

# GNR 638 MINI PROJECT - 1 REPORT

## Team:

Shubham Jadhav (200040067)

Raghav Rander (200040113)

Yash Kumar (200040168)

## Dataset used:

[CUB 200 2011](#) dataset

## Model Architecture:

Base Model used is the MobileNetV2 Model.

About MobileNetV2: The architecture of MobileNetV2 consists of a series of convolutional layers, followed by depthwise separable convolutions, inverted residuals, bottleneck design, linear bottlenecks, and squeeze-and-excitation (SE) blocks.

Input	Operator	$t$	$c$	$n$	$s$
$224^2 \times 3$	conv2d	-	32	1	2
$112^2 \times 32$	bottleneck	1	16	1	1
$112^2 \times 16$	bottleneck	6	24	2	2
$56^2 \times 24$	bottleneck	6	32	3	2
$28^2 \times 32$	bottleneck	6	64	4	2
$14^2 \times 64$	bottleneck	6	96	3	1
$14^2 \times 96$	bottleneck	6	160	3	2
$7^2 \times 160$	bottleneck	6	320	1	1
$7^2 \times 320$	conv2d 1x1	-	1280	1	1
$7^2 \times 1280$	avgpool 7x7	-	-	1	-
$1 \times 1 \times 1280$	conv2d 1x1	-	k	-	-

MobileNetV2 Overall Architecture

Reason to use MobileNetV2: It is a very lightweight model and given the constraint of 10 million parameters, MobileNetv2 is one of the best choices with respect to speed of training and accuracy.

We tried other models such as EfficientNet, DenseNet, MobileNetV3 but none of them gave such good performance as MobileNetV2.

Additional things added: A Global Average Pooling (GAP) layer has been added followed by a dropout to curb overfitting. We have left 10 layers unfrozen from the pre-trained model to fine-tune them on our dataset.

### Overall Model Architecture:

Name	Type	Shape	Param
input_4	InputLayer	[(None, 224, 224, 3)]	0
Conv1	Conv2D	(None, 112, 112, 32)	864
bn_Conv1	BatchNormalization	(None, 112, 112, 32)	128
Conv1_relu	ReLU	(None, 112, 112, 32)	0
expanded_conv_depthwise	DepthwiseConv2D	(None, 112, 112, 32)	288
expanded_conv_depthwise_BN	BatchNormalization	(None, 112, 112, 32)	128
expanded_conv_depthwise_relu	ReLU	(None, 112, 112, 32)	0
expanded_conv_project	Conv2D	(None, 112, 112, 16)	512
expanded_conv_project_BN	BatchNormalization	(None, 112, 112, 16)	64
block_1_expand	Conv2D	(None, 112, 112, 96)	1536
block_1_expand_BN	BatchNormalization	(None, 112, 112, 96)	384

block_1_expand_relu	ReLU	(None, 112, 112, 96)	0
block_1_pad	ZeroPadding2D	(None, 113, 113, 96)	0
block_1_depthwise	DepthwiseConv2D	(None, 56, 56, 96)	864
block_1_depthwise_BN	BatchNormalization	(None, 56, 56, 96)	384
block_1_depthwise_relu	ReLU	(None, 56, 56, 96)	0
block_1_project	Conv2D	(None, 56, 56, 24)	2304
block_1_project_BN	BatchNormalization	(None, 56, 56, 24)	96
block_2_expand	Conv2D	(None, 56, 56, 144)	3456
block_2_expand_BN	BatchNormalization	(None, 56, 56, 144)	576
block_2_expand_relu	ReLU	(None, 56, 56, 144)	0
block_2_depthwise	DepthwiseConv2D	(None, 56, 56, 144)	1296
block_2_depthwise_BN	BatchNormalization	(None, 56, 56, 144)	576
block_2_depthwise_relu	ReLU	(None, 56, 56, 144)	0
block_2_project	Conv2D	(None, 56, 56, 24)	3456
block_2_project_BN	BatchNormalization	(None, 56, 56, 24)	96
block_2_add	Add	(None, 56, 56, 24)	0
block_3_expand	Conv2D	(None, 56, 56, 144)	3456
block_3_expand_BN	BatchNormalization	(None, 56, 56, 144)	576
block_3_expand_relu	ReLU	(None, 56, 56, 144)	0
block_3_pad	ZeroPadding2D	(None, 57, 57, 144)	0
block_3_depthwise	DepthwiseConv2D	(None, 28, 28, 144)	1296

block_3_depthwise_BN	BatchNormalization	(None, 28, 28, 144)	576
block_3_depthwise_relu	ReLU	(None, 28, 28, 144)	0
block_3_project	Conv2D	(None, 28, 28, 32)	4608
block_3_project_BN	BatchNormalization	(None, 28, 28, 32)	128
block_4_expand	Conv2D	(None, 28, 28, 192)	6144
block_4_expand_BN	BatchNormalization	(None, 28, 28, 192)	768
block_4_expand_relu	ReLU	(None, 28, 28, 192)	0
block_4_depthwise	DepthwiseConv2D	(None, 28, 28, 192)	1728
block_4_depthwise_BN	BatchNormalization	(None, 28, 28, 192)	768
block_4_depthwise_relu	ReLU	(None, 28, 28, 192)	0
block_4_project	Conv2D	(None, 28, 28, 32)	6144
block_4_project_BN	BatchNormalization	(None, 28, 28, 32)	128
block_4_add	Add	(None, 28, 28, 32)	0
block_5_expand	Conv2D	(None, 28, 28, 192)	6144
block_5_expand_BN	BatchNormalization	(None, 28, 28, 192)	768
block_5_expand_relu	ReLU	(None, 28, 28, 192)	0
block_5_depthwise	DepthwiseConv2D	(None, 28, 28, 192)	1728
block_5_depthwise_BN	BatchNormalization	(None, 28, 28, 192)	768
block_5_depthwise_relu	ReLU	(None, 28, 28, 192)	0
block_5_project	Conv2D	(None, 28, 28, 32)	6144
block_5_project_BN	BatchNormalization	(None, 28, 28, 32)	128

block_5_add	Add	(None, 28, 28, 32)	0
block_6_expand	Conv2D	(None, 28, 28, 192)	6144
block_6_expand_BN	BatchNormalization	(None, 28, 28, 192)	768
block_6_expand_relu	ReLU	(None, 28, 28, 192)	0
block_6_pad	ZeroPadding2D	(None, 29, 29, 192)	0
block_6_depthwise	DepthwiseConv2D	(None, 14, 14, 192)	1728
block_6_depthwise_BN	BatchNormalization	(None, 14, 14, 192)	768
block_6_depthwise_relu	ReLU	(None, 14, 14, 192)	0
block_6_project	Conv2D	(None, 14, 14, 64)	12288
block_6_project_BN	BatchNormalization	(None, 14, 14, 64)	256
block_7_expand	Conv2D	(None, 14, 14, 384)	24576
block_7_expand_BN	BatchNormalization	(None, 14, 14, 384)	1536
block_7_expand_relu	ReLU	(None, 14, 14, 384)	0
block_7_depthwise	DepthwiseConv2D	(None, 14, 14, 384)	3456
block_7_depthwise_BN	BatchNormalization	(None, 14, 14, 384)	1536
block_7_depthwise_relu	ReLU	(None, 14, 14, 384)	0
block_7_project	Conv2D	(None, 14, 14, 64)	24576
block_7_project_BN	BatchNormalization	(None, 14, 14, 64)	256
block_7_add	Add	(None, 14, 14, 64)	0
block_8_expand	Conv2D	(None, 14, 14, 384)	24576
block_8_expand_BN	BatchNormalization	(None, 14, 14, 384)	1536

block_8_expand_relu	ReLU	(None, 14, 14, 384)	0
block_8_depthwise	DepthwiseConv2D	(None, 14, 14, 384)	3456
block_8_depthwise_BN	BatchNormalization	(None, 14, 14, 384)	1536
block_8_depthwise_relu	ReLU	(None, 14, 14, 384)	0
block_8_project	Conv2D	(None, 14, 14, 64)	24576
block_8_project_BN	BatchNormalization	(None, 14, 14, 64)	256
block_8_add	Add	(None, 14, 14, 64)	0
block_9_expand	Conv2D	(None, 14, 14, 384)	24576
block_9_expand_BN	BatchNormalization	(None, 14, 14, 384)	1536
block_9_expand_relu	ReLU	(None, 14, 14, 384)	0
block_9_depthwise	DepthwiseConv2D	(None, 14, 14, 384)	3456
block_9_depthwise_BN	BatchNormalization	(None, 14, 14, 384)	1536
block_9_depthwise_relu	ReLU	(None, 14, 14, 384)	0
block_9_project	Conv2D	(None, 14, 14, 64)	24576
block_9_project_BN	BatchNormalization	(None, 14, 14, 64)	256
block_9_add	Add	(None, 14, 14, 64)	0
block_10_expand	Conv2D	(None, 14, 14, 384)	24576
block_10_expand_BN	BatchNormalization	(None, 14, 14, 384)	1536
block_10_expand_relu	ReLU	(None, 14, 14, 384)	0
block_10_depthwise	DepthwiseConv2D	(None, 14, 14, 384)	3456
block_10_depthwise_BN	BatchNormalization	(None, 14, 14, 384)	1536

block_10_depthwise_relu	ReLU	(None, 14, 14, 384)	0
block_10_project	Conv2D	(None, 14, 14, 96)	36864
block_10_project_BN	BatchNormalization	(None, 14, 14, 96)	384
block_11_expand	Conv2D	(None, 14, 14, 576)	55296
block_11_expand_BN	BatchNormalization	(None, 14, 14, 576)	2304
block_11_expand_relu	ReLU	(None, 14, 14, 576)	0
block_11_depthwise	DepthwiseConv2D	(None, 14, 14, 576)	5184
block_11_depthwise_BN	BatchNormalization	(None, 14, 14, 576)	2304
block_11_depthwise_relu	ReLU	(None, 14, 14, 576)	0
block_11_project	Conv2D	(None, 14, 14, 96)	55296
block_11_project_BN	BatchNormalization	(None, 14, 14, 96)	384
block_11_add	Add	(None, 14, 14, 96)	0
block_12_expand	Conv2D	(None, 14, 14, 576)	55296
block_12_expand_BN	BatchNormalization	(None, 14, 14, 576)	2304
block_12_expand_relu	ReLU	(None, 14, 14, 576)	0
block_12_depthwise	DepthwiseConv2D	(None, 14, 14, 576)	5184
block_12_depthwise_BN	BatchNormalization	(None, 14, 14, 576)	2304
block_12_depthwise_relu	ReLU	(None, 14, 14, 576)	0
block_12_project	Conv2D	(None, 14, 14, 96)	55296
block_12_project_BN	BatchNormalization	(None, 14, 14, 96)	384
block_12_add	Add	(None, 14, 14, 96)	0

block_13_expand	Conv2D	(None, 14, 14, 576)	55296
block_13_expand_BN	BatchNormalization	(None, 14, 14, 576)	2304
block_13_expand_relu	ReLU	(None, 14, 14, 576)	0
block_13_pad	ZeroPadding2D	(None, 15, 15, 576)	0
block_13_depthwise	DepthwiseConv2D	(None, 7, 7, 576)	5184
block_13_depthwise_BN	BatchNormalization	(None, 7, 7, 576)	2304
block_13_depthwise_relu	ReLU	(None, 7, 7, 576)	0
block_13_project	Conv2D	(None, 7, 7, 160)	92160
block_13_project_BN	BatchNormalization	(None, 7, 7, 160)	640
block_14_expand	Conv2D	(None, 7, 7, 960)	153600
block_14_expand_BN	BatchNormalization	(None, 7, 7, 960)	3840
block_14_expand_relu	ReLU	(None, 7, 7, 960)	0
block_14_depthwise	DepthwiseConv2D	(None, 7, 7, 960)	8640
block_14_depthwise_BN	BatchNormalization	(None, 7, 7, 960)	3840
block_14_depthwise_relu	ReLU	(None, 7, 7, 960)	0
block_14_project	Conv2D	(None, 7, 7, 160)	153600
block_14_project_BN	BatchNormalization	(None, 7, 7, 160)	640
block_14_add	Add	(None, 7, 7, 160)	0
block_15_expand	Conv2D	(None, 7, 7, 960)	153600
block_15_expand_BN	BatchNormalization	(None, 7, 7, 960)	3840
block_15_expand_relu	ReLU	(None, 7, 7, 960)	0
block_15_depthwise	DepthwiseConv2D	(None, 7, 7, 960)	8640
block_15_depthwise_BN	BatchNormalization	(None, 7, 7, 960)	3840
block_15_depthwise_relu	ReLU	(None, 7, 7, 960)	0
block_15_project	Conv2D	(None, 7, 7, 160)	153600
block_15_project_BN	BatchNormalization	(None, 7, 7, 160)	640



block_15_add	Add	(None, 7, 7, 160)	0
block_16_expand	Conv2D	(None, 7, 7, 960)	153600
block_16_expand_BN	BatchNormalization	(None, 7, 7, 960)	3840
block_16_expand_relu	ReLU	(None, 7, 7, 960)	0
block_16_depthwise	DepthwiseConv2D	(None, 7, 7, 960)	8640
block_16_depthwise_BN	BatchNormalization	(None, 7, 7, 960)	3840
block_16_depthwise_relu	ReLU	(None, 7, 7, 960)	0
block_16_project	Conv2D	(None, 7, 7, 320)	307200
block_16_project_BN	BatchNormalization	(None, 7, 7, 320)	1280
Conv_1	Conv2D	(None, 7, 7, 1280)	409600
Conv_1_bn	BatchNormalization	(None, 7, 7, 1280)	5120
out_relu	ReLU	(None, 7, 7, 1280)	0
global_average_pooling2d_3	GlobalAveragePooling2D	(None, 1280)	0
dropout_3	Dropout	(None, 1280)	0
dense_3	Dense	(None, 200)	256200

**Total params: 2,514,184**

**Trainable params: 988,680**

**Non-trainable params: 1,525,504**

Hyperparameters:

Batch\_size = 64

Dropout = 0.45

Epochs = 100

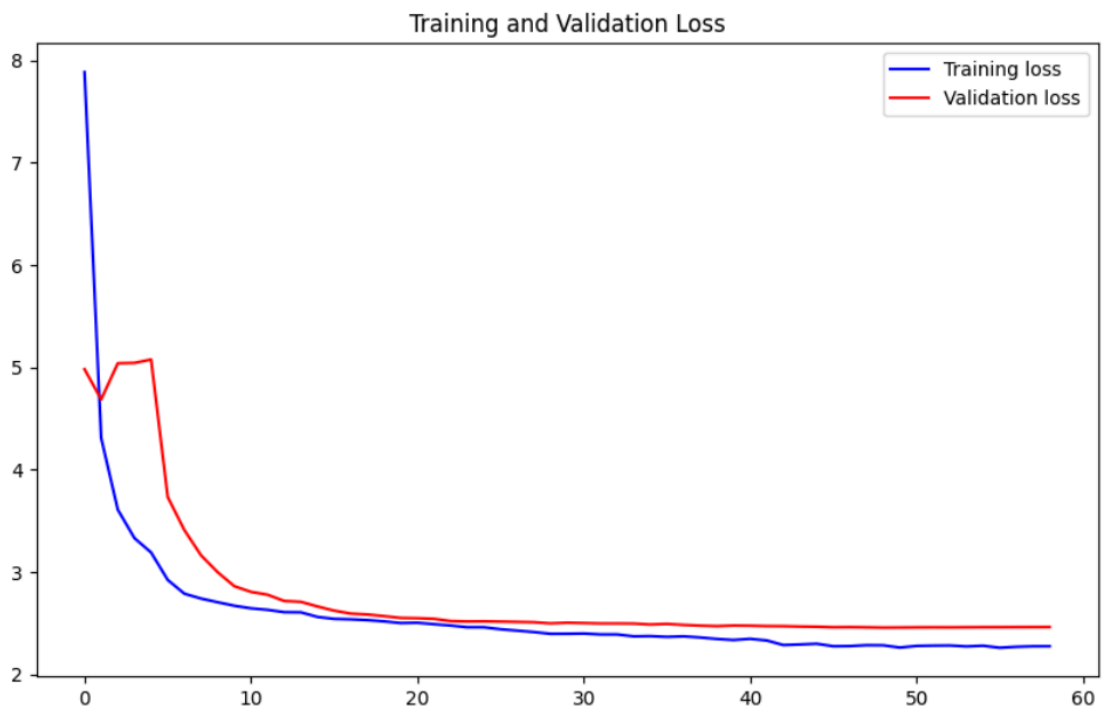
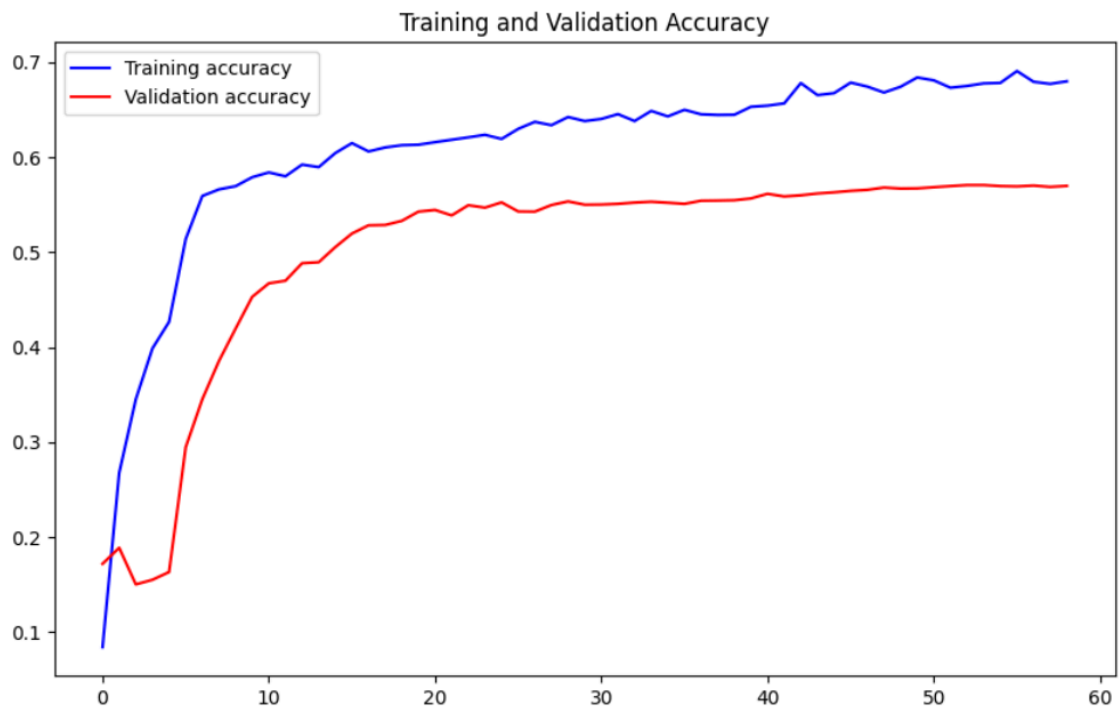
steps/epoch = (5994 // batch\_size)

validation\_steps = (5794 // batch\_size)

Min\_lr = 1e-7

Optimizer used is adam along with loss function as categorical crossentropy. Also, we have applied the early stopping criteria.

**Training:**



From the above curves, we can see that the training and validation accuracy increase and subsequently the losses decrease with each epoch.

The training stops at epoch 59 as it reaches the threshold of  $\text{min\_lr} = 1\text{e-}7$

Thus after the complete training:

Training Accuracy = 67.99%

Validation Accuracy = 56.98%

Training Loss = 2.2752

Validation Loss = 2.4627

Approximate time: ~2.5min per epoch (on GPU). Total time required ~ 2.5 hours

Considering the high number of classes, the time and accuracy trade-off and the lightweight MobileNetV2 model, these accuracy values can be considered to be satisfactory.

## **Appendix:**

The following are the files in the zip folder:

bird\_model.h5 - Model checkpoint

gnr638-2.ipynb - Contains the model training

training\_history.pkl and training\_history.csv - Contains the logs for training

train\_test.py - Contains script to split dataset into train and test

In case of any weight\_decay error while loading the model, please load and compile the model separately as follows:

```
# Load the model without compile to avoid the optimizer error
```

```
model = load_model('Bird_model.h5', compile=False)
```

```
# Manually compile the model
```

```
model.compile(optimizer='adam', loss='categorical_crossentropy',  
metrics=['accuracy'])
```

