Identification of Bird Species Using Deep Learning

T. Sunitha [1], Dr.T. Nedunchezhian [2], Dr.G. Vijaya Kumar [3], B. Neelima [4], Mrs.B.Malleswari [5]

[1, 2, 3, 4]Department of Computer Science and Engineering,

[5]Electronics and Communications Engineering,

[1, 2, 3, 4, 5] QIS College of Engineering and Technology, Ongole, Andhra Pradesh, India

thella.sunitha@qiscet.edu.in[1], nedunchezhian.t@qiscet.edu.in [2], vijayakumar.g@qiscet.edu.in[3], neelima.b@qiscet.edu.in[4], malleswari.b@qiscet.edu.in[5]

Corresponding Author Mail: qispublications@qiscet.edu.in

Article Info
Page Number: 182-193
Publication Issue:
Vol 69 No. 1 (2020)

Abstract— Many avian species are getting harder to find, and even when they are discovered, their categorization could be difficult to predict. Birds, when seen from a distance, may appear in a dazzling array of sizes, shapes, colours, and orientations. Compared to the auditory classification, the visual depiction of bird breed variety is far more extensive. The ability to tell birds apart is greatly enhanced by visual aids such as photographs. The Caltech-UCSD Birds 200 dataset serves as the basis for this method's training and validation data. In order to facilitate comparison, we employ deep convolutional neural networks (DCNN) to convert a picture to a grayscale representation and the tensor flow to build a complex autograph consisting of many nodes. As a result of analysing the various access points to the validation data, a ranking table is constructed. Perhaps it can guess the necessary flock by scanning the scoreboard and picking the bird with the greatest rating. By analysing the dataset (CUB-200-2011), we find that the algorithm achieves an accuracy of 89% for identifying bird species. In the research, Linux and the Tensorflow framework were employed.

Article Received: 10 August 2020
Revised: 15 September 2020
Accepted: 20 October 2020
Publication: 25 November 2020

Tags: Autograph, Caltech, UCSD, Grayscale, Pixels, Tensorflow

INTRODUCTION

Bird behaviour and population dynamics research has been trending upward in recent years [1]. Birds, with their lightning-fast reactions to changing weather conditions, are invaluable for environmental monitoring and species identification (such as the bugs they consume). Hand collecting data on birds is a far more time-consuming and expensive option. Scientists, government entities, and the general public need a reliable tool for analysing massive amounts of bird data. As a result, being able to identify the species of bird shown in a photograph is essential. Therefore, it is crucial to be able to correctly identify different types of birds. Many different categories may be applied to a picture of a bird to determine what kind of bird it is. People may be recognised via their audio, visual, or aural characteristics. The vocalisations of a bird may be recorded and analysed using aural computational methods. However, the numerous noises in the environment, such as bugs, real-world items, and so on, confound the interpretation of such data. More than words or moving pictures, pictures make an influence on most individuals. Therefore, visual identification of birds has surpassed the use of voice and video. Humans and computers both have a hard time sorting birds into their proper species. Taxonomy of bird species has been of interest to bird watchers for centuries, if not millennia. For a whole picture of species, scientists must examine not just their abundance in the environment, but also their ecological niche, migratory patterns, and overall ecological effect. Linnaeus' system of grouping organisms into larger taxonomic units (kingdoms, phyla, classes, orders, families, and species) provides the basis for the categorization of birds used by ornithologists today [4]. Better image labelling algorithms have resulted in artefacts being moved to databases like Caltech-UCSD, which include many more defining criteria. Lots of ground has been broken in this area by recent research. The Caltech and UCSD Birds 200 (CUB-200-2011) is a well-known database for Photos of birds, categorised into 200 groups [5]. The majority of the birds in this collection are native to North America. There are 11,788 pictures in the Caltech-UCSD Birds 200 database with information including Component Positions (15), Numeric Characteristics (312), and a Reference Image (1). In this study, we focus on the challenge of identifying a large number of subclasses inside a single class, namely birds, rather than a wide range of diverse classes. Birds already provide a challenge when trying to classify them, but the great degree of similarity amongst groups just makes things more complicated. Since birds are not rigid things, they may be distorted in a variety of ways, therefore there is also a large range of variation within categories. Previous research on classifying birds has focused on either a small subset of categories or on sound.

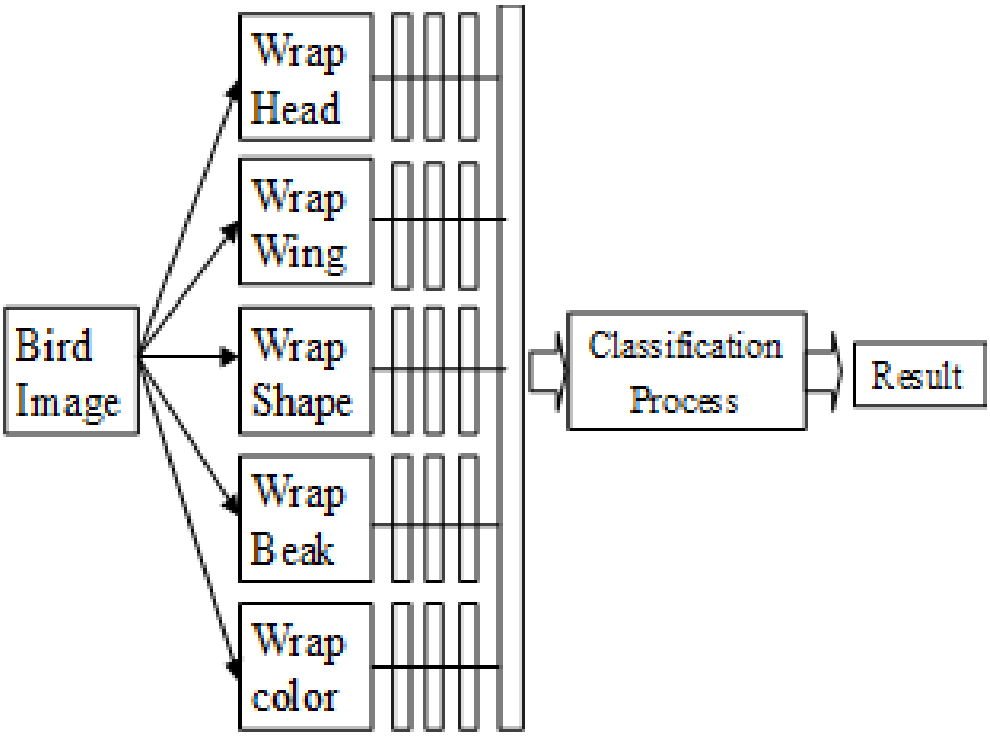

Fig. 1. Classification Process

In Fig. 1 we see the steps required to locate the bird in the image. After the photo has been submitted, it is analysed in a number of ways, including by skull and thorax type, colour, plumage, and overall appearance. Moreover, a deep convolutional neural network (DCNN) transports each configuration to collect features across several layers of the underlying network [6]. After that, we'll think about what the image really shows. The bird species will then be obtained together with the classifying result (i.e. features are aggregated to communicate to the classification model). The rules for this piece of writing are as follows: Considerations for a correct physical identification are highlighted in Section II. In Sections III and IV, we detail the procedures used when formulating the established methodology. In Section V, we get a comprehensive illustration of the flowchart.

RELATED WORKS

Both sight and sound are used extensively in the process of bird identification. The appearance of the bird, including its body, feathers, shape, attitude, colour, and so on, is very important [7]. However, considering the time of year is important since birds' feathers change as they mature, therefore this must be taken into account while assessing the criterion. The vocalisations of several birds are composed of vibrations in the air factors [8]. Distinguishing characteristics of birds include shoulder patches, wing bands (often shown by tiny lines upfield), eyebrow bands, crowns, and foreheads [9]. Identification of birds is often greatly aided by studying their beaks. Some of the most common ways to describe birds are by referring to their physical characteristics, such as their shape and posture. Due of the difficulty in changing this attribute, many experts are able to identify a bird just by looking at it. It is also possible to recognise a bird by its wing. Hooked, straight and curved, or curved heads are all distinguishable types. Recognizing an image in a shorter or longer format often requires the use of one's feet [10]. A desirable result cannot be achieved by using a single unsuitable element. Therefore, several factors need to be considered in order to get accurate outcomes. Variables such as image quality, distance between the birds and the recording equipment, and camera optical length all affect how big or small an individual bird appears in a picture [11]. Pictures are judged unfavourably because of their colour, which is made up of many different elements, as shown by a thorough examination of a large database of images [12]. It has been shown that the greater the quality of the image, the more trustworthy it is [13], [14]. The study offers a series of comparisons carried out in a CUB- 200 database consisting of more than 6,000 pictures categorised into 200 different categories [15]. In this study, they examined the usefulness of two colour systems, RGB and HSV, in classifying a wide range of animals. The quality of the final product varied between 8.82% and 0.43% if the model included more than 70% of the photos [16].

PROPOSED METHODOLOGY

The platform was created using a number of different methods. Information gathered using a database (the Berkeley Birds 200), a Deep Convolutional Neural Network (DCNN), an Unsupervised Learning Algorithm (ULA), and so on. In this study, a ULA was used to build the apparatus because to the lack of information on the definition of the input pictures. Moreover, the input data to the ULA is unlabeled; that is, just the input parameters (X) are supplied, with no corresponding result parameters. Interesting data trends are uncovered by ULA's systems. In-depth, we make use of grouping to organise data into distinct categories [4]. Large perceptrons were found with the use of DL techniques. When a photo is sent through a network of NN layers, the DL algorithms get a deeper understanding of the image.

For categorization purposes, we use NN. Information retrieval NN layers are shown in Fig. 2. Many machine learning (ML) techniques build on top of the NN. To construct NN, a weighted sum (W) and a biassed matrix are used (B). When tasked with image recognition, CNNs are often used by deep neural networks (DL). It's equipped with an activation function, a receptive field, and a number of secret nodes. The formation of synapses occurs in layers, with each successive layer making contact with the synapses in the layers below it. In this case, the resultant level serves as a predictor of the output. Sending an image to the fully connected layer causes it to provide a number of unique translations [2]. The convolution method will transfer one 3d point cloud to the next since the input image may have several inputs, such as colour, feathers, eyes, and bird beaks. Dimensions such as length, height, and thickness are all considered when dealing with 3D objects.

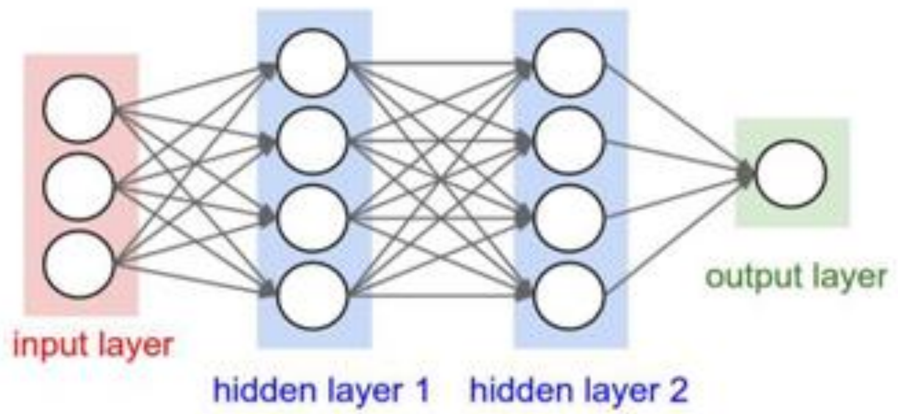

Fig. 2. NN Layers

CNN has two distinct divisions. During the extraction phase, the system learns new traits via a chain of coevolutionary events. In the segmentation process, as illustrated in Fig. 3, the extracted features are fed into a set of convolutional layers that operate as a predictor.

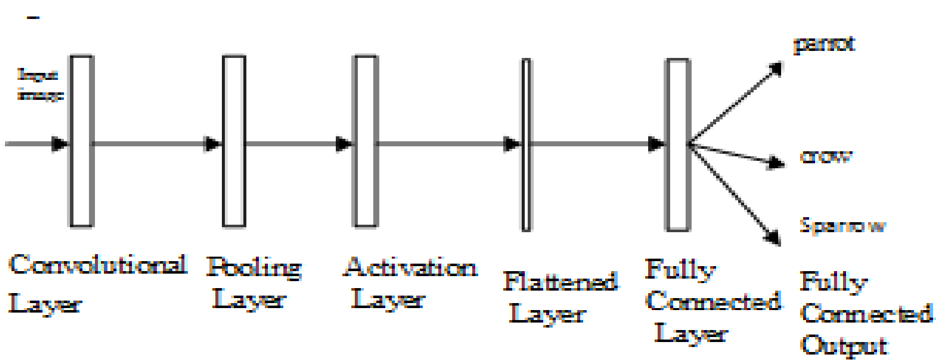

Fig. 3. CNN Layers

ISSN: 2094-0343
2326-9865

Four layers make up CNN: the hidden state, dense layer, activator overlay, and fully connected layer. In order to extract minute quantities of visual characteristics, the convolution layer is used. By employing pooling, the number of synapses in a convolutional network's previous layer may be reduced without losing any useful information. At the enabled setting, a value is calculated using an algorithm that aggregates disparate factors into a spectrum. Each and every synaptic connection in one layer is connected to each and every one in the next in a completely interconnected layer. CNN provides more trustworthiness since it accurately identifies each synapse. Two common methods for classifying images using ML are: Colors in the RGB spectrum and a grayscale It is common practise to grayscale all data. In the grey scale approach, the computer will assign values to each picture based on the image's numerical value. Images are analysed by placing each pixel into a matrix, which is then used by the computer to make decisions on what to classify. Google's Tensorflow is an open-source software platform for machine learning. This "network" of synapses is under the engineers' control, and its parameters may be tweaked to achieve the desired level of efficiency. There are several image classification packages available in Tensorflow [3]. An autograph is built using Tensorflow, and it consists of a series of compute nodes. Each computing node in the network represents both a statistical function and a connection or boundary between other nodes. Engineers can perform things thanks to ML:

Calculating using the Python software package. When information is collected and organized in one place, we call it a database. For tasks involving birds, we turn to the Berkeley Birds 200 (CUB-200-2011) dataset. Similar to the CUB-200 database, but with almost twice as many images for each categorization and extra component position labels for enhanced accuracy [8]. The following is a detailed explanation of the data set: Classes range from 200 to infinity. Altogether, there are 11,788 images. 15 Locations of Individual Parts, 312 Quantitative Features, and 1

Fig. 4. Images in Dataset

Fig. 5 depicts the actual workings of the proposed method. Creating a database that can be used to label images is a necessary step in developing such software. A training database splits up the study and exam phases. Improving recognition efficiency requires retraining the database using Google Collab's retrain.py. For this reason, the database has over 50,000 movements, since the greater the number of phases, the more exact the motion. The accuracy of the trained model is over 93%. The experimental database comprises over a thousand photos and a reliability of 80%. The database has been rigorously tested, and its reliability has been established at 75%. When users submit pictures, the pictures are briefly kept in databases. These data files are processed by an algorithm, sent to CNN, and then compared to the machine's classification findings. Each layer in a convolutional neural network (CNN) contributes to the overall network. The maximum possible categorization accuracy is achieved by taking into account a number of interrelated features, including the bird's head, chest, plumage, mouth, form, and overall appearance. DCNN uses information from several network levels to provide an output for each possible arrangement. Afterwards, deep learning (DL) and convolutional neural networks (CNN) are utilised for unsupervised picture classification. Consequently, the picture is sorted according to the amount of greyscale present in each pixel. The classifier then makes use of these characteristics when assigning labels to records. The feed will be compared to the examples used for training in order to generate potential outcomes. After the data has been sorted, a signature is generated, which is made up of relationships that form a structure. On the basis of this approach, a score card is constructed, and output is generated with the aid of the evaluation form.

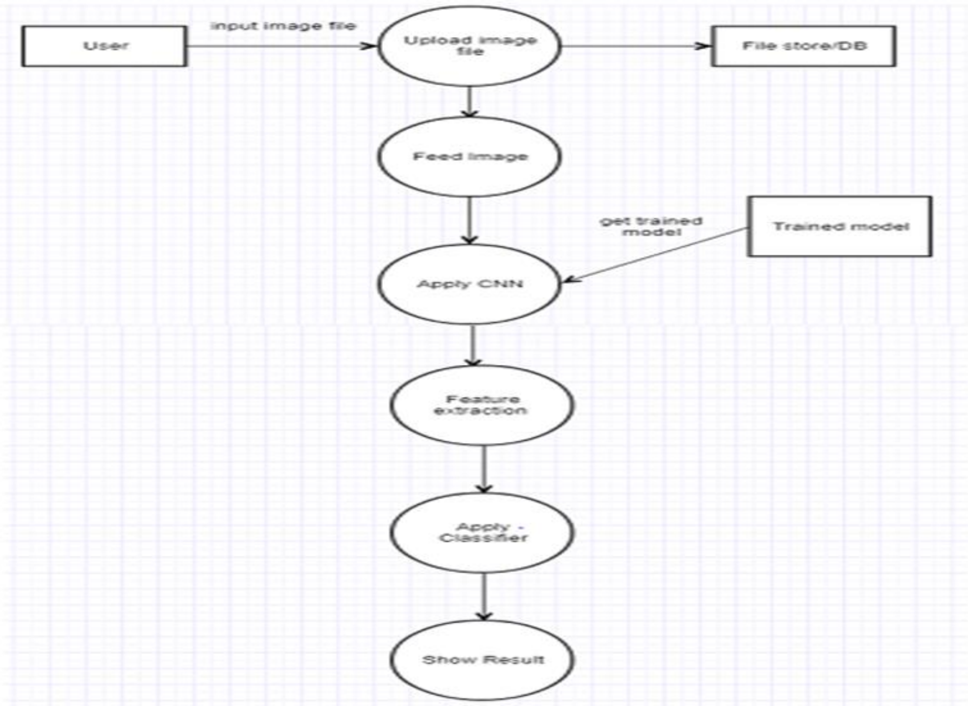

Fig. 5. Methodology employed

IV. RESULTS AND DISCUSSION

IV.Results and Discussion

The details of the suggested approach, as seen in Fig. 5, are illustrated there. Building an image-labeling database is an essential part of creating this kind of programme. Exam and study periods are separated by a training database. Retraining the database using Google Collab's retrain.py is necessary to increase recognition efficiency. Given that the bigger the number of phases, the more precise the motion, the database has approximately 50,000 motions. The trained model has an accuracy of 93%+. More than a thousand images with an accuracy of 80% are included in the experimental database. There have been many tests, and the database has been shown to be 75% reliable. When users upload photos, the images are temporarily stored in archives. An algorithm reads these data files, transfers them to CNN, and then checks CNN's classification results. A convolutional neural network (CNN) is built layer by layer, and each layer helps the network as a whole. The bird's head, chest, plumage, mouth, shape, and general look are all taken into consideration to provide the highest level of classification accuracy. Differential evolution neural networks (DCNN) use data from several network stages to provide a single result for every conceivable configuration. After that, unsupervised image categorization is performed using deep learning (DL) and convolutional neural networks (CNN). Thus, the image is ranked by the relative quantity of grey in each pixel. The classifier uses these features while labelling records. Potential results will be generated by comparing the feed to the examples used during training. A signature, made up of interconnected parts, is produced once the data has been organised. This method is used to build a report card and create results by use of an assessment form.

Fig. 6. Image input

TABLE I. EVALUATION APPROACH

S.No	Species	Score generated
1	Elegant tern	0.00921
2	Red faced cormorant	0.00929
3	Brant cormorant	0.0082
4	Pelagic cormorant	0.0085
5	White pelican	0.00807

TABLE II. COMPARATIVE ANALYSIS

S. No	Model	Accuracy
1	Pose Norm	82%
2	Part-based R-CNN	78.2%
3	Multiple granularity CNN	83%
4	Diversified visual attention network (DVAN)	80%
5	The deep LAC localization, alignment, and classification	82.7%
6	Proposed Method	89%

The results for each platform are shown in Table I and Table II, respectively. This analysis revealed that the top-scoring species was predicted to be an essential one. Figures 7 and 8 show the results that lead to this conclusion.

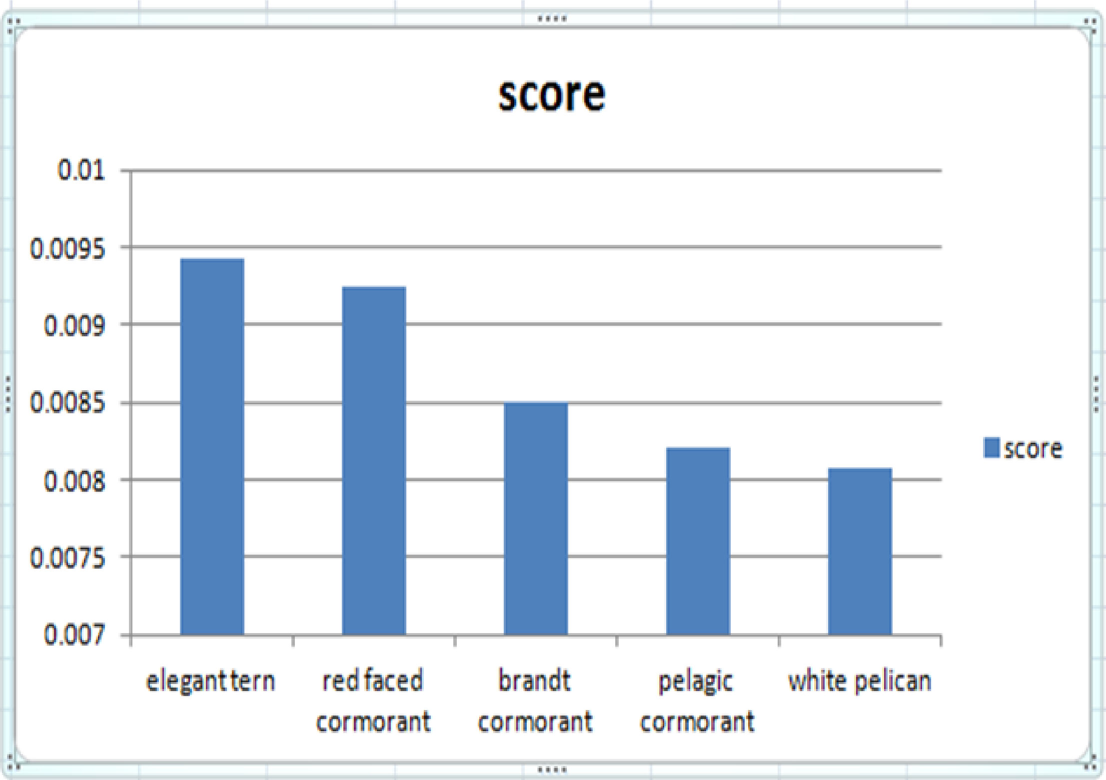

Fig. 7. Graphical representation of scores generated

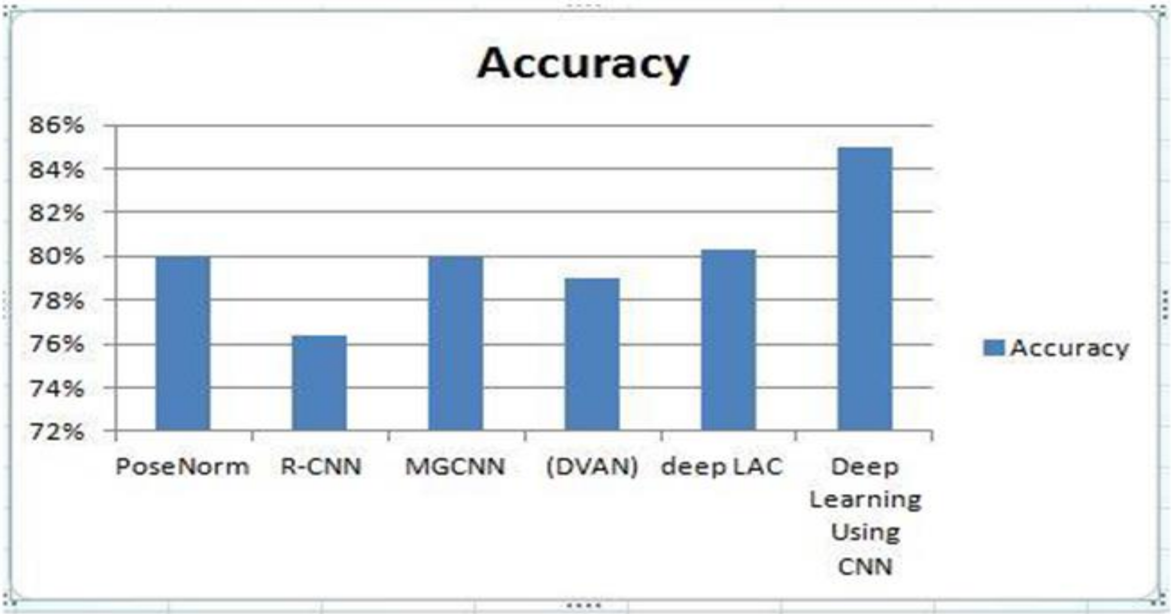

Fig. 8. Comparative analysis graph

It turns out that while working with a predetermined input, accuracy suffers. Combination techniques, on the other hand, increase the design's accuracy by considering many factors simultaneously, such as the bird's location, feathers, colour, head, limbs, and so on.

V. CONCLUSION AND FUTURE SCOPE

Caltech-UCSD Birds 200 photos were classified using a DL (Unsupervised Learning) technique to determine which bird species they belonged to. In all, there are 200 chapters with 11,788 photographs. The module is linked to a straightforward website where users may submit images for verification, and it produces accurate results. The proposed method prioritises the identification of a subset and the extraction of CNN properties from many fully connected layers. After collecting these details, the algorithm is fed the information in order to classify it. The system accurately identified the existence of 80% of bird communities using just 9% of the available data. Consider making an app for Android and iOS instead of a website. The convenience for the user is enhanced greatly. The internet may be used for this, since it could house massive quantities of data for comparison and significantly increase the available computational power (in the case of NN).

REFERENCES

1. Toth, B.P. and Czeba, B., September 2016. Large-Scale Bird Song Classification Using Convolutional Neural Networks. Working Notes formatted as CLEF (pp. 560-568).
2. Fagerlund, S., 2007. Support vector machines for avian species identification. Pages 64-64, 2007 Volume 1 Issue 1 EURASIP Journal on Applied Signal Processing.
3. Pradelle B., Meister B., Baskaran M., Springer J., and Lethin R. 2017. November. TensorFlow Computation Graph Polyhedral Optimisation. When it comes to high-performance computing, networking, storage, and analysis, the international community

gathers annually for the International Conference on High Performance Computing, Networking, Storage, and Analysis, where the 6th Workshop on Extreme-scale Programming Tools (ESPT- (SC17).

4. Cireşan, D., U. Meier, and J. Schmidhuber. Classifying images using deep neural networks that span many columns. A preprint version of this article may be found at arXiv:1202.2745.
5. Postgraduate Program in Computer Science (PPGIa) at Pontifical Catholic University of Paran'a (PUCPR), Curitiba PR, Brazil 80215-901, Andr'eia Marini, Jacques Facon, and Alessandro L. Koerich. The Use of Plumage Coloration for the Identification of Bird Species
6. Deep Learning for Image Recognition Faculty of Electronics, Telecommunications, and Information Technology at the Bucharest Polytechnic University; ANDREIPETRU BARAR; VICTOR-EMIL NEAGOE; NICU SEBE.
7. Deep Learning using Depth-wise Separated Convolutions, Xception Chollet, François Web search giant Google.com.
8. Zagoruyko and Komodakis (2016) Better convolutional neural network performance by attention transfer: a closer look. Reference: arXiv preprint arXiv:1612.03928.
9. Alexander A. Alemi, Christian Szegedy, Sergey Ioffe, and Vincent Vanhoucke
10. Stefan Kahl, Thomas Wilhelm-Stein, Hussein Hussein, Holger Klinck, Danny Kowerko, Marc Ritter, and Maximilian Eibl. Extensive Use of Convolutional Neural Networks for Bird Noise Categorization
11. T. Berg, J. Liu, S. Lee, M. L. Alexander, D. W. Jacobs, P. N. Belhumeur, and S. E. Alexander Large-Scale Birdsnap: Fine-grainedVisualCategorizationofBirds
12. Bo Zhao, Jiashi Feng Xiao Wu Shuicheng Yan. The Application of Deep Learning to Fine-grained Object Classification and Semantic Segmentation
13. Department of Electrical Engineering, ETH Zurich; Department of Engineering Science, University of Oxford; Andrew Zisserman, University of California, Berkeley College of Engineering and Physical Science, University of Oxford This paper presents BiCoS, a bi-segmentation approach to image classification.
14. International Conference on Advanced Computing Technologies and Applications (ICACTA), 2022, pp. 1–6, doi: 10.1109/ICACTA54488.2022.9753392, S. J. Basha, S. R. Madala, K. Vivek, E. S. Kumar, and T. Ammannamma
15. "Automatic Identification of Bird Species using Audio/Video Processing," 2022 International Conference for Advancement in Technology (ICONAT), 2022, pages 1–6, doi: 10.1109/ICONAT53423.2022.9725906;
16. "Automated Bird Species Identification using Audio Signal Processing and Neural Network," 2022 International Conference on Electronics and Renewable Energy, 2022, pages 1–6;
17. Qiao, Baowen, Zuofeng Zhou, Hongtao Yang, and Jianzhong Cao. 2017. "Bird Species Recognition Based on SVM Classifier and Decision Tree." In 2017 First International Conference on Electronics Instrumentation & Information Systems 1–4.
18. Sarayrah, Bayan mahmoud. 2019. "Finger Knuckle Print Recognition Using Deep Learning." Mutah University.

19. S. Al-Showarah et. al. 2020. “The Effect of Age and Screen Sizes on the Usability of Smartphones Based on Handwriting of English Words on the Touchscreen”, Mu’tah Lil-Buhuth wad-Dirasat, Natural and Applied Sciences series, Vol. 35, No. 1, 2020. ISSN: 1022-6812.
20. Triveni, G., Malleswari, G. N., Sree, K. N. S., & Ramya, M (2020). Bird Species Identification using Deep Fuzzy Neural Network Int. J. Res. Appl. Sci. Eng. Technol.(IJRASET), 8: 1214-1219.
21. C. Wah, S. Branson, P. Welinder, P. Perona, and S. Belongie 2011. “The Caltech-UCSD Birds-200-2011 Dataset.” Technical Report CNS-TR2011-001, California Institute of Technology.
22. Welinder, Peter et al. 2010. “Caltech-UCSD Birds 200.” Technical ReportCNS-TR-2010-001, California Institute of Technology, 2010.
23. Brownlee, Jason. 2016. “How To Use Classification Machine Learning Algorithms in Weka.” Retrieved from https://machinelearningmastery. com/use-classification-machine-learning-algorithms-weka/.
24. Cai, J., Ee, D., Pham, B., Roe, P., & Zhang, J. (2007, December). Sensor network for the monitoring of ecosystem: Bird species recognition. In 2007 3rd international conference on intelligent sensors, sensor networks and information 293-298. IEEE
25. Kumar, A., & Das, S. D. 2018. “Bird Species Classification Using Transfer Learning with Multistage Training”. In Workshop on Computer Vision Applications 28-38. Springer, Singapore.
26. P Ramprakash, M Sakthivadivel, N Krishnaraj, J Ramprasath. ”Host-based Intrusion Detection System using Sequence of System Calls” International Journal of Engineering and Management Research, Vandana Publications, Volume 4, Issue 2, 241-247, 2014
27. N Krishnaraj, S Smys.”A multihoming ACO-MDV routing for maximum power efficiency in an IoT environment” Wireless Personal Communications 109 (1), 243-256, 2019.
28. N Krishnaraj, R Bhuvanesh Kumar, D Rajeshwar, T Sanjay Kumar, Implementation of energy aware modified distance vector routing protocol for energy efficiency in wireless sensor networks, 2020 International Conference on Inventive Computation Technologies (ICICT),201-204
29. Ibrahim, S. Jafar Ali, and M. Thangamani. "Enhanced singular value decomposition for prediction of drugs and diseases with hepatocellular carcinoma based on multi-source bat algorithm based random walk." Measurement 141 (2019): 176-183. https://doi.org/10.1016/j.measurement.2019.02.056
30. Ibrahim, Jafar Ali S., S. Rajasekar, Varsha, M. Karunakaran, K. Kasirajan, Kalyan NS Chakravarthy, V. Kumar, and K. J. Kaur. "Recent advances in performance and effect of Zr doping with ZnO thin film sensor in ammonia vapour sensing." GLOBAL NEST JOURNAL 23, no. 4 (2021): 526-531. https://doi.org/10.30955/gnj.004020 , https://journal.gnest.org/publication/gnest_04020
31. N.S. Kalyan Chakravarthy, B. Karthikeyan, K. Alhaf Malik, D.Bujji Babbu,. K. Nithya S.Jafar Ali Ibrahim , Survey of Cooperative Routing Algorithms in Wireless Sensor Networks, Journal of Annals of the Romanian Society for Cell Biology ,5316-5320, 2021

32. Rajmohan, G, Chinnappan, CV, John William, AD, Chandrakrishan Balakrishnan, S, Anand Muthu, B, Manogaran, G. Revamping land coverage analysis using aerial satellite image mapping. Trans Emerging Tel Tech. 2021; 32:e3927. https://doi.org/10.1002/ett.3927
33. Vignesh, C.C., Sivaparthipan, C.B., Daniel, J.A. et al. Adjacent Node based Energetic Association Factor Routing Protocol in Wireless Sensor Networks. Wireless Pers Commun 119, 3255–3270 (2021). https://doi.org/10.1007/s11277-021-08397-0.
34. C Chandru Vignesh, S Karthik, Predicting the position of adjacent nodes with QoS in mobile ad hoc networks, Journal of Multimedia Tools and Applications, Springer US,Vol 79, 8445-8457,2020