Deep learning and machine learning classification technique for integrated forecasting

Vigilson Prem Monickaraj¹, Sterlin Rani Devakadacham², Nithyadevi Shanmugam³, Nithya Nandhakumar⁴, Manjunathan Alagarsamy⁵, Kannadhasan Suriyan⁶

¹Department of Computer Science and Engineering, R.M.K. College of Engineering and Technology, Tamil Nadu, India
²Department of Computer Science and Engineering, K. Ramakrishnan College of Engineering, Tamil Nadu, India
³Department of Electronics and Communication Engineering, Dr.N.G.P. Institute of Technology, Tamil Nadu, India
⁴Department of Computer Science and Engineering, K. Ramakrishnan College of Engineering, Tamil Nadu, India
⁵Department of Electronics and Communication Engineering, R.M.D. Engineering College, Tamil Nadu, India
⁶Department of Electronics and Communication Engineering, Study World College of Engineering, Tamil Nadu, India

ABSTRACT

Smart fisheries are increasingly using artificial intelligence (AI) technologies to increase their sustainability. The potential fishing zone (PFZ) forecasts several fish aggregation zones throughout the duration of the prediction in any sea. The autoregressive integrated moving average (ARIMA) and random forest model are used in the current study to provide a technique for locating viable fishing zones in deep marine seas. A significant amount of data was gathered for the database's creation, including monitoring information for Indian fishing fleets from 2017 to 2019. Using expert label datasets for validation, it was discovered that the model's detection accuracy was 98%. Our method uses salinity and dissolved oxygen, two crucial markers of water quality, to identify suitable fishing zones for the first time. In the current research, a system was created to identify and map the quantity of fishing activity. The tests use a number of parameter measurements to evaluate the contrast-enhanced computed tomography (CECT) approach to machine learning (ML) and deep learning (DL) methodologies. The findings showed that the CECT had a 94% accuracy rate compared to a convolutional neural network's 92% accuracy rate for the 80% training data and 20% testing data.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Vigilson Prem Monickaraj
R.M.K. College of Engineering and Technology
Chennai, Tamil Nadu, India
Email: vigiprem@gmail.com

1. INTRODUCTION

Many parties with interest in the issue are turning to artificial intelligence-based smart fisheries to help alleviate the issue of declining fish populations [1]. Since 2018, the United Nations (UN), the European Unions (EU), and several state governments have proclaimed a purportedly new "AI era" [2]. Since 2017, the UN has had an artificial intelligence programme for global governance. In order to make sure that the Sustainable Development Goals (SDGs) benefit everyone and promote the SDGs, AI was utilised to evaluate the SDGs. Because of its superior resources, practical living places, and rich biodiversity, the coastal marine environment is essential to India's economy. India's exclusive economic zone (EEZ), which includes islands and extends 7517 km of coastline, is a key area for research and the utilisation of shared resources. Its total area is 2.5 million km². The marine fishing sector employs about 14 million people and generates revenue by exporting to untapped markets. Despite having a harvestable potential of 3.93 million tonnes, India produces...
around 2.94 million tonnes of marine fisheries annually [3], [4]. Finding the best fishing spots still presents a challenge for anglers.

Both the maritime characteristics visible in satellite photos and data from ground truth sources have been used to forecast the fishing zone. However, the majority of study frameworks made predictions using aspects of oceanography, such as chlorophyll and sea surface temperature (SST). Machine and deep learning techniques have already helped fishing zone prediction reach new heights. Long short-term memory (LSTM), Markov models, Nave Bayes (NB) classifiers, support vector machines (SVM), and deep neural networks (DNN) are some of the methods used to estimate the location of fisheries based on various oceanographic variables. Less promising is that we still lack a reliable prognosis for Packet Filter (PF), which is a result of the study. In this work, a brand-new hybrid model called hybrid ensemble DEEPFISHNETS (HE-DFNETS) is suggested to address the aforementioned issue. The flitter bat optimised LSTM (FOLSTM) and double tier convolutional neural networks (DTCN) are combined for an effective prediction of PFZs (PEZ) utilising remote sensing pictures [5]–[8]. The results show how much better the RF model is than the GBDT model and how well it can recover the SSA. Furthermore, the accuracy of the two models often decreases at depths below 500 m.

Using SST data from satellites, SSHA data from sea surface height anomalies, and OHC700 climate data, the artificial neural network (ANN) approach has been applied to assess OHC700 in the Indian Ocean. The results suggest that the ANN method may be useful for accurately investigating OHC700 on a practically continuous basis. This interaction seeks to anticipate the sensor simplicity of chlorophyll-A (C.A.) in the upper expanse of the Arctic Ocean (AO) by assessing data from upper sea perception floats placed in the AO. It is suggested to use a model that combines stacked auto encoder (SAE), bidirectional LSTM, and wavelet change. These frameworks provide superior findings when comparing the mean absolute error (MAE) and root mean square error (RMSE) of the experiment [9]–[12].

By looking at restrictions that specify the appropriateness of frame speed increases. It is shown how ice condition may be dispersed among several pretrained conditions using a Kullback-Leibler disparity metric. The study demonstrates that factual order approaches developed using measurement data provide more precise and reliable results [13]–[18]. Automatic information systems (S-AIS), which have been offered as a revolutionary way to follow the movements of fishing fleets in almost real time, are typically installed on the majority of ocean-going boats nowadays. It is shown that PFZ may be predicted accurately utilising a data mining (DM) strategy and an algorithm developed via study of water quality indicators. Finding PFZs requires the use of two water quality indicators: salinity and dissolved oxygen [19]–[22]. For purse seiners, a multi-layered filtration system was created based on vessel speed and operating time. When measured against expert-labeled datasets, the average detection accuracies for trawler and longline fishing were 83% and 97%, respectively. Our research is the only all-inclusive strategy currently available for detecting and classifying probable fishing activity utilising the three main gear types used around the world [23]–[25].

2. RELATED WORKS

South West Atlantic digital echo-recording techniques for ANN training and testing: automated credit and categorization in the classroom. The ANNs were created to extract energetic school and bathymetric parameters from echo recordings. The best result was achieved by changing the input values for species. Correct rates may reach 96% depending on the species, network design, and quantity of schools utilised. It was recommended to do further study into this strategy as a practical tool for echogram analysis. All aquatic creatures that depend on dissolved oxygen (DO), a critical indicator of water quality, must have it to survive. The amount of DO present in water bodies in catchments may be influenced by both man-made activities and natural occurrences. The quantity of dissolved oxygen in waterbodies is affected by a number of factors, including water temperature, animal respiration and decomposition rates, the amount of oxygen replenished by photosynthesizing plants, stream flow, and aeration. Water temperature significantly affects the quantity of DO because warm water dissolves less oxygen than cold water. A number of sensors deployed on satellites and other platforms, such aeroplanes, measure the quantity of radiation reflecting off the water's surface at different wavelengths. The total suspended solids (TSS), chlorophyll a concentration, turbidity, salinity, total phosphorus (TP), Secchi disc depth (SDD), temperature, pH, and dissolved organic carbon (DOC) are just a few examples of water quality indicators that can be directly or indirectly determined using these reflections. In order to monitor and evaluate the water quality, it is crucial to include contaminants that are caused by the hydrological, biological, and chemical components of the water, among other factors. The research thus covers the commonly used airborne and spaceborne sensors for evaluating the quality of water and evaluates the utility of remotely sensed approaches for evaluating the quality of waterbodies. The characteristics of airborne and spaceborne sensors, including their spectrum, spatial, and temporal capabilities, are summarised to serve as a sensor selection guide. The research also looks at eleven water
quality measures in more depth based on the methodologies employed to quantify their concentrations. Based on a review of the literature, the study offers a range of sensors that may be used to investigate certain measurable water quality measures. All aquatic creatures that depend on dissolved oxygen (DO), a critical indicator of water quality, must have it to survive. The amount of DO present in water bodies in catchments may be influenced by both man-made activities and natural occurrences. The temperature of the water, the quantity of oxygen that organisms take in and expel via respiration and decomposition, the amount of oxygen that plants give off through photosynthetic activity, stream movement, and aeration all influence the amount of dissolved oxygen in water bodies. Water temperature significantly affects the quantity of DO because warm water dissolves less oxygen than cold water.

3. PROPOSED SYSTEM

The NOAA-AVHRR thermal-infrared channels, the Eumetsat (ESA) Met-OP satellites, the sea surface temperatures (SSTs), the chlorophyll, the Oceansat II (India) optical bands, and the MODIS Aqua (US) data are utilized to determine the potential fishing zones along the Indian coastline. It is necessary to compile PFZ texts, PFZ maps, and PFZ location information for each sector’s PFZ advertising. Because the sea is a dynamic environment, the fishing zones displayed on the maps may change depending on the location. For the purpose of assisting fisherman in foreseeing changes to the PFZ, the wind speed and direction information is also shown on the PFZ maps. Using this information, fishermen can locate the PFZ on a map, but it takes them a day to go there. The PFZ charts and text are also accessible in the local language of each location for the benefit of fisherman. The PFZ text details the site’s position (length and latitude), depth, and separation from prominent beach features (fishing landing center, light structures), as seen in Figure 1.

The processing of the data is another key factor since raw data values are seldom provided to an algorithm for usage. Thus, the information has to be modified to meet our situation. Using the number of fish gathered, this study seeks to predict the IP filter (IPF) across the given territory. Future contracts need to be negotiated, often a year in advance, and their duration may be as long as you want. Additionally, a greater amount of data processing is required due to the uniqueness of the PFZ projections. Therefore, we apply a Financial Derivative feature to the data to make it more understandable and relevant for the issue area. Since the outcome seems to have been less substantially affected, financial derivatives are utilized to encourage data fluidization. Furthermore, it seems that PFZ forecasting reduces daily volatility and facilitates pattern detection. This IPFZ forecasting technique may be tailored to any period of interest and is extremely versatile. Neural networks integrate several processing levels by using fundamental, concurrent, and components that are modelled after the human nervous system. A hidden layer or layers, an output layer, and an input layer make up this system. Neurons are able to connect the various layers together due to the number of nodes, or neurons, that make up each layer’s input. A neuron often alters its weights as it learns to either strengthen or weaken the output of the neuron. Figure 2 depicts the architecture of a feed forward neural net (FFNN).
The node $i$ becomes

$$Y_i = g_i = g\left(\sum_{j=1}^{K} w_{ji} x_j + \theta_i\right)$$

A FFNN network is created by connecting several nodes in parallel and series. Figure 3 depicts a typical network with a single hidden layer.

4. RESULTS AND DISCUSSION

The effectiveness of our strategy for segmenting and categorizing data is evaluated using the metrics for the challenge assessment. Accuracy, specificity, sensitivity, and false measure are among the segmentation assessment criteria. These are the required performance criteria,

$$\text{accuracy} = \frac{\text{tp}+\text{tn}}{\text{N}}$$

$$\text{sensitivity} = \frac{\text{tp}}{\text{tr}}$$

$$\text{specificity} = \frac{\text{tn}}{\text{fn}}$$

$$\text{precision} = \frac{\text{tp}}{\text{tp}+\text{fp}}$$
The letters $tp$, $tn$, $fp$, and $fn$, respectively, stand for the number of true positives, true negatives, false positives, and false negatives. The overall number of pieces is $N$. To accurately forecast PFZ's future, the 16 criteria and other financial derivative features are used. Utilising efficient pre-processing methods, the undesired and missing data are deleted. The neural network is the finest and most effective approach currently in use for anticipating nonlinear occurrences. In order to forecast the IPFZ, the Function Fitting Neural Network (FFNN) technique is primarily employed. In order to do the practical analysis, a number of factors including the 80%-20%, 60%-40%, and future data are employed. The findings show that the recommended FFNN obtained 90% accuracy compared to the previous neural network's 86% accuracy utilising financial derivative characteristics for the accessible dataset's 80%-20%. To safeguard the general populace's health, clean, secure drinking water must always be available. By using risk assessment and management strategies, beginning with watershed areas and ending with home storage vessels, drinking water safety may be efficiently assured. In the current investigation, risk variables impacting water quality throughout the distribution network may be found using the Wireless Session Protocol [WSP] method. The primary causes of faecal coliforms entering potable water have been identified as inconsistent water supply and inadequate distribution network maintenance. Small risks were well controlled, and immediate remedial action was implemented. However, our analysis demonstrates that in order to ensure the sustainability of the WSP, local water management teams must have ongoing assistance and direction. The comparison of the proposed technique toward other techniques are shown in Table 1. 

<table>
<thead>
<tr>
<th>Classification technique</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Precision</th>
<th>Recall</th>
<th>G-mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probabilistic neural network</td>
<td>1.0025</td>
<td>0.8598</td>
<td>0.0994</td>
<td>1.0025</td>
<td>0.9565</td>
</tr>
<tr>
<td>Function fitting neural network (FFNN)</td>
<td>1.0025</td>
<td>0.9895</td>
<td>0.0884</td>
<td>1.0025</td>
<td>0.9851</td>
</tr>
<tr>
<td>Convolutional neural network (CNN)</td>
<td>1.0025</td>
<td>0.9258</td>
<td>0.1209</td>
<td>1.0025</td>
<td>0.9882</td>
</tr>
<tr>
<td>Proposed Method</td>
<td>1.0025</td>
<td>0.9280</td>
<td>0.1541</td>
<td>1.0025</td>
<td>0.9795</td>
</tr>
</tbody>
</table>

5. CONCLUSION

Modern ecology and conservation have significant challenges in accurately monitoring the worldwide spread of varied human influences, such as fishing. Average annual fish production in India is just 2 tonnes per person. Through increasing fish exports, knowledge sharing regarding sustainable fishing methods, and high productivity, there is an opportunity to raise the living standards of Indian fishermen. The major objectives of this project are to find a solution to this issue and assist the fishermen in locating productive fishing zones and locations. Using pre-processing and normalisation techniques, missing values and undesirable data are eliminated from the collected data. IPFZ prediction will be carried out in the future using the feature selection approach and the CECF methodology. The fishing effort maps of other nations, particularly those for remote regions and the high seas, are rife with doubt. While some nations routinely monitor their coastal fisheries in their territorial waters, others do not. To prioritise and carry out international fisheries management and conservation initiatives, a fuller knowledge of the world's fishing fleets is required. 4,827 fishing settlements, 2,914 fish landing facilities, and 7589 miles of marine coastline may be found in India. In this industry, there are between 13 and 14 million workers. 14 criteria are then employed to extract the important features, together with geometrical and financial derivatives considerations. The results revealed that the suggested CECT approach had an accuracy of over 93% for the 85%-25% range and 72% for the analysis of future data.

REFERENCES


**BIOGRAPHIES OF AUTHORS**

Vigisson Prem Monickaraj completed his B. Sc. Degree in Physics from M. K. University and Post Graduate Diploma in Computer Application from M. S. University in the years 1992 and 1993 respectively. He received his M. Sc. Degree in Computer Science from Bharathidasan University in the year 1996 and M. Phil. Degree in Computer Science in the year 2003 from M. S. University. He completed his M.E. degree in Computer Science and Engineering (2006) from Anna University. He received his Ph.D. Degree in Information and Communication Engineering from Anna University in 2013. His research interests lie in the areas of mobile agents, networks, distributed system, database, and network security. He has over 26 years of teaching experience. He has around 25 publications in the international conferences and referred journals. At present, he is working as a Professor in the Department of Computer Science and Engineering at R.M.K. College of Engineering and Technology, Puduvoyal. He can be contacted at email: vigiprem@gmail.com
Sterlin Rani Devakadacham received the Bachelor of Engineering degree from St. Xavier’s Catholic College of Engineering at Chennakesvaram in 2007, Master of Engineering from S.A. Engineering College at Chennai in 2009. She is pursuing Ph.D. degree in Cloud Computing from Anna University, Chennai. She is currently an Assistant Professor at R.M.D. Engineering College, Chennai, India. She has totally 13 years of teaching experience. She has authored 3 research publications both international and national journals and presented 7 papers in conferences. Her areas of interest include cloud computing, data mining and machine learning. She is a life member of professional bodies like IFERP and MISTE. She can be contacted at email sterlinrani@gmail.com.

Nithyadevi Shanmugam received the Bachelor of Engineering degree from Sasurie College of Engineering, Vijayamanagalram in 2009, Master of Engineering from Anna University Regional center, Coimbatore 2013. She is pursuing PhD degree in Testing of Low power VLSI digital circuits from Anna University, Chennai. She is an Assistant Professor at Dr. N.G.P. Institute of Technology, Coimbatore since 2016 and earlier she had gained her experience from Info Institute of Technology, Coimbatore. She has totally 13 years of teaching experience. She has authored 8 research publications both international and national journals and presented 11 papers in international conferences. Her areas of interest include Testing of VLSI digital circuits, artificial intelligence, signal processing and VLSI for wireless communications. She is a life member of professional bodies like ISTE, IETE and IAENG. She can be contacted at email: ndsniathy88@gmail.com.

Nithya Nandhakumar is working as Assistant Professor in the Department of Computer Science Engineering at K. Ramakrishnan college of Engineering. She completed her undergraduate & post graduate in the stream of Computer Science Engineering. Currently She doing her research area in the field of artificial intelligence and block chain technology. She can be contacted at email: nithyaasrichithra@gmail.com

Manjunathan Alagarsamy received the Engineer degree in Electronics and Communication Engineering from Dr. Navalar Nedunchezhiyan College of Engineering in 2010. He received the Master degree in Embedded System Technologies from Raja College of Engineering and Technology, Madurai, Tamilnadu, India in 2013. He is currently working as an Assistant Professor in the Department of Electronics and Communication Engineering at K. Ramakrishnan College of Technology, Trichy, India. His area of interests includes embedded systems, image processing, sensors and interfacing networks and internet of things. He has published 31 articles in peer reviewed international journals and presented 6 papers in international conferences. He can be contacted at email: manjunathankrit@gmail.com.

Dr. Kannadhasan Suriyan is working as an Assistant Professor in the Department of Electronics and Communication Engineering in Study World College of Engineering, Coimbatore, Tamilnadu, India. He is Completed the Ph.D in the field of Smart Antenna for Anna University in 2022. He is Twelve years of teaching and research experience. He obtained his B.E in ECE from Sethu Institute of Technology, Kariapatti in 2009 and M.E in Communication Systems from Velammal College of Engineering and Technology, Madurai in 2013. He obtained his M.B.A in Human Resources Management from Tamilnadu Open University, Chennai. He has published around 45 papers in the reputed indexed international journals indexed by SCI, Scopus, Web of science, Major indexing and more than 146 papers presented/published in national, international journal and conferences. Besides he has contributed a book chapter also. He also serves as a board member, reviewer, speaker, session chair, advisory and technical committee of various colleges and conferences. He is also to attend the various workshop, seminar, conferences, faculty development programme, STTP and Online courses. His areas of interest are smart antennas, digital signal processing, wireless communication, wireless networks, embedded system, network security, optical communication, microwave antennas, electromagnetic compatibility and interference, wireless sensor networks, digital image processing, satellite communication, cognitive radio design and soft computing techniques. He is Member of IEEE, ISTE, IEI, IETE, CSI, IAENG, SEE, IEAE, INSC, IARDO, ISRPM, IACSIT, ICSES, SPG, SDIWC, IJSPR and EAI Community. He can be contacted at email: kannadhasan.ece@gmail.com.