Article

iDehaze: Supervised Underwater Image Enhancement and Dehazing via Physically Accurate Photorealistic Simulations

Mehdi Mousavi, Rolando Estrada and Ashwin Ashok *

Department of Computer Science, Georgia State University, Atlanta, GA 30303, USA; smousavi2@student.gsu.edu (M.M.)
* Correspondence: aashok@gsu.edu

Abstract: Underwater image enhancement and turbidity removal (dehazing) is a very challenging problem, not only due to the sheer variety of environments where it is applicable, but also due to the lack of high-resolution, labelled image data. In this paper, we present a novel, two-step deep learning approach for underwater image dehazing and colour correction. In iDehaze, we leverage computer graphics to physically model light propagation in underwater conditions. Specifically, we construct a three-dimensional, photorealistic simulation of underwater environments, and use them to gather a large supervised training dataset. We then train a deep convolutional neural network to remove the haze in these images, then train a second network to transform the colour space of the dehazed images onto a target domain. Experiments demonstrate that our two-step iDehaze method is substantially more effective at producing high-quality underwater images, achieving state-of-the-art performance on multiple datasets. Code, data and benchmarks will be open sourced.

Keywords: underwater image enhancement; underwater image dehazing; synthetic data; deep learning

1. Introduction

There is an increasing need for underwater imagery in applications ranging from unmanned underwater vehicles (UUVs) to oceanic engineering and marine biology research. However, water is denser and more dielectric than air, so capturing clear images underwater is much more challenging than on land. Specifically, since water absorbs and scatters more light than air, a submerged image sensor can capture less information about the surrounding environment, leading to a hazy, blurry image. Underwater image enhancement and restoration (UIER) seeks to remedy this by using image processing techniques to enhance the images after they have been captured, specifically by applying dehazing (to remove scattering effects) and colour correction (to reduce absorption effects) to the raw image, see as Figure 1. Current state-of-the-art image analysis uses deep learning due to its unmatched ability to learn task-relevant features. However, applying deep learning to UIER is challenging due to the dearth of data in this domain. As is well known, deep neural networks require vast amounts of (mostly labelled) data to achieve good results. Specifically, the impressive results of deep learning on challenging computer vision tasks, such as depth estimation, surface normal estimation and segmentation have leveraged (mostly free) high-quality (and sometimes synthetic) datasets [1–4]. In contrast, underwater image data is expensive to acquire due to the equipment and transportation costs involved. Capturing underwater images also requires specialized skills, and the data is far more time-consuming to label. As such, free, high-quality underwater images are scarce. In this paper, we propose a novel, two-step supervised method (Figure 2) for underwater image dehazing and colour correction. Our approach combines state-of-the-art deep learning with synthetic data generation, the latter addressing the dearth of real-image data. The key contributions of this paper are as follows:
1. Design and implementation of a unique photorealistic 3D simulation system modelled after real-world conditions for underwater image enhancement and dehazing.

2. A deep convolutional neural network (CNN) for underwater image dehazing and colour transfer, trained on pure synthetic data and a UIEB dataset.

3. A customizable synthetic dataset/simulation toolkit for training and diagnosing of underwater image-enhancement systems with robust ground truth and evaluation metrics.

Figure 1. iDehaze on the UIEB dataset [5]. Top row: hazy, raw images. Bottom row: reconstructed, dehazed through the iDehaze pipeline. Note, the detailed image reconstruction, colour transformation and sharpness of the final images. Best viewed at 4× zoom.

Figure 2. The two-step approach of iDehaze. The input image is first dehazed by a specialized dehazing model, trained on synthetic data. The resulting colours are then transferred unto a target domain by the colour model.

2. Related Work

Below, we review key prior work on underwater image enhancement and synthetic data generation for computer vision tasks.

Underwater image enhancement: Restoring an underwater image is often labelled as “dehazing” or “enhancement” and presented as a cumulative process in which the colours of the image are enhanced through a colour correction pass, and local and global contrast is altered to yield the enhanced final image. Such pipelines are often collections of linear and non-linear operations through algorithms that break down images into regions [6], or estimate attenuation and scattering [7] to approximate real scattering and correct it accordingly. However, for reasons explored further in Section 3, colour correction and dehazing are two different problems that require their own separate solutions.

Prior GAN-based approaches and synthetic data: The use of synthetic data has been the topic of several recent publications [8–16], and the application of synthetic data varies greatly depending on the method of data generation. In this context, synthetic data mostly refers to making underwater images from in-land images using different methods. One might apply a scattering filter effect [8], or make use of colour transformations to reach
the look of an underwater image. Most notably, Li et al. converted real indoor on-land images to underwater-looking images via the use of a GAN (generative adversarial network) [9,10], which sparked an avalanche of GAN-based underwater image enhancement methods [11–16]. GANs remain a subject of interest for underwater image enhancement and restoration (UIER) due to the fact that labelled, high-quality data is rare in UIER, as discussed above. While such methods can be helpful, there are caveats and challenges to GAN-based synthetic data generation and image-enhancement models. GANs in image enhancement are typically finicky in the training process as they are very sensitive to hyper-parameters, and adjustments to the learning rate and momentum, making stable GAN training an open research problem and a very common issue in GAN-based approaches [11,15,17,18]. In comparison, CNNs are feedforward models that are far more controllable in training and testing. Furthermore, features of the underwater domain might differ from features learned and generated by the GAN, causing further inaccuracies in the supervised image-enhancement models that learn from this generated data. Therefore, the most accurate method of generating synthetic data for underwater scenarios is to use 3D photorealistic simulations that allow for granular control over every variable, can be modelled after many different environments, and allow for diagnostic methods and wider ground truth annotations [19].

Lack of standardized data: Underwater image enhancement suffers from a lack of high-quality, annotated data, while there have been numerous attempts to gather underwater images from real environments [11,20], the inconsistencies between image resolution, amount of haze, and imaged objects makes the testing and training of deep learning models significantly more challenging. For example, the EUVP dataset [11] contains small images of 256 × 256 resolution, while the UFO-120 dataset [12] contains 320 × 240 and 640 × 480 images, and the UIEB dataset [5] contains images of various resolutions, ranging from 330 × 170 to 2160 × 1440 pixels. This difference between image samples, especially in image quality, haze, and imaged objects, is an issue with many learning-based systems, both in training and evaluation.

Underwater simulations: Currently, there are a handful of open-source underwater simulations available. Prior simulations exclusively developed for underwater use include UWSim (published in 2012) [21] and UUV (unmanned underwater vehicle) simulator (published in 2016) [22]. These provide tools for simulations of rigid body dynamics or sensors such as sonar or pressure sensors. However, these tools have not been recently updated or developed to support modern hardware. More importantly, these simulations do not focus on real-time, realistic image rendering with ray tracing, nor are they designed for modern diagnostic methods such as data ablation [1,23–25]. In contrast, our simulation supports real-time ray tracing, physics-based attenuation and scattering, allowing for dynamic modifications to the structure of the scenes and captured images.

3. Methods

As demonstrated in Figure 2, our proposed method separates dehazing from colour correction in a two-step process. First, we reconstruct a dehazed image from the input, then feed the resulting dehazed image to a colour-transfer model to obtain a final image. As we detail in Section 4, our dehazing model is trained on 5000 synthetic images that physically model light scattering and attenuation in water. Our model restores pixel information in areas affected by this attenuation and scattering, and the colour model—trained on a subset of the UIEB dataset—matches the colour space of the dehazed image with the target domain, finishing the process. By splitting the image-enhancement task into dehazing and colour transfer, we can effectively train deep learning models, independently control how they process the input image, and quickly update our pipeline to match a new target domain for colour transfer without the need to retrain the dehazing model.

In this section, we will cover the various parts of the iDehaze pipeline. Below, we first review the 3D simulation used to gather the training images for the dehazing model.
3.1. Simulation of Underwater Environments

Compared to on-land images, underwater images require significantly more time and effort to gather for any type of visual task. This, in turn, makes preparation of underwater data for machine/deep learning analysis very challenging due to the manual effort required for labelling (supervised) and data clean-up (unsupervised). In addition, these datasets are immutable, making it impossible to modify them after acquisition. We address this dearth of real data by generating photorealistic data using a 3D simulation environment. We then use this generated data to train our deep neural network to dehaze underwater images. Our simulation is made in Unreal Engine 4 using real-time ray tracing. Unreal Engine 4 is a 3D content creation program made by Epic games, and is often the tool of choice for creating extremely photorealistic images [26]. We modelled an underwater environment to match the properties of the target domain. In particular, our environment contains dynamic swimming fish, inanimate objects, dynamic aquatic plants, wreckage and boulders. The lighting of the simulation is achieved by real-time global illumination via ray tracing, this makes our underwater scenes very realistic since objects are shaded and lit realistically for every pixel in the image. We use a global pixel shader to model the attributes of light propagation underwater. Specifically, we model the exponential signal decay based on the Beer–Lambert law [27], where light is exponentially scattered away depending on the optical depth and attenuation coefficient:

\[
\text{FragColor}(r, g, b) = \text{Lerp} \left( \lambda(r, g, b), 1 - \theta(r, g, b) e^{\Delta(r, g, b) \cdot \mu} \right)
\]

In Equation (1), \( \lambda \) is the ray-traced normal image, \( \theta \) is the scattered wavelengths, and \( \Delta \) is the pixel-wise optical depth in each channel \((r, g, b)\) and \( \mu \) is the molar concentration of the dielectric material (water). In the Beer–Lambert law the term \( \Delta \cdot \mu \) is called the attenuation coefficient. The \( \text{Lerp} \) function interpolates between the scattered image and the ray-traced image rendered at the GPU frame-buffer, using the Beer–Lambert law as the interpolation key. The wavelength term \( \theta \) allows us to control which wavelengths of light are scattered and which reach the camera, hence enabling the realistic modelling of different types of murky waters (see Figure 3). It is worth noting that when applied at the pixel-wise level, this equation simulates a homogenous light transport medium. An intriguing extension of this approach would be to model non-homogenous light transport media, such as water with varying temperatures or different degrees of volumetric molar concentration, which could be explored in the future (see Section 5).

Figure 3. Manipulation of the attenuation coefficient in the global pixel shader allows for creating a supervised dataset and train a deep learning model to learn to reverse the effects of attenuation and scattering in underwater images. The attenuation coefficient depends on the molar concentration of the dielectric material. For hazy images, we chose a random molar concentration above 90 mol/L. For a dehazed ground truth, we chose the molar concentration of pure water: 55.5 mol/L. Images with a high attenuation coefficient are hazy, matching the under water light characteristics.
3.2. Dehazing vs. Colour Transfer

Similar to earlier studies [28,29], we split the overall image-enhancement task into two distinct tasks: (1) image dehazing and (2) colour transfer. By splitting these two fundamentally different tasks, we are able to train specialized deep learning models to dehaze and transfer the colour to a target domain. We note that, due to the lack of accurate measurements at the image capture time in the UIEB and many other underwater image-enhancement datasets, we usually do not know the “true” colour space of the underwater images. Hence, rather than colour correction, we believe that this step is best termed colour transfer. We address this issue in our simulation by including a customizable colour checker chart to accurately gauge the correctness of the colour-transfer methods. We expect that this tool will prove useful for training underwater colour correction algorithms in the future.

4. Experiments

4.1. Neural Networks

As shown in Figure 4, we use a custom version of U-NET [30] with random dropout and a modified final layer to generate a three-channelled image. The patch-based approach in our data intake pipeline allows for the use of various sized images when training the neural network, more specifically, we cut images into uniform-sized patches with a previously set overlap value. For our experiments, we set the patch size to 384 × 384 and slid the patches by 300 pixels to cover the image. To make sure that the model learned both the image structure and to reduce the outlier prediction values and image reproduction noise, we used a hybrid loss function that allowed for controlling the amount of processing for each image:

\[
\text{HybridLoss} = (\lambda \times (1 - \text{ssim})) + ((1 - \lambda) \times \text{mse})
\]

In Equation (2), \(\lambda\) is a weight hyper-parameter that controls how much the model can deviate from the original image, \(\text{ssim}\) is the structural similarity index [31], and \(\text{mse}\) is the mean-squared error. Both models were trained with a batch size of 128 and a learning rate of 0.001. The colour model was trained for 100 epochs, and the dehazing model was trained for 50 epochs. The training procedure for each network took approximately 4 h on two Nvidia GTX 1080ti cards.

![Figure 4. Architecture of the customized U-NET used in the iDehaze pipeline. At layers 4, 6 and 8 there is a 20% chance of dropout. In the figure, double convolution refers to performing convolution, batch normalization, and ReLU twice in a row.](image)

4.2. Data Acquisition

When acquiring image data, we used a Python module to move and rotate in six directions in the 3D space and gather imaging samples. This module automatically changed the scattering wavelengths in each captured image to cover a wide array of scattering colour patterns. These wavelengths were chosen by analysing the infinity regions (i.e., where light reaching the camera is completely scattered) of the UIEB dataset [5] with the eyedropper tool in Photoshop. We gathered 5000 RGB images from various angles and scattering wavelengths, which we then used to train our dehazing model. Generating this
dataset took around 3 h with our hardware. These images were split to an 80:10:10% ratio for training, validation and testing set, respectively.

4.3. Experimental Setup

We conducted all our experiments in a Dell Precision 7920R server with two Intel Xeon Silver 4110 CPUs, two GeForce GTX 1080 Ti graphics cards, and 128 GBs of RAM. As noted above, we trained the neural networks for 100 epochs for each task (dehaze, colour transfer) and the \( \lambda \) hyper-parameter in Equation (2) was set to 0.6 for each neural network. The dehazing synthetic dataset was generated in Unreal Engine 4.26 using a Windows 10 machine with 16 GBs of RAM, an RTX 2080ti and AMD Ryzen 5600X CPU.

4.4. Metrics

We quantitatively evaluated the output of iDehaze using the most common metrics used in prior work \[11,28,29,32\], namely the underwater image quality measure (UIQM), peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM). The UIQM metric is a non-reference measure that considers three attributes of the final result: (1) UICM, image colourfulness (2) UISM, image sharpness and (3) UIConM, image contrast. Here, each attribute captures one aspect of image degradation due to signal path loss in underwater images. In Equation (3), the UIQM is measured by adding UISM, UICM, and UIConM using a ratio defined by three constants:

\[
\text{UIQM} = c_1 \times \text{UICM} + c_2 \times \text{UISM} + c_3 \times \text{UIConM}
\]

The \( c_1, c_2 \) and \( c_3 \) constant values are set to the same values as the original publication’s suggested values \[32\] and are the same across the comparisons drawn in Section 5, more specifically: \( c_1 = 0.0282 \), \( c_2 = 0.2953 \), \( c_3 = 3.5773 \).

The PSNR metric measures the approximate reconstruction quality of image \( x \) compared to the ground truth image \( y \) based on the mean-squared error (mse):

\[
\text{PSNR}(x, y) = 10 \log_{10} \left[ \frac{255^2}{\text{mse}(x, y)} \right]
\]

The SSIM metric, on the other hand, compares the image patches based on luminance, contrast and structure. In Equation (5), \( \mu \) denotes the mean, and \( \sigma \) denotes the variance, while \( \sigma_{xy} \) denotes the cross-correlation between \( x \) and \( y \). In addition, the constants \( c_1 = (255 \times 0.01)^2 \) and \( c_2 = (255 \times 0.03)^2 \) are present to ensure numerical stability \[11,31\].

\[
\text{SSIM}(x, y) = \left( \frac{2 \mu_x \mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1} \right) \cdot \left( \frac{2 \sigma_{xy} + c_2}{\sigma_x^2 + \sigma_y^2 + c_2} \right)
\]

For the UIQM, higher values indicate better-quality images. The PSNR metric is used to evaluate the reconstruction quality and noise performance, with higher values signifying better image quality. The SSIM serves as a supplementary assessment mechanism, while an SSIM value of 1 indicates identical images (which is undesirable in image enhancement as the goal is to modify the input image), it should not be excessively low either. In most cases, an SSIM value between 0.5 and 1 is considered desirable, reflecting a balance between maintaining the image structure and achieving the desired enhancements.

4.5. Datasets

We use a subset of the UIEB dataset \[5\] to train our colour model, specifically, we discard low-resolution images, and only use images that have at least 384 pixels in each dimension. The UIEB dataset is a set of 890 real underwater images captured under different lighting conditions and have diverse colour range and contrast. We chose UIEB since its reference images were obtained without synthetic techniques. We reserve 80 images to evaluate the iDehaze pipeline. For benchmarks, we chose the EUVP \[11\] and UFO-120 datasets \[12\]. EUVP is a large collection of lower-resolution underwater images, manually
captured by the authors during oceanic explorations. We evaluate our method on 515 paired test images in the EUVP dataset. Note that due to the lower resolution of the EUVP test images (256 × 256), we padded the samples with empty pixel values when feeding them into our pipeline. The UFO-120 dataset is a collection of 1500 640 × 480 underwater images and 120 test samples for evaluation. We used the 120 test samples from the UFO-12 dataset to evaluate and compare iDehaze with other systems.

5. Results

In this section, we analyse the results of our proposed approach both qualitatively and quantitatively. Table 1 shows the performance of the iDehaze system against four state-of-the-art methods on the three aforementioned underwater image datasets, while Table 2 compares the performance of iDehaze against the most recent GAN-based models on the UFO-120 and EUVP datasets. The UIEB metrics were measured with a separate test set of 80 randomly chosen images, unseen by the colour model in the iDehaze pipeline.

As shown on these tables, iDehaze outperforms all the other methods based on the UIQM metric, achieving state-of-the-art performance on the UIEB, UFO-120 and EUVP datasets. Note that our deep neural networks were not trained on each of these three datasets separately. Instead, our dehaze model was solely trained on synthetic data obtained from our UE4 simulation, and our colour model was only trained on a subset of the UIEB dataset. As such, the result on the UFO-120 and EUVP prove that (1) our 3D simulation is able to generate realistic data that matches real-world data and (2) that our two-step pipeline is able to learn features that can generalize to a wide variety of data with no additional training. Finally, we note that iDehaze also achieved state-of-the-art SSIM on the EUVP dataset.

In more detail, Table 1 compares iDehaze to various state-of-the-art models in underwater image enhancement. The WaterNet model [5] was trained on the entire UIEB dataset, the [11] FUnIE-GAN method was trained on the EUVP dataset, the deep SESR was trained on the UFO-120 dataset [12], and the Shallow-UWNet used the pre-trained weights of the deep SESR model and was re-trained on a subset of the EUVP dataset [28]. iDehaze showed superior performance and stability on the EUVP and UFO-120 datasets despite not being trained on them, and it also outperformed all methods in the UIEB test set in the UIQM. For the SSIM, iDehaze narrowly beat the Shallow-UWNet, but it was less stable with a relatively higher standard deviation. Conversely, iDehaze outperformed WaterNet on the UIEB dataset with a more accurate SSIM and higher stability.

iDehaze performed relatively poorly on the EUVP and UFO-120 datasets on the PSNR metric. We postulate that such a performance drop could be due to the iDehaze pipeline trying to reconstruct every detail present in the image. Since the dehazing model was trained on clean, noiseless images, some pixel values reconstructed in the real images are noise captured by the underwater camera, irrelevant particulates in the water and compression artefacts (that should not be reconstructed) and thus hurt the PSNR score. Another reason for the low PSNR value is the presence of visual artefacts in the unseen structures of the images, we explain more about this in Section 5.1.

Finally, to see how iDehaze (a CNN-based method) compares to the state-of-the-art GAN-based models, we ran evaluations on the UFO-120 and EUVP datasets. As Table 2 shows, iDehaze outperformed all GAN-based models in the UIQM, and seemed to have a higher but less stable SSIM on the EUVP dataset.
Table 1. Performance of iDehaze in comparison to WaterNet [5], FUNIE-GAN [11], deep SESR [12], Shallow-UWNet [28]. Metrics: peak signal-to-noise ratio (PSNR), structural similarity index (SSIM), underwater image quality measure (UIQM)—↑ higher is better. The values with bold font indicates the best among the evaluated methods for each metric.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Metric</th>
<th>WaterNet</th>
<th>FUNIE-GAN</th>
<th>Deep SESR</th>
<th>Shallow-UWNet</th>
<th>iDehaze (Ours)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IUER</td>
<td>PSNR †</td>
<td>19.11 ± 3.68</td>
<td>19.13 ± 3.91</td>
<td>19.26 ± 3.56</td>
<td>18.99 ± 3.60</td>
<td>17.96 ± 2.79</td>
</tr>
<tr>
<td></td>
<td>SSIM †</td>
<td>0.79 ± 0.09</td>
<td>0.73 ± 0.11</td>
<td>0.73 ± 0.11</td>
<td>0.67 ± 0.13</td>
<td>0.80 ± 0.07</td>
</tr>
<tr>
<td></td>
<td>UIQM †</td>
<td>3.02 ± 0.34</td>
<td>2.99 ± 0.39</td>
<td>2.95 ± 0.39</td>
<td>2.77 ± 0.43</td>
<td>3.28 ± 0.33</td>
</tr>
<tr>
<td>EUVP</td>
<td>PSNR †</td>
<td>24.43 ± 4.64</td>
<td>26.19 ± 2.87</td>
<td>25.30 ± 2.63</td>
<td>27.39 ± 2.70</td>
<td>23.01 ± 1.97</td>
</tr>
<tr>
<td></td>
<td>SSIM †</td>
<td>0.82 ± 0.08</td>
<td>0.82 ± 0.08</td>
<td>0.81 ± 0.07</td>
<td>0.83 ± 0.07</td>
<td>0.84 ± 0.09</td>
</tr>
<tr>
<td></td>
<td>UIQM †</td>
<td>2.97 ± 0.32</td>
<td>2.84 ± 0.46</td>
<td>2.95 ± 0.32</td>
<td>2.98 ± 0.38</td>
<td>3.11 ± 0.36</td>
</tr>
<tr>
<td>UFO-120</td>
<td>PSNR †</td>
<td>23.12 ± 3.31</td>
<td>24.72 ± 2.57</td>
<td>26.46 ± 3.13</td>
<td>25.20 ± 2.88</td>
<td>17.55 ± 1.86</td>
</tr>
<tr>
<td></td>
<td>SSIM †</td>
<td>0.73 ± 0.07</td>
<td>0.74 ± 0.06</td>
<td>0.78 ± 0.07</td>
<td>0.73 ± 0.07</td>
<td>0.72 ± 0.07</td>
</tr>
<tr>
<td></td>
<td>UIQM †</td>
<td>2.94 ± 0.38</td>
<td>2.88 ± 0.41</td>
<td>2.98 ± 0.37</td>
<td>2.85 ± 0.37</td>
<td>3.29 ± 0.26</td>
</tr>
</tbody>
</table>

5.1. Discussion

Below, we discuss the main takeaways from our experimental results.

Qualitative results: Figures 5–8 show qualitative results from the output of several different image-enhancement methods. In Figure 5, the images had to be resized in previous works due to limitations in their implementation, such limitations do not exist in the iDehaze pipeline, which can dehaze images at their original aspect ratio.

UIER metrics: The underwater image-enhancement field has a big challenge with metrics. Namely, it can be very difficult to express the relationships between metrics, such as UIQM and SSIM, when they are looked at in isolation. First, it’s important to state that our dehazing model learns to specifically deal with haze, and is trained on specialized data that isolates that feature in its image and ground truth. Therefore, it removes significant amounts of haze in the mid to high range in deep underwater images. This makes the output of the dehazing model noticeably sharper, and its structure noticeably clearer than the input, and even more different than the ground truth. “Enhancing” an image means changing its structure, and that will cause the SSIM value to inevitably drop. It can also be noted that a model that enhances an image but has a very low SSIM is not desirable; because the enhanced image needs to be substantially similar to the input image. Our pipeline dramatically changes the amount of haze in the input images, which will cause the SSIM score to decrease. We argue that to accurately evaluate the performance of any image-enhancement model, the SSIM/PSNR values should be considered in tandem with the UIQM and qualitative results. Furthermore, the results of such experiments should be calculated with the exact same constants (c1, c2 and c3) and ideally the exact same code to be accurately comparable.

Strengths of the patch-based approach: In our pipeline, we split the images at the input of the U-NET [30] (the CNN used in the iDehaze pipeline) into patches and reconstruct them together at the output. Because of this, iDehaze is not sensitive to the input image size during training and can accept various sizes and image qualities as the input—an important feature when the availability of high-resolution, labelled real underwater images is limited. This also multiplies the available training data by a large factor. However, if this approach is used for inference, stitching the image patches together can create a patchwork texture in some images, which appears from time to time in the iDehaze image outputs. It is possible to remedy this by having large patch sizes, large overlaps between patches, and averaging the overlap prediction values at reconstruction. Our approach at inference exploits the flexible nature of the U-NET; therefore, we pad the images to the nearest square resolution divisible by 16, and feed the entire image to the network, resulting in clean inference output images with no patchwork issues or artefacts.

The use of compressed images: A frustrating fact about the available image datasets in the UIER field is the use of compressed image formats. More specifically, the JPG and JPEG
file formats use lossy compression to save disk space. Image compression can introduce artefacts that while invisible to the human eye, will affect the neural network’s performance. Hopefully, as newer and more sophisticated image datasets are gathered in the UIER field, the presence of compressed images will eventually fade away. To take a step in the correct direction, the iDehaze synthetic dataset uses lossless 32-bit PNG images and will be freely available for public use.

**Figure 5.** Qualitative comparison between iDehaze (rightmost image) and RedChannel [33], GDCP [34], blurriness and light absorption (UIBLA) [35], Fusion-Based [36], FunIE-GAN method [11] and UWCNN [37]. Each of these six images (a–f) were obtained from [29].

**Figure 6.** Sample inputs and outputs from various datasets (UIEB [5], EUVP [11], UFO-120 [12]). Features learned from the dehazing simulation transfer well to other domains in image restoration, as iDehaze brings back lost information in the scattered areas and dark parts of the images.
Table 2. Quantitative analysis on the EUVP and UFO-120 across various GAN-based models. Best values are highlighted in bold. Sample outputs from the mentioned models are found in Figures 7 and 8. ↑ higher is better. The values with bold font indicates the best among the evaluated methods for each metric.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Metric</th>
<th>Deep SESR</th>
<th>FUnIE-GAN</th>
<th>FUnIE-GAN-UP</th>
<th>UGAN</th>
<th>UGAN-P</th>
<th>iDehaze (Ours)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EUVP</td>
<td>PSNR</td>
<td>25.30 ± 2.63</td>
<td>26.19 ± 2.87</td>
<td>25.21 ± 2.75</td>
<td>26.53 ± 3.14</td>
<td><strong>26.53 ± 2.96</strong></td>
<td>23.01 ± 1.97</td>
</tr>
<tr>
<td></td>
<td>SSIM</td>
<td>0.81 ± 0.07</td>
<td>0.82 ± 0.08</td>
<td>0.78 ± 0.06</td>
<td>0.80 ± 0.05</td>
<td>0.80 ± 0.05</td>
<td><strong>0.84 ± 0.09</strong></td>
</tr>
<tr>
<td></td>
<td>UIQM</td>
<td>2.95 ± 0.32</td>
<td>2.84 ± 0.46</td>
<td>2.93 ± 0.45</td>
<td>2.89 ± 0.43</td>
<td>2.93 ± 0.41</td>
<td><strong>3.11 ± 0.36</strong></td>
</tr>
<tr>
<td>UFO-120</td>
<td>PSNR</td>
<td><strong>26.46 ± 3.13</strong></td>
<td>24.72 ± 2.57</td>
<td>23.29 ± 2.53</td>
<td>24.23 ± 2.96</td>
<td>24.11 ± 2.85</td>
<td>17.55 ± 1.86</td>
</tr>
<tr>
<td></td>
<td>SSIM</td>
<td>0.78 ± 0.07</td>
<td>0.74 ± 0.06</td>
<td>0.67 ± 0.07</td>
<td>0.69 ± 0.07</td>
<td>0.69 ± 0.07</td>
<td>0.72 ± 0.07</td>
</tr>
<tr>
<td></td>
<td>UIQM</td>
<td>2.98 ± 0.37</td>
<td>2.88 ± 0.41</td>
<td>2.60 ± 0.45</td>
<td>2.54 ± 0.45</td>
<td>2.59 ± 0.43</td>
<td><strong>3.29 ± 0.26</strong></td>
</tr>
</tbody>
</table>

Figure 7. Qualitative analysis on the EUVP dataset [11]. Other methods include UGAN and UGAN with gradient penalty (UGAN-P) [13], deep SESR [12] and FUnIE-GAN [11].
Figure 8. Qualitative analysis on the UFO-12 dataset [12]. Other methods include UGAN and UGAN with gradient penalty (UGAN-P) [13], deep SESR [12] and FUnIE-GAN [11].

5.2. Future Work

The two-step approach of iDehaze can often effectively enhance underwater images by bringing out details and restoring lost information, especially for images with significant scattering, wide optic depth, and relatively uniform optic depth. However, in certain cases, the full iDehaze pipeline may appear “over-processed” to the human eye. In such cases, the colour model alone produces a more aesthetically pleasing result. An interesting future direction for this research could be the development of an automated system that selects between the color model output and the full iDehaze pipeline based on psychometric criteria. In addition, investigating the effects of non-homogeneous media, changes in water temperature, and changes in density on the performance of dehazing systems could provide valuable insights and contribute to the advancement of this field.

6. Conclusions

In this paper, we presented iDehaze, a state-of-the-art image dehazing and colour transfer pipeline. Our proposed system includes a 3D simulation toolkit capable of generating millions of customizable, unique photorealistic underwater images with physics-based scattering and attenuation enabled by real-time ray tracing. In our pipeline, we break down the larger task of underwater image enhancement into two steps: dehazing and colour transfer. Our experiments demonstrate that iDehaze is capable of reconstructing clear images from raw, hazy inputs, achieving a state-of-the-art SSIM score on the EUVP dataset and state-of-the-art UIQM scores for the UIEB, UFO-120 and EUVP datasets despite not being trained on the latter two datasets at all. These results showcase the strengths of a carefully curated, physically modelled synthetic dataset made using 3D digital content.
creation tools. Our synthetic dataset, benchmarks and code will be released upon publication. We believe the image-enhancement research community will greatly benefit from the availability of free, high-quality, high-resolution labelled training data for underwater image dehazing and restoration tasks.

**Author Contributions:** Conceptualization, M.M. and A.A.; methodology, M.M.; software, M.M.; validation, M.M.; investigation, A.A.; resources, A.A. and R.E.; writing—original draft preparation, M.M. and R.E.; writing—review and editing, M.M. and A.A.; funding acquisition, A.A. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was partially funded by National Science Foundation (NSF) under the grant number 2000475.

**Data Availability Statement:** The data will be made available upon request. Please email author Dr. Ashwin Ashok to request data.

**Conflicts of Interest:** The authors declare no conflict of interest.

### References


37. Li, C.; Anwar, S.; Porikli, F. Underwater scene prior inspired deep underwater image and video enhancement. *Pattern Recognit.* **2020**, *98*, 107038. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.