



## Article

# A Comprehensive Assessment of the Pansharpening of the Nighttime Light Imagery of the Glimmer Imager of the Sustainable Development Science Satellite 1

Hui Li <sup>1,2,3</sup> , Linhai Jing <sup>1,2,\*</sup>, Changyong Dou <sup>1,2</sup> and Haifeng Ding <sup>1,2</sup>

<sup>1</sup> International Research Center of Big Data for Sustainable Development Goals, Beijing 100094, China; lihui@radi.ac.cn (H.L.); doucy@aircas.ac.cn (C.D.); dinghf@aircas.ac.cn (H.D.)

<sup>2</sup> Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China

<sup>3</sup> Hainan Key Laboratory of Earth Observation, Hainan Aerospace Information Research Institute, Wenchang 571399, China

\* Correspondence: jinglh@aircas.ac.cn

**Abstract:** The Sustainable Development Science Satellite 1 (SDGSAT-1) satellite, launched in November 2021, is dedicated to providing data detailing the “traces of human activities” for the implementation of the United Nations’ 2030 Agenda for Sustainable Development and global scientific research. The glimmer imager (GI) that is equipped on SDGSAT-1 can provide nighttime light (NL) data with a 10 m panchromatic (PAN) band and red, green, and blue (RGB) bands of 40 m resolution, which can be used for a wide range of applications, such as in urban expansion, population studies of cities, and economics of cities, as well as nighttime aerosol thickness monitoring. The 10 m PAN band can be fused with the 40 m RGB bands to obtain a 10 m RGB NL image, which can be used to identify the intensity and type of night lights and the spatial distribution of road networks and to improve the monitoring accuracy of sustainable development goal (SDG) indicators related to city developments. Existing remote sensing image fusion algorithms are mainly developed for daytime optical remote sensing images. Compared with daytime optical remote sensing images, NL images are characterized by a large amount of dark (low-value) pixels and high background noises. To investigate whether daytime optical image fusion algorithms are suitable for the fusion of GI NL images and which image fusion algorithms are the best choice for GI images, this study conducted a comprehensive evaluation of thirteen state-of-the-art pansharpening algorithms in terms of quantitative indicators and visual inspection using four GI NL datasets. The results showed that PanNet, GLP\_HPM, GSA, and HR outperformed the other methods and provided stable performances among the four datasets. Specifically, PanNet offered UIQI values ranging from 0.907 to 0.952 for the four datasets, whereas GSA, HR, and GLP\_HPM provided UIQI values ranging from 0.770 to 0.856. The three methods based on convolutional neural networks achieved more robust and better visual effects than the methods using multiresolution analysis at the original scale. According to the experimental results, PanNet shows great potential in the fusion of SDGSAT-1 GI imagery due to its robust performance and relatively short training time. The quality metrics generated at the degraded scale were highly consistent with visual inspection, but those used at the original scale were inconsistent with visual inspection.

**Keywords:** nighttime light imagery; pansharpening; glimmer imager (GI); sustainable development science satellite 1 (SDGSAT-1)



**Citation:** Li, H.; Jing, L.; Dou, C.; Ding, H. A Comprehensive Assessment of the Pansharpening of the Nighttime Light Imagery of the Glimmer Imager of the Sustainable Development Science Satellite 1. *Remote Sens.* **2024**, *16*, 245. <https://doi.org/10.3390/rs16020245>

Academic Editor: Bailang Yu

Received: 25 October 2023

Revised: 30 December 2023

Accepted: 5 January 2024

Published: 8 January 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

At the United Nations (UN) Sustainable Development Summit on 25 September 2015, the 193 UN member states formally adopted the 2030 Agenda for Sustainable Development, which announced 17 Sustainable Development Goals (SDGs) [1]. The SDGs aims to move toward a sustainable development path by thoroughly addressing the social, economic, and

environmental dimensions of development in an integrated manner between 2015 and 2030. Since the implementation of the UN 2030 Agenda for Sustainable Development, it has been facing huge challenges such as a lack of data, insufficient research on indicator systems, and uneven development between different countries [2]. The essence of sustainable development is the harmonious coexistence of humans and nature. Therefore, how to improve the understanding of the mechanisms and evolutionary laws of the interaction between human activities and the natural environment and deepen the in-depth knowledge of the multidimensional development of society, the economy, and the environment and the intrinsic connection with the Earth system are critical scientific issues facing the progress towards sustainable development.

Space technology, represented by Earth observations from space, has the capability to continuously observe critical elements of the surface environment and socioeconomic development indicators in a macroscopic, objective, dynamic, and high-precision manner. This enables a deeper understanding of the interaction mechanisms between human activities and the Earth's environment and can play a major role in monitoring, assessing, and analyzing sustainable development goals. Big Earth data can be used to support SDG indicators through three significant aspects. First, Big Earth data are used to compensate for data gaps and provide new data sources for the monitoring and assessment of SDGs. Second, new methodologies are developed based on Big Earth data technologies, and new models are constructed to monitor and assess SDG indicators. Finally, case studies on monitoring and assessment of SDG indicators using Big Earth data are provided to support SDGs globally and regionally, which provides a practical contribution to decision making.

Remote sensing of nighttime light (NL) offers a unique opportunity to directly observe human activity from space [3]. The number and quality of NL remote sensing sensors has greatly increased since the early 2000s, enabling a large number of applications such as tracking urbanization and socioeconomic dynamics, evaluating armed conflicts and disasters, investigating fisheries, assessing greenhouse gas emissions and energy use, and analyzing light pollution and health effects [4–10]. The earliest NL products were derived from the Defense Meteorological Satellite Program's Operational Line-scan System (DMSP-OLS). With a spatial resolution of 2.7 km, the DMSP-OLS products have been available since 1992 [10]. The NL imagery generated by the Day–Night Band (DNB) of the Visible Infrared Imaging Radiometer Suite (VIIRS) sensor carried on the Suomi National Polar Orbiting Partnership (SNPP) satellite has been available since April 2012. The VIIRS/DNB routinely provides panchromatic global imagery with a 742 m spatial resolution [11]. In comparison to DMSP/OLS images, the VIIRS/DNB data have a better spatial resolution and lower light detection limits ( $2 \times 10^{-11} \text{ W/cm}^2/\text{sr}$  vs.  $5 \times 10^{-10} \text{ W/cm}^2/\text{sr}$  in US-DMSP), which is especially important for analyzing dimly lit areas. Additionally, the VIIRS/DNB data also do not exhibit bright light saturation, which is one of the major shortcomings of the DMSP/OLS data collections [3,10]. DMSP-OLS and VIIRS-DNB played an indispensable role in large-scale NL studies, which included urbanization, socioeconomic activities, and environmental changes [12].

Besides coarse-spatial-resolution NL images from DMSP-OLS and VIIRS/DNB, NTL data with a higher resolution are photographs taken by astronauts on the International Space Station (ISS). The astronaut photos are the earliest multispectral images and have a spatial resolution ranging from 5 to 200 m, providing more details of the Earth. However, technical challenges in radiometric calibration and uneven temporal and spatial distributions of these original photos hinder a wide application of ISS images. There are also some commercial satellites providing fine-spatial-resolution NL imageries, such as EROS-B and JL1-3B. Launched in 2013, EROS-B provides 0.7 m spatial resolution NL imageries with a spectral band wavelength range of 0.5–0.9  $\mu\text{m}$  and a dynamic range of 10 bits. However, as the NTL images of EROS-B are panchromatic, the lighting type cannot be identified from these data [13]. JL1-3B was launched in 2017 and provides multispectral (red, green, and blue) NTL imageries at a spatial resolution of 0.92 m with a capability to detect light as low as  $7 \times 10^{-7} \text{ W}\cdot\text{cm}^{-2}\cdot\text{sr}^{-1}$  [12]. With the advantages of a submeter spatial resolution and

multispectral information, as well as its on-board radiance calibration, new capabilities of lighting types, road detection, land use, and urban nightscape patterns are promising in future studies [10,14]. Another data source of NL imagery is cubesats, such as Luojia-1 launched in 2018. Luojia-1 provides NL images with a spatial resolution of 130 m and a dynamic range up to 14 bits, enabling us to accurately map urban dynamics, monitor the construction of infrastructure, and retrieve  $PM_{2.5}$  concentrations at a moderate spatial resolution [10].

A new source of NL data is the Sustainable Development Science Satellite 1 (SDGSAT-1), launched in November 2021. SDGSAT-1 is the first global scientific satellite dedicated to the implementation of the UN 2030 Agenda for Sustainable Development and the needs of global scientific research [15]. SDGSAT-1 is equipped with three payloads including a thermal infrared spectrometer (TIS), a glimmer imager (GI), and a multispectral imager (MI). The day–night coordinated observations from the three payloads provide a detailed description of the traces of human activities. They will provide data support for the study of indicators characterizing the interactions between humans and nature, as well as monitoring, evaluation, and scientific research on the achievement of the SDGs on a global scale. They also provide spatial data to countries along the Belt and Road to contribute to global scientific needs.

The GI sensor of SDGSAT-1 provides 10 m nighttime imagery with a single panchromatic (PAN) band and 40 m nighttime color imagery with the red, green, and blue bands. The sensor has significantly improved spectral and spatial resolution compared to the existing DMSP-OLS, SNPP-VIIRS-NDB, and Luojia-1 products. The GI can provide worldwide products that are free of charge to researchers. The GI nighttime light (NL) imagery reflects information about road networks and residential areas and other information that is closely related to the distributions of populations and cities. Some studies tried to transform panchromatic NL images to RGB images to help generate more information on humans' presence on Earth [11]. In contrast, the GI of SDGSAT-1 provides directly 30-m RGB images containing spectral information that is helpful for identifying the type of lights. Moreover, image fusion techniques can be employed to fuse the 40 m RGB image with the 10 m PAN image to produce 10 m RGB NL images. The fused products can be used to distinguish information on the type, intensity, and spatial distribution of nighttime lights and to identify the spatial distribution of road networks and residential areas, which can provide fine-resolution NL data for improving the accuracy of SDG indicator monitoring.

Existing remote sensing image fusion algorithms are mainly developed for the fusion of daytime optical remote sensing images. Current fusion algorithms can be broadly classified into component substitution (CS) methods [16–24], methods based on multiresolution analysis (MRA) [25–36] and variational optimization (VO) [37–41], and deep learning (ML) algorithms [42–51]. CS methods transform the MS image into a new domain, such as an intensity–hue–saturation (IHS) color space. One of the components is replaced by the original PAN band, and the new components are then transformed back to the original domain. Typical CS methods include the intensity–hue–saturation (IHS), principal component analysis (PCA), Brovey transform (BT), PRCAS, and Gram–Schmidt (GS). The CS methods are easy to implement, and the generated fused MS images yield high spatial quality. However, the CS methods suffer from spectral distortions, since the local dissimilarities between the PAN and MS channels, which are caused by different spectral response ranges between PAN and MS bands, are not considered by them. Fortunately, the spectral range of the PAN band of SDGSAT-1 GI sensor covers the entire ranges of the RGB bands. MRA methods extract spatial details from the PAN band using a multiresolution decomposition, such as a Laplacian filter or wavelet. The details are then injected into the upsampled MS bands. Although the MRA-based methods better preserve spectral information of the original MS images than the CS-based methods, they may cause spatial distortions, such as ringing or aliasing effects, leading to shifts or blurred contours and textures [52]. VO methods address the pansharpening problem through the optimization of suitable variational models. However, the problem to be solved is clearly ill posed. The target image is usually estimated under the assumption of proper co-registered PAN and MS

images. This limits the application of such methods, as in most cases, we need to deal with registration issues [41]. ML methods use designed convolutional neural networks (CNNs) to learn a pansharpened image or the residual between an upsampled MS image and a pansharpened MS image. ML methods generally yield outstanding performances in terms of both spatial detail enhancement and spectral fidelity, but they require that training and testing data have similar properties, indicating that these methods have a relatively weak generalization ability [53]. Additionally, some studies have focused on exploring optimal image fusion methods for a certain sensor, such as WorldView-2 (WV-2) [54–57]. Several studies have compared the performances of commonly used pansharpening algorithms on WV-2 imagery, which has four additional multispectral bands that are absent in earlier satellites such as GeoEye-1 and Pleiades.

Different from daytime optical remote sensing images, NL images are characterized by a large number of dark pixels and have significantly fewer spatial details. In addition, NL imagery commonly suffers from background noise. GI NL images record the spatial distributions of nighttime lights of road networks and residential and commercial districts in cities and towns. It is dark in areas outside cities, such as rural areas. Therefore, there are a large number of dark pixels in SDGSAT-1 NL imagery, which increases the indeterminacy of the performances of some statistically based pansharpening methods. Additionally, some SDGSAT-1 NL imagery of Level 4 also faces the problems of stripe noise and misalignment between the MS and PAN bands, which is one of the most critical factors directly affecting the quality of fused images. Therefore, it is necessary to verify the effectiveness of traditional optical remote sensing image fusion methods on the fusion of SDGSAT-1 NL imagery. Furthermore, there is an urgent need to confirm which method or which type of fusion algorithms should be selected for SDGSAT-1 NL data. Determining these questions is of great importance for obtaining high-quality fused GI NL images and for better serving the monitoring and assessment of urbanization-related SDG indicators.

Therefore, in this study, thirteen state-of-the-art pansharpening methods, including five CS methods, five MRA-based methods, and three CNN-based methods were comprehensively evaluated through quantity indices and visual inspection of fused products of four SDGSAT-1 GI NL images. The experimental results of this work can provide valuable references for the selection of optimal pansharpening algorithms to generate high-quality 10 m RGB NL images used for monitoring and assessing SDGs. The results will also give helpful directions for developing new pansharpening methods for SDGSAT-1. The remaining parts of this study are organized as follows: Section 2 introduces the relevant parameters of SDGSAT-1 GI, the experimental datasets, the fusion algorithms, and the quantitative evaluation metrics of the fused images; Section 3 presents the results of the fusion experiments; and Sections 4 and 5 are the discussion and conclusion, respectively.

## 2. Materials

### 2.1. SDGSAT-1 Night Light Sensor Parameters

The spectral band settings for the SDGSAT-1 GI are shown in Table 1. The sensor consists of four bands with a spectral response range of 424–910 nm. The specific band ranges and bandwidths for the PAN, red (R), green (G), and blue (B) bands are shown in Table 1. The spatial resolution is 10 m for the PAN band and 40 m for the R, G, and B bands. Each band is recorded with no less than 12 bits, and the minimum detection limit for the bands is set at  $\sim 1 \times 10^{-5} \text{ W}/(\text{m}^2 \cdot \text{sr})$ .

**Table 1.** Band parameters of the GI of SDGSAT-1.

Band	Center Wavelength (nm)	Wavelength Range (nm)	Bandwidth (nm)	Spatial Resolution (m)	SNR (Observed Objects with a Reflectance Higher than 0.2)
Panchromatic	680.72	444–910	466	10	Lights on city trunk roads $\geq 30$ ; Urban residential district $\geq 20$ ; Polar moon light $\geq 10$ .
Blue	478.87	424–526	102	40	Lights on city trunk roads $\geq 15$ ; Urban residential district $\geq 10$ .
Green	561.20	506–612	96	40	
Red	734.25	600–894	294	40	

## 2.2. Datasets

Four Level 4 GI NL products were used in this work. The Level 4 products are ortho-corrected products using ground control points and digital elevation models based on the Level 1 products, which are standard products after relative radiation correction, band alignment, high dynamic range (HDR) fusion, and rational polynomial coefficient (RPC) processing based on Level 0 products. The Level 4 products provide three radiance bands derived from the PAN band, which are panchromatic low (PL), panchromatic high (PH), and HDR. The PL and PH are obtained using two versions of gain bias values for the PAN band, whereas the HDR is the average of the PL and PH bands. We used the HDR band to fuse with the three color bands in this work to produce pansharpened images with more details in this work.

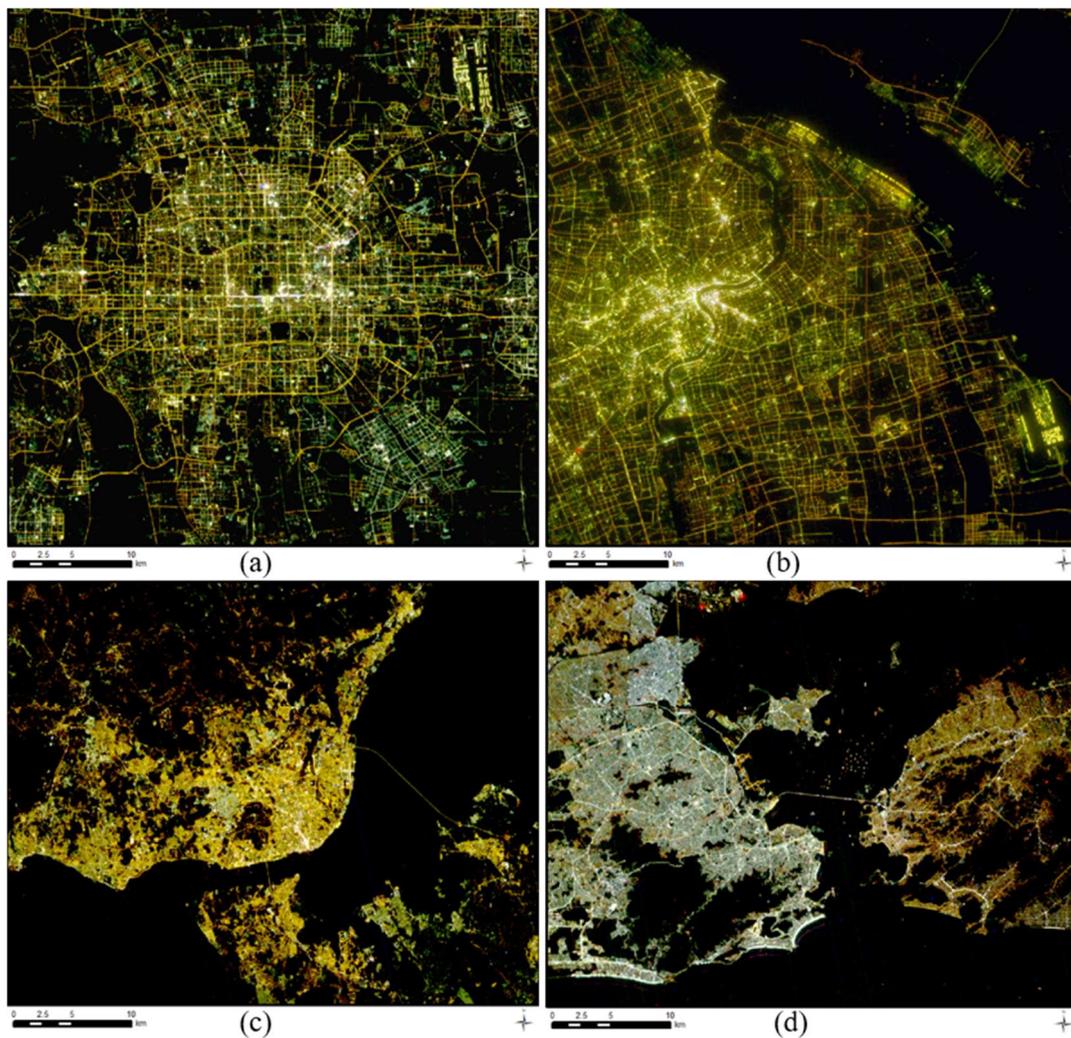
Figure 1 shows the four image scenes in Beijing, China; Rio de Janeiro, Brazil; Lisbon, Portugal; and Shanghai, China. Subarea images were selected from the four image scenes to carry out fusion experiments with quantitative evaluation of fused image quality and visual comparison analysis. Table 2 introduces the recording time and image size information of these subimages, and Table 3 shows the statistical information such as minimum, maximum, mean, and standard deviation of the four subimages. As seen from Table 3, the mean values of the images were extremely low and even lower than the corresponding standard deviation values. This is because there are a large number of dark pixels in the images, although these images cover urban areas. In addition, the mean and variance of the PAN band are significantly lower than those of the R, G, and B bands. This is because the PAN image has a larger number of dark pixels than the R, G, and B bands.

**Table 2.** Four SDGSAT-1 NL datasets used for the evaluation of pansharpening methods.

Id	Location	Sensor	Date	Image Size (MS/PAN)
1	Beijing	SDGSAT-1 GI	November 2021	$512 \times 512/2048 \times 2048$
2	Lisbon	SDGSAT-1 GI	January 2022	$512 \times 512/2048 \times 2048$
3	Shanghai	SDGSAT-1 GI	April 2022	$512 \times 512/2048 \times 2048$
4	Brazil	SDGSAT-1 GI	June 2022	$512 \times 512/2048 \times 2048$

**Table 3.** The statistics of the four SDGSAT-1GI nighttime light datasets.

Dataset	Band	Minimum	Maximum	Mean	Standard Deviation
Beijing	R	1	4426	412.26	522.81
	G	1	4643	396.93	530.18
	B	1	4411	135.11	315.24
	PAN	1	4465	35.37	153.17
Lisbon	R	1	3852	270.96	411.55
	G	1	4079	180.75	301.28
	B	1	3487	27.69	89.97
	PAN	1	4152	13.47	37.05
Shanghai	R	7	4357	236.24	339.20
	G	7	4616	219.54	343.54
	B	7	4427	67.80	176.19
	PAN	7	4420	17.17	78.74
Brazil	R	7	3789	230.14	218.47
	G	7	3935	300.94	33.92
	B	7	3962	126.52	33.06
	PAN	7	3620	19.09	32.89



**Figure 1.** The 40 m SDGSAT-1 GI color imagery of Beijing, China (a), Shanghai, China (b), Lisbon, Portugal (c), and Rio de Janeiro, Brazil (d).

The fusion experiments were performed on both the original and degraded scales according to Wald's protocol of quality assessment [58]. The fused products at the two scales were then evaluated using quantitative indexes. Specifically, the original MS and PAN images were degraded using the modulation transfer function (MTF) of the sensor and then downsampled to spatial resolutions of 160 m and 40 m, respectively, to generate the test images at the degraded scale [32]. The fused products that were generated at the degraded scale were then evaluated using the original 40 m MS image as a reference. For the original scale, the 40 m MS image was fused with the 10 m PAN image to generate 10 m MS images. The fused images were evaluated using quality index without a reference image along with visual inspection. The details of the quality metrics are introduced in Section 2.4.

### 2.3. Fusion Algorithms

The thirteen pansharpener methods considered in this work are shown in Table 4. Five CS methods and five MRA-based methods, which belong to the conventional pansharpener methods, and three CNN-based pansharpener methods were evaluated in this work. The details of these methods are briefly introduced in this subsection.

### 2.3.1. Conventional Pansharpening Methods

Conventional pansharpening methods include CS- and MRA-based methods and methods based on sparse representation. Here, we introduce some details of the CS- and MRA-based fusion algorithms considered in this work.

#### (1) CS methods

The CS algorithms use a simulated component obtained by a weighted sum of the MS bands to replace the original PAN band. It is given that  $M$  and  $P$  represent the original low-resolution MS image and high-resolution PAN image, respectively.  $\tilde{M}$  and  $\hat{M}$  represent the upsampled MS and the fused MS images, respectively. The CS fusion can be formulized as

$$\begin{aligned}\hat{M}_i &= \tilde{M}_i + g_i(P - I_L), \quad i = 1, \dots, N \\ I_L &= \sum_{i=1}^N w_i \tilde{M}_i,\end{aligned}\quad (1)$$

in which the subscript  $i$  indicates the  $i$ th spectral band,  $g_i$  is the injection gains of the  $i$ th band,  $w_i$  is the weight of the  $i$ th MS band, and  $N$  is the number of MS bands, while  $I_L$  is a low-resolution simulation of the original PAN band. In this work, we evaluated the performances of the Intensity Hue Saturation (IHS) [16], principal component analysis (PCA) [17], adaptive Gram–Schmidt (GSA) [19,20], haze- and ratio-based (HR) [21], and reduce misalignment impact (RMI) [22] for the fusion of SDGSAT-1 GI imagery. The IHS and PCA methods were selected because they were the earliest methods developed for pansharpening. The GSA, HR, RMI were considered due to their outstanding performances in previous works [57].

For the IHS method, the  $I_L$  can be obtained by setting the coefficients  $w_i$  to  $1/N$ , whereas the injection coefficients  $g_i$  are all equal to 1.

The PCA method transforms the spectral bands to new uncorrelated spectral directions through an orthonormal projection matrix. For the PCA method, the  $I_L$  is the first component, which presents the largest variance and contains abundant spatial information. The coefficient  $g_i$  is the first column of the backward transformation matrix.

The GSA is an improved version of the GS method, and the weights  $w_i$  for GSA are estimated by the linear regression coefficients between the original MS band and the degraded version of the PAN band. The  $g_i$  for GSA is the ratio of the covariance between the two bands  $\tilde{M}_i$  and  $I_L$  to the variance of  $I_L$ . For HR, the  $I_L$  is a low-pass version of the original PAN image. For RMI, the  $I_L$  is obtained using a weighted sum of the MS bands, where the weights  $w_i$  are estimated using a similar approach to the GSA method. On the other hand, both the HR and RMI methods are based on the assumption that the ratio of an HSR MS band to a low-spatial-resolution (LSR) MS band is equal to the ratio of an HSR PAN image to a synthetic LSR PAN image.

#### (2) MRA-based methods

The MRA techniques inject the spatial details obtained through a multiresolution decomposition of the original PAN band into the upsampled MS bands. Similarly, a general formulation for MRA-based methods can be given by (2)

$$\hat{M}_i = \tilde{M}_i + g_i(P - P_L), \quad i = 1, \dots, N \quad (2)$$

where  $P_L$  is the low-frequency component of the PAN band.  $P_L$  can be derived for the PAN band through different approaches, such as low-pass filter, Laplacian pyramid, and wavelet decomposition.

We considered the high-pass filter (HPF) [25], “à trous” wavelet transform (ATWT) [26,28] and generalized Laplacian Pyramid (GLP) [29,30,32] methods with different detail injection models in this work. These methods were employed as they are classic methods that have obtained stable performances in previous comparisons [59]. The HPF uses a low-pass filter, typically a box mask with uniform weights for average filtering, to obtain  $P_L$ , and the coefficient

$g_i$  is equal to 1. The ATWT method uses the “À trous” wavelet decomposition to obtain  $P_L$ . Similarly,  $g_i$  is equal to 1. For the GLP methods,  $P_L$  is generated by a Gaussian low-pass filter that matches the MTF of the SDGSAT-1 GI sensor to achieve accurate estimation of the spatial degradation model [32]. Moreover, several injection models can be used to obtain the injection coefficients  $g_i$ . For the GLP fusion scheme, the simplest approach is the additive injection model, for which the coefficients  $g_i$  are all equal to 1 [60]. For GLP-based fusion with a context-based decision model (GLP\_CBD), the coefficient  $g_i$  is the ratio of the covariance between two bands  $\tilde{M}_i$  and  $I_L$  to the variance of  $I_L$ , which ensures that the spectral vector of a pixel in the fused image is parallel to that of the corresponding pixel in the upsampled MS image [32,52,60]. For the GLP-based fusion with the high-pass model (GLP\_HPM), the coefficient  $g_i$  is the ratio of the  $i$ th MS band and  $P_L$  [32,61].

**Table 4.** Fusion algorithms considered for the fusion of SDGSAT-1 GI data.

Category	Method
Conventional algorithms	CS IHS [16], PCA [17], GSA [19,20], RMI [22], HR [21]
	MRA-based HPF [25], ATWT [26,28], GLP [29,32,60], GLP_HPM [32,61], GLP_CBD [32,52,60]
CNN-based pansharpening	A-PNN [46], PanNet [44], Z-PNN [50]

### 2.3.2. CNN-Based Pansharpening Methods

The first CNN-based pansharpening (PNN) method, proposed in 2016 [42], was improved from an architecture proposed for single nature image with super-resolution tasks [62]. In recent years, deep learning has become a particularly popular solution for pansharpening. A large number of CNN-based methods were proposed in recent years [43–50]. The PNN method adopts a three-layer CNN architecture. The input of the model is the combination of the original PAN band and the upsampled MS bands, whereas the output is the pansharpened MS bands. The proposed deep residual pansharpening neural network (DRPNN) [43] employs a residual learning approach and an 11-layer CNN architecture. The deeper architecture contributes to the sharper edges obtained by the DRPNN method. A target-adaptive version of the PNN method (A-PNN) adopts residual learning and a target-adaptive fine-tuning step to improve the training efficiency and robustness over a wide distribution of data [46]. The PanNet method uses the ResNet structure along with a spectra-mapping strategy for spectral preservation [44]. The input of PanNet is the original PAN image and the original MS image. Most existing CNN-based pansharpening methods, such as PNN, A-PNN, and PanNet, use degraded datasets at a lower resolution for the training of model parameters, which are used for the fusion of the original dataset. Recently, an improved version of the PNN method, namely, Z-PNN or Zoom-PNN, employed a full-resolution training framework [50]. The Z-PNN method trains the model using a dataset of the original resolution with a newly defined loss, including a spectral component and a spatial component. The spectral loss component enforces the spectral consistency between the pansharpened image and the original low-resolution MS image, whereas the spatial loss component maximizes the spatial correlation between each fused band and the PAN band. The PNN and its improved versions are commonly used in comparisons. The target-adaptive PNN [46] and Z-PNN [50], as well as PanNet [44], were considered in terms of exploring their performances on GI images, as they have yielded outstanding performances in previous works [53].

The A-PNN implemented in Theano [46] and the Z-PNN implemented in PyTorch [50] were used in this work. Another image scene recorded on 15 April 2022 was trained to obtain pre-trained models for A-PNN and Z-PNN. For A-PNN, 1000 epochs were considered for the fine-tuning phase using the degraded version of the test datasets. For Z-PNN, the datasets at the original scale were then trained for 3000 epochs based on the pre-trained model. The PanNet was implemented in Keras, and the model was trained using the degraded version of the test datasets. The three networks were trained using a

graphic station equipped with an NVIDIA GeForce RTX 2080 Ti GPU, which has 11 GB memory.

## 2.4. Fusion Image Quality Evaluation Index

### 2.4.1. Quality Indices with a Reference Image

Pansharpened remote sensing images at the degraded scale are commonly evaluated using quality indices such as ERGAS, SAM, UIQI, and SCC. The introductions of the abbreviations are presented in Table 5.

The ERGAS is defined as a weighted sum of the root of the mean square error (RMSE) [63]. The ERGAS is calculated by (3)

$$\text{ERGAS} = \frac{100}{R} \sqrt{\frac{1}{N} \sum_{k=1}^N \left( \frac{\text{RMSE}(\hat{M}_k, M_k)}{u(M_k)} \right)^2} \quad (3)$$

where  $u(M_k)$  is the mean of the  $k$ th band of the reference image  $M$ ;  $R$  is the spatial resolution ratio between the MS and PAN bands. The optimal value for ERGAS is 0.

SAM measures the spectral similarity between a fused product and the corresponding reference image [64]. Let two spectral vectors  $V = \{V_1, V_2, \dots, V_M\}$  and  $\hat{V} = \{\hat{V}_1, \hat{V}_2, \dots, \hat{V}_M\}$  present the reference spectral pixel and the fused spectral pixel, respectively; their spectral angle SAM is defined as (4)

$$\text{SAM} = \arccos\left(\frac{\langle V, \hat{V} \rangle}{|V| \cdot |\hat{V}|}\right) \quad (4)$$

where  $\langle X, Y \rangle$  stands for the inner product of the two vectors  $X$  and  $Y$ , and  $|X|$  stands for the modulus of a vector  $X$ . The smaller the spectral angle, the higher the similarity between the two vectors. The optimal value for SAM is 0.

The UIQI is a comprehensive measure that considers intensity, contrast, and local correlation. As UIQI is for a single band, the multiband extensions of UIQI, including Q4 [65] and Q2<sup>n</sup> [66], were proposed and then widely used in pansharpening. As the SDGSAT-1 nighttime light imagery considered in this work has only three bands, the UIQI index is employed. The UIQI index is defined as (5)

$$Q = \frac{4\sigma_{x,y} \cdot \bar{x} \cdot \bar{y}}{(\sigma_x^2 + \sigma_y^2) \left[ (\bar{x})^2 + (\bar{y})^2 \right]} = \frac{\sigma_{x,y}}{\sigma_x \sigma_y} \cdot \frac{2\bar{x} \cdot \bar{y}}{(\bar{x})^2 + (\bar{y})^2} \cdot \frac{2\sigma_x \cdot \sigma_y}{\sigma_x^2 + \sigma_y^2} \quad (5)$$

where  $\sigma_{x,y}$  denotes the covariance between the fused image  $x$  and the reference image  $y$ ,  $\bar{x}$  and  $\bar{y}$  are the means, and  $\sigma_x^2$  and  $\sigma_y^2$  are the variances of  $x$  and  $y$ , respectively. The dynamic range of UIQI is  $[-1, 1]$ , and the best value is achieved if  $x = y$  for all pixels. Specifically, the UIQI value in this work refers to the average of the UIQI values of the three fused bands.

SCC assesses the correlation between spatial details that are presented in two images. Similar to the procedure proposed by Otazu et al. [31], the spatial information presented in the two images to be compared is extracted by using a Laplacian filter; then, the correlations between the two filtered images are calculated band by band. A high SCC value indicates that many of the spatial details of the reference image are presented in the fused image. The optimal value for SCC is 1. Specifically, an overall correlation coefficient of the two edge images with three bands is calculated in this study.

Some published studies have also considered PSNR and SSIM [51,67]. As a spatial quality index, SSIM measures the structural similarity between a fused product and a reference image. The PSNR is defined based on the mean square error (MSE); the higher the PSNR value is, the higher the similarity of the two images is. However, some studies expressed that MSE and PSNR are not very well matched to perceived visual quality [68].

**Table 5.** Quality indices used at the reduced and original scales for the evaluation of fused products.

Scale	Index	Details	Ideal Value
Reduced scale	ERGAS [63]	Erreur Relative Globale Adimensionnelle de Synthèse	0
	SAM [64]	Spectral angle mapper	0
	UIQI [65]	Universal Image Quality Index	1
	SCC [31]	Spatial correlation coefficient	1
	SSIM [59]	Structural similarity index	1
	PSNR	Peak signal-to-noise ratio	-
Original scale	QNR [69]	Quality with no-reference index	1
	$D_\lambda$ [69]	Spectral distortion index	0
	$D_s$ [69]	Spatial distortion index	0
	$D_\lambda^k$ [70,71]	Khan's spectral consistency index	0
	HQNR [70,71]	Hybrid QNR based on $D_\lambda^k$ and $D_s$	1
	$D_\rho$ [51]	Spatial consistency index	1

#### 2.4.2. Quality Indices without a Reference Image

The evaluation of pansharpened products at the original scale is a challenging problem due to the lack of a reference image. The most widely used quality index for fused products obtained at the original scale is the quality with no reference (QNR) index. The QNR is a combination of two separate metrics measuring spectral ( $D_\lambda$ ) and spatial ( $D_s$ ) distortions [69].  $D_\lambda$  measures the intrarelationship changes between the pansharpened MS bands and those between the original MS bands. The value of  $D_\lambda$  is estimated as (6)

$$D_\lambda \triangleq \sqrt[p]{\frac{1}{N(N-1)} \sum_{l=1}^N \sum_{\substack{r=1 \\ r \neq l}}^N \left| Q(\hat{M}_l, \hat{M}_r) - Q(\tilde{M}_l, \tilde{M}_r) \right|^p} \quad (6)$$

where  $Q(\hat{M}_l, \hat{M}_r)$  is the UIQI index calculated from the  $l$ th and  $r$ th band of the fused image, and  $Q(\tilde{M}_l, \tilde{M}_r)$  is the UIQI index of the  $l$ th and  $r$ th band of the upsampled MS image  $\tilde{M}$ .  $N$  is the number of spectral bands.  $p$  is a positive integer exponent chosen to emphasize large spectral differences and is typically set to one. The index is proportional to the  $p$ -norm of the difference matrix, being equal to 0, if and only if the two matrices are identical.

$D_s$  measures interrelationship changes between the MS and Pan bands and is estimated as (7)

$$D_s \triangleq \sqrt[q]{\frac{1}{N} \sum_{l=1}^N \left| Q(\hat{M}_l, P) - Q(\tilde{M}_l, \tilde{P}) \right|^q} \quad (7)$$

in which  $P$  is the original PAN image, and  $\tilde{P}$  is a spatially degraded version of the PAN image, obtained by filtering with a low-pass filter, having normalized the frequency cutoff at the resolution ratio between MS and PAN, followed by decimation. The index  $D_s$  attains its minimum (equal to zero) when the two vectors are identical.

As a joint index based on  $D_\lambda$  and  $D_s$ , weighted by the coefficients  $\alpha$  and  $\beta$ , the QNR index is defined as (8)

$$QNR \triangleq (1 - D_\lambda)^\alpha \cdot (1 - D_s)^\beta \quad (8)$$

The highest value of QNR is one, and it is obtained when the spectral distortion  $D_\lambda$  and spatial distortion  $D_s$  are both zero.

Both  $D_\lambda$  and  $D_s$  do not measure the discrepancy between the fused product and original MS and PAN images directly. Due to the lack of reference images, an improved spectral distortion index, i.e.,  $D_{\lambda'}^k$ , was proposed by using the  $Q2^n$  index to compare the

degraded version of the fused image obtained using the MTF filter followed by decimation with the original MS image [70]. The  $D_\lambda^k$  index is defined as (9)

$$D_\lambda^k \triangleq 1 - Q2^n(\hat{M}_L, M) \quad (9)$$

where  $\hat{M}_L$  is the degraded version of the fused image  $\hat{M}$ , and  $M$  is the original MS image. The optimal value of  $D_\lambda^k$  is close to zero when  $\hat{M}_L$  is identical to  $M$ .

The hybrid QNR (HQNR) index is the combination of the spectral distortion  $D_\lambda^k$  and the spatial distortion index  $D_s$  used by QNR. The HQNR index is defined as (10)

$$HQNR \triangleq (1 - D_\lambda^k)^\alpha \cdot (1 - D_s)^\beta \quad (10)$$

The optimal value of HQNR is also close to 1 when both  $D_\lambda^k$  and  $D_s$  are close to zero.

Similarly, an improved spatial distortion index, referred to as  $D_s^F$ , used UIQI to measure the interrelationships between the details of the PAN image and those of the MS images across resolution scales [70,71]. However,  $D_s^F$  was sensitive to the spatial filters used to generate high-frequency images [71].

The  $D_\rho$  index is another spatial consistency index that is proposed to obtain closer correlation with human judgment than those of  $D_s$  [51]. This index computes the average local correlation between the pansharpened image and the PAN. Let  $X_{ij}^\sigma$  indicate a  $\sigma \times \sigma$  patch of image  $X$ , centered on location  $(i, j)$ . We compute the correlation field  $\rho_{P\hat{M}}^\sigma$  given by the local correlation coefficients between  $P$  and each band  $b$  of  $\hat{M}$ , as shown in (11). Then, we reduce it to its average value over space and spectral bands. The final index is then defined as (12)

$$\rho_{P\hat{M}}^\sigma(i, j, l) = \text{corrcoef}(P_{i,j}^\sigma, \hat{M}(l)_{i,j}^\sigma) \quad (11)$$

$$D_\rho \triangleq 1 - \overline{\rho_{P\hat{M}}^\sigma} \quad (12)$$

The choice of the parameter  $\sigma$  is of critical importance. It was suggested to choose the resolution ratio as the value of  $\sigma$ . The optimal value for  $D_\rho$  is 0, corresponding to perfect correlation.

The jointed indexes QNR and HQNR and the spatial index  $D_\rho$  were considered in this work. The values of the spatial index  $D_s$  and the spectral indexes  $D_\lambda$  and  $D_\lambda^k$  were presented together to evaluate the effectiveness of these indexes. The toolboxes provided in [51,71] were employed to calculate the indexes.

### 3. Experiments and Results

#### 3.1. Results for the Beijing Dataset

The quality indices of the fused images of the degraded and original scales of the Beijing dataset are shown in Tables 6 and 7, respectively. Figures 2 and 3 show the original and fused images of the two-scale datasets, respectively. The fused products of the degraded datasets are shown to evaluate the consistency between visual inspection and quality indexes.

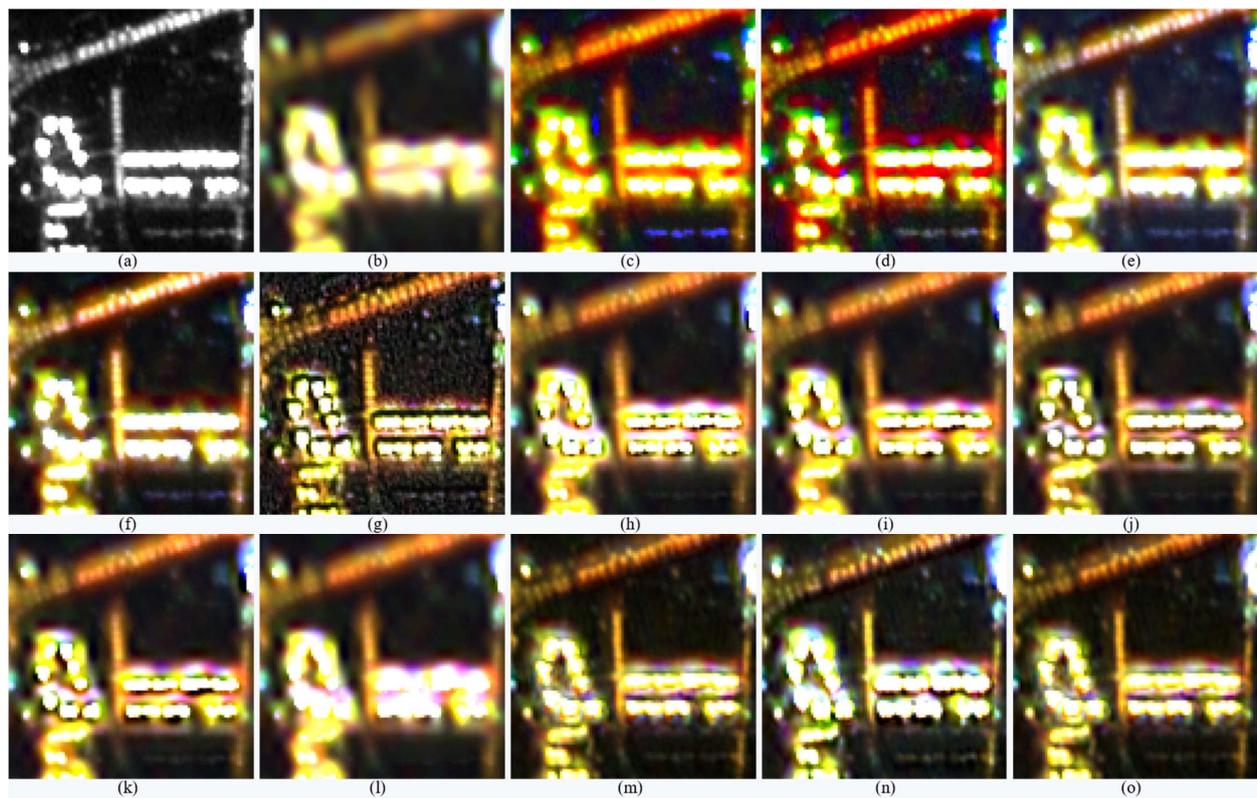
**Table 6.** Quality indices of fused products of the Beijing dataset at the reduced scale. The symbol  $\downarrow$  denotes the lower the index value, the better the performance;  $\uparrow$  means the reverse.

Method	ERGAS $\downarrow$	SAM $\downarrow$	UIQI $\uparrow$	SCC $\uparrow$	SSIM $\uparrow$	PSNR $\uparrow$
IHS	25.749	11.539	0.582	0.522	0.658	24.117
PCA	23.637	9.625	0.584	0.528	0.665	23.910
GSA	21.987	9.777	0.819	0.589	0.870	25.541
RMI	19.307	8.532	0.807	0.624	0.871	26.208
HR	17.289	8.907	0.854	0.592	0.895	26.770
HPF	18.494	8.818	0.792	0.522	0.853	26.166
ATWT	17.550	8.856	0.812	0.539	0.869	26.560
GLP	17.384	8.876	0.816	0.543	0.873	26.599
GLP_HPM	16.360	8.584	0.846	0.593	0.890	27.012
GLP_CBD	17.990	8.785	0.776	0.522	0.838	26.100
A-PNN	12.159	8.803	0.930	0.784	0.964	30.813
PanNet	9.287	8.199	0.952	0.864	0.974	33.029
Z-PNN	17.312	9.886	0.866	0.629	0.919	27.191
EXP	25.309	8.535	0.609	0.205	0.751	23.530

**Table 7.** Quality indices of fused products of the Beijing dataset at the original scale. The symbol  $\downarrow$  denotes the lower the index value, the better the performance;  $\uparrow$  means the reverse.

Method	$D_\lambda$ $\downarrow$	$D_s$ $\downarrow$	QNR $\uparrow$	$D_\lambda^k$ $\downarrow$	HQNR $\uparrow$	$D_p$ $\downarrow$
IHS	0.465	0.017	0.526	0.417	0.573	0.254
PCA	0.351	0.019	0.636	0.456	0.533	0.228
GSA	0.035	0.012	0.953	0.215	0.775	0.141
RMI	0.011	0.016	0.973	0.223	0.765	0.158
HR	0.211	0.003	0.787	0.239	0.759	0.149
HPF	0.025	0.009	0.966	0.027	0.964	0.296
ATWT	0.035	0.009	0.956	0.029	0.962	0.275
GLP	0.035	0.009	0.956	0.033	0.958	0.270
GLP_HPM	0.031	0.010	0.959	0.033	0.957	0.285
GLP_CBD	0.020	0.008	0.972	0.024	0.968	0.440
A-PNN	0.109	0.082	0.817	0.491	0.467	0.365
PanNet	0.097	0.087	0.825	0.490	0.465	0.271
Z-PNN	0.081	0.088	0.838	0.441	0.509	0.094

**Figure 2.** The degraded PAN, original MS images, and fused products of the degraded Beijing dataset. (a) PAN image at 40 m, (b) original MS image at 40 m, and fused products of IHS (c), PCA (d), GSA (e), RMI (f), HR (g), HPF (h), ATWT (i), GLP\_HPM (j), GLP (k), GLP\_CBD (l), A-PNN (m), PanNet (n), and Z-PNN (o).



**Figure 3.** The original PAN, upsampled MS images, and fused products of the original-scale Beijing dataset. (a) The PAN image of 10 m, (b) the upsampled version of the 40 m MS image, and the fused products of IHS (c), PCA (d), GSA (e), RMI (f), HR (g), HPF (h), ATWT (i), GLP\_HPM (j), GLP (k), GLP\_CBD (l), A-PNN (m), PanNet (n), and Z-PNN (o).

As shown in Table 6, the three CNN-based methods yielded higher UIQI, SCC, SSIM, and PSNR values than all the other methods, indicating their excellent performance at the degraded scale. The higher SCC and SSIM values indicate that they achieved better preservation of spatial details. A lower SAM value means less spectral distortion. Among the three CNN-based methods, PanNet provided the highest UIQI, SCC, SSIM, and PSNR values and the lowest RASE, ERGAS, SAM, and RMSE values, indicating that PanNet showed outstanding performances in terms of both the spatial detail enhancement and spectral fidelity. The ERGAS and SAM values of the Z-PNN were higher than some of the traditional fusion methods, indicating a relatively high spectral distortion. According to the visual inspection of Figure 2, the fused products of the three CNN-based methods were very close to the original MS image in both spectral fidelity and spatial detail enhancement resolution, which is consistent with the quantitative evaluation indexes. Among the CS-based methods, GSA, HR, and RMI had higher UIQI values than IHS and PCA. The former also yielded significantly higher SCC and SSIM values than the latter, indicating a better spatial consistency. As shown in Figure 2, the fused products of the former showed lower spectral distortions than those of the latter, which is consistent with the relatively high UIQI that was provided by the former. Among the MRA-based methods, GLP\_HPM proved to have the highest UIQI value, followed by GLP, ATWT, and HPF, whereas GLP\_CBD yielded the lowest UIQI value. According to the visual inspection, the fused image of GLP\_HPM was clearer than the other products. The fused product of GLP\_CBD looked blurred due to a lack of details, which is consistent with the low SCC and SSIM values of GLP\_CBD.

As seen from Table 7, the MRA-based method provided the highest QNR and HQNR values and the lowest  $D_{\lambda}^k$  values. However, the  $D_{\rho}$  values of MRA-based methods were relatively high, indicating a poor preservation of spatial details. Among the MRA methods, GLP\_CBD yielded a higher QNR and HQNR and lower  $D_{\lambda}^k$ . However, the fused product

of GLP\_CBD shown in Figure 3 was the most blurred, indicating that the evaluations in terms of QNR, HQNR, and  $D_\lambda^k$  were inconsistent with the visual comparison.

Among the CS-based methods, GSA and RMI yielded relatively high QNR and HQNR values, along with relatively low  $D_\lambda^k$  and  $D_\rho$  values. According to a visual inspection of Figure 3, the fused products of GSA and RMI presented lower spectral distortions and more spatial details than those of PCA and IHS. Offering the lowest  $D_s$  and  $D_\rho$  values, the HR-fused image also showed rich spatial details, but it had noticeable spectral distortions.

According to the visual inspection, the fused images of the three CNN-based methods showed unnoticeable spectral distortions, as well as more spatial details than the MRA-based methods. The QNR and HQNR of Z-PNN were the highest among the CNN-based methods, while the  $D_\rho$  was the lowest, which is also significantly lower than those of all the other methods. However, PanNet yielded better visual effects than Z-PNN and A-PNN. Therefore, it can be inferred that the performance evaluated in terms of  $D_\rho$  did not match the visual inspection.

Combining quantitative and visual comparisons on both the degraded and the original scales, PanNet offered more robust performances than the other methods. At the degraded scale, for PanNet, HR, and GLP\_HPM, the performances evaluated using quantitative indexes were consistent with the visual inspection. For the original scale, PanNet, GLP\_HPM, RMI, and HR yielded better visual effects than the other methods.

### 3.2. Results for the Brazil Dataset

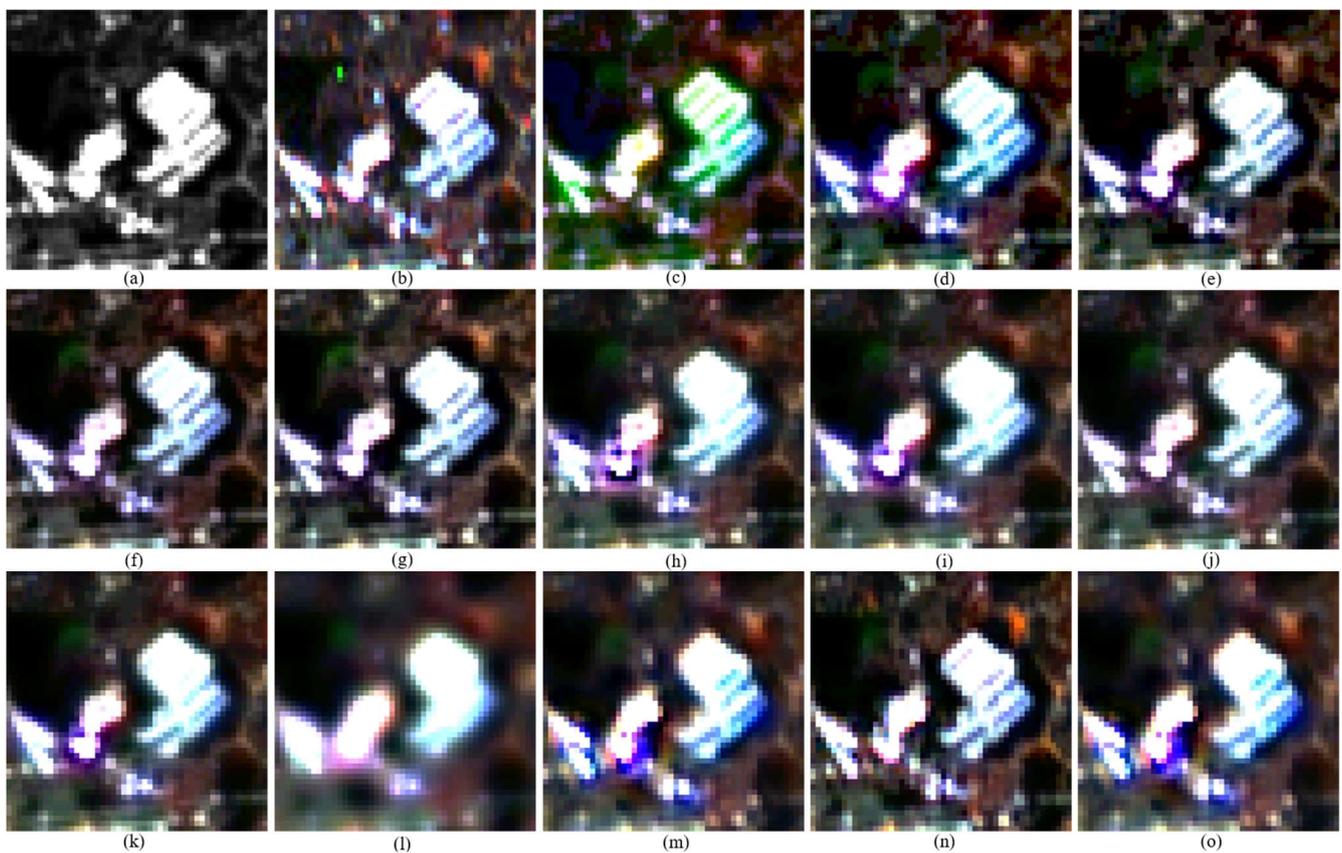
The quality indices of the fused images of the degraded and original scales of the Brazil dataset are shown in Tables 8 and 9, respectively. Figures 4 and 5 show the original images and fused images generated from the degraded and original scales, respectively.

**Table 8.** Quality indices of fused products of the Brazil dataset at the reduced scale. The symbol ↓ denotes the lower the index value, the better the performance; ↑ means the reverse.

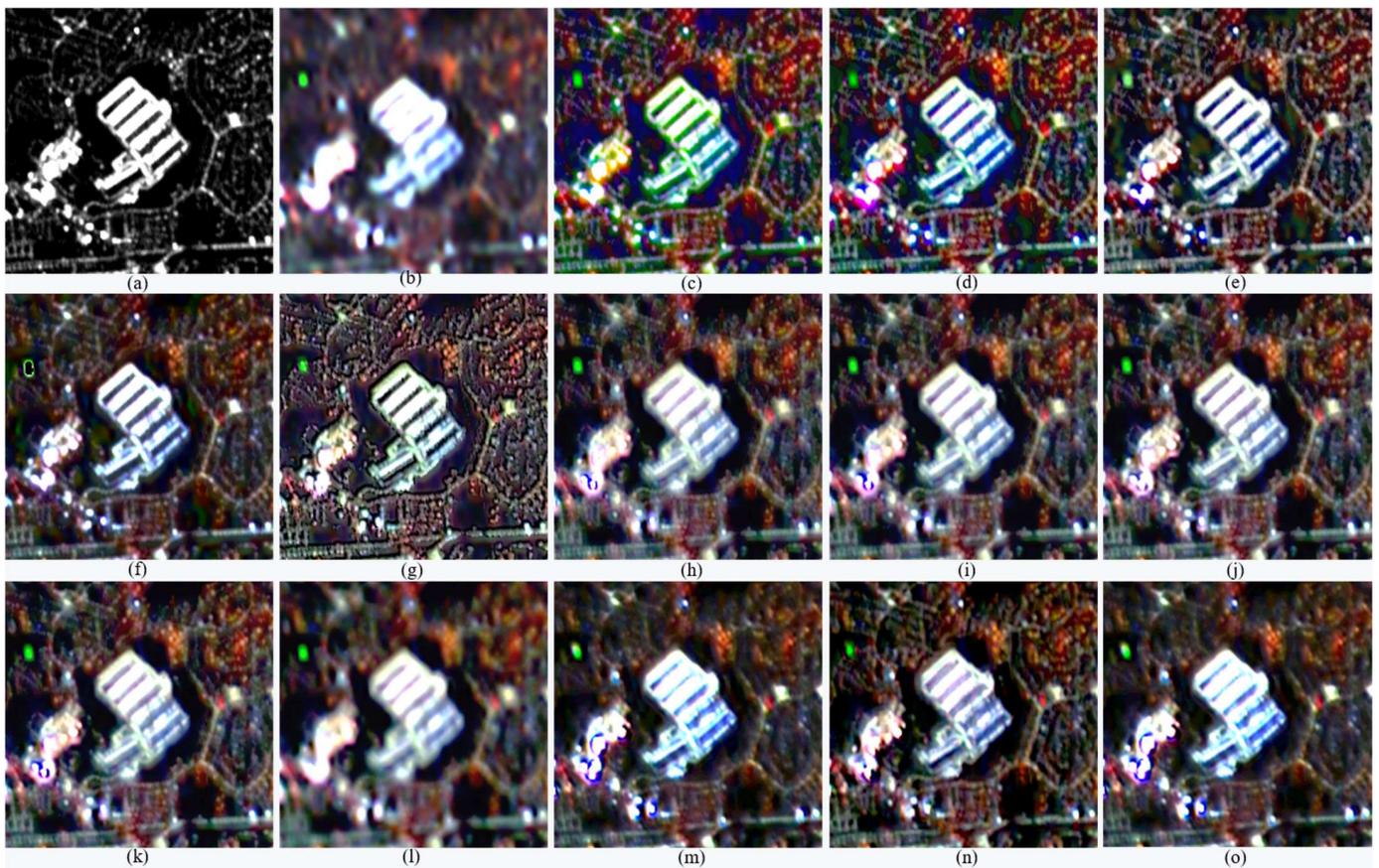
Method	ERGAS ↓	SAM ↓	UIQI ↑	SCC ↑	SSIM ↑	PSNR ↑
IHS	15.120	11.002	0.646	0.557	0.780	29.163
PCA	14.034	9.333	0.666	0.586	0.795	29.259
GSA	14.044	9.877	0.774	0.597	0.876	29.615
RMI	15.120	11.002	0.646	0.557	0.780	29.163
HR	11.196	8.952	0.797	0.589	0.917	31.294
HPF	11.305	8.943	0.794	0.587	0.914	31.210
ATWT	10.318	8.741	0.821	0.633	0.928	31.901
GLP	13.894	8.865	0.671	0.516	0.849	28.655
GLP_HPM	11.779	8.690	0.790	0.593	0.906	30.853
GLP_CBD	16.009	8.690	0.579	0.182	0.682	22.255
A-PNN	17.151	18.711	0.701	0.493	0.802	29.214
PanNet	6.134	8.410	0.928	0.866	0.976	36.884
Z-PNN	11.780	9.766	0.786	0.583	0.907	11.780
EXP	14.034	9.333	0.666	0.586	0.795	29.259

**Table 9.** Quality indexes of fused products of the Brazil dataset at the original scale. The symbol  $\downarrow$  denotes the lower the index value, the better the performance;  $\uparrow$  means the reverse.

Method	$D_\lambda \downarrow$	$D_s \downarrow$	QNR $\uparrow$	$D_\lambda^k \downarrow$	HQNR $\uparrow$	$D_\rho \downarrow$
IHS	0.093	0.008	0.900	0.466	0.530	0.288
PCA	0.066	0.010	0.924	0.449	0.546	0.277
GSA	0.097	0.003	0.901	0.377	0.621	0.257
RMI	0.083	0.005	0.912	0.322	0.675	0.271
HR	0.212	0.001	0.787	0.261	0.738	0.391
HPF	0.054	0.006	0.941	0.042	0.952	0.363
ATWT	0.067	0.006	0.928	0.048	0.946	0.332
GLP	0.069	0.006	0.926	0.048	0.946	0.331
GLP_HPM	0.061	0.006	0.933	0.047	0.948	0.325
GLP_CBD	0.013	0.004	0.983	0.039	0.957	0.598
A-PNN	0.008	0.014	0.978	0.481	0.512	0.479
PanNet	0.051	0.029	0.921	0.555	0.432	0.443
Z-PNN	0.040	0.035	0.926	0.496	0.487	0.312



**Figure 4.** The degraded PAN image, the original MS image, and fused products of the degraded Brazil dataset. (a) The PAN image of 40 m, (b) the original MS image of 40 m, and the fused products of IHS (c), PCA (d), GSA (e), RMI (f), HR (g), HPF (h), ATWT (i), GLP\_HPM (j), GLP (k), GLP\_CBD (l), A-PNN (m), PanNet (n), and Z-PNN (o).



**Figure 5.** The original PAN, upsampled MS images, and fused products of the original-scale Brazil dataset. (a) The PAN image of 10 m, (b) the upsampled version of the 40 m MS image, and the fused products of IHS (c), PCA (d), GSA (e), RMI (f), HR (g), HPF (h), ATWT (i), GLP\_HPM (j), GLP (k), GLP\_CBD (l), A-PNN (m), PanNet (n), and Z-PNN (o).

As shown in Table 8, PanNet provided the highest UIQI, SCC, SSIM, and PSNR values and the lowest ERGAS and SAM values among all the methods. The fused product of PanNet shown in Figure 4 was very close to the original MS image in both spectral and spatial and showed more spatial details than those of A-PNN and Z-PNN. Therefore, the visual comparison is completely consistent with the quantitative evaluation index. Among the CS methods, GSA and HR provided higher UIQI values than the other three methods. The GSA-fused product was relatively close to the original MS image, whereas the IHS-fused image showed noticeable spectral distortions. Among the MRA methods, ATWT provided the highest UIQI, followed by HPF, GLP\_HPM, and GLP, whereas GLP\_CBD yielded the lowest UIQI. The SCC of GLP\_CBD was also significantly lower than the other methods, indicating the lack of spatial details in the fused image. The fused product of GLP\_CBD was also blurred, which is consistent with the low SCC value of the method.

As seen from Table 9, the MRA-based methods yielded relatively high QNR and HQNR, relatively low  $D_s$ , and the lowest  $D_\lambda^k$  values. Among the three MRA-based methods, GLP\_CBD provided the highest QNR and HQNR, the lowest  $D_\lambda^k$ , but also the highest  $D_\rho$  values, which theoretically equates to a poor performance in spatial detail preservation. However, the fused product of GLP\_CBD shown in Figure 5 was the most blurred, indicating that the evaluation using QNR, HQNR, and  $D_\lambda^k$  might be inconsistent with the visual inspection. Among the CS-based methods, PCA and RMI provided the highest QNR, HR yielded the highest HQNR, whereas GSA had the lowest  $D_\rho$ . However, according to the visual inspection, the GSA-fused image balanced better between spectral fidelity and spatial details compared with the fused products of PCA and RMI. The HR-fused image showed

the richest spatial details, although it provided the highest  $D_\rho$  value, which theoretically indicates a lack of spatial details. Among the CNN-based methods, A-PNN yielded a higher QNR and HQNR than the other methods, and it also provided the highest  $D_\rho$  value. The three CNN-based fused products showed unnoticeable spectral distortions and more spatial details than the MRA-based methods. The fused image of PanNet showed a better visual effect in terms of both spectral and spatial parameters than A-PNN and Z-PNN, indicating that the assessment based on QNR, HQNR, and  $D_\rho$  values was inconsistent with the visual inspection.

Combining quantitative and visual comparisons on both the degraded and original scales, PanNet, GSA, and HR performed better than the other methods in the fusion of the two scales. The MRA-based fusion products were relatively spatially blurred, and the assessment based on QNR, HQNR, and  $D_\rho$  values was inconsistent with the visual inspection.

### 3.3. Results for the Lisbon Dataset

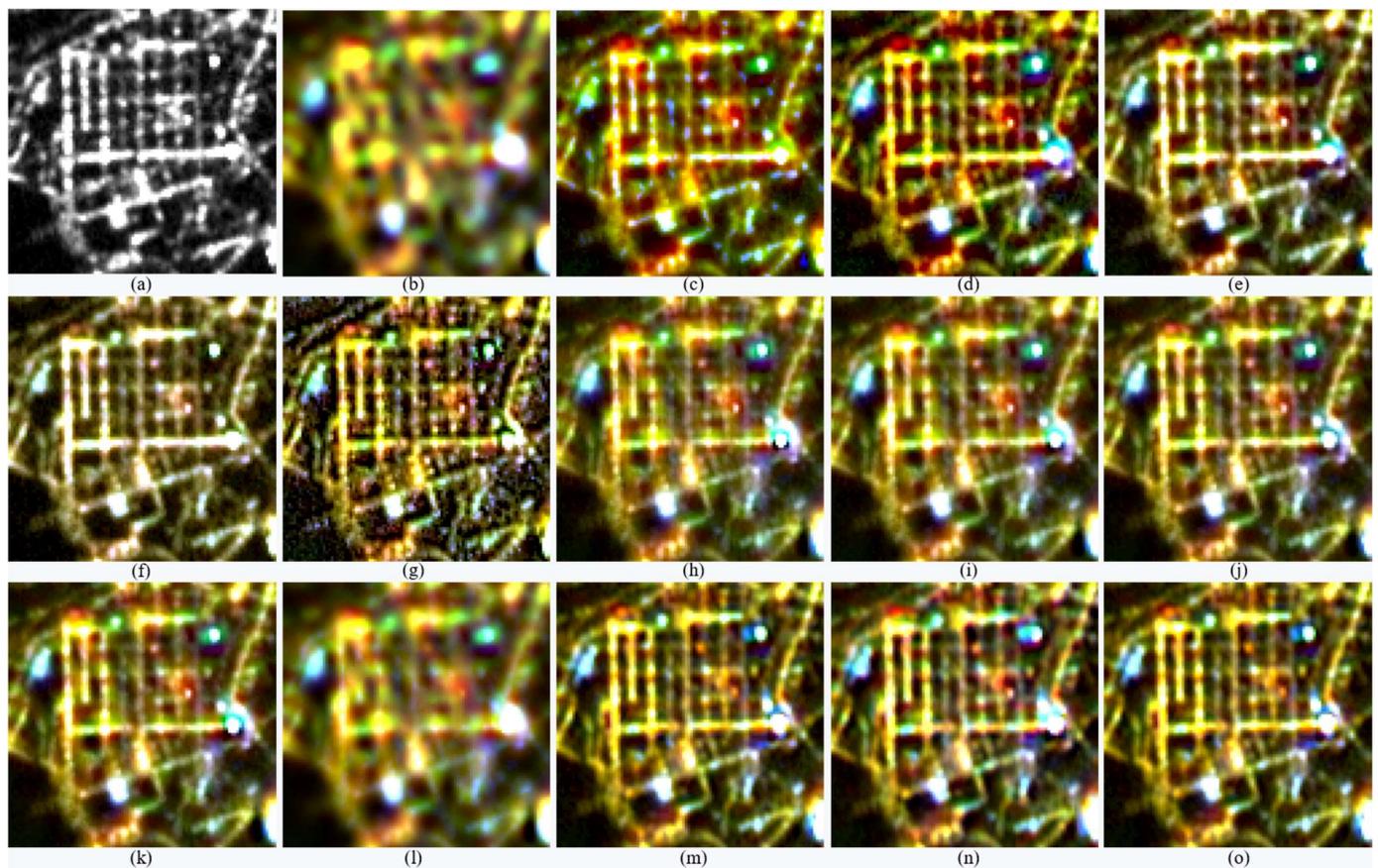
The quality indices of the fused images of the degraded and original Lisbon datasets are shown in Tables 10 and 11, respectively. Figure 6 shows the original Pan and upsampled MS images and the fused products of the original-scale dataset.

**Table 10.** Quality indices of fused products of the Lisbon dataset at the reduced scale. The symbol  $\downarrow$  denotes the lower the index value, the better the performance;  $\uparrow$  means the reverse.

Method	ERGAS $\downarrow$	SAM $\downarrow$	UIQI $\uparrow$	SCC $\uparrow$	SSIM $\uparrow$	PSNR $\uparrow$
IHS	37.421	13.805	0.619	0.694	0.735	29.684
PCA	28.072	15.258	0.653	0.756	0.796	29.666
GSA	24.244	15.455	0.839	0.781	0.967	33.083
RMI	24.502	13.717	0.770	0.777	0.946	32.594
HR	20.541	14.247	0.848	0.779	0.964	33.260
HPF	25.456	14.117	0.807	0.738	0.945	31.841
ATWT	24.749	14.312	0.823	0.751	0.956	32.459
GLP	24.403	14.033	0.827	0.752	0.959	32.633
GLP_HPM	19.011	13.829	0.862	0.799	0.971	34.197
GLP_CBD	26.766	13.803	0.712	0.594	0.882	28.706
A-PNN	18.777	14.531	0.880	0.823	0.982	35.663
PanNet	14.219	12.492	0.907	0.888	0.987	37.185
Z-PNN	25.077	16.884	0.781	0.721	0.943	32.384
EXP	34.892	13.717	0.601	0.194	0.839	26.317

**Table 11.** Quality indices of fused products of the Lisbon dataset at the original scale. The symbol  $\downarrow$  denotes the lower the index value, the better the performance;  $\uparrow$  means the reverse.

Method	$D_\lambda$ $\downarrow$	$D_s$ $\downarrow$	QNR $\uparrow$	$D_\lambda^k$ $\downarrow$	HQNR $\uparrow$	$D_\rho$ $\downarrow$
IHS	0.066	0.006	0.929	0.612	0.385	0.168
PCA	0.036	0.003	0.961	0.486	0.512	0.154
GSA	0.092	0.004	0.904	0.342	0.656	0.106
RMI	0.073	0.004	0.923	0.385	0.612	0.166
HR	0.433	0.003	0.565	0.451	0.547	0.153
HPF	0.079	0.005	0.916	0.097	0.899	0.180
ATWT	0.088	0.005	0.908	0.111	0.885	0.156
GLP	0.088	0.004	0.908	0.150	0.846	0.155
GLP_HPM	0.081	0.004	0.915	0.146	0.850	0.160
GLP_CBD	0.011	0.003	0.985	0.063	0.934	0.565
A-PNN	0.057	0.114	0.836	0.792	0.185	0.498
PanNet	0.061	0.115	0.830	0.793	0.183	0.416
Z-PNN	0.069	0.218	0.728	0.780	0.172	0.192



**Figure 6.** The original PAN, upsampled MS images, and fused products of the original-scale Lisbon dataset. (a) The PAN image of 10 m, (b) the upsampled version of the 40 m MS image, and the fused products of IHS (c), PCA (d), GSA (e), RMI (f), HR (g), HPF (h), ATWT (i), GLP\_HPM (j), GLP (k), GLP\_CBD (l), A-PNN (m), PanNet (n), and Z-PNN (o).

As seen from Table 10, PanNet provided higher UIQI, SCC, SSIM, and PSNR values and significantly lower ERGAS and SAM values than those of all the other methods. This indicates that the fused product of PanNet shows small spectral distortion as well as good preservation of spatial details. Among the CS-based methods, GSA and HR offered higher UIQI, SCC, and SSIM values than IHS, PCA, and RMI. Among the MRA-based methods, GLP\_HPM yielded the highest UIQI, SCC, and SSIM values, whereas GLP\_CBD offered the lowest UIQI. The SCC of GLP\_CBD was significantly lower than for the other methods, indicating a lack of spatial details. This is very similar to the results for the Brazil dataset.

As shown in Table 11, GLP\_CBD offered higher QNR and HQNR values than the other two MRA-based methods, due to the significantly lower  $D_\lambda$  and  $D_\lambda^k$  values provided by the method. However, GLP\_CBD also yielded a significantly higher  $D_\rho$  value than the other methods, indicating a relatively low spatial consistency with the original PAN image. The fused image of GLP\_CBD was the most spatially blurred, which is consistent with the high  $D_\rho$  value. Among the CS-based methods, PCA yielded the highest QNR, whereas GSA offered the highest HQNR and the lowest  $D_\lambda^k$  and  $D_\rho$  values. However, the fused images of GSA and RMI showed noticeable spectral distortions, and that of PCA yielded a better visual effect. The three CNN-based methods offered relatively low QNR and HQNR values, due to their relatively high  $D_s$  and  $D_\lambda^k$  values. The QNR and HQNR of A-PNN were higher than those of PanNet and Z-PNN, but the  $D_\rho$  of the former was also higher than the latter. According to visual inspection, the three fused products showed richer spatial details than those of the MRA methods. A-PNN and Z-PNN showed slightly better visual effects than PanNet.

According to the quantitative metrics at the degraded scale, PanNet, A-PNN, GLP\_HPM, HR, and GSA performed better than the other methods. According to the visual inspection, A-PNN and Z-PNN outperformed the other methods at the original scale.

### 3.4. Results for the Shanghai Dataset

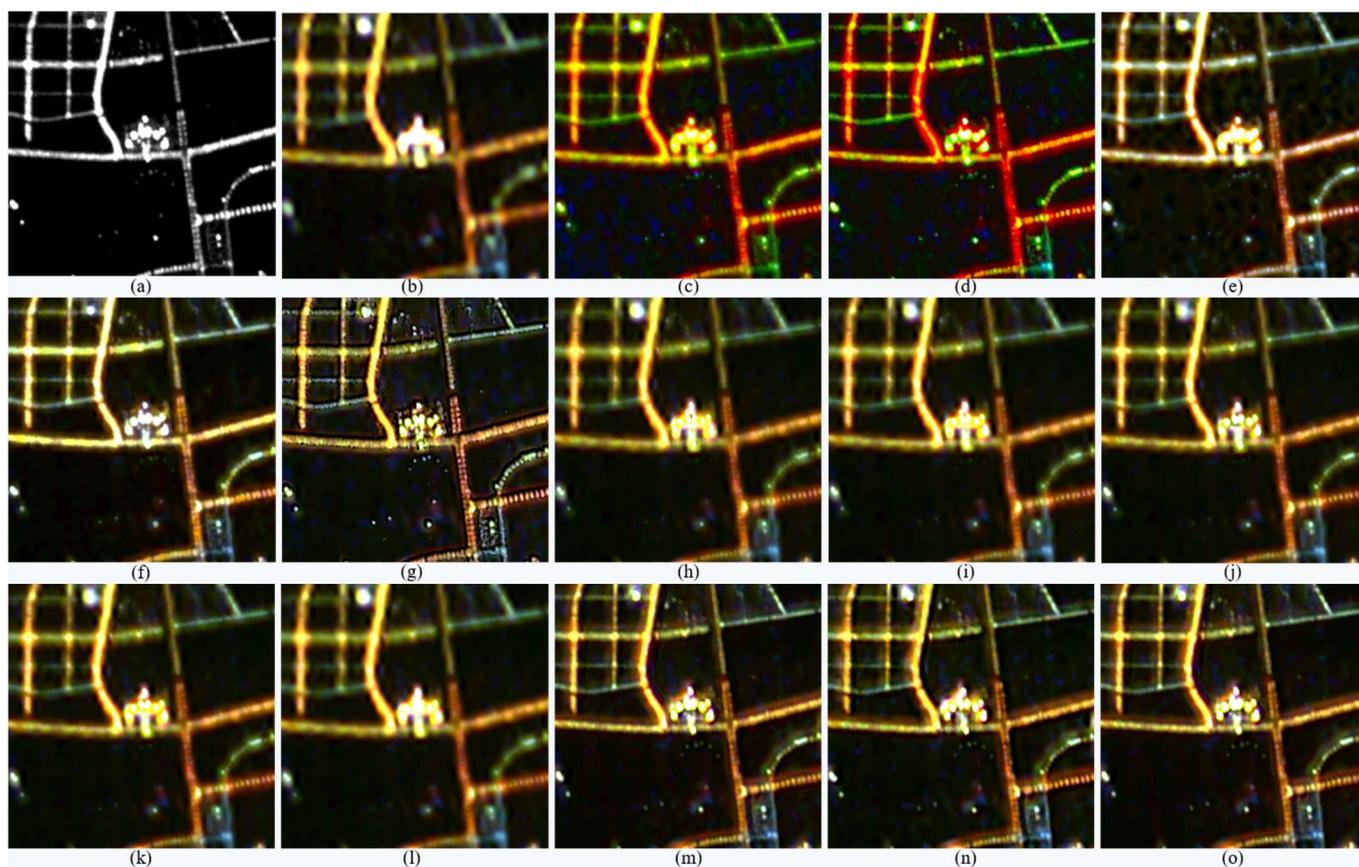
The quality indices of the fused images of the degraded and original Shanghai datasets are shown in Tables 12 and 13, respectively. Figure 7 shows the original Pan and upsampled MS images and the fused products of the original-scale dataset.

**Table 12.** Quality indices of fused products of the Shanghai dataset at the reduced scale. The symbol ↓ denotes the lower the index value, the better the performance; ↑ means the reverse.

Method	ERGAS ↓	SAM ↓	UIQI ↑	SCC ↑	SSIM ↑	PSNR ↑
IHS	31.057	13.215	0.450	0.574	0.709	27.916
PCA	28.578	10.710	0.452	0.595	0.721	27.690
GSA	22.660	10.901	0.785	0.675	0.912	30.767
RMI	21.720	9.306	0.792	0.725	0.927	31.132
HR	19.713	10.026	0.870	0.783	0.954	31.103
HPF	22.560	9.535	0.747	0.578	0.901	29.942
ATWT	21.205	9.565	0.764	0.600	0.910	30.409
GLP	20.800	9.582	0.769	0.604	0.914	30.507
GLP_HPM	17.409	9.388	0.816	0.688	0.935	31.881
GLP_CBD	22.593	9.519	0.714	0.548	0.884	29.356
A-PNN	13.787	9.126	0.903	0.825	0.893	35.811
PanNet	10.194	8.572	0.924	0.900	0.907	38.363
Z-PNN	19.206	11.432	0.818	0.702	0.808	32.077
EXP	31.666	9.306	0.577	0.209	0.833	27.460

**Table 13.** Quality indices of fused products of the Shanghai dataset at the original scale. The symbol ↓ denotes the lower the index value, the better the performance; ↑ means the reverse.

Method	$D_\lambda$ ↓	$D_s$ ↓	QNR ↑	$D_\lambda^k$ ↓	HQNR ↑	$D_p$ ↓
IHS	0.551	0.010	0.444	0.442	0.552	0.618
PCA	0.430	0.011	0.563	0.491	0.504	0.606
GSA	0.053	0.006	0.941	0.291	0.704	0.561
RMI	0.041	0.008	0.952	0.260	0.735	0.571
HR	0.226	0.002	0.772	0.229	0.769	0.560
HPF	0.017	0.004	0.979	0.032	0.964	0.635
ATWT	0.025	0.004	0.971	0.031	0.965	0.625
GLP	0.024	0.004	0.971	0.033	0.963	0.621
GLP_HPM	0.023	0.004	0.973	0.033	0.963	0.615
GLP_CBD	0.014	0.004	0.983	0.032	0.964	0.714
A-PNN	0.095	0.053	0.857	0.677	0.306	0.638
PanNet	0.083	0.049	0.872	0.681	0.303	0.594
Z-PNN	0.111	0.075	0.822	0.677	0.299	0.439



**Figure 7.** The original PAN, upsampled MS images, and fused products of the original-scale Shanghai dataset. (a) The PAN image of 10 m, (b) the upsampled version of the 40 m MS image, and the fused products of IHS (c), PCA (d), GSA (e), RMI (f), HR (g), HPF (h), ATWT (i), GLP\_HPM (j), GLP (k), GLP\_CBD (l), A-PNN (m), PanNet (n), and Z-PNN (o).

As seen from Table 12, PanNet offered the highest UIQI, SCC, SSIM, and PSNR and the lowest ERGAS and SAM among all the methods, indicating an outstanding performance in terms of both spectral fidelity and spatial detail enhancement. Among the CS methods, the UIQI, SCC, and SSIM values of HR were significantly higher than those of the other four methods. Among the MRA-based methods, GLP\_HPM offered a higher UIQI, SCC, and SSIM, whereas GLP\_CBD yielded the lowest. The SCC and SSIM values of the GLP\_CBD method were significantly lower than those of the other methods, indicating a poor performance in spatial detail preservation.

As seen from Table 13, the QNR values of the MRA-based fusion methods were higher than those of the CS- and CNN-based methods. However, the fused products of the CNN-based methods (Figure 7) showed more spatial details than those of the MRA-based methods. Among the MRA-based methods, GLP\_CBD yielded the highest QNR value, due to a  $D_s$  of 0.004, which is significantly lower than for the other methods. However, the  $D_\rho$  of GLP\_CBD was also the highest, which theoretically indicates that the fused product showed low spatial consistency with the original PAN image. According to the visual inspection, the fused product of GLP\_CBD was very blurred, which is consistent with the high  $D_\rho$  value provided by the method. Among the CS-based methods, RMI offered the highest QNR, whereas HR yielded the highest HQNR. According to the visual inspection, the fused products of IHS, PCA, and GSA showed spectral distortions, and the spatial details in the RMI-fused product were very close to those in the original PAN image. The HR-fused image showed more spatial details than that of RMI, which is consistent with the low  $D_s$  that HR offered. Among the CNN-based methods, PanNet offered the highest QNR, A-PNN yielded the highest HQNR, and Z-PNN provided the lowest  $D_\rho$ . According

to the visual inspection, spectral distortions of the three fused products were not distinct, and the spatial details were close to those presented in the original PAN image.

Combining the quantitative indicators of the two scales with visual comparisons, the GLP\_HPM outperformed the other MRA-based methods, whereas RMI and HR outperformed the other CS-based methods. The CNN-based methods yielded different performances across the two scales.

## 4. Discussion

### 4.1. Comparisons of Pansharpening Methods for SDGSAT-1 GI Data

The pansharpening methods yielded relatively robust performances at the degraded scale among the four datasets. The CNN-based methods yielded better performances than the other methods at the degraded scale according to the quality metrics, as well as the visual inspection. The three CNN methods also achieved more robust and better visual effects than the MRA methods at the original scale. This is mainly due to the powerful learning ability of CNN-based models, which successfully achieve the enhancement of spatial details while ensuring spectral consistency. Similar to the fusion of optical remote sensing images, the CNNs showed remarkable advantages over traditional methods and great potential in the pansharpening of the GI imagery of SDGSAT-1. PanNet yielded a more stable and excellent performance than A-PNN and Z-PNN, mainly due to the deeper network structure employed by PanNet. As introduced in Section 2, pre-trained models that are trained using additional datasets were used for both A-PNN and Z-PNN. In contrast, only the test image itself was used for the training of the PanNet model, which saves a large amount of extra work and training time. Consequently, PanNet has great potential in the fusion of SDGSAT-1 GI imagery.

A concern about PanNet is the feasibility and generalizability of the model when used for other datasets, which is a challenge when considering remote sensing images that are recorded in different seasons from multiple satellite platforms with different spatial resolutions. Actually, PanNet was approved with outstanding generalization ability [53]. In this work, we primarily focused on the effectiveness of the CNN models working on SDGSAT-1 NL images. Different from daytime multispectral images, NL images show much simpler backgrounds except for nighttime lights. Consequently, we are optimistic about using the trained PanNet models on other GI NL image scenes covering a different place. Moreover, the performance of PanNet can be further improved through fine-tuning training using the degraded version of the original dataset, which requires just a short running time. Additionally, the trained PanNet models of the four datasets used in this work can be used as pre-trained models for the fusion of other image scenes to further reduce the training time.

Among the CS-based methods, GSA outperformed IHS and PCA in terms of SAM and UIQI at the degraded scale, as well as visual inspection. This may be because both methods are robust to misalignment between the MS and PAN bands [18]. The HR-fused images were richer in spatial details and more suitable for generating fusion images that needed to be used for the exhibition. However, HR-fused images may suffer from spatial distortions, such as ringing artefacts, when the MS and PAN bands are not perfectly aligned. The fused images of IHS and PCA usually had higher spectral distortions and blurrier spatial details. This is very similar to the fusion of optical remote sensing images.

The GLP\_HPM method yielded more robust and significantly better visual effects and higher UIQI, SCC, SSIM, and PNSR values at the degraded scale than the other MRA methods. For the GLP\_HPM method, the injection weight  $g_i$  is the ratio of the  $i$ th MS band and  $P_L$ , which limits the spectral distortions and ensures the injection of enough spatial details. The GLP\_CBD yielded very blurred details concerning lights and networks, although it had relatively high QNR and HQNR values.

#### 4.2. Evaluation of Quality Indices Used for Fusion Products for SDGSAT-1 GI Data

The fused products of the degraded Beijing and Brazil datasets are shown in Figures 2 and 4, respectively, to evaluate the consistency between the visual inspection and the quantitative metrics. The pansharpening method, i.e., PanNet, provided the highest UIQI, SCC, PNSR values and the lowest SAM and ERGAS values, and this also corresponded to an outstanding visual effect. Generally, the performances of the state-of-the-art pansharpening in terms of quality indexes at the degraded scale were highly consistent with the visual inspection.

However, it was found that the quantitative metrics generated from the fused products of the four SDGSAT-1 GI datasets at the original scale were completely different from the performances in terms of visual inspection. For example, the fused products of GLP\_CBD provided very high QNR and HQNR values for the four datasets but yielded very blurred details, which achieved the poorest visual effect. Both  $D_s$  and  $D_\rho$  measure the spatial similarity between a pansharpened product and the original PAN band. The closer the values are to 0, the better the spatial consistency is. Tables 6, 8, 10 and 12 show that the fused products of the MRA methods, RMI, GSA, and HR achieved very similar  $D_s$  values but showed significant visual differences. Some of them had very blurred details regarding lights and road networks. Similarly, a fused image with a relatively low  $D_\rho$  value may even show blurred details. This indicates that the QNR-like indexes were not suitable for the pansharpened SDGSAT-1 GI images.

These different performances of the quality indices are mainly due to the differences between the NL data and the optical multispectral data. The most significant difference is that the NL images contain a large number of dark pixels and lack spatial details of ground objects. Additionally, some objects shown in the GL imagery, such as the 10 m Pan image, may even be discontinuous. Currently, the quality index for the assessment of fused images at the original scale is still an open issue for optical remote sensing imagery. However, it is necessary to consider other metrics measuring the overall quality of pansharpened GI images considering the difference between daytime optical images and NL images.

## 5. Conclusions

This work assessed the performances of thirteen state-of-the-art pansharpening methods on GI NL imagery provided by SADSAT-1 to provide a reference for selecting an optimal pansharpening method for GI data. The fused products of four GI datasets from SDGSAT-1 at both degraded and original scales were compared and analyzed by visual inspection and quantitative indicators. The following conclusions were obtained:

According to the experimental results, the three CNN-based methods (A-PNN, PanNet, and Z-PNN) yielded relatively stable and outstanding performances for the fusion of the four datasets at both the degraded and original scales. Specifically, PanNet offered UIQI values ranging from 0.907 to 0.952 for the four datasets, and the PanNet-fused products that were generated at the degraded scale yielded better visual effects in terms of spectral fidelity and spatial detail enhancement. Among the CS and MRA methods, GSA, HR, and GLP\_HPM provided UIQI values ranging from 0.77 to 0.856 for the four datasets and outperformed other methods in terms of the visual inspection and quantitative index comparison. If only the visual effect is considered, HR-fused images showed the richest spatial details. The CNN-based methods also yielded comparable visual effects at the original scale. PanNet has great potential in the fusion of SDGSAT-1 GI imagery due to its robust performance on the four datasets and relatively short training time.

The quality metrics of the degraded scale were highly consistent with the visual inspection. However, the quality indexes used at the original scale were inconsistent with the visual inspection, especially spatial indexes such as  $D_s$  and  $D_\rho$ . Although many efforts have been made to achieve full-resolution quality assessment, it is a challenge due to the absence of a ground truth image as a reference. It is urgent to explore quality metrics at full resolution, measuring the overall quality of pansharpened GI images considering the difference between daytime optical images and NL images. Although PanNet showed

great potential in the pansharpening of GI imagery, how to obtain a model with a good generalization ability is still a problem that remains to be explored. Further improvements can also be made through exploring other advanced CNN models using advanced loss functions and measuring similarities at full resolution. As the alignment between the MS and PAN bands is a crucial factor of image fusion, it would be very useful for developing a pansharpening method integrating accurate registration and image fusion.

**Author Contributions:** H.L. drafted the manuscript and was responsible for the research design, experiments, and the analysis. L.J. provided technical guidance and reviewed the manuscript. C.D. and H.D. processed the GI images and reviewed the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the International Research Center of Big Data for Sustainable Development Goals (CBAS2022RP04) and the National Key Research and Development Program of China (2022YFC3800701).

**Data Availability Statement:** The data acquired by SDGSAT-1 are available to the scientific community globally free of charge through the SDGSAT-1 Open Science Program ([www.sdgsat.ac.cn](http://www.sdgsat.ac.cn) (accessed on 8 September 2022)).

**Acknowledgments:** It is acknowledged that the SDGSAT-1 GI data are kindly provided by the International Research Center of Big Data for Sustainable Development Goals (CBAS). The authors would like to thank the editors and anonymous reviewers for their detailed review, valuable comments and constructive suggestions.

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

## References

1. SDGeHandbook. Available online: <https://unstats.un.org/wiki/display/SDGeHandbook?preview=/34505092/106497383/SDGeHandbook-111121-2121-805.pdf> (accessed on 5 June 2022).
2. Indicators List. Available online: <https://unstats.un.org/sdgs/indicators/indicators-list/> (accessed on 5 June 2022).
3. Levin, N.; Kyba, C.C.M.; Zhang, Q.; de Miguel, A.S.; Román, M.O.; Li, X.; Portnov, B.A.; Molthan, A.L.; Jechow, A.; Miller, S.D.; et al. Remote sensing of night lights: A review and an outlook for the future. *Remote Sens. Environ.* **2020**, *237*, 111443. [CrossRef]
4. Zhou, Y.; Smith, S.J.; Elvidge, C.D.; Zhao, K.; Thomson, A.; Imhoff, M. A cluster-based method to map urban area from DMSP/OLS nightlights. *Remote Sens. Environ.* **2014**, *147*, 173–185. [CrossRef]
5. Liu, X.; de Sherbinin, A.; Zhan, Y. Mapping urban extent at large spatial scales using machine learning methods with VIIRS nighttime light and MODIS daytime NDVI data. *Remote Sens.* **2019**, *11*, 1247. [CrossRef]
6. Zhang, G.; Guo, X.; Li, D.; Jiang, B. Evaluating the potential of LJ1-01 nighttime light data for modeling socio-economic parameters. *Sensors* **2019**, *19*, 1465. [CrossRef]
7. Liu, H.; Luo, N.; Hu, C. Detection of county economic development using LJ1-01 nighttime light imagery: A comparison with NPP-VIIRS data. *Sensors* **2020**, *20*, 6633. [CrossRef]
8. Yao, F.; Wu, J.; Li, W.; Peng, J. A spatially structured adaptive two-stage model for retrieving ground-level PM<sub>2.5</sub> concentrations from VIIRS AOD in china. *ISPRS J. Photogramm. Remote Sens.* **2019**, *151*, 263–276. [CrossRef]
9. Zhang, G.; Shi, Y.; Xu, M. Evaluation of LJ1-01 nighttime light imagery for estimating monthly PM<sub>2.5</sub> concentration: A comparison with NPP-VIIRS nighttime light data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 3618–3632. [CrossRef]
10. Zhao, M.; Zhou, Y.; Li, X.; Cao, W.; He, C.; Yu, B.; Li, X.; Elvidge, C.D.; Cheng, W.; Zhou, C. Applications of satellite remote sensing of nighttime light observations: Advances, challenges, and perspectives. *Remote Sens.* **2019**, *11*, 1971. [CrossRef]
11. Rybnikova, N.; Portnov, B.A.; Mirkes, E.M.; Zinoviyev, A.; Brook, A.; Gorban, A.N. Coloring panchromatic nighttime satellite images: Comparing the performance of several machine learning methods. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–15. [CrossRef]
12. Zheng, Q.; Weng, Q.; Huang, L.; Wang, K.; Deng, J.; Jiang, R.; Ye, Z.; Gan, M. A new source of multi-spectral high spatial resolution night-time light imagery—jl1-3b. *Remote Sens. Environ.* **2018**, *215*, 300–312. [CrossRef]
13. Levin, N.; Johansen, K.; Hacker, J.M.; Phinn, S. A new source for high spatial resolution night time images—The EROS-B commercial satellite. *Remote Sens. Environ.* **2014**, *149*, 1–12. [CrossRef]
14. Van Doren, B.M.; Horton, K.G.; Dokter, A.M.; Klinck, H.; Elbin, S.B.; Farnsworth, A. High-intensity urban light installation dramatically alters nocturnal bird migration. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 11175–11180. [CrossRef]
15. User Guide of SDGSAT-1 (Released on July 2022). Available online: <http://124.16.184.48:6008/downresouce> (accessed on 9 July 2022).
16. Carper, W.J.; Lillesand, T.M.; Kiefer, R.W. The use of intensity-hue-saturation transformations for merging SPOT panchromatic and multispectral image data. *Photogramm. Eng. Remote Sens.* **1990**, *56*, 459–467.

17. Shettigara, V.K. A generalized component substitution technique for spatial enhancement of multispectral images using a higher resolution dataset. *Photogramm. Eng. Remote Sens.* **1992**, *58*, 561–567.
18. Tu, T.M.; Lee, Y.C.; Chang, C.P.; Huang, P.S. Adjustable intensity-hue-saturation and Brovey transform fusion technique for IKONOS/QuickBird imagery. *Opt. Eng.* **2005**, *44*, 116201. [[CrossRef](#)]
19. Aiazzi, B.; Baronti, S.; Selva, M. Improving component substitution pansharpening through multivariate regression of MS + PAN data. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 3230–3239. [[CrossRef](#)]
20. Aiazzi, B.; Baronti, S.; Selva, M.; Alparone, L. MS + PAN image fusion by an enhanced gram-schmidt spectral sharpening. In *New Developments and Challenges in Remote Sensing*; Bochenek, Z., Ed.; Millpress: Rotterdam, The Netherlands, 2007; pp. 113–120.
21. Jing, L.; Cheng, Q. Two improvement schemes of pan modulation fusion methods for spectral distortion minimization. *Int. J. Remote Sens.* **2009**, *30*, 2119–2131. [[CrossRef](#)]
22. Jing, L.; Cheng, Q. An image fusion method for misaligned panchromatic and multispectral data. *Int. J. Remote Sens.* **2011**, *32*, 1125–1137. [[CrossRef](#)]
23. Zhong, S.; Zhang, Y.; Chen, Y.; Wu, D. Combining component substitution and multiresolution analysis: A novel generalized BDSB pansharpening algorithm. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 2867–2875. [[CrossRef](#)]
24. Li, H.; Jing, L.; Tang, Y.; Ding, H. An improved pansharpening method for misaligned panchromatic and multispectral data. *Sensors* **2018**, *18*, 557. [[CrossRef](#)]
25. Chavez, P.S.; Sides, S.C.; Anderson, J.A. Comparison of three different methods to merge multiresolution and multispectral data. *Photogramm. Engineering Remote Sens.* **1991**, *57*, 295–303.
26. Shensa, M.J. The discrete wavelet transform: Wedding the a trous and Mallat algorithms. *IEEE Trans. Signal Process.* **1992**, *40*, 2464–2482. [[CrossRef](#)]
27. Aiazzi, B.; Alparone, L.; Barducci, A.; Baronti, S.; Pippi, I. Multispectral fusion of multisensor image data by the generalized Laplacian pyramid. *IEEE Int. Geosci. Remote Sens. Symp.* **1999**, *2*, 1183–1185.
28. Nunez, J.; Otazu, X.; Fors, O.; Prades, A.; Pala, V.; Arbiol, R. Multiresolution-based image fusion with additive wavelet decomposition. *IEEE Trans. Geosci. Remote Sens.* **1999**, *37*, 1204–1211. [[CrossRef](#)]
29. Aiazzi, B.; Alparone, L.; Baronti, S.; Pippi, I.; Selva, M. Generalised Laplacian pyramid-based fusion of MS + P image data with spectral distortion minimisation. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2002**, *34*, 1–4.
30. Aiazzi, B.; Alparone, L.; Baronti, S.; Garzelli, A. Context-driven fusion of high spatial and spectral resolution images based on oversampled multiresolution analysis. *IEEE Trans. Geosci. Remote Sens.* **2002**, *40*, 2300–2312. [[CrossRef](#)]
31. Otazu, X.; Gonzalez-Audicana, M.; Fors, O.; Nunez, J. Introduction of sensor spectral response into image fusion methods: Application to wavelet-based methods. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 2376–2385. [[CrossRef](#)]
32. Aiazzi, B.; Alparone, L.; Baronti, S.; Garzelli, A. MTF-tailored multiscale fusion of high-resolution ms and pan imagery. *Photogramm. Eng. Remote Sens.* **2006**, *72*, 591–596. [[CrossRef](#)]
33. Amolins, K.; Zhang, Y.; Dare, P. Wavelet based image fusion techniques—An introduction, review and comparison. *ISPRS J. Photogramm. Remote Sens.* **2007**, *62*, 249–263. [[CrossRef](#)]
34. Hong, G.; Zhang, Y. Comparison and improvement of wavelet-based image fusion. *Int. J. Remote Sens.* **2008**, *29*, 673–691. [[CrossRef](#)]
35. Bruzzone, L.; Aiazzi, B.; Alparone, L.; Baronti, S.; Garzelli, A.; Selva, M. Advantages of laplacian pyramids over “à trous” wavelet transforms for pansharpening of multispectral images. *Proc. SPIE Image Signal Process. Remote Sens. XVIII* **2012**, *853704*, 12–21.
36. Cheng, J.; Liu, H.; Liu, T.; Wang, F.; Li, H. Remote sensing image fusion via wavelet transform and sparse representation. *ISPRS J. Photogramm. Remote Sens.* **2015**, *104*, 158–173. [[CrossRef](#)]
37. Cao, K.; Zhang, H.; Chen, J.; Zhang, W.; Yu, L. Variational model-based very high spatial resolution remote sensing image fusion. *J. Appl. Remote Sens.* **2014**, *8*, 83565. [[CrossRef](#)]
38. Xiao, Y.; Fang, F.; Zhang, Q.; Zhou, A.; Zhang, G. Parameter selection for variational pan-sharpening by using evolutionary algorithm. *Remote Sens. Lett.* **2015**, *6*, 458–467. [[CrossRef](#)]
39. Zhang, G.; Fang, F.; Zhou, A.; Li, F. Pan-sharpening of multi-spectral images using a new variational model. *Int. J. Remote Sens.* **2015**, *36*, 1484–1508. [[CrossRef](#)]
40. Liu, P.; Xiao, L.; Tang, S. A new geometry enforcing variational model for pan-sharpening. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9*, 5276–5289. [[CrossRef](#)]
41. Duran, J.; Buades, A.; Coll, B.; Sbert, C.; Blanchet, G. A survey of pansharpening methods with a new band-decoupled variational model. *ISPRS J. Photogramm. Remote Sens.* **2017**, *125*, 78–105. [[CrossRef](#)]
42. Masi, G.; Cozzolino, D.; Verdoliva, L.; Scarpa, G. Pansharpening by convolutional neural networks. *Remote Sens.* **2016**, *8*, 594. [[CrossRef](#)]
43. Wei, Y.; Yuan, Q.; Shen, H.; Zhang, L. Boosting the accuracy of multispectral image pansharpening by learning a deep residual network. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 1795–1799. [[CrossRef](#)]
44. Yang, J.; Fu, X.; Hu, Y.; Huang, Y.; Ding, X.; Paisley, J. PanNet: A deep network architecture for pan-sharpening. In *Proceedings of the 2017 IEEE International Conference on Computer Vision, Venice, Italy, 24–27 October 2017*; pp. 1753–1761.
45. Liu, Y.; Chen, X.; Wang, Z.; Wang, Z.J.; Ward, R.K.; Wang, X. Deep learning for pixel-level image fusion: Recent advances and future prospects. *Inf. Fusion* **2018**, *42*, 158–173. [[CrossRef](#)]

46. Scarpa, G.; Vitale, S.; Cozzolino, D. Target-adaptive CNN-based pansharpening. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 5443–5457. [[CrossRef](#)]
47. Li, Z.; Cheng, C. A CNN-based pan-sharpening method for integrating panchromatic and multispectral images using Landsat 8. *Remote Sens.* **2019**, *11*, 2606. [[CrossRef](#)]
48. Jiang, M.; Shen, H.; Li, J.; Yuan, Q.; Zhang, L. A differential information residual convolutional neural network for pansharpening. *ISPRS J. Photogramm. Remote Sens.* **2020**, *163*, 257–271. [[CrossRef](#)]
49. Vitale, S.; Scarpa, G. A detail-preserving cross-scale learning strategy for CNN-based pansharpening. *Remote Sens.* **2020**, *12*, 348. [[CrossRef](#)]
50. Ciotola, M.; Vitale, S.; Mazza, A.; Poggi, G.; Scarpa, G. Pansharpening by convolutional neural networks in the full resolution framework. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–17. [[CrossRef](#)]
51. Scarpa, G.; Ciotola, M. Full-resolution quality assessment for pansharpening. *Remote Sens.* **2022**, *14*, 1808. [[CrossRef](#)]
52. Alparone, L.; Wald, L.; Chanussot, J.; Thomas, C.; Gamba, P.; Bruce, L.M. Comparison of pansharpening algorithms: Outcome of the 2006 GRS-S data-fusion contest. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 3012–3021. [[CrossRef](#)]
53. Deng, L.; Vivone, G.; Paoletti, M.E.; Scarpa, G.; He, J.; Zhang, Y.; Chanussot, J.; Plaza, A. Machine learning in pansharpening: A benchmark, from shallow to deep networks. *IEEE Geosci. Remote Sens. Mag.* **2022**, *10*, 279–315. [[CrossRef](#)]
54. Ghosh, A.; Joshi, P.K. Assessment of pan-sharpened very high-resolution worldview-2 images. *Int. J. Remote Sens.* **2013**, *34*, 8336–8359. [[CrossRef](#)]
55. Jawak, S.D.; Luis, A.J. A spectral index ratio-based antarctic land-cover mapping using hyperspatial 8-band worldview-2 imagery. *Polar Sci.* **2013**, *7*, 18–38. [[CrossRef](#)]
56. Maglione, P.; Parente, C.; Vallario, A. Pan-sharpening worldview-2 IHS, brovey and zhang methods in comparison. *Int. J. Eng. Technol.* **2016**, *8*, 673–679.
57. Li, H.; Jing, L.; Tang, Y. Assessment of pansharpening methods applied to Worldview-2 imagery fusion. *Sensors* **2017**, *17*, 89. [[CrossRef](#)]
58. Wald, L.; Ranchin, T.; Mangolini, M. Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images. *Photogramm. Eng. Remote Sens.* **1997**, *63*, 691–699.
59. Vivone, G.; Alparone, L.; Chanussot, J.; Dalla Mura, M.; Garzelli, A.; Licciardi, G.A.; Restaino, R.; Wald, L. A critical comparison among pansharpening algorithms. *IEEE Trans. Geosci. Remote Sens.* **2015**, *33*, 2565–2586. [[CrossRef](#)]
60. Garzelli, A.; Nencini, F. Interband structure modeling for pan-sharpening of very high-resolution multispectral images. *Inf. Fusion* **2005**, *6*, 213–224. [[CrossRef](#)]
61. Lee, J.; Lee, C. Fast and efficient panchromatic sharpening. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 155–163.
62. Dong, C.; Loy, C.C.; He, K.; Tang, X. Image super-resolution using deep convolutional networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2016**, *38*, 295–307. [[CrossRef](#)]
63. Ranchin, T.; Aiazzi, B.; Alparone, L.; Baronti, S.; Wald, L. Image fusion—The arsis concept and some successful implementation schemes. *ISPRS J. Photogramm. Remote Sens.* **2003**, *58*, 4–18. [[CrossRef](#)]
64. Yuhas, R.; Goetz, A.; Boardman, J. Discrimination among semi-arid landscape endmembers using the spectral angle mapper (sam) algorithm. In Proceedings of the Summaries of the Third Annual JPL Airborne Geoscience Workshop, Pasadena, CA, USA, 1 June 1992; Jet Propulsion Laboratory: Pasadena, CA, USA, 1992; pp. 147–149.
65. Alparone, L.; Baronti, S.; Garzelli, A.; Nencini, F. A global quality measurement of pan-sharpened multispectral imagery. *IEEE Geosci. Remote Sens. Lett.* **2004**, *1*, 313–317. [[CrossRef](#)]
66. Garzelli, A.; Nencini, F. Hypercomplex quality assessment of multi/hyperspectral images. *IEEE Geosci. Remote Sens. Lett.* **2009**, *6*, 662–665. [[CrossRef](#)]
67. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Geosci. Remote Sens.* **2004**, *13*, 600–612. [[CrossRef](#)]
68. Yang, C.; Zhang, J.Q.; Wang, X.R.; Liu, X. A novel similarity based quality metric for image fusion. *Inf. Fusion* **2008**, *9*, 156–160. [[CrossRef](#)]
69. Alparone, L.; Alazzi, B.; Baronti, S.; Garzelli, A.; Nencini, F.; Selva, M. Multispectral and panchromatic data fusion assessment without reference. *Photogramm. Eng. Remote Sens.* **2008**, *74*, 193–200. [[CrossRef](#)]
70. Khan, M.M.; Alparone, L.; Chanussot, J. Pansharpening quality assessment using the modulation transfer functions of instruments. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 3880–3891. [[CrossRef](#)]
71. Arienzo, A.; Vivone, G.; Garzelli, A.; Alparone, L.; Chanussot, J. Full-resolution quality assessment of pansharpening: Theoretical and hands-on approaches. *IEEE Geosci. Remote Sens. Mag.* **2022**, *10*, 168–201. [[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.