Detection of Classifications Of Living Organisms Using Voice Recognition

***Abstract*—The diversity of living organisms within the Animal Kingdom presents challenges for researchers and biologists in terms of species identification and classification. Due to limits in human classification and recognition techniques, many undiscovered species continue to exist. This research paper solves this issue by suggesting a voice recognition-based system to describe and categorize distinct animal species. The primary goal is to develop a tool that helps zoologists and biologists recognise unidentified species and comprehend their behavioral patterns. The system uses speech recognition techniques to distinguish between the vocalizations of various species, providing easy access to data about particular creatures like birds, frogs, and more. This approach improves categorization accuracy and frequency by extracting special sound cues and using deep learning models. The potential impact of this research lies in providing a user-friendly solution for detecting and studying different species within the Animal Kingdom.**

***Index Terms*—species identification, voice recognition, categorization, behavioral patterns**

I. Introduction

The Animal Kingdom is home to a wide variety of species, each with its own unique traits and behaviors, many of which have yet to be fully understood. The task of discovering, identifying, and categorizing unknown species is of utmost importance in the field of biology and in efforts to preserve our planet's biodiversity. Conventional methods of species identification often prove insufficient in accurately classifying different organisms solely based on their physical characteristics or appearances. To overcome these limitations, this study endeavors to harness the power of voice recognition technology in order to create a robust system capable of detecting and categorizing living creatures based on their distinct vocalizations. In light of the pressing need for a fresh perspective, the research aims to devise a system capable of precisely capturing and examining the unique vocal patterns exhibited by various animal species. By harnessing the immense potential of profound learning models and data-centric methodologies, this revolutionary system endeavors to bridge the divide between uncharted species and the comprehension researchers possess regarding their conduct. Not only does this groundbreaking system facilitate the identification of species, but it also provides valuable insights into the behavioral characteristics exhibited by these magnificent creatures. This research is inspired by previous successful studies that have used voice recognition to differentiate between species, including birds, frogs, and other animals. The proposed system builds on the knowledge gained from these studies, incorporating the latest methodologies to improve accuracy and usefulness. By developing a system that can effectively extract and analyze vocal features, researchers and biologists will have a valuable tool for classifying species, studying behavior, and conducting ecological research. This paper presents a unique approach to address the difficulties of identifying species, ultimately advancing our understanding of the diverse Animal Kingdom and enhancing biological research efforts.

II. Problem Detection

When it comes to identifying and classifying the diverse species that make up the vast Animal Kingdom, researchers and biologists face a significant challenge. Despite their best efforts, innumerable species remain unknown because of the shortcomings of traditional techniques that rely on human vision. This academic study proposes a novel voice recognition-based method to address the urgent problem of species identification. This ground-breaking technique intends to address the critical issue of effectively and precisely recognising and classifying undiscovered species based on their distinctive vocalizations, especially those that have thus far evaded detection. It does this by accurately defining and categorizing diverse animal species, especially those that have escaped detection.

The classification of various species using traditional methods, which typically rely on visual or physical qualities, can be challenging, especially when several species have similar properties. Due to this restriction, there are unnamed species, which hinders scientific research, ecological understanding, and conservation efforts. The main problem is the lack of a reliable, automated system that can recognise different species by listening to their vocalizations, overcoming the challenges posed by genetic or physical differences.

The problem stems from our poor capacity to identify intricate vocalizations in humans, which is made worse by the lack of extensive databases on animal vocalizations. In order to solve this, the researchers develop a voice recognition system that makes use of cutting-edge deep learning models and data-driven approaches in order to precisely record and analyze the specific speech patterns displayed by various animal species.

The researchers also acknowledge that while there have been advancements in the field of voice recognition for specific animals, such as avian and amphibian species, there is still room for improvement in terms of accuracy, ability to adapt to new situations, and applicability to a variety of habitats. As a result, the need for a novel approach that combines profound learning, characteristic extraction, noise mitigation, and taxonomy categorization emerges.

This study paper essentially explores the fundamental problem of accurately identifying and classifying novel species within the enormous Animal Kingdom through the use of speech recognition technology. The goal is to develop a robust framework that can act as a link between as-yet-unidentified species and our understanding of their behavior, helping researchers, zoologists, and biologists in their efforts to categorize and advance ecological research and conservation initiatives.

III. Applied Methods

The study of life's diversity has always been central to biology. Traditional taxonomy used observations, anatomy, and genetics. Now, AI and biology are merging, using voice recognition to identify species by their unique sounds. This breakthrough decodes animal vocalizations, aiding new research in categorization. Different techniques are explored to identify known or unknown animals through their voices.

Firstly, in the deep learning field transfer learning and pseudo learning has been implemented to collect voices of animals for classification of specific species of animals precisely.r learning is a machine learning technique where knowledge gained from training on one task is used to improve the performance on a different but related task. It involves adapting a pre-trained model's knowledge to new, similar tasks, often leading to faster and more accurate learning.When a machine learning model looks to be learning and improving, it may really be learning from noise or unimportant patterns in the data. This is known as pseudo learning. As a result, the model may overfit and have difficulty generalizing to new data since it is not actually learning anything useful. For classifying the Frog species effectively KNN and SVM classifiers have been used.The sounds of birds has been detected using CNN system with the involvement of different configuration and hyperparameters. Convolutional Neural Networks, often known as CNNs, are a particular kind of artificial neural networks created specifically for processing and evaluating visual input, such as pictures and movies. For tasks like image identification, object detection, and picture classification, it uses convolutional layers to automatically learn and extract pertinent characteristics from the input data. Spectrograms has been applied to obtain the compatibility for a representational transformation of images of two dimensions and audio of one dimension. The STFT output is seen as an image by the spectrogram using a color map, which may subsequently be fed into an image-based pre-trained network. The multileveled ternary pattern (TP) feature generator, a multileveled and handmade feature, is used in the model. The performance of the model is further enhanced by an iterative variation of the ReliefF feature selection (IRF). IRF uses several classifiers, such as support vector machine (SVM), linear discriminant (LD), and k-nearest neighbor (kNN), to automatically select the most advantageous characteristics. The outputs of these classifiers have an impact on the results of the final classification. For noise adaptation in SED models two types of techniques have been developed mentioning bioacoustic sensors networks to track the migration of bird species. This technique will help in wildlife conservation, urban planning, and risk assessment. Automated Remote Biodiversity Monitoring Network (ARBIMON) helps to automate data acquisition, management, and species identification based on audio recordings. It plays a great role in conservation of biodiversity.

IV. Research Objective

Animal Kingdom is the vast place of existence of enormous unknown species of different types of living organisms. Most of the people are unaware of that species and their characteristics.So this creates a problem for humans to detect any unknown species for any research or experimental purposes.As a result sometimes they could not adapt that animal for not knowing their behavioral patterns.In addition biologist or zoologist sometimes fail to classify the unknown species of animals despite of having research. So this research aims to build such a system which would show the characteristics of any unknown species of any specific animal with the power of voice recognition. This system will solve the issue of the human and gives an output of the characteristics and names of that specific species.This will create an easy access for zoologist or biologist to work freely with classification of animals.

- Differentiating different voices according to various species of animal kingdom
- Making easy access for people to collect information about a specific animal for different purposes such as birds,frogs etc
- Can be used easily for extraction of characteristics of any animal of the kingdom.
- Enhancing the performance of the models with more accuracy and frequency
- Extracting the unique sound of different species easily perfectly to detect any unknown species of animal kingdom

V. Literature Review

Classification of living organisms using voice recognition can help the zoologists or biologists to do any type of experimental work using animals. So many researchers have already worked out on the voice recognition of different animals to analyze their different species including amphibians, reptiles, birds and many more.So some of the glimpse of previous works performed to analyze the different voices of animals to detect their behavioral attitudes on our atmosphere.

One of the research papers discusses a study conducted in Puerto Rico that aimed to identify the sounds produced by different bird and amphibian species. The researchers used deep learning, labeling techniques, and a customized loss function to address the issue of limited training data. By utilizing both labeled and unlabeled data, they were able to enhance the accuracy of their model. The authors believe that their approach can be applied to other problems involving sound classification and stress the importance of transfer learning and pseudo-labeling with deep learning models for

precise classifications. They acknowledge the limitations of their data and propose additional techniques to further enhance the performance of their model. Overall, the paper presents a method and framework that can be utilized for other bioacoustics problems and underscores the significance of transfer learning and pseudo-labeling in situations where labeled data is scarce [9].

Moreover, another article presents a computerized system that can identify different frog species based on their sound. The system analyzes sound samples by breaking them down into smaller parts and extracting three specific features. The purpose of this system is to make it easier for people to access and find information about frogs online. Experimental results show that the system is effective in recognizing frog species, with both the kNN and SVM classifiers achieving high accuracy rates. However, there is still room for improvement in accurately classifying frog species with unique sounds, and the researchers plan to develop new methods to address this challenge. Future work: Thus this paper intend to develop new techniques in the feature to extract the features of the sounds for some specific frog species, whose classification accuracy rate falls below average, such as Microhyla butleri and Microhyla ornate [2].

Another researcher analyzes how well people with and without dyslexia can recognize voices in different languages. The study involved 16 individuals without dyslexia and 16 individuals with dyslexia, who were matched based on different factors. The participants listened to recorded sentences in English and Mandarin and were trained to identify the voices of different speakers. They then took a test to identify the speakers without any feedback. The experiment was conducted in a quiet room, and the participants completed it at their own pace. The researchers specifically examined the voice recognition skills of native English speakers with and without dyslexia, focusing on the English and Mandarin languages. They used recordings of sentences spoken by native speakers of both languages and had the participants go through training and testing to assess their ability to recognize speakers based on their voices [1].

The research article named "Bird Sound Recognition Using a Convolutional Neural Network" discusses about classification of bird sounds using CNN system.In this CNN system different types configuration and hyperparameters have been applied.A dataset has been initialized to pre-trained CNN model by the collection of Xeno canto bird songs. It helps to generate different numbers of species of birds, their characteristics and color schemes of the bird kingdom.In the present research, transfer learning is employed to refine an already existing neural network for the purpose of identifying bird sounds.Many such pre-trained networks are taught to recognize general features in images.However, as images are represented as two-dimensional signals and audio is one dimension, a representational transformation is required for compatibility. Spectrograms, a visual representation of the magnitude provided by the Short Time Fourier Transform (STFT)[4], are utilized for this goal. Instead of conducting a single DFT over a longer signal, the STFT is a variation of the DFT that divides the signal into partially overlapping chunks and applies the DFT to each using a sliding window. As a result, an audio slice is represented as a two-dimensional spectral image with time and frequency acting as the axes. The STFT output is seen as an image by the spectrogram using a color map, which may subsequently be fed into an image-based pre-trained network. Besides the literature describes about other methodologies related to CNN such as unsupervised learning, decision tree based feature selection, recurrent CNNs or Hidden Markov models.One of the methods known as spectrograms is being implemented where raw audio is taken as input to perform sound visualization mechanism. In this research different steps have been discussed to create a model for bird sound recognition such as corpus,automated data download,sound preprocessing, pretrained networks and training the network. The tagged spectrogram pictures are compiled into TFRecord format files in order to train the classification algorithm. An inference graph, which permits assessing fresh incoming spectrograms, is the outcome of the training phase. The trained model's primary use is to categorize a single recording of a bird sound. The user records the sound and uploads it to the server, which preprocesses and transforms it into the WAV format. Depending on the threshold setting, the result will be 0 or more spectrograms, which serve as the inference graph's input. The assessment returns to the user its findings after averaging them. Furthermore, future work is focused on noise reduction for preprocessing stages through filtration of extreme frequencies in bird sounds.Planning is performed to Test different gamma values to increase the information collected for spectrograms [3].

Another experiment was mainly done on detection of different species of birds by classifying the sounds of birds with high accuracy using the deep learning models.In this study, a novel machine learning model is presented to accurately detect bird noises while preserving a lightweight design. The multileveled ternary pattern (TP) feature generator, a multileveled and handmade feature, is used in the model. High-level, low-level, and mid-level characteristics are combined in the TP feature generating network to boost classification performance. In order to create the levels, the discrete wavelet transform (DWT), which is useful for classifying bird sounds, is used.An iterative variant of the ReliefF feature selector (IRF) is used to further improve the model's performance. IRF automatically chooses the most useful features, which are then used for different classifiers, such as support vector machine (SVM), linear discriminant (LD), and k-nearest neighbor (kNN). The outcomes of the final classification are influenced by the outputs of these classifiers. On a dataset with 18 distinct bird sound classes, the suggested multilayer TP and IRF-based bird sound classification approach obtained an amazing accuracy of 96.67 percent using SVM. This method presents a potential strategy for keeping a simple classification model while properly categorizing bird noises.Birds have their own acoustic sound signal.Basically sounds of birds are collected from noise environments. The automated classification of sound is quite

hard in the machine learning process.The sound of a specific bird may overlap with other sounds. So to solve these problems this research study has discussed different algorithms which will help to differentiate bird sounds according to specific species.In this research the team aims to create a mobile application for cloud-based bird species classification and detection. To achieve this, it proposes utilizing the multilevel TP and IRF method based on a newly developed deep learning model. This approach offers advanced capabilities for accurately identifying bird species. In the experimental setup, it used the gathered bird sounds as input for the proposed multilevel TP and IRF based method. The implementation of this method was carried out using the MATLAB2018a environment on a personal computer (PC). The PC used in the experiments had a system configuration consisting of 32 GB of memory, an i7 7th generation microprocessor, and a 1 TB HDD. The operating system utilized was Windows 10.1. The abstraction of the proposed multilevel TP and IRF based method was performed according to the specified system configuration and software environment. The research study focuses on addressing two primary challenges: generating low, medium, and high-level features, and utilizing an effective feature selector to achieve a highly accurate bird sound classification model. Deep neural networks have commonly been employed to tackle these challenges.Future works include presenting a new TP-based deep model for sound classification, addressing the problem of setting parameters in recurrent neural networks. This method could also classify other animals' sounds and suggest new generation patterns-based classification models [8].

Furthermore, another research study mainly focuses on finding the solution of issues faced by the non-human animals in voice perception to normalize the similar voice differences. The non-human animals can identify humans through voice characteristics.This paper aims to address elements of human speech as part of the debate of species-shared perceptual systems.Traditional theories of speech perception do not place much emphasis on perception or voice recognition, but cannot be disregarded.As a result, speaker normalization describes our capacity to despite significant auditory differences between speakers, phonologically similar utterances (Johnson, 2005). An effective illustration of the enormous variation in the voice signal brought on by variations in vowel production varies amongst speakers. The first and second formant frequencies (F1 and F2) can be used to distinguish vowels with accuracy; However, the F1 and F2 values of vowels produced by different speakers (and particularly different genders) are highly varied within a vowel category and considerably overlap between categories, to the extent that the acoustic distance within a vowel category can be just as wide as the gender difference vocal types' acoustic distance from one another. In-depth study on speaker voice recognition in animals is included in the article. It talks about research showing that animals can recognize various people based on their vocal traits, such as pitch, length, and formants. This shows that animals are able to recognize and retain particular vocalizations, which may facilitate social interactions and the identification of conspecifics. The final section of the essay explores the idea of speaker normalization, which is the idea that an animal may receive and recognize a vocalization independent of who is making it. It examines experiments showing that even when exposed to vocalizations from novel or unfamiliar speakers, animals can normalize speaker variations and still detect vocalizations based on underlying acoustic properties.

This paper concludes by proposing that non-human animals have perceptual skills pertaining to vowel discrimination, speaker voice recognition, and speaker normalization. These discoveries advance our knowledge of how animals perceive vocalizations and give information on the cognitive mechanisms underlying social interaction and communication among various species [4]. Another wonderful article discusses the use of bioacoustic sensors, also known as autonomous recording units (ARUs), for recording wildlife sounds in a scalable and non-invasive manner. The article proposes several machine listening techniques to improve the sound event detection (SED) models across different acoustic environments. To enhance the SED models, two noise adaptation techniques are introduced. The article highlights the importance of automated solutions for avian flight call detection in large-scale bioacoustic monitoring. By utilizing machine listening techniques and bioacoustic sensor networks, it becomes possible to track bird migration patterns, which can have implications for wildlife conservation, urban planning, and risk assessment. Overall, the article presents a comprehensive approach to improving the robustness and generalizability of sound event detection models in heterogeneous acoustic environments, specifically focusing on avian flight call detection in bioacoustic recordings. The researchers examined twelve different models that represent combinations of three design choices: time-frequency representation, formulation in context adaptation, and the use of artificial data augmentation. Overall, the study demonstrates the effectiveness of combining PCEN, GDA, and context adaptation techniques to enhance the generalizability of SED models in bioacoustic sensor networks. These advancements have significant implications for large-scale deployment and monitoring of avian flight calls, enabling better understanding and conservation of migratory bird populations. Conclusion : The researchers propose various methods to improve sound event detection (SED) models in diverse acoustic environments, with a focus on avian flight call detection in large-scale bioacoustic monitoring [4].

The researcher had done another experiment which discusses the use of machine learning to detect bird sounds in acoustic recordings. The authors conducted a challenge to evaluate the performance of different machine learning methods on this task. The challenge results showed that deep learning methods are able to achieve very high recognition rates, even with mismatched training data. However, there is still room for improvement, especially in terms of generalizing to new conditions. This research paper discusses the application of machine learning, specifically deep learning, in the field of bird detection and monitoring through acoustic data. The authors highlight the importance of monitoring bird populations

for conservation, research, and ecosystem management. The paper focuses on addressing the challenges associated with automatic bird detection, such as the need for manual parameter tuning, customization of template libraries, and limited generalizability to new acoustic conditions. The outcomes of the challenge demonstrate that machine learning methods, particularly deep learning, can achieve high recognition rates on remote monitoring acoustic data, even in the presence of weather noise, low signal-to-noise ratios, and various bird call types. The results represent a significant advancement in the state of the art, indicating the practical utility of machine learning in remote monitoring projects. However, the study also highlights the gap between performance in matched conditions and mismatched conditions. Overall, the research paper demonstrates the potential of machine learning, especially deep learning, in improving bird detection and monitoring through acoustic data. The findings highlight the progress made in reducing error rates and the importance of addressing the challenges associated with generalization in real-world deployment scenarios [5].

Another research paper introduces the Automated Remote Biodiversity Monitoring Network (ARBIMON), a system designed to automate data acquisition, management, and species identification based on audio recordings. The traditional methods for assessing animal species diversity and abundance are limited in terms of cost, spatial and temporal coverage, and the lack of a permanent record. The paper demonstrates the capabilities of ARBIMON by presenting data on the vocal activity patterns of birds, frogs, insects, and mammals from Puerto Rico and Costa Rica. The system proves to be effective in collecting detailed and long-term acoustical data, providing fine temporal resolution and coverage throughout the day and week. The study also highlights the benefits of automating species identification through machine learning algorithms. It offers advantages such as reduced observer bias, permanent records, real-time monitoring, and rapid response to significant events. The ARBIMON system allows users to create species-specific identification models without the need for extensive programming or math skills. Overall, ARBIMON represents a valuable tool for improving biodiversity monitoring by expanding spatial and temporal coverage, providing a permanent record of fauna activity, and enabling automated species identification. The system has the potential to contribute to our understanding of large-scale ecological processes and assist in conservation and management efforts [6].

VI. Future Work

The main goal of this research is to improve the precision of word classification. The researchers are focused on overcoming difficulties caused by specific types of disturbance, specifically those associated with frog species, by introducing a novel and distinct approach. This method not only applies to frogs but also holds promise for enhancing the classification of diverse sound-related problems.

One of the primary objectives of this research is to develop innovative methods for extracting sound characteristics from particular frog species that currently exhibit lower than desired average classification accuracy rates. The aim is to enhance the accuracy of classification algorithms through the implementation of these novel techniques, ultimately contributing to a greater understanding of frog behavior and communication.

The upcoming research phase will focus on bird sounds and will introduce a pre-processing step aimed at eliminating frequencies that are too high or too low. By improving the quality of the noise data, this noise reduction method aims to generate more precise classification results for different bird species. Consequently, this step is expected to enhance the overall accuracy of the classification process.

The recommended method is not limited to any specific type or classification of animals. It possesses the ability to detect and classify different types of animal sounds. This technique has the potential to pave the way for developing advanced models that can accurately categorize a diverse array of animal sounds by analyzing patterns in their generation.

In addition to focusing on the details of classification, the primary objective of this study is to provide a deeper understanding of the way animals perceive and interact through sound. By analyzing sound patterns and classifications, the researchers aspire to gain insights into the cognitive mechanisms that drive social interaction and communication among different species.

While theoretical knowledge often forms the basis for scientific progress, this research underscores the importance of acknowledging the challenges and intricacies associated with implementing such knowledge in real-world situations. The overarching objective is not only to minimize errors in classification but also to fully grasp and tackle the obstacles that arise when endeavoring to put theoretical concepts into practice. By doing so, we aim to achieve a deeper understanding of the complexities involved and effectively address the issues faced in practical contexts.

To conclude this study is dedicated to enhancing the accuracy of word classification through the development of innovative approaches for extracting word features, reducing noise, and constructing classification models. The main objective is to deepen our understanding of animal communication, cognition, and social interactions by specifically investigating frog and bird species. Nonetheless, the study recognizes the challenges involved in translating research outcomes into practical applications.

VII. Dataset

The research paper harnesses the potential of three distinct audio datasets to propel avian bioacoustics research and the automation of bird sound analysis.

The BirdVox-DCASE-20k dataset encompasses 20,000 ten-second audio clips gathered by autonomous recording units near Ithaca, NY, during fall 2015. Primarily aimed at avian vocalization studies and bioacoustic classification model development, the dataset showcases an approximately 50 percent prevalence of bird sounds. This dataset, a derivative of

BirdVox-full-night, caters to the ”Bird Audio Detection” challenge facilitated by DCASE. Equipped with binary ”hasbird” labels in its metadata, it encourages proper attribution to the original publication, adhering to the CC BY 4.0 license. The creation of this dataset was made feasible by grants, a Google Faculty Award, and contributions from anonymous donors.

The NIPS4Bplus dataset introduces a groundbreaking avenue for automating bird species detection and classification using audio recordings. Characterized by recordings spanning regions in France and Spain, it integrates species-indicative tags. The dataset uniquely incorporates manually appended temporal annotations, offering an unprecedented understanding of species occurrence patterns. This augmentation extends the utility of the existing NIPS4B dataset for enhanced avian bioacoustics research.

The BirdVox-full-night dataset comprises six extensive ten-hour audio recordings captured in the vicinity of Ithaca, NY, during fall 2015. This dataset provides an extensive repository of avian flight calls, amounting to around 35,402 instances. Although species-specific labeling is absent, distinctions based on center frequencies remain feasible. This dataset serves as an indispensable resource for the advancement of bioacoustic classification models. With FLAC audio files, JAMS annotation files, and GPS coordinates, it adheres to the CC BY 4.0 license while emphasizing its ”as is” nature.

Together, these datasets underscore the pivotal role of automation in avian bioacoustics research and offer substantial resources to facilitate advancements in bird sound analysis.

VIII. Conclusion

The paper being presented unveils a groundbreaking and comprehensive methodology that holds immense promise for addressing a wide range of bioacoustics challenges. A noteworthy achievement of this research lies in its ability to successfully distinguish between different frog species with exceptional precision. By harnessing the power of both kNN and SVM classifiers, the study attains remarkably high levels of accuracy in species identification. This investigation tackles two crucial problems: the development of diverse feature sets and the implementation of a feature selector to construct an exceptionally reliable model for classifying bird sounds. These efforts are aimed at overcoming the inherent difficulties associated with creating and selecting relevant features in this field.

One remarkable aspect of the research paper is the captivating array of skills exhibited by animals when it comes to acoustic perception. This encompasses their remarkable capacity to differentiate between different vowel sounds, recognize different voices, and adapt to an assortment of speakers. To enhance the precision of sound event detection models in the realm of bioacoustic monitoring, the researchers propose innovative techniques that involve thorough investigation, specifically targeting the identification of avian flight calls amid the vast expanse of this field.

The study’s main discoveries revolve around showcasing the transformative power of acoustic data in the realm of bird detection and monitoring, thanks to the utilization of machine learning techniques, specifically deep learning. This advanced system has the capacity to greatly enhance our comprehension of large-scale ecological processes and offer vital assistance for conservation and ecological management efforts. The research paper highlights the crucial importance of state-of-the-art technology in reshaping our understanding of the environment and propelling essential conservation endeavors forward.

References

[1] Perrachione, T. K., Del Tufo, S. N., & Gabrieli, J. D. E. (2011). Human Voice Recognition Depends on language ability. Science (New York, N.Y.), 333(6042), 595–595.

[2] Huang, C.-J., Yang, Y.-J., Yang, D.-X., & Chen, Y.-J. (2009). Frog classification using machine learning techniques. Expert Systems with Applications, 36(2), 3737–3743. https://doi.org/10.1016/j.eswa.2008.02.059

[3] Incze, A., Jancso, H.-B., Szilagyi, Z., Farkas, A., & Sulyok, C. (2018). Bird sound recognition using a convolutional neural network. 2018 IEEE 16th International Symposium on Intelligent Systems and Informatics (SISY), 000295–000300.

[4] Kriengwatana, B., Escudero, P., & ten Cate, C. (2015). Revisiting vocal perception in non-human animals: a review of vowel discrimination, speaker voice recognition, and speaker normalization. Frontiers in Psychology, 5. https://doi.org/10.3389/fpsyg.2014.01543

[5] Lostanlen, V., Salamon, J., Farnsworth, A., Kelling, S., & Bello, J. P. (2019). Robust sound event detection in bioacoustic sensor networks.

[6] M3 Aide, T. M., Corrada-Bravo, C., Campos-Cerqueira, M., Milan, C., Vega, G., & Alvarez, R. (2013). Real-time bioacoustics monitoring and automated species identification. PeerJ, 1, e103. https://doi.org/10.7717/peerj.103

[7] Stowell, D., Wood, M. D., Pamuła, H., Stylianou, Y., & Glotin, H. (2019). Automatic acoustic detection of birds through deep learning: The first Bird Audio Detection challenge. Methods in Ecology and Evolution, 10(3), 368–380. https://doi.org/10.1111/2041-210x.13103

[8] Tuncer, T., Akbal, E., & Dogan, S. (2021). Multileveled ternary pattern and iterative ReliefF based bird sound classification. Applied Acoustics, 176(107866), 107866. https://doi.org/10.1016/j.apacoust.2020.107866

[9] Zhong, M., LeBien, J., Campos-Cerqueira, M., Dodhia, R., Lavista Ferres, J., Velev, J. P., & Aide, T. M. (2020). Multispecies bioacoustic classification using transfer learning of deep convolutional neural networks with pseudo-labeling. Applied Acoustics, 166(107375), 107375. https://doi.org/10.1016/j.apacoust.2020.107375