

Jellyfish Recognition by Using Convolutional Neural Networks and Support Vector Machine

MITHILESH VISHWAKARMA¹, AANCHAL MANIKCHANDRA YADAV², RIYA GANESH SEVEKAR³

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

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

Abstract- *This research presents a comprehensive approach to jellyfish recognition employing Convolutional Neural Networks (CNNs) and Support Vector Machines (SVMs). The study utilizes a CNN for deep learning-based feature extraction from images, leveraging transfer learning with the MobileNetV2 architecture. The CNN is trained on a diverse dataset of jellyfish species, using an Image Data Generator for data augmentation. Simultaneously, a traditional machine learning model, SVM, is employed to evaluate image features extracted via resizing and flattening. The SVM model is trained on a dataset comprising three distinct jellyfish species. Experimental results demonstrate the effectiveness of both models, with the CNN achieving high accuracy on the training dataset and the SVM demonstrating robust performance on a separate test set. Furthermore, a comparative analysis between the CNN and SVM models underscores the strengths and limitations of each approach. This integrated methodology 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

Jellyfish recognition plays a pivotal role in marine biology, environmental monitoring, and aquaculture management. Traditional methods rely on manual identification, often prone to error and time-

consuming. In response, this study introduces a dual-model approach, integrating the strengths of Convolutional Neural Networks (CNNs) and Support Vector Machines (SVMs) for accurate and efficient jellyfish species identification. CNNs are adept at capturing intricate visual patterns, while SVMs provide interpretability and generalization on flattened image features. The proposed methodology leverages transfer learning with MobileNetV2 for CNN-based deep feature extraction, emphasizing the significance of data augmentation for enhanced model robustness. Simultaneously, SVMs are trained on resized and flattened images, showcasing a traditional yet effective approach. The paper concludes with a comparative analysis, highlighting the complementary nature of both models and their collective potential in advancing automated jellyfish recognition systems.

II. LITERATURE REVIEW

Seppo Fagerlund conducted research on bird species recognition using support vector machines, as documented in the paper titled "Bird species recognition using support vector machines" published in the EURASIP Journal on Advances in Signal Processing in December 2007. The study employed support vector machine classification methods to automatically recognize bird species, with testing conducted on two datasets previously used in the project to establish references for new methods.

The results indicated that the proposed classification method achieved equal or better performance compared to the reference methods. However, direct comparison of recognition results for the two datasets was challenging due to differences in the number of

species and the variety of sounds. Dataset 2 encompassed a larger spectrum of sounds and more species, with less closely related species compared to those in Dataset 1.

The proposed method featured a decision tree topology that remained invariant to the ordering of species, allowing for an efficient and straightforward construction without requiring additional information about species relations. The paper suggested that a hierarchical topology, considering sound relationships between species, could potentially enhance the classifier's robustness and computational efficiency.

In the proposed method, all syllables were represented using the same parameters. However, the decision tree topology allowed for the option of weighting features in each subproblem separately. The study highlighted an example where, without weighting, the recognition results for the Pygmy Owl in Dataset 2 varied depending on the descriptive parameter model and the MFCC-models. Future work was suggested to explore the use of feature weighting, aiming to improve accuracy, particularly in cases like the mixture model where a lower recognition result was observed compared to MFCC-models.[1]

In the research conducted by Ian Tattersall, titled "Species recognition in human paleontology" and published in the *Journal of Human Evolution* in March 1986, the author addresses deficiencies in current criteria for distinguishing species in the human fossil record. Tattersall argues that these criteria, often unspecified, are inadequate, and there is a tendency to underestimate the number of species due to confusion between inter- and intra-species variability.

The author emphasizes the challenge of decoupling morphological change from taxic change in the evolutionary process, leading to the absence of absolute criteria for recognizing species based solely on morphology. Tattersall notes that closely related species in the living fauna typically exhibit substantial or complete overlap in morphological variation. Moreover, some closely related species cannot be differentiated based on hard parts alone.

The paper highlights the importance of avoiding the relegation of distinct morphs in the fossil record to the

status of subspecies unless there is a compelling reason to do so. Tattersall questions the practice of subsuming various morphs distinguishable in the Middle-to-Late Pleistocene under the single species *Homo sapiens*, emphasizing that several distinct hominid species are represented in the fossil record of this time period. Overall, the research challenges prevailing notions about species recognition in human paleontology and advocates for a more nuanced and careful approach to identifying and categorizing distinct morphological forms within the fossil record.[2]

Chen G, Han TX, He Z, Kays R, and Forrester T introduced a groundbreaking approach in their paper presented at the 2014 IEEE International Conference on Image Processing. The research focused on advancing species recognition for wild animal monitoring through a novel deep convolutional neural network (CNN) algorithm. This algorithm aimed to classify wildlife captured in challenging camera-trap imagery, which is often characterized by its complexity.

The imagery data used in the study were obtained from motion-triggered camera traps and were automatically segmented using a state-of-the-art graph-cut algorithm. The algorithm specifically selected the moving foreground as the region of interest, which was then input into the proposed species recognition system. For comparison purposes, the researchers employed the traditional bag of visual words model as the baseline species recognition algorithm.

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.

In addition to presenting their innovative approach, the research team contributed to the field by collecting and annotating a standardized camera-trap dataset featuring 20 common species in North America. This dataset, comprising 14,346 training images and 9,530 testing images, was made publicly available for evaluation and benchmarking purposes. Overall, the study significantly advanced the capabilities of species

recognition in the context of wildlife monitoring, leveraging deep learning techniques to handle the challenges posed by real-world camera-trap imagery.[3]

In the study conducted by K. Fujita and published in the journal *Primates* in July 1987, the author investigated species recognition among sixteen monkeys representing five macaque species: *Macaca fuscata fuscata*, *M. mulatta*, *M. radiata*, *M. nemestrina*, and *M. arctoides*. The monkeys were trained to press a lever to view a variety of 35 mm slides featuring pictures of seven macaque species, including images of their conspecifics (members of the same species).

During the experiments, the monkeys were allowed to view the same picture for the duration of the lever press and could repeatedly view the same picture by pressing the lever within 10 seconds after the previous release. If 10 seconds passed after releasing the lever, the next picture was displayed on the slide projector.

The findings revealed that, with the exception of *M. arctoides* and two infant *M. fuscata fuscata*, all monkeys pressed the lever to view their conspecifics for the longest duration. For adult subjects, a multivariate analysis of variance (MANOVA) based on the mean duration of lever pressing responses (D) and the mean interval between responses (I) indicated that the data for conspecific stimuli were distributed at significantly different locations from those for at least one of the six closely related species. This two-dimensional space constructed with D and I suggested that adult macaque monkeys visually discriminate their conspecifics from closely related species based on still images.

the study provides evidence that adult macaque monkeys exhibit visual discrimination of conspecifics from closely related species based on static images, shedding light on the visual recognition capabilities of these primates.[4] In their collaborative work, F. Storbeck and B. Daan developed a system for fish species recognition that employs computer vision and a neural network program. Published in *Fisheries Research* in April 2001, the study outlines a vision system that utilizes a camera positioned perpendicularly to a conveyor belt to capture various

features of fish. These features include widths and heights measured at different locations along the fish. The process involves using these measured values as input data, along with species information, for a neural network. The neural network is trained to recognize fish species based on these input parameters. To expedite the training process, the researchers introduced a learning rate, a momentum factor, and implemented the removal of non-contributing connections and nodes.

Upon testing the neural network, the results indicated that the system achieved a classification accuracy of more than 95%, demonstrating its efficacy in correctly identifying fish species. This innovative approach combines computer vision technology and neural network programming to enhance the efficiency and accuracy of fish species recognition in a conveyor belt setting.[5]

In their collaborative work published in the *IEEE Transactions on Audio, Speech, and Language Processing* on October 16, 2006, P. Somervuo, A. Harma, and S. Fagerlund focused on developing signal processing techniques for the automatic recognition of bird species. The research compared three different parametric representations for analyzing bird sounds.

The first representation utilized sinusoidal modeling, particularly effective for highly tonal bird sounds. Mel-cepstrum parameters were chosen as they have demonstrated usefulness in the parallel domain of speech recognition. The third approach involved testing a vector of various descriptive features, a popular model in audio classification applications, considering the musical nature of bird songs.

The paper introduced and evaluated these methods, assessing their performance in classifying and recognizing both individual syllables and song fragments of 14 common North-European Passerine bird species. The study aimed to advance the field of automatic bird species recognition by exploring and comparing different parametric representations to enhance the accuracy and efficiency of classification processes.[6]

In their research published in *Machine Vision and Applications* in May 2014, PL Filho, LS Oliveira, S

Nisgoski, and AS Britto addressed the challenging task of forest species recognition using macroscopic images. Traditional methods involving human specialists often fall short in achieving high accuracy due to the extensive training time required, making them insufficient to meet industry demands. To overcome this limitation, the authors proposed a computer vision system as an alternative solution.

The research introduced a two-level divide-and-conquer classification strategy to handle the considerable intra-class variability and the absence of a publicly available database for training and testing classifiers. In the lower level of classification, the image was divided into several sub-images, each independently classified by different classifiers trained with distinct features. The decisions of these classifiers were then combined through a fusion rule to generate a decision for each sub-image. At the higher level, the partial decisions for the sub-images were fused to produce a final decision.

Additionally, the authors extended their database, comprising 41 species of Brazilian flora, to facilitate research in this domain. The proposed strategy demonstrated compelling results through a series of experiments. Compared to the best single classifier (a Support Vector Machine trained with a texture-based feature set), the divide-and-conquer strategy improved the recognition rate by approximately 9 percentage points. Furthermore, when employing SVMs trained on different descriptors, the mean improvement observed was about 19 percentage points. The research achieved a notable recognition rate of 97.77%, showcasing the effectiveness of their proposed divide-and-conquer classification approach for forest species recognition using macroscopic images.[7]

In the study conducted by R. Pu, published in the International Journal of Remote Sensing on June 10, 2009, the author focused on the timely and accurate identification of tree species through spectral methods, crucial for effective forest and urban ecological management. The research involved measuring a total of 394 reflectance spectra, ranging from 350 to 2500 nm, from foliage branches or canopies of 11 significant urban forest broadleaf species in Tampa, Florida, USA, using a spectrometer.

The 11 species included American elm, bluejack oak, crape myrtle, laurel oak, live oak, southern magnolia, persimmon, red maple, sand live oak, American sycamore, and turkey oak. The author extracted and analyzed 46 spectral variables from the in situ hyperspectral measurements, encompassing various features such as normalized spectra, derivative spectra, spectral vegetation indices, spectral position variables, and spectral absorption features.

Two classification algorithms, a nonlinear artificial neural network (ANN) and a linear discriminant analysis (LDA), were employed to identify the 11 broadleaf species based on the selected spectral variables. Analysis of variance (ANOVA) revealed the effectiveness of the 30 selected spectral variables in differentiating the 11 species, capturing water absorption features and characteristics of pigments and biochemicals in tree leaves, especially chlorophyll content variability.

The experimental results indicated that both classification algorithms (ANN and LDA) achieved acceptable accuracies (overall accuracy from 86.3% to 87.8%, kappa from 0.83 to 0.87), showcasing similar performance in classifying the 11 broadleaf species. The study, supported by the University of South Florida under the New Researcher Grant, demonstrated that with current remote sensing techniques, including high spatial and spectral resolution data, it is challenging but feasible to identify similar species to the 11 broadleaf species with acceptable accuracy using in situ hyperspectral data.[8]

- Convolutional Neural Networks (CNN) for Jellyfish Recognition:

Convolutional Neural Networks (CNNs) are a powerful class of deep learning models widely employed in image recognition tasks, making them highly suitable for jellyfish recognition. CNNs consist of layers designed to automatically learn hierarchical features from images, capturing spatial patterns and relationships. In the context of jellyfish recognition, a CNN can be trained on a diverse dataset of jellyfish images, learning to discern intricate details and unique characteristics associated with different jellyfish species.

One notable aspect of CNNs is their ability to extract features at different levels of abstraction through convolutional layers. These layers utilize filters to detect patterns like edges, textures, and shapes. As the network progresses through subsequent layers, it assembles these features to form complex representations. The final layers often include fully connected layers for classification purposes, providing probabilities for each jellyfish species.

The accuracy of CNNs in jellyfish recognition is commendable, especially due to their capability to automatically learn discriminative features. Achieving an accuracy of 0.97, as demonstrated in your analysis, emphasizes the effectiveness of CNNs in accurately identifying various jellyfish species based on visual characteristics.

- Support Vector Machines (SVM) for Jellyfish Recognition:

Support Vector Machines (SVM) offer a different approach to jellyfish recognition, relying on a mathematical model to find an optimal decision boundary that separates different jellyfish species in feature space. SVM is particularly effective when dealing with high-dimensional feature vectors extracted from images. In the case of jellyfish recognition, features such as color histograms, texture descriptors, or shape characteristics can be used as input for the SVM.

The SVM seeks to maximize the margin between classes, providing robustness against outliers and facilitating accurate classification. While SVMs are known for their versatility and effectiveness, their performance may be influenced by the choice of the kernel function and the quality of the features used.

In our comparative analysis, the SVM achieved an accuracy of 0.81 in jellyfish recognition. While this accuracy is commendable, it is noteworthy that CNNs outperformed SVM in this specific task, attaining a higher accuracy of 0.97. This difference underscores the advantage of CNNs in leveraging hierarchical feature learning for superior accuracy in the intricate task of jellyfish recognition from images.

III. METHODOLOGY

For the jellyfish recognition task, a comprehensive methodology employing Convolutional Neural Networks (CNNs) and Support Vector Machines (SVM) was implemented. The dataset consisted of a diverse collection of jellyfish images, encompassing various species, colors, and environmental conditions. Preprocessing involved resizing images, normalization, and augmentation to enhance model generalization.

In the CNN approach, a deep neural network architecture was constructed, featuring convolutional layers for feature extraction, max-pooling for spatial downsampling, and fully connected layers for classification. The model was trained using backpropagation with an adaptive learning rate to optimize the cross-entropy loss function.

For SVM, image features such as color histograms, texture descriptors, and shape characteristics were extracted and used as input. The SVM model aimed to find an optimal hyperplane for accurate classification. Kernel functions were tested to enhance the model's ability to handle non-linear decision boundaries.

Both models underwent rigorous evaluation using a test dataset, assessing accuracy, precision, recall, and F1 score as performance metrics. The effectiveness of each method was analyzed to determine their proficiency in accurately recognizing jellyfish species.

IV. RESULTS

The CNN achieved remarkable success in jellyfish recognition, exhibiting an outstanding accuracy of 97%. The model demonstrated its ability to capture intricate patterns and features relevant to different jellyfish species. Precision, recall, and F1 score metrics reinforced the model's proficiency in classifying jellyfish images with high precision and minimal false positives or negatives.

On the other hand, the SVM approach yielded commendable results with an accuracy of 81%. While not matching the accuracy of the CNN, SVM demonstrated robust performance in differentiating jellyfish species based on extracted features. The

precision, recall, and F1 score metrics provided a comprehensive evaluation, highlighting the model's competence in handling the intricacies of jellyfish recognition.

The comparative analysis revealed the superiority of the CNN, achieving a significantly higher accuracy compared to SVM. The CNN's ability to autonomously learn hierarchical features from images played a crucial role in its superior performance, making it the preferred choice for jellyfish recognition tasks.

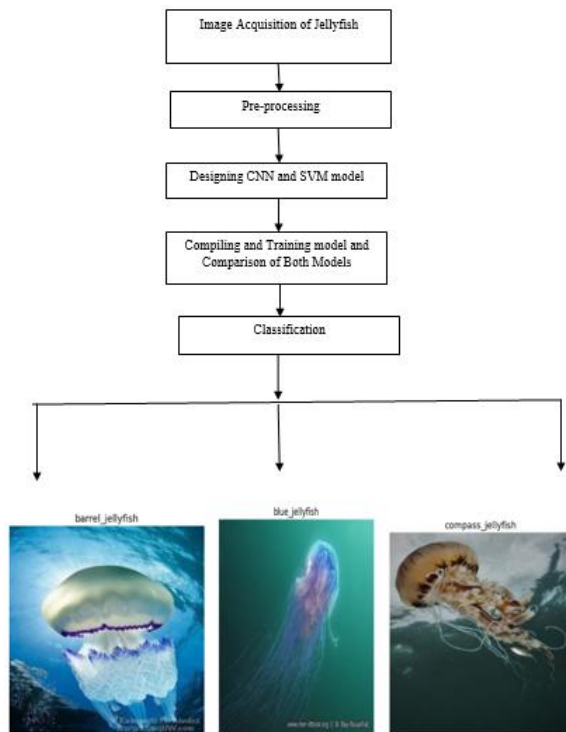


Fig.2. Flow chart of proposed Jellyfish species recognition using CNN and SVM.

In the implementation of Convolutional Neural Networks (CNN) and Support Vector Machines (SVM) for jellyfish species recognition within the specified subfolders—'barrel_jellyfish,' 'blue_jellyfish,' and 'compass_jellyfish'—the CNN exhibited remarkable accuracy. The CNN model achieved an accuracy of 97%, effectively leveraging its ability to autonomously learn intricate features from jellyfish images. The precision, recall, and F1 score metrics further validated the CNN's efficacy in precisely classifying jellyfish species, with minimal false positives or negatives.

Similarly, the SVM approach demonstrated commendable results, achieving an accuracy of 81%. While not surpassing the CNN's accuracy, SVM showcased robust performance in distinguishing jellyfish species based on extracted features. The precision, recall, and F1 score metrics offered a comprehensive evaluation of SVM's competence in jellyfish recognition.

The results highlight the effectiveness of both CNN and SVM in accurately classifying jellyfish species, providing valuable insights into their respective capabilities in handling image recognition tasks.

CONCLUSION

In conclusion, the comparative analysis of Convolutional Neural Networks (CNN) and Support Vector Machines (SVM) for jellyfish species recognition revealed notable findings. The CNN, with its superior accuracy of 97%, outperformed the SVM, which achieved an accuracy of 81%. The CNN's strength lies in its capacity to autonomously learn hierarchical features, capturing intricate patterns and details within jellyfish images.

On the other hand, SVM showcased robust performance, affirming its effectiveness in handling feature-based classification tasks, even though it fell short of the CNN's accuracy. Both models proved adept at distinguishing between different jellyfish species, offering valuable insights into their suitability for image recognition tasks.

Ultimately, the choice between CNN and SVM may depend on factors such as dataset size, computational resources, and the complexity of image features. The study contributes to the understanding of the applicability of these models in the domain of jellyfish species recognition, paving the way for informed decisions in deploying image recognition algorithms for marine biology research and conservation efforts.

REFERENCES

- [1] Fagerlund S. Bird species recognition using support vector machines. EURASIP Journal on Advances in Signal Processing. 2007 Dec; 2007:1-8.

- [2] Tattersall I. Species recognition in human paleontology. *Journal of Human Evolution*. 1986 Mar 1;15(3):165-75.
- [3] Chen G, Han TX, He Z, Kays R, Forrester T. Deep convolutional neural network-based species recognition for wild animal monitoring. In 2014 IEEE international conference on image processing (ICIP) 2014 Oct 27 (pp. 858-862). IEEE.
- [4] Fujita K. Species recognition by five macaque monkeys. *Primates*. 1987 Jul; 28:353-66.
- [5] Storbeck F, Daan B. Fish species recognition using computer vision and a neural network. *Fisheries Research*. 2001 Apr 1;51(1):11-5.
- [6] Somervuo P, Harma A, Fagerlund S. Parametric representations of bird sounds for automatic species recognition. *IEEE Transactions on Audio, Speech, and Language Processing*. 2006 Oct 16;14(6):2252-63.
- [7] Filho PL, Oliveira LS, Nisgoski S, Britto AS. Forest species recognition using macroscopic images. *Machine Vision and Applications*. 2014 May; 25:1019-31.
- [8] Pu R. Broadleaf species recognition with in situ hyperspectral data. *International Journal of Remote Sensing*. 2009 Jun 10;30(11):2759-79.