**TRIBHUVAN UNIVERSITY**

**INSTITUTE OF ENGINEERING**

ADVANCED COLLEGE OF ENGINEERING AND MANAGEMENT

DEPARTMENT OF ELECTRONICS AND COMPUTER ENGINEERING

KALANKI, KATHMANDU



**A Minor Project Final Defense Report On**

**“ BIRD SPECIES DETECTION ”**

**[CT-654]**

**Submitted By**

**Ankit Chhetri 30759**

**Ankrit Risal 30760**

**Bhaskar Subedi 30769**

**Supervisor:**

Er. Laxmi Prasad Bhatt

Er. Ukesh Thapa

A Minor Project Proposal report submitted to the Department of Electronics and Computer Engineering in the partial fulfillment of the requirements for degree of Bachelor of Engineering in Computer Engineering

Kathmandu, Nepal

10th March, 2024

ADVANCED COLLEGE OF ENGINEERING AND

MANAGEMENT

DEPARTMENT OF COMPUTER AND ELECTRONICS ENGINEERING

APPROVAL LETTER

The undersigned certify that they have read and recommended to the Institute of

Engineering for acceptance, a project report entitled “PROJECT TITLE”

submitted by:

**Ankit Chhetri 30759**

**Ankrit Risal 30760**

**Bhaskar Subedi 30769**

In partial fulfillment for the degree of Bachelor in Computer Engineering.

|  |  |
| --- | --- |
| …………………… | …………………………… |
| Er. Laxmi Prasad Bhatt  Supervisor | Er. Ukesh Thapa  Supervisor |

..................................

Associate Prof Er. Sudeep Shakya

External Examiner

..................................

Er. Laxmi Prasad Bhatt

Academic Project coordinator

Department of Computer and Electronics Engineering

10th March, 2024

## ACKNOWLEDGEMENT

We take this opportunity to express our deepest and sincere gratitude to our supervisor **Er. Laxmi Prasad Bhatt** and **Er. Ukesh Thapa,** Department Of Electronics and Computer Engineering for their insightful advice, motivating suggestions, invaluable guidance, help and support in this project selection and also for their constant encouragement and advice throughout our journey till the date.

We express our deep gratitude to **Er. Prem Chandra Roy**, Head of Department of Electronics and Computer Engineering, **Er. Bikash Acharya and Er. Dhiraj Pyakurel,** Deputy Head, Department Of Electronics and Computer Engineering, **Er. Laxmi Prasad Bhatt,** Academic Project Coordinator, Department of Electronics and Computer Engineering for their regular support, co-operation, and coordination.

The in-time facilities provided by the department throughout the Bachelors program are also equally acknowledgeable.

We would like to convey our thanks to the teaching and non-teaching staff of the Department of Electronics & Communication and Computer Engineering, ACEM for their invaluable help and support till the date. We are also grateful to all our classmates for their help, encouragement and invaluable suggestions.

Finally, yet more importantly, we would like to express our deep appreciation to our Grandparents, parents, siblings for their perpetual support and encouragement throughout the Bachelor’s degree period.

Ankit Chhetri 30759

Ankrit Risal 30760

Bhaskar Subedi 30769

## ABSTRACT

Innumerable people visit bird sanctuaries worldwide, captivated by the diverse elegance exhibited by various avian species. However, amidst this fascination lies a daunting challenge: accurately identifying these birds. Species identification often yields multiple labels, and disagreements among experts are not uncommon, highlighting the complexity of the task. Both human observers and traditional computer vision methods face limitations in their visual abilities, underscoring the need for innovative solutions.

To address this challenge, we have developed sophisticated deep learning algorithms based on Convolutional Neural Networks (CNNs) within the realm of image processing. This advanced approach enables our system to learn intricate patterns and features representative of different bird species. By meticulously training our CNN architecture with a comprehensive dataset, it becomes adept at recognizing and distinguishing among myriad avian species.

Upon receiving an input image of a bird, our model swiftly compares it with the learned features and accurately predicts the species. Furthermore, our system offers additional valuable information by displaying the Nepali scientific name corresponding to the predicted species, enhancing the user's understanding and appreciation of avian biodiversity.

This innovative model serves a diverse range of users, including general bird enthusiasts, ornithologists, researchers, and conservationists, providing them with a reliable tool for species identification. Additionally, our system offers a collaborative aspect by allowing users to contribute images that are missing from the initial dataset, thereby continually enriching and expanding its capabilities. Through this collaborative effort, we aim to foster a community-driven approach to avian research and conservation, ensuring the accuracy and comprehensiveness of our bird species identification system.

**Keywords:** *Bird species identification, deep learning, image processing, CNN*

## [TABLE](#_heading=h.46r0co2) OF CONTENTS

**Title Page**

**Approval letter i**

[**ACKNOWLEDGEMENT ii**](#_heading=h.37m2jsg)

[**ABSTRACT iii**](#_heading=h.1mrcu09)

[**TABLE**](#_heading=h.46r0co2) **OF CONTENTS**  [**iv**](#_heading=h.46r0co2)

[**LIST**](#_heading=h.2lwamvv) **OF FIGURES**  [**vii**](#_heading=h.2lwamvv)

[**LIST OF ABBREVIATIONS/ACRONYMS viii**](#_heading=h.3l18frh)

[**CHAPTER 1 1**](#_heading=h.4k668n3)

[**INTRODUCTION 1**](#_heading=h.2zbgiuw)

[1.1 Background 1](#_heading=h.1egqt2p)

[1.2 Motivation 3](#_heading=h.3ygebqi)

[1.3 Statement of the Problem 3](#_heading=h.2dlolyb)

[1.4 Project objective 3](#_heading=h.sqyw64)

[1.5 Significance of the study 3](#_heading=h.3cqmetx)

[**CHAPTER 2 4**](#_heading=h.1rvwp1q)

[**LITERATURE REVIEW 4**](#_heading=h.4bvk7pj)

[**CHAPTER 3 6**](#_heading=h.34g0dwd)

[**REQUIREMENT ANALYSIS 6**](#_heading=h.1jlao46)

3.1 Functional Requirements [6](#_heading=h.43ky6rz)

3.2 Non-Functional Requirements 6

3.3 System Requirements 7

3.3.1 Hardware Requirements [7](#_heading=h.2iq8gzs)

3.3.2 Software Requirements 7

3.4 Feasibility Study 7

3.4.1 Technical Feasibility 7

3.4.2 Operational Feasibility 8

3.4.3 Economical Feasibility 8

[**CHAPTER 4 9**](#_heading=h.xvir7l)

[**SYSTEM DESIGN AND ARCHITECTURE 9**](#_heading=h.3hv69ve)

4.1 System Architecture 9

4.2 Use Case Diagram 9

4.3 Flowchart 10

4.3 Sequence Diagram 11

[**CHAPTER 5 11**](#_heading=h.1x0gk37)

[**METHODOLOGY 12**](#_heading=h.4h042r0)

5.1 Software Development model [12](#_heading=h.2w5ecyt)

5.2 Machine learning Algorithm 13

5.3 Pre-processing 14

5.3.1 Image Resize [14](#_heading=h.1baon6m)

5.3.2 Random Horizontal Flip [14](#_heading=h.3vac5uf)

5.3.3 Random rotation 14

5.3.4 Random resized crop 15

5.3.5 Color Jitter 15

5.3.6 Normalize 15

5.3.7 Categorical Encoding 15

5.3.8 Representation and Description 15

5.4 Architecture [15](#_heading=h.2afmg28)

5.4.1 Convolutional Neural Network [15](#_heading=h.pkwqa1)

5.4.2 Feature Extraction 16

5.4.2.1 Convolutional Layer 16

5.4.2.2 Activation Layer 17

5.4.2.3 Pooling Layer 17

5.4.2.4 Fully Connected layer 17

5.4.2.5 Classification 17

5.5 Optimizer [18](#_heading=h.2afmg28)

5.6 Cross Validation 18

5.7 Loss Function 19

[**CHAPTER 6 20**](#_heading=h.39kk8xu)

**RESULT AND ANALYSIS** [**20**](#_heading=h.1opuj5n)

6.1 Dataset [20](#_heading=h.48pi1tg)

6.2 Result 20

6.3 Result Obtained 21

[**CHAPTER 7 23**](#_heading=h.haapch)

**CONCLUSION, LIMITATIONS AND FUTURE ENHANCEMENT**  [**23**](#_heading=h.319y80a)

7.1 Conclusion [23](#_heading=h.48pi1tg)

7.2 Limitation  [23](#_heading=h.48pi1tg)

7.3 Future Enhancement [23](#_heading=h.48pi1tg)

[**CHAPTER 8 24**](#_heading=h.1gf8i83)

[**REFERENCES 24**](#_heading=h.2fk6b3p)

## LIST OF FIGURE

**Title Page**

4.1 System Architecture 9

4.2 Use Case Diagram 9

4.3 Flowchart 10

4.4 Sequence Diagram 11

5.1 Prototype Model 13

5.2 Machine Learning Algorithm 14

5.3.1 Block Diagram of System 16

5.3.2 Activation Function 16

5.3.3 Convolutional operation 18

5.5 Cross validation 19

6.1 Train and validation losses and accuracy graph 20

6.2 User Interface 21

6.3 Upload Photo 21

6.4 Bird Recognition 22

## [LIST OF ABBREVIATIONS/ACRONYMS](#_heading=h.3l18frh)

CNN Convolutional Neural Network

IDE Integrated Development Environment

IUCN International Union for Conservation of Nature

ReLU Rectified Linear Unit

RGB Red Green Blue

SGD Stochastic Gradient Descent

VGG Net Visual Geometry Group Network

## CHAPTER 1

## INTRODUCTION

Deep Learning is a Machine Learning subfield which is in turn a subfield of Artificial Intelligence. Deep learning can be visualized as a platform where artificial, human brain inspired neural networks and algorithms learn from large amounts of data. Deep Learning allows computers to solve complex problems even though they use a very diverse, unstructured, and interconnected data set Identification of bird is a difficult process that often results in confusing labels bird species identification is seen as a perplexing problem which often leads to confusion collecting and gathering information on bird’s needs tremendous human effort. Many people visit bird sanctuaries to look at the birds while they barely recognize the differences between different species of birds and their characteristics. Even while various bird species share the same basic set of components, their form and appearance might vary. Attempts to automate species identification in audio recordings have included birds, amphibians, bats, insects, fish, and marine mammals.

To predict the birds in their natural habitats, we developed an interface to extract information from bird images using the Convolutional Neural Network (CNN) algorithm. First, a vast dataset of birds were gathered and localized. Second, CNN architecture was designed similar to the VGG Net Network. Now that the network was implemented, we trained the CNN model with the bird data set using Pytorch, and subsequently the classified, trained data was stored on the disk to identify a target object. Ultimately, the client-server architecture navigates a sample bird image submitted by an end-user to retrieve information and predict the bird species from the qualified model stored on the disk. This method allows the autonomous identification of birds from the captured images and can provide important, useful knowledge about bird species.

## 1.1 Background

According to the International Union for Conservation of Nature (IUCN), over 1300 species of birds are globally at risk of extinction. In Nepal, a country renowned for its rich biodiversity, the conservation status of avian species is of particular concern. Out of the 878 bird species recorded in Nepal, a significant proportion, totaling 167 species (19%), have been assessed as nationally threatened. This alarming statistic encompasses 67 species (40%) categorized as Critically Endangered, 38 species (23%) as Endangered, and 62 species (37%) as Vulnerable [6]. Additionally, 62 species have been designated as Near Threatened, indicating their susceptibility to future threats. The plight of birds in Nepal is further underscored by the sobering fact that nine species (1% of the total threatened) are considered Regionally Extinct, with no recorded sightings since the 19th century. Furthermore, 22 species (2.5% of the total) are labeled as Data Deficient, highlighting gaps in our understanding of their conservation status and urgent need for further research [6].

Birds fulfill essential ecological roles that underpin the functioning of ecosystems. They contribute to vital processes such as pollination, seed dispersal, insect control, and nutrient cycling, thereby maintaining ecosystem health and resilience. The loss of bird populations can have far-reaching ecological repercussions, disrupting ecosystem dynamics and compromising biodiversity. Moreover, birds serve as crucial indicators of ecosystem health, with declines in bird populations often signaling broader environmental problems.

In addition to their ecological significance, birds hold economic importance for communities worldwide. Birdwatching tourism, in particular, has emerged as a lucrative industry, generating revenue and employment opportunities in many regions. In Nepal, the diverse avifauna attracts birdwatchers and nature enthusiasts from around the globe, contributing to local economies and fostering conservation initiatives. Furthermore, birds provide valuable ecosystem services in agricultural landscapes, including pest control and pollination, which directly benefit farmers and enhance crop yields.

Effective conservation strategies rely on accurate species identification, population monitoring, and habitat preservation. Technological advancements, including remote sensing, GPS tracking, and artificial intelligence, offer promising tools for enhancing bird conservation efforts. These technologies enable researchers to conduct detailed assessments of bird populations, identify critical habitats, and develop targeted conservation interventions. Furthermore, citizen science initiatives empower individuals to contribute valuable data and actively participate in conservation endeavors.

### 1.2 Motivation

The development of the intelligent automated system for endangered bird species classification can have far-reaching implications beyond just conservation efforts. The use of deep learning in the industry of ornithology can help us better understand the ecology and behavior of bird species. By accurately identifying and tracking populations of different species, we can gather more data on their movements, habitat preferences, and overall behavior patterns. Thus overall improves the awareness of endangered bird species among the people.

### 1.3 Statement of the Problem

Manual identification of bird species is very tedious task as well as very unreliable as his/her knowledge may not be in-depth and limited to the local bird species. This process is a lot of time-consuming and it may contain some errors. There are lots of books that have been published for the process of helping a human incorrectly identifying bird species. The current bird species identification process involved using the bird audio which is recorded and fed into the system. Nevertheless, it requires the hundreds of hours to carefully analyzed and classify the species. Due to such a process, large scale bird identification is almost an impossible task. So, to automate the process is a more practical approach.

**1.4 Project objective**

To accurately identify the bird species through Nepali image processing using deep convolution neural networks (CNNs).

### 1.5 Significance of the study

Our research aims to contribute to the development of effective, accurate, and scalable solutions for bird identification using deep learning techniques. Such solutions can help to protect and conserve endangered bird populations and ultimately contribute to the preservation of our planet's rich biodiversity. The development of deep learning algorithms for bird identification is an exciting and rapidly growing field, and we look forward to advancing this research further in the future.

## CHAPTER 2

## LITERATURE REVIEW

Bird species identification is a fundamental task in biodiversity conservation efforts. Various automated systems have been developed to detect and classify different species of birds. Fagerlund [1] employed machine learning algorithms such as support vector machines and k-nearest neighbors for bird species detection. While these algorithms demonstrate accurate prediction of bird species, they rely on handcrafted visual features. However, the manual extraction of features from large datasets can be cumbersome and may not fully capture the complexity of avian characteristics.

To address the limitations of handcrafted features, deep learning models such as Convolutional Neural Networks (CNNs) have emerged as more efficient alternatives for image recognition tasks [2]. CNNs have shown superior accuracy and performance compared to traditional machine learning models, owing to their ability to automatically learn discriminative features directly from the raw data. Furthermore, to enhance the accuracy and efficiency of CNN architectures, optimization techniques such as the Adam optimizer have been employed [3].

The Adam optimizer [3] is a popular optimization algorithm that adjusts the weights of neural networks using individual learning rates for each parameter. This adaptive learning rate mechanism enables faster convergence and more efficient training compared to conventional optimization methods like Stochastic Gradient Descent (SGD) [4]. Additionally, to fine-tune the performance of the CNN model and prevent overfitting, cross-validation techniques are employed during training.

Cross-validation, particularly k-fold cross-validation, is an essential hyperparameter tuning approach utilized in training datasets. By splitting the dataset into k subsets and iteratively training the model on k-1 subsets while validating on the remaining subset, k-fold cross-validation helps to ensure that the model generalizes well to unseen data and mitigates the risk of overfitting [5].

In line with these advancements, our project proposes a CNN architecture optimized with the Adam optimizer for the detection of endangered Nepalese bird species. Moreover, our CNN model incorporates a feature to provide the Nepali name and conservation status of the detected species, thereby enhancing its utility for conservation efforts in Nepal. Through the integration of state-of-the-art deep learning techniques and optimization strategies, our project aims to contribute to the preservation of Nepal's avian biodiversity and facilitate more effective conservation practices.

## 

## CHAPTER 3

## REQUIREMENT ANALYSIS

Requirements analysis or requirements engineering is a process used to determine the needs and expectations of a new product. It involves frequent communication with the stakeholders and end-users of the product to define expectations, resolve conflicts, and document all the key requirements.

**3.1 Functional Requirements**

Functional requirements are specifications that describe the functions and capabilities a system or product must have to meet its intended purpose.

**Bird identification:** Our system allows ornithologists and general people detecting different species of birds with greater accuracy.

**Bird Description:** Our system gives the Nepali name and scientific name of the predicted bird.

**3.2 Non-Functional Requirements**

A Non-functional requirement is a requirement that specifies criteria that can be used to judge the operation of a system, rather than specific behaviors.

i. **Accuracy**: The system has to give an accuracy greater than 75%.

ii. **Training Time**: Estimated training time for the algorithm must be less.

iii. **Maintenance**: The system should be easy to maintain and support.

iv. **Response time**: Response time must be as quick as possible.

**3.3 System Requirements**

**3.3.1 Hardware Requirements**

i. A multi-core CPU (e.g., quad-core) is typically sufficient for basic CNN training and inference.

ii. NVIDIA GTX or AMD RX (GPU) can be used if needed.

iii. A fast SSD with sufficient capacity (e.g., 256 GB) for storing datasets, model weights, and other training-related files.

**3.3.2 Software Requirements**

* Language: Python
* Pytorch
* Web framework: Flask
* Numpy, Sklearn, Optuna and Matplotlib
* Operating System: Windows and Linux
* Frontend: HTML,CSS and Javascript
* IDE: VS code, Jupyter Notebook

## 3.4 Feasibility Study

In the feasibility study for the Bird Species Recognition project using CNN would typically include the following elements:

**3.4.1 Technical Feasibility**

Our model is capable of leveraging available hardware resources like GPUs for accelerated model training and inference. It utilizes the Pytorch framework and relevant libraries, enabling seamless integration of state-of-the-art deep learning techniques and facilitating the development and deployment of the CNN-based solution. Our team possesses the necessary expertise in deep learning, and PyTorch framework, enabling us to effectively implement and maintain the CNN-based solution, and continually improve its performance through ongoing research and development efforts.

* + 1. **Operational Feasibility**

Our model is designed to be user-friendly and intuitive, ensuring easy acceptance and adoption by wildlife researchers and conservationists who may not have extensive technical expertise in deep learning or computer vision. We provide comprehensive training materials, documentation, and ongoing support to users, empowering them to effectively utilize the bird species recognition system in their research and conservation efforts. With robust data management practices in place, our model efficiently handles data storage, retrieval, and organization tasks, ensuring data integrity, security, and compliance with regulatory requirements.

* + 1. **Economic Feasibility**

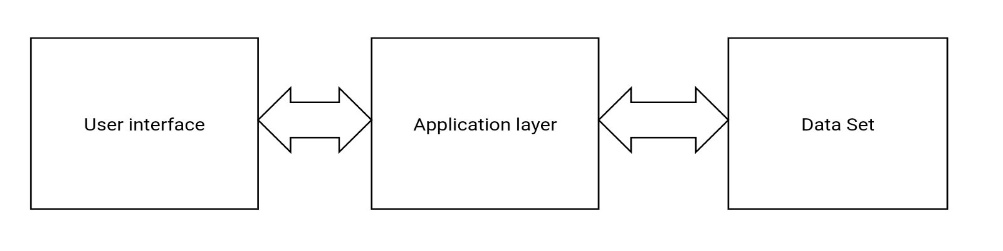
Our model offers a cost-effective solution for bird species recognition balancing initial investment with long term operational costs. Through thorough cost-benefit analysis, we've determined that the benefits of improved wildlife monitoring outweigh project costs, making it a sound investment. Our scalable approach optimizes resource utilization, maximizing return on investment for wildlife research and conservation efforts.

## CHAPTER 4

## 

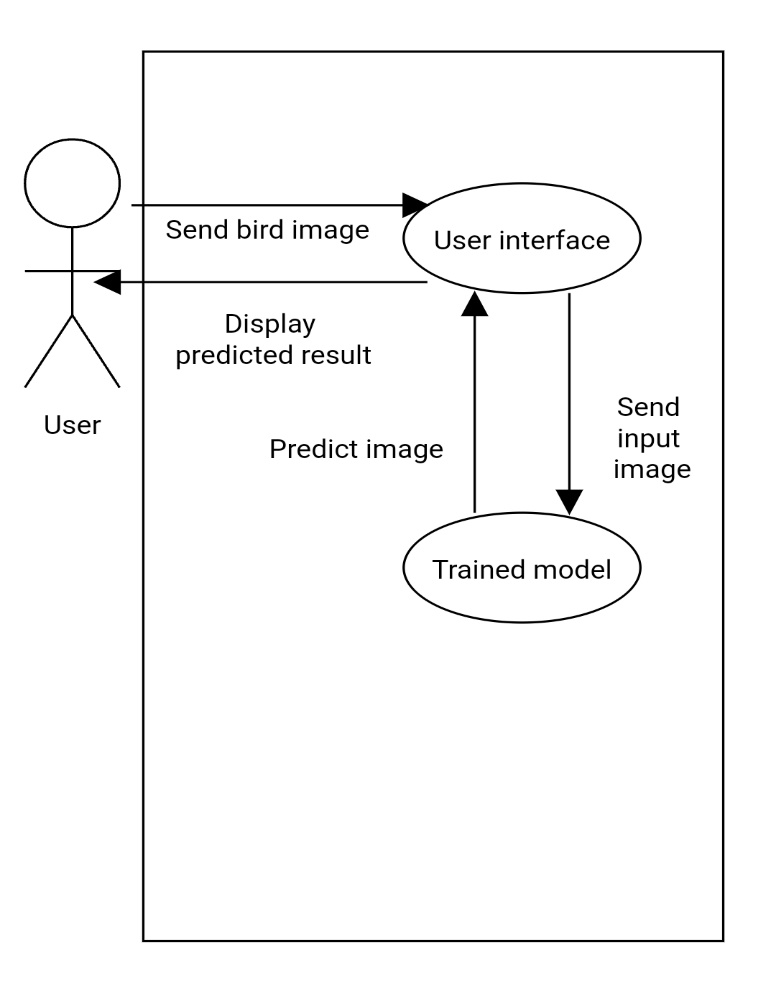
## SYSTEM DESIGN AND ARCHITECTURE

**4.1 System Architecture:-**



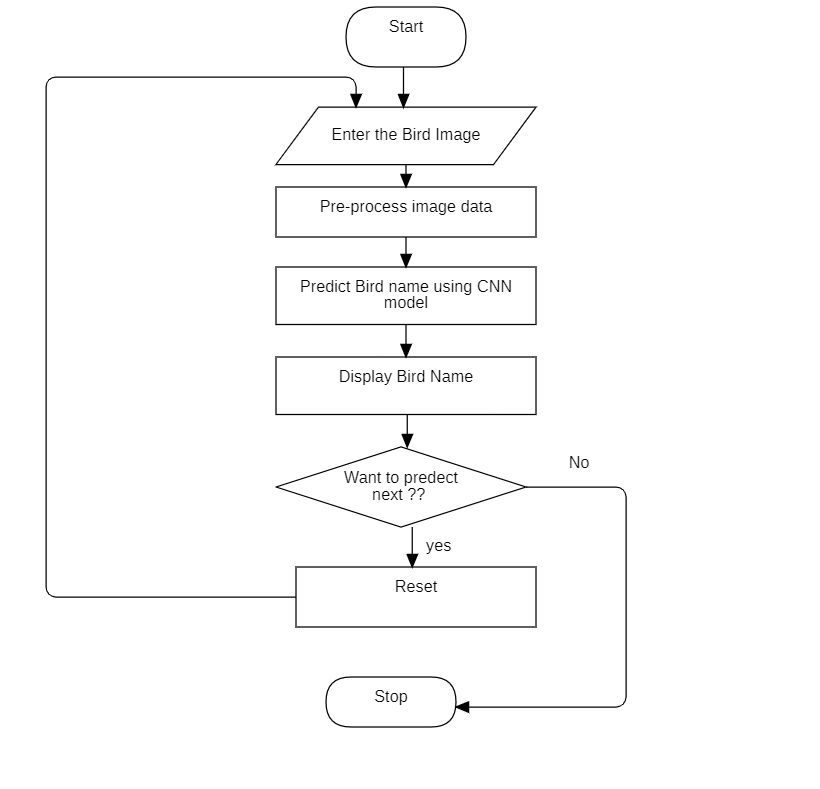
*Fig 4.1: System Architecture*

**4.2 Use Case Diagram:-**

****

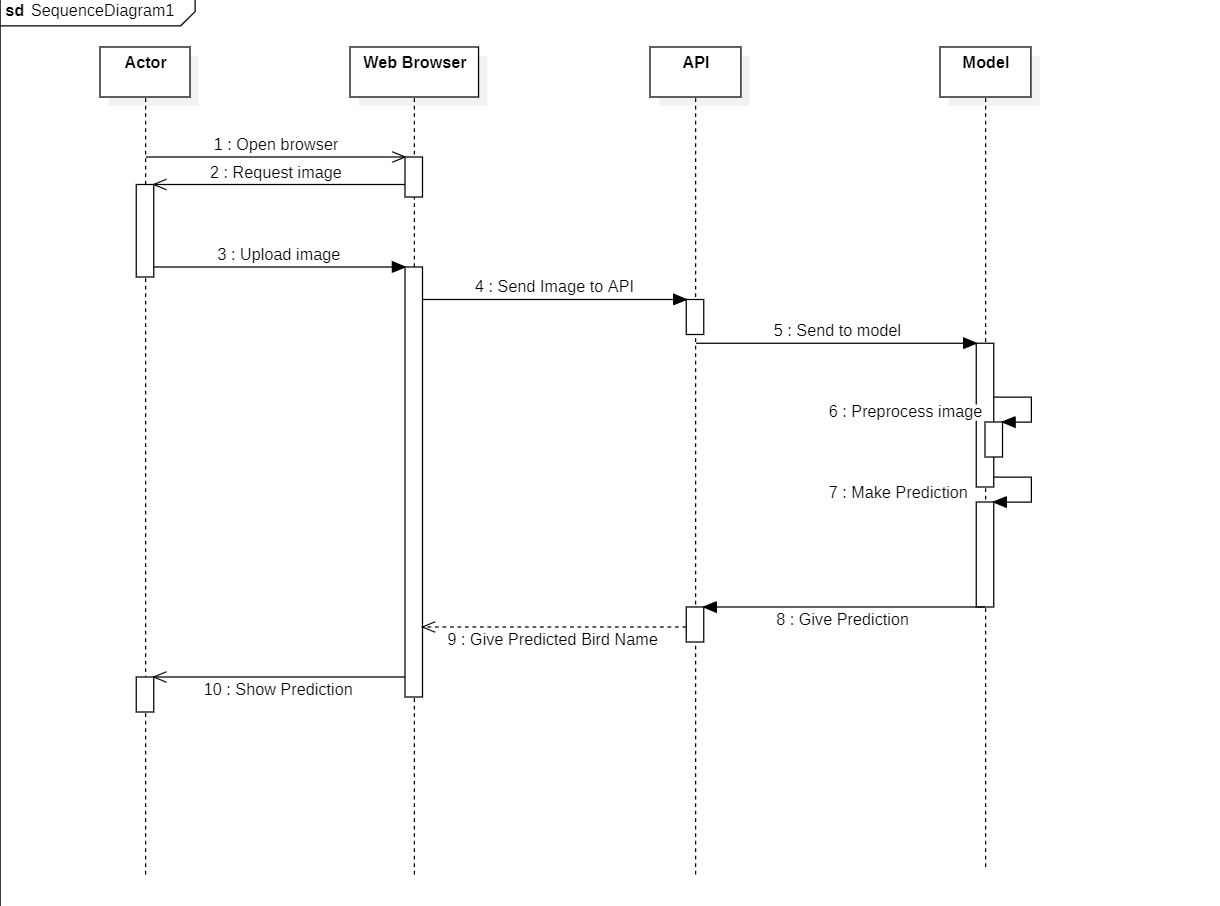
*Fig 4.2: Use Case Diagram*

**4.3 Flowchart:-**



*Fig 4.3: Flowchart*

**4.4 Sequence Diagram:-**



*Fig 4.5: Sequence Diagram*

## CHAPTER 5

## 

## METHODOLOGY

The proposed project aims to develop a system that can classify different species of birds.

### 5.1 Software Development Model:

Prototyping is defined as the process of developing a working replication of a product or system that has to be engineered. It offers a small-scale facsimile of the end product and is used for obtaining customer feedback.

In our project, a prototype of the end product was first developed, tested, and will be refined as per customer feedback repeatedly till a final acceptable prototype is achieved which forms the basis for developing the final product.

**Step 1:** Initial Requirement: Initially all the requirements needed were evaluated such as data collection, preprocessing and then goes to next step.

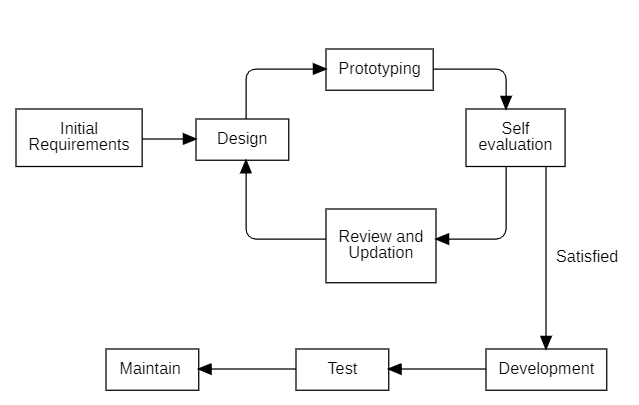
**Step 2:** Quick Design: This is the second step in Prototyping Model. This model covers the basic design of the requirement through which a quick overview can be easily described.

**Step 3:** Build a Prototype: This step helps in building an actual prototype from the knowledge gained from prototype design.

**Step 4:** Initial User Evaluation: This step describes the preliminary testing where the investigation of the performance model occurs, as the customer will tell the strength and weaknesses of the design.

**Step 5:** Refining Prototype: If any feedback is given by the user, then improving the client’s response to feedback and suggestions, the final system is approved.

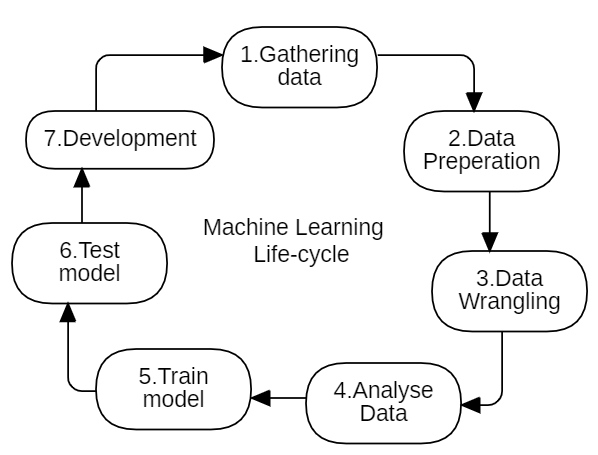
**Step 6:** Implement Product and Maintain: This is the final step in the phase of the Model where the final system is tested and distributed to production, here program is run regularly to prevent failures.



*Fig 5.1: Prototype model*

**5.2 Machine learning Algorithm:**

* The following steps are followed in machine learning life cycle:
  1. Gathering data: 38 Species bird dataset was downloaded.
  2. Data preparation: Images were split into train, test and validation.
  3. Data wrangling: Data was augmented, pre-processed and changed into input form.
  4. Analyze data: Classification model was selected for the project.
  5. Train model: Model was trained with input data.
  6. Test model: Model was tested using unseen data.
  7. Deployment: A complete system was built by integrating with UI and API.



*Fig 5.2: Machine learning algorithm*

#### 5.3 Pre-processing:

Preprocessing is the removal of systematic noise from data and is usually necessary before image classification and analysis. The goal of preprocessing is to attempt to decrease undesirable variety in images due to lighting, scale, deformation, etc. Some of the common preprocessing techniques used in our model are:

#### 5.3.1 Image Resize:-

Each image get resized to fixed size of (150,150) pixel in while training dataset.

**5.3.2 Random Horizontal Flip:-**

We implemented random flips of the images horizontally with a probability of 0.5 to induce diversity to training data, enhancing the model’s robustness and generalization.

#### 5.3.3 Random rotation:-

The input images are rotated by maximum 15 degree for augmentation of dataset by introducing variation of images.

**5.3.4 Random resized crop:-**

We randomly crop and resize the input images to a size of 150x150 pixels, with a random scale and aspect ratio for improving model’s ability to learn invariant features.

**5.3.5 Color Jitter:-**

The color of the input images are adjusted randomly by adjusting brightness, contrast and saturation.

**5.3.6 Normalize:-**

Normalize the pixel values of the input images using specific mean and standard deviation values to standardize the pixel intensities across all images in the dataset.

**5.3.7 Categorical Encoding:-**

Convert the name of bird to numeric value (labels)

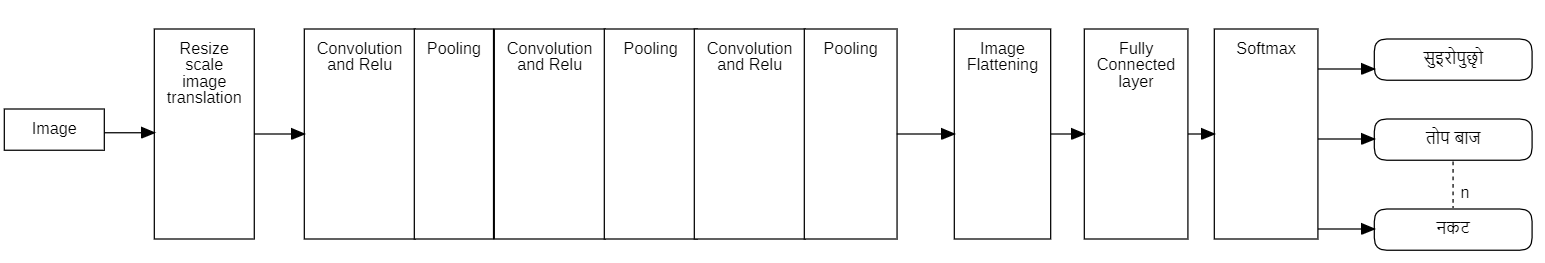
**5.3.8 Representation and Description:-**

Mapped their unique labels into their respective names in Nepali as well as English.

### 5.4 Architecture: -

**5.4.1 Convolutional Neural Network (CNN): -**

Convolutional neural network (CNN) is designed for image classification. It consists of three convolutional layers followed by batch normalization, ReLU activation, max-pooling, and dropout regularization. The model processes input images through these layers to extract features and make predictions about the class labels. It utilizes a feedforward architecture, with information flowing from inputs to outputs, and employs convolution and pooling layers specifically tailored for image analysis tasks. The model is trained on a dataset containing images of various classes, enabling it to learn and recognize patterns associated with different objects or categories. Once trained, the model can be used to classify new images, providing valuable insights and predictions based on the learned features. So, by the image, we can obtain information and predict the species from the trained model with the help of Figure 1.



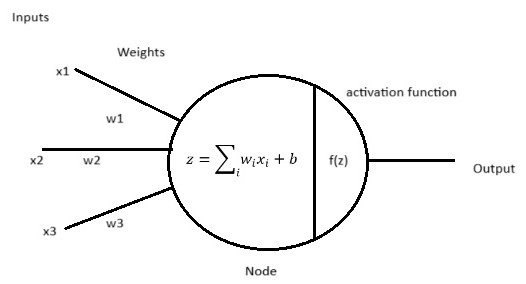
*Fig 5.3.1.: Block diagram of system*

**5.4.2 Feature Extraction:-**

Features are found whenever the network performs a sequence of convolutional and pooling procedures.

**5.4.2.1 Convolutional layer:-**

Convolutional layers, pooling layers, and fully connected layers are the three main types of convolutional neural network layers. Convolutional layers perform operations on their inputs before passing the result to the next layer. It applies a set of learnable filters known as the kernels of size 3x3 to the input images. It slides over the input image data and computes the dot product between kernel weight and the corresponding input image patch. The output of this layer is referred to as feature maps. Here we use a total of 12 filters for this layer we’ll get an output volume of dimension 32 x 32 x 12.



*Fig5.3.2: Activation Function*

**5.4.2.2 Activation layer:-**

By adding an activation function to the output of the preceding layer, activation layers add nonlinearity to the network. It will apply an element-wise activation function to the output of the convolution layer.

One common activation function is ReLU.

R(z)=max(0,z)

ReLU is half rectified (from bottom).

The ReLU activation function is differentiable at all points except at zero. For values greater than zero, we just consider the max of the function. All the negative values default to zero, and the maximum for the positive number is taken into consideration.

For the computation of the backpropagation of neural networks, the differentiation for the ReLU is relatively easy. The only assumption we will make is the derivative at the point zero, which will also be considered as zero. This is usually not such a big concern, and it works well for the most part. The derivative of the function is the value of the slope. The slope for negative values is 0.0, and the slope for positive values is 1.0.

**5.4.2.3 Pooling layer:-**

Local or global pooling layers of convolutional networks integrate the output of groups of neurons in one layer into individual neurons in the next layer. It assists in minimizing the size of the feature maps and increases the model's resistance to slight changes in the input pictures.

We use max pooling in our neural network. Max pooling is a downs sampling technique commonly used in convolutional neural networks (CNNs) to reduce the spatial dimensions of an input volume.

**5.4.2.4 Fully Connected layer:-**

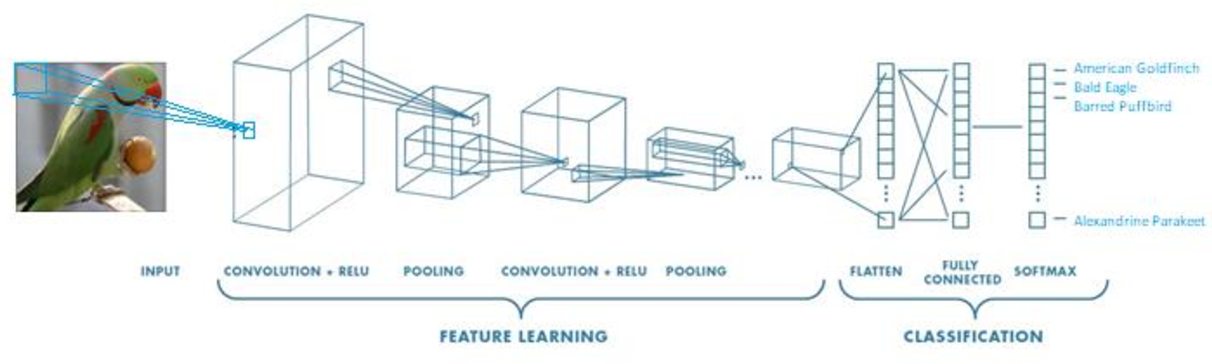
Connected layers allow every neuron in one layer to further communicate with every neuron in every other layer. It takes the input from the previous layer and computes the final classification or regression task. We use softmax function as:

Softmax layer translate the number into probability distribution.

**5.4.2.5 Classification:-**

Putting anything or someone right into a certain organization or system primarily based on unique standards is the definition of classifying.

The classification module takes the pre-processed images and predicts the bird species. The results are then displayed to the user. The accuracy of the classification depends on the quality of the dataset, the training algorithm used, and the size of the training dataset.



*Fig 5.3.3: Convolutional operation*

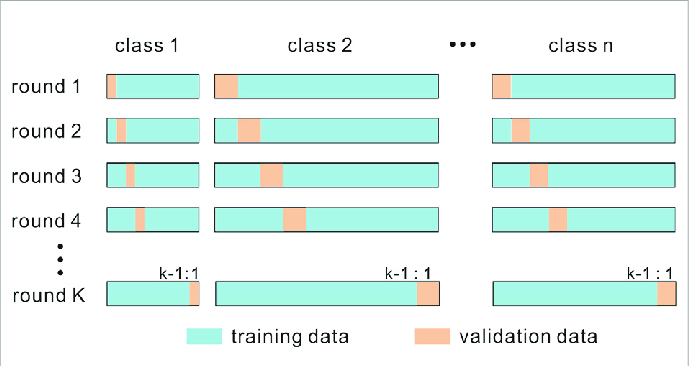
*Source:* *https://www.semanticscholar.org/paper/IMAGE-BASED-BIRD-SPECIES-IDENTIFICATION-USING-DEEP-Siva-Reddy/9a8bcba6b02a8973f9f99217bfce6b67aa5980e7*

#### 5.5 Optimizer

The optimizer plays a crucial role in training neural networks. In this implementation, the Adam optimizer is utilized, which is an adaptive learning rate optimization algorithm. It efficiently computes adaptive learning rates for each parameter based on their first and second-order moments of gradients. The optimizer is initialized with a learning rate of 0.0001 and a weight decay of 0.001 to encourage regularization during training. These hyper parameters can be adjusted based on the specific requirements of the training task and the characteristics of the dataset.

**5.6 Cross validation:**

Cross-validation is a robust technique used to assess the performance and generalization ability of machine learning models. In this implementation, k-fold cross-validation is employed, where the dataset is divided into k subsets or folds. The model is trained k times, each time using k-1 folds for training and the remaining fold for validation. This process helps in obtaining a more reliable estimate of the model's performance by reducing bias and variance. Stratified k-fold cross-validation is used here to ensure that each fold maintains the same class distribution as the original dataset, thus avoiding potential biases. Early stopping is also implemented, which monitors the validation loss and terminates training if no improvement is observed for a certain number of epochs, thereby preventing overfitting and improving efficiency.



*Fig 5.5: Cross validation*

#### 5.7 Loss Function:

Our model employs the Cross-Entropy Loss function as its primary measure of performance during training. This loss function, tailored for multi-class classification tasks, quantifies the disparity between predicted class probabilities and true class labels. By penalizing incorrect classifications more heavily, Cross-Entropy Loss incentivizes the model to learn discriminative features essential for accurate predictions. This choice aligns with the nature of the image classification task undertaken by our model, facilitating efficient training and robust performance across diverse classes. Adjusting the loss function according to specific task requirements and dataset characteristics can significantly impact the model's convergence and final performance.

## CHAPTER 6

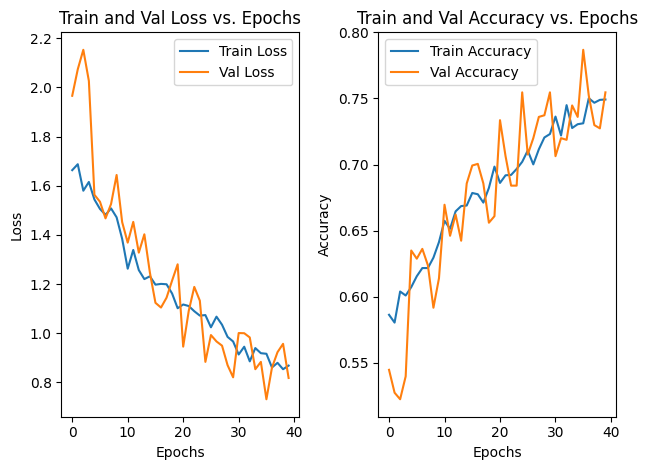
## RESULT AND ANALYSIS

### 6.1 Dataset:

We gained our dataset from internet as endangered bird species dataset. This dataset contains the data of 38 birds’ species with their scientific names and about 8073 images are there in training dataset and 402 images for testing dataset. We changed the name of each birds to unique labels so that we can map them to their respective Nepali names.

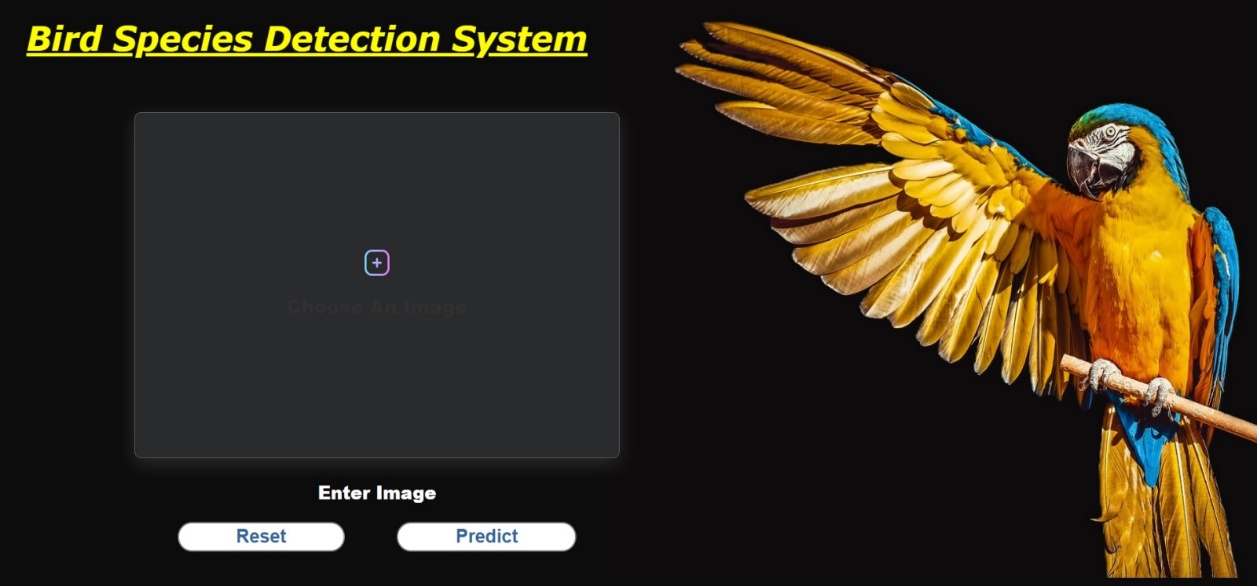
### 6.2 Result:

Based on our codes and model design we have obtained different results on dataset. We have got the train accuracy as about 73% and Validation accuracy as 76%. Here is the accuracy graph of our model.

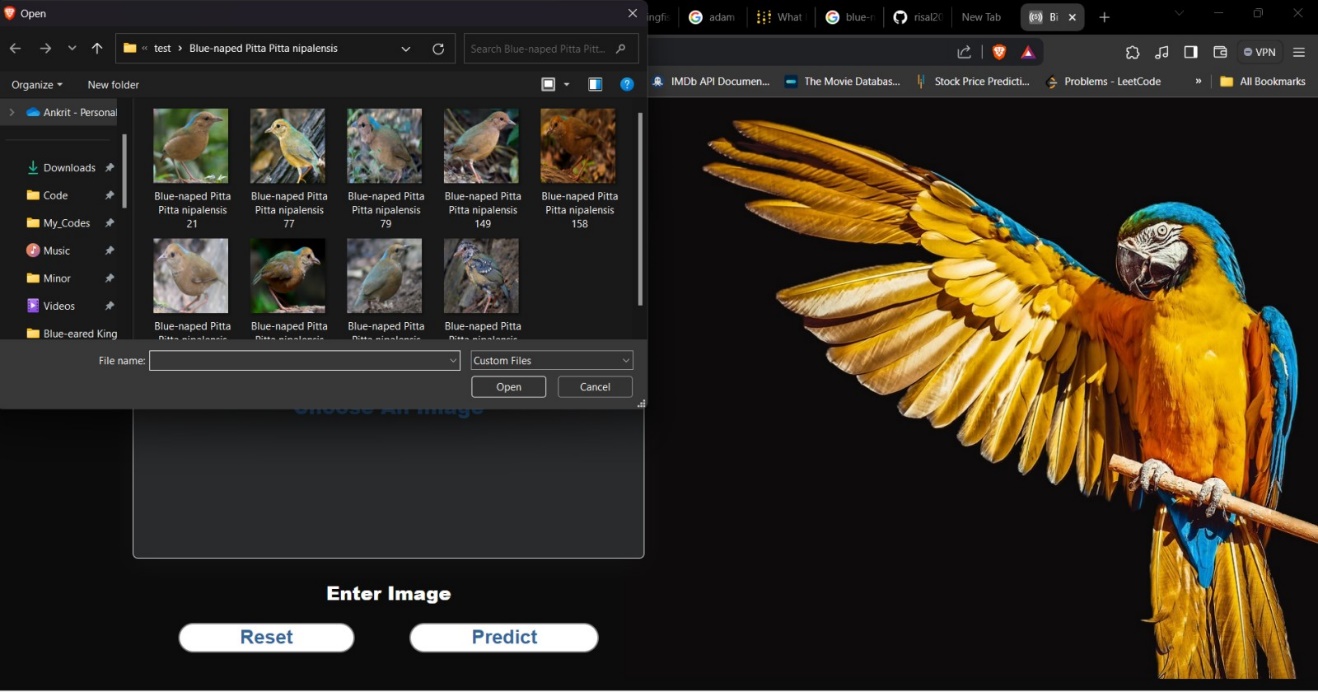


*Fig 6.1: Train and validation losses and accuracy graph*

**6.3 Result Obtained:**



*Figure 6.2: User interface*

****

*Figure 6.3 Upload photo*

**

*Figure 6.4 Bird Recognition*

**CHAPTER 7**

## CONCLUSION, LIMITATIONS AND FUTURE ENHANCEMENT

**7.1 Conclusion**

A Convolutional Neural Network (CNN) has been developed to accurately classify bird species from user-input images. Using this model we have achieved an accuracy of 77%. In this model we have integrated Cross-validation which helps to achieve higher accuracy and hyper-parameters tuning is done to identify the parameter values that is suitable for our model performance.

**7.2 Limitations**

There are a few drawbacks regarding bird species classification in our system:

* Our model is based on a limited dataset (38 bird species only).
* Cannot provide the correct information if more than one bird is present in a single picture.
* Can give wrong information if the image is not clear (blured).

## 7.3 Future Enhancements

There are serval possible future enhancements that could be considered for Bird Species Recognition System:

* Collecting and using a larger, more recent dataset to train and evaluate the models.
* Integrate information from audio recording or environmental data.
* Combine prediction from variation of CNN models to improve performance.

## REFERENCES

[1] Seppo Fagerlund,” *Bird Species Recognition Using Support Vector Machines*”, EURASIP Journal on Advances in Signal Processing volume 2007, Article number: 038637 (2007), Accessed: Dec. 19, 2023. [Online].

[2] LeCun, Y. Boser, B. Denker, J. S. Henderson, D. Howard, R. E. Hubbard, W. Jacker, L. D. (June 1990). "*Handwritten digit recognition with a back-propagation network*"

[3] Diederik P. Kingma, Jimmy Ba ,” *Adam: A Method for Stochastic Optimization*”, Published as a conference paper at the 3rd International Conference for Learning Representations, San Diego, 2015. Accessed: Dec. 20, 2023. [Online].

[4] Sebastian Ruder,” *An overview of gradient descent optimization algorithms*”, Insight Centre for Data Analytics, NUI Galway, Accessed: Dec. 22, 2023. [Online].

[5] Payam Refaeilzadeh, Lei Tang & Huan Liu, “*Cross Validation*”, Encyclopedia of Database Systems pp 532–538, Accessed: Dec. 19, 2023. [Online].

[6] “*Volume 1 The Status of Nepal’s Birds: The National Red List Series.*” Accessed: Dec. 19, 2023. [Online]. Available: https://cms.zsl.org/sites/default/files/2023-02/ZSL%20National%20Red%20List%20of%20Nepal%27s%20Birds%20Volume%201%202016.pdf

[7] D. Siri, S. Desu, K. Alladi, D. Swathi, S. Singh, and V. Srilakshmi, “*Automated System for Bird Species Identification Using CNN*,” E3S Web of Conferences, vol. 430, p. 01057, 2023, doi: https://doi.org/10.1051/e3sconf/202343001057.

[8]R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, “*Convolutional Neural networks: an Overview and Application in Radiology*,” Insights into Imaging, Jun. 2018, doi: https://doi.org/10.1007/s13244-018-0639-9.

[9] J. Brownlee, “*A Gentle Introduction to k-fold Cross-Validation*,” Machine Learning Mastery, Oct. 04, 2023. https://machinelearningmastery.com/k-fold-cross-validation/

‌