Journal of *Marine Science and Engineering*

MDPI

*Article*

# Restoration of Missing Patterns on Satellite Infrared Sea Surface Temperature Images Due to Cloud Coverage Using Deep Generative Inpainting Network

**Song-Hee Kang [1], Youngjin Choi [2] and Jae Young Choi [1,*]**

[1] Pattern Recognition and Machine Intelligence (PMI) Lab., Division of Computer & Electronic Systems Engineering, Hankuk University of Foreign Studies, Yongjin 17035, Korea; thdgml9776@hanmail.net
[2] GeoSystem Research Inc., Gunpo 15807, Korea; yjchoi@geosr.com
* Correspondence: jychoi@hufs.ac.kr; Tel.: +82-31-330-4906

**Abstract:** In this paper, we propose a novel deep generative inpainting network (GIN) trained under the framework of generative adversarial learning, which is optimized for the restoration of cloud-disturbed satellite sea surface temperature (SST) imagery. The proposed GIN architecture can achieve accurate and fast restoration results. The proposed GIN consists of rough and fine reconstruction stages to promote the details and textures of missing (clouded) regions in SST images. We also propose a nov el preprocessing strategy that replaces the land areas with the average value of daily oceanic surface temperatures for improving restoration accuracy. To learn the proposed GIN, we developed a novel approach that combines multiple loss functions well suited for improving the restoration quality over missing SST information. Our results show that the difference in temperature between restored and actual satellite image data was no larger than 0.7 °C in monthly average values, which suggests excellent resilience against the missing sea surface temperature data. The proposed GIN has a faster restoration time and is feasible for real-time ocean-related applications. Furthermore, the computational cost of restoring SST images is much lower than the popular interpolation methods.

**Keywords:** sea surface temperature; daily-averaged SST; NOAA-19 infrared data; loss of sea surface temperature; restoration; deep generative adversarial network

check for updates

**Citation:** Kang, S.-H.; Choi, Y.; Choi, J.Y. Restoration of Missing Patterns on Satellite Infrared Sea Surface Temperature Images Due to Cloud Coverage Using Deep Generative Inpainting Network. *J. Mar. Sci. Eng.* **2021**, *9*, 310. https://doi.org/10.3390/jmse9030310

Academic Editor: Fausto Pedro García Márquez

Received: 15 January 2021
Accepted: 4 March 2021
Published: 11 March 2021

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

**Copyright:** © 2021 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 (https://creativecommons.org/licenses/by/4.0/).

## 1. Introduction

Satellite remote sensing information of sea surface temperature (SST) has been available for more than 40 years. The advantage of geostationary satellite measurements is their high spatiotemporal resolution [1,2]. Raw satellite data products are processed at five levels from Level 0 to Level 4. Raw data (Level 0) are calibrated, geolocated (Level 1), and then converted into geophysical parameters (Level 2). Practically used Level 3 and Level 4 data are produced by space-mapping and the composite analysis of measurement data [3,4].

However, it is necessary to implement highly complicated processes to obtain Level 4 data. Since infrared measurements are disturbed by cloud coverage, lower-level data provide SSTs only for limited areas. Guan and Kawamura [5] showed that the annual-mean availability of the National Oceanic and Atmospheric Administration, Advanced Very High Resolution Radiometer (NOAA AVHRR) infrared measurements is 48% in the ocean south of Japan. To remove the cloud disturbance, the National Institute of Fisheries Science (NIFS) developed a correction system using data assimilation with in situ observations [6]. SST data in blank regions due to cloud coverage are restored by complex processes [7]. The observational data and SST anomalies are used to complement the blank regions. Since direct observations are insufficient in most cases, five-day data are used for interpolation. The operation of the correction system takes a couple of days, which makes the real-time operation impossible. Instead of using infrared measurements of SST which are obstructed

by clouds, microwave radiometry could offer a solution to the cloud occlusion problem as microwaves penetrate clouds with little attenuation, giving a clear view of the sea surface under all weather conditions except rain [8]. However, research work [8] showed the limitations for the use of microwave SST as follows. Much coarser resolution (about 50 km) is a major restriction of using microwave SST, compared to infrared SST (about 1 to 10 km) [8]. The EU plans to launch a new microwave radiometer satellite in 2025, named CIMR (Copernicus Imaging Microwave Radiometer) project [9]. CIMR will provide improved SST resolution (about 15 km), but still much lower than recently used infrared sensors (about 0.5 to 1 km). Moreover, the microwave estimates of SST are contaminated by land within about 50 km from coast [8]. For these reasons, the infrared SST has been widely used rather than the microwave SST; hence, restoration of infrared SST images in the cloud areas (in an accurate and efficient way) is very important, not only in the coastal areas, but also in the open oceans.

A generative adversarial network (GAN) is a class of deep learning methods proposed by Goodfellow et al. [10]. GANs consist of two neural networks: generator and discriminator networks. The structure of the discriminator network is the same as that of the convolutional neural network (CNN) while the generator network has a reverse-ordered structure of the discriminator network. Given a training set, the GAN learns to generate new data with the same statistics as the training data. Recently, GANs have been applied to the analysis of various satellite images. Morales et al. [11] detected shadow areas in Peru. Enomoto et al. [12] proposed multicontext GAN (MCGAN) that can remove clouds and predict the colors under the clouds. However, the method proposed by [12] is effective only when the cloud is sufficiently thin to be penetrated by the near-infrared band. Kadow et al. [13] applied deep learning-based image inpainting techniques for reconstructing missing climate information; however, their approach is limited to reconstructing global temperature patterns of only an irregular hole in images.

In this study, we propose a novel GAN architecture well suited for removing cloud occlusion (disturbances) and restoring sea surface temperatures to values close to observational SST images. Dong et al. [14] used the deep convolutional GAN (DCGAN) to recover SST images by removing cloud occlusion and showed that the use of GAN reveals better results than conventional interpolation algorithms. However, Dong et al. [14] only considered artificial occlusion regions rather than realistic missing regions during the evaluation of their method. Hirahara et al. [15] developed the adversarial physical model loss of the GAN to incorporate the data assimilation images, which represent the physical characteristics of SST images, into the restoration process. However, their approach is limited in restoring the SST images closer to data assimilation images (not the real SST images) and thus in restoring original satellite images. Shibata et al. [16] suggested using partially occluded images during the training of the GAN to increase the number of available training images, but one drawback to this approach is the computation of the loss values only from the nonoccluded region of training images assuming that noise-free data are present in the nonoccluded region of SST images. However, note that satellite images often contain noise during observations and hence one cannot always obtain noise-free data in nonoccluded regions. Therefore, the discriminator of the GAN [16] classifies noisy images as real images, which makes the generator reconstruct noisy occluded regions of SST images.

Although the idea of using GAN for SST image restoration has been previously introduced by [14–16], our proposed method, named generative inpainting network (GIN), was developed in a novel way. Key technical contributions of this work can be summarized as follows. First, we developed a novel preprocessing technique using a land mask to exclude land areas during the restoration stage; in addition, every pixel value in the land area of SST images was filled with the average of the RGB values of the ocean surface area to improve the restoration quality of images (Section 2.1). Second, to train our GIN for restoring the input SST images as realistically as possible, we defined and effectively combined multiple loss functions; in particular, we developed a novel "sea surface loss"

that enables the application of the information collected from the surrounding locations of the ocean surface regions (no land region) to the restoration of missing regions (Section 2.3). Third, our GIN is composed of two-stage networks [17,18], so-called coarse network and refinement network, to obtain a realistic and coherent completed SST image; such two-stage network architecture has proven effective for restoring missing regions with highly complex and arbitrary shapes. Our proposed GIN employs a multistage transfer learning framework, which consists of pretraining and fine-tuning stages. Satellite image data collected for 20 years were used for training, and temporal characteristics were considered to train the deep network properly. Various quantitative performance measures of the GIN model were used to evaluate the errors between the images via the conventional system of NIFS and our results.

The remainder of the paper is organized as follows. Section 2 contains a description of the dataset and deep network structures. Section 3 describes experimentation setup. Restoration results and the quantitative comparison are presented and discussed in Section 4. Section 5 presents the summary and conclusion.

## 2. Method

### 2.1. Random Masks and Preprocessing

Image inpainting (a.k.a. image completion or image hole-filling) is a task of synthesizing alternative contents in missing regions. We extend image inpainting techniques to the restoration (in the remainder of this paper, "restoration" and "reconstruction" will be used interchangeably) (or reconstruction) of temperature information captured in the form of an ocean satellite image. The goal of using random masks is to make randomly generated missing regions in each training image during the learning stage. In our method, a "random mask" is defined as a binary image with 256 × 256 pixel size. Each pixel of a random mask image has the value 1 for the position (missing region) of the pixel to be inpainted and 0 for the position (no missing region) to give RGB information used for temperature restoration. A random mask has randomly generated rectangles (composed of a value of "1") whose size is randomly determined in the range from 32 × 32 to 256 × 256 pixels. The random mask denoted by **R** is defined as follows:

$$\mathbf{R}(x,y) = \left\{ \begin{array}{c} 1 \text{ if } (x,y) \text{ is included in missing region} \\ 0 \text{ otherwise} \end{array} \right\} \quad (1)$$

The "ocean satellite image" used for training contains some "land area" at a fixed location. Note that the land area should not be used for restoring temperature because the goal of restoration is to reconstruct temperature information only for the ocean areas, and the land area represented by black pixel values in Figure 1a can adversely affect the restoration results. To solve this problem, we developed an effective preprocessing approach. We used a "land mask" to identify the land region, which makes it possible to exclude the land areas during the restoration stage. In addition, to improve restoration quality, every pixel value in the land area was replaced by the average of the RGB values of the ocean surface area (Figure 2). Based on the experimental results, the restoration quality is considerably improved by the proposed preprocessing strategy.

### 2.2. Generative Inpainting Network (GIN)

A generative inpainting network (GIN) was developed to recover the missing regions in ocean satellite images, referred to as image inpainting. As shown in Figure 3, the input of the GIN is an image with white pixels that indicate the missing regions, as well as a binary mask indicating the pixel location of the missing regions as input pairs, and the output of the GIN is the final restored image. Note that the binary mask enables the GIN to restore the missing region without modifying the nonoccluded region by selecting the pixels to be restored.

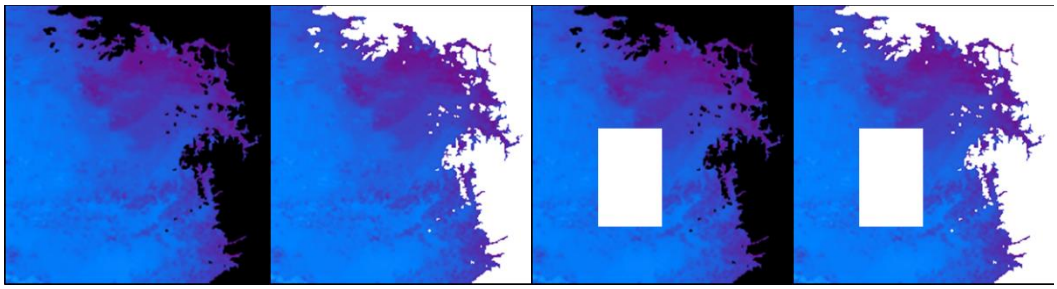



**Figure 1.** Examples of (**a**) ground truth of ocean satellite images, (**b**) land areas indicated by the land mask (white pixel values), (**c**) training image overlapped with the random mask to be restored, and (**d**) training image overlapped with both the land mask and random mask.

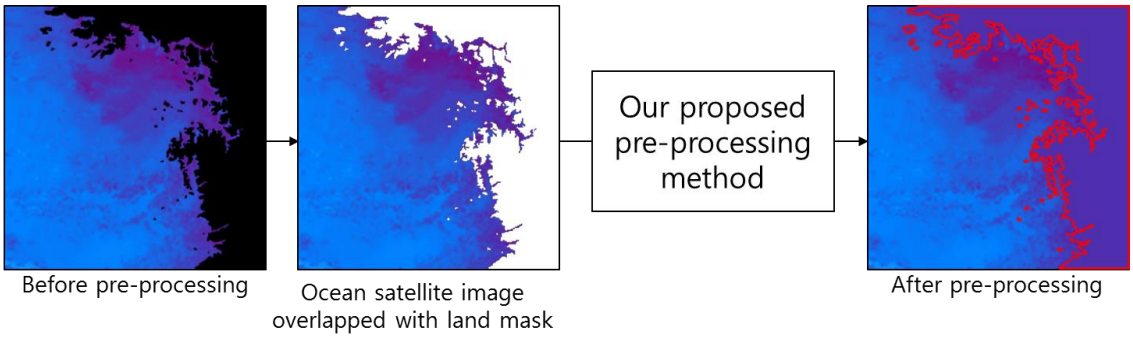



**Figure 2.** Illustration of the proposed preprocessing approach where every pixel value in the land area is replaced by the average of the RGB pixel values of the ocean surface area.

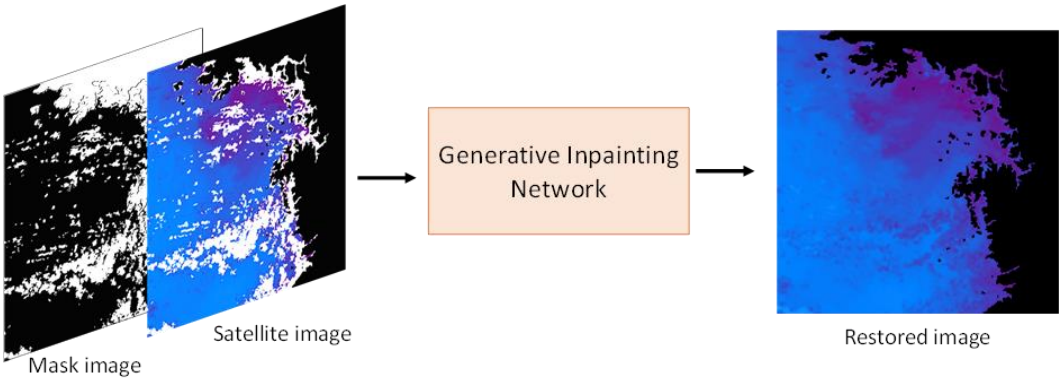



**Figure 3.** Proposed GIN model takes a pair of occluded SST and corresponding binary mask image as input and outputs a restored image without modifying the nonoccluded region.

The GIN model is designed for restoring a missing region with various sizes, complex shapes, and arbitrary positions because the proposed GIN inputs a pair of binary masks corresponding to the missing region and proceeds with image restoration for that region. The size of the input image is 256 × 256 pixels containing a part of the missing regions randomly sampled from the entire missing regions with a rectangular shape. This approach allows the GIN model to process input images with different sizes and multiple missing regions.

Figure 4 presents the overall process of constructing the GIN for restoring satellite sea surface temperature data. For the robust and stable restoration of missing regions, the GIN model adopts the state-of-the-art two-stage coarse-to-fine network architecture [17,18]. It consists of a coarse network that makes an initial coarse reconstruction and a refinement

network that takes the coarse reconstruction as input and then predicts the refined results with details and textures. The coarse network is trained with the reconstruction loss explicitly, while the refinement network is trained with the reconstruction as well as GAN losses. For improved generalization capability, the GIN model is first pretrained using the places2 dataset [19], and then each of the 12 different GIN models is fine-tuned using one of the corresponding 12 different monthly training data to learn the temperature characteristics and patterns in each month.

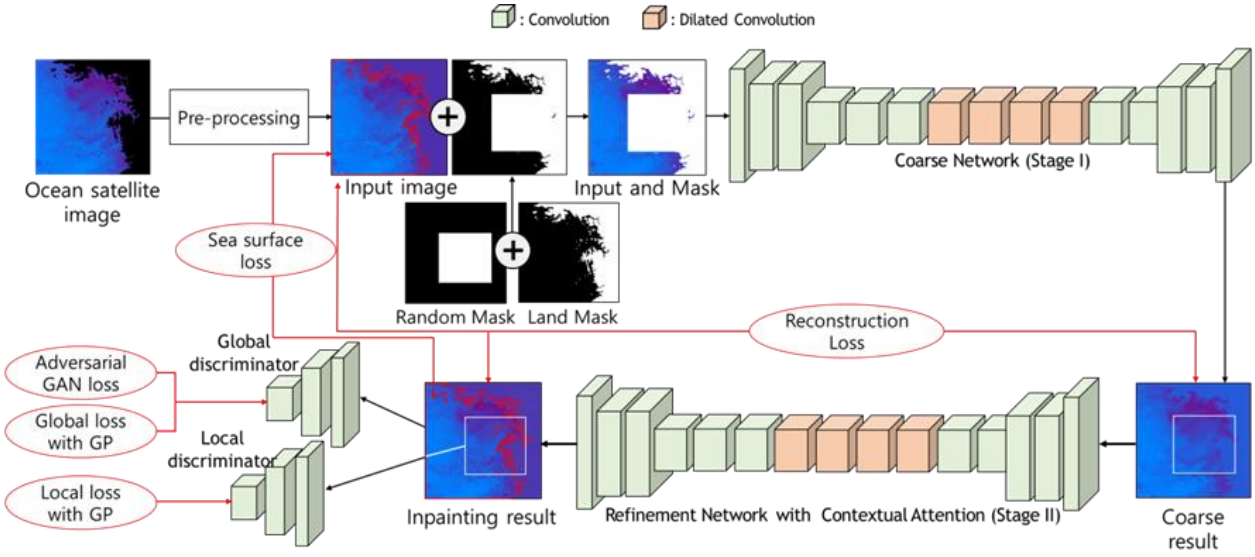



**Figure 4.** Overall process of constructing our generative inpainting network (GIN) to restore satellite sea surface temperature. It consists of a coarse network and refinement network, as well as two auxiliary discriminant networks that are used only for training the coarse and refinement networks; hence, they are not used during the testing. The coarse network takes the entire ocean satellite image as input, while the refinement network takes only a local region of the inpainted region as input. Note that the coarse network is trained with reconstruction and sea surface losses explicitly, while the refinement network is trained with adversarial GAN loss, and global and local losses with gradient penalty.

For the layer implementation of our inpainting network, we used mirror padding [20] for all convolution layers and removed batch normalization layers [21], which encourage spatial and/or color coherency. We also used Exponential Linear Units (ELUs) [22,23] as activation functions and clipped the output filter values. Our GIN is trained end-to-end with the adversarial GAN loss, reconstruction loss, and three Wasserstein GAN losses, two of which account for the global image while one accounts for the local patch of the missing regions. This allows for the use of both global and local feature representations for GIN training.

### 2.3. Loss Function

To implement the proposed GIN model, we developed a modified version of the Wasserstein GAN (WGAN) [24] with a gradient penalty. To train the GIN, we attached the WGAN loss to both global and local outputs of the second-stage refinement network to enhance the consistency of global and local regions of reconstructed ocean satellite images via image inpainting. Reference [24] reports that the WGAN loss outperforms general GAN losses for image generation tasks, and it also works well when combined with the $L_1$ reconstruction loss computed between the ground-truth (GT) and reconstructed images.

In order to train the GIN for restoring the input image realistically, five different loss functions are effectively combined: adversarial GAN loss, two gradient penalty losses, sea surface loss, and reconstruction loss. Note that an adversarial GAN loss is used for training stability, and three gradient penalty and reconstruction losses are used to improve the realism of the restoration result. Using the combination of five loss functions allows

the stable training of the high-performance network for image restoration (completion). Before introducing the combined loss function, we describe each of the five loss functions as follows.

As for the adversarial GAN loss, we adopted the WGAN loss that uses the earth mover's distance $W(\mathbf{P}_r, \mathbf{P}_g)$ to compute data distribution $\mathbf{P}_r$ and generated distribution $\mathbf{P}_g$. To compute the minimal cost of transporting mass from $\mathbf{P}_g$ to $\mathbf{P}_r$, the WGAN loss (i.e., adversarial GAN loss) is constructed using the Kantorovich–Rubinstein duality [25] to define the minimax objective between generator G and discriminator D.

$$L_{\text{adv_GAN}} = \min_{G} \max_{D \in D} E_{\mathrm{x} \sim \mathbf{P}_r}[D(\mathrm{x})] - E_{(\widetilde{\mathrm{x}} \sim \mathbf{P}_g}[D(\widetilde{\mathrm{x}})] \tag{2}$$

where $\mathcal{D}$ is the set of 1-Lipschitz functions and $\mathbf{P}_g$ is the model distribution implicitly defined by $\widetilde{\mathbf{x}} = \mathrm{G}(\mathbf{z})$, $\mathrm{z} \sim p(\mathrm{z})$ (input Z to the generator is sampled from some simple noise distribution $p$ such as the uniform distribution). In Equation (2), $L_{\text{adv_GAN}}$ is defined to minimize $W(\mathbf{P}_r, \mathbf{P}_g)$ by using the minimization of the value function with respect to the generator parameters under an optimal discriminator.

To recognize the global consistency of the full image, we used an improved loss function with a gradient penalty term defined as follows:

$$L_{\text{global_gp}} = \lambda_1 \mathbf{E}_{\hat{\mathrm{x}} \sim \mathrm{P}_{\hat{\mathrm{x}}}} (\|\nabla_{\hat{\mathbf{x}}} D(\hat{\mathbf{x}})\|_2 - 1)^2 \tag{3}$$

where $\hat{\mathrm{x}}$ is defined as $\hat{\mathrm{x}} = (1 - t)\mathrm{x} + t\widetilde{\mathrm{x}}$; $\hat{\mathrm{x}}$ is sampled from the straight line between the points sampled from distributions $\mathbf{P}_r$ and $\mathbf{P}_g$. For image inpainting, we should restore only the missing local regions; thus, the gradient penalty should be applied only to the pixels inside the missing local regions. This can be implemented by the multiplication of gradients and the input mask, leading to the following loss function:

$$L_{\text{local_gp}} = \lambda_2 \mathbf{E}_{\hat{\mathrm{x}} \sim \mathrm{P}_{\hat{\mathrm{x}}}} (\|\nabla_{\hat{\mathbf{x}}} D(\hat{\mathrm{x}}) \otimes \mathrm{m}\|_2 - 1)^2 \tag{4}$$

where m is the input random mask with value "1" for filling the pixels within the missing regions and "0" otherwise, $\otimes$ denotes the multiplication operation, and $\hat{\mathrm{x}} = (1 - t)\mathrm{x} + t\widetilde{\mathrm{x}}$. In our implementation, $\lambda_1 = \lambda_2 = 10$ shows the best restoration performance. Note that the ocean surface region should only be used for restoration, and the gradient penalty should only be applied to the pixels inside the ocean surface region. This can be implemented by the multiplication of gradients and the land mask lm, leading to the following "sea surface loss" function:

$$L_{\text{sea_surface}} = \lambda_3 \mathbf{E}_{\hat{\mathrm{x}} \sim \mathrm{P}_{\hat{\mathrm{x}}}} (\| (\mathbf{x} - \hat{\mathbf{x}}) \otimes (1 - \mathbf{lm}) \|) \tag{5}$$

where lm is the land mask with the value of "0" for the pixels inside the ocean surface regions and "1" otherwise, which allows for the consideration of the reconstruction of the missing regions contained within the ocean surface region. Note that $\lambda_3$ is set to 2 in our implementation. As the reconstruction loss, we adopted a weighted sum of pixel-wise $\ell_1$ loss [22] named "Wasserstein-1 distance." In primal space, Wasserstein-1 distance is based on the ground distance:

$$W(\mathrm{P}_r, \mathrm{P}_g) = \inf_{\Upsilon \in \prod(\mathrm{P}_r, \mathrm{P}_g)} \mathbf{E}_{(\mathrm{x},\mathrm{y}) \sim \Upsilon} (\|\mathrm{x} - \mathrm{y}\|) \tag{6}$$

where $\prod(\mathrm{P}_r, \mathrm{P}_g)$ denotes the set of all joint distributions $\Upsilon(\mathrm{x}, \mathrm{y})$ whose marginal values are $\mathbf{P}_r$ and $\mathbf{P}_g$, respectively. Intuitively, the pixel-wise reconstruction loss $W(\mathbf{P}_r, \mathbf{P}_g)$ directly regresses holes to the current GT image. The $W(\mathbf{P}_r, \mathbf{P}_g)$ reconstruction loss measures the pixel-wise distance, which makes the training possible and makes the optimization process more stable.

The proposed loss function is defined by combining the aforementioned five loss functions:

$$L = L_{\text{adv_GAN}} + \alpha(L_{\text{global_gp}} + L_{\text{local_gp}}) + \beta L_{\text{sea_surface}} + W(P_r, P_g) \quad (7)$$

where $\alpha$ and $\beta$ are the hyperparameters that control the relative importance of the global and local losses with the penalty term and sea surface loss, respectively. The $\alpha = 10$ and $\beta = 2$ values were used during the training of the GIN model. The training was performed using the proposed loss function in Equation (7) and the stochastic gradient descent (SGD) algorithm [20].

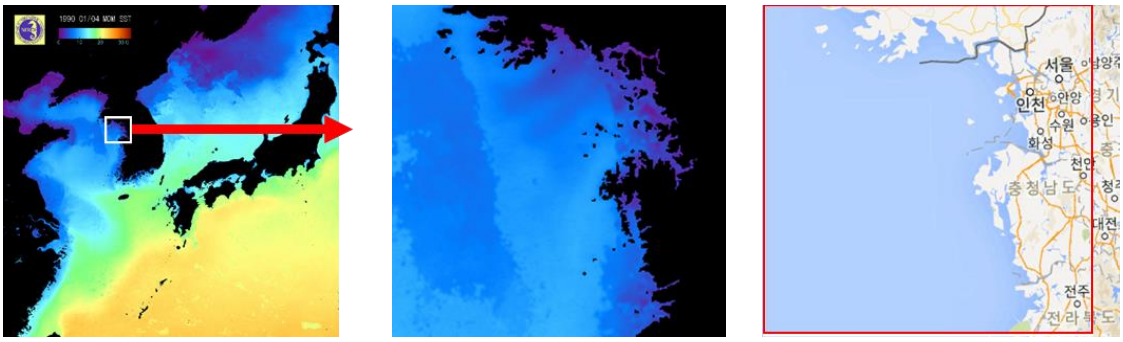
### 2.4. Implementation

The GAN architecture proposed in [17] was used as the backbone network for constructing the GIN model. This backbone network used the place2 dataset [19] with 256 × 256 resolution and was trained by generating a random mask of size ranging from 64 × 64 to 128 × 128 pixels. As suggested in [17], we strictly followed the rule of setting hyperparameters associated with the backbone network: learning rate 0.001, COARSE_L1_ALPHA 1.2, L1_LOSS_ALPHA 1.2, AE_LOSS_ALPHA 1.2, BATCH_SIZE 16 (for further details, please refer to [17]). The GIN model was implemented on TensorFlow v1.14.0, CUDNN 5.1, CUDA 8.0 and run on hardware with CPU Intel i7-6850K, memory 128GB, GPU NVIDIA GTX TITAN X (Pascal).

## 3. Experimentation

### 3.1. Datasets

The SST images used in our study were produced by the National Institute of Fisheries Science (NIFS), which provides a satellite-based sea surface temperature field since 1989 [26]. The SST image was obtained by synthesizing data received from the AVHRR (Advanced Very High Resolution Radiometer) sensors of NOAA (National Oceanic and Atmospheric Administration) with a spatial resolution of 1 km. We collected daily SST images (one image per day) from 1989 to 2018; each SST image had 2000 × 2000 pixels, in which each pixel corresponded to a 1 km square region corresponding to a 0.01 degree square in latitude and longitude. In addition, we used SST images in the area of longitudes 124.20° N to 127.30° N and latitudes 35.63° E to 38.18° E, which is surrounded by the west coast of Seoul and Jeonju, Korea. As shown in Figure 5, to make a dataset, we cropped 256 × 256 pixel regions from these SST images with 2000 × 2000 pixel resolution. In this way, we constructed a dataset consisting of 700~800 images per month (total number of training and testing images was 13,589 and 4532, respectively).




(a) (b) (c)

**Figure 5.** Example image used to train and test our deep network model. (**a**) 256 × 256 pixel region (white box) was cropped from each collected SST image of 2000 × 2000 pixels, (**b**) enlarged cropped image for better visualization, (**c**) cropped region is shown on a map.

For training, we used the images of 256 × 256 pixels with the random mask with sizes ranging from 64 × 64 to 128 × 128 pixels described in Section 2.1. We evaluated the

GIN model using satellite images (not used in the training data) with arbitrary shapes and locations of the missing sea surface regions. Since the sea surface temperature information can be significantly changed depending on the month, we constructed 12 different deep network models (i.e., GIN) (one per month). The number of training and test images used to create a network model in each month is shown in Table 1. To test the GIN model in terms of temperature restoration, matching graphs, which map the RGB color values to temperature (°C), were used (Figure 6). By using the matching graphs, the original $256 \times 256 \times 3$ RGB color image can be converted into the temperature value matrix with the size of $256 \times 256$ pixels.

**Table 1.** Number of training and test images used to construct a deep network model (i.e., GIN) in each month.

| Month | No. of Training Images | No. of Testing Images |
|---|---|---|
| January | 762 | 463 |
| February | 711 | 448 |
| March | 810 | 340 |
| April | 759 | 312 |
| May | 769 | 316 |
| June | 718 | 385 |
| July | 773 | 442 |
| August | 797 | 396 |
| September | 720 | 339 |
| October | 739 | 305 |
| November | 743 | 364 |
| December | 759 | 422 |

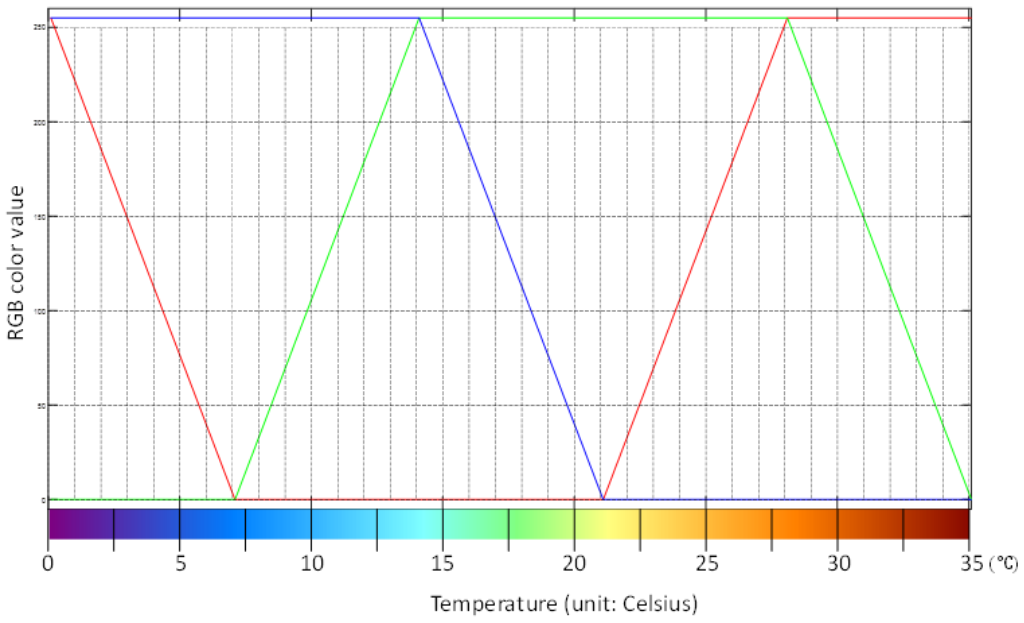



**Figure 6.** Matching graph that maps RGB color values to the corresponding temperature value to convert an original ocean satellite image into the temperature value matrix.

### 3.2. Restoration Error Metrics

To objectively evaluate the restoration quality of the GIN model, we employed commonly used restoration error metrics (or measures) such as average (AVG), standard deviation (SD), mean absolute error (MAE), and root mean square percent error (RMSPE). To obtain these metrics, we first obtained the temperature values (unit: °C) from the satellite images using an RGB matching color graph (Figure 6) and then computed the differences in temperature between restored image X and GT (actual) image Y. The following AVG metric represents the average value calculated over all temperature differences (deviations) between X and Y.

$$\mathrm{AVG} = \frac{1}{\sum_{i=1}^{256}\sum_{j=1}^{256} \mathrm{M}(i,j)} \sum_{i=1}^{256}\sum_{j=1}^{256} (\mathrm{X}(i,j) - \mathrm{Y}(i,j)) \times \mathrm{M}(i,j) \tag{8}$$

where $\mathrm{M}(i,j)$ is a binary image that has the value 1 on the pixel locations of missing regions and 0 otherwise, which allows for the measurement of the restoration error only for the missing regions in X. The associated SD is expressed as follows:

$$\mathrm{SD} = \sqrt{\frac{1}{\sum_{i=1}^{256}\sum_{j=1}^{256} \mathrm{M}(i,j)} \sum_{i=1}^{256}\sum_{j=1}^{256} (\mathrm{X}(i,j) - \mathrm{Y}(i,j))^2 \times \mathrm{M}(i,j) - \mathrm{AVG}^2} \tag{9}$$

The MAE is obtained as follows:

$$\mathrm{MAE} = \frac{1}{\sum_{i=1}^{256}\sum_{j=1}^{256} \mathrm{M}(i,j)} \sum_{i=1}^{256}\sum_{j=1}^{256} |\mathrm{X}(i,j) - \mathrm{Y}(i,j)| \times \mathrm{M}(i,j) \tag{10}$$

where MAE represents the average over the absolute error values between the X and Y images in the missing regions. We also used RMSPE as a metric. Note that by squaring the errors, we can obtain more accurate results because negative and positive errors do not cancel out each other. Moreover, the square root of the mean square error (MSE) values is robust against extreme values owing to the square root used in error computation.

$$\mathrm{RMSPE} = \sqrt{\frac{1}{\sum_{i=1}^{256}\sum_{j=1}^{256} \mathrm{M}(i,j)} \sum_{i=1}^{256}\sum_{j=1}^{256} \left(\frac{\mathrm{X}(i,j) - \mathrm{Y}(i,j)}{\mathrm{X}(i,j)}\right)^2 \times \mathrm{M}(i,j) \times 100} \tag{11}$$

## 4. Results and Discussion

To demonstrate the effectiveness of the GIN model in terms of restoration capability, we observe the MAE values computed on DaySST images using Equation (10). The mean and standard deviation of the MAE values for each month are shown in Table 2. The restoration error (i.e., temperature difference) is 0.7 °C or lower in all months and at the same time, standard deviation is considerably small. This observation indicates that the GIN model is capable of correctly restoring the missing regions of the ocean satellite images.

Table 3 shows the restoration error metrics including AVG, SD, MAE, and RMSPE values (defined in Section 3.2) for the example SST images shown in Figure 7. In Table 3, most of the AVG, MAE, and SD values obtained for each pixel of the temperature differences between ground-truth (GT) and restored images are smaller than 1 °C. This confirms the effectiveness of the proposed GIN model in terms of restoration ability. The AVG value shows the overall trend in temperature recovery. If AVG is negative, then the value of the restored temperature is lower than the GT temperature at many pixel locations, and if it is positive, then the restored temperature is higher than the GT temperature. If there

is a bias during the restoration process, the AVG value will have a constant negative or positive value for most months. Furthermore, in Table 3, the RMSPE value is expressed as a percentage of reconstruction sensitive to the mean value. Overall, it is lower in both summer and winter, while higher RMSPE values are obtained for the months of March and November. These higher RMSPE errors might be due to the fact that based on our observation, sea surface temperatures for these two months have a high-variance compared to corresponding average temperatures (i.e., the distribution of temperatures in both months is highly uneven compared to the other months); this results in larger RMSPE values based on its definition in Equation (11). In light of this fact, comparing the other error metrics used, worse restoration performance measured via RMSPE is produced for March and November.

To further validate the restoration ability of the GIN model, we computed the daily average temperature for each month on both GT and restored images. The converted temperature values only coming from the ocean surface area except for the land area within each image were averaged, yielding daily average temperature values. Using a set of daily average temperature values, we computed the "mean and median temperature" values of the GT and restored images for each month (Table 4). Note that the difference in the mean (and/or median) temperature values between the GT and restored images is smaller than 0.5 °C for all 12 months. This finding demonstrates that the GIN model works well for restoring temperature information of the missing regions of the satellite images of sea surface temperature.

Figure 7 shows the examples of DayMOM (GT) images, DaySST (test) images that need to be restored, and the restored images by using the GIN model and difference maps. In DaySST images, pixel values with white color indicate the missing regions of the SST images, and the difference maps were created by computing the difference in the pixel values between DayMOM and restored images. Except for the land area (indicated by the pixel values with black color in Figure 7), the color bar on the top left corner was used to match the RGB values with temperature for the ocean surface areas of DayMOM, DaySST, and our restored images described in Figure 7a–c, respectively. For the difference maps in Figure 7d,e, the map color bar on the top right corner was used to match the RGB values with temperature. In the RGB matching color bar for the satellite SST data, there is a little color difference between DayMOM and restored images. Especially in April or August, although data for most of the ocean surface are missing, the proposed GIN model can learn the patterns of missing areas, which allows for the restoration of the correct sea level temperature.

The proposed GIN model was developed based on the state-of-the-art structure of the GAN proposed in [17]. Compared to the GAN [17], the GIN model can be improved by incorporating a novel preprocessing step and a joint loss function described in Sections 2.1 and 2.3, respectively. Table 5 presents a comparative MAE value for each month to demonstrate the improvement in the GIN over the GAN. Based on the MAE values, the proposed GIN showed higher performance than the state-of-the-art deep inpainting network [17]. In addition, as a qualitative comparison, Figure 8 shows the restored images generated by the GIN and GAN models [17]. As shown in Figure 8c, when the GAN is used [17], the restoration quality considerably deteriorates because RGB pixel values within the land area have a great influence on temperature restoration. However, the proposed GIN is capable of reducing the influence of the land area in the process of restoration due to the following reasons: (1) the use of preprocessing where each pixel value in the land area is replaced by the average of the RGB values of the ocean surface area, and (2) the use of joint loss function (specifically, sea surface loss) that allows for the application of pixels only in the ocean surface region during the restoration process. The observations in Table 5 and Figure 8 confirm the superiority of the GIN over the state-of-the-art GAN architectures in restoring SST images.

**Table 2.** Mean and corresponding standard deviations of mean absolute error (MAE) values (computed using Equation (10)) for each month. Note that MAE denotes the average difference in temperature between restored and ground-truth images (unit: Celsius, °C)

| Month | Jan. | Feb. | Mar. | Apr. | May. | Jun. | Jul. | Aug. | Sep. | Oct. | Nov. | Dec. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Mean ± SD * | 0.57 ± 0.41 | 0.44 ± 0.31 | 0.52 ± 0.38 | 0.67 ± 0.49 | 0.62 ± 0.44 | 0.70 ± 0.46 | 0.66 ± 0.40 | 0.60 ± 0.37 | 0.56 ± 0.38 | 0.45 ± 0.28 | 0.52 ± 0.33 | 0.61 ± 0.34 |

* SD: Standard Deviation.

**Table 3.** Restoration error metrics (°C) such as AVG, SD, MAE, and RMSPE for 12 months.

| Error \ Month | Jan. | Feb. | Mar. | Apr. | May. | Jun. | Jul. | Aug. | Sep. | Oct. | Nov. | Dec. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AVG | −0.25 | 0.54 | 0.35 | 0.88 | –0.06 | –0.07 | 0.53 | 0.56 | 0.63 | 0.80 | 0.27 | 0.01 |
| SD | 0.61 | 0.50 | 0.54 | 0.81 | 0.54 | 1.01 | 0.51 | 0.57 | 0.83 | 0.82 | 0.56 | 0.71 |
| MAE | 0.69 | 0.72 | 0.64 | 1.11 | 0.61 | 0.74 | 0.70 | 0.75 | 1.08 | 1.06 | 0.73 | 0.75 |
| RMSPE | 0.09 | 20.24 | 20.25 | 11.77 | 13.24 | 5.66 | 4.06 | 4.6 | 4.23 | 4.23 | 27.28 | 8.67 |

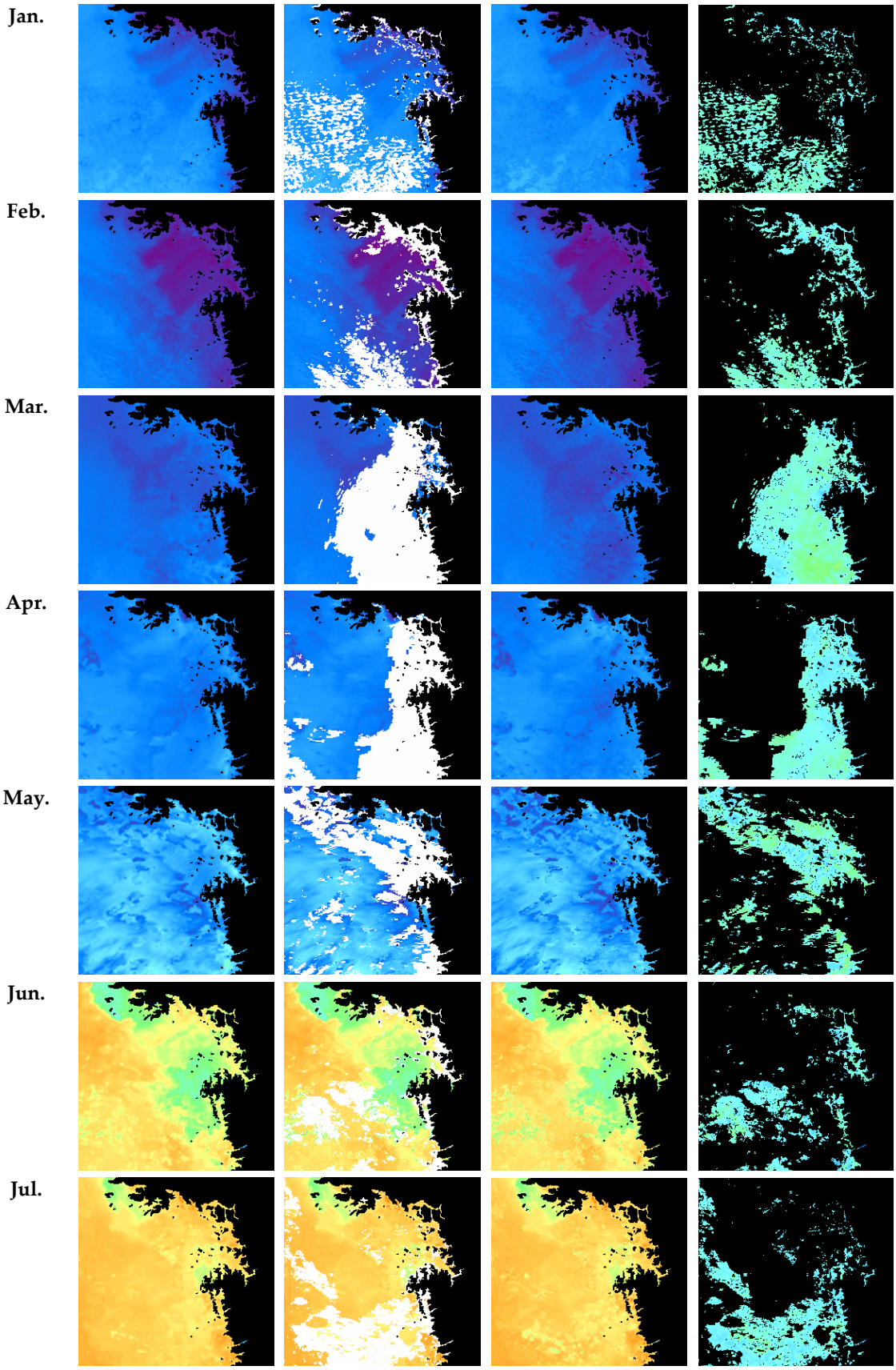



**Figure 7.** *Cont.*

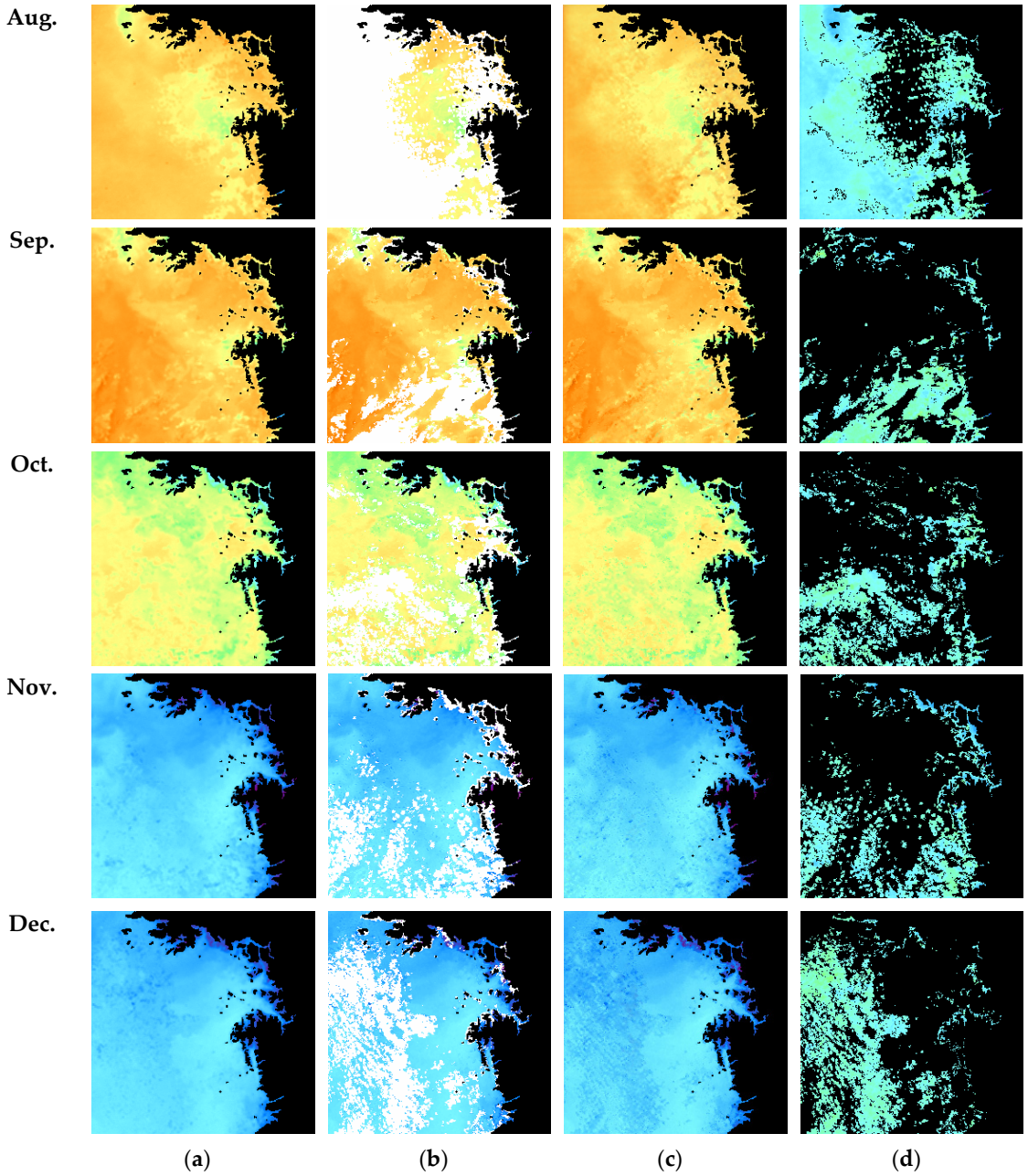



**Figure 7.** Visualized examples to demonstrate the restoration capability of the proposed GIN model. (**a**) DayMOM (ground-truth) images, (**b**) DaySST (test) images with missing (occluded) regions, (**c**) restored images created using our GIN model, (**d**) difference map images representing difference in pixel values computed between the DayMOM and restored images. White pixels indicate the occluded region by clouds and black pixels indicate land.

Figure 9 shows the plot of the daily ocean surface temperature data for a year to investigate the temperature similarity between the DayMOM (GT) and restored images. To compute the surface temperature values in Figure 8, for each DayMOM (or restored) image, we obtained $256 \times 256$ temperature values and then averaged them, yielding an "average temperature value". To this end, 4532 DaySST test images and corresponding DayMOM images were used. The differences in temperature on DayMOM images and restored images are quite marginal (at most 1.16 when all days of a year are considered (Figure 9). This observation indicates that the proposed GIN model works well for correctly restoring the sea surface temperature data regardless of the date (or season) the satellite sea surface images were taken.

**Table 4.** Mean and median temperature (°C) of ground-truth and restored images for each month.

| Month | Mean Temperature (GT) | Mean Temperature (Restored) | Median Temperature (GT) | Median Temperature (Restored) |
|---|---|---|---|---|
| Jan. | 6.5622 | 6.1424 | 6.5234 | 6.1171 |
| Feb. | 5.0895 | 4.7951 | 5.0673 | 4.7895 |
| Mar. | 5.141 | 4.7721 | 4.994 | 4.7581 |
| Apr. | 7.4216 | 6.9414 | 7.5475 | 6.964 |
| May | 12.3856 | 12.0729 | 12.3276 | 12.0007 |
| Jun. | 18.6566 | 18.2727 | 19.1618 | 18.7691 |
| Jul. | 23.2746 | 22.9665 | 23.1312 | 22.6566 |
| Aug. | 25.6947 | 25.3524 | 25.7025 | 25.3794 |
| Sep. | 23.3008 | 22.9532 | 23.3974 | 23.1224 |
| Oct. | 18.7751 | 18.5498 | 18.8778 | 18.5855 |
| Nov. | 13.9851 | 13.6752 | 14.1451 | 13.8896 |
| Dec. | 9.6363 | 9.2279 | 9.4895 | 9.1494 |

**Table 5.** MAE (°C) comparison between the proposed GIN and the state-of-the-art GAN [17] to demonstrate the effectiveness of our novel preprocessing method and joint loss function.

| Month | Proposed GIN | Baseline GAN [17] |
|---|---|---|
| Jan. | $0.57 \pm 0.41$ | $2.26 \pm 0.22$ |
| Feb. | $0.44 \pm 0.31$ | $4.24 \pm 0.63$ |
| Mar. | $0.52 \pm 0.38$ | $0.95 \pm 0.20$ |
| Apr. | $0.67 \pm 0.49$ | $1.49 \pm 0.57$ |
| May. | $0.62 \pm 0.44$ | $2.25 \pm 0.16$ |
| Jun. | $0.70 \pm 0.46$ | $2.37 \pm 0.19$ |
| Jul. | $0.66 \pm 0.40$ | $3.35 \pm 1.07$ |
| Aug. | $0.60 \pm 0.37$ | $1.25 \pm 0.18$ |
| Sep. | $0.56 \pm 0.38$ | $3.87 \pm 1.35$ |
| Oct. | $0.45 \pm 0.28$ | $4.75 \pm 0.42$ |
| Nov. | $0.52 \pm 0.33$ | $3.62 \pm 0.95$ |
| Dec. | $0.61 \pm 0.34$ | $4.13 \pm 0.68$ |

The most popular approach to reconstruct the missing regions of an image is the inverse distance weighting (IDW) interpolation method [27]. The IDW determines the values using the weights that linearly combine information from nearby areas of the missing regions [27]. The IDW restores the missing regions by averaging the neighboring eight pixels around to interpolate the center pixel. For the restoration error, the MAE of the GIN model is 0.7811, and the MAE for the IDW method is 0.7857. The *t*-test [28] used for the statistical comparison of the difference in the MAE values yields a two-tailed *p*-value of 0.048 so that the improvement in MAE achieved by the GIN is statistically significant at the level of 0.05. As shown in Figure 10, an image with $256 \times 256$ pixels is divided into smaller patches, each with $16 \times 16$ pixels, yielding a total of 256 patch images. To evaluate the finer restoration capability, we computed a patch-level MAE where the temperature difference for a particular restored patch (compared to a corresponding GT patch) was calculated and then all the temperature differences obtained using 256 patches were averaged. As a

result, patch-level MAE for the GIN and IDW methods is 0.9535 and 1.4781, respectively; the improvement in MAE is statistically significant ($p < 0.001$) at the level of 0.05 via $t$-test. This further confirms the superiority of the GIN model over one of the most popular image inpainting methods.

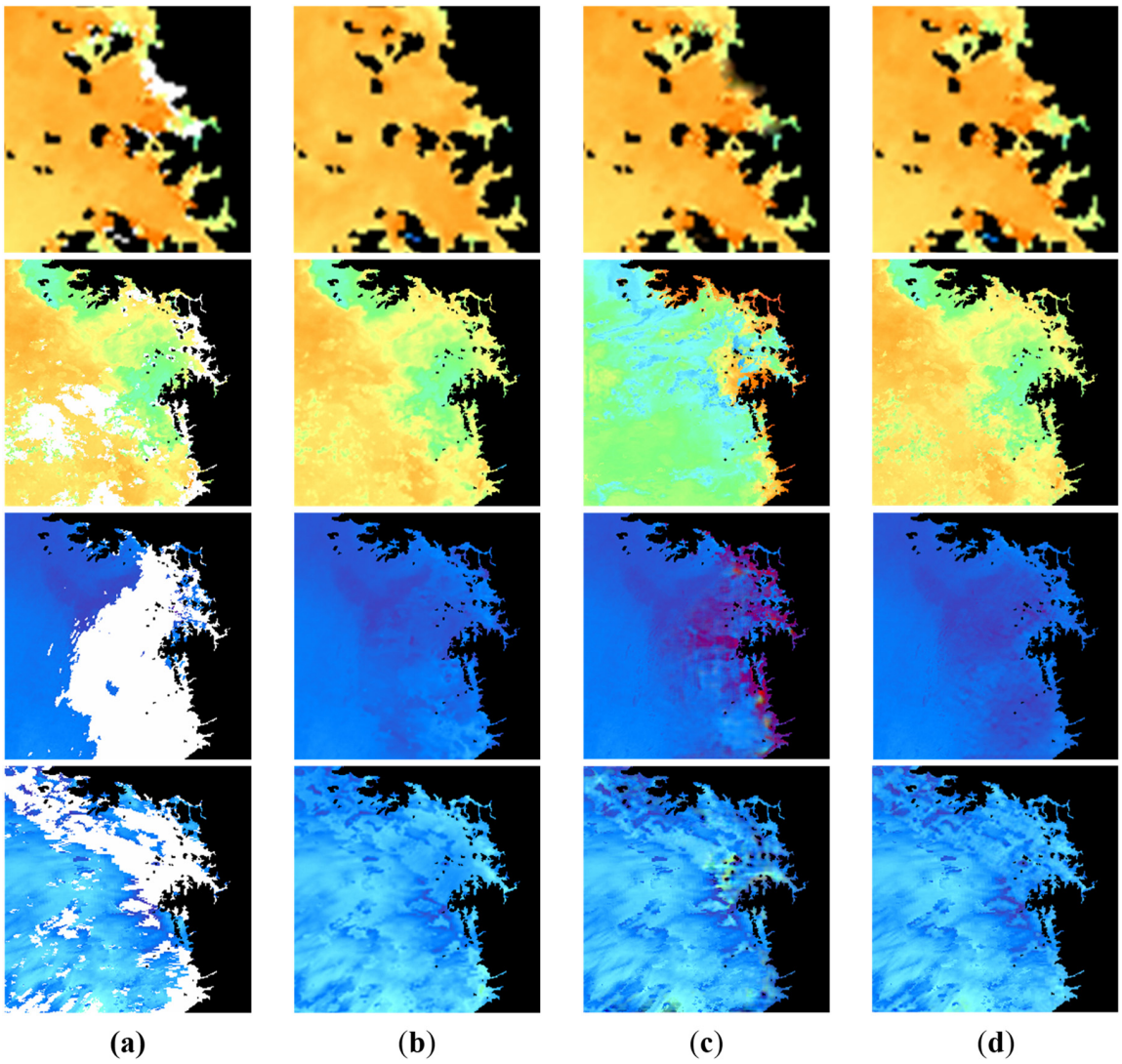



**Figure 8.** Effectiveness of the proposed GIN model compared to the state-of-the-art GAN [17] in Table 5. (**a**) DaySST image with occluded (missing) regions expressed using white color, (**b**) DayMOM image, (**c**) restored image generated using GAN model [17], and (**d**) restored image generated using our GIN model.

To further verify that the proposed GIN is able to produce a finer restoration result than IDW, specific restored patch images enclosed by the red boxes in Figure 10a–c are presented in Figure 10d. The comparison of the patch images placed at the first and second rows in Figure 10d (especially regions pointed at by the green and red arrows, respectively) shows that the GIN model can restore the patch image close to the DayMOM (GT) image while IDW is less capable of correct restoration because IDW tends to fill the missing regions with the global mean value, leading to the poor result of restoring the pixel values with a strange color mixed with gray. In addition, when the restored regions indicated by the red arrow are considered, the GIN model allows for reliably restoring the suddenly changed temperature values that are difficult to be restored via the IDW method as observed in Figure 10d.

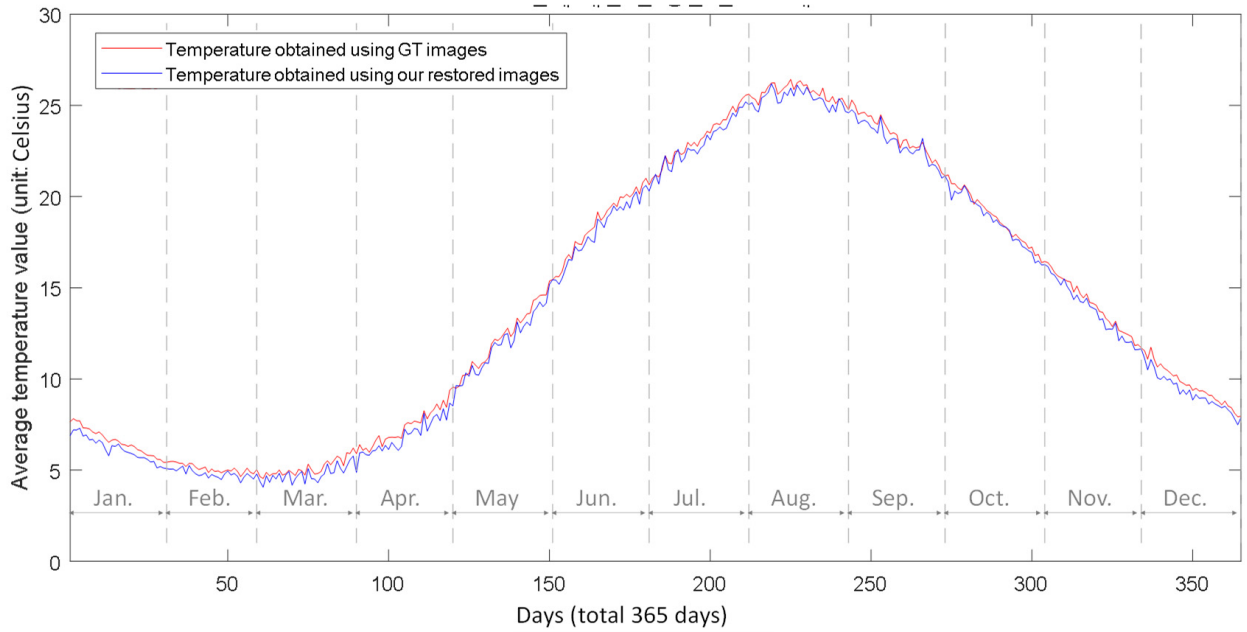



**Figure 9.** Plot of daily ocean surface temperature values within one year (365 days) for demonstrating restoration ability of the proposed GIN.

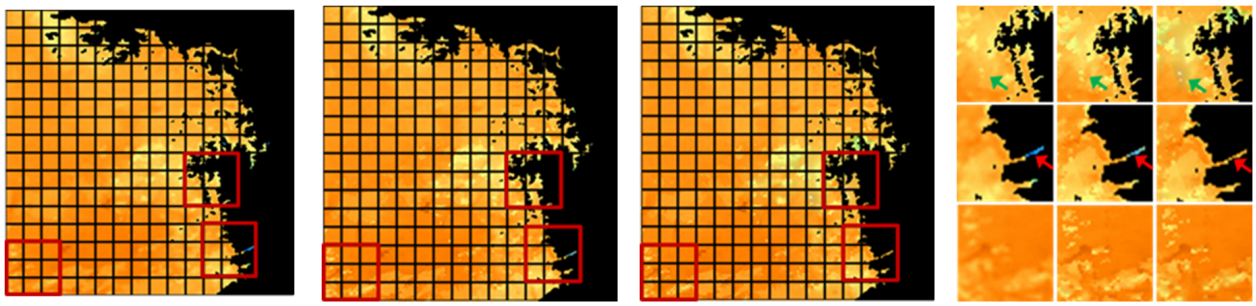



**Figure 10.** Comparison of temperature restoration between IDW [27] interpolation and GIN model. (**a**) DayMOM (ground-truth) image divided into 16 × 16 patch images, (**b**) restoration result by using GIN, (**c**) restoration result by using IDW interpolation, (**d**) DayMOM image patch (left), restored patch by using GIN (middle), restored patch by using IDW (right) for three local regions enclosed by red boxes in (**a**–**c**), respectively.

To demonstrate the restoration capability of our GIN model for SST images with large occlusion areas (patterns), partially occluded images were created by applying a randomly chosen occlusion mask to the ground-truth test images [15,16]. We evaluated the restoration accuracy with eight levels of occlusion rates, that is, 10~20%, 20~30%, 30~40%, 40~50%, 50~60%, 60~70%, 70~80%, and 80~90%. Figure 11 shows the variation in MAE with respect to the eight occlusion rates. The results in Figure 11 show the robustness to a large extent against variations in occlusion rates (i.e., the size of missing areas). For the highest occlusion rate (80~90%), our GIN model is capable of restoring SST images with quite low MAE values (at most only MAE of 0.76 °C for the month of November). This confirms the feasibility of our GIN in terms of correct restoration over challenging and real-life SST images with complex and large occlusion patterns.

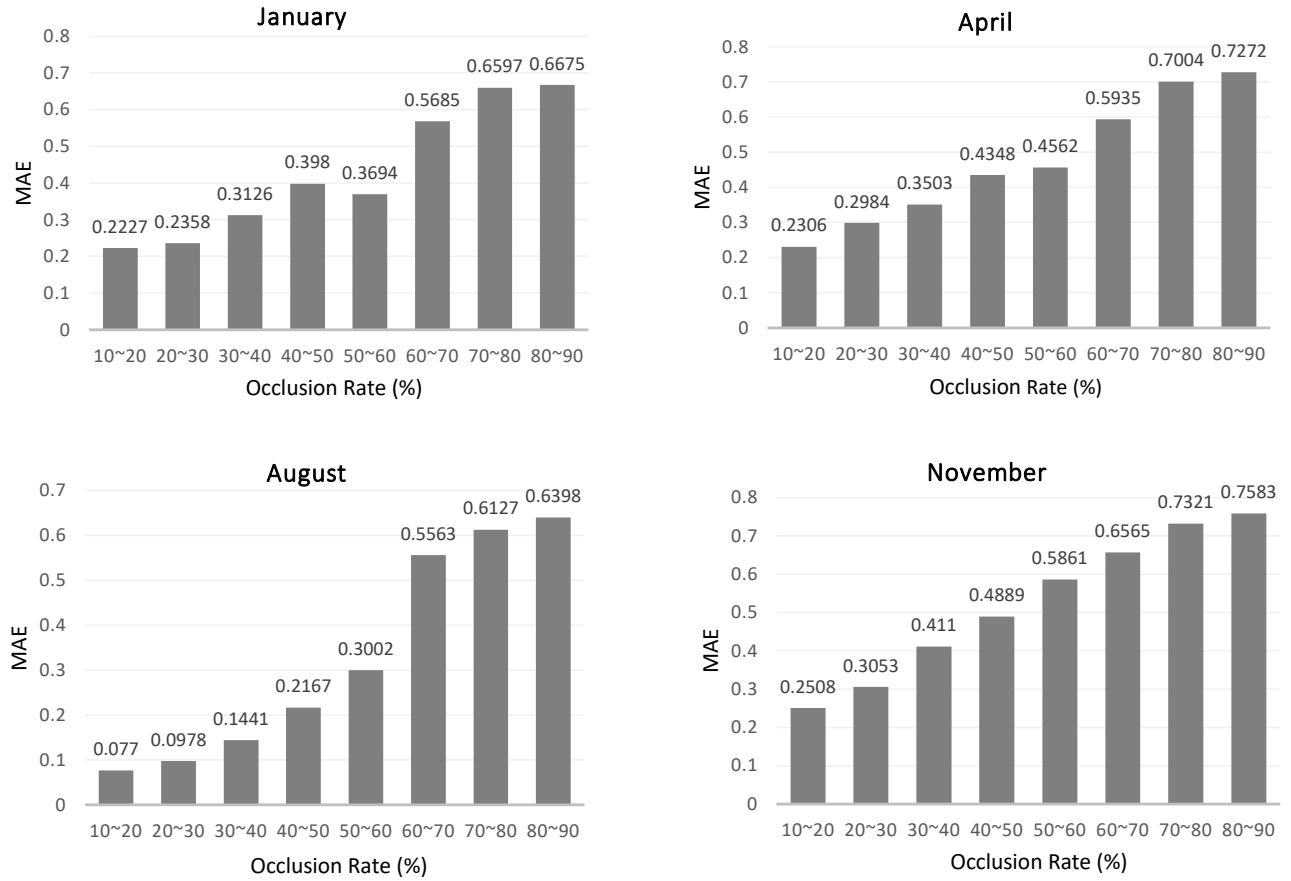



**Figure 11.** Reconstruction accuracy in MAE (°C) with respect to the eight different occlusion rates for four different months, namely January, April, August, and November.

To compare the overall run-time performance needed to restore one SST (test) image with the resolution of 256 × 256 pixels, we used implementation hardware with CPU Intel i7-6850K, memory 128 GB, GPU NVIDIA GTX TITAN X (Pascal). As a result, the GIN model requires only 4 s for completing the restoration process, while the IDW method takes more than 10 min to restore the image. The restoration process for the GIN is sufficiently fast for ocean-related applications such as weather forecasting, fisheries, and marine transportation. It takes several days to train and produce a GIN model each month using the aforementioned hardware settings, whereas the IDW requires no training because it restores images using only the nearest pixel value of the occluded region. However, the training step can be executed "off-line" because the restoration time of the GIN is much shorter than the IDW. The proposed GIN model is quite useful because it has higher restoration accuracy and much shorter processing time despite the higher computational burden in the training stage.

## 5. Conclusions

The proposed deep learning method successfully restored the satellite images of sea surface temperature which have missing values due to clouds or sea fog. The GIN method provides pragmatic advantages over currently used data composition methods. A major advantage is its immediacy that yields seamless SST grid data. The algorithm also allows an efficient training time and accuracy by using a coarse network and a fine-tuning process. The GIN is computationally efficient when using a large training dataset. In addition to the coarse network to define the overall patterns, we used a specifically designed fine-tuning procedure. In this dualized training process, we used four additional loss functions with the existing GAN loss function. The commonly used loss functions of deep learning only consider the differences in the RGB images. For the application of deep learning methods in

physical oceanography, it is necessary to develop loss functions and network architectures reflecting the physical characteristics in terms of oceanography.

We restored the spatial patterns based on the training dataset, but some data had noise owing to the use of an inappropriate composition method or insufficient data. For example, in regions where the sky is covered by clouds for over a week, the SST would be generated by using the climatology data that may have inappropriate patterns and values. During the data assimilation process that synthesizes the in situ observational data on satellite images, the SST composite data has a bias of about 0.38 °C and RMSE of 0.55 to 0.82 °C [29], which are slightly smaller than those in this study. For this purpose, modification of the training satellite dataset by advanced data assimilation schemes and synthesis of additional in situ observation data may generate semi-real-time satellite grid SST data as accurate as traditionally obtained data.

Finally, since SST images are considered sequential data, temporal trends are also helpful to restore SST images. In our future work, we will attempt to improve the accuracy by considering sequential transitions.

**Author Contributions:** Conceptualization, S.-H.K., Y.C., and J.Y.C.; methodology, S.-H.K., Y.C., and J.Y.C.; formal analysis, S.-H.K., Y.C., and J.Y.C.; investigation, S.-H.K., Y.C., and J.Y.C.; writing—original draft preparation, S.-H.K., Y.C., and J.Y.C.; writing—review and editing, J.Y.C.; visualization, S.-H.K.; supervision, J.Y.C.; funding acquisition, Y.C. and J.Y.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was funded by Korea Ministry of Oceans and Fisheries and Hankuk University of Foreign Studies Research Fund.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Acknowledgments:** This research was a part of the project titled "Technology development projects based on demand for maritime industries" of the Korea Institute of Marine Science & Technology Promotion. The authors would also like to thank the National Institute of Fisheries Science for extensive efforts in building satellite data.

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

## References

1. Wu, X.; Menzel, W.P.; Wade, G.S. Estimation of sea surface temperature using GOES-8/9 radiation measurements. *Bull. Am. Meteorol. Soc.* **1999**, *80*, 1127–1138. [CrossRef]
2. Guan, L.; Kawamura, H. Merging satellite infrared and microwave SSTs: Methodology and evaluation of the new SST. *J. Oceanogr.* **2004**, *60*, 905–912. [CrossRef]
3. Wolfe, R.E.; Roy, D.P.; Vermote, E. MODIS land data storage, gridding, and compositing methodology: Level 2 grid. *IEEE Trans. Geosci. Remote Sens.* **1998**, *36*, 1324–1338. [CrossRef]
4. Dash, P.; Ignatov, A.; Martin, M.; Donlon, C.; Brasnett, B.; Reynolds, R.W.; Grumbine, R. Group for High Resolution Sea Surface Temperature (GHRSST) analysis fields inter-comparisons—Part 2: Near real time web-based level 4 SST Quality Monitor (L4-SQUAM). *Deep Sea Res. Part II Top. Stud. Oceanogr.* **2012**, *77*, 31–43. [CrossRef]
5. Guan, L.; Kawamura, H. SST availabilities of satellite infrared and microwave measurements. *J. Oceanogr.* **2003**, *59*, 201–209. [CrossRef]
6. Wang, Y.S.; Park, S.J.; Lee, J.C. Development of Retreatment Techniques for NOAA Sea Surface Temperature Imagery. *NIFS Annu. Rep.* **2006**, 180.
7. Woo, H.J.; Park, K.; Choi, D.Y.; Byun, D.S.; Jeong, K.Y.; Lee, E.I. Comparison of multi-satellite sea surface temperatures and in-situ temperatures from Ieodo Ocean Research Station. *J. Korean Earth Sci. Soc.* **2019**, *40*, 613–623. [CrossRef]
8. Chelton, D.B.; Wentz, F.J. Global microwave satellite observations of sea surface temperature for numerical weather prediction and climate research. *Bull. Am. Meteorol. Soc.* **2005**, *86*, 1097–1116. [CrossRef]
9. Wentz, F.J.; Gentemann, C.; Smith, D.; Chelton, D. Satellite measurements of sea surface temperature through clouds. *Science* **2000**, *288*, 847–850. [CrossRef] [PubMed]
10. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Bengio, Y. Generative adversarial nets. *Adv. Neural Inform. Process. Syst.* **2014**, 2672–2680.

11. Morales, G.; Arteaga, D.; Huamán, S.G.; Telles, J.; Palomino, W. Shadow detection in high-resolution multispectral satellite imagery using generative adversarial networks. In Proceedings of the IEEE XXV International Conference on Electronics 2018, Electrical Engineering and Computing (INTERCON), Lima, Peru, 8–10 August 2018; pp. 1–4.
12. Enomoto, K.; Sakurada, K.; Wang, W.; Fukui, H.; Matsuoka, M.; Nakamura, R.; Kawaguchi, N. Filmy cloud removal on satellite imagery with multispectral conditional generative adversarial nets. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Honolulu, HI, USA, 21–26 July 2017; pp. 48–56.
13. Kadow, C.; Hall, D.M.; Ulbrich, U. Artificial intelligence reconstructs missing climate information. *Nat. Geosci.* **2020**, *13*, 408–413. [CrossRef]
14. Dong, J.; Yin, R.; Sun, X.; Li, Q.; Yang, Y.; Qin, X. Inpainting of remote sensing sst images with deep convolutional generative adversarial network. *IEEE Geosci. Remote Sens. Lett.* **2018**, *2*, 173–177. [CrossRef]
15. Hirahara, N.; Sonogashira, M.; Kasahara, H.; Iiyama, M. Denoising and Inpainting of Sea Surface Temperature Image with Adversarial Physical Model Loss. In *Asian Conference on Pattern Recognition*; Springer: Cham, Switzerland, 2019; pp. 339–352.
16. Shibata, S.; Iiyama, M.; Hashimoto, A.; Minoh, M. Restoration of sea surface temperature satellite images using a partially occluded training set. In Proceedings of the 2018 24th International Conference on Pattern Recognition (ICPR), Beijing, China, 20–24 August 2018; pp. 2771–2776.
17. Yu, J.; Lin, Z.; Yang, J.; Shen, X.; Lu, X.; Huang, T.S. Generative image inpainting with contextual attention. In Proceedings of the IEEE conference on computer vision and pattern recognition, Salt Lake City, UT, USA, 18–22 August 2018; pp. 5505–5514.
18. Li, C.T.; Siu, W.C.; Liu, Z.S.; Wang, L.W.; Lun, D.P.K. DeepGIN: Deep Generative Inpainting Network for Extreme Image Inpainting. *arXiv* **2020**, arXiv:2008.07173.
19. Zhou, B.; Lapedriza, A.; Khosla, A.; Oliva, A.; Torralba, A. Places: A 10 million image database for scene recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *40*, 1452–1464. [CrossRef] [PubMed]
20. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
21. Ioffe, S.; Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv* **2015**, arXiv:1502.03167.
22. Iizuka, S.; Simo-Serra, E.; Ishikawa, H. Globally and locally consistent image completion. *ACM Trans. Graph.* **2017**, *36*, 1–14. [CrossRef]
23. Clevert, D.A.; Unterthiner, T.; Hochreiter, S. Fast and accurate deep network learning by exponential linear units (elus). *arXiv* **2015**, arXiv:1511.07289.
24. Gulrajani, I.; Ahmed, F.; Arjovsky, M.; Dumoulin, V.; Courville, A.C. Improved training of Wasserstein GANs. *Adv. Neural Inform. Process. Syst.* **2017**, 5767–5777.
25. Rachev, S.T. Duality theorems for Kantorovich-Rubinstein and Wasserstein functionals. *Diss. Math.* **1990**, *299*, 8–30.
26. Kang, Y.Q.; Hahn, S.D.; Suh, Y.S.; Park, S.J. Variations of SST around Korea inferred from NOAA AVHRR data. *Korean Soc. Remote Sens.* **1998**, *17*, 236–241.
27. Bartier, P.M.; Keller, C.P. Multivariate interpolation to incorporate thematic surface data using inverse distance weighting (IDW). *Comput. Geosci.* **1996**, *22*, 795–799. [CrossRef]
28. McDonald, J.H. *Handbook of Biological Statistics*, 2nd ed.; Sparky House Publishing: Baltimore, MD, USA, 2009; pp. 6–59. Available online: http://www.biostathandbook.com/index.html (accessed on 4 December 2014).
29. Kilic, L.; Prigent, C.; Aires, F.; Boutin, J.; Heygster, G.; Tonboe, R.T.; Donlon, C. Expected performances of the Copernicus Imaging Microwave Radiometer (CIMR) for an all—Weather and high spatial resolution estimation of ocean and sea ice parameters. *J. Geophys. Res. Ocean.* **2018**, *123*, 7564–7580. [CrossRef]