

# Khidmat Report

## *WWF Bird Recognition Model*

Ali Hamza  
ah05084

Ali Haider Rizvi  
xy09876

M. Usaid Rehman  
mr04302

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
1.1	Project Description . . . . .	3
1.2	About WWF Pakistan . . . . .	3
1.3	Project Background & Context . . . . .	3
1.4	Work Plan . . . . .	4
<b>2</b>	<b>Weekly Work Log</b>	<b>5</b>
2.1	Week 1: . . . . .	5
2.2	Week 2: . . . . .	6
2.3	Week 3: . . . . .	7
2.4	Week 4: . . . . .	8
2.5	Week 5: . . . . .	9
2.6	Week 6: . . . . .	10
2.7	Week 7: . . . . .	11
2.8	Week 8: . . . . .	12
<b>3</b>	<b>Conclusion</b>	<b>13</b>
<b>4</b>	<b>Technical Overview</b>	<b>14</b>
4.1	Resources . . . . .	14
4.2	Overview of CS Techniques Used . . . . .	15
4.2.1	Image Collection . . . . .	15
4.2.2	Image Preprocessing . . . . .	15
4.2.3	Neural Network Construction . . . . .	15
4.2.4	Training the Neural Network . . . . .	16
4.2.5	Testing & Fine-Tuning . . . . .	17
4.3	Results . . . . .	17
<b>A</b>	<b>Documentation</b>	<b>18</b>

# 1 Introduction

## 1.1 Project Description

This project is aimed at creating and deploying a deep learning pipeline for WWF Pakistan to classify images of three different species of birds – namely: common myna, house crow, and the house sparrow. This project would serve as a proof-of-concept for a larger model that WWF can use to classify a larger number of birds using a mobile application.

## 1.2 About WWF Pakistan

The World Wildlife Fund for Nature (WWF) is an international non-governmental organization founded in 1961 that primarily works in the areas of wildlife conservation and preservation. WWF also works in the domains of food security, renewable energy, ocean and forest preservation, and many others<sup>1</sup>.

WWF Pakistan is the Pakistan chapter of the World Wildlife Fund, working on similar problems as WWF International but in a more local context. WWF has more than five regional offices throughout the country, and has conducted several successful projects pertaining to freshwater conservation, water security, forests, climate change and others. What they are most known for, however, is wildlife conservation — the area that we have assisted WWF in.

## 1.3 Project Background & Context

Keeping track of biodiversity is a challenge in Pakistan due to a lack of resources, personnel and governmental support. Pakistan has a plethora of wildlife and a rich and diverse ecosystem. WWF Pakistan has helped in the conservation of several endangered species such as the Indus River Dolphin and the Snow Leopard — both of which are now experiencing a rise in numbers.

However, wildlife preservation and protection extends beyond just the conservation of endangered species. Maintaining biodiversity is extremely important for wildlife protection and to maintain a healthy ecosystem and environment. Unfortunately, a lack of biodiversity is a major crisis that is unfolding across the globe. In order to combat this, it is imperative that researches have updated data about species populations, habitats, interactions etc.

An important group of animals that contribute to biodiversity are birds. Therefore, it is vital that data pertaining to bird population is collected in Pakistan. To assist in this endeavor, WWF aims to enlist the help of citizen scientists who could help identify and locate birds and consequently update WWF's database.

Our project aims to implement a deep learning pipeline that would classify images of three different species of birds — house crow, common myna, and the house sparrow.

---

<sup>1</sup><https://wwf.panda.org/>

The pipeline would take an image, preprocess it to a form that the model can work with, and then classify it using a deep neural network — telling us what species the bird in the image belongs to.

This pipeline serves as a proof-of-concept for a larger model that would be classifying hundreds of species of birds. This model will fit into a planned mobile application that citizen scientists can use to spot birds and then capture their image. Not only will the application classify the species of the bird but it will also update WWF’s database, giving them access to more data about bird populations and their locations.

The development of a larger deep learning model and an application is beyond the scope of our *Khidmat* due to time restraints. We have successfully implemented a computer vision pipeline for preprocessing and a deep neural network for classification for three species of birds.

## **1.4 Work Plan**

## 2 Weekly Work Log

### 2.1 Week 1:

Item	Activity	Time	ID
------	----------	------	----

The total time spent on the Khidmat this week is as follows.

ID	Total Hours
st1	
st2	
st3	

## 2.2 Week 2:

Item	Activity	Time	ID
------	----------	------	----

The total time spent on the Khidmat this week is as follows.

ID	Total Hours
st1	
st2	
st3	

### 2.3 Week 3:

Item	Activity	Time	ID
------	----------	------	----

The total time spent on the Khidmat this week is as follows.

ID	Total Hours
st1	
st2	
st3	

## 2.4 Week 4:

Item	Activity	Time	ID
------	----------	------	----

The total time spent on the Khidmat this week is as follows.

ID	Total Hours
st1	
st2	
st3	



## 2.5 Week 5:

Item	Activity	Time	ID
------	----------	------	----

The total time spent on the Khidmat this week is as follows.

ID	Total Hours
st1	
st2	
st3	

## 2.6 Week 6:

Item	Activity	Time	ID
------	----------	------	----

The total time spent on the Khidmat this week is as follows.

ID	Total Hours
st1	
st2	
st3	

## 2.7 Week 7:

Item	Activity	Time	ID
------	----------	------	----

The total time spent on the Khidmat this week is as follows.

ID	Total Hours
st1	
st2	
st3	

## 2.8 Week 8:

Item	Activity	Time	ID
------	----------	------	----

The total time spent on the Khidmat this week is as follows.

ID	Total Hours
st1	
st2	
st3	

### 3 Conclusion

## 4 Technical Overview

### 4.1 Resources

The resources that we used for this project can be categorized into two main categories: images, and technologies. Our first step was to acquire a set of images that would function as our dataset. We collected around 550 images for each of the three species of birds — the house crow, common myna, and the house sparrow. We used several different websites and online databases to find these images, we list these websites below:

1. <https://search.macauleylibrary.org/catalog?taxonCode=myna&mediaType=p&q=Common%20Myna>
2. [https://ebird.org/media/catalog?taxonCode=commyn&mediaType=p&sort=rating\\_rank\\_desc&q=Common%20Myna%20-%20Acridotheres%20tristis](https://ebird.org/media/catalog?taxonCode=commyn&mediaType=p&sort=rating_rank_desc&q=Common%20Myna%20-%20Acridotheres%20tristis)
3. [https://ebird.org/media/catalog?taxonCode=houcrol&sort=rating\\_rank\\_desc&mediaType=p&regionCode=](https://ebird.org/media/catalog?taxonCode=houcrol&sort=rating_rank_desc&mediaType=p&regionCode=)
4. [https://search.macauleylibrary.org/catalog?taxonCode=houcrol&mediaType=p&region=Pakistan%20\(PK\)&regionCode=PK&q=House%20Crow%20-%20Corvus%20splendens](https://search.macauleylibrary.org/catalog?taxonCode=houcrol&mediaType=p&region=Pakistan%20(PK)&regionCode=PK&q=House%20Crow%20-%20Corvus%20splendens)
5. <https://www.kaggle.com/gpiosenska/100-bird-species>
6. <https://search.macauleylibrary.org/catalog?taxonCode=houspa&mediaType=p&q=House%20Sparrow>
7. [https://ebird.org/media/catalog?taxonCode=houspa&mediaType=p&sort=rating\\_rank\\_desc&q=House%20Sparrow%20-%20Passer%20domesticus](https://ebird.org/media/catalog?taxonCode=houspa&mediaType=p&sort=rating_rank_desc&q=House%20Sparrow%20-%20Passer%20domesticus)

All collected images can be viewed at [https://drive.google.com/drive/folders/18k-roE-VJSB1dcrhvN1y\\_EosVF7KbOdY?usp=sharing](https://drive.google.com/drive/folders/18k-roE-VJSB1dcrhvN1y_EosVF7KbOdY?usp=sharing).

To preprocess our images, convert them into a usable dataset, and to create our deep learning model, we had to rely on several different preexisting tools. We list the tools that we used in this project below:

- **Python** — The primary programming language that we used for this project. Most of the other tools that we used are different libraries & modules for in Python.
- **OpenCV** — A computer vision module in Python to preprocess the images.
- **TensorFlow & Keras** — A Python library that we used to construct, and then train and test our neural network.

- **Matplotlib** — A plotting library for Python, which we used to visualize our
- **Sci-kit learn** — A Python library for machine learning that we used to construct our training and testing set.
- **Google Colab** — We used Colab notebooks to write code and then run code using GPU-accelerated computation that is made available using Google Colaboratory.

## 4.2 Overview of CS Techniques Used

This project was an amalgamation of several different aspects of computer science techniques. The process can be summarized into 5 main steps:

1. Image Collection
2. Image Preprocessing
3. Construction of Neural Network
4. Training the Neural Network
5. Testing and Optimization

We briefly expand upon these steps in the following sections. For a more technical description, please view the documentation ([LINK HERE](#)).

### 4.2.1 Image Collection

We used the sources listed at [7](#) to collect images of the three species of birds. We downloaded approximately 650 images for each birds, which we then filtered down to around 550 to maintain uniform image quality. Based on our results, we believe the number of images that we have, i.e. the size of our dataset is adequate.

### 4.2.2 Image Preprocessing

The next step was to pre-process these images to convert them into a usable dataset.

### 4.2.3 Neural Network Construction

Artificial neural networks form the basis for solving most artificial intelligence problems. Image classification is one such problem. Neural networks represent connections between artificial neurons, also called units. These connections are represented as weights in the neural network. These weights are also called parameters and learned by the neural network during supervised learning.

There are several types of neural networks that are optimal for solving different intelligence problems. Image recognition/classification has been solved using several types of neural networks since the 1980s such as multilayer perceptrons. A newer method to solve image classification problems (and problems generally pertaining to computer vision) is the Convolutional Neural Network (CNN).

Unlike other neural networks, a CNN takes spatial information about the input image into consideration during learning, allowing it to generally be more accurate. Therefore, we decided to implement a CNN for this project. We first implemented a simplistic CNN using two-dimensional convolutional layers that preexist in TensorFlow. However, we found that our results were not as satisfactory as we hoped for them to be.

We then tried implementing the Residual Neural Network (ResNet), which is built upon the same idea as CNNs but significantly improves accuracy and performance. ResNet utilize skip connections or shortcuts between layers. Specifically, we leveraged TensorFlow's built-in ResNet50 implementation, a variant of ResNet that utilizes 48 convolutional layers, a MaxPool layer and an Average Pool layer. However, since ResNet50 is a fairly deep architecture which is more useful when implementing models for data sets that are quite large, with at least several thousand examples.

After using ResNet50, we use the output of the last AvgPool layer and flatten the output. Then we use two flattened fully-connected layers and an output layer to give us our final output. In total, our neural network contained 24,715,779 trainable parameters and 53,120 non-trainable parameters. For more technical information on the neural network construction, please see [LINK](#).

#### 4.2.4 Training the Neural Network

To train neural networks, there are several different methods that optimize the cost function. The most famous of these is the stochastic gradient descent (SGD). SGD is a stochastic approximation of the simpler gradient descent optimization. Gradient descent approximates the gradient for the entire training set at each iteration. However, stochastic gradient descent approximates the gradient for one example in the training set per iteration — making it much faster.

A newer method called minibatch stochastic gradient finds the middle ground between gradient descent and SGD. It approximates the gradient in each iteration using a randomly selected minibatch of training examples. These minibatches are also effective for parallel computation using TPUs.

We used the Adaptive Moment Estimation or *Adam* optimization algorithm. Adam is an effective algorithm for problems with sparse gradients such as computer vision problems. This optimization algorithm relies on three hyperparameters:  $\alpha$ ,  $\beta_1$ ,  $\beta_2$ , and  $\epsilon$ . Adam is very effective in optimizing problems that contains a large number of trainable parameters. The value we set for  $\alpha$  was 0.001 and the  $\beta$  values were set as default, i.e. 0.9 and 0.999. The loss function for our model was the cross-entropy loss function which is quite



popular in image classification tasks.

#### **4.2.5 Testing & Fine-Tuning**

### **4.3 Results**

## A Documentation

## Khidmat Completion Form

To be completed by the external supervisor.

*Please use the space below to provide any comments you may have on the students' performance, the Khidmat program, or any other feedback you want to share with Habib University's Khidmat committee. We can also be reached at [khidmat@sse.habib.edu.pk](mailto:khidmat@sse.habib.edu.pk).*

---

I hereby certify that I supervised XXX and XXX for the Khidmat described in this report. Furthermore, that I have read and agree with the weekly updates included in this report. My signature below marks the successful completion of the work undertaken for the Khidmat.

\_\_\_\_\_  
Name and signature

&

\_\_\_\_\_

Location and date