TCCL-DenseFuse: Infrared and Water Vapor Satellite Image Fusion Model Using Deep Learning

Chang-Jiang Zhang, Jia-Xu Guo, Lei-Ming Ma, Xiao-Qin Lu, and Wen-Cai Liu

***Abstract*—This article proposes an infrared and water vapor channel satellite image fusion model (TCCL-DenseFuse) based on DenseNet. The infrared channel satellite image reflects the ground and cloud top infrared radiation or the distribution of temperature, and the water vapor channel satellite image reflects the spatial distribution of water vapor in the upper atmosphere. Studies have shown that infrared brightness temperature gradient and water vapor transport are closely related to tropical cyclone (TC) generation and evolution. In order to facilitate the fusion image obtained by the proposed fusion model to have a positive effect on TC monitoring and warning, the brightness temperature gradient and multiscale structural similarity in the satellite image are used to construct loss function of the proposed TCCL-DenseFuse model. The quality of the fused images is evaluated by seven objective quantitative indicators. In order to further verify the real application value of the proposed TCCL-DenseFuse model, fused images are also used to TC center location. Experimental results show that the proposed TCCL-DenseFuse fused satellite image not only contains rich information from both infrared and water vapor channels but also improves the accuracy of TC center positioning. The comprehensive performance of the proposed fusion model has certain advantages compared with similar fusion methods and can provide a reference for typhoon prevention and disaster mitigation.**

***Index Terms*—Deep learning, infrared image, satellite image fusion, tropical cyclone (TC) center positioning, water vapor image.**

I. Introduction

SATELLITE image data commonly used in meteorological services are mainly derived from polar-orbiting and geostationary satellites. Because of their good temporal resolution and wide coverage, geostationary satellites are widely used in meteorological disaster monitoring and forecasting. Geostationary meteorological satellites include infrared, water vapor, visible light, and other channels, and the characteristics of tropical cyclone (TC) clouds are reflected in the images of different channels. The infrared channel image is obtained from the infrared radiation intensity of the target, whose temperature is transformed into shades of color in the image. The lower the temperature, the brighter the cloud is on the image. The water vapor channel image is the radiation intensity image obtained in the water vapor band of the nonatmospheric window region, and it is expressed similar to the infrared channel image, where white and black represent parts with high and low humidity, respectively [1]. The visible light channel image detects the scattering and reflection of sunlight by the surface and clouds. The albedo of the surface of the object and the altitude angle of the sun determine the brightness value of the visible light image. Because the imaging principles of satellite images differ by channel, as do the weather characteristics reflected by them, the monitoring and early warning of meteorological disasters based on satellite image data of a single channel may not be satisfactory. Using image fusion technology to fuse satellite images from different channels, satellite images containing the different weather characteristics of each channel can be obtained, and the rich weather information in the fusion images can be used to realize the automatic analysis of meteorological tasks. The infrared channel satellite image reflects the distribution of infrared radiation or temperature on the ground and cloud tops, and the water vapor channel satellite image reflects the spatial distribution of water vapor in the upper atmosphere. The two reflect atmospheric characteristics from different angles and have a large amount of TC cloud information. Studies have shown that identification of severe convective weather, cloud classification, and TC positioning and intensity determination are all related to the infrared brightness temperature gradient and the spatial distribution of water vapor [2], [3], [4]. Therefore, this article attempts to fuse the satellite images of infrared and water vapor channels and discusses the influence of the fusion images on the accuracy of TC center positioning.

The main purpose of this article is to propose a model for fusion of infrared and water vapor channel satellite images, aiming to generate a fusion image containing rich dual-channel information. This is consistent with the goal of traditional image fusion methods. Traditional image fusion methods generally focus on obtaining a fusion image with rich information but seldom pay attention to the practical application value of the fusion image in some fields after using subjective evaluation and some quantitative indicators to objectively evaluate the quality

Manuscript received 21 March 2023; revised 6 May 2023; accepted 14 May 2023. Date of publication 19 May 2023; date of current version 30 May 2023. This work was supported in part by the National Natural Science Foundation of China under Grant 42075140 and Grant 41575046 and in part by the Project of Commonwealth Techniques and Application Research of Zhejiang Province, China, under Grant LGF20D050004. *(Chang-Jiang Zhang and Jia-Xu Guo contributed equally to this work.) (Corresponding author: Chang-Jiang Zhang.)*

Chang-Jiang Zhang is with the School of Electronic and Information Engineering (School of Big Data Science), Taizhou University, Taizhou 318000, China (e-mail: zcj74922@zjnu.edu.cn).

Jia-Xu Guo is with the College of Physics and Electronic Information Engineering, Zhejiang Normal University, Jinhua 321004, China (e-mail: gjx970208@zjnu.edu.cn).

Lei-Ming Ma is with the Shanghai Central Meteorological Observatory, Shanghai 200030, China (e-mail: malm@typhoon.org.cn).

Xiao-Qin Lu is with the Shanghai Typhoon Institute of the China Meteorological Administration, Shanghai 200030, China (e-mail: 1176523687@qq.com).

Wen-Cai Liu is with the College of Physics and Electronic Information Engineering, Zhejiang Normal University, Jinhua 321004, China (e-mail: wencai_liu@zjnu.edu.cn).

Digital Object Identifier 10.1109/JSTARS.2023.3277842

of the fusion image. The development of fusion methods should also focus on their applications, such as highlighting valuable information in a specific application [5]. Therefore, this article discusses the application value of fusion image in some specific fields on the basis of producing rich information fusion image. In this article, based on the objective evaluation of the quality of fused images by using common quantitative indicators in the field of image fusion, the influence of fused images on the positioning accuracy of TC center is further discussed. As a kind of severe weather, a TC's central positioning is an important part of disaster analysis and forecasting. At present, most of the research on tropical cyclone center location (TCCL) methods based on satellite images is basically based on single infrared satellite image combined with image processing, machine learning and other technologies to study the corresponding model. The fusion image obtained by using the fusion model proposed in this article can provide source image, which contains rich information of infrared and water vapor channels, for the existing TC center positioning algorithms based on satellite images. Therefore, it has nothing to do with whether the TC center positioning method used is novel or not. The reason why the fusion image is used for TC center positioning experiment in this article is just to show that the fusion image obtained by our fusion method not only contains rich information of dual channel satellite images but can also be valuable in practical application.

Satellite image fusion methods are mainly based on multiscale analysis [6], [7], [8], compressed sensing (CS) [9], [10], [11], intensity-hue-saturation (IHS) transform [12], [13], [14], and neural networks [15], [16], [17], [18], [19], [20], as described follows.

Multiscale analysis methods include wavelet, shearlet, curvelet, contourlet, and Tetrolet transforms. The image is decomposed by multiscale analysis to obtain high- and low-frequency information, which is fused according to specific rules to obtain multiscale coefficients, which are used for inverse transformation to obtain the final fusion image. This type of fusion method requires the selection of multiscale analysis methods and determination of fusion rules.

The premise of the CS-based method is that the signal has sparseness but satellite images do not, and sparse transformation must be used. The idea is to exploit the ability of CS to reconstruct the original signal from a small number of sampled values, a process similar to multiscale analysis. The satellite image is sparsely represented, the measurement matrix is used to reduce the dimension of the signal observation, the observations are fused according to certain rules, and the signal is reconstructed to obtain a fused image. This method has four problems: sparse representation of the signal, design of the measurement matrix, determination of fusion rules, and choice of the reconstruction algorithm.

Traditional image fusion based on an IHS transform can improve the spatial resolution of the image and better preserve spatial details, but it will bring about a large spectral distortion. The idea is to transform the image from the RGB band to IHS space to obtain I, H, and S components; fuse the I component through a specific fusion rule to obtain the I' component; reverse the I', H, and S components; and transform to the RGB band to get a fused image. The key problem of this method is the rule design when fusing I components.

Shallow [15], [16], [17] and deep [18], [19], [20] neural networks have been applied to satellite image fusion. Most methods based on shallow neural networks use a pulse-coupled neural network (PCNN) [21], which usually incorporates multiscale transformation. PCNN is a simplified neural network model based on cat vision principle. PCNN does not need to learn or train, can extract effective information from complex background, and has features such as synchronous pulse release and global coupling. Methods based on deep neural networks include convolutional neural networks [18], [19] and generative adversarial networks [20], where the idea is to design an image fusion network framework, including a feature extraction network, feature reconstruction network, and loss function; feed the images into the network to start training, where network parameters are continuously adjusted according to the fusion results; and generate a fusion image that meets the requirements through much learning and training. The difficulty of these methods lies in the design and parameter adjustment of the network framework and loss function. They have advantages in extracting image features while avoiding the manual design of fusion rules, and the subjective and objective evaluation indicators of fusion images perform well. Hence, they are of interest for satellite image fusion.

The proposed satellite image fusion method TCCL-DenseFuse is more pertinent and purposeful than existing methods, and factors affecting the TC monitoring and warning are used to guide the training process of the neural network. In particular, the purpose of this article is not to design a TC center location method based on deep learning but to propose an effective dual channel satellite image fusion method based on deep learning.

To solve the above situation, we propose a satellite image fusion model based on DenseNet, TCCL-DenseFuse. Our main contributions are as follows.

1) The proposed TCCL-DenseFuse network consists of only seven convolutional layers and is basically a lightweight model, and the training is fast.
2) The image fusion model proposed in this article not only can obtain the fusion image containing rich information of infrared and water vapor channels but can also focus on the application value of the fusion image in some specific fields, such as improving the accuracy of TC center positioning.
3) Loss function is constructed by combining the brightness temperature gradient information and multiscale structural similarity in the satellite image. It is beneficial to the application in the field of TC monitoring and warning.

II. Methods

In this section, we discuss the network structure for the fusion of infrared and water vapor channel satellite images and the design of the loss function in the training process.

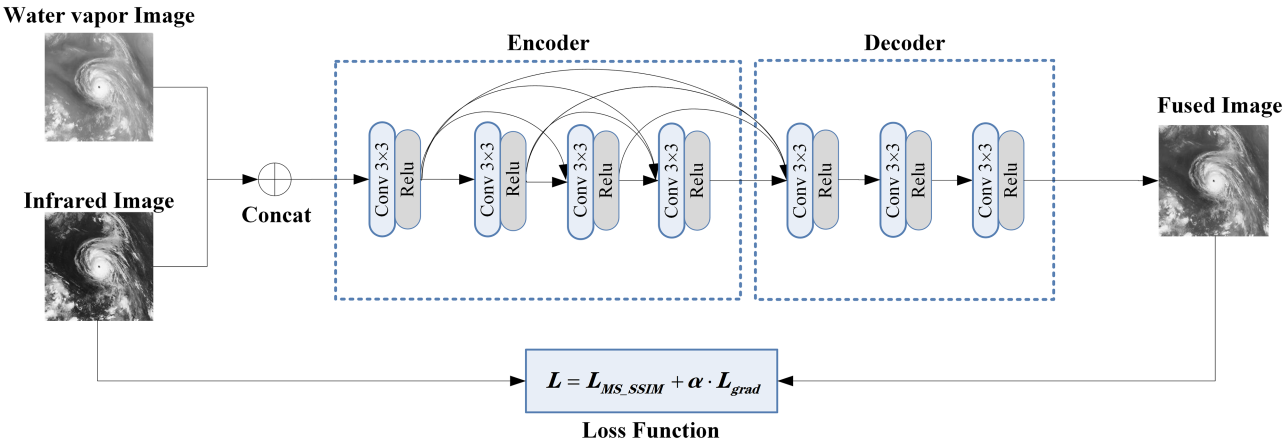


Fig. 1. TCCL-DenseFuse network structure.

A. Network Structure

Fig. 1 shows the TCCL-DenseFuse satellite image fusion model, whose framework includes an encoder and decoder. It has a simpler network structure than deep learning-based methods, enabling the completion of image fusion with only seven convolutional layers. DenseNet can ensure maximum information transmission between layers, strengthen feature transfer, and use features more effectively to retain a large amount of valuable information in the fusion image. DenseNet saves parameters and computing, the trained model is relatively small, and training is relatively fast.

The fusion process of TCCL-DenseFuse is as follows. The infrared and water vapor images are added in the channel dimension and input to the encoder, which outputs the extracted image features to the decoder, which reconstructs the fused image with the same size as the input image. The fusion image and source images are sent to the loss function. The smaller the loss, the closer the TC cloud structure of the fusion image is to the source images, and the more TC cloud structure information there is in the fusion image. The TCCL-DenseFuse network generates a fusion image that meets our goal.

1) Encoder: The encoder consists of four convolutional layers in the form of DenseNet, where the output of each layer is retained as the input to each of the following layers. Each convolutional layer has a 3 × 3 convolution kernel and a ReLU activation function. DenseNet is used to extract satellite image features, which can maximize the use of image information, ensure a high-quality fused image, and avoid gradient disappearance or explosion.

2) Decoder: The decoder consists of three convolutional layers, each composed of a convolutional layer with a 3 × 3 convolution kernel and a ReLU activation function, mainly for feature reconstruction and image fusion. The input of the decoder is the splicing of the infrared and water vapor image features output by the encoder on the channel, and the output is the fusion image. It is worth noting that there is no downsampling in the TCCL-DenseFuse network, and hence, no unnecessary loss of information.

It is worth noting that the convolution layers in this article are all composed of a 3 × 3 convolution kernel and a ReLU activation function. As we all know, TC is a complex nonlinear system. Multiple 3 × 3 convolution kernels have more layers of nonlinear functions than one larger convolution kernel, which can increase the nonlinear representation of the model. Common activation functions usually use ReLU, sigmoid, and tanh. Sigmoid and tanh are floating-point operations, while ReLU uses if–else operations, so ReLU requires little computation. Moreover, the saturation interval of ReLU is relatively small, so gradient dispersion problem is not easy to occur. At the same time, ReLU activation function has very strong nonlinear characteristics. Therefore, the small convolution kernel and ReLU activation function are beneficial to extract the structural characteristics of TC spiral rain bands and the variation characteristics of brightness temperature in infrared and water vapor channels. Therefore, the proposed TCCL-DenseFuse model is proper to extract rich nonlinear features from infrared and water vapor images.

B. Loss Functions

The loss function plays a crucial role in neural network training, and its composition directly affects the focus and quality of image fusion. The loss function in this article takes into account factors related to the TC center positioning. The loss function [see (1)] can be used to obtain high-quality fused images while improving the accuracy of TC center positioning

$$L = L_{\text{MS\_SSIM}} + \omega \cdot L_{\text{grad}} \tag{1}$$

where MS_SSIM [22] denotes the multiscale structure similarity index measurement operation, and grad denotes gradient. Because the TC center positioning method in this article is related to the brightness temperature gradient, which is closely related to the structural information of the TC, $L_{\text{MS\_SSIM}}$ is introduced into the loss function and is defined as

$$L_{\text{MS\_SSIM}} = L_{\text{ms\_ssim}}^{\text{IR1}} + L_{\text{ms\_ssim}}^{\text{IR3}} \tag{2}$$

$$L_{\text{ms\_ssim}}^{\text{IR1}} = 1 - \text{MS\_SSIM}\left(I_{\text{ir1}}, I_f\right) \tag{3}$$

$$L_{\text{ms\_ssim}}^{\text{IR3}} = 1 - \text{MS\_SSIM}\left(I_{\text{ir3}}, I_f\right). \tag{4}$$

Here, $L_{\text{ms\_ssim}}^{\text{IR1}}$ and $L_{\text{ms\_ssim}}^{\text{IR3}}$ represent the losses of structural similarity between the infrared and water vapor channel satellite image and the fusion image, respectively. Here, I_{ir1}, I_{ir3}, and I_f represent infrared, water vapor, and fused images, respectively.

This article pays great attention to preserving the gradient information of the source images, which will have a significant impact on the subsequent application of the fused image. L_{grad} is defined as

$$L_{\text{grad}} = \frac{1}{HW}\sum_{i=1}^{H}\sum_{j=1}^{W}\left(\max\left(\left|\nabla I_{\text{ir1}}\right|, \left|\nabla I_{\text{ir3}}\right|\right) - \left|\nabla I_f\right|\right). \tag{5}$$

III. Experimental Details

We introduce the experimental setup, including datasets, evaluation metrics, and training. To confirm the value of α, we perform ablation experiments, then compare the performance of our fusion model with other existing fusion models, and set the parameters according to the parameters used in these studies.

A. Experimental Setup

1) Dataset: The infrared and water vapor channel satellite images studied in this article are from China's Fengyun-2 (FY-2) and Fengyun-4 (FY-4) geostationary satellites, where the wavelengths of the infrared and water vapor satellite images are 10.2–11.5 μm and 6.3–7.6 μm, respectively. The training set consists of 3396 FY-2 satellite images from 07/11/2010 to 12/30/2014 and 3382 FY-4 satellite images from 04/22/2017 to 11/04/2017 and 12/30/2017 to 11/27/2018; and the test set consists of 340 FY-2 satellite images from 07/31/2005 to 03/25/2010 and 377 FY-4 satellite images from 11/08/2017 to 12/24/2017.

2) Software, Hardware Environments, and Model Parameter: All experiments are performed on an Nvidia GeForce RTX 2080 Ti GPU and a 2.67-GHz Intel Xeon X5650 CPU. The model is trained in a PyTorch framework; the source image is normalized to the range of 0–1; we trained the models for 50 epochs; the batch size is 16; the initial learning rate is 0.001 and adjusted exponentially; gamma is set to 0.95; and the Adam optimizer guides the model training. To make the fused image contain more gradient information, ω is set to 13.

3) Evaluation Indicators: The performance of the fusion image is evaluated both subjectively and objectively. Subjective evaluation involves the perception of the human eye, and quality will be affected by subjective factors. Objective evaluation simulates the human visual system through relevant indicators. This article adopts the fusion image positioning error (PE, km) and seven popular evaluation indicators: peak signal-to-noise ratio (PSNR, dB), correlation coefficient (CC), and feature mutual information (FMI) [23], standard deviation (SD) [24], noise in images ($N^{AB/F}$) [25], edge information retention ($Q^{AB/F}$) [26], and image structural similarity (SSIM) [27]. This article uses the method presented in [2] to locate the TC center. Note that the larger the other indicators except for $N^{AB/F}$ and PE, the better the quality of the fused image. The smaller the $N^{AB/F}$, the less noise interference there is in the fused image, and the better the fused image. The smaller the PE, the more accurate the TC center positioning using the fused image.

PE is the TC center positioning error value, and it is defined as

$$\text{PE} = \sqrt{\left(x_f - x_r\right)^2 + \left(y_f - y_r\right)^2} \cdot \text{SR} \tag{6}$$

where x_f and y_f denote estimated TC center coordinates, x_r and y_r denote ground truth TC center coordinates, and SR denotes the spatial resolution of geostationary satellite. Spatial resolution of FY-2 and FY-4 is 5 km and 4 km, respectively. A lower PE indicates that TC center positioning using fusion image is more accurate. In this article, ground truth TC center coordinates are derived from the best track data, which is published by China Meteorological Administration (CMA).

PSNR is the ratio of peak value power to noise power in the fused image and, thus, reflects the distortion during the fusion process. It is defined as

$$\text{PSNR} = 10\log_{10}\frac{r^2}{\text{MSE}} \tag{7}$$

$$\text{MSE} = \left(\text{MSE}_{AF} + \text{MSE}_{BF}\right)/2 \tag{8}$$

$$\text{MSE}_{XF} = \frac{1}{MN}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\left(X\left(i,j\right) - F\left(i,j\right)\right)^2 \tag{9}$$

where r is the peak value of the fused image and is set as 256 in this article. MSE is the mean square error, which measures the dissimilarity between the fused image and source images. A larger PSNR indicates that more source image information is contained in the fused image.

CC is the degree of correlation between the fused and source images

$$\text{CC} = \frac{r_{AF} + r_{BF}}{2} \tag{10}$$

$$r_{XF} = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}\left(X\left(i,j\right) - \bar{X}\right)\left(F\left(i,j\right) - \mu\right)}{\sqrt{\sum_{i=1}^{M}\sum_{j=1}^{N}\left(X\left(i,j\right) - \bar{X}\right)^2\left(\sum_{i=1}^{M}\sum_{j=1}^{N}\left(F\left(i,j\right) - \mu\right)^2\right)}} \tag{11}$$

where $\bar{X}$ is the average brightness value of the source image X, and μ is the average brightness value of the fused image F. A larger CC indicates that the fused image is more similar to the source image and has a better fusion effect.

FMI measures the amount of feature information transferred from the source image to the fused image and is defined as

$$\text{FMI} = \text{MI}_{A,F} + \text{MI}_{B,F} \tag{12}$$

$$\text{MI}_{X,F} = \sum_{x,f} p_{X,F}\left(x,f\right)\log\frac{p_{X,F}\left(x,f\right)}{p_X\left(x\right)p_F\left(f\right)} \tag{13}$$

where $p_{X,F}(x,f)$ denotes the joint histogram of the source image X and the fused image F. $p_X(x)$ and $p_F(f)$ denote the

marginal histograms of X and F, respectively. A larger FMI indicates that there is abundant information transferred from source images to the fused image.

SD reflects the intensity distribution and contrast of the fused image. It is defined as

$$\mathrm{SD} = \sqrt{\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}(x_{i,j}-\mu)^2} \tag{14}$$

where μ is the average brightness value of the image x. Human attention is more likely to be drawn to areas with high contrast. A larger SD indicates that the visual effect of the fused image is good.

A fusion artifact introduced by a fusion process into the fused image can lead to a benign object being classified as a threat or a valid target [25]. To quantify the fusion artifacts, we present the fusion artifact measurement indicator $N^{AB/F}$, which is defined as

$$N^{AB/F} = \frac{\sum_{\forall i}\sum_{\forall j} AM_{i,j}\left[\left(1-Q_{i,j}^{AF}\right)w_{i,j}^{A}+\left(1-Q_{i,j}^{BF}\right)w_{i,j}^{B}\right]}{\sum_{\forall i}\sum_{\forall j}\left(w_{i,j}^{A}+w_{i,j}^{B}\right)} \tag{15}$$

where $AM_{i,j} = \begin{cases} 1, & g_{i,j}^{F} > g_{i,j}^{A}, \text{ and } g_{i,j}^{F} > g_{i,j}^{B} \\ 0, & \text{otherwise} \end{cases}$ indicates the locations of fusion artifacts where fused gradients are stronger than input. $g_{i,j}^{A}$, $g_{i,j}^{B}$, and $g_{i,j}^{F}$ are the edge strength of source image A, B, and fused images F, respectively. $Q_{i,j}^{AF}$ and $Q_{i,j}^{BF}$ are the gradient information preservation estimates of source images A and B, respectively. $w_{i,j}^{A}$ and $w_{i,j}^{B}$ are the perceptual weights of source images A and B, respectively. A lower $N^{AB/F}$ indicates that the fused image contains less noise.

$Q^{AB/F}$ measures the amount of edge information that is transferred from source images to the fused image and is defined as

$$Q^{AB/F} = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N} Q^{AF}(i,j)\,w^{A}(i,j) + Q^{BF}(i,j)\,w^{B}(i,j)}{\sum_{i=1}^{M}\sum_{j=1}^{N}\left(w^{A}(i,j)+w^{B}(i,j)\right)} \tag{16}$$

$$Q^{XF}(i,j) = Q_{g}^{XF}(i,j)\,Q_{a}^{XF}(i,j) \tag{17}$$

where w^{X} denotes the weight, which expresses the importance of each source image to the fused image. Q_{g}^{XF} and Q_{a}^{XF} denote the edge strength and orientation values, respectively. A large $Q^{AB/F}$ indicates that considerable edge information is transferred to the fused image.

SSIM denotes the structural similarity between source images and fused image and is defined as

$$\mathrm{SSIM} = \mathrm{SSIM}_{A,F} + \mathrm{SSIM}_{B,F} \tag{18}$$

$$\mathrm{SSIM}_{X,F} = \sum_{x,\,f}\frac{2\mu_x\mu_f + C_1}{\mu_x^2+\mu_f^2+C_1}\cdot\frac{2\sigma_x\sigma_f + C_2}{\sigma_x^2+\sigma_f^2+C_2}\cdot\frac{\sigma_{xf}+C_3}{\sigma_x\sigma_f+C_3} \tag{19}$$

where μ_x and μ_f denote the mean values of the source and fused images, respectively. σ_x and σ_f denote the SD. σ_{xf} denotes the covariance of the source and fused images. C_1, C_2, and C_3 are the parameters used to make the algorithm stable. A larger SSIM indicates that the fused image is more similar to the source images.

B. Ablation Experiments

To determine the value of θ in the loss function, we conduct the following experiments: $\theta = 1$, 6, 12, 18, and 24. According to the TC standard (GB/T19201-2006) issued by the National Meteorological Center of China, TC is divided into tropical depressions (TD: 10.8–17.1 m/s), tropical storms (TS: 17.2–24.4 m/s), severe tropical storms (STS: 24.5–32.6 m/s), typhoons (TY: 32.7–41.4 m/s), severe typhoons (STY: 41.5–50.9 m/s), and super typhoons (SuperTY: $\geq$51.0 m/s), so we randomly select a group of infrared and water vapor satellite images in each intensity level for image fusion using our model. The fusion results are shown in Fig. 2.

From Fig. 2, it can be seen that the fused image when $\theta = 1$ is similar to the water vapor image; when $\theta = 6$, the contrast of the fused image is significantly improved. When $\theta = 12,\ 18,$ and 24, the fused image is closer to $\theta = 6$, and the images are relatively similar overall. Visually, when $\theta = 1$, the contrast of the image is low, and the fusion effect is poor. When α takes other values, the contrast is better, and the fusion effect seems to be better. We also quantitatively analyze 1113 fused images (test set) with different θ values, and the measurement results of the evaluation indicators and PE are given in Table I.

The best metric is highlighted in red, the second-best metric is indicated in blue italics, and third-best metric is underlined. All subsequent tables use the same representation for the best, second-best, and third-best metrics. It can be seen from Table I that on the seven common indexes, each θ value has one or two best value. Among the PE measurement results, $\theta = 12$ is the best, and $\theta = 6$ is the second best. This article focuses on generating the fusion image containing rich information can be expected to improve the positioning accuracy of TC center. The results show that $\theta = 12$ is more consistent with these requirements; it has the best PE value, and other indicators are not much different from the best indicators, which are close, so we choose $\theta = 12$ as the final parameter of loss function in our model (TCCL-DenseFuse).

C. Comparative Experiments

The fusion results of our TCCL-DenseFuse model are compared with those of SDNet [28], RFN-Nest [29], SEDRFuse [30], DIF-Net [31], and IFCNN [32]. To verify the effectiveness of the TCCL-DenseFuse model, we first analyze the entire test set, and then classify the test set according to the categories of TC and draw conclusions through in-depth discussion of each category. Finally, to further verify the performance of the proposed TCCL-DenseFuse model, two TC cases from the test set are analyzed and discussed.

1) Discussion of the Entire Test Set: We randomly select a group of infrared and water vapor satellite images in each TC intensity level for fusion. The fusion results are shown in Figs. 3–8.

From Figs. 3 to 8, it can be seen that the TC main body of the fusion results of RFN-Nest, DIF-Net, and IFCNN are bright and

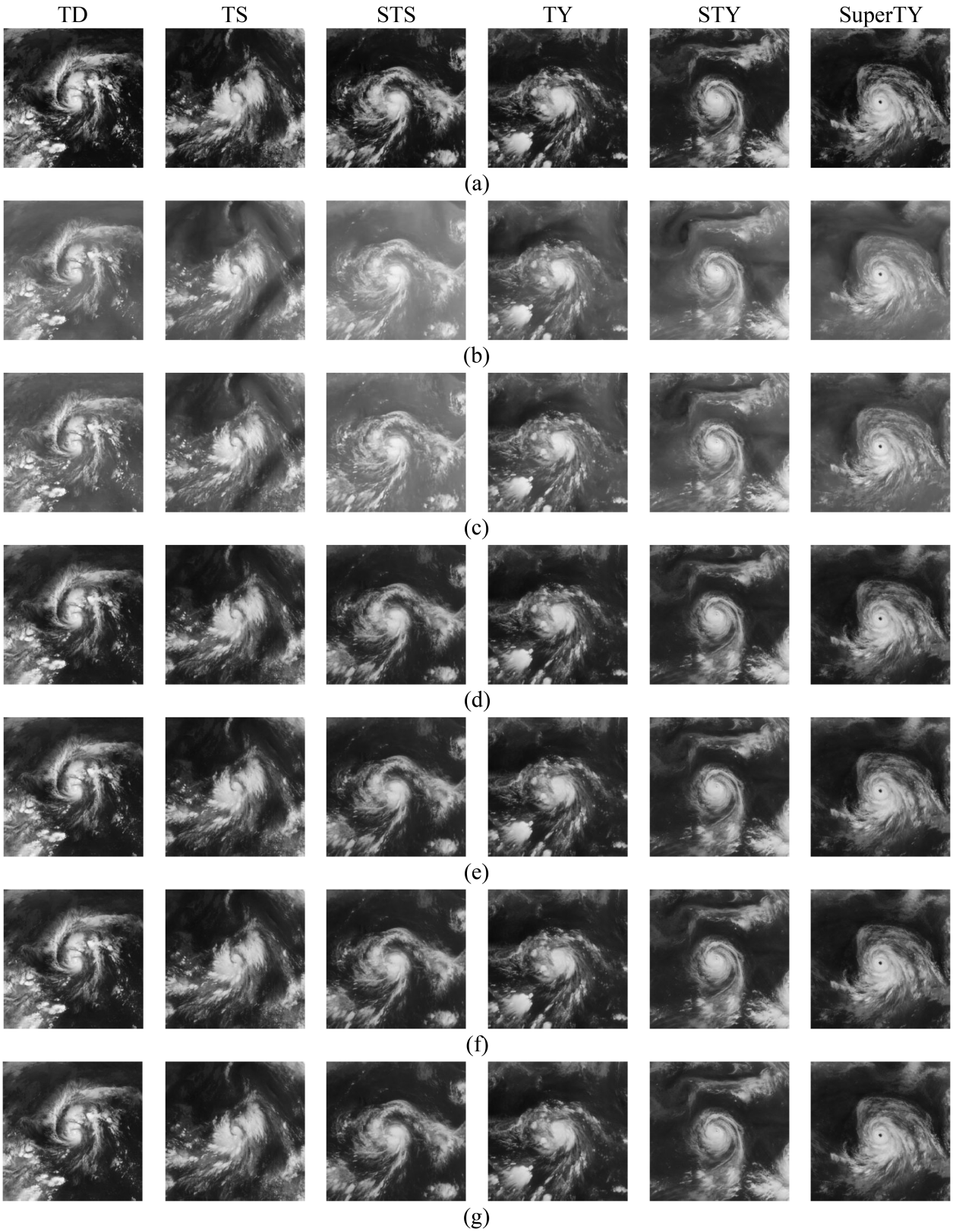


Fig. 2. Fusion results of each TC intensity level. (a) Infrared image. (b) Water vapor image. (c) Fusion image ($\theta = 1$). (d) Fusion image ($\theta = 6$). (e) Fusion image ($\theta = 12$). (f) Fusion image ($\theta = 18$). (g) Fusion image ($\theta = 24$).

TABLE I
EVALUATION INDEX MEASUREMENT RESULTS OF ABLATION EXPERIMENTS

Model	PSNR	CC	FMI	SD	$N^{AB/F}$	$Q^{AB/F}$	SSIM	PE
Infrared image	—	—	—	—	—	—	—	262.8289
Water vapor image	—	—	—	—	—	—	—	242.1628
Our model (θ = 1)	23.1439	0.9618	0.8902	40.8558	0.0021	0.6361	0.8961	246.4406
Our model (θ = 6)	23.3994	0.9640	0.8864	46.0002	0.0030	*0.6569*	0.8974	239.6160
Our model (θ = 12)	23.4018	0.9637	0.8865	*46.0154*	0.0030	0.6573	0.8975	233.7738
Our model (θ = 18)	*23.4309*	0.9638	0.8868	46.0303	0.0030	0.6555	*0.8977*	*237.8548*
Our model (θ = 24)	23.5250	*0.9639*	*0.8869*	45.5300	*0.0029*	0.6558	0.8983	244.4478

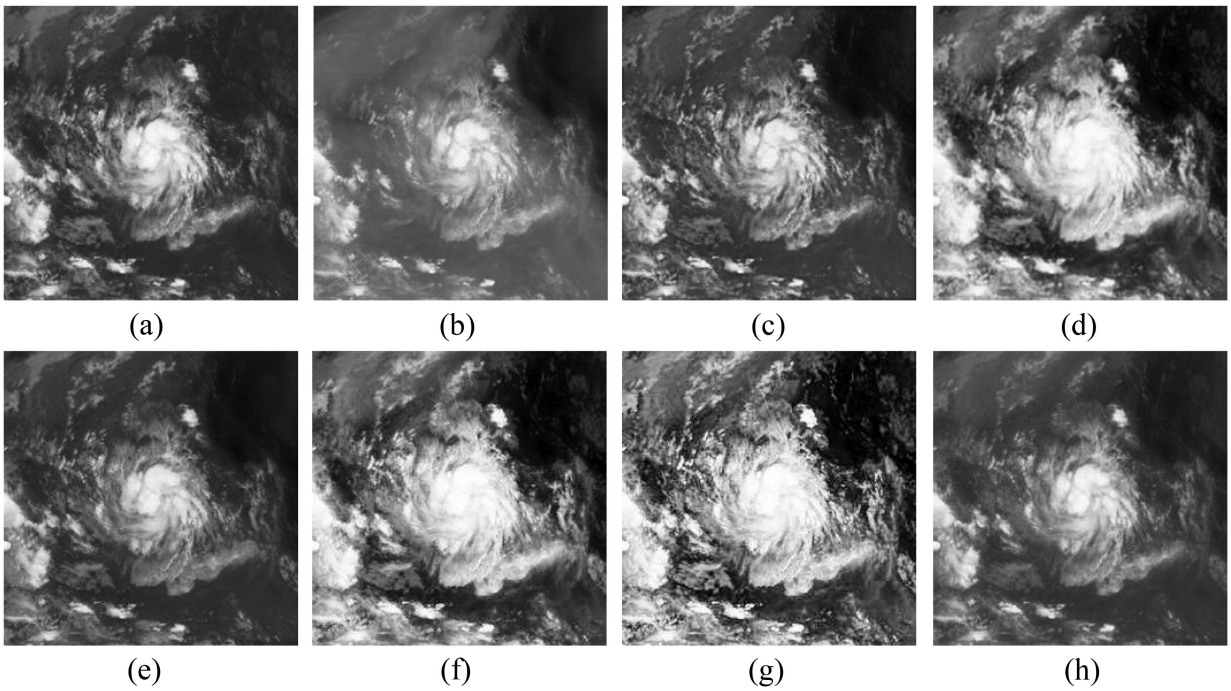


Fig. 3. Fusion results of FY-4 satellite TD images at 21:00 on November 12, 2017. (a) Infrared image. (b) Water vapor image. (c) SDNet. (d) RFN-Nest. (e) SEDRFuse. (f) DIF-Net. (g) IFCNN. (h) TCCL-DenseFuse.

tend toward 255, while those of SDNet and SEDRFuse are similar. Compared with other five image fusion models, the global contrast of the fused image using our TCCL-DenseFuse model is not very good. From the background, gray-level values of fused images using RFN-Nest, DIF-Net, and IFCNN tend to be 0, followed by other models, and the fused image using TCCL-DenseFuse tends to be an infrared image. On the whole, the fused images of RFN-Nest, DIF-Net, and IFCNN have better contrast but a little too much contrast to enhance the spiral rainband part of TC, which has some lost details or gradient information of brightness temperature. TCCL-DenseFuse performs similarly to several other models, but on individual fused images, it is inferior. This is because the human eye is good at capturing the situation where the difference between the foreground and background is large. To avoid this interference and evaluate the image quality more objectively, we quantitatively evaluate the fusion results using different fusion models by eight evaluation indicators. The eight metrics curves of the 1113 fused images (TC test images) obtained by the existing five similar image fusion models and our TCCL-DenseFuse model are shown in Fig. 9.

In the upper left or upper right boxes of Fig. 9(a)–(h), the decimal after each model name represents the average value of each model under different evaluation metrics on TC test images. In Fig. 9, the abscissa of each graph adopts the interval of the numbers in parentheses below the respective graph, and

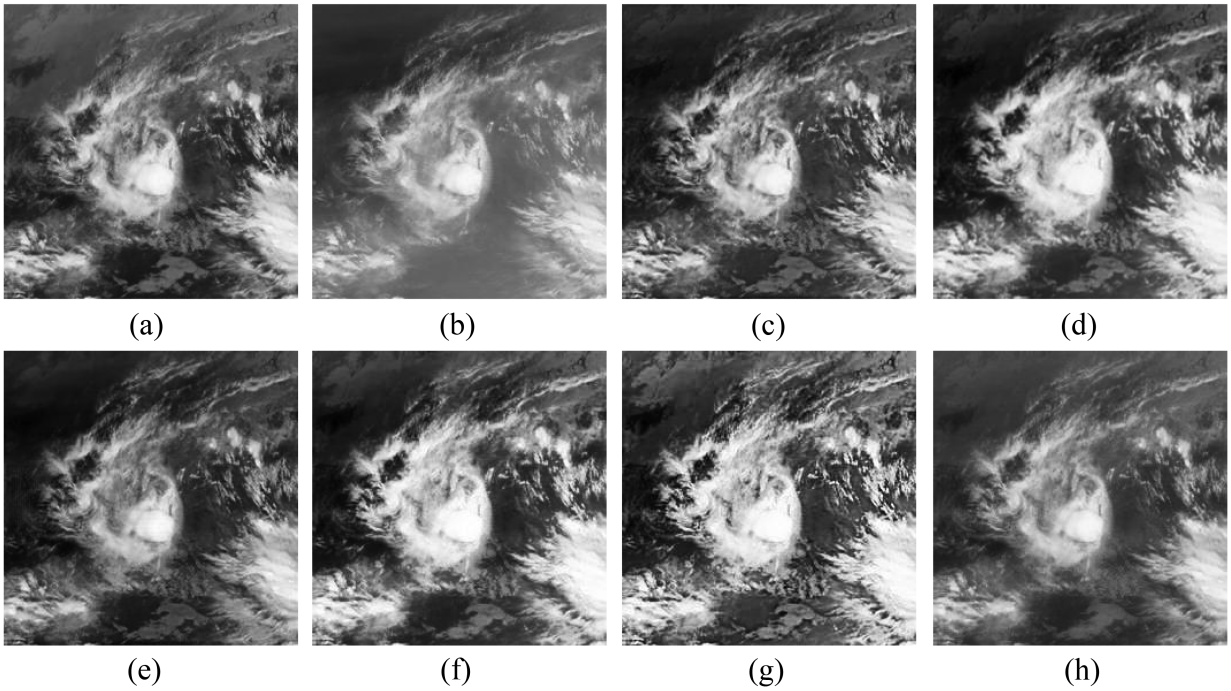


Fig. 4. Fusion results of FY-4 satellite TS images at 23:00 on November 11, 2017. (a) Infrared image. (b) Water vapor image. (c) SDNet. (d) RFN-Nest. (e) SEDRFuse. (f) DIF-Net. (g) IFCNN. (h) TCCL-DenseFuse.

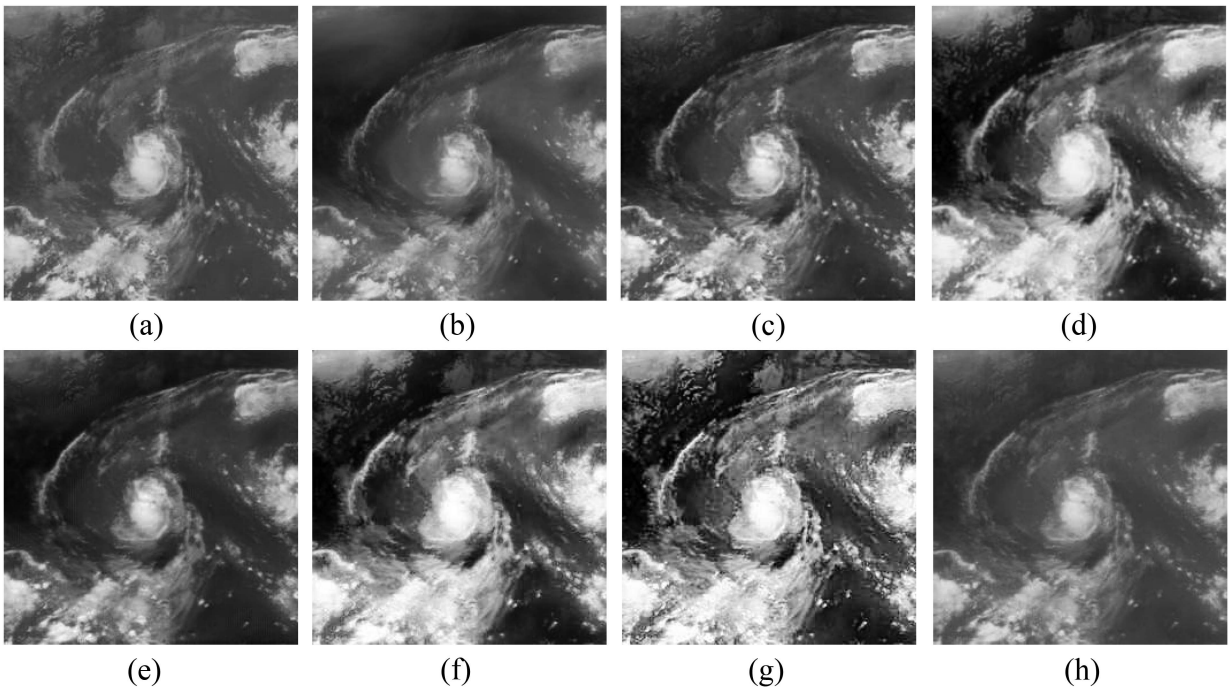


Fig. 5. Fusion results of FY-2 satellite STS images at 9:00 on July 25, 2006. (a) Infrared image. (b) Water vapor image. (c) SDNet. (d) RFN-Nest. (e) SEDRFuse. (f) DIF-Net. (g) IFCNN. (h) TCCL-DenseFuse.

the ordinate is the number of TC images corresponding to each interval range. This means that the higher the right side and the lower the left side of PSNR, CC, FMI, SD, $Q^{AB/F}$, and SSIM curves are, the better the model effect is. The higher the left side of the $N^{AB/F}$ and PE curves, the lower the right side, the better the model effect is.

From Fig. 9, it can be seen that our TCCL-DenseFuse model is significantly ahead of other models on PSNR, $N^{AB/F}$, SSIM, and

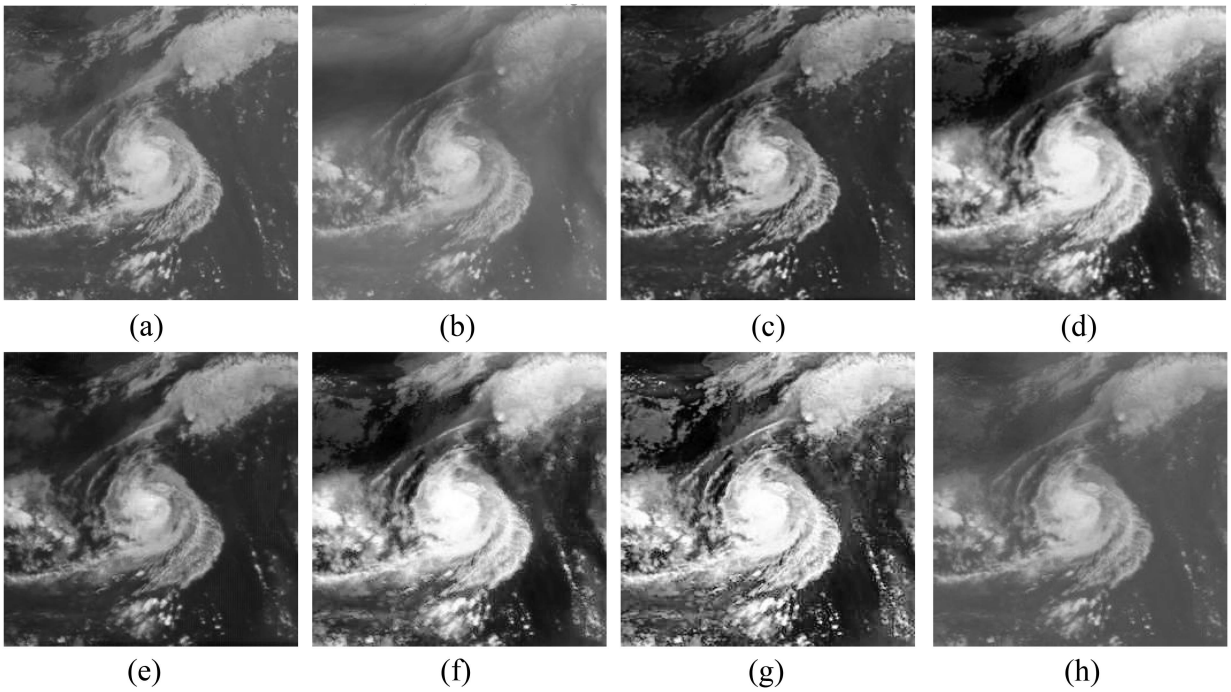


Fig. 6. Fusion results of FY-2 satellite TY images at 12:00 on May 4, 2009. (a) Infrared image. (b) Water vapor image. (c) SDNet. (d) RFN-Nest. (e) SEDRFuse. (f) DIF-Net. (g) IFCNN. (h) TCCL-DenseFuse.

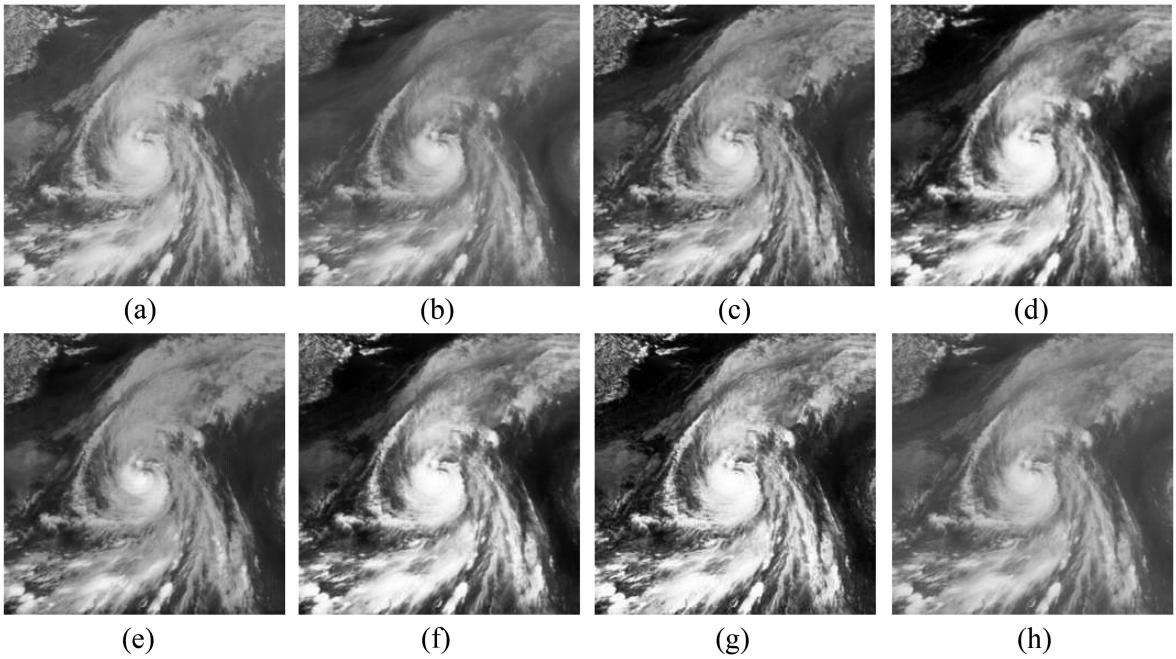


Fig. 7. Fusion results of FY-2 satellite STY images at 12:00 on May 16, 2006. (a) Infrared image. (b) Water vapor image. (c) SDNet. (d) RFN-Nest. (e) SEDRFuse. (f) DIF-Net. (g) IFCNN. (h) TCCL-DenseFuse.

PE, slightly ahead on $Q^{AB/F}$, slightly inferior to SDNet on FMI and CC, and the worst on SD. The fluctuations of each model can be seen well from Fig. 9. Fig. 9(a)–(g) are commonly used indicators for objective evaluation of image fusion at present. In addition, in order to further verify the value of the proposed fusion model in practical application, the obtained fusion image is used to locate the TC center to test its influence on the positioning accuracy of the TC center [see Fig. 9(h)]. From Fig. 9(h), the higher the left side of the PE curve and the lower the right side, the better the TC center localization effect of the model.

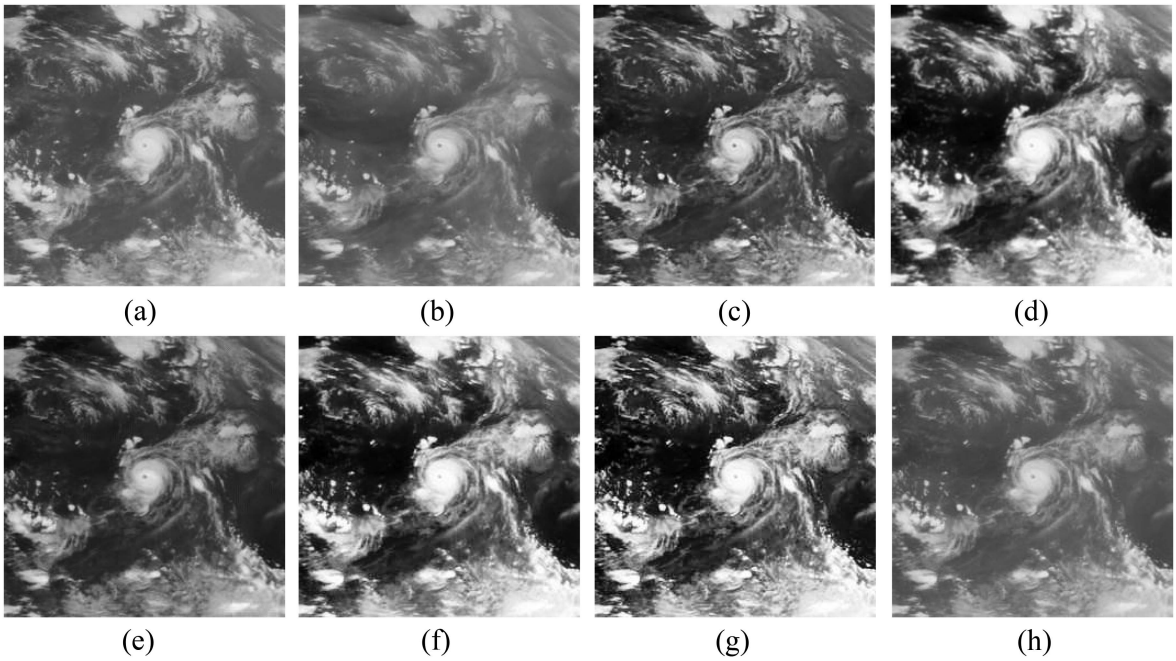


Fig. 8. Fusion results of FY-2 satellite superTY images at 0:00 on September 27, 2008. (a) Infrared image. (b) Water vapor image. (c) SDNet. (d) RFN-Nest. (e) SEDRFuse. (f) DIF-Net. (g) IFCNN. (h) TCCL-DenseFuse.

TABLE II
AVERAGE EVALUATION METRICS OF 1113 TC FUSED IMAGES

Method	PSNR	CC	FMI	SD	$N^{AB/F}$	$Q^{AB/F}$	SSIM	PE
Infrared image	—	—	—	—	—	—	—	262.8289
Water vapor image	—	—	—	—	—	—	—	*242.1628*
SDNet	18.8605	0.9636	0.8919	55.3836	0.0923	0.6434	0.8155	248.3938
RFN-Nest	17.0606	0.9696	0.8847	57.4125	*0.0243*	0.6529	0.8090	243.6182
SEDRFuse	17.4384	0.9565	0.8767	57.1056	0.0482	0.6453	0.7974	251.5469
DIF-Net	*19.1447*	*0.9670*	0.8863	*62.2563*	0.0461	*0.6558*	*0.8450*	249.0691
IFCNN	15.4712	0.9269	0.8089	73.9047	0.2699	0.4470	0.6140	258.1872
TCCL-DenseFuse	23.4018	0.9637	*0.8865*	46.0154	0.0030	0.6573	0.8975	233.7738

When the PE is less than 200 km, TCCL-DenseFuse occupies the biggest number in several PE ranges. When the PE is greater than 200 km, TCCL-DenseFuse occupies the least number in most PE ranges. This fully demonstrates the superior performance of our TCCL-DenseFuse model on PE. To compare the various models more intuitively, we show the average evaluation indicators of each model in Table II.

As can be seen from Table II, the proposed TCCL-DenseFuse model has five best quantitative values (PSNR, $N^{AB/F}$, $Q^{AB/F}$, SSIM, PE), one second-best value (FMI), one third value (CC), and a worst value (SD). The best PSNR and $N^{AB/F}$ mean that there is less noise interference in the fused images, and the overall fusion performance of our TCCL-DenseFuse model is good. Owing to the loss function designed in this article, our TCCL-DenseFuse model has the best $Q^{AB/F}$ and SSIM, which means that the fused image contains a large amount of gradient information and rich structural information, which directly affects the TC center location accuracy. The FMI and CC values of TCCL-DenseFuse are second and third but only 0.0054 and 0.0059 away from the best metric value, respectively. This means that TCCL-DenseFuse largely preserves the information in the two-channel source images, which facilitates the

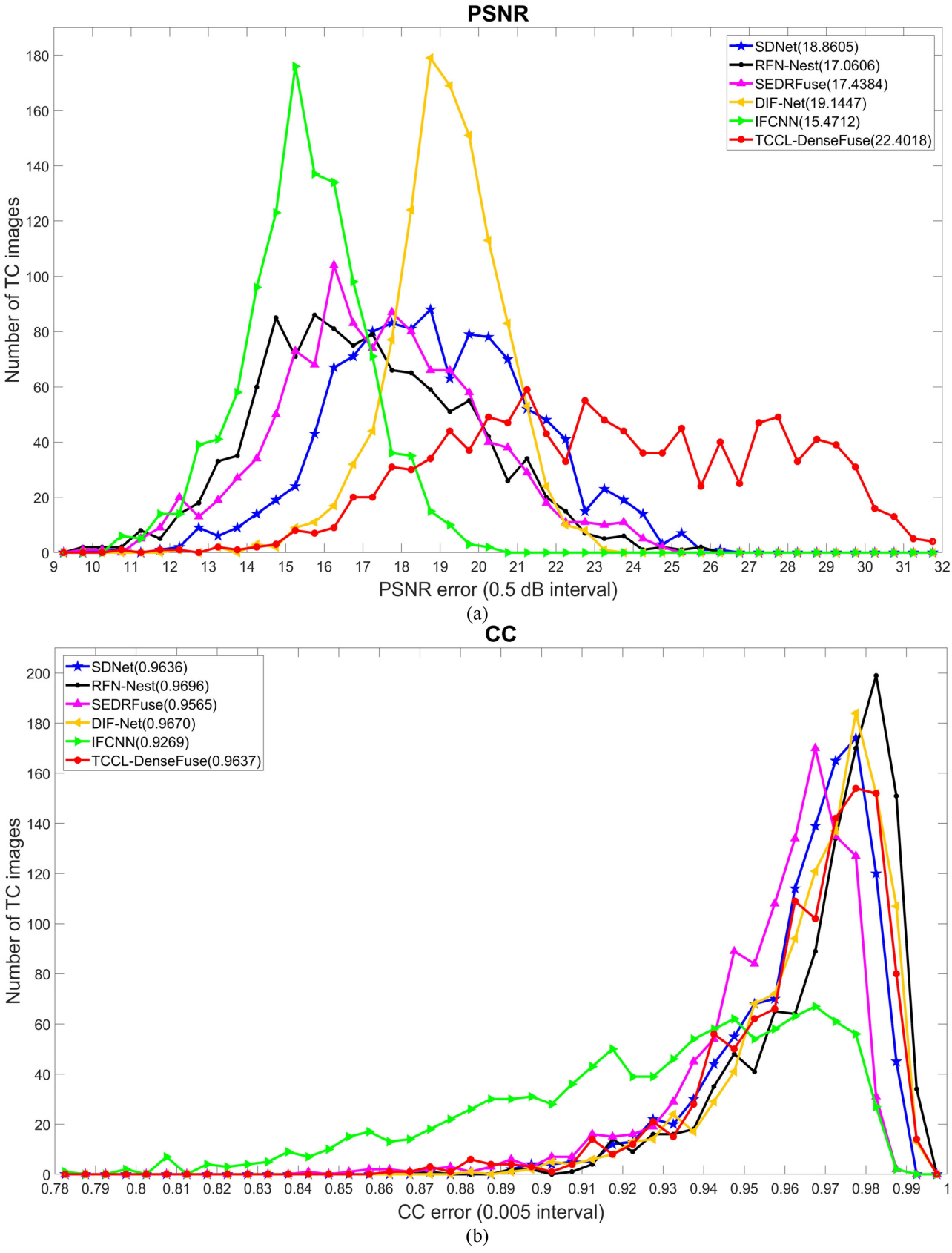
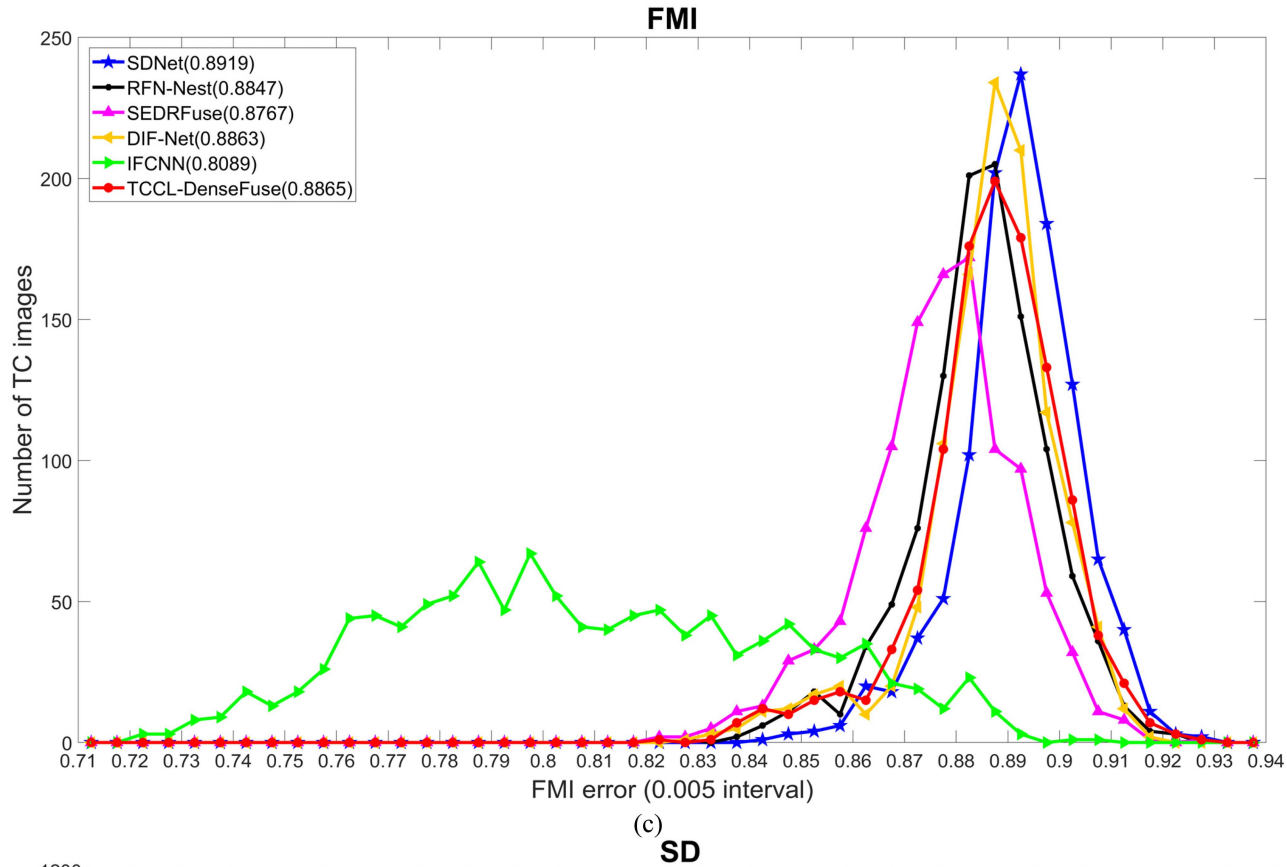
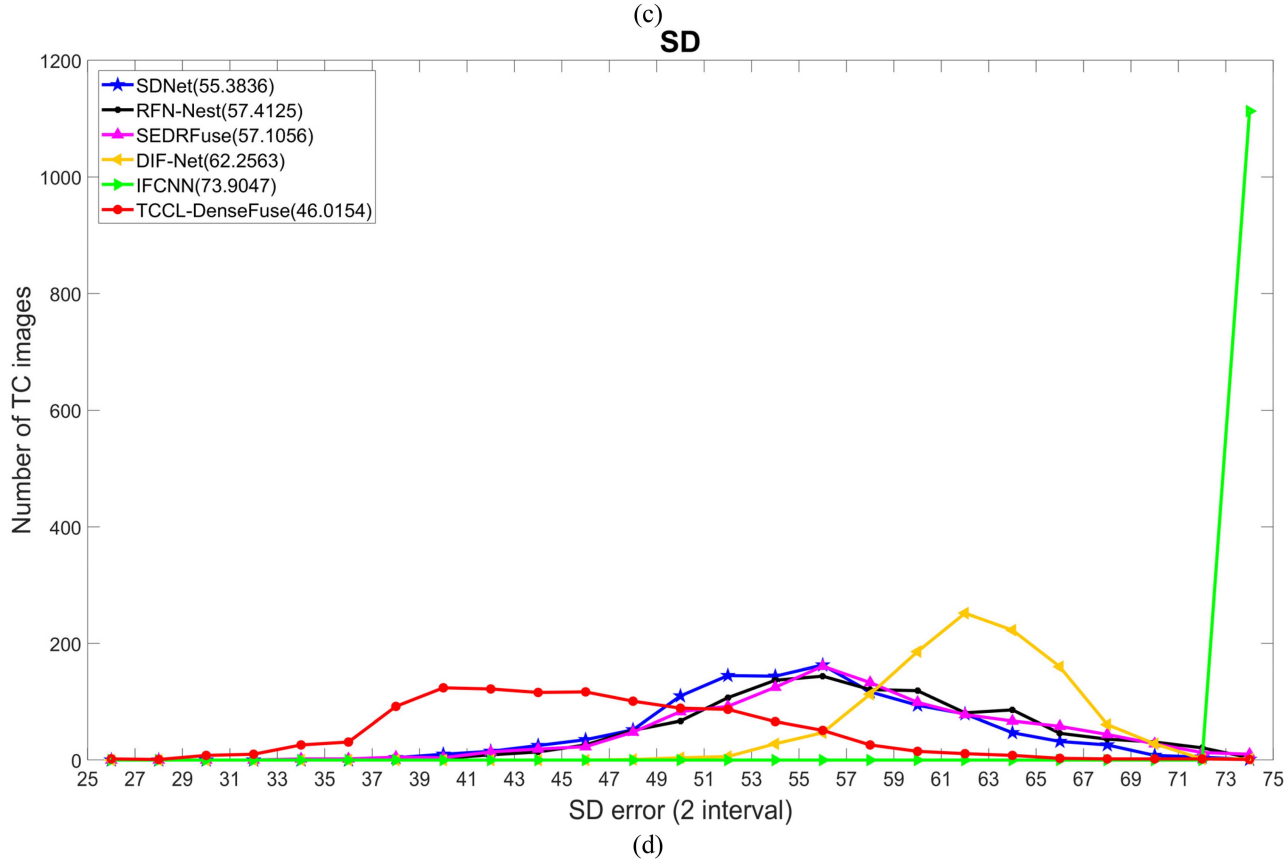


Fig. 9. Eight curves using different image fusion models on TC test images. (a) PSNR. (b) CC.

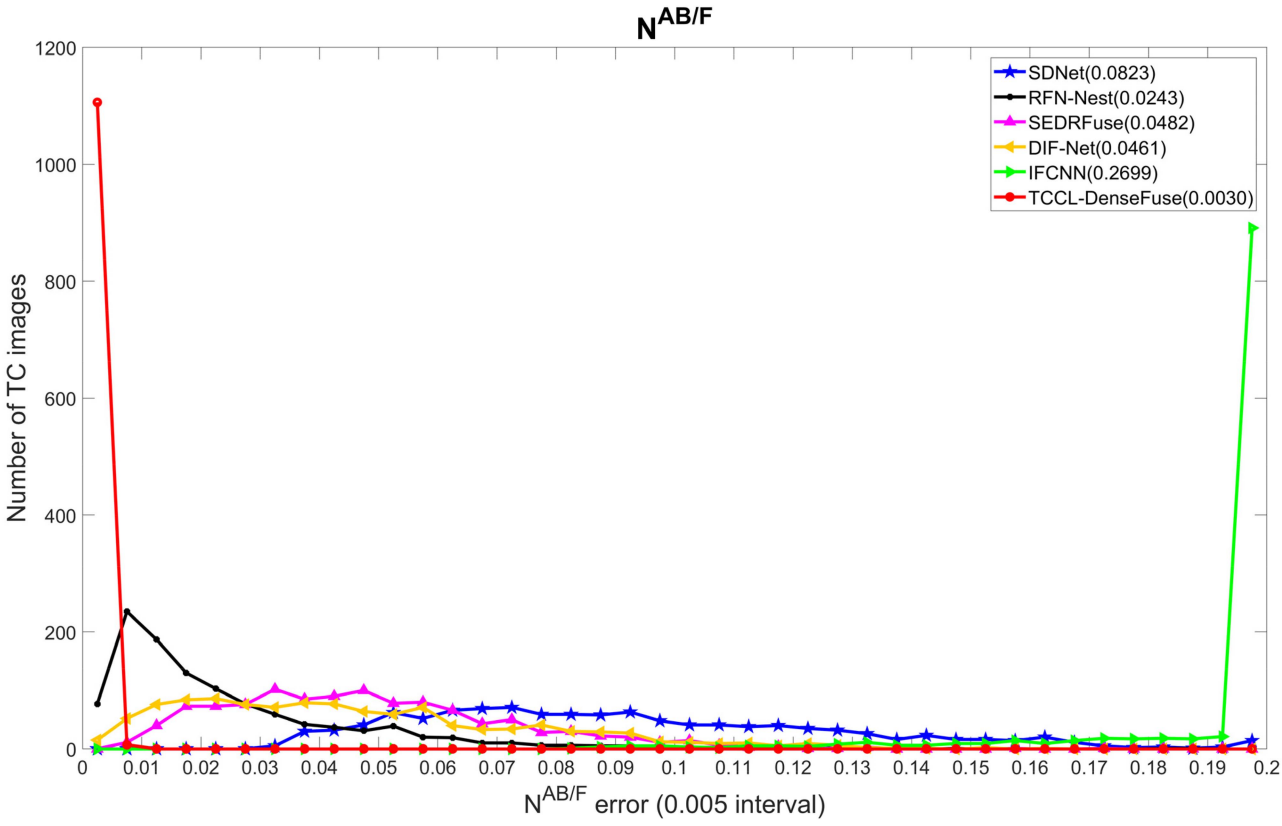


(c)

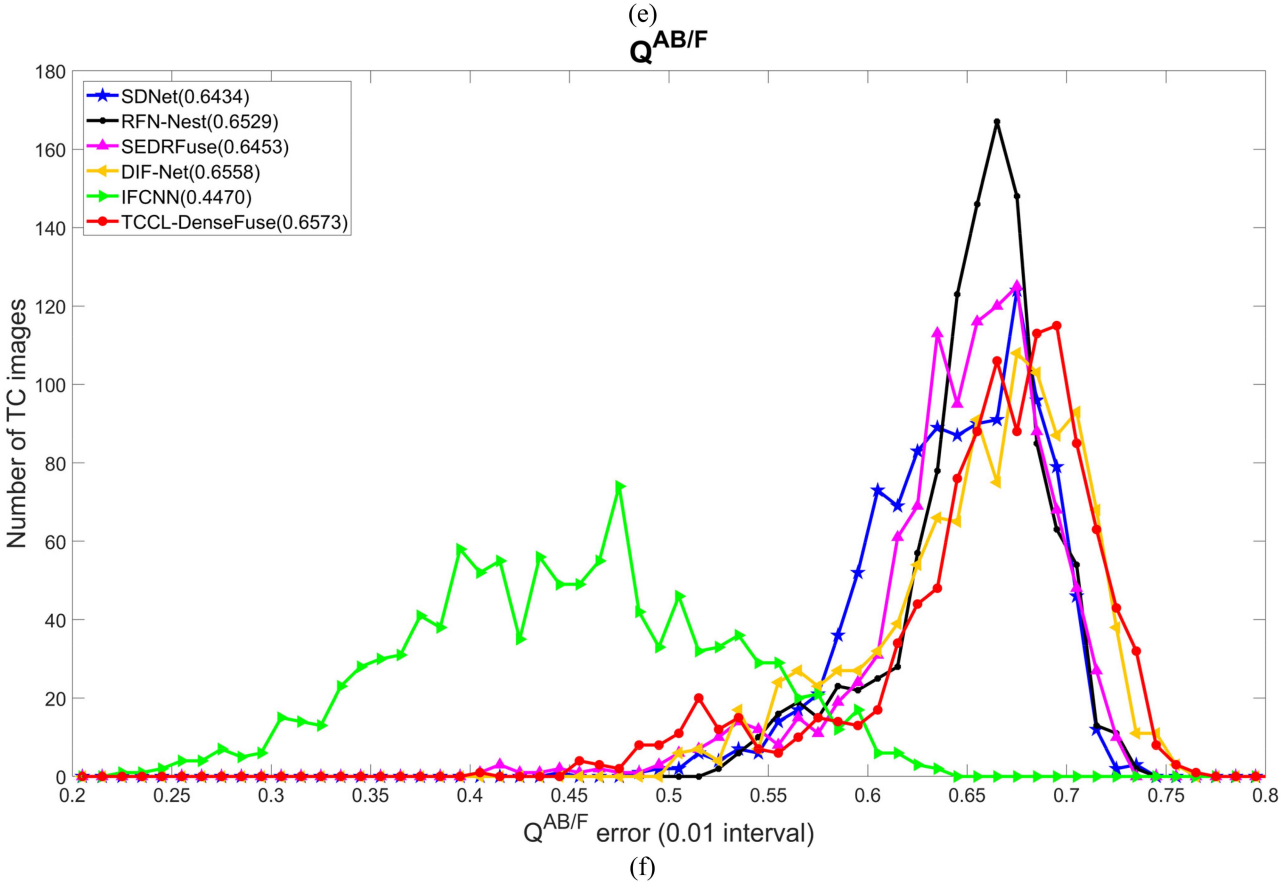


(d)

Fig. 9. (*Continued..*) Eight curves using different image fusion models on TC test images. (c) FMI. (d) SD.

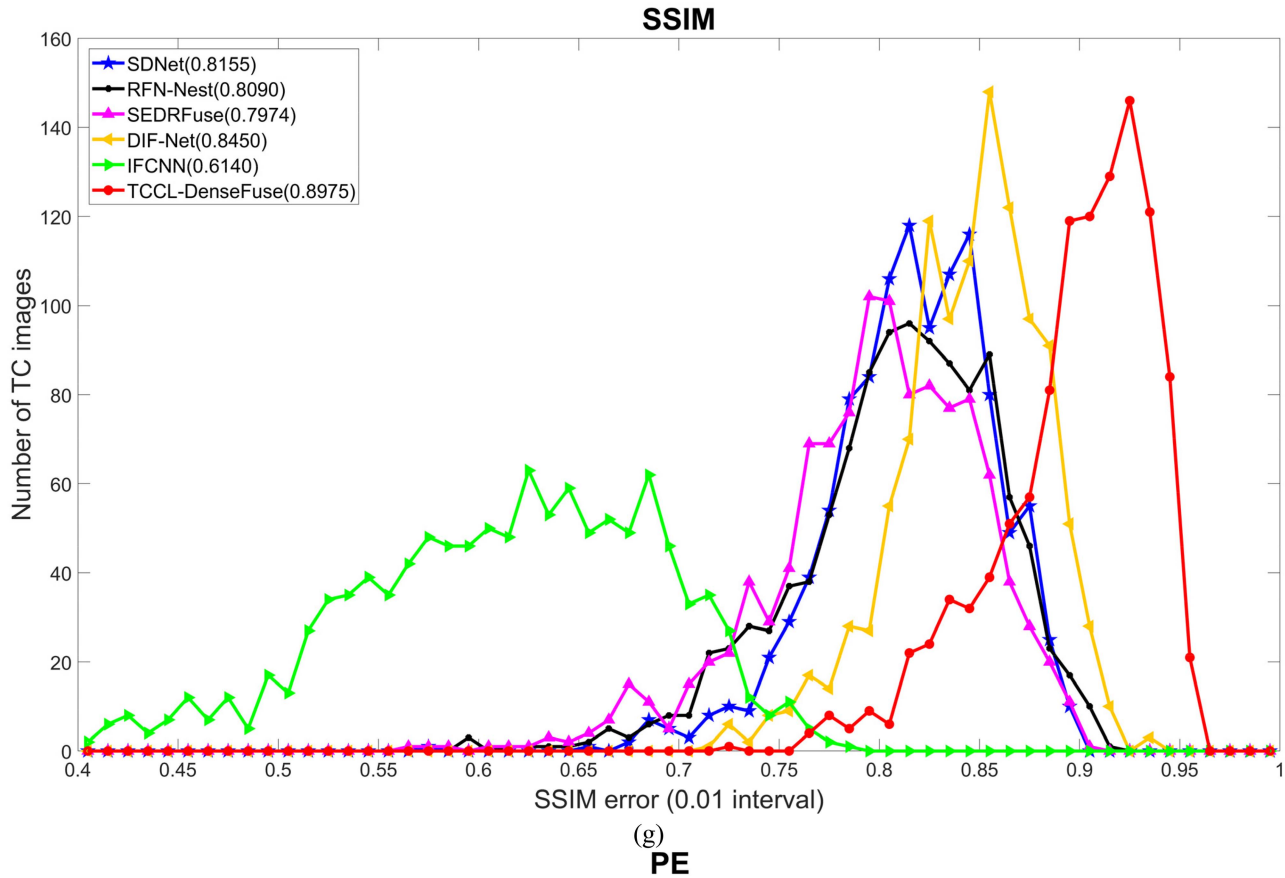


(e)

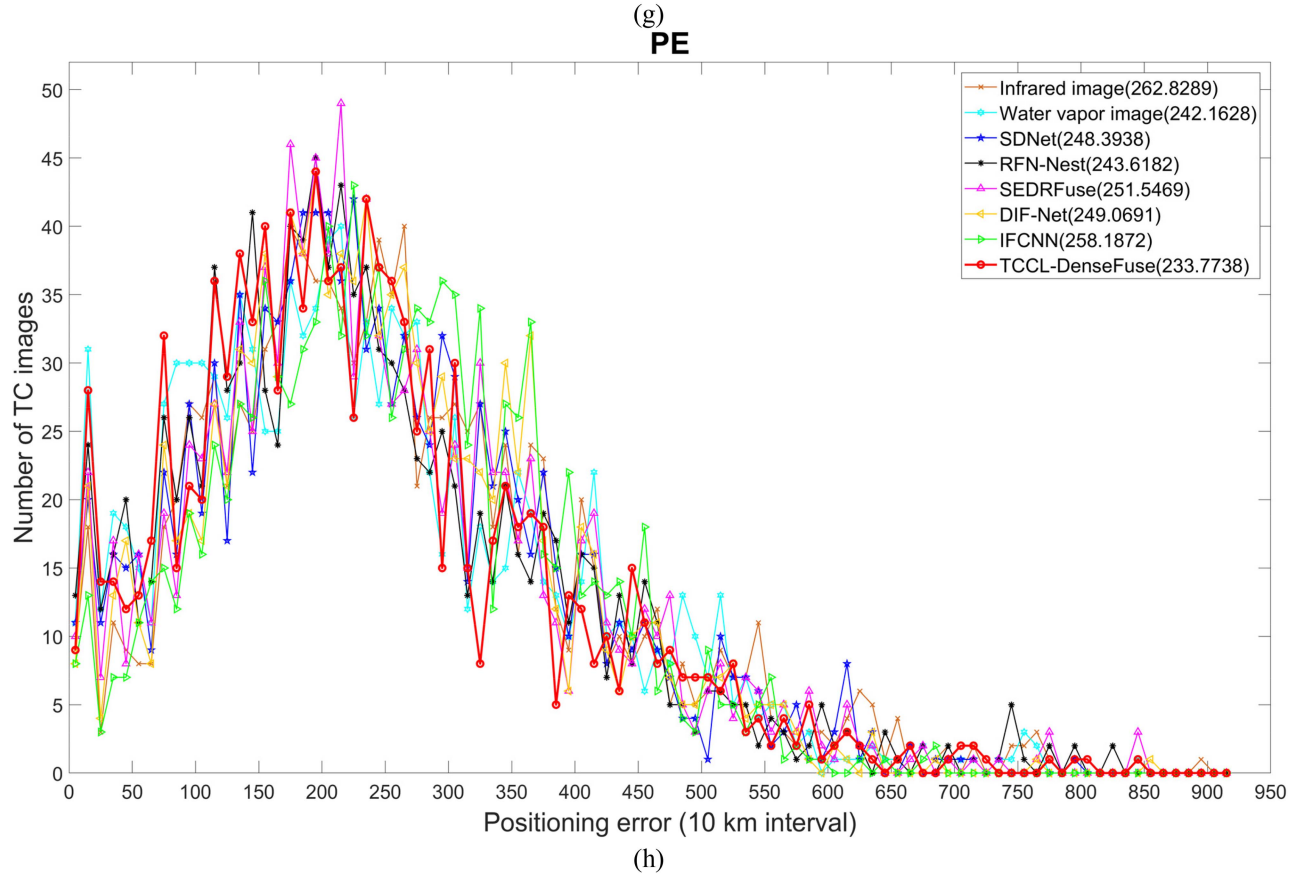


(f)

Fig. 9. (*Continued..*) Eight curves using different image fusion models on TC test images. (e) $N^{AB/F}$. (f) $Q^{AB/F}$.



(g)



(h)

Fig. 9. (*Continued.*.) Eight curves using different image fusion models on TC test images. (g) SSIM. (h) PE.

TABLE III
PE OF EACH MODEL IN EACH TC CATEGORY ON TEST IMAGES

Model	TD	TS	STS	TY	STY	SuperTY
Infrared image	257.0042	266.1774	274.9773	288.4279	249.0583	236.7684
Water vapor image	232.0300	253.7801	*262.8449*	244.7933	242.0773	*195.8002*
SDNet	226.2821	254.1423	283.9908	277.0555	219.2131	211.6336
RFN-Nest	232.0945	253.5403	275.0759	*239.3918*	236.5271	202.3029
SEDRFuse	234.8281	*251.5348*	285.3550	291.5182	228.5511	199.2272
DIF-Net	231.1554	259.6305	264.5922	289.5296	*214.0823*	241.1164
IFCNN	233.9288	270.4459	290.1945	278.3916	249.8930	214.6880
TCCL-DenseFuse	*229.7258*	244.0753	256.2979	235.1365	200.6174	188.1466

subsequent application of fused images. In terms of SD indicators, as previously analyzed qualitatively, RFN-Nest, DIF-Net, and IFCNN rank in the top three and are significantly ahead of other models. The worst SD means that the fused image of TCCL-DenseFuse does not have high contrast, and the gray value difference between the foreground and background is not large.

2) Category Discussion: To further explore the influence of different TC categories on PE index, we divide the test set into six categories according to the TC intensity grade classification criteria. There are 314 TD images, 409 TS images, 121 STS images, 160 TY images, 59 STY images, and 50 SuperTY images on the test set of this article, and PE of each model is given in Table III.

As can be seen from Table III, TCCL-DenseFuse has the minimum PE on TS, STS, TY, STY, and SuperTY, second-minimum PE on TD. Because TD is weak TC, their TC clouds are incomplete, and the structure of the spiral rain band is not clear, resulting in TCCL-DenseFuse ranking second on PE. Our model does not perform well in the localization for TD, and although the PE is smaller than that of single-channel satellite images, it is not the best. On TS, only the PE of our model is significantly ahead of that of the water vapor channel, and SDNet, RFN-Nest, and SEDRFuse are very close to that of the water vapor channel. On STS, only our model is better than that of the water vapor channel. This may be because STS belongs to relatively weak TCs, whose TC clouds are generally small, and the spiral rain bands are not clear or complete. TY, STY, and SuperTY are strong TCs. Their TC cloud structures are relatively complete, the overall outline is clearly visible, and some of them even have clear eyes. The STS is at the boundary between strong and weak TCs, and its TC clouds vary in size, so it is difficult to find the region to be detected close to the center of the TC. TY presents the similar situation as TS, that is, the PE of our model and RFN-Nest is smaller than that of the water vapor channel. TCCL-DenseFuse is significantly better than other models in the PE of the STY and SuperTY categories. This is because the spiral rain band or eye of a strong TC is clear. The fusion TC image obtained by our TCCL-DenseFuse model can be used to find the region closest to the TC center, so the PE is small.

To see the PE of each model in each category, we ranked PE obtained by different methods in each TC category, and then drew the ranking results into a line graph, as shown in Fig. 10.

In Fig. 10, the decimal after each model's name is its average ranking in each of the six TC categories, with the order of the average ranking in parentheses. This means that the smaller the decimal after a model, the higher the order in brackets, and the better the effect of TC center positioning. From Fig. 10, it can be seen that TCCL-DenseFuse has the minimum PE on TS, STS, TY, STY, and SuperTY, and the second-minimum PE on TD. The TC center positioning performance of all TC categories using the fusion image of our TCCL-DenseFuse model is the most stable, and other models fluctuate in a wide range. TCCL-DenseFuse has the smallest fluctuation and ranks first on average, which shows the stability of our model and its good positioning performance for various categories of TCs.

In practical applications, we need to consider the spatial resolution of the satellite imagery. This article adopts the spatial resolution of FY-2 (5 km) and FY-4 (4 km), which means that if the PE using two models is within 4 or 5 km, it can be considered that they are ranked in the same place. To make a fair comparison, the spatial resolution of FY-4 (4 km) is used to determine whether the two models rank in the same place in terms of the PE. We reranked PE obtained by different models in each TC category, and then drew the reranking results into a line graph, as shown in Fig. 11.

After the spatial resolution of satellite images is taken into account, the average ranking of all models is higher. From Fig. 11, it can be clearly seen that the TC center positioning performance of all TC categories using the fusion image of our TCCL-DenseFuse model is the first best. The advantage of our TCCL-DenseFuse model is obvious, and the overall

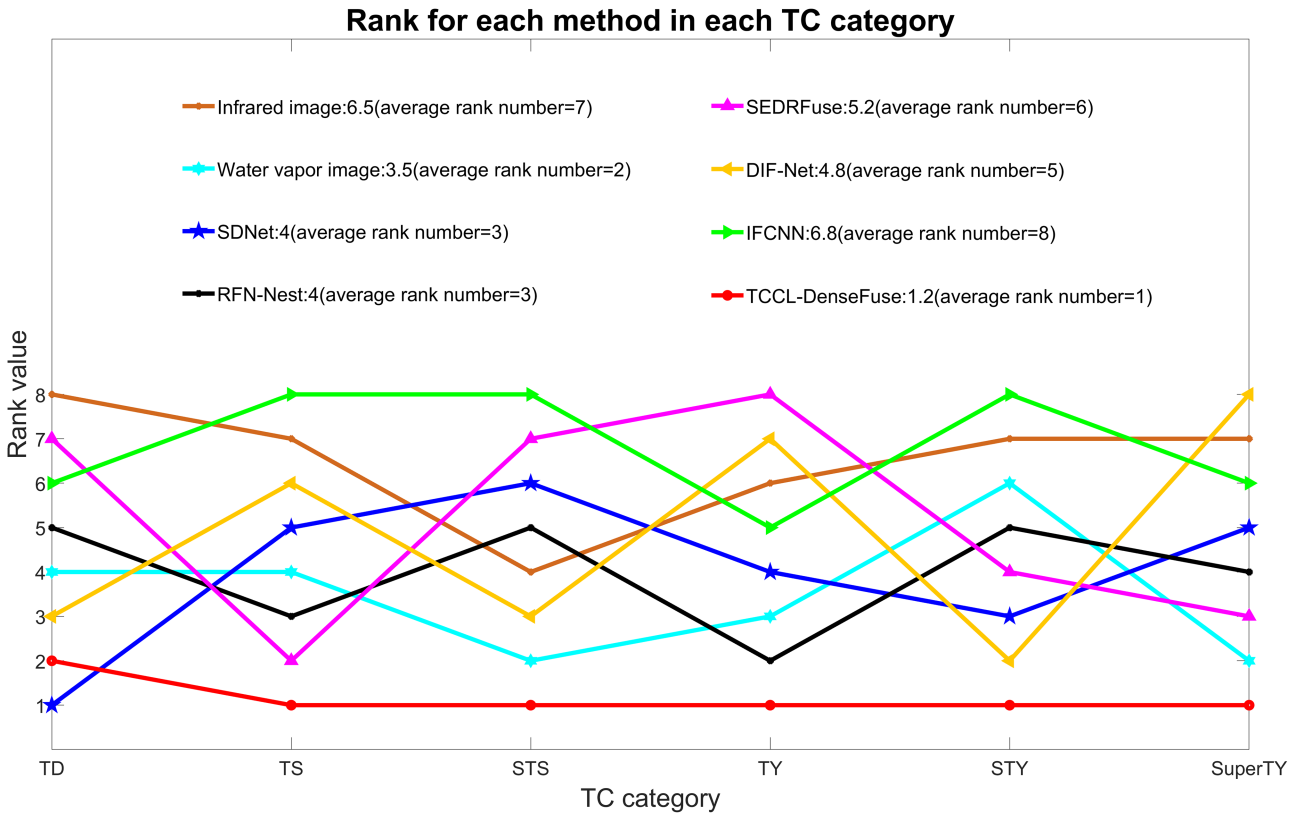


Fig. 10. Ranking curve for each TC category using different models. The name of each model is followed by its corresponding average ranking in parentheses.

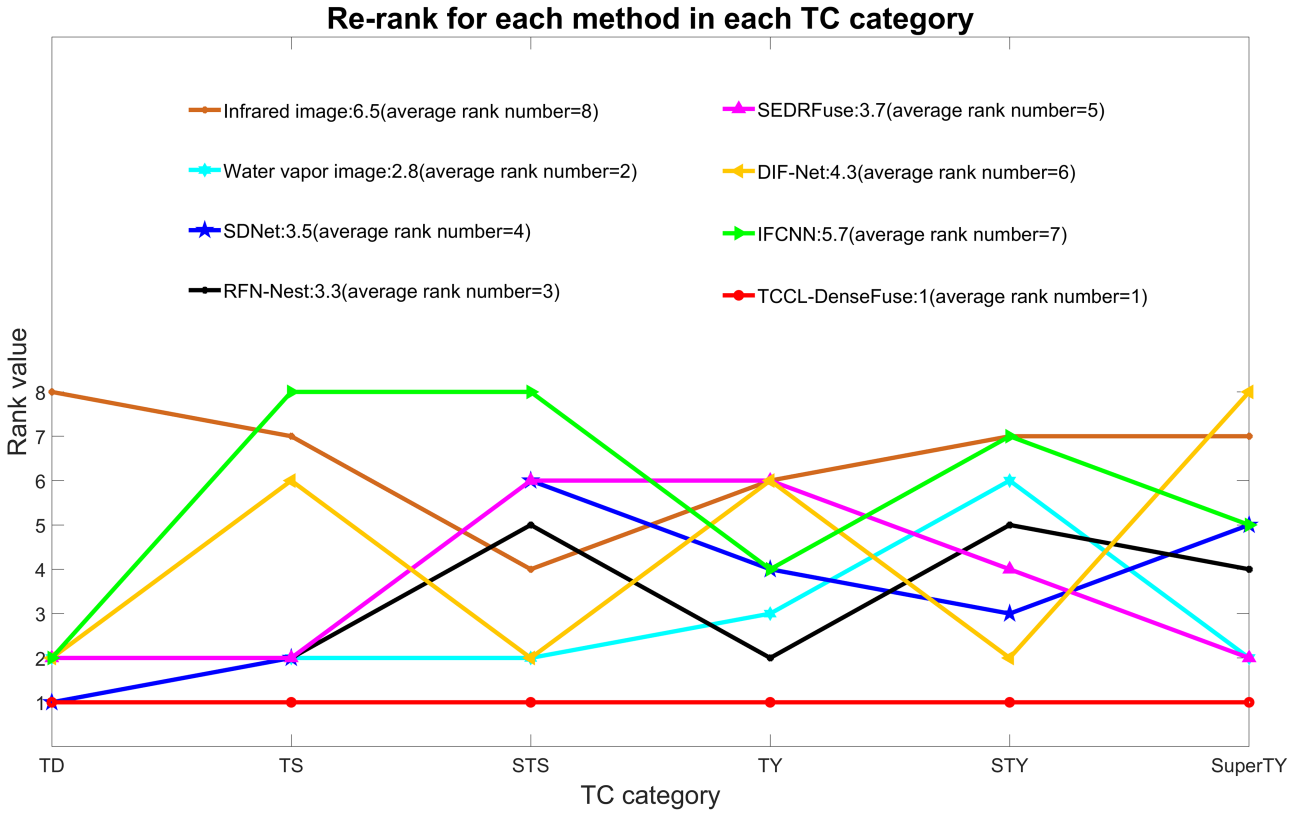


Fig. 11. Reranking curve for each TC category using different models. The name of each model is followed by its corresponding average ranking in parentheses.

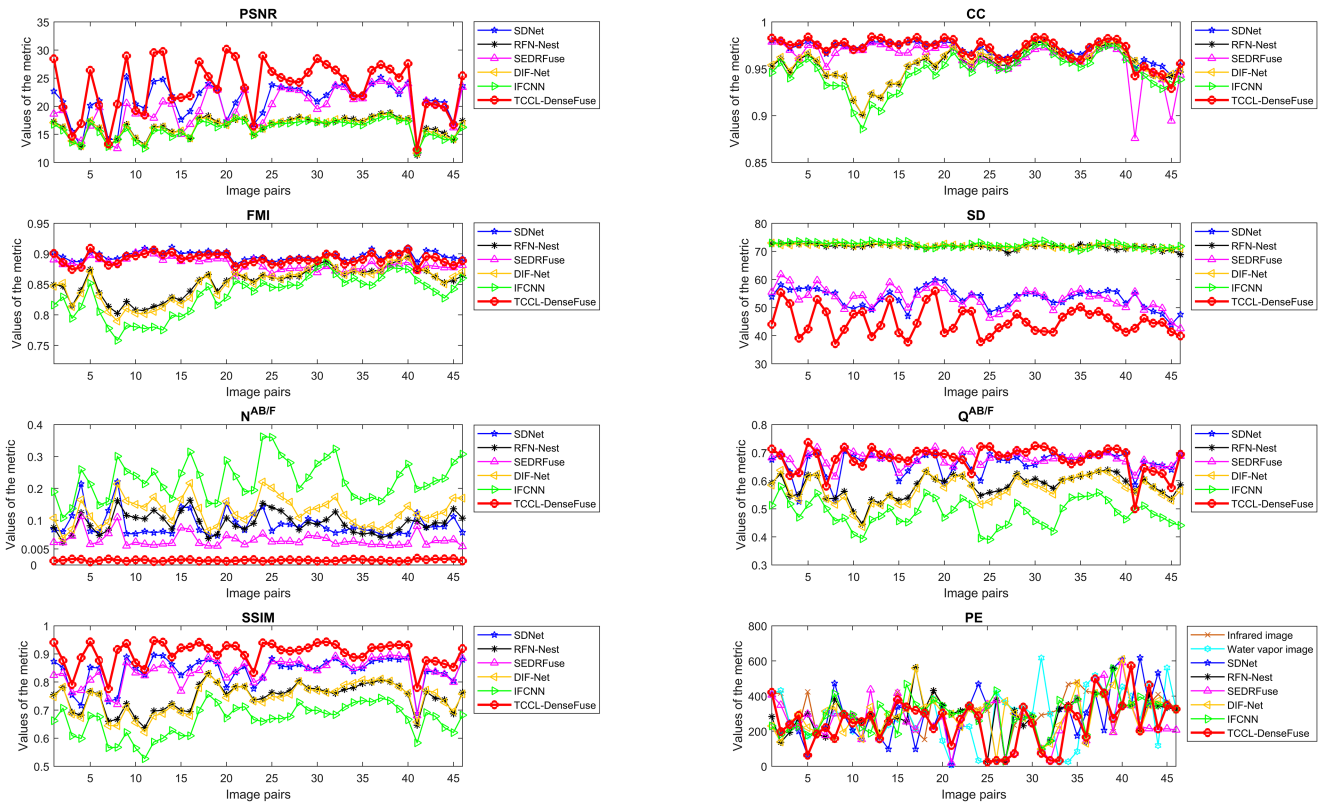


Fig. 12. Quantitative comparison results of case 1 (superTY Chanchu).

performance of our model is fully verified. Using the fusion image obtained by our model to locate the TC center can achieve a good positioning effect for all categories of TCs.

3) TC Case Discussion: Two TC cases from the TC test set are explored to further verify the performance of our TCCL-DenseFuse model. Case 1 comes from FY-2 geostationary satellite and case 2 comes from FY-4 geostationary satellite.

a) Case 1: Super Typhoon (SuperTY) Chanchu (No.0601): Case 1 is a super typhoon Chanchu (No.0601) from 12:00 on May 9, 2006 to 12:00 on May 18, 2006. Super typhoon Chanchu includes a total of 46 satellite images, including TS, STS, TY, STY, and the SuperTY category. Eight objective evaluation indicator curves are shown in Fig. 12.

From Fig. 12, it can be seen that TCCL-DenseFuse performs well on PSNR, $N^{AB/F}$, and SSIM, with obvious advantages. TCCL-DenseFuse is equivalent to SDNet, SEDRFuse on CC, FMI, and $Q^{AB/F}$. The SD is still the worst. On PE, most of the TCCL-DenseFuse results are below 200 km, and there are only five images above 400 km. We average the evaluation indicators of the case 1 (SuperTY Chanchu), and the results are given in Table IV.

From Table IV, it can be seen that our model has six best quantitative values on PSNR, CC, $N^{AB/F}$, $Q^{AB/F}$, SSIM, and PE; one second-best value on FMI and a worst value on SD. TCCL-DenseFuse's PE leads SEDRFuse by 15 km, which is also a big lead.

b) Case 2: Tropical Storm (TS) Nuri (No.2002): Case 2 is tropical storm Nuri (No. 2002) from 00:00 on June 10, 2020 to 12:00 on June 14, 2020. Tropical storm Nuri includes a total of 24 satellite images, including TD and TS categories. Eight objective evaluation indicator curves are given in Fig. 13.

From Fig. 13, it can be seen that TCCL-DenseFuse performs well on PSNR, $N^{AB/F}$, and SSIM with obvious advantages. TCCL-DenseFuse is equivalent to SDNet, SEDRFuse on CC and FMI, and TCCL-DenseFuse is equivalent to DIF-Net on $Q^{AB/F}$. The SD is still the worst. In terms of PE, most of the TCCL-DenseFuse results are below 300 km. We average the evaluation indicators of case 2 (TS Nuri), and the results are given in Table V.

From Table V, it can be seen that our TCCL-DenseFuse model has five best quantitative values on PSNR, $N^{AB/F}$, $Q^{AB/F}$, SSIM, and PE; one third best value on CC; one fourth best value on FMI; and a worst value on SD. Only our model is better than that of the water vapor channel, which shows that its TC center positioning performs well in case 2. In sum, our TCCL-DenseFuse model is largely ahead of other models in both cases (the ranking is the first), showing its good robustness to different TC categories.

Note that TCCL-DenseFuse with low contrast achieves the best performance in the evaluation index of PE. However, DIF-Net with better contrast ranks second on PE, which shows that the accuracy of TC center positioning does not relate much to the image contrast. To further explore the reasons for the improvement of the TC center positioning accuracy in this

TABLE IV
AVERAGE METRICS OF CASE 1 (SUPERTY CHANCHU)

Model	PSNR	CC	FMI	SD	$N^{AB/F}$	$Q^{AB/F}$	SSIM	PE
Infrared image	—	—	—	—	—	—	—	304.6675
Water vapor image	—	—	—	—	—	—	—	282.5761
SDNet	*20.9691*	*0.9709*	0.8970	53.3470	0.0879	0.6622	*0.8406*	279.3183
RFN-Nest	16.3566	0.9539	0.8538	71.6961	0.0931	0.5797	0.7463	288.9606
SEDRFuse	19.7530	0.9614	0.8833	53.2375	*0.0389*	*0.6779*	0.8378	*264.5203*
DIF-Net	16.2767	0.9553	0.8522	*71.9835*	0.1294	0.5700	0.7428	295.5221
IFCNN	15.9345	0.9467	0.8331	72.3651	0.2294	0.4856	0.6644	300.3306
TCCL-DenseFuse	23.4692	0.9712	*0.8920*	45.0353	0.0030	0.6780	0.9013	249.4475

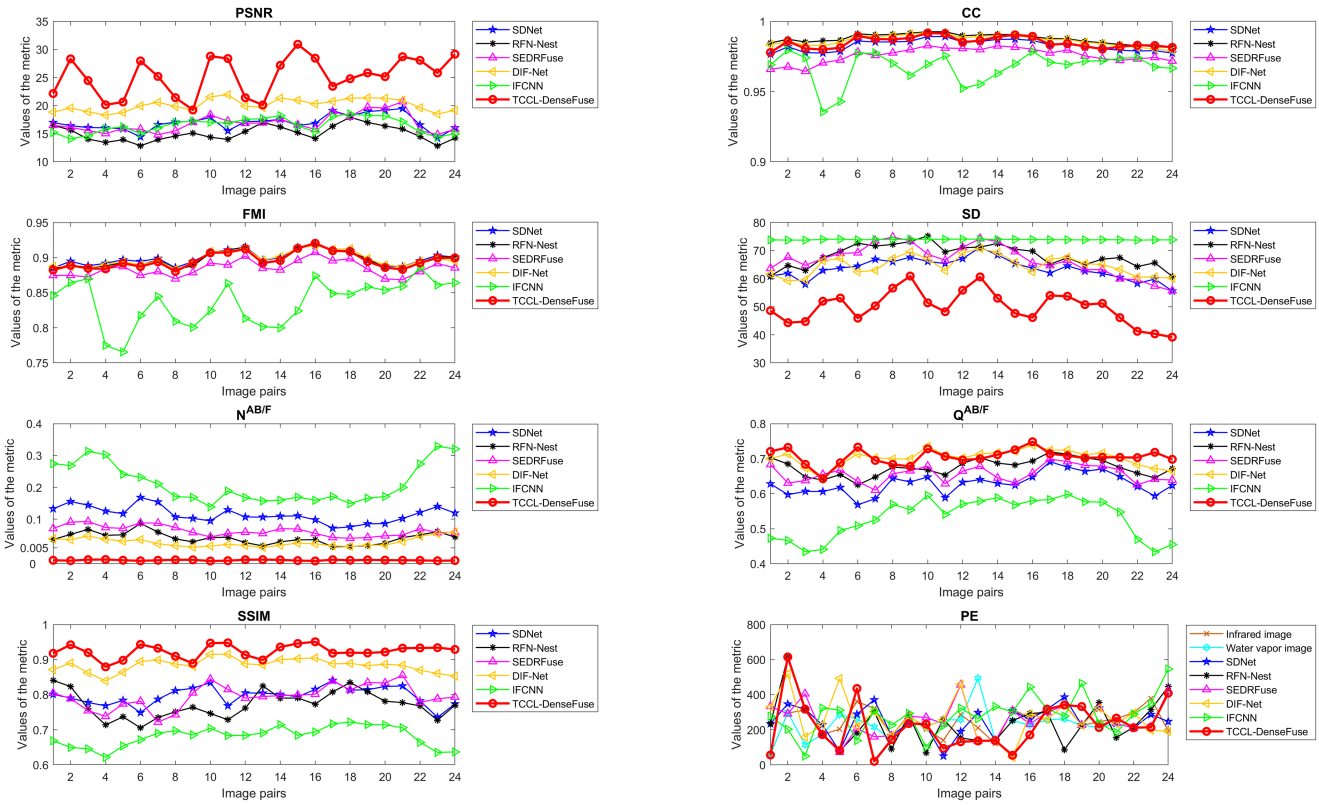


Fig. 13. Quantitative comparison results of case 2 (TS Nuri).

article, we perform pseudocolor processing on Figs. 3–8. To extrude difference in fusion images using different models, only nine colors are used to produce the pseudocolor fusion images. The results are shown in Figs. 14 –19.

It can be seen from Figs. 14(a) and (b) to 19(a) and (b) that the contours of the spiral rain band and TC inner core are mainly concentrated in the yellow and orange regions. The spiral rain band structure in the fusion TC image obtained by our TCCL-DenseFuse model is relatively clear and complete, and the brightness temperature gradient levels of TC inner core are preserved well [see Figs. 14(h)–19(h)]. However, the fusion TC image obtained by other models is too finely segmented and disorderly owing to the excessively enhanced contrast, as shown in Figs. 14(c)–(g) and 19(c)–(g). From Figs. 14 to 19, it

TABLE V
AVERAGE METRICS OF CASE 2 (TS NURI)

Model	PSNR	CC	FMI	SD	$N^{AB/F}$	$Q^{AB/F}$	SSIM	PE
Infrared image	—	—	—	—	—	—	—	277.0247
Water vapor image	—	—	—	—	—	—	—	*244.2608*
SDNet	16.9545	0.9829	0.8993	63.5573	0.1151	0.6287	0.7966	268.3201
RFN-Nest	15.0525	0.9877	0.8973	*68.2542*	0.0401	0.6752	0.7718	248.5358
SEDRFuse	16.8592	0.9758	0.8844	66.7884	0.0635	0.6563	0.7955	274.6287
DIF-Net	*20.1077*	*0.9865*	*0.8980*	64.7078	*0.0277*	*0.7027*	*0.8840*	254.0397
IFCNN	16.4658	0.9674	0.8363	73.9176	0.2126	0.5305	0.6813	283.2539
TCCL-DenseFuse	25.2365	0.9850	0.8962	49.8360	0.0022	0.7053	0.9242	238.0739

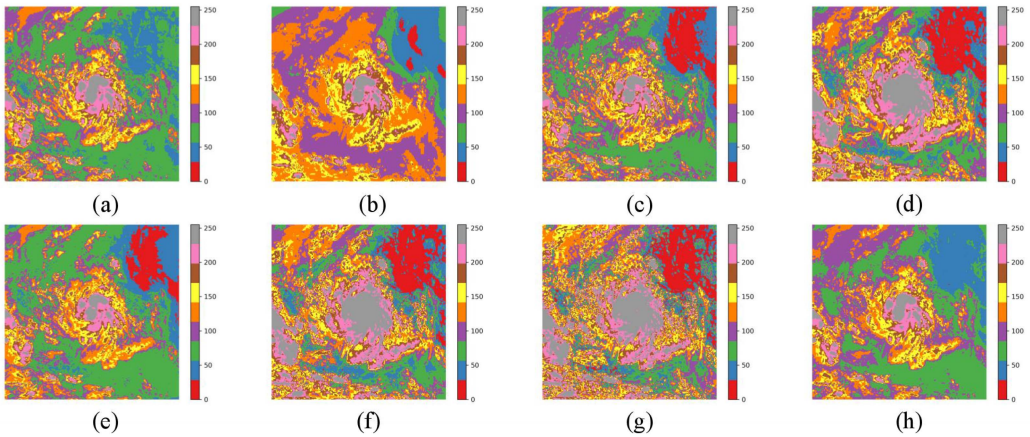


Fig. 14. TD. (a) Infrared image (PE = 206.5527 km). (b) Water vapor image (PE = 178.3704 km). (c) SDNet (PE = 195.2434 km). (d) RFN-Nest (PE = 212.2074 km). (e) SEDRFuse (PE = 195.2434 km). (f) DIF-Net (PE = 215.1279 km). (g) IFCNN (PE = 185.9247 km). (h) TCCL-DenseFuse (PE = 105.1095 km).

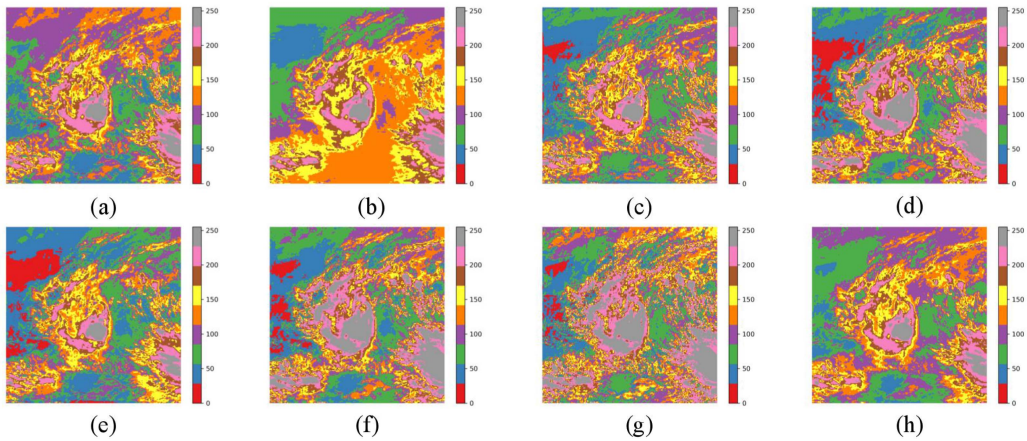


Fig. 15. TS. (a) Infrared image (PE = 149.9066 km). (b) Water vapor image (PE = 234.3502 km). (c) SDNet (PE = 149.9066 km). (d) RFN-Nest (PE = 149.9066 km). (e) SEDRFuse (PE = 149.9066 km). (f) DIF-Net (PE = 271.4259 km). (g) IFCNN (PE = 271.4259 km). (h) TCCL-DenseFuse (PE = 48.0833 km).

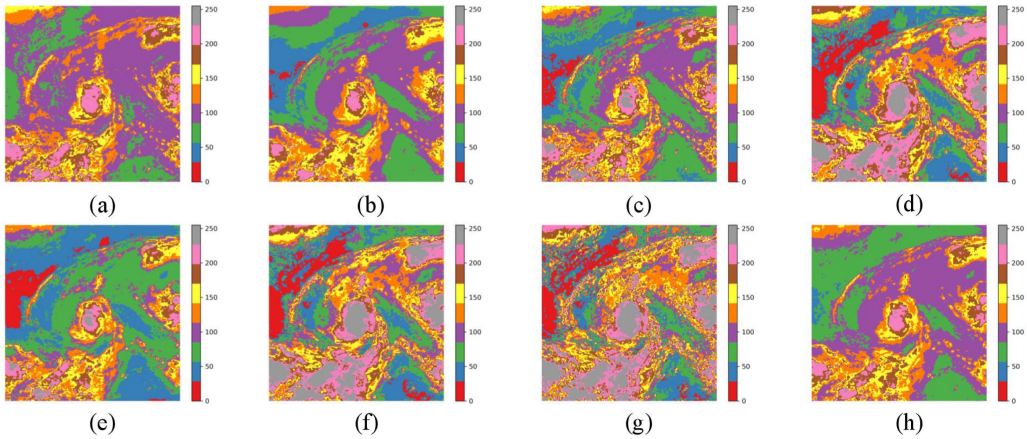


Fig. 16. STS. (a) Infrared image (PE = 161.9027 km). (b) Water vapor image (PE = 390.4645 km). (c) SDNet (PE = 161.9027 km). (d) RFN-Nest (PE = 283.0415 km). (e) SEDRFuse (PE = 633.3344 km). (f) DIF-Net (PE = 283.0415 km). (g) IFCNN (PE = 262.7975 km). (h) TCCL-DenseFuse (PE = 70.7990 km).

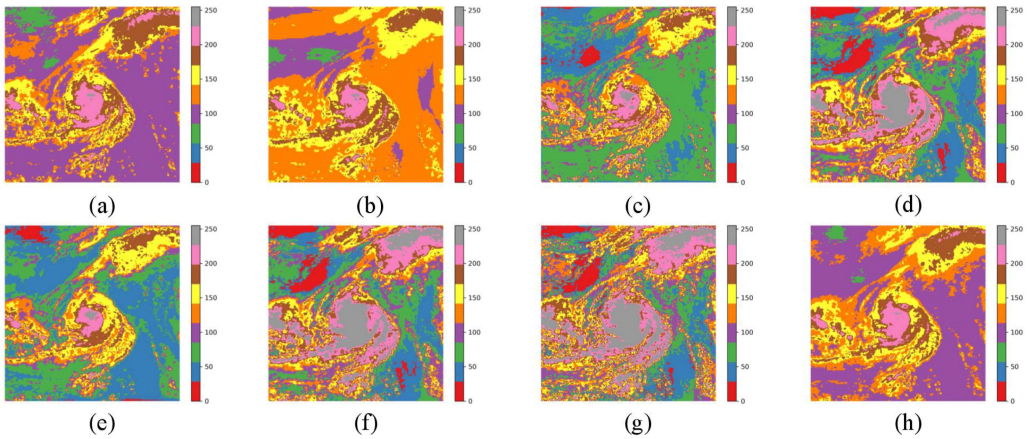


Fig. 17. TY. (a) Infrared image (PE = 204.6033 km). (b) Water vapor image (PE = 316.8793 km). (c) SDNet (PE = 208.7163 km). (d) RFN-Nest (PE = 187.3833 km). (e) SEDRFuse (PE = 196.3734 km). (f) DIF-Net (PE = 266.7630 km). (g) IFCNN(PE = 445.0421 km). (h) TCCL-DenseFuse (PE = 90.8983 km).

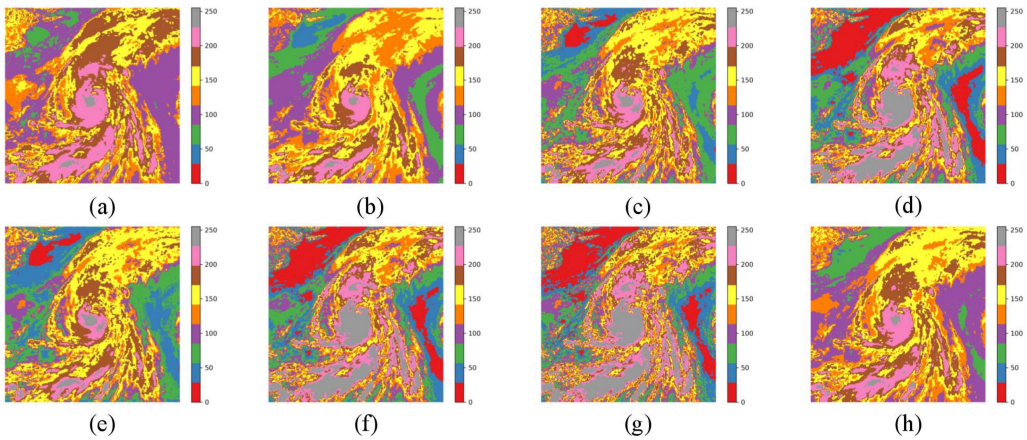


Fig. 18. STY. (a) Infrared image (PE = 292.5107 km). (b) Water vapor image (PE = 611.2385 km). (c) SDNet (PE = 97.7880 km). (d) RFN-Nest (PE = 97.7880 km). (e) SEDRFuse (PE = 97.7880 km). (f) DIF-Net (PE = 97.7880 km). (g) IFCNN (PE = 97.7880 km). (h) TCCL-DenseFuse (PE = 74.2462 km).

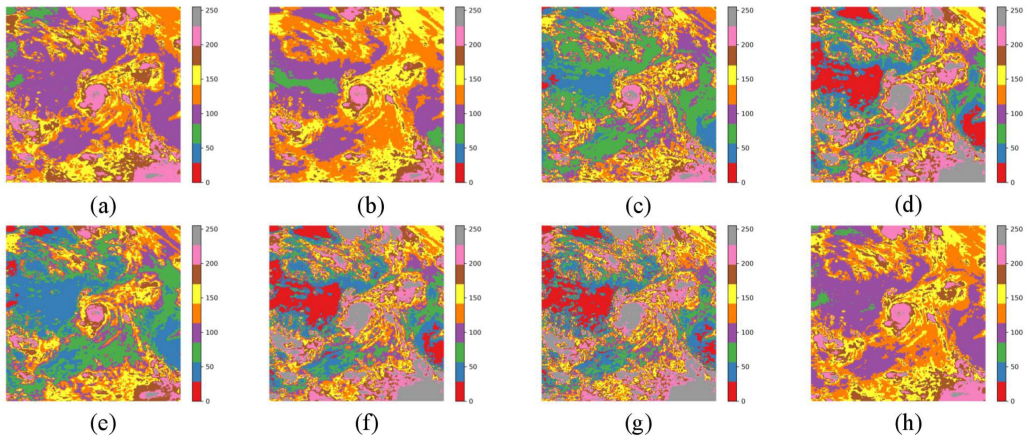


Fig. 19. SuperTY. (a) Infrared image (PE = 80.7001 km). (b) Water vapor image (PE = 526.6047 km). (c) SDNet (PE = 80.7001 km). (d) RFN-Nest (PE = 419.1807 km). (e) SEDRFuse (PE = 50.4743 km). (f) DIF-Net (PE = 419.6576 km). (g) IFCNN (PE = 339.0612 km). (h) TCCL-DenseFuse (PE = 12.7476 km).

can be seen that the gray value of the TC main body in the fusion image obtained using RFN-Nest, DIF-Net, and IFCNN models has little change, and the color is mainly gray, which is because of the excessive enhancement of the cloud structure and the loss of some details of the rain band and inner core gradient. Therefore, when other spiral-like clouds except for TC clouds appear in the fused image, the excessive enhancement leads to the wrong location being detected, and the PE is relatively large. When the fusion image is used to locate the TC center, the wrong region to be detected is found, leading to a large TC PE. This indicates that the brightness temperature gradient level near the TC inner core affects the determination of the region to be detected, thus affecting the TC center location result. This fully illustrates the importance of finding a region to be detected close to the center of the TC.

As can be seen from Figs. 14 to 19, there are some modes that are only a few tens of kilometers apart from the PE of our model. This may be because the area to be detected, including the TC center, is found, but the gradient levels near the TC center are not as rich as those of TCCL-DenseFuse. The PE of other models is too large, while the PE of our model is relatively small. For the fusion image with no-eye region and only one TC, the clear and complete spiral rain band indicates that the brightness temperature gradient of the fusion TC image changes greatly, which is conducive to finding the region to be detected closer to the TC center. For the no-eye TC, the richer the brightness temperature gradient level in the TC inner core region containing the TC center, the better the TC center positioning. However, for the fusion image with an obvious eye and only one TC, whether the TC inner core brightness temperature gradient level is rich does not have a significant impact on the final TC center location result owing to the great variation in the gradient around the TC eye. This is why for the PE using SEDRFuse, the water vapor images in Figs. 18 and 19 are small and close to the PE values of our model. Generally, eye TC images mostly correspond to the TC with high intensity, while there are relatively few moments with high intensity in the whole TC life cycle. Therefore, no-eye TC images are more common in practical applications. As mentioned above, the change in the gradient level near the no-eye TC center has a great impact on the accuracy of TC center positioning.

In conclusion, the fusion image used to improve the positioning accuracy of the TC center should have relatively complete main structures, such as spiral rain bands and rich brightness temperature gradient levels in the TC inner core region. If the fusion image obtained by a model has rich brightness temperature gradient levels near the TC center, but the detection region, including the TC center, is not accurate, a large PE is obtained. If the region to be detected near the TC center can be found in the fusion image obtained using a certain model, a small PE can also be obtained even if there are fewer brightness temperature gradient levels near the TC center. However, when two models can find the region containing the TC center to be detected, the richness of the brightness temperature gradient level near TC center plays an important role in the positioning accuracy of the TC center. Therefore, according to the TC center positioning method, to obtain a small PE, it is important to find the region to be detected containing the TC center in the fusion image. Second, the brightness temperature gradient level of the region to be detected also has a great influence on the final TC center positioning, especially for no-eye TCs.

In the entire test set, category and case discussions, in view of the common evaluation indicators and PEs, the comprehensive performance of the proposed TCCL-DenseFuse model is the best, which shows that it not only generates information-rich fusion images but also improves the accuracy of TC center positioning. This also means that the proposed neural network structure is effective for the fusion of the infrared image and water vapor image.

IV. Conclusion

The proposed TCCL-DenseFuse satellite image fusion model can generate information-rich images while improving TC center positioning accuracy. To retain the information of both the infrared and water vapor channels, we add them in the channel dimension and use DenseNet to extract image features, which can avoid neural network overfitting and reduce the number

of network parameters. When deep learning is used for image fusion, the loss function plays an important role. In order to try to contain as much information as possible in the fusion image, to ensure that the fusion image contains as much information of the source images as possible, we introduce multiscale structural similarity into the loss function, and to better preserve the gradient information of the source image, we also introduce gradient loss.

In particular, the focus of this article is to propose an effective fusion model of infrared and water vapor channel satellite images, rather than using deep learning technology combined with satellite images to solve the problem of TC center positioning. The proposed TCCL-DenseFuse model can obtain the fusion image containing rich source image information just like most of the current image fusion methods. At present, most image fusion methods often evaluate the quality of image fusion based on subjective evaluation combined with some objective indicators, and rarely discuss the practical application of image fusion. Originally, the purpose of image fusion is to obtain better results in some fields of practical application, rather than just to obtain a so-called fusion image containing rich information. Different from most current image fusion methods, experimental results show that the fusion image obtained by our TCCL-DenseFuse model is valuable in some practical applications, such as improving the accuracy of TC center positioning, providing a reference for typhoon prevention and disaster mitigation. In the future, perhaps this purpose-specific fusion method will be more promising in some practical applications.

Of course, our TCCL-DenseFuse model also has some disadvantages. For example, the fusion image obtained by our model cannot improve the center positioning accuracy of TC with weak intensity (such as TD) very well. In the future, we will modified our image fusion model to improve the center positioning accuracy of weak TC. In fact, both the TC center location and intensity estimation are important in making TC forecasts. Therefore, improving the TC intensity estimation accuracy using fused images is our next goal.

Acknowledgment

The authors would like to thank LetPub (www.letpub.com) for its linguistic assistance during the preparation of this article.

References

[1] C. Zhang, Y. Chen, C. Duanmu, and H. Feng, "Multi-channel satellite cloud image fusion in the tetrolet transform domain," *Int. J. Remote Sens.*, vol. 35, no. 24, pp. 8138–8168, 2014, doi: 10.1080/01431161.2014.980918.

[2] C. Zhang, X. Wang, and L. Xue, "DAGDU: Center location of tropical cyclones using deviation angle gradient distribution uniformity from infrared satellite images," *Int. J. Remote Sens.*, vol. 42, no. 24, pp. 9222–9240, 2021, doi: 10.1080/01431161.2021.1993462.

[3] M. F. Pineros, E. A. Ritchie, and J. S. Tyo, "Objective measures of tropical cyclone structure and intensity change from remotely sensed infrared image data," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 11, pp. 3574–3580, Nov. 2008.

[4] N. Jaiswal and C. M. Kishtawal, "Objective detection of center of tropical cyclone in remotely sensed infrared images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 1031–1035, Apr. 2013, doi: 10.1109/JSTARS.2012.2215016.

[5] J. Ma, Y. Ma, and C. Li, "Infrared and visible image fusion methods and applications: A survey," *Inf. Fusion.*, vol. 45, pp. 153–178, 2019, doi: 10.1016/j.inffus.2018.02.004.

[6] R. J. M. Daza, E. S. U. Cardona, F. Pineda, and O. Ardila, "Assessment of satellite-image fusion using Fast haar Wavelet transform and Contourlet," in *Proc. 9th Iberian Conf. Inf. Syst. Technol.*, 2014, pp. 1–6, doi: 10.1109/CISTI.2014.6877022.

[7] F. Qin and L. Wang, "Multiscale remote sensing image fusion algorithm based on variational segmentation," *Acta Electronica Sinica*, vol. 48, no. 6, pp. 1084–1090, 2020, doi: 10.3969/j.issn.0372-2112.2020.06.006.

[8] R. G. Tambe, S. N. Talbar, and S. S. Chavan, "Satellite image fusion using undecimated rotated wavelet transform," *Int. J. Comput. Sci. Eng.*, vol. 24, no. 2, pp. 171–184, 2021, doi: 10.1504/IJCSE.2021.115103.

[9] Z. Chi, "Research on satellite remote sensing image fusion algorithm based on compression perception theory," *J. Comput. Methods Sci.*, vol. 21, no. 2, pp. 341–356, 2021, doi: 10.3233/JCM-204411.

[10] F. Li, S. Hong, and L. Wang, "A new satellite image fusion method based on distributed compressed sensing," in *Proc. IEEE 25th Int. Conf. Image Process.*, 2018, pp. 1882–1886, doi: 10.1109/ICIP.2018.8451648.

[11] B. Sathyabama, S. G. S. Sankari, and S. Nayagara, "Fusion of satellite images using compressive sampling matching pursuit (CoSaMP) method," in *Proc. 4th Nat. Conf. Comput. Vis., Pattern Recognit., Image Process. Graph.*, 2013 pp. 1–4, doi: 10.1109/NCVPRIPG.2013.6776256.

[12] S. Saeed and M. Xavier, "Satellite image fusion by using a combination of IHS and HPM methods," *Proc. Int. Soc. Opt. Eng.*, vol. 10214, 2017, Art. no. 102141L, doi: 10.1117/12.2266038.

[13] M. Ghadjati, A. Benazza-Benyahia, and A. Moussaoui, "Satellite image fusion using an iterative IHS-based approach," in *Proc. Mediterranean Middle-East Geosci. Remote Sens. Symp.*, 2020, pp. 133–136, doi: 10.1109/M2GARSS47143.2020.9105197.

[14] Y. Tian, J. Wu, and X. Tian, "Satellite image fusion using RPCA combined IHS transform. Optics and precision engineering," *Opt. Precis. Eng.*, vol. 23, pp. 504–508, 2015, doi: 10.3788/OPE.20152313.0504.

[15] F. Cheng et al., "Non-subsampled Shearlet transform remote sensing image fusion combined with parameter-adaptive PCNN," *Acta Geodaetica et Cartographica Sinica*, vol. 50, no. 10, pp. 1380–1389, 2021, doi: 10.11947/j.AGCS.2021.20200589.

[16] X. Jin et al., "Remote sensing image fusion method in CIELab color space using nonsubsampled Shearlet transform and pulse coupled neural networks," *J. Appl. Remote Sens.*, vol. 10, no. 2, 2016, Art. no. 025023, doi: 10.1117/1.JRS.10.025023.

[17] K. Zhang, Y. Huang, and C. Zhao, "Remote sensing image fusion via RPCA and adaptive PCNN in NSST domain," *Int. J. Wavelets, Multiresolution Inf. Process.*, vol. 16, no. 5, 2018, Art. no. 1850037, doi: 10.1142/S0219691318500376.

[18] W. Li, C. Yang, Y. Peng, and X. Zhang, "A multi-cooperative deep convolutional neural network for spatiotemporal satellite image fusion," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 10174–10188, 2021, doi: 10.1109/JSTARS.2021.3113163.

[19] H. Song, Q. Liu, G. Wang, R. Hang, and B. Huang, "Spatiotemporal satellite image fusion using deep convolutional neural networks," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 3, pp. 821–829, Mar. 2018, doi: 10.1109/JSTARS.2018.2797894.

[20] X. Luo, X. Tong, and Z. Hu, "Improving satellite image fusion via generative adversarial training," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 8, pp. 6969–6982, Aug. 2021, doi: 10.1109/TGRS.2020.3025821.

[21] R. Eckhorn, H. J. Reitbock, M. Arndt, and P. Dicke, "A neural network for feature linking via synchronous activity: Results from cat visual cortex and from simulations," *Can. J. Microbiol.*, vol. 46, no. 8, pp. 759–763, 1989.

[22] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *Proc. 37th Asilomar Conf. Signals, Syst. Comput.*, 2003, pp. 1398–1402, doi: 10.1109/ACSSC.2003.1292216.

[23] M. B. A. Haghighat, A. Aghagolzadeh, and H. Seyedarabi, "A non-reference image fusion metric based on mutual information of image features," *Comput. Elect. Eng.*, vol. 37, no. 5, pp. 744–756, 2011, doi: 10.1016/j.compeleceng.2011.07.012.

[24] Y.-J. Rao, "In-fibre Bragg grating sensors," *Meas. Sci. Technol.*, vol. 8, no. 4, 1997, Art. no. 355, doi: 10.1088/0957-0233/8/4/002.

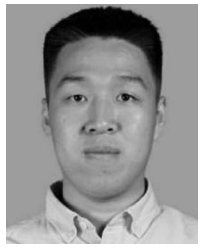
[25] B. S. Kumar, "Multifocus and multispectral image fusion based on pixel significance using discrete cosine harmonic wavelet transform," *Signal Image Video Process.*, vol. 7, no. 6, pp. 1125–1143, 2013, doi: 10.1007/s11760-012-0361-x.

[26] G. Piella and H. Heijmans, "A new quality metric for image fusion," in *Proc. Int. Conf. Image Process.*, 2003, pp. 173–176, doi: 10.1109/ICIP.2003.1247209.
[27] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004, doi: 10.1109/TIP.2003.819861.
[28] H. Zhang and J. Ma, "SDNet: A versatile squeeze-and-decomposition network for real-time image fusion," *Int. J. Comput. Vis.*, vol. 129, no. 10, pp. 2761–2785, 2021, doi: 10.1007/s11263-021-01501-8.
[29] H. Li, X. Wu, and J. Kittler, "RFN-Nest: An end-to-end residual fusion network for infrared and visible images," *Inf. Fusion*, vol. 73, pp. 72–86, 2021, doi: 10.1016/j.inffus.2021.02.023.
[30] L. Jian, X. Yang, Z. Liu, G. Jeon, M. Gao, and D. Chisholm, "SEDRFuse: A symmetric encoder–Decoder with residual block network for infrared and visible image fusion," *IEEE Trans. Instrum. Meas.*, vol. 70, 2021, Art. no. 5002215, doi: 10.1109/TIM.2020.3022438.
[31] H. Jung, Y. Kim, H. Jang, N. Ha, and K. Sohn, "Unsupervised deep image fusion with structure tensor representations," *IEEE Trans. Image Process.*, vol. 29, pp. 3845–3858, 2020, doi: 10.1109/TIP.2020.2966075.
[32] Y. Zhang et al., "IFCNN: A general image fusion framework based on convolutional neural network," *Inf. Fusion*, vol. 54, pp. 99–118, 2020, doi: 10.1016/j.inffus.2019.07.011.

Chang-Jiang Zhang was born in Heilongjiang Province, China, in 1974. He received the Ph.D. degree in control theory and control engineering from the Beijing Institute of Technology, Beijing, China, in 2004.

He was the middle-aged Discipline Leader of Zhejiang Normal University, China, in 2010. From 2012 to 2013, he was a Research Fellow with Cornell University. From 2009 to 2021, he was a Professor with the Electronic and Information Engineering Department, College of Physics and Electronic Information, Zhejiang Normal University. He is currently a Professor with the Electronic and Information Engineering Department, School of Electronic and Information Engineering (School of Big Data Science), Taizhou University, Taizhou, China, where he is also the Head of the Laboratory of Intelligent Vision and Meteorological Big Data. He is the author of 3 books, has authored or coauthored more than 100 articles, and has more than 20 inventions. He is the Principal Investigator of more than 20 foundations and projects. His research interests include image processing, machine learning, multiscale geometry analysis, medical informatics, and applied meteorology.

Dr. Zhang was a recipient of the First Prize, Second Prize, and Third Prize by the Education Department of Zhejiang Province, China. He was awarded as the Second Level of New Century 151 Talent Project of Zhejiang Province, China.



Jia-Xu Guo was born in Heilongjiang Province, China, in 1997. He is currently working toward the M.S. degree in electronics information with Zhejiang Normal University, Jinhua, China.

His research interests include image processing and deep learning.

Lei-Ming Ma was born in China in 1975. He received the Ph.D. degree in meteorology from Nanjing University, Nanjing, China, in 2009.

He is currently the Director, a Research Fellow, and a Professor with Shanghai Central Meteorological Observatory, Shanghai, China. He is a committee member of Numerical Forecasting Committee, China Meteorological Society. He is an author of more than 20 articles. He is the Principal Investigator of more than ten foundations and projects. His research interests include numerical forecasting theory and methods for tropical cyclone.

Dr. Ma was a recipient of Science and Technology Progress Award of Shanghai City, National Ocean Innovation Award of China, and National Outstanding Young Meteorologist of China.

Xiao-Qin Lu was born in China in 1975. She received the M.S. degree in meteorology from East China Normal University, Shanghai, China, in 2003.

She is currently a Research Fellow with Shanghai Typhoon Institute, China Meteorological Administration, Shanghai, China. Her research interests include research on objective analysis technology of tropical cyclone structure and application of tropical cyclone multisource data mining.

Wen-Cai Liu was born in Shandong Province, China, in 1998. She is currently working toward the M.S. degree in electronics information with Zhejiang Normal University, Jinhua, China.

Her research interests include polar code and image processing.