electronics

Article

Hyperspectral Image Shadow Enhancement Using Three-Dimensional Dynamic Stochastic Resonance and Classification Based on ResNet

Xuefeng Liu [1,†], Yangyang Kou [1,†] and Min Fu [2,3,*]

[1] College of Automation and Electronic Engineering, Qingdao University of Science and Technology, Qingdao 266061, China; 02937@qust.edu.cn (X.L.); 2021040011@mails.qust.edu.cn (Y.K.)
[2] Sanya Oceanographic Institution, Ocean University of China, Sanya 572024, China
[3] College of Electronic Engineering, Ocean University of China, Qingdao 266100, China
* Correspondence: fumin@ouc.edu.cn
† These authors contributed equally to this work.

Abstract: Classification is an important means of extracting rich information from hyperspectral images (HSIs). However, many HSIs contain shadowed areas, where noise severely affects the extraction of useful information. General noise removal may lead to loss of spatial correlation and spectral features. In contrast, dynamic stochastic resonance (DSR) converts noise into capability that enhances the signal in a way that better preserves the image's original information. Nevertheless, current one-dimensional and 2D DSR methods fail to fully utilize the tensor properties of hyperspectral data and preserve the complete spectral features. Therefore, a hexa-directional differential format is derived in this paper to solve the system's output, and the iterative equation for HSI shadow enhancement is obtained, enabling 3D parallel processing of HSI spatial–spectral information. Meanwhile, internal parameters are adjusted to achieve optimal resonance. Furthermore, the residual neural network 152 model embedded with the convolutional block attention module is proposed to diminish information redundancy and leverage data concealed within shadow areas. Experimental results on a real-world HSI demonstrate the potential performance of 3D DSR in enhancing weak signals in HSI shadow regions and the proposed approach's effectiveness in improving classification.

Keywords: remote sensing image; residual neural network; shadow region; three-dimensional dynamic stochastic resonance

check for updates

Citation: Liu, X.; Kou, Y.; Fu, M. Hyperspectral Image Shadow Enhancement Using Three-Dimensional Dynamic Stochastic Resonance and Classification Based on ResNet. *Electronics* **2024**, *13*, 500. https://doi.org/10.3390/electronics13030500

Academic Editor: Chiman Kwan

Received: 20 December 2023
Revised: 20 January 2024
Accepted: 21 January 2024
Published: 24 January 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The rapid development of remote sensing images has sparked a growing interest in utilizing satellite technology for monitoring the Earth's surface. Nowadays, remote sensing techniques have advanced significantly and can be applied to various task scenarios, requiring high precision to ensure meaningful results [1].

The hyperspectral image (HSI) is a relatively new and continuously evolving remote sensing technology. It was initially developed for Earth observation purposes [2]. It captures the spectral information of objects within hundreds of narrow and contiguous spectral bands. Compared to traditional color imaging, HSI provides richer spectral information, enabling the discovery of hidden features of objects that are imperceptible to the human eye. Consequently, it enhances the accuracy of classification, detection, tracking processes, and material identification. By integrating imaging with spectroscopy, HSI can cater to the diverse needs of various industries. Utilizing HSI in remote sensing applications can aid in anomaly detection [3–6], space exploration [7], fire detection [8], food safety and quality control [9–13], boundary monitoring [14], chemical agent detection [15], archaeological preservation and authentication [16,17], change detection [18], clinical diagnosis [19], and mineral exploration [20]. Among these applications, target detection [21] and land cover classification [22,23] are particularly widespread.

The noise present in shadow regions of HSI can diminish the information extraction capability of images and disrupt the output of nonlinear systems. However, under certain conditions, noise can actually regularize the system's output, amplify signal amplitudes, and improve image quality. In the field of image processing, Ryu et al. [24] enhanced fingerprint images with low perceptual quality by introducing Gaussian noise to the signal, thus improving feature extraction. Rallabandi et al. [25] reduced inherent noise in magnetic resonance images by adding external noise in the Fourier domain. Maragatham et al. [26] utilized stochastic resonance (SR) to add noise to dark regions, thereby enhancing low-contrast images. The results demonstrate that appropriately adding noise to noisy images benefits signal enhancement.

Dynamic stochastic resonance (DSR) is a phenomenon that amplifies weak input signals in nonlinear systems by adding noise. It can be applied to address the shadow regions formed due to uneven illumination in HSIs [27]. The unique aspect of this method lies in using noise that would typically degrade image quality to enhance image perceptibility. It has been widely applied in various fields, primarily for processing time series signals. When noise resonates with a bistable nonlinear system and a periodic signal, the noise that typically degrades image quality is transformed into capability that enhances the signal, thereby improving the system's output. The DSR method proposed by Wang [28] requires converting three-dimensional (3D) HSIs into one-dimensional (1D) sequences, performing multiple iterations on each element of the sequence, and applying the Discrete Cosine Transform (DCT) to reconstruct the 1D sequence into a two-dimensional (2D) matrix. This process is complex and does not fully utilize the inter-pixel correlation in HSIs. Liu proposed the two-dimensional dynamic stochastic resonance (2D DSR) [29] for spatial domain enhancement of HSI by directly inputting spatial information into bistable nonlinear systems for DSR. Although this approach enhances spatial pixel correlation, it loses spectral information.

One of the most important tasks in HSI analysis is classification, aiming to assign each pixel in HSI to a specific class. HSI classification is challenging due to several reasons. Firstly, shadow regions in HSI contain fewer samples, potentially leading to inaccurate predictions due to insufficient training data. This poses challenges for effective data processing and analysis. Secondly, HSI data are high-dimensional and highly correlated. These characteristics make traditional machine learning methods less effective and can diminish the performance of classification algorithms.

Müller et al. [30] introduced a seeded-region growing algorithm grounded in form features to effectively extract and segment buildings. Leveraging the characteristic right-angle construction of buildings, the utilization of form features aids in distinguishing buildings from natural elements, thereby enhancing the precision of building recognition. Moreover, given the commonplace occurrence of shadows adjacent to buildings, shadow extraction serves as a valuable indicator for determining building locations. However, it is worth noting that the method primarily considers shadow regions with straight boundaries, limiting its applicability solely to aerial images. HSI, on the other hand, encompasses various irregular targets, including natural objects such as trees, which exhibit irregular shadows. Consequently, this restricts the method's suitability for image segmentation within HSI datasets. Furthermore, the method's primary focus lies in segmenting buildings and does not extend its processing capabilities to targets within shadows.

Convolutional Neural Networks (CNNs) have gained widespread application in image classification owing to their robust feature extraction capabilities. Maltezos et al. [31] introduced an efficient deep learning framework based on CNNs, leveraging height information derived from point clouds in orthorectified images and dense image matching for accurate building extraction. Experimental results demonstrate that the combination of raw image data and height information enhances the resilience of automatic building detection. However, this approach proves less effective in scenarios featuring multiple targets. In a related effort, Makantasis et al. [32] in 2015 proposed an enhanced CNN and multilayer perceptron capable of autonomously constructing hierarchical high-level spectral–spatial

features for classification. They introduced a random principal component analysis method to mitigate similarities among pixels with comparable spectral features, thereby refining classification accuracy in hyperspectral imaging (HSI). This improvement is achieved by iteratively updating training parameters through the backpropagation method, optimizing values for enhanced performance. While this method significantly expedites the training and prediction processes, it effectively harnesses both spatial and spectral correlations. It is worth noting, however, that the presence of noisy shadow regions in HSI may lead to the removal of valuable spectral features, potentially impacting classification outcomes.

Singh et al. [33] introduced a high-precision model, Ps-ProtoPNet, based on ResNet-34, achieving an accuracy of 99.79% in fault classification. However, it is only applicable to large datasets of unmanned aerial vehicle captured images. There are certain limitations for the HSI shadow regions with fewer samples and strong noise. Zhao et al. [34] proposed a cycle-consistent adversarial network for shadow compensation. This unsupervised method automatically transfers spectra from shadow to non-shadow regions, reducing the need for training samples. Nonetheless, the partial compensation effect at shadow boundaries proved unsatisfactory, significantly impacting subsequent classification accuracy in experiments on the selected dataset. The method presented in this paper effectively addresses or mitigates the issue. It only requires a small sample of shadow region images for information extraction, which offers a computational time advantage.

To address the challenges of shadow enhancement and classification in HSI, in this study, we propose a DSR-based HSI enhancement algorithm called three-dimensional dynamic stochastic resonance (3D DSR). We embed the Convolutional Block Attention Module (CBAM) into the Residual Neural Network 152 (ResNet152) model to enhance classification performance. Firstly, to preserve the characteristics of the 3D tensor, we derive the differential equations and difference solution equations for the 3D DSR model based on the model derivation of DSR, enabling the direct input of 3D data. To ensure appropriate processing of shadow regions with suitable parameters, we employ a dynamically adjusted iterative approach to enhance the contrast of shadow regions. A threshold is defined to quickly select the optimal output. Next, the enhanced image and original components are fused. Finally, the enhanced image is input into the optimized ResNet-152 model for classification, and its performance is compared with other state-of-the-art architectures. Experimental results demonstrate that the proposed method has lower computational complexity, effectively exploits shadow information, and improves the classification accuracy of HSI by facilitating correct classification of target pixels.

The remaining sections of this paper are organized as follows: Section 2 provides a review of the DSR algorithm and ResNet152, explaining in detail the model and main implementation process of the proposed method. This includes the derivation of the 3D DSR model equation based on DSR and the embedding of CBAM into ResNet152. Section 3 presents a detailed record of the experimental results, while Section 4 analyzes and discusses these results. Lastly, Section 5 concludes this work.

2. Materials and Methods

2.1. Materials

2.1.1. Dynamic Stochastic Resonance

Stochastic resonance (SR) is a concept proposed by Benz et al. [35] in 1981 to explain the glacial cycle. Its principle involves adding appropriate noise to the original sequential signal to amplify the signal's amplitude and increase the probability of surpassing a threshold at peak moments. The basic requirements for SR to occur in a system are low-amplitude signals, an energy threshold, and the presence of external or intrinsic noise sources.

DSR is an iterative form of SR that can describe the variations in pixel values during the image enhancement process. The 3D grayscale HSI can be represented as a function S = f (x, y, z), where x, y, and z represent spatial coordinates along the x-axis, y-axis, and z-axis, respectively. It is a continuous function. During image processing, it is necessary to sample a continuous grayscale image. Image sampling is the process of converting a

continuously varying image in space into a set of discrete points. In the context of 3D space, the continuous image is divided into an m × n × k grid, where each grid corresponds to a pixel. Similar to Benzi's explanation of the double-well model for interpreting the glacial ages, the pixel values in the image are macroscopically modeled as discrete kinetic parameters, akin to the position of particles in a double-well system. For a low-contrast image, the analogy states that the pixel is initially in a weak signal state (because of low-intensity value since image is low contrast, i.e., a subthreshold signal). The addition of the optimum amount of noise affects its transition to the strong signal state (high contrast), much like a particle making a transition from one well to another. Discrete image pixels are treated as discrete particles, with grayscale values corresponding to the positions of physical particles in Brownian motion. The contrast of the image is represented by a double-well potential. The positions of the particles can be analogous to the states of intensity values.

For the stochastic resonance of the signal, applying a weak periodic force to the system involves loading a sinusoidal signal of an excitation amplitude A (smaller than the transition threshold) into a bistable nonlinear system, extracting a signal of one period as the excitation signal. HSI itself can be sampled as a signal sequence, where the shadow region exhibits properties of strong noise and weak signal. Thus, the grayscale value of each pixel can serve as the input of a weak periodic force. When the excitation value is zero, it is assumed that the current system is in the equilibrium position of the left potential well as shown in Figure 1. At this point, the HSI is in its original state. As the excitation value increases, the particles will move upward from the equilibrium position but will not transition beyond the highest point of the potential well. The effect of noise will cause more and more instantaneous excitation values of the system to exceed the transition threshold. Under certain conditions, this can drive the particle across the highest point of the potential well towards another well, initiating a transition. When the particle transitions to another potential well and stabilizes, the image reaches its enhanced state in terms of contrast. Therefore, the stable value of the initial state and the stable value after transition can, respectively, represent the low contrast state and the enhanced state. Finding the optimal performance measure can be considered as the process of particles transitioning from one well to the other, i.e., from the input state (low-contrast state) to the enhanced state (high-contrast state). When the iteration count exceeds the optimal value, the particles transition from a stable inter-well state to an oscillating state. The movement of particles in the bistable nonlinear system is depicted in Figure 1, where ΔU represents the barrier height and has a size of $a^2/4b$ (a and b are system parameters).

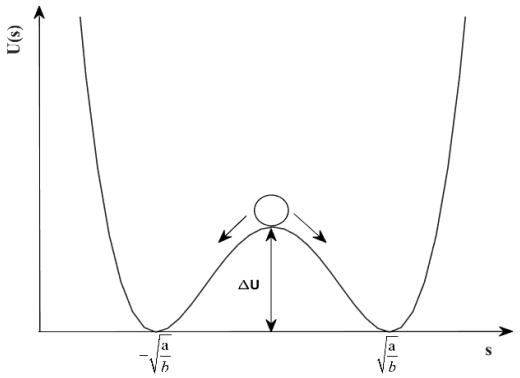


Figure 1. Potential well function U(s) for a bistable nonlinear system.

The change in state of a pixel under noise can be modeled by the Brownian motion of a particle in a double well as follows:

$$\frac{\mathrm{ds}^2(\mathrm{t})}{dt^2} = -\Omega\frac{\mathrm{ds}(\mathrm{t})}{dt} - \frac{dU(s)}{ds} + p(t) + \sqrt{D}\lambda(t), \tag{1}$$

where t represents time, Ω denotes damping factor, s(t) represents the position of the particle in the double well, $p(t)$ represents the periodic input signal, D represents the applied noise intensity, and λ represents the noise term. U(s) is the bistable fourth-order potential function in the SR theory. Under the overdamped condition, the second-order term is zero, and Equation (1) can be simplified to the above equation in Equation (2).

$$\begin{cases} \frac{\mathrm{ds(t)}}{dt} = -\frac{dU(s)}{ds} + p(t) + \sqrt{D}\lambda(t) \\ U(s) = -a\frac{s^2}{2} + b\frac{s^4}{4} \end{cases}, \tag{2}$$

where a and b represent two parameters of the bistable nonlinear system, both of which are greater than 0.

When an external periodic signal with amplitude A and frequency ω is input to the system, combined with the Langevin equation, Equation (2) will be transformed into

$$\frac{\mathrm{ds}}{dt} = as - bs^3 + A\sin(\omega t) + \sqrt{D}\lambda(t). \tag{3}$$

Some HSIs lack sufficient illumination and good perceptual quality, especially in poorly lit shadow regions, which are considered to be noisy. Therefore, when using DSR to process HSIs containing shadow areas, it is common to induce SR using the noise inherent in the image itself, referred to as internal noise-induced resonance. However, when dealing with images that have proper, uniform illumination and good perceptual quality, external noise needs to be added to the system to induce SR, known as external noise-induced resonance. The inherent noise in the grayscale distribution of images caused by a lack of illumination, together with the periodic input signal, constitute the driving force I(t). Equation (3) can be equated as

$$\frac{\mathrm{ds}}{dt} = as - bs^3 + \mathrm{I}(t). \tag{4}$$

By adjusting the system parameters a and b, the height of the potential barrier can be modified. When the system is subjected to a noisy input signal, the particle transitions from initial slight oscillations to vigorous oscillations. As the noise reaches a certain threshold, the particle is able to overcome the potential barrier and reach the other potential well, completing the process of enhancing weak signals.

When enhancing shadow areas in HSIs, Equation (4) should be transformed into an iterative form, namely DSR, to be applied

$$\mathrm{s}(n+1) = s(n) + h[as(n) - bs^3(n) + I(t)], \tag{5}$$

where $\mathrm{s}(n+1)$ represents the output of the system after n iterations, n denotes the number of iterations. The variable h represents the step size of the iteration, with $\mathrm{h} = 1/\mathrm{f_s}$, where $\mathrm{f_s}$ is the sampling frequency.

2.1.2. Deep Residual Neural Network

Deep learning-based methods [36,37] have shown great potential in the field of image classification. CNN, as one representative approach, exhibits strong feature extraction capabilities and has achieved remarkable performance in various tasks, especially in image classification. However, as the depth of CNNs increases, they often suffer from the issues of gradient vanishing and overfitting, making the network difficult to train and optimize. To address these problems, He et al. [38] introduced the concept of the Residual Neural Network (ResNet) in 2015. The design of ResNet aims to facilitate the training of extremely deep neural networks and overcomes the degradation problem of deep neural networks. They introduced the concept of residual learning, which involves learning residual functions instead of directly learning the underlying mapping. ResNets also introduced skip connections, allowing information to flow directly between non-adjacent layers in the

network, making it easier to optimize deeper networks. The use of residual connections in ResNet allows the preservation of original features, providing a greater advantage in handling shadow regions in HSI. This can further enhance the accuracy and generalization capabilities of the model.

ResNet152 is the deepest and most powerful variant in the ResNet architecture, consisting of 152 layers. Its model has been widely used in various computer vision applications, including image recognition, object detection, and semantic segmentation. However, HSI classification methods based on ResNet generate feature maps that contain a significant amount of spatial information redundancy, leading to decreased classification accuracy. Therefore, we optimize the network structure of the ResNet152 classification model by incorporating CBAM to improve classification accuracy.

Attention modules have been shown to enhance the feature extraction and generalization performance of neural networks [39,40]. CBAM can be directly embedded into current mainstream CNN network structures, enhancing the feature extraction capabilities without significantly increasing computational and parameter requirements. CBAM improves the weighting of important information by explicitly modeling the correlation between channel and spatial positions.

The module consists of two parts: the channel attention module (CAM) and the spatial attention module (SAM). CAM captures the importance of each feature channel, while SAM captures the importance of each spatial position. The network structure of CBAM is illustrated in Figure 2.

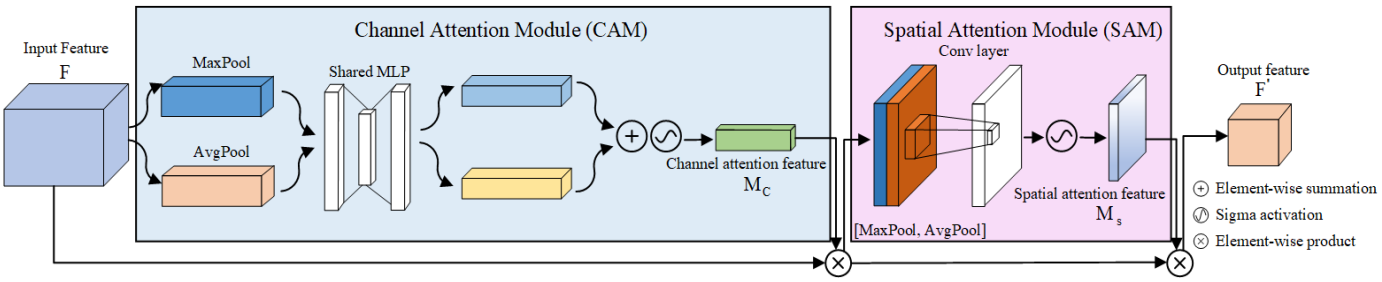


Figure 2. Main structure of Convolutional Block Attention Module.

In the CAM, global average pooling and global max pooling are applied to the input feature map to generate two feature vectors. These feature vectors are then passed through two fully connected layers (FC) for dimensionality reduction and non-linear transformation, resulting in channel correlations and reducing the number of parameters in the network. The output of the fully connected layers is further processed by a sigmoid activation function, generating a channel-wise attention map. This map is used to reweight the feature channels in the input feature map, enhancing focus on crucial channel features.

In the SAM, a similar process is followed, but the input feature map undergoes convolutional operations. This operation produces a spatial attention map, which is used to reweight the spatial positions in the input feature map, enhancing focus on crucial spatial features.

2.2. Proposed Approach Based on Three-Dimensional Dynamic Stochastic Resonance and ResNet152 Embedded with CBAM

The method of handling HSI by DSR involves sampling images at regular intervals to form a sequence undergoing resonance. In 2D DSR, images are sequentially processed in the spatial dimension. Both of these methods, however, do not account for the spatial–spectral characteristics of three-dimensional HSI. Therefore, in this paper, the derivation of 3D DSR is presented to address this issue. The commonly employed HSI classification method, known as the residual network, can be utilized to evaluate the capability of enhancing information features in the shadow region of HSI by 3D DSR. In this paper, the

introduction of an attention mechanism into the ResNet152 is proposed to extract crucial features, thereby achieving a balance between time and accuracy.

2.2.1. Derivation of Three-Dimensional Dynamic Stochastic Resonance

Currently, within the domain of HSI processing, DSR is constrained to employing, at most, 2D data as input for nonlinear systems, thereby overlooking valuable spectral information. In this work, the input of DSR's 1D nonlinear system is elevated to a 3D counterpart that incorporates spectral parameters.

An advancement is achieved by deriving hexa-directional difference equations based on the bidirectional differencing method. Three-dimensional DSR achieves the direct input of 3D hyperspectral data, eliminating the need for the sequential sampling of spatial information into a one-dimensional signal sequence and the subsequent reconstruction of the one-dimensional sequence into a two-dimensional matrix, as required by DSR and 2D DSR. Simultaneously, the introduction of the six-directional difference equation enables parallel processing of spatial and spectral information from six different endpoints, simplifying the computational process and enhancing efficiency. The extension of DSR and the incorporation of hexa-directional difference equations offer one way of thinking about the direct processing of 3D data. By leveraging the tridimensional nature of HSI, our methodology enables the comprehensive utilization of spectral–spatial information, effectively addressing the limitations inherent in conventional DSR model.

Generally, in 2D DSR, the system output is solved using bidirectional differencing, which can be expressed as

$$\begin{cases} s(x,y) = \Delta t_x \cdot \left[as(x,y-1) - bs^3(x,y-1) + I(x,y-1)\right] + s(x,y-1) \\ s(x,y) = \Delta t_y \cdot \left[as(x-1,y) - bs^3(x-1,y) + I(x-1,y)\right] + s(x-1,y) \end{cases}, \tag{6}$$

where (x, y) represents the spatial position of the pixel, Δt_x and Δt_y are the sampling intervals in the horizontal and vertical directions, and $s_{x,y}$ represents the output pixel value. The differencing is performed in parallel in both the left-to-right and top-to-bottom directions.

To achieve DSR in 3D space, we extend Equation (6) into six-direction differencing, simultaneously enhancing the input in the spectral dimension forwards and backwards, and in the spatial dimensions horizontally and vertically. Our derived six-direction difference equation suitable for solving the 3D DSR problem is shown in Equation (7).

$$\begin{cases} s(x,y,z)(n+1) = \Delta t_x \cdot \left[as(x,y-1,z)(n) - bs^3(x,y-1,z)(n) + I(x,y-1,z)\right] + s(x,y-1,z)(n) \\ s(x,y-1,z)(n+1) = \Delta t_x \cdot \left[as(x,y,z)(n) - bs^3(x,y,z)(n) + I(x,y,z)\right] + s(x,y,z)(n) \\ s(x,y,z)(n+1) = \Delta t_y \cdot \left[as(x-1,y,z)(n) - bs^3(x-1,y,z)(n) + I(x-1,y,z)\right] + s(x-1,y,z)(n) \\ s(x-1,y,z)(n+1) = \Delta t_y \cdot \left[as(x,y,z)(n) - bs^3(x,y,z)(n) + I(x,y,z)\right] + s(x,y,z)(n) \\ s(x,y,z)(n+1) = \Delta t_z \cdot \left[as(x,y,z-1)(n) - bs^3(x,y,z-1)(n) + I(x,y,z-1)\right] + s(x,y,z-1)(n) \\ s(x,y,z-1)(n+1) = \Delta t_z \cdot \left[as(x,y,z)(n) - bs^3(x,y,z)(n) + I(x,y,z)\right] + s(x,y,z)(n) \end{cases}, \tag{7}$$

where (x, y, z) represents the position of a pixel in 3D space, and Δt_z represents the sampling interval in the spectral direction.

Equation (7) represents the iterative formula for 3D hyperspectral shadow data enhancement. Through this computation, the 3D parallel processing of hyperspectral shadow data is achieved.

2.2.2. ResNet152 Embedded with CBAM

Directly classifying HSI using ResNet152 can result in redundancy of spatial and spectral information, which can negatively impact classification accuracy. After embedding CBAM into ResNet152, the feature maps can be dynamically calibrated through spatial and channel attention mechanisms, reducing information redundancy and enhancing the discriminative power of the model. Average pooling and max pooling can be employed to aggregate information from feature maps, which is then fed into a shared network. This process compresses the spatial dimensions of the input feature map, merging them

through element-wise summation to generate spatial and channel attention maps. These attention maps are multiplied with the input feature map for adaptive feature optimization, thereby reweighting channel and spatial features in the input feature map. The output of the convolutional layer undergoes a channel attention module first, yielding weighted results. Subsequently, it passes through a spatial attention module before obtaining the final optimized result through further weighting. This enables ResNet152 to focus on important channels and spatial regions. Additionally, CBAM effectively captures both local and global contextual information, thereby improving the accuracy of feature extraction. By leveraging the attention mechanism of CBAM, the classification performance of ResNet152 can be enhanced, enabling it to handle more complex classification tasks.

Figure 3 illustrates the main structure of CBAM-ResNet152. ResNet152 consists of three main parts: the input part, convolutional part, and output part. The convolutional part is represented as Phase1 to Phase4 in the figure. Each convolutional part consists of residual blocks and downsampling block. A residual block is depicted in Figure 4, where I denotes the input data, H(I) represents the identity mapping of I, and the output U is obtained by adding H(I) and I through a skip connection. A Rectified Linear Unit (ReLU) is chosen as the activation function. Two CBAMs are inserted into the middle layers of the convolutional part, specifically between the 13th and 14th layers in Phase2, as well as between the 14th and 15th layers. This choice is made because early convolutional layers have larger spatial feature maps and fewer channels, making it challenging for the extracted features to be representative. On the other hand, later convolutional layers have excessive channels, which may lead to overfitting, and they have a significant impact on the classification decision as they are closer to the fully connected (FC) layer. In the FC layer, we employ the softmax function to display the probability distribution of the classes.

2.2.3. Main Flow of the Proposed Algorithm

HSI is a 3D tensor, meaning it possesses the characteristics of a three-dimensional data structure. The DSR method enhances it by sampling as a sequential signal, thereby reducing the consistency and correlation with the original data. In order to preserve more image details, this paper proposes a 3D DSR method that directly takes the 3D image as input without disrupting the correlation of spatial–spectral pixels.

CBAM, through its channel and spatial attention mechanisms, can easily be integrated into the existing ResNet152 architecture, emphasizing important information while suppressing redundant information, thus improving HSI classification.

The main workflow of the algorithm is illustrated in Figure 5, and its main steps are as follows:

1. Perform linear normalization on the original HSI to transform the original grayscale distribution from [0, 255] to [0, 1]. Create a shadow region mask using ground truth and apply it to obtain the 3D shadow region of the original data.

2. Use the 3D shadow data as input to the 3D DSR model and enhance the image in both the spectral and spatial dimensions using the established 3D DSR model.

3. Adjust n, Δt_x, Δt_y, Δt_z, a, and b. Establish an iterative loop, compare the average value of the output image with the threshold multiplied by the input image's mean value to determine the optimal number of iterations and obtain the system output with the best performance metrics.

4. Extract the 3D non-shadow region of the original data using the mask and merge it with the enhanced shadow region to obtain the enhanced HSI.

5. Two CBAMs are embedded between the intermediate convolutional layers of ResNet152. Specifically, arrange them in the middle of the 13th and 14th layers in Phase2, as well as between the 14th and 15th layers. Each CBAM consists of a CAM and a SAM. CAM models channel dependencies, while SAM models spatial dependencies.

6. Use standard skip connections to connect the output of each CBAM to the input of the next layer in the ResNet152 architecture, ensuring the flow of information in the CBAM-enhanced layer of the network.

7. Train the modified ResNet152 model using the optimization algorithm of Stochastic Gradient Descent on a real-world hyperspectral dataset.

8. Evaluate the performance of the ResNet152 model embedded in CBAM on the test image set using overall classification accuracy (OA) and Kappa.

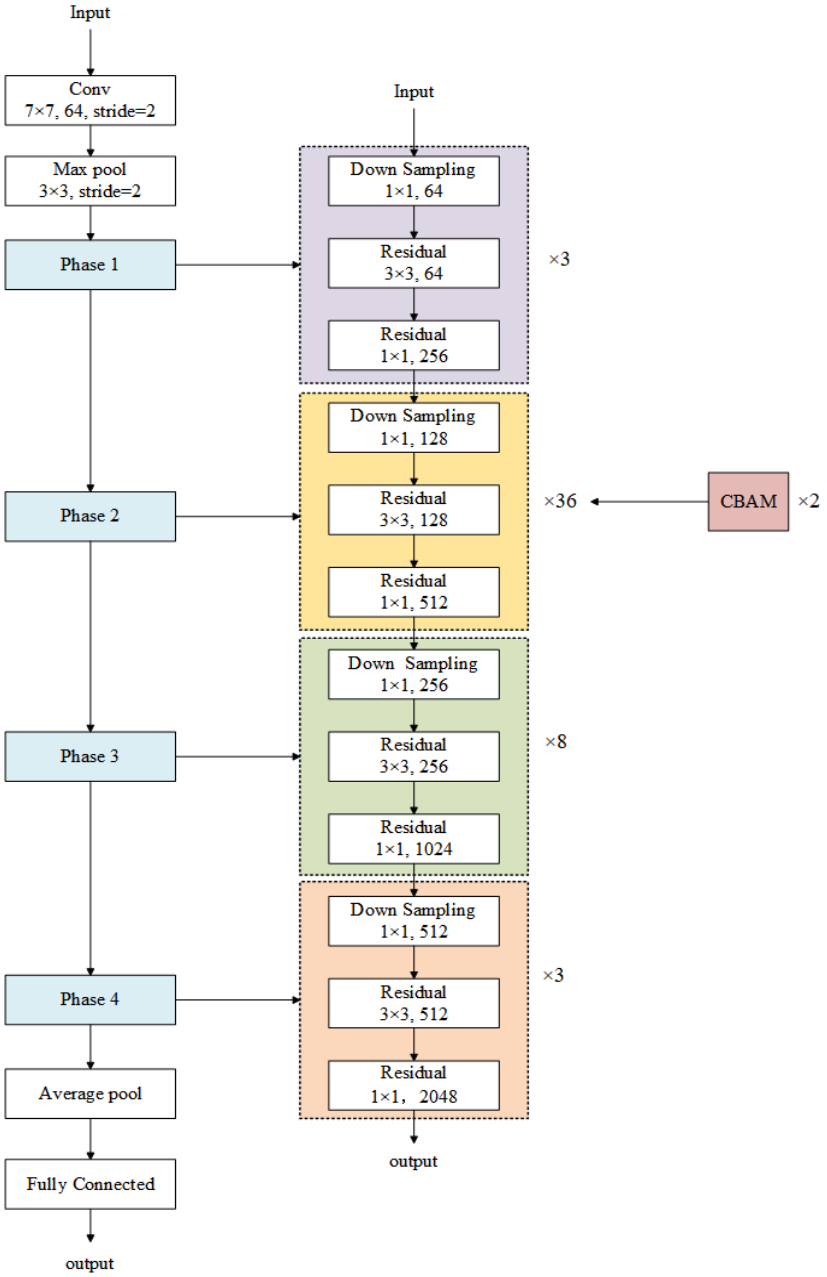


Figure 3. Main structure of CBAM-ResNet152.

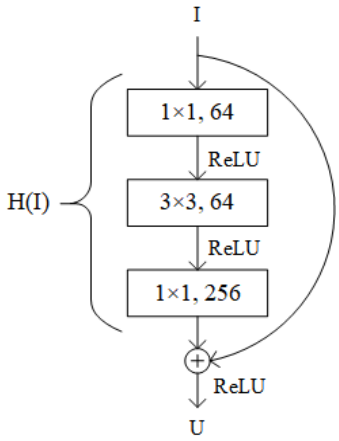


Figure 4. A residual block in ResNet152.

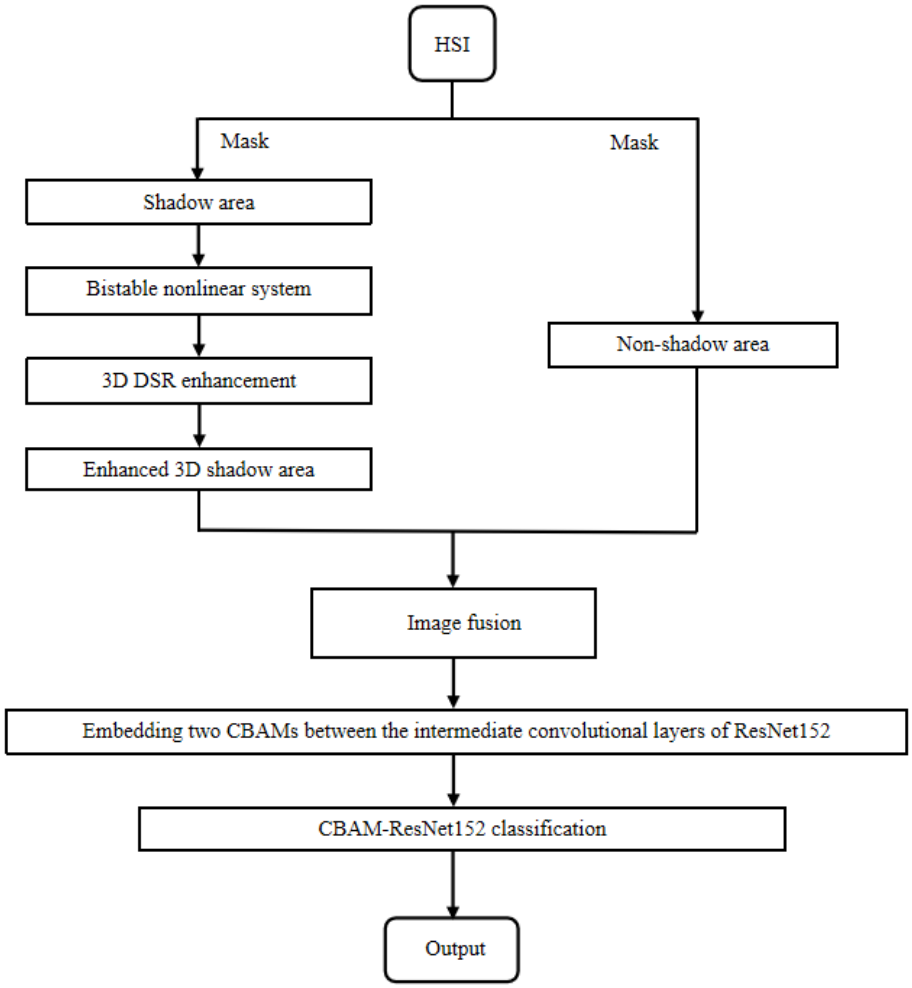


Figure 5. The main procedure of the proposed method.

3. Results

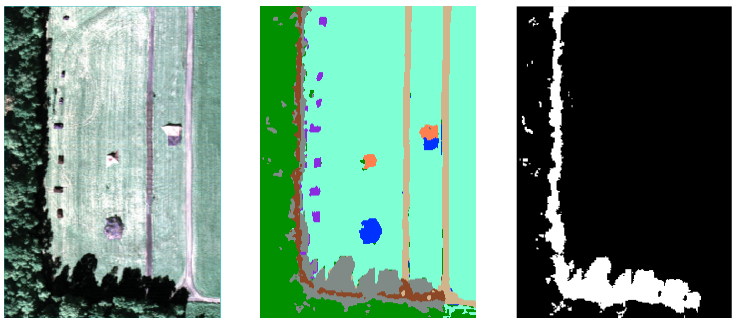
The current work was implemented on a computer with an Intel Core i5-6200U CPU and NVIDIA RTX 2080ti GPU, running at a clock speed of 2.30 GHz and equipped with 8 GB of memory. The algorithm was implemented using MATLAB 2021b and Python 3.7 in a Ubuntu 16.04 operating system.

3.1. Experimental Dataset

We selected the Hyperspectral Digital Imagery Collection Experiment (HYDICE) as our real-world hyperspectral dataset for the experiments. The HYDICE dataset contains continuous and distinct shadow regions that meet the requirements of our study. Figure 6a shows a scene from the HYDICE, Figure 6b presents the ground truth for this scene, and Figure 6c displays the shadow extraction mask created based on the ground truth. The spatial pixel size of the dataset is 316 × 216, and it consists of 148 spectral bands. The spatial and spectral resolutions are 0.75 m and 10 nm, respectively. The HYDICE comprises eight land cover classes: grassland, grassland under shadow, road, road under shadow, architecture, pits, car, and trees. To accurately differentiate each land cover class after classification, we represent them with eight distinct colors, as shown in Table 1.

Table 1. Sample categories and numbers of HYDICE.

No.	Color	Classes	Number of Pixels
1		Grassland	41,480
2		Tree	13,562
3		Road	4220
4		Road under shadow	2180
5		Architecture	671
6		Pits	404
7		Car	642
8		Grassland under shadow	5169
	All classes		68,256

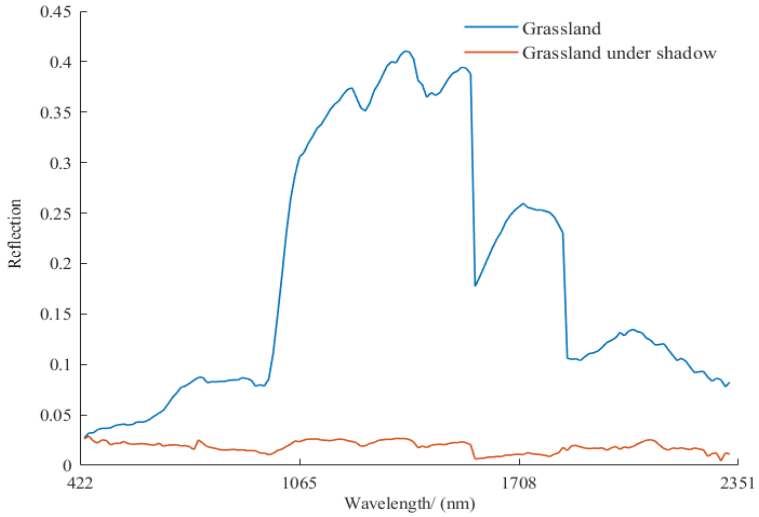


(**a**) (**b**) (**c**)

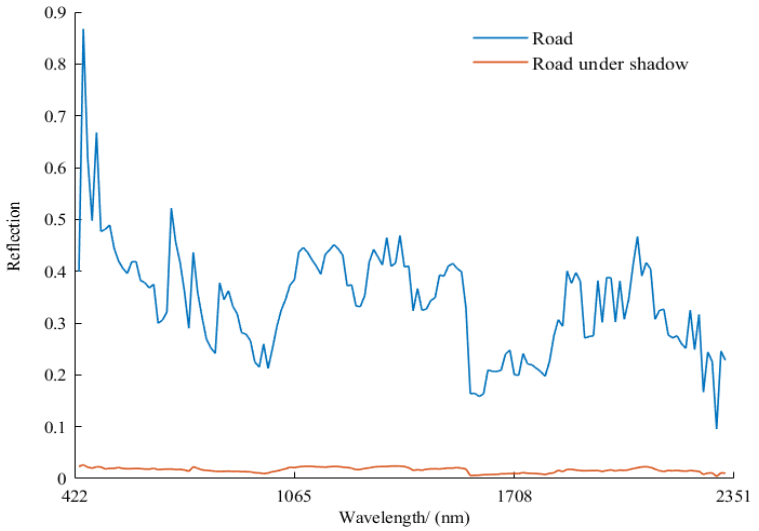
Figure 6. (**a**) A scene of HYDICE. (**b**) Ground truth. (**c**) Shadow extract mask.

3.2. Preprocessing

Figure 7a,b present the spectral curves of grassland and road in the shadow and non-shadow regions of the HYDICE, respectively. It can be observed that the spectral reflectance of grassland and road in the non-shadow region is higher compared to that in the shadow region, indicating higher illumination. Therefore, a method that enhances the shadow region is not suitable for the non-shadow region, as it would cause excessive enhancement. This is why a mask is applied to extract the shadow region and process it separately before enhancement.



(**a**)



(**b**)

Figure 7. Spectral profiles of different targets in the shadow and non-shadow regions. (**a**) Grassland in the shadow and non-shadow regions. (**b**) Road in the shadow and non-shadow regions.

3.3. Performance Evaluation

3.3.1. Evaluation Metrics for Enhanced Images

Performance metrics such as Peak Signal-to-Noise Ratio (PSNR), Mean Squared Error (MSE), and Structural Similarity Index Measure (SSIM) are not suitable for HSIs without a reference image. Therefore, we selected contrast enhancement measure (CEM) and discrete entropy (DE) as no-reference performance metrics, based on contrast and perceptual quality, respectively.

The shadow regions in HSI exhibit low brightness and contrast, making it difficult to visually discern the objects covered by shadows. CEM calculates the contrast difference based on the global standard deviation and mean values of the original and enhanced images [41]. A higher CEM indicates greater grayscale differences between shadow pixels, which leads to more discernible information. The formula for CEM is defined as follows:

$$CEM = \frac{Q_{out}}{Q_{\text{in}}}, \tag{8}$$

where Q_{out} and Q_{in} represent the contrast quality indices of the enhanced and original images, respectively. The formula for calculating Q is given by [42]:

$$Q = \frac{\sigma^2}{\mu}, \tag{9}$$

where σ and μ are the standard deviation and mean values of the images, respectively.

DE defines the informational content of the image [43]. It can be represented using the following equation:

$$DE = -\sum_{i=0}^{h-1} p(v)\log_2 p(v), \tag{10}$$

where $p(v)$ represents the probability of the grayscale value v occurring in the entire image dataset, and h denotes the total number of grayscale levels in the image. A higher entropy value indicates better image quality.

3.3.2. Evaluation Metrics for Image Classification

The effectiveness of HSI enhancement can be quantified by the accuracy of the classified image. OA and Kappa are selected as metrics to evaluate the accuracy of the classified image. OA and Kappa are defined as follows:

$$OA = \frac{1}{E}\sum_{r=1}^{g} x_r \times 100\% \tag{11}$$

$$Kappa = \frac{E\sum\limits_{r=1}^{g} x_r - \sum\limits_{r=1}^{g}(x_{uu}\cdot x_{pp})}{E^2 - \sum\limits_{r=1}^{g}(x_{uu}\cdot x_{pp})} \times 100\%, \tag{12}$$

where E and g represent the total number of samples and categories in the HSI, subscript r denotes the pixel position, x_r is the number of correctly classified test samples, x_{uu} is the number of true samples per category, and x_{pp} is the number of predicted samples per category. A higher OA indicates a larger number of pixels correctly classified in the image, while a larger Kappa implies a smaller error compared to random ordering classification.

3.4. 3D DSR Algorithm to Enhance HSI

3.4.1. Parameter Adjustment

In this study, we refer to the iterative nature of DSR to enhance HSI. According to the proposed 3D DSR model, we need to iterate in six directions simultaneously, which requires finding the optimal values for Δt_x, Δt_y, Δt_z, a, b and n.

A higher iteration count not only requires more time but may also lead to the problem of excessive saturation. However, if the iteration count is insufficient, the image cannot be maximally improved. Conventional image enhancement methods often use performance metrics as the termination condition for iteration, which can be time-consuming. To shorten the required time, we define a dynamic threshold to determine the optimal iteration count. The threshold starts with an initial value and increases continuously until the performance metrics and classification accuracy of the output image no longer significantly improve. After each iteration cycle, the average value of the 3D image output by the system is calculated and compared with the product of the defined threshold and the mean value of the original data, enabling the rapid selection of the optimal system output. The calculation of the average value is shown in Equation (13), and the specific workflow is illustrated in Figure 8. Let Threshold be the threshold value and n be the iteration count. I_{mean} and O_{mean}, respectively, represent the mean values of the input image and the enhanced image. The relationship between the size of Threshold and CEM and OA is shown in Table 2. In

this study, the optimal threshold is chosen as 10 times the average grayscale value of the shadow region in the original image, and a total of seven iterations are required.

$$\text{mean} = \frac{1}{E}\sum_{i=1}^{l}\sum_{j=1}^{w}\sum_{k=1}^{b} T(i,j,k), \tag{13}$$

where l, w, and b represent the width, height, and number of bands of the HSI, respectively. $T(i,j,k)$ represents the grayscale value of the pixel at position (i, j, k).

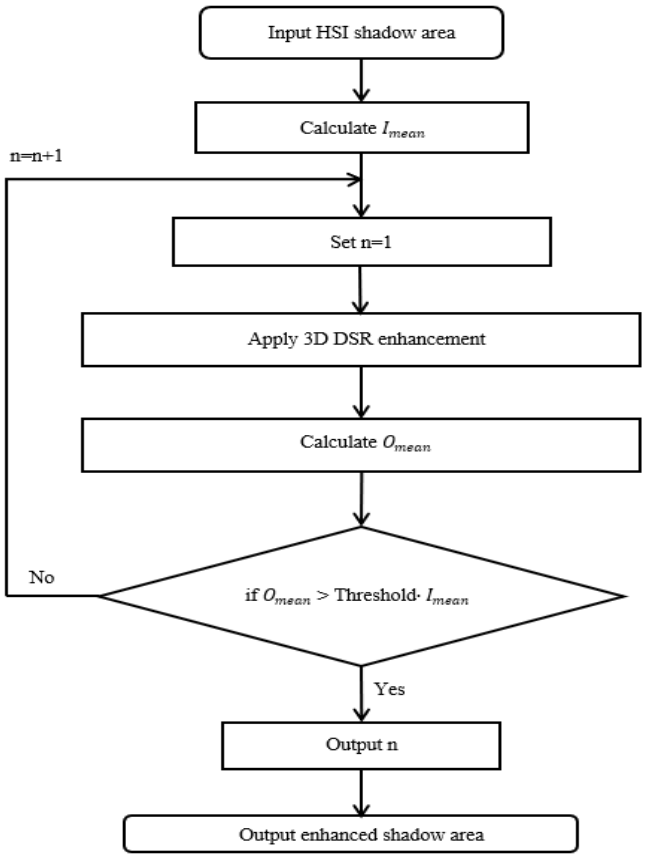


Figure 8. Main steps of parameter adjustment.

Table 2. Impact of dynamic threshold changes on CEM and OA.

Threshold	2	4	6	8	10	12	14	16
CEM	4.1255	6.9890	9.1089	11.6760	14.1348	10.8765	8.9898	6.3457
OA (%)	97.1898	97.2832	97.4180	97.5467	97.6012	97.4887	97.4002	97.2438

The restoring force is the gradient of the bistable potential well function U(s). To find the maximum possible value of a periodic signal, let the input periodic signal $M = C\sin\omega t = -\frac{dU(s)}{ds}$, which yields

$$\begin{cases} M = -\frac{dU(s)}{ds} = -as + bs^3 \\ \frac{dM}{ds} = -a + 3bs^2 = 0 \end{cases} \tag{14}$$

Solving for Equation (14), we obtain $s = \sqrt{(a/3b)}$. Thus the maximum possible force that keeps the bistable system stable is $M = \sqrt{(4a^3/27b)}$. Let the sine term $\sin\omega t$ be the

maximum value of 1, and for mathematical simplicity, we assume the amplitude C is also 1. Thus, we have

$$1 < \sqrt{\frac{4a^3}{27b}}. \tag{15}$$

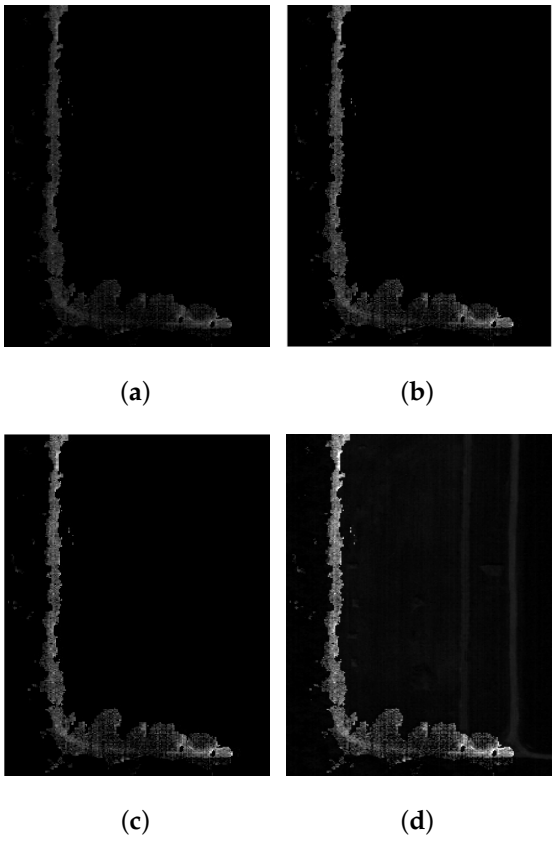
Therefore, enhancing weak signals using a bistable system requires $b < (4a^3/27)$.

Referring to the parameter setting principles used in S-DSR [35] and DWT-DSR [27], we ultimately select the system parameters of the 3D DSR model as a = 0.01 and b = $4a^3/27 \times 10^{-5}$.

After determining the values of a, b, and n, based on the univariate principle, we separately optimize the width sampling interval Δt_x, height sampling interval Δt_y, and spectral sampling interval Δt_z. The optimal values of Δt_x, Δt_y and Δt_z are finally determined to be 0.01.

3.4.2. Enhancement Results

The parameters of the 3D DSR model are set to their optimal values, with sampling frequencies of 100 in the horizontal, vertical, and spectral directions. Figure 9 shows the original shadow region image and the shadow region images enhanced using DSR, 2D DSR, and 3D DSR in the first band. Figure 10 displays the fused image of the non-shadow region from Figure 9 and the original image in the first band.



(**a**) (**b**)

(**c**) (**d**)

Figure 9. Image of the shadow area in the first band. (**a**) Original data. (**b**) Data after DSR enhancement. (**c**) Data after 2D DSR enhancement. (**d**) Data after 3D DSR enhancement.

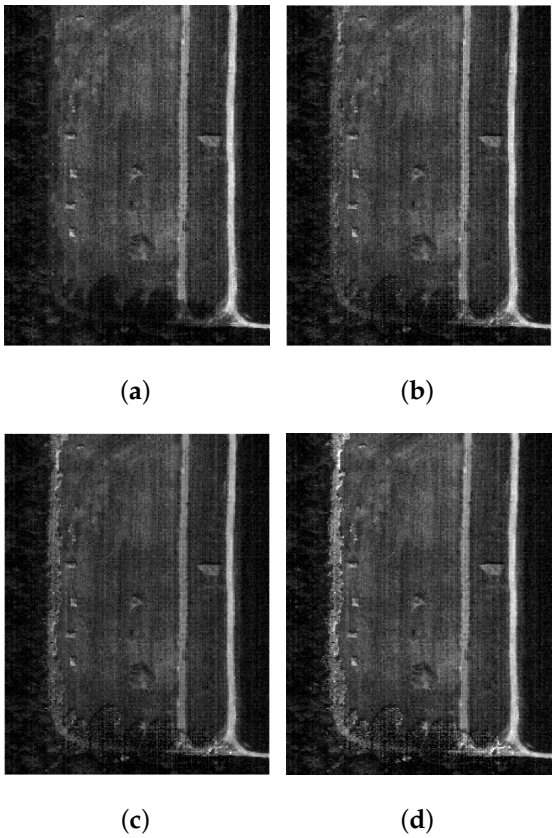


Figure 10. The fused image of the first band. (**a**) Original data. (**b**) Data after DSR enhancement. (**c**) Data after 2D DSR enhancement. (**d**) Data after 3D DSR enhancement.

Kan et al. [44] proposed an attention-based Octave dense network. This approach includes a new feature extractor with multiscale separable convolution to extract spatial–spectral features and a novel dense denoising block aided by Octave kernel and attention mechanism to extract high-frequency features containing more noise and details for noise suppression. The method can effectively reduce the low frequency redundancy and reduce the computational cost. The experimental results show that the OA and Kappa of Pavia University dataset after AODN denoising can reach 97.04% and 0.9609. Sidorov et al. [45] extended the deep prior algorithm to the HSI domain and implemented it using 3D convolution, which can efficiently recover single-hyperspectral-image contaminated by Gaussian noise, and show comparable performance to that of trained CNN even without training on any dataset. These methods are based on the classical Gaussian assumption, but the noise in the shadow region of real HSI is mostly random noise, so the results are less satisfactory in the case of multiple images with more complex noise composition.

To validate the effectiveness of the proposed method for shadow region enhancement, DSR and 2D DSR are selected as the comparative methods. The CEM, DE, and Time are computed for these methods before and after enhancement, as shown in Table 3.

Table 3. Evaluation indicators before and after enhancement by different methods.

Method	Original Data	DCT	WNNM	K-SVD	HSI-DeNet	DSR	2D DSR	3D DSR
CEM	0.0151	0.0108	2.0445	3.1709	6.8249	4.7184	10.3411	14.1348
DE	6.6967	6.4274	6.7014	6.7030	6.7811	6.7227	6.9413	7.1028
Time (s)	/	23	3688	5240	102	45	58	60

From the visual representations in Figures 9 and 10, it can be observed that DSR effectively improves the brightness and contrast of the HSI shadow region, while the

image enhanced with 3D DSR exhibits richer and clearer details. Furthermore, comparing the values in Table 3, the image enhanced with 3D DSR contains more information and achieves higher classification accuracy. Therefore, from both qualitative and quantitative perspectives, it can be concluded that 3D DSR outperforms 2D DSR and DSR in enhancing the information representation capability of the shadow region, thus contributing to the improvement of subsequent HSI classification tasks.

3.5. Classification Results

After enhancing the shadow region using the 3D DSR method, the enhanced 3D image can be classified using CBAM-ResNet152 to verify the improvement of HSI classification. The internal parameters used in the experiments are shown in Table 4. These parameters are crucial internal parameters within the classification model, and their values can influence the final classification accuracy.

Table 4. Parameters of the CBAM-ResNet152 model.

Epoch	Batch Size	Test Ratio	Learning Rate	Optimizer	Loss Function
100	312	0.80	0.001	Adam	Binary Cross-Entropy

To objectively evaluate the effectiveness and superiority of 3D DSR and CBAM-ResNet152 in processing HSIs, this paper selects six enhancement algorithms in the HSI field, namely K-SVD, DCT, WNNM, HSI-DeNet, DSR, and 2D DSR, as comparative methods for 3D DSR. In 2018, Chang et al. introduced HSI-DeNet, an image denoising method [46] that distinguishes itself by its capacity to directly learn layer-specific filters without compromising the spectral–spatial structure. This methodology exhibits commendable performance in both speed and capability. HSI-DeNet excels in preserving the spatial–spectral correlation inherent in HSI and demonstrates enhanced efficacy in eliminating random and streak noise. Its suitability for image recovery is noteworthy, particularly in scenarios involving degraded images, albeit with optimal results observed only within specific noise level parameters. Additionally, commonly used models in the field of HSI classification, such as 2D-CNN, HybridSN, 3D-ResNet, 3D-CNN, and ResNet152, are combined as comparative methods for CBAM-ResNet152. Of the samples, 20% are chosen as training data and 80% as test data. The accuracy of all categories, OA, and Kappa are calculated and averaged. The experimental results are shown in Tables 5–10.

Table 5. Accuracy of each category after 2D-CNN classification.

Method	Original Data	DCT	WNNM	K-SVD	HSI-DeNet	DSR	2D DSR	3D DSR
Field	0.9900	0.9900	0.9900	0.9900	0.9900	0.9900	0.9910	0.9900
Tree	0.9565	0.9605	0.9685	0.9650	0.9665	0.9650	0.9655	0.9655
Road	0.9610	0.9625	0.9705	0.9675	0.9675	0.9685	0.9675	0.9680
Road under shadow	0.8110	0.7985	0.7980	0.8020	0.8030	0.8060	0.8095	0.8115
Grassland under shadow	0.9215	0.8650	0.8525	0.9205	0.9230	0.9055	0.9355	0.9520
Architecture	0.8440	0.8625	0.8340	0.9245	0.9000	0.9010	0.9015	0.9020
Pits	0.7620	0.7585	0.7835	0.8530	0.8145	0.8095	0.8145	0.8265
Car	0.8715	0.8680	0.8705	0.8545	0.8700	0.8695	0.8480	0.8545
OA (%)	96.4372	96.5554	96.5782	96.5108	96.5431	96.5145	96.6033	96.6984
Kappa (%)	93.8634	94.0665	94.1129	93.9883	94.0220	94.0065	94.1480	94.2902

Table 6. Accuracy of each category after 3D-Resnet classification.

Method	Original Data	DCT	WNNM	K-SVD	HSI-DeNet	DSR	2D DSR	3D DSR
Field	0.9900	0.9885	0.9905	0.9900	0.9910	0.9900	0.9910	0.9910
Tree	0.9550	0.9675	0.9710	0.9785	0.9755	0.9745	0.9820	0.9845
Road	0.9600	0.9645	0.9490	0.9670	0.9550	0.9565	0.9570	0.9585
Road under shadow	0.7915	0.8030	0.7595	0.8350	0.8405	0.8395	0.8415	0.8430
Grassland under shadow	0.9205	0.9475	0.9410	0.9315	0.9400	0.9485	0.9500	0.9555
Architecture	0.8925	0.8955	0.9005	0.8970	0.8960	0.8415	0.8815	0.9025
Pits	0.7545	0.8780	0.7550	0.7785	0.7830	0.8055	0.7680	0.7740
Car	0.8655	0.8840	0.8355	0.8655	0.8700	0.8880	0.8880	0.8880
OA (%)	96.0322	96.5066	96.2681	96.6236	96.7051	96.5319	96.9832	97.3838
Kappa (%)	93.1632	93.9750	93.5688	94.1908	94.3230	94.0171	94.8016	95.3904

Table 7. Accuracy of each category after HybridSN classification.

Method	Original Data	DCT	WNNM	K-SVD	HSI-DeNet	DSR	2D DSR	3D DSR
Field	0.9800	0.9850	0.9800	0.9900	0.9895	0.9900	0.9880	0.9895
Tree	0.9705	0.9840	0.9800	0.9870	0.9860	0.9870	0.9860	0.9875
Road	0.9730	0.9555	0.9700	0.9655	0.9655	0.9665	0.9700	0.9715
Road under shadow	0.8105	0.8420	0.7285	0.8510	0.8465	0.8620	0.8680	0.8705
Grassland under shadow	0.9030	0.9170	0.9400	0.8985	0.9035	0.9390	0.9445	0.9490
Architecture	0.8615	0.8780	0.9305	0.8955	0.8985	0.8980	0.9015	0.9050
Pits	0.8440	0.8800	0.8650	0.7980	0.8020	0.8010	0.8095	0.8205
Car	0.8520	0.8810	0.8025	0.8850	0.8890	0.8880	0.8905	0.8945
OA (%)	96.2603	96.7680	96.3574	97.2183	97.2286	97.1156	97.2752	97.3901
Kappa (%)	93.5468	94.4272	93.7090	95.2090	95.2301	95.0217	95.2990	95.4106

Table 8. Accuracy of each category after 3D-CNN classification.

Method	Original Data	DCT	WNNM	K-SVD	HSI-DeNet	DSR	2D DSR	3D DSR
Field	0.9900	0.9900	0.9900	0.9900	0.9900	0.9900	0.9900	0.9900
Tree	0.9725	0.9720	0.9700	0.9810	0.9790	0.9770	0.9890	0.9885
Road	0.9700	0.9695	0.9715	0.9670	0.9710	0.9705	0.9735	0.9760
Road under shadow	0.8155	0.8510	0.7585	0.8520	0.8580	0.8630	0.8740	0.8860
Grassland under shadow	0.9195	0.8970	0.9110	0.8990	0.9230	0.9325	0.9435	0.9645
Architecture	0.9020	0.8845	0.9300	0.8965	0.8970	0.9165	0.8945	0.9100
Pits	0.7635	0.6940	0.8390	0.7085	0.7245	0.7695	0.7700	0.7810
Car	0.8925	0.9030	0.8125	0.9140	0.9100	0.9145	0.9095	0.9125
OA (%)	96.5047	96.8012	96.2831	97.4586	97.4720	97.1829	97.5146	97.7340
Kappa (%)	93.9265	94.4875	93.5896	95.6154	95.6294	95.1367	95.6849	96.1011

Table 9. Accuracy of each category after ResNet152 classification.

Method	Original Data	DCT	WNNM	K-SVD	HSI-DeNet	DSR	2D DSR	3D DSR
Field	0.9900	0.9900	0.9900	0.9900	0.9900	0.9900	0.9900	0.9900
Tree	0.9775	0.9745	0.9700	0.9830	0.9840	0.9795	0.9895	0.9895
Road	0.9710	0.9695	0.9710	0.9695	0.9720	0.9720	0.9740	0.9765
Road under shadow	0.8165	0.8525	0.7595	0.8540	0.8695	0.8635	0.8755	0.8870
Grassland under shadow	0.9005	0.8985	0.9120	0.8995	0.9195	0.9335	0.9445	0.9660
Architecture	0.9000	0.9000	0.9025	0.9020	0.9035	0.9080	0.9100	0.9100
Pits	0.8070	0.6800	0.8350	0.8245	0.8210	0.8215	0.8210	0.8255
Car	0.9005	0.9030	0.9025	0.9140	0.9145	0.9175	0.9195	0.9195
OA (%)	96.7055	96.7045	96.8034	97.4842	97.5206	97.2048	97.6176	97.8751
Kappa (%)	94.3235	94.3210	94.4893	95.6650	95.7471	95.1830	95.8943	96.2994

Table 10. Accuracy of each category after CBAM-ResNet152 classification.

Method	Original Data	DCT	WNNM	K-SVD	HSI-DeNet	DSR	2D DSR	3D DSR
Field	0.9900	0.9900	0.9900	0.9900	0.9900	0.9900	0.9900	0.9900
Tree	0.9780	0.9765	0.9785	0.9880	0.9895	0.9950	0.9900	0.9905
Road	0.9770	0.9780	0.9785	0.9745	0.9760	0.9790	0.9790	0.9795
Road under shadow	0.8215	0.8725	0.7990	0.8790	0.8790	0.8795	0.8815	0.8825
Grassland under shadow	0.9400	0.9135	0.9480	0.9270	0.9490	0.9500	0.9595	0.9605
Architecture	0.9100	0.9095	0.9420	0.9260	0.9245	0.9265	0.9280	0.9285
Pits	0.7870	0.7405	0.8400	0.8285	0.8280	0.8265	0.8330	0.8345
Car	0.9060	0.9130	0.9185	0.9260	0.9250	0.9275	0.9285	0.9295
OA (%)	96.9013	96.9002	97.3084	97.5773	97.6154	97.6020	97.8472	97.9983
Kappa (%)	94.7391	94.6938	95.4890	95.8200	95.9737	96.3142	96.4075	96.6502

4. Discussion

4.1. Discussion of 3D DSR Effect on Shadow Enhancement

In order to compare the impact of the 3D DSR method with the DSR and 2D DSR methods on the reflectance of HSI spectral curves and fidelity, we calculated the average grayscale values of 100 pixels in different regions to determine the reflectance of the spectral curves. Figure 11 presents the spectral curves of the road and grassland under different conditions. According to Figure 11, it can be observed that the 3D DSR method effectively improves the reflectance of the spectral curves compared to other methods, and it also exhibits higher fidelity in preserving the spectral characteristics.

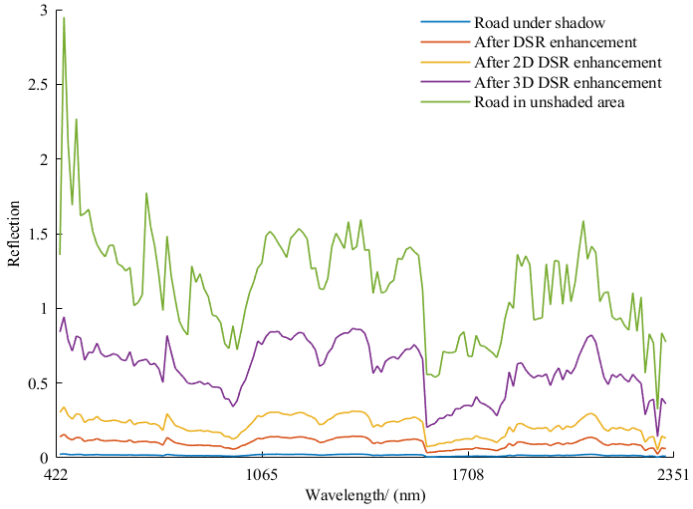
4.2. Discussion of CBAM-ResNet152 Effect on Classification

Firstly, by vertically comparing the results in the columns of Tables 5–10, it can be observed that when the training sample data are set at 20%, the images enhanced by 3D DSR exhibit relatively higher OA and Kappa coefficients compared to the considered comparative methods across different classification models, demonstrating strong robustness. Particularly, when classifying grassland and road pixels in shadow regions, the target pixels achieve more accurate classification after being enhanced by 3D DSR. In the shadow region, high noise also contains more image details, HSI-DeNet performs better in denoising but inevitably loses some image details. According to the classification results shown in Figure 12, the 3D DSR-enhanced images are closer to the ground truth after undergoing CBAM-ResNet152 classification. Therefore, 3D DSR based on 3D data augmentation can improve the feature expression capability of HSIs, better preserve the original image information, and enhance subsequent classification accuracy, playing a crucial role in the field of image enhancement.

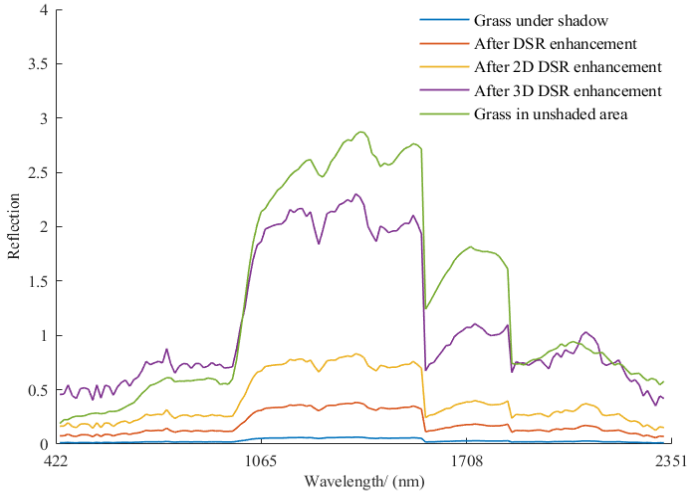
Secondly, by horizontally comparing the OA and Kappa values in Tables 5–10, it can be observed that the CBAM-ResNet152 based on the CNN model outperforms 2D-CNN, HybridSN, 3D-ResNet, 3D-CNN, and ResNet152 in HSI classification. When comparing the classification accuracy of different categories among the various classification models, although they all belong to CNN-based classification algorithms in deep learning, the ResNet-152 method embedded with CBAM performs better on grassland and road, which are two categories of small-sample in shadow regions. This confirms that CBAM-ResNet152 is more suitable for classifying small-sample hyperspectral data.

Lastly, comparing the classification accuracy in Tables 9 and 10, it can be observed that CBAM improves the performance of ResNet-152, resulting in an average increase of 0.12% and 0.36% in the OA and Kappa coefficients, respectively. CBAM achieves better classification results for ResNet152 by reevaluating the feature maps through channel and spatial attention modules. Additionally, the classification accuracy of architecture, pits, and car also improves, further validating the advantages of CBAM-ResNet152 in small-sample classification.

Table 11 shows the running time of all classification methods. From the table, it can be observed that the increase in runtime for CBAM-ResNet152, relative to the improvement in accuracy, is acceptable. In fact, the runtime increase is even lower than that of 3D-CNN. Considering all the runtime durations of HSI enhancement methods discussed in Table 3, it is evident that the enhancement method DCT, coupled with the classification method 3D-Resnet, has the shortest processing time. However, apart from surpassing WNNM in accuracy, DCT has lower precision than other enhancement methods. The classification accuracy of 3D-Resnet, although higher than that of 2D-CNN, is lower compared to other classification methods. WNNM, combined with the classification method 3D-CNN, exhibits the longest runtime, and the combined accuracy is the lowest among them. Upon evaluating the metrics in Table 3 and the combined accuracy of all enhancement methods with classification methods in Tables 5–10, it becomes apparent that the accuracy of 3D DSR, when combined with CBAM-ResNet152, surpasses that of other considered methods and is higher than that of DCT and 3D-Resnet. Overall, the average classification accuracy of images enhanced by 3D DSR is 0.3% higher than that of the closest 2D DSR, with a runtime only 2 s longer than 2D DSR, falling within an acceptable range.



(**a**)



(**b**)

Figure 11. Comparison of spectral profiles before and after enhancement. (**a**) Road. (**b**) Grass.

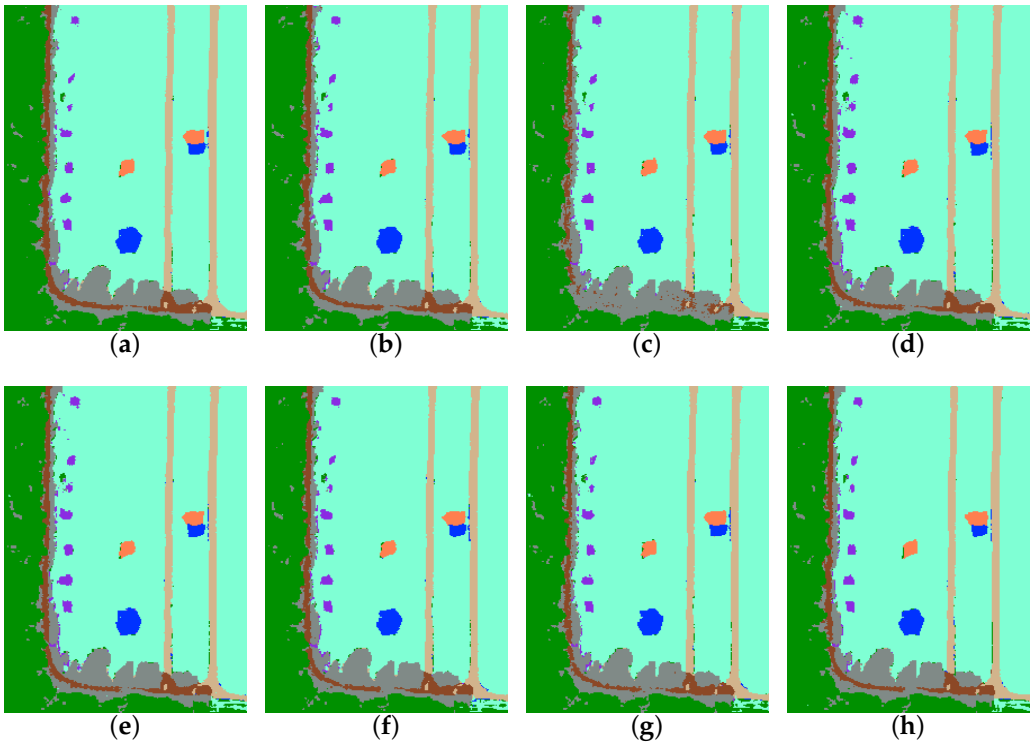


Figure 12. CBAM-ResNet152 classification results: (**a**) shows the classification results of the original data, and (**b**–**g**) present the classification results of the enhanced images using DCT, WNNM, K-SVD, HSI-DeNet, DSR, and 2D DSR, respectively; and (**h**) shows the classification results after 3D DSR enhancement.

Table 11. Running Time Comparison of Different Classification Methods.

Method	2D-CNN	3D-Resnet	HybridSN	3D-CNN	ResNet152	CBAM-ResNet152
Time (min)	51.75	51.38	53.20	61.46	60.11	60.02

In conclusion, the proposed method based on 3D DSR and CBAM-ResNet152 demonstrates excellent performance in both image enhancement and classification, enhancing the image representation of HSI and exhibiting potential for extracting information from small-sample HSI containing different targets.

5. Conclusions

Due to uneven illumination and object occlusion, shadow regions inevitably exist in HSI. The presence of strong noise and weak signals in these shadow regions poses challenges for image classification. The 3D DSR model proposed in this study overcomes the limitation of traditional DSR models, which can only handle 1D signal sequences, enabling the direct processing of 3D HSI. The hexa-directional difference equation derived based on the bidirectional difference method can fully leverage the spatial–spectral characteristics of HSI, effectively enhancing signals in shadow regions, thereby improving the accuracy of HSI classification.

In the preliminary phase, the system model of 3D DSR is formulated based on the DSR model. The derivation of the hexa-directional difference format addresses the system output, followed by the acquisition of the iterative equation for enhancing HSI shadow. The utilization of the 3D data of HSI as the system input is directly incorporated in this process and directly obtains enhanced 3D hyperspectral shadow data without data reconstruction. Subsequently, the internal parameters of 3D DSR are fine-tuned to attain the optimal

resonance state. In the final stage, HSI from a real-world scenario is classified using CBAM-ResNet152 to verify the 3D DSR method for image enhancement. The inclusion of CBAM enables classification model to focus on important information and reduce the redundancy of spatial–spectral information.

Experimental results highlight that 3D DSR enhances the output of model, improving the perceptual quality of HSI shadow regions. It enriches the information within the enhanced images and preserves more details from the original images. The utilization of CBAM emphasizes crucial spatial and channel information, thereby enhancing ResNet152's performance in HSI classification. Consequently, the generated model proves to be more accurate and robust, facilitating improved handling of complex and challenging HSI classification problems.

In conclusion, the proposed 3D DSR method exhibits significant potential in the realm of HSI enhancement and classification. This incremental innovation holds promise for advancing hyperspectral image analysis and contributing to a deeper understanding and exploitation of intricate three-dimensional data structures.

Author Contributions: Conceptualization, X.L.; data curation, Y.K.; formal analysis, X.L.; funding acquisition, M.F.; investigation, Y.K., X.L. and M.F.; methodology, X.L. and Y.K.; project administration, X.L. and M.F.; resources, X.L. and Y.K.; software , Y.K. and X.L.; supervision, X.L. and M.F.; validation, X.L. and Y.K.; visualization, Y.K. and X.L.; writing—original draft, Y.K. and X.L.; writing—review and editing, Y.K., X.L., and M.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China under Grant 61971253 and the Shandong Provincial Natural Science Foundation ZR2020MF011.

Data Availability Statement: The datasets presented in this article are not readily available because the data are part of an ongoing study. Requests to access the datasets should be directed to X.F. Liu (02937@qust.edu.cn).

Acknowledgments: The authors would like to thank the Editors and Reviewers for their valuable comments.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Zuo, D.; Li, Y.; Qiu, S.; Jin, W.; Guo, H. A Spectral Enhancement Method Based on Remote-Sensing Images for High-Speed Railways. *Electronics* **2023**, *12*, 2670. [CrossRef]
2. Goetz, A.F.; Vane, G.; Solomon, J.E.; Rock, B.N. Imaging spectrometry for earth remote sensing. *Science* **1985**, *228*, 1147–1153. [CrossRef] [PubMed]
3. Wang, J.; Xia, Y.; Zhang, Y. Anomaly Detection of Hyperspectral Image via Tensor Completion. *IEEE Geosci. Remote Sens. Lett.* **2021**, *18*, 1099–1103. [CrossRef]
4. Wang, W.; Li, S.; Qi, H.; Ayhan, B.; Kwan, C.; Vance, S. Identify anomaly componentbysparsity and low rank. In Proceedings of the 2015 7th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), Tokyo, Japan, 2–5 June 2015; pp. 1–4.
5. Qu, Y.; Qi, H.; Ayhan, B.; Kwan, C.; Kidd, R. DOES multispectral/hyperspectral pansharpening improve the performance of anomaly detection. In Proceedings of the 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, USA, 23–28 July 2017; pp. 6130–6133.
6. Raza Shah, N.; Maud, A.R.M.; Bhatti, F.A.; Ali, M.K.; Khurshid, K.; Maqsood, M.; Amin, M. Hyperspectral anomaly detection: A performance comparison of existing techniques. *Int. J. Digit. Earth.* **2022**, *15*, 2078–2125. [CrossRef]
7. Nakhostin, S.; Clenet, H.; Corpetti, T.; Courty, N. Joint Anomaly Detection and Spectral Unmixing for Planetary Hyperspectral Images. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 6879–6894. [CrossRef]
8. Thangavel, K.; Spiller, D.; Sabatini, R.; Amici, S.; Sasidharan, S.T.; Fayek, H.; Marzocca, P. Autonomous Satellite Wildfire Detection Using Hyperspectral Imagery and Neural Networks: A Case Study on Australian Wildfire. *Remote Sens.* **2023**, *15*, 720. [CrossRef]
9. Huang, H.; Liu, L.; Ngadi, M.O. Recent developments in hyperspectral imaging for assessment of food quality and safety. *Sensors* **2014**, *14*, 7248–7276. [CrossRef] [PubMed]
10. Pu, H.; Wei, Q.; Sun, D.W. Recent advances in muscle food safety evaluation: Hyperspectral imaging analyses and applications. *Crit. Rev. Food Sci. Nutr.* **2023**, *63*, 1297–1313. [CrossRef]

11. Fu, X.; Chen, J. A review of hyperspectral imaging for chicken meat safety and quality evaluation: Application, hardware, and software. *Compr. Rev. Food Sci. Food Saf.* **2019**, *18*, 535–547. [CrossRef]
12. Soni, A.; Dixit, Y.; Reis, M.M.; Brightwell, G. Hyperspectral imaging and machine learning in food microbiology: Developments and challenges in detection of bacterial, fungal, and viral contaminants. *Compr. Rev. Food Sci. Food Saf.* **2022**, *21*, 3717–3745. [CrossRef]
13. Lu, Y.; Saeys, W.; Kim, M.; Peng, Y.; Lu, R. Hyperspectral imaging technology for quality and safety evaluation of horticultural products: A review and celebration of the past 20-year progress. *Postharvest Biol. Technol.* **2020**, *170*, 111318. [CrossRef]
14. Xu, Y.; Du, B.; Zhang, L. Simultaneous Segmentation and Edge Detection for Hyperspectral Image via a Deep Supervised and boundary-constrained Network. In Proceedings of the IGARSS 2019—2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019; pp. 3860–3863. [CrossRef]
15. Zhong, Y.; Ru, C.; Wang, S.; Li, Z.; Cheng, Y. An online, non-destructive method for simultaneously detecting chemical, biological, and physical properties of herbal injections using hyperspectral imaging with artificial intelligence. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2022**, *264*, 120250. [CrossRef] [PubMed]
16. Alexopoulou, A.; Kaminari, A.A.; Moutsatsou, A. Multispectral and hyperspectral studies on greek monuments, archaeological objects and paintings on different substrates. achievements and limitations. In Proceedings of the Transdisciplinary Multispectral Modeling and Cooperation for the Preservation of Cultural Heritage: First International Conference, TMM_CH 2018, Athens, Greece, 10–13 October 2018; Revised Selected Papers, Part II 1; pp. 443–461.
17. Picollo, M.; Cucci, C.; Casini, A.; Stefani, L. Hyper-spectral imaging technique in the cultural heritage field: New possible scenarios. *Sensors* **2020**, *20*, 2843. [CrossRef] [PubMed]
18. Zerrouki, N.; Harrou, F.; Sun, Y.; Hocini, L. A Machine Learning-Based Approach for Land Cover Change Detection Using Remote Sensing and Radiometric Measurements. *IEEE Sens. J.* **2019**, *19*, 5843–5850. [CrossRef]
19. Leon, R.; Martinez-Vega, B.; Fabelo, H.; Ortega, S.; Melian, V.; Castaño, I.; Carretero, G.; Almeida, P.; Garcia, A.; Quevedo, E.; et al. Non-invasive skin cancer diagnosis using hyperspectral imaging for in-situ clinical support. *J. Clin. Med.* **2020**, *9*, 1662. [CrossRef] [PubMed]
20. Tripathi, M.K.; Govil, H. Evaluation of AVIRIS-NG hyperspectral images for mineral identification and mapping. *Heliyon* **2019**, *5*, e02931. [CrossRef] [PubMed]
21. Bar, D.E.; Wolowelsky, K.; Swirski, Y.; Figov, Z.; Michaeli, A.; Vaynzof, Y.; Abramovitz, Y.; Ben-Dov, A.; Yaron, O.; Weizman, L.; et al. Target detection and verification via airborne hyperspectral and high-resolution imagery processing and fusion. *IEEE Sens. J.* **2010**, *10*, 707–711. [CrossRef]
22. Moharram, M.A.; Sundaram, D.M. Dimensionality reduction strategies for land use land cover classification based on airborne hyperspectral imagery: A survey. *Environ. Sci. Pollut. Res.* **2023**, *30*, 5580–5602. [CrossRef]
23. Sun, H.; Zheng, X.; Lu, X. A supervised segmentation network for hyperspectral image classification. *IEEE Trans. Image Process.* **2021**, *30*, 2810–2825. [CrossRef]
24. Ryu, C.; Kong, S.G.; Kim, H. Enhancement of feature extraction for low-quality fingerprint images using stochastic resonance. *Pattern Recognit. Lett.* **2011**, *32*, 107–113. [CrossRef]
25. Rallabandi, V.S.; Roy, P.K. Magnetic resonance image enhancement using stochastic resonance in Fourier domain. *Magn. Reson. Imaging* **2010**, *28*, 1361–1373. [CrossRef] [PubMed]
26. Maragatham, G.; Roomi, S.M.M. An automatic contrast enhancement method based on stochastic resonance. In Proceedings of the 2013 IEEE Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT), Tiruchengode, India, 4–6 July 2013; pp. 1–7.
27. Chouhan, R.; Jha, R.K.; Biswas, P.K. Enhancement of dark and low-contrast images using dynamic stochastic resonance. *IET Image Process.* **2013**, *7*, 174–184. [CrossRef]
28. Liu, X.; Wang, H.; Meng, Y.; Fu, M. Classification of hyperspectral image by CNN based on shadow area enhancement through dynamic stochastic resonance. *IEEE Access* **2019**, *7*, 134862–134870. [CrossRef]
29. Liu, Q.; Fu, M.; Liu, X. Shadow Enhancement Using 2D Dynamic Stochastic Resonance for Hyperspectral Image Classification. *Remote Sens.* **2023**, *15*, 1820. [CrossRef]
30. Müller, S.; Zaum, D.W. Robust building detection in aerial images. *ISPRS Arch.* **2005**, *36*, 143–148.
31. Maltezos, E.; Doulamis, N.; Doulamis, A.; Ioannidis, C. Deep convolutional neural networks for building extraction from orthoimages and dense image matching point clouds. *J. Appl. Remote. Sens.* **2017**, *11*, 042620. [CrossRef]
32. Makantasis, K.; Karantzalos, K.; Doulamis, A.; Doulamis, N. Deep supervised learning for hyperspectral data classification through convolutional neural networks. In Proceedings of the 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Milan, Italy, 26–31 July 2015; pp. 4959–4962.
33. Singh, G.; Stefenon, S.F.; Yow, K.C. Interpretable visual transmission lines inspections using pseudo-prototypical part network. *Mach. Vision Appl.* **2023**, *34*, 41. [CrossRef]
34. Zhao, M.; Yan, L.; Chen, J. Hyperspectral image shadow compensation via cycle-consistent adversarial networks. *Neurocomputing* **2021**, *450*, 61–69. [CrossRef]
35. Benzi, R.; Sutera, A.; Vulpiani, A. The mechanism of stochastic resonance. *J. Phys. A Math. Theor.* **1981**, *14*, L453. [CrossRef]
36. Oyedotun, O.K.; Khashman, A. Deep learning in vision-based static hand gesture recognition. *Neural. Comput. Appl.* **2017**, *28*, 3941–3951. [CrossRef]

37. Molchanov, P.; Gupta, S.; Kim, K.; Kautz, J. Hand gesture recognition with 3D convolutional neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Boston, MA, USA, 7–12 July 2015; pp. 1–7.
38. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
39. Benmouna, B.; Pourdarbani, R.; Sabzi, S.; Fernandez-Beltran, R.; García-Mateos, G.; Molina-Martínez, J.M. Attention Mechanisms in Convolutional Neural Networks for Nitrogen Treatment Detection in Tomato Leaves Using Hyperspectral Images. *Electronics* **2023**, *12*, 2706. [CrossRef]
40. Zhou, F.; Deng, H.; Xu, Q.; Lan, X. CNTR-YOLO: Improved YOLOv5 Based on ConvNext and Transformer for Aircraft Detection in Remote Sensing Images. *Electronics* **2023**, *12*, 2671. [CrossRef]
41. Asha, C.; Singh, M.; Suresh, S.; Lal, S. Optimized Dynamic Stochastic Resonance framework for enhancement of structural details of satellite images. *Remote Sens. Appl. Soc. Environ.* **2020**, *20*, 100415. [CrossRef]
42. Singh, M.; Verma, A.; Sharma, N. Optimized multistable stochastic resonance for the enhancement of pituitary microadenoma in MRI. *IEEE J. Biomed. Health Inform.* **2017**, *22*, 862–873. [CrossRef] [PubMed]
43. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [CrossRef]
44. Kan, Z.; Li, S.; Hou, M.; Fang, L.; Zhang, Y. Attention-based octave network for hyperspectral image denoising. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *15*, 1089–1102. [CrossRef]
45. Sidorov, O.; Hardeberg, J. Deep hyperspectral prior: Single-image denoising, inpainting, super-resolution. In Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops, Seoul, Republic of Korea, 27–28 October 2019.
46. Chang, Y.; Yan, L.; Fang, H.; Zhong, S.; Liao, W. HSI-DeNet: Hyperspectral image restoration via convolutional neural network. *IEEE Trans. Geosci. Remote Sens.* **2018**, *57*, 667–682. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.