



Article

Spectral Reconstruction from Thermal Infrared Multispectral Image Using Convolutional Neural Network and Transformer Joint Network

Enyu Zhao ¹, Nianxin Qu ¹, Yulei Wang ^{1,*}  and Caixia Gao ²

¹ Information Science and Technology College, Dalian Maritime University, Dalian 116026, China; zhaoenyu@dlmu.edu.cn (E.Z.); qvnx@dlmu.edu.cn (N.Q.)

² Key Laboratory of Quantitative Remote Sensing Information Technology, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China; gaocaixia@aoe.ac.cn

* Correspondence: wangyulei@dlmu.edu.cn

Abstract: Thermal infrared remotely sensed data, by capturing the thermal radiation characteristics emitted by the Earth's surface, plays a pivotal role in various domains, such as environmental monitoring, resource exploration, agricultural assessment, and disaster early warning. However, the acquisition of thermal infrared hyperspectral remotely sensed imagery necessitates more complex and higher-precision sensors, which in turn leads to higher research and operational costs. In this study, a novel Convolutional Neural Network (CNN)–Transformer combined block, termed CTBNet, is proposed to address the challenge of thermal infrared multispectral image spectral reconstruction. Specifically, the CTBNet comprises blocks that integrate CNN and Transformer technologies (CTB). Within these CTBs, an improved self-attention mechanism is introduced, which not only considers features across spatial and spectral dimensions concurrently, but also explicitly extracts incremental features from each channel. Compared to other algorithms, the proposed method more closely aligns with the true spectral curves in the reconstruction of hyperspectral images across the spectral dimension. Through a series of experiments, this approach has been proven to ensure robustness and generalizability, outperforming some state-of-the-art algorithms across various metrics.

Keywords: spectral reconstruction; thermal infrared; hyperspectral; Transformer



Citation: Zhao, E.; Qu, N.; Wang, Y.; Gao, C. Spectral Reconstruction from Thermal Infrared Multispectral Image Using Convolutional Neural Network and Transformer Joint Network.

Remote Sens. **2024**, *16*, 1284.

<https://doi.org/10.3390/rs16071284>

Academic Editor: Kevin Tansey

Received: 5 March 2024

Revised: 28 March 2024

Accepted: 3 April 2024

Published: 5 April 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Thermal infrared remotely sensed technology is a method that utilizes spectral information from the thermal infrared band to observe the Earth's surface [1,2]. By capturing the thermal infrared radiation emitted from the Earth's surface, it gathers information about the composition and state of surface materials, having advantages, such as nighttime observation and rich spectral information, thus contributing to the development of Earth sciences [3]. Over the past few decades, hyperspectral remote sensing has made significant progress and has played a crucial role in various fields, such as environmental monitoring [4–6], agriculture [7–9], geological surveys [10–12], and target detection [13–15]. However, it also faces limitations, such as lower spatial resolution [16] and higher development and maintenance costs [17], which restrict its application in detailed or small-scale feature analysis for land classification and target recognition. Therefore, expanding the acquisition methods for thermal infrared hyperspectral imagery is necessary.

Recently, a category of algorithms, commonly referred to as “spectral reconstruction” or “spectral super-resolution” has been capable of deriving the corresponding hyperspectral images from multispectral images through a series of computations. Specifically, the multispectral image contains the features of the entire spectrum, although there is only one radiance value in a certain spectral range, while a hyperspectral image can acquire the radiance values of many narrow bands in the same spectral range. From an information

point of view, both represent the features of the same spectrum, which provides theoretical support for realizing spectral reconstruction.

In the past decade, due to the rapid development and application of machine learning in the field of imagery and remote sensing [18], many algorithms have been proposed to reconstruct visible hyperspectral images from RGB images [19]. While spectral reconstruction methods for visible light images have been proven feasible and achieved significant success, these methods cannot be directly applied to the thermal infrared domain. Firstly, visible light and thermal infrared spectra occupy different regions of the electromagnetic spectrum, covering wavelength ranges of approximately 400–700 nm and 3–14 μm , respectively. This wavelength difference results in significant variations in their absorption and scattering properties of materials. In the thermal infrared region, the radiation received by the sensor is mainly thermal radiation emitted by the substance, and its magnitude is related to the emissivity and temperature [20], while in the visible region, the radiation received by the sensor originates more from the reflection of the substance, and its magnitude is related to the observation geometry and reflectivity [21]. Secondly, the types of sensors used for visible light and thermal infrared imaging differ, leading to variations in their response characteristics and noise levels. This means that algorithms designed for visible light may not effectively handle specific noise or signal features in thermal infrared images.

To address the aforementioned challenges, this study introduces a supervised thermal infrared spectral reconstruction method, a Convolutional Neural Network (CNN)–Transformer combined block, named CTBNet, to compute the features in both spectral and spatial dimensions. This algorithm can reconstruct thermal infrared hyperspectral images in the range of 8.061 μm to 11.217 μm from multispectral thermal infrared images spanning 8.0 μm to 11.4 μm . Specifically, the novel block is designed to incorporate an improved self-attention mechanism utilizing multiscale convolutional operations in spectral and spatial dimensions, and this block is named CTB. The input data are first subjected to linear mapping to expand the spectral dimensions, followed by processing through several CTBs, with each block computing features in both the spatial and spectral dimensions. Experimentally, this model achieved commendable results. The main contributions of this paper are as follows:

1. In this study, a supervised deep learning algorithm is proposed to fill the gap in the spectral reconstruction of thermal infrared images. It overcomes the challenges associated with the acquisition of thermal infrared hyperspectral images and their low spatial resolution, and it can provide data support for other related studies in thermal infrared hyperspectral remote sensing.
2. This study introduces a CTB module, which incorporates an enhanced self-attention mechanism focusing on the spatial local features and the variation trends of spectral curves. This improvement significantly enhances the performance of thermal infrared spectral reconstruction. Experiments demonstrate that CTBNet possesses good robustness and stability for dealing with data noise and applications involving different sensors.

The structure of the article is as follows: Section 1 covers thermal infrared hyperspectral technology and spectral reconstruction algorithms; Section 2 introduces the related works; a spectral reconstruction algorithm for thermal infrared multispectral images is detailed in Section 3; the construction of the dataset is shown in Section 4; results are presented in Section 5; Section 6 discusses the details on model improvements and robustness; the conclusions are displayed in the final section.

2. Related Work

2.1. HSI Reconstruction

In past studies, spectral reconstruction algorithms can generally be divided into two categories: traditional algorithms that utilize prior knowledge, and data-driven deep learning algorithms [22–24]. Traditional methods rely on prior knowledge, including spectral correlation, sparsity, and spatial similarity. Fu et al. [25] introduced a method based on

non-negative sparse representation, constructing a sparse coding dictionary to reconstruct HSI from RGB images, and subsequently proposed several improved algorithms. Fotiadou et al. [26] proposed a coupled dictionary learning model considering spatial features, achieving commendable results. Recently, data-driven deep learning methods have seen exponential development in the field of image processing, with many deep learning-based HSI reconstruction algorithms being proposed. Zhao et al. [27] combined the HSCNN-R model to reconstruct hyperspectral images from a single RGB image, applying it to assess the quality parameters of tomatoes. Miao et al. [28] designed a two-stage generative network by integrating U-net into a self-attention GAN framework to achieve HSI reconstruction. In addition, in the spectral reconstruction of satellite images, Zhu et al. [29] proposed a self-supervised algorithm that employs a spectral dimension masking autoencoder architecture combined with random masking pretraining and fixed masking fine-tuning strategies, achieving spectral reconstruction of satellite images. Furthermore, Han et al. [30] proposed a spectral reconstruction method using the cluster-based multibranch backpropagation neural network and applied this to the HJ-1A satellite. Overall, data-driven deep learning algorithms have gradually begun to play a role in the HSI reconstruction field and hold great potential.

2.2. Transformer

The Transformer model is a widely utilized deep learning architecture in the domain of Natural Language Processing (NLP), introduced by Vaswani et al. [31,32]. The heart of the Transformer is the self-attention mechanism, which allows the model to consider all other words in the sequence while processing each individual word, significantly enhancing the model's ability to handle long-range dependencies. Subsequent research has demonstrated the Transformer's exceptional performance not only within the realm of NLP but also in image processing. Liu et al. [33] proposed the Swin Transformer algorithm and extensively applied it to the field of computer vision. Furthermore, Du et al. [34] designed a Transformer model integrated with convolutional operations, achieving the reconstruction of remote sensing images. However, these methods have been primarily applied to images within the visible light spectrum. Due to factors, such as imaging mechanisms, their performance degrades when tasked with the reconstruction of thermal infrared remote sensing images.

2.3. CNN–Transformer

The CNN–Transformer model, as a novel hybrid paradigm in recent years, has demonstrated excellent performance in the domain of computer vision. It melds the local perception prowess of CNN with the global associative processing strength of Transformers, endowing CNN–Transformer models with potent capabilities and versatility. DETR [35], epitomizing the CNN–Transformer models, utilizes CNN for feature extraction and leverages the encoder–decoder mechanism of Transformers to directly predict the category and bounding box of each object in images in an end-to-end manner. This synergy not only boosts the efficiency of object detection but also enhances the performance of the algorithm. With the advancement of CNN–Transformer technology, an increasing number of refined algorithms have emerged. Chen et al. [36] introduced an efficient dual-pathway Transformer structure for building extraction, which reduces time costs while maintaining performance.

3. Methodology

3.1. Problem Formulation

The precondition for reconstructing multispectral thermal infrared images into thermal infrared hyperspectral images is that they inherently contain spectral information within that wavelength range. Specifically, for a given multispectral remote sensing band, its radiance can be represented as follows:

$$R(\text{band } i) = \frac{\int_{\lambda_1}^{\lambda_2} R(\lambda) \text{SRF}(\lambda) d(\lambda)}{\int_{\lambda_1}^{\lambda_2} \text{SRF}(\lambda) d(\lambda)} \quad (1)$$

where $R(\text{band}i)$ denotes the radiance of band i , $R(\lambda)$ represents the radiance at wavelength λ , $SRF(\lambda)$ signifies the response value of band i at wavelength λ , λ_1 is the starting wavelength of band i , and λ_2 is the ending wavelength of band i . As illustrated in Equation (1), the radiance of a specific band in a multispectral remote sensing image is derived from the integration of spectral radiance within a specific range, encompassing all the information within that spectral range. This facilitates the possibility of band reconstruction for thermal infrared hyperspectral remotely sensed images. Since thermal infrared remotely sensed images primarily reflect the thermal radiation of the Earth's surface and atmosphere, their radiances are coupled with the information of the land surface and atmosphere, as the following formula indicates:

$$B(T) = \varepsilon B(T_s)\tau + R_{atm}^{\uparrow} + (1 - \varepsilon)R_{atm}^{\downarrow}\tau \quad (2)$$

where $B(T)$ denotes the radiance received by the sensor, B represents the Planck function, ε is the surface emissivity, τ stands for the atmospheric transmittance, and T_s signifies the land surface temperature, while R_{atm}^{\uparrow} and R_{atm}^{\downarrow} correspond to the upwelling and downwelling atmospheric radiation, respectively.

3.2. Network Architecture

The overall architecture of the network, as depicted in Figure 1, is primarily composed of five CTBs. Specifically, the input of the network is a multispectral image with dimensions of 100 pixels in width and height and a channel count of 4. It is worth noting that this study only uses 4-channel input data as an example, and the algorithm does not limit the number of channels of input data. Initially, this input undergoes a linear mapping process to expand its channel count to 110. The expanded data are sequentially input into a type of CTB which has been specifically designed for the spectral reconstruction. The output generated by the final CTB is a hyperspectral cube with dimensions of 100 pixels in both width and height, and it comprises 110 spectral channels.

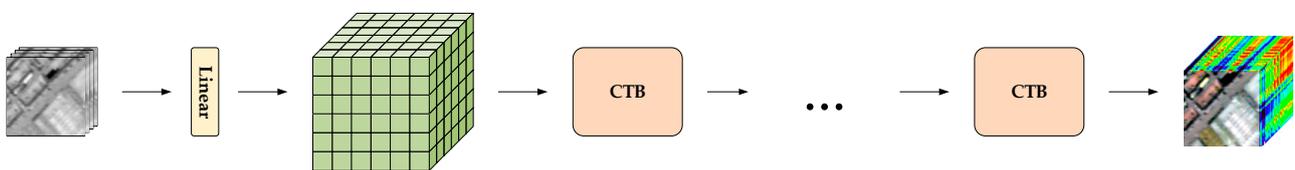


Figure 1. CTBNet structure.

3.3. CTB Structure

An enhanced self-attention mechanism has been proposed and applied within the CTB, differing from the traditional approach of employing linear mappings to obtain \mathbf{Q} , \mathbf{K} , and \mathbf{V} . CTB utilizes two-dimensional and one-dimensional convolutions to derive \mathbf{K} and \mathbf{V} , respectively, and employs convolutions to construct a Feedforward Neural Network (FNN), as shown in Figure 2. The CTB is an essential module in CTBNet, comprising an enhanced self-attention mechanism, an FNN, and a spectral forget unit. In brief, in Figure 2, FNN serves to enhance the model's fitting ability; K Map indicates the key mapping method within the improved self-attention mechanism, employing two-dimensional multiscale convolution to extract features in the spatial dimension; V Map denotes the value mapping in the improved self-attention mechanism, which involves the computation of spectral increments and one-dimensional multiscale convolution in the spectral dimension.

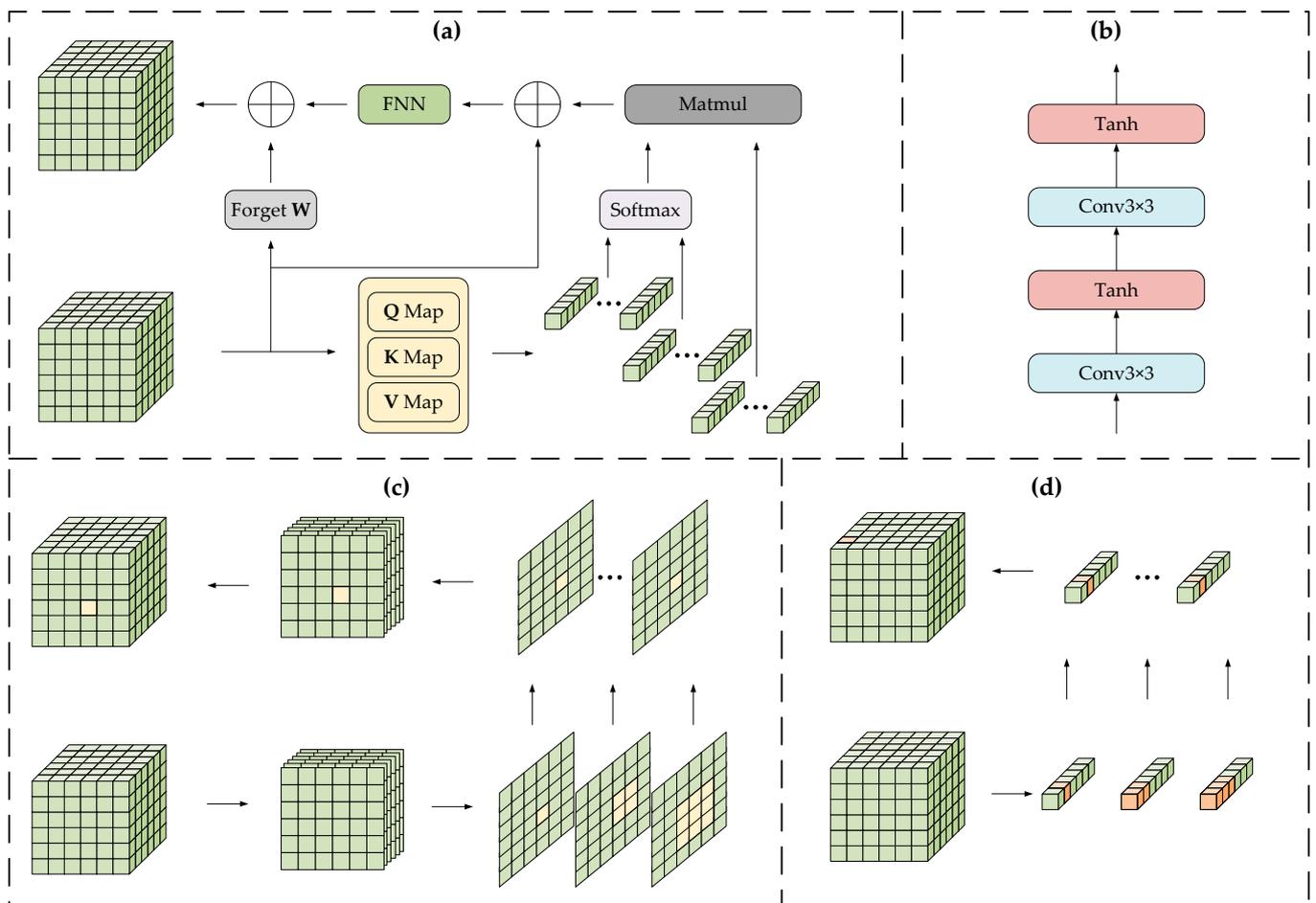


Figure 2. CTB structure. (a) overall structure of CTB; (b) composition of FNN; (c) process of K Map; (d) process of V Map.

Specifically, in the original self-attention mechanism, \mathbf{Q} , \mathbf{K} , and \mathbf{V} are obtained through linear mappings of the input features, as shown in the following equation:

$$\mathbf{Q} = \mathbf{X}\mathbf{W}^{\mathbf{Q}}; \mathbf{K} = \mathbf{X}\mathbf{W}^{\mathbf{K}}; \mathbf{V} = \mathbf{X}\mathbf{W}^{\mathbf{V}} \quad (3)$$

where \mathbf{X} denotes the input features, while $\mathbf{W}^{\mathbf{Q}}$, $\mathbf{W}^{\mathbf{K}}$, and $\mathbf{W}^{\mathbf{V}}$ represent the learnable parameter matrices. The method of self-attention based on linear mappings exhibits superior performance in extracting global features. However, this approach may overlook local characteristics within an image, such as textural details. To thoroughly extract both global and local information from thermal infrared hyperspectral images, a multi-scale convolution-based K mapping has been proposed, which is formulated as follows:

$$\mathbf{K} = \tanh(\text{Conv}2\text{d}_{1 \times 1}(\mathbf{X}) + \text{Conv}2\text{d}_{3 \times 3}(\mathbf{X}) + \text{Conv}2\text{d}_{5 \times 5}(\mathbf{X})) \quad (4)$$

where, $\tanh(\cdot)$ represents the activation function, which is a hyperbolic tangent function with a mean of zero. Its advantage is that it can accelerate model convergence and introduce nonlinear relationships. $\text{Conv}2\text{d}_{1 \times 1}(\cdot)$, $\text{Conv}2\text{d}_{3 \times 3}(\cdot)$, and $\text{Conv}2\text{d}_{5 \times 5}(\cdot)$ represent two-dimensional convolutions with kernel sizes of 1×1 , 3×3 and 5×5 , respectively, as depicted in the K Map section of Figure 2. The yellow portions in the figure illustrate the multi-scale convolution operations. During this process, the three different sizes of convolution operations individually compute local feature information of varying receptive field sizes and then aggregate these features, thereby enhancing the model's capacity to represent spatial information. Furthermore, replacing linear mappings with multi-scale

convolutions to generate \mathbf{K} can reduce the number of parameters in the model. This reduction aids in model convergence and can decrease computational overhead.

In the CTB algorithm, improvements have been made not only to the method of extracting spatial features from thermal infrared hyperspectral images but also to the strategy for extracting spectral features. Specifically, a thermal infrared hyperspectral curve possesses two critical attributes: the radiance value of each channel in the spectrum and the overall trend of the spectral curve. The former reflects the energy emitted or reflected by an object at specific wavelengths, while the latter illustrates how radiance values change with wavelength, encapsulating the substance and state information of the target pixel. Consequently, a \mathbf{V} mapping, focused on the spectral trend, is employed within the self-attention mechanism. In essence, the spectral variation trend is initially represented by calculating the spectral increment between adjacent channels in the hyperspectral data, followed by the extraction of features using multi-scale spectral-dimension convolutions. The computation of spectral increment is as follows:

$$\Delta x_i = x_i - x_{i+1}, x_i \in \mathbf{X}, \Delta x_i \in \Delta \mathbf{X} \quad (5)$$

where x_i represents the feature value of band i for the input data, Δx_i is the difference between the band i and band $i + 1$, and $\Delta \mathbf{X}$ represents the trend of channel changes across the entire spectral dimension. Based on this, \mathbf{V} can be represented as follows:

$$\mathbf{V} = \tanh(\text{Conv1d}_1(\mathbf{X}) + \text{Conv1d}_3(\mathbf{X}) + \text{Conv1d}_5(\mathbf{X})) \quad (6)$$

where $\text{Conv1d}_1(\cdot)$, $\text{Conv1d}_3(\cdot)$, and $\text{Conv1d}_5(\cdot)$ represent one-dimensional convolutions with kernel sizes of 1, 3, and 5, respectively, as shown in the orange section of the \mathbf{V} Map in Figure 2. These convolutions are performed only along the spectral dimension. Combined with spectral increment information that encapsulates spectral trends, this method can more accurately fit the spectral curve features, effectively avoiding the spectral smoothing issue in reconstructed images.

In the previous discussion, it was explained how multi-scale 2D convolutions are utilized to obtain \mathbf{K} , which focuses on local spatial features, and how spectral increments are used to derive \mathbf{V} , which focuses on spectral trends. Regarding the acquisition of \mathbf{Q} , a linear mapping incorporating global features is used, as indicated in Equation (3). After obtaining \mathbf{Q} , \mathbf{K} , and \mathbf{V} , the attention score matrix can be represented as follows:

$$\text{AttentionScore}^h = \frac{\mathbf{Q}^h (\mathbf{K}^h)^T}{\sqrt{d_{\mathbf{K}}}} \quad (7)$$

where \mathbf{Q}^h , \mathbf{K}^h , and \mathbf{V}^h are the query, key, and value for head h , respectively, and $d_{\mathbf{K}}$ represents the dimensionality of the key. The transposition of \mathbf{K}^h is denoted as $(\mathbf{K}^h)^T$. After the attention scores are computed, the output can be represented as follows:

$$\text{HeadOutput}^h = \text{softmax}(\text{AttentionScore}^h) \mathbf{V}^h \quad (8)$$

where $\text{softmax}(\cdot)$ refers to an activation function, which serves to normalize the attention score matrix. Finally, the outputs of all attention heads are concatenated, and the formula for this operation is as follows:

$$\text{MultiHeadOutput} = \text{Concat}(\text{HeadOutput}^1, \text{HeadOutput}^2, \dots, \text{HeadOutput}^H) \mathbf{W}^O \quad (9)$$

where H is the total number of heads, $\text{Concat}(\cdot)$ denotes splicing the output of all heads in some dimension, and \mathbf{W}^O denotes a linear mapping of the results of multi-head attention.

Overall, the data input to the CTB, denoted as $\mathbf{X} \in \mathbb{R}^{H \times W \times C}$, is transformed into \mathbf{Q} , \mathbf{K} , $\mathbf{V} \in \mathbb{R}^{HW \times C}$ after undergoing various mappings. Subsequently, the output of the

multi-head attention is computed and fed into the FNN. This output is then connected to the input of the CTB via a residual link.

In the residual connections, a spectral forget unit is included, which determines what information should be forgotten or discarded from the input of the CTB. This mechanism helps prevent unnecessary accumulation of information, thus avoiding the indefinite influence of past information on future states. The formula for the spectral forget unit is as follows:

$$\mathbf{X}_{\text{forget}} = \text{Concat}(\mathbf{X}_1 \cdot \mathbf{W}, \mathbf{X}_2 \cdot \mathbf{W}, \dots, \mathbf{X}_{H \times W} \cdot \mathbf{W}) \quad (10)$$

where \mathbf{W} is a learnable weight matrix of size $(1, 110)$. which, during computation, engages in an inner product with the spectral dimension of the data to yield weighted spectral data.

In essence, the CTB features two network branches. The first employs an enhanced self-attention mechanism to extract features from global, local spatial, and spectral trend perspectives, subsequently applying an FNN for further non-linear enhancement of the data. In the second network branch, the spectral forget unit processes the CTB input data to facilitate forgetting and discarding of information, which is then merged with the output from the attention branch through residual connections. The initial thermal infrared hyperspectral images undergo linear mapping to produce relatively coarse hyperspectral images, which are then sequentially refined and reconstructed using multiple CTBs, culminating in the acquisition of accurate thermal infrared hyperspectral images.

4. Dataset

Hyperspectral imagery obtained through in situ measurements was utilized in this study. This imagery was captured in January 2021, located at a longitude of 120.28 degrees East and a latitude of 29.16 degrees North, within Zhejiang Province, China. The spatial resolution of the hyperspectral image is 1 m, with a spectral range spanning from 8.061 to 11.217 μm . The image dimensions comprise 8000 pixels in length and 300 pixels in width, encompassing a total of 110 spectral channels. Included within the image are various surface features, such as lakes, man-made structures, and vegetation, as illustrated in Figure 3.

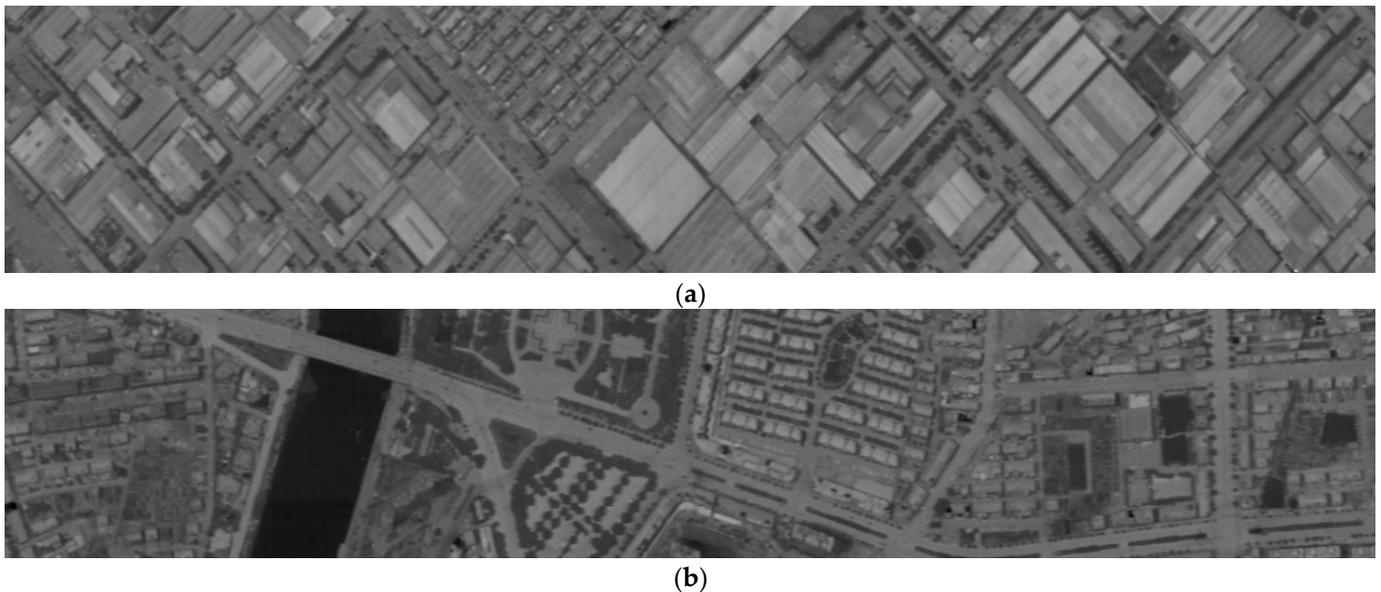


Figure 3. A portion of the hyperspectral data: (a) including various man-made structures; (b) encompassing rivers, vegetation, highways, and other elements.

Due to the absence of concurrently measured thermal infrared multispectral and hyperspectral remotely sensed data, in this part, simulated multispectral spectral response functions and multispectral data are produced based on the Zhejiang thermal infrared hyperspectral image. The spectral response functions partition the wavelength range of the hyperspectral data into four equal intervals, representing four channels. The central wavelengths for these channels are, respectively, 8.5 μm , 9.3 μm , 10.1 μm , and 10.9 μm , with each channel having a bandwidth of 1 micrometer, as illustrated in Figure 4.

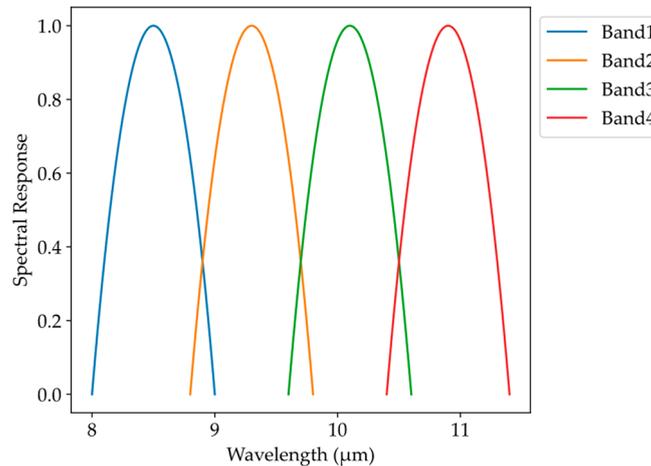


Figure 4. Simulated spectral response functions.

After the simulation, the original image with dimensions of 300 by 8000 pixels was cropped with overlap to produce a total of 4356 images, each with dimensions of 100 by 100 pixels. Ninety percent of these images were utilized to compile the training dataset, while the remaining ten percent of images were allocated for the creation of the testing dataset.

5. Results

In all experiments, CTBNet was configured with the following settings: the network comprises five blocks, and each one employs an improved attention mechanism with four attention heads. During the training phase, the batch size is set to 2, and the learning rate is initialized at 0.0001, with a cosine annealing scheme for decay. The Adam optimizer is used for the optimization, with the number of training epochs set to be 100, and the mean relative absolute error (MRAE) is used as the loss function of the network. Based on these settings, the equipment used consists of a computer with an Intel Xeon(R) Platinum 8352V CPU and an NVIDIA RTX 4090 GPU, running on a Linux operating system and programmed in Python. The total runtime for the training process is approximately 10 h. To evaluate the effectiveness of the algorithm, five performance metrics are selected: peak signal-to-noise ratio (PSNR) in decibels (dB), structural similarity index (SSIM), spectral angle mapper (SAM) in radians, root mean square error (RMSE) in Kelvin (K), and MRAE.

To illustrate the thermodynamic characteristics of the reconstructed hyperspectral images, radiances are converted to brightness temperature, as shown in the following Equation (11):

$$M_{\lambda}(T) = \frac{2\pi hc}{\lambda^5} \frac{1}{e^{\frac{hc}{\lambda kT}} - 1} \quad (11)$$

where h represents the Planck constant and has a value of 6.626×10^{-34} J·s; k represents the Boltzmann's constant and has a value of 1.3806×10^{-23} J/K; c represents the speed of light and has a value of 2.998×10^8 m/s; λ represents the wavelength in meters; T represents the thermodynamic temperature in Kelvin.

To demonstrate the efficacy of CTBNet, it was quantitatively compared with several state-of-the-art (SOTA) methods, namely MST++ [37], HRNet [38], HSCNN+ [39], HDNet [40], HINet [41], and Restormer [42], as shown in Table 1. Notably, MST++ was the winning algorithm at NTIRE2022.

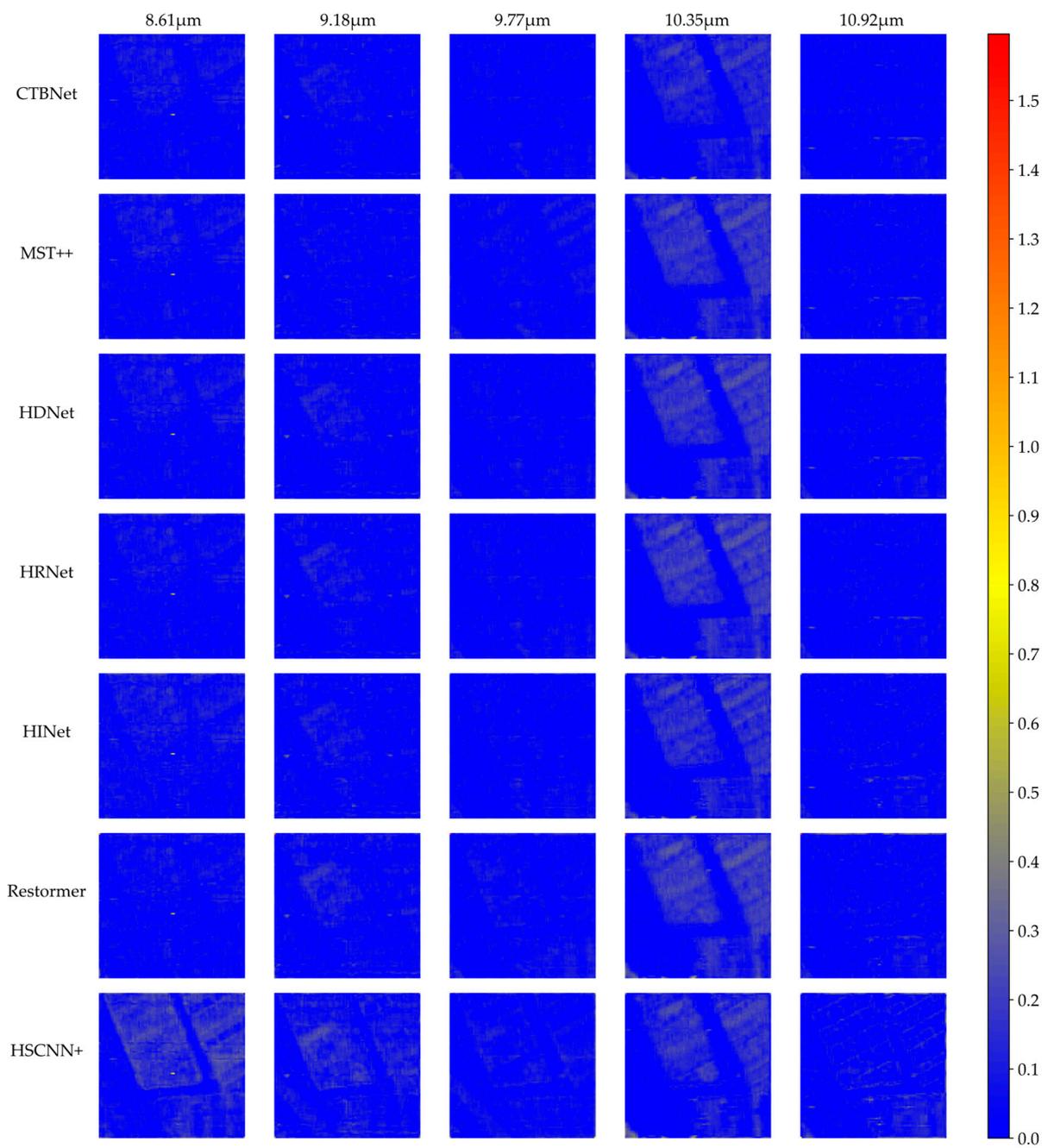
Table 1. Comparison of CTBNet with other methods.

	PSNR (dB)	SSIM	SAM (rad)	RMSE (K)	MRAE
MST++	60.93172	0.98797	0.00091	0.31442	0.00068
HRNet	61.37429	0.98842	0.00085	0.29734	0.00063
HSCNN+	59.1247	0.98611	0.00109	0.4063	0.00083
Restormer	60.14287	0.98824	0.00095	0.35053	0.00071
HDNet	61.058	0.98855	0.00088	0.30924	0.00065
HINet	60.97999	0.9879	0.00089	0.30969	0.00066
CTBNet	61.48676	0.98864	0.00084	0.29338	0.00062

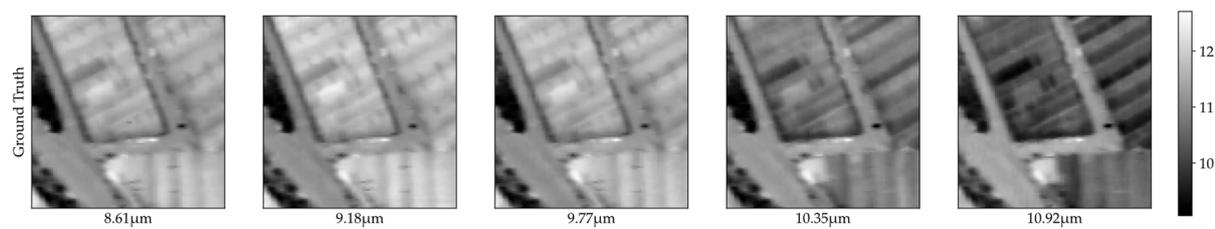
In Table 1, through the detailed comparative experiments, the performance of CTBNet and other SOTA methods on image quality assessment metrics is meticulously compared. From these metrics, a comprehensive understanding of the overall performance of each method is obtained. Firstly, in terms of the PSNR metric, CTBNet leads to a result of 61.48676 dB, indicating its superiority in reducing noise in the image reconstruction process. Secondly, the SSIM index, which measures the structural similarity between two images, also highlights CTBNet's advantage with a result of 0.98864. This indicates that CTBNet effectively preserves the structural features and details of images. On the SAM metric, CTBNet achieved a result of 0.00084 rad, reflecting the spectral similarity of the reconstructed image. A lower SAM value indicates that the spectral characteristics of the reconstructed image are closer to the original, thus proving CTBNet's excellent performance in spectral fidelity. For RMSE and MRAE, which directly measure the error in pixel values, CTBNet leads with results of 0.29338 K and 0.00062, respectively, further demonstrating its capability in error control.

Overall, the CTBNet method achieved the best results across all evaluation metrics, especially notable in the PSNR metric for image signal-to-noise ratio, where its advantage is most significant. Its performance in the RMSE and SSIM metrics also demonstrates outstanding capabilities, collectively indicating CTBNet's comprehensive superiority in image reconstruction quality, particularly in aspects, such as noise suppression, structural preservation, and error control.

It is well known in the field of thermal infrared remote sensing that the type of land cover significantly affects the thermal infrared radiance observations of pixels. This phenomenon is particularly evident between different types of land surfaces, such as land and water, due to significant differences in emissivity and temperature. Land surfaces, depending on their vegetation cover, soil moisture, and other surface characteristics, exhibit varied thermal infrared radiation properties. In contrast, water bodies, due to their inherent physical properties, typically show higher emissivity and more uniform temperature distribution. These differences not only manifest in the spectral dimension but also have a significant impact on the spatial dimension, thereby posing a higher challenge for spectral reconstruction algorithms due to the complexity of surface types. To further showcase the advantages of the algorithm, radiance error maps for two images of different land surface types, namely land and water, were selected for presentation, with the land area shown in Figure 5 and the water area in Figure 6.



(a)



(b)

Figure 5. Comparison of results for land region. (a) Error map for land region. (b) Truth map for the land region.

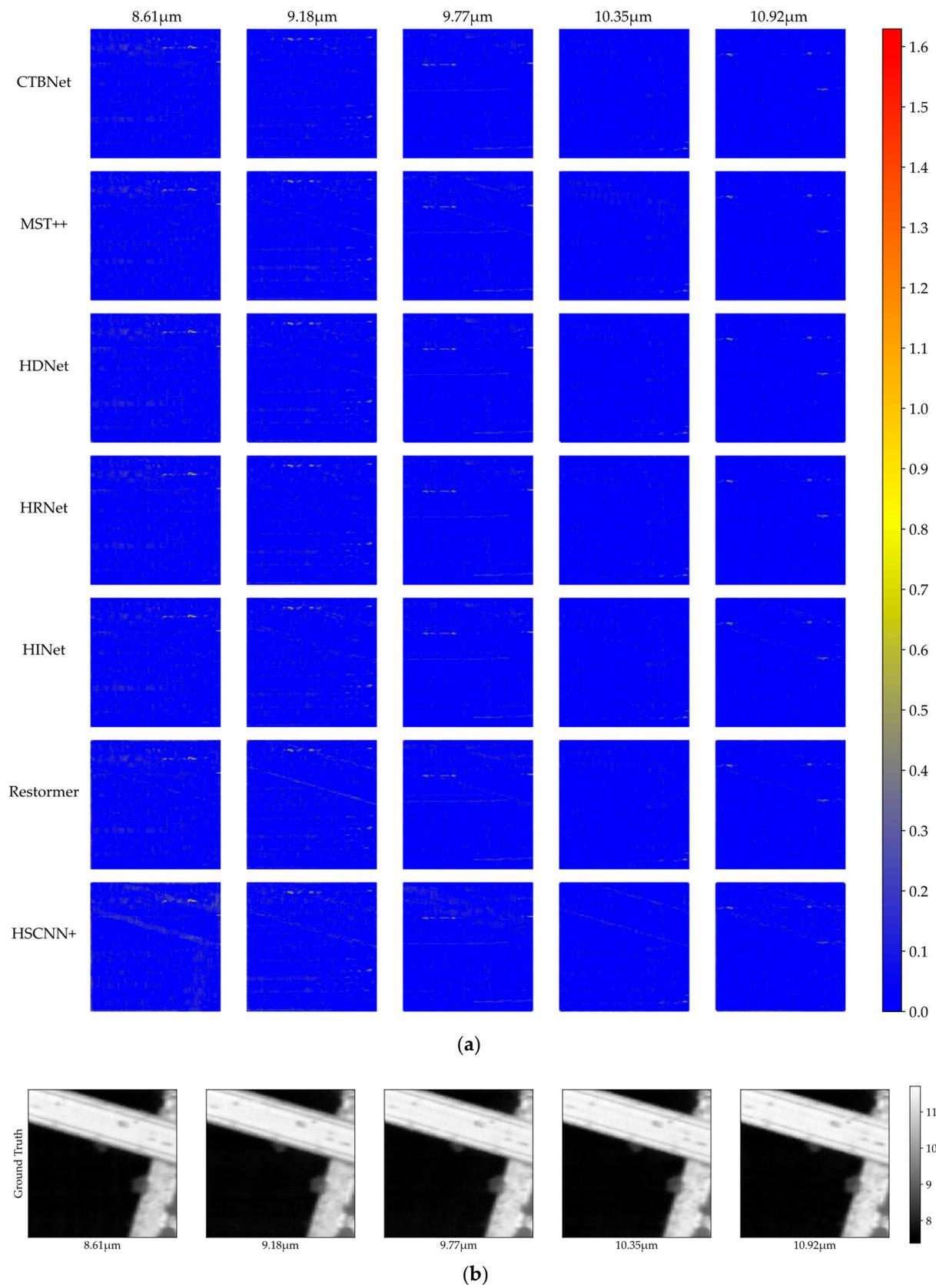


Figure 6. Comparison of results for water region. (a) Error map for the water region (b) True value for the water region.

From Figures 5 and 6, it is evident that the reconstruction errors of the thermal infrared hyperspectral images are more concentrated in areas with man-made structures, such as the buildings on the right side of Figure 5 and the bridge in Figure 6. Moreover, compared to water bodies, the errors in the land areas are more pronounced across various spectral reconstruction algorithms. This phenomenon is likely due to the greater diversity and complexity of land cover types relative to water bodies. The diversity of land surfaces includes different vegetation types, soil moisture levels, topographical features, and man-made structures, introducing a high degree of variability into the spectral data and consequently increasing the difficulty of spectral reconstruction. In this context, the CTBNet algorithm exhibits relatively smaller errors in both land and water areas, demonstrating its efficiency and applicability in spectral reconstruction tasks. This architecture is capable of better learning and simulating the complexity of land cover types and their impact on spectral data, thereby preserving more details and features in the reconstruction process.

To more intuitively assess the performance of the CTBNet algorithm in the spectral dimension, radiance spectral curve comparisons for two specific examples are presented, as shown in Figure 7.

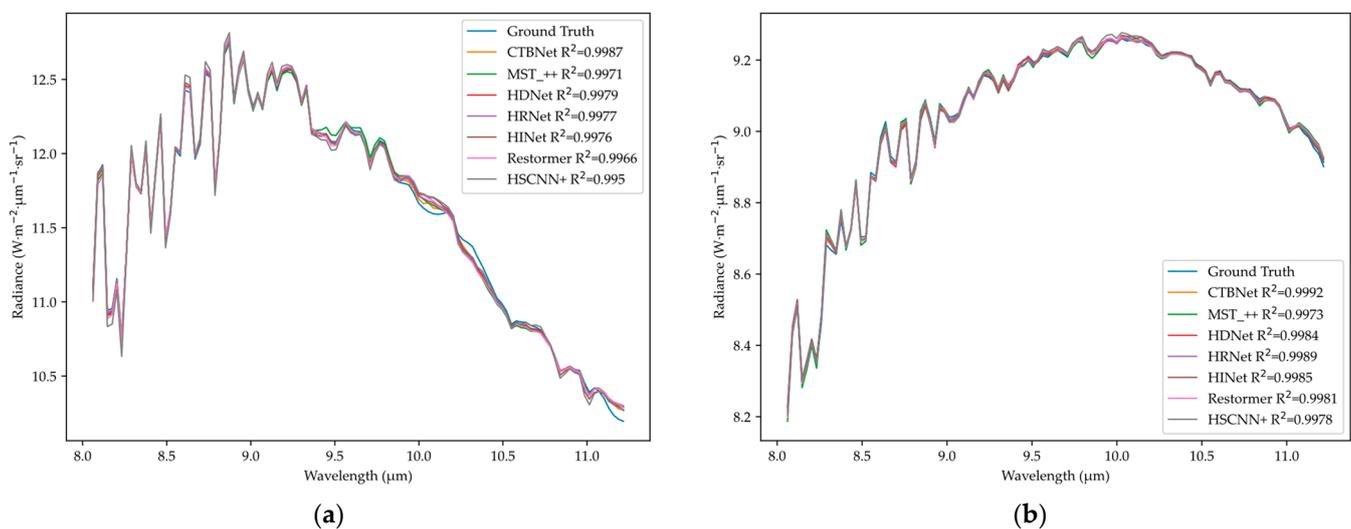


Figure 7. Radiance spectral curve. (a) for the land region and (b) for the water region.

From the comparison in Figure 7, the performance advantage demonstrated by the CTBNet algorithm can be clearly seen. Specifically, CTBNet achieves a higher R² value, a statistical metric that directly reflects the congruence between the algorithm's reconstructed spectral curve and the true data (Ground Truth). A higher R² indicates that CTBNet can more accurately simulate the real spectral characteristics, thereby providing more reliable reconstruction results across various test scenarios. Further observation of the spectral detail comparison reveals that the CTBNet-reconstructed spectral curve closely follows the trend of the real data across the entire wavelength range, including those spectral feature bands that are particularly critical for the identification of surface materials. This close match is evident not only in the major absorption and reflection peaks but also in the minor fluctuations of the spectral curve, which are crucial for precise material identification and classification.

Overall, CTBNet exhibits excellent performance in the task of thermal infrared spectral reconstruction. It not only precisely reconstructs spatial and spectral details but also effectively captures the trends of spectral variation, avoiding such issues as spectral smoothing.

6. Discussion

6.1. Ablation Study

To investigate the effectiveness of the improved self-attention mechanism within CTBNet, ablation experiments were designed in this part. Specifically, three variations were tested: a Transformer model using the original self-attention mechanism, named Nom-Transformer, a model employing the improved self-attention mechanism but without the spectral forget unit, named New-Attention, and a model incorporating the spectral forget unit but not the improved self-attention mechanism, named Nom-Attention. The results of the ablation experiments are shown in Table 2.

Table 2. Comparison of ablation experiments.

	PSNR (dB)	SSIM	SAM (rad)	RMSE (K)	MRAE
CTBNet	61.48676	0.98864	0.00084	0.29338	0.00062
New-Attention	59.93308	0.98728	0.00093	0.36361	0.00074
Nom-Attention	61.42158	0.98854	0.00085	0.29546	0.00063
Nom-Transformer	57.79278	0.98686	0.00098	0.46456	0.00109

From Table 2, it is evident that the CTBNet demonstrates the best performance across all evaluation metrics, with a PSNR value of 61.48676 dB and an SSIM value of 0.98864, indicating superior image reconstruction quality. The New-Attention model performs better than the Nom-Transformer on most metrics, suggesting that the improved self-attention mechanism positively impacts performance. The performance of the Nom-Attention model in terms of PSNR and SSIM is close to that of CTBNet but slightly inferior in SAM, RMSE, and MRAE metrics. This might indicate that while the spectral forget unit helps performance, its combination with the improved self-attention mechanism yields the best results. Overall, employing the improved self-attention mechanism and the spectral forget unit both resulted in better outcomes than the original Transformer model, with the combination of the two enhancements achieving superior image reconstruction results.

6.2. Sensitivity Analysis

6.2.1. Influence of Instrument Noise

To more closely align with real-world application scenarios, Gaussian noise was intentionally added to the dataset. The primary purpose of this modification is to simulate the impact of instrument noise on thermal infrared multispectral data. Instrument noise, an inevitable part of the actual measurement process, primarily arises from the electronic systems of instruments, instability in detector performance, and environmental factors, manifesting as random fluctuations within the data [43–45]. By introducing noise in the form of a Gaussian distribution to the simulated data, the aim is to assess and enhance the adaptability and accuracy of the reconstruction algorithm when faced with real measurement errors, thereby ensuring the reliability and robustness of the algorithm when processing real-world data. Gaussian noise at a level of 0.1% was added to the data, and the reconstruction results after incorporating this noise are shown in Table 3.

Table 3. Influence of instrument noise on results.

	PSNR (dB)	SSIM	SAM (rad)	RMSE (K)	MRAE
Without_noise	61.48676	0.98864	0.00084	0.29338	0.00062
With_noise	61.11164	0.98793	0.00088	0.30397	0.00066

The results presented in Table 3 reveal that the incorporation of Gaussian noise leads to a marginal reduction in various performance metrics, though the extent of this reduction remains limited. These outcomes affirm that, despite the challenging conditions introduced by instrument noise, the deep learning methodology proposed in this study is capable of conducting effective spectral reconstruction. It demonstrates commendable adaptability

and high precision, providing reliable technical support for the analysis and processing of spectral data in real-world application contexts.

6.2.2. Influence of Spectral Response Function

Given the differences in spectral response functions among various sensors, these discrepancies could significantly impact the quality of multispectral data reconstruction. To assess our algorithm's capability to adapt to different multispectral sensors and thereby verify its robustness, a sensitivity analysis experiment was specifically designed. This experiment aimed to simulate the spectral response functions of the thermal infrared bands of the MODIS sensor to explore how such differences affect the performance of the algorithm in reconstructing hyperspectral images. The simulated spectral response range is shown in Figure 8. Subsequently, multispectral data generated based on the simulated MODIS response functions were used as input for the algorithm, to test its adaptability and flexibility to differences in sensor spectral characteristics, with the results shown in Table 4.

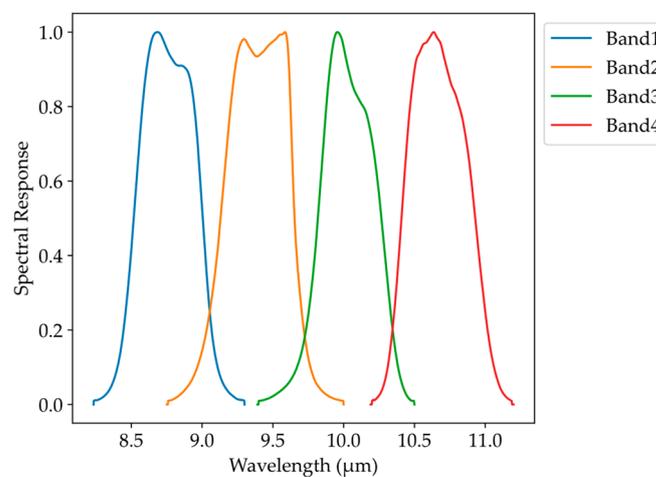


Figure 8. Spectral response function of simulated MODIS.

Table 4. Influence of simulated MODIS spectral response function on results.

	PSNR (dB)	SSIM	SAM (rad)	RMSE (K)	MRAE
Ideal data	61.48676	0.98864	0.00084	0.29338	0.00062
MODIS data	61.24813	0.98838	0.00089	0.31294	0.00065

The results in Table 4 indicate that multispectral data generated using the MODIS spectral response functions have a minor impact on the outcomes of spectral reconstruction, yet the performance remains commendable. This phenomenon may be attributed to the simulated MODIS spectral response functions being more complex and less smooth compared to the spectral response functions used in Section 4, which could affect the model's training and convergence. At the same time, since this impact is relatively minor, it also suggests that the algorithm possesses strong robustness in real-world applications with different types of multispectral sensors.

In this section, the various components of the CTBNet algorithm were explored, examining their roles in enhancing the reconstruction of thermal infrared multispectral images. Ablation studies confirmed the efficacy of the improved self-attention mechanism and spectral forget unit, providing valuable insights for subsequent thermal infrared hyperspectral image reconstruction efforts. Furthermore, two sensitivity analysis experiments demonstrated the robustness of the CTBNet algorithm against data noise and its adaptability to different sensor characteristics, thereby broadening the applicability of the algorithm and offering alternative solutions for the acquisition of thermal infrared hyperspectral data. For example, in current thermal infrared hyperspectral research, many algorithms are based

on the European Space Agency's IASI sensor [46–48], which has a relatively low spatial resolution and cannot meet the needs of small-scale studies. In future research, given adequate data support, this algorithm could be employed to reconstruct multispectral data with higher temporal and spatial resolution into hyperspectral data, thereby fulfilling research requirements.

7. Conclusions

This paper presents an improved spectral reconstruction algorithm, CTBNet, which combines CNNs and Transformers, aimed at addressing the challenges of thermal infrared multispectral image reconstruction. By incorporating an improved self-attention mechanism and spectral forget unit into the algorithm, it successfully enhances the accuracy of multispectral to hyperspectral image reconstruction while maintaining efficient computational performance. Experimental results demonstrate that CTBNet surpasses some SOTA methods across multiple performance evaluation metrics, particularly excelling in key indicators, such as PSNR, SSIM, SAM, RMSE, and MRAE. These achievements not only prove the immense potential of deep learning in the field of spectral reconstruction but also highlight the advantages of combining CNNs and Transformers in processing thermal infrared images. Furthermore, through ablation experiments and sensitivity analysis, the crucial roles of the improved self-attention mechanism and spectral forget unit in enhancing algorithm performance, as well as the algorithm's robustness to different noise conditions and spectral response function variances, were further verified.

Thermal infrared hyperspectral remote sensing technology is a crucial tool in the field of Earth sciences, and the success of CTBNet signifies that more thermal infrared hyperspectral data may be available for future scientific researches, thus contributing to the advancement of Earth sciences. For example, in environmental monitoring, the spectral reconstruction technique of CTBNet may be used for high temporal resolution monitoring of climate change, pollutant dispersion, and changes in wildlife habitats. In the agricultural sector, this technology may generate high spatial resolution thermal infrared hyperspectral images, which can then assist in monitoring crop health, assessing the impact of drought, and guiding irrigation management.

Despite the significant accomplishments of CTBNet in various aspects, we recognize that there is still some room for further optimization and application expansion of the algorithm. Future research could explore the integration of more advanced deep learning technologies, such as Graph Geural Getworks (GNNs), and more complex Transformer variants, to further improve reconstruction quality and the algorithm's generalization capability. Optimizing the algorithm's real-time performance and application on larger-scale datasets will be key to achieving broader applications.

Author Contributions: Conceptualization, E.Z. and N.Q.; Methodology, E.Z., Y.W. and N.Q.; Validation, C.G., Y.W. and N.Q.; Formal analysis, E.Z.; Investigation, C.G.; Data curation, N.Q.; Writing—original draft preparation, E.Z. and N.Q.; Writing—review and editing, E.Z., C.G. and Y.W.; Visualization, E.Z. and N.Q.; Supervision, C.G. and Y.W.; Project administration, C.G., Y.W. and E.Z.; Funding acquisition, E.Z. and C.G. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the General Program of National Nature Science Foundation of China under Grant 42271355 and 42271395.

Data Availability Statement: Data are contained within this article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Li, Z.-L.; Wu, H.; Duan, S.-B.; Zhao, W.; Ren, H.; Liu, X.; Leng, P.; Tang, R.; Ye, X.; Zhu, J.; et al. Satellite Remote Sensing of Global Land Surface Temperature: Definition, Methods, Products, and Applications. *Rev. Geophys.* **2023**, *61*, 1–77. [[CrossRef](#)]
2. Zhu, X.; Cao, L.; Wang, S.; Gao, L.; Zhong, Y. Anomaly Detection in Airborne Fourier Transform Thermal Infrared Spectrometer Images Based on Emissivity and a Segmented Low-Rank Prior. *Remote Sens.* **2021**, *13*, 754. [[CrossRef](#)]

3. Liu, X.; Li, Z.-L.; Li, Y.; Wu, H.; Zhou, C.; Si, M.; Leng, P.; Duan, S.-B.; Yang, P.; Wu, W.; et al. Local temperature responses to actual land cover changes present significant latitudinal variability and asymmetry. *Sci. Bull.* **2023**, *68*, 2849–2861. [[CrossRef](#)] [[PubMed](#)]
4. Zhang, D.; Zhu, Z.; Zhang, L.; Sun, X.; Zhang, Z.; Zhang, W.; Li, X.; Zhu, Q. Response of Industrial Warm Drainage to Tide Revealed by Airborne and Sea Surface Observations. *Remote Sens.* **2023**, *15*, 205. [[CrossRef](#)]
5. Wang, Y.; Chen, X.; Wang, F.; Song, M.; Yu, C. Meta-Learning Based Hyperspectral Target Detection Using Siamese Network. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5527913. [[CrossRef](#)]
6. Li, Y.; Li, Z.-L.; Wu, H.; Zhou, C.; Liu, X.; Leng, P.; Yang, P.; Wu, W.; Tang, R.; Shang, G.-F.; et al. Biophysical impacts of earth greening can substantially mitigate regional land surface temperature warming. *Nat. Commun.* **2023**, *14*, 121. [[CrossRef](#)] [[PubMed](#)]
7. Maes, W.H.; Steppe, K. Perspectives for Remote Sensing with Unmanned Aerial Vehicles in Precision Agriculture. *Trends Plant Sci.* **2019**, *24*, 152–164. [[CrossRef](#)] [[PubMed](#)]
8. Kuai, L.; Kalashnikova, O.V.; Hopkins, F.M.; Hulley, G.C.; Lee, H.; Garay, M.J.; Duren, R.M.; Worden, J.R.; Hook, S.J. Quantification of Ammonia Emissions with High Spatial Resolution Thermal Infrared Observations from the Hyperspectral Thermal Emission Spectrometer (HyTES) Airborne Instrument. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2019**, *12*, 4798–4812. [[CrossRef](#)]
9. Wang, Y.; Zhu, Q.; Ma, H.; Yu, H. A Hybrid Gray Wolf Optimizer for Hyperspectral Image Band Selection. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5527713. [[CrossRef](#)]
10. Liu, H.; Wu, K.; Xu, H.; Xu, Y. Lithology Classification Using TASI Thermal Infrared Hyperspectral Data with Convolutional Neural Networks. *Remote Sens.* **2021**, *13*, 3117. [[CrossRef](#)]
11. Fahlen, J.E.; Brodrick, P.G.; Thompson, D.R.; Herman, R.L.; Hulley, G.; Cawse-Nicholson, K.; Green, R.O.; Green, J.J.; Hook, S.J.; Miller, C.E. Joint VSWIR-TIR retrievals of earth's surface and atmosphere. *Remote Sens. Environ.* **2021**, *267*, 112727. [[CrossRef](#)]
12. Black, M.; Riley, T.R.; Ferrier, G.; Fleming, A.H.; Fretwell, P.T. Automated lithological mapping using airborne hyperspectral thermal infrared data: A case study from Anchorage Island, Antarctica. *Remote Sens. Environ.* **2016**, *176*, 225–241. [[CrossRef](#)]
13. Wang, Y.; Chen, X.; Zhao, E.; Song, M. Self-Supervised Spectral-Level Contrastive Learning for Hyperspectral Target Detection. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 5510515. [[CrossRef](#)]
14. Koz, A.; Efe, U. Geometric- and Optimization-Based Registration Methods for Long-Wave Infrared Hyperspectral Images. *Remote Sens.* **2021**, *13*, 2465. [[CrossRef](#)]
15. Qi, M.; Cao, L.; Zhao, Y.; Jia, F.; Song, S.; He, X.; Yan, X.; Huang, L.; Yin, Z. Quantitative Analysis of Mixed Minerals with Finite Phase Using Thermal Infrared Hyperspectral Technology. *Materials* **2023**, *16*, 2743. [[CrossRef](#)] [[PubMed](#)]
16. Li, T.; Gu, Y. Progressive Spatial-Spectral Joint Network for Hyperspectral Image Reconstruction. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5507414. [[CrossRef](#)]
17. Gerhards, M.; Schlerf, M.; Mallick, K.; Udelhoven, T. Challenges and Future Perspectives of Multi-/Hyperspectral Thermal Infrared Remote Sensing for Crop Water-Stress Detection: A Review. *Remote Sens.* **2019**, *11*, 1240. [[CrossRef](#)]
18. Wang, Y.; Wang, L.; Yu, C.; Zhao, E.; Song, M.; Wen, C.-H.; Chang, C.-I. Constrained-Target Band Selection for Multiple-Target Detection. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 6079–6103. [[CrossRef](#)]
19. Zhang, J.; Su, R.; Fu, Q.; Ren, W.; Heide, F.; Nie, Y. A survey on computational spectral reconstruction methods from RGB to hyperspectral imaging. *Sci. Rep.* **2022**, *12*, 11905. [[CrossRef](#)] [[PubMed](#)]
20. Zhao, E.; Gao, C.; Han, Q.; Yao, Y.; Wang, Y.; Yu, C.; Yu, H. An Operational Land Surface Temperature Retrieval Methodology for Chinese Second-Generation Huanjing Disaster Monitoring Satellite Data. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2022**, *15*, 1283–1292. [[CrossRef](#)]
21. Gade, R.; Moeslund, T.B. Thermal cameras and applications: A survey. *Mach. Vis. Appl.* **2014**, *25*, 245–262. [[CrossRef](#)]
22. Zou, C.; Wei, M. Cluster-based deep convolutional networks for spectral reconstruction from RGB images. *Neurocomputing* **2021**, *464*, 342–351. [[CrossRef](#)]
23. Huang, L.; Luo, R.; Liu, X.; Hao, X. Spectral imaging with deep learning. *Light-Sci. Appl.* **2022**, *11*, 61. [[CrossRef](#)] [[PubMed](#)]
24. Qu, Q.; Pan, B.; Xu, X.; Li, T.; Shi, Z. Unmixing Guided Unsupervised Network for RGB Spectral Super-Resolution. *IEEE Trans. Image Process.* **2023**, *32*, 4856–4867. [[CrossRef](#)] [[PubMed](#)]
25. Fu, Y.; Zheng, Y.; Zhang, L.; Huang, H. Spectral Reflectance Recovery From a Single RGB Image. *IEEE Trans. Comput. Imaging* **2018**, *4*, 382–394. [[CrossRef](#)]
26. Fotiadou, K.; Tsagkatakis, G.; Tsakalides, P. Spectral Super Resolution of Hyperspectral Images via Coupled Dictionary Learning. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 2777–2797. [[CrossRef](#)]
27. Zhao, J.; Kechasov, D.; Rewald, B.; Bodner, G.; Verheul, M.; Clarke, N.; Clarke, J.L. Deep Learning in Hyperspectral Image Reconstruction from Single RGB images-A Case Study on Tomato Quality Parameters. *Remote Sens.* **2020**, *12*, 3258. [[CrossRef](#)]
28. Miao, X.; Yuan, X.; Pu, Y.; Athitsos, V. lambda-Net: Reconstruct Hyperspectral Images From a Snapshot Measurement. In Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Republic of Korea, 27 October–2 November 2019.
29. Zhu, L.; Wu, J.; Biao, W.; Liao, Y.; Gu, D. SpectralMAE: Spectral Masked Autoencoder for Hyperspectral Remote Sensing Image Reconstruction. *Sensors* **2023**, *23*, 3728. [[CrossRef](#)] [[PubMed](#)]
30. Han, X.; Zhang, H.; Xue, J.H.; Sun, W. A Spectral-Spatial Jointed Spectral Super-Resolution and Its Application to HJ-1A Satellite Images. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 5505905. [[CrossRef](#)]
31. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention is All you Need. In Proceedings of the Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017.

32. Rogers, A.; Kovaleva, O.; Rumshisky, A. A Primer in BERTology: What We Know About How BERT Works. *Trans. Assoc. Comput. Linguist.* **2020**, *8*, 842–866. [[CrossRef](#)]
33. Liu, Z.; Lin, Y.; Cao, Y.; Hu, H.; Wei, Y.; Zhang, Z.; Lin, S.; Guo, B. Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, QC, Canada, 10–17 October 2021.
34. Du, D.; Gu, Y.; Liu, T.; Li, X. Spectral Reconstruction from Satellite Multispectral Imagery Using Convolution and Transformer Joint Network. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 5515015. [[CrossRef](#)]
35. Carion, N.; Massa, F.; Synnaeve, G.; Usunier, N.; Kirillov, A.; Zagoruyko, S. End-to-End Object Detection with Transformers. In Proceedings of the Computer Vision—ECCV 2020: 16th European Conference, Glasgow, UK, 23–28 August 2020.
36. Chen, K.; Zou, Z.; Shi, Z. Building Extraction from Remote Sensing Images with Sparse Token Transformers. *Remote Sens.* **2021**, *13*, 4441. [[CrossRef](#)]
37. Cai, Y.; Lin, J.; Lin, Z.; Wang, H.; Zhang, Y.; Pfister, H.; Timofte, R.; Gool, L.V. MST++: Multi-stage Spectral-wise Transformer for Efficient Spectral Reconstruction. In Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), New Orleans, LA, USA, 19–20 June 2022.
38. Zhao, Y.; Po, L.-M.; Yan, Q.; Liu, W.; Lin, T. Hierarchical Regression Network for Spectral Reconstruction from RGB Images. In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Seattle, WA, USA, 14–19 June 2020.
39. Shi, Z.; Chen, C.; Xiong, Z.; Liu, D.; Wu, F. HSCNN+: Advanced CNN-Based Hyperspectral Recovery from RGB Images. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Salt Lake City, UT, USA, 18–22 June 2018.
40. Hu, X.; Cai, Y.; Lin, J.; Wang, H.; Yuan, X.; Zhang, Y.; Timofte, R.; Gool, L.V. HDNet: High-resolution Dual-domain Learning for Spectral Compressive Imaging. In Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, 18–24 June 2022.
41. Chen, L.; Lu, X.; Zhang, J.; Chu, X.; Chen, C. HINet: Half Instance Normalization Network for Image Restoration. In Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Online, 19–25 June 2021.
42. Zamir, S.W.; Arora, A.; Khan, S.; Hayat, M.; Khan, F.S.; Yang, M.-H. Restormer: Efficient Transformer for High-Resolution Image Restoration. In Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, 18–24 June 2022.
43. Sun, W.; Ren, K.; Meng, X.; Yang, G.; Xiao, C.; Peng, J.; Huang, J. MLR-DBPFN: A Multi-Scale Low Rank Deep Back Projection Fusion Network for Anti-Noise Hyperspectral and Multispectral Image Fusion. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5522914. [[CrossRef](#)]
44. Zhang, H.; Cai, J.; He, W.; Shen, H.; Zhang, L. Double Low-Rank Matrix Decomposition for Hyperspectral Image Denoising and Destriping. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5502619. [[CrossRef](#)]
45. Zhuang, L.; Ng, M.K.; Liu, Y. Cross-Track Illumination Correction for Hyperspectral Pushbroom Sensor Images Using Low-Rank and Sparse Representations. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 5502117. [[CrossRef](#)]
46. Capelle, V.; Hartmann, J.-M. Use of hyperspectral sounders to retrieve daytime sea-surface temperature from mid-infrared radiances: Application to IASI. *Remote Sens. Environ.* **2022**, *280*, 113171. [[CrossRef](#)]
47. Lan, X.; Zhao, E.; Leng, P.; Li, Z.-L.; Labeled, J.; Nerry, F.; Zhang, X.; Shang, G. Alternative Physical Method for Retrieving Land Surface Temperatures from Hyperspectral Thermal Infrared Data: Application to IASI Observations. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–12. [[CrossRef](#)]
48. Ricciardelli, E.; Paola, F.; Cimini, D.; Larosa, S.; Mastro, P.; Masiello, G.; Serio, C.; Hultberg, T.; August, T.; Romano, F. A Feedforward Neural Network Approach for the Detection of Optically Thin Cirrus From IASI-NG. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 4104217. [[CrossRef](#)]

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