ISSN [Online]: 2583-2654

# Bird Species Identification using Audio Signals with Machine Learning & Neural Network Approach

*Mrs. Mercy Flora Pritheba, Pradeep S, S Rajesh kumar, Sachin R, Sanjay P*
*Department of CSE, Dr.T. Thimmaiah Institute of Technology, KGF, Karnataka, India*

***Abstract:*** The avian biodiversity plays a crucial role in maintaining ecological balance and understanding their behavior is pivotal for conservation efforts. This project proposes a novel approach for Bird Species Identification using Audio, leveraging advanced Machine Learning techniques and Neural Networks. The primary focus lies on employing the Ridge classifier in conjunction with an Artificial Neural Network (ANN) to achieve high accuracy and robustness in classification tasks. Spectrogram extraction technique will be employed to convert audio samples into numerical data, facilitating the analysis of frequency-time patterns essential for bird species differentiation. The Ridge classifier will be initially employed for its superior handling of multicollinearity, ensuring effective feature selection and noise reduction. Subsequently, an ANN will be implemented to learn intricate patterns within the spectrograms, enhancing classification precision. Through this innovative methodology, we aim to contribute to the field of ornithology and ecological research by providing a reliable and efficient tool for automated bird species identification.

***Keywords: Machine learning, Pre-process, Automatic identification, Spectrogram and Classification.***

## I. INTRODUCTION

The Earth's biodiversity faces unprecedented threats, making conservation efforts more critical than ever. Among the myriad of species inhabiting our planet, birds play a crucial role in maintaining ecosystem balance and biodiversity. However, monitoring avian populations for conservation purposes has traditionally been a labor-intensive task, often relying on visual identification methods that are time- consuming, expensive, and impractical in certain habitats. In recent years, the integration of technology and artificial intelligence has provided innovative solutions to address these challenges. This aims to leverage the power of machine learning and artificial neural networks (ANN) to develop an intelligent system capable of automatically identifying bird species through their distinctive audio signals.

Bioacoustics, the study of sound in the natural world, has emerged as a powerful tool for monitoring and understanding biodiversity. Birds, in particular, are known for their diverse and characteristic vocalizations, which vary across species and serve various purposes, including communication, mating rituals, and territorial defense. Audio signals, therefore, present a rich source of information that can be harnessed for non-intrusive and efficient bird species identification.

This project's main goal is to develop a reliable and precise system that can recognise different bird species based only on their distinctive auditory signatures.. The proposed approach involves a comprehensive pipeline, including data collection, preprocessing, feature extraction, and the application of machine learning algorithms, particularly artificial neural networks. By developing and implementing this system, we aim to contribute to the field of bioacoustics and provide a valuable tool for researchers, conservationists, and wildlife enthusiasts.

ISSN [Online]: 2583-2654

The first phase of the project entails the collection of a diverse dataset of bird audio recordings. These recordings will capture the vocalizations of various bird species in different environments and conditions, ensuring the model's adaptability to real-world scenarios. The dataset will be curated to include a wide range of avian species, considering factors such as habitat, geographical location, and time of day to enhance the model's generalization capabilities.

Subsequently, the collected audio data will undergo rigorous pre-processing to remove noise, standardize formats, and extract relevant features. Feature extraction is a crucial step in transforming raw audio signals into meaningful representations that can be effectively used by machine learning algorithms. Commonly employed features include spectrograms, mel-frequency cepstral coefficients (MFCCs), and other time-frequency representations that capture the unique characteristics of each bird species' vocalizations.

The system's core function is the classification of bird species using artificial neural networks and machine learning. Using the pre-processed dataset as training data, the model will discover which bird species are associated with particular acoustic attributes. The performance of the model could be enhanced by looking into transfer learning strategies. This would make it possible for the system to identify new and undiscovered species more accurately by utilising the information gained from one set of species.

The significance of this project extends beyond the realm of scientific research. The developed system has practical applications in ecological monitoring, habitat management, and conservation efforts. By automating the process of bird species identification, researchers and conservationists can efficiently gather data on avian populations, assess the health of ecosystems, and implement targeted conservation strategies. Additionally, the non-intrusive nature of audio- based monitoring minimizes disturbance to wildlife, making it an ethical and sustainable approach to biodiversity research.

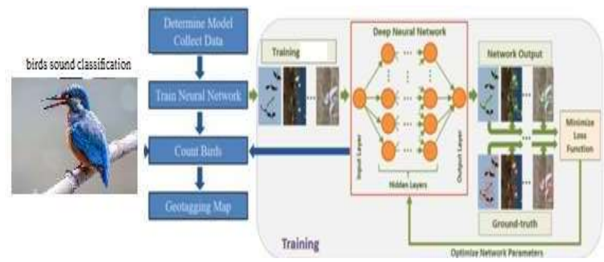


**Fig 1: The complete workflow of the automatic bird monitoringsystem.**

## II RELATED WORK

Chih-Yuan et al. [1] Recent advancements in deep learning have significantly enhanced the accuracy of bird species identification from audio recordings, benefiting bird conservation efforts. In the BirdCLEF2019 competition, participants were tasked with designing a system capable of recognizing 659 bird species from a staggering 50,000 audio recordings. The challenges encompassed memory management, recognizing numerous bird species, and dealing with signal-to-noise ratio disparities between training and testing sets. To meet this challenge, we employed two prominent convolutional neural network architectures and the inception model. The inception model excelled with a classification mean average precision (c-mAP) of 0.16, securing the second position among five successful submissions. This progress underscores the potential for deep learning in advancing avian research and conservation.

Mehyadin et al. [2] the bird classifier employs advanced machine learning technology, utilizing a specialized method to record and categorize bird calls. By focusing solely on bird sound recordings, species identification becomes more streamlined. The system additionally offers resources for species classification, enabling automated species detection and training the machine in species recognition. Undesirable noises are effectively filtered and organized into datasets, each undergoing noise suppression and a distinct classification process. The Mel-frequency cepstral

ISSN [Online]: 2583-2654

coefficient (MFCC) is employed and tested with various algorithms like Naïve Bayes, J4.8, and Multilayer perceptron (MLP) for species classification. J4.8 achieves the highest accuracy at 78.40%, outperforming others in both accuracy and processing time (39.4 seconds)..

Dharaniya R et al. [3] Bird species identification using Images holds paramount significance in preserving biodiversity and sustaining ecosystems. Birds play multifaceted roles, contributing to agriculture, shaping landscapes, and even fostering coral reef growth. In the realm of ecology, the accurate identification and observation of bird species are pivotal. Thanks to the progress in machine learning, automating bird species classification has become more accessible. Technological advancements have spurred the creation of diverse bird species classification systems, often integrated into user-friendly web or mobile applications for result presentation. This paper delves into the specifics of implementing bird species identification through the utilization of a Convolutional Neural Network (CNN).

Stefan Kahl et al. [4] the task of identifying bird species in audio recordings presents a formidable research challenge. This paper outlines a method for classifying bird sounds on a large scale, specifically addressing the LifeCLEF 2017 bird identification task. By employing a range of convolutional neural networks, we extracted features from visual representations of field recordings. The training dataset, BirdCLEF 2017, encompassed 36,496 audio recordings spanning 1500 distinct bird species. Our approach yielded notable results with a mean average precision of 0.605 (official score) and 0.687 when exclusively considering foreground species.

In this research [5] by Acconcjaioco et al., the introduction of one-shot learning in computational bioacoustics addresses the challenge of bird species identification in dynamic environments with incomplete class dictionaries. The proposed framework employs a Siamese Neural Network on logMel spectrograms, allowing for the real-time detection and incorporation of novel classes. Through evaluations on diverse bird datasets, the study demonstrates the state-of-the-art performance of the framework in both stationary and non- stationary conditions, utilizing a standard audio representation. The method's efficacy is attributed to its simultaneous consideration of similarities and dissimilarities to known classes. The research extends its application to analogous problems, explores optimal data quantities for non-stationary performance, and investigates strategies for optimal class selection during training.

Samruddhi Bhor et al. [6] have designed a model for automatic bird species identification using Convolutional Neural Network, TF-Slim Tensor files, and spectrograms. Bird species can be categorized depending on the spectrogram image created from their voices using this method. Furthermore, this technique allows for a greater number of classes to be worked on when identifying and classifying bird species, resulting in more accurate findings and classifications.



Thiago L. F. Evangelista et al. [7] have created a machine learning algorithm-based automated segmentation method for audio signals in order to identify bird species. It deals with the segmentation phase of the problem of automatically identifying bird species. In their upcoming study, they want to use the automatic segmentation process in accordance with the framework put forth in "Hierarchical Classification of Bird Species using their Audio Recorded Songs" by C. N. Silla Jr. and C. A. A. Kaestner for the hierarchical categorization of bird species.

Vemula Omkarini et al. [8] have designed an automated model based on deep neural networks that automatically identifies the species of a bird given as the test dataset. This design employs Neural Network, Image Classification, DNN, and

Machine Learning algorithms.

Ilyas Potamitis et al. [9] have developed a technique that uses tools and software to automatically identify bird noises in extensive real-field recordings. To confirm the presence of a species in a large corpus of audio recordings in a reasonable amount of time, taxon-specific call and song detectors based on statistical models must be used. They provide their processing code and a large corpus of audio recordings so that other researchers can more easily conduct and assess similar experiments utilising real-field data. Future advancements in this and other computational ecology techniques will lead to the establishment of suitable conservation plans for creatures facing extinction.

Marcelo T. Lopes et al. [10] have designed a model for automatic bird species identification from bird audio recorded songs using signal processing, pattern recognition, machine learning, and bird species identification. For future scope, the idea is to create a simple equipment with some built-in processing capacity to collect acoustic information that will be transmitted remotely to a central monitoring station. The collected data would allow ornithologists to monitor environmental conditions and study specific bird species

Dan Stowell et al. [11] Bird Detection in Audio: A Survey and a Challenge. This survey has described current approaches to automatic bird detection in audio, including the current level of generality. They have introduced a challenge giving researchers an opportunity to create a step change in these directions. A wide variety of methodological options remains open to further study, such as recent innovations indeep learning, or meta-algorithms that can automatically select detectors or combine their outputs.

XIE Jiang-jian DING Chang-qing et al. [12] have created a model for identifying bird species that extracts features using the VGG-16 model (pre-trained on ImageNet). A Softmax layer and two fully-connected hidden layers make up the classifier. Deep learning, multi-channel, and spectrogram features serve as the foundation for this model.

Zhang et al. [13] have designed a model based on Automatic Recognition of Bird Species by Their Sounds. In this paper smart sampling approach is used for bird classification. Future work will focus on developing more descriptive features to differentiate species compositions in audio files. Segmenting clips into shorter durations provides one way to characterize more detailed bird vocalizations.

## III METHODOLOGY

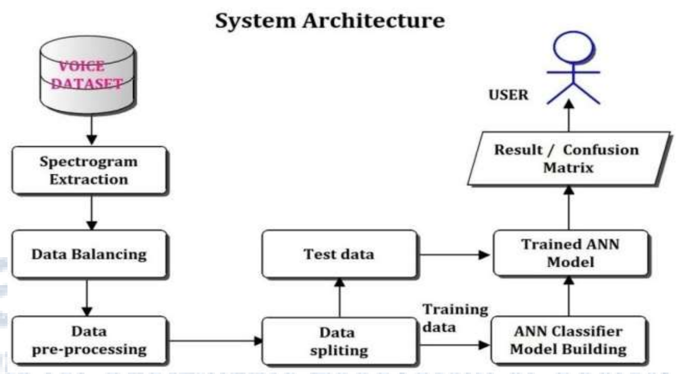


**Fig 2: System Architecture**

The proposed architecture of the system designed for Bird Species Identification using an Artificial Neural Network (ANN) Classifier encompasses several essential components. Initially, the system acquires audio recordings of bird species, initiating a preprocessing stage that involves tasks such as data balancing and feature extraction. This typically takes the form of generating balanced spectrograms. The features extracted serve as input for the ANN classifier, the foundational element of the system.

Inspired by neural structures in the human brain, the ANN classifier undergoes a dedicated training phase to learn intricate patterns. During this phase, labeled training data is supplied to the ANN, allowing it to adjust internal parameters for precise

ISSN [Online]: 2583-2654

species classification. Following training, the system's performance is assessed using separate validation and test sets to evaluate its generalization capabilities.

To enhance results, post-processing techniques may be applied, and an interactive user system can provide a platform for user engagement. The system then presents identified bird species, accompanied by pertinent information, serving as a valuable tool for various applications ranging from research to conservation efforts. Continuous monitoring and refinement mechanisms are in place to ensure the system's ongoing accuracy and effectiveness over time.

### *A. Dataset Description*

Gather a diverse dataset of bird audio signals, ensuring a wide representation of species.

Preprocess the audio data by extracting relevant features, such as spectrograms, to be used as input for machine learning models.

### *B. Data Labeling*

Annotate the audio files with their corresponding bird species labels

- Manually label the audio dataset with accurate bird species information for supervised learning.
- Ensure a balanced distribution of samples across different bird species.

### *C. Data Pre-processing*

- Preprocess the audio data for feature extraction. This may include tasks like noise reduction, resampling, and filtering.
- Noise removal: Short-Time Fourier Transform (STFT) is a technique used for analyzing the frequency content of a signal over short, overlapping time windows. It provides a time-varying representation of the signal's frequency content, making it useful for applications like audio processing, speech analysis, and signal processing. The STFT is a modification of the standard Fourier Transform that allows for the analysis of non-stationary signals.

Here is the STFT technique:

#### *1. Windowing:*

Divide the signal into short, overlapping time segments or windows. The choice of window size and overlap depends on the characteristics of the signal and the analysis requirements.

#### *2. Apply Fourier Transform to Each Window:*

Apply the Fourier Transform to each windowed segment of the signal.

The Fourier Transform converts the signal from the time domain to the frequency domain, providing information about the frequency content at different points in time.

#### *3. Time-Frequency Representation:*

The result is a time-frequency representation of the signal, where each point in the spectrogram corresponds to the magnitude or power of a specific frequency component at a particular time.

#### *4. Overlapping Windows:*

Overlap the windows to ensure that the time-frequency representation is smooth and captures changes in the signal over time.

Overlapping helps mitigate the effects of windowing on abrupt changes in the signal.

The mathematical representation of the STFT for a signal x(t) is given by:

$$X(f,\tau) = \int_{-\infty}^{\infty} x(t).\, w(t-\tau).\, e^{-j2\pi f t}\, dt$$

Where:

$X(f,\tau)$ is the STFT of the signal.

-x(t) is the input signal.

-w(t) is the window function.

- t is the time index.
- f is the frequency.

In practice, the STFT is often computed using the Fast Fourier Transform (FFT) algorithm for efficiency. The resulting time- frequency representation (spectrogram) can be used for various purposes, including audio analysis, feature extraction, and noise removal.

ISSN [Online]: 2583-2654

### *D. Feature Extraction*

The process of creating spectrograms, which are visual depictions of a signal's frequency spectrum as it evolves over time, involves converting the audio signals. Mel-frequency MFCCs, or Cepstral Coefficients: Take out characteristics that are frequently utilised in audio and voice processing. Statistical Features: Take the audio data and extract statistical features such as mean, standard deviation, etc.

***Audio signal features that are required and to be extracted using python's librosa package:***

1. Chroma STFT
2. RMSE (root mean square energy)
3. Spectral centroid
4. Roll off
5. Zero crossing rate
6. Mel- frequency ceptral coefficients (20)

### *E. Algorithm used*

#### *1. Ridge Classifier*

The Ridge Classifier is a linear classification algorithm widely used in machine learning. It extends logistic regression byadding a regularization term, known as the L2 penalty, to the loss function. This penalty helps prevent overfitting and stabilizes the model's coefficients, particularly when dealing with multicollinear features. The Ridge Classifier is valuable for tasks like bird species identification, as it effectively handles datasets with correlated variables. By striking a balance between bias and variance, it provides robust and accurate predictions in scenarios where traditional logistic regression may struggle

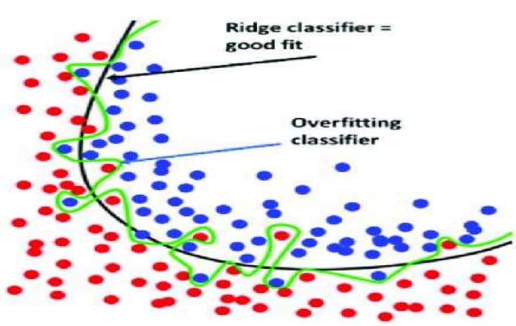


**Fig 3: Ridge Classifier**

***Training Ridge Classifier:***

- Implement a Ridge classifier for initial bird species identification.
- Train the classifier using the pre-processed audio data andevaluate its performance.
- Split the Data: Use train_test_split to split your dataset intotraining and testing sets.
- Standardize Features: Standardize the features using StandardScaler. This step is optional but often helps improve the performance of linear models.
- Create and Train the Ridge Classifier: Use the Ridge class.

rom scikit-learn to create a Ridge Classifier. Set the regularization strength (alpha) as needed. Train the classifierusing the training data.

-Make Predictions: Use the trained classifier to make predictions on the test set.

-Evaluate Performance: Use metrics like accuracy, precision, recall, and F1-score to evaluate the performance of your Ridge Classifier.

- Replace X and y with your actual feature matrix and label array. Additionally, fine-tune the alpha parameter based on the data.

#### *2. Artificial Neural Network*

A computational model that mimics the composition and operation of the human brain is known as an artificial neural network, or ANN. It is made up of layers of interconnected nodes (neurons), which are arranged as an input layer, hidden layers, and output layer. Every node-to-node link has a weight attached to it that changes during training. Because ANNs are so adept at identifying complex patterns in data, they are used in a wide range of applications, including pattern recognition, regression, and classification. They are useful in several fields, such as audio-based bird species identification in this instance, picture and speech recognition, and natural

ISSN [Online]: 2583-2654

language processing.

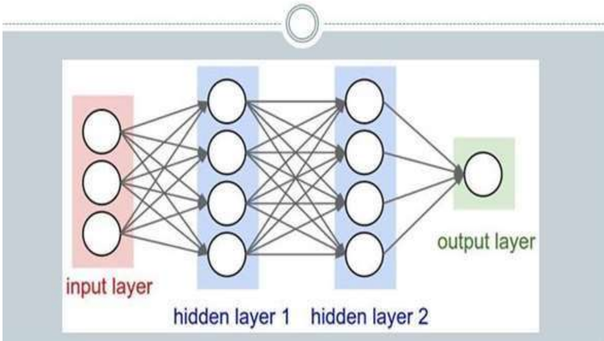


**Fig 4: Artificial Neural Network**

- Implement and train the ANN using the pre-processed audio dataset.
- Fine-tune hyper parameters and optimize the network for accurate bird species classification.

1. Model Architecture: adjust the number of neurons, layers, and activation functions.
2. Dropout Layer: The dropout layer is included for regularization, which helps prevent overfitting.
3. Early Stopping: Early stopping is employed to monitor the validation loss and stop training when it starts to increase, preventing overfitting.
4. Compile the Model: Specify the optimizer, loss function, and evaluation metric.
5. Train the Model: Train the model using the training data and monitor its performance on the validation set.
6. Evaluate the Model: Evaluate the final model on the test set and print the accuracy.

- adjust hyper parameters, the model architecture, or incorporate more advanced techniques based on the specific dataset and requirements.
- Input layer : Let X be the input vector with n features. The input layer simply passes the input to the hidden layer:

$$Z^{[1]} = X$$

- Hidden layer: For the hidden layer, we compute the weighted sum of inputs and apply an activation function :

$$A^{[1]} = W^{[1]} \cdot Z^{[1]} + b^{[1]}$$

Here:

- $W^{[1]}$ is the weight matrix for the hidden layer.
- $b^{[1]}$ is the bias vector for the hidden layer.
- $A^{[1]}$ is the linear activation of the hidden layer.

Then, we apply an activation function $g^{[1]}$ element-wise:

$$Z^{[2]} = g^{[1]}(A^{[1]})$$

- Output layer : For the output layer, we repeat a similar process:

$$A^{[2]} = W^{[2]} \cdot Z^{[2]} + b^{[2]}$$

Here:

- $W^{[2]}$ is the weight matrix for the output layer.
- $b^{[2]}$ is the bias vector for the output layer.
- $A^{[2]}$ is the linear activation of the output layer.

Finally, we apply a suitable activation function g[2] (softmax for multi-class classification)

$$\hat{Y} = g^{[2]}(A^{[2]})$$

- Loss-function:

The choice of the loss function depends on the task. For classification problems, cross-entropy loss is often used. The loss function is typically denoted as $L(\hat{Y}, Y)$, where $\hat{Y}$ is the predicted output and $Y$ is the true label.

- Training:

During training, the network's parameters (weights and biases) are updated using optimization algorithms like gradient descent. The goal is to minimize the loss function by adjusting the parameters.

### *F. Model Evaluation and Validation:*

- Evaluate the performance of both the Ridge classifier and ANN using metrics like accuracy, precision, recall, and F1 score. Validate models with a separate test dataset to assess their generalization capabilities.

ISSN [Online]: 2583-2654

### G Evaluation Metrics

Confusion matrix, accuracy score, precision, recall, sensitivity, and F1 score are employed in the evaluation procedure. A table-like structure known as a confusion matrix has true values and anticipated values, often known as true positive and true values. It is defined in four parts: the first one is true positive (TP) in which the values are identified as true and, in reality, itwas true also. untrue positive (FP) is the second type, when values that are labelled as true are actually untrue. False negative (FN) is the third type, when the value was true but was reported as negative. The number in the fourth example was negative and was accurately detected as such; it is a true negative (TN). The table is shown in Figure 5.

A positive, a negative, TP, FN, FP, TN, and FN are the false negative, false positive, and true positive, respectively, in Figure 5. Then an accuracy score is used to evaluate the performance of a model. Its definition is the division of the true positive and true negative values by the true positive and true negative and the false positive and false negative. The formula is

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad ..........(3)$$

After accuracy, there is a Precision metric that measures the accuracy of positive predictions made by the model. The formula is

$$precision = \frac{TP}{FP+TP} \quad ................(4)$$

Next is sensitivity, which is the percentage of real positive cases that were correctly predicted to be positive (or true positive). Another name for sensitivity is recall. Stated otherwise, the likelihood of an individual being unhealthy is high. The formula is

$$Recall = \frac{TP}{TP+FN} \quad ..........(5)$$

The harmonic mean of recall and precision is the F1 score. It offers a harmony between recall and precision, and the formula is

$$F1 = 2\,X\,\frac{\text{Precision} \times \text{RecallPrecision}}{\text{Precision} + \text{RecallPrecision}}$$

| | | Predicted value P | Predicted value N |
|---|---|---|---|
| True value | P | TP | FN |
| | N | FP | TN |

**Figure 5: Confusion Matrix**

## IV RESULT AND ANALYSIS

The validation accuracy serves as a crucial metric to evaluate the overall performance of the models in correctly classifying birds species using their audio. The ANN model achieved a validation accuracy of 96.50%,

The classification report provides a more detailed analysis of the model's performance by presenting precision, recall, and F1 scores for each class.

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| Acrocephalus | 0.92 | 0.82 | 0.87 | 200 |
| Bubo | 0.81 | 0.95 | 0.88 | 200 |
| Caprimulgus | 1.00 | 0.92 | 0.96 | 200 |
| Emberiza | 0.91 | 0.94 | 0.92 | 200 |
| Ficedula | 0.81 | 0.70 | 0.75 | 200 |
| Glaucidium | 0.98 | 0.91 | 0.94 | 200 |
| Hippolais | 0.68 | 0.81 | 0.74 | 200 |
| accuracy | | | 0.86 | 1400 |
| macro avg | 0.87 | 0.86 | 0.86 | 1400 |
| weighted avg | 0.87 | 0.86 | 0.86 | 1400 |

**Fig 6: Classification report of RC model**

ISSN [Online]: 2583-2654

***The Classification report of ANN is given below:***

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| Acrocephalus | 0.98 | 1.00 | 0.99 | 200 |
| Bubo | 1.00 | 0.99 | 0.99 | 200 |
| Caprimulgus | 0.99 | 1.00 | 1.00 | 200 |
| Emberiza | 1.00 | 1.00 | 1.00 | 200 |
| Ficedula | 0.88 | 0.96 | 0.92 | 200 |
| Glaucidium | 1.00 | 1.00 | 1.00 | 200 |
| Hippolais | 0.98 | 0.88 | 0.92 | 200 |
| accuracy | | | 0.97 | 1400 |
| macro avg | 0.98 | 0.97 | 0.97 | 1400 |
| weighted avg | 0.98 | 0.97 | 0.97 | 1400 |

**Fig 7: classification report of ANN model**

For the Ridge classifier model, the precision for a bird species Acrocephalus class is 0.92, indicating that 96% of the result is classified correctly. The recall for this class is 0.82.

For the ANN model, the precision for the same bird species is 98%. The recall for this class is 100% that indicates it predicted the class precisely.

The ANN model outperforms Ridge classifier model as the ANN model predicted acrocephalus 98% correctly but the Ridge classifier has an accuracy of 92%.

A result is the final consequence of actions or events expressed qualitatively or quantitatively. Performance analysis is an operational analysis, is a set of basic quantitative relationship between the performance quantities.

The result analysis and the accuracy score of the ANN model is calculated for model assessment. The accuracy of ANN model is given:

Result Analysis

```
In [24]: class_labels = ['Acrocephalus', 'Bubo', 'Caprimulgus', 'Emberiza', 'Ficedula', 'Glaucidium', 'Hippolais']
```

Accuracy Score

```
In [25]: ann_model_accuracy = accuracy_score(y_true=true_labels, y_pred=prediction)
         print("Validation accuracy of ArtificialNeuralNetwork model is {:.2f}%".format(ann_model_accuracy*100))

Validation accuracy of ArtificialNeuralNetwork model is 99.64%
```

**Fig 8: result analysis and accuracy**

The confusion matrix provides a visual representation of the model's performance, showing the number of correctly and incorrectly classified instances for each class.

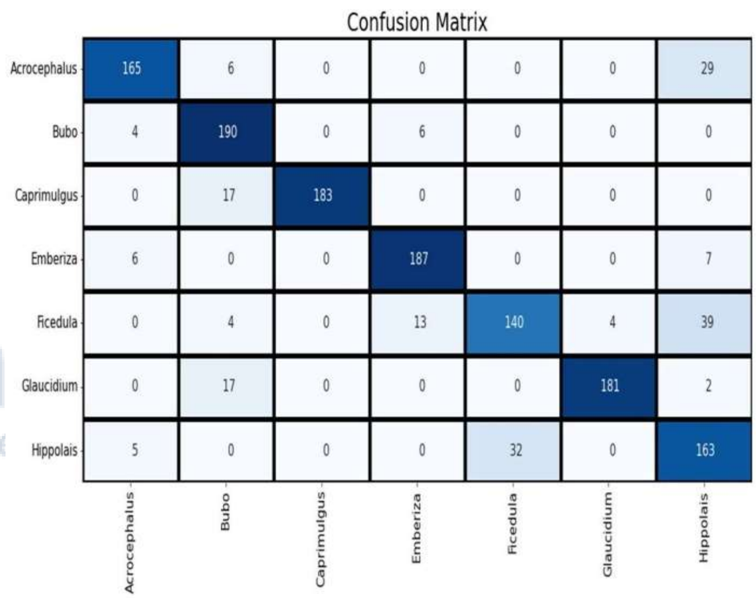


**Fig: 9: Confusion matrix of Ridge classifier model**

Consider the bird species acrocephalus, out of 200 instances the model classified 165 correctly and misclassified 35 others. Similarly for the bird species bubo the model classified 190 instances out 200 instances correctly and misclassified the other 10.

The confusion matrix of ANN model given below and the matrix provides a visual representation of the model's performance, showing the number of correctly and incorrectly classified instances for each class.

ISSN [Online]: 2583-2654

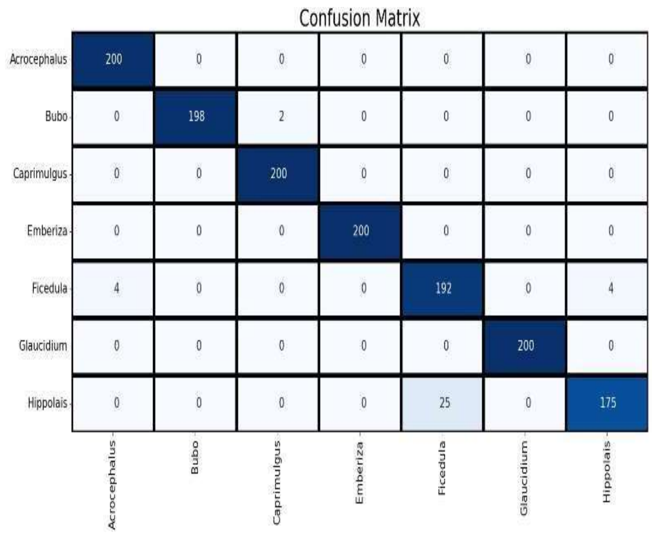


**Fig 10: confusion matrix of ANN model**

The training and testing accuracy plots provide numerical insights into the models' performance during the training process.

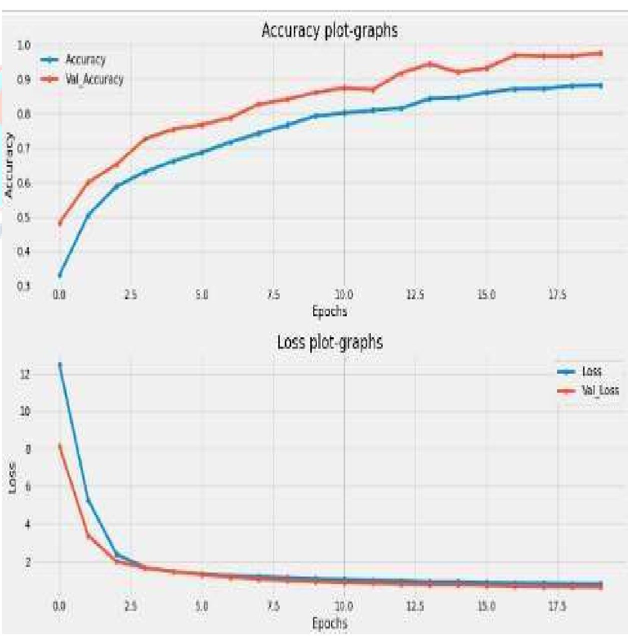


**Fig11: The Accuracy plot of ANN model**

The accuracy plot, on the left, shows the model's accuracy on the training data (red line) and validation data (blue line) as the number of epochs increases. In general, as the number of epochs increases, the model's accuracy on the training data increases. This is because the model is continually learning and improving its ability to classify videos correctly based on the training data. However, it is important that the model also generalizes well to unseen data. The validation accuracy helps to assess this. Ideally, the validation accuracy should also increase as the number of epochs increases. If the validation accuracy starts to decrease, it can be a sign that the model is overfitting to the training data.

The training loss (red line) and validation loss (blue line) grow with the number of epochs in the loss plot, which is displayed on the right. The degree to which the model's predictions agree with the data's actual labels is expressed as loss. Better performance is indicated by a lower loss. Reducing the loss function is the aim of the training. The training loss should generally reduce as the number of epochs rises. It's critical that the validation loss drops, just like the accuracy plot does. Overfitting may be indicated if the validation loss begins to rise.

The bird species prediction is done by giving the appropriate user input in the model testing process.

```
user_input_path = "user_input/caprimulgus (2).wav"
```

**Fig 12: user input**

The path and the name of the audio file is passed as an input and the model predicts which class it belongs. The output for the given input is shown in fig 13.

```
1/1 [==============================] - 0s 416ms/step
2
Caprimulgus
0.99118245
```

**Fig 13: output**

ISSN [Online]: 2583-2654

```
user_input_path = "user_input/Bubo (2).wav"

input_data = audio_to_signal(user_input_path)
input_data = [float(x) for x in input_data]

input_data = np.array([input_data])
scaled_data = scaler.transform(input_data)

prediction = model.predict(scaled_data, verbose=1)

class_label = np.argmax(prediction)
class_name = class_labels[class_label]
probability = prediction[0][class_label]

print(class_label)
print(class_name)
print(probability)

1/1 [==============================] - 0s 186ms/step
1
Bubo
0.99176544
```

**Fig 14: prediction test for bubo (bird species)**

The output for the given input is predicted as the bird species caprimulgus with an accuracy of 99% that is it predicted precisely.

## V. CONCLUSION

This paper addresses the need for efficient and non-intrusive bird species identification to aid in the conservation of avian biodiversity. Traditional visual identification methods are labor-intensive and impractical, especially in challenging habitats. Leveraging the power of technology and artificial intelligence, the project focuses on utilizing bioacoustics, specifically bird vocalizations, as a rich source of information for species identification. The primary objective is to develop an intelligent system employing machine learning and artificial neural networks (ANN) to automatically identify bird species based on their unique audio signatures.

It follows a comprehensive workflow, including data collection, pre-processing, feature extraction, and the application of machine learning algorithms. A diverse dataset of bird audio recordings is collected, considering various factors to enhance the model's adaptability. The pre- processing phase involves noise removal and feature extraction, crucial for transforming raw audio signals into meaningful representations. The core of the system lies in the ANN classifier, trained on the pre-processed dataset to associate specific audio features with corresponding bird species.

The system's significance extends beyond scientific research, offering practical applications in ecological monitoring, habitat management, and conservation efforts. The report outlines the system's functional and non-functional requirements, ensuring real-time responsiveness, scalability, accuracy, reliability, usability, and maintainability.

The design methodology includes detailed system architecture, UML diagrams illustrating use cases, data flow at different levels, and a sequence diagram depicting the deep learning algorithm's application. The module description outlines key steps, including data collection, labelling, classifier training, ANN implementation, model evaluation, and integration into a user-friendly system for bird species identification. The project aims to provide a valuable tool for researchers, conservationists, and wildlife enthusiasts, contributing to the broader field of bioacoustics and facilitating effective conservation strategies.

## VI FUTURE SCOPE

In conclusion, the Bird Species Identification project presents a pioneering endeavour at the intersection of bioacoustics, machine learning, and conservation biology. The proposed system addresses the limitations of traditional, visually reliant methods for monitoring avian biodiversity by leveraging the distinctive audio signatures of bird species. Through a comprehensive pipeline involving data collection, pre-processing, and the application of machine learning, particularly artificial neural networks, the project aims to provide an efficient and non-intrusive solution for automated bird species identification.

The potential impact extends beyond scientific research, offering a valuable tool for ecological monitoring, habitat management, and conservation efforts. As we strive to safeguard the rich tapestry of bird species on our planet, this project signifies a step toward innovative and ethical approaches that harness technology for the greater good of biodiversity conservation.

The future scope includes expanding the system to recognize a broader range of avian species, incorporating real-time monitoring capabilities, and refining the model's adaptability to diverse environmental conditions. Integration with citizen science initiatives and collaboration with conservation organizations can further enhance the system's impact on avian biodiversity monitoring and research.

***REFERENCES***

*[1] Koh, Chih-Yuan, et al. "Bird Sound Classification Using Convolutional Neural Networks." Conference and Labs ofEvaluation Forum (Clef) (working notes), 9-12 September,Lagano, Switzerland, ISSN: 2038-4726, 2019.*

*[2] Mehyadin, Aska E, et al. "Birds sound classification based on machine learning algorithms." Asian Journal of Research in Computer Science, Volume 9 [Issue 4] (2021),pp 1-11, ISSN: 2581-8260, 2021.*

*[3] Martinsson, John. "Bird species identification using convolutional neural networks." Published by Journal "Computational intelligence and its recording", Volume 19, Issue [1-2], pp 33, ISSN: 1613-0073, (2017).*

*[4] Kahl, Stefan, et al. "Large-Scale Bird Sound Classification using Convolutional Neural Networks." Conference and Labs of Evaluation Forum (CLEF)(working notes): - 1866, ISSN: 2038-4726, (2017).*

*[5] Acconcjaioco, Michelangelo, and Stavros Ntalampiras."One-shot learning for acoustic identification of bird species in non-stationary environments." 2020 25th International Conference on Pattern Recognition (ICPR). IEEE, ISSN: 1051-465, 2021.*

*[6] Samruddhi Bhor et al." automatic bird speciesidentification using Convolutional Neural Network, TF- Slim Tensor files, and spectrograms" Proceedings of the International Conference on Electronics and Renewable Systems (ICEARS 2022) IEEE Xplore, Part Number:CFP22AV8-ART; ISBN: 978-1-6654-8425-1Thiago L. F. Evangelista1 et al. "Automatic Segmentation of Audio Signals for Bird Species Identification using machine learning algorithms. 2014 IEEE International Symposium on Multimedia. IEEE, 2014.*

*[7] Vemula Omkarini et al." Automated Bird Species Identification Using Neural Networks" Annals of R.S.C.B. Vol. 25, Issue 6, 2021, pp 5402 - 5407 Received 25 April 2021; Accepted 08 May 2021, ISSN:1583-6258, 2021.*

*[8] Ilyas Potamitis a et al. "Automatic bird sound detection in long real-field recordings: Applications and tools" Applied Acoustics, 2014, vol 80, pp 1-9, ISSN: 0003-682X,2014.*

*[9] Marcelo T. Lopes et al. "Automatic Bird Species Identification for Large Number of Species" IEEE International Conference on Systems, Man, and Cyberneticsin Anchorage, Alaska, October 2011, Volume 14, and pp 745–750, ISSN: 1062-922X, 2011.*

*[10] Dan Stowell et al. "The title of the document is "Bird Detection in Audio: A Survey and a Challenge". IEEE International Workshop on Machine Learning for Signal Processing, SEPT. 13–16, 2016, SALERNO, ITALY. J, vol.2, pp. e488, 2016.*

*[11] XIE Jiang-jian1 DING Chang-qing et al. "Audio-only Bird Species Automated Identification Method with Limited Training Data Based on Multi-Channel Deep Convolutional Neural Networks", "the Fundamental Research Funds for the Central Universities (NO. 2017JC14)", 2017.*

*[12] zhang et al. "Automatic Recognition of Bird Species byTheir Sounds", published by Elsevier B.V. in the journal Ecological Informatics. Volume 34. pp 108–116. The ISSNof the journal Ecological Informatics: 1574-9541, 2016.*