Underwater Image Enhancement and Object Detection Using Edge Preserving and Multiscale Contextual Neural Network

Suvedha.J

ECE Department,

E.G.S Pillay engineering college, Nagapattianam.

Suvedha475@gmail.com.

ABSTRACT

The underwater observation situations cause great challenges to the problem of object detection from the low-resolution underwater images. In this paper, we introduce an efficient technique to enhance the images captured underwater and degraded because of the medium scattering and absorption. It builds on the blending of 2 images that are directly derived from a color compensated and white-balanced version of the original degraded image. After enhancing the underwater image, aims to detect object that present in the underwater by using novel edge preserving and multiscale contextual neural network.we focused mainly on detection of an object in the underwater that they are used to separate them an object from the background by using a combination of automatic contrast stretching followed by image arithmetic operation, global threshold, and minimum filter. Our methodology could be a single image approach that doesn't need specialized hardware or knowledge about the underwater conditions or scene structure. our enhanced images are characterized by better exposedness of the dark region, improved global contrast and edge sharpness and our salient object detection achieves both clear detection boundary and multi-scale contextual robustness simultaneously thus achieves an optimized performance.

Index Terms— **Underwater enhancement,Salient object detection, edge preserving, multi- scale context, RGB-D saliency detection, object mask.**

I. INTRODUCTION

Underwater surroundings offer several rare attractions like marine animals and fishes, superb landscape, and mysterious shipwrecks. Besides underwater photography, underwater imaging has also been an important supply of interest in numerous branches of technology and research , like review of underwater infrastructures and cables , detection of non-natural objects , management of underwater vehicles , marine biology analysis , and archaeology .Technology advances in manned and remotely operated submersibles allow people to collect images and videos from a wide range of the undersea world. Waterproof cameras have become popular, allowing people to easily record underwater creatures while snorkelling and diving. These images or videos often suffer from color distortion and low contrast due to the propagated light attenuation with distance from the camera, primarily resulting from absorption and scattering effects. Therefore, it is desirable to develop an effective method to restore color and enhance contrast for these images.Different from common images, underwater images suffer from poor visibility ensuing from the attenuation of the propagated light, primarily due to absorption and scattering effects. The absorption considerably reduces the light energy, whereas the scattering causes changes within the light-weight propagation direction. They result in foggy look and distinction degradation,making distant objects misty. practically, in common ocean water images, the objects at a distance of quite ten meters are almost undetectable, and also the colors are faded as a result of their composing wavelengths are cut according to the water depth. There are several makes an attempt to revive and enhance the visibility of such degraded images. Since the deterioration of underwater scenes results from the mixture of multiplicative and additive processes traditional enhancing techniques like gamma correction, histogram equalization seem to be powerfully restricted for such a task. salient object detection, which aim to detect object that most attracts people’s attention through out an image,has been wide exploited in recent years. , this paper introduces a completely unique approach to remove the haze in underwater images based on one image captured with a traditional camera.

As illustrated in Fig. 1, our approach builds on the fusion of multiple inputs, however derives the 2 inputs to mix by correcting the distinction and by sharpening a white-balanced version of one native input image. The white equalization stage aims at removing the colour solid induced by underwater light scattering, therefore as to provide a natural look of the sub-sea images. The multi-scale implementation of the fusion method leads to an artifact-free blending it's also been wide used for several computer vision tasks, like semantic segmentation [1], object tracking [2], [3] and image classification [4], [5]. Traditional saliency ways aim to get a heat map which provides every pixel a relative value of its level of saliency [6]–[8]. In recent years, the fashion moves

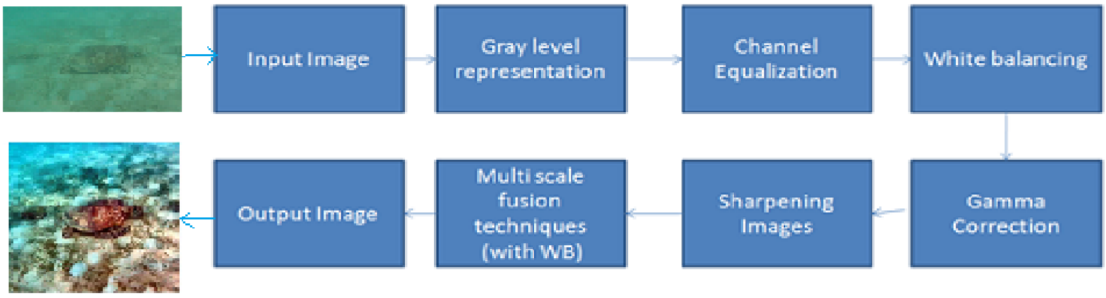

Fig. 1. Enhancement Method overview

to salient object detection that generates pixel-wise binary label for salient and non-salient objects [9]–[11].

In examination with the warmth map, the binary label would any profit segmentation primarily based applications like linguistics segmentation [1], and therefore attracts additional attention.To achieve a high accuracy for binary labeling, there ar in the main 2 requirements: initial, multi-scale discourse reliability; and second, sharp boundary between salient and non- salient objects. And therefore the clear boundary aims to separate the salient object and background clearly and to spotlight the total object uniformly. sadly, none of the prevailing strategies attain each necessities at the same time. Traditional bottom-up strategies mainly consider priors or assumptions and handmade features. for instance, center-surround difference [6], [12], individuality previous [13], [14] and backgroundness previous [15], [16]. These strategies can't think about high-level semantic discourse rela- tions and don't deliver the goods a satisfying accuracy.Recently, the deep Convolutional Neural Network (CNN) has attracted wide attention for its superior performance. CNN primarily based strategies are often divided into region-based networks and pixel-based networks. Region-based strategies aim to extract options of every region (or patch), so predict its strikingness score. However, existing region-based strategies lack of representing context info to model the link between regions and international scenes.In this paper, we propose a novel edge preserving and multi-scale contextual network for salient object detection. The proposed framework achieves both clear boundary and multi-scale contextual robustness simultaneously for the first time. As illustrated in Fig. 2, the planned structure, named RexNet, is principally composed by 2 components, the RegionNet and therefore the contextNet. First, the RegionNet is impressed by the quick R-CNN framework [17]. quick R-CNN is recently planned for object detection and achieves superior performance as a result of the convolutional options of entire image square measure shared and features of every patch (or RoI) square measure extracted via the RoI pooling layer. we have a tendency to extend quick R-CNN to salient object detection by introducing mask-based RoI pooling and formulating salient object detection as a binary region classification task. The image is 1st divided into regions and square measure used as input of RegionNet, the RegionNet then predicts prominence score of every region end-to-end to create prominence map of the complete image. Since the regions square measure divided by edge-preserved ways, prominence map generated by our network is of course with sharp boundaries.

Second, the Contex Net aims to provide strongly reliable multi-scale contextual data. totally different from most pre- vious works that consider context by increasing region window at an exact layer, during this paper, we tend to take into account to model context via multiple abstraction scales. This is primarily based on the observation that different layers of CNN represent different levels of linguistics [18], [19], considering context of various levels is also additional sufficient . we tend to succeed this by taking benefits of dense image prediction. For all max-pooling layers of RegionNet, we tend to attach multiple convolutional layers to predict strikingness map of various levels. Then all levels of strikingness map are consolidated with Region Net to get the ultimate strikingness map. Our technique generates strikingness map with correct location whereas keeping fine object boundaries. apart from the effectiveness, our projected frameworks is economical, since we tend to take benefits of regions by extending the economical quick R-CNN framework, that predicts strikingness score of regions by only 1 forwarding. we tend to conjointly extend our technique to RGB-D strikingness by applying depth refinement. Experiments on two RGB-D benchmark datasets demonstrate that the projected RexNet outperforms alternative ways by an outsized margin. the most contributions of this paper ar three-fold. First, we tend to projected RegionNet that generates strikingness score of regions expeditiously and preserves object boundaries. Second, multi-scale abstraction context is taken into account and hooked up to RegionNet to boost salient object detection performance. Third, we tend to extend our technique to RGB-D strikingness datasets and use depth information to any refine saliency maps.

II. RELATED WORK

In this section, we introduce traditional salient detection methods and the recent CNN based methods. In addition, we introduce some related works that integrate multi- scale context information and related to salient object detection.

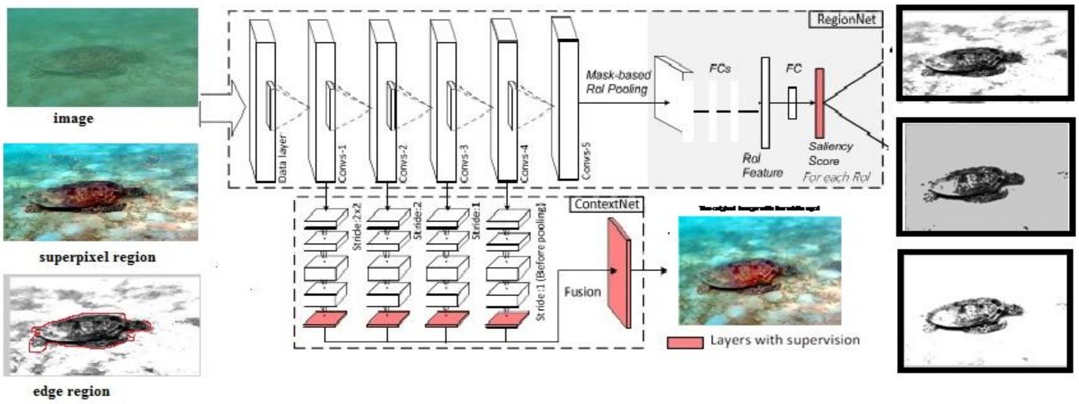

Fig.2. proposed architecture

A. Traditional Methods

Salient object detection was 1st exploited by Itti et al. [6], and later attracted wide attention within the computer vision society. ancient strategies largely admit prior assumptions and most are un-supervised. Center-surround difference that assumes that salient regions differs from their close regions is a vital previous in early analysis. Itti et al. [6] 1st planned center-surround distinction at totally different scales to compute salience. Liu et al. [12] propose center-surround bar chart that defines salience because the difference between center region and its close region. Li et al. [20] propose cost-sensitive SVM to be told and see salient regions that are totally different from their close regions. These strategies cannot offer sharp boundary for salient region as a result of they're supported parallelogram regions, that is barely ready to generate coarse and fuzzy boundary.

Global distinction based mostly ways are later proposed, e.g., Cheng et al. [10] and Yan et al. [21]. In [10], image is 1st divided into superpixels. Then salience price of every region is outlined because the distinction with all alternative regions. The distinction is weighted by abstraction distance so that close regions have bigger impact on it. To contend with objects with complicated structures, Yan et al. [21] propose a stratified model that analyzes salience cues from multiple scales supported local contrast and so infers the ultimate salience values of regions by optimizing them in a very tree model. Following them, several ways utilizing bottom-up priors square measure projected, readers are encouraged to seek out a lot of details in a very recent survey paper by Borji et al. [11].

B. CNN Based Methods

Deep Convolutional Neural Network (CNN) has attracted a lot of attention for its outstanding performance in the high-level linguistics. Here, we have a tendency to mention are few representative work. These work will be divided into 2 classes in step with their treatment of input images: region-based strategies and pixel-based strategies. Region-based strategies formulate salient object detection as a part classification task, namely, extract- ing options of regions and predict their strikingness score. whereas pixel-based strategies directly predict strikingness map pixels-to- pixels with CNN..

1) Region-Based Methods: Wang et al. [22] propose to find salient object by desegregation each native estimation and world search with 2 trained networks DNN-L and DNN-G. Zhao et al. [23] take into account world and native context by swing a worldwide and a closer-focused superpixel-centered window to extract options of every superpixel, severally, and so mix them to predict prominence score. Li et al. [24] propose multi-scale deep options by extracting options of every region at 3 scales and so fuse them to get its prominence score. These works area unit region-based that targeted on extracting options of regions and fuse larger scale of regions as context to predict prominence score of every region. This fusion is usually applied at only 1 layer and doesn't come through a optimum performance. additionally, the networks extract options of 1 region for every forwarding that is extremely long.

2) Pixel-Based Methods: Recently,CNN has been applied to pixels-to-pixels dense image prediction, like linguistics segmentation and salience prediction. Long et al. [25] propose absolutely a convolutional network that is trained end-to-end and pixels-to-pixels by introducing fully convolutional layers and a skip design. Chen et al. [26] propose a coarse-to-fine manner within which the primary CNN generates coarse map victimisation the complete image as input and so the second CNN takes the coarse and local patch as input to come up with fine-grained salience map. Li et al. [27] propose a multi-task model supported totally convolutional network. In [27], salience detection task is in conjunction with object segmentation task that is useful for perceiving objects. A Laplacian regularized regression is then applied to refine salience map. However, whereas end-to-end dense salience prediction is economical, the ensuing salience maps ar coarse and with indistinct object boundaries thanks to the presence of Convolutional layers with massive receptive fields and pooling layers.

C. RGB-D Salient Object Detection

RGB-D saliency is an emerging topic and most RGB-D saliency methods are based on fusing depth priors with RGB saliency priors. Ju *et al.* [28] propose RGB-D saliency method based on anisotropic center-surround difference, in which saliency is measured as how much it outstands from sur- roundings. Peng *et al.* [29] propose depth salience with multi- discourse distinction so fuse it with look cues via a multi-stage model. Ren et al. [30] propose normalized depth previous and global-context surface orientation previous supported depth info so fuse them with RGB region distinction priors. Depth distinction could cause false positives in background region, to handle it, in [31], Feng et al. propose native background enclosure feature supported the observation that salient objects tend to be domestically ahead of close regions. To the simplest of our data, existing RGB-D salient object detection square measure all mistreatment hand-loomed options and also the performance isn't optimized.

D. Multi-Scale Context

Multi-scale context has been proved to be useful for image segmentation task [23], [24], [32], [33]. Hariharan *et al.* [32] proposed hypercolumns for object segmentation and fine- grained localization, in which they defined “hypercolumn” at a given input location as the outputs of all layers at that location. Features of different layers are combined and then be used for classification. Zhao *et al.* [23] proposed multi- context network which extracts features of a given superpixel at global and local scale, and then predict saliency value of that superpixel. Li *et al.* [24] proposed to extract features at three scales: bounding box, neighbourhood rectangular and the entire image. Liu *et al.* [33] proposed to use recurrent con- volutional layers (RCLs) [34] iteratively to integrate context information and to refine saliency maps. At each step, the RCL takes coarse saliency map from last step and feature map at lower layer as input to predict a finer saliency map. In this way, context information is integrated iteratively and the final saliency map is more accurate than that predicted from global context.

The proposed *ContextNet* differs from those at two aspects. First, the *ContextNet* is a holistically-nested architecture [36] which predicts saliency map at each branch and fuse them finally. Second, we propose *EdgeLoss* as a supervision which makes the boundary of segmentation result more clear.

E. Fixation Prediction and Semantic Segmentation

Fixation prediction [6]–[8], [35] aims to predict the regions people may pay attention to, and semantic segmen- tation [25], [36] aims to segment objects of certain classes in images. They are topics related to salient object detection, but they also have significant differences. Fixation prediction aims to predict *regions* which most attract people’s attention, while salient object detection focuses on segmenting the most attrac- tive *objects*. For semantic segmentation, saliency detection is a class-agnostic task, whether an object is salient or not is largely depend on its surroundings, while semantic segmenta- tion mainly focuses on segmentation objects of certain classes (*e.g.* 20 classes in PASCAL VOC dataset). So compared with semantic segmentation, context information is more important for saliency detection, and this is the main motivation of our *ContextNet*.

III. PROPOSED ALGORITHM

Input image may under the format of jpg or bmp,Various types of formats such as gif, tif, png etc., Input image should be in the format of RGB. In gray level block the input image is given in the RGB format and the image is converted to Gray scale for the processing speed.

A.Enhancement process

Our image enhancement approach fig 1adopts a two step strategy, combining white balancing and image fusion, to improve underwater images without resorting to the explicit inversion of the optical model. In our approach, white balancing aims at compensating for the color cast caused by the selective absorption of colors with depth, while image fusion is considered to enhance the edges and details of the scene, to mitigate the loss of contrast resulting from back- scattering. We now focus on the white-balancing stage. White-balancing aims at improving the image aspect, primarily by removing the undesired color castings due to various illumination or medium attenuation properties. we built on the multi-scale fusion principles to propose a single image underwater dehazing solution. Image fusion has shown utility in several applications such as image compositing [37], multispectral video enhancement [38], defogging [39] and HDR imaging [40]. , fig 1 describes Input image may under the format of jpg or bmp,Various types of formats such as gif, tif, png etc., Input image should be in the format of RGB. In gray level block the input image is given in the RGB format and the image is converted to Gray scale for the processing speed .The syntax used for the representation is rgb2gray(‘input image,‘format’’).This representation initialize

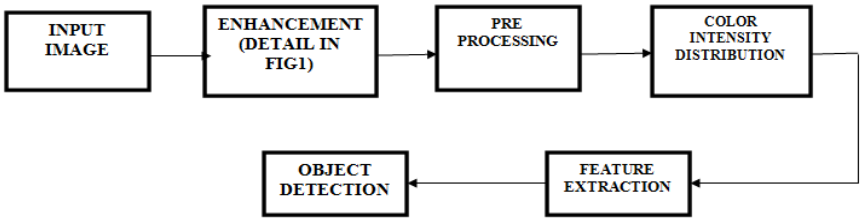

Fig. 3. Block diagram for proposed method

the original image file format. Methods to perform display gamma correction in computing. The pixel's intensity values in a given image file. The binary pixel values are stored in the file in such way that they represent the light intensity via gamma-compressed values instead of a linear encoding.Sharpening an image increases the contrast between bright and dark regions to bring out features.The sharpening process is basically the application of a high pass filter to an image. The following array is a kernel for a common high pass filter used to sharpen an image: We explore a method of multiscale decomposition on an image through the use of high and low pass filters, producing details and approximations respectively.Here, the low pass filter is a normal distribution and the high pass filter is a DoG: Derivative-of-Gaussian

B. Pre-processing

Perform the classification by pre processing the images using Otsu's threholding.With the implementation of Otsu's thresholding, a grey level histogram is created from the gray scale image for noise removal, if a pixel of the grayscale is greater than the threshold value, it is considered to be white, else declared as black. The image provides a sure foreground with the object in focus.

C.Color intensity distribution

Color intensity distribution is done by linear contrast stretching. One of the simplest piece wise linear functions is a contrast stretching transformation. It attempts to improve an image by stretching the range of intensity values it contains tomake full of possible values. Low contrast image can result from poor elimination, lack of dynamic range from preprocessing output.Adjust the image intensity values with a linear contrast stretching.

D. Feature extraction

In image process, during which algorithms are used to observe and isolate numerous desired portions or shapes (feature) of a digitized image or video stream. it's notably important within the space of optical character recognition. The common methods of image feature extraction can be divided into three categories: point feature extraction, linear feature extraction and region feature extraction [41-43]. The region is the pixels set which have a gray scale correlation on the image, where pixels have similar properties, for example the gray value, texture, and so on. However, image feature extraction algorithm is also divided into feature extraction algorithm based on corner and feature extraction algorithm based on invariant technology. The former mainly includes Harris algorithm and SUSAN algorithm [44-45], and the latter mainly includes Scale Invariant Feature Transform (SIFT) [46] and Speeded Up Robust Feature (SURF) [47] algorithm. Image point feature is the special feature of the significant point in the image. Corner is the most common feature point. It has been widely applied in the field of computer vision. Most importantly, the number of corners is much less than the number of all pixels in the image, which meet our need of extract feature points. Corner detection algorithm currently can be summarized into three categories: gray level corner detection [48], binary corner detection and corner detection through curvature scale space [49]. For example, CSS corner detection method [50] is a corner detection method based profile curve. The main idea of SURF algorithm is calculating the Hessian matrix, looking for local maxima to locate the position of the feature point, which is close to the SIFT algorithm in terms of light transform, but in terms of fuzzy invariance, rotational invariance and robustness exceed SIFT algorithm, and which is three times faster in calculation speed than SIFT algorithm, so it is an more excellent feature extraction algorithm

E. Object detection

The object localization is obtained due to automatic edge preserving , histogram equalization and image arithmetic.
STEP 1:Adjust the feature extracted image intensity values with a edge preserving. (L)
STEP2:Perform theHistogram equalization as (H)
STEP 3: Brighten most of the details, Obtain the image R1=L+H
STEP 4:Highlight all the object and its borders in the image. Obtain the image R2=L-H.
STEP5:Remove almost all the other components. Obtain the image R3=R1+R2.
STEP 6: Implement, three times, 3-by-3 minimum filter on the image R3.
STEP7: Convert R3 to binary image using the threshold from step 6.
Step 8: obtain the result

IV. EXPERIMENTAL RESULT

Fig. 4. Input image

Fig. 5. Enhanced image

Fig 6. Preprocessed image

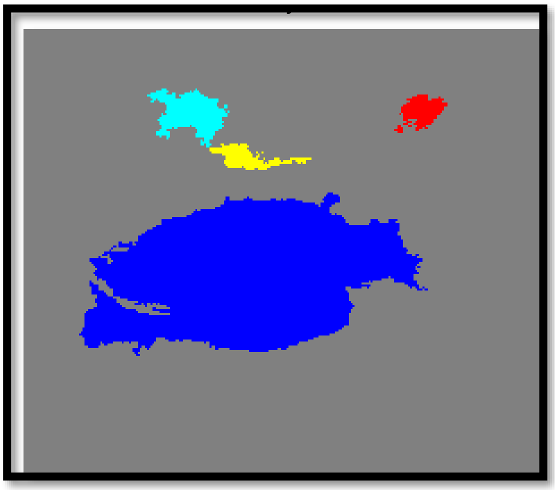

Fig. 7. Color intensity distributed image

Fig. 8. Feature extracted image

Fig.9. Brighten most of the details

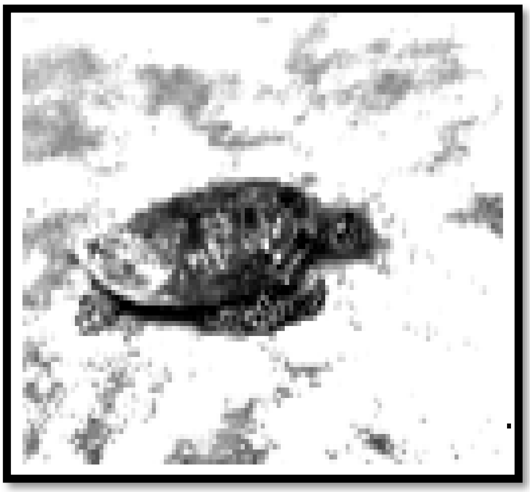

Fig .10.Highlights all the object in the image

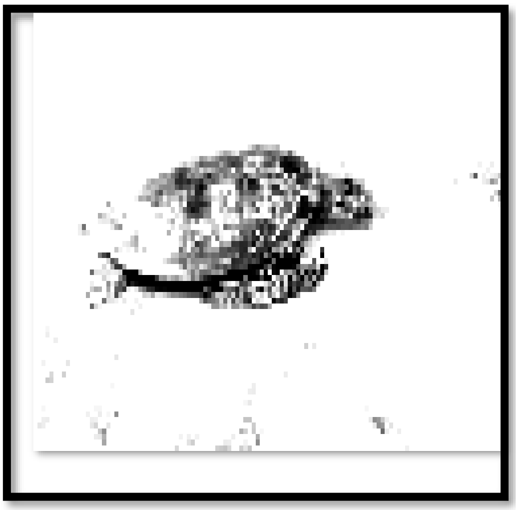

Fig.11. Remove all other components

Fig. 12. Global thresholding

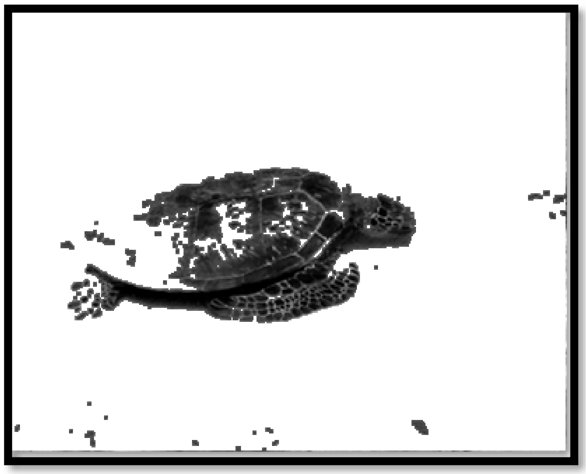

Fig. 13. Final detected image

V. CONCLUSION

In this paper we proposed an alternative approach to enhance underwater images. Our strategy builds on the fusion principle and does not require additional information than the single original image. We have shown in our experiments that our approach is able to enhance a wide range of underwater images with high accuracy, being able to recover important faded features and edges and our salient object detection by using novel edge preserving and multiscale contextual neural network achieves clear detection boundary and multi-scale contextual robustness simultaneously thus achieves an optimized performance. we focused mainly on detection of an object in the underwater that they are used to separate them an object from the background by using a combination of automatic contrast stretching followed by image arithmetic operation, global threshold, and minimum filter.Thus produced better result of underwater salient object detection .

REFERENCE

[1] Y. Wei *et al.* (2015). "STC: A simple to complex framework for weakly-supervised semantic segmentation." [Online]. Available: https://arxiv.org/abs/1509.03150

[2] V. Mahadevan and N. Vasconcelos, "Saliency-based discriminant track- ing," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2009, pp. 1007–1013.

[3] S. Hong, T. You, S. Kwak, and B. Han, "Online tracking by learning discriminative saliency map with convolutional neural network," in *Proc. Int. Conf. Mach. Learn.*, 2015, pp. 597–606.

[4] B. Lei, E.-L. Tan, S. Chen, D. Ni, and T. Wang, "Saliency-driven image classification method based on histogram mining and image score," *Pattern Recognit.*, vol. 48, no. 8, pp. 2567–2580, 2015

[5] B. Li, W. Xiong, O. Wu, W. Hu, S. Maybank, and S. Yan, "Horror image recognition based on context-aware multi-instance learning," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 5193–5205, Dec. 2015.

[6] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 11, pp. 1254–1259, Nov. 1998

[7] L. Zhang, M. H. Tong, T. K. Marks, H. Shan, and G. W. Cottrell, "SUN: A Bayesian framework for saliency using natural statistics," *J. Vis.*, vol. 8, no. 7, p. 32, Dec. 2008.

[8] N. Murray, M. Vanrell, X. Otazu, and C. A. Parraga, "Saliency estimation using a non-parametric low-level vision model," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2011, pp. 433–440.

[9] Y. Zhai and M. Shah, "Visual attention detection in video sequences using spatiotemporal cues," in *Proc. ACM MM*, 2006, pp. 815–824.

[10] M.-M. Cheng, G.-X. Zhang, N. J. Mitra, X. Huang, and S.-M. Hu, "Global contrast based salient region detection," in *Proc. CVPR*, 2011, pp. 569–582.

[11] A. Borji, M.-M. Cheng, H. Jiang, and J. Li. (2014). "Salient object detec- tion: A survey." [Online]. Available: https://arxiv.org/abs/1411.5878

[12] T. Liu, J. Sun, N.-N. Zheng, X. Tang, and H.-Y. Shum, "Learning to detect a salient object," in *Proc. CVPR*, Jun. 2007, pp. 1–5.

[13] K. Shi, K. Wang, J. Lu, and L. Lin, "Pisa: Pixelwise image saliency by aggregating complementary appearance contrast measures with spatial priors," in *Proc. CVPR*, 2013, pp. 2115–2122.

[14] P. Jiang, H. Ling, J. Yu, and J. Peng, "Salient region detection by UFO: Uniqueness, focusness and objectness," in *Proc. ICCV*, 2013, pp. 1976–1983.

[15] Y. Wei, F. Wen, W. Zhu, and J. Sun, "Geodesic saliency using background priors," in *Proc. ECCV*, 2012, pp. 29–42.

[16] W. Zhu, S. Liang, Y. Wei, and J. Sun, "Saliency optimization from robust background detection," in *Proc. CVPR*, 2014, PP 2814–2821.

[17] R. Girshick, "Fast R-CNN," in *Proc. ICCV*, 2015, pp. 1440–1448.

[18] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolu- tional networks," in *Proc. ECCV*, 2014, pp. 818–833.

[19] Y. Li, J. Yosinski, J. Clune, H. Lipson, and J. Hopcroft, "Convergent learning: Do different neural networks learn the same representations?" in *Proc. ICLR*, 2016, pp. 196–212.

[20] X. Li, Y. Li, C. Shen, A. Dick, and A. Van Den Hengel, "Contextual hypergraph modeling for salient object detection," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 3328–3335.

[21] Q. Yan, L. Xu, J. Shi, and J. Jia, "Hierarchical saliency detection," in*Proc. CVPR*, 2013, pp. 1155–1162.

[22] L. Wang, H. Lu, X. Ruan, and M.-H. Yang, "Deep networks for saliency detection via local estimation and global search," in *Proc. CVPR*, 2015, pp. 3183–3192.

[23] R. Zhao, W. Ouyang, H. Li, and X. Wang, "Saliency detection by multi- context deep learning," in *Proc. CVPR*, 2015, pp. 1265–1274

[24] G. Li and Y. Yu, "Visual saliency based on multiscale deep features," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 5455–5463

[25] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 3431–3440.

[26] T. Chen, L. Lin, L. Liu, X. Luo, and X. Li, "DISC: Deep image saliency computing via progressive representation learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 6, pp. 1135–1149, Jun. 2016.

[27] X. Li *et al.*, "DeepSaliency: Multi-task deep neural network model for salient object detection," *IEEE Trans. Image Process.*, vol. 25, no. 8, pp. 3919–3930, Aug. 2016.

[28] R. Ju, L. Ge, W. Geng, T. Ren, and G. Wu, "Depth saliency based on anisotropic center-surround difference," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2014, pp. 1115–1119.

[29] H. Peng, B. Li, W. Xiong, W. Hu, and R. Ji, "RGBD salient object detection: A benchmark and algorithms," in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 92–109.

[30] J. Ren, X. Gong, L. Yu, W. Zhou, and M. Y. Yang, "Exploiting global priors for RGB-D saliency detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2015, pp. 25–32.

[31] D. Feng, N. Barnes, S. You, and C. McCarthy, "Local background enclosure for RGB-D salient object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2343–2350.

[32] B. Hariharan and P. Arbeláez, R. Girshick, and J. Malik, "Hypercolumns for object segmentation and fine-grained localization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 447–456.

[33] N. Liu and J. Han, "DHSNet: Deep hierarchical saliency network for salient object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 678–686.

[34] M. Liang and X. Hu, "Recurrent convolutional neural network for object recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 3367–3375.

[35] J. Zhang and S. Sclaroff, "Saliency detection: A Boolean map approach," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 153–160.

[36] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "Semantic image segmentation with deep convolutional nets and fully connected crfs," in *Proc. Int. Conf. Learn. Represent.*, 2015, pp. 1–2.

[37] M. Grundland, R. Vohra, G. P. Williams, and N. A. Dodgson, "Cross dissolve without cross fade: Preserving contrast, color and salience in image compositing," *Comput. Graph. Forum*, vol. 25, no. 3, pp. 577–586, 2006.

[38] E. P. Bennett, J. L. Mason, and L. McMillan, "Multispectral bilateral video fusion," *IEEE Trans. Image Process.*, vol. 16, no. 5, pp. 1185–1194, May 2007.

[39] C. O. Ancuti, C. Ancuti, and P. Bekaert, "Effective single image dehazing by fusion," in *Proc. IEEE ICIP*, Sep. 2010, pp. 3541–3544.

[40] T. Mertens, J. Kautz, and F. Van Reeth, "Exposure fusion: A simple and practical alternative to high dynamic range photography," *Comput. Graph. Forum*, vol. 28, no. 1, pp. 161–171, 2009

[41] Pauly M, Keiser R, Gross M. Multiscale Feature Extraction on PointSampled Surfaces[C]//Computer graphics forum. Blackwell Publishing, Inc, 2003, 22(3): 281-289.

[42] Nevatia R, Babu K R. Linear feature extraction and description[J]. Computer Graphics and Image Processing, 1980, 13(3): 257-269.

[43] Zaharescu A, Boyer E, Varanasi K, et al. Surface feature detection and description with applications to mesh matching[C]//Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on. IEEE, 2009: 373-380.

[44] Chen J, Zou L, Zhang J, et al. The comparison and application of corner detection algorithms[J]. Journal of Multimedia, 2009, 4(6): 435-441.

[45] Smith S M, Brady J M. SUSAN—A new approach to low level image processing[J]. International journal of computer vision, 1997, 23(1): 45- 78.

[46] Ng P C, Henikoff S. SIFT: Predicting amino acid changes that affect protein function[J]. Nucleic acids research, 2003, 31(13): 3812-3814.

[47] Bay H, Tuytelaars T, Van Gool L. Surf: Speeded up robust features[M]//Computer vision–ECCV 2006. Springer Berlin Heidelberg, 2006: 404-417.

[48] Zheng Z, Wang H, Teoh E K. Analysis of gray level corner detection[J]. Pattern Recognition Letters, 1999, 20(2): 149-162.

[49] [Mokhtarian F, Suomela R. Robust image corner detection through curvature scale space[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 1998, 20(12): 1376-1381.

[50] Bay H, Ess A, Tuytelaars T, et al. Speeded-up robust features (SURF)[J]. Computer vision and image understanding, 2008, 110(3): 346-359.