

HSI Fusion Method Based on TV-CNMF and SCT-NMF Under the Background of Artificial Intelligence

Dapeng Zhao¹, Yapeng Zhao², Xuexia Dou^{3*}
Henan Polytechnic, Zhengzhou 450000, China^{1,3}
Henan Polytechnic University, Jiaozuo 454150, China²

Abstract—The fusion of hyper-spectral images has important application value in fields such as remote sensing, environmental monitoring, and agricultural analysis. To improve the quality of reconstructed images, an HSI fusion method based on fully variational coupled non-negative matrix factorization and sparse constrained tensor factorization techniques is proposed. Spectral sparsity description is enhanced through sparse regularization, image spatial characteristics are captured using differential operators, and convergence is improved by combining proximal optimization with augmented Lagrangian methods. The experiment outcomes on the AVIRIS and HYDICE datasets indicate that the proposed method achieves peak signal-to-noise ratios of 38.12 dB and 37.56 dB, respectively, and reduces spectral angle errors to 3.98° and 4.12°, respectively, significantly better than the other two comparative methods. The contribution of each module is further verified through ablation experiments. The complete algorithm performs the best in all indicators, verifying the synergistic effect of sparse regularization, total variation regularization, and coupled factorization strategies. In HSI fusion tasks under various complex lighting and noise conditions, the performance of the proposed algorithm is particularly excellent, fully demonstrating its robustness and applicability in complex scenes. The method proposed by the research effectively improves the fusion quality of HSI, providing an efficient and robust solution for the analysis and application of HSI.

Keywords—HSI; NMF; sparse regularization; SCT; augmented Lagrangian method

I. INTRODUCTION

Hyper-spectral images (HSI) have continuous spectral bands and HSI resolution, providing powerful information support for land classification, target recognition, environmental monitoring, and other fields [1]. For example, HSI can be used to finely distinguish vegetation types, monitor water pollution, detect mineral distribution, and identify target materials, all of which are difficult to achieve with traditional optical images [2]. However, existing HSI devices are limited by optical and sensor technologies, making it difficult to simultaneously capture both HSI and high spatial resolution images [3]. This constraint renders the acquisition of HSI that boasts both high spectral and spatial resolution a significant technical hurdle—one that remains incompletely overcome. To address this issue, the fusion technology of HSI and multi-spectral images (MSI) has emerged, with the goal of combining the advantages of both to generate fused images with HSI and high spatial resolution. By combining the advantages of HSI and MSI, fusion technology

can generate fused images that retain both spectral details of HSI and spatial clarity of MSI, thus breaking through the hardware limitations of a single device. Traditional methods often struggle to achieve a balance between spectral fidelity and spatial resolution. Therefore, an HSI fusion method combining non-negative matrix factorization (NMF) and tensor factorization techniques has been proposed in this study. NMF can effectively extract hidden features from data while maintaining the non-negativity of the data, which enables it to better preserve spectral information when processing HSI. Secondly, tensor factorization can capture multidimensional interactions in HSI data and uncover deep information. In addition, the study enhances image spatial smoothness by introducing total variation (TV) regularization and optimizes spectral characteristics using sparsity constraint (SC), which can effectively preserve spectral and spatial information while reducing noise interference. Finally, the use of proximal alternating optimization (PAO) and augmented Lagrangian methods significantly improves the convergence speed and computational efficiency of the model. Therefore, the proposed method is suitable for solving the problem of HSI fusion, which can effectively improve the fusion quality of HSI and provide an efficient and robust solution for the analysis and application of HSI.

The first section of the study furnishes an exhaustive account of the specific principles and implementation of the proposed HSI fusion method. The second section presents the experimental setup, performance evaluation indicators, and comparative experimental results on different datasets, and analyzes the ablation experiment to verify the effectiveness of each module in the method. Finally, the third section summarizes the main achievements of the research and explores the potential applications and future development directions of the method.

II. RELATED WORK

In recent times, artificial intelligence learning technology has made significant progress in the field of HSI integration. Dian et al. proposed a zero sample learning technique to improve the clarity of HSI and accurately measure the spectral and spatial characteristics of imaging sensors. This technology also achieved dimensionality reduction of HSI data, optimized model size and storage requirements, while maintaining fusion accuracy. Experiments showed that this method exhibited significant performance in both efficiency and accuracy [4]. Wu

et al. proposed a deep interpretable network that effectively integrates HSI and MSI through advanced coupling matrix factorization constraints. This network alternately processed HSI and MSI data through two branch subnets, predicting abundance and membership matrices respectively. The experimental results showed that this method was superior to existing model driven and data-driven fusion techniques in visual analysis and quality assessment [5]. Although this method could capture complex nonlinear relationships, it performed poorly in preserving spectral information, especially under high noise conditions. Zhao et al. raised an HSI classification method based on local feature decoupling and hybrid attention module. This method utilized gradient oriented histogram algorithm to preprocess HSI data, achieving preliminary nonlinear decoupling. Experimental results demonstrated that, compared to other transformer-based and traditional HSI classification methods, the proposed approach exhibited superior classification performance. [6]. Yang et al. combined tensor theory with deep learning and proposed a new unsupervised deep tensor network for the fusion of HSI and MSI. They designed a tensor filtering layer and constructed a coupled tensor filtering module based on it. This module worked in conjunction with the projection module to jointly train HSI and MSI in an unsupervised end-to-end manner. The effectiveness of this method was verified through experiments on simulated and actual remote sensing datasets [7]. However, when dealing with ultra-high resolution HSI, the demand for computing resources significantly increased, making it difficult to meet the requirements of real-time processing.

NMF is a commonly-used matrix factorization approach in signal processing, image processing, text mining, and other fields. It can effectively extract hidden features from data while maintaining its non-negativity [8]. Sun et al. proposed an adaptive graph regularized NMF model with global constraints for data representation. This model utilized the self-representation characteristics of data to construct an adaptive graph, in order to more accurately capture the relationships between samples, and used graph factorization techniques to simplify the model and enhance its discriminative power. Finally, the effectiveness of the model was validated on multiple databases through experiments [9]. However, using fixed regularization parameters was difficult to adapt to changes in different input conditions, such as lighting variations, noise levels, etc., resulting in limited generalization ability in practical applications. Yu et al. proposed a robust asymmetric NMF clustering method for directed networks. This method took into account the non-Gaussian nature of real-world network errors and assumed that the errors follow a heavier tail Cauchy distribution. The experiment outcomes showed that this method performed better than traditional NMF and other clustering methods in both real and artificial networks [10]. Yuan et al. raised a new model for embedding multi-view attribute networks with integrated manifold regularization. This model captured the Riemannian geometry structure of the network by introducing manifold regularization, which compensated for the shortcomings of traditional NMF in information capture. Non-negative coefficient matrices were obtained using NMF, and the amount of information embedded in the network was enhanced by combining cooperative regularization and manifold regularization. Through experimental verification on multiple

real datasets, the model performed better than current advanced algorithms in node classification tasks [11].

In summary, existing HSI fusion is difficult to achieve a good balance between spectral fidelity and spatial resolution. Moreover, deep learning methods have a strong dependence on large-scale annotated data and high computational complexity, which affects their practicality. Based on this, an innovative HSI fusion method combining NMF and tensor factorization was proposed. By introducing TV regularization to enhance the spatial smoothness of images and using SC to optimize spectral characteristics, both spectral and spatial information are effectively preserved while reducing noise interference. In addition, the use of PAO and augmented Lagrangian methods significantly improves the convergence speed and computational efficiency of the model. The research aims to achieve high spectral resolution while providing high spatial resolution and image quality.

III. METHODS AND MATERIALS

A. HSI Mixed Pixel Factorization Technology Based on Linear Spectral Mixing Model

The core concept of NMF is to factorize the image data collected by sensors into two nonnegative matrices, enabling the identification of endmember spectra and estimation of corresponding abundances without assuming pure pixels. Coupled NMF (CNMF) is a classic technique for HSI fusion. During the fusion process, CNMF alternately processes spectral images, extracting endmember matrices and abundance matrices until the algorithm converges [12]. Finally, these two matrices are multiplied to obtain high-resolution HSI, and the fusion process is shown in Fig. 1 [13].

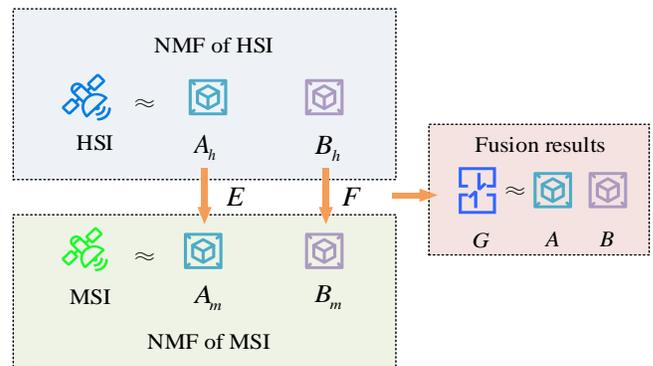


Fig. 1. CNMF fusion process.

In Fig. 1, the HSI is factorized into a base matrix A_h and a coefficient matrix B_h . The factorization of MSI involves breaking it down into a base matrix A_m and a coefficient matrix B_m . By using the correlation matrices E and F , joint constraints between the two data sources are enforced, and the fusion result G is obtained as the product of the base matrix A and the coefficient matrix B , completing the spectral image fusion process. The cost function of the CNMF algorithm is shown in Eq. (1) [14].

$$CNMF(A, B) = \|X - A_h B_h\|^2 + \|Y - A_m B_m\|^2 \quad (1)$$

In (1), X and Y respectively represent the input HSI and MSI. B_h and A_m represent the abundance matrix of spatial downsampling and the endmember matrix of spectral downsampling, respectively, as expressed in Eq. (2).

$$\begin{cases} B_h = BF \\ A_m = EA \end{cases} \quad (2)$$

In Eq. (2), non-negative tensor factorization (NTF) is an extension of NMF that factorizes high-dimensional tensor data. It is specially designed for factorizing and modeling high-dimensional tensor data. Tensors can be seen as high-dimensional forms of matrices, widely used to describe multidimensional data structures such as three-dimensional image sequences, video data, and complex relationships in sensor networks. Unlike traditional matrices, tensors can preserve the multidimensional structural characteristics of data, making them more suitable for analyzing and modeling data with high-dimensional interaction relationships. Two commonly-used tensor factorization models are shown in Fig. 2 [15].

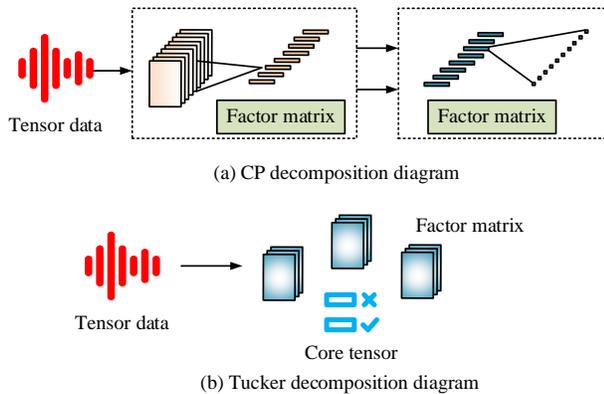


Fig. 2. Tensor factorization model.

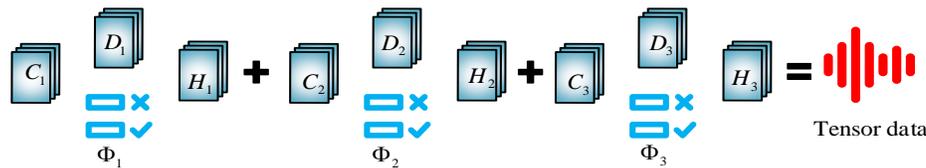


Fig. 3. BTM structure.

B. HSI Fusion Algorithm Based on TV-CNMF and SCT-NMF

To further enhance the effectiveness of HSI fusion, a mixed pixel factorization technique combining linear spectral mixing model is studied, and tensor factorization method is adopted to propose an HSI fusion algorithm based on TV-CNMF and sparse constrained tensor (SCT)-NMF to enhance image detail preservation and denoising effects. Tensor factorization is a method of multi-linear algebra, which is a high-dimensional extension of matrix factorization and can effectively handle three-dimensional and above data structures, such as video and audio. It can capture multidimensional interactions in data and

Fig. 2 (a) and (b) show typical CANDECOMP/PARAFAC (CP) factorization and Tucker factorization, respectively. CP factorization factorizes a tensor into a linear combination of rank one tensors, each consisting of the outer product of three vectors. Tucker factorization, on the other hand, represents a tensor as the product of a core tensor and multiple factor matrices. The core tensor is used to capture the global relationships of the tensor, while the factor matrix is used to represent the feature information in each dimension. These two factorization methods are used for feature extraction and pattern recognition in multidimensional data analysis. The mathematical expression for CP factorization is shown in Eq. (3) [16].

$$y = \sum_{r=1}^R \omega_r (a_r \cdot b_r \cdot c_r) \quad (3)$$

In Eq. (3), y represents the target tensor, R is the rank of the factorization, ω_r is the weight coefficient of the r th rank component, $(a_r \cdot b_r \cdot c_r)$ represents the r th rank tensor, which is a tensor generated by the outer product of three vectors. The mathematical expression for Tucker factorization is shown in Eq. (4).

$$y = \Phi_1 C_2 D_2 H \quad (4)$$

In Eq. (4), Φ represents the core tensor and C, D, H represents the factor matrix. Further extension of Tucker factorization leads to block term factorization (BTM), as shown in Fig. 3 [17]

In Fig. 3, each part is generated by modular multiplication of a kernel tensor and three factor matrices. The factorization result can be seen as representing the tensor as a combination of multiple low rank tensor blocks, with the core tensor describing the relationships within the blocks and the factor matrix describing the characteristics of the tensor in various dimensions. This factorization method is used to represent complex tensor structures more finely and is suitable for handling multi-modal or multidimensional data.

uncover deep information, which is difficult to achieve through matrix factorization. Based on the CNMF model, this study enhances the sparsity description of HSI through sparse regularization of shortest endmember distance and abundance, and captures the segmentation smoothness of the image using differential operators to reduce the impact of noise on the fusion effect. A TV-CNMF HSI fusion algorithm, TV-CNMF, is proposed. The objective function of the algorithm is shown in Eq. (5) [18].

$$\frac{1}{2} \min_{A, B} CNMF(A, B) + \lambda_1 \|AP\|_F^2 + \lambda_2 \|B\|_1 + \lambda_3 \|D_x B\|_1 + \lambda_3 \|D_y B\|_1 \quad (5)$$

In (5), λ_1 , λ_2 and λ_3 both represent regularization parameters, where λ_1 controls the strength of TV regularization and is used to enhance the spatial smoothness of the image. λ_2 controls the intensity of sparse constraints to optimize spectral sparsity. λ_3 regulates the coupling relationship between HSI and MSI to balance the information from both data sources. To determine the optimal values of these hyper-parameters, a grid search method is employed in the study. The value range of λ_1 is [0.1, 1, 10], the value range of λ_2 is [0.01, 0.1, 1], and the value range of λ_3 is [0.001, 0.01, 0.1]. Through grid search, the impact of different parameter combinations on model performance is systematically evaluated, and the optimal parameter combination ($\lambda_1=1, \lambda_2=0.1, \lambda_3=0.01$) is selected. However, this function involves constrained optimization, and the model contains numerous unknowns, making it quite difficult to solve directly. Therefore, it is necessary to transform Eq. (5) into an unconstrained optimization function and introduce auxiliary variables as constraint terms, as shown in Eq. (6).

$$\frac{1}{2}\|X - AB_h\| + \frac{1}{2}\|Y - A_m B\| + \lambda_1 \|AP\|_F^2 + \lambda_2 \|B\|_1 + \lambda_3 \|D_x B\| + \lambda_3 \|D_y B\| + l_{R^+}(A) + l_{R^+}(B) \quad (6)$$

In Eq. (6), $l_{R^+}(A)$ represents $A \geq 0$, $l_{R^+}(B)$ represents $B \geq 0$. After normalizing the objective function, A and B_h are optimized respectively, and the updated formula for B_h is obtained as shown in Eq. (7).

$$B_h \leftarrow B_h \frac{A^T X}{A^T A B_h} \quad (7)$$

In Eq. (7), \leftarrow represents the update rule. For the optimization problem of A , an auxiliary variable is first introduced, and the resulting sub-problem is shown in Eq. (8).

$$\arg \min_A \frac{1}{2}\|X - AB_h\|_F^2 + \lambda_1 \|AP\|_F^2 + l_{R^+}(I_1) \quad (8)$$

In Eq. (8), I_1 represents the auxiliary variable. Eq. (8) is derived to obtain the augmented Lagrangian function as shown in Eq. (9).

$$L(I_1, A, Z_1) = \frac{1}{2}\|X - AB_h\|_F^2 + \lambda_1 \|AP\|_F^2 + l_{R^+}(I_1) + \frac{\rho}{2}\|A - I_1 - Z_1\|_F^2 \quad (9)$$

In Eq. (9), Z_1 represents the Lagrange multiplier and ρ represents the penalty parameter. Using the augmented

Lagrangian alternative direction method of multipliers (ADMM) for solving, the result is shown in Eq. (10).

$$\begin{cases} A^{k+1} = \arg \min_A (A, I_1^k, Z_1^k) \\ I_1^{k+1} = \arg \min_{I_1} (A^{k+1}, I_1, Z_1^k) \\ Z_1^{k+1} = \arg \min_{Z_1} (A^{k+1}, I_1^{k+1}, Z_1) \end{cases} \quad (10)$$

In Eq. (10), k is the number of iterations. Similarly, the solution process of B_h is shown in Fig. 4.

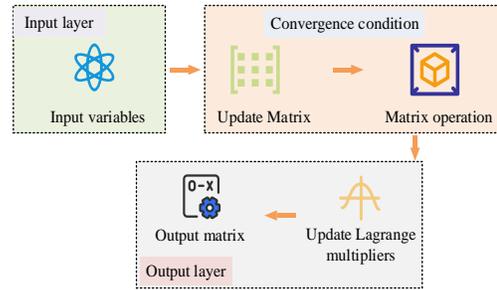


Fig. 4. The process of ADMM solving B_h .

In Fig. 4, one variable is optimized at a time while the other variables are kept constant. By iterating through this loop, each variable is continuously optimized, and the optimal value of the objective function is ultimately approached. The final TV-CNMF algorithm flow is shown in Fig. 5.

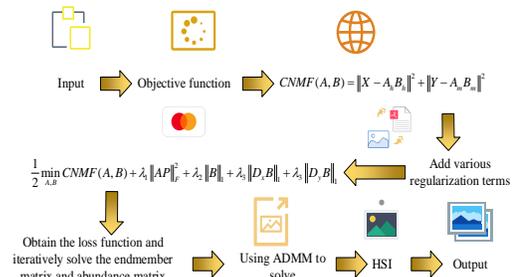


Fig. 5. TV-CNMF algorithm process.

In Fig. 5, the input includes HSI with low spatial resolution and MSI with high spatial resolution, and the error function is minimized using the CNMF model. Among them, regularization terms are introduced to optimize the sparsity and smoothness of matrices A and B . By alternately optimizing the loss function, the fused high-resolution HSI is finally obtained, and the ADMM algorithm is used to complete the optimization solution. When processing HSI fusion, although matrix factorization algorithm achieves good results, it may damage the spatial structure and spectral correlation of HSI, and fail to fully utilize all structural information of the images. The third-order tensor factorization is more suitable for HSI, reducing information loss, but there are shortcomings in preserving details. To this end, an SC-based image fusion algorithm called SCT-NMF is proposed, which combines NTF and NMF to effectively protect data structure information, explore spatial details, and enhance solution stability. Firstly, the BTM and NMF models are

combined, and the sparsity of semi norm constrained abundance is introduced to construct an efficient fusion model, as represented in Eq. (11).

$$\min_{C,D,H,V} \frac{1}{2} \left\| \mathcal{X}_H - \sum_{r=1}^R P_1 C_r (P_2 D_r)^T o_r \right\|_F^2 + \frac{1}{2} \left\| \mathcal{X}_M - \sum_{r=1}^R C_r D_r^T (P_3 o_r) \right\|_F^2 \quad (11)$$

In Eq. (11), \mathcal{X}_H and \mathcal{X}_M respectively represent the image data of HSI and MSI, V represents the sparse coding matrix, o_r is the endmember vector, and P represents the projection matrix. Due to the fact that the fusion model is a non-convex optimization problem, it is necessary to fix one variable while keeping the other variables constant during solving, so that the optimization problem for each fixed variable is convex. Therefore, the PAO method can be used to solve the variables C, D, H in the model, and the outcomes are represented in Eq. (12).

$$\begin{cases} C = \arg \min_C g(C, D, H) + \mu \|C - C^{pre}\|_F^2 \\ D = \arg \min_D g(C, D, H) + \mu \|D - D^{pre}\|_F^2 \\ H = \arg \min_H g(C, D, H) + \mu \|H - H^{pre}\|_F^2 \end{cases} \quad (12)$$

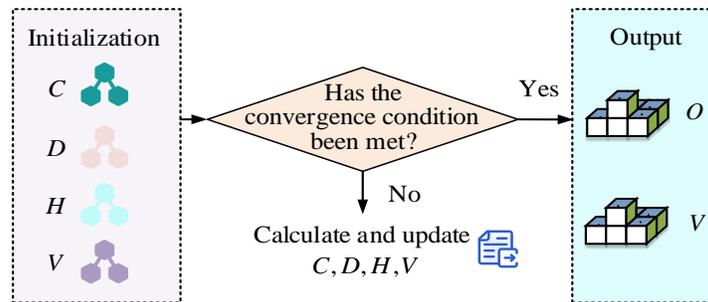


Fig. 6. SCT-NMF algorithm process.

In Fig. 6, the SCT-NMF algorithm initializes the factor matrix C, D, H and iteratively updates each matrix to minimize the objective function. The core idea of this algorithm is to use sparse regularization and tensor factorization techniques, combined with constraint conditions, to optimize the objective function until the convergence conditions are met.

Finally, the algorithm outputs a sparse matrix O and a sparse encoding matrix V , representing the feature representation and sparsity encoding of the data, respectively. Finally, by introducing the sparse encoding matrix of SCT-NMF into the objective function of TV-CNMF, both spatial details and SCs can be optimized simultaneously. Combining TV regularization with sparse encoding matrix can make the image fusion process more robust, reduce the influence of noise, and improve the effect of detail preservation. In the fused image results, TV-CNMF can provide strong denoising and smoothing effects, while SCT-NMF can better protect the structural information

In Eq. (12), μ is a penalty parameter used to control the difference between the current iteration value and the previous iteration value. μ_1 is used to control the difference between the newly obtained C and the C^{pre} obtained in the previous iteration. μ_2 is used to control the difference between the newly obtained D and the D^{pre} obtained in the previous iteration. μ_3 is used to control the difference between the newly obtained H and the H^{pre} obtained in the previous iteration. $g(C, D, H)$ is the objective function, representing the coupling relationship of C, D, H . C^{pre} , D^{pre} , and H^{pre} represent the values of the previous iteration of C, D, H . The SCT-NMF algorithm flow is shown in Fig. 6.

and details of the image. In order to evaluate the computational efficiency of TV-CNMF and SCT-NMF algorithms, the complexity of TV-CNMF and SCT-NMF is analyzed using Big-O representation. Assuming the size of the input HSI image is $M \cdot N \cdot L$, where M and N are spatial dimensions, L is spectral dimension, and r is the factorized rank. The main computational cost of TV-CNMF comes from the optimization of matrix factorization and regularization terms. The time complexity of matrix factorization is $O(MNr)$, and the time complexity of TV regularization is $O(MrN)$. Therefore, the overall time complexity of TV-CNMF is $O(MNr + MrN + MN) \cdot k$. The main computational overhead of SCT-NMF comes from tensor factorization and optimization of sparse constraints. The time complexity of tensor factorization is $O(MNrL)$, and the time complexity of sparse constraints is $O(MN)$. Therefore, the overall time complexity of SCT-NMF is $O(MNrL \cdot k + MN)$.

IV. RESULTS

A. Performance Testing of HSI Fusion Algorithm Based on TV-CNMF and SCT-NMF

To confirm the performance of the raised fusion algorithm, it was tested against the Residual Selective Kernel Attention-based U-net (RSKAU-net) [19] and the Efficient Phase-induced Gabor Cube Selection and Weighted Fusion (EPCS-WF) method [20]. Two datasets, AVIRIS and HYDICE, were selected, and Peak Signal-to-Noise Ratio (PSNR), Spectral Angle Mapper (SAM), and Root Mean Square Error (RMSE) were used as indicators. The outcomes are in Table I.

According to the data in Table I, the proposed algorithm achieved a PSNR of 38.12 dB on the AVIRIS dataset, outperforming RSKAU-net (35.42 dB) and EPCS-WF (36.78 dB) by 7.6%, indicating superior image reconstruction quality. SAM was 3.98°, which was 25.5% lower than RSKAU net and exhibited stronger spectral fidelity; and the RMSE was 0.031, lower than other algorithms, indicating the minimum reconstruction error. On the HYDICE dataset, the proposed algorithm achieved a PSNR of 37.56 dB, 7.7% higher than RSKAU-net. In addition, the SAM of the algorithm proposed in

the research was 4.12°, which was 27.3% lower than RSKAU net. The RMSE was 0.033, which also outperformed other algorithms. Overall, the algorithm proposed in the research performed the best in terms of PSNR, SAM, and RMSE. It outperformed RSKAU net and EPCS-WF in terms of spectral fidelity, spatial resolution, and pixel-level error in reconstructed images, demonstrating excellent spectral and spatial property preservation capabilities. It was suitable for HSI fusion and analysis tasks. The loss function changes of the three algorithms on different datasets are shown in Fig. 7.

TABLE I PERFORMANCE TEST RESULTS OF FUSION ALGORITHM

Database	Algorithm	PSNR	SAM	RMSE
AVIRIS	RSKAU-net	35.42	5.34	0.042
	EPCS-WF	36.78	4.56	0.038
	Ours	38.12	3.98	0.031
HYDICE	RSKAU-net	34.87	5.67	0.045
	EPCS-WF	35.91	4.89	0.041
	Ours	37.56	4.12	0.033

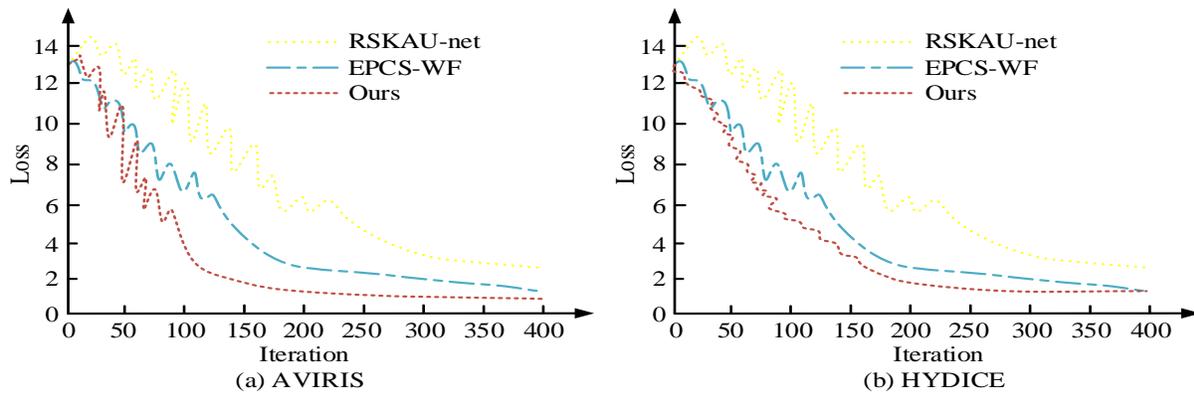


Fig. 7. Changes in loss function on different datasets.

Fig. 7(a) and Fig. 7(b) respectively show the trend of the loss function of three algorithms with iteration times on the AVIRIS and HYDICE datasets. Overall, the loss value gradually decreased as the number of iterations increased, and each algorithm eventually tended to converge. On the AVIRIS dataset, the initial loss of RSKAU net was relatively high, the decrease was slow, and the final loss value was significantly higher than other algorithms. The descent rate of EPCS-WF was slightly faster, but the final loss was still higher than the algorithm proposed by the research. The loss value of the algorithm proposed by the research decreased the fastest and was significantly better than other algorithms at the 100th iteration, resulting in the lowest loss value in the end. On the HYDICE dataset, both RSKAU net and EPCS-WF converged slowly, and the final loss value was higher than the algorithm proposed in the study. The algorithm proposed by the research not only had the fastest convergence speed and the best optimization performance on two datasets, but also had the lowest final loss value, indicating that its fusion performance and optimization effect were better than other algorithms. Meanwhile, it had good robustness and universality, and was suitable for HSI fusion tasks. Through ablation experiments, the

specific contributions of each module to the algorithm's performance were gradually verified. Firstly, the complete algorithm consisted of multiple key components, including CNMF, sparse regularization TV, and ADMM. In ablation settings, the complete model was considered the baseline model, which included all modules. The algorithm that removed sparse regularization and only retained coupling tensor factorization and TV regularization was referred to as A1. The algorithm that removed TV regularization and only retained sparse regularization and coupling tensor factorization was referred to as A2. The algorithm that removed coupling factorization and replaced it with independent NMF was referred to as A3. The algorithm that removed all regularization terms and only used coupling factorization was referred to as A4. The algorithm that replaced the optimization strategy with a simple multiplication update method was referred to as A5. The ablation experiment results in the two datasets are shown in Table II, using PSNR, SAM, and RMSE as evaluation indicators.

TABLE II RESULTS OF ABLATION EXPERIMENT

Database	Algorithm	PSNR (dB)	SAM (°)	RMSE
AVIRIS	A1	38.12	3.98	0.031
	A2	36.45	4.76	0.038
	A3	36.23	4.91	0.042
	A4	35.15	5.45	0.048
	A5	34.52	5.72	0.052
HYDICE	A1	37.56	4.12	0.033
	A2	35.98	4.78	0.041
	A3	35.62	4.96	0.045
	A4	34.88	5.67	0.052
	A5	34.15	5.94	0.054

According to Table II, in the AVIRIS dataset, A1 had the best PSNR, SAM, and RMSE of 38.12dB, 3.98°, and 0.031, respectively. The PSNR of A2 and A3 decreased to 36.45dB and 36.23dB respectively, SAM increased to 4.76° and 4.91°, and RMSE increased to 0.038 and 0.042, indicating that sparse regularization and TV regularization played an important role in spectral fidelity and reconstruction quality. The PSNR of A4 decreased to 35.15dB, SAM increased to 5.45°, and RMSE reached 0.048. The PSNR of A5 further decreased to 34.52dB, SAM increased to 5.72°. RMSE was 0.052, indicating that the lack of regularization and coupling factorization significantly reduced model performance. In the HYDICE dataset, A1 had the best PSNR, SAM, and RMSE of 37.56dB, 4.12°, and 0.033, respectively. The PSNR of A2 and A3 decreased to 35.98dB and 35.62dB, respectively, while SAM increased to 4.78° and 4.96°, and RMSE increased to 0.041 and 0.045. The PSNR of A4 was 34.88dB, SAM was 5.67°, and RMSE was 0.052. Finally, in A5, PSNR decreased to 34.15dB, SAM increased to 5.94°, and RMSE was 0.054. Overall, A1 performed the best on both datasets, validating the effectiveness and importance of sparse regularization, TV regularization, and coupled factorization. The algorithm A5, which removed all regularization terms, performed the worst, further demonstrating the importance of regularization for model optimization and robustness.

B. Analysis of the Effect of HSI Fusion Algorithm based on TV-CNMF and SCT-NMF

To verify the application effect of the proposed algorithm, simulation experiments were conducted to compare and analyze the applicability of the algorithm under different lighting conditions and noise levels. Selecting SAM, Spectral Correlation Coefficient (SCC), and Spectral Mean Square Error (SMSE) as indicators, the results are shown in Table III.

In Table III, under illumination conditions, the SAM of the proposed algorithm in high light environments was 3.89°, significantly better than RSKAU net and EPCS-WF. SCC was 0.965, significantly higher than RSKAU net and EPCS-WF, while SMSE was the lowest, only 0.031, indicating that it could better preserve spectral information under high light conditions and had high spectral fidelity and low error. Under low light conditions, the SAM of the proposed algorithm was 4.32°, SCC

was 0.952, and SMSE was 0.036, which were also superior to RSKAU net and EPCS-WF, demonstrating good spectral fidelity and robustness. At the noise level, the SAM of the proposed algorithm in high noise environments was 4.89°, and the SCC was 0.941, both significantly better than RSKAU net and EPCS-WF. The SMSE was the lowest, at 0.038, demonstrating strong noise resistance. Under low noise conditions, the proposed algorithm had a SAM of 5.02°, SCC of 0.952, and SMSE of 0.041, which were also superior to the other two compared algorithms and maintained advantages in spectral fidelity and error control.

TABLE III APPLICABILITY UNDER DIFFERENT LIGHTING CONDITIONS AND NOISE LEVELS

Condition	Algorithm	SAM (°)	SCC	SMSE
High light	RSKAU-net	5.34	0.912	0.042
	EPCS-WF	4.56	0.932	0.038
	Ours	3.89	0.965	0.031
Low light	RSKAU-net	6.21	0.892	0.051
	EPCS-WF	5.02	0.915	0.045
	Ours	4.32	0.952	0.036
High noise	RSKAU-net	6.78	0.876	0.056
	EPCS-WF	5.89	0.903	0.048
	Ours	4.89	0.941	0.038
Low noise	RSKAU-net	7.12	0.896	0.058
	EPCS-WF	6.21	0.912	0.041
	Ours	5.02	0.952	0.063

In Fig. 8(a) and Fig. 8(b) respectively show the CPU usage comparison of three methods under different lighting and noise conditions. In Fig. 8(a), under high light conditions, the RSKAU net method had the highest CPU utilization rate of 56.6%. The EPCS-WF method was 52.2%. The lowest calculation efficiency of the method proposed by the research was 50.6%, indicating that it had the highest computational efficiency. Under low light conditions, the CPU utilization rates of RSKAU net and EPCS-WF methods were 51.1% and 45.2%, respectively, while the proposed method remained the lowest at 45.5%, demonstrating good computational efficiency and stability. In Fig. 8(b), under high noise conditions, the EPCS-WF method had the highest CPU utilization rate of 60.5%. The RSKAU net method was 51.7%. The method proposed by the research was 53.3%, demonstrating higher resource utilization efficiency. Under low noise conditions, the CPU utilization rates of RSKAU net and EPCS-WF methods were 46.6% and 56.3%, respectively, while the proposed method was 59.8%, still performing well in terms of computational performance and resource utilization. The Salinas dataset and Chikusei dataset were selected for the study, with 100 randomly selected samples each, to evaluate the running time of the three algorithms in HSI fusion tasks. The results are shown in Fig. 9.

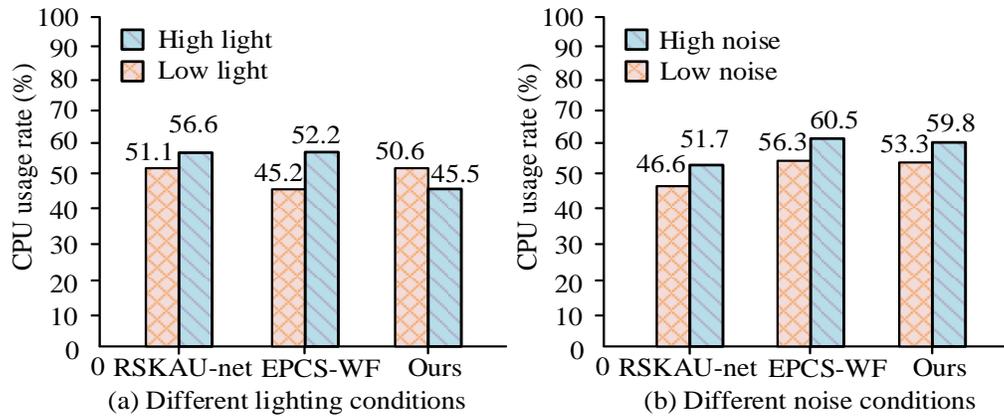


Fig. 8. Comparison of CPU usage under different lighting and noise conditions.

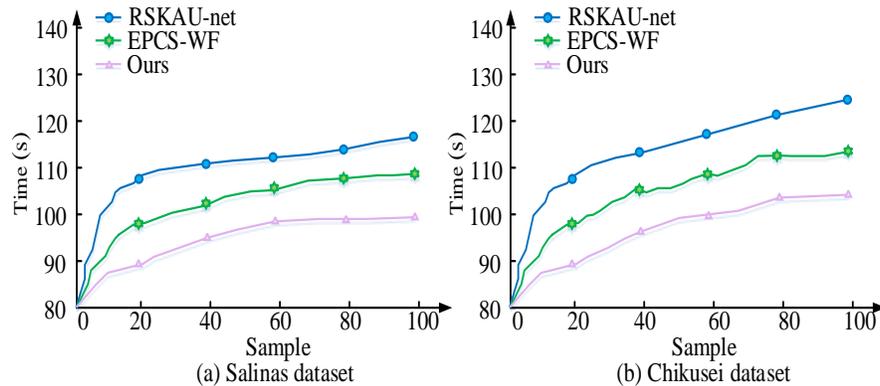


Fig. 9. Comparison of running time under different sample sizes.

Fig. 9(a) and Fig. 9(b) show the trend of running time for different algorithms on different datasets, respectively. In Fig. 9(a), the running time of the proposed method was consistently lower than that of RSKAU net and EPCS-WF, indicating better computational efficiency. Although the running time of EPCS-WF was relatively low, as the sample size increased, the running time gradually increased, and the gap between the proposed method and the research gradually widened. The running time of RSKAU net was the highest, and the growth trend was the most significant, indicating that its computational cost was relatively high when processing large-scale samples. In Fig. 9(b), the proposed method also showed the lowest running time, and the growth trend was relatively flat, indicating that the proposed method not only had advantages in computational efficiency, but also had good adaptability and stability on different datasets. The performance of EPCS-WF on this dataset was also relatively close to the proposed method, but still slightly higher. The running time of RSKAU net increased the fastest with an increase in sample size, further confirming its shortcomings in computational efficiency. Based on the results of the two datasets, the proposed method outperformed the other two methods in terms of running time, indicating that it could not only provide high-quality fusion results but also achieve them at a lower computational cost when processing HSI fusion tasks. This is particularly important for practical applications where computing resources are limited. The qualitative visual comparison results of the three methods in different scenarios are shown in Fig. 10.

From Figs 10(a) to 10(d) show the original images of four scene images and the fusion results of three methods, respectively. As shown in the figure, in the sky scene, the proposed method could better capture the color gradient of the sky at sunset while maintaining the delicate texture of the clouds, while other methods may result in unnatural color transitions or loss of texture details. Under the EPCS-WF method, there were phenomena of exposure and distortion in the image. In the farmland scene, the method proposed by the research not only clearly displayed the outline of the farmland, but also preserved rich details of the soil and vegetation. In contrast, RSKAU net was not accurate enough in color reproduction, while EPCS-WF lacked detailed representation. For architectural scenes, both RSKAU net and the methods proposed by the research could display the subtle textures of windows and walls while maintaining the clarity of the building structure, while EPCS-WF still needed to improve the richness of details. Finally, in the street scene, the proposed method could better restore the true colors of the street and trees, while maintaining high contrast and clarity of the image, resulting in a good display of the texture of the street and the details of the trees. However, color distortion occurred in the fused images of RSKAU net, while EPCS-WF lacked attention to detail. Overall, the method proposed in the study could effectively preserve the spectral and spatial information of HSI during fusion, while reducing noise interference and improving image quality.

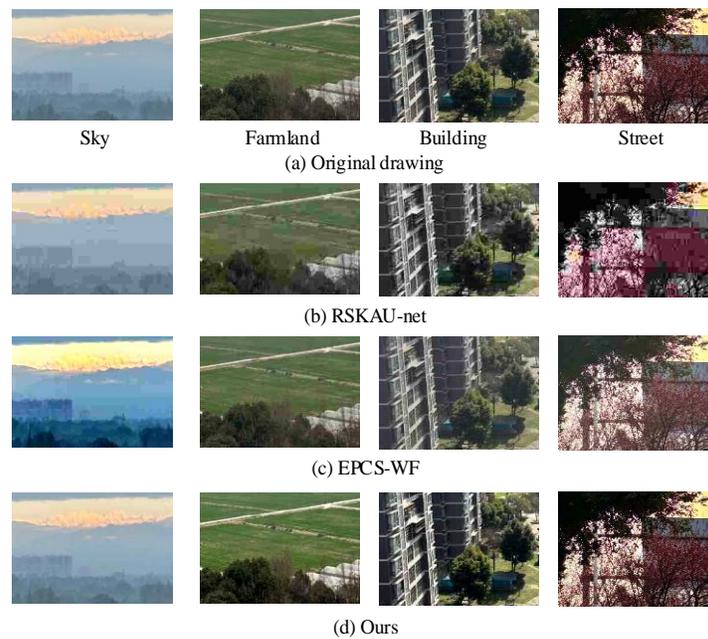


Fig. 10. Qualitative visual comparison of images in different scenarios.

V. DISCUSSION

An HSI fusion algorithm based on TV-CNMF and SCT-NMF was proposed in this study. By combining TV regularization and SCT, the spectral and spatial characteristics were optimized, and the convergence and computational efficiency of the model were improved through efficient optimization strategies. The experimental results on the AVIRIS and HYDICE datasets showed that the proposed method significantly outperformed RSKAU net and EPCS-WF in terms of PSNR, SAM, and RMSE. On the AVIRIS dataset, the PSNR of the proposed method reached 38.12 dB, which was 7.6% higher than RSKAU net, and the SAM decreased to 3.98° , which was 25.5% lower than RSKAU net. On the HYDICE dataset, PSNR reached 37.56 dB, an increase of 7.7% compared to RSKAU net, and SAM decreased to 4.12° , a decrease of 27.3%. In addition, the method proposed in the research performed particularly well under high noise and low light conditions, further verifying its robustness and applicability. Compared with reference [7], although the unsupervised deep tensor network proposed by Yang J et al. performed well in HSI and MSI fusion tasks, its robustness in handling complex lighting and noise conditions still needed to be improved. The method proposed by the research significantly improved robustness under high noise and low light conditions by introducing adaptive regularization strategies and efficient optimization algorithms, while reducing dependence on large-scale annotated data. Through ablation experiments, the key contributions of sparse regularization, TV regularization, and coupled factorization strategies to model performance were identified. The complete model performed the best on all indicators, demonstrating that the synergistic effect of each module significantly improved the fusion quality of images. Compared with existing algorithms, the proposed method not only had significant advantages in spectral fidelity and spatial resolution but also demonstrated lower computational costs and higher practical application potential. The method proposed in the

research was suitable for HSI fusion tasks under complex lighting and noise conditions and could significantly improve the fusion quality of images, providing an effective solution for the analysis and application of HSI.

VI. CONCLUSION

In summary, an HSI fusion algorithm based on TV-CNMF and SCT-NMF was proposed, which significantly improved the spectral fidelity and spatial resolution of images by combining TV regularization and SCT. The experimental results showed that this method exhibited strong robustness and applicability under complex lighting and noise conditions. In addition, this method had broad practical application potential in fields such as satellite imaging and medical imaging. For example, in satellite imaging, this method could be used to process hyperspectral data in real-time, improving the accuracy of land cover classification and target recognition. In medical imaging, this method could be used for multi-modal image fusion to assist in disease diagnosis and treatment planning. However, there are still some limitations to the research, as the performance of the method may decrease in high noise or high dynamic environments, especially in extreme noise conditions or scenarios where the target is moving rapidly. Although the method proposed by the research improved computational efficiency, the time complexity and computational resource requirements were still high when processing ultra-high resolution HSI, which may limit its real-time performance in practical applications. Future research can further explore deep learning-based enhancement methods, such as designing deep neural network modules to optimize the selection of regularization parameters or enhance feature extraction capabilities, to further improve fusion performance. Meanwhile, adaptive regularization techniques can be studied to dynamically adjust regularization parameters based on the characteristics of input data, to improve the robustness of the algorithm under different lighting, noise, and dynamic conditions. In addition, for

the processing requirements of ultra-high resolution HSI, the time complexity and computational resource utilization of the algorithm can be further optimized, and the possibility of distributed computing or hardware acceleration can be explored.

REFERENCES

- [1] S. N. Chaudhri and S. K. Roy, "A discriminatory groups-based supervised band selection technique for hyperspectral image classification," *Remote Sens. Lett.*, vol. 15, no. 2, pp. 111–120, Jan 2024.
- [2] J. Purohit and R. Dave, "Leveraging deep learning techniques to obtain efficacious segmentation results," *Arch. Adv. Eng. Sci.*, vol. 1, no. 1, pp. 11–26, Jan 2023.
- [3] G. Sersa, U. Simoncic, and M. Milanic, "Imaging perfusion changes in oncological clinical applications by hyperspectral imaging: a literature review," *Radiol. Oncol.*, vol. 56, no. 4, pp. 420–429, Dec 2022.
- [4] R. Dian, A. Guo, and S. Li, "Zero-shot hyperspectral sharpening," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 10, pp. 12650–12666, Oct 2023.
- [5] X. Wu, S. Xiao, W. Dong, J. Qu, T. Zhang, and Y. Li, "Coupled matrix factorization constrained deep Hyperspectral and MSI fusion," *IEEE Sens. J.*, vol. 24, no. 5, pp. 6392–6404, March 2024.
- [6] Y. Zhao, W. Bao, X. Xu, and Y. Zhou, "Hyperspectral image classification based on local feature decoupling and hybrid attention SpectralFormer network," *Int. J. Remote Sens.*, vol. 45, no. 5, pp. 1727–1754, Feb 2024.
- [7] J. Yang, L. Xiao, Y. Q. Zhao, and J. C. W. Chan, "Unsupervised deep tensor network for hyperspectral–MSI fusion," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 9, pp. 13017–13031, Sept 2023.
- [8] J. Yu, B. Pan, S. Yu, and M. F. Leung, "Robust capped norm dual hypergraph regularized non-negative matrix tri-factorization," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 20, no. 7, pp. 12486–12509, May 2023.
- [9] Y. Sun, J. Wang, J. Guo, Y. Hu, and B. Yin, "Globality constrained adaptive graph regularized non-negative matrix factorization for data representation," *IET Image Process.*, vol. 16, no. 10, pp. 2577–2592, May 2022.
- [10] Y. Yu, J. Baek, A. Tosyali, and M. K. Jeong, "Robust asymmetric non-negative matrix factorization for clustering nodes in directed networks," *Ann. Oper. Res.*, vol. 341, no. 1, pp. 245–265, Feb 2024.
- [11] W. Yuan, X. Li, and D. Guan, "Multi-View attributed network embedding using manifold regularization preserving non-negative matrix factorization," *IEEE Trans. Knowl. Data Eng.*, vol. 36, no. 6, pp. 2563–2571, June 2023.
- [12] K. Milligan, K. Scarrott, J. L. Andrews, A. G. Brolo, J. J. Lum, and A. Jirasek, "Reconstruction of raman spectra of biochemical mixtures using group and basis restricted non-negative matrix factorization," *Appl. Spectrosc.*, vol. 77, no. 7, pp. 698–709, Apr 2023.
- [13] Y. Yang, D. Li, Y. Lv, F. Kong, and Q. Wang, "Multispectral and hyperspectral images fusion based on subspace representation and nonlocal low-rank regularization," *Int. J. Remote Sens.*, vol. 45, no. 9, pp. 2965–2984, Apr 2024.
- [14] Z. Chen, Q. Xiao, T. Leng, Z. Zhang, D. Pan, Y. Liu, and X. Li, "Multi-constraint non-negative matrix factorization for community detection: orthogonal regular sparse constraint non-negative matrix factorization," *Complex Intell. Syst.*, vol. 10, no. 4, pp. 4697–4712, Apr 2024.
- [15] H. Chen, G. Yang, and H. Zhang, "Hider: A hyperspectral image denoising transformer with spatial–spectral constraints for hybrid noise removal," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 7, pp. 8797–8811, July 2022.
- [16] T. Zhang, Z. Liang, and Y. Fu, "Joint spatial-spectral pattern optimization and hyperspectral image reconstruction," *IEEE J. Sel. Top. Signal Process.*, vol. 16, no. 4, pp. 636–648, June 2022.
- [17] X. T. Nguyen and G. S. Tran, "Hyperspectral image classification using an encoder-decoder model with depthwise separable convolution, squeeze and excitation blocks," *Earth Sci. Inform.*, vol. 17, no. 1, pp. 527–538, Dec 2024.
- [18] M. A. Haq, U. Garg, M. A. R. Khan, and V. Rajinikanth, "Fusion-Based deep learning model for Hyperspectral images classification," *Comput. Mater. Contin.*, vol. 72, no. 1, pp. 939–957, Feb 2022.
- [19] J. Deng and B. Yang, "Hyperspectral and MSI fusion via residual selective kernel attention-based U-net," *Int. J. Remote Sens.*, vol. 45, no. 5, pp. 1699–1726, Feb 2024.
- [20] R. L. Cai, C. Y. Liu, and J. Li, "Efficient phase-induced gabor cube selection and weighted fusion for hyperspectral image classification," *Sci. China Technol. Sci.*, vol. 65, no. 4, pp. 778–792, March 2022.