Super-resolution on Remote Sensing Images Based on Modified VDSR Model

Yuting Yang[a,b], Kin-Man Lam[a], Junyu Dong[b], Xin Sun[b], and Muwei Jian[c]

[a]Department of Electronic and Information Engineering The Hong Kong Polytechnic University, 11 Hong Hum, Kowloon, HK 999077;
[b]Institude of Advanced Ocean Study, Ocean University of China, 238 Lao Shan, Qingdao, China 266100;
[c]School of Computer Science and Technology, Shandong University of Finance and Economics, 7366 Li Xia, Jinan, China 250000

ABSTRACT

High-resolution ocean remote sensing images are of vital importance in the research field of ocean remote sensing. However, the available ocean remote sensing images are composed of averaged data, whose resolution is lower than the instant remote sensing images. In this paper, we propose a very deep super-resolution learning model to upsample the remote sensing images. In our research, we target satellite-derived sea surface temperature (SST) images, a typical kind of ocean remote sensing image, as a specific case study on SST images. In this paper, we propose a novel model architecture based on the very deep super-resolution (VDSR) model, to further enhance its performance. Furthermore, we evaluate the peak signal-to-noise ratio (PSNR) and perceptual loss trained on the natural images and SST frames. We designed and applied our model to the China Ocean SST database, the Ocean SST database, and the Ocean-Front databases, all containing remote sensing images captured by advanced very high resolution radiometers (AVHRR). Experimental results show that our model performs better than the state-of-the-art models on SST frames.

Keywords: Deep learning, super-resolution, sea surface temperature, ocean fronts

1. INTRODUCTION

Sea surface features are an important factor in a wide range of oceanography application fields, such as air-sea interaction, severe weather prediction, and ocean engineering. It has been a hot topic for many years, bringing together many remote sensing sea surface feature images, thus giving us new insights in small-scale processes at a larger area coverage. The sea surface features include, but are not limited to, sea surface temperature, sea surface chlorophyll and sea surface height. Based on these feature images, we can further research some important marine environmental characteristics, such as vortices and ocean-fronts. However, research on sea surface features is limited by the resolution of the weekly, or monthly, averaged remote sensing images, and the cloud contamination of the instant remote sensing images. Thus, searching for ways to increase the resolution of remote sensing images and eliminate the cloud contamination are of great value for various ocean-related research works.

Ocean-fronts are one of the most well-studied ocean mesoscale characteristics. They are open streamlines that separates water masses with different physical properties, for example, water masses that contain distinct amounts of salt, heat, and chemicals. They can be transported by currents, and mixed with surrounding water masses. Usually, jets and strong currents will engender an ocean front, and the existence of an ocean-front makes it difficult for particles to cross. Thus, the water masses separated by long-living ocean-fronts can maintain their biogeochemical properties for a long time. They provide the best living environments for various kinds of fish species and form a complicated ecological system.

enkmlam@polyu.edu.hk
dongjunyu@ouc.edu.cn; phone: +86 133 2639 2372

To locate the ocean-fronts, we use one of the state-of-the-art ocean-front detection methods, named Microcanonical Multiscale Formalism (MMF), in our algorithm. It has been used to extract the ocean-fronts in sea surface temperature (SST) images in our former work.[1] Under the microcanonical framework, the critical transitions in oceanographic satellite data are analysed by calculating the Singularity Exponents (SE). This method transforms SST images into clean and simple line drawings of ocean-fronts. The method has been validated by oceanographers and has an advantage over traditional edge extraction methods in dealing with satellite oceanographic images.[2]

In this paper, we focus on increasing the resolution of remote sensing SST images. The state-of-the-art methods can be approximately divided into two categories. One embodies the traditional methods based on PCA (Principal Components Analysis) or analog schemes, and the other follows the data-driven approach, such as neural networks and machine learning methods. Ducournau A. and Fablet R.[3] is the first to use a convolutional neural network to upscale SST images, namely, super-resolution convolutional neural network (SRCNN).[4] Their experimental results showed an obvious improvement in the super-resolution performance compared to traditional methods and conventional CNNs. However, shallow CNNs have been surpassed by deep neural networks in many tasks.

Deep neural networks are a key breakthrough in the field of computer vision and pattern recognition. For the past decade, deep networks have enabled machines to classify images as accurately as humans. The most notable innovation is the concept of "residuals learning", used to effectively train a neural network. It redirects the flow of information in deep neural networks and increase the network depth to retain more information. For this reason, we consider adopting the state-of-the-art deep residual neural network for image super-resolution, namely, the very deep super-resolution network (VDSR),[5] as our baseline model. We also further improve its performance by enhancing its ability to learn features with different levels of abstraction. VDSR has 20 convolutional layers, much deeper than SRCNN which contains 3 convolutional layers only.

The enhanced deep super-resolution network (EDSR) is a more complex and deeper residual network, proposed after VDSR. Its performance on natural images, for example the Set 291, is better than that of the VDSR model. In this paper, we take EDSR model as a comparison method. Since the performance of EDSR on SST images is not better than VDSR, we decided to focus on enhancing VDSR model rather than EDSR model.

The structure of this paper is organized as follows. In Section 2, we present the structure of the modified VDSR model. Then, in Section 3, we introduce our database used for evaluation, and our experiment design and results. A detailed analysis on the experiment results is also given. In Section 4, we draw a conclusion and our future research plan.

2. THE MODIFIED VDSR MODEL

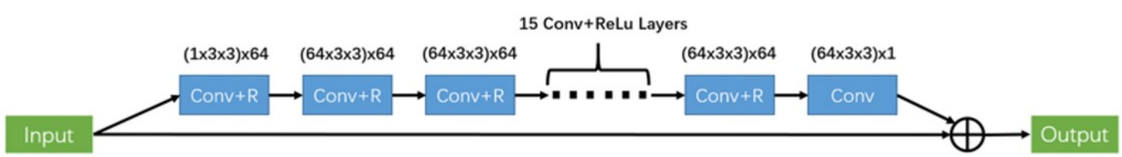

Figure 1. The VDSR model architecture. "Conv+R" means a convolutional layer followed by a ReLU unit.

In this paper, a modified VDSR model is proposed for SST image super-resolution. As shown in Fig. 1, the VDSR model consists of 20 layers. This model can generate features with different levels of abstraction. When the number of layers increases, the degree of abstraction of the learned features will also increase. Our modified VDSR model is trained and tested on data processed by a data transformation method. The data transformation method can make use of the data correlation of the remote sensing images and normalize the data to the range of 0 - 255. The data transformation method will be proposed in the near future. In this paper, we only focus on the modified VDSR model.

As is well known in the deep-learning field, the low-level features contain the most reliable information of the input images. We can obtain more abstract features, when the network gets deeper. Therefore, we propose a

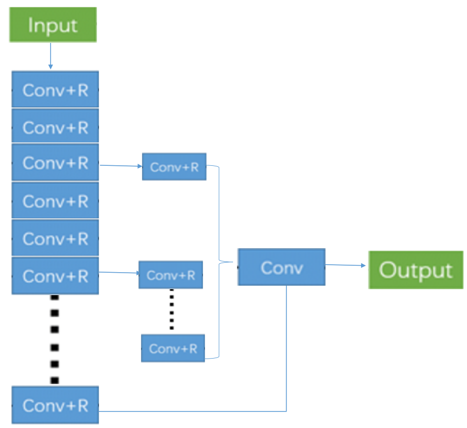

Figure 2. The proposed VDSR model architecture. Another layer of convolutional and ReLU blocks is added to the VDSR model to learn features of different levels of abstraction.

network structure, as shown in Fig. 2, which enables the VDSR model to learn features with different degrees of abstraction. Compared with the original VDSR model structure, our model adds additional convolutional layers and ReLU layers between the VDSR residual block and the final convolutional layer. The principle of the proposed network is to enhance the learning ability of VDSR by integrating the features from the different convolutional layers at different levels.

2.1 Databases

In this paper, we consider the images from the periods between January and April in 2007, 2008, 2009, and 2010, as the training data, and the images from January to April, in 2012, as the evaluation data. These images are from the China Ocean SST database.[1,6] This database contains a total of 600 daily averaged advanced very high-resolution radiometers (AVHRR) SST images, covering China's coastal waters (112.5E-135E,10N-40N), with a spatial resolution of approximately 5 km. The waters, considered in the experiments, are highly dynamic, and contain complex turbulence structures. This is because the waters are in a part of the Kuroshio Current system, which can convey a great amount of mass and power, such as heat, from low to middle latitude waters.[7] The other three databases used in our experiments are given as follows:

The ocean SST database: This database contains 1500 SST images, captured from January 2008 to December 2011. Each pixel in the images has a radius of 2.5km. It covers the ocean from longitude 155 to 180 degrees east, and latitude 30 to 50 degrees north. It is an important database for evaluating the performance of the proposed method.

The SST ocean-front database: This database consists of 1500 ocean-front images processed by the MMF method on the Ocean SST database.

The Set291: This database contains 291 natural images, and each image represents a natural scene.

2.2 Evaluation setting

We evaluate the model performance in terms of PSNR and Perceptual Loss. The pixel-wise loss between the reconstructed and the ground-truth images are averaged to evaluate the reconstruction quality in terms of PSNR. For the Perceptual Loss, it is defined as follows:

$$L_X = \frac{1}{W_{i,j}H_{i,j}} \sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} (\phi_{i,j}(I^G)_{x,y} - (\phi_{i,j}(M(I^B))_{x,y})^2 \tag{1}$$

where the model M is trained for the super-resolution task. I^B represents the input image interpolated by the bicubic interpolation and I^G is the corresponding ground-truth super-resolved image. $\phi_{i,j}$ denotes the feature maps extracted by convolutional layers, the layer depth is designated by i and j, where i is the index of the convolutional layer and j is the depth in the layer. The VGG19 network is pretrained on the ImageNet database.[8] The third convolution operation of the third convolutional layer, labled as $VGG_{3,3}$, is used to extract the perceptual feature for the input image. $W_{i,j}$ and $H_{i,j}$ represent the height and width of the feature maps. It is worth noting that those deeper layers extract higher-level or more abstract features.[9–11] The perceptual loss function pays attention to the content and the overall spatial structure information of the input images,[10] rather than their texture and exact shapes, while PSNR measures the pixelwise difference.

We evaluate the PSNR and perceptual loss of the models trained on the Set 291 and China Ocean SST databases. For the China Ocean SST database, we tested the model performance on both the coastal waters and ocean waters.

3. ANALYSIS ON THE EXPERIMENTAL RESULTS

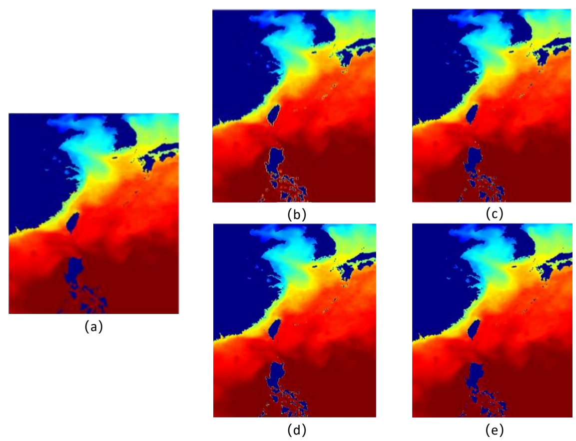

Figure 3. Performance comparison using different methods on the China Ocean SST database. a) the ground-truth SST image, b) the bicubic interpolated SST image, c) super-resolved SST image using the VDSR model trained on Set 291, d) super-resolved SST image using the modified VDSR model trained on Set 291, and e) super-resolved SST image using the modified VDSR model trained on the China Ocean SST database.

In this paper, a novel method is proposed based on VDSR for image super-resolution. Our model considers the feature maps from different layers, so it can deal with the complex situations at any location of SST images for image super-resolution. This includes the most complex and also very important coastal waters.[12] Fig. 3 compares the performance of our method with the traditional bicubic interpolation method and the VDSR network. All of these methods work well on ocean waters, while the most obvious differences lie in the coastal waters. Specifically, the Set 291 database does not have coastal waters, but the SST database does contain coastal waters. The different methods have a significant difference, in term of reconstruction performance, on coastal waters. Whenever an image contains coastal waters, the bicubic method cannot reconstruct the detailed features over the coastal waters. For comparison, when trained on the SST database, the proposed network outperforms all of the other methods, by producing the cleanest reconstruction images. Fig. 4 provides a zoom in on the coastal waters generated by the different methods. As can be observed, the proposed model trained on the SST database produces the clearest coastline, much better than the VDSR network.

Finally, Tables 1 and 2 tabulate the performance of the bicubic method, VDSR model, and the proposed model on the SST images. As shown in Table 1, our proposed model trained on Set 291 achieves a higher PSNR and Perceptual Loss than those of the VDSR model and the bicubic method on ocean waters of the China Ocean database. The upscaling factor used in the experiments is set at 3. When our model is trained on the China Ocean SST database, a slightly higher PSNR value and a much lower Perceptual Loss can be achieved, compared to the other methods. As shown in Table 2, our proposed model trained on SST images has its PSNR 3.86dB and 4.49dB higher, and its perceptual loss 47.88 and 15.18 lower, than those of the VDSR model and the bicubic method on coastal waters.

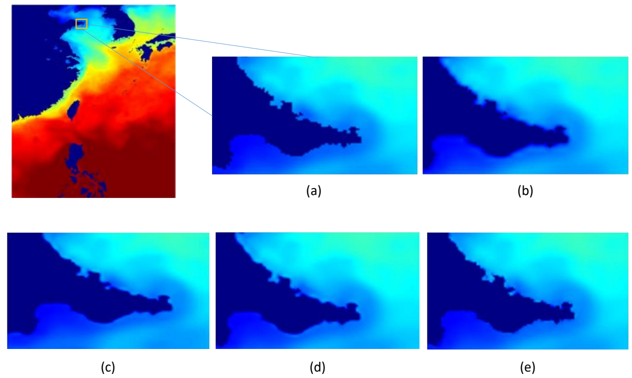

Figure 4. Zoom in on the reconstruction details of coastal waters. a) ground-truth SST image b) results from the bicubic interpolation method c) coastal waters reconstructed by VDSR model d) details reconstructed by the proposed model trained on the Set 291 e) details reconstructed by the proposed model trained on China Ocean SST database.

Table 1. The PSNR and Perceptual Loss using different methods on the China Ocean SST databases (ocean waters).

	PSNR	PERCEPTUAL LOSS
BICUBIC(291)	36.83	5.34
VDSR(291)	45.68	3.05
MODIFIED VDSR(291)	46.29	6.92
MODIFIED VDSR(SST)	46.94	3.79

Table 2. The PSNR and Perceptual Loss using different methods on the China Ocean SST database (coastal waters).

	PSNR	PERCEPTUAL LOSS
BICUBIC	17.04	21.07
VDSR	17.67	53.77
MODIFIED VDSR	21.53	5.89

Table 3. The PSNR using different methods on the Ocean SST and SST ocean-front databases.

	OCEAN SST DATASET	FRONT DATASET
BICUBIC	43.15	27.40
EDSR	31.85	34.55
VDSR	45.89	34.78
MODIFIED VDSR	46.04	34.77

Table 3 shows the experiment results using the bicubic method, EDSR model, VDSR model, and our proposed model for upsampling ocean-fronts. In this experiment, we set the downsampling scale to 3 for both the Ocean SST database and the Ocean-Front database. The upsampling factor is also set to 3. Then, these images are fed to the networks to reconstruct the upscaled SST and ocean-front images. The MMF method is then applied to the upscaled SST images to locate the ocean-fronts. By comparing the ocean-front regions based on the different methods, we can conclude that the best method for the SST ocean-front super-resolution task is the modified VDSR model. Furthermore, it can be noticed that the VDSR method can achieve a better performance than the EDSR method.

4. CONCLUSION AND FUTURE WORK

In this paper, we have proposed a modified VDSR model for remote-sensing image super-resolution, such as sea surface temperature (SST) images. Our model integrates features from different layers of the VDSR model, so low-level and high-level features are learned for reconstructing complex structures, such as the coastal waters. Experiment results have proven the superior performance achieved by our model, when compared to other state-of-the-art deep networks for image super-resolution. This work may inspire further research on new deep neural networks for super-resolution in the geophysical fields, especially for zooming in on the sea-surface features in remote sensing images. For this research, we have produced the China Ocean SST database, the Ocean SST database and the Ocean-Front database, which may promote more research on finding a better solution for the super-resolution of different types of remote-sensing ocean images.

In the future, we plan to further improve the performance of our deep model for image super-resolution and use real low-resolution and high-resolution images from the Low-resolution Ocean SST database and High-resolution Ocean SST database, which are captured by AVHRR, rather than generated by using downsampling or upsampling methods. We believe that the future work will be more challenging and meaningful.

ACKNOWLEDGMENTS

This paper was supported by The Hong Kong Polytechnic University, Hong Kong, under the project SBoS.

REFERENCES

[1] Yang, Y., Dong, J., Sun, X., Lguensat, R., Jian, M., and Wang, X., "Ocean front detection from instant remote sensing sst images," *IEEE Geoence and Remote Sensing Letters* **PP**(99), 1–5 (2017).

[2] Pont, O., Turiel, A., and Yahia, H., "Singularity analysis of digital signals through the evaluation of their unpredictable point manifold," *International Journal of Computer Mathematics* **90**(7-8), 1693–1707 (2013).

[3] Ducournau, A. and Fablet, R., "Deep learning for ocean remote sensing: an application of convolutional neural networks for super-resolution on satellite-derived sst data," in [*2016 9th IAPR Workshop on Pattern Recogniton in Remote Sensing (PRRS)*], 1–6, IEEE (2016).

[4] Dong, C., Loy, C. C., He, K., and Tang, X., "Image super-resolution using deep convolutional networks," *IEEE Trans Pattern Anal Mach Intell* **38**(2), 295–307 (2016).

[5] Kim, J., Kwon Lee, J., and Mu Lee, K., "Accurate image super-resolution using very deep convolutional networks," in [*Proceedings of the IEEE conference on computer vision and pattern recognition*], 1646–1654 (2016).

[6] Yang, Y., Dong, J., Sun, X., Lima, E., and Wang, X., "A cfcc-lstm model for sea surface temperature prediction," *IEEE Geoence and Remote Sensing Letters* **PP**(2), 1–5 (2017).

[7] Tseng, Y. H., Shen, M. L., Jan, S., Dietrich, D. E., and Chiang, C. P., "Validation of the kuroshio current system in the dual-domain pacific ocean model framework," *Progress in Oceanography* **105**(OCT.), 102–124 (2012).

[8] Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Fei-Fei, L., "Imagenet: A large-scale hierarchical image database," in [*2009 IEEE conference on computer vision and pattern recognition*], 248–255, Ieee (2009).

[9] Zeiler, M. D. and Fergus, R., "Visualizing and understanding convolutional networks," in [*European conference on computer vision*], 818–833, Springer (2014).

[10] Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., et al., "Photo-realistic single image super-resolution using a generative adversarial network," in [*Proceedings of the IEEE conference on computer vision and pattern recognition*], 4681–4690 (2017).

[11] Dong, J., Yin, R., Sun, X., Li, Q., Yang, Y., and Qin, X., "Inpainting of remote sensing sst images with deep convolutional generative adversarial network," *IEEE Geoscience and Remote Sensing Letters* **16**(2), 173–177 (2018).

[12] Holthuijsen, L. H., [*Waves in oceanic and coastal waters*], Cambridge university press (2010).