FACIAL EMOTION RECOGNITION

by: Jefferson Qiu
General Assembly DSI20

01 INTRODUCTION

02 THE DATASET

03 THE MODELS

04 RECOMMENDATION

FACIAL EMOTION RECOGNITION

01.

INTRODUCTION

WHY BOTHER?

FACIAL EXPRESSION

Facial Expression plays a crucial part of human communication

DIGITAL INTERACTION

Increase in human-machine interaction

EMOTION RECOGNITION

Growing needs for emotion recognition tools

POSSIBLE APPLICATIONS

STUDENT MONITORING

Assess children emotion during remote study, i.e. attention, confusion, frustration, etc.

CUSTOMER SERVICE

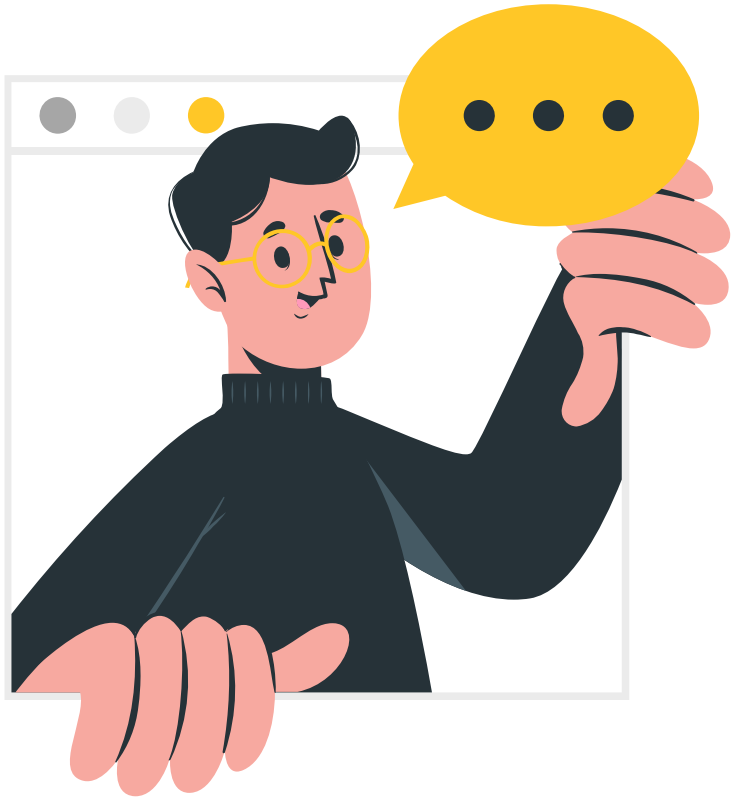
Assess CS quality based on customer’s facial expression

MARKET RESEARCH

A/B testing based on correspondent's expression

PERSONAL WELL-BEING

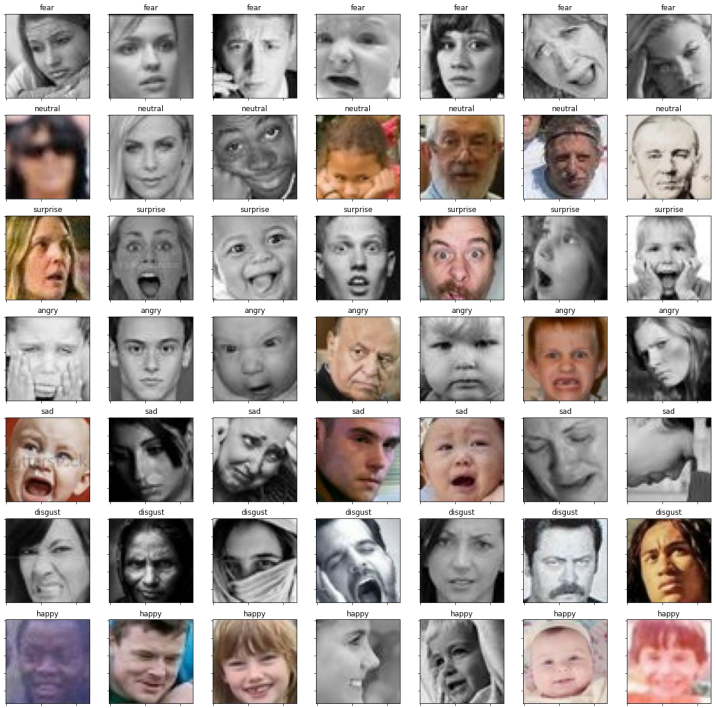
Daily emotion monitoring



02.

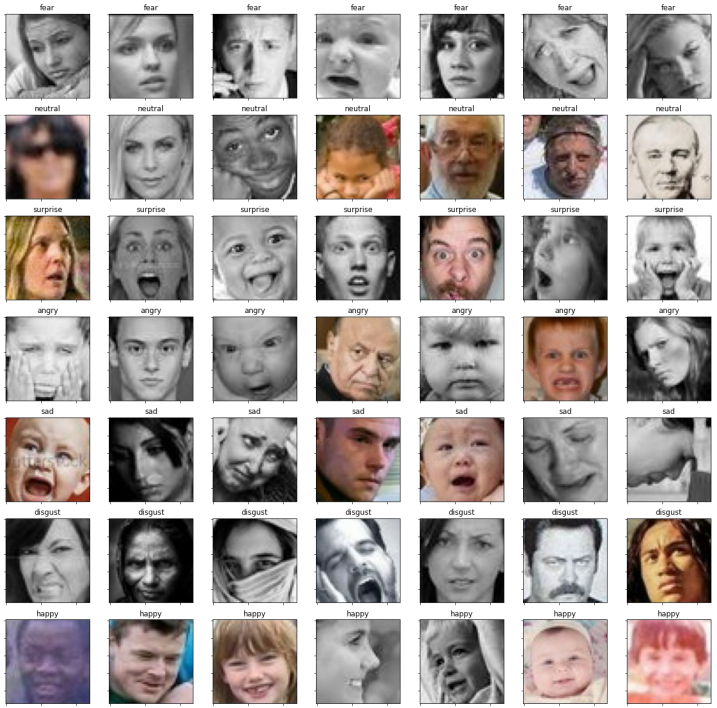
THE DATASET

FACIAL EXPRESSION DATASET



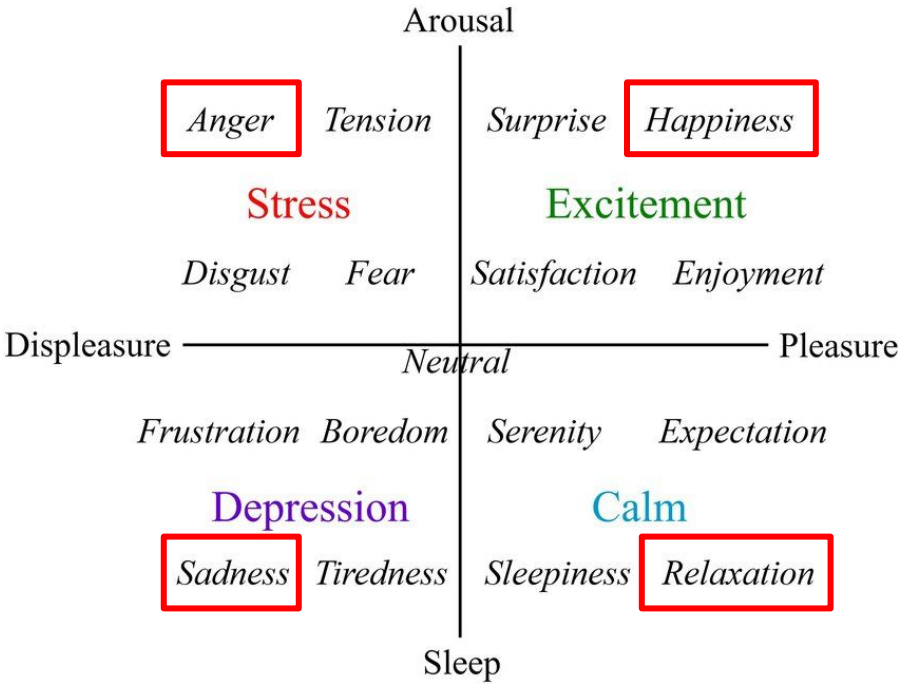
- From Kaggle
- Consists of:
  - >90k training set
  - >17k validation sets
  - >17k test set
- Divided into **7 expressions**:
  - Neutral : 33%
  - Happy : 31%
  - Sad : 12%
  - Surprise : 8%
  - Angry : 6%
  - Fear : 4%
  - Disgust : 3%

FACIAL EXPRESSION DATASET



- Many ambiguity (?)

EMOTION QUADRANT



Obtain the most distinctive emotions to avoid ambiguity:

- Angry
- Sad
- Happy
- Relaxation

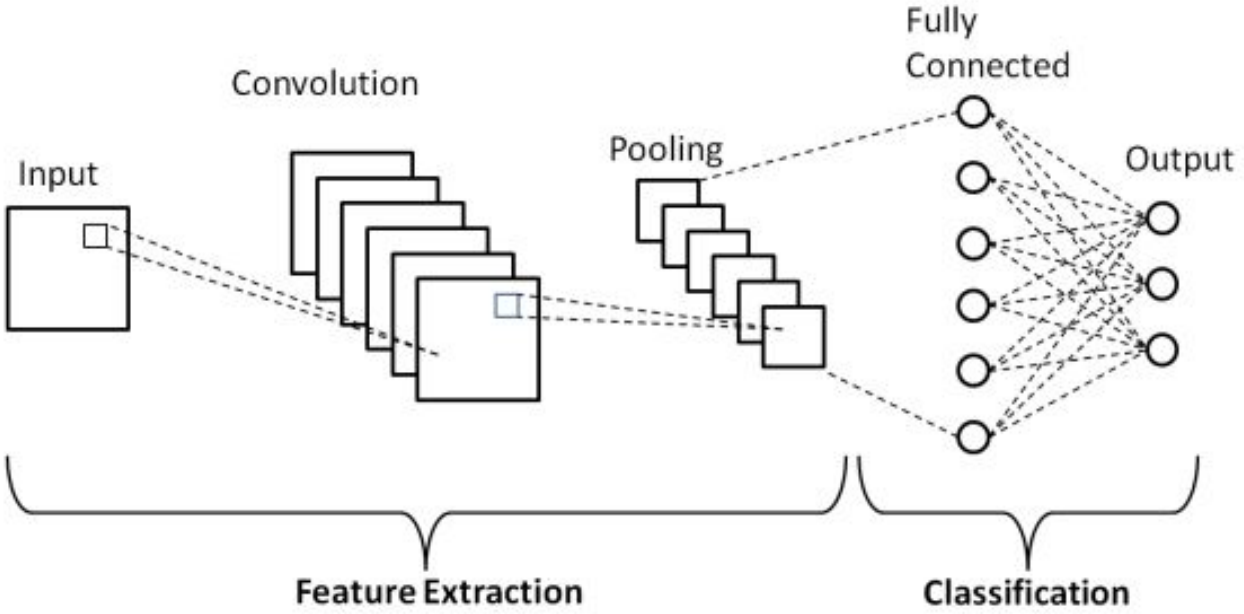
Source: https://www.researchgate.net/publication/335191634_Quantum_Structure_for_Modelling_Emotion_Space_of_Robots

FACIAL EXPRESSION DATASET

03.

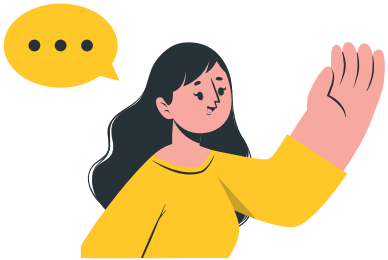
THE MODEL

BASIC OF CONVOLUTIONAL NEURAL NETWORK



Source : https://www.researchgate.net/figure/Schematic-diagram-of-a-basic-convolutional-neural-network-CNN-architecture-26_fig1_336805909

THE MODEL BUILDING



PRETRAINED MODEL

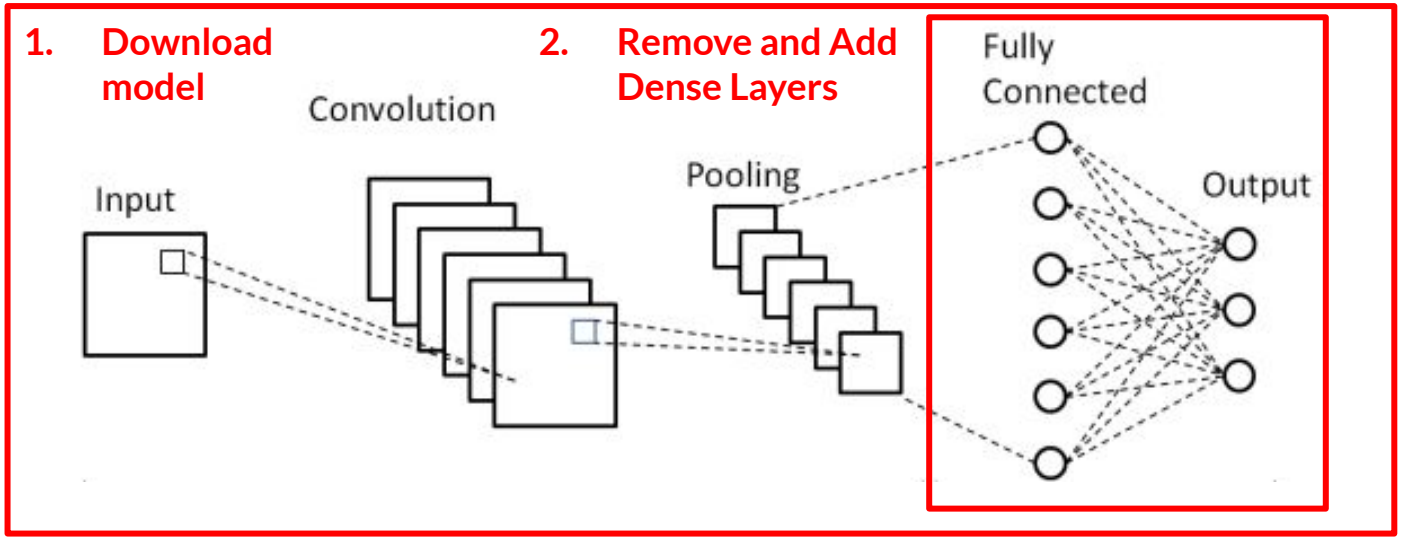
- Build model based on well-established model.
- Fine-tuned to the existing dataset.

DIY MODEL

- Build from scratch
- Fine-tuned by random search on a set of variable search space

PRETRAINED MODELS

4. Fine tuning:
Unfreeze last block



3. Initial training — 10 epochs

5. Final training — Early Stopping, Reduce Learning Rate on Plateau

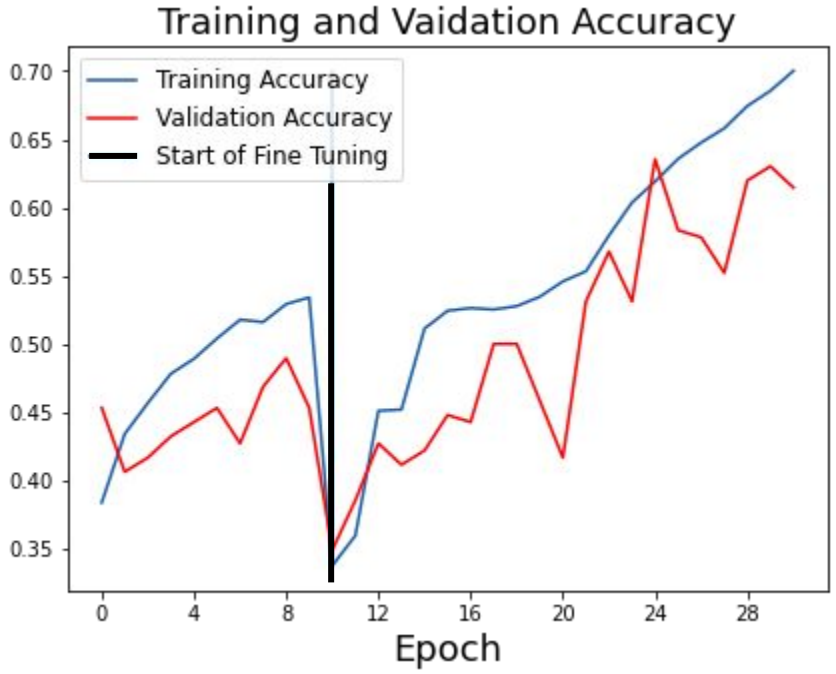
VGG16

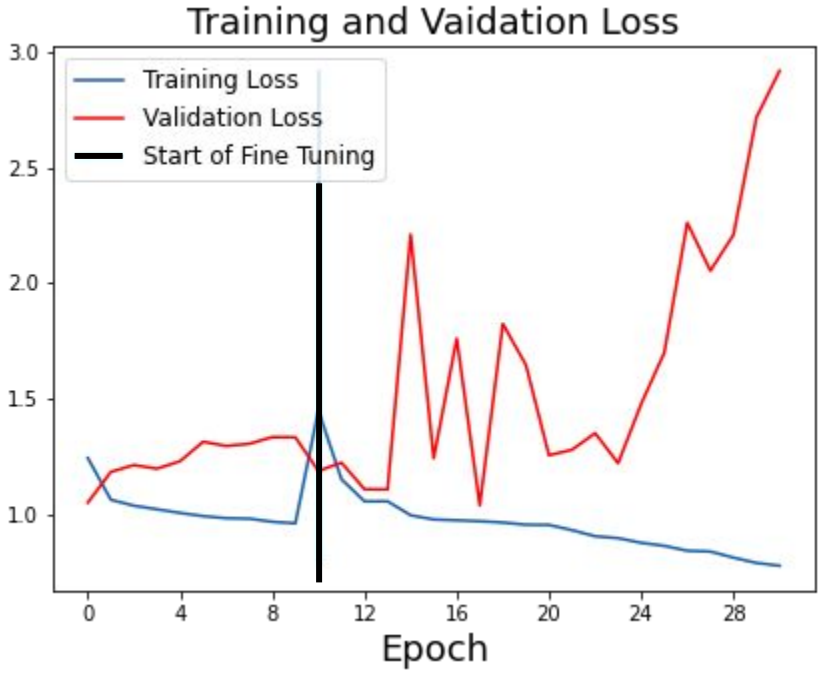
Layer (type)	Output Shape	Param #
input_11 (InputLayer)	[(None, 48, 48, 3)]	0
block1_conv1 (Conv2D)	(None, 48, 48, 64)	1792
block1_conv2 (Conv2D)	(None, 48, 48, 64)	36928
block1_pool (MaxPooling2D)	(None, 24, 24, 64)	0
block2_conv1 (Conv2D)	(None, 24, 24, 128)	73856
block2_conv2 (Conv2D)	(None, 24, 24, 128)	147584
block2_pool (MaxPooling2D)	(None, 12, 12, 128)	0
block3_conv1 (Conv2D)	(None, 12, 12, 256)	295168
block3_conv2 (Conv2D)	(None, 12, 12, 256)	590080
block3_conv3 (Conv2D)	(None, 12, 12, 256)	590080
block3_pool (MaxPooling2D)	(None, 6, 6, 256)	0
block4_conv1 (Conv2D)	(None, 6, 6, 512)	1180160
block4_conv2 (Conv2D)	(None, 6, 6, 512)	2359808
block4_conv3 (Conv2D)	(None, 6, 6, 512)	2359808
block4_pool (MaxPooling2D)	(None, 3, 3, 512)	0
block5_conv1 (Conv2D)	(None, 3, 3, 512)	2359808
block5_conv2 (Conv2D)	(None, 3, 3, 512)	2359808
block5_conv3 (Conv2D)	(None, 3, 3, 512)	2359808
block5_pool (MaxPooling2D)	(None, 1, 1, 512)	0
global_average_pooling2d_9 (	(None, 512)	0

Total params: 14,714,688
Trainable params: 0
Non-trainable params: 14,714,688

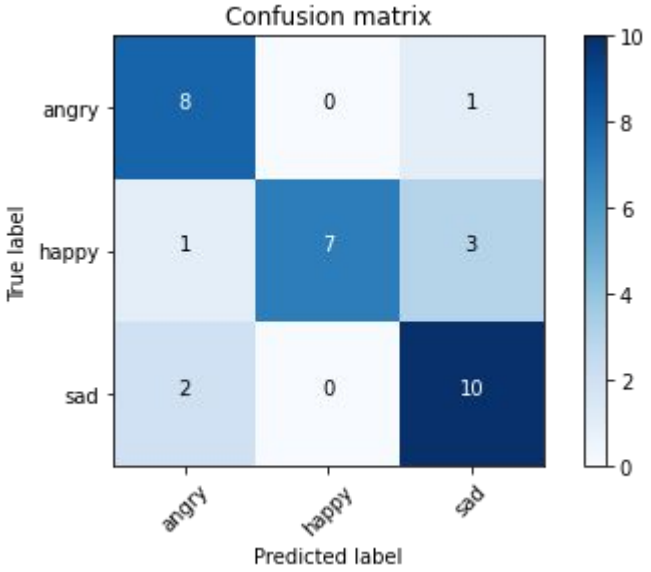
VGG16

Best Validation Accuracy : 63.5%
with Validation Loss : 1.47





VGG16



Predicted as Sad

Predicted as Angry

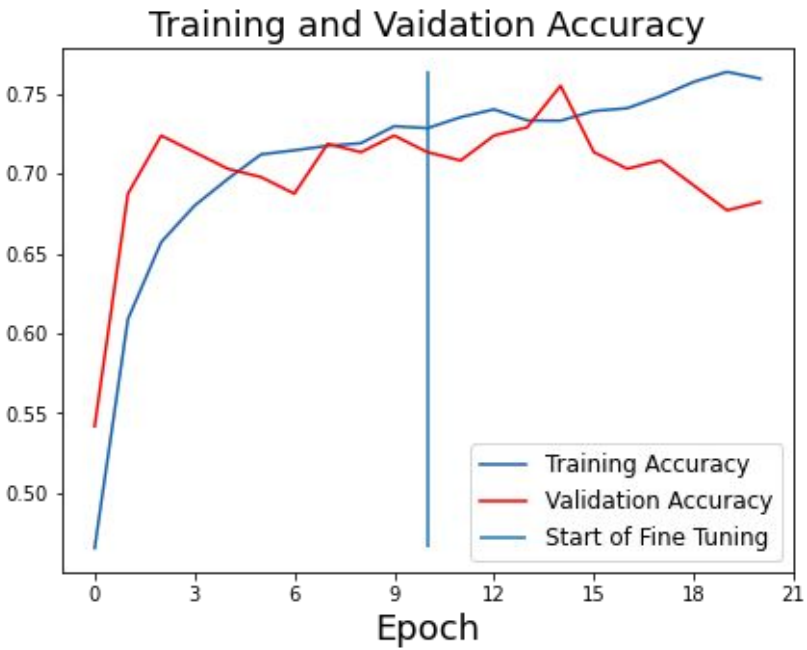
VGGFACE

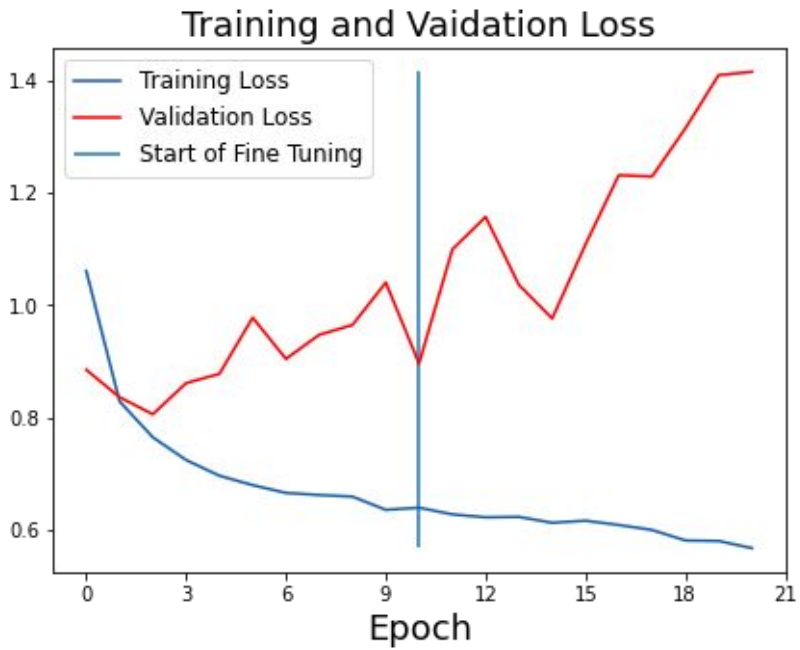
Layer (type)	Output Shape	Param #
input_19 (InputLayer)	[(None, 224, 224, 3)]	0
conv1_1 (Conv2D)	(None, 224, 224, 64)	1792
conv1_2 (Conv2D)	(None, 224, 224, 64)	36928
pool1 (MaxPooling2D)	(None, 112, 112, 64)	0
conv2_1 (Conv2D)	(None, 112, 112, 128)	73856
conv2_2 (Conv2D)	(None, 112, 112, 128)	147584
pool2 (MaxPooling2D)	(None, 56, 56, 128)	0
conv3_1 (Conv2D)	(None, 56, 56, 256)	295168
conv3_2 (Conv2D)	(None, 56, 56, 256)	590080
conv3_3 (Conv2D)	(None, 56, 56, 256)	590080
pool3 (MaxPooling2D)	(None, 28, 28, 256)	0
conv4_1 (Conv2D)	(None, 28, 28, 512)	1180160
conv4_2 (Conv2D)	(None, 28, 28, 512)	2359808
conv4_3 (Conv2D)	(None, 28, 28, 512)	2359808
pool4 (MaxPooling2D)	(None, 14, 14, 512)	0
conv5_1 (Conv2D)	(None, 14, 14, 512)	2359808
conv5_2 (Conv2D)	(None, 14, 14, 512)	2359808
conv5_3 (Conv2D)	(None, 14, 14, 512)	2359808
pool5 (MaxPooling2D)	(None, 7, 7, 512)	0
global_average_pooling2d_12	(None, 512)	0

Total params: 14,714,688
Trainable params: 0
Non-trainable params: 14,714,688

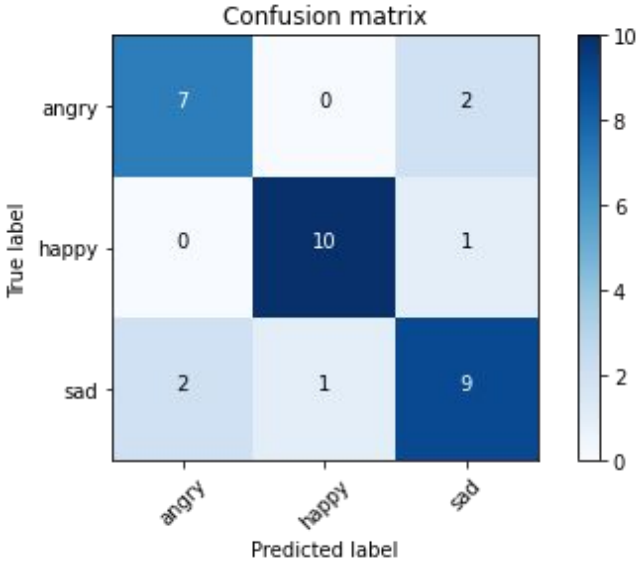
VGGFACE

Best Validation Accuracy : 75.5%
with Validation Loss : 0.97





VGGFACE



Predicted as Happy

Predicted as Angry

Predicted as Sad

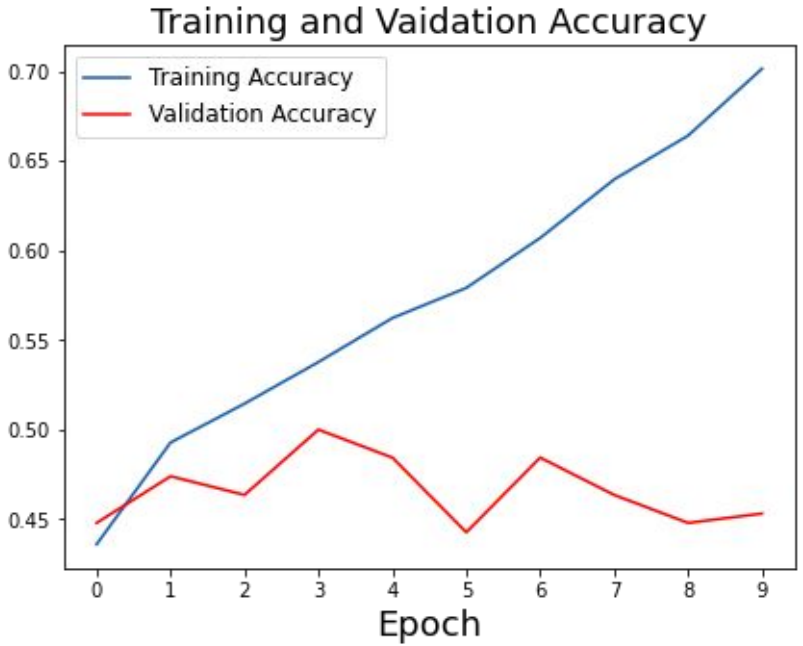
RESNET50

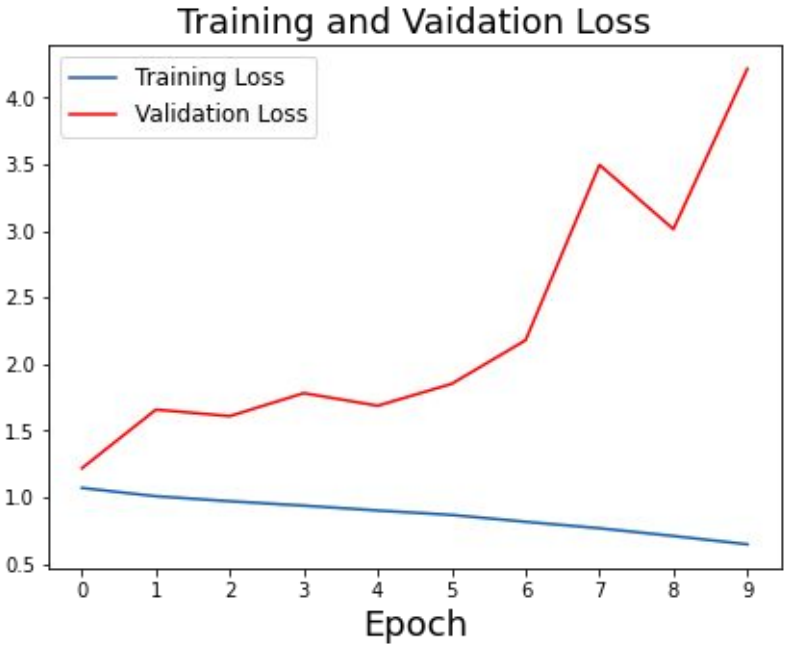
Layer (type)	Output Shape	Param #	Connected to
input_22 (InputLayer)	[(None, 48, 48, 3)]	0	
conv1_pad (ZeroPadding2D)	(None, 54, 54, 3)	0	input_22[0][0]
conv1_conv (Conv2D)	(None, 24, 24, 64)	9472	conv1_pad[0][0]
conv1_bn (BatchNormalization)	(None, 24, 24, 64)	256	conv1_conv[0][0]
conv1_relu (Activation)	(None, 24, 24, 64)	0	conv1_bn[0][0]
pool1_pad (ZeroPadding2D)	(None, 26, 26, 64)	0	conv1_relu[0][0]
pool1_pool (MaxPooling2D)	(None, 12, 12, 64)	0	pool1_pad[0][0]
conv2_block1_1_conv (Conv2D)	(None, 12, 12, 64)	4160	pool1_pool[0][0]
conv2_block1_1_bn (BatchNormali	(None, 12, 12, 64)	256	conv2_block1_1_conv[0][0]
conv2_block1_1_relu (Activation	(None, 12, 12, 64)	0	conv2_block1_1_bn[0][0]
conv2_block1_2_conv (Conv2D)	(None, 12, 12, 64)	36928	conv2_block1_1_relu[0][0]
conv2_block1_2_bn (BatchNormali	(None, 12, 12, 64)	256	conv2_block1_2_conv[0][0]
conv2_block1_2_relu (Activation	(None, 12, 12, 64)	0	conv2_block1_2_bn[0][0]
conv2_block1_0_conv (Conv2D)	(None, 12, 12, 256)	16640	pool1_pool[0][0]
conv2_block1_3_conv (Conv2D)	(None, 12, 12, 256)	16640	conv2_block1_2_relu[0][0]
conv2_block1_0_bn (BatchNormali	(None, 12, 12, 256)	1024	conv2_block1_0_conv[0][0]
conv2_block1_3_bn (BatchNormali	(None, 12, 12, 256)	1024	conv2_block1_3_conv[0][0]
conv2_block1_add (Add)	(None, 12, 12, 256)	0	conv2_block1_0_bn[0][0] conv2_block1_3_bn[0][0]
conv2_block1_out (Activation)	(None, 12, 12, 256)	0	conv2_block1_add[0][0]
conv5_block2_1_conv (Conv2D)	(None, 2, 2, 512)	1049088	conv5_block1_out[0][0]
conv5_block2_1_bn (BatchNormali	(None, 2, 2, 512)	2048	conv5_block2_1_conv[0][0]
conv5_block2_1_relu (Activation	(None, 2, 2, 512)	0	conv5_block2_1_bn[0][0]
conv5_block2_2_conv (Conv2D)	(None, 2, 2, 512)	2359808	conv5_block2_1_relu[0][0]
conv5_block2_2_bn (BatchNormali	(None, 2, 2, 512)	2048	conv5_block2_2_conv[0][0]
conv5_block2_2_relu (Activation	(None, 2, 2, 512)	0	conv5_block2_2_bn[0][0]
conv5_block2_3_conv (Conv2D)	(None, 2, 2, 2048)	1050624	conv5_block2_2_relu[0][0]
conv5_block2_3_bn (BatchNormali	(None, 2, 2, 2048)	8192	conv5_block2_3_conv[0][0]
conv5_block2_add (Add)	(None, 2, 2, 2048)	0	conv5_block1_out[0][0] conv5_block2_3_bn[0][0]
conv5_block2_out (Activation)	(None, 2, 2, 2048)	0	conv5_block2_add[0][0]
conv5_block3_1_conv (Conv2D)	(None, 2, 2, 512)	1049088	conv5_block2_out[0][0]
conv5_block3_1_bn (BatchNormali	(None, 2, 2, 512)	2048	conv5_block3_1_conv[0][0]
conv5_block3_1_relu (Activation	(None, 2, 2, 512)	0	conv5_block3_1_bn[0][0]
conv5_block3_2_conv (Conv2D)	(None, 2, 2, 512)	2359808	conv5_block3_1_relu[0][0]
conv5_block3_2_bn (BatchNormali	(None, 2, 2, 512)	2048	conv5_block3_2_conv[0][0]
conv5_block3_2_relu (Activation	(None, 2, 2, 512)	0	conv5_block3_2_bn[0][0]
conv5_block3_3_conv (Conv2D)	(None, 2, 2, 2048)	1050624	conv5_block3_2_relu[0][0]
conv5_block3_3_bn (BatchNormali	(None, 2, 2, 2048)	8192	conv5_block3_3_conv[0][0]
conv5_block3_add (Add)	(None, 2, 2, 2048)	0	conv5_block2_out[0][0] conv5_block3_3_bn[0][0]
conv5_block3_out (Activation)	(None, 2, 2, 2048)	0	conv5_block3_add[0][0]
avg_pool (GlobalAveragePooling2	(None, 2048)	0	conv5_block3_out[0][0]

Total params: 23,587,712
Trainable params: 0
Non-trainable params: 23,587,712

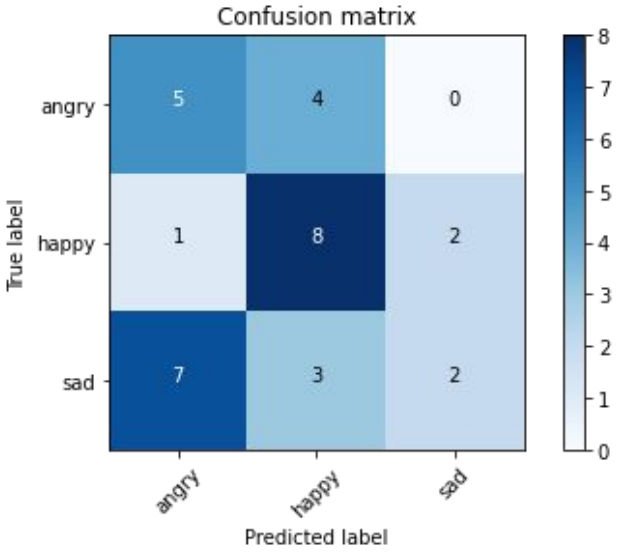
RESNET50

Best Validation Accuracy : 50.0%
with Validation Loss : 1.79





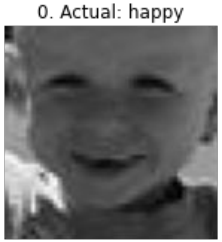
RESNET50



Predicted as Sad

Predicted as Happy

Predicted as Angry



SUMMARY OF PRETRAINED MODELS

01

VGG16

Best accuracy:
63.5%

02

VGGFACE

Best accuracy:
75.5%

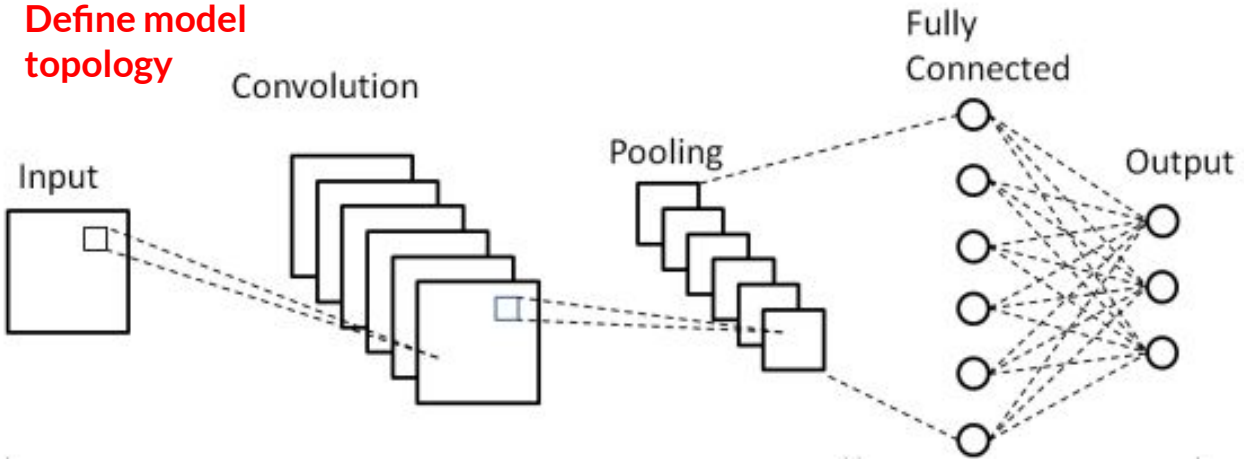
03

RESNET50

Best accuracy:
50.0%

CUSTOM MODELS

1. Define model topology



2. Training

Early Stopping,
Reduce Learning Rate on Plateau

CUSTOM MODEL 1

Layer (type)	Output Shape	Param #
conv_layer_1a (Conv2D)	(None, 46, 46, 64)	1792
conv_layer_1b (Conv2D)	(None, 44, 44, 64)	36928
activation (Activation)	(None, 44, 44, 64)	0
pool_layer_1 (MaxPooling2D)	(None, 22, 22, 64)	0
conv_layer_2a (Conv2D)	(None, 20, 20, 128)	73856
conv_layer_2b (Conv2D)	(None, 18, 18, 128)	147584
activation_1 (Activation)	(None, 18, 18, 128)	0
pool_layer_2 (MaxPooling2D)	(None, 9, 9, 128)	0
conv_layer_3a (Conv2D)	(None, 7, 7, 256)	295168
conv_layer_3b (Conv2D)	(None, 5, 5, 256)	590080
activation_2 (Activation)	(None, 5, 5, 256)	0
pool_layer_3 (MaxPooling2D)	(None, 2, 2, 256)	0
global_average_pooling2d (Gl	(None, 256)	0
dense (Dense)	(None, 128)	32896
dropout (Dropout)	(None, 128)	0
dense_1 (Dense)	(None, 32)	4128
dropout_1 (Dropout)	(None, 32)	0
dense_2 (Dense)	(None, 3)	99

Total params: 1,182,531
Trainable params: 1,182,531
Non-trainable params: 0

1st CNN Block

2nd CNN Block

3rd CNN Block

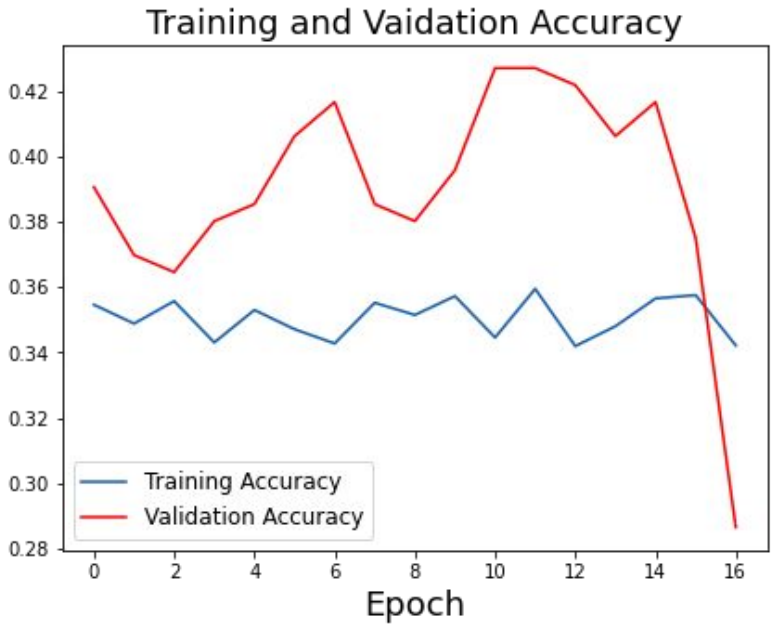
Global Avg Pooling

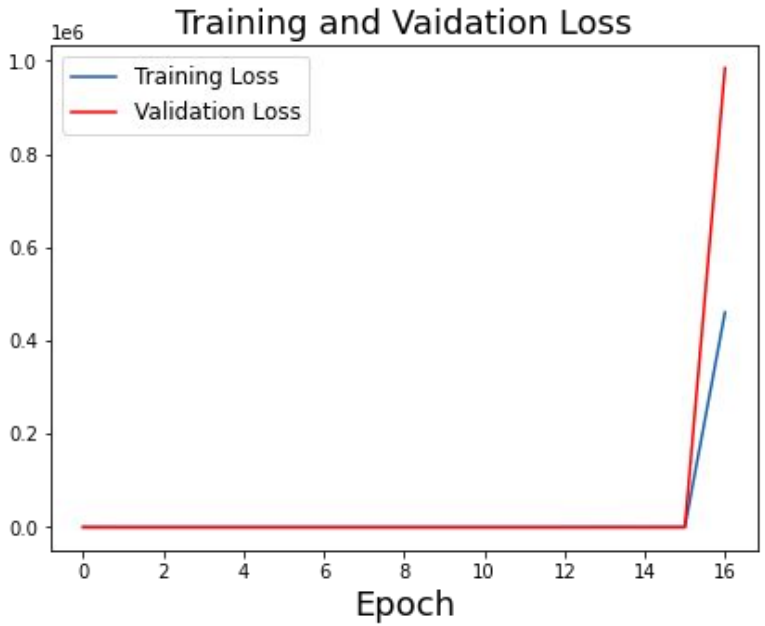
2 layers Fully Connected Network

Prediction Layer

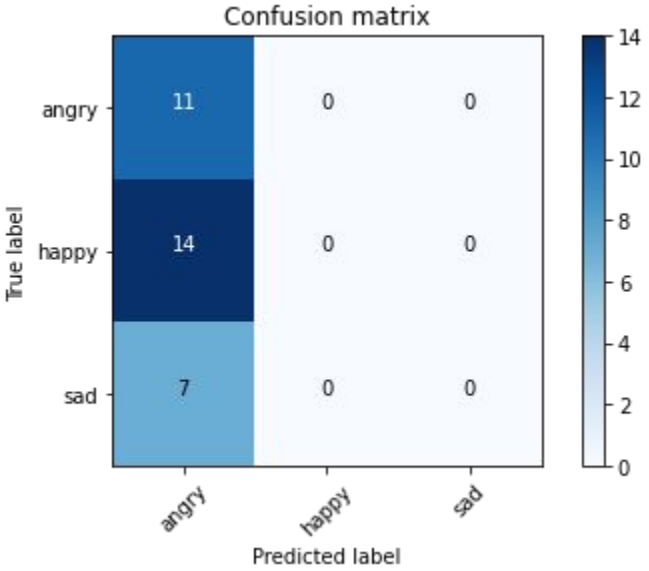
CUSTOM MODEL 1

Best Validation Accuracy : 42.7%
with Validation Loss : 2.09





CUSTOM MODEL 1



Predicted as Angry

CUSTOM MODEL 2

Layer (type)	Output Shape	Param #
conv_layer_1a (Conv2D)	(None, 46, 46, 64)	1792
conv_layer_1b (Conv2D)	(None, 44, 44, 64)	36928
batch_normalization (BatchNo	(None, 44, 44, 64)	256
activation_6 (Activation)	(None, 44, 44, 64)	0
pool_layer_1 (MaxPooling2D)	(None, 22, 22, 64)	0
conv_layer_2a (Conv2D)	(None, 20, 20, 128)	73856
conv_layer_2b (Conv2D)	(None, 18, 18, 128)	147584
batch_normalization_1 (Batch	(None, 18, 18, 128)	512
activation_7 (Activation)	(None, 18, 18, 128)	0
pool_layer_2 (MaxPooling2D)	(None, 9, 9, 128)	0
conv_layer_3a (Conv2D)	(None, 7, 7, 256)	295168
conv_layer_3b (Conv2D)	(None, 5, 5, 256)	590080
batch_normalization_2 (Batch	(None, 5, 5, 256)	1024
activation_8 (Activation)	(None, 5, 5, 256)	0
pool_layer_3 (MaxPooling2D)	(None, 2, 2, 256)	0
global_average_pooling2d_2 (	(None, 256)	0
dense_6 (Dense)	(None, 128)	32896
dropout_4 (Dropout)	(None, 128)	0
dense_7 (Dense)	(None, 32)	4128
dropout_5 (Dropout)	(None, 32)	0
dense_8 (Dense)	(None, 3)	99

Total params: 1,184,323
Trainable params: 1,183,427
Non-trainable params: 896

1st CNN Block

2nd CNN Block

3rd CNN Block

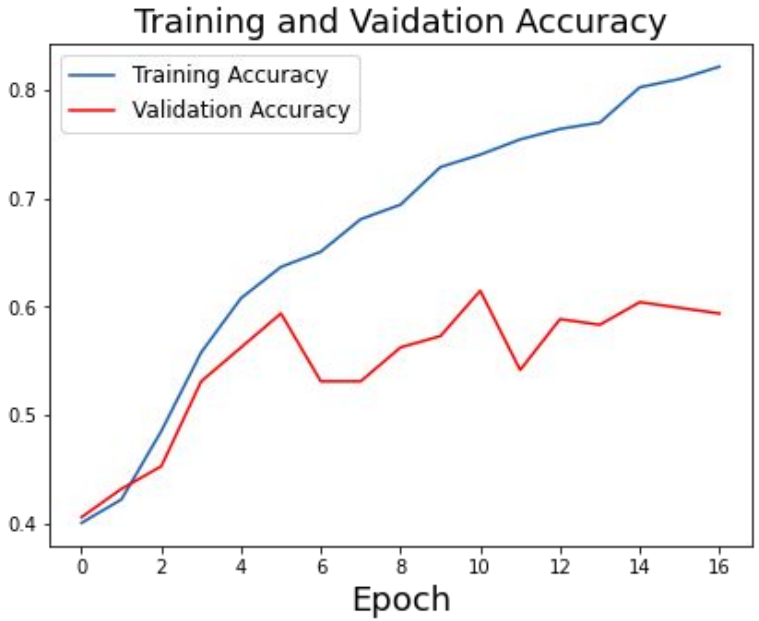
Global Avg Pooling

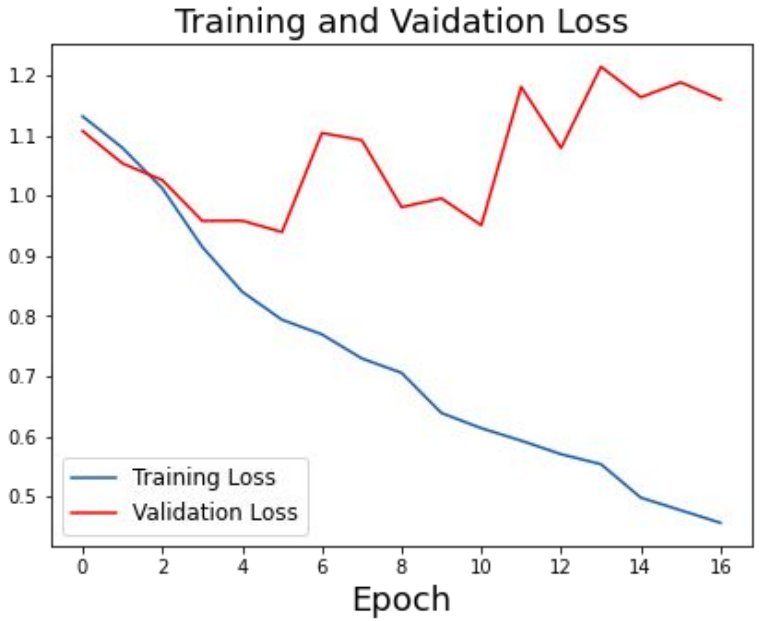
2 layers Fully Connected Network

Prediction Layer

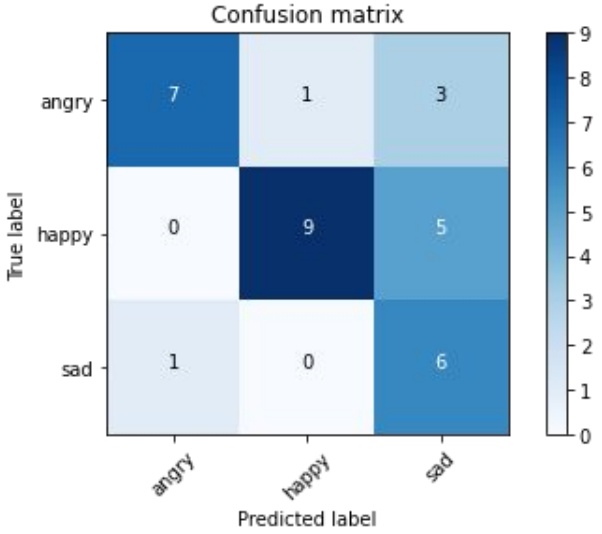
CUSTOM MODEL 2

Best Validation Accuracy : 61.9%
with Validation Loss : 2.09





CUSTOM MODEL 2



Predicted as Happy

Predicted as Angry

Predicted as Sad

HYPERPARAMETER TUNING

Layer (type)	Output Shape	Param #
conv_layer_1a (Conv2D)	(None, 46, 46, 64)	1792
conv_layer_1b (Conv2D)	(None, 44, 44, 64)	36928
batch_normalization (BatchNo	(None, 44, 44, 64)	256
activation_6 (Activation)	(None, 44, 44, 64)	0
pool_layer_1 (MaxPooling2D)	(None, 22, 22, 64)	0
conv_layer_2a (Conv2D)	(None, 20, 20, 128)	73856
conv_layer_2b (Conv2D)	(None, 18, 18, 128)	147584
batch_normalization_1 (Batch	(None, 18, 18, 128)	512
activation_7 (Activation)	(None, 18, 18, 128)	0
pool_layer_2 (MaxPooling2D)	(None, 9, 9, 128)	0
conv_layer_3a (Conv2D)	(None, 7, 7, 256)	295168
conv_layer_3b (Conv2D)	(None, 5, 5, 256)	590080
batch_normalization_2 (Batch	(None, 5, 5, 256)	1024
activation_8 (Activation)	(None, 5, 5, 256)	0
pool_layer_3 (MaxPooling2D)	(None, 2, 2, 256)	0
global_average_pooling2d_2 (	(None, 256)	0
dense_6 (Dense)	(None, 128)	32896
dropout_4 (Dropout)	(None, 128)	0
dense_7 (Dense)	(None, 32)	4128
dropout_5 (Dropout)	(None, 32)	0
dense_8 (Dense)	(None, 3)	99

Total params: 1,184,323
Trainable params: 1,183,427
Non-trainable params: 896

1. 2. 3. 4. 5. 6. 7.

List of hyperparameter to tune:

1. Filter numbers for Convolution block 1
2. Filter numbers for Convolution block 2
3. Filter numbers for Convolution block 3
4. Unit numbers for Dense Layer 1
5. Unit numbers for Dense Layer 2
6. Dropout value 1
7. Dropout value 2

CUSTOM MODEL 3

Layer (type)	Output Shape	Param #
conv_layer_1a (Conv2D)	(None, 46, 46, 32)	896
conv_layer_1b (Conv2D)	(None, 44, 44, 32)	9248
batch_normalization_3 (Batch	(None, 44, 44, 32)	128
activation_3 (Activation)	(None, 44, 44, 32)	0
pool_layer_1 (MaxPooling2D)	(None, 22, 22, 32)	0
conv_layer_2a (Conv2D)	(None, 20, 20, 256)	73984
conv_layer_2b (Conv2D)	(None, 18, 18, 256)	590080
batch_normalization_4 (Batch	(None, 18, 18, 256)	1024
activation_4 (Activation)	(None, 18, 18, 256)	0
pool_layer_2 (MaxPooling2D)	(None, 9, 9, 256)	0
conv_layer_3a (Conv2D)	(None, 7, 7, 256)	590080
conv_layer_3b (Conv2D)	(None, 5, 5, 256)	590080
batch_normalization_5 (Batch	(None, 5, 5, 256)	1024
activation_5 (Activation)	(None, 5, 5, 256)	0
pool_layer_3 (MaxPooling2D)	(None, 2, 2, 256)	0
global_average_pooling2d_1 (	(None, 256)	0
dense_3 (Dense)	(None, 256)	65792
dropout_2 (Dropout)	(None, 256)	0
dense_4 (Dense)	(None, 64)	16448
dropout_3 (Dropout)	(None, 64)	0
dense_5 (Dense)	(None, 3)	195

Total params: 1,938,979
Trainable params: 1,937,891
Non-trainable params: 1,088

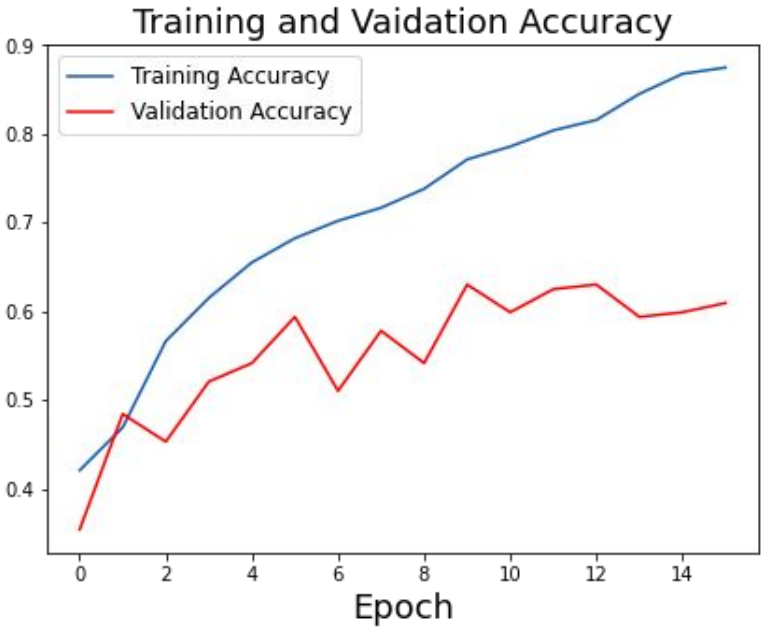
1. 2. 3. 4. 5. 6. 7.

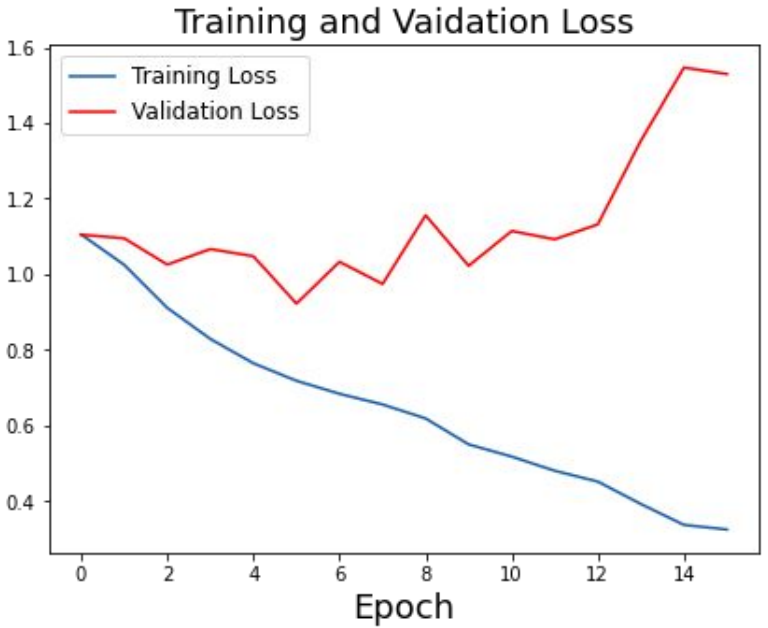
Best hyperparameters:

1. Filter numbers for Convolution block 1 : 32
2. Filter numbers for Convolution block 2 : 256
3. Filter numbers for Convolution block 3 : 256
4. Unit numbers for Dense Layer 1 : 256
5. Unit numbers for Dense Layer 2 : 64
6. Dropout value 1 : 0.25
7. Dropout value 2 : 0.1

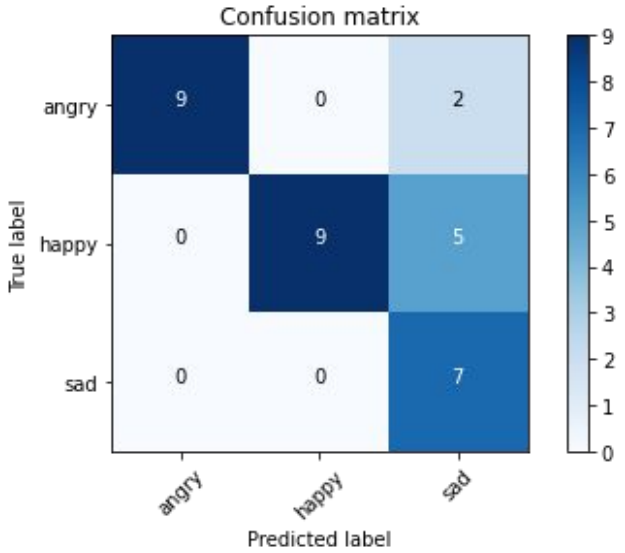
CUSTOM MODEL 3

Best Validation Accuracy : 63.0%
with Validation Loss : 2.09





CUSTOM MODEL 3



Predicted as Sad

SUMMARY OF CNN MODELS

01

VGG16

Best accuracy:
63.5%

02

VGGFACE

Best accuracy:
75.5%

03

RESNET50

Best accuracy:
50.0%

04

CUST. MODEL 1

Best accuracy:
42.7%

05

CUST. MODEL 2

Best accuracy:
61.4%

06

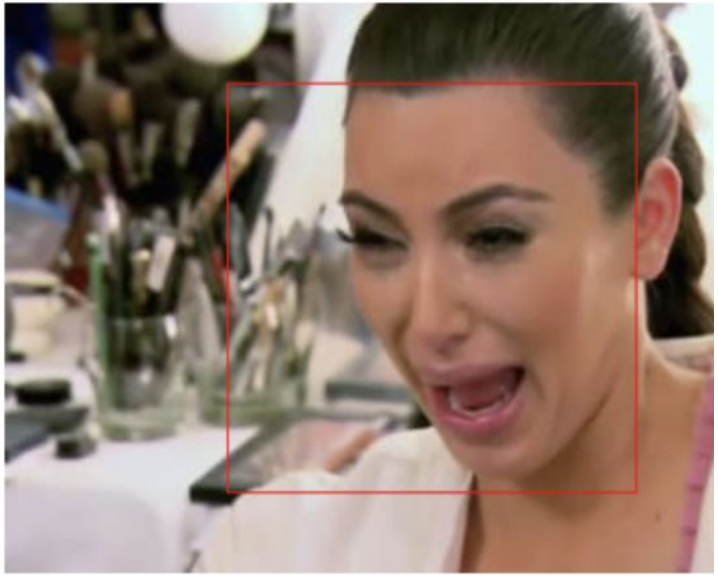
CUST. MODEL 3

Best accuracy:
63.0%

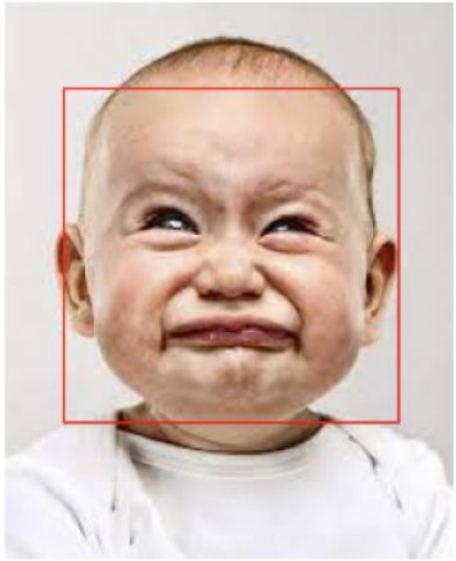
REAL-WORLD PERFORMANCE

SAD FACES

```
VGG16                          > sad    : 68.3%
VGGFace                        > sad    : 71.3%
Custom 2: Using BN             > sad    : 70.7%
Custom 3: After Random Search  > sad    : 69.6%
```

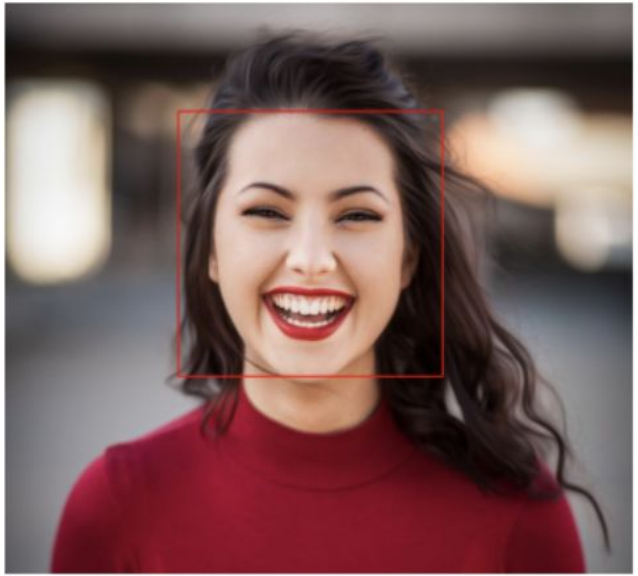


```
VGG16                          > sad    : 66.0%
VGGFace                        > sad    : 97.7%
Custom 2: Using BN             > happy  : 45.6%
Custom 3: After Random Search  > happy  : 75.3%
```



HAPPY FACES

```
VGG16                              > happy : 77.5%
VGGFace                            > happy : 100.0%
Custom 2: Using BN                 > happy : 99.7%
Custom 3: After Random Search      > happy : 99.9%
```



```
VGG16                              > sad   : 68.4%
VGGFace                            > happy : 99.5%
Custom: Using BN                   > happy : 53.4%
Custom: After Random Search        > happy : 89.9%
```

ANGRY FACES

```
VGG16                          > angry : 99.1%
VGGFace                        > angry : 100.0%
Custom: Using BN               > angry : 68.2%
Custom: After Random Search    > angry : 81.7%
```

```
VGG16                          > angry : 87.9%
VGGFace                        > angry : 100.0%
Custom 2: Using BN             > angry : 65.1%
Custom 3: After Random Search  > angry : 96.4%
```

UNDERSTANDING THE MODEL

Layer (type)	Output Shape	Param #
conv_layer_1a (Conv2D)	(None, 46, 46, 64)	1792
conv_layer_1b (Conv2D)	(None, 44, 44, 64)	36928
batch_normalization_3 (Batch	(None, 44, 44, 64)	256
activation_6 (Activation)	(None, 44, 44, 64)	0
pool_layer_1 (MaxPooling2D)	(None, 22, 22, 64)	0
conv_layer_2a (Conv2D)	(None, 20, 20, 128)	73856
conv_layer_2b (Conv2D)	(None, 18, 18, 128)	147584
batch_normalization_4 (Batch	(None, 18, 18, 128)	512
activation_7 (Activation)	(None, 18, 18, 128)	0
pool_layer_2 (MaxPooling2D)	(None, 9, 9, 128)	0
conv_layer_3a (Conv2D)	(None, 7, 7, 256)	295168
conv_layer_3b (Conv2D)	(None, 5, 5, 256)	590080
batch_normalization_5 (Batch	(None, 5, 5, 256)	1024
activation_8 (Activation)	(None, 5, 5, 256)	0
pool_layer_3 (MaxPooling2D)	(None, 2, 2, 256)	0
global_average_pooling2d_14	(None, 256)	0
dropout_9 (Dropout)	(None, 256)	0
dense_3 (Dense)	(None, 128)	32896
dropout_10 (Dropout)	(None, 128)	0
dense_4 (Dense)	(None, 32)	4128
dropout_11 (Dropout)	(None, 32)	0
dense_5 (Dense)	(None, 3)	99

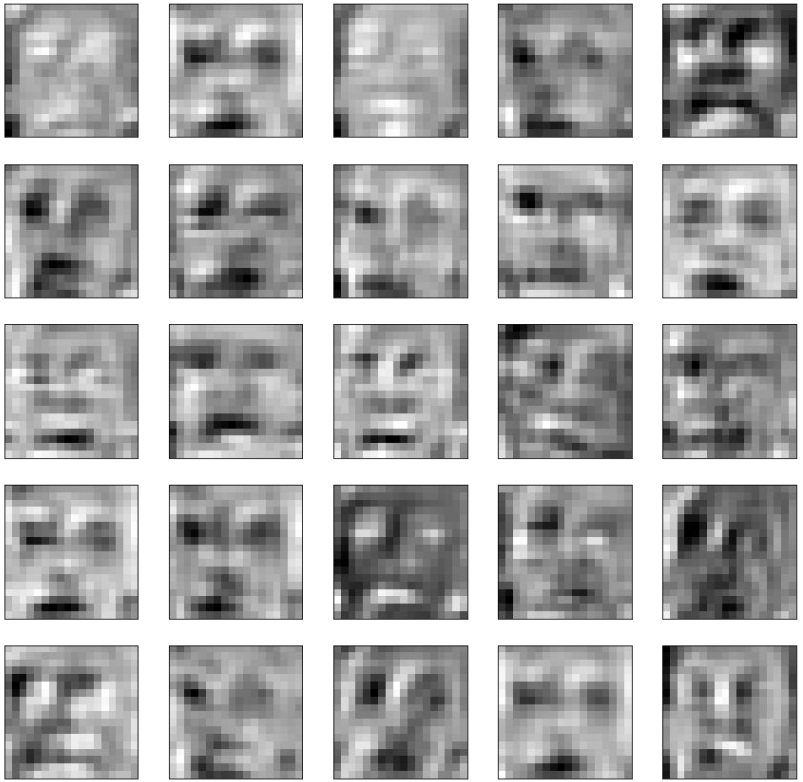
Total params: 1,184,323
Trainable params: 1,183,427
Non-trainable params: 896

CUSTOM MODEL 2: EARLY LAYER

Layer (type)	Output Shape	Param #
conv_layer_1a (Conv2D)	(None, 46, 46, 64)	1792
conv_layer_1b (Conv2D)	(None, 44, 44, 64)	36928
batch_normalization_3 (Batch	(None, 44, 44, 64)	256
activation_6 (Activation)	(None, 44, 44, 64)	0
pool_layer_1 (MaxPooling2D)	(None, 22, 22, 64)	0
conv_layer_2a (Conv2D)	(None, 20, 20, 128)	73856
conv_layer_2b (Conv2D)	(None, 18, 18, 128)	147584
batch_normalization_4 (Batch	(None, 18, 18, 128)	512
activation_7 (Activation)	(None, 18, 18, 128)	0
pool_layer_2 (MaxPooling2D)	(None, 9, 9, 128)	0
conv_layer_3a (Conv2D)	(None, 7, 7, 256)	295168
conv_layer_3b (Conv2D)	(None, 5, 5, 256)	590080
batch_normalization_5 (Batch	(None, 5, 5, 256)	1024
activation_8 (Activation)	(None, 5, 5, 256)	0
pool_layer_3 (MaxPooling2D)	(None, 2, 2, 256)	0
global_average_pooling2d_14	(None, 256)	0
dropout_9 (Dropout)	(None, 256)	0
dense_3 (Dense)	(None, 128)	32896
dropout_10 (Dropout)	(None, 128)	0
dense_4 (Dense)	(None, 32)	4128
dropout_11 (Dropout)	(None, 32)	0
dense_5 (Dense)	(None, 3)	99

Total params: 1,184,323
Trainable params: 1,183,427
Non-trainable params: 896

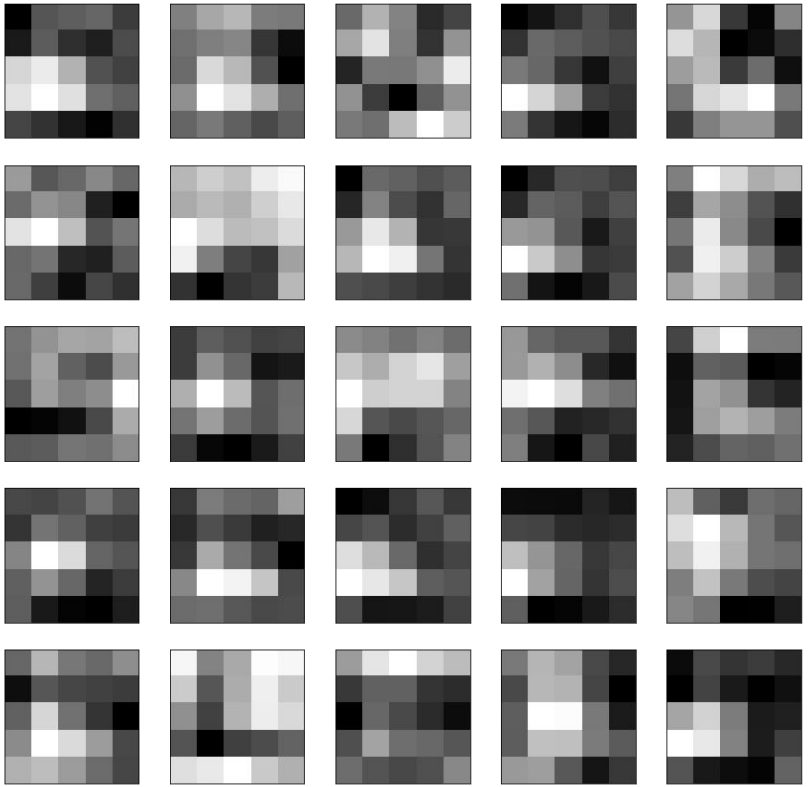
CUSTOM MODEL 2: MIDDLE LAYER



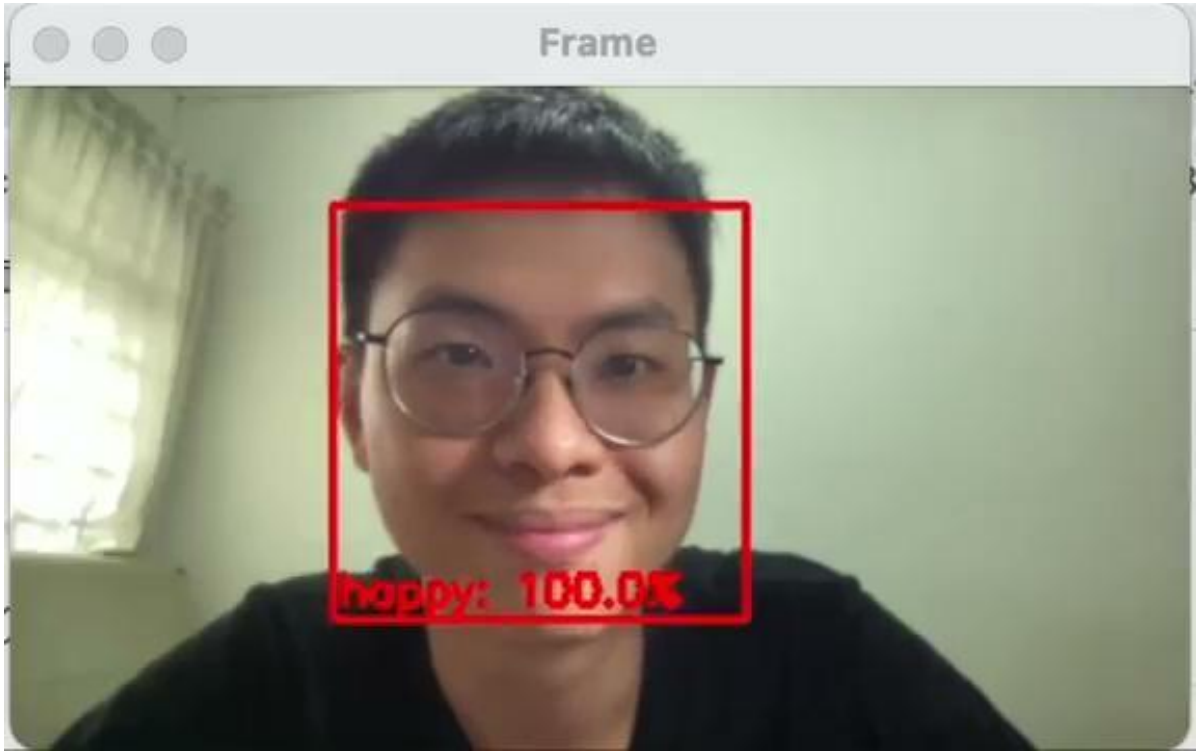
Layer (type)	Output Shape	Param #
conv_layer_1a (Conv2D)	(None, 46, 46, 64)	1792
conv_layer_1b (Conv2D)	(None, 44, 44, 64)	36928
batch_normalization_3 (Batch	(None, 44, 44, 64)	256
activation_6 (Activation)	(None, 44, 44, 64)	0
pool_layer_1 (MaxPooling2D)	(None, 22, 22, 64)	0
conv_layer_2a (Conv2D)	(None, 20, 20, 128)	73856
conv_layer_2b (Conv2D)	(None, 18, 18, 128)	147584
batch_normalization_4 (Batch	(None, 18, 18, 128)	512
activation_7 (Activation)	(None, 18, 18, 128)	0
pool_layer_2 (MaxPooling2D)	(None, 9, 9, 128)	0
conv_layer_3a (Conv2D)	(None, 7, 7, 256)	295168
conv_layer_3b (Conv2D)	(None, 5, 5, 256)	590080
batch_normalization_5 (Batch	(None, 5, 5, 256)	1024
activation_8 (Activation)	(None, 5, 5, 256)	0
pool_layer_3 (MaxPooling2D)	(None, 2, 2, 256)	0
global_average_pooling2d_14	(None, 256)	0
dropout_9 (Dropout)	(None, 256)	0
dense_3 (Dense)	(None, 128)	32896
dropout_10 (Dropout)	(None, 128)	0
dense_4 (Dense)	(None, 32)	4128
dropout_11 (Dropout)	(None, 32)	0
dense_5 (Dense)	(None, 3)	99

Total params: 1,184,323
Trainable params: 1,183,427
Non-trainable params: 896

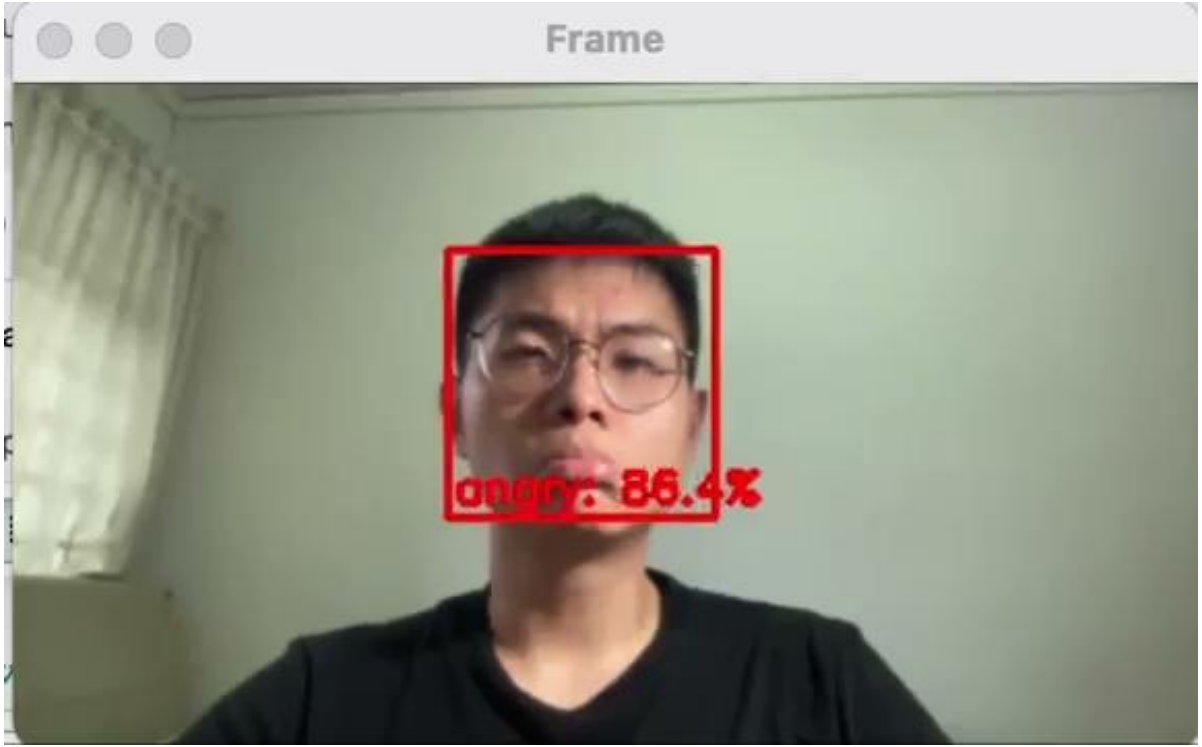
CUSTOM MODEL 2: LAST LAYER



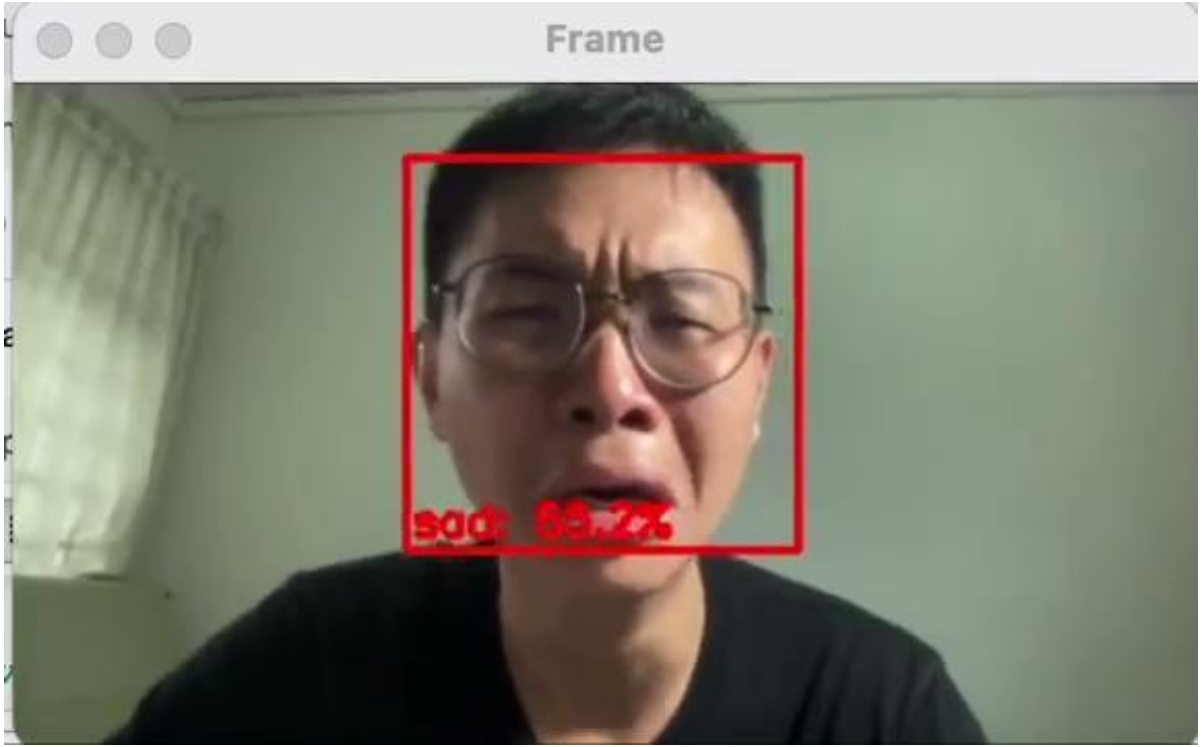
ANGRY-HAPPY



ANGRY-SAD



MIXED EXPRESSIONS



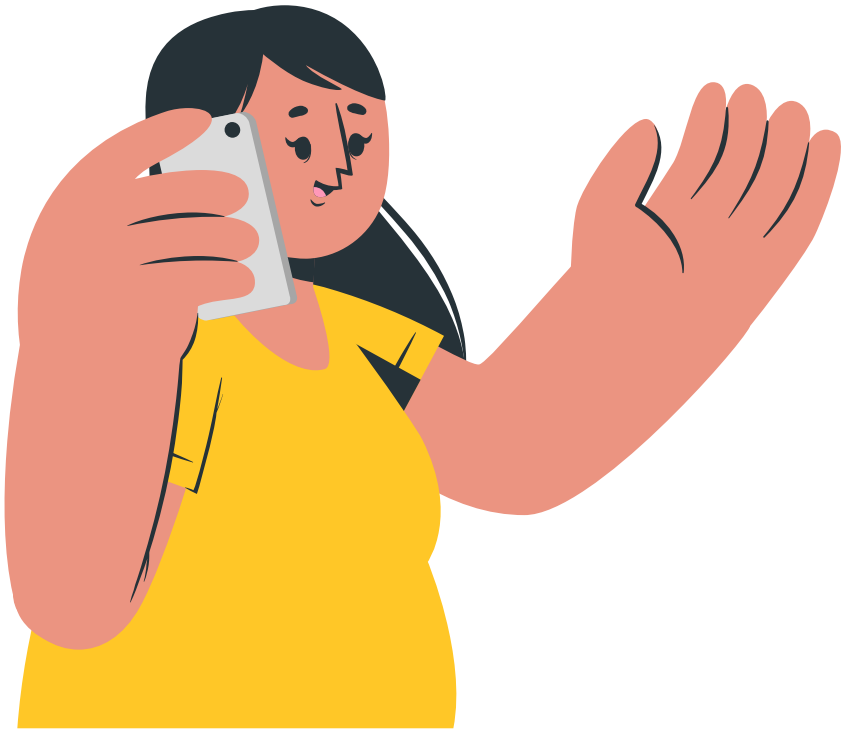
SUMMARY

- Best pre-trained model for emotion detection is VGGFace
  - Validation Accuracy : 75.5%
- Best custom-built model is found after implementing Batch Normalization and Fine-tuning
  - Validation Accuracy : 63.0%
- The model able to capture facial landmarks such as mouth, eyes, and eyebrows shape and location also head positions

RECOMMENDATIONS

- Train with better and clearer dataset
- Expand limit to more emotions to be more useful
  - Include complex ones
- Trained on masked images to accommodate on the current pandemic situation

THANKS!



DOES ANYONE HAVE ANY QUESTIONS?

CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, and infographics & images by **Freepik** and illustrations by **Stories**