

Pansharpening with multi-CAE: impact of patch size and overlapping pixels on spectral and spatial distortion

Ahmad Al Smadi¹, Ahed Abugabah², Mutasem Khilail Alsmadi³, Ala Alsanabani⁴, Atif Mehmood⁵,
Ahmad Mohammad Al-Smadi⁶

¹Department of Data Science and Artificial Intelligence, Zarqa University, Zarqa, Jordan

²College of Technological Innovation, Zayed University, Abu Dhabi, UAE

³Department of Management of Information Systems, College of Applied Studies and Community Service, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia

⁴School of Artificial Intelligence, Xidian University, Xian, China

⁵Department of Computer Science, Zhejiang Normal University, Jinhua, China

⁶Department of Applied Science, Ajloun University College, Al-Balqa Applied University, Ajloun, Jordan

Article Info

Article history:

Received Nov 10, 2023

Revised Jan 12, 2024

Accepted Feb 2, 2024

Keywords:

Adaptive intensity-hue-saturation

Convolutional-autoencoder

Multispectral

Pansharpening

Remote sensing

ABSTRACT

A novel technique utilizing a convolutional autoencoder (CAE) is introduced with the aim of enhancing the spatial resolution of multispectral (MS) images while concurrently mitigating spectral distortion. First, an original panchromatic (PAN) image is constructed from its spatially degraded version. Then, the relationship between the original PAN image and its degraded version is utilized to reconstruct the high-resolution MS image; in addition, an intensity component of MS image, which is obtained using an adaptive intensity-hue-saturation (AIHS), is reconstructed by utilizing the aforementioned relationship. Two types of remote sensing datasets are adopted, and the effect of the patch size with the overlapping pixel on spectral and spatial distortion is considered. After training CAE, the low-resolution MS image and its intensity component are given to the trained network as input to obtain the MS image and intensity component with better details. Eventually, the fused image is obtained by using a component substitution (CS) framework. Experimental findings corroborate that the proposed method yields superior outcomes compared with several existing approaches, demonstrating advantages in both objective metrics and visual fidelity.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Ahmad Al Smadi

Department of Data Science and Artificial Intelligence, Zarqa University

Zarqa 13100, Jordan

Email: aalsmadi@zu.edu.jo

1. INTRODUCTION

In remote sensing images, panchromatic (PAN) images have higher spatial resolution than multispectral (MS) images. MS images have multiple spectral bands, offering high spectral resolution, and their spatial resolution can vary depending on the sensor used, with some MS sensors providing moderate to high spatial resolution. Due to the limitation of technology in earth surveillance satellites [1], a high spatial resolution multispectral (HRMS) image cannot be given directly by satellites such as QuickBird and GeoEye. However, the high spectral resolution simplifies identifying the objects while the high spatial resolution facilitates locating them. Pansharpening aims to the acquisition of a HRMS image [2]–[4]. The pansharpening process is one branch of image fusion, and image processing is used in many fields, such as image segmentation [5] and classification [6], image reconstruction, and medical applications [7], [8]. A wide variety of pansharpening studies have been introduced in the literature, which can be grouped into three groups [9]: i) component

substitution (CS) [10], [11]; ii) multiresolution analysis (MRA) [12], [13]; and iii) sparse representation (SR) [14], [15] based methods. Both CS and MRA-based methods share a common framework in which an extracted detail map is incorporated into a resampled MS image, albeit with differing methodologies for detail map extraction [16]. In CS-based methods, the extracted detail map is obtained by computing the disparity between the PAN image and the linear combination of upsampled MS bands.

The CS based-methods such as intensity hue saturation (IHS) [17], principle component analysis (PCA) [18], gram-schmidt adaptive (GSA) [19], and brovey [20] often suffer from remarkable spatial and spectral distortions in the fused images. The MRA based-methods such as indusion [21], wavelet transform [22], and generalized Laplacian pyramid (MTF-GLP) approach [23] leverages the detailed map by computing the variance between the PAN image and its low-resolution counterpart via PAN image decomposition. Nevertheless, these methods are susceptible to spatial distortion issues. The SR-based method generates HRMS images by some dictionaries and sparse optimization algorithms, and it has shown effectiveness in pansharpening. However, implementation can be challenging as constructing the dictionary requires a substantial amount of raw HMS images [14]. Recently, the convolutional neural network (CNN) architecture was used for pansharpening in [24]–[26]. Azarang and Ghassemian [24] used the denoising autoencoder (DAE) design to enhance fusion results. Rao *et al.* [25] used the deep residual network (DRN) architecture to achieve a pansharpening task. More recently, convolutional autoencoder (CAE) [26] was considered as part of a CS-based method where HRMS images were computed by training the CAE from the PAN image and its degraded version. The researchers [27], [28], used a deep CAE for pansharpening in which they fixed 8×8 patches with five overlapping pixels. Further development of deep learning-based pansharpening has been introduced in [29], [30].

In this study, an advanced approach to pansharpening is proposed, leveraging the CAE architecture as the cornerstone of the method. The CAE is trained by utilizing the PAN image and its degraded version. Then, the low-resolution MS is fed to the trained CAE in order to enhance the MS image. The intensity component is also reconstructed using the same trained CAE. The proposed method is applied to both degraded and real datasets, following which it is systematically compared against a selection of existing methods. The contributions of our method are summarized as: i) a novel deep learning pansharpening method is presented in which the CAE architecture is trained to enhance the low-resolution MS image and its intensity component and ii) the effect of the patch size with the overlapping pixel on spectral and spatial distortion is studied. It is noted that the patch size with the overlapping pixel can be different for different datasets. Thus, the best outcomes for each one are listed.

The paper is organized as: related work for the proposed method is presented in section 2. The proposed method is detailed in section 3. The experimental results are presented and discussed in section 4. Finally, the conclusion of this work is presented in section 5.

2. RELATED WORK

Image fusion has played a significant role in recent years by enhancing the quality and interpretability of multi-source images. The fusion of MS and PAN images has emerged as a crucial challenge in remote sensing applications [29], such as land cover categorization, object recognition, and change detection. A range of solutions, from traditional methodologies to cutting-edge deep learning strategies, has been proposed to address this problem [30]. In this review, we focus on CAEs and discuss the existing literature on MS and PAN image fusion.

Saxena *et al.* [31] proposed a CNN-based model for pansharpening, aiming to derive the fine grained components of the image. Their approach utilized a CNN to extract the PAN detail image, which was then suitable for the MRA-based pansharpening scheme. This process helped minimize spatial and spectral distortions. The authors conducted qualitative and quantitative analyses using GeoEye-1 and ikonos satellite images. Huang *et al.* [32] presented a deep neural network (DNN)-based pansharpening technique that effectively modeled complex relationships between variables through multiple levels of nonlinearity. Li *et al.* [33] developed a method for pansharpening called detail injection-based CAE (DiCAE). Their approach involved injecting MS details as PAN details and integrating them using an injection gain.

Previous studies [26]–[28] utilized the CAE model to explore the spatial and spectral relationships between PAN and MS images. Azarang *et al.* [26] trained a CAE on a set of PAN and MS images to generate high-resolution MS images. In another study [27], a combination of the non-subsampled contourlet transform (NCCT) and CAE was employed for pansharpening. The NCCT decomposed PAN and MS images into space and frequency bands, while the CAE learned how these bands could be combined. Fusion involved integrating detail maps from the CAE with low-resolution images from the MS. In contrast, a different approach was taken in [28], where CAE and the multi-scale guided filter (MGGF) were used together. The MGGF was utilized to smooth out the CAE's generated detail maps, and the low-resolution MS images were added to the smoothed-out detail maps. However, none of these methods extensively investigated the influence of patch size and the number

of overlapping pixels on spectral and spatial warping. In our proposed method, we employ a learned CAE to enhance low-resolution MS image patches and intensity image patches. Additionally, we investigate the impact of patch size and the number of overlapping pixels on spectral and spatial warping, which previous methods have not thoroughly explored.

2.1. Adaptive IHS

The conventional IHS combination strategy is easy to implement, fast, and has high spatial goals. Except it can be applied to just three bands. The IHS formulation is given by:

$$I = \sum_{i=1}^n w_i \widetilde{MS}_i \quad (1)$$

where w_i denotes the weight of the i th MS band, n represents the number of MS bands, \widetilde{MS} , I indicate the up-sampled MS image and its intensity component, respectively. Therefore, Rahmani *et al.* [10] proposed an adaptive computation method in order to reduce the spectral distortion, which can be (2):

$$\underset{w_i}{\operatorname{argmin}} \|P - \sum_{i=1}^n w_i \widetilde{MS}_i\|^2 \quad (2)$$

where P represents the PAN image.

2.2. CAEs

CAEs are considered unsupervised learning of convolutional filters. Therefore, any input would be used to extract features once a CAE has been learned [34]. Then the extracted features are utilized to perform many tasks, such as image classification, and image enhancement [35]. The CAE involves two stages named the encoding stage and the decoding stage. The encoding stage of constructing the k th feature map is given by:

$$f^k = \sigma(x * w^k + b^k) \quad (3)$$

where f^k represents the feature maps of the input x , b^k represents the bias, $*$ denotes the 2D convolution, w^k indicates k th convolution filters, and σ denotes an activation function. To reconstruct the input by decoding stage, which can be expressed as (4):

$$\tilde{x} = \sigma(f^k * \tilde{w}^k + c^k) \quad (4)$$

where \tilde{x} is the reconstructed input, c^k represents the bias, and \tilde{w}^k indicates k th convolution filters. The mean square error (MSE) loss function can be used in order to update the weights during the training of CAE.

$$\mathcal{L}(\tilde{x}, x) = \frac{1}{2} \|\tilde{x} - x\|_2^2 \quad (5)$$

3. METHOD

The flowchart of our method is illustrated in Figure 1, and the fusion steps are outlined as:

- The intensity image is first obtained using AIHS by (2).
- Then, the PAN image and its degraded version are used for training our CAE, the architecture of the CAE-based method is illustrated in Figure 2, where the CAE network is utilized to learn the relationship between the output original PAN patches $p \times p$ with overlapping pixel ω and its degraded version patches as input. The effect of a patch size with an overlapping pixel will be introduced in the experiments section.
- After training our CAE, the trained CAE network is used to enhance the low-resolution MS image patches and intensity image patches. The low-resolution MS and intensity images are partitioned into patches $p \times p$ are then fed to the trained CAE in order to reconstruct the high-resolution MS image \widetilde{MS}_i and enhanced intensity image \hat{I} , by utilizing the relationship between the original PAN image and its degraded version.
- The MS image bands have different forms/levels of detail; therefore, to derive more proper injection gains, the injection gains are computed using both reconstructed intensity and MS images. Next, these injection gains are injected into corresponding MS images.
- Finally, the high-resolution MS image is constructed by (7).

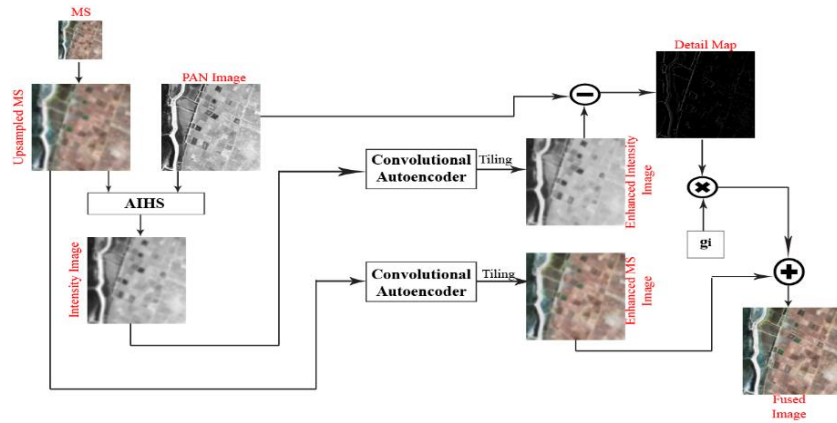


Figure 1. Flowchart of image fusion based on multi CAE

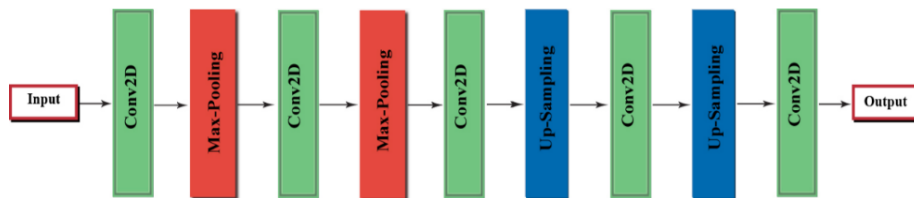


Figure 2. The architecture of the CAE

Mathematically, the aforementioned steps can be expressed as (6):

$$\{\tilde{P}_i\}_{i=1}^p = Decoding \left(Encoding \left(\{\hat{P}_i\}_{i=1}^p \right) \right) \tag{6}$$

where $\{\tilde{P}_i\}_{i=1}^p$ indicates the output patches of CAE and $\{\hat{P}_i\}_{i=1}^p$ represents the degraded version PAN image patches of the original PAN image patches $\{P_i\}_{i=1}^p$. The MSE between $\{P_i\}_{i=1}^p$ and $\{\tilde{P}_i\}_{i=1}^p$ is used to update the weights at each iteration. Thus, the back-propagation algorithm is utilized for training. The fusion process is expressed as (7):

$$HRMS_i = \widehat{MS}_i + g_i(P - \hat{I}) \tag{7}$$

where $HRMS_i$, \widehat{MS}_i denote the fused image and the reconstructed MS image in the i_{th} band, respectively. P , \hat{I} denote the PAN image and enhanced intensity image, respectively, and g_i represents the i_{th} injection gain of the detailed map that is expressed as (8):

$$g_i = \frac{covariance(\widehat{MS}_i, \hat{I})}{variance(\hat{I})} \tag{8}$$

4. RESULTS AND DISCUSSION

4.1. Datasets

In our experiment, we utilize datasets acquired by QuickBird and GeoEye satellites to evaluate our method. The experiments are carried out on four datasets, including two full-reference datasets with PAN and MS image sizes of 256×256 and 64×64 , respectively, as well as two no-reference datasets with PAN and MS image sizes of 512×512 and 128×128 , respectively.

4.2. Quality assessment of fusion results

For the degraded evaluation that is based on Wald’s protocol, six well-known indexes are utilized to assess the spatial and spectral qualities of the outcomes [36]. These indexes include, the correlation coefficient (CC), spectral angle mapper (SAM), root mean square error (RMSE), erreur relative globale adimensionnelle de synthse (ERGAS), universal image quality index (UIQI), and the relative average spectral error (RASE). For the no-reference data sets (real data), the objective evaluation is based on the spectral distortion index D_λ .

the spatial distortion index D_s , and the quality with no reference QNR. Further, our method is compared with some other existing methods, such as band-dependent spatial detail (BDSD), adaptive CS using partial replacement (PRACS), and fractional-order differentiation (FDIF).

4.3. Effect of patch size with overlapping pixels on datasets

The effect of patch size $p \times p$ with overlapping pixels ω on the values of quality indexes is analyzed in this part. We consider patch sizes 4, 8, and 16 for our experiment. Overlapping pixels should be less than the patch size. Tables 1 and 2 exhibit the numerical results of quality indexes concerning the changing patch size with overlapping pixels. Overall, the patch size with overlapping pixels 4×4 and 3 has performed better than the others for QuickBird. The patch size with overlapping pixels 8×8 and 7 have served better than the others for GeoEye. It is worth noting that the patch size with overlapping pixels has a direct correlation with the size and quantity of image features. Thus, the value of patch size with overlapping pixels is set according to the best-obtained result for each kind of dataset.

Table 1. Effect of patch size with overlapping pixels on QuickBird

$p \times p$	ω	CC	UIQI	RMSE	RASE	SAM	ERGAS
4×4	1	0.96596	0.97393	12.575	8.2965	2.8796	2.1191
4×4	3	0.96856	0.976	12.082	7.9717	2.6283	2.0379
8×8	1	0.96124	0.96993	13.516	8.9177	3.2526	2.2783
8×8	3	0.95783	0.96738	14.198	9.3677	3.2352	2.4078
8×8	5	0.96459	0.9729	12.895	8.5082	2.8095	2.183
8×8	7	0.96822	0.97569	12.102	7.9845	2.8277	2.0376
16×16	3	0.95377	0.96516	14.544	9.5956	3.4056	2.4638
16×16	5	0.95783	0.96589	14.449	9.5329	3.6864	2.4395
16×16	7	0.95028	0.96156	15.414	10.17	3.6956	2.6171
16×16	9	0.96302	0.97167	13.062	8.6183	3.0555	2.2067
16×16	12	0.96255	0.97077	13.369	8.8209	3.0335	2.2643
16×16	15	0.96843	0.97504	12.259	8.0883	2.8633	2.0655

Table 2. Effect of patch size with overlapping pixels on GeoEye

$p \times p$	ω	CC	UIQI	RMSE	RASE	SAM	ERGAS
4×4	1	0.95306	0.95331	17.992	16.519	6.1715	4.0873
4×4	3	0.94559	0.95216	18.747	17.212	6.2608	4.2798
8×8	1	0.94188	0.94682	19.855	18.23	7.1187	4.4835
8×8	3	0.95216	0.9582	17.779	16.323	6.1369	4.0488
8×8	5	0.95541	0.96106	17.289	15.874	5.8811	3.9407
8×8	7	0.95674	0.962	16.849	15.469	5.5989	3.8974
16×16	3	0.93113	0.93661	21.636	19.864	8.0888	4.8854
16×16	5	0.93306	0.93776	21.096	19.369	7.9247	4.7848
16×16	7	0.92062	0.90749	24.265	22.278	8.972	5.4696
16×16	9	0.92885	0.91784	22.95	21.071	8.3703	5.1767
16×16	12	0.95365	0.95948	17.755	16.301	6.087	4.048
16×16	15	0.95029	0.95702	17.724	16.273	6.1695	4.1015

4.4. Fusion results

Figures 3(a)-(l) depicts the fusion results for the degraded QuickBird-1 datasets, alongside the objective comparison in Table 3. Our proposed technique (Figure 3(l)) outperformed alternatives (Figure 3(c) to Figure 3(k)) by preserving the reference image's (Figure 3(a)) hues and incorporating superior spatial details from the PAN image (Figure 3(b)). This observation is consistent with the objective comparison presented in Table 3, where our method achieved the highest quality indexes throughout.

Figures 4(a)-(l) presents the fusion outcomes for the degraded QuickBird-2 datasets, accompanied by an objective comparison in Table 4. The reference-MS image (Figure 4(a)) was used as the true-image, in order to compare it with the fusion result (Figure 4(b)). While FDIF and CAE methods (Figures 4(j) and (k)) achieved relatively favorable results compared to others (Figures 4(c)-(i)), our proposed method (Figure 4(l)) produced a higher resolution image and achieved superior outcomes according to quality measures in Table 4.

Figures 5(a)-(l) showcase the fusion outcomes of the real QuickBird dataset. Figures 5(a) and 5(b) display the up-sampled MS image and PAN image, respectively. While most methods enhance spectral resolution, some introduce spatial artifacts (e.g., indusion (Figure 5(h)), GSA (Figure 5(e)), and PRACS (Figure 5(g))). The BDSD method (Figure 5(f)) suffers from particularly significant spatial and spectral distortions. Although the CAE method (Figure 5(k)) achieves good results, the proposed method (Figure 5(l)) excels in preserving both spatial and spectral details. This superiority is further confirmed by the objective comparison metrics in Table 5.

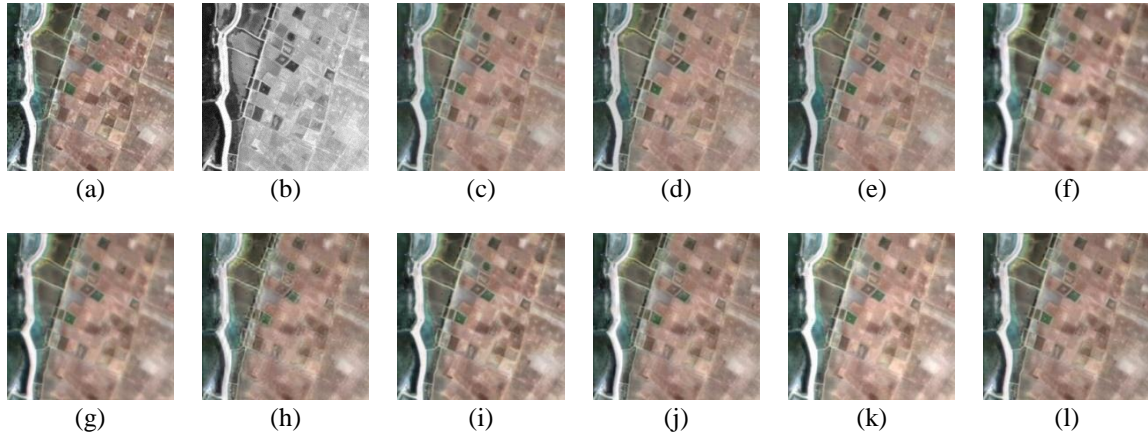


Figure 3. Fusion outcomes of the degraded QickBird-1 dataset; (a) reference image, (b) PAN image, (c) PCA, (d) AIHS, (e) GSA, (f) BSDS, (g) PRACS, (h) indusion, (i) MTF-GLP, (j) FDIF, (k) CAE, and (l) proposed

Table 3. Objective comparison of proposed method vs. other methods for degraded QuickBird-1 dataset

Method	ERGAS	SAM	RASE	RMSE	UIQI	CC
Proposed	2.4239	3.55	9.1642	13.385	0.96865	0.95104
AIHS [10]	2.7063	3.7497	10.301	15.045	0.95778	0.93518
PCA [18]	2.9025	3.6063	11.234	16.408	0.94871	0.93743
GSA [19]	3.1952	3.6982	12.476	18.221	0.93744	0.92289
Indusion [21]	3.1483	3.5524	12.302	17.967	0.93911	0.91806
MTF-GLP [23]	3.2217	3.6399	12.534	18.306	0.93984	0.91144
CAE [26]	4.1479	3.3501	16.402	23.955	0.95735	0.94467
BSDS [37]	2.9444	3.7728	11.253	16.435	0.95204	0.92229
PRACS [38]	2.7466	3.8317	10.265	14.992	0.95991	0.93038
FDIF [39]	2.5878	3.5933	9.8041	14.319	0.96123	0.94303

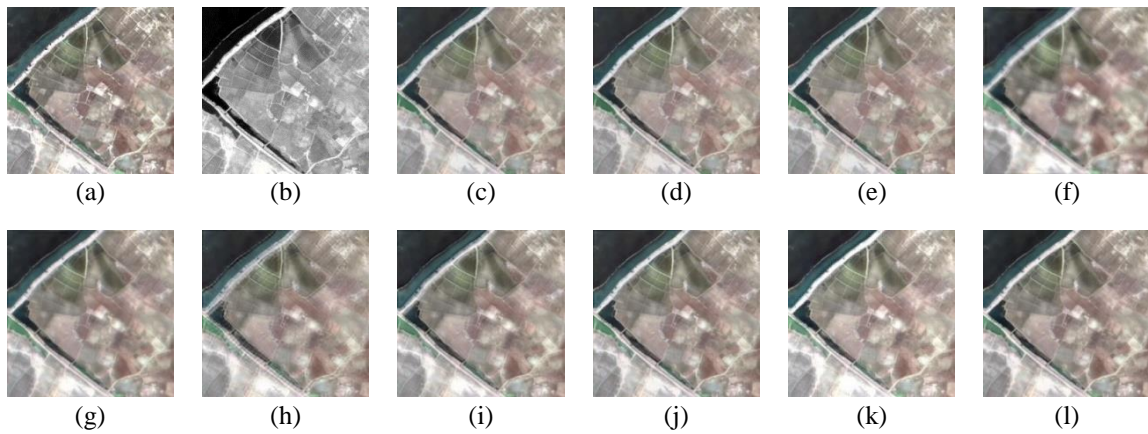


Figure 4. Fusion outcomes of the degraded QickBird-2 dataset: (a) reference image, (b) PAN image, (c) PCA, (d) AIHS, (e) GSA, (f) BSDS, (g) PRACS, (h) indusion, (i) MTF-GLP, (j) FDIF, (k) CAE, and (l) proposed

Table 4. Objective comparison of proposed method vs other methods for degraded QuickBird-2 dataset

Method	ERGAS	SAM	RASE	RMSE	UIQI	CC
Proposed	2.0379	2.6283	7.9717	12.082	0.976	0.96856
AIHS [10]	2.2727	2.8699	8.8843	13.465	0.96837	0.96171
PCA [18]	2.557	2.9418	10.066	15.256	0.95906	0.95507
GSA [19]	2.6728	2.9789	10.538	15.972	0.95527	0.95062
Indusion [21]	2.9657	2.9083	11.714	17.754	0.94457	0.93623
MTF-GLP [23]	2.9018	2.9938	11.42	17.309	0.94967	0.93661
CAE [26]	2.2271	2.7686	8.7657	13.286	0.97157	0.96366
BSDS [37]	2.7548	3.0727	10.828	16.412	0.9547	0.9425
PRACS [38]	2.2505	3.0473	8.7685	13.29	0.97082	0.96067
FDIF [39]	2.2135	2.9248	8.6698	13.14	0.96982	0.96515

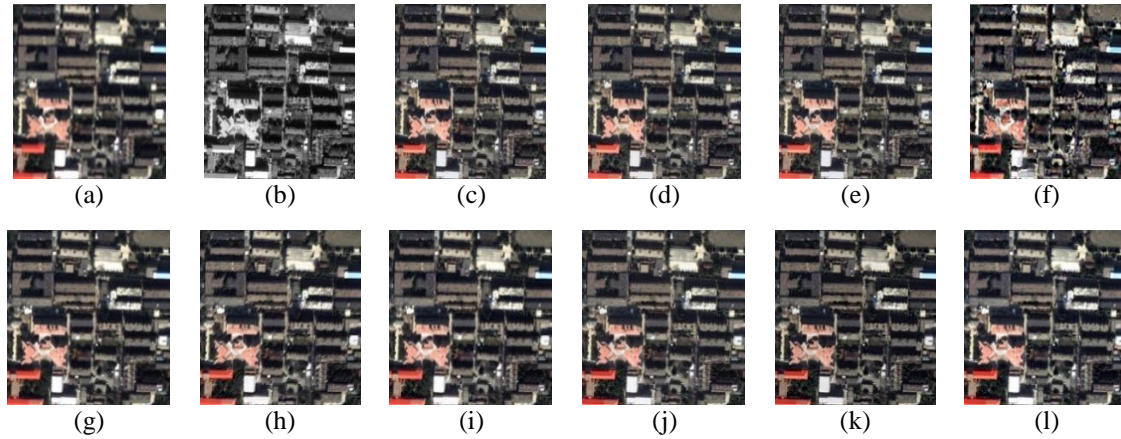


Figure 5. Fusion outcomes of the real QuickBird dataset; (a) upsampled MS, (b) PAN image, (c) PCA, (d) AIHS, (e) GSA, (f) BDSD, (g) PRACS, (h) indusion, (i) MTF-GLP, (j) FDIF, (k) CAE, and (l) proposed

Table 5. Objective comparison of proposed method vs other methods for real QuickBird dataset

Method	D_s	D_d	QNR
Proposed	0.032745	0.014321	0.98529
AIHS [10]	0.052273	0.027263	0.97038
PCA [18]	0.080653	0.034207	0.97399
GSA [19]	0.080168	0.033785	0.97399
Indusion [21]	0.048852	0.028099	0.9752
MTF-GLP [23]	0.045615	0.035905	0.97937
CAE [26]	0.057	0.04463	0.98467
BDSD [37]	0.10022	0.1104	0.9471
PRACS [38]	0.047739	0.022022	0.98394
FDIF [39]	0.59769	0.025623	0.90076

Figures 6(a)-(l) illustrates the fusion outcomes of the real GeoEye dataset, showcasing notable enhancements in spatial details across all methods (Figures 6(c)-(l)). Figures 6(a) and (b) show the up-sampled MS image and PAN image, respectively. Nevertheless, the PCA and AIHS methods exhibit spectral distortion, as shown in Figures 6(c) and 6(d). The BDSD fusion method (Figure 6(f)) shows significant distortions. The fusion result of the proposed method (Figure 6(l)) demonstrates superior performance compared to the other methods (Figures 6(c)-(k)). For an objective evaluation, the proposed method exhibits superior performance in terms of quality indices, as detailed in Table 6. Additionally, Table 7 presents a comparative analysis of the computational time required for the proposed method vs other techniques. It is noteworthy that the proposed method necessitates more time for pansharpening image acquisition compared to others, attributable to the utilization of deep multi-CAE.

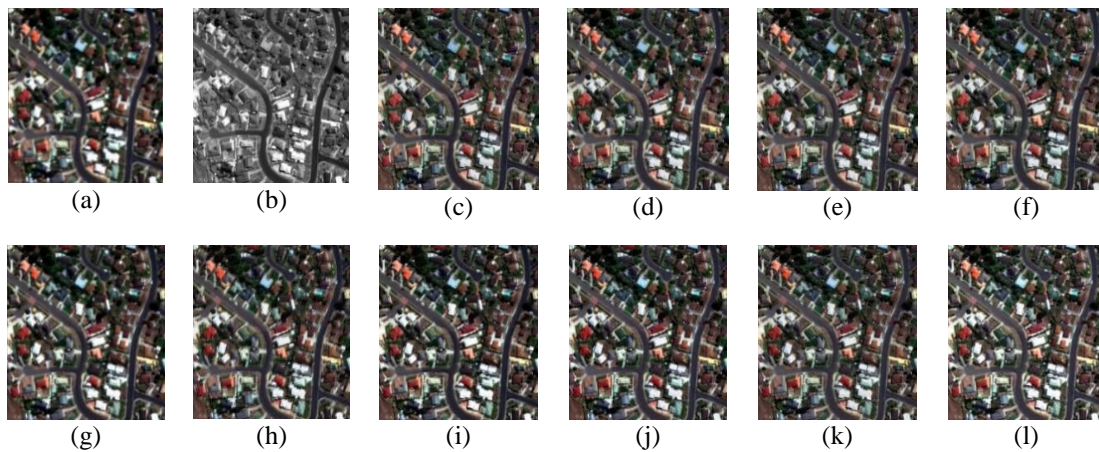


Figure 6. Fusion outcomes of the real GeoEye dataset: (a) upsampled MS, (b) PAN image, (c) PCA, (d) AIHS, (e) GSA, (f) BDSD, (g) PRACS, (h) indusion, (i) MTF-GLP, (j) FDIF, (k) CAE, and (l) proposed

Table 6. Objective comparison of proposed method vs other methods for real GeoEye dataset

Method	D_s	D_λ	QNR
Proposed	0.0602	0.013907	0.9775
AIHS [10]	0.1259	0.0449	0.9604
PCA [18]	0.1597	0.0592	0.9512
GSA [19]	0.1342	0.0411	0.9588
Indusion [21]	0.081795	0.065	0.9607
MTF-GLP [23]	0.1274	0.1158	0.9483
CAE [26]	0.11723	0.054282	0.9737
BDSB [37]	0.1487	0.1496	0.9341
PRACS [38]	0.0833	0.0257	0.9756
FDIF [39]	0.7555	0.048	0.7213

Table 7. Comparison of processing times for pansharpening methods

Method	Time (sec)
Proposed	13.28
AIHS [10]	0.17
PCA [18]	0.06
GSA [19]	0.57
Indusion [21]	0.07
MTF-GLP [23]	0.22
CAE [26]	11.59
BDSB [37]	0.20
PRACS [38]	0.19
FDIF [39]	1.90

5. CONCLUSION

Amidst the ongoing evolution of remote sensing image fusion techniques, pansharpening technology has emerged as a prevalent methodology employed across remote sensing satellite platforms. Since the remote sensing satellites collect a PAN image with high spatial resolution and MS image with a high spectral resolution, therefore, in this paper, we developed a pansharpening method using a CS framework and multi-CAE. The high-resolution MS and intensity images were estimated after obtaining the relationship between the PAN image and its degraded version using a CAE. The patch size with overlapping pixels differs for each remote-sensing satellite. The experimental findings obtained from both degraded and authentic datasets demonstrated that our proposed method exhibited superior fusion quality when juxtaposed with other relevant methodologies. Future research will investigate whether the proposed technique can be scaled to other forms of image fusion. A further important implication is to improve the injected semantic detail map for superior outcomes.

ACKNOWLEDGEMENTS

This work is supported by the Office of Research, Zayed University, under fund ref: RIF R22050 and PRFA: R22147.




REFERENCES

- [1] F. Zhang, K. Zhang, and J. Sun, "Multiscale Spatial-Spectral Interaction Transformer for Pan-Sharpener," *Remote Sensing*, vol. 14, no. 7, p. 1736, 2022, doi: 10.3390/rs14071736.
- [2] K. Zhang, F. Zhang, and S. Yang, "Fusion of multispectral and panchromatic images via spatial weighted neighbor embedding," *Remote Sensing*, vol. 11, no. 5, p. 557, 2019, doi: 10.3390/rs11050557.
- [3] A. Al Smadi and A. Abugabah, "Intelligent information systems and image processing: A novel pan-sharpening technique based on Multiscale decomposition," in *ACM International Conference Proceeding Series*, New York, NY, USA: ACM, pp. 208–212, 2018, doi: 10.1145/3301506.3301523.
- [4] A. Alsmadi, S. Yang, and K. Zhang, "Pansharpening via Deep Guided Filtering Network," *International Journal of Image Processing and Visual Communication*, vol. 5, no. 1, pp. 2319–1724, 2018.
- [5] H. C. Luo, H. K. Zhou, and B. Sun, "Image Segmentation with Multi-feature Fusion in Compressed Domain based on Region-Based Graph," *International Arab Journal of Information Technology*, vol. 20, no. 2, pp. 159–169, 2023, doi: 10.34028/iajit/20/2/2.
- [6] H. Elmannai and A. D. Al-Garni, "Classification using semantic feature and machine learning: Land-use case application," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 19, no. 4, pp. 1242–1250, 2021, doi: 10.12928/TELKOMNIKA.v19i4.18359.
- [7] B. Mansouri, A. Khobzaoui, M. Damou, M. Chetoui, and A. Boudkhal, "Adaptive segmentation algorithm based on level set model in medical imaging," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 21, no. 5, pp. 1130–1138, 2023, doi: 10.12928/TELKOMNIKA.v21i5.22365.
- [8] J. Susan and P. Subashini, "Deep Learning inpainting Model on Digital and Medical Images-A Review," *International Arab Journal of Information Technology*, vol. 20, no. 6, pp. 919–936, 2023, doi: 10.34028/iajit/20/6/9.
- [9] G. Vivone et al., "A critical comparison among pansharpening algorithms," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 5, pp. 2565–2586, 2015, doi: 10.1109/TGRS.2014.2361734.
- [10] S. Rahmani, M. Strait, D. Merkurjev, M. Moeller, and T. Wittman, "An adaptive IHS pan-sharpening method," *IEEE Geoscience and Remote Sensing Letters*, vol. 7, no. 4, pp. 746–750, 2010, doi: 10.1109/LGRS.2010.2046715.
- [11] J. Liu, Y. Hui, and P. Zan, "Locally Linear Detail Injection for Pansharpening," *IEEE Access*, vol. 5, pp. 9728–9738, 2017, doi: 10.1109/ACCESS.2017.2710226.
- [12] Y. Gao, C. Song, C. Yang, M. Wang, and S. Yang, "Pansharpening with joint local low rank decomposition and hierarchical geometric filtering," *IEEE Access*, vol. 7, pp. 130578–130589, 2019, doi: 10.1109/ACCESS.2019.2940482.
- [13] J. Nunez, X. Otazu, O. Fors, A. Prades, V. Pala, and R. Arbiol, "Multiresolution-based image fusion with additive wavelet decomposition," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 3, pp. 1204–1211, 1999, doi: 10.1109/36.763274.
- [14] K. Rong, L. Jiao, S. Wang, and F. Liu, "Pansharpening based on low-rank and sparse decomposition," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 12, pp. 4793–4805, 2014, doi: 10.1109/JSTARS.2014.2347072.
- [15] S. Panchal and R. A. Thakker, "Improved Image Pansharpening Technique using Nonsampled Contourlet Transform with Sparse




- Representation,” *Journal of the Indian Society of Remote Sensing*, vol. 45, no. 3, pp. 385–394, 2017, doi: 10.1007/s12524-016-0608-z.
- [16] P. P. Singh and R. D. Garg, “A Hybrid Approach for Information Extraction from High Resolution Satellite Imagery,” *International Journal of Image and Graphics*, vol. 13, no. 2, p. 1340007, 2013, doi: 10.1142/S021946781340007X.
- [17] T. M. Tu, S. C. Su, H. C. Shyu, and P. S. Huang, “A new look at IHS-like image fusion methods,” *Information Fusion*, vol. 2, no. 3, pp. 177–186, 2001, doi: 10.1016/S1566-2535(01)00036-7.
- [18] Z. Wang, D. Ziou, C. Armenakis, D. Li, and Q. Li, “A comparative analysis of image fusion methods,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 6, pp. 1391–1402, 2005, doi: 10.1109/TGRS.2005.846874.
- [19] B. Aiazzi, S. Baronti, and M. Selva, “Improving component substitution pansharpening through multivariate regression of MS+Pan data,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 10, pp. 3230–3239, 2007, doi: 10.1109/TGRS.2007.901007.
- [20] A. R. Gillespie, A. B. Kahle, and R. E. Walker, “Color enhancement of highly correlated images. II. Channel ratio and ‘chromaticity’ transformation techniques,” *Remote Sensing of Environment*, vol. 22, no. 3, pp. 343–365, 1987, doi: 10.1016/0034-4257(87)90088-5.
- [21] M. M. Khan, J. Chanussot, L. Condat, and A. Montanvert, “Indusion: Fusion of multispectral and panchromatic images using the induction scaling technique,” *IEEE Geoscience and Remote Sensing Letters*, vol. 5, no. 1, pp. 98–102, 2008, doi: 10.1109/LGRS.2007.909934.
- [22] Z. Wang, S. Li, Y. Lv, and K. Yang, “Remote sensing image enhancement based on orthogonal wavelet transformation analysis and pseudo-color processing,” *International Journal of Computational Intelligence Systems*, vol. 3, no. 6, pp. 745–753, 2010, doi: 10.1080/18756891.2010.9727737.
- [23] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, and M. Selva, “MTF-tailored multiscale fusion of high-resolution MS and pan imagery,” *Photogrammetric Engineering and Remote Sensing*, vol. 72, no. 5, pp. 591–596, 2006, doi: 10.14358/PERS.72.5.591.
- [24] A. Azarang and H. Ghassemian, “A new pan *Image Analysis, IPRIA 2017*, IEEE, pp. 1–6, 2017, doi: 10.1109/PRIA.2017.7983017.
- [25] Y. Rao, L. He, and J. Zhu, “A residual convolutional neural network for pan-sharpening,” in *RSIP 2017 - International Workshop on Remote Sensing with Intelligent Processing, Proceedings*, IEEE, pp. 1–4, 2017, doi: 10.1109/RSIP.2017.7958807.
- [26] A. Azarang, H. E. Manoochehri, and N. Kehtarnavaz, “Convolutional Autoencoder-Based Multispectral Image Fusion,” *IEEE Access*, vol. 7, pp. 35673–35683, 2019, doi: 10.1109/ACCESS.2019.2905511.
- [27] A. Al Smadi, S. Yang, Z. Kai, A. Mehmood, M. Wang, and A. Alsanabani, “Pansharpening based on convolutional autoencoder and multi-scale guided filter,” *Eurasip Journal on Image and Video Processing*, vol. 2021, no. 1, p. 25, 2021, doi: 10.1186/s13640-021-00565-3.
- [28] A. Al Smadi, S. Yang, A. Abugabah, A. A. Alzubi, and L. Sanzogni, “A Pansharpening Based on the Non-Subsampled Contourlet Transform and Convolutional Autoencoder: Application to QuickBird Imagery,” *IEEE Access*, vol. 10, pp. 44778–44788, 2022, doi: 10.1109/ACCESS.2022.3169698.
- [29] Z. Wang, Y. Ma, and Y. Zhang, “Review of pixel-level remote sensing image fusion based on deep learning,” *Information Fusion*, vol. 90, pp. 36–58, 2023, doi: 10.1016/j.inffus.2022.09.008.
- [30] P. Wang, H. Yao, B. Huang, H. Leung, and P. Liu, “Multiresolution Analysis Pansharpening Based on Variation Factor for Multispectral and Panchromatic Images From Different Times,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1–17, 2023, doi: 10.1109/TGRS.2023.3252001.
- [31] N. Saxena, G. Saxena, N. Khare, and M. H. Rahman, “Pansharpening scheme using spatial detail injection-based convolutional neural networks,” *IET Image Processing*, vol. 16, no. 9, pp. 2297–2307, 2022, doi: 10.1049/ipr2.12384.
- [32] W. Huang, L. Xiao, Z. Wei, H. Liu, and S. Tang, “A new pan-sharpening method with deep neural networks,” *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 5, pp. 1037–1041, 2015, doi: 10.1109/LGRS.2014.2376034.
- [33] M. Li, J. Li, Y. Liu, and F. Liu, “Detail Injection-Based Convolutional Auto-Encoder for Pansharpening,” *Journal of Remote Sensing (United States)*, vol. 2022, 2022, doi: 10.34133/remotesensing.0004.
- [34] S. Dolgikh, “Spontaneous concept learning with deep autoencoder,” *International Journal of Computational Intelligence Systems*, vol. 12, no. 1, pp. 1–12, 2018, doi: 10.2991/ijcis.2018.25905178.
- [35] J. Masci, U. Meier, D. Cireşan, and J. Schmidhuber, “Stacked convolutional auto-encoders for hierarchical feature extraction,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, pp. 52–59, 2011, doi: 10.1007/978-3-642-21735-7_7.
- [36] G. Vivone, M. D. Mura, A. Garzelli, and F. Pacifici, “A Benchmarking Protocol for Pansharpening: Dataset, Preprocessing, and Quality Assessment,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 6102–6118, 2021, doi: 10.1109/JSTARS.2021.3086877.
- [37] A. Garzelli, F. Nencini, and L. Capobianco, “Optimal MMSE pan sharpening of very high resolution multispectral images,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 1, pp. 228–236, 2008, doi: 10.1109/TGRS.2007.907604.
- [38] J. Choi, K. Yu, and Y. Kim, “A new adaptive component-substitution-based satellite image fusion by using partial replacement,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 1, pp. 295–309, 2011, doi: 10.1109/TGRS.2010.2051674.
- [39] A. Azarang and H. Ghassemian, “Application of fractional-order differentiation in multispectral image fusion,” *Remote Sensing Letters*, vol. 9, no. 1, pp. 91–100, 2018, doi: 10.1080/2150704X.2017.1395963.

BIOGRAPHIES OF AUTHORS






Ahmad Al Smadi    is an assistant professor in the Department of Data Science and Artificial Intelligence at Zarqa University. He earned his Ph.D. in Computer Science and Technology from Xidian University, China. Previously, he was a research assistant at Zayed University, UAE, focusing on information systems, computer vision, machine learning, and deep learning. He can be contacted at email: aalsmadi@zu.edu.jo.






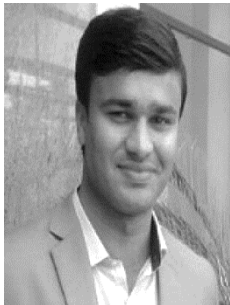
Ahed Abugabah    is an associate professor at the College of Technological Innovation, Zayed University. He earned his degrees in information systems in Australia and has extensive experience in the airline industry, particularly in Aircraft Engineering and Supply Chain Management. Before joining Zayed University in 2016, he served as an Associate Dean and University Council Member at the American University, UAE. His research interests encompass information systems, enterprise applications and development, healthcare information systems, and RFID in healthcare. He can be contacted at email: ahed.abugabah@zu.ac.ae.






Mutasem Khilail Alsmadi    is currently an associate professor at the Faculty of Applied Studies and Community Service, Department of Management of Information Systems, Imam Abdurrahman Bin Faisal University. He received his B.S. degree in Software Engineering in 2006 from Philadelphia University, Jordan, his M.Sc. degree in Intelligent Systems in 2007 from University Utara Malaysia, Malaysia, and his Ph.D. in Computer Science from The National University of Malaysia. He has published more than one hundred papers in the image processing and algorithm optimization areas. His research interests include artificial intelligence, pattern recognition, algorithms optimization, and computer vision. He can be contacted at email: mksalsmadi@gmail.com.






Ala Alsanabani    received the B.S degree in Computer Science from Tamar University, Yemen, in 2008, and the M.S degree in Information Technology from Open Malaysian University, in 2015. He is currently pursuing the Ph.D. degree in Computer Vision with the School of Computer Science and Technology, Xidian University, China. His research interests include surveillance image and video analysis, computer vision, and machine learning. He can be contacted at email: ala.alsanabani@gmail.com.



Atif Mahmood    with a Ph.D. from Xidian University, is an assistant professor at NUML, Pakistan. He is also a postdoc associate at Zhejiang Normal University, China, specializing in machine learning, image processing, and computer vision. He can be contacted at email: atifedu151@yahoo.com.



Ahmad Mohammad Al-Smadi    is a lecturer at Al-Balqa' Applied University (BAU), Jordan, holding a distinguished Doctoral degree in Management Information Systems from Jinan University (2016). With over 20 years of teaching experience, his research focuses on Information System Management, Business Intelligence, Decision Support Systems, and Artificial Intelligence. He has authored 10+ publications in peer-reviewed journals and conferences. He can be contacted at email: amhs1966@bau.edu.jo.