



## Article

# Raindrop-Aware GAN: Unsupervised Learning for Raindrop-Contaminated Coastal Video Enhancement

Jinah Kim <sup>1,†</sup> , Dong Huh <sup>2,†</sup>, Taekyung Kim <sup>2</sup>, Jaeil Kim <sup>2,\*</sup> , Jeseon Yoo <sup>1</sup> and Jae-Seol Shim <sup>1</sup>

<sup>1</sup> Korea Institute of Ocean Science and Technology, 385 Haeyang-ro, Yeongdo-gu, Busan 49111, Korea; jakim@kiost.ac.kr (J.K.); jyoo@kiost.ac.kr (J.Y.); jsshim@kiost.ac.kr (J.-S.S.)

<sup>2</sup> Kyungpook National University, 80 Daehak-ro, Buk-gu, Daegu 41566, Korea; herick@knu.ac.kr (D.H.); paperrune@knu.ac.kr (T.K.)

\* Correspondence: jaeilkim@knu.ac.kr; Tel.: +82-53-950-7283

† J. Kim and D. Huh contributed equally to this work.

Received: 24 September 2020; Accepted: 16 October 2020; Published: 21 October 2020



**Abstract:** We propose an unsupervised network with adversarial learning, the *Raindrop-aware GAN*, which enhances the quality of coastal video images contaminated by raindrops. Raindrop removal from coastal videos faces two main difficulties: converting the degraded image into a clean one by visually removing the raindrops, and restoring the background coastal wave information in the raindrop regions. The components of the proposed network—a generator and a discriminator for adversarial learning—are trained on unpaired images degraded by raindrops and clean images free from raindrops. By creating raindrop masks and background-restored images, the generator restores the background information in the raindrop regions alone, preserving the input as much as possible. The proposed network was trained and tested on an open-access dataset and directly collected dataset from the coastal area. It was then evaluated by three metrics: the peak signal-to-noise ratio, structural similarity, and a naturalness-quality evaluator. The indices of metrics are 8.2% (+2.012), 0.2% (+0.002), and 1.6% (−0.196) better than the state-of-the-art method, respectively. In the visual assessment of the enhanced video image quality, our method better restored the image patterns of steep wave crests and breaking than the other methods. In both quantitative and qualitative experiments, the proposed method more effectively removed the raindrops in coastal video and recovered the damaged background wave information than state-of-the-art methods.

**Keywords:** coastal video enhancement; raindrop removal; background information recovery; generative adversarial network; unsupervised learning

## 1. Introduction

Coastal area plays an important role in national economy, commerce, and recreation. However, these regions are currently threatened by climate change, sea-level rise, beach erosion, extreme storms, and coastal urbanization [1]. Studying coastal area is a complicated task. Coastal research is strongly interrelated with hydrodynamics, morphodynamics, and anthropogenic interactions, and is also linked to geological, meteorological, hydrological, and biological processes [2]. Furthermore, these processes and their complex interactions vary on temporal scales from seconds to decades and on spatial scales from centimeters to tens of kilometers. Although many processes and interactions have been elucidated over the past few decades, the remaining scientific challenges require advancements in simulation and observation [3].

Remote sensing can observe the large-scale variability over a long-term period, and provides coastal measurements even in extreme events. In particular, land-based remote-sensing equipment such as cameras and video systems can provide long-term, sub-meter-scale consecutive images with high

temporal resolution. They also cover large spatial areas during extreme events. Conventional image processing techniques extract the shoreline position and infer the sub-surface morphology [4,5], from coastal video images, providing measurements for long-term studies of coastal behavior [6–8].

In recent years, deep learning has greatly benefited computer vision problems such as object detection, motion tracking, action recognition, and semantic segmentation [9]. Deep learning for computer vision overcomes the limitations of conventional image processing and analysis techniques by learning large amounts of data. Consequently, the performance is determined by the quantity and quality of the training data.

An essential problem in coastal studies is the nearshore wave process, which especially in extreme events is highly nonlinear and uncertain. Deep learning-based computer vision is expected to identify this phenomenon in numerous video images obtained from coastal video systems. Unfortunately, during extreme events, video image data are often degraded by bad weather conditions such as rain, wind, and fog. Raindrops on the camera lens especially degrade the visibility of video images. The degraded video data are difficult to use as training data for learning the coastal wave behavior due to low visibility of background wave information.

In addition, annotation of the video is not only very labor-intensive, but ambiguous labeling of contaminated video images degrades learning performance. Thus, supervised learning approach with paired dataset, including images and corresponding labels, is difficult to apply coastal videos. Therefore, unsupervised learning from unlabeled data is a promising direction for the coastal analysis of accumulated video data. It means that obtaining a larger amount of quality-controlled data are essential in learning-based approaches with an unsupervised manner.

This paper proposes an unsupervised learning framework with a generative adversarial network (GAN)-based video generation (*Raindrop-aware GAN*) to enhance a raindrop-contaminated coastal video. To detect and correct the degraded video images attentively, the generator produces a mask image for raindrop regions with various sizes and shapes together with the background-restored image through an encoder-decoder architecture with dilated convolution blocks and long skip connections. A mask sparsity loss function guides the generator to create the mask image by focusing only on the contaminated regions. In addition, the generator is also trained using a clean image set, which does not necessarily correspond to the contaminated image set, in order to identify raindrop-free regions. This behavior is achieved by a regularization loss function representing image differences between the clean images and their reconstructed images by the generator. The adversarial training of the generator with a discriminator helps to create undistorted images that follow a clean image distribution.

The novel contributions of this paper are: (1) constructing unsupervised *Raindrop-aware GAN* architecture to correct the raindrop-degraded coastal videos by learning behavior of propagating coastal waves; (2) developing new loss functions for adversarial learning of *Raindrop-aware GAN* and attentively correcting the raindrop-degraded coastal video imagery; and (3) validating applicability of the proposed method to real-world coastal monitoring videos acquired at Anmok beach in South Korea.

The remainder of the paper is organized as follows. Section 2 briefly introduces the related work on raindrop detection and removal by deep neural networks. Section 3 introduces our GAN-based method and Section 4 describes the three datasets used in our experiments. Section 5 details our data configuration and experiments, and introduces the state-of-the-art models employed in the performance comparison. Section 6 presents the quantitative and qualitative comparison results. Sections 7 and 8 discuss the applicability of our proposed method to reconstruct raindrop-contaminated wave scenes and concludes the paper with directions of future work, respectively.

## 2. Related Work

Raindrop-removal techniques can be categorized into two main groups: image processing methods that mathematically model-driven and data-driven approaches.

Garg and Nayar [10] introduced a physics-based motion blur model and a correlation model, which captures the photometry and dynamics of rain by detecting the rain streaks in sequential



images. The rain streaks are identified by their common movements and appearances in the two models. Roser and Geiger [11] proposed a photometric raindrop model that detects the spherical shapes of raindrops. However, neither of these techniques restores the background information. The region-completion method of You et al. [12] accomplishes both raindrop detection and background restoration using the neighboring pixel information. Although this method restores the raindrop-occluded pixels in the frequency domain, it is limited to static raindrops with spherical shapes.

Lauded for their excellent performance in image processing and translation, convolutional neural networks (CNNs) have been introduced to various data-driven approaches for raindrop removal. CNN-based techniques can be classified into supervised learning and unsupervised learning, depending on whether or not clean images corresponding to distorted images are used as the ground truth in the training phase.

Employing a supervised CNN restored the background of raindrop-distorted regions in a patch-wise manner [13]. High performance is achieved only on images contaminated by scattered small-sized sprinkles. Qian et al. [14] investigated raindrop removal by supervised image-to-image translation methods with adversarial learning, and proposed the attentive generative adversarial network (*Attentive GAN*) [14]. *Attentive GAN* generates a raindrop mask using a long short-term memory (LSTM)-based network, which guides the subsequent generator network to restore the background in the distorted regions. *Attentive GAN* achieves higher similarities between the restored images and the corresponding clean images than conventional methods such as *Pix2Pix* [15]. However, *Attentive GAN* requires an additional ground-truth dataset for supervising the raindrop-mask generation. More recently, Peng et al. [16] introduced a fully-CNN with skip connections and concurrent channel and spatial attention modules. Their method removes raindrops from single images, but supervised approaches invariably require paired datasets including clean images. Backgrounds with dynamically changing subjects are not easily acquired as clean images with accurate correspondence to distorted images. For this reason, supervised methods are usually evaluated on a simulated dataset.

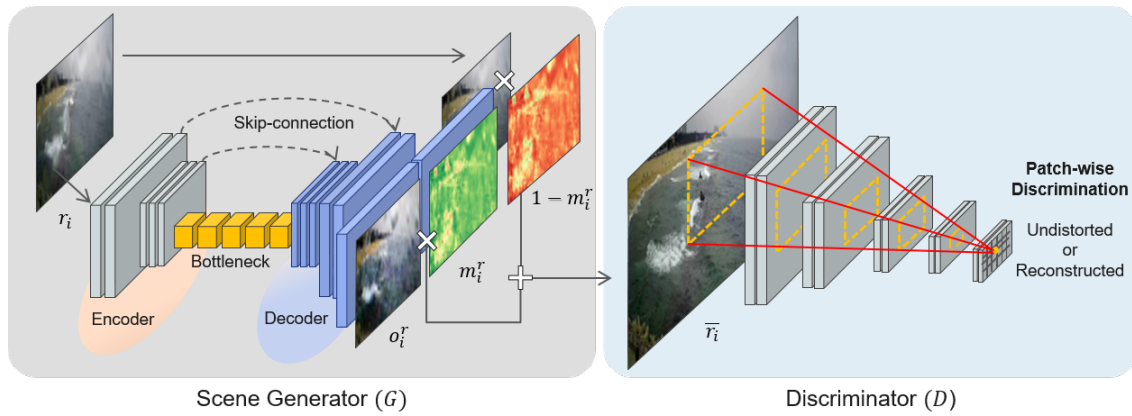
Unsupervised image-to-image translation with adversarial learning has been drawing attention in the field of remote sensing, owing to their advantages in using unlabeled and unpaired data set for training. Recently, a cycle-consistent GAN has been exploited to transcode synthetic-aperture-radar (SAR) images into optical images for building change detection using optical-like features [17,18]. GAN architectures have also proven their advantages using discriminative features of the discriminator in hyperspectral image classification [19,20]. Unsupervised image-to-image translation has been employed for domain adaptation in aerial image segmentation [21]. In raindrop removal, to avoid the paired-data requirement, Uzun and Temizel [22] and Wei et al. [23] proposed unsupervised and semi-supervised approaches using GAN architectures to map the distorted images to a target distribution of clean images. However, as shown in Qian et al. [14], image-to-image translation without attention to raindrops often fails the background reconstruction task in undistorted regions. On the contrary, we present an unsupervised GAN with a raindrop-restoration focus and regularization for improved background learning.

### 3. Methodology

#### 3.1. Raindrop-Aware GAN

The proposed raindrop-removal framework (*Raindrop-aware GAN*) is an unsupervised image-to-image translation method with adversarial learning, which transforms a raindrop-distorted image into a clean image domain. As shown in Figure 1, the proposed network consists of two sub-networks: (1) a conditional generative model called the scene generator, which restores the backgrounds in the raindrop regions, and (2) a discriminator that captures the distributions of the distorted and clean image samples in a patch-wise manner. The scene generator attentively restores the backgrounds by focusing on the raindrop regions determined by a raindrop mask, while preserving

the input image as much as possible. The adversarial learning in the scene generator restores the natural appearance of the coastal waves in the backgrounds. The restoration level is sufficient to deceive the discriminator.



**Figure 1.** Architecture of the proposed *Raindrop-aware GAN*.

### 3.2. Scene Generator for Raindrop-Aware Background Reconstruction

Given an input image ( $r_i$ ) contaminated by raindrops, the scene generator ( $G$ ) produces a background-restored image ( $o_i^r$ ) and a raindrop mask ( $m_i^r$ ). The mask represents the raindrop regions on a scale from zero (background) to one (raindrops). Our scene generator recognizes the raindrop regions without any supervision, and transforms the distorted images into a distribution of clean images while preserving the backgrounds of the raindrop regions. The reconstruction is formulated as follows:

$$\bar{r}_i = \{(1 - m_i^r) \odot r_i\} + m_i^r \odot o_i^r \quad (1)$$

where  $i$  is the image index of coastal videos,  $r$  indicates the distorted images, and  $\bar{r}_i$  is the final output of the generator through the fusion of the background-restored image ( $o_i^r$ ) and the input image using the raindrop mask. The scene generator is implemented as a fully-convolutional network with skip connections.

Figure 1a shows the structure of the scene generator. It is composed of three parts: (1) an encoder that learns the significant features from the inputs; (2) a bottleneck that builds a deep architecture with a large receptive field; and (3) a decoder that generates the background-restored image and a raindrop mask with the same input resolution.

The encoder consists of a series of convolution blocks. Each block comprises  $3 \times 3$  convolutional layers followed by a rectified linear unit (RELU) activation function [24]. For downsampling operations in the feature encoding, the second and last convolution blocks are performed with a stride of 2. Thus, the spatial resolution of the encoder output is  $w/4 \times h/4$ , where  $w$  is the width, and  $h$  is the height of the input image, respectively. The bottleneck part is a series of residual blocks with dilated convolution layers [25]. Each residual block combines a shortcut connection by addition and a sequence of two  $3 \times 3$  dilated convolutions followed by RELU functions. The dilated convolution efficiently increases the receptive field while maintaining spatial resolution. The bottleneck part with multiple residual blocks ensures the robustness of *Raindrop-aware GAN* when detecting raindrop regions of various sizes. In the decoder, transposed convolutional layers with a stride of 2 are used for upsampling operations, and the convolutional blocks are used for reconstruction. Among the four output channels of the decoder, three channels are allocated to the background-restored image, and one channel is reserved for the raindrop mask. To minimize the loss of image details by frequent convolution and downsampling operations, we employ long skip-connections that deliver the early-layer features of the encoder to the decoder by concatenation.

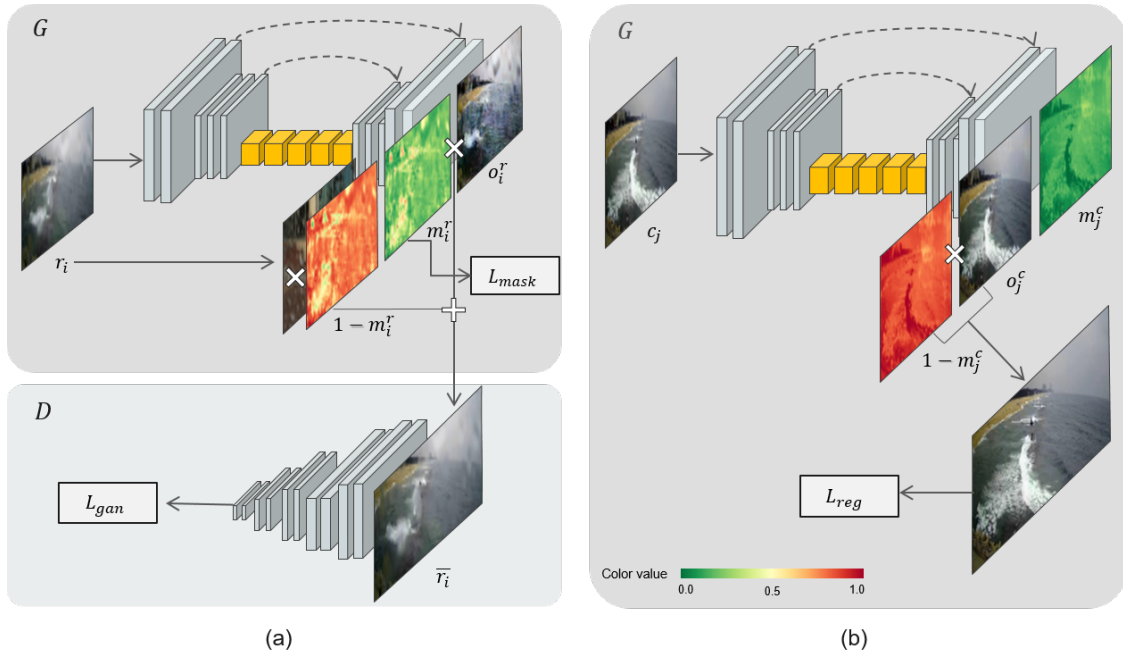
The scene generator is trained to minimize three loss functions:  $\mathcal{L}_{gen}$  for adversarial learning against the discriminator,  $\mathcal{L}_{mask}$  for enforcing the sparsity of the raindrop mask, and  $\mathcal{L}_{reg}$  for background learning.  $\mathcal{L}_{gen}$  and the discriminator loss are discussed to describe the adversarial learning between the generator and discriminator networks in the next Section 3.3. The total loss for training the scene generator is also described along with  $\mathcal{L}_{gen}$ . In the reconstruction operation (Equation (1)), the mask sparsity loss ( $\mathcal{L}_{mask}$ ) is an  $l_1$  loss that prevents easy saturation (1.0) of the mask, and which drives the focus to the regions distorted by raindrops. The loss function ( $\mathcal{L}_{mask}$ ) is given by

$$\mathcal{L}_{mask} = \mathbb{E}_{r_i \in \mathbb{I}_{rain}} [\| m_i^r \|_1] \quad (2)$$

where  $\mathbb{I}_{rain}$  is the set of training instances of distorted images. Because the raindrop discrimination is unsupervised, we regularize the scene generator to produce a zero-valued mask in the undistorted background regions while correctly reconstructing the backgrounds. To achieve this task, we compute the regularization loss ( $\mathcal{L}_{reg}$ ) of selected clean images ( $c_j$ ) that are randomly sampled from the training set:

$$\mathcal{L}_{reg} = \mathbb{E}_{c_j \in \mathbb{I}_{clean}} [\| (1 - m_j^c) \odot o_j^c - c_j \|_1] \quad (3)$$

where  $m_j^c$  is the mask,  $o_j^c$  is the background-restored version of image  $c_j$ , and  $\mathbb{I}_{clean}$  is the training set of clean images.  $c$  indicates the outputs of the scene generator from the clean images. Figure 2 shows the forward propagation of the input samples in the scene generator when computing the loss functions of unsupervised training.



**Figure 2.** Forward propagations of the input in the scene generator of *Raindrop-aware GAN* when computing (a) the loss functions  $\mathcal{L}_{mask}$  and  $\mathcal{L}_{gan}$  for the sparse mask and adversarial learning, respectively; and (b) the loss function  $\mathcal{L}_{reg}$  for background learning.

### 3.3. Discriminator for Regional Raindrop Pattern Recognition

During adversarial learning, the discriminator ( $D$ ) is trained to distinguish clean images from the images reconstructed by the scene generator. As a classifier, the discriminator learns the distributions of the real and fake samples; meanwhile, the scene generator attempts to deceive it by producing more realistic images. Because the raindrop distortions are observed locally, we construct the discriminator in the *PatchGAN* architecture [15]. The discriminator is composed of a series of convolution blocks (convolutional layers followed by a RELU activation function). After multiple downsampling

operations through convolution blocks with a stride of 2, the output dimension of the discriminator is  $w/32 \times h/32 \times 1$ . Each element of the output feature represents one patch of the given input (see the discriminator in Figure 1). The loss function of the discriminator is given by

$$\mathcal{L}_D = \mathbb{E}_{r_i \in \mathbb{I}_{rain}} [\|D(\bar{r}_i)\|_2^2] + \mathbb{E}_{c_j \in \mathbb{I}_{clean}} [\|D(c_j) - 1\|_2^2] \quad (4)$$

where  $\bar{r}_i$  is the image reconstructed by Equation (1). To deceive the discriminator, the scene generator is trained using the following loss function:

$$\mathcal{L}_{gen} = \mathbb{E}_{r_i \in \mathbb{I}_{rain}} [\|D(\bar{r}_i) - 1\|_2^2] \quad (5)$$

The total loss of the scene generator is then computed as

$$\mathcal{L}_G = \omega_1 \mathcal{L}_{reg} + \omega_2 \mathcal{L}_{mask} + \omega_3 \mathcal{L}_{gen} \quad (6)$$

where  $\omega_1$ ,  $\omega_2$ , and  $\omega_3$  are hyperparameters that control the importance of each loss function. As the  $\omega_1$  and  $\omega_2$  values increase, the masks more rapidly approach to zero; conversely, increasing  $\omega_3$  drives more areas of the mask toward one.

Algorithm 1 performs the unsupervised training of *Raindrop-aware GAN* on the training sets of the distorted and clean images. The Shuffle ( $\cdot$ ) function in Algorithm 1 shuffles the data indices of the training sets, and the Next, ( $\cdot$ ) function sequentially calls the data instances of the next indices. In the testing stage, single video frames, extracted from coastal videos, are given to the trained generator. Note that the block of frame sequences can be input to the generator as well. This approach can help reduce the processing time for entire videos. Since the generator learns the backgrounds to regenerate them using the loss  $\mathcal{L}_{reg}$ , extracted frame sequences including undistorted images can be given without any further processing.

---

**Algorithm 1** Unsupervised training procedure for the *Raindrop-aware GAN*

---

**Input:**  $\mathbb{I}_{rain}, \mathbb{I}_{clean}, N$  (Number of epochs),  $\omega_1, \omega_2, \omega_3$  (loss weights)

**for** 0 to  $|N| - 1$  **do**

$\mathbb{I}'_{rain} = \text{Shuffle}(\mathbb{I}_{rain})$

$\mathbb{I}'_{clean} = \text{Shuffle}(\mathbb{I}_{clean})$

**while** True **do**

**if**  $i = |\mathbb{I}'_{rain}| - 1$  **then**

**break**

**end if**

$r_i = \text{Next}_i(\mathbb{I}'_{rain})$

$c_j = \text{Next}_j(\mathbb{I}'_{clean})$

      Compute  $\bar{r}_i$  using Equation (1) with  $r_i$  as an input

      Compute  $\mathcal{L}_{mask}$  and  $\mathcal{L}_{gan}$  of G with  $r_i$  as an input

      Compute  $\mathcal{L}_{reg}$  of G with  $c_i$  as an input

      Compute  $\mathcal{L}_D$  of D with  $c_i$  and  $\bar{r}_i$  as an input

      Update parameters of D using  $\mathcal{L}_D$

      Update parameters of G using Equation (6)

**end while**

**end for**

**Output:** Trained networks G and D

---

#### 4. Dataset

Three kinds of datasets were used in this study: (1) the open-access dataset, *Raindrop1119* [14]; (2) *Anmok paired dataset*; and (3) *Anmok unpaired dataset*.

#### 4.1. Raindrop1119

Qian et al. [14] created a dataset of raindrops attached to a glass window or lens, and released it for use in raindrop removal studies in *Attentive GAN*. Their dataset includes 1119 image pairs, one degraded by raindrops, the other free from raindrops in the same background scene (corresponding ground-truth image of degraded one). The outdoor background scenes and the raindrop sizes and distributions are varied among the image pairs. This dataset, hereafter referred to as *Raindrop1119*, contains 891 training data and 58 test data. Samples from *Raindrop1119* are shown in Figure 3.



**Figure 3.** Samples selected from the *Raindrop1119* dataset [14] (a) images degraded by raindrops; and (b) their corresponding ground-truth images.

#### 4.2. Coastal Wave Monitoring at Anmok Beach

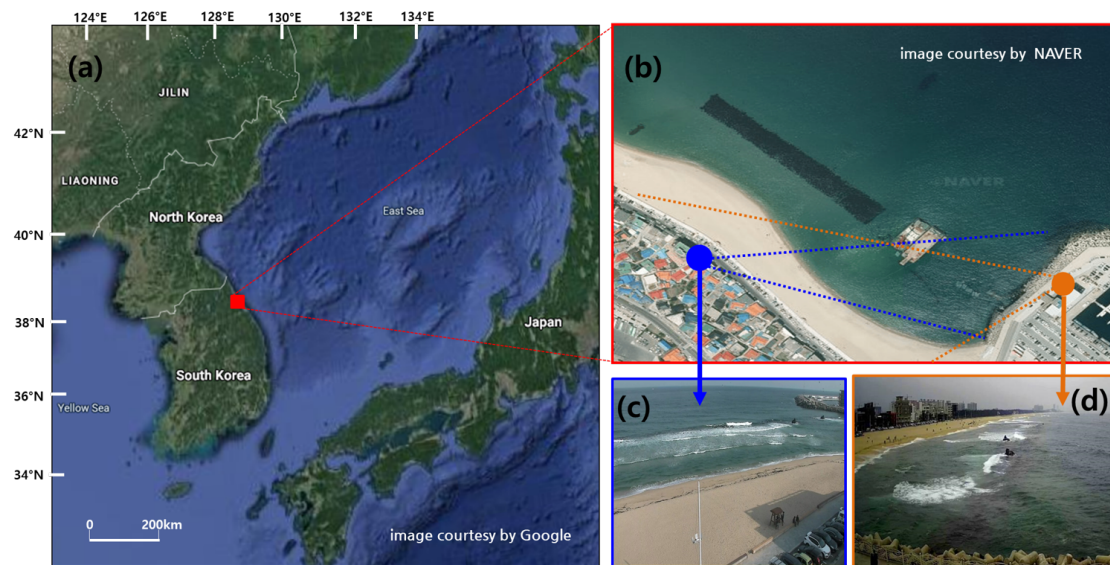
The video enhancement study is conducted at Anmok beach, a straight, almost 4 km-long strips dominated by micro-tidal waves. The beach is located on the east coast of South Korea (Figure 4a,b), which has eroded in the last few decades. To understand the associated physical process, a video monitoring system using general Closed Circuit Television (CCTV) has been installed, and the video data captured in 2016 and 2017 have been stored. The camera locations and an image of the camera view are indicated in (b) and (c) of Figure 4, respectively. In this study, we additionally installed CCTVs to acquire raindrop-contaminated coastal video images paired with their clean images (see (d) in Figure 4) for a fair comparison with the supervised approaches using ground-truth. Details of the paired coastal video acquisition are described in the following section.

##### 4.2.1. Anmok Paired Dataset

Similarly to *Raindrop1119*, pairs of raindrop images at Anmok beach were generated for fairly comparing the proposed unsupervised method with the state-of-the-art supervised baseline models. The camera location is indicated in Figure 4b, and samples of the acquired video images are shown in Figure 5. An image of the camera view is displayed in Figure 4d. Hereafter, this dataset is referred to as the *Anmok paired dataset*. The images in Figure 5a are degraded by raindrops, whereas those in Figure 5b are raindrop-free. The images (ground-truth) are paired against the same background.

The two CCTVs were connected for simultaneously recording the raindrop-contaminated video images and their corresponding raindrop-free images. Under inclement weather conditions, the CCTV lens-distortion pattern caused by raindrops on the instrument was replicated by a (20 × 15 cm) transparent glass plate attached to the front of the CCTV lens, and the raindrop pattern was simulated by a sprayer. Clear images without distortion were acquired as the corresponding synchronous CCTV images taken at the same time as the contaminated images, and with the same glass plate attached. The collected images contained 13 different raindrops and raindrop deformation images collected at 1-min intervals. The samples showed the spatial and temporal changes in the raindrops over approximately 20 min. From each video, we randomly extracted video frames (16,841 frames total) and divided them into two sets according to acquisition time: 13,605 data for training and 3236 data for verification. The training and test sets are not overlapped and not extracted from the same videos.





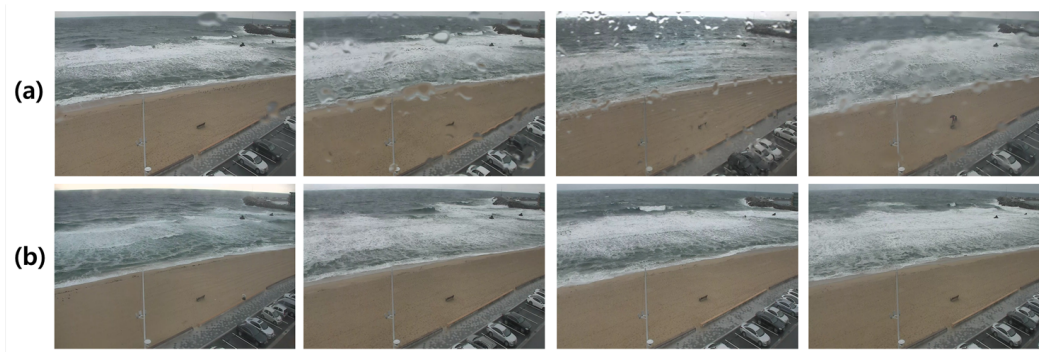
**Figure 4.** Study area of Anmok Beach located in north-eastern South Korea (a). The CCTV video system was installed at two locations (b); one for the unpaired raindrops dataset (blue circle) and one for the paired raindrop dataset (green circle). The locations for collecting the unpaired- and paired-image datasets are shown in (c,d), respectively. Dark pixels in (b) indicate the area where the submerged breakwater is installed.



**Figure 5.** Samples of our *Anmok paired dataset*: (a) images degraded with raindrops and (b) their corresponding ground-truth images. The dark pixels at the point where the wave breaks indicate the area where the submerged breakwater is installed.

#### 4.2.2. Anmok Unpaired Dataset

To evaluate the applicability of the proposed method, we also acquired coastal video images from the video monitoring system (see Figure 4c), installed at Anmok beach. Samples of the acquired video images are shown in Figure 6. Hereafter, this dataset is called *Anmok unpaired dataset*. Unlike the *Raindrop1119*, Figure 6b is free from raindrops but do not correspond to the raindrop-contaminated images in Figure 6a. The spatial and temporal resolutions of the video system are  $1920 \times 1080$  and 30 frames per second (fps), respectively. From the video clips acquired in November 2016 and 2017, we randomly selected 233 and 360 clips for training and validating the proposed networks, respectively (training:validation ratio = 8:2). All videos were recorded in daytime and cover different wave-breaking and light conditions. Each video clip is approximately 10 min long. For unsupervised learning, which excludes the labeling task, all video frames recorded at 30 fps were used without downsampling.



**Figure 6.** Samples of our *Anmok unpaired dataset*: (a) images degraded by raindrops; and (b) clean (raindrop-free) images not corresponding to those in (a).

## 5. Experiments

### 5.1. Implementation Details

The kernel parameters of the proposed networks were randomly initialized by *He* initialization [26] and trained by an Adam optimizer [27]. In both the scene generator and discriminator, the parameters  $\beta_1$ ,  $\beta_2$ , and  $\epsilon$  of the Adam optimizer were set to 0.5, 0.999, and 0.001, respectively. The learning rate ( $\eta$ ) was 0.0001, and the training datasets were augmented with random horizontal flipping. The size of the images input to the networks was fixed to  $720 \times 480$ . The hyperparameters  $\omega_1$ ,  $\omega_2$ , and  $\omega_3$  were empirically set to 1.0, 0.01, and 0.5 respectively through hyperparameter tuning process using a tuning set, which was randomly selected from the *Raindrop1119* dataset. The size of the tuning set was 10% of the training set. The hyperparameters were applied equally to all experiments.

All experiments were conducted on a workstation, equipped with a single TITAN XP (12 GB), Intel i9 CPU, and 32 GB main memory, with a batch size of 16. Training time depends on the size of the training data. We trained the proposed networks with the observation of the validation loss to avoid overfitting.

### 5.2. Baseline Models with Supervised Learning

In this experiment, the proposed method was competed against *Pix2Pix* [15] and *Attentive GAN* [14]. *Pix2Pix* is a supervised image-to-image translation method based on adversarial learning. The generator of *Pix2Pix* is trained to minimize the  $l_1$  loss between the reconstructed images and their corresponding ground truth images, along with the adversarial loss against the discriminator. The generator of *Attentive GAN* is composed of two sub-networks: an attentive network based on convolutional LSTM that generates the attention mask of the given input, and a fully-convolutional network that produces a de-rained image from the attention mask and input. *Attentive GAN* requires the binary raindrop masks of all training instances for supervised learning. The hyperparameters of *Pix2Pix* and *Attentive GAN* were tuned using the validation sets.

### 5.3. Evaluation Metrics

The accuracy of the background restoration on each method was evaluated by the peak signal-to-noise ratio (PSNR) [28], the structural similarity index (SSIM) [29], and the natural image quality evaluator (NIQE) [30].

The PSNR is the ratio of the maximum possible power of a signal and the power of the corrupting noise. The SSIM metric captures the perceived quality loss between two image sequences. As reconstructed images are more similar to their paired clean images than the raw images, the PSNR of a reconstructed image should be high and the SSIM should approach 1.0. The PSNR and SSIM are employed to measure the image quality with respect to the ground-truth, undistorted images.

Meanwhile, the NIQE is a non-referenced image-quality score that assesses the perceptual image-quality enhancement. The NIQE measures the distance between the statistics-based features in an image of a natural scene and the same features obtained from image databases. The features are modeled as multidimensional Gaussian distributions, called the space domain natural scene statistic (NSS) model. The NIQE can assess images with arbitrary distortions such as blurring, ringing, and noise. A lower NIQE score indicates a better image quality.

## 6. Results

### 6.1. Comparison with Baseline Models Using the Raindrop1119

For a fair comparison with the supervised-learning baseline methods, we evaluated the accuracy and image-quality improvements on *Raindrop1119*'s test dataset. Table 1 lists the PSNR, SSIM, and NIQE scores calculated from the images reconstructed by *Pix2Pix*, *Attentive GAN*, and the proposed *Raindrop-aware GAN*. The average PSNR and SSIM between the distorted images and their corresponding clean images (ground-truth) in *Raindrop1119* were 24.078 and 0.850, respectively. The *Pix2Pix* method generated images with lower perceptual quality (NIQE = 12.296) than the distorted images. The *Attentive GAN* and our *Raindrop-aware GAN* improved all three metrics of the raindrop-removal performance. In this experiment, the performance of *Attentive GAN* was evaluated using the trained model provided by Qian et al. [14].

**Table 1.** Accuracy and image-quality assessment of the two baseline models and our proposed method, which were trained only training data of *Raindrop1119* dataset, with test data of *Raindrop1119* dataset.

	Distorted Images	<i>Pix2pix</i>	<i>Attentive GAN</i>	<i>Raindrop-Aware GAN</i>
PSNR	24.078	27.725	31.618	28.277
SSIM	0.850	0.876	0.920	0.882
NIQE	10.524	12.296	10.171	10.233

Figure 7 shows the reconstruction results of the two baseline models and *Raindrop-aware GAN* on *Raindrop1119*'s test data. As evidenced by its best performance scores, *Attentive GAN* successfully removed most of the raindrops, whereas the *Pix2Pix* reconstruction left gray stains in the raindrop regions. *Raindrop-aware GAN* also generated high-quality images but could not remove tiny raindrops from some test examples.

According to these results, direct supervised learning of the raindrop masks is advantageous for recognizing small-sized and scattered patterns. The acquisition process of the raindrop mask in Qian et al. [14] used empirically derived thresholds for the paired sets in *Raindrop1119*. It is important to note that *Raindrop1119* was acquired in a city with a limited field of view and artificial raindrops. The mask acquisition method using simple thresholds is difficult to apply in practical cases such as coastal videos having moving backgrounds like waves.





**Figure 7.** Test examples from the *Raindrop1119* dataset: (a) distorted examples; (b) their corresponding clean images (ground-truth); and the reconstruction results of (c) *Pix2Pix*; (d) *attentive GAN*; and (e) the proposed *Raindrop-aware GAN*.

## 6.2. Performance Evaluation on the Anmok Paired Dataset

To evaluate the effectiveness of the baselines and the proposed model on actual coastal videos, we acquired a paired dataset of the wave videos at Anmok Beach. As described above, the dataset was captured by two connected CCTV cameras placed side by side. While recording the coastal waves, we sprayed water droplets on one camera multiple times to simulate the image distortions caused by raindrops. The geometric differences between the corresponding images acquired by the cameras were minimized by rigid registration. To simulate various raindrop-distortion patterns, we acquired 15 images splashed with artificial raindrops. Each video clip was filmed for approximately one minute (on average) to capture the wave movements. From the video datasets of both cameras, we randomly sampled 16,841 paired frames at the same time points, and divided them into a 13,605-image training set and a 3236-image test set.

In this experiment, we re-trained the baselines and *Raindrop-aware GAN* that were previously trained on *Raindrop1119*, and investigated their validity on the newly acquired datasets. *Attentive GAN* performed poorly after fine-tuning because the threshold-based method in Qian et al. [14] could not acquire the correct mask images for localizing the raindrops.

Table 2 shows the experimental results. After unsupervised re-training and fine-tuning, *Raindrop-aware GAN* outperformed *Pix2Pix* and *Attentive GAN*, yielding the highest performance indicators of raindrop removal (PSNR = 26.505, SSIM = 0.940, NIQE = 11.878). The NIQE was of limited value in comparing the image-quality scores of coastal video images. Although the PSNR and SSIM scores reflected the degraded performance of *Attentive GAN*, the image reconstruction of *Attentive*

GAN yielded a lower (i.e., better) average NIQE score than the distorted images. These results may be caused by complex patterns with many white particles observed during wave breaking.

**Table 2.** Accuracy and image quality assessment of the tested methods on *Anmok paired dataset*.

	Distorted Images	<i>Pix2Pix</i>	<i>AttentiveGAN</i>	Pre-Trained <i>Raindrop-Aware</i> GAN	Fine-Tuned <i>Raindrop-Aware</i> GAN
PSNR	25.872	25.134	24.493	25.354	26.505
SSIM	0.938	0.930	0.938	0.932	0.940
NIQE	12.820	15.158	12.074	11.961	11.878

In the visual assessment of image quality, the fine-tuned *Raindrop-aware* GAN better restored the image patterns of steep wave crests and breaking (see Figure 8) than the other methods. Black stains and blurred regions appeared in the images reconstructed by *Pix2Pix*, *Attentive* GAN, and the pre-trained *Raindrop-aware* GAN. These results indicate that wave-specific patterns in coastal video images are difficult to restore without additional training processes. Whereas supervised learning methods must acquire paired images with and without raindrops, our unsupervised approach can utilize all coastal videos acquired from outdoor visual-sensing systems.

### 6.3. Application to the *Anmok Unpaired Dataset*

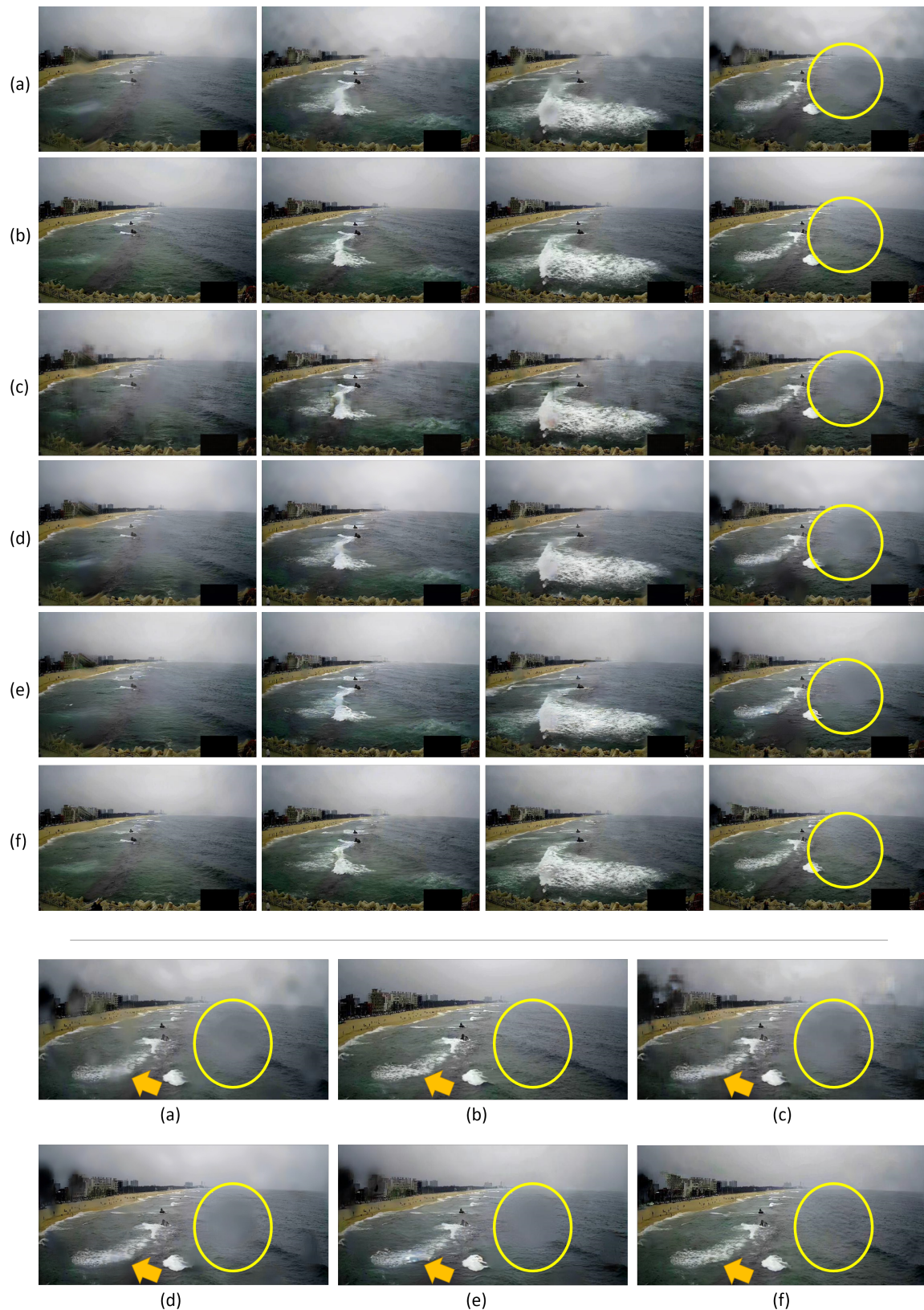
To validate the proposed method in practice, we assessed the image qualities of the baselines (the state-of-the-art models) and *Raindrop-aware* GAN on the *Anmok unpaired dataset*. From the coastal videos acquired over two months, we collected separate video clips under wet and dry conditions and then randomly sampled 12,000 video frames from the two datasets. Note that no temporal correspondence exists between the raining and non-raining video datasets. In the validation experiment, we took the *Raindrop-aware* GAN trained on *Raindrop1119* and fine-tuned it on the *Anmok unpaired dataset*.

In this experiment, we measured only the NIQE because the image sets were not paired. Table 3 gives the image quality scores of the reconstructed images. *Raindrop-aware* GAN outperformed *Pix2Pix* and *Attentive* GAN; moreover, the fine-tuned *Raindrop-aware* GAN obtained clearer boundaries of the propagated waves and wave-like patterns in the reconstructed images than the other methods (see Figure 9).

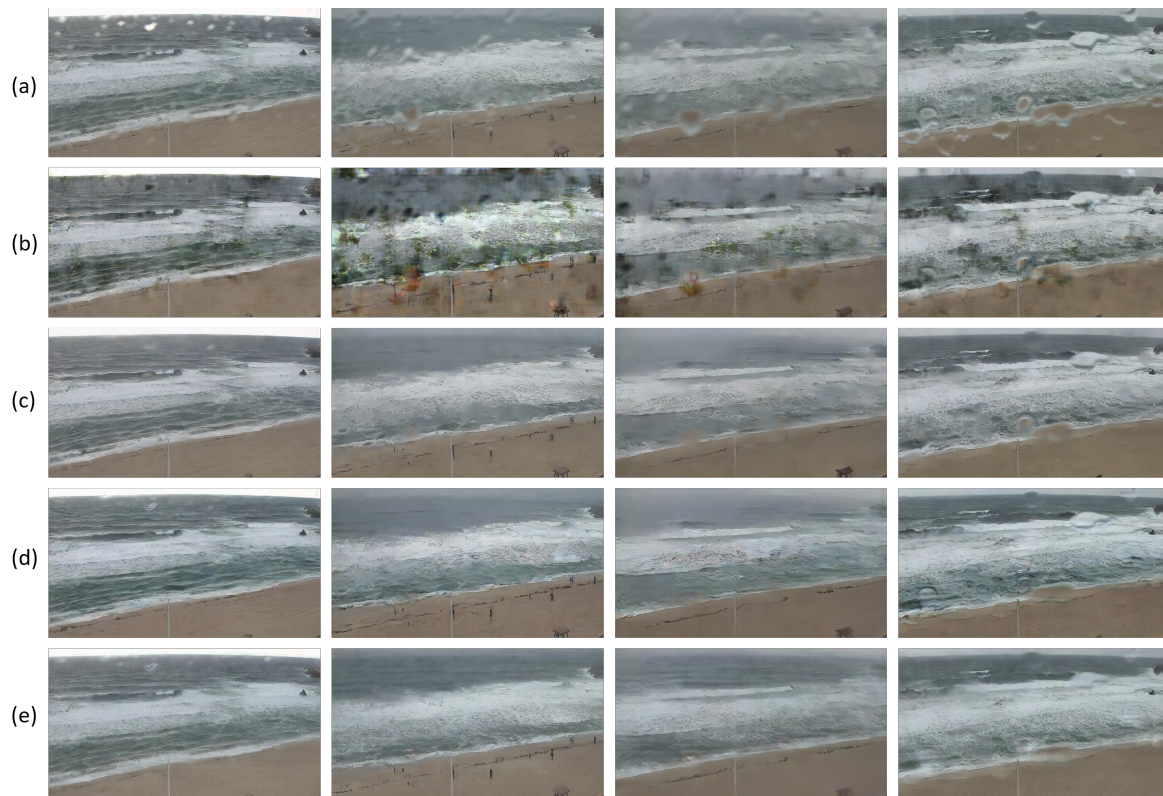
**Table 3.** Image-quality assessment on raindrop-contaminated video images taken from a CCTV that monitors coastal areas at Anmok Beach (*Anmok unpaired dataset*).

	Distorted Images	<i>Pix2Pix</i>	<i>Attentive</i> GAN	Pre-Trained <i>Raindrop-Aware</i> GAN	Fine-Tuned <i>Raindrop-Aware</i> GAN
NIQE	17.498	16.336	12.648	16.083	13.333





**Figure 8.** Test examples from the *Anmok paired dataset*: (a) distorted examples; (b) their corresponding clean images; and the reconstruction results of (c) *Pix2Pix*; (d) *Attentive GAN*; (e) pre-trained *Raindrop-aware GAN*; and (f) fine-tuned *Raindrop-aware GAN*. The right most examples are displayed in larger size in bottom rows. Circles and arrows indicate areas of steep wave crests and breaking, respectively.



**Figure 9.** Test examples of the CCTV monitoring coastal areas at Anmok Beach: (a) Distorted examples; and the reconstructed images of (b) *Pix2Pix*; (c) *Attentive GAN*; (d) pre-trained *Raindrop-aware GAN*; and (e) fine-tuned *Raindrop-aware GAN*.

## 7. Discussion

The proposed method provides a learning-based approach to enhance raindrop-contaminated coastal video. To see how the enhancement of the coastal image from video monitoring system is helpful for video-based coastal dynamic research beyond the evaluation metric mainly used in deep learning-based image enhancement, we would like to examine the applicability of the proposed method through timestack images.

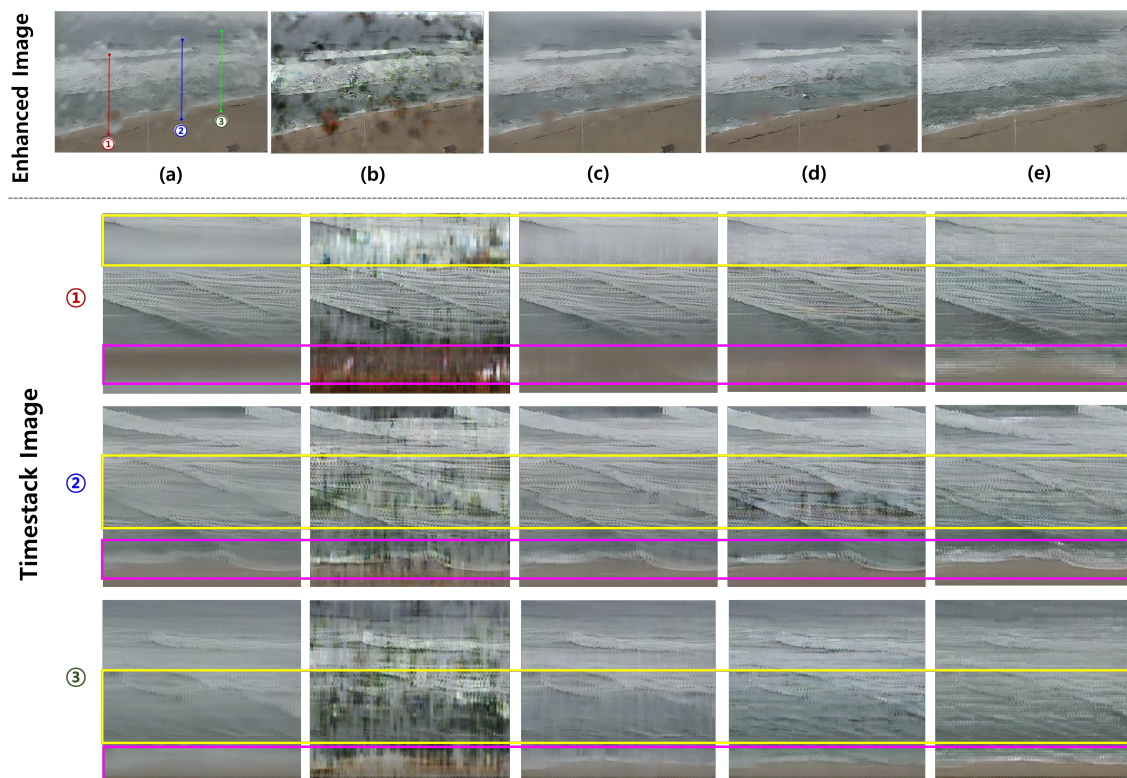
A temporally stacked image, called a timestack image, is an image in which the intensity of an array of pixels is plotted along time. It is very useful to monitor long-term shoreline and bathymetry evolution and to track and estimate individual breaking waves on the surf and swash zone. (a)–(e) in the upper part of the Figure 10 show the raindrop-contaminated image, the reconstructed images obtained using *Pix2Pix*, *Attentive GAN*, pre-trained *Raindrop-aware GAN*, and fine-tuned *Raindrop-aware GAN*, respectively. The lower part of the Figure 10 shows a group of timestack image created by accumulating consecutive 300 frames of images for 6 s for the five types from (a) to (e) of images on the upper part, and for each the three cross-shore line transect marked ①, ②, and ③.

Looking at the timestack image created along the first line transect ①, the refraction due to large raindrops in the sand side (pink box area) and the contaminated sea area (yellow box area) are overall well reconstructed when using the fine-tuned *Raindrop-aware GAN* in the timestack image placed in the last rightmost column. In the timestack image created along the ② and ③ line transect, it is clear that the white foam of the breaking waves on the sand side is best reconstructed in the proposed method (pink box area) and the crest of the breaking wave is clearly displayed in the sea area (yellow box area) in the timestack images placed in the last rightmost column.

The timestack images placed in the second column from the right vertically show the results of *Raindrop-aware GAN* trained with *Raindrop1119* dataset only. This is slightly inferior in reconstruction compared to the images located in the rightmost column. It shows that the best reconstruction



performance can be obtained when the coastal video is used for fine-tuning even in the same model architecture of *Raindrop-aware GAN*.



**Figure 10.** Test examples with time-stack images: The upper part shows the enhanced images of raindrop-contaminated image (a) by applying *Pix2Pix* in (b); *Attentive GAN* in (c); pre-trained *Raindrop-aware GAN* in (d); and fine-tuned *Raindrop-aware GAN* in (e). The lower part shows the timestack images acquired at the line transect marked ①, ②, and ③ in (a).

By creating timestack image and visually assessment it, we can confirm that the performance of the proposed method is the best and it also has high applicability in studying nearshore wave dynamics with video remote sensing, in particular data preparation step, such as breaking wave height estimation from coastal video [31,32], video sensing of nearshore bathymetry evolution [33,34], nearshore wave transform with video imagery [35], shoreline response and resilience through video monitoring [36–39], wave run-up prediction [40,41], and nearshore wave tracking through coastal video [42,43].

## 8. Conclusions

We performed unsupervised learning with a GAN-based video generation method that enhances coastal video images contaminated by raindrops. Unlike recent approaches based on supervised learning, which require the pairing of degraded images with clean (ground-truth) images, the proposed raindrop-removal network *Raindrop-aware GAN* is an unsupervised learning method. *Raindrop-aware GAN* attentively corrects the degraded region with minimal changes to the raindrop-free areas in the contaminated image. For this purpose, it learns the raindrop region and its surroundings, and then generates a mask image mapped with the spatial-attention information. The scene generator in the proposed method is expanded by adding the mask generated by background learning, and is supplemented with a discriminator that distinguishes the raindrop and raindrop-free regions in a patch-wise manner using adversarial learning.

To evaluate its reference performance, the proposed network was pre-trained on the open dataset *Raindrop1119*. For unsupervised learning, the paired dataset (the paired clean and degraded images)

was ignored and the whole dataset was shuffled in random order. Via transfer learning, the pre-trained network was applied to coastal video images of Anmok Beach. The Anmok video dataset was continuously acquired over a long period by CCTVs. To quantitatively verify the proposed network, we collected an additional dataset of clean images paired with raindrop-contaminated images.

The PSNR, SSIM, and NIQE performance indices of *Raindrop-aware GAN* were 8.2% (+2.012), 0.2% (+0.002), and 1.6% (−0.196) better than those of *Attentive GAN*, respectively, and 5.5% (+1.151), 1.1% (+0.008), and 21.6% (−0.083) better than those of *Pix2Pix*, respectively. In addition, the raindrops were removed from the raindrop-contaminated Anmok video images, and the coastal wave information of the background was well restored in the removal areas.

The images from cameras, video systems, and other land-based remote-sensing technologies are severely degraded by bad weather conditions during extreme events. In such situations, unsupervised learning-based data-driven modeling is essential. Therefore, our proposed method is expected to assist the data-preparation stage of vision-based remote-sensing studies.

However, whether or not the method correctly restores the movement of propagating waves in continuous time is difficult to quantify. To correct this uncertainty, we will encode not only the spatial features, but also the temporal features in an extended version of our network. Moreover, we intend to modify the architecture of the scene generator for recognizing temporal changes in the raindrop regions. For this purpose, we will employ a recurrent sub-network.

**Author Contributions:** Conceptualization, J.K. (Jaeil Kim) and J.K. (Jinah Kim); methodology, J.K. (Jaeil Kim), D.H., and J.K. (Jinah Kim); software, D.H. and T.K.; computing resource, J.K. (Jaeil Kim); validation, D.H. and J.K. (Jinah Kim); formal analysis, D.H. and J.K. (Jaeil Kim); data curation, D.H., T.K., and J.K. (Jaeil Kim), J.K. (Jinah Kim) and J.Y.; writing—original draft preparation, J.K. (Jinah Kim), J.K. (Jaeil Kim), and D.H.; writing—review and editing, J.K. (Jinah Kim) and J.K. (Jaeil Kim); visualization, D.H. and J.K. (Jinah Kim); funding acquisition, J.K. (Jinah Kim), J.Y., and J.-S.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Development of the ‘AI-based Coastal Disaster Modeling Platform and Sea-fog Prediction System’, the basic research program funded by the Korea Institute of Ocean Science and Technology (PE99842) and ‘Construction of Ocean Research Stations and their Application Studies’, the project funded by the Ministry of Oceans and Fisheries, Republic of Korea. In addition, this work also supported by Electronics and Telecommunications Research Institute (ETRI) grant funded by ICT R&D program of MSIT/IITP [No. 2020-0-00843, Development of low power satellite multiple access core technology based on LEO cubesat for global IoT service].

**Acknowledgments:** This research was the result of a study on the “HPC Support” Project, supported by the ‘Ministry of Science and ICT’ and NIPA. We thank the anonymous reviewers whose comments/suggestions helped improve and clarify this manuscript.

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

## References

1. Carter, R.W.G. *Coastal Environments: An Introduction to the Physical, Ecological, and Cultural Systems of Coastlines*; Elsevier: Amsterdam, The Netherlands, 2013.
2. Davidson-Arnott, R.; Bauer, B.; Houser, C. *Introduction to Coastal Processes and Geomorphology*; Cambridge University Press: Cambridge, MA, USA, 2019.
3. Elko, N.; Feddersen, F.; Foster, D.; Hapke, C.; McNinch, J.; Mulligan, R.; Ozkan Haller, H.T.; Plant, N.; Raubenheimer, B. The future of nearshore processes research. *Shore Beach* **2015**, *83*, 13.
4. Aarninkhof, S.; Ruessink, B.; Roelvink, J. Nearshore subtidal bathymetry from time-exposure video images. *J. Geophys. Res. Ocean.* **2005**, *110*, C6. [\[CrossRef\]](#)
5. Plant, N.G.; Holland, K.T.; Haller, M.C. Ocean wavenumber estimation from wave-resolving time series imagery. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 2644–2658. [\[CrossRef\]](#)
6. Holman, R.A.; Stanley, J. The history and technical capabilities of Argus. *Coast. Eng.* **2007**, *54*, 477–491. [\[CrossRef\]](#)
7. Holman, R.; Haller, M.C. Remote sensing of the nearshore. *Annu. Rev. Mar. Sci.* **2013**, *5*, 95–113. [\[CrossRef\]](#)
8. Splinter, K.D.; Harley, M.D.; Turner, I.L. Remote sensing is changing our view of the coast: Insights from 40 years of monitoring at Narrabeen-Collaroy, Australia. *Remote Sens.* **2018**, *10*, 1744. [\[CrossRef\]](#)

9. Voulodimos, A.; Doulamis, N.; Doulamis, A.; Protopapadakis, E. Deep learning for computer vision: A brief review. *Comput. Intell. Neurosci.* **2018**, *2018*, 7068349. [[CrossRef](#)]
10. Garg, K.; Nayar, S.K. Detection and removal of rain from videos. In Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Washington, DC, USA, 27 June–2 July 2004.
11. Roser, M.; Geiger, A. Video-based raindrop detection for improved image registration. In Proceedings of the 2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops, Kyoto, Japan, 27 September–4 October 2009; pp. 570–577.
12. You, S.; Tan, R.T.; Kawakami, R.; Ikeuchi, K. Adherent Raindrop Detection and Removal in Video. In Proceedings of the The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Portland, OR, USA, 23–28 June 2013.
13. Eigen, D.; Krishnan, D.; Fergus, R. Restoring an image taken through a window covered with dirt or rain. In Proceedings of the IEEE international Conference on Computer Vision, Sydney, Australia, 1–8 December 2013; pp. 633–640.
14. Qian, R.; Tan, R.T.; Yang, W.; Su, J.; Liu, J. Attentive generative adversarial network for raindrop removal from a single image. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 2482–2491.
15. Isola, P.; Zhu, J.Y.; Zhou, T.; Efros, A.A. Image-to-image translation with conditional adversarial networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1125–1134.
16. Peng, J.; Xu, Y.; Chen, T.; Huang, Y. Single-image raindrop removal using concurrent channel-spatial attention and long-short skip connections. *Pattern Recognit. Lett.* **2020**, *131*, 121–127. [[CrossRef](#)]
17. Fuentes Reyes, M.; Auer, S.; Merkle, N.; Henry, C.; Schmitt, M. Sar-to-optical image translation based on conditional generative adversarial networks—Optimization, opportunities and limits. *Remote Sens.* **2019**, *11*, 2067. [[CrossRef](#)]
18. Saha, S.; Bovolo, F.; Bruzzone, L. Building Change Detection in VHR SAR Images via Unsupervised Deep Transcoding. *IEEE Trans. Geosci. Remote Sens.* **2020**. [[CrossRef](#)]
19. Zhong, Z.; Li, J.; Clausi, D.A.; Wong, A. Generative adversarial networks and conditional random fields for hyperspectral image classification. *IEEE Trans. Cybern.* **2019**, *50*, 3318–3329. [[CrossRef](#)] [[PubMed](#)]
20. Saha, S.; Solano-Correa, Y.T.; Bovolo, F.; Bruzzone, L. Unsupervised Deep Transfer Learning-Based Change Detection for HR Multispectral Images. *IEEE Geosci. Remote Sens. Lett.* **2020**. [[CrossRef](#)]
21. Benjdira, B.; Bazi, Y.; Koubaa, A.; Ouni, K. Unsupervised domain adaptation using generative adversarial networks for semantic segmentation of aerial images. *Remote Sens.* **2019**, *11*, 1369. [[CrossRef](#)]
22. Uzun, Ü.; Temizel, A. Cycle-Spinning GAN for Raindrop Removal from Images. In Proceedings of the 2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), Taipei, Taiwan, 18–21 September 2019; pp. 1–6.
23. Wei, Y.; Zhang, Z.; Zhang, H.; Qin, J.; Zhao, M. Semi-DerainGAN: A New Semi-supervised Single Image Deraining Network. *arXiv* **2020**, arXiv:2001.08388.
24. Nair, V.; Hinton, G.E. Rectified linear units improve restricted boltzmann machines. In Proceedings of the 27th International Conference on Machine Learning (ICML-10), Haifa, Israel, 21–24 June 2010; pp. 807–814.
25. Yu, F.; Koltun, V. Multi-scale context aggregation by dilated convolutions. *arXiv* **2015**, arXiv:1511.07122.
26. He, K.; Zhang, X.; Ren, S.; Sun, J. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 1026–1034.
27. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
28. Huynh-Thu, Q.; Ghanbari, M. Scope of validity of PSNR in image/video quality assessment. *Electron. Lett.* **2008**, *44*, 800–801. [[CrossRef](#)]
29. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [[CrossRef](#)]
30. Mittal, A.; Soundararajan, R.; Bovik, A.C. Making a “completely blind” image quality analyzer. *IEEE Signal Process. Lett.* **2012**, *20*, 209–212. [[CrossRef](#)]
31. Gal, Y.; Browne, M.; Lane, C. Long-term automated monitoring of nearshore wave height from digital video. *IEEE Trans. Geosci. Remote Sens.* **2013**, *52*, 3412–3420. [[CrossRef](#)]



32. Andriolo, U.; Mendes, D.; Taborda, R. Breaking Wave Height Estimation from Timex Images: Two Methods for Coastal Video Monitoring Systems. *Remote Sens.* **2020**, *12*, 204. [\[CrossRef\]](#)
33. Holman, R.; Plant, N.; Holland, T. cBathy: A robust algorithm for estimating nearshore bathymetry. *J. Geophys. Res. Ocean.* **2013**, *118*, 2595–2609. [\[CrossRef\]](#)
34. Thuan, D.H.; Almar, R.; Marchesiello, P.; Viet, N.T. Video sensing of nearshore bathymetry evolution with error estimate. *J. Mar. Sci. Eng.* **2019**, *7*, 233. [\[CrossRef\]](#)
35. Andriolo, U. Nearshore Wave Transformation Domains from Video Imagery. *J. Mar. Sci. Eng.* **2019**, *7*, 186. [\[CrossRef\]](#)
36. Abessolo Ondo, G.; Bonou, F.; Tomety, F.S.; Du Penhoat, Y.; Perret, C.; Degbe, C.G.E.; Almar, R. Beach response to wave forcing from event to inter-annual time scales at Grand Popo, Bénin (Gulf of Guinea). *Water* **2017**, *9*, 447. [\[CrossRef\]](#)
37. Valentini, N.; Saponieri, A.; Danisi, A.; Pratola, L.; Damiani, L. Exploiting remote imagery in an embayed sandy beach for the validation of a runup model framework. *Estuarine, Coast. Shelf Sci.* **2019**, *225*, 106244. [\[CrossRef\]](#)
38. Da Silva, P.G.; Coco, G.; Garnier, R.; Klein, A.H. On the prediction of runup, setup and swash on beaches. *Earth-Sci. Rev.* **2020**, *2020*, 103148. [\[CrossRef\]](#)
39. Jôia Santos, C.; Andriolo, U.; Ferreira, J.C. Shoreline Response to a Sandy Nourishment in a Wave-Dominated Coast Using Video Monitoring. *Water* **2020**, *12*, 1632. [\[CrossRef\]](#)
40. Vousedoukas, M.I.; Wziatek, D.; Almeida, L.P. Coastal vulnerability assessment based on video wave run-up observations at a mesotidal, steep-sloped beach. *Ocean Dyn.* **2012**, *62*, 123–137. [\[CrossRef\]](#)
41. Atkinson, A.L.; Power, H.E.; Moura, T.; Hammond, T.; Callaghan, D.P.; Baldock, T.E. Assessment of runup predictions by empirical models on non-truncated beaches on the south-east Australian coast. *Coast. Eng.* **2017**, *119*, 15–31. [\[CrossRef\]](#)
42. Kim, J.; Kim, J.; Kim, T.; Huh, D.; Caires, S. Wave-Tracking in the Surf Zone Using Coastal Video Imagery with Deep Neural Networks. *Atmosphere* **2020**, *11*, 304. [\[CrossRef\]](#)
43. Sit, M.; Demiray, B.Z.; Xiang, Z.; Ewing, G.J.; Sermet, Y.; Demir, I. A comprehensive review of deep learning applications in hydrology and water resources. *Water Sci. Technol.* **2020**. [\[CrossRef\]](#)

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).