IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 8, 2022, accepted April 17, 2022, date of publication April 22, 2022, date of current version May 2, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3169698

A Pansharpening Based on the Non-Subsampled Contourlet Transform and Convolutional Autoencoder: Application to QuickBird Imagery

AHMAD AL SMADI[1], SHUYUAN YANG[1], (Senior Member, IEEE), AHED ABUGABAH[2], (Member, IEEE), AHMAD ALI ALZUBI[3], AND LOUIS SANZOGNI[4]

[1]School of Artificial Intelligence, Xidian University, Xi'an 710071, China
[2]College of Technological Innovation, Zayed University, Abu Dhabi, United Arab Emirates
[3]Computer Science Department, Community College, King Saud University, Riyadh 11437, Saudi Arabia
[4]Department of Business Strategy and Innovation, Nathan Campus, Griffith University, Nathan, QLD 4111, Australia

Corresponding author: Ahed Abugabah (ahed.abugabah@zu.ac.ae)

This work was supported by the Researchers Supporting Project, King Saud University, Riyadh, Saudi Arabia, under Grant RSP-2021/395.

ABSTRACT This paper presents a pansharpening technique based on the non-subsampled contourlet transform (NSCT) and convolutional autoencoder (CAE). NSCT is exceptionally proficient at presenting orientation information and capturing the internal geometry of objects. First, it's used to decompose the multispectral (MS) and panchromatic (PAN) images into high-frequency and low-frequency components using the same number of decomposition levels. Second, a CAE network is trained to generate original low-frequency PAN images from their spatially degraded versions. Low-resolution multispectral images are then fed into the trained convolutional autoencoder network to generate estimated high-resolution multispectral images. Third, another CAE network is trained to generate original high-frequency PAN images from their spatially degraded versions. The result of low-pass CAE is fed to the trained high-pass CAE to generate estimated high-resolution multispectral images. The final pan-sharpened image is accomplished by injecting the detailed map of the spectral bands into the corresponding estimated high-resolution multispectral bands. The proposed method is tested on QuickBird datasets and compared with some existing pan-sharpening techniques. Objective and subjective results demonstrate the efficiency of the proposed method.

INDEX TERMS Pansharpening, multispectral image fusion, convolutional autoencoder, NSCT, remote sensing.

I. INTRODUCTION

The earth observatory satellite can provide two types of images: panchromatic images with high spatial resolution but low spectral resolution, and multispectral images with low spatial resolution but high spectral resolution. Most satellites earth observation applications require high spectral and spatial resolution images, such as change detection and mapping [1]. These satellites are unable to produce images with high spatial and spectral resolutions due to physical and technical constraints [2]. Over the last several years, image fusion has witnessed increasing use in applications such as pansharpening [3], image dehazing [4], and medical treatment [5]. Pansharpening is an integration process for spatial detail, extracted from a PAN image into an MS image, which yields a composite image with high spatial and spectral resolution. Many pansharpening techniques have been proposed up to now. The basic technique is the high-pass filtering technique, which consists of injecting high-frequency components from a PAN image into an MS image with high spatial resolution using spatial filtering techniques, such as intensity-hue-saturation (IHS) [6], principal component analysis (PCA) [7], and Brovey transform (BT) [8], etc. Multiresolution-based approaches are another type of pansharpening technology. Multiresolution decomposition provides a simple hierarchical structure for merging images with varying spatial resolutions. As such, Laplacian pyramids, discrete wavelet transform, curvelet transform,

The associate editor coordinating the review of this manuscript and approving it for publication was Gangyi Jiang.

contourlet transform, and non-subsampled contourlet transform [9]–[11] are the most commonly used multi-scale decomposition techniques in pansharpening. However, many studies have reviewed the performance of the aforementioned pansharpening categories [8], [12]–[14], which reflect the advancement of the pansharpening strategies. The NSCT is a completely shift-in-variation, multi-scale, and multi-direction development that decays the images at various goal levels and directions. This manner permits infusing the subtleties of the data given from the PAN image into the MS image [15]. Furthermore, in recent years, deep learning-based pansharpening methods [16], [17], and [18] have been proposed, which can be viewed as a modern era of pansharpening strategies. In [19], the CAE architecture was used as part of a component substitution-based method yielding enhanced fusion outcomes. In [20], the authors introduced a pansharpening method based on a convolutional autoencoder and a guided filter (CAE-GF). Our work centers around the pansharpening method dependent on the non-subsampled contourlet transform and the deep convolutional autoencoder. The NSCT is used to decompose the source images in order to obtain the high-pass subband and low-frequency subband. Then, two deep convolutional autoencoders are trained, one from the low frequency of the original PAN image and its re-sampled version, and the second one from the high frequency of the original PAN image and its re-sampled version. After training with the deep convolutional autoencoder, the upsampled MS image is fed to the network as input.

In comparison to currently existing approaches, our paper makes the following contributions: 1) A new deep learning-based method for pansharpening based on the convolutional autoencoder and non-subsampled contourlet transform is introduced. 2) By learning the nonlinear mapping of high- and low-pass subband panchromatic images and their degraded versions, our CAE network can successfully inject spatial information. 3) The proposed method's pansharpening result is then compared to a representative set of existing strategies using both full-reference and no-reference index measures. Thus, we can find that the proposed method outperforms its counterparts.

The rest of the paper is organized as follows: Section II elaborates on the proposed fusion method. Section III analyzes the comparative experimental results and their discussion. Finally, the conclusion is drawn in section IV.

II. DEVELOPED FUSION METHOD FOR PANSHARPENING

A. BASIC PANSHARPENING FRAMEWORK

The following equation represents the basic pansharpening framework of component substitution based-approaches [21].

$$\mathrm{MS}_i^{\mathrm{HR}} = \mathrm{MS}_i^{\mathrm{LR}} + \mathrm{g}_i \times (\mathrm{PAN} - I) \tag{1}$$

where $\mathbf{MS_i^{HR}}$ and $\mathbf{MS_i^{LR}}$ represent the high-resolution and low-resolution MS image in the $\mathbf{i_{th}}$ band, respectively. $\mathbf{g_i}$ represents the $\mathbf{i_{th}}$ injection gain of the detailed map, which

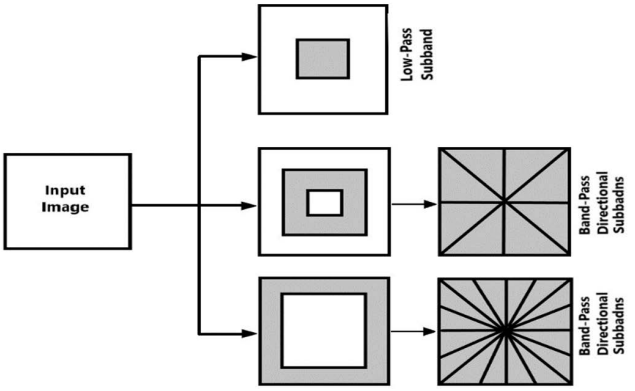


FIGURE 1. The decomposition framework of Non-subsampled Contourlet Transform (NSCT) [23].

is expressed as:

$$\mathrm{g}_i = \frac{cov\left(\mathrm{MS}_i^{\mathrm{LR}}, I\right)}{var(I)} \tag{2}$$

where **cov** indicates the covariance between images, and **var** is the variance of the intensity component image. **I** denotes the intensity component of MS image, which is expressed as:

$$I = \sum\nolimits_{i=1}^{n} w_i \mathrm{MS}_i^{\mathrm{LR}} \tag{3}$$

where $\mathbf{w_i}$ denotes the weights, and **n** is the number of spectral bands. However, the optimal weights can be computed by solving the following optimization problem [22].

$$w_i^* = \arg \left\| \mathrm{PAN} - \sum\nolimits_{i=1}^{n} w_i \mathrm{MS}_i^{\mathrm{LR}} \right\|^2 \tag{4}$$

B. NONSUBSAMPLED CONTOURLET TRANSFORM (NSCT)

An increasing number of pansharpening techniques have been proposed based on the Laplace pyramid, wavelet transform, and contourlet transform multiresolution approaches. Wavelet transforms are often used in pansharpening because of their characteristics, such as multiple resolutions, localizations, critical scans, and limited orientation [8]. However, it is not possible to determine the smoothness along the contour [2]. The contourlet transform seems to overcome this shortcoming [10]. Indeed, the contourlet transform is a multiresolution transform that provides a proficient directional portrayal and considers wavelet properties. The contourlet transform was used for image fusion and pansharpening [2] and [8]. The contourlet transformation without subsampling is a shift-invariant version of the contourlet transform. The decomposition framework of NSCT is shown in Fig. 1. This paper uses the NSCT to decompose the source images to obtain the high-pass and low-frequency subbands.

C. CONVOLUTIONAL AUTOENCODER (CAE)

An autoencoder is a type of unsupervised learning that takes an input picture and attempts to recreate it. The convolutional autoencoder is a type of convolutional neural network that generates output image patches based on the input image

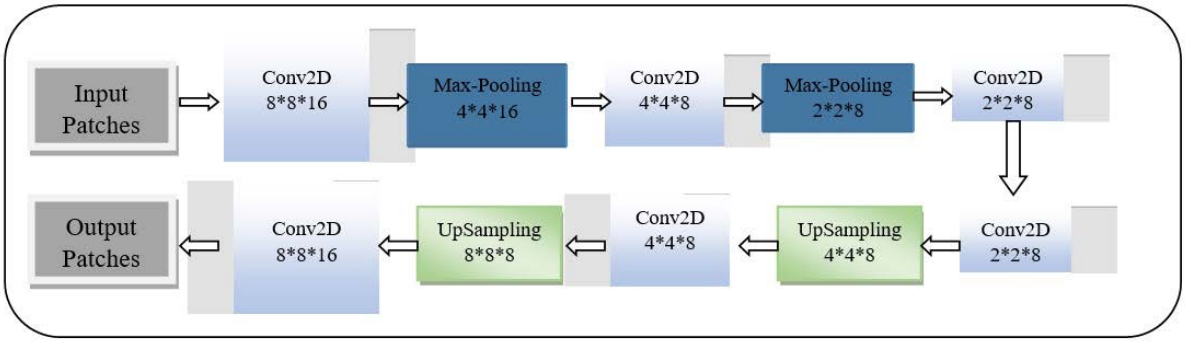


FIGURE 2. The network structure of CAE.

patches. The design of a convolutional autoencoder works in two stages: encoding and decoding. The encoding stage accounts for half of the network and includes convolution and max-pooling layers. The deconvolution and upscaling layers are used in the decoding stage to recreate the input image from its degraded image. The CAE architecture is depicted in Fig. 2.

D. PROPOSED PANSHRPENING METHOD

The developed pansharpening scheme consists of several steps after decomposing the source images using NSCT into high-pass and low-pass subbands. Our method sets the patch size with overlapping pixels, 8×8, 5 as indicated in [19], [20]. The high-pass subband and low-pass subband and their degraded version are partitioned into 8×8 patches with 5 overlapping pixels. The high-pass CAE network is trained using the back-propagation algorithm to learn the relationship between a high-pass subband panchromatic image and its degraded version. The low-pass CAE network is trained to learn the relationship between a low-pass subband panchromatic image and its degraded version. In order to improve the spatial information of the low-resolution MS bands, the MS bands are then fed to the low-pass CAE network. The intensity component is also fed to the low-pass CAE network for enhancement. Then, the result of the low-pass CAE is fed to the high-pass CAE in order to get the estimated MS image. The pseudo-code for the developed pansharpening scheme is summarized in Algorithm 1. The complete procedure of the developed pansharpening scheme is described in Fig. 3.

First, Table 1 demonstrates the abbreviations that would use in Mathematical equations. So, let the original low-pass PAN patches, marked by $\{X_i\}_{i=1}^{N}$, and spatially degraded low-pass PAN patches, marked by $\{x_i\}_{i=1}^{N}$, form the target and the input of the low-pass CAE network, respectively. On each iteration, the output corrections for the low-pass CAE network are calculated as:

$$\left\{\tilde{X}_i\right\}_{i=1}^{N} = Decoder(Encoder\left(\{x_i\}_{i=1}^{N}\right)) \tag{5}$$

TABLE 1. List of abbreviations.

Symbol	DefiNITION
PAN	Panchromatic
MS	Multispectral
X	Low-pass Original PAN
x	Low-pass Degraded PAN
Y	High-pass Original PAN
y	High-pass Degraded PAN
N	Number of Patches

Similarly, the original high-pass PAN patches, marked by $\{Y_i\}_{i=1}^{N}$, and spatially degraded high-pass PAN patches, marked by $\{y_i\}_{i=1}^{N}$, form the target and the input of the low-pass CAE network, respectively. On each iteration, the output corrections for the high-pass CAE network are calculated as:

$$\left\{\tilde{Y}_i\right\}_{i=1}^{N} = Decoder(Encoder\left(\{y_i\}_{i=1}^{N}\right)) \tag{6}$$

Both low-pass CAE and high-pass CAE use the Mean Square Error (MSE) between the original patches and their reconstructed versions to update the weights.

$$Loss\left(\left\{\tilde{X}_i\right\}_{i=1}^{N}, \{X_i\}_{i=1}^{N}\right) = \frac{1}{2}\sum_{i=1}^{N}\left\|\tilde{X}_i - X_i\right\|_2^2 \tag{7}$$

$$Loss\left(\left\{\tilde{Y}_i\right\}_{i=1}^{N}, \{Y_i\}_{i=1}^{N}\right) = \frac{1}{2}\sum_{i=1}^{N}\left\|\tilde{Y}_i - Y_i\right\|_2^2 \tag{8}$$

In the testing part, the up-sampled MS patches are fed to the low-pass CAE; afterward, they are fed to the high-pass CAE. In addition, the intensity components patched are fed to the low-pass CAE. Because of the similitude in the spectral attributes of PAN and MS images, the prepared networks are relied upon to work on the spatial/spectral data of the upsampled MS bands. Indeed, not just the estimated version saves the spectral information of MS bands, but it also conveys more spatial data in correlation with the information of input patches. Afterward, tiling the assessed

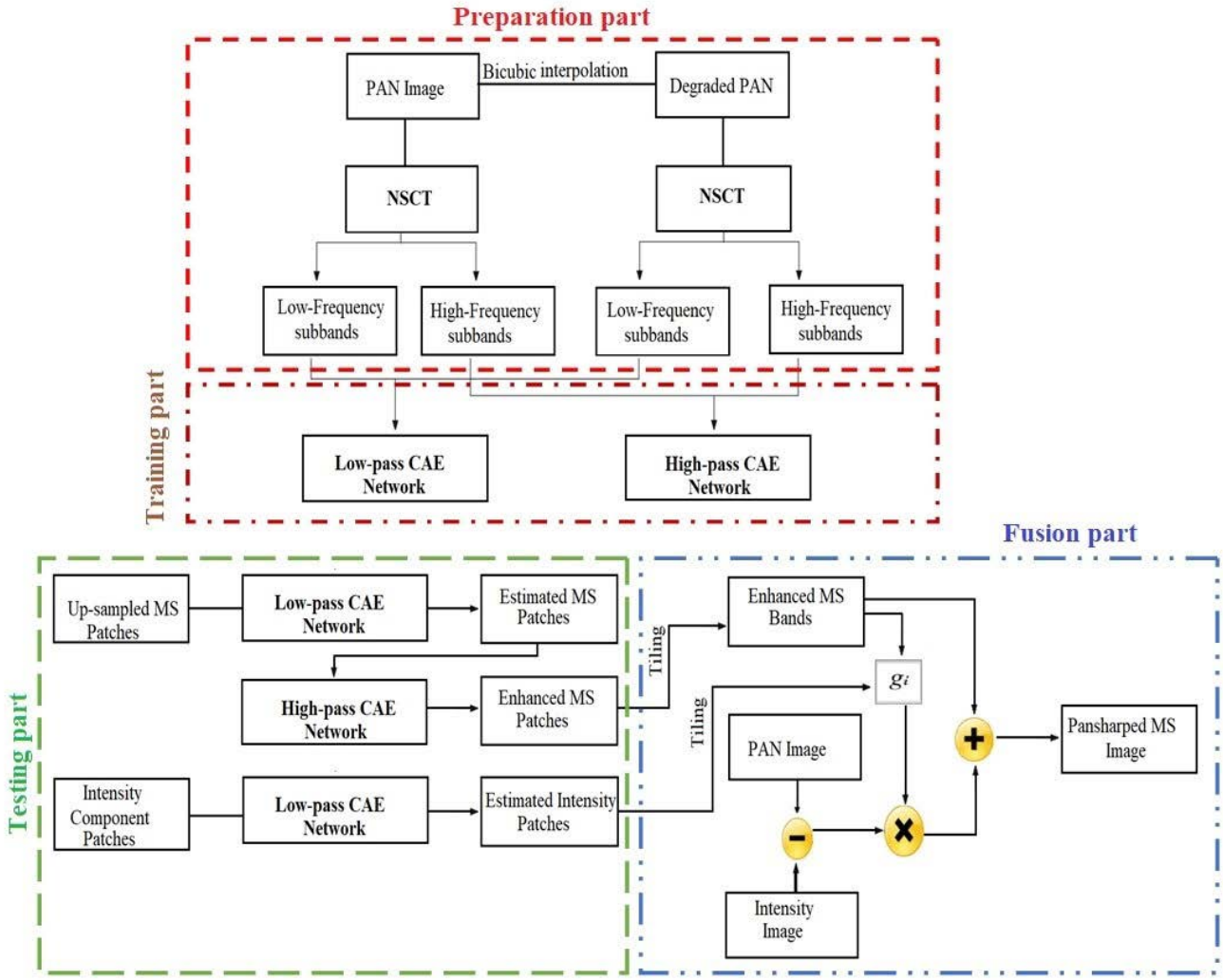


FIGURE 3. Block diagram of pansharpening proposed method.

high-resolution MS bands, the pansharpening is done through the following equation:

$$\mathrm{HRMS}_i = \hat{M}S_i + \mathrm{g}_i \times (\mathrm{PAN} - I) \tag{9}$$

where HRMS_i represents the i_{th} estimated high resolution $\hat{M}S_i$ band obtained from the trained both CAE networks. Here, the injection gains g_i are used to obtain the refined detail map between the estimated $\hat{M}S_i$ and the estimated intensity image $\hat{I}_i$. It's carried out as follows

$$\mathrm{g}_i = \frac{cov\left(\hat{M}S_i, \hat{I}\right)}{var(\hat{I})} \tag{10}$$

III. EXPERIMENTAL RESULTS AND ANALYSIS

In this section, experiments have been conducted on QuickBird satellite images to evaluate the performance of the proposed technique. Table 2 shows twelve different image fusion methods for comparison, including BT-H, BDSD-PC, GSA, PRACS, MF, FE-HPM, SR-D, PWMBF, AIHS, PNN, CAE, and CAE-GF. We used the available toolbox [14]. In this toolbox, several quality/distortion indices were also reported. Therefore, eight image quality assessments were utilized to assess the quality of the pansharpening image.

A. QUALITY ASSESSMENTS

1) The correlation coefficient (CC) is expressed as (11), as shown at the bottom of the next page,

where *Ref* and *Fus* are the reference image and the fused image with size $M \times N$, respectively. $\bar{R}ef$ and $\bar{F}us$ are the means of the reference image and fused image, respectively. The CC demonstrates the comparability of spectral features between the pan-sharped and the reference images [33]. The two images correspond when CC is near 1.

2) Universal image quality indexes (UIQI)

The UIQI measures the similarity between the fused image and the reference image; in other words, it measures the structure distortion degree between two images [34]. The quality of the fused image is better if the value of UIQI is

Algorithm 1 Pansharpening Scheme Based on NSCT and CAE

Input: PAN image, MS image
Output: Fused image $HRMS_i$
Step 1: Preparation part
1. Degrading the PAN image using Bi-cubic interpolation.
2. Appling the NSCT algorithm to decompose the PAN image and its degraded version to the high-pass subband and low-frequency subband.
Step 2: Training part
1. The low-pass CAE network is trained using (Eq. (5)).
2. The high-pass CAE network is trained using (Eq. (6)).
Step 3: Testing part
1. Feeding the low-resolution MS bands to the trained low-pass CAE network.
2. Then, low-pass CAE results are fed to the high-pass CAE to get the estimated MS image.
3. Feeding the intensity component to the low-pass CAE network for enhancement.
4. Tiling the enhanced MS image bands and the enhanced intensity component.
Step 3: Fusion part
1. Using (Eq. (9)) to get the fused image band $HRMS_i$.

TABLE 2. Brief description of the comparison methods used to compare the performance of fusion results.

Method	Description
BT-H	Brovey data fusion with haze correction [24]
BDSD-PC	Band-Dependent Spatial-Detail [25]
GSA	Gram-Schmidt Adaptive [26]
PRACS	Partial Replacement Adaptive CS [27]
MF	Morphological Pyramid Decomposition using Half-Gradient [28]
FE-HPM	The High Pass Modulation injection model and the Estimated Filter via Deconvolution [29]
SR-D	Sparse Representation of injected Details [30]
PWMBF	Model-based Fusion using PCA and Wavelet [31]
AIHS	Adaptive Intensity Hue Saturation [32]
PNN	Pansharpening by Convolutional Neural Networks [12]
CAE	Convolutional Autoencoder-Based Multispectral Image Fusion [19]
CAE-GF	Convolutional Autoencoder and multi-scale Guided Filter [20]

closer to 1. It is calculated by

$$UIQI = \frac{\sigma_{FusRef} \times (2\mu_{Fus}\mu_{Ref}) \times (2\sigma_{Fus}\sigma_{Ref})}{(\sigma_{Fus}\sigma_{Ref}) \times (\mu_{Fus}^2 + \mu_{Ref}^2) \times (\sigma_{Fus}^2 + \sigma_{Ref}^2)} \quad (12)$$

where σ_{FR} denotes the covariance between the pan-sharped and the reference images, it collects three different operators, loss of correlation, luminance distortion, and contrast distortion. The term σ_{Fus} and σ_{Ref} are the standard deviations of the pan-sharped image and the reference image, respectively.

3) An image quality index $Q4$ index

$Q4$ is a comprehensive measure to the 4-bands image of the Q index [35], which is formulated as

$$Q4 = E(Q4_{B\times B}) \quad (13)$$

where E(.) represents the quaternion acquired via averaging the pixel quaternion within a B×B block.

4) The root means square error (RMSE)

The RMSE gives the dissimilarity between the pan-sharped and the reference images depending on the changes in the pixels' value [36]. Note that a smaller RMSE value shows that the pan-sharped image is closer to the reference image. It is expressed by

$$RMSE = \frac{1}{MN}\sqrt{\sum_{i,j=1}^{MN}(Fus_{(i,j)} - Ref_{(i,j)})^2} \quad (14)$$

5) The Relative Average Spectral Error (RASE)

The RASE gives the average performance of image fusion methods for each spectral band [37]. It is calculated by

$$RASE = \frac{100}{\mu}\sqrt{\frac{1}{N}\sum_{i=1}^{N} RMSE} \quad (15)$$

where μ is the average radiation value of N bands.

6) The erreur relative global adimensionnelle de synthèse (ERGAS)

The ERGAS represents the difference between the pan-sharped and the reference images. In other words, it gives a global picture of the quality of a fused image [38]. It is defined as

$$ERGAS = 100\frac{h}{l}\sqrt{\frac{1}{N}\sum_{i=1}^{N}(\frac{RMSE_i}{MEAN(R_i)})^2} \quad (16)$$

where h, l are the spatial resolution of the PAN image and the MS image, respectively, N marks the number of bands of the fused image, the MEAN(i) denotes the mean value of i_{th} band of the reference MS images and the RMSE(i) represents the RMSE between the i_{th} band of the fused image and the i_{th} band of the MS reference images. The smaller ERGAS value indicates that a small spectral distortion in the pan-sharped image.

7) The spectral angle mapper (SAM)

The SAM displays the spectral distortion between the pan-sharped and the reference images by considering the absolute angles between the two vectors in consideration [39], which is constructed from the pan-sharped and reference images. The ideal fused image should be zero. The SAM is expressed as

$$SAM = arccos(\frac{u_{Ref}, u_{Fus}}{\|u_{Ref}\|_2 \cdot \|u_{Fus}\|_2}) \quad (17)$$

$$CC = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}[Fus(i,j) - \bar{F}us][Ref(i,j) - \bar{R}ef]}{\sqrt{\sum_{i=1}^{M}\sum_{j=1}^{N}[Fus(i,j) - \bar{F}us]^2 \sum_{i=1}^{M}\sum_{j=1}^{N}[Ref(i,j) - \bar{R}ef]^2}} \quad (11)$$

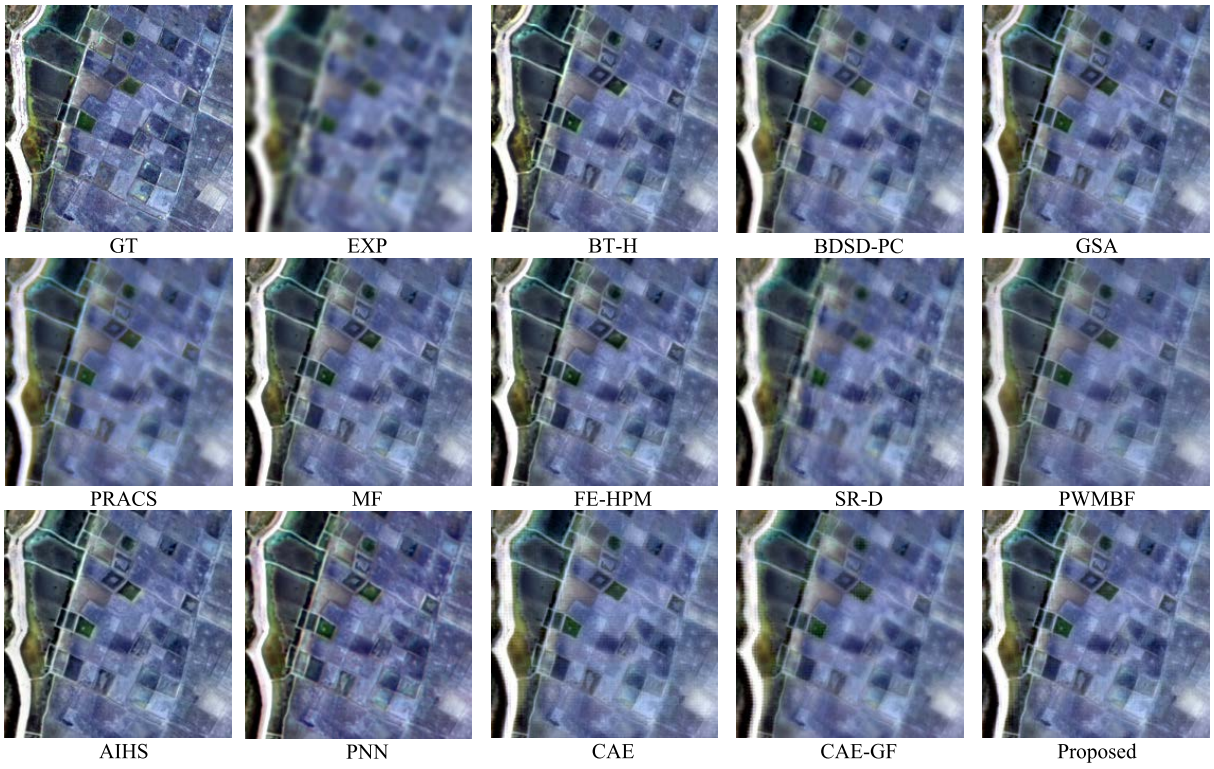


FIGURE 4. Pansharpening outcomes for a sample degraded QuickBird-1 dataset.

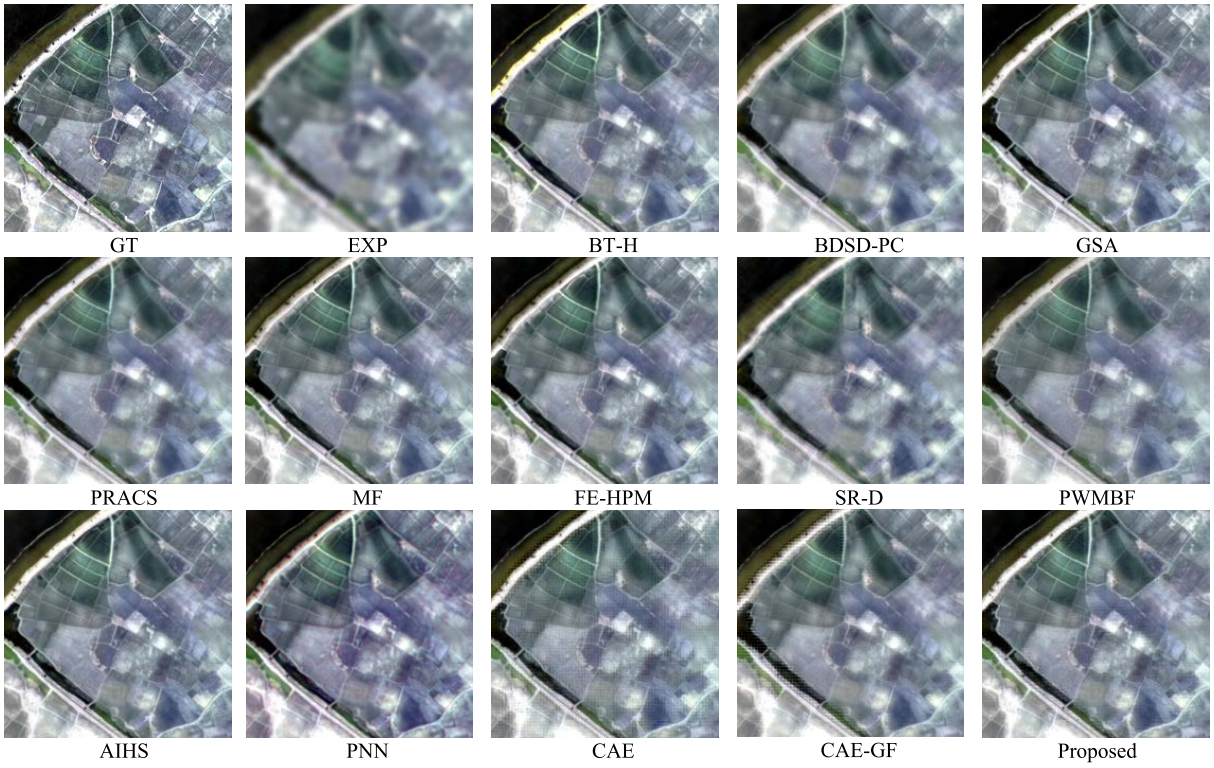


FIGURE 5. Pansharpening outcomes for a sample degraded QuickBird-2 dataset.

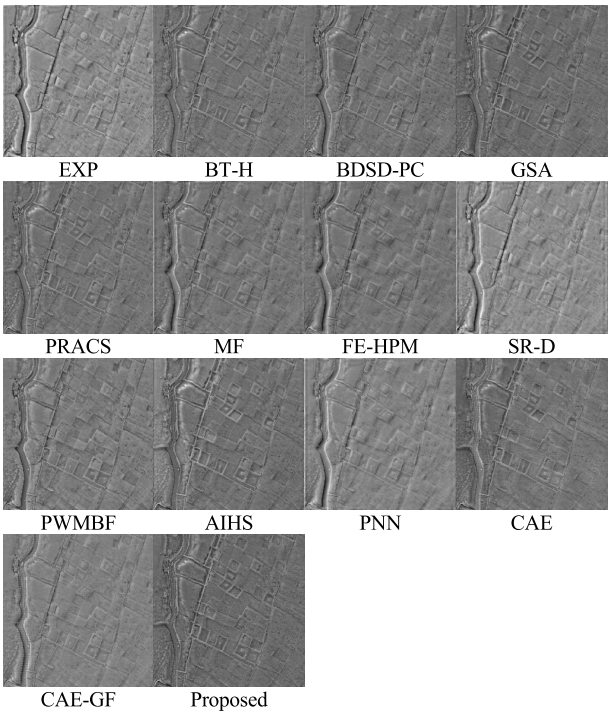


FIGURE 6. Error maps between the Pansharpening outcomes and the reference image for the degraded QuickBird-1 dataset.

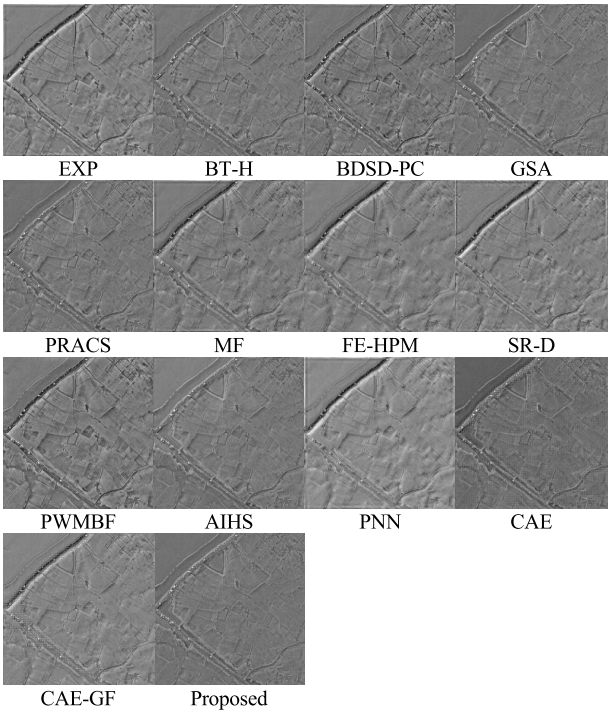


FIGURE 7. Error maps between the Pansharpening outcomes and the reference image for the degraded QuickBird-2 dataset.

8) The spatial correlation coefficient (sCC)

sCC value reflects the spatial quality of the fused results. The fused image is compared to the ground truth image using the correlation coefficient between the spatial details extracted from the two images. The greater the value, the more spatial information the fused image has from the PAN image [14], [40].

TABLE 3. Comparison of quantitative assessments for full reference quickbird-1 dataset, the best values are bolded.

	CC	UIQI	RMSE	RASE	SAM	ERGAS	Q4	sCC
GT	1.000	1.000	0.000	0.000	0.000	0.000	1.000	1.000
EXP	0.824	0.852	25.374	16.694	3.521	4.175	0.674	0.700
BT-H	0.927	0.956	14.549	9.573	3.524	2.544	0.875	0.797
BDSD-PC	0.918	0.947	15.875	10.445	3.487	2.730	0.865	0.791
GSA	0.932	0.956	14.274	9.391	3.570	2.480	0.882	0.07
PRACS	0.917	0.950	15.360	10.106	3.684	2.708	0.857	0.18
MF	0.914	0.942	16.803	11.055	3.478	2.858	0.858	0.796
FE-HPM	0.902	0.932	18.257	12.012	3.486	3.084	0.845	0.795
SR-D	0.838	0.872	24.392	16.048	3.558	4.023	0.735	0.751
PWMBF	0.924	0.942	15.967	10.505	3.473	2.743	0.839	0.777
AIHS	0.923	0.949	15.289	10.059	3.574	2.646	0.870	0.810
PNN	0.899	0.922	21.857	14.965	4.868	3.787	0.818	0.889
CAE	0.945	0.957	23.955	16.402	3.350	4.148	0.897	0.915
CAE-GF	0.946	0.964	14.096	9.651	**3.326**	2.536	0.887	0.913
Proposed	**0.951**	**0.969**	**13.273**	**9.088**	3.374	**2.378**	**0.898**	**0.919**

For real dataset experiments (Full Resolution), we used the quality without reference (QNR) index to define the comprehensive quality of the fusion result without needing the reference image [41]. The QNR measure consists of a spectral distortion index D_λ and a spatial distortion index D_s, respectively. The following equations show the estimation of QNR, D_S, and D_λ.

$$QNR = (1 - D_\lambda)\,(1 - D_s) \tag{18}$$

where $(1-D_\lambda)$ represents the spectral quality and $(1-D_s)$ represents the spatial quality (19) and (20), as shown at the bottom of page 9.

The ideal value for the QNR index is 1, while D_λ and D_s are 0.

B. FUSION RESULTS

Two pairs of reduced resolution QuickBird satellite datasets were examined and evaluated according to the Wald's protocol; the sizes of the MS image and the PAN image are 64×64 and 256×256, respectively. Moreover, two pairs of No-reference QuickBird datasets were examined, and the sizes of the MS image and the PAN image are 256×256 and 1024×1024, respectively. The pansharpening outcomes for reduced resolution QuickBird datasets are provided in Figs. 4–5, and the objective quality indexes are reported in Table 3–4. By visually comparing the pansharpening outcomes, the first remark is that our method is very close to the ground truth image compared with others, which can better preserve the colors and the spatial information of the PAN image are also better injected. By objectively comparing, the proposed method outperforms relatively in comparison with SOTA methods. Due to the difficulties in visual comparison, we provided the error maps where the distinction between the pansharpening image and the ground truth can be inspected to see the consistency and edges.

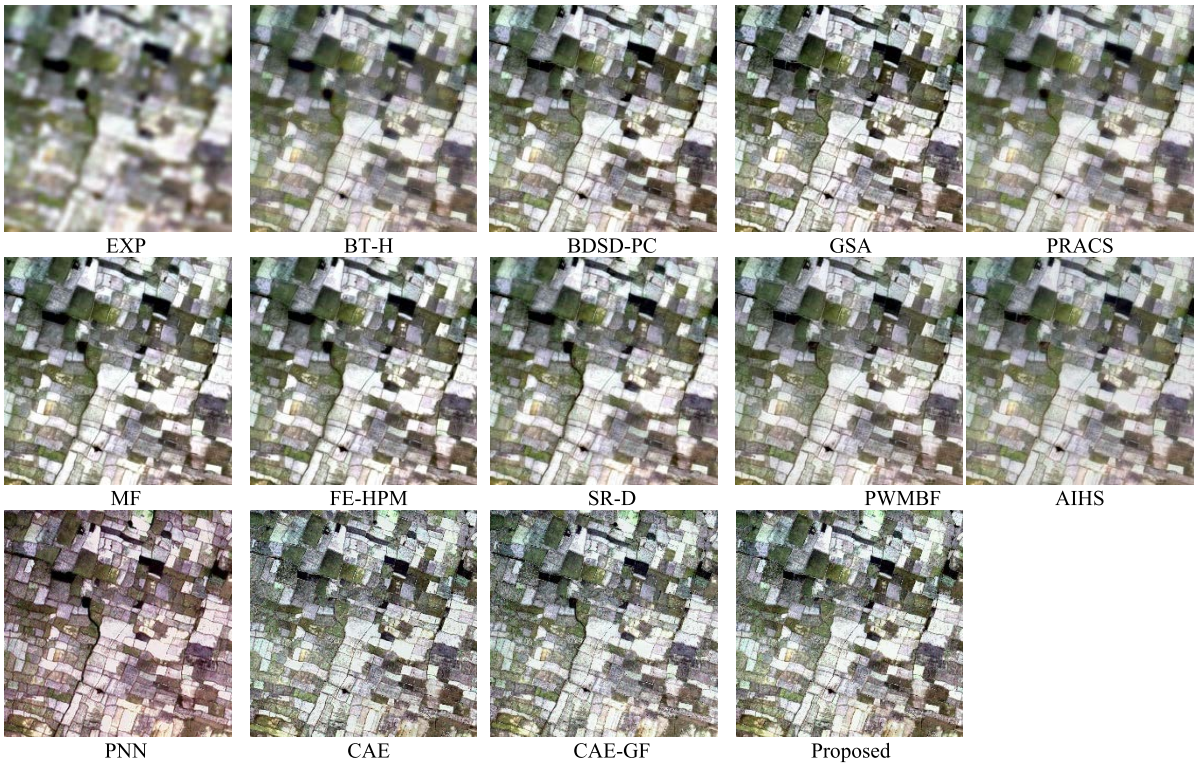


FIGURE 8. Pansharpening outcomes for a sample real QuickBird-1 dataset.

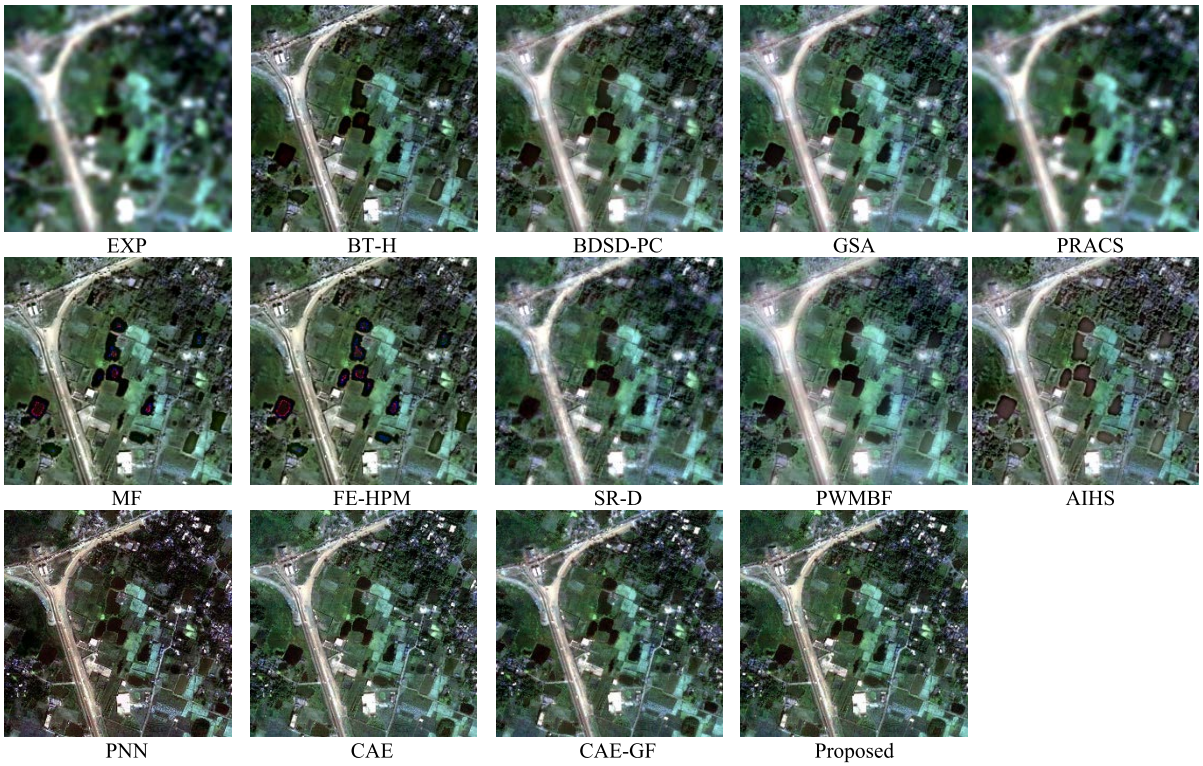


FIGURE 9. Pansharpening outcomes for a sample real QuickBird-2 dataset.

TABLE 4. Comparison of quantitative assessments for full reference quickbird-2 dataset, the best values are bolded.

	CC	UIQI	RMSE	RASE	SAM	ERGAS	Q4	sCC
GT	1.000	1.000	0.000	0.000	0.000	0.000	1.000	
EXP	0.869	0.883	23.571	15.47	3.001	3.898	0.703	0.788
BT-H	0.948	0.962	13.998	9.188	3.113	2.378	0.900	0.922
BDSD-PC	0.943	0.953	15.217	9.988	2.956	2.551	0.873	0.905
GSA	0.958	0.968	12.988	8.524	2.974	2.187	0.911	0.933
PRACS	0.954	0.966	13.241	8.691	3.021	2.244	0.897	0.926
MF	0.939	0.951	15.754	10.34	2.977	2.637	0.881	0.922
FE-HPM	0.925	0.939	17.641	11.58	2.981	2.946	0.865	0.917
SR-D	0.883	0.903	22.089	14.49	2.964	3.668	0.782	0.859
PWMBF	0.944	0.952	15.165	9.953	2.908	2.551	0.868	0.913
AIHS	0.955	0.963	13.348	8.761	2.862	2.254	0.911	0.933
PNN	0.928	0.938	19.765	13.04	3.895	3.293	0.841	0.909
CAE	0.961	0.966	18.999	12.54	2.871	3.164	0.896	0.925
CAE-GF	0.964	0.972	13.286	8.766	2.769	2.227	0.906	0.916
Proposed	**0.966**	**0.973**	**12.783**	**8.434**	**2.757**	**2.147**	**0.912**	**0.943**

TABLE 5. Comparison of quantitative assessments for no-reference quickbird datasets, the best values are bolded.

	Full Resolution QuickBird-1			Full Resolution QuickBird-2		
	D_λ	D_s	QNR	D_λ	D_s	QNR
EXP	0.000	0.1145	0.8855	0.0000	0.1144	0.8856
BT-H	0.015	0.0430	0.9426	0.1447	0.1763	0.7045
BDSD-PC	0.014	0.0522	0.9344	0.0878	0.1158	0.8066
GSA	0.012	0.0374	0.9509	0.1301	0.1471	0.7419
PRACS	0.009	0.0398	0.9518	0.0524	0.0550	0.8954
MF	0.012	0.0250	0.9636	0.2092	0.2241	0.6136
FE-HPM	0.011	0.0274	0.9615	0.2151	0.2087	0.6211
SR-D	0.018	0.0165	0.9659	0.0804	0.0922	0.8349
PWMBF	0.012	0.0556	0.9326	0.1346	0.1354	0.7482
AIHS	0.010	0.0617	0.9287	0.1165	0.1186	0.7787
PNN	0.231	0.1957	0.8976	0.0737	0.0427	0.8868
CAE	0.001	0.16027	0.9620	0.0942	0.0885	0.8256
CAE-GF	**0.0008**	0.1289	0.9699	0.0591	0.0579	0.8864
Proposed	0.0064	**0.0165**	**0.9772**	**0.0274**	**0.0413**	**0.9324**

It can be observed in Figs. 6–7; our strategy accomplished superior pansharpening in most image areas among other techniques.

For full resolution QuickBird datasets, the pansharpening outcomes are provided in Figs. 8–9, the objective indexes assessment of the pansharpening results are listed in Table 5. Visually, the proposed method preserves the PAN image's edges and spatial details better than others. The comparison of quality indexes demonstrates that our method achieves better results than other methods.

C. AVERAGE PROCESSING TIME

The proposed method's average consumption time compared to that of the other techniques on the aforementioned datasets is reported in Table 6. For deep learning methods, we only considered the testing phase processing time. On a MacBook with an Intel Core i5-5257U CPU (2.7 GHz/8G), we conducted the experiments using MATLAB R2016a and Jupyter Notebook. From Table 6, it can be seen that the deep learning methods require the most time for fusion. Additionally, the proposed technique requires more time to acquire the fusion results than other methods due to the deep CAE and mechanism of NSCT (decomposition and reverse construction). Despite that, the proposed method outperforms others in terms of fusion results.

TABLE 6. Average time comparison of the different pansharpening methods.

Method	Time (sec)
BT-H	0.15
BDSD-PC	0.31
GSA	0.18
PRACS	0.19
MF	0.24
FE-HPM	0.47
SR-D	1.70
PWMBF	0.65
AIHS	0.17
PNN	11.97
CAE	11.59
CAE-GF	12.57
Proposed	14.98

IV. CONCLUSION

In this paper, we have proposed a pansharpening method based on CAE and NSCT. The major contributions of this paper are twofold. First, the low-pass CAE network was trained to learn the relationship between a low-pass subband panchromatic image and its degraded version. Meanwhile, the spatial information of the low-resolution MS bands improved. Second, the high-pass CAE network was trained to learn the relationship between a high-pass subband panchromatic image and its degraded version. The proposed method has fused the images while preserving spectral information of MS images and carried more spatial information. The experimental results have shown that the proposed method outperformed others in both aspects, subjective and objective. Furthermore, the CAE and NSCT could be extended to the existing state-of-the-art-based image fusion methods.

$$D_\lambda = \sqrt{\frac{1}{N(N-1)} \sum_{l=1}^{N} \sum_{r=1, r\neq l}^{N} |Q(MS_l, MS_r) - Q(Fus_l - Fus_r)|} \tag{19}$$

$$D_s = \sqrt{\frac{1}{N} \sum_{l=1}^{N} |Q(Fus_l, PAN_{LR}) - Q(MS_l, PAN_{HR})|} \tag{20}$$

REFERENCES

[1] A. A. Smadi, S. Yang, A. Mehmood, A. Abugabah, M. Wang, and M. Bashir, "Smart pansharpening approach using kernel-based image filtering," *IET Image Process.*, vol. 15, no. 11, pp. 2629–2642, Sep. 2021.

[2] V. P. Shah, N. H. Younan, and R. L. King, "An efficient pan-sharpening method via a combined adaptive PCA approach and contourlets," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 5, pp. 1323–1335, May 2008.

[3] F. D. Javan, F. Samadzadegan, S. Mehravar, A. Toosi, R. Khatami, and A. Stein, "A review of image fusion techniques for pan-sharpening of high-resolution satellite imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 171, pp. 101–117, Jan. 2021.

[4] Z. Zhu, H. Wei, G. Hu, Y. Li, G. Qi, and N. Mazur, "A novel fast single image dehazing algorithm based on artificial multiexposure image fusion," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–23, 2020.

[5] Z. Zhu, M. Zheng, G. Qi, D. Wang, and Y. Xiang, "A phase congruency and local Laplacian energy based multi-modality medical image fusion method in NSCT domain," *IEEE Access*, vol. 7, pp. 20811–20824, 2019.

[6] M. C. El-Mezouar, N. Taleb, K. Kpalma, and J. Ronsin, "A new intensity-hue-saturation fusion technique with color distortion reduction for IKONOS imagery," in *Proc. IEEE 5th Int. Conf. Sci. Electron., Technol. Inf. Telecommun. (SETIT)*, Mar. 2009, p. 408.

[7] P. Chavez, S. C. Sides, and J. A. Anderson, "Comparison of three different methods to merge multiresolution and multispectral data landsat TM and SPOT panchromatic," *Photogramm. Eng. Remote Sens.*, vol. 57, no. 3, pp. 295–303, Mar. 1991.

[8] I. Amro, J. Mateos, M. Vega, R. Molina, and A. K. Katsaggelos, "A survey of classical methods and new trends in pansharpening of multispectral images," *EURASIP J. Adv. Signal Process.*, vol. 2011, no. 79, pp. 1–22, Sep. 2011.

[9] I. Amro and J. Mateos, "Multispectral image pansharpening based on the contourlet transform," in *Information Optics and Photonics*, New Yark, NY, USA: Springer, 2010, pp. 247–261.

[10] M. N. Do and M. Vetterli, "The contourlet transform: An efficient directional multiresolution image representation," *IEEE Trans. Image Process.*, vol. 14, no. 12, pp. 2091–2106, Dec. 2005.

[11] M. C. El-Mezouar, K. Kpalma, N. Taleb, and J. Ronsin, "A pan-sharpening based on the non-subsampled contourlet transform: Application to worldview-2 imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 5, pp. 1806–1815, May 2014.

[12] C. Chen, "Twenty-five years of pansharpening: A critical review and new developments," in *Signal and Image Processing for Remote Sensing*, Boca Raton, FL, USA: CRC, 2012, pp. 554–601, doi: 10.1201/B11656-31.

[13] H. Ghassemian, "A review of remote sensing image fusion methods," *Inf. Fusion*, vol. 32, pp. 75–89, Nov. 2016.

[14] G. Vivone, M. D. Mura, A. Garzelli, and F. Pacifici, "A benchmarking protocol for pansharpening: Dataset, preprocessing, and quality assessment," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 6102–6118, 2021.

[15] S. Mazari, M. C. El Mezouar, and K. Belloulata, "A pan-sharpening method based on NSCT and formulated as compressive sensing problem," *Geocarto Int.*, vol. 31, no. 10, pp. 1142–1157, Nov. 2016.

[16] W. Huang, L. Xiao, Z. Wei, H. Liu, and S. Tang, "A new pan-sharpening method with deep neural networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 5, pp. 1037–1041, May 2015.

[17] G. Masi, D. Cozzolino, L. Verdoliva, and G. Scarpa, "Pansharpening by convolutional neural networks," *Remote Sens.*, vol. 8, no. 7, p. 594, Jul. 2016.

[18] Q. Yuan, Y. Wei, X. Meng, H. Shen, and L. Zhang, "A multi-scale and multi depth convolutional neural network for remote sensing imagery pan-sharpening," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 3, pp. 978–989, Feb. 2018.

[19] A. Azarang, H. E. Manoochehri, and N. Kehtarnavaz, "Convolutional autoencoder-based multispectral image fusion," *IEEE Access*, vol. 7, pp. 35673–35683, 2019.

[20] A. A. L. Smadi, S. Yang, Z. Kai, A. Mehmood, M. Wang, and A. Alsanabani, "Pansharpening based on convolutional autoencoder and multi-scale guided filter," *EURASIP J. Image Video Process.*, vol. 2021, no. 1, pp. 1–20, Dec. 2021.

[21] G. Vivone, L. Alparone, J. Chanussot, M. D. Mura, A. Garzelli, G. A. Licciardi, R. Restaino, and L. Wald, "A critical comparison among pan-sharpening algorithms," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2565–2586, Dec. 2014.

[22] A. Garzelli, "Pansharpening of multispectral images based on nonlocal parameter optimization," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 4, pp. 2096–2107, Apr. 2015.

[23] A. L. da Cunha, J. Zhou, and M. N. Do, "The nonsubsampled contourlet transform: Theory, design, and applications," *IEEE Trans. Image Process.*, vol. 15, no. 10, pp. 3089–3101, Oct. 2006.

[24] S. Lolli, L. Alparone, A. Garzelli, and G. Vivone, "Haze correction for contrast-based multispectral pansharpening," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 12, pp. 2255–2259, Dec. 2017.

[25] G. Vivone, "Robust band-dependent spatial-detail approaches for panchromatic sharpening," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 6421–6433, Sep. 2019.

[26] B. Aiazzi, S. Baronti, and M. Selva, "Improving component substitution pansharpening through multivariate regression of MS+Pan data," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3230–3239, Oct. 2007.

[27] J. Choi, K. Yu, and Y. Kim, "A new adaptive component-substitution-based satellite image fusion by using partial replacement," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 1, pp. 295–309, Jan. 2011.

[28] R. Restaino, G. Vivone, M. D. Mura, and J. Chanussot, "Fusion of multispectral and panchromatic images based on morphological operators," *IEEE Trans. Image Process.*, vol. 25, no. 6, pp. 2882–2895, Jun. 2016.

[29] G. Vivone, M. Simões, M. D. Mura, R. Restaino, J. M. B. Dias, G. A. Licciardi, and J. Chanussot, "Pansharpening based on semi blind deconvolution," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 4, pp. 1997–2010, Sep. 2014.

[30] M. R. Vicinanza, R. Restaino, G. Vivone, M. D. Mura, and J. Chanussot, "A pansharpening method based on the sparse representation of injected details," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 1, pp. 180–184, Jan. 2015.

[31] F. Palsson, J. R. Sveinsson, M. O. Ulfarsson, and J. A. Benediktsson, "Model-based fusion of multi- and hyperspectral images using PCA and wavelets," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2652–2663, May 2015.

[32] S. Rahmani, M. Strait, D. Merkurjev, M. Moeller, and T. Wittman, "An adaptive IHS pan-sharpening method," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 4, pp. 746–750, Oct. 2010.

[33] X. X. Zhu and R. Bamler, "A sparse image fusion algorithm with application to pan-sharpening," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 5, pp. 2827–2836, May 2013.

[34] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Process. Lett.*, vol. 9, no. 3, pp. 81–84, Mar. 2002.

[35] A. Garzelli and F. Nencini, "PAN-sharpening of very high resolution multispectral images using genetic algorithms," *Int. J. Remote Sens.*, vol. 27, no. 15, pp. 3273–3292, Aug. 2006.

[36] Y. Yang, S. Tong, S. Huang, and P. Lin, "Multifocus image fusion based on NSCT and focused area detection," *IEEE Sensors J.*, vol. 15, no. 5, pp. 2824–2838, May 2015.

[37] T. Ranchin and L. Wald, "Fusion of high spatial and spectral resolution images: The ARSIS concept and its implementation," *Photogramm. Eng. Remote Sens.*, vol. 66, no. 1, pp. 49–61, Jan. 2000.

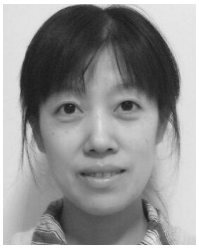
[38] L. Wald, "Quality of high-resolution synthesised images: Is there a simple criterion?" in *Proc. 3rd Conf. Fusion Earth Data, Merging Point Meas., Raster Maps Remotely Sensed Images (SEE/URISCA)*, 2000, pp. 99–103.

[39] R. H. Yuhas, A. F. Goetz, and 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*, vol. 1, 1992, pp. 147–149.

[40] X. Otazu, M. González-Audícana, O. Fors, and J. Núñez, "Introduction of sensor spectral response into image fusion methods. Application to wavelet-based methods," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 10, pp. 2376–2385, Oct. 2005.

[41] L. Alparone, B. Aiazzi, S. Baronti, A. Garzelli, F. Nencini, and M. Selva, "Multispectral and panchromatic data fusion assessment without reference," *Photogramm. Eng. Remote Sens.*, vol. 74, no. 2, pp. 193–200, Feb. 2008.

AHMAD AL SMADI received the B.S. degree in computer engineering from Yarmouk University, Irbid, Jordan, in 2014, and the M.S. degree in intelligent information processing from Xidian University, Xi'an, China, in 2018, where he is currently pursuing the Ph.D. degree in computer science and technology with the School of Artificial Intelligence. His research interests include AI, machine learning, computers networks, and information systems.



SHUYUAN YANG (Senior Member, IEEE) received the B.S. degree in electrical engineering and the M.S. and Ph.D. degrees in circuit and system from Xidian University, Xi'an, China, in 2000, 2003, and 2005, respectively. She is currently a Professor with the Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education, Xidian University. Her research interests include intelligent signal processing, machine learning, and image processing.

AHED ABUGABAH (Member, IEEE) received the degrees in information systems from Zayed University. He is currently a Professor of information systems. He also works at the College of Technological Innovation, Zayed University. Before joining Zayed University, he worked in higher education in Australia. His research interests include information systems, enterprise applications and development, machine learning and data mining in healthcare, information systems, and RFID in healthcare.

AHMAD ALI ALZUBI received the Ph.D. degree in computer networks engineering from the National Technical University of Ukraine (NTUU), in 1999. He is currently a Full Professor with King Saud University (KSU). His current research interests include computer networks, cloud computing, big data, and data extracting. He is also a Certified Member of the Board of Assessors of Quality Management System (BOA-QMS), KSU. He also worked as a Consultant and a member of the Saudi National Team for measuring E-Government in Saudi Arabia, for a period of three years.

LOUIS SANZOGNI is currently a Senior Lecturer with the Department of Business Strategy and Innovation, Nathan Campus, Griffith University, Australia. His research interests include diffusion of innovation and information systems adoption, including business and organizational intelligence technologies and diffusion of innovation modeling, information access, knowledge acquisition, and the technology that supports information access and knowledge acquisition. With more than 30 years' industry and academic experience, he is widely published in the area of business and organizational intelligence technologies and also teaches business processes, enterprise mobility, and operations management.