Journal of Theoretical and Applied Information Technology
31st August 2026. Vol.104. No.16

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

IMPROVING IMAGE FUSION QUALITY WITH APC-FUSENET: AN AUTOENCODER AND PCNN-BASED METHOD

BATTIREDDY SURENDARA[1], S. ABIRAMI[2], GURRAM SUNITHA[3]

[1,2]Department of Computer Science and Engineering, Annamalai University, Chidambaram, India

[3]Department of Computer Science and Engineering, Mohan Babu University, Tirupati, India

E-mail: [1]Surendara.nyp7@gmail.com*, [2]abiramiaucse@gmail.com, [3]gurramsunitha@gmail.com

ABSTRACT

The objective of image fusion is to develop a single composite image that effectively incorporates the most relevant and informative details from each image respectively. In this research, a new multi-focus image fusion method, APC-FuseNet model is proposed. The model integrates an autoencoder with a Pulse-Coupled Neural Network (PCNN) to improve the quality and focus of the fused images. The methodology commences with loading and preprocessing of Multi Spectral (MS) and Panchromatic (PAN) images to ensure compatibility. Subsequently, Bilateral and Laplacian of Gaussian (LoG) filters are applied to the PAN image to preserve edges and suppress noise. The PAN image is then separated into detail, coarse, and base layers. A pre-trained autoencoder is used to extract compressed features. The Multi-Scale Morphological Gradient (MSMG) is utilized to enhance prominent characteristics, while a PCNN analyzes the MSMG outcomes to generate binary maps for the purpose of merging selective region. The ultimate merged image is acquired by integrating these outcomes into the HSV color space and then converting it back to Red Green Blue (RGB). The APC-FuseNet model's effectiveness is demonstrated through visual inspection and metric analysis, thereby revealing its superior performance in ERGAS (2.30564), UIQI (0.99521), CC (0.99634), PSNR (44.012), and SSIM (0.92017) compared to other methods. This confirms its ability to generate high-quality, focused fused images.

Keywords: *Autoencoder, Deep Learning, Multi Focus, Panchromatic Satellite, Pulse-Coupled Neural Network, Remote Sensing.*

1. INTRODUCTION

Multi-focus satellite images are essential in remote sensing applications as they provide extensive data for a wide range of environmental, agricultural, and urban studies [1-4]. These images typically include multispectral (MS) images, which record data across multiple wavelengths, and panchromatic (PAN) images, which offer high-resolution spatial details in a single band [5-8]. MS images are highly valuable because their rich spectral information enables the identification of various materials and vegetation types [9-10]. In contrast, PAN images are crucial due to their exceptional spatial resolution, which provides detailed visual information [11-14]. By merging these two types of images, the overall data quality and usability can be greatly enhanced, thereby making image fusion a critical process in remote sensing [15-18].

Image fusion is essential because MS and PAN images possess complementary characteristics. MS images provide rich spectral information, while PAN images offer high spatial clarity [19]. Contemporary image fusion techniques include traditional methods such as the Brovey Transform (BT), Principal Component Analysis (PCA), as well as more sophisticated approaches like Deep Residual Networks (DRN), Generative Adversarial Networks (GANs), and various forms of spatiotemporal fusion. Despite these advancements, challenges remain in achieving optimal fusion outcomes, especially in maintaining a proper equilibrium between spectral and spatial quality.

1.1 Problem Statement and Objectives

In image fusion tasks, particularly those involving Panchromatic (PAN) and multispectral (MS) images, maintaining spatial detail while effectively reducing noise remains a major challenge. Traditional image fusion methods often suffer from information loss, poor noise suppression, and inadequate feature representation, resulting in suboptimal image quality. To address these issues, the Autoencoder and Pulse-Coupled Fused Network (APC-FuseNet) is designed to extract meaningful features, preserve essential details, and suppress noise, thereby improving the overall quality and accuracy of the fused image.

Thank you for the valuable comment. Based on the reviewer's comment, we have revised the introduction section to clearly define the fusion problem, state three research questions, specify the scope of the proposed claim, identify the evidence used for evaluation, and describe the intended relevance of the findings. The contribution of APC-FuseNet is described in terms of the staged integration and defined roles of its computational components, while the individual components are recognized as established methods.

1.2 Research Perspective, Questions, and Testable Claim

This study interprets PAN–MS fusion as a constrained information-preservation problem: spatial detail from the panchromatic image should be introduced without materially degrading the multispectral content. Readers should therefore expect an explicit processing path from noise-aware decomposition to feature encoding, activity-based selection, reconstruction, and evaluation. The findings can be used to judge whether this hybrid path is a suitable preprocessing option for subsequent remote-sensing analysis and to identify which stages require controlled ablation or adaptation.

RQ1: Does APC-FuseNet produce fused Landsat 7 images with low global error and high spectral, correlation, signal, and structural quality according to ERGAS, UIQI, CC, PSNR, and SSIM?

RQ2: How does the division of roles among bilateral/LoG filtering, layer decomposition, autoencoder encoding, MSMG, and PCNN region selection explain the observed fusion behavior?

RQ3: What practical use and limitations follow from the observed results, given the dataset size and the different evaluation settings used in prior literature?

The testable claim is that, for the selected Landsat 7 image pairs and the stated evaluation procedure, APC-FuseNet will produce lower ERGAS and higher UIQI, CC, PSNR, and SSIM than the contextual values listed in Tables 1–5. This claim is deliberately restricted to the evaluated setting; it can be challenged through same-dataset reimplementation, external sensor testing, ablation, and downstream-task evaluation.

The novelty lies in the ordered integration of bilateral and LoG preprocessing, three-layer PAN decomposition, autoencoder representation, MSMG activity extraction, PCNN selection, and HSV reconstruction. The individual operations are established; the research question is whether their complementary roles improve PAN–MS quality.

1.3 Contributions

The key contributions of this research are described as follows:

- The proposed APC-FuseNet combines the advantages of autoencoders for extracting features and Pulse-Coupled Neural Networks (PCNNs) for enhancing fusion quality.
- Image preprocessing is performed using Bilateral and Laplacian of Gaussian (LoG) filters, and the PAN image is decomposed into detail, coarse, and base layers, to efficiently eliminate noise.
- The pre-trained autoencoder, along with MSMG and PCNN algorithms, is employed to produce fused image with improved performance metrics compared to existing techniques.

The organization of this paper is as follows: Section 2 provides a comprehensive analysis of recent works, Section 3 provides a detailed explanation of the proposed system and the APC-FuseNet. Section 4 reports the findings and provides an in-depth examination of the performance metrics. Section 5 concludes the paper with a summary of the results.

2. RELATED WORKS

Li et al. [20] developed the Deep Residual Network (DRN) for combining satellite images, demonstrating the efficacy of deep learning in improving the quality of fused satellite images by utilizing residual connections to capture intricate details and minimize information loss. Nevertheless, the intricacy of the model and the need for substantial computational resources restricted its usefulness in real-time or resource-limited settings.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Reddy and Shashi [21] utilized the Brovey Transform (BT) to fuse images and demonstrated its utility in integrating spectral and spatial data, effectively improving the visual clarity of the merged images. However, the relatively lower UIQI suggested that this method might not consistently maintain high-quality indices, which could potentially have impacted the accuracy of subsequent analyses based on the fused images.

Wang et al. [22] conducted a study on a spatiotemporal fusion (STF) method that successfully tackled the difficulties of integrating spatial and temporal data to generate superior fused images, showcasing notable enhancements in capturing temporal variations. Although the fused images had some positive aspects, the moderate SSIM indicated that the structural integrity of these images was occasionally compromised which could have affected applications that rely on accurate spatial details.

Wang et al. [23] introduced the Inter-Image Variability process technique. The DIFIV method successfully addressed the differences in variability between hyperspectral and multispectral images, resulting in improved quality and usefulness of the fused outputs. Nevertheless, the comparatively lower Peak Signal-to-Noise Ratio (PSNR) suggested that noise levels could still be problematic, potentially impacting the overall visual sharpness of the images.

Brezini and Deville [24] utilized Sparsity-Based Unmixing to fuse images, leveraging the sparsity of image data to produce fused image outputs of superior quality, characterized by exceptional noise reduction and preservation of fine details. However, the substantial computational requirements of sparse coding techniques could hinder their effectiveness, particularly when dealing with extensive image dataset or real-time applications.

Vakilian and Saradjian [25] proposed a method called spatiotemporal sparse representation (STSR) that successfully combined spatial and temporal information, resulting in highly accurate fused images. However, the relatively high ERGAS value suggested that radiometric accuracy was still limited, which is essential for accurate quantitative analyses.

Azarang and Kehtarnavaz [26] utilized a Generative Model to merge images, which exhibited robust capabilities in producing high-quality fused images by leveraging deep learning to learn intricate data distributions. Nevertheless, generative models frequently required substantial amounts of training data and computational resources, potentially constraining their accessibility and scalability.

Tan et al. [27] introduced a Dual-Channel PCNN for the purpose of image fusion. This approach employed pulse-coupled neural networks to optimize the fusion process, effectively capturing intricate image details. Although the approach was innovative, the relatively low UIQI and high ERGAS indicated potential difficulties in maintaining consistent image quality.

Each of these methods had distinct advantages and difficulties. Although deep learning techniques such as DRN and generative models demonstrated exceptional performance, they frequently entailed substantial computational expenses. While traditional techniques such as the BT and sparse representations offered easier solutions, but they did not always maintain high-quality indices. The selection of the method was determined by the specific demands of the application, taking into account the trade-off between quality, computational resources, and ease of implementation.

3. METHODOLOGY

The proposed APC-FuseNet image fusion begins by importing the MS and PAN images from the LANDSAT 7 dataset, as shown in Figure. 1. The images are preprocessed to ensure compatibility, which includes resizing the MS image to match the PAN image and converting both to the HSV color space to extract the intensity component. The PAN image is then processed using a Bilateral filter to produce a refined version that has less noise. This helps in separating the image into different frequency components while preserving important structural details. Subsequently, the PAN image is enhanced using a Laplacian of Gaussian (LoG) filter, which combines the edge-detection capability of the Laplacian filter with the noise-suppressing properties of the Gaussian filter, thereby effectively capturing intricate details and high-frequency data.

The PAN image is decomposed into three layers: the detail layer captures intricate details and high-frequency data; the coarse layer captures medium-frequency data; and the base layer represents the low-frequency content, which is enhanced using guided filtering. A pre-trained autoencoder and PCNN are employed to analyze the MSMG outcomes, generating binary maps that highlight regions of increased activity or significance. These maps selectively combine areas from the input matrices, preserving the most relevant features from both images. Ultimately, the combined results are integrated into the HSV color space by replacing the intensity element with the fused component. The

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

image is then converted back into the RGB color space to obtain the final fused output. The efficacy of the fusion process is evaluated by visually examining the initial MS, PAN, and fused images. This analysis demonstrates that APC-FuseNet improves the fusion process by employing autoencoders and PCNNs to generate a superior, concentrated output image.

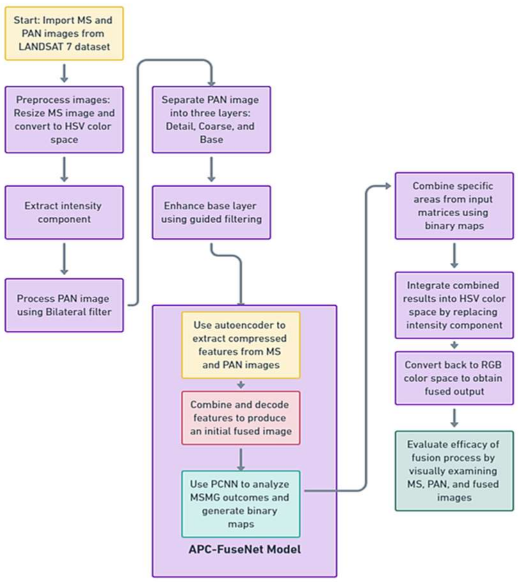

Figure 1: Proposed System block diagram for Image fusion

The multi-focus image fusion process in elaboration with mathematical equations is provided below. The workflow begins by loading the multispectral (MS) image and panchromatic (PAN) image using the proposed APC-FuseNet framework. Let I_MS denote the multispectral (MS) image and I_PAN denote the panchromatic (PAN) image. The images undergo preprocessing to ensure compatibility, which involves resizing the MS image to match the PAN image and converting both images to the HSV color space to extract the intensity component. This conversion is expressed mathematically in the following Equation (1):

$$I_{PAN}: I_{MS} \rightarrow \widetilde{I_{MS}} \quad (1)$$

The preprocessed images $\widetilde{I_{MS}}$ and I_PAN are converted into the HSV color space to extract the intensity component I_V. The PAN image is subsequently processed using a Bilateral filter to generate a refined version with reduced noise. The mathematical representation of denoising process using Bilateral filter is given in Equation (2):

$$I_{BF} = BilateralFilter(I_{PAN}, \sigma_s, \sigma_r) \quad (2)$$

where σ_s represents the spatial standard deviation and σ_r denotes the range standard deviation. This process facilitates the separation of the image into distinct frequency components while maintaining the integrity of edges, thus ensuring the preservation of crucial structural information. Subsequently, the PAN image undergoes the application of the LoG filter, which serves to accentuate edges and textures and is expressed in Equation (3):

$$I_{LoG} = \nabla^2 G_\sigma(I_{PAN}) \quad (3)$$

where $\nabla^2 G_\sigma$ represents the Laplacian of the Gaussian function with standard deviation σ. This filter enhances structural information by merging the edge detection ability of the Laplacian filter with the noise-reducing characteristics of the Gaussian filter, resulting in the accurate representation of intricate details and high-frequency data. The PAN image is decomposed into three layers: the detail layer S, which is obtained by subtracting the LoG filtered image from the original PAN image and captures fine details and high-frequency information; the coarse layer L, which is obtained by subtracting the Bilateral filtered image from the LoG filtered image and captures mid-frequency information; and the base layer B, which represents the low-frequency content and is obtained by further refining the Bilateral filtered image using guided filtering which is mathematically formulated in Equations (4) – (6):

$$S = I_{PAN} - I_{LoG} \quad (4)$$
$$L = I_{LoG} - I_{BF} \quad (5)$$
$$B = GuidedFilter(IBF, r, \epsilon) \quad (6)$$

A pre-trained autoencoder is used to extract compressed features (F) from a dataset of paired MS and PAN images which is done using Equations (7) and (8):

$$F_{MS} = Encoder(\widetilde{I_{MS}}) \quad (7)$$
$$F_{PAN} = Encoder(I_{PAN}) \quad (8)$$

The encoded features extracted from the MS and PAN images are merged into a unified representation, which is then decoded to produce the initial fused image, as expressed in Equation (9):

$$F_{merged} = Merge(F_{MS}, F_{PAN}) \quad (9)$$

To further enhance the quality of the initial fused image, MSMG is applied to both the decomposed detail and base layers, emphasizing important features at different scales. This process is expressed in Equation (10):

$$I_{fused,init} = Decoder(F_{merged}) \quad (10)$$

Subsequently, a Pulse-Coupled Neural Network (PCNN) is initialized with predetermined parameters and used to process the MSMG results derived from the input matrices. The MSMG is applied to both the detail and base layers, as described by the mathematical formulations in Equations (11) – (13):

$$G_S = MSMG(S) \quad (11)$$
$$G_B = MSMG(B) \quad (12)$$
$$PCNN(G) = \{Bi, j\}i, j \quad (13)$$

This process produces binary maps $\{Bi, j\}$ that identify regions of increased activity or significance. These binary maps are employed to selectively merge regions from the input matrices, ensuring the preservation of the most pertinent characteristics from both images. Ultimately, the merged outcomes are restored by incorporating them into the HSV color space, substituting the intensity element of the original image with the fused element. Subsequently, the image is converted back to the RGB color space to acquire the final fused output. The effectiveness of the fusion process is assessed by visualizing the original MS, PAN, and fused images. The APC-FuseNet methodology utilizes autoencoders to extract features and PCNNs to enhance the fusion process, resulting in a high-quality and focused output image.

3.1 Proposed Fusion Model

APC-FuseNet is a specialized framework designed to enhance the fusion of satellite images, producing superior and more focused results. This is achieved by integrating an autoencoder and a PCNN model respectively. Satellite images frequently encounter problems such as noise, blurriness, and insufficient detail. The autoencoder effectively extracts salient features from the input images, capturing crucial patterns and structures while reducing noise. The PCNN, renowned for its resilient image segmentation capabilities, improves the fusion process by preserving vital details and intensifying the focus. By integrating the images, APC-FuseNet ensures that the resulting fused images maintain high spatial and spectral quality. This makes APC-FuseNet ideal for applications involving satellite imagery, where accuracy and sharpness are of utmost importance.

The architecture of the proposed APC-FuseNet model consists of two primary components such as: an autoencoder and a PCNN. Figure 2. illustrate the architecture of proposed model. The autoencoder is composed of an encoder and a decoder. The encoder includes multiple convolutional layers with Rectified Linear Unit (ReLU) activation functions, each followed by max-pooling layers to reduce the spatial dimensions while preserving essential features. The decoder mirrors the encoder through the use of transposed convolutional layers, gradually reconstructing the image to its original dimensions. Skip connections are employed to establish connections between corresponding layers in the encoder and decoder, guaranteeing the preservation of intricate details. The autoencoder's output is then fed into the PCNN, which incorporates layers dedicated to image segmentation and detail enhancement. The PCNN employs iterative pulse-coupling mechanisms that adaptively modify themselves according to the intensity values with a specific emphasis on important image characteristics, resulting in an improved fusion quality.

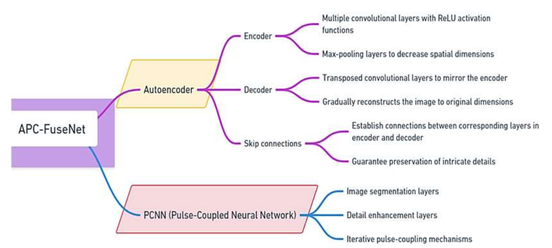

Figure 2: Proposed APC-FuseNet model Architecture

The key innovation of the APC-FuseNet model resides in its effective integration of an autoencoder and a PCNN, surpassing the capabilities of conventional approaches. While standard autoencoders are proficient at feature extraction, they may not sufficiently retain the intricate details required for high-quality image fusion. Conversely, a standalone PCNN, although effective for segmentation, lacks the extensive feature extraction capabilities of an autoencoder. APC-FuseNet combines these two techniques to leverage their respective strengths. The autoencoder efficiently extracts salient features and suppresses noise, while the PCNN further enhances these features to produce a fused image that is both highly detailed and focused. This hybrid approach yields enhanced image quality compared to using either a conventional PCNN or autoencoder, as well as other conventional image fusion techniques.

The APC-FuseNet method exhibits substantial enhancements over image fusion techniques, such as wavelet transform or simple averaging methods. While wavelet transform techniques are effective for multi-resolution analysis, they frequently encounter artifacts and results in loss of detail. Simple averaging techniques are inadequate for capturing the intricate relationships among image features. The proposed APC-FuseNet model addresses these limitations by combining the feature extraction capabilities of a deep learning-based autoencoder, with the detail enhancement techniques of a PCNN. This hybrid approach produces merged images that accurately preserve the original inputs while enhancing clarity and fine details, making APC-

FuseNet as a preferred option for satellite image fusion tasks.

3.2 Pseudo Code for Proposed APC-FuseNet Model

The proposed APC-FuseNet is an advanced image fusion algorithm designed for image processing tasks, with a particular focus on image fusion and segmentation. It integrates an autoencoder with a PCNN model to extract, reconstruct, and enhance images effectively. Algorithm 1 outlines the workflow the proposed APC-FuseNet based image fusion model.

Algorithm: APC-FuseNet Model

```
  # Step 1: Define the Encoder
  function Encoder(input):
    layer1 = Conv2D(input, filters=64, kernel_size=3, activation=ReLU)
    layer1_pooled = MaxPooling2D(layer1, pool_size=2)

    layer2 = Conv2D(layer1_pooled)
    layer2_pooled = MaxPooling2D(layer2, pool_size=2)

    layer3 = Conv2D(layer2_pooled, filters=256, kernel_size=3, activation=ReLU)
    layer3_pooled = MaxPooling2D(layer3, pool_size=2)

    layer4 = Conv2D(layer3_pooled, filters=512, kernel_size=3, activation=ReLU)
    layer4_pooled = MaxPooling2D(layer4, pool_size=2)

    return [layer1, layer2, layer3, layer4, layer4_pooled]

  # Step 2: Define the Decoder with Skip Connections
  function Decoder(encoded_layers):
    layer4_pooled = encoded_layers[4]

    layer4_up = Conv2DTranspose(layer4_pooled)
    layer4_up = Concatenate(layer4_up, encoded_layers[3])

    layer3_up = Conv2DTranspose(layer4_up)
    layer3_up = Concatenate(layer3_up, encoded_layers[2])

    layer2_up = Conv2DTranspose(layer3_up)
    layer2_up = Concatenate(layer2_up, encoded_layers[1])

    layer1_up = Conv2DTranspose(layer2_up)
    layer1_up = Concatenate(layer1_up, encoded_layers[0])

    output = Conv2DTranspose(layer1_up, filters=1, kernel_size=3, activation=Sigmoid)

    return output

  # Define the Pulse Coupled Neural Network (PCNN)
  function PCNN(input):
    layer1 = PCNNLayer(input, filters=64)
    layer2 = PCNNLayer(layer1, filters=128)
    layer3 = PCNNLayer(layer2, filters=256)
    output = PCNNLayer(layer3, filters=1)

    return output

  # Define the APC-FuseNet Model
  function APC_FuseNet(input_image):
    encoded_layers = Encoder(input_image)
    decoded_image = Decoder(encoded_layers)
    final_output = PCNN(decoded_image)

    return final_output

  # Example usage
  input_image = LoadImage('path_to_image')
  output_image = APC_FuseNet(input_image)
  SaveImage(output_image, 'path_to_save_image')
```

4. RESULTS AND ANALYSIS

Fusion quality is assessed using five complementary metrics, where lower ERGAS indicates lower normalized global radiometric error, while UIQI, CC, and SSIM indicate greater agreement as their values approach 1, and higher PSNR indicates lower reconstruction error relative to the signal range. These interpretation directions follow the conventions reported in the cited fusion literature [20–27], with metric values influenced by the sensor combination, image content, reference-image construction, and preprocessing procedures rather than by a fixed acceptance threshold. APC-FuseNet is evaluated using all the five metrics, while values reported in other studies are considered contextual references rather than equivalent experimental results.

For this study, we employed the USGS Landsat 7 Collection 2 Tier 1 TOA refers to the dataset of Landsat 7 satellite [28] imagery provided by the

Journal of Theoretical and Applied Information Technology
31st August 2026. Vol.104. No.16

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

United States Geological Survey (USGS). This dataset is classified as Tier 1, indicating high quality and reliability. TOA stands for Top of Atmosphere, which refers to the measurements taken by the satellite sensor at the outermost layer of the Earth's atmosphere. The reflectance dataset provides a vast amount of geospatial data, including both multispectral (MS) and panchromatic (PAN) images, which are well-suited for in-depth analysis of the Earth's surface. The dataset, covering the time period from May 28, 1999, to January 19, 2024, is extremely valuable for applications such as environmental monitoring, land cover change detection, and vegetation analysis.

For our investigation, we selected 48 specific images from this dataset to examine how they influence the performance of vegetation indices when combined with UAV (Unmanned Aerial Vehicle) imagery. The images were exemplary specimens, covering a wide range of geographical regions and historical periods, enabling us to assess the impact of image fusion on the accuracy and reliability of vegetation index measurements. The multispectral images captured by Landsat 7 provide comprehensive data across multiple wavelengths, which is crucial for evaluating vegetation health. Additionally, the high-resolution panchromatic images offer enhanced spatial detail. This investigation leverages the advanced capabilities of the Earth Engine platform, which facilitates the analysis and visualization of large-scale scientific dataset. The platform is freely available for research, educational, and nonprofit purposes, making it a highly accessible tool for conducting extensive geospatial analyses.

200*200 MS image

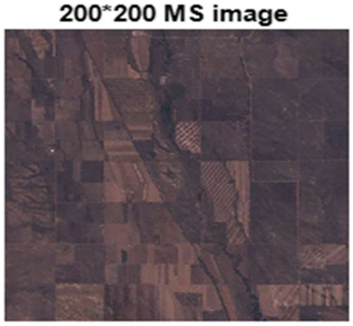

Figure 3: 200 x 200 Multispectral Image input -I

Figure 3 displays a multispectral image input with dimensions of 200 x 200, referred to as Input-I. This image offers comprehensive spectral data across numerous bands, which is essential for diverse remote sensing applications. The multispectral image acquires data at various wavelengths, allowing for the examination of characteristics that are not discernible in conventional optical images. The provided input is an essential element in the fusion process, as it adds valuable spectral information that improves the overall quality of the fused output.

Figure 4 shows a 400 x 400 panchromatic image input, also referred to as Input-I. This image delivers high level of spatial resolution, capturing fine details in a single band, usually within the visible spectrum. The panchromatic image plays a crucial in the fusion process by providing precise and distinct spatial information. Its higher resolution of the panchromatic image compared to the multispectral image helps enhance the spatial characteristics of the fused image, resulting in improved clarity and finer detail.

400*400 PAN image

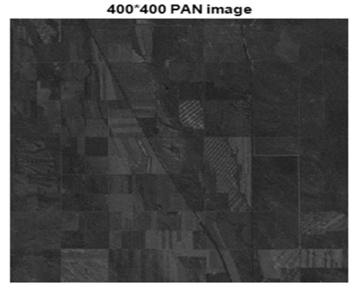

Figure 4: 400 x 400 Panchromatic Image input -I

Figure 5 displays the fused Output-I, obtained by merging the multispectral (MS) and panchromatic (PAN) images from Inputs-I. The fusion process aims to integrate the spectral diversity of the MS image with the spatial sharpness of the PAN image. The resulting fused image demonstrates improved visual quality and greater information content, rendering it suitable for more precise analysis and interpretation across a wide range of applications.

Fused image

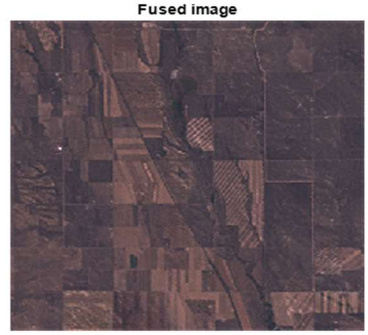

Figure 5: Fused Output –I

Figure 6 displays an additional 200 x 200 multispectral image input, referred to as input-II. Similar to Input-I, this image offers extensive spectral information that is essential for the fusion process. The uniformity of the multispectral images in terms of size and format ensures that the fusion

Journal of Theoretical and Applied Information Technology
31st August 2026. Vol.104. No.16

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

process can be consistently applied to various dataset, thereby preserving the quality and reliability of the resulting outputs.

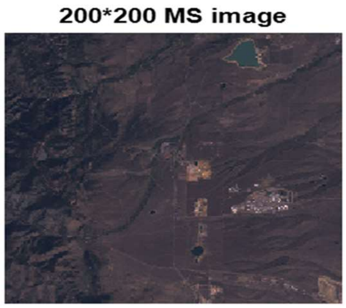

Figure 6: 200 x 200 Multispectral Image input -II

Figure 7 shows a panchromatic image input with dimensions of 400 x 400, which is referred to as input-II. Like the previous panchromatic input, this image provides essential high-resolution spatial details for the fusion process. Its larger dimensions compared to the MS image facilitate the extraction of intricate spatial features, enhancing the quality of the fused result.

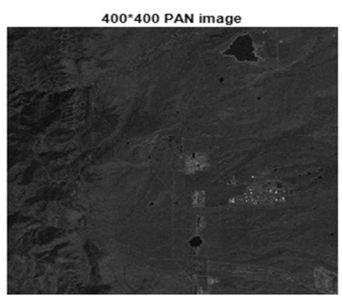

Figure 7: 400 x 400 Panchromatic Image input -II

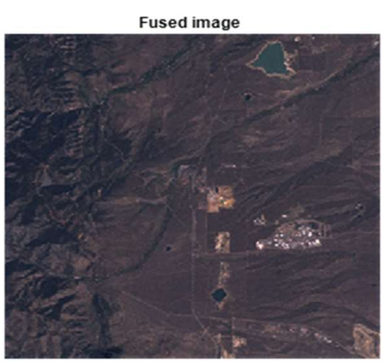

Figure 8: Fused Output –II

The fused output-II, depicted in Figure 8, is the outcome of combining the multispectral and panchromatic images from inputs-II. This composite image demonstrates the efficacy of the APC-FuseNet approach in integrating spectral and spatial data from various image sources. The resulting image exhibits improved detail and enriched information content, highlighting the capability of the fusion process to generate high-quality outputs.

4.1 Comparative Analysis

The comparative values for the existing methods presented in Tables 1–5 are obtained from their corresponding publications [20]–[27], whereas the APC-FuseNet results were obtained from the experiments conducted in the present study. The cited publications used different datasets, sensor combinations, preprocessing procedures, and experimental protocols. Therefore, the reported values are presented as contextual literature comparisons rather than as results from a controlled evaluation under identical experimental conditions.

Table 1 presents the Error Relative Global Dimensional Synthesis (ERGAS) values reported for different image-fusion methods. A lower ERGAS value indicates reduced global radiometric distortion in the fused image. APC-FuseNet obtained an ERGAS value of 2.30564 on the evaluated Landsat 7 images, which is lower than the literature-reported values included in the table. This result indicates effective radiometric preservation under the experimental conditions of the present study.

Table 1: Comparative ERGAS results for APC-FuseNet and existing image-fusion methods.

Method	ERGAS
DRN [20]	3.725
STF [22]	2.92014
DIFIV [23]	2.396
Sparsity-Based Unmixing [24]	4.966
STSR [25]	9.25
Generative Model [26]	2.36
Dual-Channel PCNN [27]	26.67
Proposed APC-FuseNet	2.30564

Table 2 compares the Universal Image Quality Index (UIQI) values of the selected methods. UIQI values approaching 1 indicate stronger agreement between the fused image and the reference information in terms of correlation, luminance, and contrast. APC-FuseNet achieved a UIQI of 0.99521, demonstrating strong preservation of image information on the evaluated dataset.

Table 2: Comparative UIQI results for APC-FuseNet and existing image-fusion methods.

Method	UIQI
DRN [20]	0.992
BT [21]	0.418
STF [22]	0.83482
DIFIV [23]	0.899

Sparsity-Based Unmixing [24]	0.9728
STSR [25]	0.9924
Dual-Channel PCNN [27]	0.74
Proposed APC-FuseNet	0.99521

Table 3 presents the Correlation Coefficient (CC) values reported for the selected image-fusion approaches. A CC value closer to 1 represents stronger correspondence between the fused image and the source information. APC-FuseNet produced a CC value of 0.99634, indicating a high degree of correlation under the present experimental conditions.

Table 3: Comparative CC results for APC-FuseNet and existing image-fusion methods.

Method	CC
DRN [20]	0.993
BT [21]	0.975
STF [22]	0.91752
Generative Model [26]	0.97
Proposed APC-FuseNet	0.99634

Table 4: Comparative PSNR results for APC-FuseNet and existing image-fusion methods.

Method	PSNR (dB)
DRN [20]	29.557
DIFIV [23]	25.834
Sparsity-Based Unmixing [24]	43.01
Proposed APC-FuseNet	44.012

Table 4 compares the Peak Signal-to-Noise Ratio (PSNR) values of APC-FuseNet and the selected existing methods. A higher PSNR generally indicates lower reconstruction error relative to the signal range. APC-FuseNet achieved a PSNR of 44.012 dB on the evaluated Landsat 7 images, demonstrating effective reconstruction of the fused image.

Table 5 presents the Structural Similarity Index Measure (SSIM) values of the selected methods. SSIM evaluates the preservation of luminance, contrast, and structural information, with values closer to 1 indicating greater structural similarity. APC-FuseNet obtained an SSIM of 0.92017, indicating effective preservation of structural content in the fused images.

Table 5: Comparative SSIM results for APC-FuseNet and existing image-fusion methods.

Method	SSIM
BT [21]	0.868
STF [22]	0.81792
Proposed APC-FuseNet	0.92017

4.2 Difference from Prior Work and Computing Contribution

The existing methods reviewed in Section 2 addresses various image-fusion settings and objectives. Residual and generative networks primarily learn end-to-end mappings [20,26], whereas sparse methods focus on unmixing and spatiotemporal reconstruction [24,25]. PCNN-based approaches use activity and boundary information for the fusion process [27]. APC-FuseNet is designed specifically for PAN–MS fusion, with the objectives of preserving spatial details, reducing noise, and maintaining spectral consistency through a structured sequence of filtering, decomposition, learned feature extraction, region selection, and HSV-based reconstruction. Table 6 compares the workflow of APC-FuseNet with the cited approaches and presents their methodological differences.

Table 6: Advantage of Proposed work over prior works

Prior method	Primary scope	Limitation relative to the present task	APC-FuseNet distinction
Progressive cascaded DRN [20]	End-to-end residual learning for remote-sensing image fusion.	Deep cascades increase training and inference cost; no PCNN-based region selection.	Uses filtered layer decomposition, autoencoder features, and PCNN selection in one staged path.
PCA/Brovey fusion [21]	Transform-based fusion of UAV and Sentinel-2A imagery for vegetation indices.	A fixed transform failed to learn image-dependent representations.	Learns compressed features and performs activity-based region selection before reconstruction.
STF [22]	Single-band and multiband prediction for spatiotemporal fusion.	Designed for temporal prediction rather than PAN–MS spatial–spectral fusion.	Targets PAN–MS fusion and evaluates spatial, spectral, and structural fidelity.
DIFIV [23]	Deep hyperspectral–multispectral fusion with inter-image variability.	Uses a different input modality and degradation setting.	Focuses on multispectral and panchromatic Landsat inputs with explicit detail/base decomposition.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Sparsity-based unmixing [24]	Endmember-bundle extraction and sparse unmixing for spectral variability.	Requires endmember estimation and iterative sparse optimization.	Uses learned encoding and PCNN activation maps without an unmixing stage.
STSR [25]	Object-based sparse representation for spatiotemporal gap filling.	Targets temporal data gaps and depends on sparse dictionary representation.	Targets same-scene spatial–spectral fusion through a deterministic reconstruction sequence.
Generative model [26]	Unsupervised generative fusion with spectral and spatial objectives.	Adversarial optimization increases training complexity and stability requirements.	Uses an autoencoder–PCNN sequence with directly interpretable preprocessing and selection stages.
Dual-channel PCNN [27]	Boundary-measured PCNN fusion in the multi-scale morphological-gradient domain.	It failed to include learned autoencoder compression before PCNN-based selection.	Adds bilateral/LoG decomposition and autoencoder feature extraction before MSMG–PCNN refinement.

The computational contribution of APC-FuseNet lies in the ordered integration of bilateral filtering, LoG filtering, autoencoder-based representation learning, MSMG-based activity estimation, and PCNN-based region selection, rather than in the individual use of these components. Bilateral and LoG filtering process noise-sensitive and high-frequency information, respectively. The autoencoder generates a compact joint representation, while MSMG estimates activity across multiple scales. The PCNN then uses the resulting activity information to generate region-selection maps for subsequent spectral reconstruction. This ordered processing provides a clearly defined role for each component within the fusion workflow.

For the 48 selected Landsat 7 image pairs, APC-FuseNet achieved ERGAS = 2.30564, UIQI = 0.99521, CC = 0.99634, PSNR = 44.012, and SSIM = 0.92017. These results are reported alongside the corresponding literature values in Tables 1–5 to provide a comparative reference for spatial, spectral, and structural quality. Since the cited studies use different datasets, experimental settings, and evaluation protocols, the reported values are interpreted as a literature-based comparison rather than as a statistically controlled comparison across methods.

4.3 Potential Applications

APC-FuseNet is used as a preprocessing stage for remote-sensing tasks that require spatial detail together with multispectral information. The proposed model is also considered for applications such as vegetation-condition mapping, land-cover classification, urban-boundary extraction, change detection, and visual assessment of areas affected by floods, fires, or storms. The fused image combines spatial information from the panchromatic image with spectral information from the multispectral image, providing a unified representation for subsequent remote-sensing analysis.

The present experiments assess image-fusion quality using standard fusion metrics, with downstream classification, detection, and vegetation-index estimation tasks not included in the evaluation. Therefore, the effect of APC-FuseNet on these application-specific tasks is not assessed in the present study. Evaluation of these effects requires task-specific experiments using an independently defined test set.

4.4 Limitations and Open Research Issues

Since the existing studies [20-27] differ in sensor combinations, fusion objectives, datasets, and evaluation protocols, direct comparison with the proposed model is limited. Additionally, variations in source images, reference-image construction, preprocessing procedures, and test scenes affect the reported fusion metrics. Therefore, the literature values reported in Tables 1–5 are retained as contextual reference values, while direct comparison requires evaluation under a common dataset and evaluation protocol.

The current study evaluates APC-FuseNet using 48 selected image pairs from a single Landsat 7 collection. The evaluation is limited to this dataset and does not include cross-sensor or cross-region datasets, repeated-run analysis, or an independent external test set. Additionally, the individual effects of the bilateral filter, LoG decomposition, autoencoder, MSMG, and PCNN are not examined separately through ablation experiments. Moreover, training cost, inference time, memory usage, parameter sensitivity, and the complete autoencoder pretraining procedure are not reported in the current evaluation.

However, several computing-related aspects are not assessed in the present study, including fusion across different spatial-resolution ratios and the computational requirements of PCNN inference. In addition, no-reference quality assessment is not considered for cases where a reference image is unavailable. Uncertainty estimation and the effects of cloud contamination and sensor-specific artifacts are also not included in the current evaluation.

5. CONCLUSION

This research evaluated a staged autoencoder–PCNN framework for preserving complementary spatial and spectral information in Landsat 7 PAN–MS image fusion. For the 48 selected image pairs, APC-FuseNet achieved ERGAS = 2.30564, UIQI = 0.99521, CC = 0.99634, PSNR = 44.012, and SSIM = 0.92017. These results indicate low spectral distortion, high image-quality consistency, strong correlation with the reference information, and preservation of structural characteristics in the evaluated fused images.

The proposed framework integrates bilateral and LoG filtering, layer decomposition, autoencoder-based feature extraction, MSMG-based activity measurement, and PCNN-based region selection as sequential processing stages. This arrangement provides separate processing roles for noise reduction, feature representation, activity estimation, and region selection within the fusion process. The resulting fused images provide a basis for subsequent spatial–spectral remote-sensing analysis.

However, the research findings are limited to the evaluated Landsat 7 image pairs and the corresponding experimental conditions. The literature-based comparison is also influenced by differences in datasets, sensor combinations, and evaluation protocols. Additionally, component-level ablation, cross-sensor evaluation, independent external testing, computational-complexity analysis, and downstream application assessment are not considered in the present study. These findings define the scope of the study and indicate the evaluation requirements for assessing APC-FuseNet under a broader range of experimental conditions.

REFERENCES:

[1] M. Zhou, J. Huang, K. Yan, D. Hong, X. Jia, J. Chanussot, and C. Li, "A general spatial-frequency learning framework for multimodal image fusion", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 47, No. 7, 2024, pp. 5281-5298.

[2] W. Wang, L.J. Deng, R. Ran, and G. Vivone, "A general paradigm with detail-preserving conditional invertible network for image fusion", *International Journal of Computer Vision*, Vol. 132, No. 4, 2023, pp. 1029-1054.

[3] P. Wang, Z. He, B. Huang, M. Dalla Mura, H. Leung, and J. Chanussot, "VOGTNet: Variational optimization-guided two-stage network for multispectral and panchromatic image fusion", *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 36, No. 5, 2024, pp. 9268-9282.

[4] B. Yang, Z. Jiang, D. Pan, H. Yu, G. Gui, and W. Gui, "LFDT-Fusion: a latent feature-guided diffusion Transformer model for general image fusion", *Information Fusion*, Vol. 113, 2025, p. 102639.

[5] X. Cao, Y. Lian, K. Wang, C. Ma, and X. Xu, "Unsupervised hybrid network of transformer and CNN for blind hyperspectral and multispectral image fusion", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 62, 2024, p. 5507615.

[6] J. Guo, W. Zhan, Y. Jiang, W. Ge, Y. Chen, X. Xu, J. Li, and Y. Liu, "MFHOD: Multi-modal image fusion method based on the higher-order degradation model", *Expert Systems with Applications*, Vol. 249, 2024, p. 123731.

[7] Y. Jia, Q. Hu, R. Dian, J. Ma, and X. Guo, "PAPS: Progressive attention-based pan-sharpening", *IEEE/CAA Journal of Automatica Sinica*, Vol. 11, No. 2, 2024, pp. 391-404.

[8] J. Pan, Y. Bai, Q. Shu, Z. Zhang, J. Hu, and M. Wang, "M-swin: Transformer-based multi-scale feature fusion change detection network within cropland for remote sensing images", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 62, 2024, p. 4702716.

[9] X. Xie, X. Zhang, X. Tang, J. Zhao, D. Xiong, L. Ouyang, B. Yang, H. Zhou, B.W.K. Ling, and K.L. Teo, "MACTFusion: Lightweight cross transformer for adaptive multimodal medical image fusion", *IEEE Journal of Biomedical and Health Informatics*, Vol. 29, No. 5, 2024, pp. 3317-3328.

[10] L. Li, Y. Shi, M. Lv, Z. Jia, M. Liu, X. Zhao, X. Zhang, and H. Ma, "Infrared and visible image fusion via sparse representation and guided filtering in laplacian pyramid domain", *Remote Sensing*, Vol. 16, No. 20, 2024, p. 3804.

[11] W. Tang, F. He, and Y. Liu, "ITFuse: An interactive transformer for infrared and visible

image fusion", *Pattern Recognition*, Vol. 156, 2024, p. 110822.

[12] Z. Wang, Y. Ma, and Y. Zhang, "Review of pixel-level remote sensing image fusion based on deep learning", *Information Fusion*, Vol. 90, 2023, pp. 36-58.

[13] Q. Wang, X. Jin, Q. Jiang, L. Wu, Y. Zhang, and W. Zhou, „DBCT-Net: A dual branch hybrid CNN-transformer network for remote sensing image fusion", *Expert Systems with Applications*, Vol. 233, 2023, p. 120829.

[14] L. Wu, X. Jiang, Y. Yin, T.C.E. Cheng, and X. Sima, "Multi-band remote sensing image fusion based on collaborative representation", *Information Fusion*, Vol. 90, 2023, pp. 23-35.

[15] J. He, Q. Yuan, J. Li, Y. Xiao, and L. Zhang, "A self-supervised remote sensing image fusion framework with dual-stage self-learning and spectral super-resolution injection", *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 204, 2023, pp. 131-144.

[16] S. Karim, G. Tong, J. Li, A. Qadir, U. Farooq, and Y. Yu, "Current advances and future perspectives of image fusion: A comprehensive review", *Information Fusion*, Vol. 90, 2023, pp. 185-217.

[17] X. Wang, X. Wang, R. Song, X. Zhao, and K. Zhao, "MCT-Net: Multi-hierarchical cross transformer for hyperspectral and multispectral image fusion", *Knowledge-Based Systems*, Vol. 264, 2023, p. 110362.

[18] C. Zhu, S. Deng, Y. Zhou, L.J. Deng, and Q. Wu, "QIS-GAN: A lightweight adversarial network with quadtree implicit sampling for multispectral and hyperspectral image fusion", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 61, 2023, p. 5531115.

[19] X. Li, Y. Li, H. Chen, Y. Peng, and P. Pan, "CCAFusion: cross-modal coordinate attention network for infrared and visible image fusion", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 34, No. 2, 2023, pp. 866-881.

[20] J. Li, Jing, H. Hongwu, L. Lilan, and L. Jin, "Remote Sensing Image Fusion Method Based on Progressive Cascaded Deep Residual Network", *Wireless Communications and Mobile Computing*, Vol. 2023, 2023, p. 7793444.

[21] A. A. Reddy and M. Shashi, "Impact of UAV and Sentinel-2A Imagery Fusion on Vegetation Indices Performance", *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. X-1/W1-2023, ISPRS Geospatial Week Cairo, Egypt 2023, September 2–7, 2023, pp. 785–792.

[22] Z. Wang, F. Shuai, and J. Zhang, "Spatiotemporal Fusion Model of Remote Sensing Images Combining Single-Band and Multi-Band Prediction", *Remote Sensing*, Vol. 15, No. 20, 2023, p. 4936.

[23] X. Wang, R. A. Borsoi, C. Richard, J. Chen, "Deep hyperspectral and multispectral image fusion with inter-image variability", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 61, 2023, p. 5510915.

[24] S.E. Brezini and Y. Deville, "Hyperspectral and Multispectral Image Fusion with Automated Extraction of Image-Based Endmember Bundles and Sparsity-Based Unmixing to Deal with Spectral Variability", *Sensors*, Vol. 23, No. 4, 2023, p. 2341.

[25] A. A. Vakilian, M. R. Saradjian, "An object-based sparse representation model for spatiotemporal image fusion", *Scientific Reports*, Vol. 12, 2022, p. 5021.

[26] A. Azarang and N. Kehtarnavaz, "A generative model method for unsupervised multispectral image fusion in remote sensing", *Signal, Image and Video Processing*, Vol. 16, 2021, pp. 63-71.

[27] W. Tan, P. Xiang, J. Zhang, H. Zhou, and H. Qin, "Remote Sensing Image Fusion via Boundary Measured Dual-Channel PCNN in Multi-Scale Morphological Gradient Domain", *IEEE Access*, Vol. 8, 2020, pp. 42540-42549.

[28] Google Earth Engine Data Catalog, "USGS Landsat 7 Collection 2 Tier 1 TOA Reflectance," dataset ID LANDSAT/LE07/C02/T1_TOA, available at: https://developers.google.com/earth-engine/datasets/catalog/LANDSAT_LE07_C02_T1_TOA (accessed August 8, 2026).