

Model-Guided Coarse-to-Fine Fusion Network for Unsupervised Hyperspectral Image Super-Resolution

Jiaxin Li^{ID}, Student Member, IEEE, Ke Zheng^{ID}, Wengu Liu, Zhi Li, Haoyang Yu^{ID}, Member, IEEE, and Li Ni^{ID}

Abstract—Fusing a low-resolution hyperspectral image (LrHSI) with an auxiliary high-resolution multispectral image (HrMSI) is a burgeoning technique to realize hyperspectral image super-resolution (HSI-SR), in which learning-based methods have dominated the mainstream direction. However, the underutilization of degradation models and strong dependence on large-scale training triplets severely impede their applicability and performance. Considering these issues, we reformulate the fusion task as a spectral mapping problem and hence propose an unsupervised model-guided coarse-to-fine (C2F) fusion network. Specifically, degradation knowledge learning (DKL) is first performed to fully excavate latent model information, which will serve as guidance for better mapping learning. Following that, a C2F fusion network is constructed with a multiscale attentional fusion (MSAF) module in the head and a C2F structure in the tail. The former is deployed to achieve a more informative compression, and the latter is adopted to capture the spectral relationship, including a spectral degradation-guided (SDG) subnetwork for group-by-group coarse reconstruction and a refinement subnetwork for intergroup correlation and dependencies. Finally, high-resolution HSI can be recovered via established spectral mapping. Extensive experiments on simulated and real datasets verify the superiority of our proposed method. The code is available at https://github.com/JiaxinLiCAS/UMC2FF_GRSL.

Index Terms—Coarse-to-fine (C2F), hyperspectral image (HSI), super-resolution, unsupervised learning.

I. INTRODUCTION

HYPERSPECTRAL images (HSIs) with subtle spectral resolution are widely applied in the remote-sensing community [1], [2], [3]. However, the current imaging system

Manuscript received 13 June 2023; revised 15 August 2023; accepted 24 August 2023. Date of publication 29 August 2023; date of current version 11 September 2023. This work was supported in part by the National Natural Science Foundation of China under Grant 42101350 and Grant 42201362 and in part by the China Postdoctoral Science Foundation under Grant 2022T150080. (Corresponding author: Ke Zheng.)

Jiaxin Li and Zhi Li are with the Key Laboratory of Computational Optical Imaging Technology, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China, and also with the College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100049, China (e-mail: lijiaxin203@mails.ucas.ac.cn; lizhi21@mails.ucas.ac.cn).

Ke Zheng is with the College of Geography and Environment, Liaocheng University, Liaocheng 252059, China (e-mail: zhengke@lcu.edu.cn).

Wengu Liu is with the College of Civil Engineering, Chongqing University, Chongqing 400045, China (e-mail: 452709890@qq.com).

Haoyang Yu is with the Center of Hyperspectral Imaging in Remote Sensing (CHIRS), Information Science and Technology College, Dalian Maritime University, Dalian 116026, China (e-mail: yuhu@dltmu.edu.cn).

Li Ni is with the Key Laboratory of Computational Optical Imaging Technology, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China (e-mail: nili@radi.ac.cn).

Digital Object Identifier 10.1109/LGRS.2023.3309854

severely constrains their spatial resolution and further jeopardizes relevant applications. By contrast, multispectral images (MSIs) with high spatial resolution contain richer texture details and hence are widely employed in fine-grained tasks, for example, 4-m MSI in the GF-2 satellite and 30-m HSI in the ZY-1 02D satellite. Therefore, fusing a low-resolution HSI (LrHSI) with an auxiliary high-resolution MSI (HrMSI) becomes a straightforward way to realize hyperspectral image super-resolution (HSI-SR).

The past decade has witnessed rapid development in the HSI-SR domain, in which traditional methods and deep-learning (DL)-based methods are two main research directions. For traditional ones, researchers treat observed HSI-MSI pairs as degraded HrHSI and hence establish the degradation model via spectral response function (SRF) and point spread function (PSF). With this observation, different hand-crafted priors, including sparsity and low-rank, are introduced to regularize the ill-posed problem [4]. Beyond matrix decomposition, a high-dimensional tensor is involved to describe data cube, including Tucker, canonical polyadic, and tensor ring decomposition [5].

Recently, DL-based methods have achieved impressive success in HSI-SR owing to their powerful learning ability [6]. Sun et al. [7] proposed a multitasking network to jointly remove noise and achieve high-quality fusion outcomes. Considering the limitation of the convolution operator, some works [8] attempt to capture long-range dependencies by designing various Transformer-based structures. However, the competitive performance of the above-mentioned approaches largely depends on sufficient training triplets, that is, LrHSI, HrMSI, and HrHSI, which renders them inherently impractical in real scenarios. To mitigate this issue, some works study spectral mixing theory or deep image before realizing the unsupervised paradigm. Li et al. [9], Gao et al. [10] and Zheng et al. [11], [12] employed the autoencoder network to fulfill linear spectral unmixing and simultaneously reconstruct satisfactory HrHSI. Besides, other researchers [13] replace vanilla priors with deep networks to regularize the unknown HrHSI, forming an end-to-end optimization scheme. Different from the aforementioned strategies, Li et al. [14] reconsidered this fusion task as an MSI-aided spectral mapping problem. However, rich information hidden in the degradation model is not fully excavated, hence leaving large room for further improvement.

Considering the above problems, we propose an unsupervised model-guided C2F fusion network, UMC2FF for short,

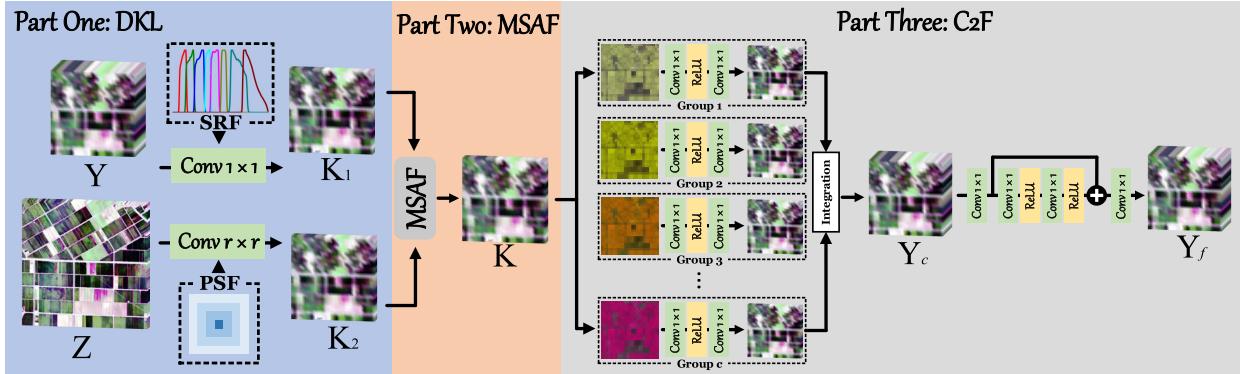


Fig. 1. Overall pipeline of the proposed UMC2FF.

to realize HSI-SR. The main contributions of our work are presented as follows.

- 1) By fully excavating degradation knowledge, we propose a model-guided network to realize unsupervised HSI-SR, which reformulates the fusion task as an MSI-aided spectral mapping problem.
- 2) By resorting to attention selective mechanism and multiscale representations, a multiscale attentional fusion (MSAF) module is proposed.
- 3) A coarse-to-fine (C2F) structure is designed to fulfill spectral mapping learning, including a spectral degradation-guided (SDG) subnetwork for a group-by-group coarse reconstruction and a refinement subnetwork for intergroup correlation and dependencies. Specifically, the knowledge of SRF is embedded into the first subnetwork to group bands by considering their spectral correlation.

II. METHODOLOGY

A. Problem Formulation

The goal of HSI-SR is to reconstruct HrHSI $\mathbf{X} \in \mathbb{R}^{HW \times C}$ by fusing observed LrHSI $\mathbf{Y} \in \mathbb{R}^{hw \times C}$ and HrMSI $\mathbf{Z} \in \mathbb{R}^{HW \times c}$, where H/h , W/w , and C/c denote the height, width, and band, respectively. Normally, $w \ll W$, $h \ll H$, and $c \ll C$ are satisfied with $r = H/h = W/w$ being the scale factor. The degradation model, which clarifies the relationship among LrHSI, HrMSI, and HrHSI, is formulated as follows:

$$\mathbf{Y} = \mathbf{P}\mathbf{X}, \quad \mathbf{Z} = \mathbf{X}\mathbf{S} \quad (1)$$

where $\mathbf{P} \in \mathbb{R}^{wh \times WH}$ and $\mathbf{S} \in \mathbb{R}^{C \times c}$ indicate PSF for spatial degradation and SRF for spectral degradation, respectively. Beyond that, a latent consistency can be established by further digging out (1)

$$\mathbf{K}_1 = \mathbf{Y}\mathbf{S}, \quad \mathbf{K}_2 = \mathbf{P}\mathbf{Z}, \quad \mathbf{K}_1 = \mathbf{K}_2. \quad (2)$$

For simplicity, we call \mathbf{K}_1 and $\mathbf{K}_2 \in \mathbb{R}^{hw \times c}$ low-resolution MSI (LrMSI). Via the same spectral degradation matrix \mathbf{S} , $\mathbf{K}_1/\mathbf{K}_2$ and \mathbf{Z} are generated from observed \mathbf{Y} and unknown \mathbf{X} , respectively. With this observation, we can indirectly capture the spectral inverse mapping from $\mathbf{K}_1/\mathbf{K}_2$ to \mathbf{Y} and then reconstruct unknown HrHSI by applying the established relationship to observed \mathbf{Z} . Based on this idea, we construct our UMC2FF network step-by-step to realize HSI-SR.

B. Overall Pipeline

The overall pipeline of UMC2FF is shown in Fig. 1, which contains three parts, including degradation knowledge learning (DKL), MSAF, and C2F mapping learning. The details of each part will be introduced in the following.

C. Degradation Knowledge Learning

According to (2), a lightweight network is designed to excavate desired degradation knowledge in preparation for the next two parts. Specifically, the spatial downsampling acts in effect as the strided $r \times r$ convolution layer, where \mathbf{P} serves as the learnable weight. Similarly, the spectral downsampling equals to 1×1 convolution layer with \mathbf{S} being the parameters of corresponding kernels. Given LrHSI \mathbf{Y} and HrMSI \mathbf{Z} , we can formulate the loss function as follows:

$$\mathcal{L}_1 = \|\text{Conv}_{1 \times 1}(\mathbf{Y}; \mathbf{S}) - \text{Conv}_{r \times r}(\mathbf{Z}; \mathbf{P})\|_1 \quad (3)$$

where $\text{Conv}_{1 \times 1}(\cdot)$ and $\text{Conv}_{r \times r}(\cdot)$ represent the spectral and spatial degradation, respectively. Once finishing the network optimization, two LrMSIs together with degradation parameters can be obtained as guidance for the construction of the following parts.

D. MSAF Module

To recover HrHSI, the spectral relationship from LrMSI to LrHSI should be established. Since two LrMSIs are jointly derived from DKL, how to fully utilize their information for better mapping learning is a key problem. Based on the selection mechanism of spectral attention, we design an MSAF module to achieve a soft fusion between \mathbf{K}_1 and \mathbf{K}_2 , as shown in Fig. 2. To encode sufficient information and achieve representative descriptors for weight learning, we first employ a three-branch layout to capture multiscale representations, aiming to cover scale variations of different objects. Here, asymmetric convolution is used to alleviate the computational burden and increase representative features. The outputs of each scale are then concatenated in pairs and fed into three individual layers for further fusion. With the summation of two vectors from global max and average pooling (GMP and GAP), the descriptor of each branch can be obtained. Lastly, we directly get the fusion weight for \mathbf{K}_1 and \mathbf{K}_2 by feeding these descriptors into 1×1 kernel and Sigmoid unit.

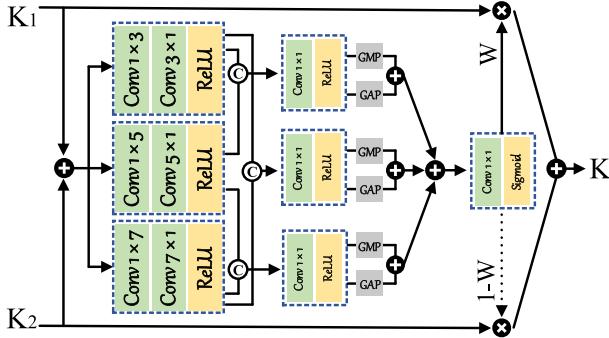


Fig. 2. Illustration for MSAF.

Therefore, we generate the fused entity via weighted averaging between \mathbf{K}_1 and \mathbf{K}_2 , which can be expressed as follows:

$$\mathbf{K} = \text{MSAF}(\mathbf{K}_1, \mathbf{K}_2) = \mathbf{W} * \mathbf{K}_1 + (1 - \mathbf{W}) * \mathbf{K}_2 \quad (4)$$

where $*$ is the channel-wise multiplication.

E. C2F Mapping Learning

Instead of directly constructing hundreds of bands synchronously, we adopt a C2F structure to progressively build the mapping relationship from fused \mathbf{K} to \mathbf{Y} , including an SDG subnetwork for a group-by-group coarse reconstruction and a refinement subnetwork for intergroup correlation and dependencies. Specifically, SRF estimated in part one provides the spectral correlation between MS and HS bands from an imaging point of view, that is, a bigger value corresponds to a higher relevance. Guided by this important degradation knowledge, the SDG subnetwork constructs c individual groups, forming c one-to-many mapping learning subproblems. In other words, each group is responsible for establishing the relationship between one band of \mathbf{K} and corresponding band sets of \mathbf{Y} , where the assignment rule is determined by the maximum response value of SRF. In each group, a simple structure with two 1×1 convolutions and a ReLU unit is introduced. According to the assignment index, the reconstructed groups are integrated together to form the coarse estimation. The whole process is formulated as follows:

$$\mathbf{Y}_c = f_{\text{coarse}}(\mathbf{K}). \quad (5)$$

However, the SDG subnetwork primarily focuses on reconstructing each group independently without considering inter-group relationships, which may lead to unexpected spectral distortion. To alleviate this problem, a refinement subnetwork is deployed to fine-tune the coarse estimation by exploiting their correlation and dependencies, which can be formulated as follows:

$$\mathbf{Y}_f = f_{\text{refine}}(\mathbf{Y}_c) \quad (6)$$

where $f_{\text{refine}}(\cdot)$ consists of four layers, with a 1×1 kernel in the head, a residual module in the body, and a 1×1 kernel in the tail. Specifically, two 1×1 kernels are utilized to separately compress and recover the spectral dimension, and the residual module is to further boost information flow

TABLE I
DATA SPECIFICATIONS

	Washington DC Mall	TianGong-1	Chikusei	LN01
Size of HrHSI	$300 \times 300 \times 191$	$240 \times 240 \times 54$	$400 \times 400 \times 110$	$360 \times 360 \times 159$
Scale factor	10	12	16	3
Size of LrHSI	$30 \times 30 \times 191$	$20 \times 20 \times 54$	$25 \times 25 \times 110$	$120 \times 120 \times 159$
Size of HrMSI	$300 \times 300 \times 8$	$240 \times 240 \times 8$	$400 \times 400 \times 8$	$360 \times 360 \times 8$

and promote learning performance. The overall C2F mapping learning is summarized as follows:

$$\mathbf{Y}_f = f_{\text{C2F}}(\mathbf{K}). \quad (7)$$

F. Training Details

HrHSI $\hat{\mathbf{X}}$ is recovered via a three-step scheme, as shown in Algorithm 1. Specifically, we first achieve DKL to obtain \mathbf{K}_1 , \mathbf{K}_2 , and degradation parameters, which will be involved in the optimization of parts two and three as guidance. By directly applying $f_{\text{C2F}}(\cdot)$ to observed \mathbf{Z} , we can obtain desired $\hat{\mathbf{X}}$. It is worth mentioning that we jointly optimize the MSAF module and C2F mapping learning, which can guarantee the optimal fusion result between \mathbf{K}_1 and \mathbf{K}_2 .

Algorithm 1 Proposed UMC2FF Algorithm

Input: Observed LrHSI \mathbf{Y} and HrMSI \mathbf{Z} .

Procedure:

Step1: Training part one:

- a) $\min \mathcal{L}_1$;
- b) Obtain \mathbf{K}_1 , \mathbf{K}_2 , and degradation parameters.

Step2: Jointly training part two and part three:

- a) $\min \|f_{\text{C2F}}(\text{MSAF}(\mathbf{K}_1, \mathbf{K}_2)) - \mathbf{Y}\|_1$.

Step3: Generate HrHSI $\hat{\mathbf{X}}$:

- a) $\hat{\mathbf{X}} = f_{\text{C2F}}(\mathbf{Z})$;

End Procedure

III. EXPERIMENTS AND RESULTS

A. Experimental Setting

Three publicly available HSI data are used in our experiments, including Washington DC Mall, TianGong-1, and Chikusei datasets. After discarding noisy and water absorption bands, some subimages are cropped as the reference for simulation experiments. Specifically, Gaussian blurring together with a downsampling operation is applied to generate LrHSI and WorldView-2 SRF is used to produce HrMSI. Beyond simulated datasets, a real dataset Liaoning-01 (LN01 for short) captured by ZY-1 02D (VNIC and AHSI cameras) over Dalian City, Liaoning Province, is used to verify the ability in practical applications. The data specifications are summarized in Table I.

Eleven state-of-the-art methods are selected as competitors, including three matrix-based methods, that is, G-SOMP+ [15], coupled spectral unmixing (CSU) [16], and coupled nonnegative matrix factorization (CNMF) [17], three tensor-based methods, that is, STEREO [18], coupled sparse tensor factorization (CSTF) [19], and SCOTT [20], and five DL-based methods, that is, DBSR [13], HyCoNet [11], CUCaNet [21], MIAE [22], and SURE [23].

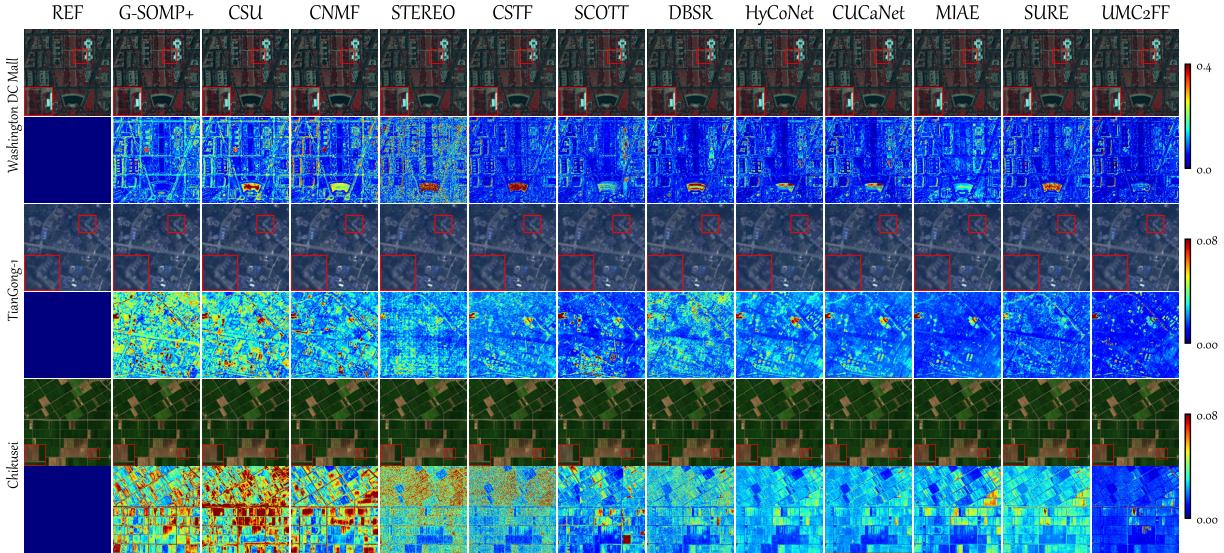


Fig. 3. Visual results of three datasets, each of which contains a pseudocolor R-G-B image and corresponding heatmap of mean relative absolute error.

TABLE II
QUANTITATIVE PERFORMANCE OF THREE DATASETS. THE BEST IS SHOWN IN BOLD

Methods	Washington DC Mall					TianGong-1					Chikusei				
	SAM	PSNR	RMSE	SSIM	UIQI	SAM	PSNR	RMSE	SSIM	UIQI	SAM	PSNR	RMSE	SSIM	UIQI
G-SOMP+	2.9933	34.2878	0.0070	0.9476	0.9840	1.7714	36.1735	0.0128	0.9715	0.9980	1.2286	38.2567	0.0100	0.9710	0.9948
CSU	2.9542	34.3353	0.0072	0.9308	0.9813	1.5471	36.8678	0.0115	0.9722	0.9983	1.3406	37.8617	0.0107	0.9476	0.9930
CNMF	2.8402	34.8903	0.0062	0.9392	0.9797	1.2508	37.0527	0.0113	0.9834	0.9988	1.0593	40.5064	0.0100	0.9867	0.9954
STEREO	2.5790	36.3611	0.0043	0.9191	0.9628	1.6001	41.5475	0.0082	0.9706	0.9991	1.0095	42.0077	0.0070	0.9731	0.9970
CSTF	1.8249	37.7773	0.0037	0.9795	0.9838	1.3019	41.3339	0.0072	0.9813	0.9993	0.8152	43.1051	0.0064	0.9772	0.9980
SCOTT	1.9705	36.6655	0.0041	0.9737	0.9809	1.2381	41.1616	0.0077	0.9828	0.9993	0.9489	44.0998	0.0063	0.9856	0.9980
DBSR	1.6278	38.1332	0.0034	0.9798	0.9820	1.3416	40.5637	0.0078	0.9800	0.9990	0.8983	45.3765	0.0059	0.9869	0.9985
HyCoNet	1.6737	38.0249	0.0035	0.9796	0.9841	1.4548	41.7587	0.0075	0.9863	0.9991	0.9101	46.0336	0.0057	0.9908	0.9985
CUCA-Net	1.7168	37.9092	0.0035	0.9794	0.9840	1.5125	41.2222	0.0087	0.9852	0.9992	0.9165	44.9185	0.0064	0.9908	0.9987
MIAE	1.3081	38.8602	0.0029	0.9724	0.9738	1.1705	43.9053	0.0063	0.9895	0.9994	0.8081	45.9483	0.0050	0.9871	0.9983
SURE	1.5861	38.6988	0.0030	0.9749	0.9843	1.2233	43.6018	0.0064	0.9820	0.9995	0.8121	44.2385	0.0057	0.9846	0.9983
UMC2FF	1.1946	39.0272	0.0025	0.9862	0.9855	1.0168	44.6499	0.0056	0.9897	0.9995	0.6565	47.4223	0.0046	0.9935	0.9993

To perform a holistic evaluation, five quantitative metrics are used for simulation experiments, including the root mean square error (RMSE), the peak signal-to-noise ratio (PSNR), the spectral angle mapper (SAM), relative dimensionless structure similarity (SSIM), and a universal image quality index (UIQI). Beyond that, we adopt visual comparison and classification performance to evaluate the real dataset due to the lack of HrHSI, where the ground-truth classification map is produced via field investigation and visual interpretation.

B. Comparison With State-of-the-Art Methods

1) *Simulated Experiments:* Table II summarizes the quantitative performance of three datasets. Specifically, CNMF obtains more stable performance compared with G-SOMP+ and CSU, and CSTF achieves satisfactory outcomes in the category of tensor representation. Compared with traditional ones, DL-based approaches achieve more competitive results. UMC2FF method outperforms all competitors in terms of three datasets, which firmly confirms our ability in spatial enhancement and spectral preservation.

The visual result illustrated in Fig. 3 further demonstrates the above-mentioned conclusions. Traditional methods are prone to spectral distortion and suffer from various artifacts. By contrast, UMC2FF obtains overall better fusion outcomes with high-quality spatial-spectral details.

2) *Real Dataset Experiments:* The results in the real dataset are shown in Fig. 4. All methods can enhance the spatial details for better visual observation. However, STEREO causes some spatial blurring and SCOTT has a local spectral distortion. HyCoNet and our UMC2FF achieve overall good recovery. Besides, our method obtains the highest overall accuracy (OA) by adopting an SVM classifier with 50 samples in each category, which confirms our potential in real applications.

C. Ablation Study

Ablation experiments are performed in the TianGong-1 dataset to demonstrate the effectiveness of the main parts in UMC2FF.

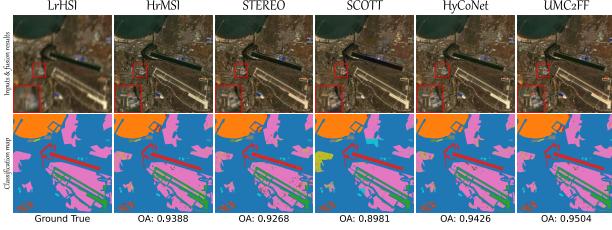


Fig. 4. Fusion outcomes and corresponding classification results in a real dataset.

TABLE III
EFFECTIVENESS OF MSAF. THE BEST IS SHOWN IN BOLD

MSAF	SAM	PSNR	RMSE	SSIM	UIQI
\times	K_1	1.0180	44.0431	0.0060	0.9874 0.9995
	K_2	1.0544	43.8576	0.0060	0.9878 0.9995
\checkmark	$K_1 + K_2$	1.0168	44.6499	0.0056	0.9897 0.9995

TABLE IV
EFFECTIVENESS OF C2F. THE BEST IS SHOWN IN BOLD

C2F	SAM	PSNR	RMSE	SSIM	UIQI
SDG	Refinement				
\times	\checkmark	1.0183	43.6610	0.0060	0.9894 0.9995
\checkmark	\times	4.5753	39.1146	0.0299	0.8125 0.9923
\checkmark	\checkmark	1.0168	44.6499	0.0056	0.9897 0.9995

1) *Effectiveness of MSAF*: To verify the influence of part two, we simply remove the MSAF module and individually feed K_1 and K_2 for mapping learning. As can be seen in Table III, the MSAF module can fully utilize multisource data and obtain performance gain compared with one input.

2) *Effectiveness of C2F*: A C2F structure is designed to realize mapping learning, and we separately remove its key components to investigate their performance. As shown in Table IV, all variants with incomplete configurations yield relatively poor performance, especially missing refinement subnetwork. As expected, the joint cooperation of the two components achieves the best results.

IV. CONCLUSION

In this letter, we propose an unsupervised network, UMC2FF, for HSI-SR. UMC2FF can fully excavate degradation model knowledge for achieving a better fusion outcome, in which the MSAF module and C2F structure are designed to establish spectral mapping. Extensive experiments in simulated and real datasets demonstrate that our method can recover detailed spatial contents and preserve spectral information.

ACKNOWLEDGMENT

The authors would like to thank Yao Liu and Prof. Haoyang Yu for providing the raw ZY-1 02D data and the Liaoning-01 dataset.

REFERENCES

- [1] D. He, Q. Shi, X. Liu, Y. Zhong, and X. Zhang, "Deep subpixel mapping based on semantic information modulated network for urban land use mapping," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 12, pp. 10628–10646, Dec. 2021.
- [2] Z. Xiong, N. Liu, N. Wang, Z. Sun, and W. Li, "Unsupervised pan-sharpening method using residual network with spatial texture attention," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5402112.
- [3] X. Zhao, W. Li, C. Zhao, and R. Tao, "Hyperspectral target detection based on weighted Cauchy distance graph and local adaptive collaborative representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5527313.
- [4] S. Li, R. Dian, and H. Liu, "Learning the external and internal priors for multispectral and hyperspectral image fusion," *Sci. China Inf. Sci.*, vol. 66, no. 4, Apr. 2023, Art. no. 140303.
- [5] M. Wang et al., "Tensor decompositions for hyperspectral data processing in remote sensing: A comprehensive review," *IEEE Geosci. Remote Sens. Mag.*, vol. 11, no. 1, pp. 26–72, Mar. 2023.
- [6] J. Li et al., "Deep learning in multimodal remote sensing data fusion: A comprehensive review," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 112, Aug. 2022, Art. no. 102926.
- [7] W. Sun et al., "MLR-DBPFN: A multi-scale low rank deep back projection fusion network for anti-noise hyperspectral and multispectral image fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5522914.
- [8] Y. Gao, M. Zhang, J. Wang, and W. Li, "Cross-scale mixing attention for multisource remote sensing data fusion and classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5507815.
- [9] J. Li, K. Zheng, Z. Li, L. Gao, and X. Jia, "X-shaped interactive autoencoders with cross-modality mutual learning for unsupervised hyperspectral image super-resolution," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5518317.
- [10] L. Gao, J. Li, K. Zheng, and X. Jia, "Enhanced autoencoders with attention-embedded degradation learning for unsupervised hyperspectral image super-resolution," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5509417.
- [11] K. Zheng et al., "Coupled convolutional neural network with adaptive response function learning for unsupervised hyperspectral super resolution," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 3, pp. 2487–2502, Mar. 2021.
- [12] K. Zheng, L. Gao, D. Hong, B. Zhang, and J. Chanussot, "NonRegSR-Net: A nonrigid registration hyperspectral super-resolution network," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5520216.
- [13] L. Zhang, J. Nie, W. Wei, Y. Li, and Y. Zhang, "Deep blind hyperspectral image super-resolution," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 6, pp. 2388–2400, Jun. 2021.
- [14] J. Li, K. Zheng, J. Yao, L. Gao, and D. Hong, "Deep unsupervised blind hyperspectral and multispectral data fusion," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.
- [15] N. Akhtar, F. Shafait, and A. Mian, "Sparse spatio-spectral representation for hyperspectral image super-resolution," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*. Cham, Switzerland: Springer, 2014, pp. 63–78.
- [16] C. Lanaras, E. Baltsavias, and K. Schindler, "Hyperspectral super-resolution by coupled spectral unmixing," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2015, pp. 3586–3594.
- [17] N. Yokoya, T. Yairi, and A. Iwasaki, "Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 2, pp. 528–537, Feb. 2012.
- [18] C. I. Kanatsoulis, X. Fu, N. D. Sidiropoulos, and W.-K. Ma, "Hyperspectral super-resolution: A coupled tensor factorization approach," *IEEE Trans. Signal Process.*, vol. 66, no. 24, pp. 6503–6517, Dec. 2018.
- [19] S. Li, R. Dian, L. Fang, and J. M. Bioucas-Dias, "Fusing hyperspectral and multispectral images via coupled sparse tensor factorization," *IEEE Trans. Image Process.*, vol. 27, no. 8, pp. 4118–4130, Aug. 2018.
- [20] C. Prévost, K. Usevich, P. Comon, and D. Brie, "Hyperspectral super-resolution with coupled Tucker approximation: Recoverability and SVD-based algorithms," *IEEE Trans. Signal Process.*, vol. 68, pp. 931–946, 2020.
- [21] J. Yao, D. Hong, J. Chanussot, D. Meng, X. Zhu, and Z. Xu, "Cross-attention in coupled unmixing nets for unsupervised hyperspectral super-resolution," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*. Cham, Switzerland: Springer, 2020, pp. 208–224.
- [22] J. Liu, Z. Wu, L. Xiao, and X.-J. Wu, "Model inspired autoencoder for unsupervised hyperspectral image super-resolution," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5522412.
- [23] H. V. Nguyen, M. O. Ulfarsson, J. R. Sveinsson, and M. D. Mura, "Deep SURE for unsupervised remote sensing image fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5412613.