International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 06 | June - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

Hyperspectral Image Denoising using Deep Learning

Santhoshini S, Dr.Goutham M A

ABSTRACT: A new hyperspectral image denoising algorithm, called the dual deep convolutional neural network (DD-CNN), is proposed in this paper. In contrast to internal denoising methods that utilize only the features from the target noisy image, the DD-CNN extensively explores the similarities between the target noisy image and the clean reference image from other bands. As external data, the reference images are selected based on the structural similarity index metric (SSIM). The DD-CNN is composed of two CNNs: one is responsible for extracting the features of the target image, and the other is responsible for extracting features from the reference image. A new activation function is proposed that activates the two types of features in the DD-CNN. Based on the dual structure and the new activation function, the external features extracted from the reference images are thoroughly integrated into the internal features of the target noise image. We experimented on different datasets with different noise levels; we also tested special cases for reference images with extra or undesirable features. The DD-CNN algorithm can effectively utilize the similarity between the external image and the target image. When the noise level is high, the advantages of the DD-CNN are obvious.

Keywords: Hyperspectral image denoising, deep dual neural network, feature learning, activation function.

I. INTRODUCTION

Hyperspectral images (HSI) - images of the same scene captured across a number of different wavelengths - are important to various civilian and military applications including remote sensing [1], agriculture [2], object tracking [3] and recognition [4]. Existing hyperspectral imaging methods often suffer from the problem of low signal-to-noise ratio (SNR) due to the fundamental limitation in physics - i.e., simultaneously acquiring many images of different spectral bands inevitably leads to shorter exposure time and narrower bandwidth. Accordingly, the acquired hyperspectral images are often corrupted by heavy noise which could severely deteriorate the performance of HSI-based segmentation and analysis algorithms. In view of the expensive cost of hardware-based solution to noise suppression, it is therefore desirable to develop effective software based HSI denoising techniques as a preprocessing step to support various HSI-related applications.

Existing HSI denoising techniques in the literature can often be viewed as the extensions of grayscale image denoising methods. Early works including dictionary learning based methods [5], BM3D [6], and sparsity based denoising methods [7], [8] simply remove the noise in a band-by-band manner and cannot exploit strong correlation across the spectral bands.

In order to exploit spatial-spectral correlations among HSI, a wavelet shrinkage method in the derivative-domain was proposed for HSI denoising in [9]; the principal component analysis (PCA) was combined with wavelet shrinkage methods in [10]; a spatial-spectral total variation model was developed for HSI denoising in [11]. This line of research has led to a multiple spectral-band conditional random fields (MSBCRF) model [12] aiming at simultaneously characterizing the spatial-spectral dependencies in a unified probabilistic framework. Other notable works include a multidimensional Wiener filtering with adaptive tensor flattening [13], a vector formulation of bilateral filtering [14] and the extension of well-known BM3D method into BM4D [15].

More recently, the class of low-rank tensor approximation methods have been extended for HSI denoising [13], [16] as well. With Tucker decomposition, the multirank of HSI can be first estimated via the alternative least square (ALS) algorithm and then the noiseless HSI can be reconstructed from multirank truncation of the core tensor [13]. To address the uniqueness issue with Tucker decomposition during the estimation of multirank, the parallel factor analysis (PARAFAC) decomposition method has been proposed for HSI denoising in [17]. A nonlocal tensor dictionary learning method was proposed in [18] where similar noisy tensors grouped by similar patches are denoised by multirank truncation. Different from multirank truncation-based tensor approximation, high order singular value decomposition (HOSVD) is exploited in [19]. After HOSVD is applied to high-order tensors formed by similar patches, the core tensor coefficients were denoised via hard-thresholding followed by Wiener filtering. Most recently, Laplacian scale mixture (LSM) models were proposed for core tensor coefficients in [20]; the LSM prior led to an adaptive thresholding algorithm in which both coefficients and thresholds can be jointly estimated from noisy data.

In this paper, a well-designed discriminative model is proposed to further improve HSI denoising performance. In the proposed method, 3-D kernels are combined with convolution to enlarge the receptive fields in the spatial and spectral bands simultaneously. This scheme supports better enforcement by capturing more pixel-wise features in multiple dimensions. A single denoising network model is trained to remove the mixed type of photon and thermal noise at various levels, rather than a single type of noise at a fixed noise level. In addition, the employed multiscale structure allows the proposed denoising model to retain subtle image details and achieve an excellent effect, outperforming other state-of-the-art denoising methods for HSIs. In more detail, the contributions of our work are summarized as follows.

1) A 3-D convolution kernel rather than a 2-D kernel is employed in HSI denoising CNN model and allows the CNN to be formulated with the 3-D kernel to extract the spectral and spatial features of the hyperspectral data simultaneously.
2) The 3-D atrous convolution kernel is introduced to capture more context information by enlarging the receptive field without increasing the number of network parameters to be trained and the computation cost.
3) The 3-D separable kernel is employed in the proposed multibranch blocks to decrease the time, which gives the advantage of fast training convergence with reduced overfitting risk.
4) In the proposed noise removal, different scale features are extracted in parallel by filters with different dilation rates, and then the information from wide and local ranges is fused for better performance. The experimental results

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 06 | June - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

demonstrate that the multiscale fusion gives a great help for performance improvement.

5) The proposed end-to-end trained CNN not only removes the mixed type of thermal and photon noises, but also achieves blind denoising over a certain range. That implies the proposed model is flexible and robust.

The remaining paper is organized as follows. The related work is introduced in Section II, the details of our proposed method are given in Section III, and the performed experiments and resulting discussions are presented.

II. RELATED WORK

A. Hyperspectral Noise Degradation Model

HSI data can be denoted by 3-D cube **Y** of size *M*×*N* ×*B*, whose degradation model can be described as

$$\mathbf{Y} = \mathbf{X} + \mathbf{V} \qquad (1)$$

where **X** is the ideal noise-free data, **V** = [*v*1, *v*2, . . ., *vB*] is the additive noise with the Gaussian distribution $\upsilon_n \sim \xi(0, \sigma_n^2)$and $1 \leq n \leq B$ and σ_n^2 mean that the noise intensity varies in the *n*th spectra. Hence, the HSI denoising process is to estimate the original data **X** from the noisy observation **Y**.

B. Analysis of the Existing HSI Denoising Methods

Up to now, there are two main types of HSI denoising methods: 1) transform-domain-based methods and 2) spatial domain-based methods. The transform-domain-based methods attempt to separate clear signals from the noisy data by various transformations, such as principal component
analysis (PCA), Fourier transform, or wavelet transform. For example, Atkinson *et al.* [11] presented an estimator utilizing discrete Fourier transform to decorrelate the signal in the spectral domain, and a wavelet transform was utilized for the spatial filtering. Othman and Qian [12] employed a hybrid spatial–spectral derivative-domain wavelet shrinkage noise removal (HSSNR) method. This method depends on the spectral derivative domain, where the noise level is elevated, and benefits from the dissimilarity of the signal nature in the spatial and spectral dimensions. The major weakness of this type of approaches is that these methods are sensitive to the selection of the transform function and cannot consider the differences in the geometrical characteristics of HSIs.

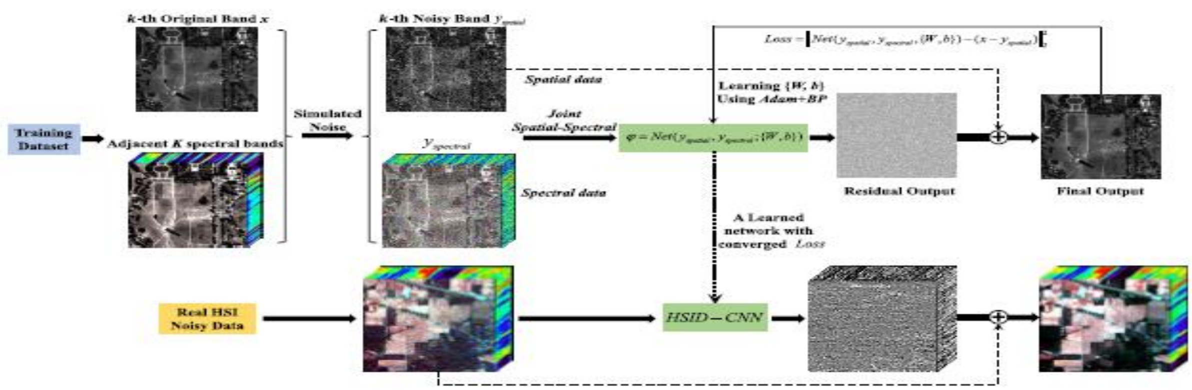

Figure 1. Flowchart of the proposed HSID-CNN method for removing noise in HSI data.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 06 | June - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

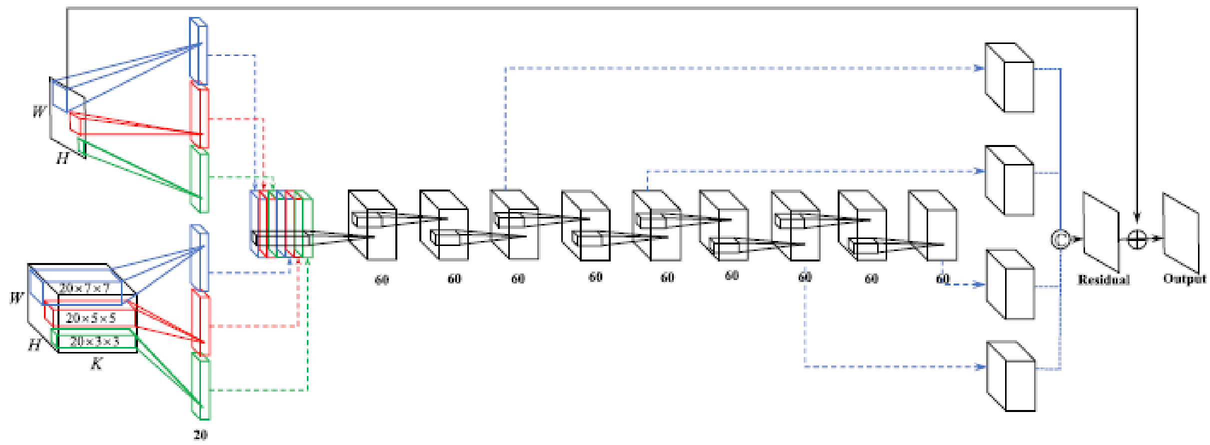

Figure 2. Structure of HSID-CNN.

Although these HSI denoising methods can achieve relatively better results, the good performance must precisely tune parameters for each HSI [22]. This generates the unintelligent and time consuming for different HSI data. Therefore, it is significant to build a fast, efficient and universal framework to adapt to different HSI data with different situations.

A. Deep Neural Networks for HSI Denoising

Several types of deep neural network models have beenproposed for HSI processing, i.e., deep brief networks,stacked autoencoder, a spectral-spatial feature-based classificationframework that jointly uses dimension reductionand deep learning techniques for spectral and spatial featureextraction, deep feature extraction and classificationmodel, spectral–spatial classification of HSI,and a deep autoencoder-based approach to identify signalfeatures.

In addition, the denoising models based on deep learninghave achieved significantly better performance than othertraditional methods in medical imaging. Because the3-D medical image with multiple successive layers to representan object body is very similar to the HSI with multiband, thesame deep learning-based denoising techniques of 3-D medicalimages can be utilized in HSI, composed of many contiguousand narrow spectral bands. Thus, it is natural to introduce theCNN to HSIs denoising task.

Also, there have been several methods that exploit CNNsto remove HSI noise. However, almost all of them directlyextend 2-D convolutions by increasing the number of spectralchannels as the same way as done for RGB images. The problem of the 2-D approach will be more precisely discussed in the following.

B. Atrous Convolution

Atrous convolution is a generalization of Kroneckerfactoredconvolutional filters that enforce exponentiallyexpanding receptive fields without reducing the resolution ofthe feature map. For pixel-level vision tasks, the image sizeof input and output should be consistent, but the networkneeds to do up-sampling to amplify the feature map sizedue to the pooling layer. The drawback of this process isthat complete recovery of missing information is hard byup-sampling. However, the receptive field of each layer willbe

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 06 | June - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

decreased if the pooling layer is removed, and thenthe output performance will be greatly reduced. The networksin and employ atrous convolution insteadof a down-sampling-based pooling layer, which not onlyavoids resolution reduction, but also covers the large-scalereceptive field.

In low-level vision tasks such as denoising, it is crucialto extract more contextinformation around each pixel andto predict each pixel according to the surrounding context.The atrous convolution with various dilated rate can expandthe desirable receptive field without increasing the number ofparameters and ensure that image information is not lost. Forthe same reasons, this strategy of capturing a wider area ofcontext information using an atrous filter is a feasible approach
in the denoising process.

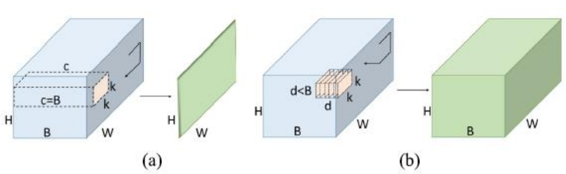

Figure 3. Difference between 2-D and 3-D convolution operations in HSI. (a) Output is a 2-D feature map after using 2-D filter convolution. The multiband are treated as multichannel. (b) Output is a 3-D cube featuremap after using 3-D filters. The spectral information in HSI is reserved.

C. 3-D Convolution

Common 2-D convolution computes a feature in the spatialdomain only, and its output is presented in the 2-D map.However, HSI is represented in both spatial and spectral
domains, so that it could be treated as a 3-D image in whichadjacent spectral bands are highly correlated. If 2-D convolutionoperations are applied to HSI as the same way as toRGB image, each output feature map from a convolution layeressentially is treated as a conversion result from the multibandinput image as in Fig. 1(a). As a result, substantial correlationinformation between adjacent bands would be lost. That isto say, the 2-D convolution reduces the spectral resolution,because it aggregates all input frequency bands into a singlechannel. Therefore, 3-D convolution is well-suited for HSI andcan preserve the spectral information as much as possible. Thisis very important for HSI denoising, which still contains richspectral information in the 3-D cube after removing noise.The proposed method performs 3-D convolution to extract thefeature maps along both spatial and spectral dimensions.

The 3-D convolution is enforced by convolving with 3-Dkernels and then outputs a volume formed by stacking multiplecontiguous spectral bands together as in Fig. 1(b).In addition, the network using 3-D filters is flexibly capableof receiving HSIs with the arbitrary number of bands, while2-D filters could only take care of the same fixed number ofbands as they are trained. In, they fixed ten channels ofinput and output layers as using the 2-D filter. It is unsuitablefor diverse HSI data, whose number of spectral bands is fromdozens to hundreds.

Recently, however, a novel method has been proposedto resolve the limitations of 2-D convolution, which takes bothinputs of a single band and its adjacent multibands throughtwo branches simultaneously. However, the number of adjacentbands for extracting spectral information is fixed, so that theadjustability in the spectral dimension could not be flexiblyconsidered.
In our opinion, the 3-D atrous convolution is a flexible andefficient method which the model enlarges the receptive fieldsand extracts multiscale information in both spatial and spectraldimensions by adjusting the dilated rate without increasing parameters.

III. PROPOSED DENOISING METHOD

In this section, the proposed 3DADCNN model is explainedin detail, including how to design network architecture, howto determine the training parameters, how to lay out themultibranch and multiscale module, and how to obtain trainingdata by adding the mixed type of noise.

A. Architecture of Denoising CNN

In the field of computer vision, most problems have beensuccessfully solved using deep neural networks. However,as the network becomes deeper and deeper, the training difficultyincreases correspondingly. Recently, however, the networkstructures with proper fusion have also received researchattention. In comparison with deep network, the improvedstructure contains much shorter paths and more channels.
We corrupt clean images with two types of photon andthermal noise. The noisy observation is $x = y + n$, where$y$ and n stand for the corresponding clean image and the
noise, respectively. As shown in Fig. 2, the proposed modeladopts residual learning to learn a mapping function $F(x) \approx n$.The output of the network is a residual image, and then weemploy a skip connection that adds noisy observation x tothe output-end and performs $\hat{y} = x + F(x)$, where $\hat{y}$ isthe predicted image. Mean squared error (MSE) between theestimated clean images $\hat{y}$ and the ground truth y is calculatedas the loss function $l(\theta)$ to optimize the network parametersθ in the training. In (1), x_i and y_i represent the ith training noise-corrupted input image and output target clean image,respectively,

$$l(\theta) = \frac{1}{2N}\sum_{i=1}^{N} \|x_i + F(x_i, \theta) - y_i\|_F^2. \qquad (1)$$

B. Random Noise Level Training

In the existing HSI denoising methods, some models are only for a fixed noise level, and some other modelsare only designed for removing Gaussian noise.Models for removing noise need to apply *a priori* transformationsuch as variance stabilization transform (VST), whichtransforms mixed Poisson and Gaussian noises to a Gaussiandistribution.

In the training stage, however, the proposed method considerstwo typical random photon and thermal noises simultaneously,corresponding to Poisson and Gaussian noises,respectively. That is, n = n p + ng, where ng and n prepresent the Gaussian and Poisson noises to be added tothe clean image, respectively. To obtain various noise levels,the standard deviation σ of zero mean Gaussian noise is randomlychosen within the interval, i.e., $\sigma \in [10, 30]$. Parameterλ associated with Poisson noise is selected from a range,i.e., $\lambda \in [1, 5]$. Given a test image corrupted with two types ofrandom noise that belongs to the ranges, the trained networkcan be utilized to denoise directly without any preestimationof them. It provides flexibility in the HSI denoising task.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 06 | June - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

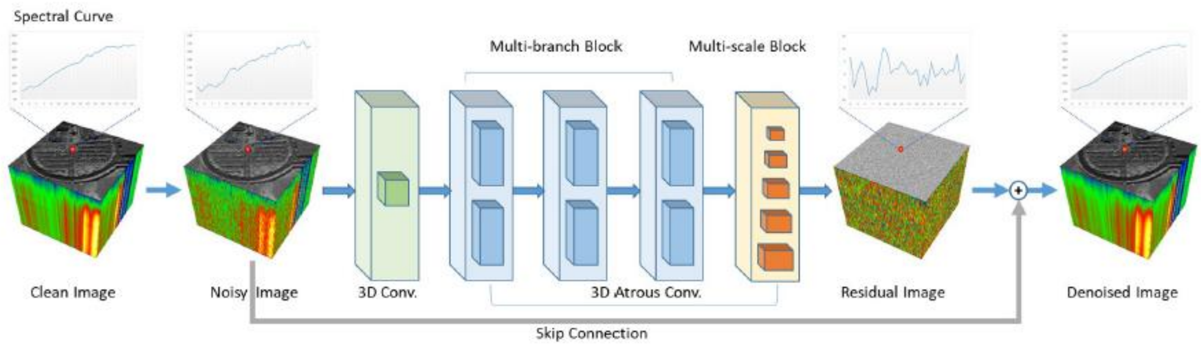

Figure 4. Architecture of out proposed network. The input is a corrupted image by random noise. The output is a residual image. The second to fourth blocks (blue blocks) adopt the multibranch structure, and the last block (orange block) is a multiscale structure.

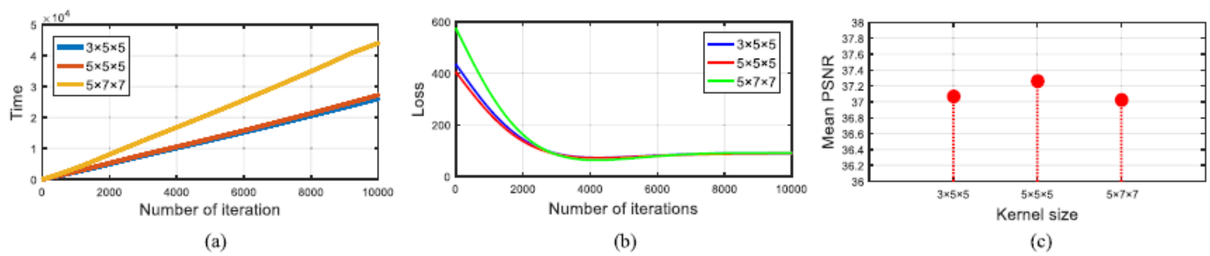

Figure 5. Comparison of different kernel sizes. (a) Training time. (b) Training loss. (c) Mean PSNR of 300 test patches.

C. Size and Number of Kernels

The time and MSE of the training on the same set of trainingdata are measured with different kernel sizes while the numberof filters at each block is fixed. Thereafter, the results for a testdata set of 300 images are compared in terms of mean peaksignal-to-noise ratio (PSNR). As shown in Fig. 3, the trainingtime increases rapidly as the filter size in the spatial domainincreases because of the computation complexity. In Fig. 3,the size of the 3-D filters is denoted by the $d \times k \times k$, where$d$ is the depth of the kernel in the spectral dimension, and$k$ is the kernel size in the spatial dimensions. In terms ofoverall performance and efficiency, 5 × 5 × 5 produces thebest result.

D. 3-D Atrous Convolution

The proposed method combines the 3-D filter kernels withatrous convolution to enlarge the receptive fields in the spatialand spectral bands simultaneously. This scheme achieves betterperformance by capturing more information from pixel-wisefeatures in multiple dimensions.For the denoising task, extracting large-scale pixel-levelfeatures is crucial to improving performance, so the kernel sizeof the convolutional should be as large as possible. However,the larger 3-D kernel dramatically increases computational
complexity, making the model training more difficult. Insteadof directly using a larger kernel, the proposed model employsa combination of 3-D

filter kernels with atrous convolution,which enlarges the receptive field without increasing thetraining parameters, which provides less chance of over-fitting.

TABLE I
STRUCTURE OF TWO BRANCHES

$B \times H \times W \times C_i$	
3D atrous Conv. + BN $d \times 1 \times k \times C_i \times C_{i+1}$	3D atrous Conv. + BN $d \times k \times 1 \times C_i \times C_{i+1}$
3D atrous Conv. + BN + ReLU $d \times k \times 1 \times C_{i+1} \times C_{i+1}$	3D atrous Conv. + BN + ReLU $d \times 1 \times k \times C_{i+1} \times C_{i+1}$
Sum $B \times H \times W \times C_{i+1}$	

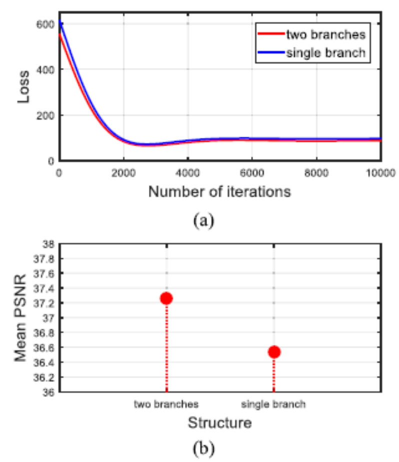

Figure 6. Comparison of single branch and two branches structure. (a) Training loss. (b) Test PSNR of mixed-type Gaussian noise with σ = 20 and Poissonnoise with λ = 3.

E. Multibranch Fusion

The proposed method employs multibranch modules inthe three blocks. As illustrated in Tables I and II 4, 3-Dseparable kernel is implemented with two parallel branches,which correspond $d \times k \times 1$ and $d \times 1 \times k$ kernels, respectively,instead of a $d \times k \times k$ kernel. The input and output featuremaps are represented as $B \times H \times W \times Ci$ and $B \times H \times W \times Ci+1$, respectively. B, H, and W denote the numberof bands, height, and width of the feature map, respectively.In addition, Ci is the number of input feature maps, and$Ci+1$ is the number of output feature maps. In our scheme,d and k are 5, respectively, so the number of parameters ofeach block is $4 \times 52 \times Ci \times Ci+1$ rather than$53 \times Ci \times Ci+1$Therefore, the time complexity of each block should be$O(4 \times M \times 52 \times Ci \times Ci+1)$ rather than$O(M \times 53 \times Ci \times Ci+1)$,where M represents the product of the number of bands,height, and width of every feature map. Our method canreduce the complexity by 20%, and has been validated in theexperiment.

F. Multiscale Fusion

In the deep neural network, the size of the receptivefield influences how much context feature information canbe utilized to a certain extent. Utilizing both local detailsand wider range features has an important influence ondenoising performance. Global average pooling is a goodmethod to describe the global contextual information, whichis commonly employed in image classification and semanticsegmentation models. However, this strategy cannotcover the necessary local information. Directly concentratinga 3-D HSI cube into a single vector may disrupt the spatialand spectral relation. Spatial pyramid polling extracts the feature map from different scales, which improves algorithmrobustness and accuracy of object recognition. In,atrous convolution is used to capture multiple scale featuresfor encoding the context information.

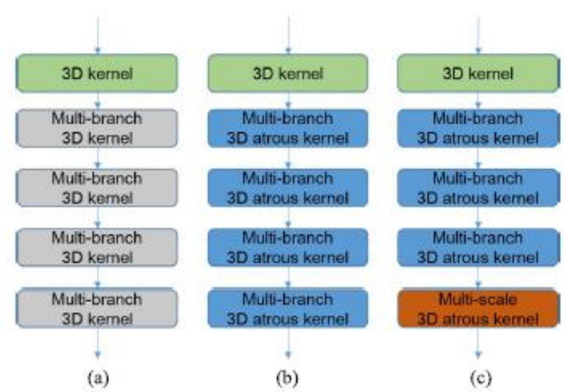

Figure 7. Architectures. (a) 3-D. (b) 3DA. (c) M-3DA.

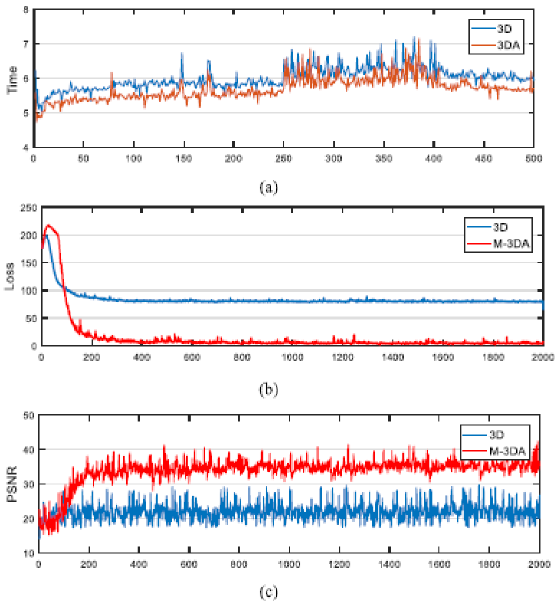

Figure 8. (a) Training time of 3-D and 3DA. (b) Training loss of 3-D and M-3DA. (c) PSNR of 3-D and M-3DA.

Fig. 6(b) uses the atrous convolutions instead of just 3-Dconvolutions in Fig. 6(a), and Fig. 6(c) equips the additionalmultiscale module. From the experimental results, in almostall epochs the computation time of model 3DA is less than3-D see Fig. 7(a)]. The multiscale fusion is employed in thelast block of the M-3DA; the overall performance has beenrapidly increased. Fig. 7(b) and (c) illustrates that the PSNRhas been improved considerably as the loss decreases. Theexperiments show that multiscale fusion based on 3-D atrousconvolution plays a significant role in the process for betterdenoising results.

IV. CONCLUSION

We have presented a new learning-based method for HIS denoising, called single denoising CNN (HSI-SDeCNN). Thismethod considers the spatial–spectral correlation present inHSIs, taking as input a full data cube instead of a single band.The main characteristics of this method are: a down samplinglayer that allows the network to be faster without losingdenoising performance, and a noise-level map that is used togive as input to the network an estimation of the amount ofnoise. The proposed method outperformed other mainstreammethods commonly adopted in HSI denoising on synthetic andreal data sets, with only one single trained model. In particular,it exhibits superior performance both in terms of denoisingcapability and computational efficiency. The performance ofthe method depends on the input noise level map *M*, thatis, the only hyperparameter that needs to be tuned. Thisparameter, as demonstrated from the results, is flexible in
handling different levels of noise.As with any new approach, there are still some future
research avenues that can be further explored. Specifically,the proposed network makes the denoising at only one levelfor all the bands. Such a level is specified by the input noise levelmap. However, in HSIs, the noise generally differs fromone band to another. For this reason, a further improvementof the method will focus on adapting the input noise level toeach specific band.

REFERENCES

[1] Wei Liu and Joonwhoan LeeA 3-D Atrous Convolution Neural Networkfor Hyperspectral Image Denoising0196-2892 © 2019 IEEE.

[2] J. Li, W. Huang, and X. Tian, “Fast detection and visualization of earlydecay in citrus using Vis-NIR hyperspectral imaging,” *Comput. Electron.Agricult.*, vol. 127, pp. 582–592, Sep. 2016.

[3] Z. Wu, Q. Wang, J. Jin, and Y. Shen, “Structure tensor total variation regularizedweighted nuclear norm minimization

for hyperspectral imagemixed denoising," *Signal Process.*, vol. 131, pp. 202–219, Feb. 2017.

[4] J. Kim, J. K. Lee, and K. M. Lee, "Accurate image super-resolutionusing very deep convolutional networks," in *Proc. CVPR*, Jun. 2016,pp. 1646–1654.

[5] Y. Li, J. Hu, X. Zhao, W. Xie, and J. Li, "Hyperspectral image super resolutionusing deep convolutional neural network," *Neurocomputing*,vol. 266, pp. 29–41, Nov. 2017.

[6] J. Kim, J. K. Lee, and K. M. Lee. (2016). "Deeply-recursive convolutionalnetwork for image super-resolution."

[7] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond aGaussian Denoiser: Residual learning of deep CNN for image denoising,"*IEEE Trans. Image Process.*, vol. 26, no. 7, pp. 3142–3155,Jul. 2017.

[8] W. Shao and S. Du, "Spectral-spatial feature extraction for hyperspectralimage classification: A dimension reduction and deep learningapproach," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 8,pp. 4544–4554, Oct. 2016.

[9] Y. Chen, H. Jiang, C. Li, X. Jia, and P. Ghamisi, "Deep feature extractionand classification of hyperspectral images based on convolutionalneural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp. 6232–6251, Oct. 2016.

[10] K. G. Lore, A. Akintayo, and S. Sarkar. (Apr. 2016)."LLNet: A deep autoencoder approach to natural low-light imageenhancement."

[11] D. Wu, K, Kim, G. E. Fakhri, and Q. Li. (May 2017). "A cascadedconvolutional neural network for X-ray low-dose CT image denoising."

[12] H. Chen *et al.* (Feb. 2017). "Low-dose CT with a residualencoder-decoder convolutional neural network (RED-CNN)."

[13] Y. Han, J. Yoo, and J. C. Ye. (Nov. 2016). "Deep residual learningfor compressed sensing CT reconstruction via persistent homologyanalysis."

[14] J. Li, W. Huang, and X. Tian, "Fast detection and visualization of earlydecay in citrus using Vis-NIR hyperspectral imaging," *Comput. Electron.Agricult.*, vol. 127, pp. 582–592, Sep. 2016.

[15] D. Wu, K, Kim, G. E. Fakhri, and Q. Li. (May 2017). "A cascadedconvolutional neural network for X-ray low-dose CT image denoising."

[16] Y. Han, J. Yoo, and J. C. Ye. (Nov. 2016). "Deep residual learningfor compressed sensing CT reconstruction via persistent homologyanalysis."

[17] P. Liu and R. fang. (Aug. 2017). "Wide inference network for imagedenoising via learning pixel-distribution prior."

[18] X.-J. Mao, C. Shen, and Y.-B. Yang. (Jun. 2016). "Image restorationusing convolutional auto-encoders with symmetric skip connections."

[19] W. Xie and Y. Li, "Hyperspectral imagery denoising by deep learningwith trainable nonlinearity function," *IEEE Geosci. Remote Sens. Lett.*,vol. 14, no. 11, pp. 1963–1967, Nov. 2017.

[20] C. Peng, X. Zhang, G. Yu, G. Luo, and J. Sun. (Mar. 2017). "Largekernel matters—Improve semantic segmentation by global convolutionalnetwork."

[21] D. Liu, B. Wen, X. Liu, Z. Wang, and T. S. Huang. (Sep. 2018).“When Image denoising meets high-level vision tasks: A deep learningapproach.”

[22] K. He, X. Zhang, S. Ren, and J. Sun. (Feb. 2015). “Delving deep intorectifiers: Surpassing human-level performance on ImageNet classification.”

[23] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, “Beyond aGaussian Denoiser: Residual learning of deep CNN for image denoising,”*IEEE Trans. Image Process.*, vol. 26, no. 7, pp. 3142–3155,Jul. 2017.

[24] C. Jiang, H. Zhang, L. Zhang, H. Shen, and Q. Yuan, “Hyperspectralimage denoising with a combined spatial and spectral weighted hyperspectraltotal variation model,” *Can. J. Remote Sens.*, vol. 42, no. 1,pp. 53–72, Apr. 2016.

[25] Q. Yuan, Q. Zhang, J. Li, H. Shen, and L. Zhang, “Hyperspectralimage denoising employing a spatial–spectral deep residual convolutionalneural network,” *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 2,pp. 1205–1218, Feb. 2019.