Shallow-UWnet: Compressed Model for Underwater Image Enhancement (Student Abstract)

Ankita Naik*, Apurva Swarnakar *, Kartik Mittal*
University of Massachusetts Amherst
Amherst, MA-01002
{arnaik, aswarnakar, kartikmittal}@umass.edu

Abstract

Over the past few decades, underwater image enhancement has attracted an increasing amount of research effort due to its significance in underwater robotics and ocean engineering. Research has evolved from implementing physics-based solutions to using very deep CNNs and GANs. However, these state-of-art algorithms are computationally expensive and memory intensive. This hinders their deployment on portable devices for underwater exploration tasks. These models are trained on either synthetic or limited real-world datasets making them less practical in real-world scenarios. In this paper, we propose a shallow neural network architecture, **Shallow-UWnet** which maintains performance and has fewer parameters than the state-of-art models. We also demonstrated the generalization of our model by benchmarking its performance on a combination of synthetic and real-world datasets. The code is available at: github.com/mkartik/Shallow-UWnet/

Introduction

Underwater robotics represents a fast-growing research area. Recently, great efforts are being made in developing autonomous underwater vehicles (AUVs) deployed to tackle challenging engineering problems such as underwater archaeological exploration, garbage collection, underwater rescue operations, ocean floor exploration, and military operations. Many of these applications require real-time interpretation of images/videos for the AUV to intelligently perceive the environment and take follow-up measures. Underwater images that are degraded due to the transmission of light in water could hinder the correct interpretation of the camera input, thereby inhibiting the capability of the vehicle to interact with the environment. Thus, for the above applications, the first step before any downstream task (object recognition, object classification, or object detection) is image enhancement.

The existing literature based on very deep Convolutional Neural Networks (CNNs) and Generative Adversarial Networks (GANs) provide good solutions for automatic image enhancement problems such as color/contrast adjustment, colorization, dehazing, etc. as discussed by Yang et al. (2019). However, the high computational and memory requirement of these models makes them heavy for real-time underwater image enhancement tasks. Thus, to improve the deployability of machine learning models, reduce the compute and memory requirement, and make real-time learning faster we propose **Shallow-UWnet**.

Our approach has multiple highlights: i) Our proposed model is able to achieve similar performance on three synthetic and real-world datasets when compared to the state-of-art models (WaterNet (Li et al. 2019), FUnIE-GAN (Islam, Xia, and Sattar 2020), and Deep SESR (Islam, Luo, and Sattar 2020)). This was achieved by the utilization of Perceptual loss (Johnson, Alahi, and Fei-Fei 2016) and incorporating skip connections into the proposed architecture. ii) Also, utilizing shallow ConvBlocks made of Conv-ReLU-Dropout layers and skip connections in tandem has helped us obtain a model having a compression factor of 18.17 as compared to current state-of-art GANs.

Proposed Method

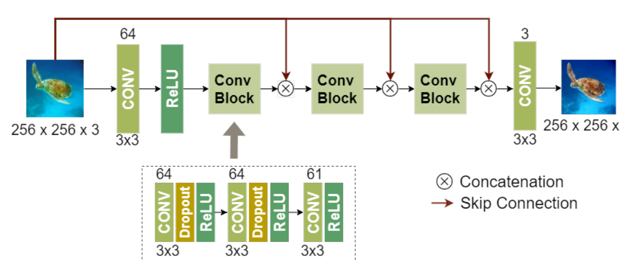

Figure 1: Model Architecture

Figure 1 shows the architecture diagram of our proposed model. The model is comprised of a fully connected convolutional network connected to three dense ConvBlocks in series. The underwater input image (256×256 (RGB)) is being concatenated to the output of each block using a skip connection.

ConvBlocks The ConvBlocks consist of two sets of convolutional layers, each followed by a dropout and ReLU activation function. The output is then passed through another

*equal contribution

Metric	Datasets	WaterNet	FUnIE-GAN	Deep SESR	Shallow-UWnet
PSNR	EUVP-Dark	24.43 ± 4.64	26.19 ± 2.87	25.30 ± 2.63	$\mathbf{27.39 \pm 2.70}$
	UFO-120	23.12 ± 3.31	24.72 ± 2.57	$\mathbf{26.46 \pm 3.13}$	25.20 ± 2.88
	UIEB	19.11 ± 3.68	19.13 ± 3.91	$\mathbf{19.26 \pm 3.56}$	18.99 ± 3.60
SSIM	EUVP-Dark	0.82 ± 0.08	0.82 ± 0.08	0.81 ± 0.07	$\mathbf{0.83 \pm 0.07}$
	UFO-120	0.73 ± 0.07	0.74 ± 0.06	$\mathbf{0.78 \pm 0.07}$	0.73 ± 0.07
	UIEB	$\mathbf{0.79 \pm 0.09}$	0.73 ± 0.11	0.73 ± 0.11	0.67 ± 0.13
UIQM	EUVP-Dark	2.97 ± 0.32	2.84 ± 0.46	2.95 ± 0.32	$\mathbf{2.98 \pm 0.38}$
	UFO-120	2.94 ± 0.38	2.88 ± 0.41	$\mathbf{2.98 \pm 0.37}$	2.85 ± 0.37
	UIEB	$\mathbf{3.02 \pm 0.34}$	2.99 ± 0.39	2.95 ± 0.39	2.77 ± 0.43

Table 1: Underwater Image Enhancement Performance Metric

Models	Compression Rate	Speed-Up
Shallow-UWnet	1	1
WaterNet	3.96	24
Deep SESR	10.17	7
FUnIE GAN	18.17	8

Table 2: Model Compression metric (Cheng et al. 2017)

set of Conv-ReLU pair which facilitates concatenation of the raw image from the skip connection. The series of ConvBlocks along with skip connection acts as a deterrent for overfitting of the network over the train data. Thus, supporting the generalization of the network.

Network Loss The model is trained using a multi-term loss function. It helps preserve the sharpness of the edges, ensure structural and texture similarity of the enhanced image, and to account for the pixel-wise loss. Total loss ($L_{TOTAL} = L_{MSE} + L_{VGG}$) is calculated using the following two loss components:

1. **MSE Loss:** The mean squared error (MSE) loss, L_{MSE} is calculated as the pixel difference between the estimated image and the clear ground truth image.
2. **VGG Perceptual Loss:** As motivated by (Johnson, Alahi, and Fei-Fei 2016) and (Li et al. 2019), we incorporated the perceptual loss, which is based on the ReLU activation layers of the pretrained 19 layers VGG network. The enhanced image and the ground truth image are passed to the last convolutional layer of the pretrained VGG network to get the feature representations. The loss, L_{VGG} is then calculated as the distance between the feature representations of the two images.

Experiments

We used EUVP - Imagenet data for train and validation of our model. Stochastic gradient descent was carried out using Adam optimizer at a learning rate of 0.0002 and dropout of 0.2. The model was benchmarked using the following datasets against state-of-art models listed in Table 1.

Datasets We benchmarked our model performance on three test datasets collected during ocean explorations under diverse conditions:

1. UFO-120 (Islam, Luo, and Sattar 2020): Distorted images were generated using style transfer. Deep SESR was trained on this dataset.
2. UIEB (Li et al. 2019): It is a near real-world dataset wherein reference images were created using 9 different algorithms and voted by volunteers. WaterNet was trained on this dataset.
3. EUVP-Dark (Islam, Xia, and Sattar 2020): Distorted images were created by a model based on GANs. FUnIE-GAN was trained on this dataset.

Evaluation Metric We used standard image enhancement evaluation metrics - PSNR, SSIM and UIQM (Yang et al. 2019). Model compression and acceleration is measured using compression ratio and speedup rate (Cheng et al. 2017).

Results and Conclusion

Our proposed Shallow-UWnet model maintains comparable quantitative performance while requiring 18 times lesser trainable parameters and makes testing 10 times faster. As seen in Table 1 our model is able to generalize on varied datasets emphasizing its real-world application.

References

Cheng, Y.; Wang, D.; Zhou, P.; and Zhang, T. 2017. A survey of model compression and acceleration for deep neural networks. *IEEE Signal Processing Magazine* .

Islam, M. J.; Luo, P.; and Sattar, J. 2020. Simultaneous Enhancement and Super-Resolution of Underwater Imagery for Improved Visual Perception. In *Robotics: Science and Systems (RSS). ArXiv:2002.01155*.

Islam, M. J.; Xia, Y.; and Sattar, J. 2020. Fast Underwater Image Enhancement for Improved Visual Perception. *IEEE Robotics and Automation Letters (RA-L)* 5(2): 3227–3234.

Johnson, J.; Alahi, A.; and Fei-Fei, L. 2016. Perceptual losses for real-time style transfer and super-resolution. In *European conference on computer vision*, 694–711. Springer.

Li, C.; Guo, C.; Ren, W.; Cong, R.; Hou, J.; Kwong, S.; and Tao, D. 2019. An underwater image enhancement benchmark dataset and beyond. *IEEE Transactions on Image Processing* 29: 4376–4389.

Yang, M.; Hu, J.; Li, C.; Rohde, G.; Du, Y.; and Hu, K. 2019. An in-depth survey of underwater image enhancement and restoration. *IEEE Access* 7: 123638–123657.