batch_normalization_2 (Batch (None, 7, 25, 50, 256)    1024
_________________________________________________________________
activation_2 (Activation)    (None, 7, 25, 50, 256)    0
...
Trainable params: 12,322,949
Non-trainable params: 1,408
_________________________________________________________________
None
```

Accuracy and Loss:



Experiment 6: Type Conv3d

Frames Considered: 10 Frames (middle)

Batch Size = 16

Activation Functions: Relu and Softmax

Kernel Size = 3,3,3

Number of Trainable Parameters: 40,437,125

Image Size = 100,100

Epochs = 30

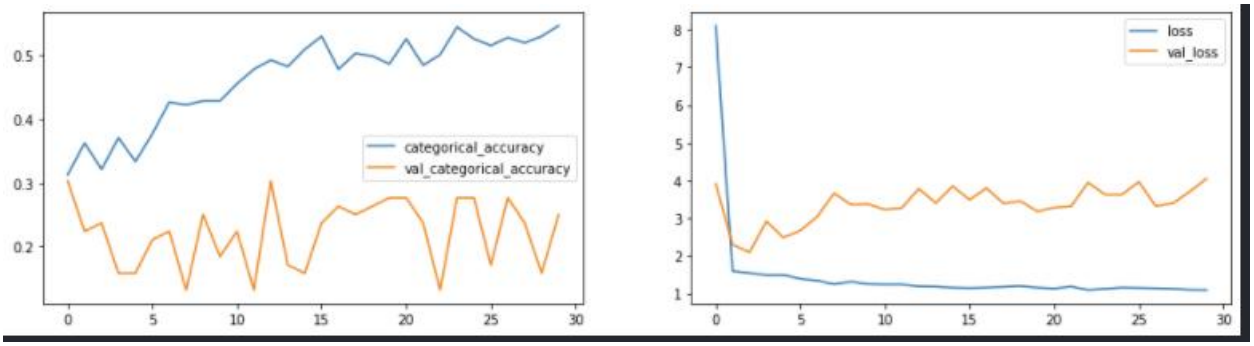
Optimiser = .SGD(lr=0.001, decay=1e-6, momentum=0.9, nesterov=True)

Epoch 00030: saving model to model_init_2022-05-2504_11_39.296931\model-00030-1.09299-0.54658-4.04598-0.25000.h5

Model Summary:

```
Output exceeds the size limit. Open the full output data in a text editor
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv3d (Conv3D)              (None, 10, 100, 100, 64)  5248
_________________________________________________________________
batch_normalization (BatchNo (None, 10, 100, 100, 64)  256
_________________________________________________________________
activation (Activation)      (None, 10, 100, 100, 64)  0
_________________________________________________________________
max_pooling3d (MaxPooling3D) (None, 5, 50, 100, 64)    0
_________________________________________________________________
conv3d_1 (Conv3D)            (None, 5, 50, 100, 128)   221312
_________________________________________________________________
batch_normalization_1 (Batch (None, 5, 50, 100, 128)   512
_________________________________________________________________
activation_1 (Activation)    (None, 5, 50, 100, 128)   0
_________________________________________________________________
max_pooling3d_1 (MaxPooling3 (None, 2, 25, 50, 128)    0
_________________________________________________________________
conv3d_2 (Conv3D)            (None, 2, 25, 50, 256)    884992
_________________________________________________________________
batch_normalization_2 (Batch (None, 2, 25, 50, 256)    1024
_________________________________________________________________
activation_2 (Activation)    (None, 2, 25, 50, 256)    0
...
Trainable params: 40,437,125
Non-trainable params: 896
_________________________________________________________________
None
```

Accuracy and Loss:



Experiment 7: Type Conv3d

Frames Considered: 18 Frames

Batch Size = 16

Activation Functions: Relu and Softmax

Kernel Size = 3,3,3

Number of Trainable Parameters: 19,157,189

Image Size = 100,100

Epochs = 30

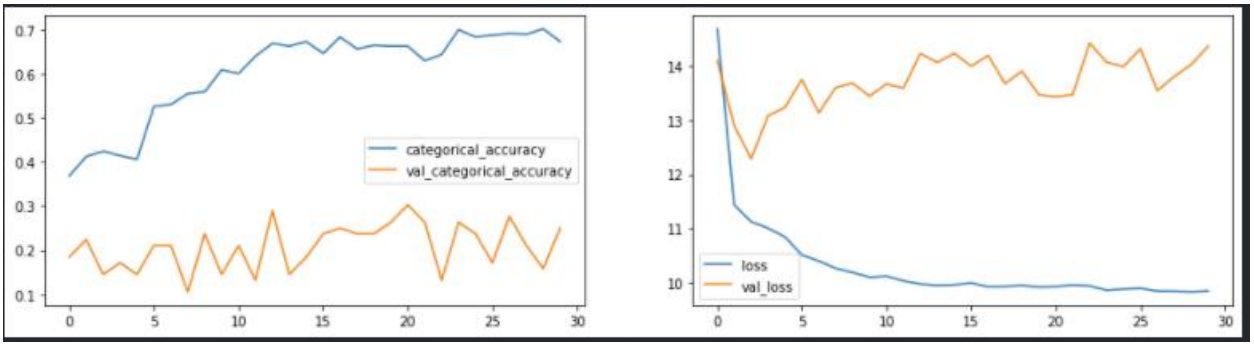
```
Optimiser = .SGD(lr=0.001, decay=1e-6, momentum=0.9, nesterov=True)
```

Epoch 00030: saving model to model_init_2022-05-2510_45_50.623547\model-00030-9.83907-0.67288-14.37144-0.25000.h5

Model Summary:

```
Output exceeds the size limit. Open the full output data in a text editor
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv3d (Conv3D)              (None, 10, 100, 100, 64)  5248
_________________________________________________________________
batch_normalization (BatchNo (None, 10, 100, 100, 64)  256
_________________________________________________________________
activation (Activation)      (None, 10, 100, 100, 64)  0
_________________________________________________________________
max_pooling3d (MaxPooling3D) (None, 5, 50, 100, 64)    0
_________________________________________________________________
conv3d_1 (Conv3D)            (None, 5, 50, 100, 128)   221312
_________________________________________________________________
batch_normalization_1 (Batch (None, 5, 50, 100, 128)   512
_________________________________________________________________
activation_1 (Activation)    (None, 5, 50, 100, 128)   0
_________________________________________________________________
max_pooling3d_1 (MaxPooling3 (None, 2, 25, 50, 128)    0
_________________________________________________________________
conv3d_2 (Conv3D)            (None, 2, 25, 50, 256)    884992
_________________________________________________________________
batch_normalization_2 (Batch (None, 2, 25, 50, 256)    1024
_________________________________________________________________
activation_2 (Activation)    (None, 2, 25, 50, 256)    0
...
Trainable params: 40,437,125
Non-trainable params: 896
_________________________________________________________________
None
```

Accuracy and Loss:



Experiment 8: Type LSTM

Frames Considered: 18 Frames

Batch Size = 16

Activation Functions: Relu and Softmax

Kernel Size = 3,3,3

Number of Trainable Parameters: 1,647,877

Image Size = 100,100

Epochs = 15

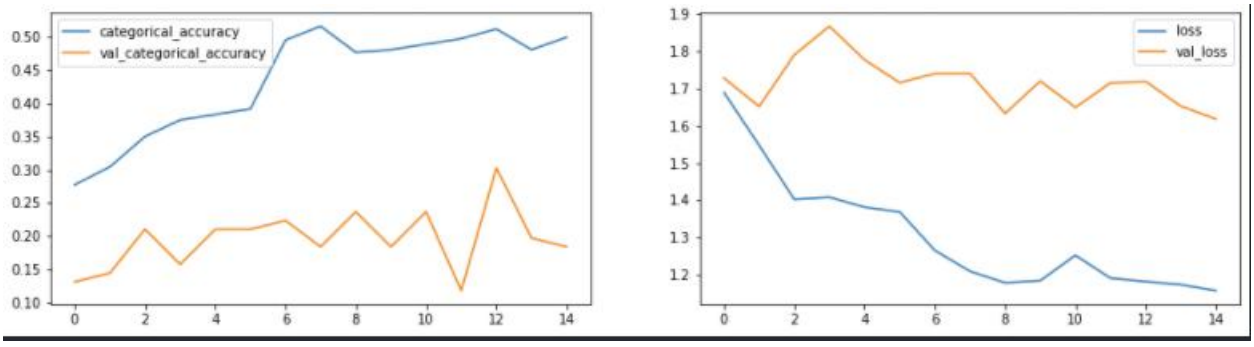
Optimiser = Adam()

Epoch 00015: saving model to model_init_2022-05-2514_00_28.262800\model-00015-1.15772-0.49896-1.61810-0.18421.h5

Model Summary:

```
Output exceeds the size limit. Open the full output data in a text editor
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
time_distributed (TimeDistri (None, 18, 100, 100, 16)  448
_________________________________________________________________
time_distributed_1 (TimeDist (None, 18, 100, 100, 16)  64
_________________________________________________________________
time_distributed_2 (TimeDist (None, 18, 50, 50, 16)    0
_________________________________________________________________
time_distributed_3 (TimeDist (None, 18, 50, 50, 32)    4640
_________________________________________________________________
time_distributed_4 (TimeDist (None, 18, 50, 50, 32)    128
_________________________________________________________________
time_distributed_5 (TimeDist (None, 18, 25, 25, 32)    0
_________________________________________________________________
time_distributed_6 (TimeDist (None, 18, 25, 25, 64)    18496
_________________________________________________________________
time_distributed_7 (TimeDist (None, 18, 25, 25, 64)    256
_________________________________________________________________
time_distributed_8 (TimeDist (None, 18, 12, 12, 64)    0
_________________________________________________________________
time_distributed_9 (TimeDist (None, 18, 12, 12, 128)   73856
_________________________________________________________________
time_distributed_10 (TimeDis (None, 18, 12, 12, 128)   512
...
Trainable params: 1,647,877
Non-trainable params: 992
_________________________________________________________________
None
```

Accuracy and Loss:



Experiment 9: Type LSTM

Frames Considered: 18 Frames

Batch Size = 16

Activation Functions: Relu and Softmax

Kernel Size = 3,3,3

Number of Trainable Parameters:999,813

Image Size = 100,100

Epochs = 15

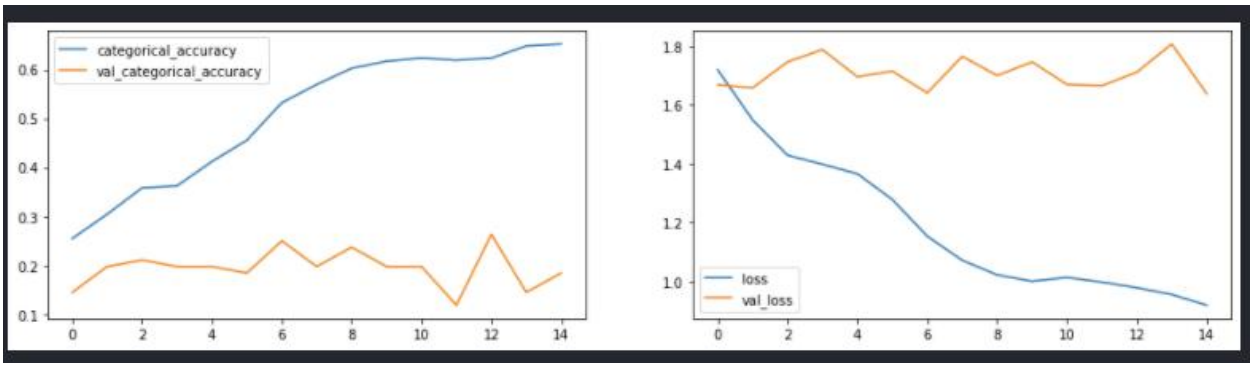
Optimiser = Adam()

Epoch 00015: saving model to model_init_2022-05-2515_35_43.990920\model-00015-0.91924-0.65217-1.63753-0.18421.h5

Model Summary:

```
Output exceeds the size limit. Open the full output data in a text editor
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
time_distributed (TimeDistri (None, 18, 100, 100, 16)  448
_________________________________________________________________
time_distributed_1 (TimeDist (None, 18, 100, 100, 16)  64
_________________________________________________________________
time_distributed_2 (TimeDist (None, 18, 50, 50, 16)    0
_________________________________________________________________
time_distributed_3 (TimeDist (None, 18, 50, 50, 32)    4640
_________________________________________________________________
time_distributed_4 (TimeDist (None, 18, 50, 50, 32)    128
_________________________________________________________________
time_distributed_5 (TimeDist (None, 18, 25, 25, 32)    0
_________________________________________________________________
time_distributed_6 (TimeDist (None, 18, 25, 25, 64)    18496
_________________________________________________________________
time_distributed_7 (TimeDist (None, 18, 25, 25, 64)    256
_________________________________________________________________
time_distributed_8 (TimeDist (None, 18, 12, 12, 64)    0
_________________________________________________________________
time_distributed_9 (TimeDist (None, 18, 12, 12, 128)   73856
_________________________________________________________________
time_distributed_10 (TimeDis (None, 18, 12, 12, 128)   512
...
Trainable params: 999,813
Non-trainable params: 480
_________________________________________________________________
None
```

Accuracy and Loss:



Experiment 10: Type GRU

Frames Considered: 18 Frames

Batch Size = 16

Activation Functions: Relu and Softmax

Kernel Size = 3,3,3

Number of Trainable Parameters:1,925,893

Image Size = 100,100

Epochs = 30

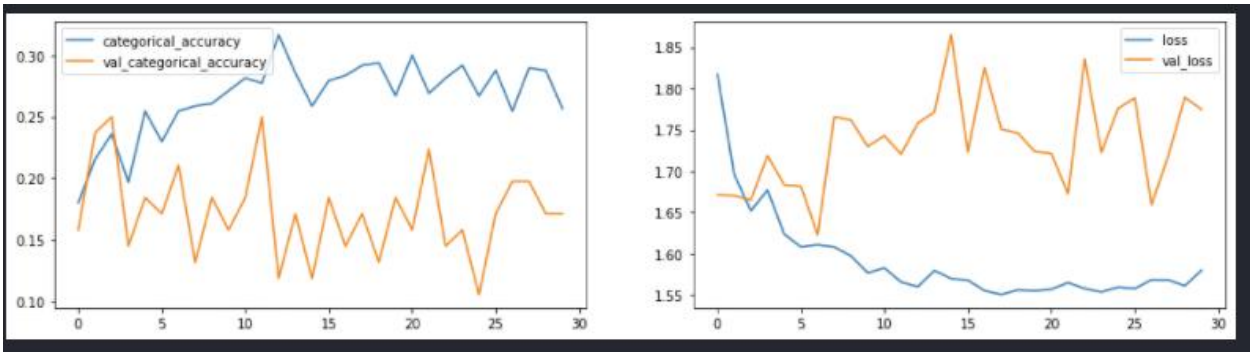
Optimiser =SGD(lr=0.001, decay=1e-6)

Epoch 00030: saving model to model_init_2022-05-2519_53_48.104918\model-00030-1.58027-0.25673-1.77449-0.17105.h5

Model Summary:

```
Output exceeds the size limit. Open the full output data in a text editor
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
time_distributed (TimeDistri (None, 18, 100, 100, 16)  448
_________________________________________________________________
time_distributed_1 (TimeDist (None, 18, 100, 100, 16)  64
_________________________________________________________________
time_distributed_2 (TimeDist (None, 18, 50, 50, 16)    0
_________________________________________________________________
time_distributed_3 (TimeDist (None, 18, 50, 50, 32)    4640
_________________________________________________________________
time_distributed_4 (TimeDist (None, 18, 50, 50, 32)    128
_________________________________________________________________
time_distributed_5 (TimeDist (None, 18, 25, 25, 32)    0
_________________________________________________________________
time_distributed_6 (TimeDist (None, 18, 25, 25, 32)    0
_________________________________________________________________
time_distributed_7 (TimeDist (None, 18, 25, 25, 64)    18496
_________________________________________________________________
time_distributed_8 (TimeDist (None, 18, 25, 25, 64)    256
_________________________________________________________________
time_distributed_9 (TimeDist (None, 18, 12, 12, 64)    0
_________________________________________________________________
time_distributed_10 (TimeDis (None, 18, 12, 12, 64)    0
...
Trainable params: 1,925,893
Non-trainable params: 480
_________________________________________________________________
None
```

Accuracy and Loss:



Experiment 11: Type GRU

Frames Considered: 18 Frames

Batch Size = 16

Activation Functions: Relu and Softmax

Kernel Size = 3,3,3

Number of Trainable Parameters::748,549

Image Size = 100,100

Epochs = 30

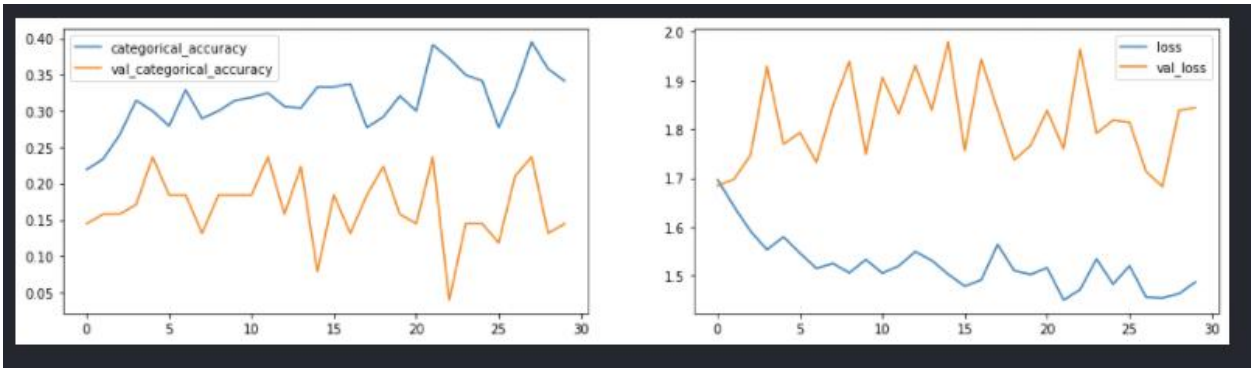
```
Optimiser =SGD(lr=0.001, decay=1e-6)
```

Epoch 00030: saving model to model_init_2022-05-2521_49_00.765099\model-00030-1.48737-0.34161-1.84481-0.14474.h5

Model Summary:

```
Output exceeds the size limit. Open the full output data in a text editor
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
time_distributed (TimeDistri (None, 30, 100, 100, 16)  448
_________________________________________________________________
time_distributed_1 (TimeDist (None, 30, 50, 50, 16)    0
_________________________________________________________________
time_distributed_2 (TimeDist (None, 30, 50, 50, 16)    64
_________________________________________________________________
time_distributed_3 (TimeDist (None, 30, 50, 50, 16)    0
_________________________________________________________________
time_distributed_4 (TimeDist (None, 30, 50, 50, 32)    4640
_________________________________________________________________
time_distributed_5 (TimeDist (None, 30, 25, 25, 32)    0
_________________________________________________________________
time_distributed_6 (TimeDist (None, 30, 25, 25, 32)    128
_________________________________________________________________
time_distributed_7 (TimeDist (None, 30, 25, 25, 32)    0
_________________________________________________________________
time_distributed_8 (TimeDist (None, 30, 25, 25, 64)    18496
_________________________________________________________________
time_distributed_9 (TimeDist (None, 30, 12, 12, 64)    0
_________________________________________________________________
time_distributed_10 (TimeDis (None, 30, 12, 12, 64)    256
...
Trainable params: 748,549
Non-trainable params: 736
_________________________________________________________________
None
```

Accuracy and Loss:


categorical_accuracy
val_categorical_accuracy
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0
5
10
15
20
25
30
loss
val_loss
2.0
1.9
1.8
1.7
1.6
1.5
0
5
10
15
20
25
30