International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

Undersea Water Objects Detection and Classification Using Optimized Region-Based Convolutional Neural Network

Kaipa Sandhya[1], Jayachandran Arumugam[2]
[1]Assistant Professor, School of Computer Science and Information Science
Presidency University,
Bangalore, India
kaipasandhya@presidencyuniversity.in
[2]Professor & Head, School of Computer Science and Information Science
Presidency University,
Bangalore, India
ajayachandran@presidencyuniversity.in

Abstract— Underwater autonomous operation is becoming increasingly crucial to avoid the hazardous in the environment of high-pressure deep-sea due to the significance of underwater investigation. The most crucial piece of technology for underwater-based task is intelligent computer vision. In an underwater environment, underwater vision requires good image quality, and illumination with better classification of sea objects. This work presented a novel technique of Optimized Region-based Convolutional Neural Network (ORCNN). For this work, the Gaussian filter is used to remove the noise and enhance the image quality during pre-processing and the Improved Affinity Propagation Clustering (IAPC) model segments the objects. After that, the Region-based Convolutional Neural Network (RCNN) model classifies various objects such as urchins, seagrass, fishes, and rocks in which the RCNN parameters are tuned via a light spectrum optimizer algorithm (LSOA). can also input the segmentation prediction images and the labeled images into the discriminant convolutional network and improve the segmentation accuracy of underwater images by further enhancing the essential characteristics of learning data through the confrontation training of generators and discriminators. The experimental results demonstrate that higher performance is provided by the newly developed ORCNN -Net predictive model when compared to other comparative algorithms while considering the negative and positive metrics..

Keywords- Image enhancement, Gaussian filter, Light spectrum optimizer algorithm ,Region-based convolutional neural network.

I. INTRODUCTION

Object detection [1], also known as video conversion technology aids in identifying and drawing attention to objects in a video or image. Additionally, it can give underwater machines perceptual data about their surroundings, tragically, compared to conventional visual processing, submerged robotic vision remains in its infancy. In order to be more specific, object detection builds a network of interrelated elements around the objects it has discovered, which enables us to pinpoint where those objects are in a surroundings [2]. Finding instances of each object in picture aspects or in everyday circumstances, classifying them, and looking at their key characteristics for within-the-moment estimations are the basic goals of object detection. The picture that was input is divided into many segments, and each of these segments is then taken as one continuous frame. In this construction, a multiple-layer per is built from the binary pictures using a fully-connected network [3]. Underwater vision technology has several unique hurdles in terms of picture augmentation, recognition of objects, and navigation because of the water-resistance specifications for images and the chemical composition of light propagation underwater.

To advance underwater machines' sensory and perception capabilities, underwater imagery technology [4] must be improved, this would help the autonomous submerged vehicle investigation and operate much more. Sonar pictures have historically been the subject of substantial study and use. Three categories of acoustic processing of pictures exist echo-based analyzing, based on characteristics processing, and pixel-based synthesis [5]. Typically, sonar pictures are utilized to locate the object of desire. Sonar pictures are less appropriate for underwater modification than visual data. An underwater robot's visual observing equipment can be incorporated as a standalone element and is essential for monitoring the water's conditions up close. Unintentional capture of a variety of creatures, especially sea turtles in existence, predatory creatures [6], and other aquatic creatures, can harm the environment and cause a reduction in the population of such species.

Since current investigations suggest that a great number of sea creatures are killed in fishing operations each year, this accidental capture, known as bycatch, poses a concern in

particular to green species. Additionally, by destroying the fishing equipment or forcing relocation away from lucrative areas, these catastrophes result in huge financial losses for coastal towns. The challenging process of converting visible light into a meaningful image is supported by a variety of brain regions and increasingly sophisticated cognitive processes [7]. Sensation includes the gathering, handling, and assessment of knowledge. This entails the perception of specific data that enters throughout our perceptual operations, gets organized into our pre-existing patterns and cycles, and is then assessed in light of our prior expertise. Evaluation, authorization, a downtrend, and confirmation are the four essential components that make up main activities. Social perception [8] is based on assessments, which come from the interplay of three basic sources: people, situations, and behavior.

Effective observation is facilitated by a consciousness of how a system of thought develops as a result of suppressed and changed sensory input. The fundamental objective of identifying objects is to find representations of each item in digital pictures or real-world situations, segregate them, and analyze their essential properties for in-the-moment estimations, an organization's entire data infrastructure includes object detection. An underwater transmission of light system's effectiveness in ocean water is primarily constrained by maritime instability, which can be characterized as oscillations in the coefficient of refractive brought on by changes in both salinity and temperature. Particular challenges with underwater object recognition include difficult-to-detect tiny and numerous targets, poor camera and image quality, and a lack of computer power in submarines. This study's main contribution is highlighted as follows:

To employ Gaussian filtering to pre-process the images of the sea objects in order to correct illumination, improve contrast and ignore noise.

To segment the precise object of an image from the pre-processed image using an improved affinity propagation clustering model.

Optimized region with convolutional neural network (RCNN) model identify and categorize underwater objects namely rocks, fishes, seagrass and urchins.

The remainder of the paper is organized as follows: Section 2 summarized the literature reviews, and Section 3 describes the proposed work for the classification of underwater object. Section 4 discussed the section on the results, and Section 5 concludes the paper.

II. LITERATURE ANALYSIS

Lu et al. [9] have presented a deep neural network (DNN) to address the difficulty, where the overall functionality that may be achieved relied exclusively on the offsite instructional data packages. The suggested system has the potential of identifying many combined water kinds and using efficient identification on signals gathered after just a limited number of waters were used for offline education. The performance is enhanced and reliable to provide superior analysis of features. However, it is typically not reliable enough because of the challenging underwater conditions.

Sung et al. [10] have described deep learning-based object detection which is required for fully automated underwater activities like ship exploration, mine clearance, and landmark-based guidance. It also used augmented data to build the database with a wider variety of photos. Several of the experiment's items have symmetrical shapes, so they could still look identical even after rotation. The suggested approach may be helpful for the preliminary analysis and information enhancement of ultrasonic pictures. Nevertheless, it is challenging to create an authentic ultrasound generator.

Yeh et al. [11] have implemented an improved visual attention (IVA) module that may be included in any current Convolutional Neural Networks to help classify images. the learned feature representations into a single channel or image and the remainder communicates in the last layer of convolution to reinforce the key area characteristic patterns and attenuate the inconsequential characteristic patterns for a starting point picture for classification purposes. The suggested module incorporated improves the efficiency of classification significantly. Hence, it is inadequate to perform in large-scale networks.

Dos Santos et al. [12] evaluated cross-domain and cross-view image matching, using a deep neural network to determine if these photographs were taken in an identical location by comparing a color aerial image with a submerged acoustic image. The technique can be used in somewhat organized situations like harbors and canals. As a result of the comparison, places may be recognized using different types of sensors. Both numerical and qualitative findings demonstrate the capability of our technology to match real-world information from various places. Thus, the issue of excess fitting arises when a neural network's complexity is increased.

Wang et al. [13] have developed a real-time lightweight object detector (RLOD) intended for Graphics processing units placed in underwater machines that conserve electricity. Submerged artificially intelligent grasping capability in terms of both acceleration and sensitivity is encountered. To increase effectiveness in detection and velocity, the highlighted circular network and thick linkages are introduced into the intermittent measurement network. It demonstrates the efficacy of the suggested picture modification and recognition of object techniques. Therefore, the allocation problem is to be decreased to track and identify items.

Demir et al. [14] suggested a Convolutional neural network (CNN) to determine the level of assurance for a particular frame, according to classification structure. To reach a final determination for the identification of seashells, the result of this system for many pictures is analyzed along with ecological

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

variables such as the clarity of the water and ambient light level. It is feasible and the consumption is reduced. Still, the unpredictability or diminishing gradient issue makes the acquisition process more challenging.

Kvasić et al. [15] highlighted convolutional neural networks in developing a stable and trustworthy ultrasound image analysis technique employing multilayer neural systems for the identification and monitoring of actual human explorers. It is also in charge of constructing a frame across the item and determining its precise location within the picture if it appears in the picture or report. This method cut an aperture out of an image that was considerably less than the actual size of the picture and sent it successively to a neural network for the forecast. Therefore, this method produced extremely costly computations.

Sung et al. [16] a convolutional neural network to locate and eliminate crossover distortion Utilizing beam-based comparisons between a generated picture and a real acoustic picture, an intended item is seen from numerous perspectives. From the ultrasonic perspective of view, the form of the item alters dramatically in the sonar frame. As a result, these algorithms, particularly in practical situations, show inadequate precision and frequent false positives. Thus, restoration of the perspective is thus an impossible obligation.

III. PROPOSED METHODOLOGY

The proposed framework for sea object detection and classification is outlined in Figure 1. This involves three major steps namely pre-process the image, segment the correct sea objects and finally classification to categorize the fishes, sea grasses, urchins and rocks. The following sub-sections elaborates the above steps with respect to various techniques.

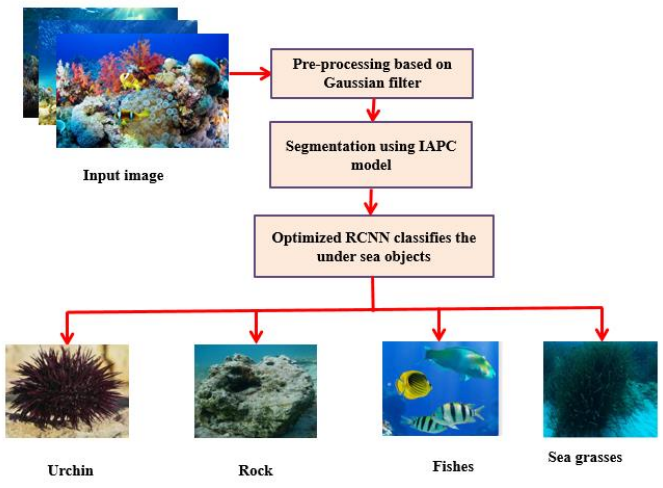


Figure 1: Proposed workflow diagram

A. Gaussian filter for pre-processing

Gaussian filtering is a commonly used image filtering technique and its weight defined as,

$$M_{j,k}\beta\exp\left(\left\|y_j - y_k\right\|\right), \qquad j \neq k \qquad (1)$$

The rapid decay distance is $M_{j,k}$. To over smooth sea water object images, the well-known technique to remove the noise is Gaussian filter that preserves the images low frequency components and it filter the image blur [17]. The minimum and maximum pixel intensity is y_j and is y_k to improve the image contrast.

B. Segmentation using IAPC

The segmentation of captured images is done using the improved affinity propagation clustering (IAPC) model. The sub-segments were made by using the IAPC and using each image's pixel as a potential exemplar. The key benefit of using the IAPC is that it automatically selects sub-segments based on the particular problem, eliminating the necessity for earlier sub-segment determination. The sub-segmented sections are produced using the correlation between each pixel to start the segmentation of the images [18]. To identify the precise exemplar of pixel j, the index of similarity has to be determined for each pixel of k. In addition to being the cause of the pixel correlation, the calculated Euclidean distance is also simple to find in the search space. Due to the fact that we have mentioned a negative Euclidean distance, the similarity index is given below:

$$Similarity\ index\ (j,l) = -\left\|a_j - a_l\right\|^2 \qquad (1)$$

Pixels jth and kth positions are identified as a_j and a_l. IAPC method makes use of the message-passing methodology after each pixel's similarity index has been completed. The initial value of the pixel accessibility is set to zero at first, and then the assessment of duties is performed. This formula effectuates the availabilities of up-gradation.

$$B(j,l) = \min imum\left\{0, r(j,l) + \sum_{k\neq\{i,l\}} \max imum\{0, r(j,l)\}\right\} \qquad (2)$$

The self-availability assessment differs from that of equation (2) and is described below,

$$B(j,l) = \sum_{k\neq l} \max imum\{0, r(j,l)\} \qquad (3)$$

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

Additionally, by taking into account the methods that follow, the numerical fluctuations that occur throughout the damping process can be avoided.

$$r(j,l) = \vartheta \times r(j,l)_{old} + (1-\vartheta) \times r(j,l) \quad (4)$$

$$B(j,l) = \vartheta \times B(j,l)_{old} + (1-\vartheta) \times B(j,l) \quad (5)$$

From this, is the damping factor that falls between [0, 1]. The damping factor is calculated as the value that falls between 0 and 1. Therefore, if l=j, the pixel l can be used as an example. Up till the assessed exemplar remained constant for more than two iterations, the previously described process is repeated. As a result, the suggested IAPC efficiently segments images of marine objects.

C. Classification of under seawater objects

Light Spectrum Optimizer (LSO) The LSO is relied on the metrological values and for that the some assumptions are made [19]. Let us assume LSO parameters such as (a) each colorful ray are used for the determination of candidate solutions. (b) The dispersion of the light rays' lies between the 40o to 42o, (c) the best dispersion is used for the selection of global best solution of population light rays, (d) the reflection incorporated with refraction controlled arbitrarily, (e) initial scattering stages are managed with the help of fitness value of current solutions. The initialization of the LSOA is started with the arbitrary population of white lights and can be interpreted as,

$$\overrightarrow{a^0} = V_L + QU_1(V_U - V_L) \quad (6)$$

Here the random vector QU_1 flexibly shift uniformly in the ranges of 0 to 1 with the initialized solution $\overrightarrow{a^0}$. The upper and lower constraints are V_U and V_L with the problem dimension of d. The mechanism of exploitation, exploration, scattering of colorful rays along with the direction of rainbow spectrum are used for this stage [20].

(i) Rainbow spectrum direction:

The inner refraction ($\vec{a}_{nX}$) and reflection ($\vec{a}_{nY}$) along with the outer refraction ($\vec{a}_{nZ}$) and their respective normal vectors are outlined as follows,

$$\vec{a}_{nX} = \frac{\overrightarrow{a_t^q}}{Nor\left(\overrightarrow{a_t^q}\right)} \quad (7)$$

$$\vec{a}_{nY} = \frac{\overrightarrow{a_t^r}}{Nor\left(\overrightarrow{a_t^r}\right)} \quad (8)$$

$$\vec{a}_{nZ} = \frac{\overrightarrow{a^*}}{Nor\left(\overrightarrow{a^*}\right)} \quad (9)$$

The randomly chosen value when the iteration t is selected is $\overrightarrow{a_t^p}$ with the current solution of $\overrightarrow{a_t^r}$. The normalized vector value can be evaluated using the global solution is $\overrightarrow{a^*}$ operated with $Nor(\cdot)$. The normalized vector can be interpreted as,

$$Nor\left(\vec{a}\right) = \sqrt{\sum_{l=0}^{D} a_l^2} \quad (10)$$

The dimensionality D along with the input vector $\vec{a}$ is used to evaluate the incident light ray and is formulated as,

$$\overrightarrow{a_{L0}} = \frac{A_{mean}}{Nor(A_{mean})} \quad (11)$$

A_{mean} is the mean value of current population with their respective solution $\overrightarrow{a_j}(j = 1,....,N)$. For the incident light lay the population size is N with the value $\overrightarrow{a_{L0}}$.

(ii) Generation of new colorful ray:

The solutions r of arbitrarily generated are lies under the range of 0 to 1 and the new candidate solution is formulated as,

$$\vec{a}_{T+1} = \vec{a}_T + \delta QU_1^n\left(\vec{a}_{K1} - \vec{a}_{K3}\right) \times \left(\vec{a}_{q1} - \vec{a}_{q3}\right) \quad (12)$$

The current and generated candidate solutions are denoted as $\vec{a}_T$, $\vec{a}_{T+1}$ and QU_2^n are the uniform arbitrary vectors with the scaling factor of δ .

Scattering or colorful rays:

The enhancement of the exploitation operator is effectuated with the arbitrarily chosen from the current population.

$$\vec{a}_{T+1} = \vec{a}_T + QU_3\left(\vec{a}_{q1} - \vec{a}_{q2}\right) + QU_4^n \times (Q < \beta) \times \left(\vec{a}^* - \vec{a}_T\right) \quad (13)$$

The updated current solution and respective best solutions at new location are given as,

$$\vec{a}_{T+1} = 2\cos(\pi \times q_1)\left(\overrightarrow{a^*}\right)\left(\overrightarrow{a_T}\right) \quad (14)$$

Some of the solutions of LSOA is expand outside and for that two ways are proposed to transform the infeasible solutions to the feasible one. (i) Setting lower constraints over the dimensions to find the smaller value and the upper one to find the higher values. (ii) Creating arbitrary values inside the search boundaries. The hybrization of these two improves the convergence rate.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

D. LSOA based RCNN for the classification underwater objects

The framework of the proposed RCNN is employed majority with region proposal network (RPN) [21]. The entire seawater object detection accuracy is escalated with the utilization of LSOA with the mitigated computational complexity. This is because, the easy implementation with reduced time, flexible and easily converge able. Hence, for these reasons the proposed LSOA based RCNN could be utilized for the seawater object detection. The RPN can be used to generate the candidate regions along with the relied sliding window. The convolutional layer and sliding window are adopted for the measurement of 256 D Eigen vector. Utilizing these, center point, l anchor boxes, and convolutional feature map have attained for each location [22].

For the estimation of full connection score the values used are candidate region obtained using LSOA and the Centre location coordinates (a, b). Meanwhile, there is higher impact over the entire network with the anchor boxes width, size and length. For the tuning of these parameters we used higher search ability algorithm known as LSOA. While performing the underwater object detect there occurs loss and classification loss and the detection loss are given as $Clas_{loss}$ and Det_{loss} correspondingly with the loss function of the RPN.

$$Loss(R_j, S_j) = \frac{1}{G_{Clas}} \sum_j Clas_{loss}(R_j, R_j^*) + \frac{1}{G_{Det}} \sum_j S_J^* Det_{loss}(S_j, S_j^*) \quad (15)$$

$$Clas_{loss}(R_j, R_j^*) = -\log\left[(R_j, R_j^*) + (1 - R_j)(1 - R_j^*)\right] \quad (16)$$

$$Dec_{loss}(T_j, T_j^*) = \sum_J Smooth_{loss_1}(T_j, T_j^*) \quad (17)$$

$$Smooth_{loss_1}(a) = \begin{cases} 0.5a^2, & |a| < 1 \\ |a| - 0.5, & |a| \geq 1 \end{cases} \quad (18)$$

The anchor coordinates for the whole dataset of seawater object detection are N_{Det} and N_{clas} with the sample label of R_j^* and R_j. The bounding boxes' observed and real coordinates are taken as S_j^* and S_j respectively.

- (i) Evaluation of Fitness function

The classification of objects as sea urchins, sea grasses, fishes, and rocks are effectuated with proposed approach in which the LSOA fine tunes the RCNN parameters and provided optimal output. The Fitness function of proposed approach can be interpreted as below,

$$Fitness = \frac{Number\ of\ misclassified\ objects}{Total\ number\ of\ objects\ present\ in\ the\ seawater\ images} \times 100\% \quad (19)$$

The proposed approach ensures lower error rate with higher classification accuracy of objects as sea urchins, sea grasses, fishes, and rocks. The proposed approach schematic diagram is interpreted in figure 2.

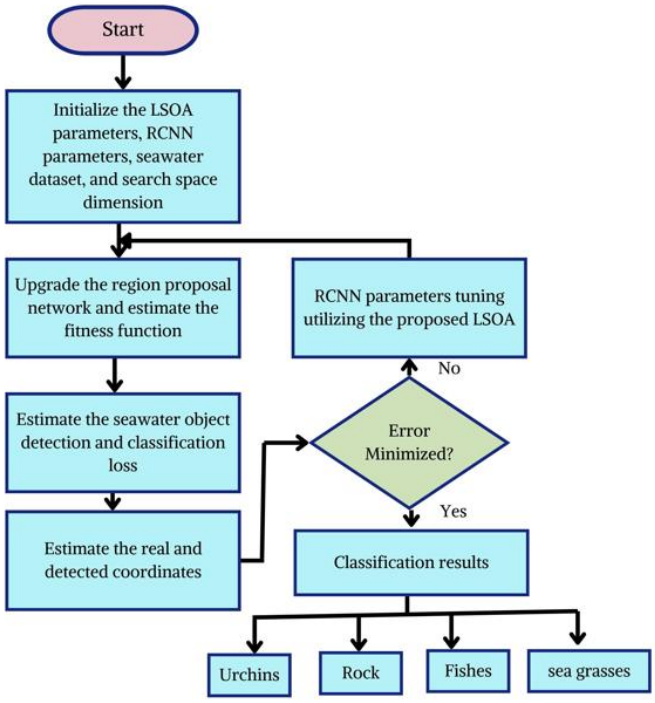


Figure 2: Proposed schematic overlay based on LSOA-RCNN for the classification of objects from the seawater images

IV EXPERIMENTAL RESULT AND DISCUSSION

The proposed work's effectiveness evaluation using an experimental setup and dataset. The remainder of this section clarifies the specifics. We used the MATLAB 2008a simulator operating on a PC with an NVIDIA GeForce GTX 1080 and Intel 5 core processor.

A. Evaluation Metrics

Several popular metrics are adopted to evaluate the performances, including the Precision, Specificity, and Accuracy. Dice Similarity Coefficient (DSC), Intersection over Union (IoU). It is given in the following Equations

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

$$\Pr = \frac{TP}{TP + FP}$$
$$Sp = \frac{TN}{TN + FP}$$
$$Acc = \frac{TP + TN}{TP + FP + TN + FN}$$
$$F1 = 2.\frac{Se.\Pr}{Se + \Pr}$$

Where T P, TN, FP, and FN denote true positive, true negative, false positive and false negative, respectively.

$$\text{IoU} = \frac{A \cap B}{A \cup B},$$
$$\text{DSC} = \frac{2\,|A \cap B|}{|A| + |B|},$$

Where A and B represent pixel sets for the ground truths and their detection results, respectively

B. EXPERIMENTAL RESULTS

Segnet [22] uses an encoder-decoder method and up-pools pixels to recover their spatial position. While it increases the resolution of segmentation, an excessive number of network training parameters results in significant computational costs. In order to improve the segmentation results, EncNet [24] presents a context encoding module that selectively highlights the feature maps associated with the categories while capturing the scene contextual semantics. To effectively balance segmentation speed and accuracy, DFANet [26] uses Xception as the backbone structure and embeds high-level context into the encoder. A dual-attention network with a self-attention mechanism is used by DANet [27]. It considerably improves segmentation outcomes and strengthens the discriminative capacity of scene segmentation feature representation

Experiment 1, transfer learning techniques: In this experiment, the datasets are first pre- processed based on the Keras application (i.e., normalizing). Then, different transfer learning techniques are used, namely TR1 and TR2. In transfer learning type 1 (TR1), the layers selected from the Xception architecture are considered to be feature extractors and froze during the training phase. But in transfer learning type 2 (TR2), as mentioned in Section 3.2., the weights taken from the Xception architecture are considered as initial weights and re-trained. The average loss and DSC curves are shown in Figure 4. As can be seen in this figure, by changing the type of transfer learning from TR1 to TR2, not only the average DSC has increased from 0.8304 to 0.8726 on the validation data but also the average loss decreased from 0.1709 to 0.1273. As a result, due to the low similarity between the SUIM and UIEBD datasets, updating the Xception pre-trained architecture weights improves the performance of the Xcep-ResUNet architecture.

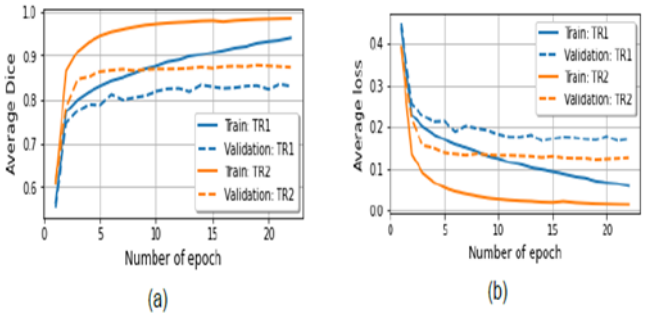


Fig.4. The average loss and DSC curves for training and validation datasets.

Experiment 2, preprocessing techniques and architectures: The experimental results of the proposed ORCNN method using in SUIM dataset [23] dataset are given in Table 2and the UIEBD dataset [24].is given in Table 3.

Table 2. Comparison of the different classification methods in the SUIM dataset.

Method	IoU	DSC	Acc	Se	Sp	Pr
UNet [27]	0.9328	0.9646	0.9903	0.9659	0.9948	0.9657
ResUNet [30]	0.9189	0.9562	0.9883	0.9627	0.9929	0.955
DeepLabV3 [29]]	0.937	0.9666	0.9910	0.9755	0.9941	0.9601
ResUNet++ [28]	0.9326	0.9646	0.9675	0.9748	0.9545	0.9656
DFANet [26]	0.8681	0.9281	0.989	0.9711	0.9903	0.8926
SegNet [25]	0.9100	0.9400	0.9534	0.9691	0.9423	0.9543
Proposed ORCNN	**0.9471**	**0.9725**	**0.9928**	**0.9774**	**0.9954**	**0.9691**

Comparing the proposed method with other modified ResUNet architectures [25-30] in Table 2 shows that the proposed method performs well on the SUIM test dataset. This is because there is a relatively large gap between the quantitative criteria obtained by the architectures designed in [25-30]. However, the segnet [25] and DFANet [26] architectures have better performance than the proposed ORCNN, as the formers have filters of different sizes that can retrieve spatial information more accurately. On the other hand, as reported in Table 3, the proposed architecture achieves good performance compared to other ResUNet-based architectures due to the use of the transfer learning technique.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

Table 3. Comparison of the different classification methods in the UIEBD dataset.

Method	IoU	DSC	Acc	Se	Sp	Pr
UNet [27]	0.9111	0.9507	0.9464	0.9353	0.9354	0.9439
ResUNet [30]	0.8727	0.9248	0.9801	0.9044	0.9939	0.9559
DeepLabV3 [29]]	0.8980	0.9444	0.9833	0.9483	0.9900	0.9455
ResUNet++ [28]	0.9290	0.9354	0.9439	0.9354	0.9439	0.9354
DFANet [26]	0.8873	0.9392	97.21	0.9895	0.9354	0.9439
SegNet [25]	0.9290	0.9354	0.9439	0.9700	0.9100	0.9567
Proposed ORCNN	**0.9351**	**0.9651**	**0.9895**	**0.9711**	0.9944	0.9612

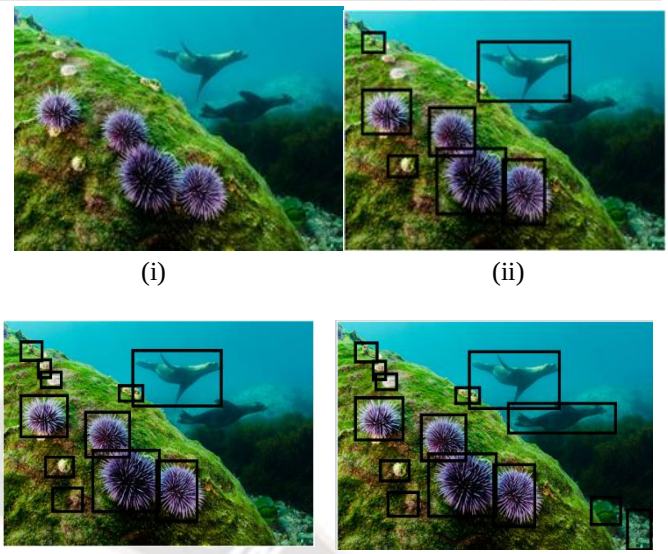
The performance comparisons of the proposed method on UIEBD data set are summarized in Table 4.

Table 4 Performance comparison of the proposed ORCNN method on UIEBD dataset.

Metrics	Net	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5	Overall
IoU	UNet	0.9058	0.8608	0.8539	0.8734	0.8994	0.8787
	ResUnet++	0.8998	0.8672	0.8593	0.869	0.8921	0.8775
	SegNet	0.8467	0.8203	0.8149	0.8165	0.8423	0.8281
	ORCNN	**0.9155**	**0.8757**	**0.8696**	**0.8790**	**0.9077**	**0.8895**
DSC	UNet	0.9492	0.9238	0.9201	0.9297	0.9459	0.9337
	ResUnet++	0.9455	0.9274	0.9230	0.9267	0.9413	0.9328
	SegNet	0.9125	0.8975	0.8939	0.8930	0.9098	0.9013
	ORCNN	**0.9546**	**0.9325**	**0.9293**	**0.9328**	**0.9503**	**0.9399**

C. Results Visualization

This study used the Enhancing Underwater Visual Perception (EUVP) dataset that contained matched and mismatched image samples of higher and worse quality images. The object detection in the seawater images are performed using our proposed approach and compared with state-of-art works such as RLOD, and CNN. For sample-1, we have compared the proposed work with state-of-art works and listed in figure 3. In figure 3(i) illustrates the sample image-1, 3(ii) shows the result using the RLOD approach, 3(iii) shows the outcome of CNN approach, and 3(iv) interprets the proposed approach result.

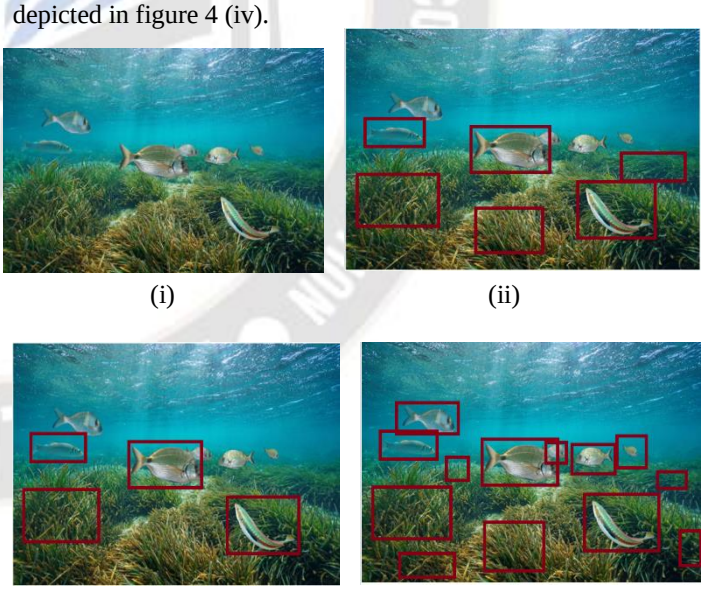


(i) (ii)

(iii) (iv)

Figure 3: Object detection from seawater images (i) Sample images -1, (ii) RLOD, (iii) CNN, and (iv) Proposed

Figure 4 illustrates the comparative outcomes of proposed and state-of-art works such as RLOD and CNN for sample image-2. Figure 4(i) depicts the sample image-2, figure 4 (ii) depicts the outcome of RLOD approach, figure 4(iii) illustrates the result of approach known as CNN and proposed approach outcome is depicted in figure 4 (iv).



(i) (ii)

(iii) (iv)

Figure 4: Object detection from seawater images (i) Sample images-2, (ii) RLOD, (iii) CNN, and (iv) Proposed

D. Comparative results

We used statistical metrics including accuracy, specificity, and sensitivity to evaluate the performance of the proposed technique. The performances are evaluated against existing

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

works like DNN [9], IVA [11], RLOD [13], CNN [14], and the proposed method. Figure 5 shows how performance is estimated depending on how well marine organisms including sea grasses, rocks, fish, and urchins are classified with respect to accuracy. The proposed method is more accurate than previous methods, with a 98.1% rate, while DNN, IVA, RLOD, and CNN have accuracy rates of 91%, 85%, 92%, and 93.29%, respectively. As a result, the proposed strategy outperforms all others.

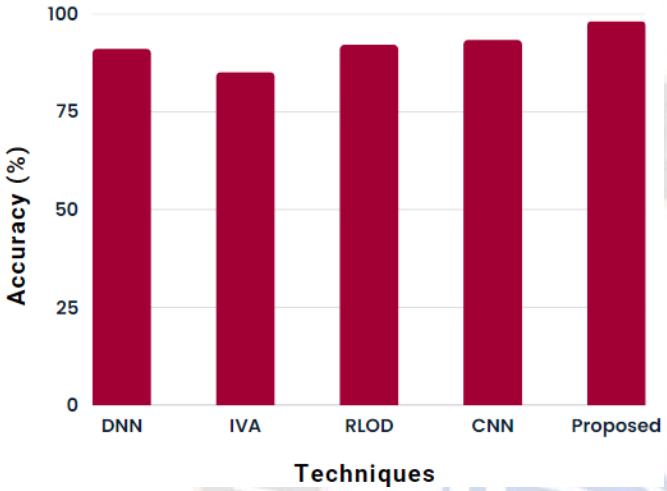


Figure 5: Comparison of accuracy

The sensitivity-based performance evaluation is shown in Figure 6, and it is clear from the figure that the proposed approach obtains a sensitivity of 97.87% while DNN, IVA, RLOD, and CNN achieved lesser sensitivities of 92%, 90%, 93%, and 93.1%, respectively. This is because the proposed method successfully categorizes the true positive and true negative values

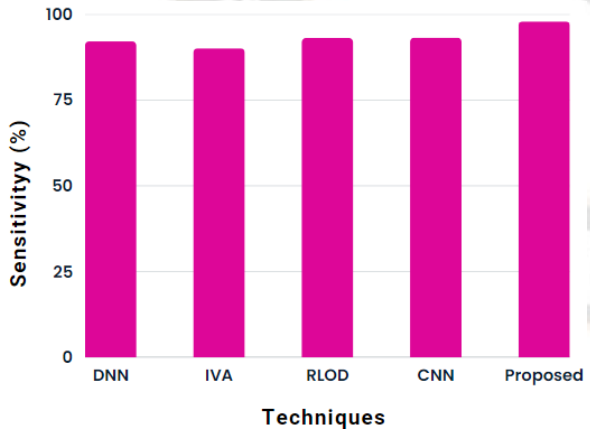


Figure 6: Comparison of sensitivity

Figure 7 displays an illustration of the specificity -based comparison analysis. The proposed method outperforms all other methods, including DNN, IVA, RLOD, CNN and proposed methods offers 92%, 90%, 93%, 93.1% and 97.61% specificity. The proposed work demonstrating superior sensitivity outcomes than he previous approaches.

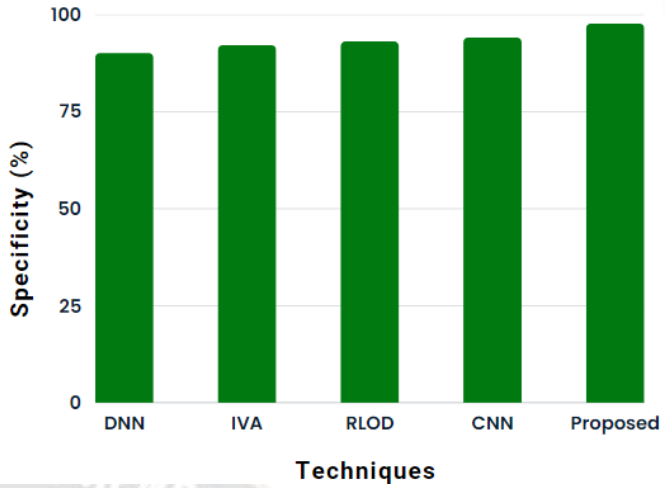


Figure 7: Comparison of specificity

V. CONCLUSION

This work proposed a novel Optimized Region based Convolutional Neural Network (ORCNN) to classifying the under seawater objects namely urchins, seagrass, fishes and rocks. Gaussian filter and IAPC model pre-process and segment the images. The execution work is handled via MATLAB software, and several types of assessment measures are compared between the proposed and existing approaches. The EUVP dataset, which included matched and mismatched image samples of better and lower quality images, was employed in this investigation. In comparison to various existing methods, including DNN, IVA, RLOD, and CNN, the proposed optimized RCNN for object classification has accuracy, specificity, and sensitivity of 98.1%, 97.61%, and 97.87%. In conclusion, the choice of a deep learning model depends on several factors, and it is not always necessary to use the deepest model to achieve the best performance on a specific task. It is important to evaluate different models and select the one that best suits the specific needs of the project.

REFERENCES

[1] Zhao, X., Sun, P., Xu, Z., Min, H. and Yu, H., 2020. Fusion of 3D LIDAR and camera data for object detection in autonomous vehicle applications. IEEE Sensors Journal, 20(9), pp.4901-4913.

[2] Pan, B., Sun, J., Leung, H.Y.T., Andonian, A. and Zhou, B., 2020. Cross-view semantic segmentation for sensing surroundings. IEEE Robotics and Automation Letters, 5(3), pp.4867-4873.

[3] Yang, J., Xi, M., Jiang, B., Man, J., Meng, Q. and Li, B., 2020. FADN: fully connected attitude detection network based on industrial video. IEEE Transactions on Industrial Informatics, 17(3), pp.2011-2020.

[4] Chen, L., Jiang, Z., Tong, L., Liu, Z., Zhao, A., Zhang, Q., Dong, J. and Zhou, H., 2020. Perceptual underwater image enhancement with deep learning and physical priors. IEEE Transactions on Circuits and Systems for Video Technology, 31(8), pp.3078-3092

[5] Guo, R., Jia, Z., Song, X., Li, M., Yang, F., Xu, S. and Abubakar, A., 2020. Pixel-and model-based microwave inversion with

supervised descent method for dielectric targets. IEEE Transactions on Antennas and Propagation, 68(12), pp.8114-8126.

[6] Romano, D., Bloemberg, J., Tannous, M. and Stefanini, C., 2020. Impact of aging and cognitive mechanisms on high-speed motor activation patterns: evidence from an orthoptera-robot interaction. IEEE Transactions on Medical Robotics and Bionics, 2(2), pp.292-296

[7] Wang, Y., Kwong, S., Leung, H., Lu, J., Smith, M.H., Trajkovic, L., Tunstel, E., Plataniotis, K.N., Yen, G.G. and Kinsner, W., 2020. Brain-inspired systems: A transdisciplinary exploration on cognitive cybernetics, humanity, and systems science toward autonomous artificial intelligence. IEEE Systems, Man, and Cybernetics Magazine, 6(1), pp.6-13.

[8] Li, L., Zhang, Q., Wang, X., Zhang, J., Wang, T., Gao, T.L., Duan, W., Tsoi, K.K.F. and Wang, F.Y., 2020. Characterizing the propagation of situational information in social media during covid-19 epidemic: A case study on weibo. IEEE Transactions on computational social systems, 7(2), pp.556-562.

[9] Lu, H., Jiang, M. and Cheng, J., 2020. Deep learning aided robust joint channel classification, channel estimation, and signal detection for underwater optical communication. IEEE transactions on communications, 69(4), pp.2290-2303.

[10] Sung, M., Kim, J., Lee, M., Kim, B., Kim, T., Kim, J. and Yu, S.C., 2020. Realistic sonar image simulation using deep learning for underwater object detection. International Journal of Control, Automation and Systems, 18, pp.523-534.

[11] Yeh, C.H., Lin, M.H., Chang, P.C. and Kang, L.W., 2020. Enhanced visual attention-guided deep neural networks for image classification. IEEE Access, 8, pp.163447-163457.

[12] Dos Santos, M.M., De Giacomo, G.G., Drews, P.L. and Botelho, S.S., 2020. Matching color aerial images and underwater sonar images using deep learning for underwater localization. IEEE Robotics and Automation Letters, 5(4), pp.6365-6370.

[13]] Wang, Y., Tang, C., Cai, M., Yin, J., Wang, S., Cheng, L., Wang, R. and Tan, M., 2020. Real-time underwater onboard vision sensing system for robotic gripping. IEEE Transactions on Instrumentation and Measurement, 70, pp.1-11.

[14] Demir, H.S., Christen, J.B. and Ozev, S., 2020. Energy-efficient image recognition system for marine life. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 39(11), pp.3458-3466.

[15] Kvasić, I., Mišković, N. and Vukić, Z., 2019, June. Convolutional neural network architectures for sonar-based diver detection and tracking. In OCEANS 2019-Marseille (pp. 1-6). IEEE.

[16] Sung, M., Cho, H., Kim, T., Joe, H. and Yu, S.C., 2019. Crosstalk removal in forward scan sonar image using deep learning for object detection. IEEE Sensors Journal, 19(21), pp.9929-9944.

[17] Perumal, S. and Velmurugan, T., 2018. Preprocessing by contrast enhancement techniques for medical images. International Journal of Pure and Applied Mathematics, 118(18), pp.3681-3688.

[18] Serdah, A.M. and Ashour, W.M., 2016. Clustering large-scale data based on modified affinity propagation algorithm. Journal of Artificial Intelligence and Soft Computing Research, 6(1), pp.23-33.

[19] Abdel-Basset, Mohamed, Reda Mohamed, Karam M. Sallam, and Ripon K. Chakrabortty. "Light spectrum optimizer: a novel physics-inspired metaheuristic optimization algorithm." Mathematics 10, no. 19 (2022): 3466.

[20] Vijay, M. M., and D. Shalini Punithavathani. "Implementation of memory-efficient linear pipelined IPv6 lookup and its significance in smart cities." Computers & Electrical Engineering 67 (2018): 1-14.

[21] Devaraj, S. Allwin, N. Muthukumaran, Stalin Jacob, Jenifer Darling Rosita, M. M. Vijay, and A. Andrew Roobert. "Performance analysis of reconstructed image using colour demosaicing algorithm for digital images." International Journal of Innovative Computing and Applications 13, no. 3 (2022): 138-149.

[22] Abbas, Syed Mazhar, and Shailendra Narayan Singh. "Region-based object detection and classification using faster R-CNN." In 2018 4th International Conference on Computational Intelligence & Communication Technology (CICT), pp. 1-6. IEEE, 2018.

[23] Islam, M. J., Edge, C., Xiao, Y., Luo, P., Mehtaz, M., Morse, C., Enan, S. S., & Sattar, J. (2020). Semantic Segmentation of Underwater Imagery: Dataset and Benchmark. IEEE International Conference on Intelligent Robots and Systems, 1769–1776. DOI: https://doi.org/https://doi.org/10.1109/IROS45743.2020.9340821

[24] Li C., Guo C., Ren W., Cong R., Hou J., Kwong S., et al. An underwater image enhancement benchmark dataset and beyond. IEEE Trans. Image Process. 2019, 29, 4376–4389.

[25] Badrinarayanan V.; Alex K.; Roberto C. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE Trans. Pattern Analy. Mach. Intel. 2017, 39, 2481–2495.

[26] Li H.; Xiong P.; Fan H.; Sun J. Dfanet: Deep feature aggregation for real-time semantic segmentation.In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, (2019).

[27] Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical Image Computing and Computer-assisted Intervention, pp. 234–241

[28] Yang Z, X. Peng, and Z. Yin, “Deeplab v3 plus-net for image semantic segmentation with channel compression,” in Proceedings of IEEE 20th International Conference on Communication Technology (ICCT). IEEE, 2020, pp. 1320–1324.

[29] Li H.; Xiong P.; Fan H.; Sun J. Dfanet: Deep feature aggregation for real-time semantic segmentation.In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, (2019).

[30] Jha D, Smedsrud PH, Riegler MA, Johansen D, De Lange T, Halvorsen P, et al., editors. Resunet++: An advanced architecture for medical image segmentation. 2019 IEEE International Symposium on Multimedia (ISM); 2019: IEEE