FLYING BIRD SPECIES RECOGNITION USING SELECTIVE MACHINE LEARNING TECHNIQUES

Thesis Submitted in

In Partial Fulfillment of the Degree of

Master of Philosophy (M.Phil)

By

K.Annalakshmi

Roll No. 18MPCSF002

Department of Computer Science

Avinashilingam Institute for Home Science and Higher Education for Women

Coimbatore

July 2019

DECLARATION

I declare that the dissertation entitled **"FLYING BIRD SPECIES RECOGNITION USING SELECTIVE MACHINE LEARNING TECHNIQUES"** submitted by me for the degree of Master of Philosophy (M.Phil.) is the record work carried out by me during the period from August 2018 to July 2019 under the guidance of **Dr.I.Elizabeth Shanthi, M.Sc., M.Phil,Ph.D**, Professor, Department of Computer Science, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore and has not formed the basis for the award of any Degree / Diploma / Associateship / Fellowship , Tittles in this University or any other University or other similar institution of Higher Learning.

K. Aef.

Signature of the Candidate

CERTIFICATE FROM THE SUPERVISOR

I certify that the dissertation entitled **"FLYING BIRD SPECIES RECOGNITION USING SELECTIVE MACHINE LEARNING TECHNIQUES"** submitted for the degree of Master of Philosophy(M.Phi) by K.Annalakshmi is the record of research work carried out by her during the period from August 2018 to July 2019 under my guidance and supervision, and that this work has not formed the basis for the award of any Degree / Diploma / Associateship / Fellowship or other Titles in this University or any other University or institution of Higher Learning.

29/7/17

Signature of the Supervisor with designation

ACKNOWLEDGEMENT

First and foremost, the investigator is extremely thankful to the **LORD ALMIGHTY** for his graces and blessings showered on her.

I take immense pleasure in thanking **Shri,Dr.P.R.KRISHNAKUMAR**, Chancellor, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for providing the conducive infrastructure for the conduct of the research study.

I would like to thank **Dr. (Mrs.) PREMAVATHY VIJAYAN M.Sc., M.Ed., Dip.Spl.Edn.(U.K)., M.Phil., Ph.D.,** Vice Chancellor, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for the support extended by her throughout the study.

I extend my heartfelt thanks **to Dr. (Mrs.) S.KOWSALYA, M.Sc., M.Phil., Ph.D.** Registrar, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for the encouragement given by her during the investigation.

I also thankful to **Dr. (Mrs.) K.UDAYA CHANDRIKA, M.Sc., M.Phil., Ph.D.,** Dean, School of Physical Science and Computer Science, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for all the encouragement.

I would like to thank **Dr. (Mrs.) V.RADHA, M.Sc., B.Ed., M.Phil., Ph.D.,** Professor and Head of the Department of Computer Science, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for her excellent advice and valuable guidance extended throughout the study.

I express my reverential gratitude to thank **Dr.(Mrs.) I ELIZABETH SHANTHI , M.Sc, M.Phil., Ph.D.,** Professor, Department of Computer Science, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for Suggesting this topic and supervising the thesis work which helped me to complete my project successfully.

I would like to express my sincere thanks to all the **STAFF MEMBERS OF DEPARTMENT OF COMPUTER SCIENCE** who were responsible for the good finish of this dissertation.

Words fail to express my deep indebtedness to my **LOVING PARENTS, HUSBAND, MY SON, FRIENDS** and **ALL WELL WISHERS** for being the motivating forces behind this dissertation and providing moral support and encouragement in carrying out work.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Identifying the species of birds is a widely-studied problem by ornithologists, and an important task in ecosystem monitoring and biodiversity preservation. Recognition and analysis of bird's species in images is a challenging task due to various appearances, backgrounds, environmental changes and moving position of the birds. In this paper, we intend to build up a flying creature's acknowledgment framework utilizing three learning-type modules. The image will be removed from the Web using the image crawl. To refine the collected images into the training dataset, preprocessing was used with the help of Gaussian filtering method. In the outlier removal, the features are extracted by applying the deep learning techniques to the bird image and final recognition of the system is done by Hidden Markov Model (HMM) classifier. The proposed system achieved an accuracy of 98%.

CHAPTER 1

INTRODUCTION

Bird species identification is a challenging task to ornithologists, and is considered as a scientific task since antiquity. There are also practical reasons to monitor birds. In order to evaluate the quality of our living environment it is important to obtain reliable information about the population of wild animals. Birds are numerous and sensitive to environmental changes; and are easier to monitor than other creatures. Therefore, the use of automated methods for bird species identification is an effective way to evaluate the quantity and diversity of the birds which appear in a region. Another concern is the security of plane flights near airports. The Brazilian Centre for Aeronautical Accident Investigation and Prevention (CENIPA) reported that in two years a total of 1.321 aerial accidents involving bird collision with planes occurs in the Brazilian airspace, resulting in financial losses to the airline companies higher than US$ 3 million. Furthermore, this type of accident is potentially dangerous: the collision of a vulture with a commercial plane is equivalent to an impact of several tones. With more information, authorities can do specific regulations to eliminate the problem. The practical reasons previously mentioned justify the study of mechanisms for bird species identification. In this paper we focus on the automatic identification of bird species using signal processing and machine learning techniques. Until few years ago it was necessary to have a direct contact with a bird in order to determine its species. This situation has changed recently, with the use of automatic computational devices. One way to indirectly identify the species of a bird is using its recorded song. This task can be accomplished using signal processing and machine learning techniques. Bird species identification based on images is also challenging due to the variation of the background and illumination because most of the bird images are collected in their natural habitat. In these images one cannot control rotation, scale and viewing angle at the time of image acquisition. Using audio records rather than bird pictures is justified by current practices. Birds are actually not easy to photograph; audio calls and songs have proven to be easier to collect and sufficiently species specific. However, the visual properties such as color, shape, size, parts, among others, are important for the bird species recognition and can be very useful as the number of species taken into account increases to hundreds or thousands of species.

1.1 Need for the work

Identifying bird species based on their calls, songs and sounds in audio recordings is an important task in wildlife monitoring. Hence there is an essential need to know more about the world's most threatened bird species. The system puts different species into categories according to a given set of criteria. The most important data are population size and range, and trends in one or other of these. Trends are impossible to measure unless some baseline has previously been set. For most species this has yet to be done. In addition, threat might be measured as known or inferred change of extent or condition of habitat. This can only be applied if the habitat requirements of the target species are reasonably well known. For the majority of bird species, and especially for many threatened species, these most basic parameters are simply unknown.

1.2 Identifying birds species from image

Birds' species identification from images is an important and challenging problem with many applications the real world such as environment protection and endangered animal rescue. There are also some other practical reasons to monitor birds. In order to evaluate the quality of our living environment it is important to obtain reliable information about the population of birds. Birds are numerous and sensitive to environmental changes; also, easier to screening than other species. In this way, the utilization of mechanized strategies for bird's species recognition is an effective method to assess the amount and assorted variety of the bird's species which show up in a region. The practical reasons recently referenced legitimize the investigation of systems for bird's species recognizable. Birds' species distinguishing proof is a notable issue to ornithologists, and is considered as a logical assignment since vestige. Ornithologists study bird's species, their reality in nature, their science, their tunes, their dispersion, and their environmental effect. Bird's species distinguishing proof dependent on images is additionally testing because of the variety of the foundation and enlightenment in light of the fact that the greater part of the new images is gathered in their normal territory in these images one cannot control rotation, scale and viewing angle at the time of image acquisition. Utilizing sound records instead of bird's images is defended by current practices. Capturing the birds' image and identification of species is a difficult task than audio identifying of species. Bird species identification from images is an important and challenging problem with many applications in the real world such as environment protection and endangered animal rescue. There are also some other practical reasons to monitor birds. In order to evaluate the quality of our living environment it is important to obtain reliable information about the population of wild animals. Birds are numerous and sensitive to environmental changes; also, and are easier to monitor than other species. Therefore, the use of automated methods for bird species identification is an effective way to evaluate the quantity and diversity of the birds which appear in a region. The practical reasons previously mentioned justify the study of mechanisms for bird species identification. Bird identification is a well-known problem to ornithologists, and is considered as a scientific task since antiquity. Ornithologists study birds; their existence in nature, their biology, their songs, their distribution, and their ecological impact. Bird classification is usually done by ornithology experts based on an animal classification system proposed by Linnaeus: Kingdom, Phylum, Class, Order, Family, and Species1. Birds are typically categorized by their shape or silhouettes and physical characteristics. Bird species classification is a challenging problem both to humans and to computational algorithms that attempts to carry out such a task in an automatic fashion. In the last years, several approaches based on bioacoustics signals have been proposed. Such approaches have achieved very interesting correct classification rates, depending on the number of bird species take into account. Besides the interesting results achieved with bioacoustics signals, this problem has been tackled also based on bird images. Compared with sound identification, visual features are not well studied for bird classification. The classification problem can be stated as given a bird image, classify its species among a fixed but large number of possibilities. The challenge of such a classification task is due to the variation in the background and illumination since most of the images are gathered on birds' natural habitat and in the birds pose since it is not possible to control rotation, scale, and angle of view while acquiring images. Visual properties (e.g., shape, color, marks, etc.) are important keys for bird recognition. Some researchers utilize these features to automatically identify birds. Figure1 images of moving birds below

Fig 1: Images of Moving bird

Several bird species classification challenges with closely related, but different, task descriptions have been held during the last few years. The interest and participation in these challenges have been high which indicates that these are relevant problems and that there is a need to solve them. The challenges are usually to predict which species are present in a set of recordings with hidden labels, called the test set, and to submit the predicted species for each test data point for evaluation against the ground truth labels. The task description can vary from predicting only the presence or absence of birds in a recording to predicting all actively singing bird species. That is, the challenges have varying degree of difficulty. In the rest of this section, we present the results of some of the most recent such challenges in chronological order. Identifying the species of a bird is a widely-studied problem to ornithologists, and an important task in ecosystem monitoring and biodiversity preservation. Recognition of bird species in images is a challenging task due to various appearances, backgrounds, and environmental changes. Despite of this, this fine-grained recognition task has received a significant amount of attention in the computer vision community because of its potential widespread applications. Compared to generic object recognition, fine-grained recognition benefits more from learning critical parts of the object that can help align objects of same class and discriminate between neighboring classes. Current state-of-the-art methods adopt CNN-based architectures that learn representations directly from the raw data and can be used to extract set of discriminative features.

1.3 Visual identification of bird species

Basic bird identification clues are acoustic and visual. The main visual clues comprise of the bird's silhouette, its plumage and coloration. However, we must take into account the time of year because bird's plumage will change during different seasons. The acoustics clues comprise the songs and calls that birds make. Furthermore, the bird's behaviour and habitat it is found in are also possible clues. Field marks, the marks that distinguishing one bird from another are also important, such as breast spots, wing bars which are described as thin lines along the wings, eye rings, or crowns, eyebrows which are described as lines over the eyes, eye lines, which are lines through the eyes, and many others. The shape of the beak is often an important clue. Size can usually be a good indicator of the species of the bird. As an example, most song birds fit into a certain size group. Bird shape and posture are the most important characteristics used to identify birds. Most experts can identify a bird from its shape or silhouette because this is the least likely characteristic to change. The tail of a bird can have many variations. The tail can be notched, long and pointed, or rounded. The legs can be long, or short. In spite of having many different features that can take into account to distinguish between bird species, from the pattern recognition point of view, relying on many of such features to identify different bird species is not trivial mainly due to the difficulty in obtaining standard bird images. The acquisition of bird images is somewhat out of the control of the researchers and this imposes an extraordinary variability which is very difficult to be handled by most of the image processing, feature extraction methods and learning algorithms.

1.3.1 Motion detection

The already made reference to technique for recognizing moving birds just functions admirably for scenes where there is no development of extensive articles since there is no real way to differentiate between moving birds and an arriving bird. Shockingly, numerous scenes that require observation have a considerable measure of development in them, for example, trace people on foot. To settle this, we require a type of movement location that can differentiate between moving birds and landing birds. One of the manners in which that birds act directly from moving articles is that flying bird are not in the image, at that point they show up and remain there until the point that they are evacuated or fall on the screen. Moving birds then again rest are not in the image, at that point they show up and after that they vanish. Be that as it may, this isn't in every case valid for course since moving birds can stop and stay in the image, for example, a stopping auto; however for most cases this holds. To recognize a moving article three edges are required. Notwithstanding, it is just conceivable to state that a question was moving in the event that it showed up in the second edge since it needs one past and one after edge where it didn't show up. There is no chance to get of knowing whether a question that showed up in the resting birds or a moving bird. Hence nothing can be said in regards to the rest and last casing, and they should just be utilized to distinguish moving birds in different edges and not be utilized for landing birds location.

1.3.2 Birds characterization

Birds characterization is normally done by ornithologists based on animal classification system. Birds are regularly sorted by their shape or outlines and physical attributes. Fledgling species order is a difficult issue both to people and to computational calculations that endeavours to do such an undertaking in a programmed manner. In the most recent years, a few methodologies dependent on bioacoustics sign have been proposed. Such approaches have achieved very interesting correct classification rates, based on the number of bird species take into account. Other than the fascinating outcomes accomplished with bioacoustics signals, this issue have been handled additionally dependent on birds' images. Contrasted and sound distinguishing proof, visual highlights are not very much concentrated for bird's order. The grouping issue can be expressed as given a bird's picture, characterize its species among a fixed yet huge number of conceivable outcomes. The test of such a grouping errand is because of the variety out of sight and light since the vast majority of the images are assembled on bird’s characteristic territory and in the birds present since it's impractical to control turn, scale, and edge of view while securing images. Visual properties (e.g., shape, shading, marks, and so on.) are significant keys for bird’s classification. A few specialists use these highlights to naturally distinguish birds.

Fundamental birds distinguishing proof pieces of information are acoustic and visual. The principle visual intimations include the birds's outline, its plumage and shading. In any case, we should consider the season since birds’ plumage will change during various seasons. The acoustics pieces of information include the tunes and calls that birds make. Moreover, the birds' conduct and environment it is found in are likewise potential hints. Field denotes, the imprints that distinctive one fowl from another are additionally significant, for example, bosom spots, wing bars which are portrayed as dainty lines along the birds, eye rings, or crowns, eyebrows which are depicted as lines over the eyes, eye lines, which are lines through the eyes, and numerous others. The state of the snout is regularly a significant intimation. Size can more often than not be a decent marker of the types of the fowl. For instance, most melody birds’ creatures fit into a specific size gathering. Birds shape and stance are the most significant attributes used to recognize birds s. Most specialists can distinguish a bird’s creature from its shape or outline since this is the most outlandish trademark to change. The tail of birds can have numerous varieties. The tail can be scored, long and pointed, or adjusted. The legs can be long, or short. Regardless of having a wide range of highlights that can be considered to recognize bird’s species, from the example acknowledgment perspective, depending on a significant number of such highlights to distinguish diverse bird’s creature species isn't unimportant principally because of the trouble in getting standard fowl images. The procurement of fowl images is to some degree out of the control of the analysts and this force a phenomenal changeability which is exceptionally hard to be taken care of by the majority of the Image handling, highlight extraction techniques and learning calculations. The Image was gathered from the Image creep. To refine the gathered images into the preparation dataset, the tainted Image is remedied and erased, the exception is evacuated, and lastly the Image is extended to acquire the refined preparing information. What's more, the highlights were spoken to as factual highlights, which give data on the area, fluctuation and presence of the conveyance of information and furthermore to guarantee order can be performed by utilizing classifiers.

Birds and their sounds are in many ways important for our culture. They can be heard even in big cities and most people can recognize at least a few most common species by their sounds. Bird song has also been an important source of inspiration for many composers, musicians, and writers. In this article we study automatic recognition of bird song. Technology for sound-based identification of bird species and even individuals would be a significant addition to the research methodology in taxonomy and monitoring of migration and population in biology. At a higher level it would also facilitate systematic research on vocal communications between birds and characterization of their sounds. There is also commercial potential for such systems because the number of active bird watchers is really large in many countries. Sounds of birds are mainly produced by syrinx, which an organ unique to birds. It is located in the intersection between the main bronchi of the lungs and the trachea, or in the trachea. There is a considerable variability in the anatomy of syrinx even in different families within the same order of birds. The function of syrinx resembles that of human vocal cords in many ways. But, it typically has much more complex structure and can produce a significantly larger variety of different sounds than glottis in mammals. While in human speech sounds are mainly produced by muscular control of the vocal tract, mouth, lips, tongue, and teeth, the syrinx is the main source of variability of sounds in birds. Only few bird species, mainly parrots, can use their tongue in a way which resembles speech production in humans.

1.4 Use of convolutional neural networks

Image order is a vital undertaking in the field of PC vision and alludes to the task of one mark, from various predefined classifications, to an info image. This assignment has turned out to be a perplexing and testing issue for PCs since deciphering pictures properties from pixel esteems is non-minor because of the numerous varieties a image can be exposed to. For instance, on account of protest discovery, a similar question can be caught in various positions or its appearance may change when presented to various light conditions. In any case, for machine vision frameworks is extremely troublesome understanding that the protest is the equivalent in spite of the diverse way it very well may be spoken to. Refined comprehension of the information, almost human-level, created over human learning and past experience, would be important to take care of this sort of issue. Since expressly programming a PC to do this is a non-minor errand Machine vision specialist have intensely drawn from Machine Learning, which enables PCs to gain for a fact, by model and by similarity naturally.

1.5 Use of Hidden Markov Model

The HMM is based on augmenting the Markov chain. A Markov chain is a model that tells us something about the probabilities of sequences of random variables, states, each of which can take on values from some set. These sets can be words, or tags, or symbols representing anything, like the weather. A Markov chain makes a very strong assumption that if we want to predict the future in the sequence then all that matters is the current state. The states before the current state have no impact on the future except via the current state. It’s as if to predict tomorrow's weather you could examine today's weather but you weren't allowed to look at yesterday's weather.

A hidden Markov model (HMM) allows us to talk about both observed events Hidden Markov model (like words that we see in the input) and hidden events (like part-of-speech tags) that we think of as causal factors in our probabilistic model. An HMM is specified by the following components:

$Q = q_1 q_2 \ldots q_N$ a set of N states

$A = a_{11} \ldots a_{ij} \ldots a_{NN}$ a transition probability matrix A, each a_{ij} representing the probability of moving from state i to state j, s.t. $\sum_{k=1}^{n} a_{ij} = 1 \quad \forall i$

$O = o_1 o_2 \ldots o_T$ a sequence of T observations, each one drawn from a vocabulary $V = v_1, v_2, \ldots, v_V$

$B = b_i(o_t)$ a sequence of observation likelihoods, also called emission probabilities, each expressing the probability of an observation o_t being generated from a state i

$\pi = \pi_1, \pi_2, \ldots, \pi_N$ an initial probability distribution over states. π_i is the probability that the Markov chain will start in state i. Some states j may have $\pi_j = 0$, meaning that they cannot be initial states. Also, $\sum_{l=1}^{n} \pi_i = 1$

HMM applications, include:

Speech recognition

Robot localization

Gene finding

User modeling

Fetal heart rate monitoring

1.6 Web crawler

Web crawlers are more or less as same as the web. Most of the web-based search engine uses the common methods of image retrieval exploit some method of accumulating the metadata such as file names, captioning, keywords or descriptions to the images constructed by human. Therefore, that retrieval can be performed over the annotation words rather than the content of the image. Images from the minute it was invented, has had an immense impact on the world we live in. Extracting the required images from the World Wide Web (WWW) is very difficult because web contains a huge number of images. To solve this problem, we need a system that can retrieve the required images needed by the user. Image Crawler is a web-based tool that collects and indexes group of web images available on the internet. This tool collects the keyword or phrase from the user to retrieve the images from the web. Then the collected keyword is applied to the different general search tools like Google, Yahoo, Bing etc. The collected web page information is stored in the temporary file till 200KB file size from the server. Then this file content will be scanned and extract the image URL's and it is compared the URL which is present in the database to avoid the duplicate downloads. The extracted URL's images are downloaded and finally stores unique image and corresponding metadata like filename, URL, size etc. in the database. Figure 2 Web Crawler Architecture below.

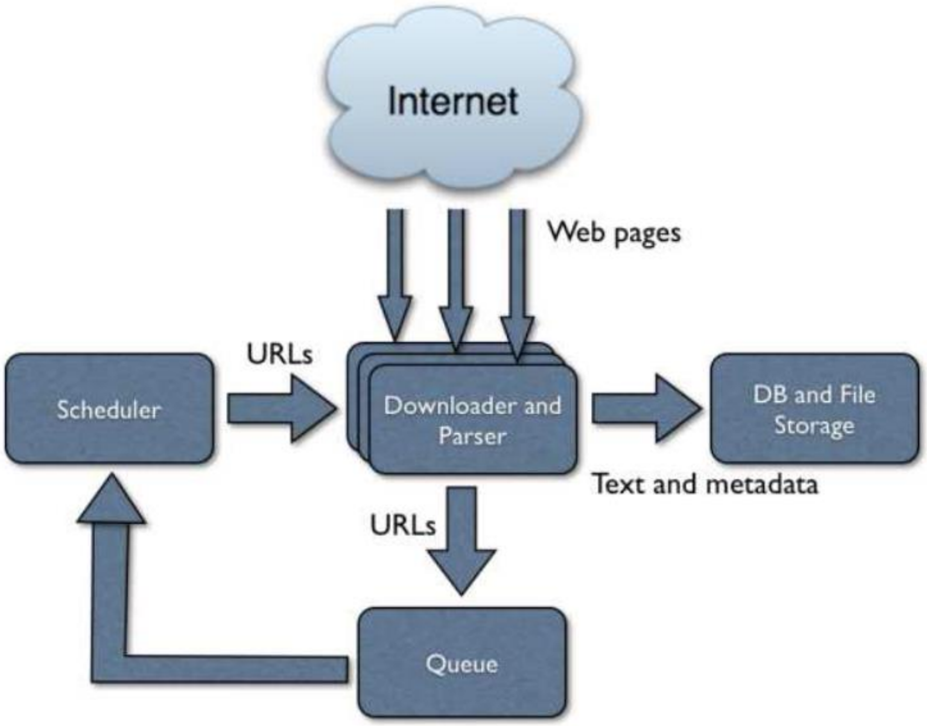


Fig 2. Web Crawler Architecture.

Web Crawler is a search-based tool where it requires only a keyword or phrase from the user to present the relevant images according to the user requirements.

1.7 Application of data mining

Data mining is that the method of analyzing knowledge from completely different views and summarizing it into helpful data that may be won't to increase revenue, cuts costs, or both. Data processing computer code is one in every of the variety of analytical tools for analyzing knowledge. It permits users to research knowledge from many alternative dimensions or angles, reason it, and summarize the relationships known. Technically, data mining and data method is that the process of finding correlations or patterns among dozens of fields in giant relative databases. Data mining, or data discovery, is that the computer assisted method of analyzing monumental sets of information so extracting which means of the info. Data processing tools predict behaviors and future trends, permitting businesses to form proactive, knowledge driven selections. Data processing tools will answer business queries that historically were too moment overwhelming to resolve. Data mining uses information from past data to analyze the outcome of a particular problem or situation that may arise. Data mining works to analyze data stored in data warehouses that are used to store that data that is being analyzed. That particular data may come from all parts of business, from the production to the management. Managers also use data mining to decide upon marketing strategies for their product. They can use data to compare and contrast among competitors. Data mining interprets its data into real time analysis that can be used to increase sales, promote new product, or delete product that is not value-added to the company. Data mining is primarily used today by companies with a strong consumer focus - retail, financial, communication, and marketing organizations. It enables these companies to determine relationships among "internal" factors such as price, product positioning, or staff skills, and "external" factors such as economic indicators, competition, and customer demographics. And, it enables them to determine the impact on sales, customer satisfaction, and corporate profits. Finally, it enables them to "drill down" into summary information to view detail transactional data. With data mining, a retailer could use point-of-sale records of customer purchases to send targeted promotions based on an individual's purchase history. By mining demographic data from comment or warranty cards, the retailer could develop products and promotions to appeal to specific customer segments.

1.8 Data mining techniques

The most important task in Data Mining is the choice of selection of the precise data mining technique. It is based on the type of business and the type of problem faced by the business. To improve the accuracy and cost effectiveness of data mining techniques a generalized approach has to be used. There are basically seven main Data Mining techniques. Even-though there were lots of other Data Mining techniques, these seven are considered as more frequently used by the business people. The techniques are Statistics, Clustering, Visualization, Decision Tree, Association Rules, Neural Networks and Classification.

1.8.1 Statistical Techniques

Statistics is a branch of mathematics which relates to the collection and description of data in the data mining technique. Even-though Statistical technique is not considered as a data mining technique by many analysts, it still helps to discover the patterns and build predictive models. For this reason some knowledge about the different statistical techniques were needed by the data analyst. Because, in day to day life the user have to deal with huge amount of data and develop important patterns from it. Statistics can support the user to a greater extent to get answers for the following questions about the data like

- What are the patterns in the database?
- What is the probability of an event to take place?
- Which patterns are more useful to the business?
- What is the high level summary that can give you a depth view of the database?

Even-though there were different forms of statistics, the most useful technique is collection and counting of data.

1.8.2 Clustering Technique

Clustering analysis is the process of classifying data that are related to each other. With the help of this the user can able to understand the comparative features between the data, which can be denoted as segmentation. For example, an insurance company can group its customers, based on their income, age, nature of policy and type of claims. Following are some of the different types of clustering methods.

1.8.3 Visualization

One of the most reliable techniques which is used to discover data patterns is, Visualization. It will be used at the beginning of the Data Mining process. For Projection Pursuit, many researches are going on these days to produce interesting projection of databases. It helps in converting poor data into good data, which lets the different kinds of Data Mining methods to be used in discovering the Hidden patterns.

1.8. 4 Induction Decision Tree Technique

A predictive model, in which each branch of the tree is viewed as a classification question and the leaves of the trees are considered as partitions of the dataset related to that particular classification. In exploration analysis, data pre-processing and prediction work, this technique can be widely used. This technique can be considered as segmentation of the source dataset which is done for a specific reason. Every single data that comes under segment has some set of similarities in their predicted information. Since the result provided by the decision tree can be easily understood by the user, it is mostly used by the statisticians to find out the appropriate database related to the problem of the business. In many aspects this technique is used for "Prediction and Data pre-processing". The foremost step in this technique is developing the tree, which depends on discovering the finest possible question that to be asked at each branch of the tree. Following are the circumstances where the decision tree stops its growth.

- If the segment contains only one record
- All the records contain identical features
- The growth is not enough to make any further spilt

1.8.5 Neural Networks

One of the important technique, which is most often used in the initial stage of the data mining technology. Neural network, finds very easy to use as they are automated to a specific extent and due to this the user is not anticipated to have much knowledge on the database. Anyhow to make use of the neural network efficiently, the users have to know the following statements.

- How the nodes are connected?
- How many processing units to be used?
- When should the training process to be stopped?

1.8.6 Use of Association Rule

To find the association between two or more items this technique will be used. In database, it assists the user to find the relation between the different variables. It finds the hidden pattern in the dataset through which the user can able to identify the variables and the occurrence rate of different variables that appear with the highest frequency. In retail industry, this technique is often used to find the patterns in sales, through which there will be increase in conversion rate and profit. Association rule offers 2 major information, they are

- Support – How often is the rule applied?
- Confidence – How often the rule is correct?

It follows the 2 step process.

- Find all the frequently occurring data sets
- Create strong association rules from the frequent data sets

Its types are as follows.

- Multilevel Association Rule
- Multidimensional Association Rule
- Quantitative Association Rule

1.8.7 Classification Techniques

One of the most commonly used data mining technique which consists of a set of pre-classified illustrations to create a model to categorize the large set of data. In deriving important information about data and metadata this technique is often used. This technique is closely related to cluster analysis and it uses decision tree or neural network system. Following are the two main processes involved in this technique

- Learning – In this process the data are analyzed by classification algorithm

- Classification – In this process the data is used to measure the precision of the classification rules

Complex structures and sequencing patterns of genes are most effectively presented in graphs, trees, cuboids, and chains by various kinds of visualization tools. Such visually appealing structures and patterns facilitate pattern understanding, knowledge discovery, and interactive data exploration. Visualization therefore plays an important role in biomedical data mining.

1.9 Current trends in bird species recognition

The important habitats for the biodiversity were found to be the sand flats, salt marshes, brackish pools, lagoons, sand dunes, mangrove areas, gravel and pebbles, rocky beaches, cliffs, rocky islands, wastelands and built up areas. Counting is central to ecological studies and conservation research in ornithology. This paper examines the present status of counting birds for various purposes in India. Though there is general awareness about proper counting techniques, as evidenced by the variety of methods employed by field biologists, reliability estimates are not available for most studies. Counts that employ large sets of volunteers, such as the Asian Waterfowl Census have not developed rigorousness, and bird-counts in community studies have ignored to perform standardization exercises such as developing species incidence curves. Studies on single species, though important, have mostly focused on species with high conservation value, but are expanding to include other species as well. Developing models that specify the associations between individuals and their habitat is essential to increase our understanding of how species use their environment, and for developing models that allow managers to predict the consequences of management actions. Florent Bled, John Sauer, et al, analyse the conservation-oriented studies of spatio-temporal dynamics are especially timely as the modelling of species distribution dynamics will be useful in developing predictions about distributional changes expected to accompany climate changes, land use changes and active land management. By understanding how species' ranges have changed over the last several decades, we can provide a basis for projections about future range changes in response to global climate change. Fortunately, a variety of hierarchical-model based approaches are now available for determining whether relationships exist between animals and environmental characteristics (i.e., to identify habitat) as well as monitor spatial or temporal changes in the populations

In Finland, modern forestry has caused the gross modification of native forest habitats and a loss of biological diversity due to intensive management regimes. Although less than 1% of the old-growth forest area remains in southern Finland, the effects of forestry on bird population long-term trends have not been comprehensively studied in this area. To fill this knowledge gap, we analysed 30 years of monitoring data of 32 common forest breeders in southern Finland. Wetlands are an important indicator of flora and fauna. Wet land provide habitat, feeding, nesting, and rearing for near about 310 sp of birds in India. Ahiran Lake wetland situated in Murshidabad district, West Bengal, India has recorded 30 species of water birds belonging to 29 genera and 12 families in the month of February 2014 to January 2015. Out of 30 species about 16 were migrants and 14 species residents. The family Anatidae represented by 7 species are dominated; Ardeidae by 5; Rallidae and Scolopacidae by 4; Laridae and Jacanidae by 2; Alcadinidae, Charadridae, Phalacrocoracidae, Anhingidae and Ciconidae by 1 are fewer in number. In case of resident Little Grebe (Tachybaptus ruficollis), Lesser Whistling-Duck (Dendrocygna javanica), White-breasted Water hen (Amaurornis phoenicurus), Purple Moorhen (Porphyrio porphyrio), Pheasant-tailed Jacana (Hydrophasianus chirurgus) are common and in migratory birds Red-crested Pochard (Rhodonessa rufina), Northern Pintail (Anas acuta), Cotton teal (Nettapus coromandelianus) are abundant species. Migratory birds displayed a definite pattern for arrival and departure from the wetland that is species specific. The peak of winter population of migratory birds was observed during the months of December and February. Numbers of water bird in Ahiran Lake are decreases by several anthropogenic threats, and are affected feeding and breeding habitat of birds.

The winter migratory birds displayed a definite pattern specific to species for arrival at and departure from the Lake. They appeared at the wetland from November and stayed up to March. The peak of winter population of migratory birds was observed during the months of December to February. The Observation through the year revealed that Northern Pintail Anas acuta, Common Pochard Aythya ferina, Common Coot, Common Moorhen, Fulica atra and Gadwall Aythya ferina arrived in late November; Ferruginous Pochard , Garganey, Grey-headed Lapwing arrived in first week of December. As far as the departure time was concerned, Gadwall generally departed in February. Birds like Red-crested Pochard, Common Coot, Common Moorhen, Ferruginous Pochard, Garganey and Northern Pintail departed in March. The summer visitors, namely Cotton Teal Nettapus coromandelianus were spotted during summer seasons from March to October.

1.10 Current techniques used in bird species recognition

Currently investigated about the identification of bird species were used by many techniques. Some of the techniques were discussed below.

- The use of two canonical classification methods, chosen due to their widespread use and ease of interpretation, namely a k Nearest Neighbour (kNN) classifier with histogram-based features and a Support Vector Machine (SVM) with time summarisation features.

- Deep convolutional neural networks (DCNN) have achieved breakthrough performance on bird species identification tasks based on spectrogram features, but a huge number of labelled samples are needed to train an excellent DCNN model. However, it is difficult to collect enough samples for certain bird species. For practical uses of bird species identification, it is significant to study a method which only requires a small sample set.

- Besides the interesting results achieved with bioacoustics signals, this problem has been tackled also based on bird images. The feature vectors are labeled with the corresponding bird species and used in a supervised learning process. Support vector machine (SVM) was chosen as the supervised learning model and the Platt's sequential minimal optimization algorithm as the learning algorithm. The choice of the SVM is due to the fact that it can efficiently perform nonlinear classification using the kernel trick which implicitly maps their inputs into high-dimensional feature spaces.

- In the classification and decision fusion module, the RAID-G features are used as the image descriptors for making final classification of the bird in an input image. they are trained by using different RAID-G data, i.e., the RAID-G of object, the RAID-G of head, and the RAID-G of body as well as the RAID-G of original image, respectively. Therefore, each classifier will give its classification results with different scores for a bird in an image.

- Mel-frequency cepstral coefficients (MFCC) have been a popular signal representation method used in many audio classifications tasks, especially in automatic speech recognition (ASR). The basis for the MFCC mel-frequency scale is derived from the human perceptual system. Perceptual systems of birds are not the same as in humans, but exhibit similar characteristics. The calculation of MFCC parameters is efficient and straightforward since they do not involve any tuning parameters. The calculation of MFCC parameters begins with the segmentation of a signal into overlapping frames.

Motivation for the work

Bird species identification is a typical pattern recognition problem and most studies include signal pre-processing feature extraction and classification sections. Bird vocalization segmentation into smaller recognition units is performed by hand or automatically. The number of species has ranged between 2 and 16 in previous studies. Hence identifying the bird species is a critical issue from this. On the other hand, sound also provide us important information about the world around us. Many animals make sounds either for communication or their living activities such as moving, flying, mating etc. Although sound also provide many confusions among the birds and also in previous research many researchers were concentrated only on sitting birds' images only not concentrates on flying birds. To identifying the moving bird species is our main motive for this thesis.

CHAPTER 2

LITERATURE REVIEW

The biodiversity conservation is one of the most crucial issues that governments and international organizations have to deal with. The protection of endangered species is a must and relies primarily on the accurate monitoring of the biodiversity and secondarily on the application of conservation actions, based on the monitored biodiversity status. The observation and monitoring of birds are of major importance for the biodiversity conservation. Science and technology have improved our quality of life, but the rapid development of some industries has given up people's future living environment and causes an adverse impact on the survival of some wild animals. Taking measures to protect environment and endangered animal is an urgent task. Because birds are numerous and sensitive to environmental changes; also, and are easier to monitor than other species, observing the behavior of birds allows us to better evaluate our living environment. Therefore, it is particularly important to identify the birds. With the development of image identification technology, using modern technology to identify birds has become an effective research method. This paper proposed a new algorithm for bird identification based on color image. Although pattern recognition has been more than half a century of history, identification for bird is still rare. In last years, several approaches have been proposed to recognize the bird. Using birds' voice for bird recognition has achieved interesting correct classification rates, but bird recognition based on image signals is facing the challenge because of illumination, different point of camera, background and so on. Bird classification can be done manually by domain experts; however, with growing amounts of data, this rapidly becomes a tedious and time-consuming process. Therefore, automatic tools which can aid in this process are needed. Several bird species identification challenges such as the Bird CLEF.

Monitoring of specific habitats and provides the information needed for biodiversity monitoring, animal species population estimation, species behavior understanding etc. The bird species recognition task falls in the category of audio pattern recognition task, and can briefly be structured in the audio acquisition stage, the audio parameterization stage and the classification stage. The recognition of the species in the bird's habitats includes interferences that are additive to the bird species vocalizations. Such interferences are the wind, the sum of the leaves, vocalizations from other animal species of the habitat, sounds produce by human activities, etc. Typically, for the recognition of species in such a noisy environment detection of audio intervals with bird vocalizations precedes the species classification stage.

Birds are important indicator species for environmental health. They spread over a wide range of landscape and reflect the dynamic change of an ecosystem. Acoustics has long been used to monitor bird species in the natural environment due to the fact that birds are vocal creatures that utilize acoustic cues for daily communication. Acoustics has several advantages in terms of conveying information. It travels long distances, costs little energy to produce and is especially useful when visual conditions are poor, such as during the night. Bird species richness studies the number of unique bird species in a particular area over a fixed period of time. Classical methods for bird species richness survey require skilled persons physically present in the field and inventory each species they see or hear

ABOUT MACHINE LEARNING AND DEEP LEARNING

Recent machine learning and deep learning research has put a lot of focus on facial recognition, which involves detecting faces within images, landmarking specific features to account for the angle at which the picture was taken, and finally determining whose face is in the picture. This is extremely valuable research and has many applications including social media, computer security, marketing, and healthcare. While this research is booming, there has been less focus on the classification of other organisms like birds, for example. As with facial recognition, there is much complexity in the classification of birds due to issues like the similarities in different subspecies of birds, challenges with the foreground or background setting of images, and lighting conditions in photos. We intend to learn and classify different bird species from a classified data set.

REVIEW OF BIRD SPECIES RECOGNITION

Pellegrini, Thomas (2017), is detecting the bird sounds in the manner of automatically audio recordings mechanism; it is very helpful to monitoring the biodiversity based on audio field recordings working with bio- and eco-acoustics. In recently organized state-of-the-art machine through learning approaches by Bird Audio Detection that involves large audio datasets. This paper uses several types of convolutional neural networks (i.e. standard CNNs, remaining nets and densely connected nets) that are described in the outline of this challenge. Performance gains were gained to data increase through time and occurrence shifting, model parameter be around during training and collaborative methods using the geometric mean. And this leading to the performance of 88.22% area under the ROC curve score and it is the best performing and most compact model. On the diverse, the efforts to increase the training dataset with models of the test set with automatic predictions used as pseudo-ground truth tags consistently stained performance. Figure 3 is explained.

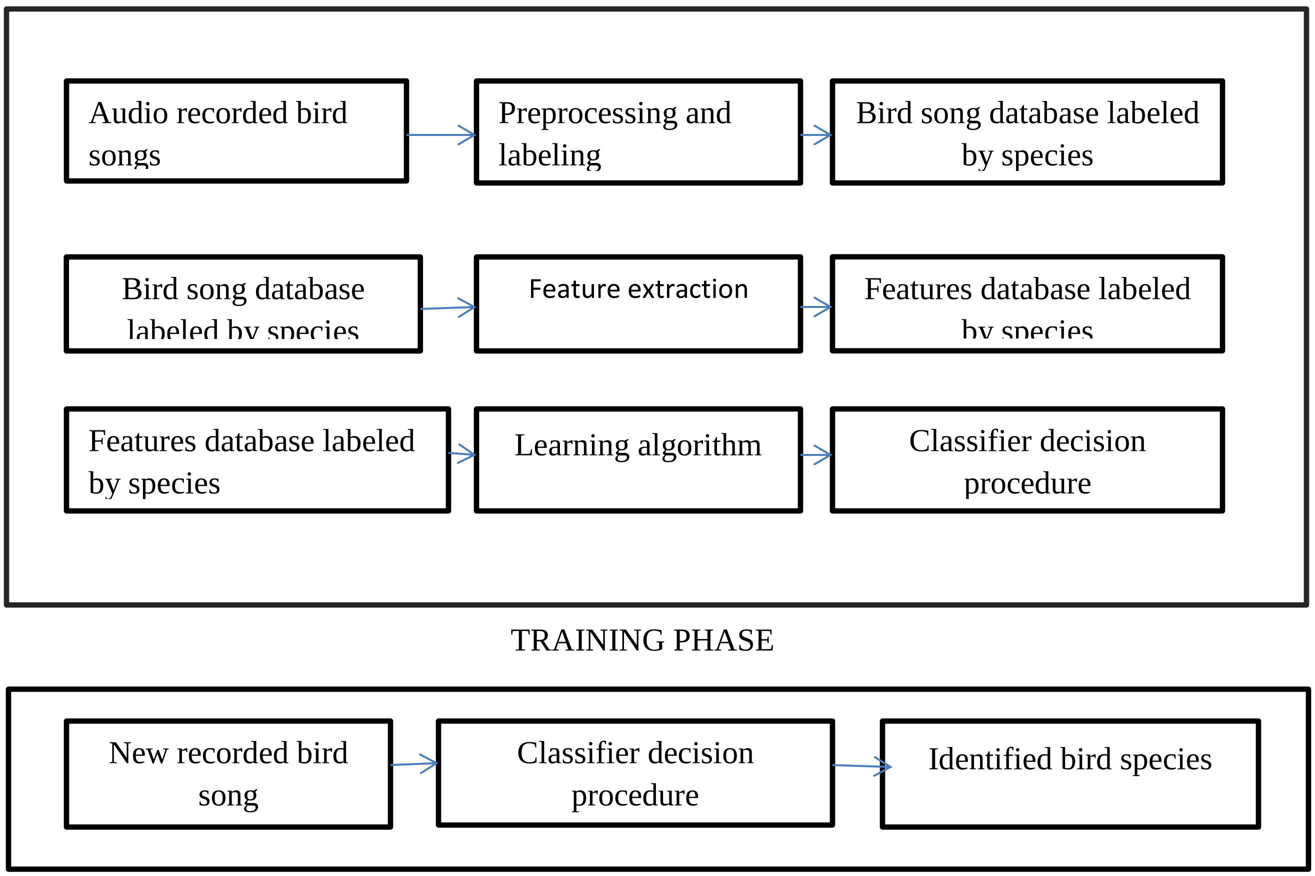
This process includes:

1) The bird song recording in the field;

2) The use of audio pre-processing techniques to improve the signal quality, because these recordings usually occur in noise environments;

3) The extraction of features from the audio signal;

4) The use of this feature in machine learning algorithms to produce a decision procedure that is able to predict the bird species.



CLASSIFICATION PHASE

Fig 3 : The general classification frame work

The practical reasons previously mentioned justify the study of mechanisms for bird species identification. In this paper authors focus on the automatic identification of bird species using signal processing and machine learning techniques. Until few time ago it was necessary to have a direct contact with a bird in order to determine its species. This situation has changed recently, with the use of automatic computational devices. One way to indirectly identify the species of a bird is using its recorded song.

Timos Papadopoulos (2018) proposed the automated bird sound identification system to produce the accurate identification of bird species. For this author trained the classification algorithms and collect the bird audio data to avoid the manual processing. Author assess the performance of the proposed bird sound recognition system in number of candidate classes, based on the encountered in real conditions. Author also use the threshold selection method to categorize bird sound from silence in the crowd-sourced of the trained data. And clean the un wanted sound that appeared in that audio. Sample data was gathered from hand-curated recordings for choosing the related application scenario hence the end user can present the result in accurate manner with no extra information, to the identification system. For this approach they use two classification methods

- k Nearest Neighbour (kNN)
- Support Vector Machine (SVM)

k Nearest Neighbour (kNN), it is for histogram-based features and another one is Support Vector Machine (SVM) with time summarisation features. Also they further investigated the use of a certainty measure, resulting from the output likelihoods of the classifiers, to improve the interpretability and consistency of the class decisions.

Marini, et al (2015) presents a new approach for the bird’s species recognition which is based on the unconstrained bird images by extracting the visual features and audio features extracted from the bird vocalizations. Local feature identification was made by using the Scale Invariant Feature Transform (SIFT) method with the help of SVM classifier. By the occurrence of the classification is not made properly hence it reject and reclassifies using Mel-frequency cepstral coefficients (MFCCs) which were used for extracting the bird song. For this analysis they put on experiments conducted on 50 dataset of bird species and audio samples and they got the result of 1.2 and 15.7 percentage points improvement using the acoustic classifier.

Sangita S.Londhe (2015) Automatic identification of bird species by their vocalization is studied in this paper. Bird sounds are divided by their function into songs and calls which are further divided into hierarchical levels of phrases, syllables and elements. It is shown that syllable is suitable unit for recognition of bird species. Diversity within different types of syllables birds are able to produce is large. Automatic recognition system for bird species used in this paper consists of segmentation of syllables, feature generation, classifier design. Support vector Machine is chosen for species identification.

Bird sounds are typically divided into categories of songs and calls depending upon their function. Generally, songs are longer and more complex than calls and occur more spontaneously. The main function of songs is related to breeding and territorial defense. Many bird species sing only during the breeding season and is generally further limited to males only. Call sounds are typically short vocalizations that carry a function, for example, an alarm, flight, or feeding. Distinguishing between songs and calls can sometimes be ambiguous and hence the separation of bird sounds into these categories is not studied in this work. Bird vocalizations are often divided into hierarchical levels of phrases, syllables, and elements.

The segmentation of a recording to individual syllables is performed using an iterative time-domain algorithm. First, a smooth energy envelope of the signal is computed and the global minimum energy is selected as the initial background noise level estimate NdB. Initial threshold TdB is set to the half of the initial noise level, which is set to the lowest signal envelope energy. Noise level and threshold are updated using the following algorithm until convergence so that the noise level is sufficiently stable. Algorithm:

1) Find syllable candidates, i.e., regions that are above syllable threshold TdB.

2) Update NdB from gaps between syllable candidates.

3) Update the threshold, e.g. TdB=NdB/2, and return to step 1.

Bird Name	**Scientific Name**	**No. of Files**	**No. of Syllables**	**No. of Training Syllables**	**No. of Testing Syllables**
Sparrow	Passer domesticus	18	777	543	234
Crow	Corvus splendens	19	422	295	127
Koel	Eudynamys scolopaceus	18	647	453	194
Red vented bulbul	Pycnonotus cafer	22	448	327	141
Common myna		18	959	671	288
Peacock	Acridotheres tristis	10	185	129	56
Indian cuckoo	Pavo cristatus	10	750	535	230

Table:1 Structure of Audio database

Salamon, et al, (2017) particularly concentrates on the automated classification of migrating birds. This classification is based on the Voice of the bird, voice would pay an attention for the enormous activity to monitor the biodiversity in nature and it is wide range of ecology field.

Flying birds produce new biological insights and conservation applications for the birds that produce the sound during relocating. This paper presents the classification technique for bird species classification while in flying bird using the voice, which produce the sound during the moving birds. Mainly concentrate on the “shallow learning” method based on unsupervised learning with a deep convolutional neural network with data growth method with both significantly outperforming an MFCC baseline.

Baowen Qiao et al (2017) Bird species recognition is a challenging problem due to the variant illumination and different view point of camera. In this paper, a new feature which is the ratio between the distance of the eye to the root of beak and the distance of the width of the beak is used to distinguish the different bird species.

- Integrated the new feature into the multi-scale decision tree and the SVM framework, a new bird species recognition algorithm is proposed to get the final recognition result.

- The Experiment results show that the proposed new feature can improve the correct classification rate about nine percent.

Iosif Mporas et al (2012) report on a recent progress with the development of an automated bioacoustics bird recognizer, which is part of a long-term project, aiming at the establishment of an automated biodiversity monitoring system at the Hymettus Mountain near Athens. In particular, employing a classical audio processing strategy, which has been proved quite successful in various audio recognition applications, we evaluate the appropriateness of six classifiers on the bird species recognition task. In the experimental evaluation of the acoustic bird recognizer, we made use of real-field audio recordings for seven bird species, which are common for the Hymettus Mountain. Encouraging recognition accuracy was obtained on the real-field data, and further experiments with additive noise demonstrated significant noise robustness in low SNR conditions.

Wei-Ho Tsai et al (2013)this study proposes an automatic bird sound identification system to help people learn to identify bird species. The system is built upon a two-stage identification framework. The first stage performs call/song classification. If an unknown sound clip is classified as a call, it is then handled by a call identifier in the second stage; otherwise, it is handled by a song identifier. Both identifiers use two acoustic features, timbre and pitch, to determine which of the bird species the sound clip belong to. In using timbre features, bird sounds are converted into MFCCs and their first derivatives and then analysed using Gaussian mixture models. In using pitch feature, we convert bird sounds into MIDI note sequences and then use bigram models to analyse the dynamic change information of the notes. Our experiments, conducted using a database including twenty common bird species, show that the proposed system can achieve 90.4% accuracy.

Liang Zhang et al (2015) analyse the Bird species richness survey is one of the most intriguing ecological topics for evaluating environmental health. Here, bird species richness denotes the number of unique bird species in a particular area. Factors affecting the investigation of bird species richness include weather, observation bias, and most importantly, the prohibitive costs of conducting surveys at large spatiotemporal scales. Thanks to advances in recording techniques, these problems have been alleviated by deploying sensors for acoustic data collection. Although automated detection techniques have been introduced to identify various bird species, the innate complexity of bird vocalizations, the background noise present in the recording and the escalating volumes of acoustic data pose a challenging task on determination of bird species richness. In this authors proposed a two-step computer-assisted sampling approach for determining bird species richness in one-day acoustic data. First, a classification model is built based on acoustic indices for filtering out minutes that contain few bird species. Then the classified bird minutes are ordered by an acoustic index and the redundant temporal minutes are removed from the ranked minute sequence. The experimental results show that our method is more efficient in directing experts for determination of bird species compared with the previous methods.

Guillermo Sarasa et al (2017) In this work, we analyze the usefulness of the normalized compression distance (NCD) as a similarity measure to bird species identification through audio samples. As a first approach we review the effect of different compression methods from 7z and CompLearn Toolkit, over subsets of bird audio samples obtained from the xeno-canto database. The performance of each compression method was measured applying hierarchical clustering and projection mapping to the distance matrix, and later on, measuring the quality of both of them. Our results are very promising and show that the identification of a bird species among multiples audio samples is possible through NCD-based-on clustering.

Min-Seok Kang et al (2018) from this paper, only the bird-species recognition keywords are input, and the Web image is refined as the learning data. We propose a method to generate a bird-species recognition model through learning a model based on refined training data. First, if you enter the name of the targeted bird breed, the image will be collected from the Web using the image crawl. To refine the collected images into the training dataset, the corrupted image is corrected and deleted, the outlier is removed, and finally the image is expanded to obtain the refined training data. In the process of modifying and deleting the images, the white-background image is deleted and the header is initialized. In the outlier removal, the features are extracted using the deep learning of the image data collected for each keyword. Then, the cluster distance of each label is measured using the K-means clustering that forms the training data according to the measured cluster distance. Lastly, the image expansion is performed to improve the training of the training data and for the learning accuracy of the refined database. To evaluate the performance of the proposed method, CUB-200 [1], an existing Bird Breed Dataset, is divided into training data and test data. In addition, the recognition-rate change according to various parameter changes is confirmed. As a result, the proposed method shows different recognition rates than the data collected from the existing institutions. We can use this proposal to provide varieties of training data and automated processes compared with the existing methods.

Martin Graciarena et al (2011) authors main goal of this work was to explore modeling techniques to improve bird species classification from audio samples. We first developed an unsupervised approach to obtain approximate note models from acoustic features. From these note models we created a bird species recognition system by leveraging a phone n-gram statistical model developed for speaker recognition applications. We found competitive performance from the note n-gram system compared to a Gaussian mixture model baseline using the same acoustic features. We found an important gain by doing score-level combination relative to the best individual system results. We verified that on most of the bird species under study there was a gain from system combination.

Peter Jancovic ; Münevver Köküer (2019)This paper investigates acoustic modeling for recognition of bird species from audio field recordings. First, the acoustic scene is decomposed into isolated segments, corresponding to detected sinusoids. Each segment is represented by a sequence of the frequency and normalized magnitude values of the sinusoid. The temporal evolution of these features is modeled using hidden Markov models (HMMs). A novel method for an unsupervised modeling of individual bird vocalization elements is proposed. The element models are initialized using HMM-based clustering and then further trained using an iterative maximum likelihood label re-assignment procedure. State duration modeling, performed in a post-recognition stage, is explored. Finally, we developed a hybrid deep neural network-hidden Markov model. The developed acoustic models are employed for bird species identification, detection of specific species, and recognition of multiple bird species vocalizing in a given recording. The detection system employs score normalization. Recognition of multiple bird species is performed based on maximizing the likelihood of a set of segments on a subset of bird species models, with penalization based on Bayesian information criterion applied. Experimental evaluations are performed on more than 37 h of sound field recordings, containing vocalizations of 48 bird species, plus more than 16 h of non-bird sound recordings. Using 3 s of the detected signal, the best system achieved: identification accuracy of 98.7%, detection with the equal error rate of 2.7%, and recognition accuracy of 97.3% and 95.4% when vocalizations of multiple bird species are present, with the number of bird species known and estimated, respectively.

Kunja Bihari Swain (2014) used a Cuckoo Search Algorithm (CSA) based neural network is proposed for noise removal from a signal. A new training function is proposed, which uses Cuckoo Search Algorithm for the training of the network. The trained network is then used to remove noise from a sinusoidal signal contaminated with white Gaussian noise. Various types of random walks and Levy flights are used in the algorithm and their effect on the final output has been analysed. A comparison is made between the present technique and already existing Levenberg-Marquardtback-propagation (LMBP). The compare-ison shows the superiority of the present approach over the Levenberg-Marquardt back-propagation.

Xuye Zhi and Chengan Guo (2018): Bird species recognition is one of the most challenging tasks in fine-grained visual categorizations (FGVC) and has attracted wide attention in recent years. In this paper, authors develop a bird recognition system that consists of three learning-type computational modules: the first one is for extracting the object and key parts from input images that is implemented by training a deep convolutional neural network (CNN) model, the second module is for feature extraction from the detected object and parts by using four other CNNs and further for computation of the high dimensional Gaussian descriptors based on the deep features, and the third module is for getting the final recognition result that is implemented by training four SVM classifiers with the Gaussian descriptors and integrating the outputs of the SVMs together with a decision fusion method.

Satya Dharanipragada (2017) studied Automatic identification of bird species by their vocalization . Bird sounds are represented with two different parametric representations:

(i) the mel-cepstrum parameters and

(ii) a set of low-level signal parameters, both of which have been found useful for bird species recognition. Recognition is performed in a decision tree with support vector machine (SVM) classifiers at each node that perform classification between two species.

Recognition is tested with two sets of bird species whose recognition has been previously tested with alternative methods. Recognition results with the proposed method suggest better or equal performance when compared to existing reference methods.

Naranchimeg (2018), In recent decade, many state-of-the-art algorithms on image classification as well as audio classification have achieved noticeable successes with the development of deep convolutional neural network (CNN). However, most of the works only exploit single type of training data. In this paper, we present a study on classifying bird species by exploiting the combination of both visual (images) and audio (sounds) data using CNN, which has been sparsely treated so far. Specifically, we propose CNN-based multimodal learning models in three types of fusion strategies (early, middle, late) to settle the issues of combining training data cross domains. The advantage of our proposed method lies on the fact that We can utilize CNN not only to extract features from image and audio data (spectrogram) but also to combine the features across modalities. In the experiment, we train and evaluate the network structure on a comprehensive CUB-200-2011 standard data set combing our originally collected audio data set with respect to the data species. We observe that a model which utilizes the combination of both data outperforms models trained with only an either type of data. We also show that transfer learning can significantly increase the classification performance.

Purohit Shrinivasacharya (2013) presented an effective image crawler to crawl the images from the WWW by using different search engines. This tool collected the images and its corresponding metadata for later uses. The crawled images were best input for the content-based image retrieval systems. It was observed that the performance this crawler was best for the CBIR system. The experiment was conducted with 1000 different text query for downloading the images from the different web sites. The enhanced reranking technique and giving the image itself as a query to extract the images from the Google and Yahoo needs to be adapted to get the attractive performances for feature work. This work is enforced mistreatment JAVA and Oracle SQL in Windows XP software. The analysis of the Image Crawler system is completed by submitting a text question to retrieve pictures from numerous classes of web pictures. We tend to conducted experiments on giving totally different keywords to extract the photographs from web. Once the keyword is submitted, it'll check its connected pictures area unit gift within the information or not. If information consists the connected pictures then it'll raise to update the information or terminate. If the choice is changing the information then it'll search the pictures within the web to gather the new images and stores its data within the information.

Hari Babu Srivastava et al (2009) The problem of detecting small/point targets in infra-red imagery is an important research area for defence applications. The challenge is to achieve high sensitivity for detection of dim point like small targets with low false alarms and high detection probability. To detect the target in such scenario, pre-processing algorithms are used to predict the complex background and then to subtract predicted background from the original image. The difference image is passed to the detection algorithm to further distinguish between target and background and/or noise. The aim of the study is to fit the background as closely as possible in the original image without diminishing the target signal. A number of pre-processing algorithms (spatial, temporal and spatio-temporal) have been reported in the literature. In this paper a survey of different pre-processing algorithm is presented. An improved hybrid morphological filter, which provides high gain in signal-to-noise plus clutter ratio (SCNR), has been proposed for detection of small/point targets.

Grant Van Horn (2015) has introduced tools and methodologies to collect high quality, large scale fine-grained computer vision datasets using citizen scientists – crowd annotators who are passionate and knowledgeable about specific domains such as birds or airplanes. Author worked with citizen scientists and domain experts to collect NABirds, a new high quality dataset containing 48,562 images of North American birds with 555 categories, part annotations and bounding boxes. Author fined that citizen scientists are significantly more accurate than Mechanical Turkers at zero cost. Author worked with bird experts to measure the quality of popular datasets like CUB-200-2011 and ImageNet and found class label error rates of at least 4%. Nevertheless, author found that learning algorithms are surprisingly robust to annotation errors and this level of training data corruption can lead to an acceptably small increase in test error if the training set has sufficient size. At the same time, author found that an expert-curated high quality test set like NABirds is necessary to accurately measure the performance of fine-grained computer vision systems. We used NABirds to train a publicly available bird recognition service deployed on the web site of the Cornell Lab of Ornithology.

Andr´eia Marini (2016), presents a novel approach for bird species classification based on color features extracted from unconstrained images. This means that the birds may appear in different scenarios as well may present different poses, sizes and angles of view. Besides, the images present strong variations in illuminations and parts of the birds may be occluded by other elements of the scenario. The proposed approach first applies a color segmentation algorithm in an attempt to eliminate background elements and to delimit candidate regions where the bird may be present within the image. Next, the image is split into component planes and from each plane, normalized color histograms are computed from these candidate regions. After aggregation processing is employed to reduce the number of the intervals of the histograms to a fixed number of bins. The histogram bins are used as feature vectors to by a learning algorithm to try to distinguish between the different numbers of bird species. Table 1 is the experimental results on the CUB-200 dataset show that the segmentation algorithm achieves 75% of correct segmentation rate. Furthermore, the bird species classification rate varies between 90% and 8%, depending on the number of classes taken into account.

Fusion Rule	Correct Classification Rate (%)			
	Number of Classes			
	2	5	17	200
MAX	85.29	47.68	19.65	6.76
SUM	86.76	49.01	22.16	7.16
PROD	88.24	49.67	22.54	7.25
WSUM	89.71	51.66	23.89	7.59
WPROD	91.18	51.66	23.70	8.03

Table:2 CLASSIFICATION RATES FOR THE FUSION

Based on the results presented in this study and the performance of the related works, authors can assert that color features are interesting alternatives for bird species identification problem, however, the best results reported have been achieved with SIFT features.

Erika Vilches (2005), In this work author explore the application of data mining techniques to the problem of acoustic recognition of bird species. Most bird song analysis tools produce a large amount of spectral and temporal attributes from the acoustic signal. The identification of distinctive features has become critical in resource constrained applications such as habitat monitoring by sensor networks. Reducing computational requirements makes affordable to run a classifier on devices with power consumption constraints, such as nodes in a sensor network. Experimental results demonstrate that considerable dimensionality reduction can be achieved without significant loss in classification efficiency. Once authors converted the entire species data sets into quantized data, authors proceeded to process the information with a decision tree algorithm. “Decision tree algorithms use full binary trees that represent the comparisons between elements that are performed by a particular sorting algorithm operating on an input of a given size.” The ID3 algorithm was used to generate the decision tree with Weka software. ID3 is a decision tree algorithm that takes all unused attributes and counts their entropy concerning the test samples to be used.

Herve Goeau (2017), The LifeCLEF bird identification challenge provides a largescale testbed for the system-oriented evaluation of bird species identification based on audio recordings. One of its main strength is that the data used for the evaluation is collected through Xeno-Canto, the largest network of bird sound recordists in the world. This makes the task closer to the conditions of a real-world application than previous, similar initiatives. The main novelty of the 2016-th edition of the challenge was the inclusion of soundscape recordings in addition to the usual xeno-canto recordings that focus on a single foreground species. This paper reports the methodology of the conducted evaluation, the overview of the systems experimented by the 6 participating research groups and a synthetic analysis of the obtained results Table 4 below.

RunNameShort	Official score: mAP(with background species)	mAP(only main species, same queries BirdCLEF2015)	mAP with background species(only queries 2016"Soundscape")
Cube Run 4	0.555(1)	0.686(1)	0.072(5)
Cube Run 3	0.536(2)	0.660(2)	0.078(4)
Cube Run 2	0.522(3)	0.649(3)	0.066
MarioTsaBerlin Run 1	0.519(4)	0.585(4)	0.137(1)
MarioTsaBerlin Run 4	0.472(5)	0.551(5)	0.129(3)
WUT Run 4	0.412	0.529	0.036
MarioTsaBerlin Run 3	0.396	0.456	0.130(3)
WUT Run 2	0.375	0.483	0.032
WUT Run 3	0.352	0.455	0.029
WUT Run 1	0.35	0.453	0.027
BME TMIT Run 2	0.338	0.426	0.053
BME TMIT Run 3	0.337	0.426	0.059
MarioTsaBerlin Run 2	0.336	0.399	0.000
BME TMIT Run 4	0.335	0.424	0.053
BME TMIT Run 1	0.323	0.407	0.054
Cube Run 1	0.284	0.364	0.020
LSIS naïve MFCC Run 1	0.149	0.183	0.037
BIG Run 1	0.021	0.021	0.004
Best run BirdCLEF 2015	-	0.454	-

Table:3 Results of the Life CLEF 2016 Bird Identification Task

Interestingly, one can see that the performance does not seem to be correlated to the number of training samples. In the same way, authors did observe that it is not correlated to the average length of the recordings in the class. This means that the high variability in performance is more related to other factors such as (i) the bird sounds variability (some birds are more audible than others), (ii) the acquisition difficulty (some birds are easier to record than others), (iii) the degree of confusion across close species. Another interesting remark is that two of the species that are not recognized at all by the CNN are comparatively pretty well recognized by the template matching kernel approach of MNB TSA. Thus, it would be interesting to study in more details the kind of audio patterns that have been matched by their method so as to understand what the CNN missed and how such patterns could be automatically learned as well. The main outcome was that after two years of resistance of bird song identification systems based on engineering features, convolutional neural networks finally managed to outperform them with a significant margin. It is noticeable that the best performing CNN did not used any fine-tuning so that it did not benefit from the transfer learning capacities of that techniques. Authors could thus expect even better performances. Also, the used CNN architecture was mostly inspired by the ones which perform the best on computer vision tasks.

Summary of this chapter

In this chapter 2 the introduction of biodiversity conservation about monitoring the birds, and the improvement of Science and technology for the helpful of wild animals, is explained. We also have and described about the bird species and its uses, Monitoring of specific habitats and bird species recognition. We have further explained about the about machine learning and deep learning. Finally, various researches review about of bird species recognition have been described in this chapter.

CHAPTER 3

RESEARCH METHODOLOGY

INTRODUCTION

This chapter discusses about the methodologies incorporated in the thesis. The dataset was retrieved from the web crawler dataset of the bird images, specially flying bird images to find out the species of the birds. This research work implemented in four phases such as:

1. Image Retrieval
2. Image Pre-Processing
3. Image Segmentation
4. Species Recognition using HMM

These phases are discussed below.

Several bird species classification challenges with closely related, but different, task descriptions have been held during the last few years. The interest and participation in these challenges have been high which indicates that these are problems that gain relevance in the present times and there is a need to solve them. The challenges usually include, predicting which species are present in a set of recordings with hidden labels, called the test set, and submitting the predicted species for each test data point for evaluation against the ground truth labels. The task description can vary from predicting only the presence or absence of birds in a recording to predicting all actively singing bird species. That is, the challenges have a varying degree of difficulty. Ornithologists study birds; their existence in nature, their biology, their songs, their distribution, and their ecological impact. Bird species identification is a recent research area, identifying the species from the image is an important and challenging issue. As many applications in the real world is struggling to protect the environment and the endangered animals, there are many reasons to consider the monitoring of the birds. Evaluating the quality of our environment is an important activity for analyzing the wild animals and also birds. There are many birds and it is sensitive to the environmental changes, and it is easy to monitor when compared to other species. In this way, the utilization of automated strategies for bird's species recognition is an effective method to evaluate the quantity and diversity of the bird's species which show up in a region. The practical reasons recently referenced legitimize the investigation of systems for bird's species recognition. Bird's species distinguishing proof is a notable issue to ornithologists, and is considered as a logical assignment since vestige. Bird's species distinguishing proof is dependent on images and involves additional testing because of the variety of the foundation and enlightenment in light of the fact that the greater part of the new images is gathered in their normal territory. In these images one cannot control rotation, scale and viewing angle at the time of image acquisition. Utilizing sound records instead of bird's images is defended by current practices. Capturing the birds' image and identifying of species is a difficult task than audio identifying of species.

Bird's classification is normally done by ornithology specialists based on an animal classification system. Birds are regularly sorted by their shape or outlines and physical attributes. Bird species classification is a difficult issue both to people and to computational calculations that attempts to carry out such an undertaking in a programmed manner. In the most recent years, a few methodologies dependent on bioacoustics signals have been proposed. Such approaches have achieved very interesting correct classification rates, based on the number of bird species take into account. Other than the fascinating outcomes accomplished with bioacoustic signals, this issue has been handled additionally based on birds' images. Contrasted and sound distinguishing proof, visual highlights are not very much concentrated for bird's order. The grouping issue can be expressed as given a bird's picture, characterize its species among a fixed yet huge number of conceivable outcomes. The test of such a grouping errand is because of the various out of sight and light since the vast majority of the images are assembled on bird's characteristic territory and in the birds presence since it's impractical to control turn, scale, and edge of view while securing images. Visual properties (e.g., shape, shading, marks, and so on.) are significant keys for bird's classification. A few specialists use these highlights to naturally distinguish birds.

Fundamental birds distinguishing proof pieces of information are acoustic and visual. The principle visual intimations include the birds 's outline, its plumage and shading. In any case, we should consider the season since birds' plumage will change during various seasons. The acoustics pieces of information include the tunes and calls that birds make. Moreover, the birds' conduct and environment it is found in are likewise potential hints. Field denotes, the imprints that distinctive one fowl from another are additionally significant, for example, bosom spots, wing bars which are portrayed as dainty lines along the birds, eye rings, or crowns, eyebrows which are depicted as lines over the eyes, eye lines, which are lines through the eyes, and numerous others. The state of the snout is regularly a significant intimation. Size can more often than not be a decent marker of the types of the fowl. For instance, most melody birds' creatures fit into a specific size gathering. Birds shape and stance are the most significant attributes used to recognize birds s. Most specialists can distinguish a bird's creature from its shape or outline since this is the most outlandish trademark to change. The tail of a birds can have numerous varieties. The tail can be scored, long and pointed, or adjusted. The legs can be long, or short.

Regardless of having a wide range of highlights that can consider to recognize bird’s species, from the example acknowledgment perspective, depending on a significant number of such highlights to distinguish diverse bird’s creature species isn't unimportant principally because of the trouble in getting standard fowl images. The procurement of fowl images is to some degree out of the control of the analysts and this forces a phenomenal changeability which is exceptionally hard to be taken care of by the majority of the Image handling, highlight extraction techniques and learning calculations. The Image was gathered from the Image creep. To refine the gathered images into the preparation dataset, the tainted Image is remedied and erased, the exception is evacuated and finally the Image is extended to acquire the refined preparing information. What's more, the highlights were spoken to as factual highlights, which give data on the area, fluctuation and presence of the conveyance of information and furthermore to guarantee order, Hidden Markov Model (HMM) classifiers can be utilized.

Bird’s species are perceived as valuable biodiversity markers. They are receptive to changes in touchy environments, while populace level changes in conduct are both unmistakable and quantifiable. Information about bird’s populace is subsequently a significant instrument for biologists in a wide scope of conditions and settings, including farmland use, marine settings, and relocation conduct. The work we present here spotlights on the characterization of winged creature species. For this we propose a bird’s species recognition system which is represented with three modules. The image was collected from the image crawl. To refine the collected images into the training dataset, the corrupted image is corrected and deleted, the outlier is removed, and finally the image is expanded to obtain the refined training data. And the features were represented as measurable highlights, which give data on the area, inconsistency and presence of the dissemination of information and furthermore to guarantee grouping can be performed by utilizing Hidden Markov Model (HMM) classifiers. The parameters of a shrouded Markov model are of two types, change probabilities and emanation probabilities.

3.1.1 Phase 1: Image Retrieval

Image recovery is the way toward seeking and recovering images from an enormous dataset or World Wide Web (WWW). As the Image develops in the database or WWW, recovery of the right images turns into a troublesome assignment and it is testing. The vast majority of the online internet searcher utilizes the normal strategies for Image recovery abuse some technique for amassing the metadata, for example, document names, inscribing, watchwords or depictions to the images built by human. In this way, that recovery can be performed over the explanation words instead of the substance of the Image. The images retrieved from icrawler (ICrawlerRPG, U.S.). The method for finding the WWW images is nothing but browsing the several webpages and extracting the related text and file name extensions to identify the bird's image.

3.1.2 Phase 2: Image Pre-Processing

The pre-processing algorithms are important in the small target detection. The convenience of different spatial just as transient channels has been examined. Regardless of the strategy picked, the pre-handling channel ought to play out the errands of upgrading the disconnected focuses. The identification and handling of the little moving targets is a significant subject in the zone of sign/picture preparing, particularly in the state of confused foundation and long range. The difficulties for detection of small/point targets arise out of the facts that the target occupies single or few pixels on the image plane, has no texture and shape information and has low signal-to-noise plus clutter ratio (SNCR). Target and background are both complex and dynamic. Pre-processing algorithms are used to enhance the SNCR to improve detection probability. For this

Gaussian Filter was used as in Fig.5. The Gaussian smoothing15 operator is a 2-D convolution operator that is used to blur images and remove detail and noise. It uses a kernel that represents the shape of a Gaussian (bell-shaped) hump. In 2-D space, an isotropic (i.e., circularly symmetric) Gaussian has the form:

$$G(m,n)=1/(2\pi\sigma^2)\, e^{((m^2+n^2)/(2\sigma^2))}$$

Gaussian smoothing uses this 2-D distribution as a point-spread function [14]. This is achieved by convolution. Since the image is stored as a collection of discrete pixels, a discrete approximation to the Gaussian function is needed to perform the convolution. The kernel can be truncated at s (s: standard deviation) limits from the mean. We represent color features by color moments and color log-polar; shape features by shape moments; and texture features by Gaussian filters and grayscale histogram.

Fig.4

Fig. 5

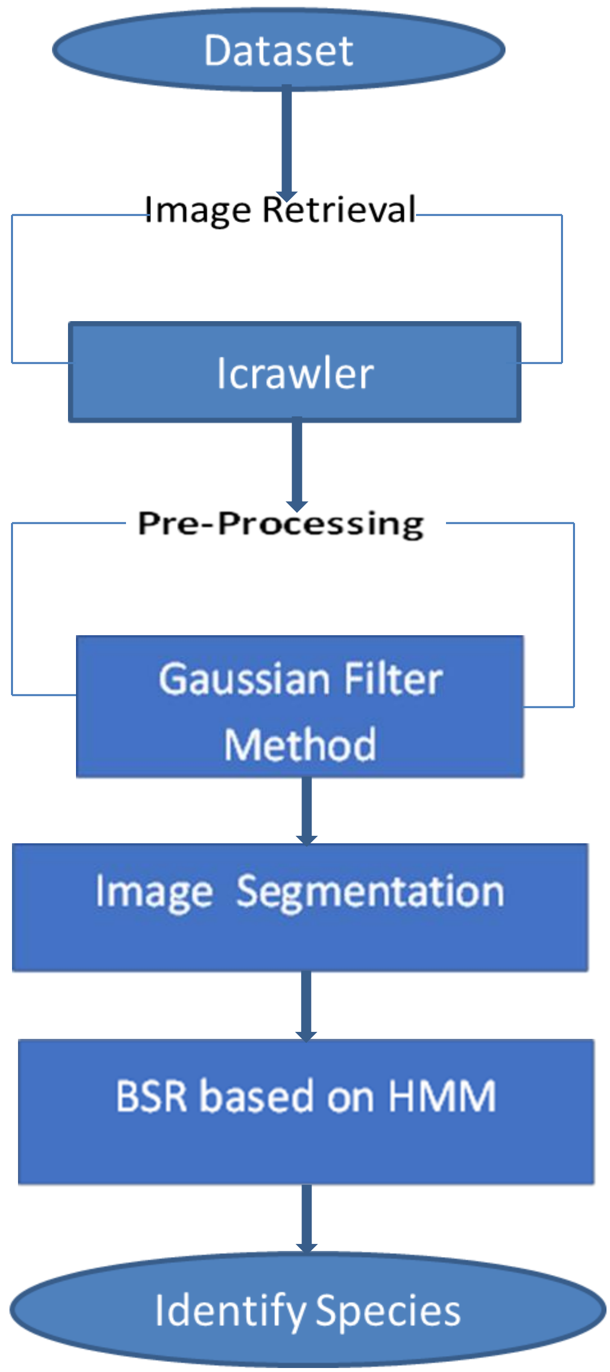
3.1.3 Phase 3: Image Segmentation

The environment where the birds are found when the image are acquired, acts as an unfavourable factor. Therefore, any attempt to eliminate the background before extracting features should be considered. Although, eliminating the background could be considered a problem as difficult as the species identification itself. The segmentation step is based on the assumption that all available images are in colors, that the birds are at the central position in the images, and that the bird's edges are far away from the image borders. Therefore, there are some strips at the image borders that can be considered as belonging to the image background. The size of these strips is chosen to be a percentage, usually between 2% and 10% of the image horizontal and vertical dimensions. First, these strips are scanned and the colors that are found into them are stored in a ranked list according to the color frequency. Next, a search procedure is carried out on the remaining of the image and the pixels that have similar colors to those found in the strips are labelled as background; otherwise they are labelled as" birds". At the end we have all pixels in the image labelled either as background or birds. To classify bird's species, color, shape and texture are important features. Color is an important feature in the image classification process. Details of the color segmentation approach them together other feature types. To obtain relevant information about the relationship of the color in the image, both the red, green and blue (RGB) and the hue, saturation and value (HSV) spaces are considered. RGB is the most popular color space used in electronic systems for sensing, representation and display of images. Birds have many kinds of gestures such as side of the camera and back to the camera, but the side of the bird's head is usually caught. In the process of identifying bird of human, beak is a particularly important factor, beak feature of many birds are so various that people can identify easily. Finding characteristics of the beak is an available method for recognizing the birds. The height of the beak and the width of the beak are typical characteristics in the characteristics of the beak in our eyes. Due to the size of different pictures is not the same, characteristics should take a relative value. The ratio of the height of beak and the width of beak (R-HBWB) as a feature to describe the beak, but it lost the information of the length because of the change of the height of the beak, how to determine what kind of beak belong to is a problem. Therefore, a new beak feature has been proposed. Compare with the width of beak, range of change of the distance between the eye and the root of beak is relatively small. The beak is described by using the ratio between the distance of the eye to the root of beak and the distance of the width of beak.

3.1.4 Phase 4: Species Recognition using HMM

The proposed bird's recognition system is illustrated in Fig.6 that is composed of two computational modules: the feature extraction, and the classification of bird's species. Given an input image as in Fig.4 the object, head and body of the birds in the image are detected/localized and then segmented through the detection module.

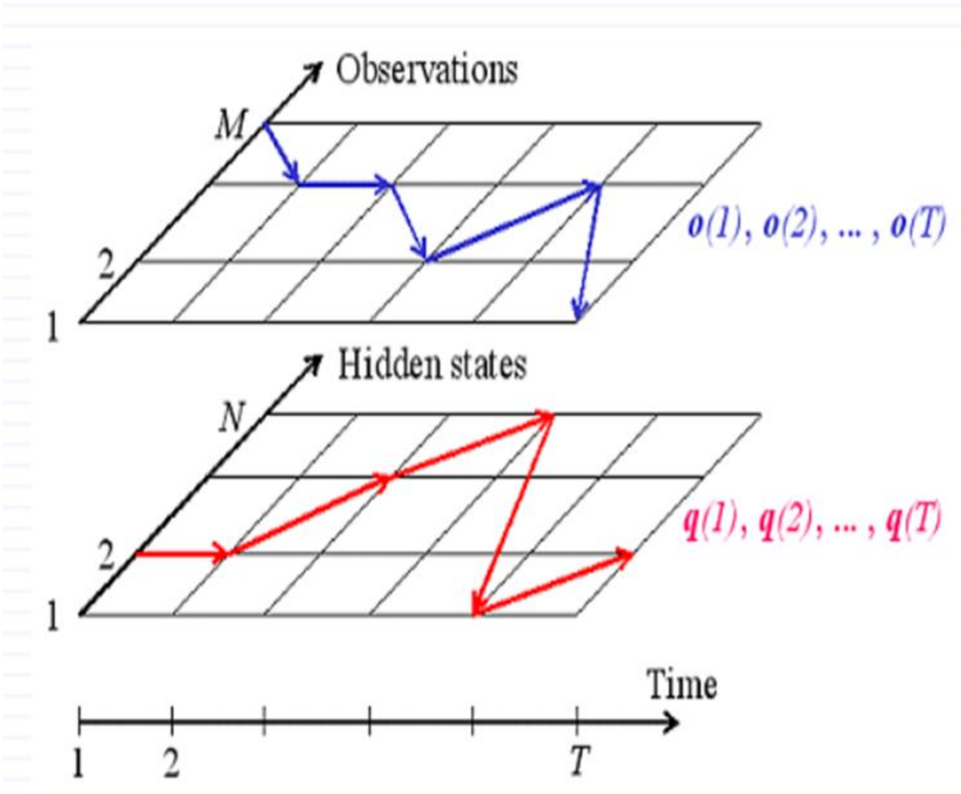
Fig. 6 shows the proposed architecture of our system.



3.2 Application of Hidden Markov Model (HMM)

In the broadest sense of the word, a hidden Markov model is a Markov process that is split into two components: an observable component and an unobservable or 'hidden' component. That is, a hidden Markov model is a Markov process (Xk, Yk) $k \geq 0$ on the state space $E \times F$, where we presume that we have a means of observing Yk, but not Xk. Adopting terminology from signal processing, we will typically refer to the unobserved component Xk as the signal process and E as the signal state space, while the observed component Yk is called the observation process and F is the observation state space. HMM produces a succession of observables by moving from inert state to idle state as per the change probabilities and radiating a discernible (from a discrete arrangement of observables, for example a limited letter set) from each idle state visited by the discharge probabilities of the state. This model used to decide the probability of the arrangement of perceptions, foresee the following perception in the grouping of perceptions, and locate the in all likelihood hidden clarification of the succession of perception. Hidden Markov Model (HMM) is a factual Markov model in which the framework being demonstrated is thought to be a Markov procedure with in secret (concealed) states.

This methodology played out an arrangement of flying creature's images, and fitted an oval, which is utilized to acquire the birds s' parts for grouping. The technique accomplished a generally excellent grouping rate; all of the previously mentioned strategies utilize single images and appearance-based models for order; be that as it may, winged creature species likewise display recognizing practices (flying, moving, presents, and so forth) which could likewise be utilized to support strong computerized distinguishing proof.



Partial observation sequence Fig.7

In the partial observation sequence as in Fig.7 Ot = {o(1), o(2), ... , o(t)} to be produced by all possible state sequences that end at the i-th state.

And the unconditional probability of the partial observation sequence is the sum of αt(i) over all N states which is denoted as q(1), q(2),….q(T).

This is especially applicable to the ID of feathered creatures in flying, particularly at separation where appearance-based highlights, for example, shading will in general weaken, while movement highlights stay recognizable. Instating the article utilizing

$\alpha 1(i) = pi\ bi(o(1))$, $i = 1, \ldots, N$,

the restrictive likelihood of the halfway perception grouping from o(t+1) as far as possible to be created by all state successions that begin at I-th state.

$\beta t(i) = P(o(t+1), o(t+2), \ldots , o(T) \mid q(t) = qi)$.

Motion-features related with flying have not yet been well-investigated for robotized visual recognizable proof. Our framework present and assess another arrangement of appearance highlight, which we show tentatively is more dependable for ordering birds in flying than existing single-picture based classifiers.

Three strategies to fuse features: early fusion, middle fusion and late fusion. Early fusion, also known as feature level fusion, is a feature combination scheme that features from multiple modalities concatenated to form a merged feature vector. Middle fusion, also called mid-level combination, combines the high-level features learned by single network. Late fusion, also called decision-level fusion, combines the decisions of the uni-modal classifier and determine the final classification. In this paper, we use concatenation to combine low and high-level features, and summation or multiplication to combine decisions of the classifiers. The feature combination layers can be trained with standard back-propagation and stochastic gradient descent.

Set of states: {S1, S2….Sn}

Process moves from one state to another generating a sequence of states: Si1,Si2,….Sik

Markov chain property: probability of each subsequent state depends only on what was the previous state: $P(SIK \mid SI1, SI2, \ldots, SIK-1) = P(SIK \mid SIK-1)$

States are not visible, but each state randomly generates one of M observations (or visible states) $\{V1,V2,\ldots VM\}$

To define hidden Markov model, the following probabilities have to be specified: matrix of transition probabilities $A=(a_{ij})$, $a_{ij}= P(s_i \mid s_j)$, matrix of observation probabilities $B=(b_i (v_m))$, $b_i(v_m)= P(v_m \mid s_i)$ and a vector of initial probabilities $\pi=(\pi_i)$, $\pi_i = P(s_i)$. Model is represented by $M=(A, B, \pi)$.

Hidden Markov models appear in a wide variety of applications. To fix some ideas one might distinguish between two main classes of applications, though many applications fall somewhere in between. On the one hand, hidden Markov models naturally describe a setting where a stochastic system is observed through noisy measurements. For example, in communications theory, one might think of X_k as a (random) signal to be transmitted through a communications channel. As the channel is noisy, the receiver observes a corrupted version Y_k of the original signal, and he might want to reconstruct as well as is possible the original signal from the noisy observations. This is the origin of the signal/observation process terminology. On the other hand, it may be the process Y_k which is ultimately of interest, while the X_k represents the influence on Y_k of certain unobservable external factors. For example, one might think of Y_k as the market price of stock, where X_k is an unobserved economic factor process which influences the fluctuations of the stock price. We are ultimately interested in modeling the observed stock price fluctuations, not in the unobservable factor process, but by including the latter one might well be able to build a model which more faithfully reflects the statistical properties of the observed stock prices. It should be noted that even though $(X_k, Y_k)_{k\geq 0}$ is Markov, typically the observed component $(Y_k)_{k\geq 0}$ will not be Markov itself. Hidden Markov models can thus be used to model non-Markov behavior (e.g., of the stock price), while retaining many of the mathematical and computational advantages of the Markov setting. This course is an introduction to some of the basic mathematical, statistical and computational methods for hidden Markov models. To set the stage for the rest of the course, we will describe in the next two sections a number of representative examples of hidden Markov models in applications taken from a variety of fields, and we will introduce the basic questions that will be tackled in the remainder of the course. Before we do this, however, we must give a precise definition of the class of models which we will be considering.

3.3 Use of CNN Architecture

Finding the best CNN architecture is a time-consuming task and often done purely by intuition. Current state-of-the-art approaches try to tackle this issue with automated hyper parameter search. We decided to reduce the amount of possible design decisions and relied on current best practices for CNN layouts. All weighted layers (except for input and output layers) use Batch Normalization, Exponential Linear Units (ELU) for unit activation and are initialized using He-initialization. We wanted large receptive fields in our first convolutional layers, which have proven to be very effective for spectrograms during our experiments. We use filter sizes of 7x7 and 5x5 for larger inputs and 3x3 kernels for smaller input sizes in deeper layers.

Although the BirdCLEF classification task with 1500 classes, class imbalances and a distorted dataset is rather complex, shallow CNN architectures with classic layouts and only a few layers seem to be more effective than more complex highway net-works with multiple tens of layers like DenseNet or ResNet. We tried different implementations of state-of-the-art convolutional networks but found them inferior to our simple CNN architectures. This might be due to the fact, that the image do-main of spectrograms is very homogenous despite more than 1500 different signal types. Most spectrograms contain only little information, leaving most pixels blank.

CNN uses hierarchical features in its processing pipeline. The features from initial layers are primitive while late layers are high-level abstract features made from combinations of lower-level features. The CaffeNet consists of five convolutional layers (with max pooling layers following the first, second and fifth convolution layer) followed by three fully connected (FC) layers and a softmax classifier. Rectified linear unit is applied to every convolutional layer and fully connected layer and local normalization is applied in first and second convolutional layer. The process through this 8-layer CNN network can be treated as a process from low to mid to high-level features. We hypothesize, that combining the features of different layers in this pipeline can lead to achieve better performance.

3.4 Applying Deep learning algorithms

Deep learning algorithms learn more about the image as it goes through each neural network layer. For classifying Neural Network is used. Figure 8 represents layers of neural networks for feature extraction. The neural network is a framework for many machine learning algorithms. Neural networks consist of vector of weights (W) and the bias (B).

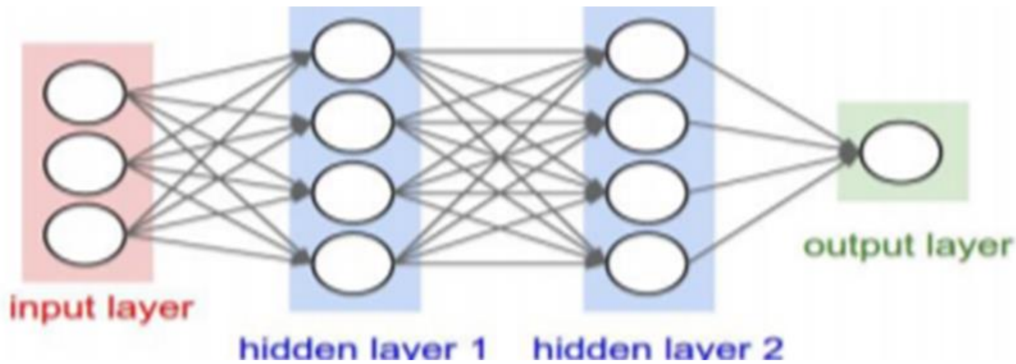


Fig.8 Structure of DNN

Fig.8 Three layers of Neural Network In deep learning, convolutional neural network (CNN) is a class of deep neural network mostly used for analyzing visual images. It consists of an input layer and output layer as well as multiple hidden layers. Every layer is made up of group of neurons and each layer is fully connected to all neurons of its previous layer. The output layer is responsible for prediction of output. The convolutional layer takes an image as input, and produces a set of feature maps as output [2]. The input image can contain multiple channels such as color, wings, eyes, beak of birds which means that the convolutional layer perform a mapping from volume to another volume. Image volumes considered are width, height, depth. The CNN have two components:

1) Feature extraction part: features are detected when network performs a series of convolutional and pooling operation.

2) Classification part: extracted features are given to fully connected layer which acts as classifier.

CNN consists of four layers: convolutional layer, activation layer, pooling layer and fully connected. Convolutional layer allows extracting visual features from an image in small amounts. Pooling is used to reduce the number of neurons from previous convolutional layer but maintaining the important information. Activation layer passes a value through a function which compresses values into range. Fully connected layer connects a neuron from one layer to every neuron in another layer. As CNN classifies each neuron in depth, so it provides more accuracy. Image classification: image classification in machine learning is commonly done in two ways:

1) Gray scale

2) Using RGB values

Normally all the data is mostly converted into gray scale. In gray scale algorithm, computer will assign values to each pixel based on how the value of the pixel is it. All the pixel values are put into an array and the computer will perform operation on that array to classify the data.

A convolutional layer takes an image as input, and produces a set of feature maps as output. The input image can contain multiple channels (e.g. RGB), which means that the convolutional layer learns a mapping from a 3D volume to another 3D volume. The layer consists of a number of convolution kernels each of which is made up of adjustable weights. The weights are adjusted during optimization using stochastic gradient decent. The feature maps are produced by performing a discrete convolution between each kernel in the convolutional layer and the input volume yielding one feature map for each kernel. Performing a convolution can be thought of as sliding a window along the input image and computing the dot product between the kernel and the values of the neighbourhood patch, or the so-called receptive field, in the image as seen through the window. The window has the same size as the kernel, and it is important to understand that the kernel has a depth extending through all the channels of the input image.

3.5 Summary

In this chapter we described about our methodology adopted. First we discussed about challenges of bird species classification, Birds classification technique and HMM model. Then discussion on the image retrieval process from the web crawler and ther methodology of pre-processing, retrieval, segmentation process for feature extraction process and about the CNN model and deep learning approach and finally HMM model is done.

CHAPTER 4

RESULTS AND DISCUSSION

This chapter described about our research work result. For this dataset were retrieved from the web crawler dataset of the bird images, specially flying bird images to find out the species of the birds.

Web crawlers

Web crawlers are more or less as same as the web. The spring of 1993 Matthew Gray [6] written a first web crawler World Wide Web named as “Wanderer” after a month the release of NCSA Mosaic, it was used since from 1993 to 1996 to accumulate statistics about the growth of the web. The David Eichmann has written the first research paper the RBSE spider containing a squat explanation of a web crawler. The Burner has published a first paper that describes the web crawler architecture, it is the original Internet Archive crawler. The Google search engine architecture was presented in the Brin and Page’s paper, this can be used as a distributed system of page-fetching method and a central database for coordinating the crawl. Brin and Page’s paper becomes the blueprint for the other crawlers. A distributed and extensible web crawler designed by Heydon and Najork described Mercator, that has become the outline for a number of other crawlers. The literature includes the other distributed crawling systems PolyBot, UbiCrawler, C-proc and Dominos. The text retrieval systems use the ranking and reranking approach to extract the best result from the search copies. Image retrieval is the process of searching and retrieving images from a huge dataset or WWW. As the image grows in the database or WWW, retrieval of the correct images becomes a difficult task and it is challenging. Most of the web-based search engine uses the common methods of image retrieval exploit some method of accumulating the metadata such as file names, captioning, keywords or descriptions to the images constructed by human. Large input sizes are not common in current image classification publications. Most approaches reduce the input size to a maximum of 256x256 pixels. Current consumer GPUs are well suited for larger inputs. On the other hand, models with large input sizes are considerably harder to train and tune, training takes significantly more time and larger inputs do not always benefit generalization. Our experiments showed that non-square, high-resolution inputs of spectrograms do indeed achieve better classification results especially for large and diverse datasets. We used stride convolutions and pooling layers to cope with large inputs. Additionally, a larger number of filters seems to be more effective than a larger number of hidden units. We found that 512 units per dense layer is sufficient even for 1500 classes. Determining the right amount of network parameters is crucial to avoid under- and overfitting. This process is also very time consuming considering the fact that less parameters might work well on small validation sets but usually underfit on larger datasets. Validation experiments should always show slight overfitting in order to have good generalization capacity when trained on more classes. Even though, models with a large number of weights did eventually overfit during our experiments with 1500 classes, so we decided to dial down the weight count.

Training

Time efficient training becomes crucial when training on 1500 classes with more than 940.000 samples. We tried to optimize our training process in order to save computation time and maintain a good overall performance at the same time. We evaluated different kinds of parameter settings and found the following to be very effective.

We performed experiments to evaluate the effectiveness or accuracy as in Fig.9, and Table 4. of our proposed appearance and motion feature sets. We assessed the appearance and movement sets autonomously, and contrasted and the outcomes acquired utilizing the HMM technique. The result is

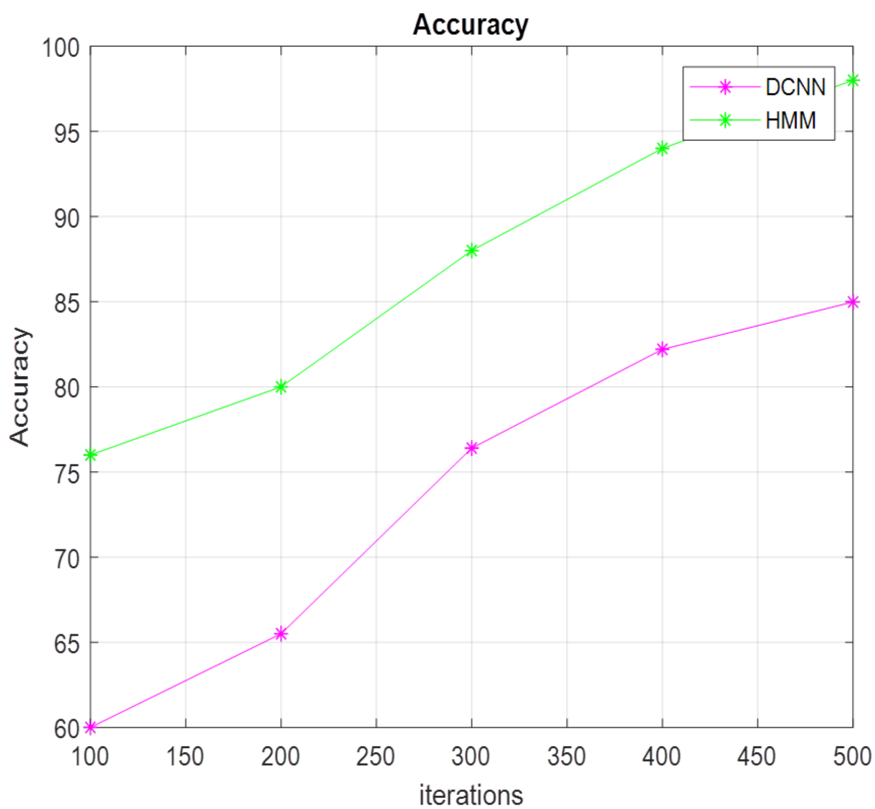


Fig.9: Comparison of Accuracy

No of Iterations	100	200	300	400	500
DCNN	60	65.05	76.04	82.2	84.98
HMM	76	80	88	94	98

Table 4: Accuracy comparison of existing and proposed method

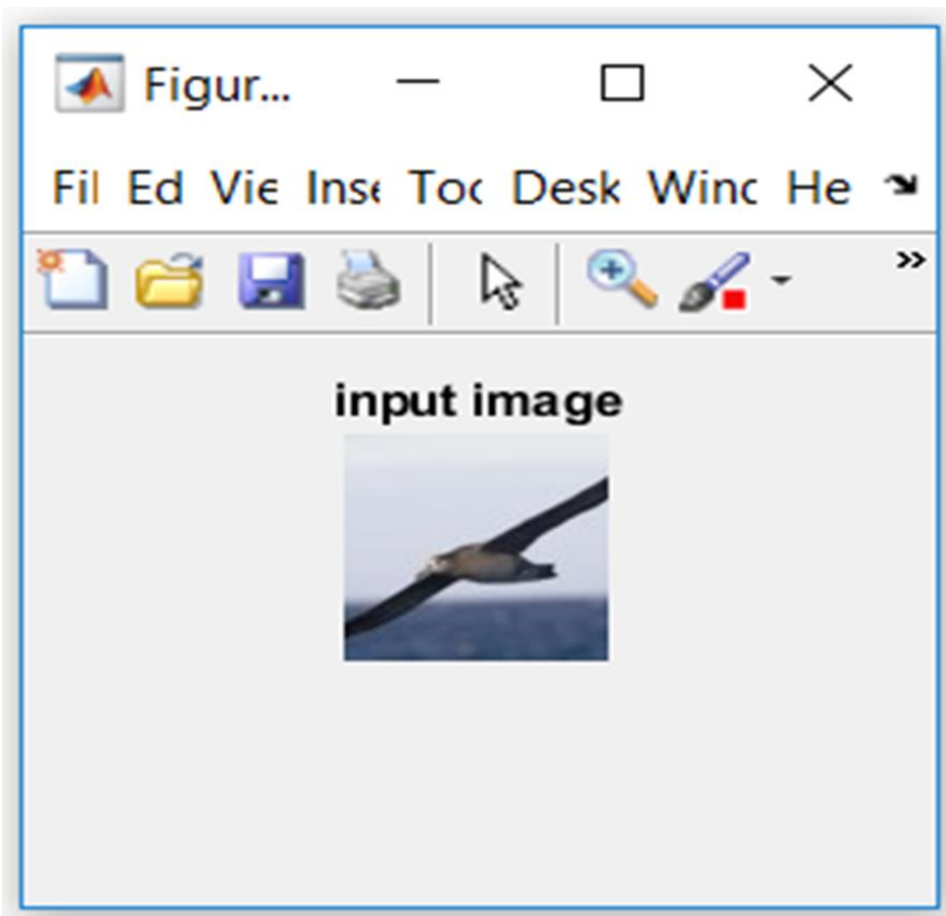
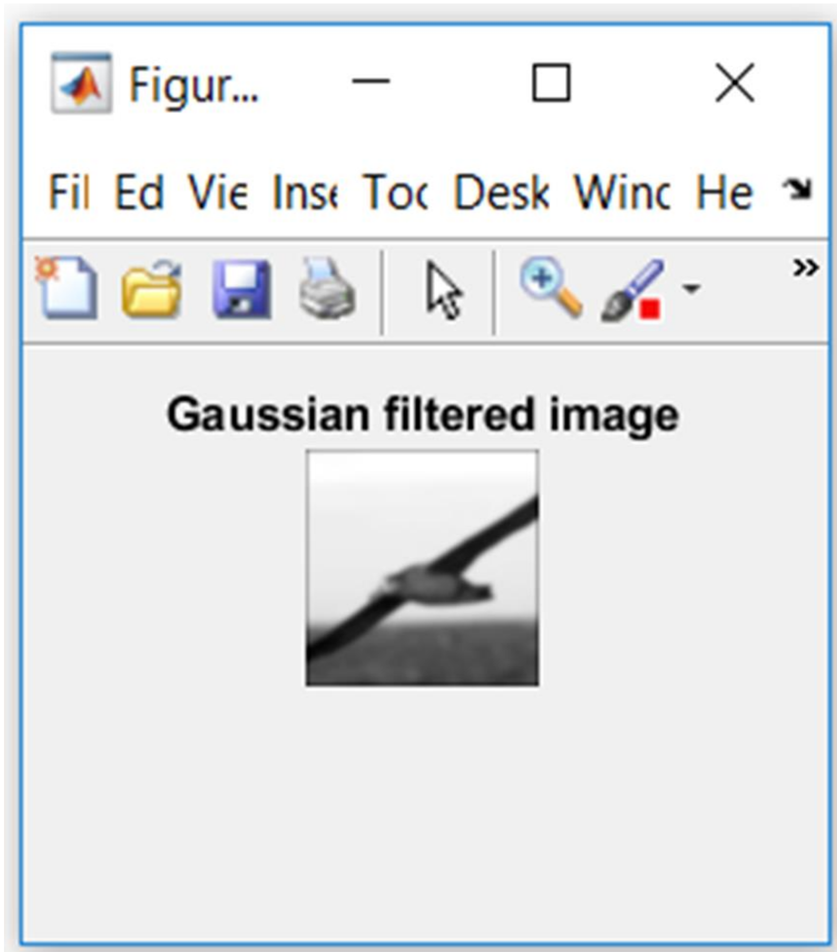


Fig 10: Input image of black footed albatross



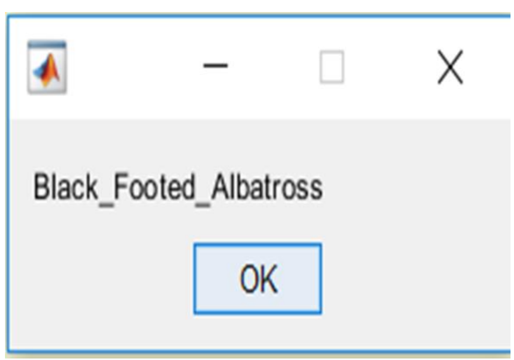


Fig 11: Gaussian filtered image and species detection for black footed albatross bird

From above shown fig 10it describes the input image of the flying bird. Capturing the flying bird image for the retrieval process. And the Fig 11 shows the filtering methods were applied to remove the noise data to project the data in clear manner and easy to extract and analyse the result in well manner it helps a lot fig 11 shown the result of filtering image for the bird of black footed albatross.

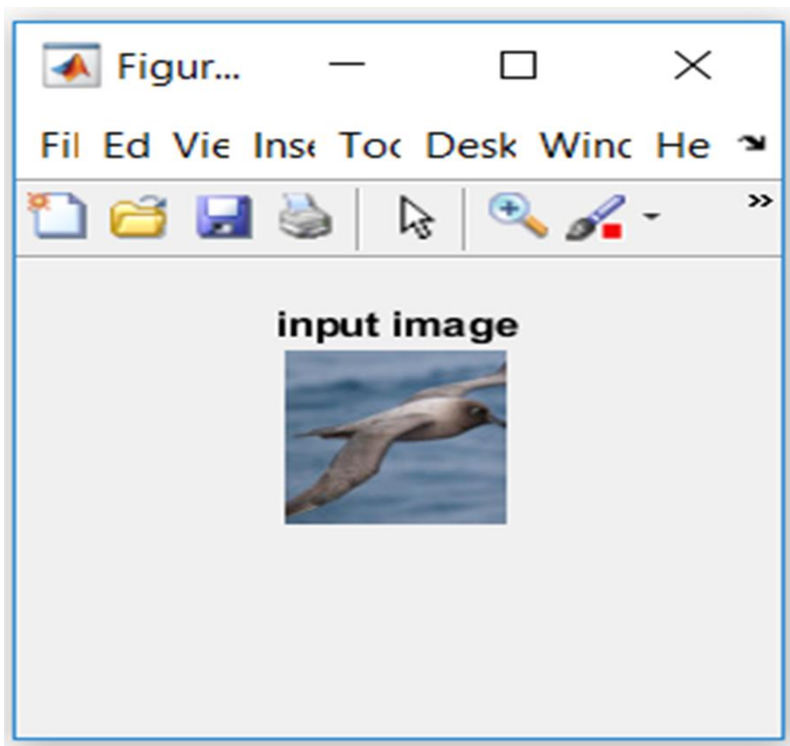
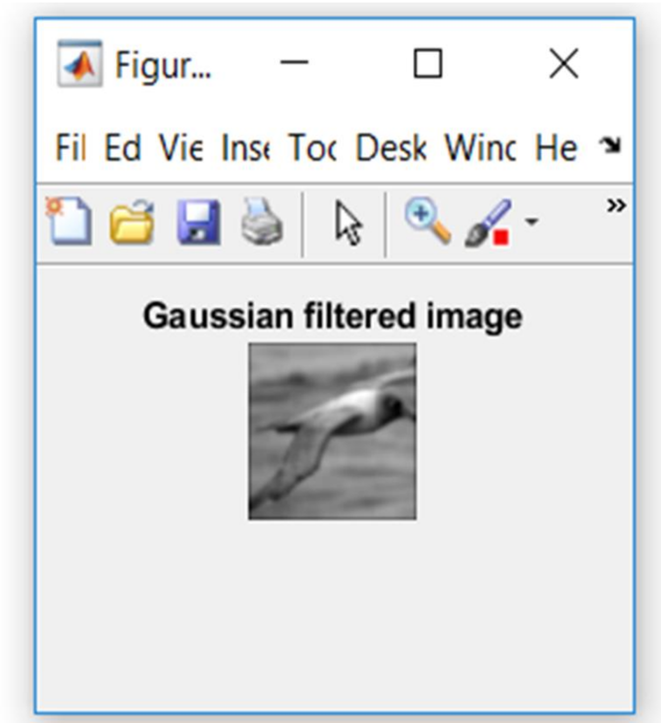


Fig 12: Input image of sooty albatross



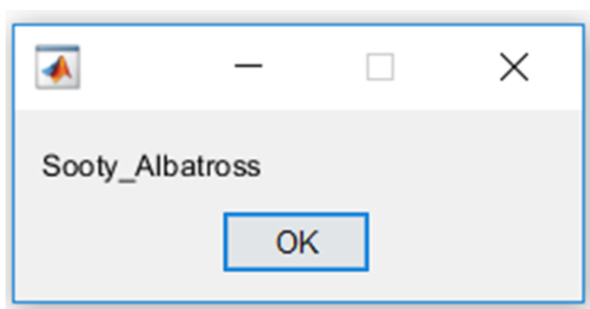


Fig 13: Gaussian filtered image and species detection for sooty albatross

In fig 12and13 shows the result of identifying the species of the bird of sooty albatross/. Filtering methods were applied to find the bird species filtering clear the noisy data

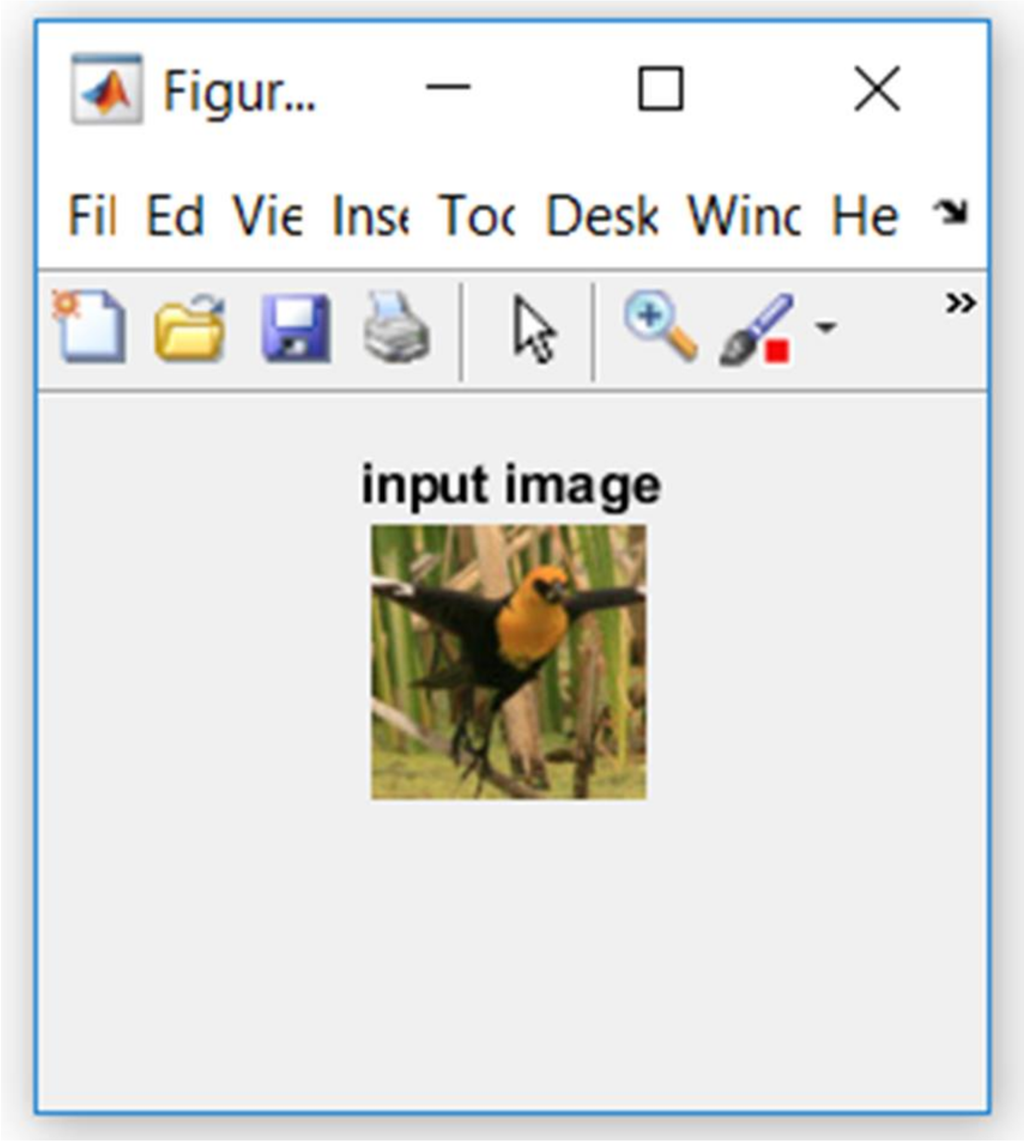
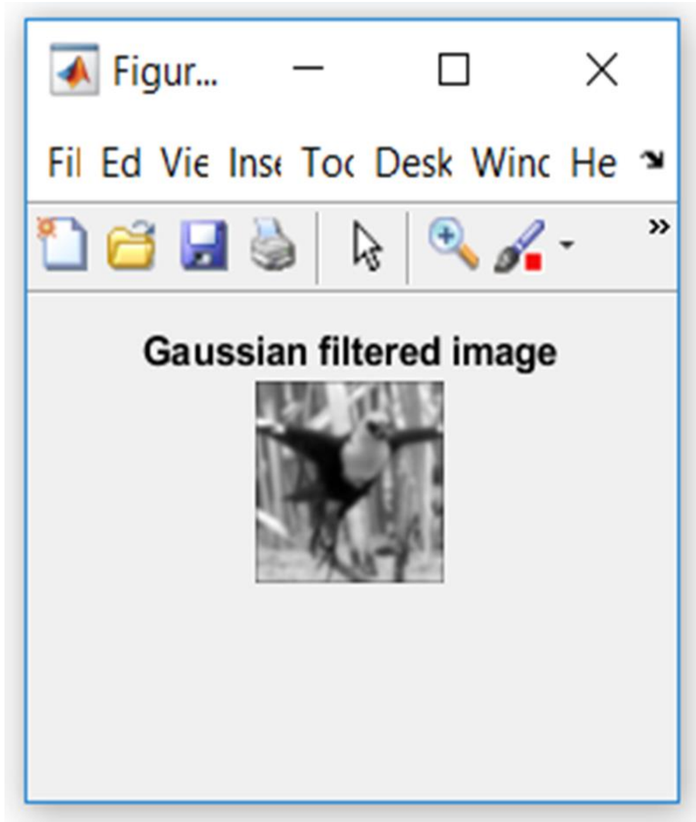


Fig 14: Input image of Japanese white eye



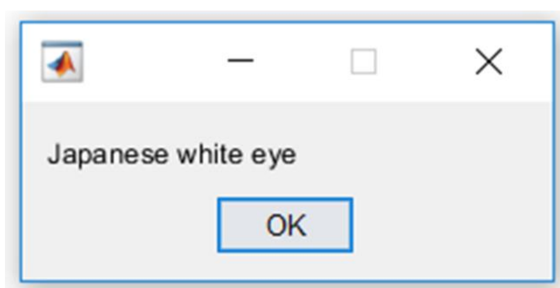


Fig 15: Gaussian filtered image and species detection for Japanese white eye

In fig 14 and15 shows the result of identifying the species of the bird of Japanese white eyes. This were applied to find the bird species.

CHAPTER 5

CONCLUSION AND FUTURE WORK

Bird species recognition is one of the most challenging tasks in recent years. Many researchers were concentrating on this work. It is interesting to identify the bird species. It is one of the useful biodiversity indicators for bird species recognition. Because birds are numerous and sensitive to environmental changes; also, and are easier to monitor than other species, observing the behaviour of birds allows us to better evaluate our living environment. Although pattern recognition has been more than half a century of history, identification for bird is still rare. In last years, several approaches have been proposed to recognize the bird. The problem of bird species identification can be defined as: given a bird image and/or bird song, assign a species among a fixed and large number of possibilities. However, the main question addressed in this paper is analyzing the bird species while moving bird also. In this paper we have described our proposed feature sets appearance and motion for species classification of flying birds, and presented supporting experimental results in which we compared our appearance features with existing one. For the outcome of accuracy, this paper uses Hidden Markov Model (HMM). HMM is a statistical Markov model in which the system being modeled is assumed to be a Markov process with unobserved (hidden) states with accuracy of 98%.

FUTURE RESEARCH WORK

- In future, automatic bird's species identification based on the motion features will be applied to find the bird species.

- Further, monitoring the health of an ecosystem in order to preserve the environment will also be considered in the future research work.

BIBLIOGRAPHY

[1]. Marini, A.A. Marini, A. J. Turatti, A. S. Britto Jr., and A. L. Koerich, “Visual and acoustic identification of bird species,” IEEE International Conference on Acoustics, Speech and Signal Processing, 2015, pp. 2309-2313.

[10]. M.T. Lopes, C.N. Silla Jr., A.L. Koerich, and C.A.A. Kaestner, “Feature set comparison for automatic bird species identification,” in IEEE Int’l Conf Sys Man Cybernetics, 2011, pp. 965–970.

[11]. Branson, Steve, Grant Van Horn, Serge Belongie, and Pietro Perona. "Bird species categorization using pose normalized deep convolutional nets." arXiv preprint arXiv:1406.2952(2014).

[12]. Sprengel, Elias, Martin Jaggi, Yannic Kilcher, and Thomas Hofmann. Audio based bird species identification using deep learning techniques. No. CONF. 2016.

[13]. Kahl, Stefan, Thomas Wilhelm-Stein, Hussein Hussein, Holger Klinck, Danny Kowerko, Marc Ritter, and Maximilian Eibl. "Large-Scale Bird Sound Classification using Convolutional Neural Networks." In CLEF (Working Notes). 2017.

[14]. Trifa, Vlad M., Alexander NG Kirschel, Charles E. Taylor, and Edgar E. Vallejo. "Automated species recognition of antbirds in a Mexican rainforest using hidden Markov models." The Journal of the Acoustical Society of America 123, no. 4 (2008): 2424-2431.

[15]. Brandes, T. Scott. "Feature vector selection and use with hidden Markov models to identify frequency-modulated bioacoustic signals amidst noise." IEEE Transactions on Audio, Speech, and Language Processing 16, no. 6 (2008): 1173-1180.

[16]. http://www.20search.com/image.php 20 SEARCH The Web's Best Image Search Engine List!

[17]. Shrinivasacharya, Purohit, and M. V. Sudhamani. "An Image Crawler For Content Based Image Retrieval System." International Journal of Research in Engineering and Technology eISSN: 2319-1163| pISSN: 2321 7308 (2013).

[18]. Vilches, Erika, Ivan A. Escobar, Edgar E. Vallejo, and Charles E. Taylor. "Data mining applied to acoustic bird species recognition." In 18th International Conference on Pattern Recognition (ICPR'06), vol. 3, pp. 400-403. IEEE, 2006.

[19]. Ngai, Eric WT, Li Xiu, and Dorothy CK Chau. "Application of data mining techniques in customer relationship management: A literature review and classification." Expert systems with applications 36, no. 2 (2009): 2592-2602.

[2]. Wielgat, R., et al. “On Using Prefiltration in HMM-Based Bird Species Recognition,” International Conference on Signals and Electronic Systems, ICSES'12, 18-21 September 2012, Wroclaw, Poland 2012 pp.1-5.

[20]. Phyu, Thair Nu. "Survey of classification techniques in data mining." In Proceedings of the International MultiConference of Engineers and Computer Scientists, vol. 1, pp. 18-20. 2009.

[21]. Wilson, Mark W., Tom Gittings, Josephine Pithon, Thomas C. Kelly, Sandra Irwin, and John O'Halloran. "Bird diversity of afforestation habitats in Ireland: current trends and likely impacts." In Biology and Environment: Proceedings of the Royal Irish Academy, pp. 55-68. Royal Irish Academy, 2012.

[22]. Possingham, Hugh P., Sandy J. Andelman, Mark A. Burgman, Rodrigo A. Medellín, Larry L. Master, and David A. Keith. "Limits to the use of threatened species lists." Trends in ecology & evolution 17, no. 11 (2002): 503-507.

[23]. Pellegrini, Thomas. "Densely connected CNNs for bird audio detection." 2017 25th European Signal Processing Conference (EUSIPCO). IEEE, 2017.

[24]. Papadopoulos, Timos, Stephen J. Roberts, and Katherine J. Willis. "Automated bird sound recognition in realistic settings." arXiv preprint arXiv:1809.01133 (2018).

[25]. Marini, Andreia, Alef J. Turatti, A. S. Britto, and Alessandro L. Koerich. "Visual and acoustic identification of bird species." In 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 2309-2313. IEEE, 2015.

[26]. Fagerlund, Seppo. "Bird species recognition using support vector machines." EURASIP Journal on Applied Signal Processing 2007, no. 1 (2007): 64-64.

[27]. Salamon, Justin, Juan Pablo Bello, Andrew Farnsworth, and Steve Kelling. "Fusing shallow and deep learning for bioacoustic bird species classification." In 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 141-145. IEEE, 2017.

[28]. Baowen Qiao ; Zuofeng Zhou ; Hongtao Yang ; Jianzhong Cao, “Bird species recognition based on SVM classifier and decision tree”, IEEE 2017

[29]. Iosif Mporas ; Todor Ganchev, “Automated Acoustic Classification of Bird Species from Real -Field Recordings”, IEEE 2012

[3]. Marini, Andréia, J. Facon, and A. L. Koerich. “Bird Species Classification Based on Color Features,” IEEE International Conference on Systems, Man, and Cybernetics, 2013, pp.4336-4341.

[30]. Sangita S.Londhe, “Bird Species Identification Using Support Vector Machine”, IJCSMC 2015

[31]. Wei-Ho Tsai ; Yeong-Yuh Xu ; Wei-Cheng Lin, “ Bird species identification based on timbre and pitch features ”, IEEE, 2013

[32]. Liang Zhang ; Michael Towsey ; Jinglan Zhang ; Paul Roe, “Computer-Assisted Sampling of Acoustic Data for More Efficient Determination of Bird Species Richness ”, IEEE 2015

[33]. Guillermo Sarasa ; Ana Granados ; Francisco B. Rodriguez, “An approach of algorithmic clustering based on string compression to identify bird songs species in xeno-canto database”, IEEE 2017

[34]. Min-Seok Kang ; Kwang-Seok Hong, “Automatic Bird-Species Recognition using the Deep Learning and Web Data Mining”, IEEE 2018

[35]. Martin Graciarena ; Michelle Delplanche ; Elizabeth Shriberg ; Andreas Stolcke, “Bird species recognition combining acoustic and sequence modelling”, IEEE 2011

[36]. Peter Jancovic ; Münevver Köküer, “Bird Species Recognition Using Unsupervised Modeling of Individual Vocalization Elements ” IEEE 2019

[37]. Kunja Bihari Swain, “Bio inspired Cuckoo Search Algorithm based neural network and its application to noise cancellation ”, IEEE 2014

[38]. Xuye Zhi and Chengan Guo, “ Bird Species Recognition Based on Deep Learning and Decision Fusion”, springer 2018

[39]. Satya Dharanipragada, “Bird Species Recognition Using Support VectorMachines”, EURASIP Journal on Advances in Signal Processing 2008

[4]. V.M.. Trifa, A.N.G. Kirschel, and C.E. Taylor, “Automated species recognition of antbirds in a Mexican rainforest using hidden Markov models,” J. Acoust. Soc. America, 123(4), pp. 2424-2431, April 2008.

[40]. NARANCHIMEG, “Cross-domain Deep Feature Combination for Bird Species Classification with Audio-visual Data”, ARXIV 2018

[41]. Purohit Shrinivasacharya, “An image crawler for content based image retrieval system”, IJRET 2013

[42]. Hari Babu Srivastava et al, “Image Pre-processing Algorithms for Detection of Small/Point Airborne Targets”, Defence Science Journal, Vol. 59, No. 2, March 2009

[43]. Grant Van Horn, “Building a bird recognition app and large scale dataset with citizen scientists:

[44]. The fine print in fine-grained dataset collection”, CVPR 2015

[45]. Andr´eia Marini, “Bird Species Classification Based on Color Features”, birding.com

[46]. Erika Vilches, “Data Mining Applied to Acoustic Bird Species Recognition”, IEEE 2005

[47]. Herve Goeau, “LifeCLEF Bird Identification Task 2016 The arrival of Deep learning”, sabiod 2017

[48]. H. Meutzner, N. Ma, R. Nickel, C. Schymura, and D. Kolossa, “Improving audio-visual speech recognition using deep neural networks with dynamic stream reliability estimates,” Acoustics, Speech and Signal Processing (ICASSP), 2017 IEEE International Conference on, pp.5320–5324, IEEE, 2017.

[49]. Torfi .A, S.M. Iranmanesh, N.M. Nasrabadi, and J. Dawson, “Coupled 3d convolutional neural networks for audio-visual recognition,” arXiv preprint, 2017.

[5]. P. Somervuo, A. Harma, and S. Fagerlund, “Parametric representations of bird sounds for automatic species recognition,” IEEE Transactions on Audio, Speech and Language Processing, vol. 14, no. 6, pp. 2252–2263, 2006.

[50]. S. Branson, G. Van Horn, S. Belongie, and P. Perona, “Bird species categorization using pose normalized deep convolutional nets,” arXiv preprint arXiv:1406.2952, 2014.

[51]. Atanbori, John, Wenting Duan, John Murray, Kofi Appiah, and Patrick Dickinson. "Automatic classification of flying bird species using computer vision techniques." Pattern Recognition Letters 81 (2016): 53-62.

[52]. Binitha, S., and S. Siva Sathya. "A survey of bio inspired optimization algorithms." International journal of soft computing and engineering 2, no. 2 (2012): 137-151.

[53]. Gerdes, John W., Satyandra K. Gupta, and Stephen A. Wilkerson. "A review of bird-inspired flapping wing miniature air vehicle designs." Journal of Mechanisms and Robotics 4, no. 2 (2012): 021003.

[54]. Duman, Ekrem, Mitat Uysal, and Ali Fuat Alkaya. "Migrating Birds Optimization: A new metaheuristic approach and its performance on quadratic assignment problem." Information Sciences 217 (2012): 65-77.

[55]. Wielgat, Robert, Tomasz Potempa, Pawel Świętojański, and Daniel Król. "On using prefiltration in HMM-based bird species recognition." In 2012 International Conference on Signals and Electronic Systems (ICSES), pp. 1-5. IEEE, 2012.

[56]. Jančovič, Peter, Münevver Köküer, and Martin Russell. "Bird species recognition from field recordings using HMM-based modelling of frequency tracks." In 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 8252-8256. IEEE, 2014.

[57]. Trifa, Vlad M., Alexander NG Kirschel, Charles E. Taylor, and Edgar E. Vallejo. "Automated species recognition of antbirds in a Mexican rainforest using hidden Markov models." The Journal of the Acoustical Society of America 123, no. 4 (2008): 2424-2431.

[58]. Kasten, Eric P., Philip K. McKinley, and Stuart H. Gage. "Ensemble extraction for classification and detection of bird species." Ecological Informatics 5, no. 3 (2010): 153-166.

[59]. Fagerlund, Seppo. "Bird species recognition using support vector machines." EURASIP Journal on Applied Signal Processing 2007, no. 1 (2007): 64-64.

[6]. Florent Bled, John Sauer, et al, “Modeling Trends from North American Breeding Bird Survey Data: A Spatially Explicit Approach” 2018

[60]. Briggs, Forrest, Balaji Lakshminarayanan, Lawrence Neal, Xiaoli Z. Fern, Raviv Raich, Sarah JK Hadley, Adam S. Hadley, and Matthew G. Betts. "Acoustic classification of multiple simultaneous bird species: A multi-instance multi-label approach." The Journal of the Acoustical Society of America 131, no. 6 (2012): 4640-4650.

[61]. Vilches, Erika, Ivan A. Escobar, Edgar E. Vallejo, and Charles E. Taylor. "Targeting input data for acoustic bird species recognition using data mining and hmms." In Seventh IEEE International Conference on Data Mining Workshops (ICDMW 2007), pp. 513-518. IEEE, 2007.

[62]. Vilches, Erika, Ivan A. Escobar, Edgar E. Vallejo, and Charles E. Taylor. "Data mining applied to acoustic bird species recognition." In 18th International Conference on Pattern Recognition (ICPR'06), vol. 3, pp. 400-403. IEEE, 2006.

[7]. Fraixedas, Sara, Andreas Lindén, and Aleksi Lehikoinen. "Population trends of common breeding forest birds in southern Finland are consistent with trends in forest management and climate change." Ornis Fennica 92, no. 4 (2015): 187.

[8]. R. Bardeli, D. Wolff, F. Kurth, M. Koch, K.-H. Tauchert, and K.-H. Frommolt, “Detecting bird songs in a complex acoustic environment and application to bioacoustic monitoring,” Pattern Recognition Letters, vol. 31, no. 12, pp. 1524–1534, 2010.

[9]. Marini .A, J. Facon, and A. L. Koerich, “Bird species classification based on color features,” in IEEE Int’l Conf Sys Man Cybernetics, Oct 2013, pp. 4336–4341.

PUBLICATION

K.Annalakshmi, Dr. I. Elizabeth Shanthi. **"Bird Species Recognition Techniques Based On Deep Learning Methods - A Survey" in Volume 8, Issue 3 (May 2019),** International Journal of Scientific Research in Computer Science Applications and Management Studies (ISSN 2319-1953).

APPENDIX

Sample Input Images