PRINT ISSN 1119-8362
Electronic ISSN 2659-1502

Full-text Available Online at
https://www.ajol.info/index.php/jasem
http://www.bioline.org.br/ja

J. Appl. Sci. Environ. Manage.
Vol. 27 (2) 243-247 February 2023

Underwater Image Enhancement for Instance Segmentation using Deep Learning Models

[*1]BELLO, RW; [1]OLUIGBO, CU; [1]PEANOCK, IG; [2]MORADEYO, OM

*[*1]Department of Mathematics and Computer Science, University of Africa, 561101 Sagbama, Bayelsa State, Nigeria*
[2]Department of Computer Science, Adeseun Ogundoyin Polytechnic, Eruwa, Oyo State, Nigeria

[*]Corresponding Author Email: sirbrw@yahoo.com
Co-Authors Email: chineduoluigbo@gmail.com; izonfadebless@gmail.com; tokenny2003@yahoo.com

ABSTRACT: Underwater instance segmentation greatly depends on color-blended underwater images. In this work, a combination of Generalized Color Fourier Descriptor (GCFD), Convolutional Neural Network (CNN) and Mask Region-based Convolutional Neural Network (Mask R-CNN) models were employed to generate a mask for each bounding-boxed Region of Interest (ROI) to obtain enhanced individual underwater segmented images from their complex background accurately. By this image enhancement approach, individual underwater instances are segmented from their complex background accurately. The Patch-based Contrast Quality Index (PCQI) evaluation of our proposed image enhancement method (GCFD) after conducting experiment on the employed datasets shows performance accuracy of 1.1336, which is higher than the 1.1126 performance accuracy achieved by the Contrast-enhancement Algorithm (CA).

DOI: **https://dx.doi.org/10.4314/jasem.v27i2.9**

Cite this paper as: BELLO, R. W; OLUIGBO, C. U; PEANOCK, I. G; MORADEYO O. M. (2023). Underwater Image Enhancement for Instance Segmentation using Deep Learning Models. *J. Appl. Sci. Environ. Manage.* 27 (2) 243-247

Dates: Received: 11 January 2023; Revised: 11 February 2023; Accepted: 14 February 2023
Published: 28st February 2023

Keywords: convolutional neural networks; deep learning; image enhancement; underwater image

Color distortion and optical transmission of underwater images vary from scenes to scenes. The various constrains employed by the existing work could not adapt to the cross-scene, resulting to the need for new constrains to enhance underwater images for further applications such as image segmentation. Moreover, color distortion (Oh and Kim, 2017; Shi et al, 2015; Barron, 2015; Bianco et al, 2015; Bianco et al, 2017; Cheng et al, 2014) and foggy imaging (Wu et al, 2021; Qin et al, 2022; Ren et al, 2018; Engin et al, 2018; Cai et al, 2016) negatively affect underwater images due to their interconnectedness, which makes it very difficult for image features extraction and the mapping of underwater images to color-modified images. Also, the assumption of local smoothness that underwater images are enhanced when there is constant in variation that exists in a small region does not hold much water due to the interference caused by the tiny texture contained in the image patches, which makes it difficult to achieve accurate result from the training process. Massive training labeled datasets for training a deep model for underwater image enhancement are not available (Wang et al, 2017). The filtering properties of the water are controlled and influenced by the water quality. Also, some reasonable amount of light is absorbed by the water molecules, thereby resulting in underwater images getting darker due to reduction in light and with respect to increase in depth; this is in addition to gradual drop-off in colors. The color drop-off takes turns depending on the

[*]Corresponding Author Email: sirbrw@yahoo.com

color wavelength. However, the blue color, due to its shortest wavelength has the ability to resist any interference on its way traveling through the depth of the water, thereby dominating the underwater images (Zhou et al, 2022). The amount of light that is reflected vibrates in a definite pattern and to some extent enters the water in another pattern, which is vertical pattern. Vertical polarization has the important characteristics of making the object less bright, thereby helping in capturing the impossible to capture deep colors. Also, the density of the seawater being denser than the density of the air up to 800 times negatively affects underwater images by neither allowing the light that goes to the water from the air to reflect back nor to enter the water (Zhou et al, 2018). In order to address this perception about underwater image enhancement for accurate instance segmentation, a combination of Generalized Color Fourier Descriptor (GCFD) (Bahri et al, 2017), Convolutional Neural Network (CNN) and Mask Region-based Convolutional Neural Network (Mask R-CNN) (He et al, 2017) models were employed to generate a mask for each bounding-boxed Region of Interest (ROI) to obtain enhanced individual underwater segmented images from their complex background accurately.

MATERIALS AND METHODS

In this section, we address the issues discussed in the previous section by proposing an approach based on image feature descriptor (GCFD) and feature extractor (CNN) for object instance segmentation (Mask R-CNN). We also present the datasets and the system specifications for performing the experiment. One hundred (100) clear images were acquired from the public domain with their pixels randomly disrupted at training and testing stages on local patches. From each image, fifty (50) patches of m×m size were randomly extracted. Five (5) groups of label were uniformly sampled for each patch to synthesize 5 underwater patches, thereby producing five thousand (5000) underwater patches for the training of Mask R-CNN model. The lack of a powerful Graphics Processing Unit (GPU) and memory has constrained the number of training and validation datasets used in this study.

The system specifications for carrying out the experiment are as follows: (1) Software; 64-bit Windows 10 Operating System, Jupyter IDE, and Open CV Python library, (2) Hardware; Intel Core i5 processor@2.4GHz CPU, 16 Gigabytes RAM, GeForce GTX 1080 Ti Graphics card, 2 Terabytes hard-disk, and 10.1 inch IPS HD Portable LCD Gaming Monitor PC display VGA HDMI interface for PS3/PS4/XBOx360/CCTV/Camera. To evaluate the model after training on the training dataset, the trained model is applied to the test dataset (benchmark video) using the pre-trained COCO weights of the Mask R-CNN model, and mean Average Precision (mAP) is used for evaluating the performance accuracy of the model. "The formula for mAP is as follows

$$\mathrm{AP} = \sum_{n=1}^{N} [\mathrm{R(n)} - \mathrm{R(n-1)}].\max \mathrm{P(n)} \quad (1)$$

Where N is the number of precision-recall points produced, P(n) and R(n) are the precision and recall with the lowest nth recall respectively.

$$m\mathrm{AP} = \frac{1}{\mathrm{N}} \sum_{i=1}^{N} AP_i \quad (2)$$

Where AP_i = the *AP* of class *i*, and *N* = the number of classes".

GCFD Approach for Underwater Image Enhancement: There are different values in the feature vectors with each feature vector holding unique information that helps in matching an object or retrieving information based on the images; this process depends on the type of feature descriptors employed. Multiple dimensions are used to represent some values in feature vectors and various elements are described by these feature vectors such as edge, shape and color of an object. Although the detected specific points in an image are described by a feature descriptor, and with a feature vector representing a set of features, there are notable issues with the image quality. The proposed approach is shown in Fig. 1.

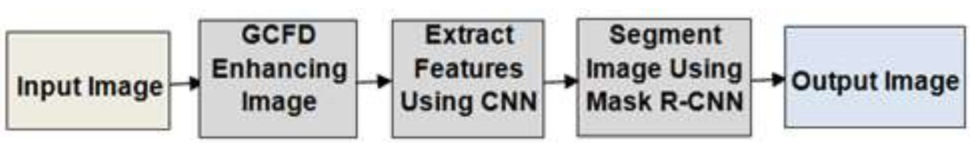

Fig. 1. Methodology for underwater image enhancement for instance segmentation

GCFD is applied to the images to reduce notable pixel problem with the existing methods in handling RGB, Hue, Saturation, Intensity and contrast based on splitting method on RGB for implicative three Fast Fourier Transforms (FFTs) and three sets of descriptors to achieve parallel and orthogonal values, with their combination generating a set of feature maps that represent the image itself. The descriptor serves multiple functions by enabling ROI detection on the image handled by the Region Proposal Network

(RPN) before aligning for classification. By this application, the Clifford Fourier transform technique is generated of each instance for mask segmentation based on different angles. The descriptors are used for the instance identification especially the image patches identification which are produced from the detected points, as each instance has several unique sets of descriptors. This operation supports the enhancement of the image without altering the image pixel, whereby the color of each image is split into separate channels of red, green and blue before fusing them as a unique vector, thereby implying three Fast Fourier Transform (FFT) and three sets of descriptors. "The combination of the two projections is defined by the following formula:

$$GCFD_B(f) = \{GCFD\|_B(f) + GCFD^{\perp}_B(f)\} \quad (3)$$

Where $GCFD_B$ (f) = Computation of GCFD, $GCFD\|_B$(f) = GCFD in parallel part, and $GCFD^{\perp}_B$(f) = GCFD in orthogonal part".

CNN and Mask R-CNN Approaches for Underwater Instance Segmentation: CNN is an integral part of the Mask R-CNN model that helps in extracting the image features for further experimental procedure and analysis. In Mask R-CNN, a Feature Pyramid Network (FPN) based Residual Network (ResNet101) (Li et al, 2016a) is used as the model backbone for the generation of region proposal and ROIAlign operation. ResNet101 of the Mask R-CNN is composed of five stages that are equivalent to five different scales of feature map, namely C1, C2, C3, C4 and C5 which are used in the establishment of the FPN feature pyramid to obtain new features, namely P1, P2, P3, P4 and P5 in that order.

The hyper parameters for the proposed Mask R-CNN in carrying out the experiment are as follows: learning rate is 0.001, weight decay is 0.0001, momentum of learning is 0.90, the minimum and maximum dimensions of the image is 500, size of batch is 200, minimum confidence of detection is 0.50, number of batches is 5, and epochs is 5, iterations per epoch is 5, steps per epoch is 1000, validation step is 5, mask shape is 28×28, and number of anchor classes (instance and background) is 2. Due to limitation of the Mask R-CNN algorithm to detection and segmentation of only patches-free objects in an image, GCFD is employed as a preprocessing of the image patches caused by light absorption and scattering to achieve good accuracy of the image edges irrespective of the patches in the image. The hybrid deep learning model is shown in Fig. 2.

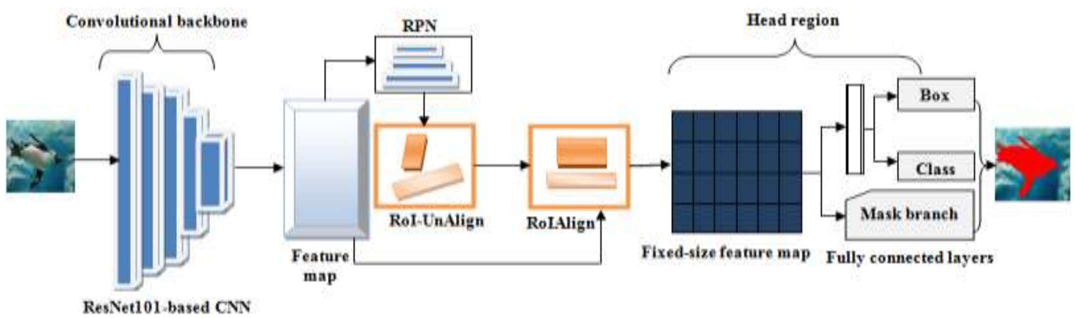

Fig. 2. Methodology for underwater instance segmentation using Mask R-CNN (Bello et al, 2021a; Bello et al, 2021b)

RESULTS AND DISCUSSION

Based on our materials and methodologies, we have devised the framework for segmenting underwater instances by initially enhancing the images to correct color distortion for contrast adjustment. The results of the image enhancement and instance segmentation are presented in this section. The performance evaluation results of GCFD and other algorithms such as Contrast-enhancement Algorithm (CA) (Li et al, 2016b) using Patch-based Contrast Quality Index (PCQI) (Wang et al, 2015), and the mAP result for Mask R-CNN on the test images are shown in Table 1 in which our proposed GCFD method achieved on a par with CA. The qualitative comparison between images before and after applying GCFD and CA is shown in Fig. 3; Fig. 3(d) shows the instance segmentation of the enhanced images in Fig. 3(c). While higher mAP of an image signifies instance precision accuracy in that image, higher PCQI signifies better contrast of an image. As shown in Fig. 3(b), the result of applying CA method on the underwater images still produced color distortion; this is improved by applying our proposed method GCFD, which removed the color distortion, whereby the contrast is increased as shown in Fig. 3(c). With the results obtained in this section, it is convincing that our proposed GCFD method for underwater image enhancement could correct color misjudgment, and Mask R-CNN for underwater instance segmentation could segment (by masking) image instances enhanced by the GCFD method for overall preservation of underwater image details

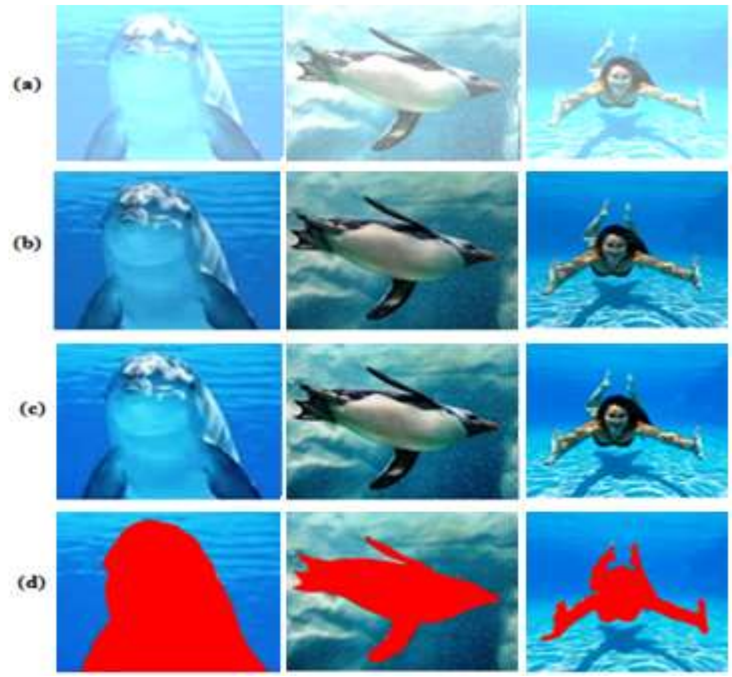

Fig. 3. Qualitative comparison of the results between (a) raw underwater images and after applying (b) CA method (c) GCFD method and (e) Mask R-CNN method. GCFD and Mask R-CNN are the proposed methods

Table 1. PCQI of GCFD and CA

Evaluator	GCFD	CA	Mask R-CNN
PCQI	1.1336	1.1126	-
mAP	-	-	0.94

Conclusion: Underwater image enhancement for instance segmentation using GCFD and hybrid deep learning models has been proposed in this paper. GCFD is applied to the images to reduce notable pixel problem with the existing methods in handling RGB, Hue, Saturation, Intensity and contrast based on splitting method on RGB for implicative three Fast Fourier Transforms (FFTs) and three sets of descriptors to achieve parallel and orthogonal values, with their combination generating a set of feature maps that represent the underwater image itself. CNN, being an integral part of the Mask R-CNN model helps in extracting the image features. Because real-time monitoring of fish, shrimp and algae is important to the aquaculture industry, our future work will intensify research effort in that area by developing a more performing underwater image restoration method using powerful Graphics Processing Unit (GPU) and memory, and further the evaluation of the proposed methods in this paper.

REFERENCES

Bahri, H; Sayadi, F; Khemiri, R; Chouchene, M; Atri, M (2017). Image feature extraction algorithm based on CUDA architecture: case study GFD and GCFD. IET Computers & Digital Techniques, 11(4): 125-132

Barron, JT (2015). Convolutional color constancy. In Proceedings of the IEEE International Conference on Computer Vision (pp. 379–387)

Bianco, S; Cusano, C; Schettini, R (2015). Color constancy using cnns. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (pp. 81–89)

Bianco, S; Cusano, C; Schettini, R (2017). Single and multiple illuminant estimation using convolutional neural networks. IEEE Transactions on Image Processing, 26(9): 4347-4362

Bello, RW; Mohamed, ASA; Talib, AZ (2021a). Enhanced mask R-CNN for herd segmentation. International Journal of Agricultural and Biological Engineering, 14(4): 238-244

Bello, RW; Mohamed, ASA; Talib, AZ (2021b). Contour extraction of individual cattle from an image using enhanced mask R-CNN instance segmentation method. IEEE Access, (9): 56984-57000

Cai, B; Xu, X; Jia, K; Qing, C; Tao, D (2016). Dehazenet: An end-to-end system for single image haze removal. IEEE Transactions on Image Processing, 25(11): 5187–5198

Cheng, D; Prasad, DK; Brown, MS (2014). Illuminant estimation for color constancy: why spatial-domain methods work and the role of the color distribution. *J. Optical Soc. America A*, 31(5): 1049–1058

Engin, D; Genç, A; Kemal Ekenel, H (2018). Cycle-dehaze: Enhanced cycleGAN for single image dehazing. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (pp. 825-833)

He, K; Gkioxari, G; Dollár, P; Girshick, R (2017). Mask R-CNN. In Proceedings of the IEEE International Conference on Computer Vision (pp. 2961-2969)

Li, K; Hariharan, B; Malik, J (2016a). Iterative instance segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3659-3667)

Li, CY; Guo, JC; Cong, RM; Pang, YW; Wang, B (2016b). Underwater image enhancement by dehazing with minimum information loss and histogram distribution prior. IEEE Transactions on Image Processing, 25(12): 5664-5677

Oh, SW; Kim, SJ (2017). Approaching the computational color constancy as a classification problem through deep learning. Pattern Recognition, (61): 405-416

Qin, X; Wang, Z; Bai, Y; Xie, X; Jia, H (2020). FFA-Net: Feature fusion attention network for single image dehazing. In Proceedings of the AAAI Conference on Artificial Intelligence, 34(7): 11908-11915

Ren, W; Ma, L; Zhang, J; Pan, J; Cao, X; Liu, W; Yang, MH (2018). Gated fusion network for single image dehazing. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 3253-3261)

Shi, W; Loy, CC; Tang, X (2016). Deep specialized network for illuminant estimation. In European Conference on Computer Vision (pp. 371-387). Springer, Cham

Wang, S; Ma, K; Yeganeh, H; Wang, Z; Lin, W (2015). A patch-structure representation method for quality assessment of contrast changed images. IEEE Signal Processing Letters, 22(12): 2387-2390

Wang, Y; Zhang, J; Cao, Y; Wang, Z (2017). A deep CNN method for underwater image enhancement. In IEEE International Conference on Image Processing (ICIP) (pp. 1382-1386)

Wu, H; Qu, Y; Lin, S; Zhou, J; Qiao, R; Zhang, Z; Xie, Y; Ma, L (2021). Contrastive learning for compact single image dehazing. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 10551-10560)

Zhou, J; Yang, T; Chu, W; Zhang, W (2022). Underwater image restoration via backscatter pixel prior and color compensation. Engineering Applications of Artificial Intelligence, (111): 104785

Zhou, Y; Wu, Q; Yan, K; Feng, L; Xiang, W (2018). Underwater image restoration using color-line model. IEEE Transactions on Circuits and Systems for Video Technology, 29(3): 907-911