Scientific Research Publishing

Open Journal of Applied Sciences, 2026, 16(5), 1940-1965
https://www.scirp.org/journal/ojapps
ISSN Online: 2165-3925
ISSN Print: 2165-3917

An Attention-Based Improved Laplacian Pyramid Network for Remote Sensing Image Pansharpening

Jiazheng Zhou

School of Electronic Information and Artificial Intelligence, Shaanxi University of Science & Technology, Xi'an, China
Email: 1985294385@qq.com

How to cite this paper: Zhou, J.Z. (2026) An Attention-Based Improved Laplacian Pyramid Network for Remote Sensing Image Pansharpening. *Open Journal of Applied Sciences*, **16**, 1940-1965.
https://doi.org/10.4236/ojapps.2026.165107

Received: May 9, 2026
Accepted: May 24, 2026
Published: May 27, 2026

Abstract

With the rapid development of Earth observation technology, remote sening images have been widely used in land resource investigation, environmental monitoring, precision agriculture, urban planning, and disaster assessment. However, due to sensor imaging mechanisms, signal to noise ratio limitations, and data transmission constraints, it is difficult for a single satellite sensor to obtain images with both high spatial resolution and rich spectral information. Panchromatic (PAN) images usually provide spatial details and clear texture structures, while multispectral (MS) images contain richer spectral information but have lower spatial resolution. Pansharpening aims to fuse PAN and MS images to generate high resolution multispectral images, which is important for interpretation and analysis tasks. Existing pansharpening methods mainly include traditional and deep learning based approaches. Traditional methods, such as component substitution, multi resolution analysis, and model based optimization, are simple and efficient, but they often suffer from spectral distortion or insufficient spatial detail recovery. In recent years, convolutional neural networks have shown strong representation ability in pansharpening tasks. Nevertheless, many existing deep learning based methods still have limitations in multi scale feature utilization and spectral fidelity preservation. Spatial details are distributed at different resolution levels, and single scale feature extraction may fail to recover high frequency textures. In addition, different spectral channels contribute unequally to fusion, while pixel wise losses cannot directly constrain spectral direction consistency. To address these problems, this paper proposes an attention based improved Laplacian pyramid network for remote sensing image pansharpening. The proposed method uses an MTF based Laplacian pyramid to decompose PAN and upsampled MS images into multi scale components. An SEFCNN adaptive fusion module is introduced at each pyramid level to enhance important spatial

spectral features through channel attention. Furthermore, a hybrid loss function combining multi scale reconstruction loss and spectral angle mapper loss is designed to improve spectral preservation. Experiments on QuickBird and WorldView-3 datasets are conducted to verify the effectiveness of the proposed method.

Keywords

Pansharpening, Laplacian Pyramid, Squeeze and Excitation Attention, Remote Sensing Images, Convolutional Neural Networks, Spectral Angle Mapper, Pan Chromatic, Multispectral

1. Introduction

With the continuous development of Earth observation technology, remote sensing images have been widely used in environmental monitoring, land resource investigation, precision agriculture, urban planning, disaster warning, ecological protection, and national defense security [1] [2]. In remote sensing applications, spatial and spectral resolutions are two important factors affecting image interpretation and analysis. However, due to sensor imaging mechanisms, optical energy collection, signal to noise ratio, data storage, and transmission cost, it is difficult for a single satellite sensor to acquire images with both high spatial resolution and rich spectral information. PAN images usually provide fine spatial structures such as edges, textures, and object boundaries, but their spectral information is limited. In contrast, low resolution multispectral (LRMS) images contain richer spectral responses across different bands, but their spatial details are relatively insufficient. Pansharpening aims to fuse the spatial information of PAN images and the spectral information of LRMS images to generate high resolution multispectral images, thereby providing more reliable data for subsequent remote sensing tasks.

Existing pansharpening methods can be broadly divided into traditional methods and deep learning based methods. Traditional methods, including component substitution, multi resolution analysis, and model-based optimization, are simple and efficient, and have been widely studied. Nevertheless, these methods often rely on fixed transformation rules or manually designed injection models, which may lead to spectral distortion or insufficient spatial detail recovery in complex scenes. In recent years, convolutional neural networks have been introduced into pansharpening and have shown stronger nonlinear representation ability than traditional approaches. Representative methods such as PNN, PanNet, MSDCNN, and PanGAN have improved fusion performance from different perspectives. Attention mechanisms and Transformer based structures have also been applied to enhance spatial spectral feature representation. However, existing deep learning based methods still have limitations in multi scale feature utilization, channel wise feature selection, and spectral fidelity preservation. Spatial details are distributed

across different scales, while different spectral channels contribute unequally to the final fusion result. Moreover, conventional pixel wise reconstruction losses mainly constrain nuSpecial description of the title. (dispensable) merical errors and cannot directly preserve the spectral direction consistency of fused images [3].

To address these problems, this paper proposes an attention based improved Laplacian pyramid network for remote sensing image pansharpening. The proposed method uses an MTF-based Laplacian pyramid to decompose PAN and upsampled MS images into multi scale components, introduces an SEFCNN adaptive fusion module to enhance important channel features, and combines multi scale reconstruction loss with spectral angle mapper loss to improve spatial detail recovery and spectral preservation. The main contributions of this study are summarized as follows. First, an MTF based Laplacian pyramid decomposition and progressive reconstruction framework is designed to extract and fuse spatial spectral information at different scales. Second, an SEFCNN adaptive fusion module is introduced to strengthen useful channel features and improve feature fusion capability. Third, a hybrid loss function combining multi scale reconstruction loss and SAM loss is constructed to enhance spectral consistency while maintaining spatial details [4] [5]. Experiments on QuickBird and WorldView-3 datasets demonstrate that the proposed method achieves favorable performance under both reduced resolution and full resolution conditions.

2. Materials and Methods

2.1. Datasets and Experimental Protocol

The proposed method was evaluated on two commonly used high resolution remote sensing datasets, namely QuickBird and WorldView-3. These two datasets provide paired panchromatic (PAN)and multispectral (MS) images with different spatial and spectral resolutions, and are therefore suitable for pansharpening research [6] [7]. The QuickBird dataset contains one PAN band and four MS bands. The spatial resolution of the PAN image is 0.6 m, while that of the low resolution multispectral (LRMS)image is 2.4 m. The WorldView-3 dataset contains one PAN band and eight MS bands, with spatial resolutions of 0.3 m and 1.2 m for PAN and LRMS images, respectively. Representative samples of the two datasets are presented to show their spatial structures and spectral characteristics (**Figure 1**). In this study, the QuickBird and WorldView-3 datasets contain 10,923 and 8806 PAN LRMS image pairs, respectively. For each dataset, 90% of the image pairs were used for training and the remaining 10% were used for testing (see **Table 1**).

Since ideal high resolution multispectral reference images are usually unavailable in real satellite imaging, both reduced resolution and full resolution experiments were conducted. The reduced resolution experiments were constructed according to Wald's protocol [8]. Specifically, the original PAN and MS images were degraded to generate simulated PAN, LRMS, and reference images. Under this

setting, the original MS images could be used as reference images, making it possible to evaluate the fusion results using full-reference quantitative metrics. The full resolution experiments directly used the original satellite images as inputs, which is closer to practical remote sensing applications. However, because no real HRMS reference image is available in this case, visual comparison and no reference metrics were adopted for evaluation [9].

The network was implemented using the TensorFlow deep learning framework and trained on an NVIDIA GeForce RTX4070 GPU. The Adam optimizer was used for parameter optimization. The initial learning rate was set to 0.0001, and both the training batch size and validation batch size were set to 32. To provide a comprehensive comparison, eight representative pansharpening methods were selected as baselines, including IHS, GS, Wavelet, MTF-GLP, PNN, MSDCNN, LPPN, and PanGAN. These methods cover component substitution based, multi resolution analysis based, and deep learning based categories, allowing the proposed method to be compared with both traditional and learning based approaches.

Table 1. Dataset and experimental settings.

Dataset	PAN Bands	MS Bands	PAN Resolution	LRMS Resolution	Image Pairs	Training/Testing Split
Quickbird	1	4	0.6 m	2.4 m	10923	90%/10%
WorldVie-3	1	8	0.3 m	1.2 m	8806	90%/10%

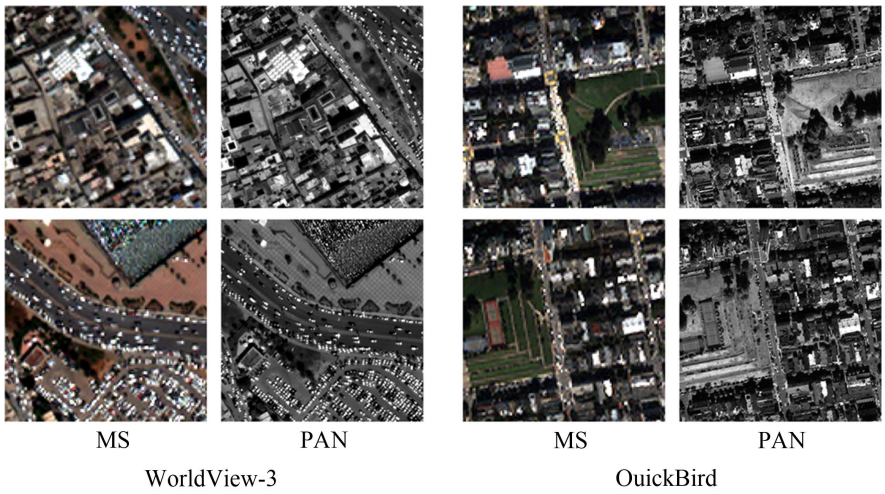

Figure 1. QuickBird and WorldView-3 benchmark datasets.

2.2. Overall Framework of the Proposed Network

This paper proposes an attention based improved Laplacian pyramid network, referred to as AILPPN, for remote sensing image pansharpening. The proposed network aims to address two main limitations of existing deep learning based methods: insufficient use of multi scale spatial features and inadequate preservation of spectral consistency [10]. The overall framework takes a low resolution multispectral (LRMS) image and a high resolution panchromatic (PAN) image as inputs.

First, the LRMS image is upsampled by a factor of four so that it has the same spatial size as the PAN image. Then, the PAN image and the upsampled MS image are decomposed into multi scale components through an MTF based Laplacian pyramid [11]. The corresponding components at each scale are fused by adaptive fusion subnetworks, and the final high resolution multispectral image is progressively reconstructed from coarse to fine scales (**Figure 2**).

The proposed AILPPN mainly consists of four parts: MTF based Laplacian pyramid decomposition, SEFCNN adaptive fusion, progressive reconstruction, and hybrid loss optimization. The pyramid decomposition module separates low frequency structural information and high frequency detail information at different resolutions. The SEFCNN module extracts and fuses PAN and MS features at each pyramid level by combining convolution, residual learning, recursive feature refinement, and channel attention. The progressive reconstruction module transfers information from lower resolution levels to higher resolution levels through upsampling and crossscale addition [12]. Finally, the hybrid loss function combines multi scale reconstruction loss and spectral angle mapper (SAM) loss to improve both spatial detail recovery and spectral preservation.

Compared with direct single scale fusion, the proposed framework has clearer multi scale representation and stronger adaptive feature selection ability. Spatial details such as edges, textures, and object boundaries can be learned at different pyramid levels, while the squeeze and excitation mechanism helps the model emphasize useful spectral spatial channels and suppress redundant responses. Therefore, AILPPN is expected to achieve a better balance between spatial enhancement and spectral fidelity.

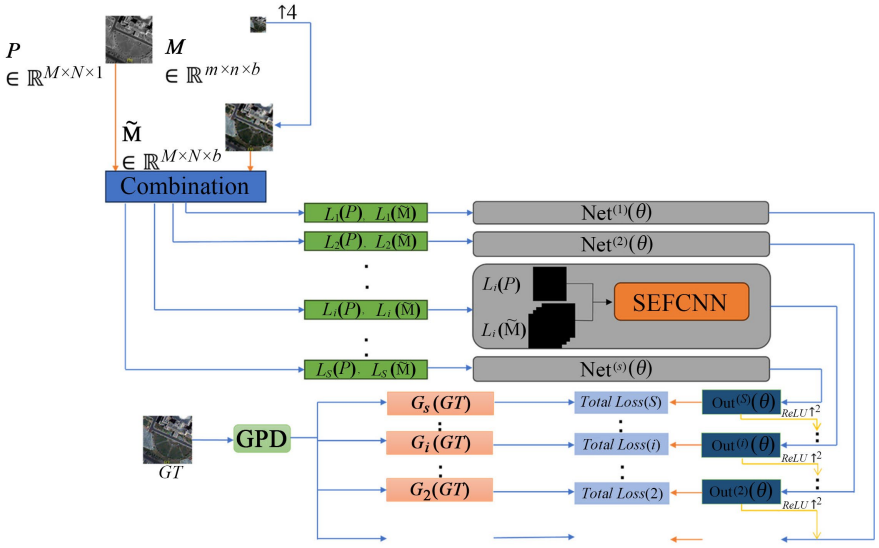

Figure 2. Diagram of the AILPPN structure.

2.3. MTF-Based Laplacian Pyramid Decomposition

The Laplacian pyramid is a classical multi scale image representation method and

has been widely used in image compression, image enhancement, image fusion, and super resolution reconstruction [13]. In pansharpening, spatial details from PAN images and spectral structures from MS images are not concentrated in a single frequency band. Edges, textures, and local structures are mainly represented as high frequency information, while scene contours and global illumination variations are related to low frequency components. If the network only performs feature extraction at one scale, it may fail to simultaneously represent global structures and local details. Therefore, multi scale pyramid decomposition is introduced to provide hierarchical inputs for subsequent feature learning.

The proposed network adopts an MTF based Laplacian pyramid rather than a standard Gaussian kernel based pyramid. The modulation transfer function (MTF) describes the response characteristics of remote sensing sensors to spatial frequency information. Since real satellite images are generated through sensor specific imaging processes, using an MTF kernel is more consistent with the degradation characteristics of remote sensing images than using a fixed Gaussian kernel. In this way, the decomposed components can better reflect the actual relationship between PAN and MS observations, which is beneficial for subsequent feature fusion and reconstruction.

In the proposed network, the PAN image and the upsampled MS image are decomposed separately through an MTF-based Laplacian pyramid. At each pyramid level, low pass filtering and downsampling are first applied to obtain a lower resolution representation. Then, the lower level image is upsampled and subtracted from the current level image to generate the corresponding high frequency residual component [14] [15]. In this way, each Laplacian pyramid level records the information that cannot be represented by the coarser scale, mainly including edges, textures, and local spatial details. In this study, a five level pyramid structure is adopted to provide sufficient multi scale representation for the pansharpening task and to maintain consistency with the overall progressive reconstruction process (**Figure 3**).

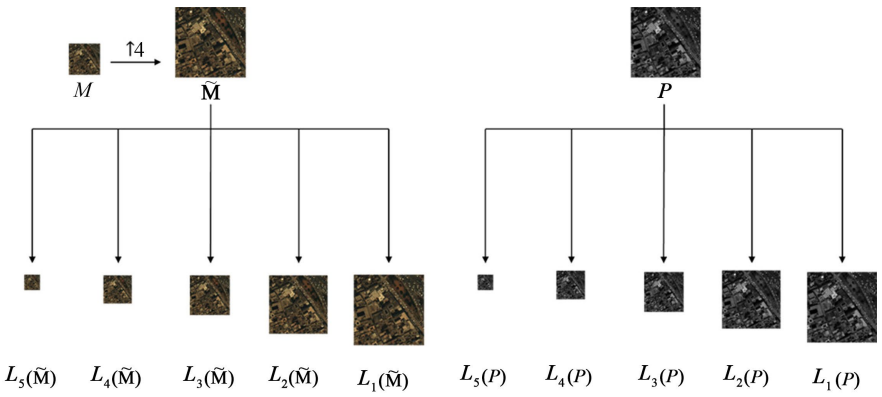

Figure 3. Diagram of Laplacian pyramid decomposition.

Let the Laplacian components of the PAN image and the upsampled MS image at the i-th scale be denoted as L_{PAN}^{i} and L_{MS}^{i}, respectively. At each scale, these

two components are concatenated along the channel dimension and then used as the input of the corresponding fusion subnetwork. This operation preserves the complementary relationship between the high resolution spatial information of the PAN image and the spectral information of the MS image. Through this decomposition strategy, the subsequent SEFCNN module can learn spatial spectral features at different scales, rather than performing fusion only at the original resolution. This helps reduce information mixing and improves the stability of multi scale feature fusion.

2.4. SEFCNN Adaptive Fusion Module

After pyramid decomposition, each scale contains one PAN component and one corresponding MS component. To effectively fuse them, this study designs an SEFCNN adaptive fusion module based on the original adaptive fusion idea of the Laplacian pyramid network. In the original structure, the PAN and MS components at each scale are concatenated and processed by convolutional layers. Residual learning and recursive blocks are then used to enhance feature representation while reducing parameter redundancy [16]. However, ordinary convolution treats all feature channels in a relatively uniform way, which may be insufficient for pansharpening because different channels contain different amounts of spatial and spectral information. Some channels may contribute more to edge restoration, while others may be more important for spectral preservation. Therefore, this study introduces the squeeze and excitation (SE) attention mechanism into the fusion module.

The SEFCNN module first receives the concatenated PAN MS component at the current pyramid scale. A convolutional layer with ReLU activation is used to extract initial features. These features contain mixed spatial and spectral responses [17]. Then, the SE attention module is embedded into the fusion process to adaptively recalibrate channel importance. Specifically, global average pooling is applied to each channel of the feature map to obtain a compact channel descriptor. This descriptor reflects the global response of each channel and provides statistical information for channel selection. After that, two fully connected layers are used to learn nonlinear channel dependencies, and a sigmoid activation function maps the learned responses into channel weights between 0 and 1. The generated weights are multiplied with the original feature maps, so that informative channels are enhanced and less useful channels are suppressed (**Figure 4** and **Figure 5**).

This channel wise recalibration is particularly useful for remote sensing image pansharpening. PAN images mainly provide spatial details, while MS images contain spectral information across different bands. During fusion, simply concatenating these components may introduce redundant responses and noise. The SE mechanism allows the network to learn which channels should receive more attention under a specific scene and scale. For example, channels related to strong edges, vegetation spectral responses, or building textures can be assigned larger weights, while channels carrying weak or noisy responses can be suppressed. In

this way, the SEFCNN module improves the selectivity of feature fusion and makes the network more capable of modeling key spatial spectral information.

After SE attention enhancement, the recalibrated features are further processed by recursive blocks [18]. The recursive block structure shares parameters across repeated feature extraction operations, which helps deepen the effective network while controlling the number of trainable parameters. Each recursive block contains convolutional operations and ReLU activation. A residual connection is preserved between the enhanced initial feature and the output of the recursive mapping. This design facilitates gradient propagation during training and helps maintain useful low level information. The output of the recursive blocks is taken as the fusion feature of the current scale.

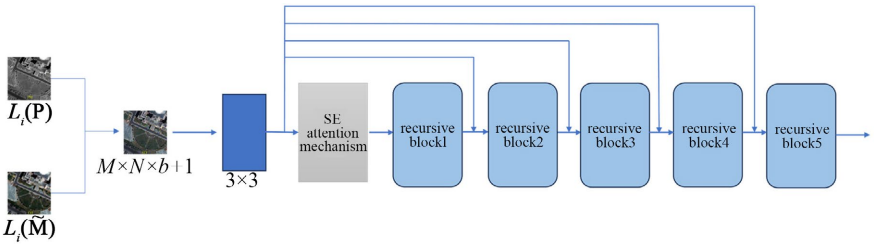

Figure 4. Architecture of the SEFCNN adaptive fusion module.

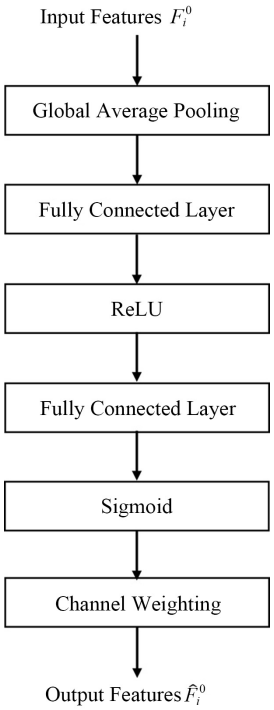

Figure 5. Architecture of the attention mechanism.

2.5. Progressive Reconstruction and Multi-Scale Output

After feature extraction and adaptive fusion at different pyramid levels, the network reconstructs the final high resolution multispectral image. Instead of directly

predicting the final result at the highest resolution, the proposed model adopts a progressive reconstruction strategy [19]. This strategy follows the basic idea of the Laplacian pyramid: coarse structural information is first recovered at low resolution levels, and spatial details are gradually supplemented at higher resolution levels. By transferring information across scales, the network reduces the difficulty of direct high resolution reconstruction and makes better use of multi scale fusion features.

At the lowest resolution level, the output is generated directly from the fusion feature produced by the SEFCNN module and activated by ReLU. From the second scale onward, the output of the previous scale is first upsampled and then added to the fusion feature of the current scale. The combined feature is then passed through ReLU activation to generate the current scale output. This process is repeated until the highest-resolution scale is reached. The final output at the highest resolution scale is regarded as the pansharpened HRMS image (**Figure 6**).

The progressive reconstruction strategy has two main benefits. First, the lower scale output contains relatively stable global structures and low frequency information. After upsampling, this information guides finer scale reconstruction, preventing the high resolution prediction from relying only on local texture cues. Second, high frequency details are gradually refined at each scale, which suits the hierarchical nature of remote sensing images. In complex scenes such as urban areas, roads, vegetation, and water bodies, spatial structures appear at multiple scales. Progressive reconstruction can recover these structures more smoothly and reduce detail loss caused by direct singlestep prediction. In this study, five scale outputs are generated. These outputs form a coarse to fine reconstruction chain and participate in multi scale loss calculation during training, improving optimization stability.

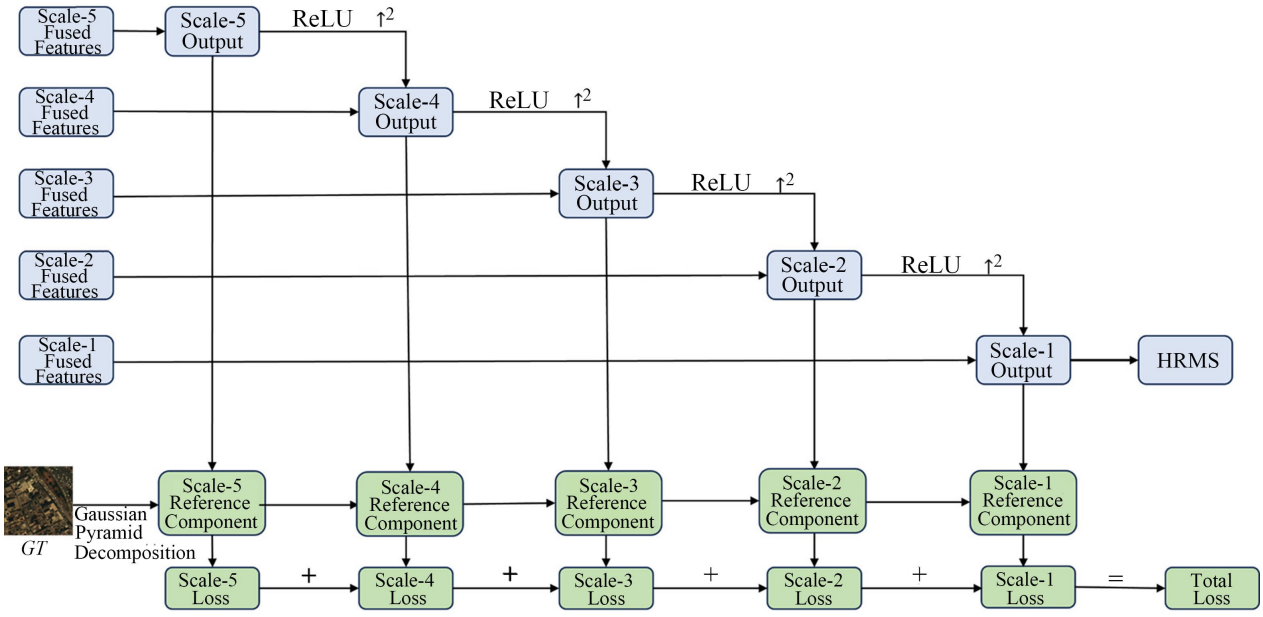

Figure 6. Architecture of the progressive reconstruction and multi-scale output module.

2.6. Hybrid Loss Function

The quality of pansharpening depends on both spatial detail reconstruction and spectral fidelity. An ideal fused high resolution multispectral image should contain clear spatial structures from the PAN image and reliable spectral information from the original MS image. If the loss function only focuses on pixel level error, the network may generate visually sharp results but still produce spectral distortion. Conversely, if it only emphasizes spectral preservation, the fused image may lose high frequency details [9]. Therefore, an effective training objective should consider both spatial reconstruction accuracy and spectral consistency.

Many image reconstruction networks use pixel wise losses, such as mean absolute error or mean squared error, to measure differences between predicted and reference images. However, in remote sensing pansharpening, pixel wise error alone is insufficient. Multispectral images contain several bands, and their relative relationships are important for land cover interpretation, vegetation analysis, object recognition, and other applications. Even with small band wise errors, the spectral direction of a pixel may change, causing spectral distortion. To address this problem, this study combines multi scale reconstruction loss and spectral angle mapper loss.

The first component of the proposed hybrid loss function is the multi scale reconstruction loss. Since AILPPN adopts a Laplacian pyramid structure and generates outputs at different levels, the network should not be supervised only at the final output level. Intermediate outputs should also be constrained. This design is consistent with progressive reconstruction. Lower scale outputs describe coarse structural and low frequency information, while higher scale outputs contain detailed textures and high frequency information. By applying multi scale supervision, the network can learn a stable coarse to fine reconstruction process.

In this study, the reconstruction loss at each scale is defined as an L1 loss. Compared with L2 loss, L1 loss is less sensitive to large local errors and is more suitable for preserving image edges and local details. Let $\hat{Y}^i$ denote the predicted output at the i-th pyramid scale, and Y^i denote the corresponding reference image or reference component at the same scale. The reconstruction loss at the i-th scale is defined as:

$$L_{rec}^{i} = \frac{1}{N_i} \sum_{p=1}^{N_i} \left| \hat{Y}_p^i - Y_p^i \right| \tag{1}$$

where N_i denotes the total number of pixels at the i-th scale, and p represents the pixel index. $\hat{Y}_p^i$ and Y_p^i represent the pixel values of the predicted image and the reference image at pixel p of the i-th scale, respectively. This formula means that the absolute difference between the predicted value and the reference value is calculated for all pixels at the current scale and then averaged. In this way, each pyramid output is directly optimized toward its corresponding reference target.

The total multi scale reconstruction loss is calculated by combining the reconstruction losses from all pyramid scales:

$$L_{ms} = \sum_{i=1}^{S} \alpha_i L_{rec}^{i} \tag{2}$$

where S is the total number of pyramid scales, and α_i is the weight of the reconstruction loss at the i-th scale. The coefficient α_i is used to control the relative importance of different pyramid levels. In general, the final high-resolution output should receive more attention, while intermediate outputs provide auxiliary supervision to improve the stability of training. Through this multi scale reconstruction constraint, the network can recover both global structures and fine local details, instead of relying only on the final output supervision.

The second component of the hybrid loss function is the spectral angle mapper loss. SAM is commonly used in remote sensing image evaluation because it measures the angle between two spectral vectors. Unlike pixel wise losses, SAM focuses on spectral direction rather than absolute numerical difference, making it suitable for constraining spectral consistency in multispectral image fusion [5] [12]. In pansharpening, each pixel can be regarded as a spectral vector composed of multiple bands. A smaller angle between the predicted and reference spectral vectors indicates better preservation of spectral characteristics.

For each pixel position p, the spectral vectors of the predicted HRMS image and the reference HRMS image are denoted as $\hat{y}_p$ and y_p, respectively. The SAM loss is defined as:

$$L_{SAM} = \frac{1}{N} \sum_{p=1}^{N} \arccos\left(\frac{\hat{y}_p^{\mathrm{T}} y_p}{\left\| \hat{y}_p \right\|_2 \left\| y_p \right\|_2 + \varepsilon} \right) \tag{3}$$

where N is the total number of pixels in the final fused image, and ε is a small constant used to avoid division by zero. The numerator represents the inner product between the predicted spectral vector and the reference spectral vector. The denominator represents the product of their Euclidean norms. Therefore, this formula calculates the spectral angle between the two vectors at each pixel and then averages the angle values over the whole image. A smaller SAM loss indicates that the predicted image has better spectral direction consistency with the reference image.

The final training objective is obtained by combining the multi scale reconstruction loss and the SAM loss:

$$L_{total} = L_{ms} + \lambda L_{SAM} \tag{4}$$

where λ is the balance coefficient used to control the contribution of SAM loss. The multi scale reconstruction loss mainly constrains the spatial and numerical reconstruction accuracy of the fused image, while the SAM loss directly constrains the spectral direction consistency. These two components are complementary. If λ is too small, the model may mainly optimize pixel level reconstruction and still produce spectral distortion. If λ is too large, the network may overemphasize spectral angle consistency and weaken spatial detail recovery. Therefore, the value of λ should be selected carefully according to experimental

performance.

2.7. Evaluation Metrics

To evaluate the proposed method comprehensively, both subjective visual comparison and objective quantitative assessment are adopted. Visual comparison focuses on whether the fused image can preserve natural color, sharp object boundaries, clear road structures, vegetation textures, and local spatial details without obvious artifacts or spectral distortion. Objective quantitative assessment provides more repeatable and comparable measurements of fusion quality.

For reduced resolution experiments, full reference evaluation metrics are used because the reference HRMS image is available under Wald's protocol. The selected metrics include peak signal to noise ratio (PSNR), structural similarity index (SSIM), correlation coefficient (CC), spectral angle mapper (SAM), relative dimensionless global error in synthesis (ERGAS), Q4. PSNR measures the global reconstruction error between the fused image and the reference image. SSIM evaluates structural similarity from luminance, contrast, and structure. CC reflects the linear correlation between the fused image and the reference image. SAM measures spectral angle distortion and is directly related to spectral fidelity. ERGAS evaluates global relative synthesis error across multiple bands. Q4 is used as a comprehensive multi band quality metric. Larger values of PSNR, SSIM, CC, and Q4 indicate better image quality, while smaller values of SAM and ERGAS indicate less spectral distortion and lower global error.

For full resolution experiments, no real HRMS reference image is available. Therefore, no reference metrics are adopted, including QNR and its two distortion components, spectral distortion and spatial distortion. QNR evaluates the overall quality of the fused image by jointly considering spectral and spatial distortions. A larger QNR value indicates better fusion quality, while smaller distortion components indicate less degradation. By using different metrics under reduced resolution and full resolution conditions, the proposed AILPPN can be evaluated more comprehensively in both simulated and real resolution scenarios.

2.8. Comparative Methods and Reproducibility Considerations

The comparison methods were selected to cover several representative categories of pansharpening algorithms. IHS and GS represent component substitution methods, which are simple and efficient but may introduce spectral distortion when the spectral response of PAN is inconsistent with that of MS bands. Wavelet and MTF-GLP represent multi resolution analysis methods, which inject spatial detail information extracted from the PAN image into the MS image. These methods usually preserve spectral information better than direct component substitution methods, but their performance depends strongly on filter design and detail injection rules. PNN, MSDCNN, LPPN, and PanGAN represent deep learning based methods. PNN verifies the feasibility of convolutional neural networks for end to end pansharpening, MSDCNN improves multi scale feature extraction by using convolutional kernels of different sizes, LPPN introduces a Laplacian pyramid reconstruction

framework, and PanGAN uses adversarial learning to improve visual quality. By comparing with these methods, the effectiveness of AILPPN can be examined in relation to both traditional fusion strategies and deep learning based models.

For fair comparison, the same training and testing partitions are used for the proposed method and learning based comparison methods when applicable. The input LRMS image is upsampled to the PAN size before being fed into the network. During training, the network parameters are updated by minimizing the hybrid loss function described above. During inference, the trained network directly takes the PAN image and the upsampled MS image as inputs and outputs the final HRMS image. The evaluation is then performed using the same metrics for all methods. This setting helps ensure that performance differences mainly come from the model structure and loss function design rather than from inconsistent data splits or evaluation procedures.

The overall methodology can therefore be summarized as a closed process consisting of data preparation, multi scale decomposition, channel attentionbased adaptive fusion, progressive reconstruction, hybrid loss optimization, and multi level evaluation. The MTF based Laplacian pyramid provides the multiscale representation, the SEFCNN module enhances channel wise feature selectivity, and SAM loss further constrains spectral direction consistency. These components are designed to work together rather than independently. The decomposition stage provides suitable hierarchical inputs, the attention module improves the quality of feature fusion at each scale, the progressive reconstruction module transfers information across scales, and the hybrid loss guides the model to maintain both spatial details and spectral fidelity. This integrated design forms the methodological basis of the proposed AILPPN.

3. Results

To evaluate the effectiveness of the proposed attention-based Laplacian pyramid pansharpening network, experiments were conducted on both QuickBird and WorldView-3 datasets under reduced-resolution and full resolution conditions. In the reduced-resolution experiments, the original multispectral images were used as reference images, and the fusion performance was evaluated using full-reference quality indices, including SSIM, PSNR, SAM, CC, ERGAS, Q4, and Q8. In the full resolution experiments, since ideal high resolution multispectral reference images were unavailable, the non-reference indices Dλ, Ds, and QNR were used to evaluate spectral distortion, spatial distortion, and overall fusion quality. The proposed method was compared with several representative pansharpening methods, including IHS, GS, Wavelet, MTF-GLP, PNN, MSDCNN, LPPN, and PanGAN.

3.1. Reduced-Resolution Experimental Results

The visual comparison results on the reduced-resolution QuickBird dataset indicate that different methods present obvious differences in spatial detail recovery and spectral information preservation (**Figure 7**). Traditional methods, including

IHS, GS, Wavelet, and MTF-GLP, can improve the spatial resolution to some extent, but they also introduce spectral distortion or insufficient detail recovery. In particular, IHS and GS tend to produce noticeable color deviation because they inject spatial details by replacing or modifying intensity related components. Wavelet and MTF-GLP preserve part of the spectral information, but the reconstructed images still appear blurred in fine textures and edge regions.

Compared with traditional methods, deep learning based methods achieve better visual quality. PNN, MSDCNN, LPPN, and PanGAN recover more spatial structures and generate images with more natural color distributions. Among them, LPPN and PanGAN show relatively strong fusion performance. However, LPPN still suffers from partial loss of spectral information in complex land-cover regions, while PanGAN produces favorable overall visual effects but does not always preserve fine structural details consistently. In contrast, the proposed AILPPN method restores clearer object boundaries, road textures, and vegetation details, while maintaining a color appearance closer to the reference image. This indicates that the combination of Laplacian pyramid decomposition, SEFCNN adaptive fusion, and SAM spectral constraint is beneficial for balancing spatial enhancement and spectral fidelity.

The quantitative results on the reduced-resolution QuickBird dataset further support the visual comparison (see **Table 2**). The proposed AILPPN obtains the best values in SSIM, PSNR, SAM, CC, and ERGAS. Specifically, it achieves an SSIM of 0.9225, PSNR of 32.2128, SAM of 0.0311, CC of 0.9298, and ERGAS of 2.0189. These results show that AILPPN not only improves spatial structure similarity and pixel level reconstruction accuracy, but also reduces spectral angle distortion and global relative error. Compared with LPPN, the proposed method improves SSIM from 0.8923 to 0.9225 and reduces ERGAS from 2.2264 to 2.0189. Compared with PanGAN [19], AILPPN obtains better results in most full reference indices, although its Q4 value is slightly lower than that of PanGAN. Overall, the reduced-resolution QuickBird results demonstrate that AILPPN has stronger comprehensive fusion ability.

The visual comparison results on the reduced-resolution WorldView-3 dataset show that the fusion difficulty increases when more multispectral bands are involved (**Figure 8**). Traditional methods still show obvious limitations. IHS, GS, Wavelet, and MTF-GLP generate fused images with relatively weak spatial detail recovery or noticeable spectral distortion. PNN preserves part of the spectral information, but its sharpening effect is insufficient, and some fine textures remain unclear. MSDCNN improves the overall reconstruction quality, but the fused image still appears relatively smooth in detailed regions. LPPN and PanGAN produce better visual results than traditional methods, but they still show different degrees of spatial information loss.

By comparison, the proposed AILPPN method presents clearer spatial structures and more stable spectral information. The edges of buildings, roads, and other ground objects are sharper, while the overall color distribution remains

close to the reference image. This suggests that the proposed multi-scale reconstruction strategy is effective for extracting spatial details from PAN images, and the SAM loss helps maintain the spectral relationship among multispectral bands.

The quantitative comparison on the reduced-resolution WorldView-3 dataset shows that AILPPN achieves the best results in most metrics (see Table 3). The proposed method obtains an SSIM of 0.9169, PSNR of 32.4928, SAM of 0.0318, CC of 0.9328, and ERGAS of 2.0189. Compared with PanGAN, the proposed method improves SSIM from 0.9021 to 0.9169 and PSNR from 32.1703 to 32.4928. It also reduces SAM from 0.0392 to 0.0318, indicating better spectral consistency. Although the Q4 value ofAILPPN is slightly lower than that of PanGAN, the proposed method performs better in the other five indices. Therefore, the results on WorldView-3 further confirm that AILPPN can effectively improve spatial resolution while maintaining spectral fidelity.

Table 2. Comparison of results of different algorithms on reduced-resolution QuickBird dataset.

algorithm	SSIM	PSNR	SAM	CC	ERGAS	Q4
IHS	0.6293	25.4355	0.1348	0.7834	5.2334	0.7239
GS	0.6372	25.5342	0.1332	0.7932	4.9324	0.7382
Wavelet	0.5732	24.9843	0.1374	0.7832	5.2355	0.7192
MTF-GLP	0.6737	26.3452	0.1294	0.8373	4.9345	0.7726
PNN	0.8327	30.2854	0.0375	0.8922	2.3544	0.8242
MSDCNN	0.8745	30.2384	0.0427	0.8929	2.5219	0.8929
LPPN	0.8923	30.8549	0.0391	0.9134	2.2264	0.9135
PanGAN	0.9175	31.8433	0.0382	0.9236	2.1938	**0.9412**
AILPPN	**0.9225**	**32.2128**	**0.0311**	**0.9298**	**2.0189**	0.9323

Table 3. Comparison of results of different algorithms on reduced-resolution WorldView-3 dataset.

algorithm	SSIM	PSNR	SAM	CC	ERGAS	Q8
IHS	0.5694	22.9238	0.1482	0.7829	5.3224	0.8844
GS	0.5734	23.4235	0.1435	0.7230	5.3945	0.8834
Wavelet	0.4934	22.7926	0.1323	0.7732	5.0183	0.8734
MTF-GLP	0.5737	24.0923	0.1453	0.8293	5.0345	0.8926
PNN	0.8583	29.3245	0.0477	0.8962	2.7329	0.9032
MSDCNN	0.8927	30.0121	0.0502	0.8833	2.7934	0.9243
LPPN	0.8978	31.0117	0.0402	0.9135	2.3179	0.9214
PanGAN	0.9021	32.1703	0.0392	0.9138	2.1343	**0.9439**
AILPPN	**0.9169**	**32.4928**	**0.0318**	**0.9328**	**2.0189**	0.9365

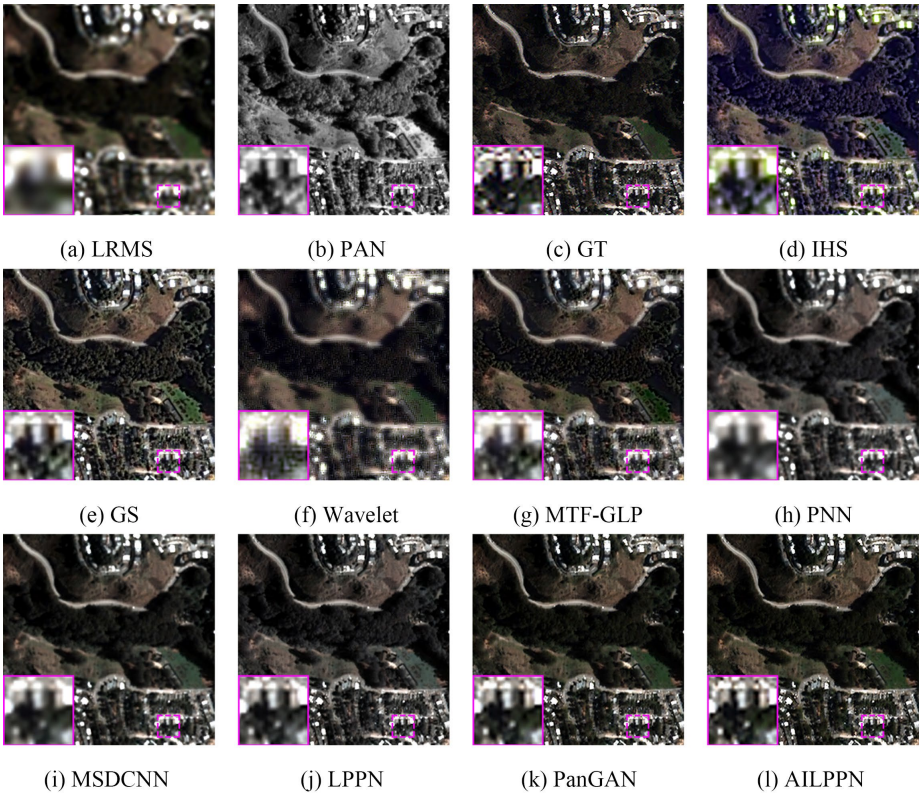

Figure 7. Sharpening results at reduced-resolution on QuickBird dataset.

Figure 8. Sharpening results at reduced-resolution on WorldView-3 dataset.

3.2. Full-Resolution Experimental Results

To further evaluate the practical applicability and robustness of the proposed method, full resolution experiments were conducted on the original QuickBird and WorldView-3 images. Unlike reduced resolution experiments, full resolution evaluation does not have ideal high-resolution multispectral reference images. Therefore, visual quality and non reference quantitative indices were jointly used for evaluation. This setting is closer to real satellite imaging conditions and can better reflect the practical fusion ability of different methods.

The full resolution pansharpening results on the QuickBird dataset show that traditional methods still have limitations in real-resolution fusion conditions (**Figure 9**). IHS and GS introduce visible color distortion, while Wavelet and MTF-GLP fail to recover sufficient high frequency details in some regions. For deep learning based methods, PNN produces relatively pale images and loses part of the spectral characteristics. MSDCNN and LPPN improve the overall visual quality, but some local structures remain blurred. PanGAN shows competitive results and can preserve main ground object structures, but some subtle details, especially edges and fine textures in complex areas, are still not fully reconstructed.

The proposed AILPPN method achieves better visual performance. The fused image contains clearer edge structures and richer spatial textures, while the color distribution remains relatively natural. In particular, roads, buildings, and vegetation areas are better distinguished, and the boundaries between different ground objects are more complete. This indicates that the proposed progressive reconstruction strategy can effectively transmit multi scale spatial details to the final high resolution output, while the SEFCNN module helps the network select more useful channel features during fusion and reduce redundant feature interference.

The quantitative results on the full-resolution QuickBird dataset also confirm the advantage of the proposed method (see **Table 4**). AILPPN achieves the lowest Dλ value of 0.0051 and the highest QNR value of 0.9520. Its Ds value is 0.0431, which is slightly higher than that of PanGAN but still better than most comparison methods. Since Dλ reflects spectral distortion and Ds reflects spatial distortion, the results indicate that AILPPN has strong spectral preservation ability and competitive spatial quality. The QNR value further demonstrates that the proposed method obtains the best overall non reference fusion quality on the QuickBird dataset, suggesting a better balance between spatial enhancement and spectral consistency.

The full-resolution results on the WorldView-3 dataset present a similar trend (**Figure 10**). Traditional methods exhibit noticeable spectral distortion and texture blurring. In several local regions, the edges of buildings and roads are not sufficiently sharp, and the color distribution is less consistent with the original multi spectral image. Among the deep learning based comparison methods, PanGAN and LPPN achieve better fusion results than PNN and MSDCNN, but there is still some information loss in edge details and complex spatial structures, especially in areas with dense buildings and mixed land cover types.

The proposed AILPPN method generates fused images with clearer spatial details and more stable spectral information. The visual result is closer to the expected high resolution multispectral image, especially in areas containing buildings, roads, and mixed land cover types. Compared with the other methods, AILPPN better preserves the continuity of object boundaries and reduces local blurring in complex spatial regions. This result suggests that the proposed network is not only effective under simulated reduced resolution conditions, but also has good adaptability to real full resolution remote sensing scenes.

The quantitative comparison on the full resolution WorldView-3 dataset shows that AILPPN obtains the best performance in all three non reference indices (see Table 5). The Dλ, Ds, and QNR values of the proposed method are 0.0050, 0.0418, and 0.9534, respectively. Compared with PanGAN, the proposed method reduces Dλ from 0.0055 to 0.0050 and Ds from 0.0452 to 0.0418, while improving QNR from 0.9495 to 0.9534. These improvements indicate that AILPPN can simultaneously suppress spectral distortion and spatial distortion. Therefore, the proposed method achieves a better balance between spectral preservation and spatial detail enhancement on the WorldView-3 dataset.

Table 4. Comparison of results of different algorithms on full-resolution QuickBird dataset.

algorithm	Ds	Dλ	QNR
IHS	0.2144	0.1143	0.6958
GS	0.2132	0.0829	0.7216
Wavelet	0.1215	0.1323	0.7623
MTF-GLP	0.1197	0.1343	0.7621
PNN	0.0061	0.0532	0.9410
MSDCNN	0.0059	0.0492	0.9452
LPPN	0.0058	0.0483	0.9462
PanGAN	0.0057	0.0427	0.9518
AILPPN	0.0051	0.0431	0.9520

Table 5. Comparison of results of different algorithms on full-resolution WorldView-3 dataset.

algorithm	Ds	Dλ	QNR
IHS	0.2013	0.1127	0.7087
GS	0.2042	0.0915	0.7230
Wavelet	0.1354	0.1228	0.7584
MTF-GLP	0.1241	0.1299	0.7621
PNN	0.0058	0.0526	0.9419
MSDCNN	0.0059	0.0492	0.9452
LPPN	0.0058	0.0474	0.9471
PanGAN	0.0055	0.0452	0.9495
AILPPN	0.0050	0.0418	0.9534

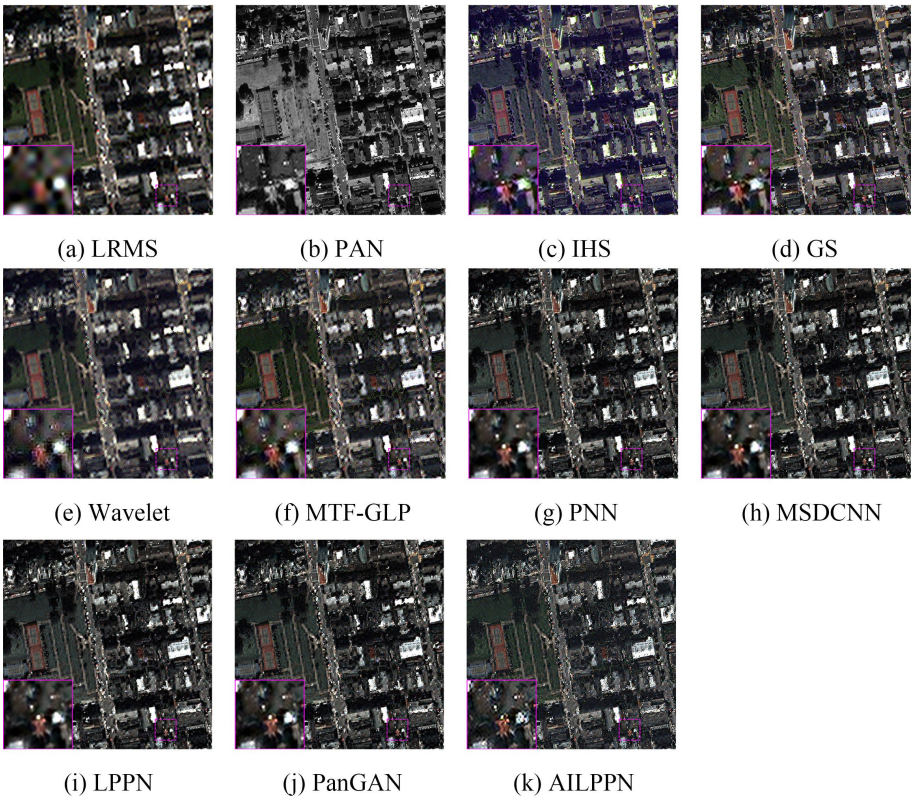

Figure 9. Sharpening results at full-resolution on QuickBird dataset.

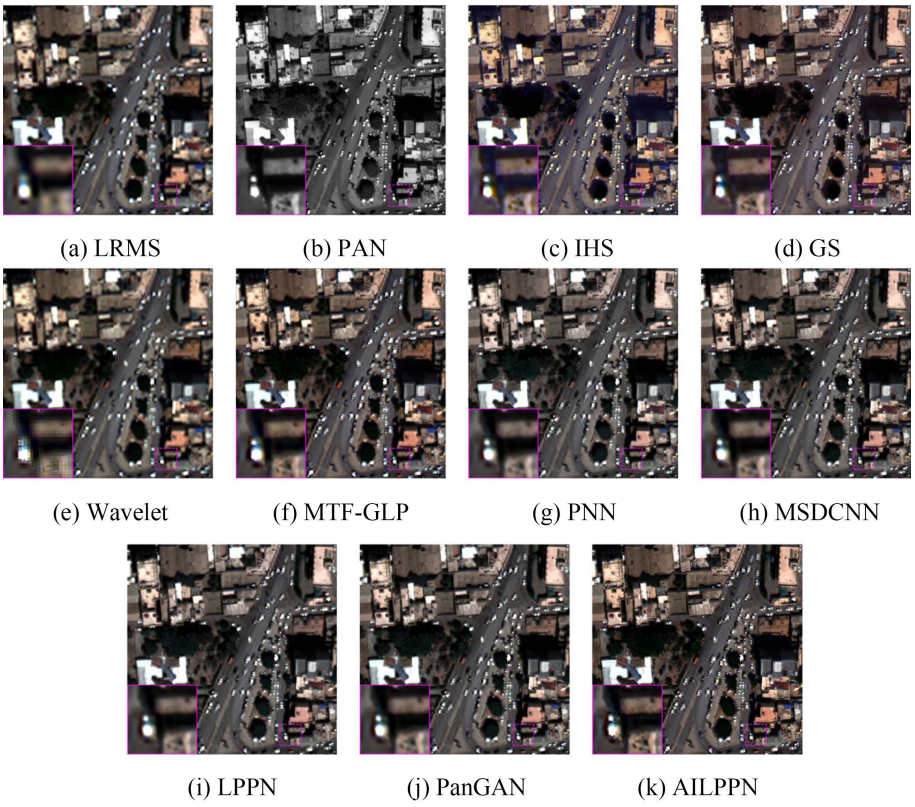

Figure 10. Sharpening results at full-resolution on WorldView-3 dataset.

3.3. Ablation Study

To verify the effectiveness of the proposed SEFCNN adaptive fusion module and SAM loss function, ablation experiments were conducted on the reduced-resolution QuickBird dataset. The baseline model corresponds to the original Laplacian pyramid pansharpening network without the SEFCNN module and SAM loss. Based on this baseline, SEFCNN and SAM loss were added separately and jointly to analyze their contributions.

The ablation results demonstrate that both the SEFCNN module and SAM loss contribute to the improvement of fusion performance (see Table 6). The baseline model obtains an SSIM of 0.8923, PSNR of 30.8549, CC of 0.9134, and ERGAS of 2.2264. After introducing the SEFCNN adaptive fusion module alone, SSIM increases to 0.9117, PSNR increases to 31.2455, and CC increases to 0.9224. This improvement indicates that the SEFCNN module can enhance feature representation by adaptively recalibrating channel responses, thereby improving the recovery of spatial and spectral features.

When only the SAM loss is added, SSIM, PSNR, and CC also improve compared with the baseline, while ERGAS decreases from 2.2264 to 2.1638. This result shows that the SAM loss can provide a more direct spectral constraint during training and reduce spectral distortion in the fused image. Compared with the SEFCNN only model, the SAM loss only model shows a more obvious reduction in ERGAS, which is consistent with its role in improving spectral fidelity.

When both SEFCNN and SAM loss are used, the complete AILPPN model achieves the best results, with SSIM of 0.9225, PSNR of 32.2128, CC of 0.9298, and ERGAS of 2.0189. These results demonstrate that the two components are complementary. The SEFCNN module mainly improves adaptive feature fusion, while the SAM loss strengthens spectral consistency. Their joint use enables the network to obtain better spatial reconstruction and spectral preservation simultaneously.

In addition, the influence of the weight coefficient of the SAM loss was further analyzed (see Table 7). When the coefficient is too small, the spectral constraint is insufficient, and the model cannot fully benefit from the SAM loss. As the coefficient increases, the performance gradually improves. The best result is achieved when the coefficient is set to 0.2, where SSIM, PSNR, CC, and ERGAS reach 0.9225, 32.2128, 0.9298, and 2.0189, respectively. However, when the coefficient is further increased to 0.3 or 0.4, the performance slightly decreases. This suggests that an excessively strong spectral constraint may weaken the network's ability to recover spatial details. Therefore, the coefficient of the SAM loss is set to 0.2 in the final model.

Overall, the ablation study confirms the rationality of the proposed network design. The SEFCNN adaptive fusion module improves feature fusion, while the SAM loss enhances spectral preservation. The complete AILPPN model achieves the best overall performance, verifying the effectiveness of combining channel attention-based adaptive fusion with spectral angle based loss optimization.

Table 6. Comparison of ablation results of different algorithms on reduced-resolution QuickBird dataset.

SEFCNN Adaptive Fusion Module	SAM Loss Function Module	QuickBird			
		SSIM	PSNR	CC	ERGAS
×	×	0.8923	30.8549	0.9134	2.2264
√	×	0.9117	31.2455	0.9224	2.2142
×	√	0.9042	31.1343	0.9188	2.1638
√	√	0.9225	32.2128	0.9298	2.0189

Table 7. Results of weight coefficient experiments for the hybrid loss module on QuickBird.

Weight Coefficient Settings of the Hybrid Loss Function	QuickBird			
	SSIM	PSNR	CC	ERGAS
0	0.9117	31.2455	0.9224	2.2142
0.01	0.9121	31.2747	0.9227	2.2049
0.05	0.9157	31.4426	0.9241	2.1589
0.1	0.9197	31.7859	0.9272	2.0868
0.2	0.9225	32.2128	0.9298	2.0189
0.3	0.9211	32.2021	0.9277	2.0723
0.4	0.9202	32.1683	0.9247	2.0837

4. Discussion

The experimental results in Section 3 show that the proposed AILPPN method achieves stable and competitive performance on both QuickBird and WorldView-3 datasets. In the reduced-resolution experiments, AILPPN obtains better results in most full reference evaluation indices, especially SSIM, PSNR, SAM, CC, and ERGAS (see Table 2 and Table 3). In the full-resolution experiments, the proposed method also achieves favorable non reference evaluation results, particularly in terms of QNR (see Table 4 and Table 5). These results indicate that AILPPN can effectively improve the spatial resolution of multispectral images while maintaining relatively stable spectral information. Therefore, this section further discusses the reasons for the performance improvement of the proposed method, the roles of the SEFCNN module and SAM loss, the difference between reduced resolution and full resolution evaluation, and the remaining limitations of the current work.

4.1. Effect of Laplacian Pyramid-Based Multi-Scale Fusion

The improvement of AILPPN is first related to its Laplacian pyramid based multi scale fusion framework. In pansharpening, the PAN image mainly provides high frequency spatial details, while the MS image contains richer spectral information. If these two types of information are fused directly at a single scale, the network may enhance spatial details but also introduce spectral distortion. This problem

can be observed in several traditional methods. For example, IHS and GS can improve image sharpness to some extent, but they tend to cause obvious color deviation. Wavelet and MTF-GLP preserve part of the spectral information, but their results still contain blurred textures and insufficient edge recovery (Figure 7 and Figure 8).

By contrast, the proposed method decomposes the input images into multiscale representations through the MTF-based Laplacian pyramid. This design allows the network to learn spatial and spectral relationships at different resolution levels. Low frequency components help preserve the overall spectral structure of the MS image, while high-frequency components provide detailed edge and texture information from the PAN image. Through progressive reconstruction, the network gradually transfers coarse scale structural information to fine scale details, which helps avoid the instability caused by direct high frequency injection. This is why the fused images generated by AILPPN show clearer road edges, building contours, and vegetation textures while maintaining a more natural color distribution (Figure 7 and Figure 8).

4.2. Role of the SEFCNN Adaptive Fusion Module

The second reason for the improvement lies in the SEFCNN adaptive fusion module. In remote sensing image fusion, different feature channels do not contribute equally to the final result. Some channels contain more important spectral information, while others are more closely related to spatial textures or may even include redundant responses. If the network treats all channels in the same way, important features may be weakened during the fusion process.

The SEFCNN module introduces channel attention into the fusion network. It adaptively assigns weights to different feature channels, enabling the network to enhance useful spatial spectral information and suppress less relevant features. This mechanism is particularly useful for WorldView-3 images, because the dataset contains more multispectral bands and therefore has a more complex spectral relationship than QuickBird. The visual comparison also supports this point. In the WorldView-3 experiments, AILPPN shows better edge restoration and more stable spectral appearance than most comparison methods (Figures 3-8 and Figure 10).

The ablation study further confirms the effectiveness of SEFCNN. When only the SEFCNN module is added to the baseline model, SSIM, PSNR, and CC are all improved (see Table 6). This shows that the channel attention mechanism is not merely an additional structural component, but actually improves the feature selection and fusion ability of the network. It helps the model focus more on useful information during multi scale reconstruction, thereby improving the overall quality of the pansharpened image.

4.3. Contribution of SAM Loss to Spectral Preservation

In addition to network structure, the loss function also plays an important role in

the proposed method. Traditional pixel-level reconstruction loss can constrain the numerical difference between the fused image and the reference image, but it does not directly describe the spectral angle relationship among different bands. For remote sensing images, spectral consistency is very important because many down stream applications, such as land cover classification, vegetation monitoring, and object interpretation, depend on stable spectral information.

For this reason, the proposed method introduces SAM loss into the hybrid loss function. The SAM loss directly measures the angle between the spectral vector of the output image and that of the reference image. A smaller spectral angle means better spectral preservation. Therefore, this loss can guide the network to maintain spectral consistency while reconstructing spatial details. The ablation results show that adding SAM loss alone can improve the baseline model and reduce ERGAS (see Table 6), indicating that the spectral constraint is effective.

However, the weight of SAM loss needs to be carefully controlled. The weight coefficient experiment shows that the model performance first improves and then decreases as the SAM loss weight increases (see Table 7). When the coefficient is too small, the spectral constraint is insufficient. When the coefficient is too large, the network may pay too much attention to spectral preservation and weaken the recovery of spatial details. The best performance is obtained when the coefficient is set to 0.2, which indicates that spatial detail enhancement and spectral fidelity should be balanced rather than optimized separately.

4.4. Full-Resolution Applicability and Remaining Limitations

The full resolution experiments further demonstrate the practical value of the proposed method. Compared with reduced resolution experiments, full resolution experiments are closer to real remote sensing applications because the original satellite images are directly used as input. However, the ideal HRMS reference image is unavailable in this situation, so non reference indices such as Dλ, Ds, and QNR are used for evaluation. The results show that AILPPN obtains the best QNR values on both QuickBird and WorldView-3 datasets (see Table 4 and Table 5). This indicates that the proposed method maintains a good balance between spectral distortion suppression and spatial detail enhancement in real resolution scenarios.

From the visual results, AILPPN also performs well in areas containing buildings, roads, vegetation, and mixed land cover objects (Figure 9 and Figure 10). Compared with traditional methods, it reduces obvious color distortion. Compared with other deep learning based methods, it recovers clearer local structures and richer spatial textures. This suggests that the proposed multi-scale reconstruction framework has good adaptability to different sensors and image contents.

Nevertheless, several limitations remain. First, the experiments are mainly conducted on QuickBird and WorldView-3 datasets. Although these two datasets are representative, they cannot fully cover all remote sensing sensors, land cover types, and imaging conditions. Second, the current evaluation mainly focuses on commonly used image quality indices and visual comparison. Future research can

further examine whether the fused images improve downstream tasks such as classification, detection, or change analysis. Third, the proposed method introduces multi scale decomposition, SEFCNN fusion, and multi-scale supervision, which improves performance but also increases model complexity. Therefore, future work should further optimize the network structure and improve computational efficiency.

Overall, the proposed AILPPN method improves pansharpening performance through the combination of MTF-based Laplacian pyramid decomposition, SEFCNN adaptive fusion, progressive reconstruction, and SAM based spectral constraint. The experimental results verify that these components work together to enhance spatial details and preserve spectral information, providing an effective framework for high quality remote sensing image pansharpening.

5. Conclusions

Taking remote sensing image pansharpening as the research object, this study proposed an attention based Laplacian pyramid pansharpening network, named AILPPN. The proposed method combines MTF-based Laplacian pyramid decomposition, SEFCNN adaptive fusion, progressive reconstruction, multi scale supervision, and a hybrid loss function with SAM spectral constraint. Through this framework, spatial details from the PAN image and spectral information from the MS image can be extracted and fused at different scales, thereby improving the balance between spatial enhancement and spectral preservation.

Experiments were conducted on QuickBird and WorldView-3 datasets under both reduced resolution and full resolution conditions. The results show that AILPPN achieves better or competitive performance compared with traditional and deep learning based pansharpening methods. In reduced resolution experiments, the proposed method improves most full reference indices, including SSIM, PSNR, SAM, CC, and ERGAS. In full resolution experiments, it also obtains favorable non reference evaluation results, especially in terms of QNR. The ablation experiments further verify that the SEFCNN module enhances adaptive feature fusion, while the SAM loss improves spectral consistency. Their joint use contributes to the best overall performance.

Overall, the contribution of this study lies in providing an effective multi-scale spatial spectral fusion framework for pansharpening. However, the current experiments are mainly limited to two datasets. Future work will further test the method on more sensors and scenes, optimize model complexity, and evaluate its influence on downstream remote sensing tasks.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] Li, J., Pei, Y., Zhao, S., Xiao, R., Sang, X. and Zhang, C. (2020) A Review of Remote

Sensing for Environmental Monitoring in China. *Remote Sensing*, **12**, Article 1130. https://doi.org/10.3390/rs12071130

[2] Fan, Z., Zhan, T., Gao, Z., Li, R., Liu, Y., Zhang, L., *et al.* (2022) Land Cover Classification of Resources Survey Remote Sensing Images Based on Segmentation Model. *IEEE Access*, **10**, 56267-56281. https://doi.org/10.1109/access.2022.3175978

[3] Sishodia, R.P., Ray, R.L. and Singh, S.K. (2020) Applications of Remote Sensing in Precision Agriculture: A Review. *Remote Sensing*, **12**, Article 3136. https://doi.org/10.3390/rs12193136

[4] Wellmann, T., Lausch, A., Andersson, E., Knapp, S., Cortinovis, C., Jache, J., *et al.* (2020) Remote Sensing in Urban Planning: Contributions Towards Ecologically Sound Policies? *Landscape and Urban Planning*, **204**, Article ID: 103921. https://doi.org/10.1016/j.landurbplan.2020.103921

[5] Jhawar, M., Tyagi, N. and Dasgupta, V. (2013) Urban Planning Using Remote Sensing. *International Journal of Innovative Research in Science, Engineering and Technology*, **1**, 42-57.

[6] Khalifeh, A., Gupta, M., Almomani, O., Khasawneh, A.M. and Darabkh, K.A. (2022) Smart Remote Sensing Network for Early Warning of Disaster Risks. In: Denizli, A., Alencar, M.S., Nguyen, T.A. and Motaung, D.E., Weds., *Nanotechnology-Based Smart Remote Sensing Networks for Disaster Prevention*, Elsevier, 303-324. https://doi.org/10.1016/b978-0-323-91166-5.00012-4

[7] Pettorelli, N., Nagendra, H., Rocchini, D., Rowcliffe, M., Williams, R., Ahumada, J., *et al.* (2017) Remote Sensing in Ecology and Conservation: Three Years On. *Remote Sensing in Ecology and Conservation*, **3**, 53-56. https://doi.org/10.1002/rse2.53

[8] Harrison, T. and Strohmeyer, M. (2022) Commercial Space Remote Sensing and Its role in National Security. Center for Strategic & International Studies.

[9] Kempeneers, P., Sedano, F., Seebach, L., Strobl, P. and San-Miguel-Ayanz, J. (2011) Data Fusion of Different Spatial Resolution Remote Sensing Images Applied to Forest-Type Mapping. *IEEE Transactions on Geoscience and Remote Sensing*, **49**, 4977-4986. https://doi.org/10.1109/tgrs.2011.2158548

[10] Nelson, P.R., Maguire, A.J., Pierrat, Z., Orcutt, E.L., Yang, D., Serbin, S., *et al.* (2022) Remote Sensing of Tundra Ecosystems Using High Spectral Resolution Reflectance: Opportunities and Challenges. *Journal of Geophysical Research: Biogeosciences*, **127**, e2021JG006697. https://doi.org/10.1029/2021jg006697

[11] Jin, X., Liu, L., Ren, X., Jiang, Q., Lee, S., Zhang, J., *et al.* (2024) A Restoration Scheme for Spatial and Spectral Resolution of the Panchromatic Image Using the Convolutional Neural Network. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **17**, 3379-3393. https://doi.org/10.1109/jstars.2024.3351854

[12] Haque, M.A., Reza, M.N., Ali, M., *et al.* (2024) Effects of Environmental Conditions on Vegetation Indices from Multispectral Images: A Review. *Korean Journal of Remote Sensing*, **40**, 319-341.

[13] Cao, Z., Deng, L., Dou, H., Wu, X. and Zhong, Y. (2024) SSDiff: Spatial-Spectral Integrated Diffusion Model for Remote Sensing Pansharpening. *Advances in Neural Information Processing Systems* 37, Vancouver, 10-15 December 2024, 77962-77986. https://doi.org/10.52202/079017-2478

[14] Jiao, Z. (2024) The Application of Remote Sensing Techniques in Ecological Environment Monitoring. *Highlights in Science, Engineering and Technology*, **81**, 449-455. https://doi.org/10.54097/7dqegz64

[15] Huang, Z., Chen, Q., Chen, Q. and Liu, X. (2018) Variational Pansharpening for Hyperspectral Imagery Constrained by Spectral Shape and Gram-Schmidt Transformation. *Sensors*, **18**, Article 4330. https://doi.org/10.3390/s18124330

[16] Ghahremani, M. and Ghassemian, H. (2015) Remote Sensing Image Fusion Using Ripplet Transform and Compressed Sensing. *IEEE Geoscience and Remote Sensing Letters*, **12**, 502-506. https://doi.org/10.1109/lgrs.2014.2347955

[17] Ciotola, M., Poggi, G. and Scarpa, G. (2023) Unsupervised Deep Learning-Based Pansharpening with Jointly Enhanced Spectral and Spatial Fidelity. *IEEE Transactions on Geoscience and Remote Sensing*, **61**, 1-17. https://doi.org/10.1109/tgrs.2023.3299356

[18] Masi, G., Cozzolino, D., Verdoliva, L. and Scarpa, G. (2016) Pansharpening by Convolutional Neural Networks. *Remote Sensing*, **8**, Article 594. https://doi.org/10.3390/rs8070594

[19] Ma, J., Yu, W., Chen, C., Liang, P., Guo, X. and Jiang, J. (2020) Pan-GAN: An Unsupervised Pan-Sharpening Method for Remote Sensing Image Fusion. *Information Fusion*, **62**, 110-120. https://doi.org/10.1016/j.inffus.2020.04.006