

# A Novel AlphaSRGAN for Underwater Image Super Resolution

Aswathy K. Cherian\* and E. Poovammal

SRM Institute of Science and Technology, Chennai, India

\*Corresponding Author: Aswathy K. Cherian. Email: aswathy2205@gmail.com

Received: 01 March 2021; Accepted: 02 April 2021

**Abstract:** Obtaining clear images of underwater scenes with descriptive details is an arduous task. Conventional imaging techniques fail to provide clear cut features and attributes that ultimately result in object recognition errors. Consequently, a need for a system that produces clear images for underwater image study has been necessitated. To overcome problems in resolution and to make better use of the Super-Resolution (SR) method, this paper introduces a novel method that has been derived from the Alpha Generative Adversarial Network (AlphaGAN) model, named Alpha Super Resolution Generative Adversarial Network (AlphaSRGAN). The model put forth in this paper helps in enhancing the quality of underwater imagery and yields images with greater resolution and more concise details. Images undergo pre-processing before they are fed into a generator network that optimizes and reforms the structure of the network while enhancing the stability of the network that acts as the generator. After the images are processed by the generator network, they are passed through an adversarial method for training models. The dataset used in this paper to learn Single Image Super Resolution (SISR) is the USR 248 dataset. Training supervision is performed by an unprejudiced function that simultaneously scrutinizes and improves the image quality. Appraisal of images is done with reference to factors like local style information, global content and color. The dataset USR 248 which has a huge collection of images has been used for the study is composed of three collections of images—high ( $640 \times 480$ ) and low ( $80 \times 60$ ,  $160 \times 120$ , and  $320 \times 240$ ). Paired instances of different sizes— $2\times$ ,  $4\times$  and  $8\times$ —are also present in the dataset. Parameters like Mean Opinion Score (MOS), Peak Signal-to-Noise Ratio (PSNR), Structural Similarity (SSIM) and Underwater Image Quality Measure (UIQM) scores have been compared to validate the improved efficiency of our model when compared to existing works.

**Keywords:** Underwater imagery; single image super-resolution; perceptual quality; generative adversarial network; image super resolution

## 1 Introduction

Procuring high quality and clear images of underwater scenes is often difficult owing to the complex nature of the ecosystem and environment present underwater [1]. Different hardware and



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

software can be used to reduce underwater image distortion but this is difficult and expensive to implement. Hence, using Super-Resolution (SR) technology is quite convenient and easy. Several algorithms have been introduced for this purpose. A sparse representation algorithm was proposed by Kumudham and Rajendran [2] who took into account the sparsity of the High Dimension (HD) data from sonar images. The image was first split into Low-Resolution (LR) dictionaries and High Definition (HD) blocks of images. A sparse coefficient along with a dictionary was then employed to represent each block, which was then used to obtain a high-definition image. But, the use of sparse representation and interpolation to achieve Super-Resolution (SR) may result in reduced image information and blurry edges. To rectify this, Lu et al. [3] introduced the use of self-similarity-based SR algorithms. The SR algorithm produces high-resolution images in a scattered form.

A generation model based on deep residual networks for Single Image Super Resolution (SISR) was provided by Islam et al. [4] along with a counter-training pipeline to analyze SISR from paired data. An objective function that oversees counter-training was developed, which evaluated the perceptive quality of the images depending on perceptive quality, color, overall content of the image was developed to supervise the training. In addition, many SR algorithms exploited adversarial training [5] to augment the perceptual quality of the SR image. SISR using GAN (SRGAN) using deep residual networks and skip connections [6] was proposed by Ledig et al. [7]. Furthermore, Liu et al. [8] recommended an image enhancement technique for underwater images using deep residual networks. Synthetic underwater images were utilized for training a Convolutional Neural Network (CNN) model with CycleGAN. ANN has been in trend in recent years and finds its applications in a variety of domains including medicine [9,10]. A model to uplift the resolution and quality of SONAR images taken underwater using Generative Adversarial Network (GAN) was introduced by Sung et al. [11]. A network consisting of 8 convolutional layers and 16 residual blocks was trained using underwater Sound Navigation and Ranging (SONAR) images studied from many angles. The results showed that this model could successfully enhance images and produce a greater Peak Signal to Noise Ratio (PSNR) than the interpolation method. Liu and Li [12] introduced an enhanced method to enhance the problem of gradient disappearance for super-resolution of images (based on gradient penalty and Wasserstein distance). Shamsolmoali et al. [13] introduced a Generative Adversarial Network method that can give comprehensive information with refined image quality and network stability and be trained in steps. Another method for generating images using the SR technique was developed by Xie et al. [14] that used time-coherent 3D volume data and a novel temporal discriminator for identification. An architecture based on residual networks that combined structural information and facial information spectrum to produce better quality SR images was also proposed by Bulat and Tzimiropoulos [15].

In this paper, an attempt to resolve the problems faced previously by introducing a super-resolution model for underwater images for real-time applications is made. The problem defined here is a translation issue (image-to-image) and assumes that a mapping that is non-linear in nature exists between the distorted images (which make up the input data) and the enhanced images (of which the output is made of). Next, an Alpha-SRGAN-based model is designed, which learns the mapping between the two images by adversarial training on the USR-248 dataset. After careful consideration of the design, implementation and experimental results and validations of the model, we make the following inputs in this contribution.

## 2 Related Works

### 2.1 GAN—Generative Adversarial Networks

Several models and techniques have been introduced for unsupervised learning. However, for image generation, GAN [5] has emerged to be the most efficient technique for unsupervised learning. Unlike previously used strategies, high-resolution and vivid images can easily be produced by GANs. In essence, GANs give a strategy to gather information from the implicit distribution of a specified target dataset; X. GAN networks are made up of two blocks of networks—a generator network and a discriminator network. GANs have been making tremendous advances in several applications like image-to-image translation [16], image SR [17], and image and video generation [18]. In the past decade, a lot of consideration has been given to improve convergence stability and the competence to generate high quality images. One such example is the Least Square Generative Adversarial Network (LSGAN) [18] which exploits the Pearson  $\chi^2$  divergence and uses the least square loss for critic output. Arjovsky et al proposed an efficient mechanism to overcome the problems faced in GANs, the Wasserstein-GAN (WGAN) [19]. Two major upgrades were made in WGAN—a weight clipping method and a contemporary aim based on the Earth Mover Distance or the Wasserstein distance. Subsequently, convergence performance of the Wasserstein distance method was found to be better than that of Jensen-Shannon divergence and Kullback-Leibler divergence in [19].

The LR-HR domain is used in DeblurGAN Yuan et al. [20] also proposed a cycle-in-cycle-based GAN model that used unpaired data to train itself. This system was brought around due to the triumph of the DualGAN [21]. But, GAN-based SISR models that were given unpaired training were found to produce inconsistent results and were susceptible to instability. As an attempt to rectify this, a model that learned the noise prior to super-resolution using DualGAN [21] was proposed and was called Super Resolution Dual Generative Adversarial Networks (SRDGAN) [22]. Similarly, generating high-quality images through tensor structures was another approach and was introduced as deep Tensor Generative Adversarial Networks (TGAN) [23]. Contrarily, RankSRGAN [24] has two parts, the SRGAN at the core and the newly introduced Ranker. The model inherits the traits of Super Resolution-GAN and aims to perform better in terms of perceptual metrics. Another model, the Image Sequence Generative Adversarial Network (ISGAN) [25], is an SR method based on aquatic image sequences obtained using multi-focus at similar angles. This method has made obtaining more details and thereby improving the resolution of the images possible.

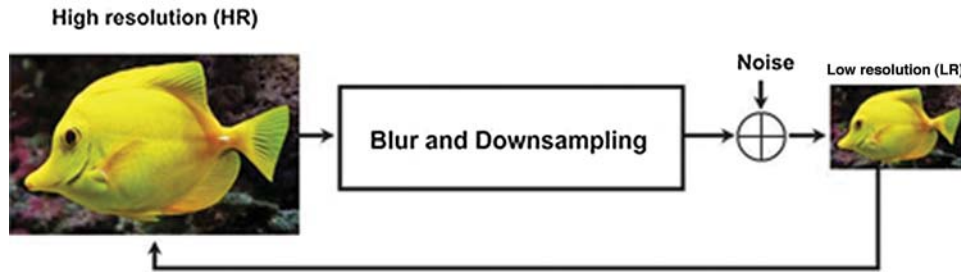
Another work by Yu et al. [26] introduces a multiscale featured fusion generator (MSFF-GAN) based on the GAN network. The network was designed such that the images generated preserved more details and information. Slow extraction of characters by the generator was resolved using a residual dense module. Yuan et al. [27] proposed a Class-condition Attention Generative Adversarial Network (CA-GAN) as an attempt to improve the quality of underwater images. They successfully built a dataset composed of simulated underwater images under varying water depth and attenuation coefficients, split into ten classes. Their model aims at producing a many-to-one mapping function.

### 2.2 SISR for Underwater Imaging

In SISR, an input of low resolution could give rise to multiple high-resolution images, and the HR space that we wish to map the low-resolution image to is typically unmalleable [28]. Consequently, this poses a huge problem since a single LR input could correspond to many HR outputs. Super-Resolution (SR) essentially is the act of reinstating versions of a picture in

high-resolution from several low-resolution images of the same view. Fig. 1 depicts the typical framework of SISR.  $y$  (low-resolution image) is defined as shown in Eq. (1)

$$y = (X \otimes K) \downarrow_S + n, \quad (1)$$



**Figure 1:** Framework of SISR to recover HR image from a LR Image

Here, the Convolution of the unknown high-resolution picture,  $X$ , and  $K$  (blurry Kernel) is given by  $X \otimes K$ .  $\downarrow_S$  is the downsampling operator with a scale factors  $s$ , the independent noise term is represented by  $N$ . Simplifying and solving Eq. (1) is quite tedious and problematic since one low-resolution image could give rise to several different high-resolution image outputs.

One of the widely studied fields in the recent years is the process of increasing the spatial-resolution of LR images, otherwise known as Single Image Super Resolution. Bicubic up-sampling, nearest neighbor and other similar interpolation methods are few solutions to this problem. Single Image Super Resolution (SISR) for terrestrial applications has been widely studied. However, this is not the case with images captured underwater. SISR techniques for improving these images have not been focused on due to the shortage of detailed and extensive data sets that effectively grasp the distortions present in underwater images. Data sets that are currently available contain synthetic images [29] and are only applicable for purposes like image enhancement [30], and object detection [31] as the resolution of these images is not more than  $256 \times 256$ . As a result, SISR models for underwater images have not been largely studied.

Nevertheless, the existing techniques are not capable of retrieving finer details in images, thus making the output blurry and of low quality. A few studies have been made in this domain, focusing mostly on rebuilding underwater images of better quality by removing noise and blurry areas [32]. Other attempts to improve fish recognition performance [33] and improve underwater image sequence have also employed SISR techniques. Though these models perform well and achieve their purpose, improvements can still be made to match the state-of-the-art performance as these methods do not still do not efficiently address the issue of lack of information in the images.

### 3 Methods and Materials

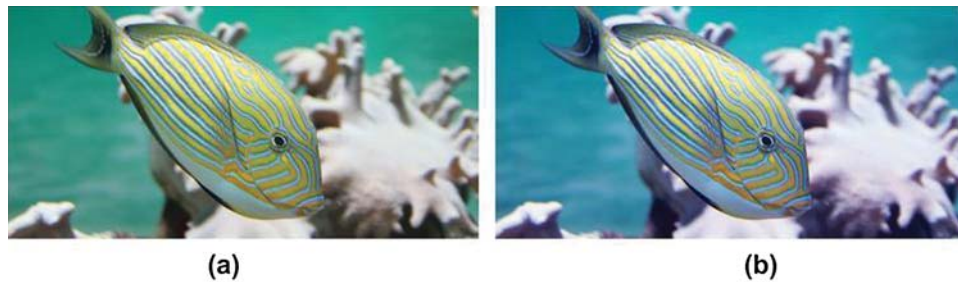
#### 3.1 Proposed Method

One dataset that contains a vast collection of paired HR-LR images is the USR-248 dataset [34]. It contains over 1050 image samples that promote SISR training on a large scale. 248 test images have also been provided in the USR-248 dataset solely for benchmark evaluation. Our Proposed Method is a two-step process—pre-processing and Alpha Super Resolution

Generative Adversarial Networks (AlphaSRGAN) structure. Pre-processing involves improving the image contrast and color-correction of the image to make training convenient. Pretraining of the images is also done to preserve network stability and to achieve better training speed. The AlphaSRGAN structure is the utilization of the dual generator method helps in maintaining the clarity and accuracy of the images produced in this step.

### 3.1.1 Preprocessing

Preprocessing is done with Contrast Limited Adaptive Histogram Equalization (CLAHE) and white balance to remove the low contrast and severe deformations present in underwater images. Next, correction of seafloor color and creation of normal underwater scenes is employed using a positive white balance. CLAHE has been employed to get an enhanced image and to improve the visibility of the aquatic animals. The results after the pre-processing stage have been presented in Fig. 2. The pre-training helps in speeding up the discriminator training process and ensuring that the generator remains stable. Prior to training, a portion of the image (high resolution) training set is inputted into the discriminator so as to preserve the efficiency/coherence and training intensity of the generator [35], simultaneously enabling early identification by the discriminator. Pre-training also ensures that the training mode does not collapse, thus preventing the continuous failure in the generation of SR images. It also makes sure that the training speed and stability of the generator and discriminator are maintained that is essential for adjustment of the training strategy.



**Figure 2:** Preprocessing results. (a) Original image (b) Pre-processed image

### 3.1.2 AlphaSRGAN Architecture

Once the input images are preprocessed and trained with the help of the AlphaSRGAN system, the generator produces the super-resolution image. Linear photometric models, affine motion and speeded up robust features (SURF) image registration are used to register the image in the generator. The image resolution is enhanced by fusing the sharpest area in every image after the registered image is inputted into the network architecture. Consecutively, decomposition of the image is done by discrete wavelet transform as the first step. The image is split into four sub-bands—High-High (HH) sub-band, High-Low (HL) sub-band, Low-High (LH) sub-band and Low-Low (LL) sub-band. Among these, the low-low sub-band retains the original features of the image and acts as the coefficient of approximation. The detailed parameters of the image are represented by the remaining sub-bands i.e., LH, HL, and HH sub-bands. Subsequently, features of the image that stand out from the rest in terms of clarity and detail are extracted and represented first with the help of Linear Discriminant Analysis (LDA). In this step, a new axis is generated on which data from both the features is projected. The information is projected such that the variance is minimized and the distance between the class means is maximized.

Variance is preferred here since the signal and noise generally have variances at the extremes, with the signal having greater variance and the noise having less variance, and because the ratio between the variances can also be depicted easily using the signal-to-noise ratio parameter. Finally, Inverse Discrete Wavelet Transform (IDWT) is used for reconstructing the image. The refused image is used to learn and achieve super-resolution by repeated processing in the generator and discriminator network.

The architecture of the different networks proposed here is based on the AlphaGAN [36] model. A latent variable  $z \sim N(0, I)$  is fed into the generator network. It contains three major parts—a fully connected layer that helps in upscaling the input tensor to size  $512 \times 2 \times 2$ , four transposed convolution layers (kernel size =  $4 \times 4$ , stride = 2, padding = 1) and the tanh activation layer. After each layer in this network, (Rectified Linear Unit (ReLU) acts as the activation function. Four convolutional networks that draw out aspects and attributes from  $32 \times 32$  inputs form the majority of the discriminator network. In the discriminator network, Leaky-ReLU is used as the activation function after each layer. After each layer, batch normalization is performed in both networks. An abs function is used in place of the last activation layer in the discriminator network. Adam optimizer [37] is used to train both the networks at learning rates of 0.0002. In the generator network, the decay rates are set at  $\beta_1 = 0.5$ ,  $\beta_2 = 0.999$ . Fig. 3 represents the Network architecture diagram AlphaSRGAN.

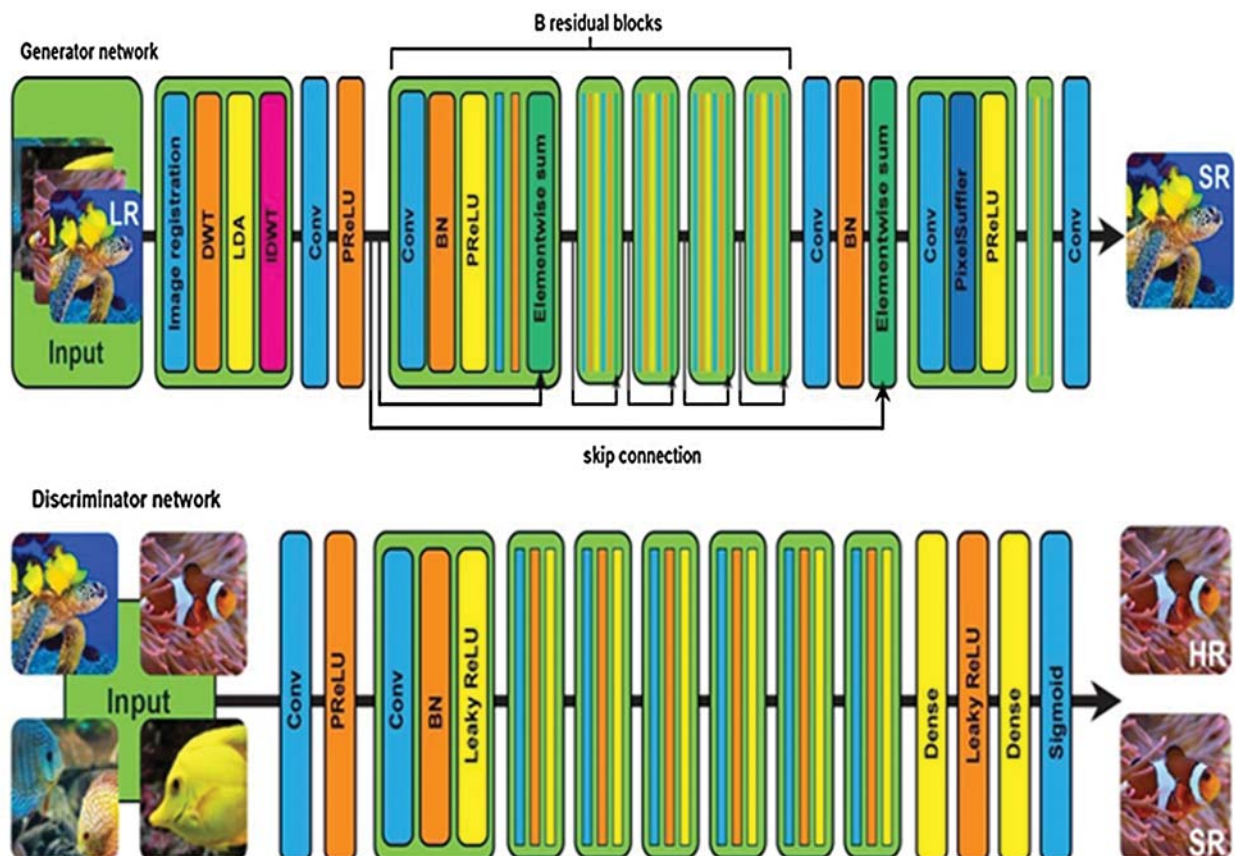


Figure 3: The Network architecture diagram AlphaSRGAN

### 3.2 Alpha-SRGAN

The discriminator's sigmoid output layer is removed and substituted by binary cross entropy loss along with the formulation for power function while using Alpha-GAN [36].

The AlphaGAN architecture resolves problems in optimization shown in Eq. (2).

$$V_{\text{alpha-GAN}}(D, G) = E_{x \sim P_{\text{data}}} [|D(x)|^a] + E_{R \sim P_r} [|D(G(R))|^b] \quad (2)$$

In addition, the proposed method introduces two more hyper-parameters, a and b balance the emphasis on  $D(x)$  and  $D(G(R))$  while training and are two-order indices, respectively. In this model, a and b are assumed to be greater than 0 so as to prevent cases like  $D(a)$  from becoming evident in the loss function that usually happens when these values are equal to 0. This assumption also improves the convergence stability of the model proposed. Otherwise, when the discriminator's output goes below unity, the value of loss becomes very large, thereby making the model unstable and difficult to converge during training. Here, the absolute value of the output from the discriminator is considered. This also prevents the output from taking an arbitrary value when a and b values are greater than 1. It is shown that the objective functions and formulas in Alpha-GAN are not associated with the alpha divergence formulation (Eq. (2)) immediately. A detailed method to find Alpha-GAN from alpha divergence will also be included.

---

#### Algorithm 1 Algorithm of AlphaGAN

---

*Input:* Batch size  $n$ , target distribution  $T_{\text{real}}$ , latent noise distribution  $L_x$ , input noise  $x$ , Adam optimizer with  $\alpha$ ,  $\beta_1 = 0.5$ ,  $\beta_2 = 0.999$ , hyper-parameters  $p$ ,  $q$  discriminator network  $D\phi$  and generator network  $G\phi$  absolute function  $af(\cdot)$ .

*Output:* Optimal generator  $G\phi$

1: while Training Alpha-GAN do

2: Sample  $a_j^S \sim T_{\text{real}}, j = 1, \dots, m$ .

3: Sample  $b_j \sim L_x, j = 1, \dots, m$ .

4:  $a_j^t \leftarrow G\phi(b_j), j = 1, \dots, m$

5:  $\theta \leftarrow \text{Adam}(-\frac{1}{m} \sum_{j=1}^n \nabla \theta [af(D_\theta((a_j^S)^p))])$

6:  $\theta \leftarrow \text{Adam}(\frac{1}{n} \sum_{i=1}^n \nabla \theta [af(D_\theta((a_i^t)^q))])$

7:  $\theta \leftarrow \text{Adam}(\frac{1}{m} \sum_{i=1}^m \nabla \theta [af(D_\theta((a_i^t)^q))])$

8: end while

9: return  $G\phi$

---

### 3.3 Content Loss

Low Resolution or High Resolution (LR/HR) content loss is the factor that stimulates the restoration of identical features as the ground truth. The representation is usually done by the generator in the form of high-level representations. Transfers of style, removal of SISR problems and enhancement of images have been effectively done. High-end attributes extricated by the final convolutional layer from a VGG-19 network that was trained previously has been used to define

the image content function  $\Phi VGG(\cdot)$ . Thereafter, the formulae for content loss for enhancement and SISR has been depicted in the Eqs. (3) and (4).

$$L_{Content}^{LR} = \|\phi VGG(E) - \phi VGG(\hat{E})\|_2 \quad (3)$$

$$L_{Content}^{HR} = \|\phi VGG(Y) - \phi VGG(\hat{Y})\|_2 \quad (4)$$

### 3.4 Adversarial Loss

The generative element of GAN has also been taken with the perceptual loss along with the content loss described in Section 3.3. Adversarial loss promotes our model to prefer responses from the vast plethora of natural images available by attempting to deceive the discriminator network. The probabilities of the discriminator  $D_{\theta_d}(G_{\theta_g}(I^{LR}))$  overall training samples has been used to define the generative loss  $L_{Gen}^{SR}$ . It can be represented as:

$$L_{Gen}^{SR} = \sum_{n=1}^N -\log \log D_{\theta_d}(G_{\theta_g}(I^{LR})) \quad (5)$$

In Eq. (5), the probability that the reconstructed image ( $G_{\theta_g}(I^{LR})$ ) is a natural high-resolution image is depicted by  $D_{\theta_d}(G_{\theta_g}(I^{LR}))$ .  $-\log \log D_{\theta_d}(G_{\theta_g}(I^{LR}))$  has been minimized instead of  $1 - \log \log D_{\theta_d}(G_{\theta_g}(I^{LR}))$  to obtain superior gradient behavior.

## 4 Experimental Results and Analysis

The Alpha-SRGAN model has been implemented using TensorFlow libraries [38]. The model has been trained on 7.5 K unpaired, and 11 K paired instances. The remaining data is used for testing and validation, respectively. Both models are trained for 60,000–70,000 interactions in batches of 8 using four NVIDIA™ GeForce GTX 1080 graphic cards. The experimental results and evaluations based on qualitative analysis, a user study, and standard quantitative metrics have been presented. The image resolutions have been compared by applying the existing algorithms such as SRGAN [7], ESRGAN [39], EDSRGAN [40], RSRGAN [24], ISGAN [25], SRDRMGAN [4] and Deep SESR [41].

### 4.1 Evaluation Metrics

#### 4.1.1 Peak Signal to Noise Ratio

PSNR is often utilized as a standard to assess images. It is a common method to quantify signal reconstruction quality during image compression. It can be evaluated by finding the Mean Square Error method (MSE) [42], which is defined as follows:

MSE for two  $m \times n$  monochromatic images called  $K$  and  $I$ , where one denoted the noise approximation of the other parameter, is given by:

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i,j) - K(i,j)]^2 \quad (6)$$



Peak signal to noise ratio can be evaluated using the following equation:

$$PSNR = 10 \log_{10} \left( \frac{MAX_1^2}{MSE} \right) = 20 \log_{10} \left( \frac{MAX_1}{\sqrt{MSE}} \right) \quad (7)$$

where MAX1 stands the maximum value that represents the color of image points, MAX1 is usually 255 when a particular sampling point is depicted using eight bits. [Tab. 1](#) shows the Evaluation Metrics Comparison of PSNR Scores for 2\*4\*8 SISR on USR-248 data set.

**Table 1:** Evaluation Metrics Comparison of PSNR Scores for 2\*4\*8 SISR on USR-248 data set

Image Scale	SRGAN [7]	ESRGAN [39]	EDSRGAN [40]	RSRGAN [24]	ISGAN [25]	SRDRMGAN [4]	Deep SESR [41]	AlphaSRGAN (Ours)
2*	28.05	26.66	27.12	25.11	26.34	28.55	27.03	29.86
4*	24.76	23.79	21.65	24.96	23.87	24.62	24.59	25.96
8*	20.14	19.75	19.87	19.89	20.19	20.25	21.62	21.89

#### 4.1.2 Structural Similarity Index

A measure of the similarity between 2 images is given by the Structural Similarity Index (SSIM). The Image and Video Engineering Laboratory based in the University of Texas, Austin, was the first one to coin the term. There are two key terms here—Structural information and distortion. These terms are usually defined with respect to the image composition. The property of the object structure that is independent of contrast and brightness is called Structural information. A combination of structure, contrast and brightness gives rise to distortion. Estimation of brightness has been done using mean values, contrast using standard deviation and structural similarity was measured with the help of covariance.

SSIM of two images, x and y can be calculated by:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (8)$$

In the [Eq. \(8\)](#), the average of x is  $\mu_x$  and the average of y is  $\mu_y$ .  $\sigma_x$  and  $\sigma_y$  gives variance. The covariance x and y are given by  $\sigma_{xy}$ .  $c_1 = (k_1L)^2$ ,  $c_2 = (k_2L)^2$  are constants. They are used to preserve stability. The pixel values' dynamic range is given by L. Generally,  $k_1$  is taken as 0.01 and  $k_2$  is given by 0.03. [Tab. 2](#) shows the Evaluation Metrics Comparison of SSIM Scores for 2\*4\*8 SISR on USR-248 data set.

**Table 2:** Evaluation Metrics Comparison of SSIM Scores for 2\*4\*8 SISR on USR-248 data set

Image Scale	SRGAN [7]	ESRGAN [39]	EDSRGAN [40]	RSRGAN [24]	ISGAN [25]	SRDRMGAN [4]	Deep SESR [41]	AlphaSRGAN (Ours)
2*	0.78	0.75	0.77	0.75	0.95	0.81	0.88	0.93
4*	0.69	0.66	0.65	0.69	0.84	0.69	0.71	0.88
8*	0.60	0.58	0.58	0.65	0.72	0.61	0.63	0.75

### 4.1.3 Mean Opinion Score

Mean Opinion Score [7], abbreviated as MOS, is a subjective Image Quality Assurance (IQA) test that is used often and involves humans to score the recognizable quality of the processed images on a scale from 1 to 5, where 1 stand for poor and 5 is for good. The final MOS score is obtained by taking the arithmetic mean of all the obtained scores. Nevertheless, MOS has quite a few drawbacks in the sense that it is based on non-linear perceived scales, variances, and biases in criteria used for rating, etc. Incidentally, there are several super-resolution models that though they have poor IQA metrics like PSNR, have preminent perceptual quality. In these cases, the MOS test proved to be much reliable and more accurate in terms of measurements in perceptual quality. [Tab. 3](#) shows the Evaluation Metrics Comparison of MOS Scores for 2\*4\*8 SISR on USR-248 data set.

**Table 3:** Evaluation Metrics Comparison of MOS Scores for 2\*4\*8 SISR on USR-248 data set

Image Scale	SRGAN [7]	ESRGAN [39]	EDSRGAN [40]	RSRGAN [24]	ISGAN [25]	SRDRMGAN [4]	Deep SESR [41]	AlphaSR GAN (Ours)
2*	2.56	1.85	2.21	1.21	1.78	2.98	2.14	3.58
4*	1.19	1.03	0.89	1.22	1.13	1.16	1.09	1.78
8*	0.80	0.65	0.67	0.69	0.73	0.76	0.84	0.95

### 4.1.4 UIQM

Underwater Image Quality Measure (UIQM) [43] has also been taken into consideration to quantify parameters like underwater image contrast, color quality, and sharpness. This umbrella term consists of three major parameters to assess the attributes of underwater images, namely, Underwater Image Sharpness Measure (UISM), Underwater Image Colorfulness Measure (UICM), along Underwater Image Contrast Measure (UIConM). Each of these parameters are used to appraise a particular aspect of the degraded underwater image carefully.

Altogether, the comprehensive quality measure for underwater images is then depicted by

$$UIQM = c_1 * UICM + c_2 * UISM + c_3 * UIConM \quad (9)$$

[Eq. \(9\)](#) relates all the three attributes mentioned effectively. The selection of the parameters  $c_1$ ,  $c_2$ , and  $c_3$  are purely based on the parameters' application. As an illustration, consider UICM; this parameter is given more weightage for applications that involve image color correction while sharpness (UISM) and contrast (UIConM) have more significance while enhancing the visibility of the images. UIQM regresses to an underwater image quality measure when these two parameters achieve null values. [Tab. 4](#) shows the evaluation metrics comparison of UIQM Scores for 2\*4\*8 SISR on the USR-248 data set.

**Table 4:** Evaluation Metrics Comparison of UIQM Scores for 2\*4\*8 SISR on USR-248 data set

Image Scale	SRGAN [7]	ESRGAN [39]	EDSRGAN [40]	RSRGAN [24]	ISGAN [25]	SRDRMGAN [4]	Deep SESR [41]	AlphaSR GAN (Ours)
2*	2.72	2.70	2.67	2.42	2.72	2.77	3.15	3.23
4*	2.42	2.38	2.40	2.55	2.35	2.48	2.96	2.85
8*	2.10	2.05	2.12	2.10	2.01	2.17	2.39	2.41

### 4.2 Performance Comparison

According to PSNR and SSIM, the comparative findings are represented in Fig. 4. The findings indicate the primacy of the research data system presented. It can be seen from the figure that in both the PSNR and SSIM evaluation indices, our proposed approach performance is momentous. By correlating the graph, the SRGAN technique obtains the lowest PSNR and SSIM score, which implies that this technique cannot recover adequate knowledge. In comparison, the ESRGAN technique has higher values of PSNR and SSIM than the SRGAN; however, this technique cannot acquire accomplished information. In further comparison with EDSRGAN technique, the results of PSNR and SSIM were slightly lower than other methods. But the technique of RSRGAN's acquired values of PSNR and SSIM was quite higher than EDSRGAN. In addition, RSRGAN method results in better images. Furthermore, while comparing with the ISGAN, SRDRMGAN and Deep SESR techniques have significantly higher PSNR and SSIM values than RSRGAN technique. While a consistent picture can be obtained by the ISGAN and SRDRMGAN methods, the missing portion of the information cannot be supplemented by these two methods.

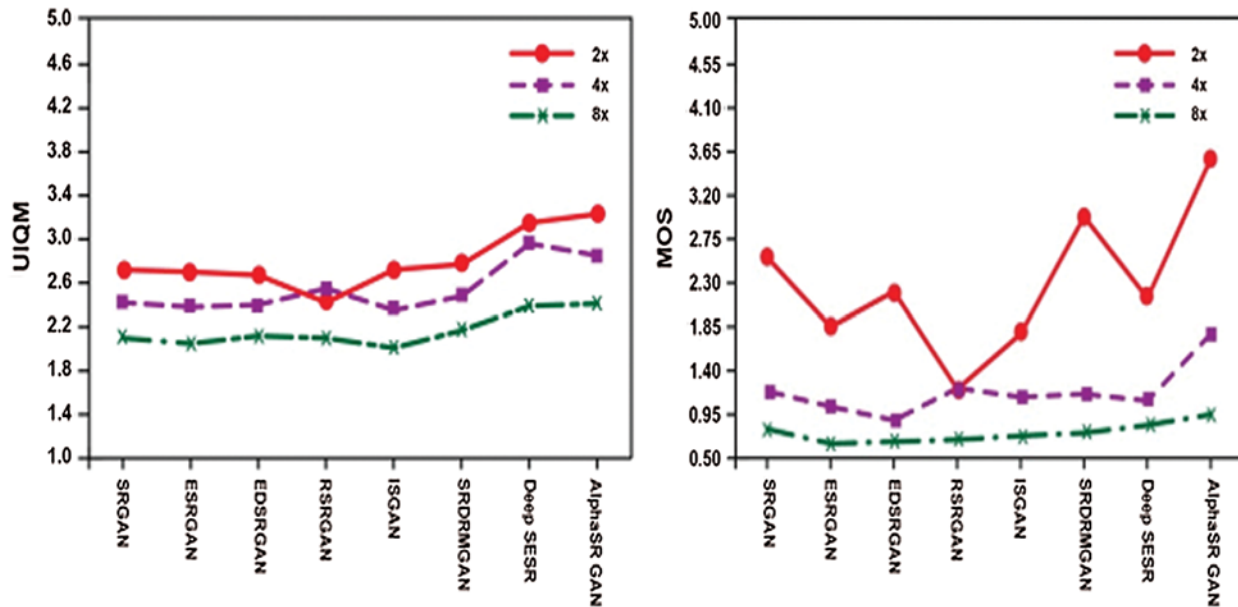


Figure 4: Evaluation index. (a) PSNR and (b) SSIM

DeepSR technique can perform two tasks simultaneously and produce a picture for a better result. Nevertheless, the compared methods, our proposed model AlphaSR GAN perform the tasks effectively in a synchronized manner and obtain an image with the best quality for pre-eminent results. To validate the performance of our proposed model, we compared it with other

techniques and the comparison results are graphically illustrated in Fig. 5; the comparison was according to UIQM and MOS values. The technique which obtained the lowest value of UIQM and MOS value is SRGAN.

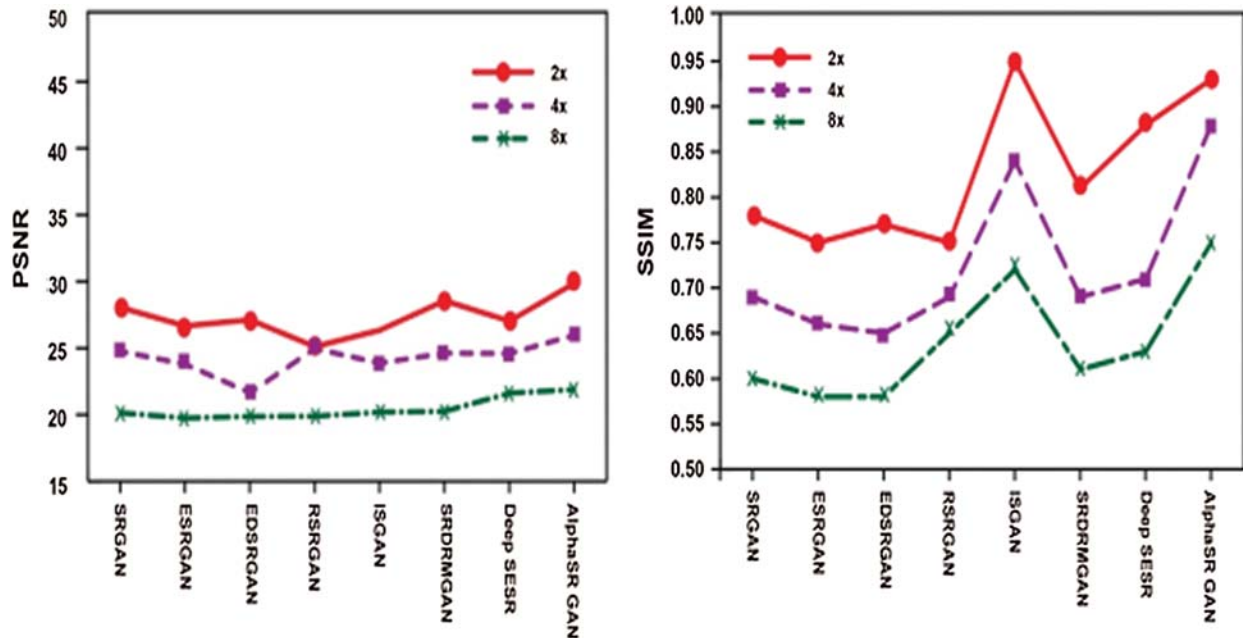
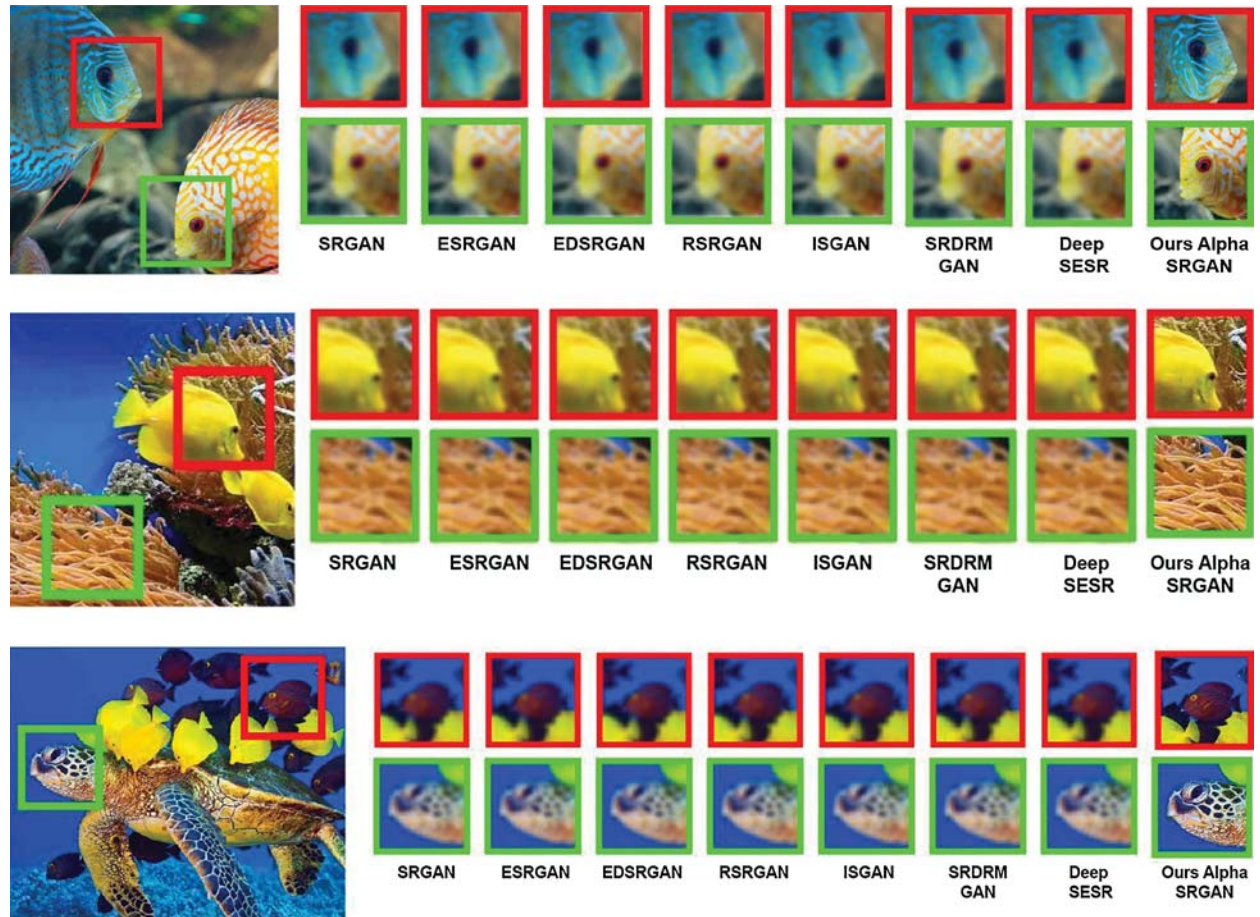


Figure 5: Evaluation index UQIM and MOS

This technique failed to obtain the required data. Whereas the ESRGAN and EDSRGAN methods have resulted in slightly higher UIQM and MOS values, but these techniques also failed to recover the complete information. In further comparison with other techniques, the RSRGAN outcome has significantly higher values of UIQM and MOS, but the picture effect was comparatively low. Then ISGAN and SRDRMGAN technique resulted in higher UIQM and MOS values similar to Deep SESR technique. However, the required part of the information was not compensated by these two techniques, even though it can obtain a clear image.

Further, with higher values of UIQM and MOS, Deep SESR produces better images. However, our proposed AlphaSRGAN model executes results in notably higher UIQM and MOS values. In addition, the tasks were efficiently performed in a coordinated manner and obtained a picture of the highest standard. Fig. 6 shows the Performance comparison of AlphaSRGAN with other related GAN for super-resolution.



**Figure 6:** Performance comparison of AlphaSRGAN with SRGAN [7], ESRGAN [39], EDSRGAN [40], RSRGAN [24], ISGAN [25], SRDRMGAN [4], Deep SESR [41]

## 5 Conclusion

In this work, a novel generative model named AlphaSRGAN image super-resolution algorithm has been introduced that amalgamates the traditional image reconstruction approaches with deep learning methods for underwater image super-resolution. In addition, several qualitative and quantitative tests have been performed on the dataset used-USR-248. The peak signal to noise ratio for image scales  $2^*$ ,  $4^*$  and  $8^*$  in our model is superior to other models. Parameters such as SSIM and UIQM also prove to be better than existing systems when compared to the proposed model. Enhanced Mean Opinion Scores have also been obtained by our model. Subsequently, the model proposed here shows the greater performance when compared to pre-existing models. AlphaSRGAN also proves to be a better alternative taking the enhanced performance, competent computational abilities, and model design into consideration. It makes the model extremely suitable for real-time applications and applications in other fields as well.

**Funding Statement:** The authors received no specific funding for this study.

**Conflicts of Interest:** The authors declare that they have no conflicts of interest to report regarding the present study.

## References

- [1] S. Anwar, C. Li and F. Porikli, “Deep underwater image enhancement,” *arXiv preprint*, arXiv: 1807.03528, 2018.
- [2] R. Kumudham and V. Rajendran, “Super resolution enhancement of underwater sonar images,” *SN Applied Sciences*, vol. 1, no. 8, pp. 852, 2019.
- [3] H. Lu, Y. Li, S. Nakashima, H. Kim and S. Serikawa, “Underwater image super-resolution by descattering and fusion,” *IEEE Access*, vol. 5, pp. 670–679, 2017.
- [4] M. J. Islam, S. SakibEnan, P. Luo and J. Sattar, “Underwater image super-resolution using deep residual multipliers,” in *2020 IEEE Int. Conf. on Robotics and Automation*, Paris, France, pp. 900–906, 2020.
- [5] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley *et al.*, “Generative adversarial networks,” in *Proc. of the 27th Int. Conf. on Neural Information Processing Systems*, MIT Press, Cambridge, MA, USA, vol. 2, pp. 2672–2680, 2014.
- [6] K. He, X. Zhang, S. Ren and J. Sun, “Deep residual learning for image recognition,” in *2016 IEEE Conf. on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA, pp. 770–778, 2016.
- [7] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham *et al.*, “Photo-realistic single image super-resolution using a generative adversarial network,” in *2017 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, pp. 105–114, 2017.
- [8] P. Liu, G. Wang, H. Qi, C. Zhang, H. Zheng *et al.*, “Underwater image enhancement with a deep residual framework,” *IEEE Access*, vol. 7, pp. 94614–94629, 2019.
- [9] R. D. Rajagopal, S. Murugan, K. Kottursamy and V. Raju, “Cluster based effective prediction approach for improving the curable rate of lymphatic filariasis affected patients,” *Cluster Computing*, vol. 22, no. 1, pp. 197–205, 2019.
- [10] R. Mehra, K. Pachpor, K. Kottilingam, and A. Saranya, “An initiative to prevent Japanese encephalitis using genetic algorithm and artificial neural network,” in *2020 Int. Conf. on Computational Intelligence*, Malaysia, IEEE, pp. 142–148, 2020.
- [11] M. Sung, J. Kim and S. Yu, “Image-based super resolution of underwater sonar images using generative adversarial network,” in *TENCON 2018–2018 IEEE Region 10 Conf.*, Jeju, Korea (South), pp. 457–461, 2018.
- [12] S. Liu and X. Li, “A novel image super-resolution reconstruction algorithm based on improved GANs and gradient penalty,” *International Journal of Intelligent Computing and Cybernetics*, vol. 12, no. 3, pp. 400–413, 2019.
- [13] P. Shamsolmoali, M. Zareapoor, R. Wang, D. K. Jain and J. Yang, “G-GANISR: Gradual generative adversarial network for image super resolution,” *Neurocomputing*, vol. 366, pp. 140–153, 2019.
- [14] Y. Xie, E. Franz, M. Chu and N. Turey, “tempoGAN: A temporally coherent, volumetric GAN for super-resolution fluid flow,” *ACM Transactions on Graphics (TOG)*, vol. 37, no. 4, pp. 95, 2018.
- [15] A. Bulat and G. Tzimiropoulos, “Super-fAN: Integrated facial landmark localization and super-resolution of real-world low resolution faces in arbitrary poses with GANs,” in *Proc. of the 2018 IEEE/CVF Conf. on Computer Vision and Pattern Recognition*, Salt Lake City, AL, USA, pp. 109–117, 2018.
- [16] J. Zhu, T. Park, P. Isola and A. A. Efros, “Unpaired image-to-image translation using cycle-consistent adversarial networks,” in *2017 IEEE Int. Conf. on Computer Vision*, Venice, Italy, pp. 2242–2251, 2017.
- [17] X. Mao, Q. Li, H. Xie, R. Y. K. Lau, Z. Wang *et al.*, “Least squares generative adversarial networks,” in *2017 IEEE Int. Conf. on Computer Vision*, Venice, Italy, pp. 2813–2821, 2017.

- [18] A. Lucas, S. Lopez-Tapia, R. Molina and A. K. Katsaggelos, "Generative adversarial networks and perceptual losses for video super-resolution," *IEEE Transactions on Image Processing*, vol. 28, no. 7, pp. 3312–3327, 2019.
- [19] M. Arjovsky, S. Chintala and L. Bottou, "Wasserstein generative adversarial networks," in *2017 Proc. of the 34th Int. Conf. on Machine Learning*, Sydney, Australia, pp. 214–223, 2017.
- [20] Y. Yuan, S. Liu, J. Zhang, Y. Zhang, C. Dong *et al.*, "Unsupervised image super-resolution using cycle-in-cycle generative adversarial networks," in *IEEE Conf. on Computer Vision and Pattern Recognition Workshops*, CA, USA, pp. 701–710, 2018.
- [21] Z. Yi, H. Zhang, P. Tan and M. Gong, "Dualgan: Unsupervised dual learning for image-to-image translation," in *Int. Conf. on Computer Vision*, Venice, Italy, pp. 2849–2857, 2017.
- [22] J. Guan, C. Pan, S. Li and D. Yu, "Srdgan: Learning the noise prior for super resolution with dual generative adversarial networks," *arXiv preprint*, arXiv: 1903.11821, 2019.
- [23] Z. Ding, X.-Y. Liu, M. Yin, W. Liu and L. Kong, "Tgan: Deep tensor generative adversarial nets for large image generation," *arXiv preprint*, arXiv: 1901.09953, 2019.
- [24] W. Zhang, Y. Liu, C. Dong and Y. Qiao, "RankSRGAN: Generative adversarial networks with ranker for image super-resolution," in *2019 IEEE/ICVF Int. Conf. on Computer Vision*, Seoul, Korea (South), pp. 3096–3105, 2019.
- [25] L. Li, Z. Fan, M. Zhao, X. Wang, Z. Wang *et al.*, "Super-resolution reconstruction of underwater image based on image sequence generative adversarial network," *Mathematical Problems in Engineering*, vol. 20, pp. 1–10, Article ID 8472875, 2020.
- [26] Y. Jin-Tao, J. Rui-Sheng, G. Li, Y. Ruo-Nan, S. Hong-Mei *et al.*, "Underwater image enhancement method based on the generative adversarial network," *Journal of Electronic Imaging*, vol. 30, no. 1, pp. 13009, 2021.
- [27] Y. Chenxi and M. Mohsen, "Attribute-aware generative design with generative adversarial networks," *IEEE Access*, vol. 8, pp. 190710–190721, 2020.
- [28] C.-Y. Yang, C. Ma and M. -H. Yang, "Single-image super-resolution: A benchmark," in *Computer Vision – ECCV 2014. Lecture Notes in Computer Science*, vol. 8692, Springer, Cham, 2014.
- [29] Y. Chen, B. Yang, M. Xia, W. Li, K. Yang *et al.*, "Model-based super-resolution reconstruction techniques for underwater imaging," in *Photonics and Optoelectronics Meetings (POEM): Optoelectronic Sensing and Imaging*, vol. 8332, pp. 83320G, 2012.
- [30] M. J. Islam, Y. Xia and J. Sattar, "Fast underwater image enhancement for improved visual perception," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 3227–3234, 2020.
- [31] M. J. Islam, M. Ho and J. Sattar, "Understanding human motion and gestures for underwater human–robot collaboration," *J. Field Robotics*, vol. 36, pp. 851–873, 2019.
- [32] C. Fabbri, M. J. Islam and J. Sattar, "Enhancing underwater imagery using generative adversarial networks," in *2018 IEEE Int. Conf. on Robotics and Automation*, Brisbane, QLD, Australia, pp. 7159–7165, 2018.
- [33] J. Sun, Z. Xu and H.-Y. Shum, "Image super-resolution using gradient profile prior," in *2008 IEEE Conf. on Computer Vision and Pattern Recognition*, Anchorage, AK, USA, pp. 1–8, 2008.
- [34] <http://irvlab.cs.umn.edu/resources/usr-248-dataset>.
- [35] A. Bulat, J. Yang and G. Tzimiropoulos, "To Learn Image Super-Resolution, Use a GAN to Learn How to Do Image Degradation First," In *Computer Vision – ECCV 2018. Lecture Notes in Computer Science*, vol. 1210, Springer, Cham, 2018.
- [36] L. Cai, Chen, N. Cai, W. Cheng and H. Wang, "Utilizing amari-alpha divergence to stabilize the training of generative adversarial networks," *Entropy*, vol. 22, no. 4, pp. 10, 2020.
- [37] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint*, arXiv: 1412.6980, 2014.
- [38] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis *et al.*, "Tensorflow: A system for large-scale machine learning," in *USENIX Symp. on Operating Systems Design and Implementation*, USA, pp. 265–283, 2016.

- [39] M. J. Islam, P. Luo and J. Sattar, “Simultaneous enhancement and super-resolution of underwater imagery for improved visual perception,” *arXiv preprint*, arXiv: 2002.01155, 2020.
- [40] X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu *et al.*, “ESRGAN: Enhanced super-resolution generative adversarial networks,” In: L. Leal-Taixé, S. Roth (Eds.) *Computer Vision – ECCV2018 Workshops. ECCV 2018. Lecture Notes in Computer Science*, Springer, Cham., vol. 11133, 2018.
- [41] K. Panetta, C. Gao and S. Agaian, “Human-visual-system-inspired underwater image quality measures,” *IEEE Journal of Oceanic Engineering*, vol. 41, no. 3, pp. 541–551, 2015.
- [42] B. Lim, S. Son, in H. Kim, S. Nah and K. Mu Lee, “Enhanced deep residual networks for single image super-resolution,” in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition Workshops, USA*, pp. 136–144, 2016.
- [43] M. Mathieu, C. Couprie and Y. LeCun, “Deep multiscale video prediction beyond mean square error,” *arXiv preprint*, arXiv: 1511.05440, 2016.