International Journal of Informatics and Communication Technology (IJ-ICT)
Vol. 14, No. 1, April 2025, pp. 240~249
ISSN: 2252-8776, DOI: 10.11591/ijict.v14i1.pp240-249

A hybrid approach of pattern recognition to detect marine animals

Vijayalakshmi Balachandran[1], Thanga Ramya Shanmugavel[2], Ramar Kadarkarayandi[3], Vijayalakshmi Kandhasamy[1]
[1]Department of Computer Science and Engineering, Ramco Institute of Technology, Rajapalayam, India
[2]Department of Computer Science and Design, RMK Engineering College, Chennai, India
[3]Department of Computer Science and Engineering, RMK College of Engineering and Technoloy, Chennai, India

Article Info

Article history:

Received Jul 11, 2024
Revised Oct 5, 2024
Accepted Nov 19, 2024

Keywords:

Classifier
Convolution neural network
Histogram of gradients
Marine animals
Pattern recognition
Support vector machine

ABSTRACT

Acquiring up-to-date knowledge about various animals will have a significant impact on effectively managing species within the ecosystem. Manually identifying animals and their traits continues to be a costly and time-consuming process. The development of a system using the most recent developments in computer vision machine learning was necessary to address the issues of detecting sharks and aquatic species in areas filled with surfers, rocks, and various other potential false positives. In the ocean most of the species are cold-blooded animals hence they cannot be tracked with thermal cameras. Ocean's dynamic environment affects simple techniques like color separation, intensity histograms, and optical flow. Hence a hybrid approach using convolutional neural network - support vector machine (CNN-SVM) classifier is proposed to perform the pattern recognition. A CNN is employed for feature extraction by using the histogram of gradients value. Subsequently, a SVM classifier is employed to identify and categorise marine species in the vicinity of the seacoast. This serves to notify individuals who engage in swimming activities in the ocean. The suggested model is evaluated against alternative machine learning approaches, and it achieves a superior accuracy of 95% compared to the others.

Corresponding Author:

Vijayalakshmi Balachandran
Department of Computer Science and Engineering, Ramco Institute of Technology
Rajapalayam, Tamilnadu, India
Email: bviji.akila@gmail.com

1. INTRODUCTION

Since the majority of marine organisms are cold-blooded, thermal cameras cannot track them. Sharks are crucial to the ecology because they keep the species lower down in the food chain alive and act as a gauge of the health of the ocean. In the oceans of the world, there are about 500 different species of shark. It comes in a range of colours and sizes, ranging from 39 feet (12 metres) to less than one metre (3 feet). The saltwater crocodiles have the strongest propensities to view people as prey among all crocodilians.

The two major marine animals that could pose a threat to swimmers are saltwater crocodiles and great white sharks. Worldwide, there have been numerous incidents of shark and saltwater crocodile attacks, many of which took place in shallow seas. The majority of assaults take place in seas close to shore, usually close to or between sandbars where sharks might get caught at low tide while feeding. Sloped areas are also possible places for attacks. Swimmers and surfers are the usual targets of the attacks, which mainly take place in the surf zone.

People who swim in the sea close to the coast are seriously threatened by both sharks and saltwater crocodiles. In order to identify these marine species using high resolution cameras and warn swimmers as they approach the coast, a pattern recognition system must be developed. The following tasks must be completed in order to spot sharks and crocodiles. Finding the suspicious shape near the coast and determining whether or not it is a harmful sea species. To save the swimmers' lives, it is important to warn them if there are sharks or crocodiles present.

The following factors make it harder to discern patterns and find marine organisms in the maritime environment:

- Use a high-resolution camera in a drone to identify sharks and crocodiles from a minimum of 60 metres above the ground.
- The ocean's incredibly dynamic environment caused by the waves, glares, and shadows.

Automatic marine animal identification may be carried out with ease thanks to the capabilities of machine learning techniques and increased processing power. This can open up a vast new array of opportunities for developing specialized warning system and data collection scenarios. By doing this, the manual work and associated expenses are decreased, and the marine ecosystems is preserved. The use of computer vision tools to research and track the marine eco system has grown in recent years. The main uses involve object detection, which is helpful for tracking fish populations [1], [2], detecting seals [3], locating whale [4] hotspots, and coral reef fish populations [5].

The unprovoked shark attacks can be avoided by using the power of artificial intelligence (AI) and machine learning (ML) algorithms to detect the occurrence of sharks in the seashore. The sharkspotter [6] classifies the detected object into one of the 16 categories, if it identifies a shark the alert generated and the swimmers can be warned. Proposed a method for detecting the dolphins and shark using the shape feature properties from the aerial images. The methodology based on the two-dimensional deformable model [7] in which the reference variable was optimized to produce low error rate and high accuracy. A deep learning algorithm was presented to identify White sharks in underwater environments, providing assistance to divers and other aficionados of underwater sports. The study used a YOLOv3 algorithm [8] which utilises convolutional neural networks (CNNs) to recognise objects, make predictions at many scales, and forecast bounding boxes using logistic regression. They tackled the concerns related to the undersea environment when identifying the species.

Deep CNN based object detection model [9] was designed to detect shark and other marine animals. It also performs the region segmentation task for shark detection which was done by fast RCNN along with VGG16 architecture that improves the accuracy ratio. Transfer learning and CNN based model [10] designed for classification of sharks. A shark alerting system [11] in which the shark detection was done using the deep neural network-based YOLO algorithm was used. It can also detect other several distinct objects (sharks, rays, surfers, paddle boarders). This alerting system is trained based on a single beach location which may perform well in that particular location and need to improve the model with by training with other location data. Proposed two algorithms for saltwater crocodile detection during daytime using multi-feature joint descriptor [12]. The first algorithm uses features color and HoG. Another algorithm makes use of color and scale-invariant feature transform feature descriptors to identify the target. An approach [13] for detecting the captured images of dugongs using a pattern recognition algorithm that make use the features form the captured image. The initial method for blob detection combines morphological procedures with color analysis. The second method employs a shape profiling methodology along with the saturation channel from the HSV color space to identify mammal. There is a need to deal with the water turbidity, wind, wave speed and period, and sunlight while doing the automatic detection. A multispectral imaging-based method [14] was proposed to find out the great white sharks in the Pacific Ocean off coast of San Diego.

The weakly supervised method [15] to detect the marine animals from the aerial images proposed which does not require to spend much time on the annotation. A multistep pipeline was proposed to map the anomaly maps to the relevant region proposal from where the objects can be detected using the patch distribution modelling method. The CNN network model [16] is proposed for the automated counting and identification of fish species. Presented a novel deep neural network model [17] for the detection of fish in underwater environments and conducted a comparative analysis using the support vector machines (SVM) classifier. A method for quickly and accurately detecting and recognising fish in underwater photographs using the fast R-CNN algorithm was proposed in [18].

Deep learning's explosive growth has recently encouraged significant theoretical advancements and real-world applications of computer-vision-based underwater object recognition methods. In this the R-CNN and mask RCNN approach [19], deep neural network model MobileNetV1 [20] was used to detect the underwater objects and marine creatures. Models in [21], [22], demonstrated a hybrid approach for real time fish species and dolphin species identification. SVM based approaches were used in various prediction task [23]. An overview of several swarm intelligence methods and their applications in image processing was presented in [24]. This provided an opportunity to identify and choose the essential optimization algorithm for the classification task.

In summary, despite the recent rise in shark attacks worldwide, relatively few efforts have been made to automate the detection procedure given the seriousness of the consequence. The majority of publications in the literature employed a standard machine learning approach, were inefficient, and had low precision. In this paper, a hybrid approach for pattern recognition and detection using histogram of gradients (HoG), CNN, SVM algorithms are proposed to find the pattern in the marine animals such as sharks and crocodiles and identify the animals reaching the sea shore and warn the swimmers and surfers near the coastal region.

2. RESEARCH METHOD

In this work a hybrid model to identify and classify the marine animal such as shark and crocodile is shown in Figure 1. The proposed model make use of the best of the functionalities from the CNN and SVM. The HoG feature extractor extracts the histogram gradients from the image and CNN supports the feature extraction and the optimized SVM performs the classification task.

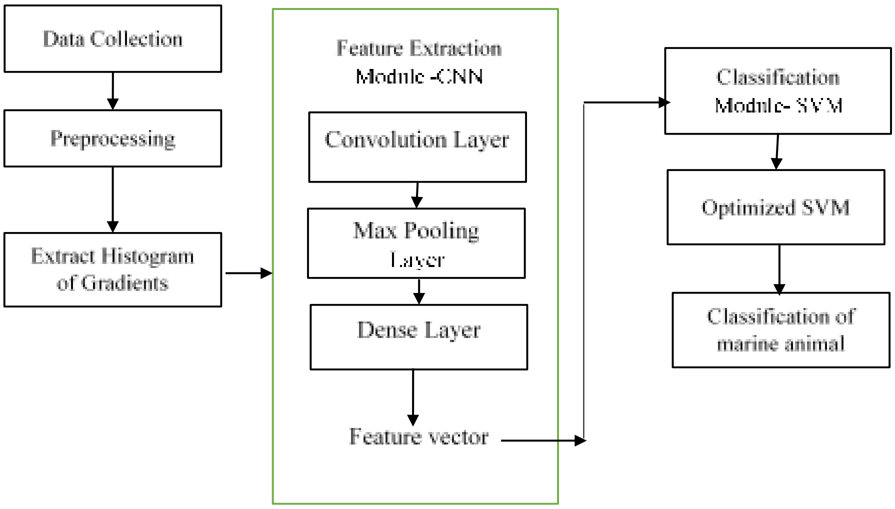

Figure 1. Pattern recognition model

2.1. Data set

The data's relevant to the training and the testing of the model was collected from various sources [25]-[27]. The image data set consists of saltwater crocodile and shark images each of 700 images. Data augmentation techniques are employed to expand the dataset, resulting in a total of 1,200 photos for each category. The dataset is partitioned into three subsets: train, test, and validation. The distribution is as follows: 70% of the data is allocated for training the model, 10% is allocated for model validation, and the remaining 20% is allocated for model testing.

2.2. Preprocessing

The standard data augmentation techniques such as flipping, rotation, shear and brightness changes are applied to the actual dataset to increase the size of the dataset. The images are received from the various sources hence to provide uniformity among the data, all the images are resized to the ratio of 1:2 (width:height). The images are converted into grey scale images before extracting the features from the images.

2.3. Feature extraction

2.3.1. Histogram of gradients feature

Using the HoG feature descriptor, one can extract features from an image to facilitate object detection. It emphasises the form and composition of an object. It extracts the gradient and orientation of the edges and these are calculated in localized portions. For each of these regions a histogram was generated which provide the important features in the image.

Calculating the gradients of a picture exposes the specific areas where the intensity of pixel gradients undergoes alterations. The kernels utilised for computing the gradients consist of a vertical gradient kernel: [-1, 0, 1] and a horizontal gradient kernel: [-1]. The gradient of the picture is determined by combining the vertical gradient (Gx) and the horizontal gradient (Gy), using the magnitude and angle obtained from the image, as specified in (1) and (2).

$$Magnitude = \sqrt{G\,{}_x^2 + G_y^2} \tag{1}$$

$$Angle(\theta) = \left| tan^{-1}(\frac{G_y}{G_x}) \right| \tag{2}$$

After the gradient is calculated a histogram of 9 bins was formed to bin the gradient magnitude values w.r.t gradient direction. Gradient levels might vary depending on lighting and foreground/background contrast hence block normalization is applied. Finally, the histogram is normalized to form the feature vector which represent the concise and succinct representation of particular patches of the images. The pictures that will be learned are converted into an array of HoG features. The chosen features may encompass the crucial information from the input data, so enabling the accomplishment of the intended job utilising a smaller quantity of data instead of the original unmodified data.

2.4. Convolutional neural network

The CNN is employed in tasks involving picture classification and detection. This is because the convolution layer of the CNN effectively reduces the high dimensionality of the data. CNN functions as a feature extractor throughout the training phase. The feature extractor employed by CNN consists of distinct neural network architectures, with the weights being established by the training process. Despite the added complexity of the learning approach, CNN provides enhanced picture recognition by using a deeper neural network feature extraction with more layers. CNN is a computational model designed to extract relevant characteristics from input images. The feature extraction network utilises the input picture as its first reference.

The CNN model’s feature extraction network is composed of a series of convolutional layers and pooling layers stacked together. The convolution layer performs the convolution operation using filters and transforms the input image. Convolution is a process in which as kernel or filter is moved over the image and transformation takes place based on the values in the filter. Feature maps are calculated based on (3):

$$G[m,n] = (f * h)[m,n] = \sum_j \sum_k h[j,k] f[m-j, n-k] \tag{3}$$

where f denotes the input image, h indicates the kernel, m and n indicates the rows and columns of the result matrix.

The pooling layer of the neural network retains the fine details in the input and enhances the process of reducing dimensions by calculating the average or lowest value, depending on the specific requirements of the application. Ultimately, a fully connected layer is added to the final layer in order to acquire knowledge from the output of the preceding layer and carry out the desired operation. The convolution and pooling layers in a CNN inherently operate on a two-dimensional plane due to the primary emphasis of CNNs on image processing.

2.5. Support vector machines

SVM can effectively handle both continuous and categorical data. SVM constructs a hyperplane in a high-dimensional space to separate different classes. The SVM constructs an optimal hyperplane through an iterative process in order to minimise error. The primary objective of SVM is to determine the maximum marginal hyperplane (MMH) that effectively divides the dataset into several groups. SVM used a hyperplane as a decision boundary to differentiate or classify between two groups. SVM is commonly employed in the tasks of image classification and object detection.

The data points that are in closest proximity to the hyperplane are referred to as support vectors. By calculating the margins, these points will more accurately delineate the dividing line. These notions are of greater importance in the construction of the classifier.

A hyperplane is a plane used to make decisions by separating a group of objects with distinct class memberships. It is defined by the set of points x that satisfies (4).

$$w.x + b = 0 \tag{4}$$

Any point x that does not satisfy the () equation must be located on one of the two sides of the hyperplane. Choose two additional hyperplanes *(H_1, H_2)* that are equally spaced from H_0, making sure that each data point is on the correct side with no data in between. Therefore, it is possible to define the hyperplanes *(H_1, H_2)* using (5) and (6).

$$(H1) \qquad w.x_i + b \geq 1 \qquad if\ y_i = 1 \tag{5}$$

$$(H2) \qquad w.x_i + b \leq 1 \qquad if\ y_i = -1 \tag{6}$$

These two hyperplanes can be combined into a single equation for a separating hyperplane using (7).

$$y_i(w.x_i + b) \geq 1 \qquad \forall_i \in \{1,2, \ldots., n\} \tag{7}$$

The optimal separating hyperplane is the solution that is farthest away from the closest data point or in other terms maximizes the margin. It can also be imagined as a parallel line cutting the margin in two halves. The main objective function is defined by (8).

$$J(w, b) = \lambda \frac{1}{2} \|w\|^2 + \frac{1}{n} \sum_{i=1}^{n} \max(0, 1 - y_i(w.x_i + b)) \tag{8}$$

The first term is basically responsible for maximizing the margin, expressed as a minimization problem with an added regularization parameter λ. The second term is the definition of a separating hyperplane is a loss function called the Hinge loss. This term is responsible for ensuring that the model predict the correct class label with enough margin.

For example, if $y_i = 1$ and x_i is correctly classified, calculating the hinge loss will result in zero since *max (0, 1–1) = 0*. However, if the class label is falsely predicted the hinge loss will result in a value greater than zero. The loss function can be optimized with gradient descent by making small steps in the opposite direction to minimize the loss.

SVM is a maximal margin classifier which is basically performing better on linearly separable data. The classification of marine animal such as shark and saltwater crocodiles can be performed using SVM since the data set is linearly separable that maximizes the decision boundary between the two classes by finding the optimal separating hyperplane. The SVM have been trained and tested with the HoG feature extracted from the image datasets and support vector classifier (SVC) is used in the SVM model. Here the classification works based on the linear data and it based on the margin classifier.

2.6. Particle swarm optimization

The Particle swarm optimization (PSO) algorithm is a computer method that draws its inspiration from the cooperative movement of natural species, including fish and birds, to accomplish a common objective. In Algorithm 1 explains the PSO involves searching across the solution space of a problem with a group of particles (representing potential solutions). Based on its own best-known answer (personal best) and the best solution found by the entire group (global best), each particle modifies its position. Particles can converge over iterations towards ideal solutions thanks to this cooperative movement.

Algorithm 1. The PSO

```
Initialize random number of particles
For each particle
Do
   Initialize particle position xi
   Initialize the initial position xi as the best known position pbesti
   Update the swarms best position gbest = pbesti, if fitness(pbesti) <fitness(gbest)
   Initialize the particles velocity as vi
Repeat until a termination criterion is met:
        Update the particle velocity
        Update the particles position
        Calculate
                Update particles best known position as pbesti=xi, if
fitness(xi)<fitness(pbesti)
                Update the swarms best known position as gbest=pbesti, if
fitness(pbesti)<fgbest)
Return the best particle of swarm
```

where x_i and v_i denotes the position and velocity of particle i. Fitness(x_i) denotes the fitness function to find the particle maximum fitness value. The parameters of the SVM model is optimized for the image classification task.

2.6.1. PSO-SVM optimization

Table 1 lists the parameters picked for optimization together with the values of its search space. The parameters used to construct the particles for optimization, each of which represents a potential solution, are used here to initialize the algorithm. In this case, a particle is specified with three positions, each of which corresponds to an optimization target parameter.

The following describes how SVM is optimized using PSO.

a) Provide the data to the SVM as training input.
b) Create the population and set the PSO parameters to their initial values.
c) Based on the PSO algorithm's results, define the SVM model's parameters.
d) The SVM model should be trained and evaluated to determine accuracy, which is the PSO algorithm's goal.
e) The accuracy value, which is the goal function for the PSO algorithm, is obtained after the SVM model has been trained and evaluated.
f) PSO assesses the objective function to determine the optimum optimal.
g) Update the PSO parameters and continue doing so until the maximum iteration is reached.
h) Provide the particle with the Gbest value, which is the SVM model's optimal value.

Table 1. Search space for PSO

Particle	Parameter	Search space
X_1	Kernal	["linear", "rbf", "poly"]
X_2	Kernel coefficient - gamma	['scale', 'auto']
X_3	Penalty parameter C	[0.1,100]

The SVM model optimized with PSO algorithm found the best parameters for classification. An RBF function is used as a kernel in the SVM model. The parameters gamma and C are chosen as 'scale' and 10. This model uses the feature vectors from the CNN model as input and perform the classification of the image.

2.7. Evaluation metrics

The marine animal classification algorithm performance can be measured with the following performance metrics given in (9)-(11). The percentage of accurate forecasts is called accuracy. Recall, also known as sensitivity, refers to the proportion of true positives, whereas precision is the percentage of actual positives among all the values projected as positive. The F1-Score is a metric that quantifies the performance of a classifier by taking the harmonic mean of its accuracy and recall and calculated using (12).

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} \tag{9}$$

$$Precision = \frac{(TP)}{(TP+FP)} \tag{10}$$

$$Recall = \frac{(TP)}{(TP+FN)} \tag{11}$$

$$F1 - score = \frac{(2*Precision*Recall)}{(Precision+Recall)} \tag{12}$$

where TP, TN, FP, and FN indicates the true positive (TP), true negative (TN), false positive (FP), and false negative (FN) respectively.

3. RESULTS AND DISCUSSION

3.1. Preprocessing

The actual or the original size image is shown in Figure 2. The actual dimensions of the image may differ. There is a need to resize the image into the required criteria in which the width is to height to be in the format of 1:2 which resembles that the size of the images can be in dimensions of (64×128) or (128×256). The resized and grayscale converted image is shown in Figure 3. The gradient and angle of the actual image are shown in Figures 4 and 5. Figure 6 displays the HoG picture of the provided scaled image. The HoG features that have been retrieved are utilised as input for the CNN model to do additional feature extraction.

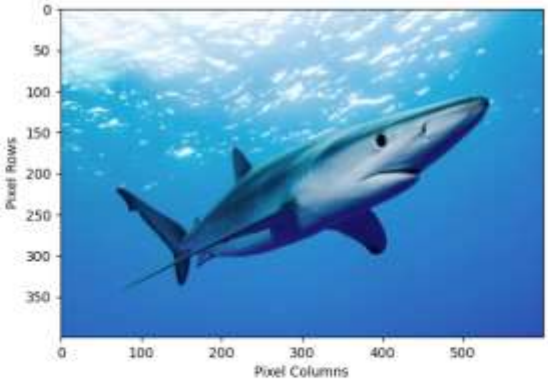

Figure 2. Actual image

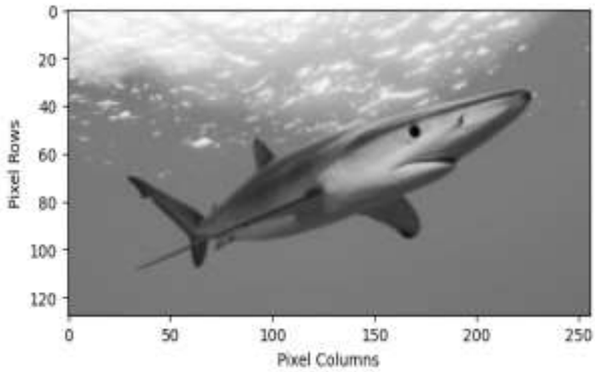

Figure 3. Resized gray scale image

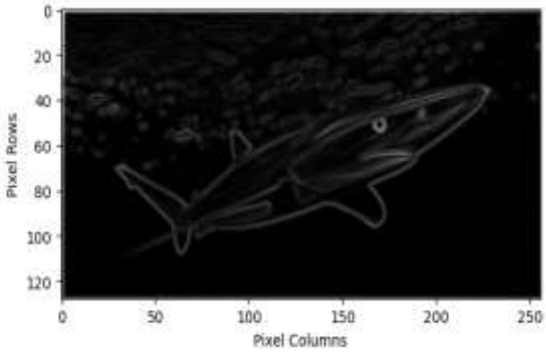

Figure 4. Gradient

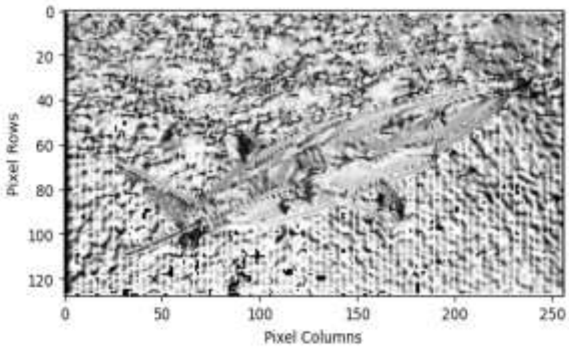

Figure 5. Angle

Figure 6. HoG

3.2. CNN model parameter

The CNN model is designed with 4 convolution layers with the number of neurons of {32, 64, 128, 32}, the kernel size of 3*3 and the activation function of relu. Next, the max pooling layer and the flatten layer are added. Followed by the dense layer with 256,512 neurons added to the model. The model is complied with the adam optimizer and the loss function is of binary_crossentry. The outcome of the CNN model is the feature vector.

3.3. Results

The model is evaluated and its performance can be improved by using the K- fold cross validation techniques. This validation technique is applied to avoid the overfitting of the model by reducing the variance of the performance estimate and it provides an opportunity to the model to learn with more amount of data. The comparison of the proposed HoG based CNN_SVM model with various machine learning models without feature extraction is given in Table 2.

Table 3 provides a detailed comparison of model performance using different configurations: HoG+SVM, HoG+CNN, and the proposed HoG+CNN+SVM model. The performance metrics include accuracy, precision, recall, and F1-score. The proposed model, HoG+CNN+SVM, shows superior performance across all metrics, with an accuracy of 95.7%, precision of 0.93, recall of 0.94, and F1-score of 0.94. This suggests that the integration of CNN and SVM with HoG features significantly enhances the model's effectiveness. In contrast, the HoG+CNN model achieves an accuracy of 91%, and the HoG+SVM model has the lowest accuracy at 89.85%. These results underscore the improved performance and robustness of the proposed HoG+CNN+SVM model for the given task.

Table 2. Comparison of the model without feature extraction

Model	Accuracy	Precision	Recall	F1-score
DT	72.2	0.62	0.72	0.69
SVM	76.80	0.66	0.75	0.70
KNN	79.72	0.69	0.79	0.72
CNN	88.80	0.74	0.84	0.77
HoG+CNN+SVM (proposed model)	95.7	0.93	0.94	0.94

Table 3. Comparison of model performance with other models

Model	Accuracy	Precision	Recall	F1-score
HoG+ SVM	89.85	0.88	0.89	0.88
HoG+CNN	91	0.89	0.90	0.89
HoG+CNN+SVM (proposed model)	95.7	0.93	0.94	0.94

Figure 7 compares three models' performance – HoG+SVM, HoG+CNN, and HoG+CNN+SVM – on four evaluation metrics: accuracy, precision, recall, and F1-score.

a) HoG+SVM: in this model, HoG is used for feature extraction while SVM is employed for classification It shows moderately high performance across all metrics, with values generally ranging around the mid-80% mark.
b) HoG+CNN: this model employs CNN along with HoG features. It performs a little better with all measures than HoG+SVM model whose performances are constantly just over 85% in all these indicators.
c) HoG+CNN+SVM (proposed model): this proposed model integrates the features of HoG, CNN, and SVM which help to outperform the other two models across all metrics, with performance percentages nearing or reaching 90%.

This shows that the proposed HoG_CNN_SVM model performs better with an accuracy of 95.7% for the classification of marine animals. This shows that feature extraction is a crucial component of the pattern recognition process for underwater photos.

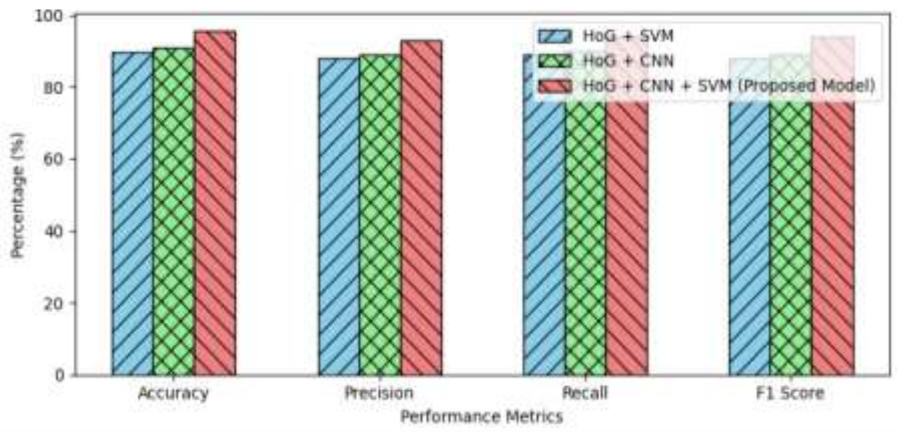

Figure 7. Performance comparison

4. CONCLUSION

A streamlined, automated, and instantaneous surveillance method is vital for identifying diverse entities (such as human activities, sizable fish, sharks, whales, and surfers) on beaches, in order to prevent unforeseen deaths and mishaps. This study introduces a feature extractor based on deep learning, which is combined with a machine learning classifier. The purpose is to automatically identify patterns and categorise marine species. This approach aims to minimise the need for human intervention and lower associated costs. The CNN model extracts the essential and significant HoG information from the image. These characteristics were provided as input to the SVM classifier in order to categorise the marine species found along the shoreline, which might potentially impact those swimming in the water. The results obtained from the suggested method demonstrate enhanced accuracy of 95% in comparison to the alternative machine learning methodology that does not involve feature extraction. Therefore, it is evident that feature extraction plays a crucial role in predicting marine creatures.

REFERENCES

[1] E. M. Ditria, M. Sievers, S. Lopez-Marcano, E. L. Jinks, and R. M. Connolly, "Deep learning for automated analysis of fish abundance: the benefits of training across multiple habitats," *Environmental Monitoring and Assessment*, vol. 192, no. 11, Oct. 2020, doi: 10.1007/s10661-020-08653-z.

[2] A. Jalal, A. Salman, A. Mian, M. Shortis, and F. Shafait, "Fish detection and species classification in underwater environments using deep learning with temporal information," *Ecological Informatics*, vol. 57, p. 101088, May 2020, doi: 10.1016/j.ecoinf.2020.101088.

[3] A. B. Salberg, "Detection of seals in remote sensing images using features extracted from deep convolutional neural networks," in *International Geoscience and Remote Sensing Symposium (IGARSS)*, Jul. 2015, vol. 2015-November, pp. 1893–1896, doi: 10.1109/IGARSS.2015.7326163.

[4] E. Guirado, S. Tabik, M. L. Rivas, D. Alcaraz-Segura, and F. Herrera, "Whale counting in satellite and aerial images with deep learning," *Scientific Reports*, vol. 9, no. 1, Oct. 2019, doi: 10.1038/s41598-019-50795-9.

[5] M. Duarte, J. Aguzzi, and E. Fanelli, "Automated video monitoring of coral reef fish populations," *Marine Ecology Progress Series*, vol. 681, pp. 105–119, 2022.

[6] N. Sharma, M. Saqib, P. Scully-Power, and M. Blumenstein, "SharkSpotter: shark detection with drones for human safety and environmental protection," in *Humanity Driven AI*, Springer International Publishing, 2022, pp. 223–237.

[7] S. Gururatsakul, D. Gibbins, and D. Kearney, "A simple deformable model for shark recognition," in *Proceedings - 2011 Canadian Conference on Computer and Robot Vision, CRV 2011*, May 2011, pp. 234–241, doi: 10.1109/CRV.2011.38.

[8] N. E. Merencilla, A. Sarraga Alon, G. J. O. Fernando, E. M. Cepe, and D. C. Malunao, "Shark-EYE: a deep inference convolutional neural network of shark detection for underwater diving surveillance," in *Proceedings of 2nd IEEE International Conference on Computational Intelligence and Knowledge Economy, ICCIKE 2021*, Mar. 2021, pp. 384–388, doi: 10.1109/ICCIKE51210.2021.9410715.

[9] N. Sharma, P. Scully-Power, and M. Blumenstein, "Shark detection from aerial imagery using region-based CNN, a study," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11320 LNAI, Springer International Publishing, 2018, pp. 224–236.

[10] J. Jenrette, Z. Y. C. Liu, P. Chimote, T. Hastie, E. Fox, and F. Ferretti, "Shark detection and classification with machine learning," *Ecological Informatics*, vol. 69, p. 101673, Jul. 2022, doi: 10.1016/j.ecoinf.2022.101673.

[11] R. Gorkin *et al.*, "Sharkeye: Real-time autonomous personal shark alerting via aerial surveillance," *Drones*, vol. 4, no. 2, pp. 1–17, May 2020, doi: 10.3390/drones4020018.

[12] M. Hemalatha, M. A. Muthiah, and B. Venkatalaksmi, "Multi-feature joint descriptor based image detection algorithm for crocodile detection," in *Proceedings of 2016 International Conference on Advanced Communication Control and Computing Technologies, ICACCCT 2016*, May 2017, pp. 527–530, doi: 10.1109/ICACCCT.2016.7831696.

[13] L. Mejias, G. Duclos, A. Hodgson, and F. Maire, *Automated marine mammal detection from aerial imagery*. To appear in MTS/IEEE OCEANS, San Diego, USA, 2013.

[14] J. Lopez, J. Schoonmaker, and S. Saggese, "Automated detection of marine animals using multispectral imaging," in *2014 Oceans - St. John's, OCEANS 2014*, Sep. 2015, pp. 1–6, doi: 10.1109/OCEANS.2014.7003132.

[15] P. Berg, D. S. Maia, M. T. Pham, and S. Lefèvre, "Weakly supervised detection of marine animals in high resolution aerial images," *Remote Sensing*, vol. 14, no. 2, p. 339, Jan. 2022, doi: 10.3390/rs14020339.

[16] A. Salman *et al.*, "Fish species classification in unconstrained underwater environments based on deep learning," *Limnology and Oceanography: Methods*, vol. 14, no. 9, pp. 570–585, May 2016, doi: 10.1002/lom3.10113.

[17] S. Villon, M. Chaumont, G. Subsol, S. Villéger, T. Claverie, and D. Mouillot, "Coral reef fish detection and recognition in underwater videos by supervised machine learning: Comparison between deep learning and HOG+SVM methods," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 10016 LNCS, Springer International Publishing, 2016, pp. 160–171.

[18] X. Li, M. Shang, H. Qin, and L. Chen, "Fast accurate fish detection and recognition of underwater images with Fast R-CNN," Oct. 2016, doi: 10.23919/oceans.2015.7404464.

[19] M. Joo Er, J. Chen, and Y. Zhang, "Marine robotics 4.0: present and future of real-time detection techniques for underwater objects," in *Industry 4.0 - Perspectives and Applications*, IntechOpen, 2023.

[20] O. Ittoo and S. Pudaruth, "Automatic recognition of marine creatures using deep learning," *International Journal of Advanced Computer Science and Applications*, vol. 15, no. 1, pp. 47–64, 2024, doi: 10.14569/IJACSA.2024.0150106.

[21] N. Shahnaz, M. M. Islam, M. A. Sarker, and M. Rahman, "Hybrid deep learning approach for marine mammal detection using acoustic data," *Journal of Marine Science and Engineering*, vol. 10, no. 2, p. 123, 2022.

[22] X. Zhang, J. W. Y. Li, and H. Xu, "A hybrid deep learning model for real-time fish species recognition in underwater videos," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 4512–4524, 2023.

[23] R. Venkatesh, S. Anantharajan, and S. Gunasekaran, "Multi-gradient boosted adaptive SVM-based prediction of heart disease," *International Journal of Computers, Communications and Control*, vol. 18, no. 5, Aug. 2023, doi: 10.15837/ijccc.2023.5.4994.

[24] M. Xu, L. Cao, D. Lu, Z. Hu, and Y. Yue, "Application of swarm intelligence optimization algorithms in image processing: a comprehensive review of analysis, synthesis, and optimization," *Biomimetics*, vol. 8, no. 2, p. 235, Jun. 2023, doi: 10.3390/biomimetics8020235.
[25] "Shark Species," *Kaggle*. https://www.kaggle.com/datasets/larusso94/shark-species (accessed Feb. 14, 2024).
[26] P. L. Falkingham and J. Milan, "Crocodile tracks and trackways: Photogrammetric models and photo datasets," *Zenodo, Dataset*, 2015. https://doi.org/10.5281/zenodo.31711.
[27] "Crocodile," *Images.cv*. https://images.cv/category/crocodile (accessed Mar. 01, 2024).

BIOGRAPHIES OF AUTHORS

Vijayalakshmi Balachandran is working as an assistant professor (SG) in the Department of Computer Science and Engineering, Ramco Institute of Technology, Rajapalayam, India. She has completed her B.E. and M.E. Degree in Computer Science and Engineering from PSR Engineering College, Anna University. Her areas of interest are image processing, data science and deep learning. She has published 13 papers in journal and international conferences. She is a life member of ISTE. She can be contacted at email: bviji.akila@gmail.com.

Thanga Ramya Shanmugavel B.E, M.S. (by Res), Ph.D., is working as an associate professor in the Department of Computer Science and Engineering, R.M.K. Engineering College, Chennai. She obtained her B.E. (CSE) from Dr. Sivanthi Aditanar College of Engineering and M.S. by Research (ICE) from Anna University, Chennai. She has obtained her Ph.D. in Information and Communication Engineering from Anna University, Chennai, in 2017. She has been in the teaching profession for the past 20 years and published more than 30 papers in various International Journals and Conference. Her areas of interest include programming languages, database management, cloud computing and data mining. She has got oracle java international certification, IBM RAD, TIVOLI, and DB2 certification and obtained her PRP certification from Wipro. She is the life member of ISTE. She can be contacted at email: thangaramya@gmail.com.

Ramar Kadarkarayandi currently working as a principal in, R.M.K. College of Engineering and Technology, Chennai. He obtained his bachelor's degree in Electronics and Communication Engineering from Government College of Engineering, Tirunelveli and Post Graduation in Applied Electronics from PSG College of Technology, Coimbatore. He obtained his Doctoral Degree in Computer Science Engineering from Manonmaniaum Sundaranar University, Tirunelveli in 2001. He has 34 years of teaching experience including 29 years of Research experience. His extensive research experience has produced 46 Ph.D. so far and 5 Scholars are pursing Ph.D. He has published 132 research articles in various international journals and 91 of them are Scopus Indexed. He can be contacted at email: kramar.einstein@gmail.com.

Vijayalakshmi Kandhasamy received her bachelor's degree and a master's in computer science and engineering from Madurai Kamaraj University, Madurai, Tamilnadu, India. She obtained Ph.D. degree in Computer Science and Engineering from Anna University, Chennai, Tamilnadu, India, in 2008. Her research interests include artfcial intelligence, networks, big data analytics, and optimization. She is currently working as a professor and head in the Department of Computer Science and Engineering at Ramco Institute of Technology, Rajapalayam, Tamil Nadu, India. She has published 30 articles in peer reviewed journals and presented papers in more than 70 national and international conferences. Totally eight scholars completed their Ph.D. under her supervision in various domains of Computer Science and Engineering. She can be contacted at email: vijayalakshmik@ritrjpm.ac.in.