IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

Bird Species Image Identification using Deep Learning

Atharva kulkarni[1], Animesh Tade[2], Suyog Sulke[3], Sameer Shah[4], Dr. Aparna Pande[5]

[1,2,3,4]Student, [5]Mentor, Computer Science and Engineering Department
[1,2,3,4,5]Nutan College of Engineering and Research, Talegaon Dabhade, Pune, India

Abstract :

These days, many inexperienced bird watchers have trouble remembering and identifying all the various bird species. Additionally, in order to save and care for diverse bird species, the general public and newly employed rescue team members lack the ability to do so. They have to go through a difficult process to locate large publications like "Birds of the Indian Subcontinent." In this study, we evaluate a deep learning-based AI model that is good at identifying birds from photographs and provide the results. One of the top Deep Learning techniques, Transfer Learning, is used in the study's Simple Web App to recognise photographs. To become more familiar with Google's InceptionV3 model,1000 photos with annotations for each of the 325 different bird species in the dataset.The article presents empirical studies that evaluate different approaches and yield insightful results. It suggests multiple possibilities for future research in the areas of object recognition and neural network-based machine learning.

Keywords: ***Deep Learning, InceptionV3, Bird Identification, CNN, ImageNet***

I. INTRODUCTION

There is currently significant attention being given to the behavior and population patterns of birds. Birds react swiftly to environmental changes, which allows humans to immediately recognise other species in the habitat (like the insects they eat). Due to this, there is growing interest in the automatic identification, There have been multiple recent research studies focused on identifying bird species through analysis of their vocalizations.

However, gathering information about birds requires a lot of human labour and is an expensive procedure.Segmenting bird vocalisations into smaller recognition units can be done manually or automatically. In earlier studies, the number of species varied between 2 and 16. Therefore, Bird species identification is essential in order to identify the species of bird that a given image belongs to. Using an image, one may identify the category that a certain species of bird belongs to.

We anticipate that by offering a complete collection of benchmarks and annotation types for one specific area, Birds-200 will aid research on subordinate classification (birds). We want to develop a level of search depth that has only hitherto been available to a small number of niche categories, including faces and pedestrians. Research will remain more manageable from a logistical and computational standpoint if it focuses on birds. Nevertheless, we think that many of the lessons discovered (in terms of annota- tion procedures, localization models, feature representations, and learning algorithms) will be transferable to other domains, such as other kinds of animals, plants.

Instead than examining the challenge of recognising a big number of discrete categories, This study investigates the task of classifying a vast number of birds into a single category, which is a more intricate task compared to simply categorizing birds. The reason for this difficulty is that bird groups often have numerous resemblances, making it challenging to differentiate between them. Additionally, as birds are flexible objects that can deform in various ways, there exists a significant amount of variation within classes. Previous research on bird classification has either focused on a limited number of classes or utilized language-based approaches.

II. LITERATURE SURVEY

The literature survey section will provide an overview of the existing research on Bird Species Image Identification. We will examine numerous studies that used rating and review systems to identify different bird species using images, and we will examine the various approaches and techniques used.

In [1], the use is highlighted. The current study looked into a technique for classifying images utilising the Birds dataset and the Deep Learning algorithm (Unsupervised Learning). 11,788 images in 200 categories make up this collection. An easily accessible website is linked to a system that allows users to input photos for identification purposes, and it provides the desired results. The proposed system utilizes a method called part detection and CNN feature extraction from multiple convolutional layers. These extracted features are merged and then passed on to the classifier for categorization. The algorithm's results indicate an 80% accuracy rate in predicting the species of birds discovered.

Following data analysis, it was discovered that using just one parameter results in lower accuracy. However, the accuracy of the project will enhance if a combination method is utilised, which takes into account elements like attitude, wings, colour, beak, legs, etc.

The [2] paper offers a machine learning method for detecting visual objects that can analyse images very quickly and have high detection rates. Computer technology makes advantage of the object recognition approach. One of the hardest and most complicated computer tasks is thought to be this one. A model using new techniques that is not only quick but also dependable has been offered in addition to numerous methods that have been used in the past. Also contrasted with numerous other types is the Easynet model. Our model generates scores at the moment of prediction based on whether the object falls into a specific category. With the help of one network evaluation, it provides predictions.

Separating the object from the background is the focus of category detection, a component of object recognition. The classification of the object into one of the already established categories is the responsibility of category recognition. It is a technique for locating certain objects in digital photos or videos.

The suggested technique is thoroughly contrasted with the current feature detection algorithm and put to the test using BSDS and remotely sensed photos. The results of the tests and comparisons demonstrate that the suggested algorithm exhibits impressive performance. It offers a fresh concept for identifying visual features using the principle of phase congruency. Even though the phase congruency algorithm has been changed to reduce noise, there is still some lack of noise while detecting image features. Therefore, more effort is needed to enhance the algorithm and effectively control noise.

The [3] The article highlights the diverse roles that humans can play in the field of computer vision's visual categorization. As experts in a specific domain, they can provide a comprehensive set of semantic components and attributes that are essential for identifying and differentiating categories, and also provide actual attribute values, such as in a field guide. Even novice users of interactive classification systems can contribute in meaningful ways.

The study proposes a system for fine-grained visual categorization that is highly effective, adaptable, and scalable. This system utilizes various computer vision techniques and similarity measures to create a unified framework based on perceptual similarity. It outperforms other picture retrieval techniques that rely on relevance feedback and vocabulary-dependent attribute-based methods. Moreover, the system can incorporate multiple metrics and intelligently prompt users with a variety of questions during the categorization process.

The dataset was divided into training and testing sets, and only training images were used to create the embedding. For the testing phase, Amazon Mechanical Turk employees were employed to select all images that clearly differed from the reference image in terms of species. The experiment ensured that no two photos from the same category were sampled for each job at the category level without replacement.The study proposes a cost-effective and efficient method for interactive fine-grained categorization that does not require subject-matter experts to provide attribute vocabulary. The system can be expanded by adding more training data and updating the perceptual similarity measure as users respond to similarity queries for new images, allowing the system to iteratively improve with additional data. In the future, these perceptual embeddings could be used to infer characteristics, components, taxonomies, etc. that would be useful for users.

In [4] Deep hierarchical neural networks can be pre-trained in unsupervised settings to enhance their ability to categorize patterns in supervised tasks. This is achieved using online back-propagation for weight adjustments after every error back-propagation step. Our research has demonstrated that such properly trained massive and deep DNNs can surpass all existing techniques in benchmarks for computer vision, including , Latin letters, Chinese characters, traffic signs, (jittered, cluttered), and . Additionally, we have shown that unsupervised initialization and pre-training are not always beneficial, especially for large datasets. Our research has also revealed that integrating multiple DNN columns in a Multi-column DNN (MCDNN) can further reduce the error rate by 30–40%.

Our MCDNN outperforms the current standard by 20–80% on numerous additional image classification datasets. Our fully supervised approach does not require any extra unlabeled data sources. While a single DNN can already achieve cutting-edge results, the merging of multiple DNNs to create MCDNNs results in even greater performance improvements. These findings represent the first instance of benchmarks for computer vision that are human-competitive.

In [5] This study proposes a novel approach to classify bird species based on the color information obtained from uncontrolled photographs. The method accounts for variations in the birds' poses, sizes, and viewpoints, as well as the diverse lighting conditions and potential occlusions by other scene elements. To identify the possible regions in the image where the bird is present, the method first employs a color segmentation algorithm to remove the background and foreground elements. Then, the image is decomposed into component planes, and from each plane, candidate regions are selected to compute normalized color histograms. Post-processing techniques are applied to reduce the histogram's intervals to a fixed number of bins. The classification algorithm uses these histograms as feature vectors to distinguish between different bird species. The experimental results on the dataset show a 75% correct segmentation rate and a classification rate ranging from 90% to 8%, depending on the number of bird species categories included.

The research focuses on the automated identification of bird species from bird photographs, based on more than 6,000 images of over 200 bird species. Two independent classification methods are used, where the first combines the feature vectors generated from the decomposed image planes and feeds them to a single classifier. The second method employs different classifiers for each feature vector. The study also investigates the scalability of the proposed approach by considering different color spaces.The study evaluates the effectiveness and impact of color-based segmentation on picture segmentation through several tests. The results show that segmentation has a significant effect on the classification outcomes, but only for colored features. Although the color segmentation successfully segments over 70% of the pixels color spaces, its impact on the classification of bird species is limited, ranging from 8% to 0.42%, especially when the number of classes is large.

In [6] the demanding dataset CUB-200 [7] of 200 bird species is enlarged as 2011. The enhanced edition features a significant increase in the number of photos per category, with almost twice as many images included. Additionally, this edition includes new annotations for part localization, which provide detailed information on bounding boxes, part positions, and attribute labels for each image.

Many Mechanical Turk users filtered the images and remarks. For multi-class classification and part localization, we present benchmarks and foundational experiments.

We anticipate that by offering a complete collection of benchmarks and annotation types for one specific area, Birds-200 will aid research on subordinate classification (birds). We want to develop a level of research depth that has so far only been applied to a small number of niche categories, like faces and pedestrians. Research will remain more manageable from a logistical and computational standpoint if it focuses on birds. Nevertheless, we anticipate that many of the lessons discovered (in terms of annotation techniques, localization models, feature representations, and learning algorithms) will apply to other domains, such as various species of animals, plants, or objects.

In [7] Object detection refers to the process of locating items in an image and categorizing them. This involves three main stages: identifying relevant regions, extracting meaningful features, and classifying objects using deep learning models. Deep learning-based approaches have gained significant attention in recent years due to their powerful learning capabilities and ability to handle challenges like occlusion, scale transformations, and background variations.This study provides a comprehensive analysis of deep learning-based object detection frameworks that address various sub-problems such as occlusion, clutter, and low resolution, with different levels of modifications to the popular R-CNN architecture.

The review starts by examining generic object detection pipelines, which form the basis for other related tasks. It then discusses three common applications of object detection: pedestrian detection, face detection, and object recognition.To fully grasp the object detection landscape, the study proposes several exciting future research directions. In addition, the review highlights the advances in neural networks and related learning systems that provide valuable insights and guidance for future growth. Object detection is a fundamental problem in computer vision that has many applications, such as image classification, and can provide useful information for the semantic understanding of images and videos.

In [8] We have developed a system to identify animals and backgrounds in images using advanced techniques. Our system uses a Deep Convolutional Neural Network (DCNN) to learn features and then applies support vector machine, k-nearest neighbour, and ensemble tree algorithms to classify the features. Our evaluation of the system on a commonly used camera-trap dataset shows that it achieves 90% accuracy.

Our approach is especially useful for images with many animals in them, which we often encounter when using camera-trap networks. To identify potential animal regions in these images, we use a multilevel graph cut method. Then, to verify whether a proposed region is actually an animal or not, we classify it as either background or animal using our DCNN-based system.

Our experiments show that our approach is effective for detecting wild animals, both during the day and at night. By incorporating DCNN features into the machine learning algorithms, we were able to achieve improved performance compared to previous methods. Overall, our system offers a reliable and efficient tool for animal detection in challenging camera-trap images.

In [9] The classification of bird species through support vector machine techniques was tested using two datasets from previous stages of the project to evaluate the new methods. The results indicated that the performance was either on par or better than the reference methodologies. However, it was difficult to compare the recognition results directly between the two datasets as dataset 2 had more species and a wider range of sounds. The species in dataset 1 were also more closely related to each other than those in dataset 2. The decision tree topology in the proposed method was independent of species ordering, and the same outcome could be achieved by rearranging the positions of species within the tree. This topology was efficient and did not require additional information about species relationships. However, a hierarchical topology based on sound similarities between species could lead to a more reliable and computationally efficient classifier. All syllables were represented using the same parameters, but the decision tree structure allowed for feature weighting in each subproblem. For example, using MFCC-models resulted in 100% accuracy for Pygmy Owl in dataset 2 without feature weighting, but the technique yielded lower recognition results (91%) in the mixed model compared to MFCC-models. Future research will explore the use of feature weighting, which could have provided 100% accuracy in the mixed model.

In [10] The aim of PlantNet, a highly interactive platform and information system, is to generate botanical data by utilizing image-based plant identification techniques. The system comprises three primary interfaces: a web-based interface, an iOS app (which is currently undergoing redevelopment), and an Android app (which is the most advanced and widely used interface). Using any of these interfaces, users can upload one or more images of a plant and receive a list of the most probable species that the plant belongs to. The graph in Figure 2 of the appendix illustrates the growing popularity of the app, particularly during the spring and summer seasons.

The architecture of the Plant Net information system is described in this publication for the first time since convolutional neural networks were first deployed in June 2015. It demonstrates how a shift from manually created image representations to the usage of deep learning in an information system can have an ongoing impact on society. Plant Net is widely utilised in a variety of professional contexts today (agriculture, education, ecotourism, etc.). It helps to foster new types of educational instruction in botany and ecology as well as the growing interest of a significant portion of society in their surroundings.

III. BACKGROUND

Generally speaking, birds can be recognised by sight or sound. The most important visual aspects include the bird's size, form, posture, colour, and wings, among others. Because a bird's wings increase as it ages, the season must also be considered while assessing the criterion.The acoustics components include the melodies and whistles that birds make.The characteristics that set one bird apart from another, such as breast spots, wing bars thin lines that run along the wings eye rings, crowns, and eyebrows are also helpful. The beak's shape is frequently significant because it helps identify a bird. The most common ways to identify birds are by their physical characteristics, like shape and posture. Recognizing a bird's silhouette is a skill possessed by most experts due to the difficulty in altering this characteristic. The tail of a bird can also serve as a distinguishing feature, with variations such as notched, elongated and pointed, or rounded. Additionally, the legs of a bird can be utilized to aid in identifying its photo.

IV. METHODOLOGY

Certain approaches are explored and tested in the development of a generic system. In this paper, we will outline the tried approaches and explain how we arrived at the most effective and precise results. The following are tried methods:

- **Deep learning:**

Deep learning is a category of machine learning techniques that employ artificial neural networks to perform representation learning. These neural networks can be trained using supervised, semi-supervised, or unsupervised learning. Convolutional neural networks are among the most widely used types of artificial neural networks in modern deep learning models

- **CNN:**

A common artificial neural network for object and image recognition and classification is the convolutional neural network, or CNN. For optimal accuracy, a variety of alignments and attributes, including the bird's head, body, colour, beak, form, and overall appearance, are taken into account.

- **Transfer Learning:**

Transfer learning generally refers to the practice of applying knowledge gained from solving one problem to another related problem. In the context of deep learning, it involves training a neural network on a similar problem, and then leveraging the learned features or layers to build a new model for a different but related problem..

V. ANALYSIS

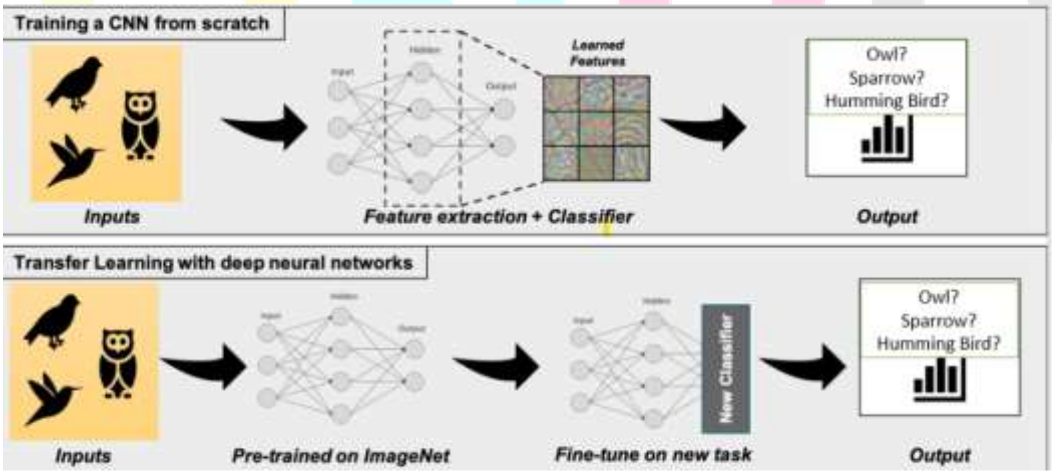

Fig. 1. Transfer Learning using neural network

Parameters			
Classifier:	**Individual Function Extractor:**	**Integrated Feature Extractor:**	**Initial Weighting:**
A pre-trained model is used directly to classify newly captured photographs into different categories.	The pre-existing model is utilized for performing image pre-processing and identifying important features, either by utilizing the entire model or by selecting a subset of it.	The process involves combining a pre-existing model, or a portion of it, with an integrated feature extractor to create a new model. However, the layers of the pre-existing model remain fixed and do not get updated during the training process.	The pre-existing layers of a trained model are utilized and fine-tuned in combination with a new model. This involves integrating either the entire pre-trained model or a subset of its layers into the new model architecture.

Bird Species Detection Method	Identified Parameters
Visual observation	Plumage color, size, shape, beak morphology
Bird call analysis	Vocalizations, song patterns
Nest monitoring	Nest structure, location
Bird banding	Individual identification through leg bands
DNA analysis	Genetic markers
Radio telemetry	Location, movement patterns
Radar	Flight patterns, presence
Thermal imaging	Heat signatures
Camera traps	Visual identification, behavior
Citizen science platforms	Sightings, photographs, audio recordings

VI. CONCLUSION

This study aims to explore the application of modern deep learning techniques in the recognition and differentiation of various bird species. Transfer Learning outperformed traditional Deep Learning Techniques like RNN, CNN, and many others to produce the greatest results. The model was trained using the Kaggle open source image dataset, which comprises of 325 species and approximately 600 images per species. Additionally, InceptionV3, one of the best model architectures, was employed to obtain the needed accuracy. Future studies will focus on creating the same technology for usage on smartphones.

REFERENCES

[1] Prof. Pralhad Gavali, Ms. Prachi Abhijeet Mhetre Bird Species Identification using Deep Learning International Journal of Engineering Research & Technology (IJERT) ISSN: 2278- 0181 Vol. 8 Issue 04, April-2019

[2] TanuDhiman Mtech (Cse) ,Object Detection and Recognition in Image Processing ,International Journal of Management, IT & Engineering Vol. 8 Issue 11, November 2018

[3] Similarity Comparisons for Interactive Fine-Grained Categorization, Catherine Wah, Grant Van Horn, Steve Branson, 13 Feb 2012

[4] Cirean, D., Meier, U. and Schmidhuber, J., 2012. Multicolumn deep neural networks image classification. arXiv preprint arXiv:1202.2745.

[5] Andreia Marini, Jacques Facon and Alessandro L. Koerich Postgraduate Program in Computer Science (PPGIa) Pontifical Catholic University of Paran´a (PUCPR) Curitiba PR, Brazil 80215–901 Bird Species

[6] The Caltech-UCSD Birds-200-2011 Dataset, Catherine Wah, Steve Branson,Peter Welinder, Pietro Perona, Serge Belongie

[7] Object Detection with Deep Learning: A Review Zhong-Qiu Zhao, Member, IEEE, Peng Zheng, Shou-tao Xu, and Xindong Wu, Fellow, IEEE

[8] Wild Animal Detection Using Deep Convolutional Neural Network Gyanendra K. Verma and Pragya Gupta on 04 June 2018.

[9] Bird Species Recognition Using Support Vector Machines, Satya Dharanipragada, Hindawi Publishing Corporation EURASIP Journal on Advances in Signa
Processing Volume 2007, Article ID 38637, 8 pages doi:10.1155/2007/38637

[10] PlantNet project, AMAP joint research unit, France , Antoine Affouard, Herve Goeau, Pierre Bonnet, June 2015