Underwater Image Enhancement Using D-Cnn

Suman Kumar Swarnkar[1*], Gurpreet Singh Chhabra[2], Abhishek Guru[3], Bhawna Janghel[4], Prashant Kumar Tamrakar[5], Upasana Sinha[6]

Abstract

Due to light dispersion and absorption, underwater photographs often exhibit colour distortion and reduced visibility. Existing techniques make use of a variety of presumptions and constraints to arrive at plausible improvements for underwater picture enhancement. The accepted assumptions may not hold true for some scenarios, which is a typical shortcoming of these methodologies. This research offers an end-to-end architecture for underwater image enhancement to solve this issue and introduces D-CNN, a CNN-based network. Color correction and haze removal are the two exercises used to train the D-CNN. With this dual training method, it is possible to concurrently learn a powerful feature representation for both tasks. The suggested learning framework considerably enhances the convergence speed and accuracy by using a pixel disruptive method to better extract the intrinsic characteristics in local patches. We create 200000 training photos based on the physical underwater imaging model to manage the training of D-CNN. Benchmark underwater photographs were used in experiments to compare D-CNN performance to those of other approaches.

KeyWords: underwater image enhancement, deep CNN, color correction, haze removal

DOI Number: 10.14704/nq.2022.20.11.NQ66212 **NeuroQuantology 2022; 20(11): 2157-2163**

1. Introduction

For applications like monitoring marine life, evaluating the geological environment, and underwater archaeology, the quality of underwater photographs is crucial. Due to the physical characteristics of underwater habitats, it is difficult to capture crisp underwater images. Underwater photos experience colour distortion and poor visibility because of light attenuation and scattering [1].

Image-based and model-based algorithms make up the two main categories of currently used underwater image enhancement techniques [2]. Image-based algorithms [3] [7] estimate the transmission map directly from collected underwater picture and utilise it for colour correction and haze removal. Model-based algorithms [8] [9] analyse underwater optical features to more accurately describe the imaging process. The two types of approaches share the fact that they both make use of different presumptions and constraints. As a result, they also have the same limitation—that the accepted assumptions could not be appropriate for certain particular settings.

In this study, we provide a complete framework for cross-scene underwater picture improvement using convolution neural networks (CNN). Our aim is to learn mapping from underwater photos to color-corrected images and transmission maps, then use the mapping for cross-scene underwater image improvement. We must overcome the following obstacles in order to do this: 1) The colour distortion and optical transmission of various underwater landscapes vary. Previous attempts made a variety of assumptions and constraints that aren't really cross-scene adaptable.For the improvement of underwater images, new learning-based strategies are required. 2) Because colour distortion and cloudy photography are typically related, it is challenging to extract the image elements needed to

Corresponding author: Suman Kumar Swarnkar
Address: [1]Shri Shankaracharya Institute of Professional Management and Technology, Raipur, Chhattisgarh, India, [2] Gandhi Institute of Technology and Management, Visakhapatnam, [3]KoneruLakshmaiah Education Foundation, Vaddeswaram, AP, India, [4] Bharti Vishwavidyalaya, Durg, [5] RSR Rungta College of Engineering and Technology, Bhilai, Chhattisgarh, 6 Guru Ghasidas Vishwavidyalaya, Bilaspur
E-mail: [1]Sumanswarnkar17@gmail.com, [2]Gurpreet.kit@gmail.com, [3]Abhishekguru0703@gmail.com, [4]bhawnajanghel2849@gmail.com, [5]prashant.tamrakar35@gmail.com, [6]drupasanasinha24@gmail.com

NEUROQUANTOLOGY | SEPTEMBER 2022 | VOLUME 20 | ISSUE 11 | PAGE 2157-2163| DOI: 10.14704/NQ.2022.20.11.NQ66212

map underwater photos to color-corrected images and transmission maps. 3). Image improvement heavily relies on the local smoothness assumption, which holds that variance is constant within a small area. However, this assumption will be hampered by the minute texture that makes up picture patches. Directly applying picture patches throughout the training phase makes it challenging to get decent results. 4). It is difficult to find large amounts of labelled data, such as pairs of clean and degraded underwater scene photos, for training a deep model for underwater image improvement.

We provide a CNN-based architecture for underwater image enhancement to overcome the aforementioned problems. The underwater imaging model from [12] is used to produce synthetic underwater pictures for the proposed deep network, known as D-CNN. The D-CNN has two branches: Color Correction Networks (CC-Net) and Haze Removal Networks (HR-Net), which may provide color-corrected images and transmission maps, respectively. The specifics of the proposed D-CNN are shown in Fig. 1.

Contributions

As far as we are aware, this is the first occasion where underwater photographs have been enhanced using CNN. The color-corrected picture and transmission map may be estimated from the input underwater image using the mapping learnt by the suggested networks. Additional labels are not required for target scenes.

A unified learning strategy is used by the proposed D-CNN, which is trained on the colour correction and haze removal tasks. By combining these two goals, it is possible to concurrently learn a robust feature representation for both tasks.

To reduce the interference of minute texture included in local patches, a pixel disrupting approach is adopted. This approach vastly increases the learning process's convergence speed and accuracy.

We synthesise 200000 underwater photos with varying degrees of colour distortion and haze using the underwater photography model from [12]. We use this synthesis dataset to train the D-CNN, which has higher generalisation capabilities on real-world underwater scenes.

2. Proposed Method

2.1. Underwater Imaging Model

The transmission of light dispersed by water-based suspended particles and the direct transmission of light reflected by object surfaces are the two basic components of the photographs taken underwater [1]. The generation of underwater visuals may be mathematically represented as in [12]:

$$I_{ij}^{\lambda} = J_{ij}^{\lambda}\eta_{ij}^{\lambda}t_{ij} + B_{surf}\eta_{ij}^{\lambda}(1-t_{ij}). \qquad (1)$$

where I stands for the underwater photos that were really seen, J for the actual scene that needed to be retrieved, for the wavelength of light, for the overall attenuation of light in the environment, Bsurf for the air light on the water's surface, and Bsurf is set to 1. The transmission map's symbol is T. By substituting the model in [12] with a homogeneous variable tij in three channels, we simplify it. Consequently, only the channel-dependent colour absorption is shown. For underwater picture enhancement, it is crucial to precisely estimate the attenuation coefficient and transmission map.

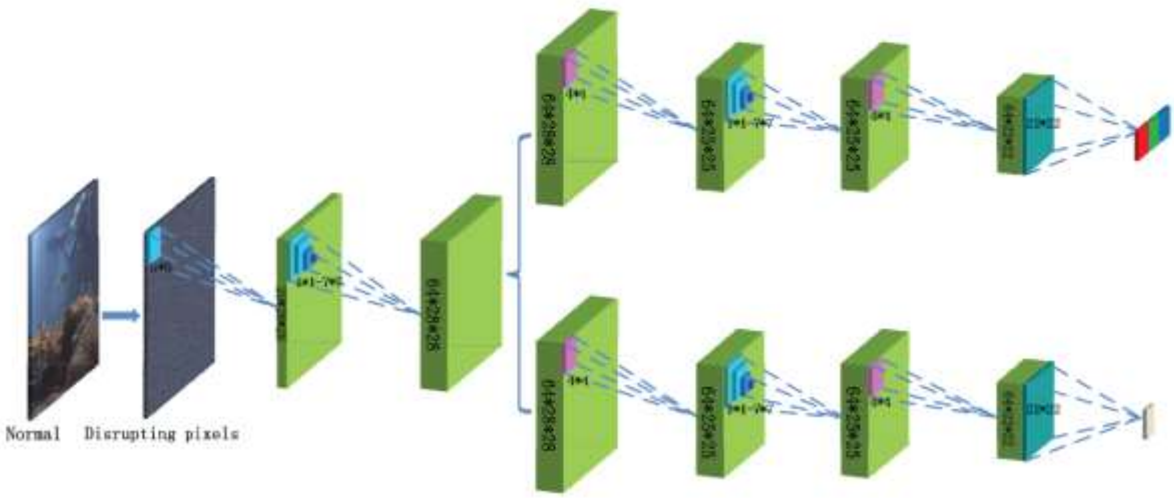

Fig. 1. The architecture of D-CNN.

2.2. Layer Design of D-CNN

In the medium, light attenuation and scattering vary from and interact with one another. As a result, the D-CNN architecture is suggested as shown in Fig. 1, which is composed of three parts: sharing networks (S-Net), CC-Net, and HRNet. Common characteristics are extracted using S-Net. The transmission map t and attenuation coefficient may be estimated via CC-Net and HR-Net, respectively.

Table 1. Configuration of the proposed network

layer	Input Size	Kernel Size	Number	Stride
Conv1	3*32*32	5*5	20	1
Conve2	20*28*28	1*1-7*7	16	1
Max-pooling	64*28*28	4*4	1	1
CC-conv3	64*25*25	1*1-7*7	16	1
Max-pooling	64*25*25	4*4	1	1
CC-conv4	64*22*22	22*22	3	1

S-Net contains two layers and take 32*32 image patches as input. The first convolution layer filters the inputs with 20 kernels each of size 5*5 with a stride of 1 pixel. The second layer contains four kinds of kernels of size 1*1, 3*3, 5*5, 7*7 for diverse receptive fields. We use sigmoid activation for normalizing the output of CC-Net to [0, 1]. The details of HR-Net and CC-Net are shown in Table 1. HR-Net has the same architecture as CC-Net except that it outputs only one transmission coefficient for haze removal.

2.3. Pixel Disrupting Strategy

In a realistic underwater environment, attenuation and scatter smoothly fluctuate. The colour distortion and haze are thus assumed to be constant in picture patches in many earlier works that address colour correction [13] and haze removal [19] issues. D-CNN also uses the local smoothness assumption, which is crucial for improving underwater images. However, interference for colour correction and haze removal may be introduced by the spatial information (such as small texture, coloured noise, etc.) present in picture patches. First, the fine texture will cause certain areas of smoothness to be disturbed. Second, microscopic roughness and noise may damage the learnt characteristics, making it difficult to adjust colours and remove haze. We suggest pixel disrupting technique as a means of suppressing this interference. By destroying every pixel in a patch, we can decrease the interference of fine texture and colour noise without affecting the haze and colour distortion features.

3. Experiments

3.1. Training Setup

The three colour channels' various attenuations caused the underwater pictures to seem blue. Additionally, directly synthesising underwater pictures in RGB colour space proved challenging. As a result, we translated (h, s, and v) from HSV to RGB colour space and evenly sampled t, h, s, and v from the HSV to RGB colour space. Four labels, one for each of the three colour channels' transmission value and attenuation coefficient, were included in a synthetic underwater picture patch. We gathered 200 clean photographs from the Internet and randomly interrupted each of their pixels. From each 32*32 picture, we randomly extracted 100 patches. We randomly selected 10 sets of labels for each patch in order to create 10 undersea patches. As a result, 200000 underwater patches in total were created for D-CNN training. In D-CNN, the bias was adjusted to 0 and the initial kernel weights for each layer were picked at random from the Gaussian distribution N (0, 0.001). The learning rate was initially set at 0.005 and decreased by 50% every 50000 iterations. Using the Euclidean Loss function, 250000 iterations were used to learn the net parameters. The pixels on nearby patches experienced sporadic disruption throughout the testing time. We utilised a guided image filter [25] to smooth the transmission maps in order to reduce the artifacts.

3.2. Model and Performance

We conducted two distinct studies to show the importance of the pixel disrupting method. The training data for one experiment was conventional training data, whereas the second experiment utilised training data with all of the pixels randomly disturbed. UIE-training Net's procedure was shown in Fig. 2 together with two different types of training data.

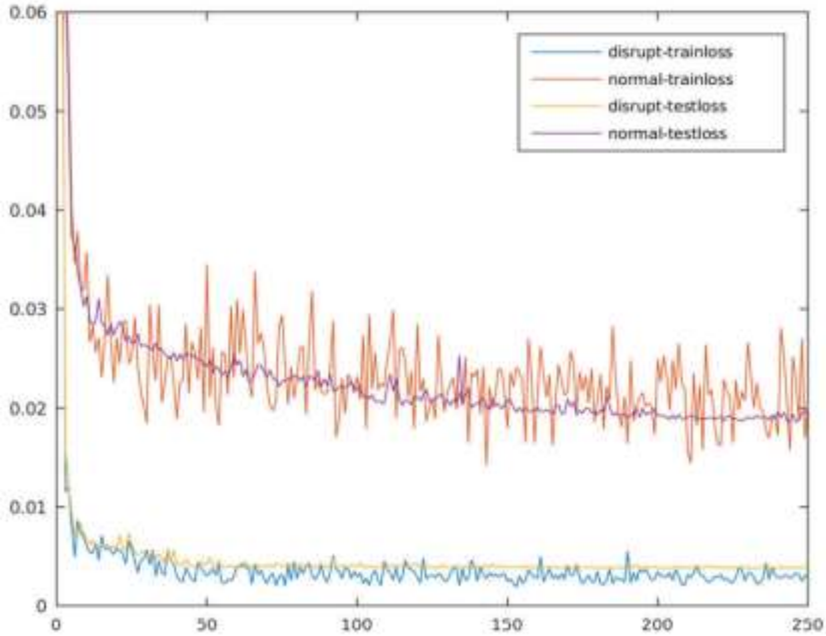

Fig. 2. The training loss and test loss with two kinds of training data.

As shown in Fig. 2, the disruptive method increased the pace of convergence by three times. Compared to utilising regular training data, the training loss and test loss were both much reduced. With common data, the D-CNN converged to a stable state after around 200000 iterations, however with pixel disrupting technique, it only required 50000 iterations. The training loss and test loss using the disruptive technique were decreased by almost 80% when compared to utilising regular data.

3.3. Quantitative results on real underwater images

Enhancing underwater photos was primarily done to fix colour distortion and boost contrast for display and analysis. We utilised the benchmark video to assess UIE-efficacy Net's and robustness in line with the earlier work in [7] [9]. In order to assess the performance, we employed Entropy and the patch-based contrast quality index (PCQI) [26], as reported in [7]. More information is contained in a picture when its entropy is higher. The better the PCQI, the more contrast there is in the picture. Wavelength Compensation and Dehazing (WCID) technique for underwater picture enhancement in [9] and The Dehazing algorithm (DA) and Contrast enhancement algorithm (CA) in [7] were used as comparative algorithms to assess performance. From the video1, 1800 randomly selected underwater photos with a resolution of 720x1280 were taken. These underwater photographs included a variety of settings, including both straightforward and difficult ones (e.g. artificial light, noisy, low contrast). Table 2 displays the test pictures' average values for entropy and PCQI. The above method's average values were taken from [7]. Table 2 demonstrates that our D-CNN outperformed other algorithms in terms of outcomes, proving that it is capable of preserving details while enhancing contrast.

Fig. 3. Subjective result comparison. (a) Raw underwater images in benchmark video. (b) WCID method. (c) DA method. (d) CA method. (e) Our contrast enhancement result.

Table2. Entropy and PCQI of different algorithms.

	WCID	DA	CA	D-CNN
ENTROPY	8.1	8.3	8.6	8.9
PCQI	0.4	0.8	1.1	1.2

3.4. Subjective results on real underwater images

We examined a variety of underwater photos for crossscene, which featured both common sights and complex settings, to further illustrate the reliability and efficiency of our D-CNN. The WCID technique, as shown in Fig. 3(b), reduced the colour cast in the raw pictures, although it did not significantly boost the contrast. As can be seen in Fig. 3(c-d), the dehazeing and final contrast enhancement results of the DA and CA algorithms still had colour distortion. As shown in Fig. 3(e), our suggested technique effectively reduced colour distortion and boosted contrast when compared to other methods. Our outcome was more detailed, less noisy, and had vibrant colours, which was in line with how people perceive things visually. Additionally, we used the suggested D-CNN to improve the calibre of underwater video, and the first row of Fig 4 shows two examples of typical frames. The outcomes of colour correction and haze removal were shown in Fig 4's second and third rows, respectively. The visually appealing findings showed how our suggested D-CNN could adapt to different settings, which was in line with the objective indices in Table 2. You might see the whole video results at http://pan.baidu.com/s/1dE4TdIh.

Fig. 4. The images from top to bottom are raw underwater images, color correction results, and contrast enhancement results of the proposed method.

4. Conclusion

In this study, we suggest a CNN-based underwater image enhancement framework called D-CNN, which consists of two subnetworks: CC-Net and HR-Net. Figure 4 shows the colour absorption coefficients that CC-Net produces for several channels. Raw underwater photographs, colour correction results, and contrast enhancement outcomes of the suggested approach are shown in the images from top to bottom. used to fix underwater photos' colour chromatic aberration. To improve the contrast of underwater photos, HRNet produces a transmission map of light attenuation. Additionally, we first provide a pixel disruptive method that effectively increases convergence speed and accuracy. The D-CNN testing procedure is being carried out using overlapping patches. The computational cost is still quite significant even though it is more effective than checking at every pixel point. Future work for us will concentrate on employing a fully-CNN implementation to increase the effectiveness of the suggested technique.

References

BL McGlamery, "A computer model for underwater camera systems," in Ocean Optics VI. International Society for Optics and Photonics, 1980, pp. 221–231.

Raimondo Schettini and Silvia Corchs, "Underwater image processing: state of the art of restoration and image enhancement methods," EURASIP Journal on Advances in Signal Processing, vol. 2010, no. 1, pp. 746052, 2010.

Stephane Bazeille, Isabelle Quidu, Luc Jaulin, and JeanPhilippeMalkasse, "Automatic underwater image preprocessing," in CMM'06, 2006, p. xx.

Kashif Iqbal, Rosalina Abdul Salam, Mohd Osman, Abdullah Zawawi Talib, et al., "Underwater image enhancement using an integrated colour model.," IAENG International Journal of Computer Science, vol. 32, no. 2, pp. 239–244, 2007.

Virendra Kumar Swarnkar, Dr. AshaAmbhaikar, Suman Kumar Swarnkar, "Big Data Security Enhancement Based Intrusion Detection System Using K-Mean Clustering of Decomposited Features" Journal of Information Technology in Industry, 2021. (ESCI)

Julia Ahl´ ˚ en, David Sundgren, and Ewert Bengtsson, "Application of underwater hyperspectral data for color correction purposes," Pattern Recognition and Image Analysis, vol. 17, no. 1, pp. 170–173, 2007.

Cosmin Ancuti, CodrutaOrnianaAncuti, Tom Haber, and Philippe Bekaert, "Enhancing underwater images and videos by fusion," in Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. IEEE, 2012, pp. 81–88.

Chong-Yi Li, Ji-Chang Guo, Run-Min Cong, Yan-Wei Pang, and Bo Wang, "Underwater image enhancement by dehazing with minimum information loss and histogram distribution prior," IEEE Transactions on Image Processing, vol. 25, no. 12, pp. 5664–5677, 2016.

Yoav Y Schechner and Yuval Averbuch, "Regularized image recovery in scattering media," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29, no. 9, 2007.

John Y Chiang and Ying-Ching Chen, "Underwater image enhancement by wavelength compensation and dehazing," IEEE Transactions on Image Processing, vol. 21, no. 4, pp. 1756– 1769, 2012.

Mahesh M Subedar and Lina J Karam, "Increased depth perception with sharpness enhancement for stereo video," in IS&T/SPIE Electronic Imaging. International Society for Optics and Photonics, 2010, pp. 75241B–75241B.

Yin Zhao, Zhenzhong Chen, Ce Zhu, Yap-Peng Tan, and Lu Yu, "Binocular just-noticeable-difference model for stereoscopic images," IEEE Signal Processing Letters, vol. 18, no. 1, pp. 19–22, 2011.

Shijie Zhang, Jing Zhang, Shuai Fang, and Yang Cao, "Underwater stereo image enhancement using a new physical model," in Image Processing (ICIP), 2014 IEEE International Conference on. IEEE, 2014, pp. 5422–5426.

Dongliang Cheng, Dilip K Prasad, and Michael S Brown, "Illuminant estimation for color constancy: why spatial-domain methods work and the role of the color distribution," JOSA A, vol. 31, no. 5, pp. 1049–1058, 2014.

Suman Kumar Swarnkar, Dr. Asha Ambhaikar, "Improved Convolutional Neural Network based Sign Language Recognition" International Journal of Advanced Science and Technology, Volume 27, Issue 1, pp 302-317, 2019. (Scopus)

Simone Bianco, Claudio Cusano, and Raimondo Schettini, "Color constancy using cnns," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2015, pp. 81–89.

Simone Bianco, Claudio Cusano, and Raimondo Schettini, "Single and multiple illuminant estimation using convolutional neural networks," arXiv preprint arXiv:1508.00998, 2015.

Jonathan T Barron, "Convolutional color constancy," in Proceedings of the IEEE International Conference on Computer Vision, 2015, pp. 379–387.

Agarwal, S., Patra, J. P., &Swarnkar, S. K. (2022). Convolutional neural network architecture based automatic face mask detection. International Journal of Health Sciences, 6(S3), 623–629. https://doi.org/10.53730/ijhs.v6nS3.5401 (Scopus)

Suman Kumar Swarnkar, Dr. Asha Ambhaikar, "Optimized Convolution Neural Network (OCNN) for Voice-Based Sign Language Recognition: Optimization & Regularization" Fifth International Conference on Information and Communication Technology, 2020. (Scopus, Springer)

Suman Kumar Swarnkar, Multimodal Biometric Face and Fingerprint Recognition Using XOR Configuration, International journal of Computing, Volume 4 Issue 4-OCT2014. ISSN [ONLINE] 2230-9039

Wu Shi, Chen Change Loy, and Xiaoou Tang, "Deep specialized network for illuminant estimation," in European Conference on Computer Vision. Springer, 2016, pp. 371–387.

SeoungWug Oh and SeonJoo Kim, "Approaching the computational color constancy as a classification problem through deep learning," Pattern Recognition, vol. 61, pp. 405–416, 2017.

RaananFattal, "Single image dehazing," ACM transactions on graphics (TOG), vol. 27, no. 3, pp. 72, 2008.

Kaiming He, Jian Sun, and Xiaoou Tang, "Single image haze removal using dark channel prior," IEEE transactions on pattern analysis and machine intelligence, vol. 33, no. 12, pp. 2341–2353, 2011.

Jing Zhang, Yang Cao, and Zengfu Wang, "A new image filtering method: Nonlocal image guided averaging," in Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on. IEEE, 2014, pp. 2460–2464.

Jing Zhang, Yang Cao, and Zengfu Wang, "Nighttime haze removal based on a new imaging model," in Image Processing (ICIP), 2014 IEEE International Conference on. IEEE, 2014, pp. 4557–4561.

Suman Kumar Swarnkar, Multimodal Biometric Face and Fingerprint Recognition Using APEX Algorithm and Feed Forward Back Propagation, International Journal Of Scientific Progress And Research, Volume 9 Issue 1-2015. ISSN [ONLINE] 2349-4689

Bolun Cai, Xiangmin Xu, Kui Jia, Chunmei Qing, and Dacheng Tao, "Dehazenet: An end-to-end system for single image haze removal," IEEE Transactions on Image Processing, vol. 25, no. 11, pp. 5187–5198, 2016.

Jing Zhang, Yang Cao, Shuai Fang, Yu Kang, and Chang Wen Chen, "Fast haze removal for nighttime image using maximum reflectance prior," in IEEE CVPR, 2017.

Kaiming He, Jian Sun, and Xiaoou Tang, "Guided image filtering," in European conference on computer vision. Springer, 2010, pp. 1–14.

BindeshwariChandel , Suman Kumar Swarnkar, "Stock Prediction Using Financial Data And News Sentimental Analysis", Journal of Xi an Shiyou University, Natural Science Edition (JXSU), ISSN: 1673-064X, Volume.16, Issue 9, pp 279-285, September 2020. (Scopus)

Shiqi Wang, Kede Ma, Hojatollah Yeganeh, Zhou Wang, and Weisi Lin, "A patch-structure representation method for quality assessment of contrast changed images," IEEE Signal Processing Letters, vol. 22, no. 12, pp. 2387–2390, 2015.