A DEEP CNN METHOD FOR UNDERWATER IMAGE ENHANCEMENT

Yang Wang   *Jing Zhang*   *Yang Cao*   *Zengfu Wang*

Department of Automation, University of Science and Technology of China

ABSTRACT

Underwater images often suffer from color distortion and visibility degradation due to the light absorption and scattering. Existing methods utilize various assumptions/constrains to achieve reasonable solutions for underwater image enhancement. However, these methods share the common limitation that the adopted assumptions may not work for some particular scenes. To address this problem, this paper proposes an end to end framework for underwater image enhancement, where a CNN-based network called UIE-Net is presented. The UIE-net is trained with two tasks, color correction and haze removal. This unified training approach enables learning a strong feature representation for both tasks simultaneously. For better extracting the inherent features in local patches, a pixels disrupting strategy is exploited in the proposed learning framework, which significantly improves the convergent speed and accuracy. To handle the training of UIE-net, we synthesize 200000 training images based on the physical underwater imaging model. Experiments on benchmark underwater images for cross-scenes show that UIE-net achieves superior performance over existing methods.

Index Terms— underwater image enhancement, deep C-NN, color correction, haze removal

1. INTRODUCTION

The quality of underwater images is significantly important for underwater computer vision, such as monitoring sea life, assessing geological environment and underwater archaeology. Capturing sharp underwater image is challenging due to the physical properties of underwater environments. Due to light attenuation and scattering, underwater images suffer from color distortion and low visibility [1].

Existing underwater image enhancement methods may be divided into two main groups: image based and model based algorithms [2]. Image based algorithms [3] [4] [5] [6] [7] estimate the transmission map directly from captured underwater image and then use it for color correction and haze removal. Model-based algorithms [8] [9] [10] [11] [12] take underwater optical properties into considered for better describing the imaging process. The common properties of the two groups of methods are they both utilize various assumptions/constrains. Therefore, they also share the same limitation that the adopted assumptions may not work well for some specific scenes.

In this paper, we propose an end to end framework for cross-scene underwater image enhancement based on the Convolution Neural Networks (CNN). Our goal is to learn mapping from underwater images to color-corrected images and transmission maps, and then to use the mapping for cross-scene underwater image enhancement. To achieve this, we need to overcome the following challenges: 1). Different underwater scenes have different color distortion and optical transmission. Previous works used various assumptions/constrains which have low adaptability for cross-scene. New learning based scheme for underwater image enhancement are needed. 2) Underwater images often suffer from color distortion and hazy imaging. Since there is an interconnection between color distortion and hazy imaging, it is difficult to extract the image features to map underwater images to color-corrected images and transmission maps. 3). Local smoothness assumption, which assumes that the variation is constant in a small region, plays an important role in image enhancement. However, the tiny texture contained in image patches will introduce interference for this assumption. It is difficult to obtain good result by directly using image patches in the training process. 4). For training a deep model for underwater image enhancement, it is not massively available for collecting the labelled data which include the pairs of clear and degraded images of underwater scenes.

To address the above challenges, we propose a CNN-based framework for Underwater Image Enhancement. The proposed deep Network (called UIE-Net) is trained with synthesized underwater images, which are generated by using the underwater imaging model in [12]. UIE-Net has two branches: Color Correction Networks (CC-Net), Haze Removal Networks (HR-Net), which can output color corrected image and transmission map, respectively. Fig.1 shows the details of proposed UIE-Net.

Contributions.

1. To the best of our knowledge, this is the first time for using the CNN to enhance the underwater images. The mapping learned by the proposed networks can be used to estimate the color-corrected image and transmission map from the input underwater image. There is no need for extra labels on target scenes.

2. The proposed UIE-Net applies a unified learning

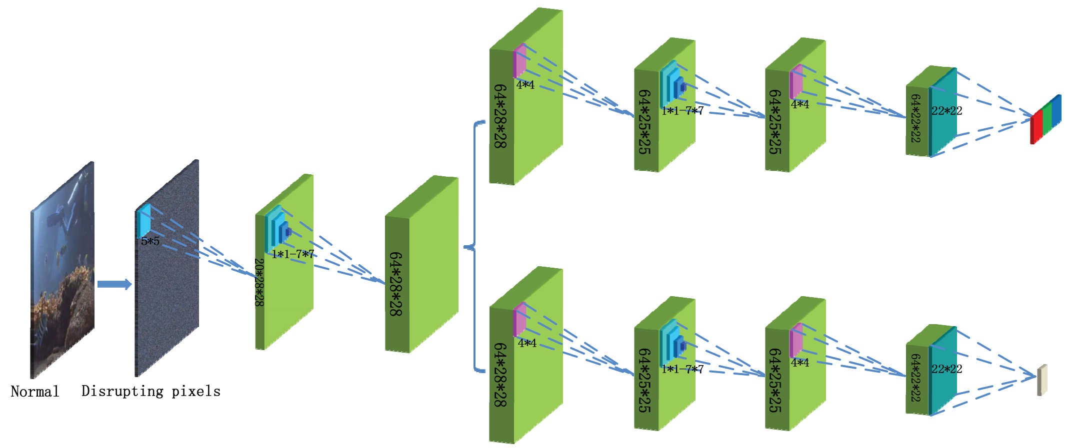


Fig. 1. The architecture of UIE-Net.

scheme, which is trained with two tasks, color correction and haze removal. Formulating these two objectives together enables learning a strong feature representation for both tasks simultaneously.

3. A pixels disrupting strategy is used to suppress the interference of tiny texture contained in local patches. This strategy significantly improves the convergent speed and accuracy of learning process.

4. By using the underwater imaging model in [12], we synthesis 200000 underwater images, which contain various degrees of color distortion and haze. We train the UIE-Net on this synthesis dataset, which however shows superior generalization ability on real scene underwater images.

2. PROPOSED METHOD

2.1. Underwater Imaging Model

The captured underwater images mainly compose two parts [1]: the direct transmission of light reflected by surface of objects, and the transmission of light scattered by suspended particles that exist in water. Mathematically, the formation of underwater images can be modeled 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 is the observed underwater images, J is the real scene to be recovered, λ is the wavelength of light, η^{λ} is the total environment light attenuation coefficient, B_{surf} is the air light on the surface of water , and B_{surf} is set to 1. T is the transmission map. We simplify the model in [12] by replacing t_{ij}^{λ} to a homogeneous variable t_{ij} in three channels. Therefore only η_{ij}^{λ} depicts the channel-dependent color absorption.

Estimating attenuation coefficient η^{λ} and transmission map t accurately is essential important for underwater image enhancement.

2.2. Layer Design of UIE-Net

The attenuation and scattering of light in the medium have both difference and contact with each other. Therefor, the framework of UIE-Net is proposed as shown in Fig.1, which consists three parts: sharing networks (S-Net), CC-Net, HR-Net. S-Net is used for common features extraction. CC-Net and HR-Net can estimate the attenuation coefficient η^{λ} and transmission map t 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

The attenuation and scatter change smoothly at practical underwater environment. Thus, many previous work use local smoothness assumption to address color correction [13] [14] [15] [16] [17] [18] and haze removal [19] [20] [21] [22] [23] [24] problems, which assumes that the color distortion and haze are constant in image patches. The local smoothness assumption is essential important for underwater image enhancement, which is also used in UIE-Net.

However, the spatial information (e.g. tiny texture, colored noisy) contained in image patches may introduce interference for color correction and haze removal. Firstly, the tiny texture will disturb the smoothness in patches. Secondly, the learned features may be affected by tiny texture and noisy, which hinder color correction and haze removal. To suppress these interference, we propose pixel disrupting strategy. We disrupt all of pixels within patches, which can suppress the interference of tiny texture and color noisy without changing any characteristics of color distortion and haze.

3. EXPERIMENTS

3.1. Training Setup

The underwater images appeared bluish because of the different attenuation in three color channels. And it was hard to synthesis underwater images directly in RGB color space. Thus, we uniformly sampled t $\in (0,1)$, h $\in (180,240)$, s $\in (0,1)$, v $\in (0,1)$, and transformed (h, s, v) from HSV to RGB color space. A synthesised underwater image patch contained four labels, which represented the transmission value and attenuation coefficient of three color channels.

We collected 200 clear images from the Internet and disrupted all of pixels randomly in these images. we extracted 100 patches randomly from every images with size of 32*32. For each patch, we uniformly sampled 10 groups of label to synthesis 10 underwater patches. Therefore, a total of 200000 underwater patches were generated for UIE-Net training. In UIE-Net, the initial kernel weights for each layer were sampled randomly from Gaussian distribution N$\sim (0, 0.001)$, and the bias in all layers were set to 0. The initial learning rate was set to 0.005, and it reduced by half for every 50000 iterations. Net parameters had been learned by 250000 iterations with Euclidean Loss function. During the testing period, the pixels were disrupted randomly on local patches. In order to suppress the artifacts in transmission maps, we used guided image filter [25] to smooth the transmission maps.

3.2. Model and Performance

To demonstrate the significance of pixel disrupting strategy, we performed two contrasting experiments. One experiment used the training data with all of pixels disrupted randomly, the other one used the normal training data. Fig.2 presented the training process of UIE-Net with two kinds of training data.

As shown in Fig.2, the disrupting strategy improved to three times faster in convergence speed. Both training loss and test loss were much lower compared with using normal training data. The UIE-Net converged to stable state after about 200000 iterations with common data, but it only needed 50000 iterations with pixel disrupting strategy. Compared with using normal data, the training loss and test loss with disrupting strategy was reduced approximately 80%.

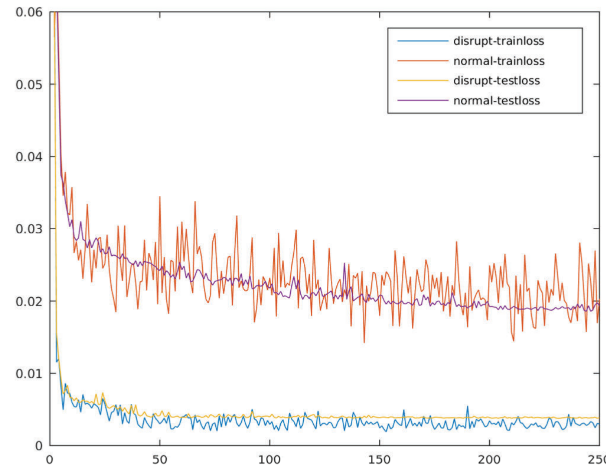


Fig. 2. The training loss and test loss with two kinds of training data.

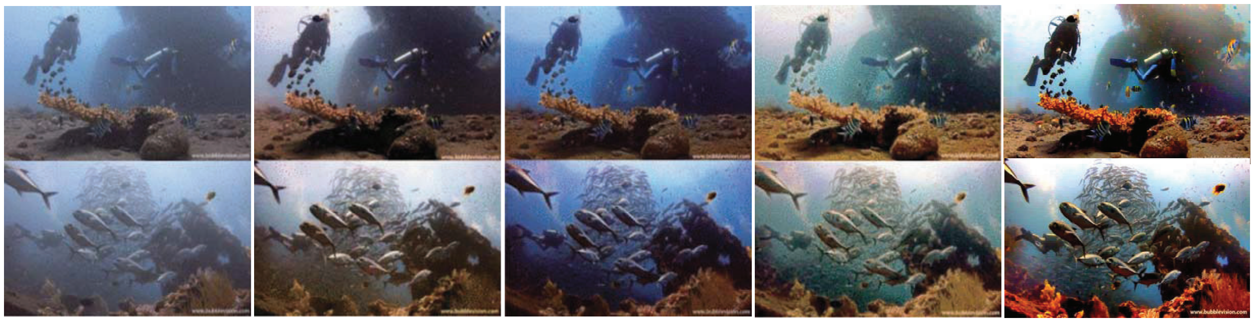
3.3. Quantitative results on real underwater images

The main goal of underwater images enhancement was to correct color distortion and increase the contrast for display and analysis. Following the previous work in [7] [9], we used the benchmark video to evaluate the effectiveness and robustness of UIE-Net. As described in [7], we used Entropy, and patch-based contrast quality index(PCQI) [26] to evaluated the performance . Higher entropy of an image mean more information contained in that image. The higher PCQI mean the better contrast of an image.

To evaluate the performance, we used following algorithms for comparisons: Wavelength Compensation and Dehazing (WCID) algorithm for underwater image enhancement in [9] ,The Dehazing algorithm (DA) and Contrast enhancement algorithm (CA) in [7].

We extracted 1800 underwater images randomly with a size of 720*1280 from the video[1]. These underwater images were cross-scene, which contained both common and challenging scenes (e.g. artificial light, noisy, low contrasty). Table 2 shows the average value of Entropy and PCQI on the test images. The average value of above method were quoted from [7]. As shown in Table 2 , our UIE-Net achieved comparable or better results than other algorithms, which convinced that our UIE-Net could enhance the contrast as well as preserve details.

[1]http://www.youtube.com/user/bubblevision



(a) (b) (c) (d) (e)

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

Table 2. Entropy and PCQI of different algorithms.

	WCID	**DA**	**CA**	**UIE-Net**
Entropy	7.6347	7.4900	7.7763	7.5735
PCQI	0.5215	0.9787	1.1126	1.1229

3.4. subjective results on real underwater images

To further demonstrate the robustness and effectiveness of our UIE-Net, we tested various underwater images for cross-scene,which contained both common scenes and challenging scene. As shown in Fig 3(b), WCID algorithm removed the color cast in the raw images, but it did not improve the contrast effectively. As shown in Fig 3(c-d), the result of DA and CA algorithms still contained the color distortion both in the dehazeing result and the final contrast enhancement result. As shown in Fig 3(e), compared with other algorithms, our proposed method removed color distortion and increased the contrast more efficiently. Our result contained more details, less noises, and vivid colors, which was consistent with human visual perception.

Further more, we applied the proposed UIE-Net to enhance the quality of underwater video, and two typical frames were shown in the first row of Fig 4. The color correction and haze removal results were shown in the second and third row of Fig 4. The visual pleasing results demonstrated the adaptability of our proposed UIE-Net on cross-scenes, which was consistent with the objective indices in Table 2. The whole video results were available on 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 paper, we propose an CNN-based framework called UIE-Net for underwater image enhancement, which contains two subnetworks: CC-Net, HR-Net. CC-Net outputs color absorption coefficients within different channels, which is used to correct the color distortion of underwater images. HR-Net outputs the transmission map of light attenuation, which is used to enhance the contrast of underwater images. Further more, we propose a pixel disrupting strategy for the first time, which improves the convergent speed as well as accuracy efficiently.

Now the testing process of UIE-Net is implemented in a manner of patches overlapping. Though it is more efficient than testing at each pixel location, the computational cost is still a little high. The future work for us will focus on improving the efficiency of the proposed approach by using a fully-CNN implementation.

Acknowledgments. This work is supported by the Natural Science Foundation of China (61472380).

5. REFERENCES

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

[2] 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.

[3] Stephane Bazeille, Isabelle Quidu, Luc Jaulin, and Jean-Philippe Malkasse, "Automatic underwater image preprocessing," in *CMM'06*, 2006, p. xx.

[4] 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.

[5] Julia Åhlén, 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.

[6] Cosmin Ancuti, Codruta Orniana Ancuti, 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.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

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

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

[17] 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.

[18] Seoung Wug Oh and Seon Joo Kim, "Approaching the computational color constancy as a classification problem through deep learning," *Pattern Recognition*, vol. 61, pp. 405–416, 2017.

[19] Raanan Fattal, "Single image dehazing," *ACM transactions on graphics (TOG)*, vol. 27, no. 3, pp. 72, 2008.

[20] 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.

[21] 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.

[22] 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.

[23] 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.

[24] 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.

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

[26] 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.