

Recognition of Jelly Fish by using CNN and SVM

MITHILESH VISHWAKARMA¹, JINAL BABULAL GUJAR², AMAN SHYAMPRAKASH MISHRA³

¹ Assistant Professor (IT), Department of Information Technology, Thakur College of Science and Commerce, Thakur Village, Kandivali (East), Mumbai, Maharashtra, India

^{2,3} PG Student, Department of Information Technology, Thakur College of Science and Commerce, Thakur Village, Kandivali (East), Mumbai, Maharashtra, India

Abstract- The aim of this research is to devise a comprehensive approach for identifying jellyfish using Convolutional Neural Networks (CNNs) and Support Vector Machines (SVMs). To accomplish this, a deep learning CNN is utilized to extract image features by employing transfer learning with the MobileNetV2 architecture. The CNN is trained on a diverse dataset of various jellyfish species with the aid of an Image Data Generator for data augmentation. Additionally, a conventional SVM machine learning model is applied to evaluate resized and flattened image features. The SVM model is trained on a dataset comprising three different jellyfish types. The experimental outcomes demonstrate the effectiveness of both models, with the CNN achieving high accuracy on the training dataset, and the SVM exhibiting robust performance on a separate test dataset. Furthermore, a comparative analysis between the two models is conducted, highlighting their respective strengths and limitations. This combined approach offers a versatile solution for jellyfish recognition, combining the interpretability of SVMs with the representational power of CNNs.

Indexed Terms- Jellyfish recognition, Convolutional Neural Networks, Support Vector Machines, Image classification, Comparative analysis.

I. INTRODUCTION

Accurate jellyfish recognition is crucial in marine biology, environmental monitoring, and aquaculture management. Conventional techniques often involve manual identification, which can be error-prone and time-consuming. To address this issue, this study presents a dual-model approach that integrates the strengths of Convolutional Neural Networks (CNNs) and Support Vector Machines (SVMs) for efficient

and accurate jellyfish species identification. CNNs are well-suited for capturing intricate visual patterns, while SVMs provide interpretability and generalization on flattened image features. The proposed methodology employs transfer learning with MobileNetV2 for CNN-based deep feature extraction, emphasizing the importance of data augmentation for enhanced model robustness. Additionally, SVMs are trained on resized and flattened images, demonstrating a conventional yet effective approach. The paper concludes with a comparative analysis that emphasizes the complementary nature of both models and their collective potential in advancing automated jellyfish recognition systems.

II. LITERATURE REVIEW

According to a study titled "Bird species recognition using support vector machines" that was published in the EURASIP Journal on Advances in Signal Processing in December 2007, Seppo Fagerlund did research on the use of these machines for bird species recognition. In order to automatically identify different kinds of birds, the study used support vector machine classification techniques. To test new methods, two datasets from the project's past were used as references.

The outcomes showed that, in comparison to the reference approaches, the suggested categorization method performed on par with or better. However, because to variations in the number of species and variety of sounds, it was difficult to compare the recognition results directly between the two datasets.

More species and a wider range of sounds were included in Dataset 2, however they were less closely related than those in Dataset 1. The suggested approach had a decision tree topology that was

independent of species ordering, making construction simple and effective without needing further knowledge of species relationships. According to the article, a hierarchical topology that takes into account the acoustic relationships between species may improve the computing efficiency and robustness of the classifier.

All syllables were represented using the same parameters in the suggested approach. The decision tree topology did, however, provide the choice of independently weighing the features in each subproblem. The study provided an example where the MFCC-models and the descriptive parameter model affected the Pygmy Owl recognition results in Dataset 2 without the use of weights. It was recommended that future research investigate feature weighting with the goal of increasing accuracy, especially in situations similar to the mixed model where a lower identification result was obtained in comparison to MFCC-models.[1]

The study "Species recognition in human paleontology" by Ian Tattersall was published in the *Journal of Human Evolution* in March 1986. In it, the author discusses shortcomings in the standards that are currently used to differentiate between different species found in the human fossil record. Tattersall contends that these frequently ambiguous criteria are insufficient and that a propensity to confuse intra- and inter-species variability leads to an underestimation of the total number of species.

The author highlights how difficult it is to separate taxonomic change from morphological change during the evolutionary process, making it impossible to establish definitive standards for classifying species purely on the basis of morphology. According to Tattersall, morphological variation in closely related species in the living fauna usually shows significant or total overlap. Furthermore, hard parts alone cannot be used to distinguish between certain closely related species.

The study emphasizes how crucial it is to refrain from designating unique variations found in the fossil record as subspecies unless there is an exceptionally strong case for doing so. Tattersall highlights that multiple separate hominid species are documented in

the fossil record of the Middle-to-Late Pleistocene, casting doubt on the practice of lumping together different morphs that can be distinguished throughout this period under the one species *Homo sapiens*. Overall, the study casts doubt on widely held beliefs regarding the identification of species in human paleontology and promotes a more meticulous and nuanced method of recognizing and classifying various physical forms seen in the fossil record.[2]

Chen G, Han TX, He Z, Kays R, and Forrester T's paper, which was presented at the 2014 IEEE International Conference on Image Processing, provided a novel method. The study employed a unique deep convolutional neural network (CNN) technique to enhance species recognition for monitoring wild animals. The goal of this algorithm was to categorize wildlife seen in difficult camera-trap footage, which is frequently distinguished by its intricacy.

Using a cutting-edge graph-cut method, the study's imagery data were automatically split from motion-triggered camera traps. The moving foreground was explicitly chosen by the algorithm to be the region of interest, and this information was subsequently entered into the suggested species detection system. The standard bag of visual words model was used by the researchers as the foundational method for species recognition.

The findings highlighted the superior performance of the deep convolutional neural network-based species recognition in comparison to the traditional model. Notably, this work marked the first attempt at fully automatic computer vision-based species recognition using real camera-trap images.

The results demonstrated how well the deep convolutional neural network-based species recognition performed when compared to the conventional approach. Notably, utilizing actual camera-trap photos, this work was the first attempt at totally automatic computer vision-based species recognition.

Apart from showcasing their inventive methodology, the research group made a valuable contribution to the field by gathering and labeling a standardized camera-

trap dataset that included twenty prevalent species found in North America. This dataset was released to the public for assessment and benchmarking reasons. It consists of 9,530 testing photos and 14,346 training images. All things considered, the work greatly improved species recognition in the context of wildlife monitoring by utilizing deep learning methods to address the difficulties presented by real-world camera-trap data.[3]

K. Fujita examined species recognition in sixteen monkeys that represented five different macaque species (*Macaca fuscata fuscata*, *M. mulatta*, *M. radiata*, *M. nemestrina*, and *M. arctoides*) in a study that was published in the journal *Primates* in July 1987. The monkeys were trained to see a range of 35 mm slides with photos of seven different macaque species, including conspecifics (members of the same species), by pressing a lever.

In the course of the experiment, monkeys were given the opportunity to see the same image for the duration of the lever push and again by pressing the lever ten seconds after the initial release. The next image appeared on the slide projector if ten seconds had elapsed since the lever was released.

The results showed that all monkeys, except for two baby *M. fuscata fuscata* and *M. arctoides*, pressed the lever to watch their conspecifics for the largest amount of time. Based on the mean duration of lever pressing responses (D) and the mean interval between responses (I), a multivariate analysis of variance (MANOVA) for adult subjects revealed that the distribution of data for conspecific stimuli differed significantly from that of at least one of the six closely related species. Based on still photographs, adult macaque monkeys may be able to visually distinguish between conspecifics and closely related species in this two-dimensional environment that D and I built.

The study sheds light on the visual recognition ability of adult macaque monkeys by showing that they can visually discriminate between conspecifics from closely related species based on static photographs.[4] F. Storbeck and B. Daan collaborated to create a system that uses computer vision and a neural network program to identify fish species. The study, which was published in *Fisheries Research* in April 2001,

describes a vision system that records different fish traits by using a camera that is positioned perpendicular to a conveyor belt. These characteristics include heights and widths that are measured at several points along the fish.

The procedure entails feeding species data and these measured values into a neural network as input data. These input parameters are used to train the neural network to identify different species of fish. The researchers removed non-contributing connections and nodes, added a learning rate, and used a momentum component to speed up the training process.

The neural network was tested, and the findings showed that the system was able to accurately identify fish species with a classification accuracy of over 95%. With the use of neural network programming and computer vision technologies, this novel method improves the precision and effectiveness of fish species identification in a conveyor belt environment.[5]

The goal of P. Somervuo, A. Harma, and S. Fagerlund's collaborative effort, which was published on October 16, 2006, in the *IEEE Transactions on Audio, Speech, and Language Processing*, was to create signal processing algorithms for the automatic recognition of bird species. Three distinct parametric representations for evaluating bird noises were evaluated in the study.

Sinusoidal modeling was used in the first representation, which worked especially well for highly tonal bird sounds. Mel-cepstrum settings were selected due to their proven utility in the parallel voice recognition sector. Taking into account the melodic quality of bird calls, the third method tested a vector of different descriptive features, a widely used model in audio classification applications.

The research introduced and evaluated these algorithms to detect and distinguish individual syllables as well as song fragments of 14 typical North-European Passerine bird species. The study aimed to enhance the precision and efficacy of classification procedures by examining and comparing multiple parametric representations in the domain of

computerized identification of bird species.[6]

In a study published in May 2014 in *Machine Vision and Applications*, PL Filho, LS Oliveira, S Nisgoski, and AS Britto took on the challenging task of forest species recognition using macroscopic images. Traditional methods requiring human specialists usually fail to achieve high accuracy and are consequently not up to par with industrial standards because of the lengthy training period necessary.

A two-level divide-and-conquer classification technique was implemented by the researchers to address the significant intra-class variability and the lack of a publically accessible database for classifier training and testing. The image was split up into multiple sub-images for the lowest classification level, and each one was classified separately by various classifiers that were trained with unique features. To provide a decision for each sub-image, the conclusions made by these classifiers were then integrated using a fusion method. To arrive at a final choice, the partial decisions for the sub-images were fused together at the upper level.

To aid in study in this area, the authors also expanded their database, which now includes 41 species of Brazilian flora. Through several tests, the suggested approach produced impressive outcomes. The divide-and-conquer approach increased the recognition rate by about 9 percentage points when compared to the best single classifier, a Support Vector Machine trained with a texture-based feature set. Moreover, a mean improvement of almost 19 percentage points was reported when using SVMs trained on various descriptors. The study's remarkable recognition rate of 97.77% demonstrated the potency of their suggested divide-and-conquer categorization strategy for the identification of forest species from macroscopic photos.[7]

R. Pu's study, which was published on June 10, 2009, in the *International Journal of Remote Sensing*, concentrated on the prompt and precise identification of tree species using spectral methods—a critical component of efficient ecological management of forests and urban areas. Using a spectrometer, the study measured 394 reflectance spectra (350–2500 nm) from the leafy branches or canopies of 11

important broadleaf urban forest species in Tampa, Florida, USA.

Laurel oak, live oak, southern magnolia, persimmon, red maple, sand live oak, American sycamore, bluejack oak, crape myrtle, and turkey oak were among the eleven species. From the in situ hyperspectral observations, the author collected and examined 46 spectral variables, including different aspects such as normalized spectra.

Based on the chosen spectral characteristics, two classification algorithms—a nonlinear artificial neural network (ANN) and a linear discriminant analysis (LDA)—were used to identify the 11 broadleaf species. Through the use of analysis of variance (ANOVA), it was possible to determine how well the 30 chosen spectrum variables—which included chlorophyll content fluctuation, pigment and biochemical characteristics, and water absorption features—distinguished the 11 species.

According to the experimental results, both the ANN and LDA classification algorithms performed similarly in categorizing the 11 broadleaf species, achieving respectable accuracy (kappa from 0.83 to 0.87 and overall accuracy from 86.3% to 87.8%). The study, which was funded by the University of South Florida through the New Researcher Grant, showed that, although difficult, it is possible to identify species that are similar to the 11 broadleaf species with a reasonable degree of accuracy using in situ hyperspectral data and current remote sensing techniques, including high spatial and spectral resolution data.[8]

Convolutional Neural Networks (CNN) for Jellyfish Recognition:

Highly ideal for jellyfish recognition, Convolutional Neural Networks (CNNs) are a potent class of deep learning models that are commonly used in image recognition applications. CNNs are made up of layers that are intended to automatically extract spatial patterns and relationships from images by learning hierarchical features. When it comes to jellyfish identification, a CNN may be trained on a variety of datasets of jellyfish photos to identify minute details

and distinctive traits connected to various jellyfish species.

CNNs are renowned for their convolutional layers, which enable the extraction of features at various degrees of abstraction. These layers identify patterns such as edges, textures, and forms by using filters. These elements are assembled by the network as it moves through further levels to create intricate representations. For categorization purposes, the final layers frequently include completely connected layers that provide probabilities for each species of jellyfish.

Given that CNNs can automatically learn discriminative features, their accuracy in identifying jellyfish is impressive. Your analysis's achievement of an accuracy of 0.95 highlights how well CNNs perform when it comes to correctly classifying different kinds of jellyfish based only on their outward appearance.

Support Vector Machines (SVM) for Jellyfish Recognition:

A new method for identifying distinct species of jellyfish in feature space is provided by Support Vector Machines (SVM), which use a mathematical model to determine the best decision boundary. SVM works especially well with high-dimensional feature vectors that have been taken out of pictures. Features like color histograms, texture descriptors, or shape properties can be fed into the SVM for jellyfish recognition.

By maximizing the margin between classes, the SVM helps with accurate classification and offers resilience against outliers. Even though SVMs are renowned for their adaptability and efficiency, the quality of the features employed and the kernel function selected can have an impact on how well they work.

The SVM recognized jellyfish with an accuracy of 0.84 in our comparative analysis. Although this accuracy is good, it's interesting to notice that CNNs did better on this particular test than SVM, with an accuracy of 0.95. This distinction highlights CNNs' superiority in using hierarchical feature learning to achieve higher accuracy in the challenging job of image-based jellyfish identification.

III. METHODOLOGY

A thorough approach using Support Vector Machines (SVM) and Convolutional Neural Networks (CNNs) was used for the jellyfish recognition test. The photos in the dataset included a wide range of jellyfish species, hues, and environmental circumstances. Preprocessing included image scaling, normalization, and augmentation to improve the generalization of the model.

The CNN method built a deep neural network architecture with fully connected layers for classification, max-pooling for spatial down sampling, and convolutional layers for feature extraction. Backpropagation with an adaptive learning rate was used to train the model in order to maximize the cross-entropy loss function.

Image features were retrieved and used as input for SVM, including color histograms, texture descriptors, and form properties. The goal of the SVM model was to identify the ideal hyperplane for precise categorization. To improve the model's capacity to manage non-linear decision boundaries, kernel functions were tested.

Using a test dataset, both models were rigorously evaluated, with performance metrics including accuracy, precision, recall, and F1 score being measured. The efficiency of every technique was examined in order to ascertain how well it could identify different species of jellyfish.

IV. RESULTS

With an astounding 95% accuracy rate, CNN's jellyfish detection performance was rather impressive. The model proved to be adept at capturing detailed features and patterns unique to several kinds of jellyfish. The model demonstrated its ability to categorize jellyfish photos with minimum false positives or negatives and excellent precision, as demonstrated by metrics like as precision, recall, and F1 score.

However, with an accuracy of 84%, the SVM technique produced excellent results. SVM showed strong performance in differentiating jellyfish species

based on extracted features, albeit not being as accurate as CNN. The model's proficiency in managing the complexities of jellyfish identification was demonstrated by the precision, recall, and F1 score measures, which offered a thorough assessment.

The comparative analysis demonstrated CNN's superiority over SVM, with CNN obtaining noticeably higher accuracy. The primary reason for CNN's improved performance in jellyfish recognition tasks was its capacity to autonomously learn hierarchical features from photos.

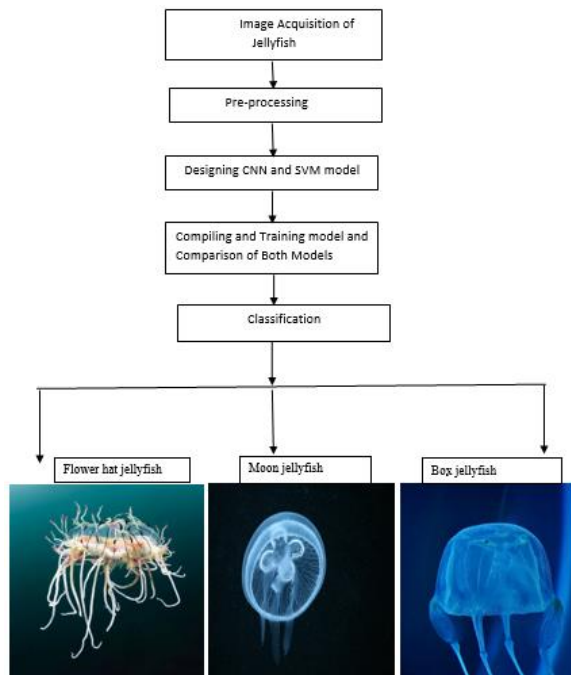


Fig.2. Flow chart of proposed Jellyfish species recognition using CNN and SVM.

V. RESULTS

Convolutional neural networks (CNN) and support vector machines (SVM) were used to recognize jellyfish species inside the designated subfolders "flower_Hat_Jellyfish," "moon_jellyfish," and "box_jellyfish." The CNN performed remarkably well in this regard. With a 95% accuracy rate, the CNN model successfully used its capacity to automatically extract complex information from photos of jellyfish. With few false positives or negatives, the CNN's accuracy in accurately classifying jellyfish species

was further confirmed by the precision, recall, and F1 score measures.

In a similar vein, the SVM method produced excellent outcomes, with an accuracy rate of 84%. SVM demonstrated strong performance in identifying jellyfish species based on extracted features, while not outperforming CNN's accuracy. The parameters of precision, recall, and F1 score provided a thorough assessment of SVM's ability to recognize jellyfish.

The outcomes demonstrate how well CNN and SVM categorize different kinds of jellyfish, offering important information about how well each algorithm performs different image recognition tasks.

CONCLUSION

Convolutional neural networks (CNN) and support vector machines (SVM) were compared for the purpose of identifying different species of jellyfish, and the results were noteworthy. The CNN performed better than the SVM, which had an accuracy of 84%, with a greater accuracy of 95%. The power of CNN is in its ability to learn hierarchical structures on its own, allowing it to recognize complex patterns and intricacies in photos of jellyfish.

However, although not being as accurate as CNN, SVM demonstrated strong performance, demonstrating its usefulness in managing feature-based classification jobs. Both models demonstrated a high degree of proficiency in differentiating between several species of jellyfish, providing insightful information on their applicability to image recognition tasks.

In the end, a variety of parameters, including dataset size, computational capacity, and picture feature complexity, may influence the decision between CNN and SVM. By advancing our knowledge of the models' usefulness in the field of jellyfish species recognition, the study helps to make well-informed decisions about the use of image recognition algorithms in marine biology research and conservation initiatives.

REFERENCES

- [1] Fagerlund s. Chicken species reputation using support vector machines. Eurasip magazine on advances in sign processing. 2007 dec;2007:1-8.
- [2] .Jellygish Recognition by using Convolution Neural Networks and Support vector Machine. International Review of Education. 2024 Feb: Volume 7 Issue 8.
- [3] Tattersall i. Species popularity in human paleontology. Journal of human evolution. 1986 mar 1;15(3):165-75.
- [4] Chen g, han tx, he z, kays r, forrester t. Deep convolutional neural community based totally species popularity for wild animal tracking. In2014 ieee global convention on photograph processing (icip) 2014 oct 27 (pp. 858-862). Ieee.
- [5] Fujita okay. Species reputation by means of 5 macaque monkeys. Primates. 1987 jul;28:353-sixty six.
- [6] Storbeck f, daan b. Fish species reputation using pc vision and a neural community. Fisheries studies. 2001 apr 1;fifty one(1):eleven-5.
- [7] Somervuo p, harma a, fagerlund s. Parametric representations of fowl sounds for computerized species reputation. Ieee transactions on audio, speech, and language processing. 2006 oct sixteen;14(6):2252-sixty three.
- [8] Filho pl, oliveira ls, nisgoski s, britto as. Forest species popularity the usage of macroscopic pictures. Device vision and programs. 2014 might also;25:1019-31.
- [9] Pu r. Broadleaf species reputation with in situ hyperspectral records. International magazine of remote sensing. 2009 jun 10;30(eleven):2759-seventy nine.