Efficient Machine Learning Approach for Subaquatic Surveillance

Ravi Shankar[1], Sunil Kumar[2]

[1]Department of Computer Science and Engineering, M. S. Ramaiah University of Applied Science, Bangalore, India
Corresponding Author Email: *ravishankar2311[at]gmail.com*

[2]Department of Computer Science and Engineering, M. S. Ramaiah University of Applied Science, Bangalore, India
Email: *skp.narayan[at]gmail.com*

Abstract: ***Deep learning has gained significant attention in recent years for its potential in categorising underwater photographs to identify various objects like fish, plankton, coral reefs, sea grass, submarines, alien objects and the movements of a deep - sea diver for subaquatic surveillance. Accurate classification is essential for the surveillance of the aqua condition and purity of water/sea bodies, as well as the preservation of endangered species living inside it. Additionally, it has importance in the fields of maritime affairs, defence and security. Thus, we have proposed a system which utilizes a deep convolutional neural network (VGG16), a type of deep learning technology, to recognize the image in order to provide underwater monitoring with great convenience. Grayscale and white balance techniques have been employed to reduce complexity and enhance the quality of underwater images before using Deep CNN. Ultimately, the experimental study confirms the successful identification of the acquired underwater images.***

Keywords: Deep learning, VGG16, Under Water Images etc.

1. Introduction

The demand for processing underwater photos has markedly risen recently. An examination of the behavior and population dynamics of many aquatic plant and animal species yields advantages for the fields of managing biodiversity, marine ecology and aquacultural economics. It will simplify the analysis of variation among subspecies and the safeguarding of subspecies at risk of extinction. Plankton, such as microorganisms and small organisms, exhibit a high degree of sensitivity towards alterations in their environment and living conditions. Consequently, examining their welfare offers a preliminary warning of climatic disasters Including pollutants and elevated temperatures. They play a vital role in the aquatic food web and facilitate the exchange of chemicals between water and the atmosphere. Plankton accounts for approximately 80% of the Earth's oxygen generation; hence, a decline in plankton is profoundly detrimental [1] [6]. Simultaneously, there is a profusion of plankton. Posidonia Oceanic, like - wise, exclusively inhabit unpolluted water and play a vital role in promoting biodiversity, mitigating beach erosion, and improving water quality. Examining the welfare of aquatic organisms can provide insights into the effects of global warming and overly abundant human activity on water ecosystems and marine life, thereby informing conservation efforts. Image processing can enhance and augment additionally, methodologies such as the analysis of water by physio - chemical means and the use of sonar for detection purposes [2]. Researchers have been driven to utilize deep learning models for the purpose of underwater image processing due to their success. CNNs have demonstrated superior predictive capabilities compared to traditional image - processing or machine learning methods, surpassing even human performance. This system offers an extensive summary of deep learning techniques employed for the classification of underwater photos. Underwater images necessitate pre - processing due to their inherent low quality. Due to the limited availability of underwater datasets and the significant imbalance in class distribution, utilizing data augmentation with transfer learning techniques is crucial.

Transfer learning decreases the computational requirements of the training system. Likewise, the constrained proportions of objects and species in underwater pictures, coupled with the paucity of datasets, necessitate the reduction of annotation efforts. The focus on the maritime environment globally has grown, with marine waste being identified as a significant contributor to the severe conditions in the marine ecosystem. Due to the growing human presence and increased waste production around coastlines and in the ocean, a significant portion of this material is being carried into the ocean and ultimately settling in the Deep Ocean [2] [3].

Our study aims to effectively classify and predict underwater photographs for which we will be implementing the Deep CNN. To enhance the accuracy and efficiency of detecting and classifying underwater objects. Image improvement and recovery applications may be utilised to improve the Visual consistency of underwater images by applying techniques such as greyscale and white balance adjustments. However, the resolution of these photographs remains limited. Super - resolution reconstruction is a widely utilized method that enhances resolution beyond the limitations of imaging systems. Gaining a comprehensive understanding of the point spread function and regularization processes has the potential to significantly improve the performance of reconstruction [3]. This study proposed a robust approach to improve the clarity of underwater images by utilizing a maximum a posteriori framework and applying regularization using the point spread function. The effectiveness of the reconstruction is evaluated using objective metrics for picture quality [4 - 10]. According to the experimental data, the proposed technique significantly enhanced the resolution and quality of underwater image detection. Single - image super - resolution has made tremendous advancements recently [2]. Interpolation processing can result in visual distortions when excessive smoothing of details occurs, particularly with a large super - resolution factor.

Most technique utilizes a very intricate convolutional network that draws inspiration from the VGG - net, a widely employed model for Image Net classification. We observed a substantial improvement in accuracy by expanding the depth of our network. The ultimate version of our model consists of a grand total of 20 weight layers [3]. Large amounts of contextual information are effectively utilized by repeatedly applying small filters in a deep network structure [5] [18]. Nevertheless, in networks with a large number of layers, the rate at which convergence occurs becomes a crucial factor to consider throughout the training process. We recommend employing a simple yet efficient training methodology. The primary focus is on training residuals and making use of very high learning rates, which are 104 times quicker than SRCNN. This is made possible with the implementation of configurable gradient cutting. Some proposed technique surpasses current methods in terms of both accuracy and aesthetic advantages. Our proposed approach utilizes a deeply recursive convolutional network (DRCN) [4], to achieve picture super - resolution (SR). The network contains a recursive layer with a maximum limit of 16 iterations. By increasing the depth of recursion, it is possible to enhance the speed of the process without the need to include additional parameters for further convolutions. Training a Deep Residual Convolutional Network (DRCN) with a standard gradient descent method is highly difficult because to the issues of exploding or disappearing gradients [5 - 7], despite the advantages it offers. In order to facilitate training, we recommend the utilization of two extensions: recursive supervision and skip - connection. Our strategy significantly outperforms previous methods. Underwater photos sometimes experience colour distortion and contrast reduction due to the absorption and dispersion of light during their journey through water [5]. In order to tackle these hurdles, we outline and resolve two subsidiary issues with the goal of enhancing the quality of underwater images. In order to rectify the colour distortion. Next, to tackle the problem of poor contrast, we provide a novel and efficient method for enhancing contrast that effectively eliminates any artefacts [6], due to the fact that the majority of operations need calculations on a per - pixel basis, the recommended method is straightforward to execute and well - suited for real - time applications. Moreover, it is not required to have previous knowledge of imaging circumstances. Experiments demonstrate that enhancing the image's colour, contrast, naturalness, and item prominence leads to improvement [14] [15].

The existing systems involves doing a survey on deep learning algorithms used for underwater picture classification which highlight the similarities and distinctions of various approaches. Underwater picture categorization is a crucial application that can determine the true effectiveness of deep learning algorithms for Subaquatic surveillance [8] [12]. In order to achieve this objective, the study aims to provide academics with information about the latest advancements in deep learning for underwater photos, as well as inspire them to further advance its boundaries which provides a comprehensive overview of deep learning methods for categorizing underwater photos. Researchers have conducted a thorough analysis of the essential parameters and underlined the similarities and differences of the underwater unclear images. We performed an extensive review of the literature concerning datasets and training, alongside the research on the construction and optimisation of Convolutional Neural Networks (CNNs) [15]. Some survey seeks to provide academics with insights into the current advancements in deep learning applied to underwater pictures. It also hopes to inspire more exploration in this field, such as the use of RCNN, UWCNN - SD [15], (NiN+SVM) and (PCANet+SVM) techniques [6] [17]. Upon analysing the optimisation of Convolutional Neural Networks (CNNs), we have identified some drawbacks of current systems. Limitations of the current system include its inefficiency in managing large amounts of data and the inadequate application of image enhancement techniques with longer duration of training and testing without removing the noise efficiently [16] [18].

In this paper we propose Deep CNN (VGG16) as despite its depth, it maintains a relatively simple and uniform architecture, with only 3x3 convolutional filters throughout the network. This simplicity ensures that the model is easy to understand, implement [11], and fine - tune for specific tasks, including underwater object detection. Additionally, its pre - trained weights highly effective for transfer learning. we have acquired various underwater photos from different sources to make exclusive dataset repository for this system. Here, we have applied the image resizing and greyscale for preprocessing [7] [13]. Next, we have utilized the white balancing enhancement to improve the quality of the photograph. Following which, we have partitioned the set of photos into test and training purposes. The train image is employed for the learning patterns whereas, the test picture is utilized for making purpose of evaluation. We need to execute the deep learning algorithm, such as the Convolutional Neural Network (VGG16) [7]. The experimental findings demonstrate the capturing of object within the boundary boxes for a certain Images are identified correctly with optimal accuracy and minimal time consumption [12 - 17].

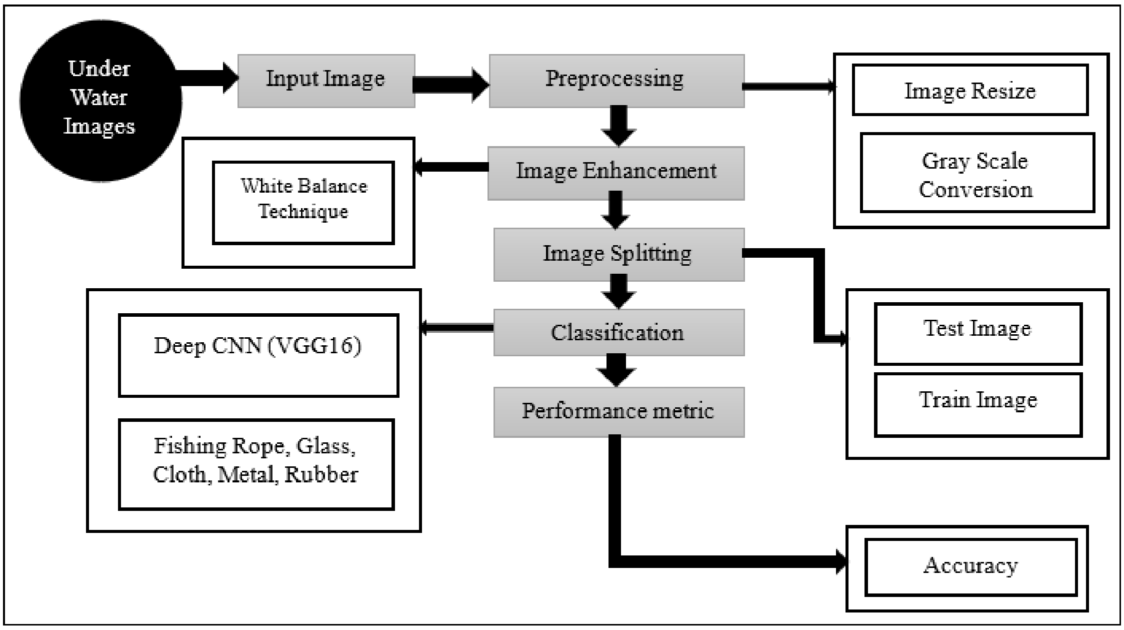

Figure 1 (a): Proposed Block Diagram

2. Theoretical Analysis

2.1 Image Selection

The various underwater picture dataset which is in the '. png', '. jpg' format was obtained from the dataset repository. The data set is read by imread () function which is present in in libraries like OpenCV or image processing libraries to load images into a program. This function parses the image file and transforms it into an array representation, suitable for computer manipulation, which can be processed by the program. A tkinter file dialogue box is a box used to pop the selected input picture in our technique.

2.2 Preprocessing of Images

Convolutional neural networks (CNNs) capture distinctive characteristics at various layers. Inconsistent resizing of photos can impede the network's ability to learn relevant features across different image sizes. Resizing is important for ensuring that comparable characteristics are identified in many photos. To alter the size of an image, employ the resize () function, with a tuple parameter consisting of two integers that specify the desired width and height of the resized image. The function preserves the original picture and generates a new picture with the modified dimensions. Gray scaling reduces the complexity of an image by converting it from a multi - channel (color) image to a single - channel (grayscale) image. This simplifies the data and makes it easier to process and analyze. To convert an image to greyscale in Python, we can use the conversion formula with the matplotlib library. Specifically, you can use the cv2. cvtColor (image, cv2. COLOR_BGR1GRAY) function from the library. Additionally, we can employ the conventional RGB to greyscale conversion method:

Image Gray = 0.5870 * G + 0.2989 * R + 0.1140 * B ….
(Equation for transforming to grayscale.)
Here B, G and R blue green red channels of the image.

2.3 Image Enhancements

In our approach, White balance solutions are essential for optimizing image quality. White balance (WB) is the procedure of eliminating false colour casts from photographs to guarantee that items that are white in actuality seem white in the photograph [7]. White balance corrects colour imbalances caused by various light sources (e. g., sunlight, fluorescent lighting, or incandescent bulbs), ensuring that whites appear white and other colours are rendered correctly. In computer vision applications, consistent colour representation across images is essential for tasks like object recognition, tracking, and classification. White balance helps achieve this consistency. Proper white balance makes images visually appealing, which is important in applications like photography, video production, and media. Facilitates Further Processing: Many image processing algorithms, such as segmentation or edge detection, perform better on colour - balanced images, as the colour differences between objects and backgrounds are more distinct. The white balance function of a digital camera guarantees that the colours of objects in the picture correspond appropriately to the light source.

White Balance Correction Formula:
For each colour channel C (Red, Green, Blue):

$$C' = C \times \frac{L_{avg}}{C_{avg}} \quad \text{………………………..1}$$

Where:
C′ is the corrected colour value for the channel C.
L_{avg} is the average luminance (or intensity) of the image.
C_{avg} is the average value of the colour channel C over the entire image.

Steps for obtaining the formula:
A. Calculate the average intensity for each of the red, green, and blue channels across the entire image:

$$R_{avg} = \frac{1}{N}\sum_{i=1}^{N} R' \quad \text{………………….. . . .2}$$

$$G_{avg} = \frac{1}{N}\sum_{i=1}^{N} G' \quad \text{……………………..3}$$

Paper ID: SR24910130120 DOI: https://dx.doi.org/10.21275/SR24910130120

$$B_{avg} = \frac{1}{N}\sum_{i=1}^{N} B' \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 4$$

Where R', G' and B' are red, green, and blue values of corresponding pixels in the image, and N is total number of pixels.

B. Compute the average luminance (or gray level) across the image:

$$L_{avg} = \frac{R_{avg}+G_{avg}+B_{avg}}{3} \ldots\ldots\ldots\ldots 5$$

C. Applying the correction to each pixel in the image which is:

$C' = C \times \frac{L_{avg}}{C_{avg}}$ ……. (by applying equation 2, 3, 4 & 5 we get equation 1)

2.4 Image Splitting

Data is essential for the machine learning process to identify and learn patterns, besides the essential training data, it is crucial to have test data to evaluate the algorithm's performance and measure its effectiveness. In our technique, we partitioned the input dataset, allocating 70% of image training data and 30% of image testing data. Data splitting is the act of partitioning available data into two equal parts, typically conducted for the aim of cross - validation. A segment of the data is employed to construct a prediction model, and another segment is used to evaluate the model's efficacy. A crucial step in the analysis of data mining models involves partitioning the data into separate training and testing sets.

2.5 Image Classification

We employ a deep learning technique, specifically VGG16 in our procedure. The Convolutional Neural Network (CNN), also referred as a ConvNet, is a specific form of deep neural network. mostly utilized for the analysis of visual data.

The filter in a Convolutional Neural Network (CNN) is an essential element that may be seen as a set of shared weights used to extract features from images. Various filters can be chosen for different features [8]. Given an input image $X = \{x_{qp}\} \in R^{M\times N}$ and the filter $W = \{W_{uv}\} \in R^{u\times v}$, suppose stride=1 and standard convolution operations yields Y = $\{y_{ij}\}$ as:

$$y_{ij} = \sum_{u=1}^{U}\sum_{v=1}^{V} w_{uv} x_{i-u+1,j-v+1} \ldots\ldots\ldots\ldots\ldots 6$$

Activation functions, including ReLU (rectified linear unit), are employed to augment the nonlinearity of convolutional neural networks (CNNs). Therefore, one can acquire feature maps. Subsequently, a crucial element called pooling is carried out. Pooling is a commonly used technique in which the size of the featured map can be reduced. Two often used types of pooling are average pooling and max pooling. Furthermore, pooling can help prevent the issue of overfitting [21] [25]. The final output image is formed by the entire connection layer after multiple convolutional layers and a pooling layer, which enables high - level reasoning [18] [22]. In this layer, the convolutional layers' local characteristics are gathered and the neurons between different layers are fully connected, similar to a regular artificial neural network (ANN). CNNs are employed in several sectors like as image and video recognition, recommender systems, image classification, and medical image analysis [9]. CNNs are widely used in tasks like facial recognition, object detection, and video analysis due to their ability to capture spatial hierarchies in images and can be used in collaborative filtering and content - based recommendation, particularly when dealing with visual data [8] [10].

VGG16 is a CNN architectural deep convolution model created at the University of Oxford in 2014 by the Visual Geometry Group (VGG) [18]. It's very efficient in tasks like image classification, object detection etc. VGG16 is composed of 16 layers that include weights, rationale behind its specific designation of suffix 16. These layers are including 13 convolutional layers and 3 dense layers. Additionally, it contains 5 max - pooling layers. The overall number of layers in the model is 21, but there are only 16 weight layers, which are the layers with learnable parameters [19]. This VGG16 architecture played significant role in the success of deep learning in computer vision, particularly in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2014, where it achieved top performance [20 - 24].

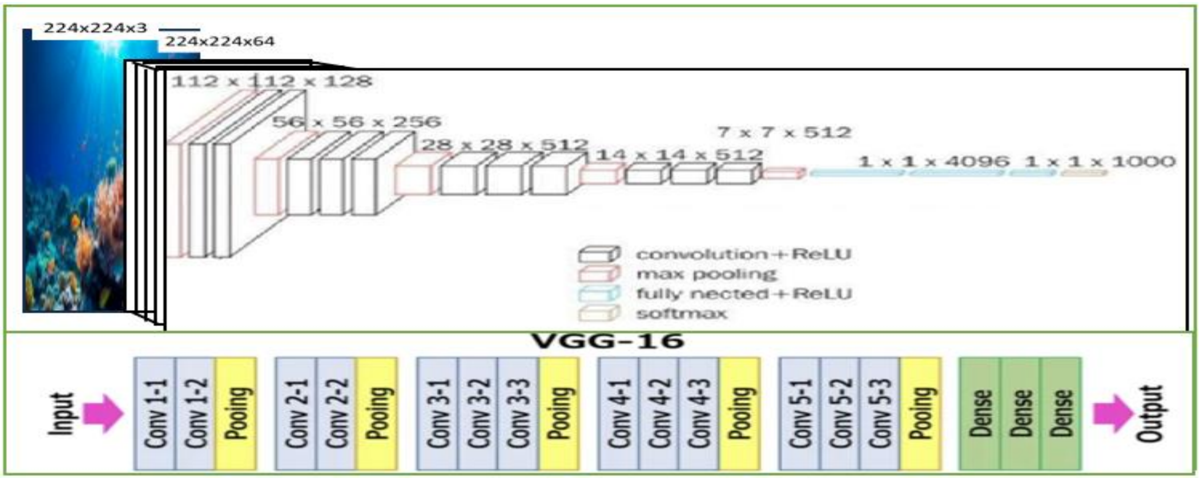

Figure 2 (a): VGG16 architecture

Paper ID: SR24910130120 DOI: https://dx.doi.org/10.21275/SR24910130120

3. Experimental Result

The ultimate outcome will be produced based on the comprehensive categorization and prediction. Underwater environments introduce various challenges like light scattering, color distortions, and low contrast. The deep and methodical architecture of VGG16 offers a notable edge in dealing with intricate image classification tasks when compared to other models as this system is able to autonomously acquire hierarchical characteristics from unprocessed data, resulting in exceptional performance in intricate undertakings like image identification and categorizations. R - CNN prioritizes region - based recommendations for object detection, but VGG16 stands out in feature extraction thanks to its deep and consistent architecture consisting of 16 layers. The increased depth of VGG16 enables it to successfully learn intricate characteristics, resulting in higher performance in tasks related to picture categorization. ResNet incorporates residual connections to tackle the problem of disappearing gradients in really deep networks. On the other hand, VGG16, which has a simpler and deeper architecture without residual connections, often achieves high accuracy in image classification with lower computational complexity. This makes VGG16 easier to implement and fine - tune. Random Forest is a conventional machine learning technique that performs well with organized data but faces difficulties when dealing with complex, unorganized input like hazy photographs. VGG16, a deep architecture specifically intended for processing picture data, utilizes deep learning to autonomously acquire hierarchical features, surpassing Random Forest in tasks such as image recognition and classification. Hence after comparing different AI algorithms for identifying underwater photographs, we concluded some findings which are shown in the table below:

Table 3.1

Ref. No.	Method/ Algorithm	Dataset	Preprocessing Techniques	Performance Metrics	Key Findings
[11]	Random Forest	Standard dataset	Feature scaling	Accuracy: 63.00%	Random Forest performance is lower, supporting CNN's effectiveness.
[14]	Simple CNN	Standard dataset	None	Accuracy: 65.00%	Demonstrates the baseline performance of simple CNN on standard datasets.
[17]	ResNet	Custom dataset	Image resizing, grayscale, white balance	Accuracy: 76.29%	ResNet showing lower accuracy, so this results into demonstrating the RCNN, VGG16 supremacy.
[25]	RCNN	Custom dataset	Image resizing, grayscale, white balance	Accuracy: 81.37%	RCNN shows moderate accuracy compared to VGG16.
- - - - - -	VGG16 (present work)	Custom dataset	Image resizing, grayscale, white balance, original image	Accuracy: 85.12% (VGG16)	VGG16 achieves the highest accuracy on the custom dataset.

Table 3.1 Showing an assessment of the model's performances based on comparison of accuracy percentage

Thus, after experimental findings shown in table 3.1, we opted to go for our proposed VGG16 as its deep architecture helps in learning robust features that are less sensitive to colour distortions and scattering, leading to better object detection performance. The effectiveness of this proposed method is assessed utilizing performance measure like accuracy. The accuracy of a classifier is a measure of its ability to properly predict the class label. Similarly, the accuracy of a predictor refers to how well it can estimate the value of a predicted attribute for fresh data.

Accuracy = (TP + TN) / (TP + TN + FP + FN) …….7
Where:

TP (True Positives): Total number of events accurately classified as positive.
TN (True Negatives): The count of events accurately classified as negative.
FP (False Positives): The total count of cases that have been inaccurately classified as positive (also referred as a Type I error.)
FN (False Negatives): The total count of events that were inaccurately classified as negative, (sometimes referred to as Type II mistake.)

For preprocessing the images, we have used Resize and grayscale approaches for consistency in feature extraction as well as simplification for edge detection and texture analysis as shown in Fig3 (b), Fig3 (c). Further we have used white balance for reducing colour temperature of light source so as to enhance the picture before convolutional process for object identification as illustrated in Figure 3 (d).

The layers in VGG16 can effectively capture spatial hierarchies, which is beneficial for detecting and localizing objects that might be partially obscured or camouflaged in underwater scenes as we can see the final output of our proposed system detecting object correctly with an adequate level of accuracy as shown in Fig3 (f).

The input to generated output sequences through our proposed approach for object detection has been given below:

International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2022): 7.942

Step1: Uploading the underwater image by selecting the particular from the dataset

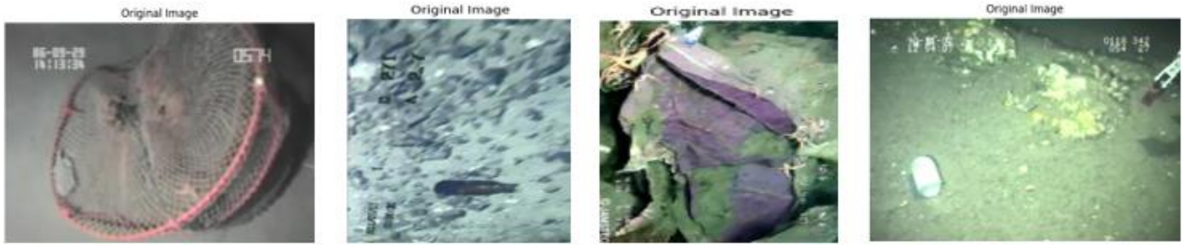

Figure 3 (a): Showing original underwater images selected for detection from dataset

Step 2: Resizing of image

Figure 3 (b): Shows the resized images used for Consistency in Feature Extraction

Step 3: Gray scaling of image

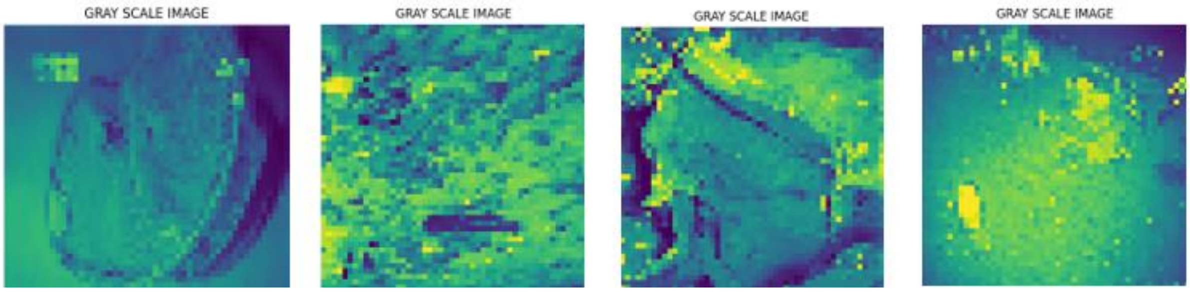

Figure 3 (c): Shows Gray Scale images utilised for reducing complexity in edge detection & texture analysis

Step 4: White balancing of image

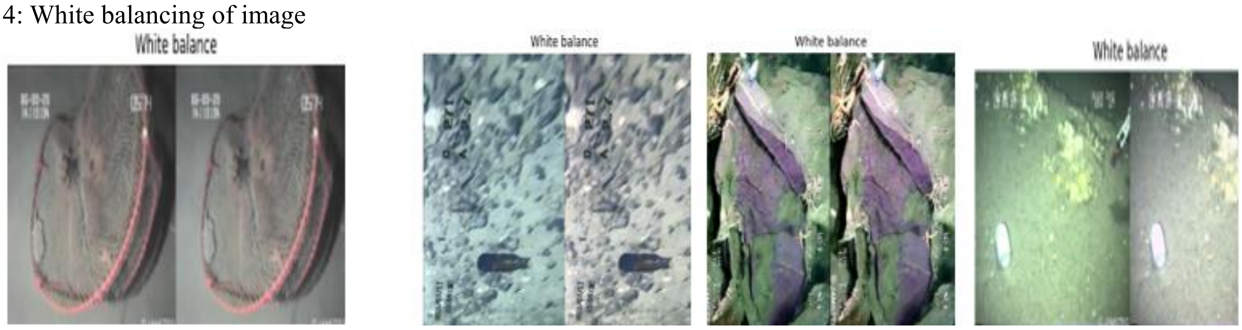

Figure 3 (d): shows White Balance Images used for reducing the colour temperature of the light source

Step 5: Object depiction in photograph enclosed by a red boundary in the balanced image

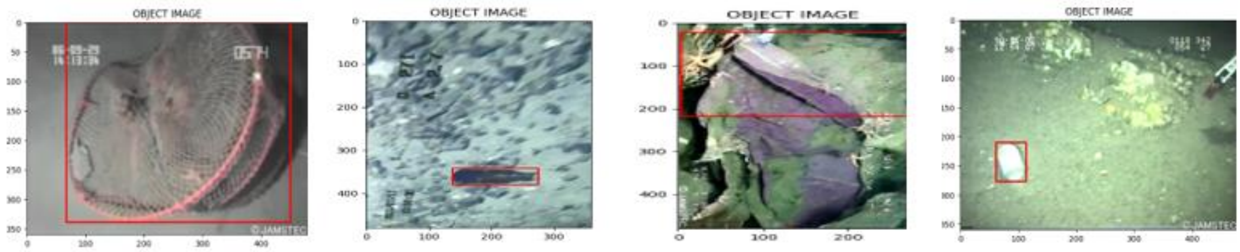

Figure 3 (e): shows object captured in red box to be detected within the balanced Images

Paper ID: SR24910130120 DOI: https://dx.doi.org/10.21275/SR24910130120

Step 6: Finaly, the conclusive outcome after applying VGG16 convolutional method

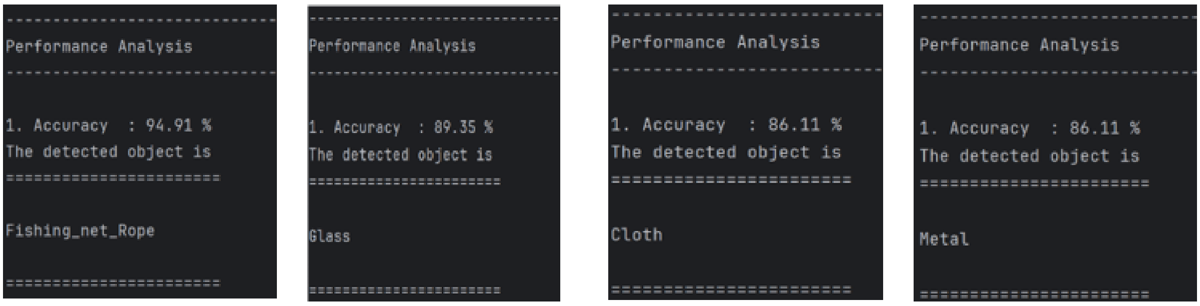

Figure 3 (f): Shows the detected object results of corresponding images after convolutional (VGG16) approach with accuracy percentage

4. Conclusion

This research introduces a deep Convolutional Neural Network (VGG16) approach for the classification and detection of underwater photographs. We conducted a comparative analysis, focussing on crucial factors and emphasising their similarities & disparities. After conducting a comprehensive analysis of the literature related to datasets and their associated works as well as topics pertaining to the design and optimisation of CNNs, we conclude this paper with a concise reference to the future research challenges. Deep learning algorithms require a significant amount of data to operate efficiently for attaining a high level of precision. Regarding the network, several of the current datasets are limited in size and discriminative among a limited number of species. There is a requirement for comprehensive, accurate and uniform datasets, encompassing various other elements. Efficient and economical methods necessitate for any research, keeping this in mind we created datasets of diverse underwater pictures from different data repository. We conclude that, here we have implemented image reducing complexities & the image enhancing approaches by using grayscale and white balance techniques following which we have applied VGG16 algorithm to get an optimal image detection with significant amount of accuracy. In future initiatives, the combination of VGG16 architecture for extracting classified features and Long Short - Term Memory networks (LSTM) for modelling sequences can be highly effective. This is especially advantageous if your data includes a temporal aspect, such as video frames.

References

[1] H. Qin, X. Li, J. Liang, Y. Peng, and C. Zhang, “Deepfish: Accurate underwater live fish recognition with a deep architecture, ” Neurocomputing, vol.187, pp.49–58, 2016.

[2] X. Li and Z. Cui, “Deep residual networks for plankton classification, ” in OCEANS. IEEE, 2016, pp.1–4.

[3] X. Guo, X. Zhao, Y. Liu, and D. Li, “Underwater Sea cucumber identification via deep residual networks, ” Information Processing in Agriculture, vol.6, no.3, pp.307–315, 2019.

[4] X. Sun, J. Shi, J. Dong, and X. Wang, “Fish recognition from low resolution underwater images, ” in CISP - BMEI, 2016, pp.471–476.

[5] M. A. M. Martija, J. I. S. Dumbrique, and P. C. Naval Jr, “Underwater gesture recognition using classical computer vision and deep learning techniques, ” Journal of Image and Graphics, vol.8, no.1, 2020.

[6] Mittal, S., Srivastava, S. and Jayanth, J. P., 2022. A survey of deep learning techniques for underwater image classification. IEEE Transactions on Neural Networks and Learning Systems, 34 (10), pp.6968 - 6982.

[7] Y. Gonzalez - Cid, A. Burguera, F. Bonin - Font, and A. Matamoros, “Machine learning and deep learning strategies to identify posidonia meadows in underwater images,” in OCEANS, 2017, pp.1–5.

[8] Zhu, S., Luo, W. and Duan, S., 2022. Enhancement of Underwater Images by CNN - Based Colour Balance and Dehazing. Electronics, 11 (16), p.2537.

[9] Qin, H., Li, X., Liang, J., Peng, Y., & Zhang, C. (2016). "Deep fish: Accurate underwater live fish recognition with a deep architecture. " Neurocomputing, 187, 49 - 58.

[10] Perez, J., Attanasio, A. C., Nechyporenko, N., & Sanz, P. J. (2017). "A deep learning approach for underwater image enhancement. " In Biomedical Applications Based on Natural and Artificial Computing: Proceedings of the International Work - Conference on the Interplay Between Natural and Artificial Computation, IWINAC 2017 (pp.183 - 192). Springer International Publishing.

[11] Mampitiya, L. I., Nalmi, R. and Rathnayake, N., 2022, July. Performance comparison of sea fish species classification using hybrid and supervised machine learning algorithms. In 2022 Moratuwa Engineering Research Conference (MERCon) (pp.1 - 6). IEEE.

[12] Chen, X., Yuan, M., Fan, C., Chen, X., Li, Y., & Wang, H. (2023). "Research on an underwater object detection network based on dual - branch feature extraction. " Electronics, 12 (16), 3413.

[13] Jian, M., Yang, N., Tao, C., Zhi, H., & Luo, H. (2024). "Underwater object detection and datasets: a survey. " Intelligent Marine Technology and Systems, 2 (1), 9.

[14] Akhtarshenas, A. and Toosi, R., 2022. An open - set framework for underwater image classification using autoencoders. SN Applied Sciences, 4 (8), p.229. Wang, B., Ma, C., Liu, B., Liu, N. and Zhu, J., 2023. Filter Pruning For CNN With Enhanced Linear Representation Redundancy. arXiv preprint arXiv: 2310.06344.

[15] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). "Going deeper with convolutions. " In

Proceedings of the IEEE conference on computer vision and pattern recognition (pp.1 - 9).
[16] Gonzalez - Cid, Y., Burguera, A., Bonin - Font, F., & Matamoros, A. (2017, June). "Machine learning and deep learning strategies to identify posidonia meadows in underwater images. " In OCEANS 2017 - Aberdeen (pp.1 - 5). IEEE.
[17] Mahmood, A., Bennamoun, M., An, S., Sohel, F. and Boussaid, F., 2020. ResFeats: Residual network based features for underwater image classification. Image and Vision Computing, 93, p.103811.
[18] Dakhil, R. A. and Khayeat, A. R. H., 2022. Review on deep learning technique for underwater object detection. arXiv preprint arXiv: 2209.10151.
[19] Zhao, L. and Lee, S. W., 2023. Multi - Domain Rapid Enhancement Networks for Underwater Images. Sensors, 23 (21), p.8983.
[20] Dimitrovski, I., Kitanovski, I., Kocev, D. and Simidjievski, N., 2023. Current trends in deep learning for Earth Observation: An open - source benchmark arena for image classification. ISPRS Journal of Photogrammetry and Remote Sensing, 197, pp.18 - 35.
[21] Johansen, T. H. and Sørensen, S. A., 2020. Towards detection and classification of microscopic foraminifera using transfer learning. arXiv preprint arXiv: 2001.04782.
[22] Guterres, B., Sbrissa, K., Mendes, A., Meireles, L., Novoveska, L., Vermeulen, F., Martinez, J., Garcia, A., Lain, L., Smith, M., & Drews, P. (2023, July). "HAB detection within aquaculture industry: A case study in the Atlantic Area. " In 2023 IEEE 21st International Conference on Industrial Informatics (INDIN) (pp.1 - 6). IEEE.
[23] Paraschiv, M., Padrino, R., Casari, P., Bigal, E., Scheinin, A., Tchernov, D. and Fernández Anta, A., 2022. Classification of underwater fish images and videos via very small convolutional neural networks. Journal of Marine Science and Engineering, 10 (6), p.736.
[24] Drews - Jr, P., Souza, I. D., Maurell, I. P., Protas, E. V. and C. Botelho, S. S., 2021. Underwater image segmentation in the wild using deep learning. Journal of the Brazilian Computer Society, 27, pp.1 - 14.
[25] Dulhare, U. N. and Ali, M. H., 2023. Underwater human detection using faster R - CNN with data augmentation. Materials today: proceedings, 80, pp.1940 - 1945.