

Bird Classification

Data Processing

Divided the .csv file into the train, test, and validation sets and performed image augmentations such as normalizing the pixels, shear, zoom, horizontal flip, vertical flip.

Training Models and Methodology

Transfer learning was employed using various popular pre-trained models such as Resnet, VGG16, InceptionV3, and Xception. Xception net and InceptionV3 worked the best however, Xception worked slightly better than InceptionV3.

Hyperparameters:

Optimizer: Adam optimizer was used with a learning rate of 0.001(Inception) and 0.0001(Xception) because it is very versatile, it works well for both deep and shallow networks, it is good for sparse tensors as well.

Activation Function: Rectified Linear Unit (ReLU) was used because it is very less computation hungry and is able to add non-linearity to the network.

A **batch size** of 32 was used so that sufficient features can be learned and also be able to fit them in the memory. (In April 2018, Yann Lecun even tweeted "Friends don't let friends use mini-batches larger than 32").

The **image size** used was 224x224 for faster training and at the same time keeping the image quality good enough. The model reached local minima after around 20 iterations hence, **20 epochs** were used.

NOTE: Optimal value of each parameter has been selected by trial and error, tried a lot of different values.

Accuracy: 0.9707, F1 Score: 0.9698, Precision Score: 0.971, Recall Score: 0.964

Confusion Matrix

