Information Fusion 93 (2023) 227–242

Contents lists available at ScienceDirect

Information Fusion

journal homepage: www.elsevier.com/locate/inffus

Panchromatic and multispectral image fusion for remote sensing and earth observation: Concepts, taxonomy, literature review, evaluation methodologies and challenges ahead

Kai Zhang [a], Feng Zhang [a], Wenbo Wan [a], Hui Yu [b,*], Jiande Sun [a,*], Javier Del Ser [c,d], Eyad Elyan [e], Amir Hussain [f]

[a] School of Information Science and Engineering, Shandong Normal University, China
[b] School of Creative Technologies, University of Portsmouth, UK
[c] TECNALIA, Basque Research and Technology Alliance (BRTA), 48160 Derio, Spain
[d] University of the Basque Country (UPV/EHU), 48013 Bilbao, Spain
[e] School of Computing, Robert Gordon University, UK
[f] School of Computing, Edinburgh Napier University, UK

ARTICLE INFO

Keywords:
Image fusion
Pan-sharpening
Image quality evaluation
Multispectral image
Panchromatic image

ABSTRACT

Panchromatic and multispectral image fusion, termed pan-sharpening, is to merge the spatial and spectral information of the source images into a fused one, which has a higher spatial and spectral resolution and is more reliable for downstream tasks compared with any of the source images. It has been widely applied to image interpretation and pre-processing of various applications. A large number of methods have been proposed to achieve better fusion results by considering the spatial and spectral relationships among panchromatic and multispectral images. In recent years, the fast development of artificial intelligence (AI) and deep learning (DL) has significantly enhanced the development of pan-sharpening techniques. However, this field lacks a comprehensive overview of recent advances boosted by the rise of AI and DL. This paper provides a comprehensive review of a variety of pan-sharpening methods that adopt four different paradigms, i.e., component substitution, multiresolution analysis, degradation model, and deep neural networks. As an important aspect of pan-sharpening, the evaluation of the fused image is also outlined to present various assessment methods in terms of reduced-resolution and full-resolution quality measurement. Then, we conclude this paper by discussing the existing limitations, difficulties, and challenges of pan-sharpening techniques, datasets, and quality assessment. In addition, the survey summarizes the development trends in these areas, which provide useful methodological practices for researchers and professionals. Finally, the developments in pan-sharpening are summarized in the conclusion part. The aim of the survey is to serve as a referential starting point for newcomers and a common point of agreement around the research directions to be followed in this exciting area.

1. Introduction

With the rapid development of remote sensing technologies, more and more satellites are launched and a large number of remote sensing images are collected by various imaging sensors. The information about Earth is recorded in different observation modes by these remote sensing images, which provide abundant data for the interpretation of the observed scene. Until now, these images have been extensively applied to many fields, such as resource exploration [1], environmental survey [2], and battlefield reconnaissance [3], and have achieved great success.

As two important attributes of remote sensing images, spatial and spectral resolutions have significant influences on the interpretation accuracy of the observed scene. Although the spatial and spectral resolutions of remote sensing images are continuously improved, a high spatial and spectral resolution cannot be achieved simultaneously for these images. This is caused by the intrinsic trade-off between spatial and spectral resolutions of imaging sensors [4].

In an imaging system, the spectral resolution can be improved by increasing the number of bands and reducing the width of bands. But when the energy of incident light is fixed, the energy received

* Corresponding authors.
E-mail addresses: zhangkainuc@sdnu.edu.cn (K. Zhang), 2019010100@stu.sdnu.edu.cn (F. Zhang), wanwenbo@sdnu.edu.cn (W. Wan), hui.yu@port.ac.uk (H. Yu), jiandesun@sdnu.edu.cn (J. Sun).

https://doi.org/10.1016/j.inffus.2022.12.026
Received 23 November 2022; Received in revised form 24 December 2022; Accepted 28 December 2022
Available online 2 January 2023

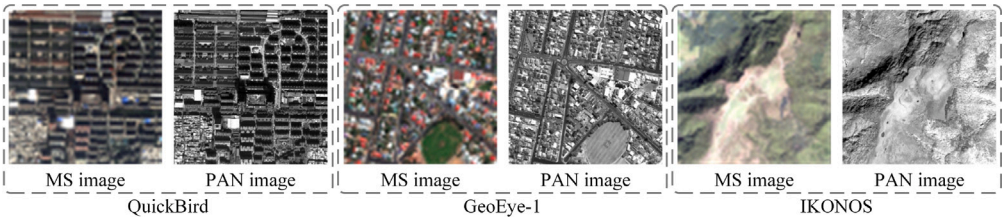

Fig. 1. Multispectral (MS) and panchromatic (PAN) image pairs from different satellites.

by the narrowband will be reduced, which results in the decrease of the signal-to-noise ratio of the remote sensing images. To satisfy a specified signal-to-noise ratio in this case, one option is to improve the size of light-sensing units. As a result, the spatial resolution of these images will degrade. On the other hand, the spatial resolution mainly depends on the density of light-sensing units on imaging sensors. By reducing the size of light-sensing units, the spatial resolution can be enhanced. However, the high density of light-sensing units means the improvement in terms of manufacturing costs and difficulty of imaging sensors.

With the intrinsic constraint, the panchromatic (PAN) image can provide a high spatial resolution. But it is composed of only one band. The multispectral (MS) image possesses several spectral bands and contains abundant spectral information. The spatial details in the MS image are inferior to those in the PAN image. Fig. 1 shows some MS and PAN image pairs from different satellites. We can find that the PAN image is more clear than the MS image. But the former is a grayscale image. Besides, Table 1 illustrates the spatial and spectral resolutions of some optical satellites, which can capture PAN and MS images simultaneously. In Table 1, the spatial resolution is measured in meters. It can be seen that the spectral resolution of the MS image is higher than that of the PAN image. Conversely, the PAN image is superior to the MS image in terms of the spatial resolution. PAN and MS images show the spatial and spectral properties of the observed scenes, respectively.

Table 1
Spatial and spectral resolutions of different satellites.

Satellite	Spectral bands		Spatial resolution (m)	Launch time
QuickBird	MS	4	2.44 m	2001
	PAN	1	0.61 m	
GeoEye-1	MS	4	2 m	2008
	PAN	1	0.5 m	
IKONOS	MS	4	4 m	1999
	PAN	1	1 m	
Pleiades-1A	MS	4	2 m	2011
	PAN	1	0.5 m	
Pleiades-1B	MS	4	2 m	2012
	PAN	1	0.5 m	
WorldView-2	MS	8	2 m	2009
	PAN	1	0.5 m	
WorldView-3	MS	8	1.24 m	2014
	PAN	1	0.31 m	
WorldView-4	MS	4	1.24 m	2016
	PAN	1	0.31 m	
GF-1	MS	4	8 m	2013
	PAN	1	2 m	
GF-2	MS	4	4 m	2014
	PAN	1	1 m	

It is obvious that the downstream tasks, such as target detection [5] and classification [6,7], cannot be efficiently achieved merely by using the PAN or the MS image. Because the lack of high spatial or spectral resolution limits the interpretation accuracy of the observed scene. To realize the required accuracy of the downstream tasks, there is an increased demand for remote sensing images with a high spatial and spectral resolution. Therefore, it is vital to improve the spatial and spectral resolutions of remote sensing images for more accurate earth observation.

To this end, image fusion is adopted to integrate the complementary information in different kinds of images for more comprehensive descriptions of the observed scene [8]. In 2006, IEEE Geoscience and Remote Sensing Society (GRSS) began to organize the data fusion contest to boost this field [9]. Up to now, the data fusion contest has been successfully held for 17 times [10]. There are many different fusion tasks in GRSS to exploit the multisource and multimodal remote sensing data. For example, synthetic aperture radar and optical remote sensing image fusion [11] explores the physical properties and surface characteristics of the observed scene, which has shown great potential in object identification. For hyperspectral image and light detection and ranging fusion [12], spectral features and range information are combined to efficiently improve the accuracy of semantic segment. Light detection and ranging exhibits strongly heterogeneous characteristics compared with the hyperspectral image and so an elaborate combination scheme needs to be considered decently. Multitemporal fusion [13] takes advantage of the time series of multisensor images, which extends MS or hyperspectral images from 3D cubes to 4D data through the introduction of the time variable. It enables to capture the change information in a short time or long time series. In the spatiotemporal fusion task, the high spatial and temporal resolution image can be created by merging the low spatial resolution but high temporal resolution image and the high spatial resolution but low temporal resolution images, which can provide continuous monitoring for regional changes. In recent years, to obtain the high spatial resolution hyperspectral image, the fusion of MS/PAN and hyperspectral images [14–16] has attracted more attention, which enriches the spatial information in the hyperspectral image. Through the models mentioned above, abundant properties of the observed scene can be reflected more completely by the fused image than the image acquired by a single sensor.

Among these tasks that integrate the complementary information from different images, we specifically focus on the fusion of PAN and MS images, also termed pan-sharpening [17]. Generally, as shown in Fig. 2, image fusion can be classified into four levels: sensor/pixel level, feature level, confidence level, and decision level. Sensor/pixel-level fusion directly integrates the original pixel information in low spatial resolution multispectral images (LR MS) and PAN images, which aims at generating a fused image with a high spatial and spectral resolution. Sensor/pixel-level fusion also can be regarded as one kind of preprocessing of LR MS and PAN images, which can improve the resolutions of images to describe the observed scenes more accurately. In feature-level fusion [18–20], the features of objects in LR MS and PAN images are extracted. Then, these features are combined for the classification or detection of the objects in observed scenes. As for confidence-level fusion [21,22], the confidence scores rather than the decision results from LR MS and PAN images are merged. Then, the final decision result is derived from the fused confidence scores. Decision-level fusion [23,24] combines the decision results, such as labels or locations

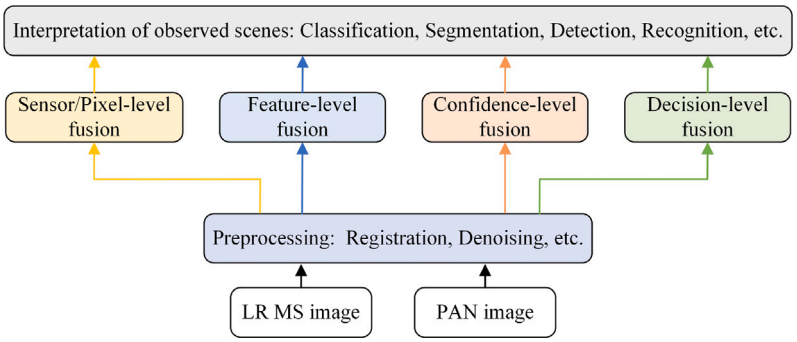

Fig. 2. Fusion levels of low spatial resolution multispectral image (LR MS) and panchromatic (PAN) images in the interpretation of observed scenes.

of objects, from LR MS and PAN images to produce a final decision, in which voting, statistical, and fuzzy logic schemes are adopted to fuse these results. In recent years, sensor/pixel-level image fusion is boosted a lot due to the development of artificial intelligence. There have been new developments and trends about the sensor/pixel-level fusion of MS and PAN images. Therefore, this paper mainly focuses on the sensor/pixel-level fusion to integrate the fine spatial details in the PAN image and the abundant spectral information in the MS image. After pan-sharpening, the high spatial resolution (HR) MS image is finally generated.

The concept and benefits of pan-sharpening are illustrated in Fig. 3. The fused HR MS image is characterized by both high spatial and spectral resolutions, which are helpful for the efficient interpretation of the observed scene. For example, Rayegani et al. [25] applied the fused images to the change detection of natural ecosystems, in which two spatial and spectral indexes were designed to select proper fusion methods. Lottering et al. [26] used the abundant texture information in the WorldView-2 pan-sharpened images. In this method, an artificial neural network was considered to infer vegetation defoliation from the fused images. Qu et al. [27] used the fused images for anomaly detection and the results showed that these images boosted the accuracy. Specifically, in [27], the original detection rate of LR MS images was 0.3 and the detection rate was improved to 0.9 by using the fused images. Du et al. [28] compared the change detection performance between LR MS images and fused images. The results demonstrated that the Kappa coefficient was raised by 0.2 and the overall accuracy was improved by 10% after implementing change detection on the fused images.

Over the past decades, many pan-sharpening methods have been proposed and they can generate satisfactory results with impressive performance. In general, these pan-sharpening methods can be categorized into four categories: component substitution (CS) methods, multiresolution analysis (MRA) methods, degradation model (DM) based methods, and deep neural network (DNN) based methods. In recent years, the rise of DNNs has brought new opportunities for the quality improvement of the fused image, and state-of-the-art fusion results have been produced by DNN-based pan-sharpening methods.

The success of DNN-based pan-sharpening methods benefits from the powerful learning capabilities and a large number of training datasets. However, for the pan-sharpening task, the training datasets must be prepared elaborately according to the Wald protocol [29]. Since DNNs are trained on the handcrafted reduced-resolution data, they cannot reach the anticipated effects on the full-resolution data. The mapping between source images and the fused image is so complex that the learned DNNs on the reduced-resolution data cannot be generalized to the full-resolution data. In addition to this, the scarcity and complicated construction of the reduced-resolution training data promote some DNNs to leverage the full-resolution data for training.

For existing reviews, they mainly summarized CS- and MRA-based pan-sharpening methods in detail while ignoring the developments in the other two fields, such as fusion methods based on DM or DNN.

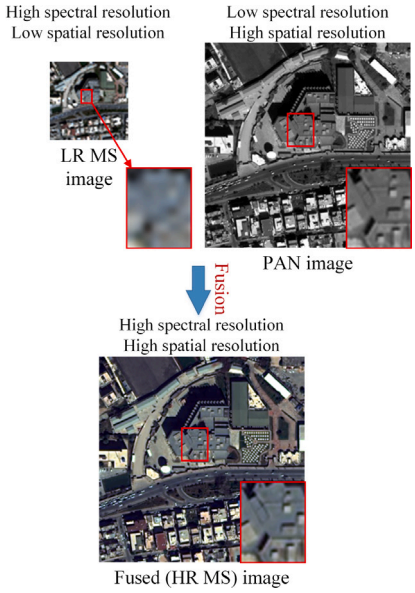

Fig. 3. Concept and benefits of low spatial resolution multispectral image (LR MS) and panchromatic (PAN) image fusion.

For example, Thomas et al. [30] reviewed the methods based on CS and MRA from the perspective of remote sensing physics. Vivone et al. [31] provided a critical and extensive comparison of different CS- and MRA-based methods on the datasets from different satellites, which promoted the standardized implementations of many available pan-sharpening codes. Vivone et al. [32] further presented an overview of pan-sharpening on the basis of [31]. In addition to the review of pan-sharpening, Meng et al. [33] published a large-scale dataset to the community, which improved the availability of HR remote sensing images. However, most of these surveys do not involve the taxonomy of DM- and DNN-based methods, nor do they conduct an in-depth analysis of the formulations of pan-sharpening methods.

Besides, we also summarize the evaluations of fused images to reflect the field more comprehensively. Nowadays, it has become increasingly difficult to visually distinguish between the fused images of different DNN-based pan-sharpening methods. On the one hand, as one important part of the pan-sharpening task, the evaluation of the fused image provides evidence for the choice of DNNs. Recently, some novel evaluation indexes and tools are proposed to calculate the similarity between the fused image and the reference image or assess the spatial and spectral features in the no-reference case. The gap between evaluation indexes and visual performance is reduced. Unfortunately, the analysis of evaluation indexes is not discussed in the reviews mentioned above. On the other hand, the quality of the fused image should be further evaluated on the downstream tasks, because

Table 2
Main symbol list.

Symbol	Description
LR MS	Low spatial resolution multispectral image
PAN	Panchromatic image
HR MS	High spatial resolution multispectral image
CS	Component substitution
MRA	Multiresolution analysis
DM	Degradation model
DNN	Deep neural network
$\mathbf{L}_b$	The bth band in the LR MS image
P	PAN image
$\mathbf{H}_b$	The bth band in the HR MS image, also fused image
$\widetilde{\mathbf{L}}_b$	The bth band in LR MS image upsampled to the size of PAN image
$\tilde{\mathbf{P}}$	The PAN image synthesized by the upsampled bands in the LR MS image
r	Spatial resolution ratio between LR MS and PAN images
N	The number of bands in MS image
X	The matrix form of LR MS image via rearrangement with matching size
Y	The matrix form of PAN image via rearrangement with matching size
Z	The matrix form of the fused image via rearrangement with matching size
D	The spatial downsampling matrix
B	The spatial blurring matrix
S	The spectral degradation matrix

pan-sharpening is just a pre-processing step in the interpretation of the observed scene. Therefore, it will be more meaningful if we can compare the interpretation performance of the same task on the fused images produced by different pan-sharpening methods. It is our firm belief that only the task-oriented performance assessment of the fused image boosts the practical process of pan-sharpening methods.

In this paper, we have analyzed about 220 papers for pan-sharpening and provide scientific readers with a comprehensive review of the state-of-the-art pan-sharpening methods, especially DM- and DNN-based methods. Meanwhile, the purpose of this review is to contribute valuable insights into the quality evaluation of the fused image. Finally, dedicated discussions on datasets, evaluations, and future trends are presented to support related researchers by understanding the limitations, difficulties, and challenges in this field.

The remainder of the paper is organized as follows. In Section 2, we present a taxonomical review of the pan-sharpening methods falling into the four categories. Section 3 describes the evaluation indexes in reference and no-reference cases, respectively. Section 4 introduces the limitations, difficulties, and challenges. Finally, conclusions are drawn in Section 5.

2. Taxonomy of pan-sharpening methods

In this section, some classical methods and their variants are also reviewed for a comprehensive literature survey. According to their formulations, different kinds of pan-sharpening methods are organized into a taxonomy.

2.1. Notions

For convenience, Table 2 reports the symbols with their corresponding description, where we denote scalars as lowercase italics letters, vectors as boldface lowercase letters, and matrices as boldface capitals. In DM-based methods, source and fused images are usually rearranged as matrices through image partition and vectorization, or other preprocessing methods to match the spatial and spectral degradation models. For other acronyms and symbols used in the following parts, they will be defined upon need.

2.2. Taxonomy

Taking into account all the reviewed papers, we divide the pan-sharpening methods in a hierarchy of classes. Fig. 4 demonstrates the hierarchical taxonomy of methods proposed for pan-sharpening. In the hierarchy, these methods can be generally split into four categories: CS-based methods, MRA-based methods, DM-based methods, and DNN-based methods.

For the methods based on CS, the spatial and spectral information in the MS image is supposed to be separable so that the spatial component in the up-sampled LR MS image can be directly replaced by the corresponding PAN image. Following the principle, the up-sampled LR MS image is firstly projected into a new space by some suitable transformations, which can produce a spatial component. Then, the spatial component is substituted entirely or partially by the histogram-matched PAN image. Finally, the fused HR MS image is generated by the corresponding inverse transformation. Within this category, there are some classical and common transformations, such as intensity-hue-saturation (IHS) transformation [34], principal component analysis (PCA) [35], and Gram–Schmidt (GS) transformation [36].

For MRA-based methods, it is assumed that the missing spatial details in the LR MS image can be inferred from the PAN image. The formulation is also embodied as *Amélioration de la Résolution Spatiale par Injection de Structures* [37], a French name. Then, some efficient tools, such as MRA, are employed to extract the spatial details, also termed high frequencies, from the PAN image. Next, these details are injected into the LR MS image through some fusion rules. Various MRA tools [38] offer different choices for researchers to specifically model the spatial information in the PAN and LR MS images. Building on the success of MRA, some MRA-like approaches are developed for pan-sharpening, including support value transformation [39], support tensor transformation [40], and morphological filters [41].

In the third category, LR MS and PAN images are thought of as the degradation results of the HR MS image in spatial and spectral domains, respectively. Then, the pan-sharpening task is regarded as the image restoration problem. Naturally, the fused image is estimated by solving the inverse problems derived from the spatial and spectral degradation models. Since the ill-condition of the degradation models, priors existing in source images or the HR MS image are mined for the regularization of its solution space. The vast majority of DM-based pan-sharpening methods utilize different forms of sparsity to generate the fusion result [42]. Besides, the gradient prior [43] and low-rank prior [44] are also considered in the pan-sharpening task due to their success in low-level computer vision tasks, such as image superresolution [45] and denoising [46].

Recently, DNN-based pan-sharpening methods have gained popularity thanks to the powerful capabilities of nonlinear learning [47]. Leveraging on the success in computer vision tasks, a large variety of DNNs are adopted and trained in a supervised manner for pan-sharpening. On the one hand, supervised learning makes DNN-based methods outperform the three kinds of methods mentioned above. On the other hand, outstanding performance benefits from a large volume of data and greater computing power. With diverse DNNs springing up, such as generative adversarial network [48] and transformer [49], there is a great deal of interest in improving the fusion performance and satisfying high flexibility and feasibility in a real scenario. However, it is difficult to categorize DNN-based methods according to the type of their networks, since network types grow exponentially. Compared to most low-level tasks, the most significant distinguishing factor of pan-sharpening is characterized by its dual-source input. Considering the stages and patterns of the dual-source information combination in LR MS and PAN images, DNN-based pan-sharpening methods can be grouped into three subcategories: source image concatenation (SIC), feature concatenation (FC), and feature fusion (FF).

To avoid excessive division, each pan-sharpening method is categorized according to its characteristics that are distinct from its remaining

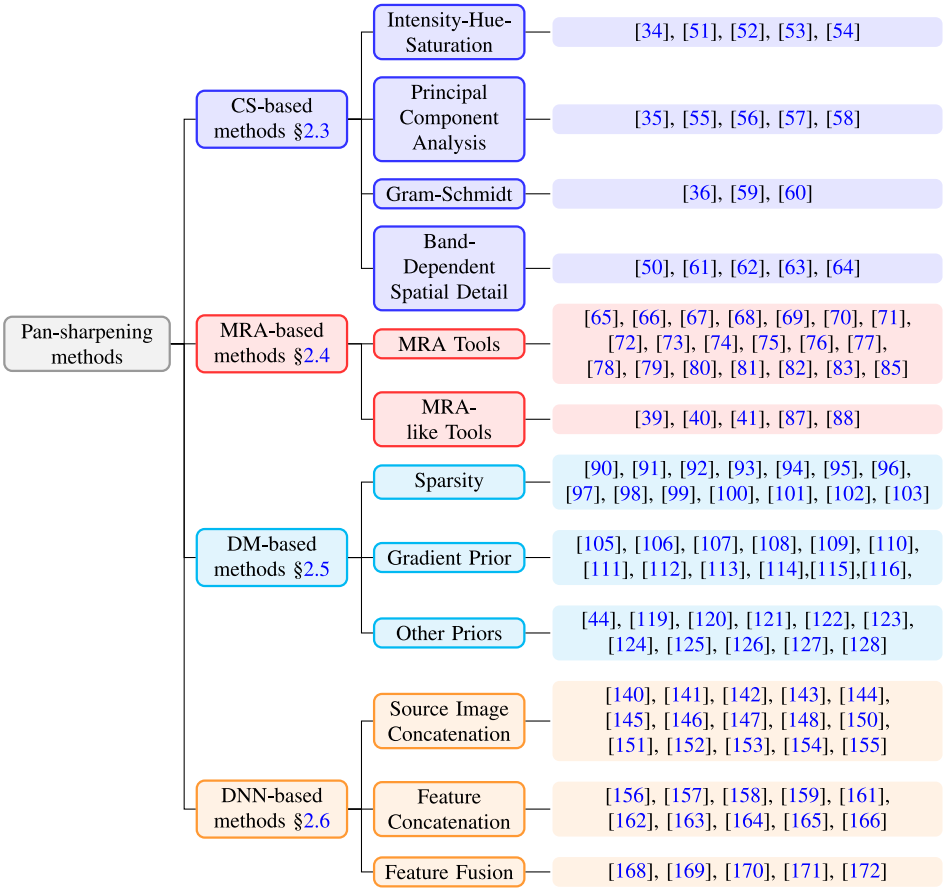

Fig. 4. Taxonomy of literature about pan-sharpening.

counterparts. In this way, mess categories will be prevented and the taxonomy depicted in Fig. 4 becomes more meaningful and informative. These four categories will be described in detail in the following subsections.

2.3. CS-based methods

In this section, we introduce the pan-sharpening methods based on CS in detail. The key to these techniques lies in the accurate estimation of the spatial component from the LR MS image. However, owing to the high inseparability of the spatial and spectral information in the LR MS image, it is difficult to synthesize the spatial component, which is well-matched with the corresponding PAN image. In addition, the mismatches in terms of the spectral range between LR MS and PAN images cannot be ignored. Fig. 5 displays the spectral response ranges of MS and PAN images from different satellites. One can see that the spectral range of the MS image cannot exactly cover that of the PAN image, which will inevitably introduce spectral distortions into the fused image. Therefore, issues about spectral response ranges and the synthesis of the spatial component are vital to reduce the spectral distortions for the methods based on CS.

Compared to other kinds of methods, CS-based methods behave well in terms of the enhancement of the spatial details in the fused image. They are characterized by their simple operation and are easy to implement. Although these methods are not expected to obtain a competitive performance with respect to other kinds of methods, such as DM- and DNN-based methods, their simplicity and effectiveness still make them worthy of further study and attractive for researchers.

Mathematically, the pan-sharpening methods based on CS are summarized by Vivone et al. [31] and can be generally written as:

$$\mathbf{H}_b = \widetilde{\mathbf{L}_b} + g_b\left(\mathbf{P} - \widetilde{\mathbf{P}}\right) \tag{1}$$

where $\widetilde{\mathbf{P}}$ is the spatial component derived from the LR MS image through a specific transformation. g_b is the injection coefficient for the bth band of the LR MS image. In Eq. (1), $\widetilde{\mathbf{P}}$ can be synthesized by some transformations. The general form of the synthesis of $\widetilde{\mathbf{P}}$ is defined as:

$$\widetilde{\mathbf{P}} = \sum_{b=1}^{B} w_b \widetilde{\mathbf{L}}_b \tag{2}$$

where w_b is the weight corresponding to $\widetilde{\mathbf{L}}_b$. Obviously, the adopted transformations and injection gains are two crucial factors that affect the fusion performance of the CS-based methods. According to the adopted transformation, the methods in this category can be grouped into four sub-classes: IHS [34], PCA [35], GS [36], and band-dependent spatial detail (BDSD) [50]. An overview of these methods is presented as follows.

- *IHS*: This kind of methods use IHS to produce the intensity component of the LR MS image, which is then replaced by the PAN image. In IHS, the intensity component is obtained by averaging all bands in the LR MS image, which means the weight w_b in Eq. (2) for each band is $\frac{1}{3}$. Unfortunately, although the synthesis of the intensity component is straightforward, the MS image is usually made up of 4 or 8 bands. Therefore, this kind of methods will face serious limitations if the transformation can only deal with the MS image with 3 bands. To broaden the application of IHS, Tu et al. [51] took the high reflectance of vegetated regions in the near-infrared band into consideration and set the weights of both red and near-infrared bands as $\frac{1}{3}$. The weights for green and blue bands are decided by their spectral response ranges, which are set as $\frac{a}{3}$ and $\frac{b}{3}$. For a and b, $a+b=1$. In the above-mentioned two cases, the differences between $\mathbf{P}$ and $\widetilde{\mathbf{P}}$ are directly combined

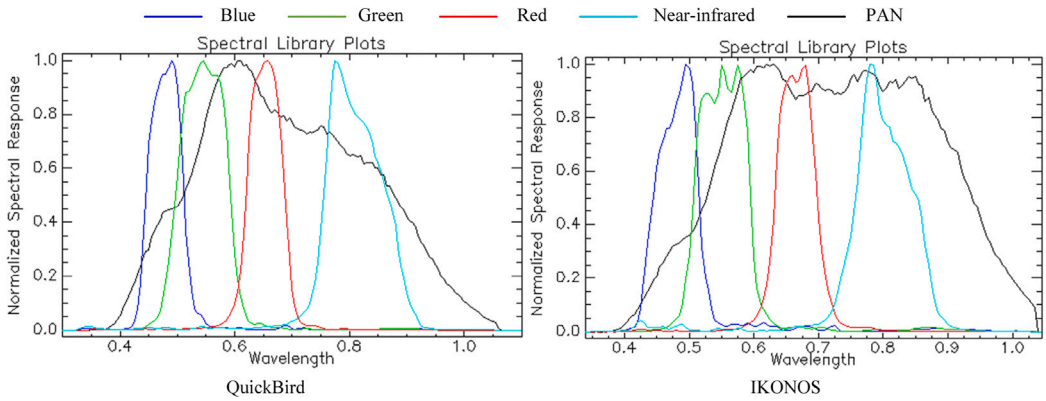

Fig. 5. Spectral responses of QuickBird and IKONOS.

with the upsampled image $\widetilde{\mathbf{L}}_b$, and thus the injection gains are 1 for each band.

Compared to the injection gains independent of the image content, the adaptive estimation of injection gains will alleviate the spectral distortions in the fused image. For instance, Rahmani et al. [52] proposed an adaptive IHS (AIHS) method, in which the weights were calculated by approximating $\mathbf{P}$ through the combination of the bands $\widetilde{\mathbf{L}}_b$, and the injection gains for all bands were computed from the edge maps derived from $\mathbf{P}$. Subsequently, an improved adaptive IHS (IAIHS) method was presented in [53], where the injection gain g_b for each band was jointly determined by the edge maps from $\mathbf{P}$ and $\widetilde{\mathbf{L}}_b$. Thus, the spectral information can be preserved better in the fused image by using the method in [53]. Based on AIHS and IAIHS, Ghahremani et al. [54] presented nonlinear IHS, in which the spatial component $\widetilde{\mathbf{P}}$ was synthesized locally.

- *PCA*: PCA [35] is another classical transformation often used in CS-based pan-sharpening methods. According to the formulation of PCA, the variance of the first principal component (1st PC) is the largest among all PCs. It indicates that the 1st PC contains the most information and so the 1st PC is regarded as the spatial component of the LR MS image. Specifically, the 1st PC of the LR MS image is selected and substituted by the PAN image, which is different from the intensity component generated by IHS. Although the 1st PC includes abundant spatial details, it does not mean that the 1st PC is the most similar to the PAN image.

  In this sense, Yang et al. [55] adopted adaptive PCA to find the most similar PC to the PAN image, in which the similarity between PCs and the PAN image was measured by the correlation coefficient. In [56], the spatial PCA and the spectral PCA were combined to infer a more similar spatial component. Besides, the spatial PCA was also considered in [57] and indirectly applied to the spatial details of $\mathbf{P}$. To extract reasonable spatial details, modulation transfer function (MTF) based filters were implemented on $\mathbf{P}$. Subsequently, PCA was also employed for pan-sharpening in [58] and the chromatic components were conditionally filtered according to the structures in the PAN image, which adjusted the injection gains through nonlocal optimization.
- *GS*: GS [36] is also a common transformation used for pan-sharpening. Based on the synthetic spatial component $\widetilde{\mathbf{P}}$, GS is designed to obtain orthogonal vectors from the up-sampled LR MS image. After the projection of GS, all components are pairwise orthogonal. Then, the spatial component $\widetilde{\mathbf{P}}$ is replaced by the PAN image. Finally, the fused image is obtained through the inverse transformation on the new components. Clearly, the production of $\widetilde{\mathbf{P}}$ plays an important role in the reconstruction of the fused image.

  Aiazzi et al. [59] minimized the mean square error to obtain the optimal weights for the generation of $\widetilde{\mathbf{P}}$. Wang et al. [60] considered the particle-swarm-optimization algorithm to minimize a non-differentiable function and achieved adaptive weights. Actually, researchers can have many kinds of methods to estimate the combined weights. However, the inflexible orthogonal projection limits the room for improvement of GS-based methods. Therefore, the number of variants of GS-based method is less than those of other kinds of transformations.
- *BDSD*: According to the former three sub-classes, we can see that content-dependent weights lead to fewer spectral distortions in the fused image. Under the circumstances, BDSD [50] further extended the general formulation in Eq. (1). In BDSD, the weight for each band $\widetilde{\mathbf{L}}_b$ was adaptively estimated from the corresponding downsampling PAN and MS images by the least square criterion. Compared to the global computing strategy, the approximating precision will be high by using the weights estimated from local regions, because of the differences in terms of the spectral response of different objects. Further, the partition strategy considering image content has a better performance in terms of the spectral distortion than the approaches that directly segment the image into patches.

  Based on the idea, many variants of BDSD are advanced. For instance, Garzelli [61] further proposed an improved version C-BDSD to deal with the preservation of spatial consistency, which utilized K-means clustering to find more consistent regions. In C-BDSD, the intensity and local standard deviation of the PAN image were regarded as the feature for clustering. In [62], spatial features were extracted by the Gabor filter bank, and then the fuzzy c-means algorithm was employed to achieve clustering. Besides, Imani [63] also established an improved methods based on BDSD, named CR-BDSD, which used the collaborative representation of $\widetilde{\mathbf{L}}_b$ instead of $\widetilde{\mathbf{L}}_b$ to efficiently smooth the spatial details in $\widetilde{\mathbf{L}}_b$. Recently, Vivone [64] presented three robust fusion methods based on BDSD, in which outlier removal, bi-square regression, and physical constraints were combined with the objective function to tackle different cases.

2.4. MRA-based methods

Although CS-based methods have been widely studied and integrated into various interpretation tasks of the observed scene, it is difficult to ignore the spectral distortions in the fusion results of these methods. The reason for this is that the spatial component of the LR MS image is directly replaced by the PAN image. A viable solution is only injecting the spatial details needed into the LR MS image. In this case, MRA enters the vision of the pan-sharpening research for the modeling of spatial details in source images. In MRA-based methods, only spatial details from the PAN image are injected into the LR MS image, so they behave better in terms of the preservation of the spectral information.

Summarily, the formulation of MRA-based methods is defined as:

$$\mathbf{H}_b = \widetilde{\mathbf{L}_b} + g_b\left(\mathbf{P} - \widehat{\mathbf{P}}\right) \tag{3}$$

Table 3
Summary of related MRA-based pan-sharpening methods.

Tool	Reference	Gain description
GLP	Vivone [65]	Full scale regression
	Vivone [66]	Reduced scale regression
	Restaino [67]	Multiple linear regression
	Addesso [68]	Data-driven multivariate regression
Wavelet	Otazu [69]	Spectral response, physical constraint
	Vivone [70]	Spectral response, physical constraint
	Lu [71]	Local covariance/local variance
	Kallel [72]	Covariance/variance
	Yang [73]	Fuzzy logic and maximum selection
Contourlet	Shah [74]	1
	El-Mezouar [75]	Local energy of coefficients
	Upla [76]	1
	Li [77]	Sparse coefficient
Curvelet	Dong [78]	1
	Devulapalli [79]	Adaptive neuro-fuzzy inference
Framelet	Shi [80]	Covariance intersection
	Wang [81]	Random walks
	Zhao [82]	SFIM
Shearlet	Shi [83]	Pulse-coupled neural network
	Moonon [84]	Local energy
MRA-like	Zheng [39]	1
	Xing [40]	Low frequency
	Restaino [41]	Morphological operators
	Yang [85]	Multiscale guided filter
	Yin [86]	Nonlocal means filter

where $\widehat{\mathbf{P}}$ is the low frequency of $\mathbf{P}$. It can be observed that the fusion performance of this category is also impacted by the extraction of high frequencies and the estimation of the injection gain g_b.

Thanks to the variety of available MRA tools, many efficient pan-sharpening methods are proposed. In Table 3, we list the representative pan-sharpening methods based on MRA, where high-frequency extraction tools and the priors of the gains are described briefly. These methods can be characterized according to whether the adopted tools are derived from MRA. MRA-based methods directly utilize the existing tools to extract the spatial details in source images. For the methods inspired by the framework of MRA, they are termed as MRA-like counterparts. The following parts present the representative methods for each sub-class.

- *MRA*: As mentioned above, MRA has a great capacity to model the high frequencies in source images. The development of MRA-based pan-sharpening methods is accompanied by the emergence of different modeling tools. There are several representative tools for the extraction of spatial details, including generalized Laplacian pyramid (GLP), Wavelet, Contourlet, Curvelet, Framelet, and Shearlet. In Table 3, we list some pan-sharpening methods based on these tools.
  To efficiently approximate the spatial information of images at different scales, the pyramidal decomposition model was investigated intensively. For instance, the Laplacian pyramid (LP) is extended to GLP for pan-sharpening. Vivone et al. [65] used GLP for high-frequency extraction and the injection gains were adjusted at the full scale by regression, which could avoid the differences in injection coefficients at different scales. Within the framework of GLP, many researchers [66–68] turned their attention to the efficient calculation of injection coefficients. As a classical MRA tool, wavelet is also considered. Otazu et al. [69] introduced spectral response and physical properties of the observed scenes into wavelet-based methods, which could produce better fusion results. Then, Vivone et al. [70] further revisited AWLP [69] for reproducible results. Similar to the improvement of other MRA-based methods, the inherent properties among bands of the MS image were exploited in [71–73] to develop more advanced wavelet-based variants and produce better fusion results.
  A wave of pan-sharpening methods is achieved on the strength of more advanced MRA tools beyond wavelet. In [74], Shah et al. used the nonsubsampled contourlet for the enhancement of the spatial information in the fused image. Moreover, the nonsubsampled contourlet was also combined with other approaches to preserve the spatial details better, including local energy [75], Markov prior [76], and sparse autoencoder [77]. To capture the orientation of spatial details in the image, curvelet transform was also considered in [78] to improve the representation capability for edges in LR MS and PAN images. A high-frequency fusion rule derived from adaptive neuro-fuzzy inference was used to integrate the coefficients of LR MS and PAN images after the curvelet transform in [79].
  Different from the contourlet and curvelet, the framelet is also explored for pan-sharpening [80–82]. The framelet favors a high number of vanishing moments, which leads to a sparse representation in the transform domain. With flexible direction feature representation, Shi et al. [83] proposed an image fusion method based on the shearlet, where the high-frequency coefficients were fused by a pulse-coupled neural network [87]. In [84], nonsubsampled shearlet transform was used for pan-sharpening, in which high frequencies and low frequencies were fused by the local energy and sparse representation [88].
- *MRA-like*: In view of the fact that MRA-based pan-sharpening methods preserve the spectral information better, some MRA-like tools are also developed, such as support value transformation [39], support tensor transformation [40], and morphological filters [41] mentioned above. Finding efficient filters is crucial to the MRA-like based pan-sharpening methods. In [85], multiscale guided filters were used to obtain the high frequencies from the PAN image. Besides, Yin et al. [86] proposed multiscale nonlocal means filters to decompose the spatial information in images, which could reduce the spatial distortions in the fused results.

For MRA-based methods, the spectral information in the fused image can be preserved well because only the high frequencies, such as spatial details, are injected into the LR MS image. However, it is sensitive to the spatial correspondence. When the spatial information from the PAN image is not matched with that from the LR MS image, local dissimilarities will appear. Besides, some spatial artifacts are introduced into the fused image owing to the excessive injection of spatial details from the PAN image.

2.5. DM-based methods

CS- and MRA-based pan-sharpening methods both utilize some transformation to infer the missing information in the LR MS image from the PAN image. However, the coupling of the spatial and spectral information makes them difficult to avoid distortions in the fused image. The framework of the image restoration brings the efficient preservation of the spatial and spectral information possible. In this context, DM-based methods employ various optimization algorithms to jointly restore the desired HR MS image from the LR MS and PAN images.

Specifically, the spatial and spectral degradation models can be defined as:

$$\mathbf{X} = \mathbf{DBZ} + n \tag{4}$$

$$\mathbf{Y} = \mathbf{SZ} + n \tag{5}$$

where $\mathbf{D}$ and $\mathbf{B}$ are the spatial downsampling and blurring matrices. In Eq. (5), $\mathbf{S}$ denotes the spectral downsampling matrix derived from the spectral response function of imaging sensors. n is the additive noise. Obviously, the solutions in Eqs. (4) and (5) are ill-posed and satisfactory

fusion results are difficult to be obtained due to a lot of potential solutions. Therefore, various priors are introduced to regularize the solution space of the fused image, which can be formulated as:

$$\min_{\mathbf{Z}} \frac{1}{2} \|\mathbf{X}-\mathbf{DBZ}\|_F^2 + \frac{\lambda}{2} \|\mathbf{Y}-\mathbf{SZ}\|_F^2 + \alpha R(\mathbf{Z}) \tag{6}$$

where λ and α are the tradeoff parameters. $R(\mathbf{Z})$ stands for the regularization term, which can efficiently embed many priors. Thus, the fusion of LR MS and PAN images is achieved from the perspective of image restoration tasks [89].

Researchers deeply analyze the priors existing in image, such as sparsity, gradient prior, and low-rank prior, for pan-sharpening and many DM-based methods are developed. The following sections introduce these latent priors briefly.

- *Sparsity*: As a prevalent prior, sparsity is exploited widely for the fusion of LR MS and PAN images due to its flexible formulation and decent properties. Inspired by the compressed sensing theory [88], a seminal framework was advanced in [90], in which the spatial and spectral degradation matrices were viewed as the measurement matrices. Then, the pan-sharpening problem was solved by the basis pursuit algorithm [91]. Although the unavailability of the HR dictionary hindered the fusion of LR MS and PAN images at full scale, the proposed method in [90] brought a burst in this field. Subsequently, some improved versions [92–95] were presented to enhance the utility of the pan-sharpening framework based on compressed sensing.
  Except for the compressed sensing framework, researchers also explored different forms of sparsity to ensure its efficiency and flexibility. Zhu et al. [96] estimated the sparse representation of the MS image on the dictionary constructed by the PAN image. The mapping between LR and HR patches in [96] was then solved by [97] from the point of manifold learning. The sparse priors in the texture domain were analyzed in [98–100], and modeled by different formulations. Zhang et al. [101] adopted the convolution sparse coding [102,103] to cope with the damage in terms of global structure and spatial consistency caused by the standard sparse representation.
- *Gradient Prior*: Owing to better spatial information preservation, the priors about gradients in MS and PAN images are also considered to regularize the model in Eq. (6). As an efficient prior in the gradient domain, total variation (TV) [104] is introduced into the pan-sharpening task [105]. In computer vision tasks, TV has been applied to image restoration and produced remarkable performance. When TV is employed for pan-sharpening, it is often adjusted for better fusion results according to the properties in source or HR MS images. Researchers generally find new domains [43,106] to investigate the attributes of TV. Some methods [107–109] replaced the L_1 or L_2 norm imposed on TV with other norms by analyzing the probability distributions of the images.
  In addition to the TV prior, some methods assume that the spatial and spectral relationships among source images and HR MS images can be inherited into their counterparts in the gradient domain. For instance, Fang et al. [109] constructed a spatial term for the characterization of the spectral degradation relationship in the gradient field. Fu et al. [110] designed a similar term to that in [109] for the preservation of spatial information. To remove the cloud in remote sensing images, Meng et al. [111] proposed a joint fusion and missing information reconstruction method and formulated the spectral information into the variation framework. Wang et al. [112] also used the spectral degradation model in PAN and HR MS images to ensure the geometric information consistency of the pan-sharpened image in the gradient field.
  Moreover, some methods obtained the fusion results by minimizing the gradient differences between the PAN image and HR MS images. In [113], Chen et al. designed a dynamic gradient sparsity term to make the gradients in each band of the HR MS image coincident with those in the PAN image. Further, Liu et al. [114] assumed that the second-order gradients in the HR MS image should also be similar to those in the PAN image. By dividing images into different components, Liu et al. [115] imposed the gradient difference prior on the cartoon and structure components of HR MS and PAN images. Tian et al. [116] obtained the final pan-sharpened image by minimizing the gradient differences between the desired fused image and the coarse HR MS image produced by sparse representation.
- *Other priors*: The richness of the prior information in images makes DM-based pan-sharpening methods retain their appeal to researchers. As a typical structural sparse prior, low-rank properties in images are also used to regularize the fusion model in Eq. (6). For example, Yang et al. [44] reformulated the pan-sharpening problem via robust PCA [117], in which the spatial and spectral correlations in all bands of HR MS images were modeled by a low-rank matrix. Meanwhile, an alternative form of low-rank matrix factorization, Go Decomposition [118] was used in [119] to extract the low-rank components in the LR MS image. Some methods [120–123] combined the low-dimension constraints with other priors and imposed them on the desired HR MS image.
  Besides, the local, nonlocal, and non-negative properties in images are also considered to reconstruct the fused image better. For example, Wang et al. [124] presented a pan-sharpening method based on sparse representation and local similarity [125], in which local autoregressive parameters learned from the PAN image were shared with the HR MS image. Khademi et al. [126] expressed the local prior by Markov random field to pass the spatial information of the PAN image into the fused image. The nonlocal similarity in the PAN image is captured in [127] to facilitate the solution space of the fusion result. In [128], Zhang et al. proposed a coupled sparse non-negative matrix factorization model for the fusion of LR MS and PAN images, by which the non-negativity of pixel values in images can be naturally assured.

Although DM-based pan-sharpening methods achieve satisfactory HR MS images, their usability is cursed by high computational complexity and the generalization of the adopted prior. On the one hand, this kind of method is solved by iterative optimization algorithms. The involved iterations take a longer time for the calculation of the optimal solution. On the other hand, DM-based methods are heavily dependent on the priors in Eq. (6). However, these priors are valid only under specific assumptions, which limit their performance and generalization. Moreover, as the launched satellites age, the MTFs of imaging sensors gradually change. In this situation, some reconstruction errors may be introduced owing to the misestimated spatial and spectral degradation matrices.

2.6. DNN-based methods

The past ten years have witnessed the great success of DNN in various fields [129–135]. For example, Zhang et al. [136] investigated a feed-forward denoising convolutional neural network to handle Gaussian noises with unknown levels, in which residual learning [137] was utilized to boost the performance. Dong et al. [138] constructed a deep convolution neural network to learn the end-to-end mapping between LR and HR images and they analyzed the network structures and parameter settings to generate satisfactory superresolution results. For image deblurring, Ren et al. [139] presented two generative networks to reconstruct clean images and blur kernels.

The explosion in the computer vision field naturally spreads to the fusion of LR MS and PAN images. As mentioned in Section 2.2, it would be unrealistic to divide the DNN-based pan-sharpening methods according to the network structures adopted by them. When DNNs in

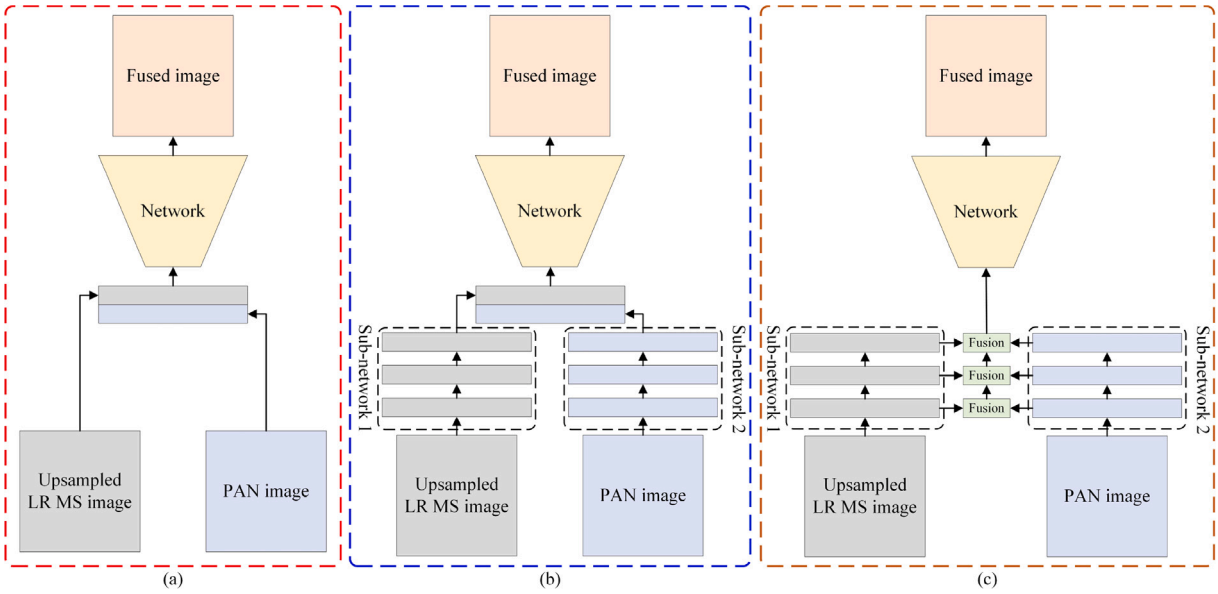

Fig. 6. Schematic diagrams for (a) SIC, (b) FC, and (c) FF, which are DNN-based pan-sharpening methods.

the computer vision field are transferred to the pan-sharpening task or new DNNs are designed for this task, the overriding consideration is how to deal with the dual-source input: LR MS and PAN images. Therefore, we capture the fundamental difference between pan-sharpening and other remote sensing image processing or computer vision tasks and divide DNN-based methods into three sub-categories: source image concatenation (SIC), feature concatenation (FC), and feature fusion (FF), whose general diagrams are shown in Fig. 6. The following is a brief presentation of the representative DNN-based methods:

- *SIC*: For the dual-source input, the straightforward way is combining LR MS and PAN images into one, which is source image concatenation (SIC). SIC implies that the LR MS and PAN images are directly concatenated as an image with $N+1$ channels, which are then fed into DNNs. Masi et al. [140] first adopted a simple and effective three-layer architecture, termed PNN, to learn the mapping between the concatenated image and the HR MS image. Subsequently, Scarpa et al. [141] explored different architectural and training variations and extended PNN as a target-adaptive method to ensure the desired performance.
  Owing to the simplicity of SIC, this kind of method is more concerned about the architecture design of DNNs. The frequently used architectures include residual network [142–146], U-Net [147–149], and multiscale network [150–152]. Some researchers [153,154] integrated DNNs into the *Amélioration de la Résolution Spatiale par Injection de Structures* framework to make full use of the advantages of the two formulations. The attention network was also considered in [155] to learn the nonlocal similarity in source images. For the DNN-based methods that adopt SIC for pre-processing, they cannot efficiently exploit the complementary information in LR MS and PAN images, and the abundant information cannot be suppressed by DNNs.
- *FC*: Compared to the concatenation in original images, FC deals with the information from different sub-networks in the feature domain. In this way, some redundant information is filtered by the pre-processing transformation or sub-networks in advance, which may avoid spatial and spectral distortions in the fused image. For example, Imani [156] stacked the feature maps of the LR MS and PAN images in the frequency domain as the input of a single layer convolution network, which applied 3D Gabor filters and shearlet transform to the domain projection. In [157], a Laplacian pyramid was constructed to model the spatial information at different frequencies.
  To obtain an adaptive transformation for feature extraction, the hand-crafted projection is replaced by the corresponding sub-networks. In [158,159], the outputs of two sub-networks were bundled together in the feature domain and fed into the generative adversarial network (GAN) [160]. An encoder–decoder conditional GAN was designed in [161] to generate more spatial details in the fused image. TFNet [162] is a typical two-stream network to deal with the dual-source input for pan-sharpening. Based on the two-stream architecture, the multiscale property [163,164], feedback connections [165], and dense connections [166] are introduced to extract more subtle feature maps by the sub-networks. Although good performance is achieved for FC, it cannot remove the redundancy among feature maps from different sub-networks. Moreover, the computational complexity may increase owing to the concatenation operation.
- *FF*: FC is a simple and direct operation to deal with the feature maps from different sub-networks. Thus, it is widely adopted by many DNN-based methods. Compared to the simplicity of FC, FF is introduced and intended to merge the complementary information of LR MS and PAN images in the feature domain by different fusion rules [167]. In this sub-category, addition, subtraction, and multiplication between feature maps are often introduced into different levels of the reconstruction network. For example, Zhang et al. [168] demonstrated a bidirectional pyramid network, in which the feature maps of the PAN image were directly added to those of the LR MS image from coarse levels to fine levels. Similarly, Luo [169] constructed a series of stacked fusion units containing feature addition for the generation of the HR MS image. In [170], the addition and subtraction of feature maps were used in combination. Uezato et al. [171] considered the multiplication operation to integrate the feature maps in which the semantic features in the guided decoder were used to facilitate the corresponding decoder.
  Besides the arithmetic operations in the feature domain, Diao et al. [172] utilized two sub-networks with different attention mechanisms [173,174] to extract the spatial and spectral features, which are then integrated by a graph attention module to emphasize the informative feature maps. FF can eliminate the redundancy among the features from different sub-networks. However, it is sometimes difficult to devise proper strategies for

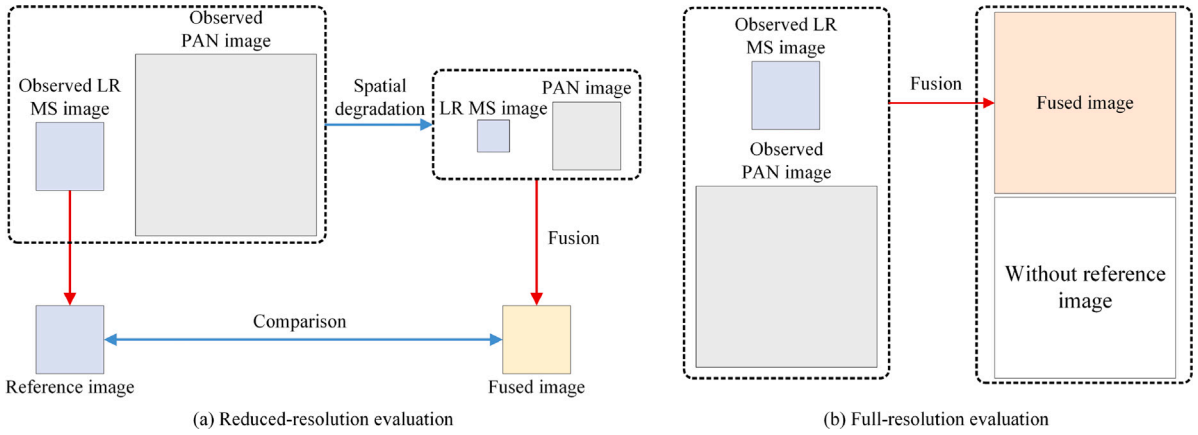

Fig. 7. Evaluation modes of the fused images.

FF. The difficulty of these methods is exploring reasonable and interpretable fusion strategies beyond the arithmetic operations mentioned above.

Apart from the diagrams in Fig. 5, some hybrid methods [175–177] are also proposed, which integrate SIC, FC, and FF together to exploit the spatial and spectral information in LR MS and PAN images. For example, Chen et al. [175] adopted an iterative residual network to produce the fused image, in which SIC and FC are simultaneously used to improve the quality of the fused image progressively. The rapid development of DNNs has created a boom in the pan-sharpening task. At present, there are many other attempts to tackle the pan-sharpening problem [178–180]. Very recently, a recently invented neural network, transformer [181], has already hit the charts in many computer vision tasks. The self-attention mechanism in the transformer is so powerful that it soon catches the attention of many researchers [182–188] in this field. The pan-sharpening driven by the optimization model is another development trend [189–194]. In these methods, the iterative optimization algorithms were unfolded as a network with deep architecture, and a specific network is equipped to learn the priors in images. While the use of DNNs injects new vitality into this field, it also brings challenges and problems, such as generalization. The related issues will be further discussed in Section 4.

3. Image quality evaluation

As a typical inverse problem, the fused image obtained by pan-sharpening cannot be directly compared with the reference image in a real-world scenario. The absence of the reference image makes the quantitative evaluation of the fusion result difficult. Naturally, two kinds of image evaluation protocols are considered: reduced-resolution evaluation and full-resolution evaluation. Fig. 7 displays the evaluation modes in these two cases. Generally, the quality of the fused image is evaluated according to Wald's protocol [29]. Furthermore, Palsson et al. [195] verified that the properties in Wald's protocol could provide a reliable evaluation.

Generally, the fused image is evaluated from three perspectives: spatial information, spectral information, and overall performance. The spatial information in the fused image contains abundant texture and shape characteristics for target detection and recognition [196,197]. The spectral information in the fused image is responsible for the record of the land cover [198,199]. Therefore, a satisfactory pan-sharpening method should achieve high accuracy in terms of spatial and spectral information. Meanwhile, the overall performance of this method will be high. A fair and comprehensive evaluation can provide reliable evidence for the selection of pan-sharpening methods. The following sections describe the spatial indexes, spectral indexes, and overall indexes in reduced-resolution and full-resolution cases.

3.1. Reduced-resolution evaluation

In reduced-resolution cases, the reference image has to be prepared in advance for comparison. To achieve reference-based image evaluation, Wald's protocol [29] is often used as a guideline to synthesize the LR MS and PAN images to be fused by spatial degradation. Then, the original MS image is regarded as the reference image. According to the consistency property of this protocol, the fused image should be identical to the original MS image in the reduced-resolution evaluation. Many indexes are used to measure the similarity between the fused image and the reference image.

- *Spatial Indexes*: Spatial indexes aim to assess the abundance of textures or edges in the fused image. In general, the gradient amplitude of textures or edges in an image is larger. Thus, the average amplitude of gradients (AG) is used in some literature to reflect the spatial quality of the fused image. Besides, the covariance and the standard variance also record the spatial variations in the fused image. Then, Wang et al. [200] employed a universal image quality index (UIQI) to model the spatial distortions in the fused image, which consisted of covariances, standard variances, and means of the fused image and the reference image. The correlation coefficient (CC) [103] and spatial correlation coefficient (SCC) [201] were introduced for the evaluation of the spatial information. Thanks to the achievement in the natural image assessment, the structural similarity (SSIM) index is sometimes utilized as a means for the perception of the structure information in the fused image [202].
- *Spectral Indexes*: For the evaluation of the spectral information, the spectral vector of each pixel in the fused image is compared to that of the corresponding pixel in the reference image. For example, Yuhas et al. [203] proposed a spectral angle mapper (SAM) to calculate the angle between the spectral vectors of the fused images and those of the reference images. The reconstruction errors can be evaluated by both the root-mean-square error (RMSE) and the relative average spectral error (RASE) [204]. According to the formulations of RMSE and RASE, we can know that RMSE is more sensitive to larger errors than RASE. Chang proposed the spectral information divergence (SID) [205] to analyze the spectral differences in the fused image.
- *Overall Indexes*: In overall indexes, both spatial and spectral information is assessed. They provide a global tradeoff between the spatial and spectral information of the fused image. Q4 [206] is a classical overall index, which adopts hypercomplex numbers, or quaternions to measure the similarity of the four-band MS image. *Erreur Relative Globale Adimensionnelle de Synthése* (ERGAS) [207] was presented to compute the amount of spectral and spatial distortions in the fused image. In ERGAS, RMSE and mean values of each band are combined with the spatial resolution ratio.

Although the above-mentioned indexes measure the spatial and spectral properties of the fused image from different perspectives, the commonly approved indexes only are SAM, Q4, and ERGAS. However, the similarity between the fused image and the reference image is measured repeatedly in some indexes to some extent. For example, RMSE is involved in ERGAS. Due to the spatial and spectral property differences in reduced-resolution and full-resolution cases, the reference-based evaluation indexes will introduce a bias with the invalid hypothesis, although these indexes can be easily calculated.

3.2. Full-resolution evaluation

In the full-resolution evaluation case, the spatial and spectral information in the fused image is compared to that in PAN and LR MS images, respectively. In addition, some no-reference indexes are exclusively calculated according to the spatial and spectral properties in the fused image. For full-resolution evaluation, the difficulty lies in the absence of the reference image. The development of full-resolution indexes revolves mainly around the improvements of the quality with no-reference (QNR) index and extensions of reduced-resolution indexes.

- *QNR and Its Variants*: In [208], Alparone et al. proposed three indexes for spatial and spectral information assessment. The first index, spectral distortion index D_λ, is derived from the differences of UIQI from the fused image, which measures the spectral differences with the LR MS image. The spatial distortion index D_s is calculated according to quality index differences between the PAN image and its corresponding spatially degraded version. Then, these two indexes are jointly introduced into the overall index, QNR, to reflect the global quality of the fused image. Then, Khan et al. [209] introduced MTFs matched by filters into D_λ and D_s. Thus, the spectral and spatial information in the fused image can be extracted more reasonably. Palubinskas [210] also proposed a joint quality measure (JQM), which consisted of QNRs in full- and reduced-cases.
- *Extensions of Reduced-resolution Indexes*: Given the availability of the reference-based indexes, their improved versions are applied to the full-resolution evaluation. For instance, Vivone et al. [211] recast the full-resolution evaluation task into a sequential Bayesian framework, which inferred the Q4 value by reproducing the fusion results at the reduced scale. Subsequently, Vivone et al. [212] reduced the computational complexity of the index in [211] by reformulating the state estimation as a convex combination problem. Besides, Carlé et al. [213] utilized linear and quadratic polynomials to fit the multiscale distortion measurements in reduced resolution and then extrapolated them to full-resolution values.

In full-resolution evaluation, the most commonly used metrics are D_λ, D_s and QNR in [208]. However, it is demonstrated that these indexes are neither consistent with the visual performance of the fused image nor reliable. Some new paradigms are proposed for a more objective assessment. Zhou et al. [214] learned a benchmark multivariate Gaussian (MVG) model for full-resolution quality assessment. In the MVG model, the spatial and spectral features were extracted from 360-generated images. Then, the distance between the benchmark MVG and the MVG fitted from the fused image can be taken as the quality assessment. However, the generated images in [214] cannot include all kinds of land covers. It limits the generalization of the benchmark MVG. Thus, it is necessary to establish a more reasonable protocol for full-resolution evaluation.

4. Limitations, difficulties and challenges

We have presented researches on pan-sharpening and discussed the literatures on the quality assessment of the fused image. Though the fusion of LR MS and PAN images has been extensively studied and achieved good performance, there are still great challenges for pan-sharpening. In this section, we share our insights on challenges and new trends for pan-sharpening.

4.1. Dataset

As pan-sharpening enters the deep learning era, more and more data are required for the training of the newly proposed DNN-based methods. Currently, there are no public and universally approved remote sensing image datasets for comparison among different pan-sharpening methods. Actually, it is difficult to establish an all-around dataset for pan-sharpening, which involves the land cover diversity and seasonal changes, and imaging differences of source images. When preparing an all-round dataset, we must take into account the following limitations and difficulties:

- *High Diversity of the Observed Scenes*: The dataset to be constructed should contain as many kinds of land covers as possible. Typical land cover contains grassland, woodland, farmland, waters, rural, and urban areas. For a specific DNN only trained on the rural dataset, the fusion performance will degrade when it is tested on the LR MS and PAN images from one urban dataset. It is difficult to balance the fusion performance for the LR MS and PAN images with various kinds of land covers. Empirically, the spectral information for vegetation areas cannot be preserved well for most of the pan-sharpening methods.

- *Seasonal Variations*: Similarly, the dataset should also consider the feature differences of the surfaces caused by seasonal variations. We cannot ignore the spectral differences of vegetation areas between summer and winter when designing the pan-sharpening scheme. For DNN-based pan-sharpening methods, the spectral changes of the land covers lead to a domain shift of the original dataset. Then, the spectral information of the fused image is distorted.

- *Large Differences in Satellite Imaging Sensors*: Due to the instrumental differences, the spectral responses of different satellites have significant differences. The distinct spectral responses mean that the mapping relationships from source images to the HR MS image will be different for the datasets from two satellites. In addition, the MTFs of different satellites are also not the same. Therefore, if we employ the DNNs trained on the QuickBird dataset to fuse the LR MS and PAN images from the GeoEye-1 satellite, satisfying results may not be achieved.

To sum up, DNN-based pan-sharpening methods are data-hungry. Only larger datasets containing all kinds of images ensure that DNN-based pan-sharpening methods produce good fusion results. Ultimately, the construction of larger datasets is used to alleviate the generalization issues of DNN-based methods. If one kind of source images is not present in the constructed dataset, the fusion results of this kind images produced by DNN-based methods may be inferior to those of the traditional methods. Therefore, it is urgently needed to construct a larger dataset for the sufficient training of DNNs. The all-around dataset can also provide an operational scheme to verify the generalization ability of DNNs on land covers, seasons, and imaging sensors.

4.2. Quality evaluation

As an indispensable preprocessing step, pan-sharpening can efficiently improve the spatial resolution of the MS image. However, the goal is to achieve a meaningful interpretation of the observed scene from the MS image, in which the significant downstream tasks are object detection [215] and image segmentation [216]. So, the evaluation of the fused image should be conducted from two aspects: image-oriented evaluation and task-oriented evaluation. Because of the difficulties of task-oriented evaluation, it is often neglected. We consider the following aspects as the research trends for the evaluation of the fused image.

- *Image-oriented Evaluation*: Generally, the quality of the fused image can be easily evaluated in the reduced-resolution case due to the existing reference HR MS image. Whereas, the synthesis of LR MS and PAN images is decided by MTFs, which may introduce some potential errors. Thus, it is more necessary to evaluate the image quality in the full-resolution case. Unfortunately, the frequently-used D_λ, D_s, and QNR in [208] sometimes are not consistent with the visual performance of the fused image. Different from natural images, the radiometric resolution of remote sensing images is high. The no-reference evaluation indexes for natural images cannot be simply transferred to the assessment of remote sensing images. The design of the no-reference image indexes should comprehensively consider the spatial, spectral, and radiometric properties in remote sensing images.

- *Task-oriented Evaluation*: It is desirable that the quality of the fused image is investigated through downstream tasks. For example, Bovolo et al. [217] investigated the effects on the change detection task of different pan-sharpening methods. As an indispensable preprocessing step, pan-sharpening can efficiently improve the spatial resolution of the MS image. So, the downstream tasks can be performed on the pan-sharpened images of different methods, and then the accuracy of the related tasks can be used as a standard for the quality comparison of the fused image. Compared to the image-oriented evaluation, the task-oriented evaluation approaches will make pan-sharpening methods more practical.

Currently, there are no available object detection or image segmentation datasets for task-oriented evaluation. Thus, a task-oriented dataset becomes particularly important. According to Sections 4.1 and 4.2, we can conclude that this field is desperate for an all-around dataset of high reorganization.

4.3. DNNs for pan-sharpening

- *Training*: For most of the DNN-based pan-sharpening methods, an end-to-end strategy is usually adopted. In these methods, LR MS and PAN images are regarded as the input of DNNs and then DNNs are trained to approximate the corresponding HR MS image. But the HR MS image for training is not acquirable. Then, the original LR MS and PAN images are spatially degraded for the generation of training image pairs. In the generated dataset, the degraded LR MS and PAN images are fed into DNNs and the original MS image is the corresponding reference image. Although good performance has been achieved by end-to-end training, it is difficult to neglect the differences in spatial and spectral properties caused by the spatial resolution ratio between the reduced-resolution and the full-resolution cases. For example, the spatial details in the full-resolution dataset cannot be inferred from the reduced-resolution dataset. Therefore, it is still an open and promising topic to train DNNs on full-resolution data or jointly train them on reduced-resolution and full-resolution data. For instance, Liu et al. [218] tried to combine supervised and unsupervised training together to simultaneously learn the spatial and spectral information in reduced-resolution and full-resolution data.

- *New Paradigms*: In general, unsupervised DNNs are unable to produce competitive results compared to supervised DNNs. Naturally, the spatial and spectral degradation models among the source images and the HR MS image are utilized to facilitate the training of DNNs. Meanwhile, the paradigm will introduce some errors caused by degradation models, which limits the performance of DNNs trained in an unsupervised manner. Recently, unrolling techniques [219,220] have been used in many fields and gained good performance by combining spatial and spectral observation models. Thus, the spatial and spectral model-driven DNNs may be a great opportunity to improve the performance of the pan-sharpened images. Besides, new DNNs are flourishing, such as transformer [181], graph neural network [221], and zero-reference GAN [222]. The evolving DNNs will further improve the quality of the fused image and provide more solutions to the existing issues in this field.

5. Conclusions

Pan-sharpening can efficiently integrate information of LR MS and PAN images. The fused image provides a more comprehensive and reliable description for the observed scene, which is beneficial to the subsequent interpretation tasks, such as target detection and recognition. However, it is difficult to realize in the LR MS images or PAN images. Thus, the fusion of LR MS and PAN images is widely studied and various methods are proposed. In this paper, we briefly arrange these pan-sharpening methods into four categories: CS-based methods, MRA-based methods, DM-based methods, and DNNs-based methods.

For the first category, proper projection methods and injection coefficients play important roles in their formulations. Different transformation methods are tried to separate the spatial and spectral information in the MS image. Spectral physical properties are considered to compute more accurate injection coefficients. In the second category, MRA is viewed as a high-frequency extractor for the estimation of spatial details, where various MRA tools focus on different image information. Injection coefficients are calculated by different global or local models. By combining different priors in LR MS and PAN images, the spatial and spectral degradation models can be regularized suitably in DM-based methods. Generally, various optimization algorithms are employed to solve their final fusion models because of the introduction of well-designed constraints. In the fourth category, massive attention is paid to the architecture design for DNNs to deal with the dual input caused by LR MS and PAN images. In addition, unsupervised training is also explored in this field because HR MS images are not available for training.

Besides, we present the quality evaluation of the fused image, which can be assessed by reduced-resolution and full-resolution indexes. In the reduced-resolution case, the reference image is necessary. Q4, SAM, and ERGAS are usually used for evaluation. For the fused image at full resolution, popular metrics are D_λ, D_s, and QNR. However, it is proved that these metrics are not proper in some literature. Thus, more reasonable full-resolution metrics should be explored further. Additionally, we also list promising directions regarding datasets, image evaluation, and new paradigms.

CRediT authorship contribution statement

Kai Zhang: Conceptualization, Investigation, Formal analysis, Software, Writing – review & editing. **Feng Zhang:** Investigation, Formal analysis, Writing – review & editing. **Wenbo Wan:** Investigation, Formal analysis, Writing – review & editing. **Hui Yu:** Conceptualization, Methodology, Formal analysis, Writing – review & editing, Supervision. **Jiande Sun:** Conceptualization, Methodology, Investigation, Formal analysis, Validation, Writing – review & editing, Supervision, Project administration. **Javier Del Ser:** Validation, Methodology, Writing – review & editing. **Eyad Elyan:** Conceptualization, Methodology, Writing – review & editing. **Amir Hussain:** Conceptualization, Methodology, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This work was supported in part by the Natural Science Foundation of China (61901246), the China Postdoctoral Science Foundation, China Grant (2019TQ0190, 2019M662432), the Scientific Research Leader Studio of Ji'nan (2021GXRC081), and Joint Project for Smart Computing of Shandong Natural Science Foundation, China (ZR2020LZH015). Amir Hussain acknowledges the support of the UK Engineering and Physical Sciences Research Council (EPSRC)-Grants Ref. EP/M026981/1, EP/T021063/1, EP/T024917/1. Hui Yu acknowledges the support of Royal Society, UK (NIF/R1/180909). J. Del Ser would like to thank the Spanish Centro para el Desarrollo Tecnologico Industrial (CDTI, Ministry of Science and Innovation) through the "Red Cervera" Programme (AI4ES project), as well as by the Basque Government, Spain through the ELKARTEK program and the Consolidated Research Group MATHMODE (Ref. IT1456-22).

References

[1] P. Han, C. Ma, Q. Li, P. Leng, S. Bu, K. Li, Aerial image change detection using dual regions of interest networks, Neurocomputing 349 (2019) 190–201.
[2] S. Wang, L. Liu, L. Qu, C. Yu, Y. Sun, F. Gao, J. Dong, Accurate ulva prolifera regions extraction of UAV images with superpixel and CNNs for ocean environment monitoring, Neurocomputing 348 (2019) 158–168.
[3] G. Huang, Z. Wan, X. Liu, J. Hui, Z. Wang, Z. Zhang, Ship detection based on squeeze excitation skip-connection path networks for optical remote sensing images, Neurocomputing 332 (2019) 215–223.
[4] L. Wang, Z. Xiong, G. Shi, W. Zeng, F. Wu, Adaptive nonlocal sparse representation for dual-camera compressive hyperspectral imaging, IEEE Trans. Pattern Anal. Mach. Intell. 39 (10) (2017) 2104–2111.
[5] B. Hou, Z. Ren, W. Zhao, Q. Wu, L. Jiao, Object detection in high-resolution panchromatic images using deep models and spatial template matching, IEEE Trans. Geosci. Remote Sens. 58 (2) (2020) 956–970.
[6] C. Paris, L. Bruzzone, D. Fernandez-Prieto, A novel approach to the unsupervised update of land-cover maps by classification of time series of multispectral images, IEEE Trans. Geosci. Remote Sens. 57 (7) (2019) 4259–4277.
[7] X. Tu, X. Shen, P. Fu, T. Wang, Q. Sun, Z. Ji, Discriminant sub-dictionary learning with adaptive multiscale superpixel representation for hyperspectral image classification, Neurocomputing 409 (2020) 131–145.
[8] Z. Wang, D. Ziou, C. Armenakis, D. Li, Q. Li, A comparative analysis of image fusion methods, IEEE Trans. Geosci. Remote Sens. 43 (6) (2005) 1391–1402.
[9] L. Alparone, L. Wald, J. Chanussot, C. Thomas, P. Gamba, L.M. Bruce, Comparison of pansharpening algorithms: Outcome of the 2006 GRS-S data-fusion contest, IEEE Trans. Geosci. Remote Sens. 45 (10) (2007) 3012–3021.
[10] R. Hnsch, C. Persello, G. Vivone, J. Navarro, A. Boulch, S. Lefevre, B. Saux, Data fusion contest 2022 (DFC2022), in: IEEE Dataport, 2022.
[11] Y. Byun, J. Choi, Y. Han, An area-based image fusion scheme for the integration of SAR and optical satellite imagery, IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. 5 (1) (2012) 125–134.
[12] H. Wang, C. Glennie, Fusion of waveform LiDAR data and hyperspectral imagery for land cover classification, ISPRS J. Photogramm. Remote Sens. 108 (2015) 1–11.
[13] H. Song, B. Huang, Spatiotemporal satellite image fusion through one-pair image learning, IEEE Trans. Geosci. Remote Sens. 51 (4) (2013) 1883–1896.
[14] K. Zhang, M. Wang, S. Yang, Multispectral and hyperspectral image fusion based on group spectral embedding and low-rank factorization, IEEE Trans. Geosci. Remote Sens. 55 (3) (2017) 1363–1371.
[15] K. Zhang, M. Wang, S. Yang, L. Jiao, Spatial-spectral-graph-regularized low-rank tensor decomposition for multispectral and hyperspectral image fusion, IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. 11 (4) (2018) 1030–1040.
[16] F. Zhang, K. Zhang, Superpixel guided structure sparsity for multispectral and hyperspectral image fusion over couple dictionary, Multimedia Tools Appl. 79 (2020) 4949–4964.
[17] I. Amro, J. Mateos, M. Vega, R. Molina, A.K. Katsaggelos, A survey of classical methods and new trends in pansharpening of multispectral images, EURASIP J. Advan. Signal Process. 79 (2011) 1–22.
[18] X. Liu, L. Li, F. Liu, B. Hou, S. Yang, L. Jiao, GAFNet: Group attention fusion network for PAN and MS image high-resolution classification, IEEE Trans. Cyber. 52 (10) (2022) 10556–10569.
[19] Y. Liao, H. Zhu, L. Jiao, X. Li, N. Li, K. Sun, X. Tang, B. Hou, A two-stage mutual fusion network for multispectral and panchromatic image classification, IEEE Trans. Geosci. Remote Sens. 60 (2022) 5413218.
[20] Y. Tan, S. Xiong, Y. Li, Automatic extraction of built-up areas from panchromatic and multispectral remote sensing images using double-stream deep convolutional neural networks, IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. 11 (11) (2018) 3988–4004.
[21] F. Giacco, C. Thiel, L. Pugliese, S. Scarpetta, M. Marinaro, Uncertainty analysis for the classification of multispectral satellite images using SVMs and SOMs, IEEE Trans. Geosci. Remote Sens. 48 (10) (2010) 3769–3779.
[22] J. Zhang, Multi-source remote sensing data fusion: status and trends I, J. Image Data Fusion 1 (2010) 5–24.
[23] B. Hu, Q. Li, G.B. Hall, A decision-level fusion approach to tree species classification from multi-source remotely sensed data, ISPRS Open J. Photogram. Remote Sens. 1 (2021) 100002.
[24] H. Zhao, S. Liu, Q. Du, L. Bruzzone, Y. Zheng, et al., GCFnet: Global collaborative fusion network for multispectral and panchromatic image classification, IEEE Trans. Geosci. Remote Sens. 60 (2022) 5632814.
[25] B. Rayegani, S. Barati, H. Goshtasb, H. Sarkheil, J. Ramezani, An effective approach to selecting the appropriate pan-sharpening method in digital change detection of natural ecosystems, Ecolog. Infor. 53 (2019) 100984.
[26] R. Lottering, O. Mutanga, K. Peerbhay, R. Ismail, Detecting and mapping gonipterus scutellatus induced vegetation defoliation using WorldView-2 pan-sharpened image texture combinations and an artificial neural network, J. Appli. Remote Sensing 13 (1) (2019) 014513.
[27] Y. Qu, H. Qi, B. Ayhan, C. Kwan, R. Kidd, DOES multispectral/hyperspectral pansharpening improve the performance of anomaly detection? in: IEEE IGARSS, 2017, pp. 1–4.
[28] P. Du, S. Liu, J. Xia, Y. Zhao, Information fusion techniques for change detection from multi-temporal remote sensing images, Inf. Fusion 14 (2013) 19–27.
[29] L. Wald, T. Ranchin, M. Mangolini, Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images, Photogramm. Eng. Remote Sens. 63 (6) (1997) 691–699.
[30] C. Thomas, T. Ranchin, L. Wald, J. Chanussot, Synthesis of multispectral images to high spatial resolution: A critical review of fusion methods based on remote sensing physics, IEEE Trans. Geosci. Remote Sens. 46 (5) (2008) 1301–1312.
[31] G. Vivone, L. Alparone, J. Chanussot, M.D. Mura, A. Garzelli, G.A. Licciardi, R. Restaino, L. Wald, A critical comparison among pansharpening algorithms, IEEE Trans. Geosci. Remote Sens. 53 (5) (2015) 2565–2586.
[32] G. Vivone, M.D. Mura, A. Garzelli, R. Restaino, et al., A new benchmark based on recent advances in multispectral pansharpening: Revisiting pansharpening with classical and emerging pansharpening methods, IEEE Geosci. Remote Sens. Mag. 9 (1) (2021) 53–81.
[33] X. Meng, Y. Xiong, F. Shao, H. Shen, W. Sun, et al., A large-scale benchmark data set for evaluating pansharpening performance: Overview and implementation, IEEE Geosci. Remote Sens. Mag. 9 (1) (2021) 18–52.
[34] W.J. Carper, T.M. Lillesand, R.W. Kiefer, The use of intensity-hue-saturation transformations for merging SPOT panchromatic and multispectral image data, Photogramm. Eng. Remote Sens. 56 (4) (1990) 459–467.
[35] P.S. Chavez, S.C. Slides, J.A. Anderson, Comparison of three different methods to merge multiresolution and multispectral data: Landsat TM and SPOT panchromatic, Photogramm. Eng. Remote Sens. 57 (3) (1991) 295–303.
[36] C.A. Laben, B.V. Brower, Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening, 2000, U.S. Patent 6 011 875, Jan. 4.
[37] T. Ranchin, L. Wald, Fusion of high spatial and spectral resolution images: The ARSIS concept and its implementation, Photogramm. Eng. Remote Sens. 66 (1) (2000) 49–61.
[38] S. Mallat, A Wavelet Tour of Signal Processing, 3rd ed., Elsevier, 2008.
[39] S. Zheng, W. Shi, J. Liu, J. Tian, Remote sensing image fusion using multiscale mapped LS-SVM, IEEE Trans. Geosci. Remote Sens. 46 (5) (2008) 1313–1322.
[40] Y. Xing, M. Wang, S. Yang, K. Zhang, Pansharpening with multiscale geometric support tensor machine, IEEE Trans. Geosci. Remote Sens. 56 (5) (2018) 2503–2517.
[41] R. Restaino, G. Vivone, M. Dalla Mura, J. Chanussot, Fusion of multispectral and panchromatic images based on morphological operators, IEEE Trans. Image Process. 25 (6) (2016) 2882–2895.
[42] S. Li, X. Kang, L. Fang, J. Hu, H. Yin, Pixel-level image fusion: A survey of the state of the art, Inf. Fusion 33 (2017) 100–112.
[43] C. Ballester, V. Caselles, L. Igual, J. Verdera, A variational model for P+XS image fusion, Int. J. Comput. Vis. 69 (1) (2006) 43–58.
[44] S. Yang, K. Zhang, M. Wang, Learning low-rank decomposition for pan-sharpening with spatial–spectral offsets, IEEE Trans. Neural Netw. Learn. Syst. 29 (8) (2018) 3647–3657.
[45] C. Ren, X. He, T.Q. Nguyen, Single image super-resolution via adaptive high-dimensional non-local total variation and adaptive geometric feature, IEEE Trans. Image Process. 26 (1) (2017) 90–106.
[46] H. Wang, Y. Cen, Z. He, Z. He, R. Zhao, F. Zhang, Reweighted low-rank matrix analysis with structural smoothness for image denoising, IEEE Trans. Image Process. 27 (4) (2018) 1777–1792.
[47] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (7553) (2015) 436–444.

[48] A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, A.A. Bharath, Generative adversarial networks: An overview, IEEE Signal Process. Mag. 35 (1) (2018) 53–65.

[49] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, et al., Attention is all you need, in: NIPS, 2017, pp. 5998–6008.

[50] A. Garzelli, F. Nencini, L. Capobianco, Optimal MMSE pan sharpening of very high resolution multispectral images, IEEE Trans. Geosci. Remote Sens. 46 (1) (2008) 228–236.

[51] T. Tu, P. Huang, C. Hung, C. Chang, A fast intensity-hue-saturation fusion technique with spectral adjustment for IKONOS imagery, IEEE Geosci. Remote Sens. Lett. 1 (4) (2015) 180–184.

[52] S. Rahmani, M. Strait, D. Merkurjev, M. Moeller, T. Wittman, An adaptive IHS pan-sharpening method, IEEE Geosci. Remote Sens. Lett. 7 (4) (2010) 746–750.

[53] Y. Leung, J. Liu, J. Zhang, An improved adaptive intensity-hue-saturation method for the fusion of remote sensing images, IEEE Geosci. Remote Sens. Lett. 11 (5) (2014) 985–989.

[54] M. Ghahremani, H. Ghassemian, Nonlinear IHS: A promising method for pan-sharpening, IEEE Geosci. Remote Sens. Lett. 12 (11) (2016) 1606–1610.

[55] S. Yang, M. Wang, L. Jiao, Fusion of multispectral and panchromatic images based on support value transform and adaptive principal component analysis, Inf. Fusion 13 (2012) 177–184.

[56] H. Shahdoosti, H. Ghassemian, Combining the spectral PCA and spatial PCA fusion methods by an optimal filter, Inf. Fusion 27 (2016) 150–160.

[57] Y. Kim, M. Kim, J. Choi, Y. Kim, Image fusion of spectrally nonoverlapping imagery using SPCA and MTF-based filters, IEEE Geosci. Remote Sens. Lett. 14 (12) (2017) 2295–2299.

[58] J. Duran, A. Buades, Restoration of pansharpened images by conditional filtering in the PCA domain, IEEE Geosci. Remote Sens. Lett. 16 (3) (2019) 442–446.

[59] B. Aiazzi, S. Baronti, M. Selva, Improving component substitution pansharpening through multivariate regression of MS+Pan data, IEEE Trans. Geosci. Remote Sens. 45 (10) (2007) 3230–3239.

[60] W. Wang, L. Jiao, S. Yang, Novel adaptive component-substitution-based pansharpening using particle swarm optimization, IEEE Geosci. Remote Sens. Lett. 12 (4) (2015) 781–785.

[61] A. Garzelli, Pansharpening of multispectral images based on nonlocal parameter optimization, IEEE Trans. Geosci. Remote Sens. 53 (4) (2015) 2096–2107.

[62] H. Shahdoosti, N. Javaheri, Pansharpening of clustered MS and pan images considering mixed pixels, IEEE Geosci. Remote Sens. Lett. 14 (6) (2017) 826–830.

[63] M. Imani, Band dependent spatial details injection based on collaborative representation for pansharpening, IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. 11 (12) (2018) 4994–5004.

[64] G. Vivone, Robust band-dependent spatial-detail approaches for panchromatic sharpening, IEEE Trans. Geosci. Remote Sens. 57 (9) (2019) 6421–6432.

[65] G. Vivone, R. Restaino, J. Chanussot, Full scale regression-based injection coefficients for panchromatic sharpening, IEEE Trans. Image Process. 27 (7) (2018) 3418–3430.

[66] G. Vivone, S. Marano, J. Chanussot, Pansharpening: Context-based generalized Laplacian pyramids by robust regression, IEEE Trans. Geosci. Remote Sens. 58 (9) (2020) 6152–6167.

[67] R. Restaino, G. Vivone, P. Addesso, J. Chanussot, A pansharpening approach based on multiple linear regression estimation of injection coefficients, IEEE Geosci. Remote Sens. Lett. 17 (1) (2020) 102–106.

[68] P. Addesso, G. Vivone, R. Restaino, J. Chanussot, A data-driven model-based regression applied to panchromatic sharpening, IEEE Trans. Image Process. 29 (2020) 7779–7793.

[69] X. Otazu, M. Gonzalez-Audicana, O. Fors, J. Nunez, Introduction of sensor spectral response into image fusion methods. Application to wavelet-based methods, IEEE Trans. Geosci. Remote Sens. 43 (10) (2005) 2376–2385.

[70] G. Vivone, L. Alparone, A. Garzelli, S. Lolli, Fast reproducible pansharpening based on instrument and acquisition modeling: AWLP revisited, Remote Sens. 11 (2019) 2315.

[71] X. Lu, J. Zhang, T. Li, Y. Zhang, Pan-sharpening by multilevel interband structure modeling, IEEE Geosci. Remote Sens. Lett. 13 (6) (2016) 892–896.

[72] A. Kallel, Pansharpening: MTF-adjusted pansharpening approach based on coupled multiresolution decompositions, IEEE Trans. Geosci. Remote Sens. 53 (6) (2015) 3124–3145.

[73] Y. Yang, C. Wan, S. Huang, H. Lu, W. Wan, Pansharpening based on low-rank fuzzy fusion and detail supplement, IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. 13 (2020) 5466–5479.

[74] V. Shah, N. Younan, R. King, Pansharpening: An efficient pan-sharpening method via a combined adaptive PCA approach and contourlets, IEEE Trans. Geosci. Remote Sens. 46 (5) (2008) 1323–1335.

[75] M. El-Mezouar, K. Kpalma, N. Taleb, J. Ronsin, A pan-sharpening based on the non-subsampled contourlet transform: Application to worldview-2 imagery, IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. 47 (5) (2014) 1806–1815.

[76] K. Upla, M. Joshi, P. Gajjar, An edge preserving multiresolution fusion: Use of contourlet transform and MRF prior, IEEE Trans. Geosci. Remote Sens. 53 (6) (2015) 3210–3220.

[77] H. Li, F. Liu, S. Yang, K. Zhang, X. Su, L. Jiao, Refined pan-sharpening with NSCT and hierarchical sparse autoencoder, IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. 9 (12) (2016) 5715–5725.

[78] L. Dong, Q. Yang, H. Wu, H. Xiao, M. Xu, High quality multi-spectral and panchromatic image fusion technologies based on curvelet transform, Neurocomputing 159 (2015) 268–274.

[79] S. Devulapalli, R. Krishnan, Synthesized pansharpening using curvelet transform and adaptive neuro-fuzzy inference system, J. Appli. Remote Sensing 13 (3) (2019) 034519.

[80] Y. Shi, X. Yang, T. Cheng, Pansharpening of multispectral images using the nonseparable framelet lifting transform with high vanishing moments, Inf. Fusion 20 (2014) 213–224.

[81] J. Wang, X. Yang, R. Zhu, Random walks for pansharpening in complex tight framelet domain, IEEE Trans. Geosci. Remote Sens. 57 (7) (2019) 5121–5134.

[82] Y. Zhao, B. Wu, A framelet-based SFIM method to pan-sharpen THEOS imagery, J. Indian Soc. Remote Sens. 47 (8) (2019) 1417–1429.

[83] C. Shi, Q. Miao, P. Xu, A novel algorithm of remote sensing image fusion based on shearlets and PCNN, Neurocomputing 117 (2013) 47–53.

[84] A. Moonon, J. Hu, S. Li, Remote sensing image fusion method based on nonsubsampled shearlet transform and sparse representation, Sens. Imaging 16 (2015) 23.

[85] Y. Yang, W. Wan, S. Huang, F. Yuan, Remote sensing image fusion based on adaptive IHS and multiscale guided filter, IEEE Access 4 (2016) 4573–4582.

[86] H. Yin, S. Li, Pansharpening with multiscale normalized nonlocal means filter: A two-step approach, IEEE Trans. Geosci. Remote Sens. 53 (10) (2015) 5734–5745.

[87] K. Zhan, J. Shi, H. Wang, Y. Xie, Q. Li, Computational mechanisms of pulse-coupled neural networks: A comprehensive review, Arch. Computa. Methods Eng. 24 (2017) 573–588.

[88] D.L. Donoho, Compressed sensing, IEEE Trans. Inform. Theory 52 (4) (2006) 1289–1306.

[89] M. Ghulyani, M. Arigovindan, Fast roughness minimizing image restoration under mixed Poisson-Gaussian noise, IEEE Trans. Image Process. 30 (2021) 134–149.

[90] S. Li, B. Yang, A new pan-sharpening method using a compressed sensing technique, IEEE Trans. Geosci. Remote Sens. 49 (2) (2011) 738–746.

[91] S. Chen, D. Donoho, M. Saunders, Atomic decomposition by basis pursuit, SIAM Rev. 43 (1) (2001) 129–159.

[92] S. Li, H. Yin, L. Fang, Remote sensing image fusion via sparse representations over learned dictionaries, IEEE Trans. Geosci. Remote Sens. 51 (9) (2013) 4779–4789.

[93] C. Jiang, H. Zhang, H. Shen, L. Zhang, A practical compressed sensing-based pan-sharpening method, IEEE Geosci. Remote Sens. Lett. 9 (4) (2012) 629–633.

[94] M. Cheng, C. Wang, J. Li, Sparse representation based pansharpening using trained dictionary, IEEE Geosci. Remote Sens. Lett. 11 (1) (2014) 293–297.

[95] M. Ghahremani, H. Ghassemian, A compressed-sensing-based pan-sharpening method for spectral distortion reduction, IEEE Trans. Geosci. Remote Sens. 54 (4) (2016) 2194–2206.

[96] X. Zhu, R. Bamler, A sparse image fusion algorithm with application to pan-sharpening, IEEE Trans. Geosci. Remote Sens. 51 (5) (2013) 2827–2836.

[97] K. Zhang, F. Zhang, S. Yang, S. Lolli, Fusion of multispectral and panchromatic images via spatial weighted neighbor embedding, Remote Sens. 11 (2019) 557.

[98] W. Wang, L. Jiao, S. Yang, K. Rong, Distributed compressed sensing-based pan-sharpening with hybrid dictionary, Neurocomputing 155 (2015) 320–333.

[99] L. Deng, G. Vivone, W. Guo, M. Dalla Mura, J. Chanussot, A variational pan-sharpening approach based on reproducible kernel Hilbert space and Heaviside function, IEEE Trans. Image Process. 27 (9) (2018) 4330–4344.

[100] S. Ayas, E. Gormus, M. Ekinci, An efficient pan sharpening via texture based dictionary learning and sparse representation, IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. 11 (7) (2008) 2448–2460.

[101] K. Zhang, M. Wang, S. Yang, L. Jiao, Convolution structure sparse coding for fusion of panchromatic and multispectral images, IEEE Trans. Geosci. Remote Sens. 57 (2) (2019) 1117–1130.

[102] B. Wohlberg, Efficient algorithms for convolutional sparse representations, IEEE Trans. Image Process. 25 (1) (2016) 301–315.

[103] K. Zhang, F. Zhang, Z. Feng, J. Sun, Q. Wu, Fusion of panchromatic and multispectral images using multiscale convolution sparse decomposition, IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. 14 (2021) 426–439.

[104] L. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, Physica D 60 (1992) 259–269.

[105] F. Palsson, J. Sveinsson, M. Ulfarsson, A new pansharpening algorithm based on total variation, IEEE Geosci. Remote Sens. Lett. 11 (1) (2014) 318–322.

[106] M. Lotfi, H. Ghassemian, A new variational model in texture space for pansharpening, IEEE Geosci. Remote Sens. Lett. 15 (8) (2018) 1269–1273.

[107] P. Liu, L. Xiao, S. Tang, A new geometry enforcing variational model for pan-sharpening, IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. 9 (12) (2016) 5726–5739.

[108] L. Deng, M. Feng, X. Tai, The fusion of panchromatic and multispectral remote sensing images via tensor-based sparse modeling and hyper-Laplacian prior, Inf. Fusion 52 (2019) 76–89.

[109] F. Fang, F. Li, C. Shen, G. Zhang, A variational approach for pan-sharpening, IEEE Trans. Image Process. 22 (7) (2013) 2822–2834.
[110] X. Fu, Z. Lin, Y. Huang, X. Ding, A variational pan-sharpening with local gradient constraints, in: IEEE CVPR, 2019, pp. 10265–10274.
[111] X. Meng, H. Shen, Q. Yuan, H. Li, L. Zhang, W. Sun, Pansharpening for cloud-contaminated very high-resolution remote sensing images, IEEE Trans. Geosci. Remote Sens. 57 (5) (2019) 2840–2854.
[112] T. Wang, F. Fang, F. Li, G. Zhang, High-quality bayesian pansharpening, IEEE Trans. Image Process. 28 (1) (2019) 227–239.
[113] C. Chen, Y. Li, W. Liu, J. Huang, SIRF: Simultaneous satellite image registration and fusion in a unified framework, IEEE Trans. Image Process. 24 (11) (2015) 4213–4224.
[114] P. Liu, L. Xiao, J. Zhang, B. Naz, Spatial-hessian-feature-guided variational model for pan-sharpening, IEEE Trans. Geosci. Remote Sens. 54 (4) (2016) 2235–2253.
[115] P. Liu, Liang Xiao, Multicomponent driven consistency priors for simultaneous decomposition and pansharpening, IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. 12 (11) (2019) 4589–4605.
[116] X. Tian, Y. Chen, C. Yang, X. Gao, J. Ma, A variational pansharpening method based on gradient sparse representation, IEEE Signal Process. Lett. 27 (2020) 1180–1184.
[117] J. Wright, A. Ganesh, S. Rao, Y. Peng, Y. Ma, Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization, in: NIPS, 2009, pp. 2080–2088.
[118] T. Zhou, D. Tao, Godec: Randomized low-rank and sparse matrix decomposition in noisy case, in: ICML, 2011, pp. 33–40.
[119] K. Rong, L. Jiao, S. Wang, F. Liu, Pansharpening based on low-rank and sparse decomposition, IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. 7 (12) (2014) 4793–4805.
[120] F. Palsson, M. Ulfarsson, J. Sveinsson, Model-based reduced-rank pansharpening, IEEE Geosci. Remote Sens. Lett. 17 (4) (2020) 656–660.
[121] X. He, L. Condat, J. Bioucas-Dias, J. Chanussot, Junshi Xia, A new pansharpening method based on spatial and spectral sparsity priors, IEEE Trans. Image Process. 23 (9) (2014) 4160–4164.
[122] P. Liu, L. Xiao, T. Li, A variational pan-sharpening method based on spatial fractional-order geometry and spectral-spatial low-rank priors, IEEE Trans. Geosci. Remote Sens. 56 (3) (2018) 1788–1820.
[123] F. Zhang, H. Zhang, K. Zhang, Y. Xing, J. Sun, Q. Wu, Exploiting low-rank and sparse properties in strided convolution matrix for pansharpening, IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. 14 (2021) 2649–2661.
[124] W. Wang, L. Jiao, S. Yang, Fusion of multispectral and panchromatic images via sparse representation and local autoregressive model, Inf. Fusion 20 (2014) 73–87.
[125] Z. Fu, Y. Zhao, Y. Xu, L. Xu, J. Xu, Gradient structural similarity based gradient filtering for multi-modal image fusion, Inf. Fusion 53 (2020) 251–268.
[126] G. Khademi, H. Ghassemian, Incorporating an adaptive image prior model into Bayesian fusion of multispectral and panchromatic images, IEEE Geosci. Remote Sens. Lett. 15 (6) (2018) 917–921.
[127] J. Duran, A. Buades, B. Coll, C. Sbert, A nonlocal variational model for pansharpening image fusion, SIAM J. Imag. Sci. 7 (2) (2014) 761–796.
[128] K. Zhang, M. Wang, S. Yang, Y. Xing, R. Qu, Fusion of panchromatic and multispectral images via coupled sparse non-negative matrix factorization, IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. 9 (12) (2016) 5740–5747.
[129] A. Krizhevsky, I. Sutskever, G.E. Hinton, ImageNet classification with deep convolutional neural networks, in: NIPS, 2012, pp. 1106–1114.
[130] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (2015) 436–444.
[131] H. Zhang, H. Xu, X. Tian, J. Jiang, J. Ma, Image fusion meets deep learning: A survey and perspective, Inf. Fusion 76 (2021) 323–336.
[132] M. Abdar, F. Pourpanah, S. Hussain, D. Rezazadegan, et al., A review of uncertainty quantification in deep learning: Techniques, applications and challenges, Inf. Fusion 76 (2021) 243–297.
[133] J. Luengo, R. Moreno, I. Sevillano, D. Charte, et al., A tutorial on the segmentation of metallographic images: Taxonomy, new MetalDAM dataset, deep learning-based ensemble model, experimental analysis and challenges, Inf. Fusion 78 (2022) 232–253.
[134] A.D. Martinez, J.D. Ser, E. Villar-Rodriguez, E. Osaba, et al., Lights and shadows in evolutionary deep learning: Taxonomy, critical methodological analysis, cases of study, learned lessons, recommendations and challenges, Inf. Fusion 67 (2021) 161–194.
[135] F. Piccialli, V.D. Somma, F. Giampaolo, S. Cuomo, G. Fortino, A survey on deep learning in medicine: Why, how and when? Inf. Fusion 66 (2021) 111–137.
[136] K. Zhang, W. Zuo, Y. Chen, D. Meng, L. Zhang, Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising, IEEE Trans. Image Process. 26 (7) (2017) 3142–3155.
[137] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: IEEE CVPR, 2016, pp. 770–778.
[138] C. Dong, C. Loy, K. He, X. Tang, Image super-resolution using deep convolutional networks, IEEE Trans. Pattern Anal. Mach. Intell. 38 (2) (2016) 295–307.
[139] D. Ren, K. Zhang, Q. Wang, Q. Hu, W. Zuo, Neural blind deconvolution using deep priors, in: IEEE CVPR, 2020, pp. 3338–3347.
[140] G. Masi, D. Cozzolino, L. Verdoliva, G. Scarpa, Pansharpening by convolutional neural networks, Remote Sens. 8 (2016) 594.
[141] G. Scarpa, S. Vitale, D. Cozzolino, Target-adaptive CNN-based pansharpening, IEEE Trans. Geosci. Remote Sens. 56 (8) (2018) 5443–5457.
[142] J. Yang, X. Fu, Y. Hu, Y. Huang, X. Ding, J. Paisley, PanNet: A deep network architecture for pan-sharpening, in: IEEE ICCV, 2017, pp. 5449–5457.
[143] X. Fu, W. Wang, Y. Huang, X. Ding, J. Paisley, Deep multiscale detail networks for multiband spectral image sharpening, IEEE Trans. Neural Netw. Learn. Syst. 32 (5) (2021) 2090–2104.
[144] Y. Wei, Q. Yuan, H. Shen, L. Zhang, Boosting the accuracy of multispectral image pansharpening by learning a deep residual network, IEEE Geosci. Remote Sens. Lett. 14 (10) (2017) 1795–1799.
[145] M. Jiang, H. Shen, J. Li, Q. Yuan, L. Zhang, A differential information residual convolutional neural network for pansharpening, ISPRS J. Photogram. Remote Sens. 163 (2020) 257–271.
[146] T. Benzenati, A. Kallel, Y. Kessentini, Two stages pan-sharpening details injection approach based on very deep residual networks, IEEE Trans. Geosci. Remote Sens. 59 (6) (2021) 4984–4992.
[147] W. Yao, Z. Zeng, C. Lian, H. Tang, Pixel-wise regression using U-net and its application on pansharpening, Neurocomputing 312 (2018) 364–371.
[148] X. Wu, T. Huang, L. Deng, T. Zhang, Dynamic cross feature fusion for remote sensing pansharpening, in: IEEE ICCV, 2021, pp. 14687–14696.
[149] W. Diao, F. Zhang, H. Wang, W. Wan, J. Sun, K. Zhang, HLF-net: Pansharpening based on high- and low-frequency fusion networks, IEEE Geosci. Remote Sens. Lett. 19 (2022) 5003705.
[150] Q. Yuan, Y. Wei, X. Meng, H. Shen, L. Zhang, A multiscale and multidepth convolutional neural network for remote sensing imagery pan-sharpening, IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. 11 (3) (2018) 978–989.
[151] J. Hu, P. H, X. Kang, H. Zhang, S. Fan, Pan-sharpening via multiscale dynamic convolutional neural network, IEEE Trans. Geosci. Remote Sens. 59 (3) (2021) 2231–2243.
[152] D. Lei, Y. Huang, L. Zhang, W. Li, Multibranch feature extraction and feature multiplexing network for pansharpening, IEEE Trans. Geosci. Remote Sens. 60 (2022) 5402613.
[153] L. He, Y. Rao, J. Li, J. Chanussot, A. Plaza, J. Zhu, B. Li, Pansharpening via detail injection based convolutional neural networks, IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. 12 (4) (2019) 1188–1204.
[154] L. Deng, G. Vivone, C. Jin, J. Chanussot, Detail injection-based deep convolutional neural networks for pansharpening, IEEE Trans. Geosci. Remote Sens. 59 (8) (2021) 6995–7010.
[155] D. Lei, H. Chen, L. Zhang, W. Li, NLRNet: An efficient nonlocal attention resnet for pansharpening, IEEE Trans. Geosci. Remote Sens. 60 (2022) 5401113.
[156] M. Imani, Texture feed based convolutional neural network for pansharpening, Neurocomputing 398 (2020) 117–130.
[157] C. Jin, L. Deng, T. Huang, G. Vivone, Laplacian pyramid networks: A new approach for multispectral pansharpening, Inf. Fusion 78 (2022) 158–170.
[158] J. Ma, W. Yu, C. Chen, P. Liang, X. Guo, J. Jiang, Pan-GAN: An unsupervised pan-sharpening method for remote sensing image fusion, Inf. Fusion 62 (2020) 110–120.
[159] X. Liu, Y. Wang, Q. Liu, PSGAN: A generative adversarial network for remote sensing image, in: IEEE ICIP, 2018, pp. 873–877.
[160] I. Goodfellow, J. Pouget-Abadie, M. Mirza, et al., Generative adversarial nets, in: NIPS, 2014, pp. 2672–2680.
[161] Z. Shao, Z. Lu, M. Ran, L. Fang, J. Zhou, Y. Zhang, Residual encoder–decoder conditional generative adversarial network for pansharpening, IEEE Geosci. Remote Sens. Lett. 17 (9) (2020) 1573–1577.
[162] X. Liu, Q. Liu, Y. Wang, Remote sensing image fusion based on two-stream fusion network, Inf. Fusion 55 (2020) 1–15.
[163] J. Wei, Y. Xu, W. Cai, Z. Wu, J. Chanussot, Z. Wei, A two-stream multiscale deep learning architecture for pan-sharpening, IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. 13 (2020) 5455–5465.
[164] F. Ozcelik, U. Alganci, E. Sertel, G. Unal, Rethinking CNN-based pansharpening: Guided colorization of panchromatic images via GANs, IEEE Trans. Geosci. Remote Sens. 21 (4) (2021) 3486–3501.
[165] S. Fu, W. Meng, G. Jeon, A. Chehri, R. Zhang, X. Yang, Two-path network with feedback connections for pan-sharpening in remote sensing, Remote Sens. 12 (2020) 1674.
[166] D. Wang, Y. Li, L. Ma, Z. Bai, J. Chan, Going deeper with densely connected convolutional neural networks for multispectral pansharpening, Remote Sens. 11 (2020) 2608.
[167] K. Zhang, A. Wang, F. Zhang, W. Diao, J. Sun, L. Bruzzone, Spatial and spectral extraction network with adaptive feature fusion for pansharpening, IEEE Trans. Geosci. Remote Sens. 60 (2022) 5410814.
[168] Y. Zhang, C. Liu, M. Sun, Y. Ou, Pan-sharpening using an efficient bidirectional pyramid network, IEEE Trans. Geosci. Remote Sens. 57 (8) (2019) 5549–5563.
[169] S. Luo, S. Zhou, Y. Feng, J. Xie, Pansharpening via unsupervised convolutional neural networks, IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. 13 (2020) 4295–4310.
[170] Y. Yang, W. Tu, S. Yang, H. Lu, W. Wan, L. Gan, Dual-stream convolutional neural network with residual information enhancement for pansharpening, IEEE Trans. Geosci. Remote Sens. 60 (2022) 5402416.

[171] T. Uezato, D. Hong, N. Yokoya, W. He, Guided deep decoder: Unsupervised image pair fusion, in: ECCV, 2020, pp. 87–102.
[172] W. Diao, F. Zhang, H. wang, J. Sun, K. Zhang, Pansharpening via triplet attention network with information interaction, IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. 15 (2022) 3576–3588.
[173] D. Lei, P. Chen, L. Zhang, W. Li, MCANet: A multidimensional channel attention residual neural network for pansharpening, IEEE Trans. Geosci. Remote Sens. 60 (2022) 5411916.
[174] X. Su, J. Li, Z. Hua, Attention-based and staged iterative networks for pansharpening of remote sensing images, IEEE Trans. Geosci. Remote Sens. 60 (2022) 5412521.
[175] S. Chen, H. Qi, K. Nan, Pansharpening via super-resolution iterative residual network with a cross-scale learning strategy, IEEE Trans. Geosci. Remote Sens. 60 (2022) 5407016.
[176] H. Zhang, H. Wang, X. Tian, J. Ma, P2Sharpen: A progressive pansharpening network with deep spectral transformation, Inf. Fusion 91 (2023) 103–122.
[177] D. Wang, P. Zhang, Y. Bai, Y. Li, MetaPan: Unsupervised adaptation with meta-learning for multispectral pansharpening, IEEE Geosci. Remote Sens. Lett. 19 (2022) 5513505.
[178] W. Huang, L. Xiao, Z. Wei, H. Liu, S. Tang, A new pan-sharpening method with deep neural networks, IEEE Geosci. Remote Sens. Lett. 12 (5) (2015) 1037–1041.
[179] Y. Xing, M. Wang, S. Yang, L. Jiao, Pan-sharpening via deep metric learning, ISPRS J. Photogram. Remote Sens. 145 (2018) 165–183.
[180] H. Xu, J. Ma, Z. Shao, H. Zhang, J. Jiang, X. Guo, SDPNet: A deep network for pan-sharpening with enhanced information representation, IEEE Trans. Geosci. Remote Sens. 59 (5) (2021) 4120–4134.
[181] K. Han, Y. Wang, H. Chen, X. Chen, et al., A survey on vision transformer, IEEE Trans. Pattern Anal. Mach. Intell. Early Access (2022) 1–20.
[182] F. Zhang, K. Zhang, J. Sun, Multiscale spatial–spectral interaction transformer for pan-sharpening, Remote Sens. 14 (2022) 1736.
[183] S. Li, Q. Guo, A. Li, Pan-sharpening based on CNN+ pyramid transformer by using no-reference loss, Remote Sens. 14 (2022) 624.
[184] M. Zhou, X. Fu, J. Huang, F. Zhao, A. Liu, R. Wang, Effective pan-sharpening with transformer and invertible neural network, IEEE Trans. Geosci. Remote Sens. 60 (2022) 5406815.
[185] M. Zhou, J. Huang, Y. Fang, X. Fu, A. Liu, Pan-sharpening with customized transformer and invertible neural network, in: AAAI, 2022, pp. 1–9.
[186] X. Meng, F. Shao, Z. Hua, Vision transformer for pansharpening, IEEE Trans. Geosci. Remote Sens. 60 (2022) 5409011.
[187] K. Zhang, Z. Li, F. Zhang, W Wan, Jiande Sun, Pan-sharpening based on transformer with redundancy reduction, IEEE Geosci. Remote Sens. Lett. 19 (2022) 5513205.
[188] H. Zhou, Q. Liu, Y. Wang, Panformer: A transformer based model for pan-sharpening, 2022, pp. 1–6, arxiv.
[189] X. Tian, K. Li, W. Zhou, J. Ma, VP-net: An interpretable deep network for variational pansharpening, IEEE Trans. Geosci. Remote Sens. 60 (2022) 5402716.
[190] S. Xu, J. Zhang, Z. Zhao, K. Sun, J. Liu, C. Zhang, Deep gradient projection networks for pan-sharpening, in: IEEE CVPR, 2021, pp. 1366–1375.
[191] H. Yin, PSCSC-net: A deep coupled convolutional sparse coding network for pansharpening, IEEE Trans. Geosci. Remote Sens. 60 (2022) 5402016.
[192] X. Cao, Y. Chen, W. Cao, Proximal PanNet: A model-based deep network for pansharpening, in: AAAI, 2022, pp. 1–9.
[193] Y. Feng, J. Liu, K. Chen, B. Wang, Z. Zhao, Optimization algorithm unfolding deep networks of detail injection model for pansharpening, IEEE Geosci. Remote Sens. Lett 19 (2022) 5001305.
[194] H. Yin, Panchromatic side sparsity model-based deep unfolding network for pansharpening, IEEE Trans. Geosci. Remote Sens. 60 (2022) 5406715.
[195] F. Palsson, J. Sveinsson, M. Ulfarsson, J. Benediktsson, Quantitative quality evaluation of pansharpened imagery: Consistency versus synthesis, IEEE Trans. Geosci. Remote Sens. 54 (3) (2016) 1247–1259.
[196] J. Han, J. Ding, J. Li, G. Xia, Align deep features for oriented object detection, IEEE Trans. Geosci. Remote Sens. 60 (2022) 5602511.
[197] X. Wang, D. Zhu, G. Li, X. Zhang, Y. He, Proposal-copula-based fusion of spaceborne and airborne SAR images for ship target detection, Inf. Fusion 77 (2022) 247–260.
[198] B. Rasti, P. Ghamisi, Remote sensing image classification using subspace sensor fusion, Inf. Fusion 64 (2020) 121–130.
[199] H. Zhu, W. Ma, L. Li, L. Jiao, S. Yang, B. Hou, A dual-branch attention fusion deep network for multiresolution remote-sensing image classification, Inf. Fusion 58 (2020) 116–131.
[200] Z. Wang, A. Bovik, A universal image quality index, IEEE Signal Process. Lett. 9 (3) (2002) 81–84.
[201] J. Zhou, D.L. Civco, J.A. Silander, A wavelet transform method to merge landsat TM and SPOT panchromatic data? Int. J. Remote Sens. 19 (4) (1998) 743–757.
[202] Z. Wang, A. Bovik, H. Sheikh, E. Simoncelli, Image quality assessment: From error visibility to structural similarity, IEEE Trans. Image Process. 13 (4) (2004) 600–612.
[203] R.H. Yuhas, A.F. Goetz, J.W. Boardman, Discrimination among semi-arid landscape endmembers using the spectral angle mapper (SAM) algorithm, in: Proc. Summaries 3rd Annu. JPL Airborne Geosci. Workshop, 1992, pp. 147–149.
[204] M. Choi, A new intensity-hue-saturation fusion approach to image fusion with a tradeoff parameter, IEEE Trans. Geosci. Remote Sens. 44 (6) (2006) 1672–1682.
[205] C.-I. Chang, Spectral information divergence for hyperspectral image analysis, in: IEEE IGARSS, 1999, pp. 509–511.
[206] L. Alparone, S. Baronti, A. Garzelli, F. Nencini, A global quality measurement of pan-sharpened multispectral imagery, IEEE Geosci. Remote Sens. Lett. 1 (4) (2004) 313–317.
[207] L. Wald, Quality of high resolution synthesized images: Is there a simple criterion? in: Proc. 3rd Conf. Fusion Earth Data, 2000, pp. 99–105.
[208] L. Alparone, B. Aiazzi, S. Baronti, A. Garzelli, F. Nencini, M. Selva, Multispectral and panchromatic data fusion assessment without reference, Photogramm. Eng. Remote Sens. 74 (2) (2008) 193–200.
[209] M. Khan, L. Alparone, J. Chanussot, Pansharpening quality assessment using the modulation transfer functions of instruments, IEEE Trans. Geosci. Remote Sens. 47 (11) (2009) 3880–3891.
[210] G. Palubinskas, Joint quality measure for evaluation of pansharpening accuracy, Remote Sens. 7 (2015) 9292–9310.
[211] G. Vivone, R. Restaino, J. Chanussot, A Bayesian procedure for full-resolution quality assessment of pansharpened products, IEEE Trans. Geosci. Remote Sens. 56 (8) (2018) 4820–4834.
[212] G. Vivone, P. Addesso, J. Chanussot, A combiner-based full resolution quality assessment index for pansharpening, IEEE Geosci. Remote Sens. Lett. 16 (3) (2019) 437–441.
[213] R. Carl, L. Santurri, B. Aiazzi, S. Baronti, Full-scale assessment of pansharpening through polynomial fitting of multiscale measurements, IEEE Trans. Geosci. Remote Sens. 53 (12) (2015) 6344–6355.
[214] B. Zhou, F. Shao, X. Meng, R. Fu, Y. Ho, No-reference quality assessment for pansharpened images via opinion-unaware learning, IEEE Access 7 (2019) 40388–40401.
[215] G. Cheng, P. Zhou, J. Han, Learning rotation invariant convolutional neural networks for object detection in VHR optical remote sensing images, IEEE Trans. Geosci. Remote Sens. 54 (12) (2016) 7405–7415.
[216] M. Wang, Z. Dong, Y. Cheng, D. Li, Optimal segmentation of high-resolution remote sensing image by combining superpixels with the minimum spanning tree, IEEE Trans. Geosci. Remote Sens. 56 (1) (2018) 228–238.
[217] F. Bovolo, L. Bruzzone, L. Capobianco, A. Garzelli, S. Marchesi, F. Nencini, Analysis of the effects of pansharpening in change detection on VHR images, IEEE Geosci. Remote Sens. Lett. 7 (1) (2010) 53–57.
[218] Q. Liu, X. Meng, F. Shao, S. Li, Supervised-unsupervised combined deep convolutional neural networks for high-fidelity pansharpening, Inf. Fusion 89 (2023) 292–304.
[219] Y. Yang, J. Sun, H. Li, Z. Xu, ADMM-CSNet: A deep learning approach for image compressive sensing, IEEE Trans. Pattern Anal. Mach. Intell. 42 (3) (2020) 521–538.
[220] L. Wang, C. Sun, M. Zhang, Y. Fu, H. Huang, DNU: Deep non-local unrolling for computational spectral imaging, in: IEEE CVPR, 2020, pp. 1661–1671.
[221] Y. Xie, Z. Xu, J. Zhang, Z. Wang, et al., Self-supervised learning of graph neural networks: A unified review, IEEE Trans. Pattern Anal. Mach. Intell. Early Access 1–20.
[222] W. Diao, F. Zhang, J. Sun, Y. Xing, K. Zhang, L. Bruzzone, ZeRGAN: Zero-reference GAN for fusion of multispectral and panchromatic images, IEEE Trans. Neural Netw. Learn. Syst. Early Access 1–16.