COL775
Assignment 1

Sarthak | 2020CS10379

March 2024

Drive Link containing the trained models : Drive

1   Part A

1.1   ResNet

In order to create the ResNet model, I first created the fundamental Residual Block. It applies Convolution, BatchNormalization, ReLu twice upon a forward pass. Twice because one residual block comprises of two individual blocks in our architecture. Also, it adds the downsample after the second application of Batch Norm.

After this, I defined the ResNet class which takes n and r as input and constructs the Network defined in the architecture by stacking the Residual Blocks together. The whole Network is divided into 3 layers, the first one having a stride of 1 while the other 2 having stride of 2 (Implying downsampling). Also, the number of filters in these 3 layers are 16,32 and 64 respectively. Finally, there is a Fully Connected Block of size r.
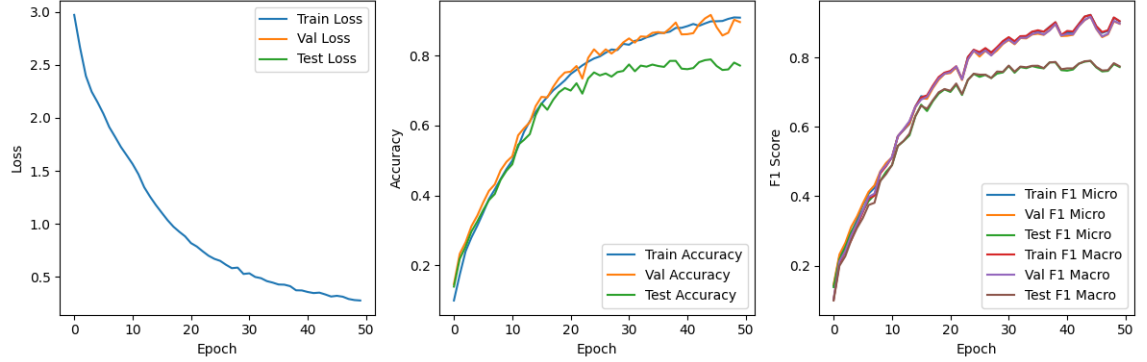
The accuracy and loss as a function of number of epochs for SGD is shown below :
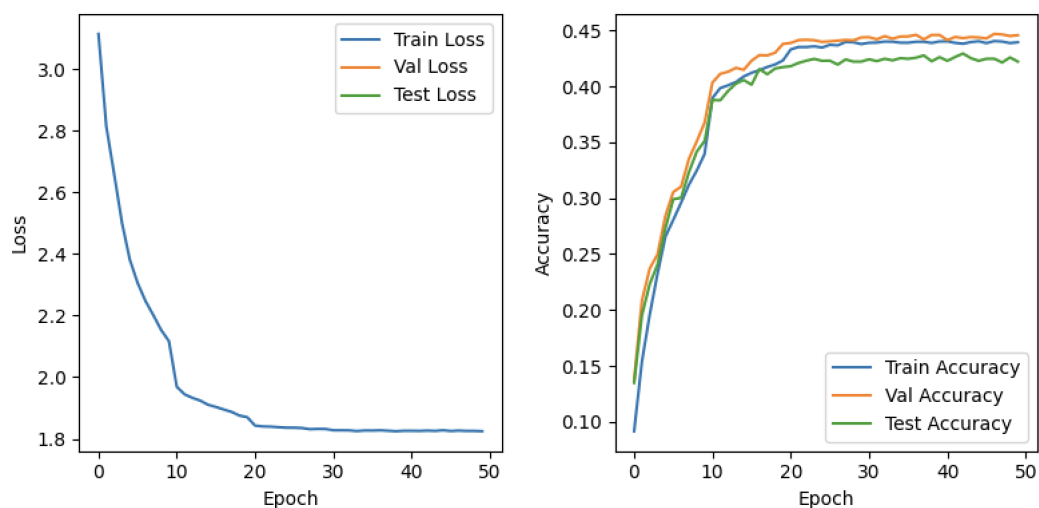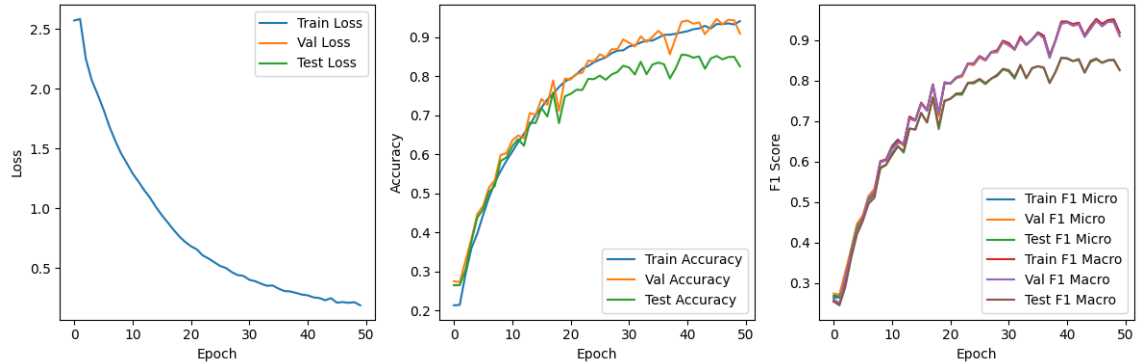
Since the accuracy was very low, hence we use Adam Optimizer instead of SGD, and also tune the Hyperparameters to have a learning rate of 10^{-3}. With such modifications, the Accuracy, Loss and F1 statistics with respect to Number of Epoch is given below :
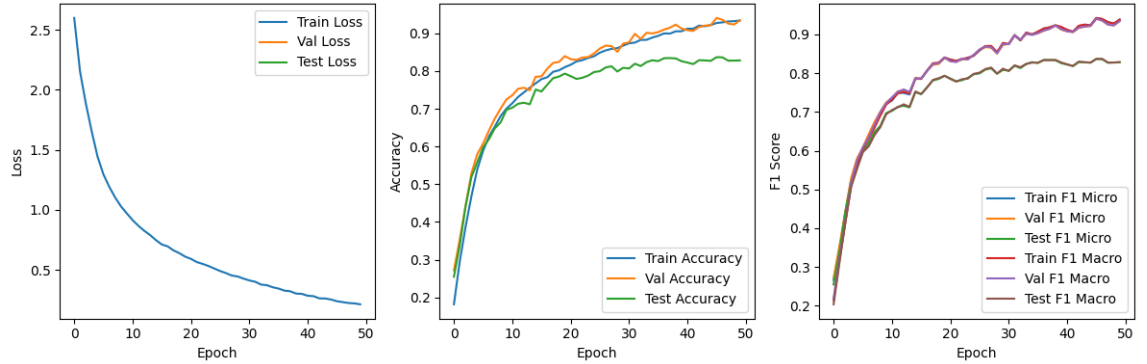

2
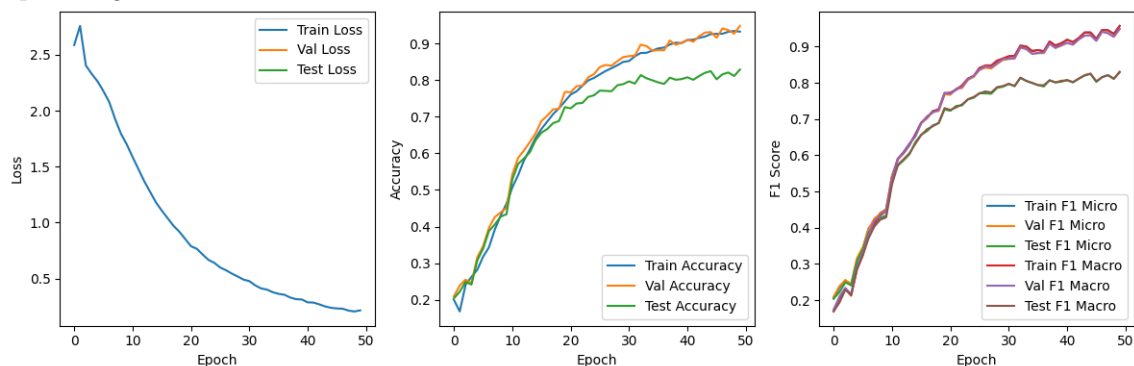
1.2   ResNet with No Normalization
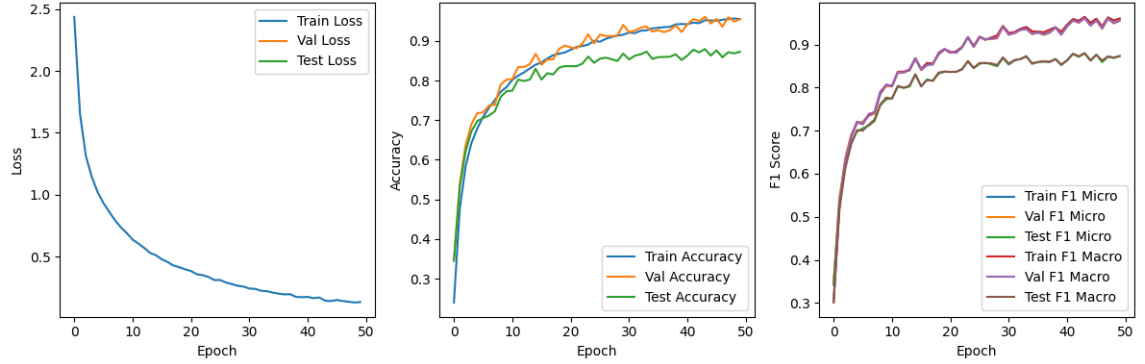

1.3   ResNet with Custom Batch Normalization

We can clearly see that the Custom Batch Normalization is very much similar to the Normal ResNet with inbuilt Batch Normalization (**1.1**). Hence, we can conclude that they are similar.
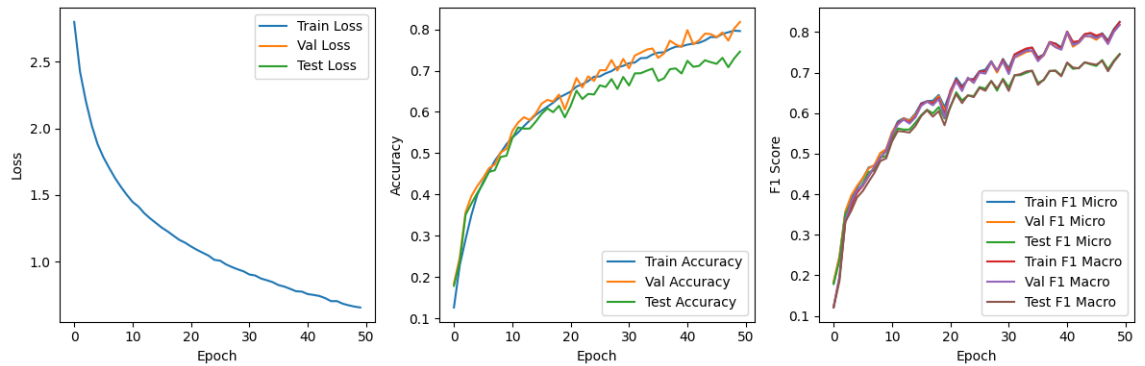
1.4   ResNet with Instance Normalization
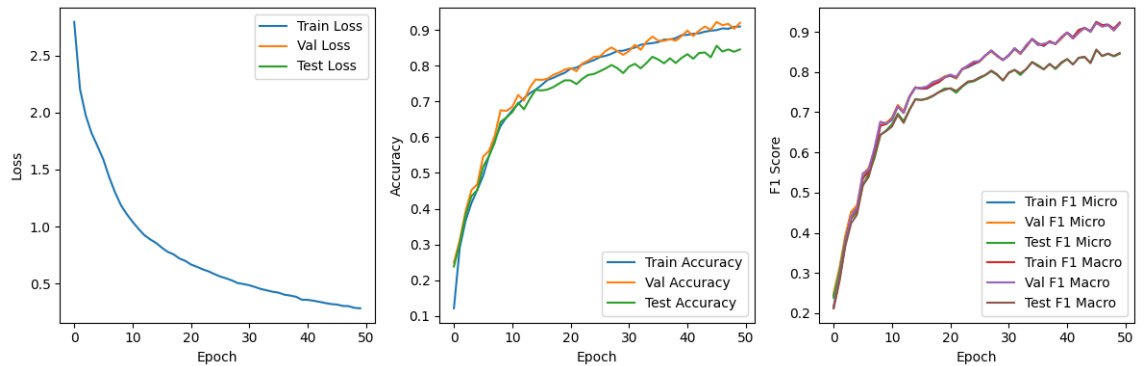
1.5 ResNet with Batch-Instance Normalization

1.6 ResNet with Layer Normalization
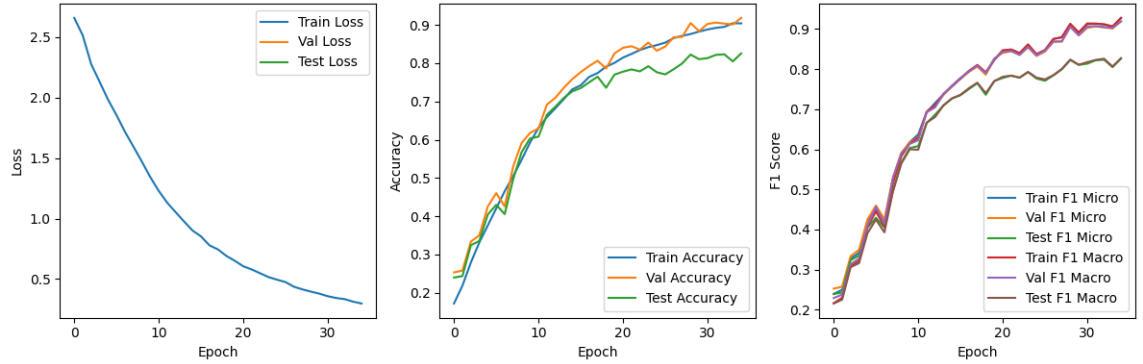

1.7 ResNet with Group Normalization

For Group size of 4 :

The ResNet with Group Normalization is one of the most robust and efficient model leading to more than 80% of testing accuracy.
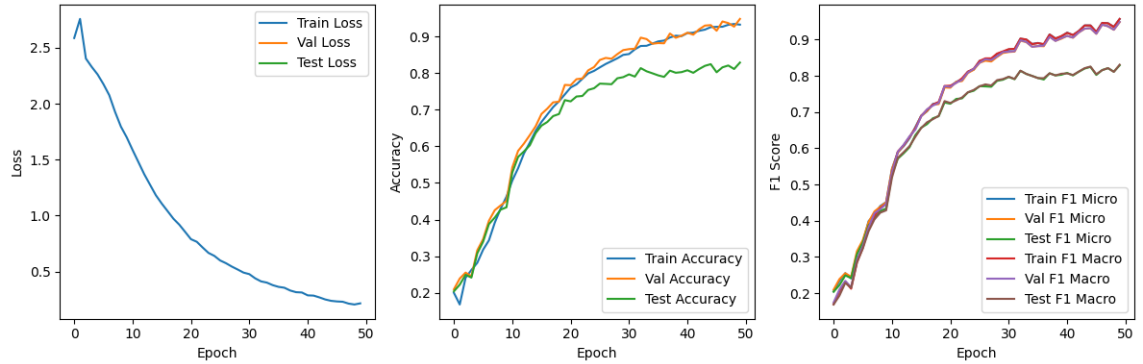
1.8   Batch Size Comparison

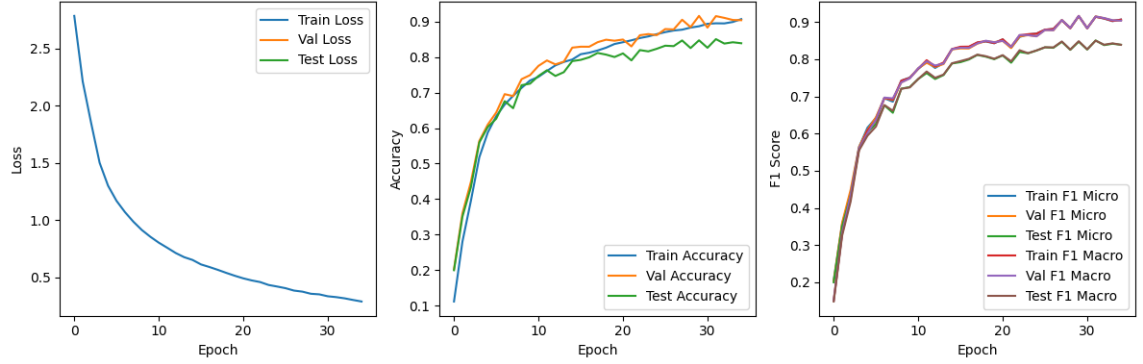### 1.8.1   Batch Normalization with Batch size 8

Clearly batch size of 8 performs much better than a batch size of 32 or 128, since
it reaches better accuracy with only around 35 epochs. But it also takes more
time to train such a model with smaller batch size.

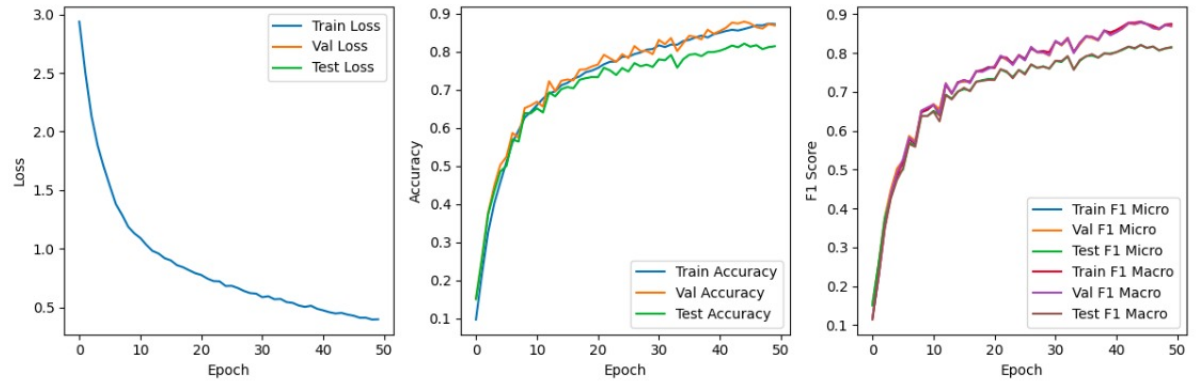### 1.8.2   Batch Normalization with Batch Size 128

A model with a larger batch size employing batch normalization performs worse
than one with smaller batch size as can be see with a relatively poorer testing
accuracy.

1.8.3 Group Normalization with Batch Size 8

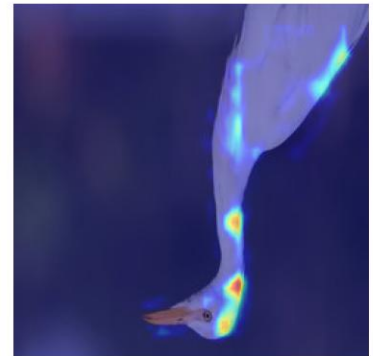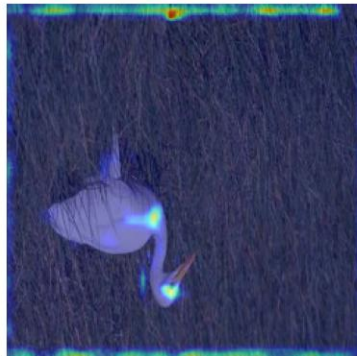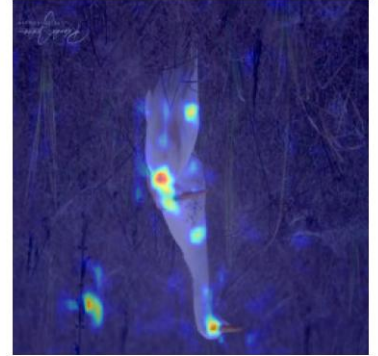### 1.8.4 Group Normalization with Batch Size 128

Here, we are able to visualise very insignificant differences between batch sizes of 8 and 128 in Group Normalization since it normalizes the data by grouping channels together rather than batches of inputs.
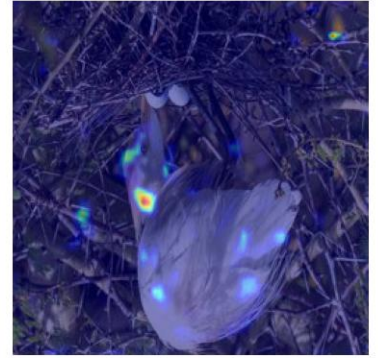
1.9 GradCAM

1.10 Cattle Egret

1.10.1 Correct

From the heatmap, it is evident that the neck, eyes and head are significant for the prediction of Cattle Egret as they are distinguished by their long necks, round heads and white body.
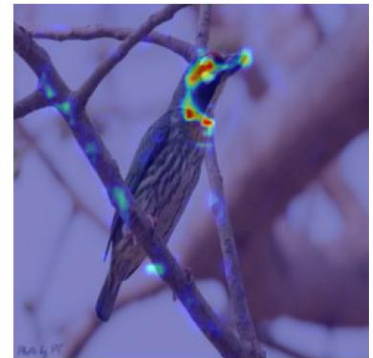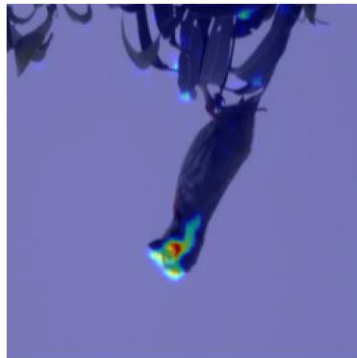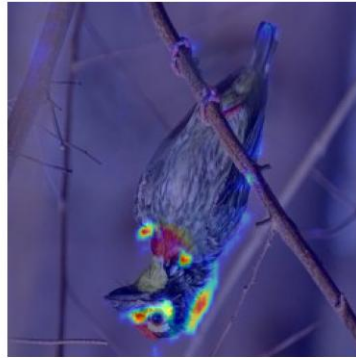
1.10.2 Incorrect

In all the incorrect predictions, none of the heatmap is around the head or the nect. In some of the images it is near the claws hence leading to failure of recognition.
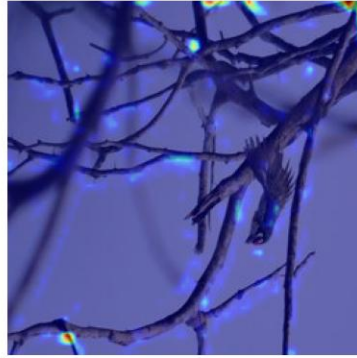
1.11 Coppersmith Barbet

1.11.1 Correct

From all the correct predictions, it can be inferred that the face of Coppersmith Barbet is significant for predictions as all heatmaps contain the face.
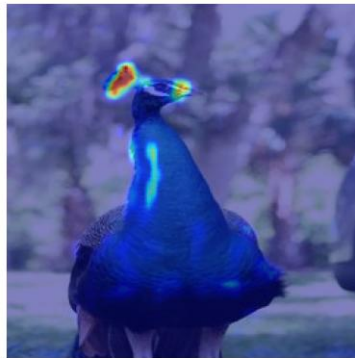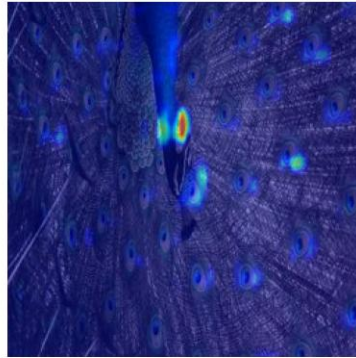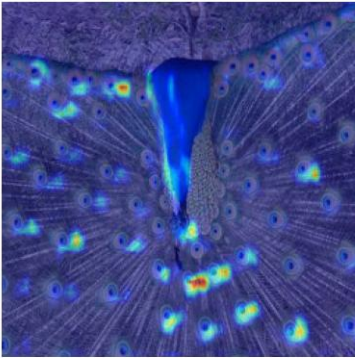
1.11.2 Incorrect

Again we can visualise that the incorrect predictions are due to not properly recognising one of the most important aspect of the bird which is face. Also, in most of these failed images, it is evident that the bird is far away and hence the face is either not visible or not distinguishable.
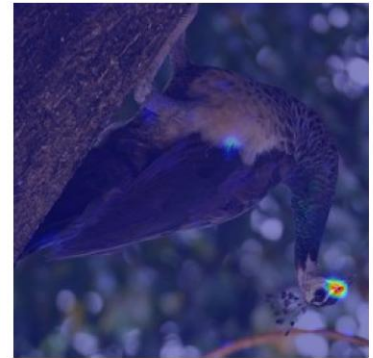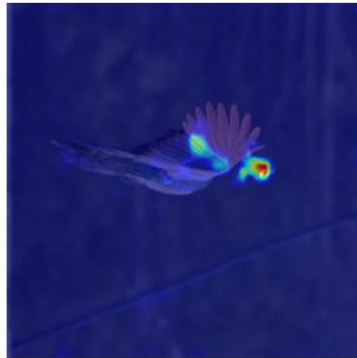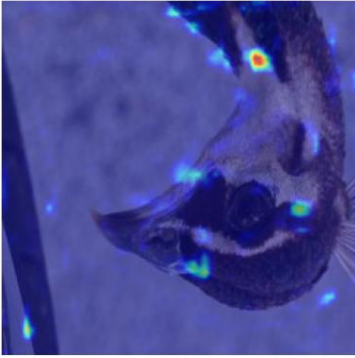
1.12   Indian Peacock

1.12.1   Correct

Indian Peacock is distinguishable by its long blue neck and colorful feathers which are captured in these heatmaps.
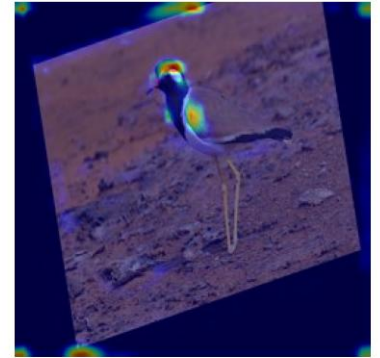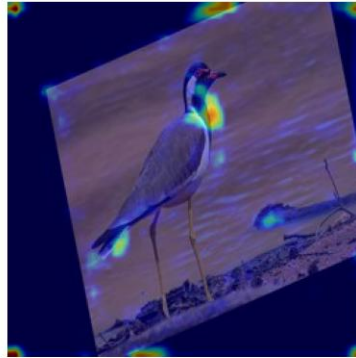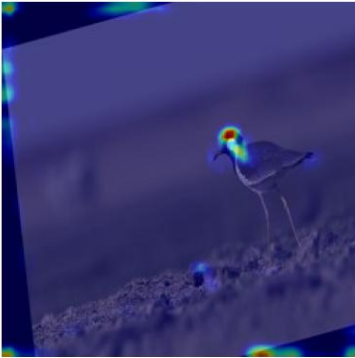
1.12.2 Incorrect

Again we can see that neither the neck, nor the feathers are highlighted in the heatmaps of the false predictions. Hence, the model was unable to indentify peacock from the incorrect images.

1.13 Red Wattled Lapwing

1.13.1 Correct

The Red Wattled Lapwing is characterised by it's neck and upper body, especially the back of it's head. All these are highlighted in the correct heatmap.
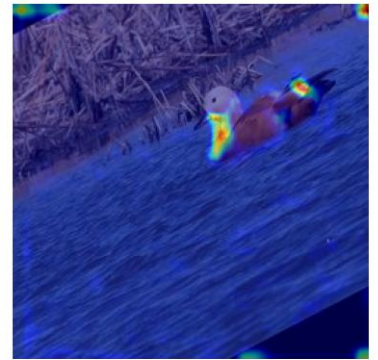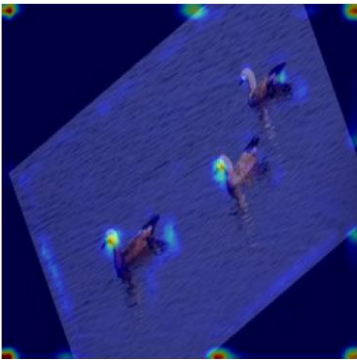
1.13.2 Incorrect

While in the incorrect heatmap, we can see that none of the important features are highlightened, leading to incorrect identification.
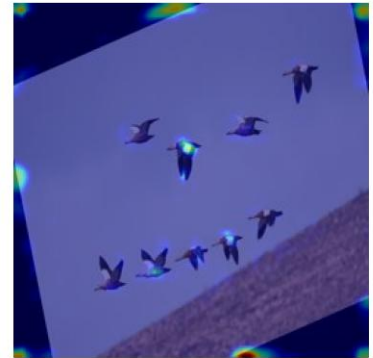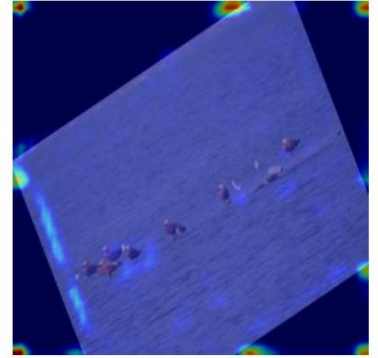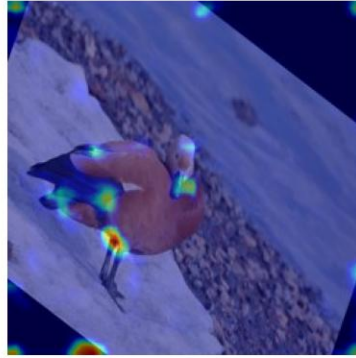
1.14 Ruddy Shelduck

1.14.1 Correct

The important attributes of a Ruddy Shelduck is it's tail and head/face along with some parts of the neck. All the correct image predictions allow us to conclude these remarks.
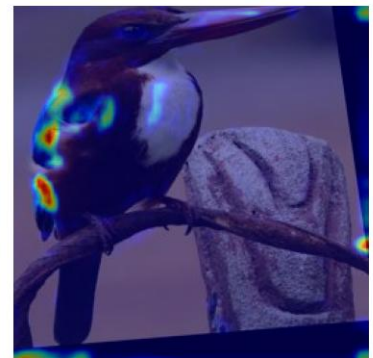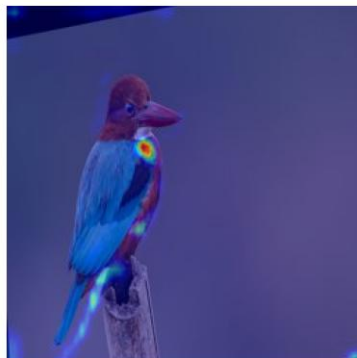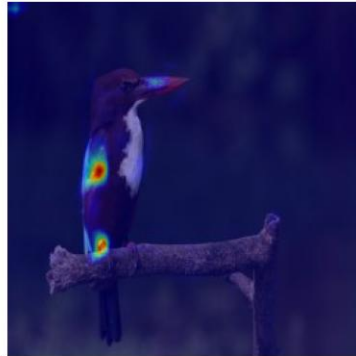
1.14.2   Incorrect

None of the important attributes are highlightened leading to the incorrect predictions in these images.
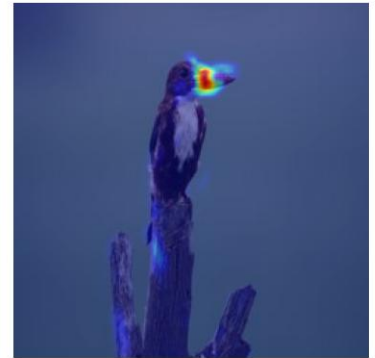
1.15   White Breasted Kingfisher

1.15.1   Correct

White breasted kingfishers have a peculiar stomach and breasts which help in identification of the bird. These parts are highlightened in the heatmaps showcasing the model's capability and techniques similar to humans.
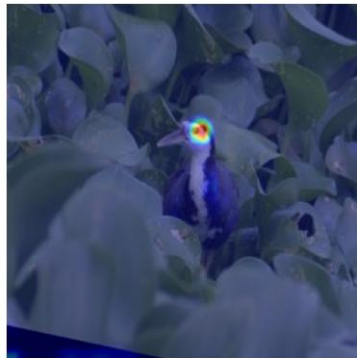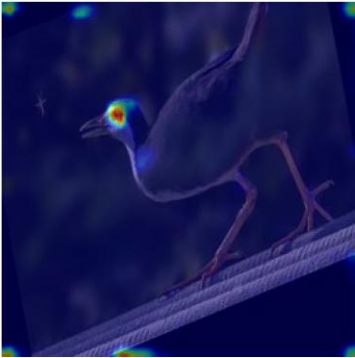
1.15.2 Incorrect

These images of White breasted kingfisher which were predicted incorrectly have their heatmaps highlightened near the beaks, claws and/or some other nearby object. Leading to trouble in identification of the species.
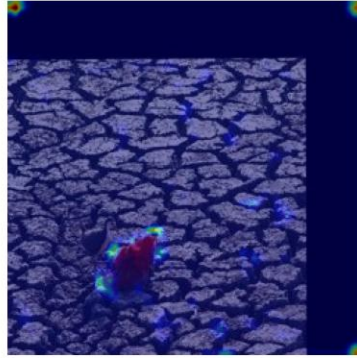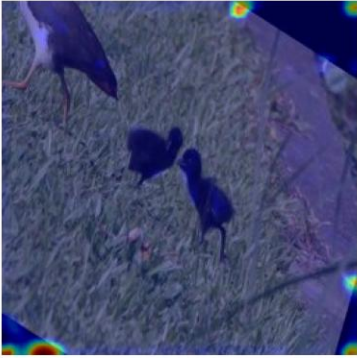
1.16 White Breasted Waterhen

1.16.1 Correct

On the other hand, white breasted waterhens are characterized by their beautiful long faces and eye which can be concluded from the heatmap images.

1.16.2 Incorrect

2 Part B