

This is a repository copy of Towards lightweight hyperspectral image super-resolution with depthwise separable dilated convolutional network.

White Rose Research Online URL for this paper: <u>https://eprints.whiterose.ac.uk/id/eprint/226158/</u>

Version: Accepted Version

Proceedings Paper:

Muhammad, U., Laaksonen, J. and Mihaylova, L. orcid.org/0000-0001-5856-2223 (2025) Towards lightweight hyperspectral image super-resolution with depthwise separable dilated convolutional network. In: Proceedings of the 2025 IEEE Statistical Signal Processing Workshop (SSP). 2025 IEEE Statistical Signal Processing Workshop (SSP), 08-11 Jun 2025, Edinburgh, United Kingdom. Institute of Electrical and Electronics Engineers (IEEE), pp. 91-95. ISBN 9798331518011

https://doi.org/10.1109/SSP64130.2025.11073307

© 2025 The Authors. Except as otherwise noted, this author-accepted version of a journal article published in Proceedings of the 2025 IEEE Statistical Signal Processing Workshop (SSP) is made available via the University of Sheffield Research Publications and Copyright Policy under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here: https://creativecommons.org/licenses/

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



Towards Lightweight Hyperspectral Image Super-Resolution with Depthwise Separable Dilated Convolutional Network

Usman Muhammad¹, Jorma Laaksonen¹, and Lyudmila Mihaylova² ¹ Department of Computer Science, Aalto University, Finland ² School of Electrical and Electronic Engineering, University of Sheffield, United Kingdom

Abstract-Deep neural networks have demonstrated highly competitive performance in super-resolution (SR) for natural images by learning mappings from low-resolution (LR) to highresolution (HR) images. However, hyperspectral super-resolution remains an ill-posed problem due to the high spectral dimensionality of the data and the scarcity of available training samples. Moreover, existing methods often rely on large models with a high number of parameters or require the fusion with panchromatic or RGB images, both of which are often impractical in real-world scenarios. Inspired by the MobileNet architecture, we introduce a lightweight depthwise separable dilated convolutional network (DSDCN) to address aforementioned challenges. Specifically, our model leverages multiple depthwise separable convolutions, similar to the MobileNet architecture, and further incorporates a dilated convolution fusion block to make the model more flexible for the extraction of both spatial and spectral features. In addition, we propose a custom loss function that combines mean squared error (MSE), an L2 norm regularization-based constraint, and a spectral angle-based loss, ensuring the preservation of both spectral and spatial details. The proposed model achieves very competitive performance on two publicly available hyperspectral datasets, making it well-suited for hyperspectral image superresolution tasks. The source codes are publicly available at: https://github.com/Usman1021/lightweight.

Index Terms—Remote-sensing, dilated convolution fusion, hyperspectral imaging, lightweight model, loss function.

I. INTRODUCTION

Hyperspectral imaging systems capture surface information across numerous spectral bands, providing richer spectral details than multispectral or natural images. This enables precise characterization of ground objects [1]-[4]. Moreover, the wide spectrum of information makes them extremely valuable for a variety of applications, including anomaly detection, surveillance, environmental monitoring, and satellite image scene classification [5]–[9]. However, hyperspectral images often suffer from low resolution due to environmental factors and sensor energy limitations [10]. Deep learning-based super-resolution (SR) methods have shown remarkable success, leveraging large models to reconstruct high-resolution images from degraded low-resolution inputs [11], [12]. Nonetheless, due to advancements in deep learning models, the number of parameters in deep learning networks has increased significantly. As a result, numerous lightweight super-resolution models have gained significant attention in recent years [13], focusing on reducing model parameters and computational complexity through various strategies.

In particular, knowledge distillation [14] is widely used for model compression, in which a large model (the teacher network) is first trained on the original data and then used to guide the intermediate feature representations of a smaller model (the student network). For instance, Gao et al. [15] employ a lightweight student SR model to acquire knowledge from a deeper teacher SR network. Other model compression techniques, such as pruning [16] and low-rank factorization [17], have also proven effective for developing super-resolution models. For instance, Zhan et al. [16] integrate neural architecture search with a layer-wise pruning strategy to develop efficient super-resolution model. Wu et al. [17] propose a convex regularizer for low-rank solutions, addressing unconstrained cases with proximal gradient and a custom PG-like algorithm. In addition, some quantization-based methods [18], [19] leverage low-bit representations to enhance the inference speed of super-resolution models. However, these models often exhibit a significant performance gap compared to state-of-theart super-resolution methods [20].

In order to reduce the size of the neural network, convolutional filters play a crucial role, and significant efforts have been made to maximize their potential. For instance, MobileNet [21] model introduces depthwise separable convolution, demonstrating great potential for network decomposition. Ahn et al. [22] employ group convolution instead of depthwise convolution to make model efficiency more adaptable. Kim et al. [23] propose a deeply recursive convolutional network, arguing that increasing recursion depth can enhance performance without introducing additional parameters for extra convolutions. Dilated convolution is another modified version of standard convolutions that utilizes a distinctive grid structure [24]. It has been widely adopted to expand the receptive field of kernels and enhance global information aggregation, leading to notable improvements over previous methods. However, the classification accuracy of these models may be compromised during compression, as a large number of image features are omitted due to the simplified convolution operations [25].

Our approach is inspired by MobileNet [21], which reduces the number of parameters and computational costs by intro-



Fig. 1. A detailed overview of the proposed model. Shallow features are initially extracted using a lightweight separable convolutional layer with ReLU activation. The gray color represent depthwise convolutional blocks for deep feature extraction. The blue block denotes the dilated fusion module, capturing spatial and spectral features in parallel. The pink block performs upsampling to reconstruct the high-resolution output.

ducing the concept of depthwise separable convolution. In particular, we aim to combine depthwise separable convolutions with dilated convolutional fusion blocks [26] to develop a novel super-resolution model called the depthwise separable dilated convolution network (DSDCN) for hyperspectral images. Specifically, the dilated convolutional fusion block consists of two different types of convolutions: (1) dilated convolution (spatial branch) and (2) pointwise convolution (spectral branch). The outputs of the dilated convolution and pointwise convolution are then concatenated to effectively integrate spatial and spectral information. By doing so, we make our model suitable for hyperspectral super-resolution tasks while minimizing the compromise on model parameters compared to the MobileNet architecture. To further improve the performance of the proposed method, we employ a custom loss function that integrates mean squared error (MSE), an L2 regularizationbased constraint, and a spectral angle-based loss to ensure highfidelity reconstruction. In summary, our contributions are threefold:

- We present a novel lightweight DSDCN that combines depthwise separable convolutions, residual connections, and dilated convolutional fusion to enhance spatial resolution and preserve spectral integrity..
- 2) A custom loss function is introduced, integrating mean squared error (MSE), an L2 regularization constraint, and a spectral angle-based loss to enhance high-fidelity reconstruction.
- Experiments on two hyperspectral datasets are conducted across various resolution degradation-restoration scenarios (2×, 4×, and 8× downsampling), demonstrating competitive performance on both datasets.

II. METHODOLOGY

Fig. 1 shows the band grouping and the three main components of the model: (1) depthwise separable convolutions with residual connections, (2) dilated convolution fusion, and (3) upsampling. We first define the band grouping along with other components. The following subsections detail this and the custom loss used in the model.

A. Band Grouping

Since hyperspectral images consist of hundreds of spectral bands, processing all bands together can be computationally challenging and may introduce redundancy due to the high spectral correlation between bands. To mitigate this, we utilize band grouping [27], which involves partitioning adjacent bands into overlapping groups for seamless integration with our proposed model. Specifically, hyperspectral bands are organized into overlapping subgroups by defining a fixed group size with a designated overlap, ensuring that consecutive subgroups share common bands.

B. Depthwise Separable Convolution

Given a low-resolution hyperspectral image $X \in \mathbb{R}^{H \times W \times B}$, where H, W, and B denote the spatial height, width, and number of spectral bands, respectively, our goal is to reconstruct a high-resolution image $\hat{X} \in \mathbb{R}^{\alpha H \times \alpha W \times B}$ with an upscaling factor α . To achieve this, we design a deep neural network $\mathcal{F}(X;\theta)$ that learns the LR-to-HR mapping efficiently while preserving spectral integrity.

We begin by extracting spatial and spectral features using depthwise separable convolutions. Specifically, we empirically employ three such blocks, where each block decomposes a standard convolution into two sequential operations [21]:

$$X'_d = W_d * X,\tag{1}$$

where W_d denotes the depthwise convolution kernel applied independently to each spectral band. Next, a pointwise convolution is applied to combine features across channels:

$$X'_p = W_p * X'_d,\tag{2}$$

where W_p is a 1×1 convolution kernel that processes each spatial location across all channels. To improve training stability

and mitigate vanishing gradients, we incorporate a residual connection by applying a separate 1×1 convolution to the original input X [28]:

$$X_{\rm res} = W_s * X,\tag{3}$$

where W_s is the projection kernel for aligning dimensions. The final output of the block is obtained by fusing the residual and transformed features:

$$C_{\rm out} = X'_p + X_{\rm res},\tag{4}$$

where C_{out} is the output feature map of the depthwise separable convolution block.

C. Dilated Convolution Fusion

To integrate both local and global spatial dependencies while maintaining a lightweight structure, we employ a *dilated fusion block* composed of three parallel 3×3 convolutions with increasing dilation rates $r = \{1, 2, 3\}$. Given an input feature map $\mathbf{F}_{in} \in \mathbb{R}^{H \times W \times C_{in}}$, feature extraction is performed as [26]:

$$\mathbf{F}_i = \operatorname{ReLU}(\mathbf{Q}_i * \mathbf{F}_{in}), \quad i \in \{1, 2, 3\},$$
(5)

where each $\mathbf{Q}_i \in \mathbb{R}^{3 \times 3 \times C_{\text{in}} \times C}$ is a convolutional kernel with dilation rate $r_i \in \{1, 2, 3\}$, and C is the number of output channels per branch. The feature maps \mathbf{F}_1 , \mathbf{F}_2 , and \mathbf{F}_3 are obtained by applying ReLU-activated convolutions with kernels \mathbf{Q}_1 , \mathbf{Q}_2 , and \mathbf{Q}_3 , respectively, to the same input \mathbf{F}_{in} , and are then concatenated along the channel dimension.

$$\mathbf{F}_{\text{concat}} = \text{Concat}(\mathbf{F}_1, \mathbf{F}_2, \mathbf{F}_3), \tag{6}$$

The concatenated features are then passed through a final 1×1 convolution to fuse the multi-scale information:

$$\mathbf{F}_{\text{out}} = \text{ReLU}(\mathbf{Q}_f * \mathbf{F}_{\text{concat}}),\tag{7}$$

where $\mathbf{Q}_f \in \mathbb{R}^{1 \times 1 \times 3C \times C}$ is the fusion kernel. The output $\mathbf{F}_{out} \in \mathbb{R}^{H \times W \times C}$ represents the final feature map enriched with contextual spatial features.

D. Upsampling with Transpose Convolution

To increase spatial resolution, we employ an upsampling block based on transpose convolution [29]. Given a lowresolution feature map $\mathbf{U}_{in} \in \mathbb{R}^{H \times W \times C_u}$, we apply a learnable transpose convolution defined as:

$$\mathbf{U}_{\rm tr} = \mathbf{K}_{\rm tr} * \mathbf{U}_{\rm in},\tag{8}$$

where $\mathbf{K}_{tr} \in \mathbb{R}^{k \times k \times C_u \times C_o}$ is the transpose convolution kernel, and $\mathbf{U}_{tr} \in \mathbb{R}^{\alpha H \times \alpha W \times C_o}$ is the upsampled output. Here, C_u is the number of input channels, and C_o is the number of output channels after upsampling. The upsampled output is then passed through a non-linearity:

$$\mathbf{U}_{act} = \operatorname{ReLU}(\mathbf{U}_{tr}). \tag{9}$$

Furthermore, we employ a skip connection by projecting and upsampling the input feature map:

$$\mathbf{U}_{skip} = \mathbf{K}_{skip} * \mathrm{Up}(\mathbf{U}_{in}), \tag{10}$$

 TABLE I

 QUANTITATIVE RESULTS AND MODEL COMPLEXITY

Ablation study on PaviaC (2×)							
Model Variant	MSSIM↑	MPSNR↑					
DSDCN w/o band grouping	0.9482	35.017					
DSDCN w/s 16	0.9558	35.447					
DSDCN + w/s 32	0.9578	36.434					
DSDCN w/s 48	0.9502	35.861					
DSDCN + w/o custom loss	0.9507	36.127					
Model Complexity							
Model	Scale	Parameters					
ERCSR [30]	4	1.59M					
MCNet [31]	4	2.17M					
PDENet [11]	4	2.30M					
CSSFENet [20]	4	1.61M					
DSDCN (Ours)	4	0.96M					

where $\mathbf{K}_{\text{skip}} \in \mathbb{R}^{1 \times 1 \times C_u \times C_o}$ is a projection kernel. The final output is obtained by summing the skip and upsampled features:

$$\mathbf{U}_{\text{out}} = \mathbf{U}_{\text{act}} + \mathbf{U}_{\text{skip}},\tag{11}$$

where $\mathbf{U}_{\text{out}} \in \mathbb{R}^{\alpha H \times \alpha W \times C_o}$ is the resulting feature map.

E. Custom Loss Function

We employ a custom loss function that combines three components: mean squared error (MSE), spectral angle mapper (SAM) loss, and ℓ_2 loss. This combination ensures both pixel-wise accuracy and spectral fidelity in hyperspectral reconstruction. The total loss is defined as:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{MSE}} + \lambda_1 \cdot \mathcal{L}_{\text{SAM}} + \lambda_2 \cdot \mathcal{L}_{\ell_2}, \qquad (12)$$

where: $\mathcal{L}_{\text{MSE}} = \frac{1}{N} \sum_{i=1}^{N} (y_{\text{true}}^{(i)} - y_{\text{pred}}^{(i)})^2$ minimizes pixel-wise differences, $\mathcal{L}_{\text{SAM}} = \frac{1}{N} \sum_{i=1}^{N} \cos^{-1} \left(\frac{\langle y_{\text{true}}^{(i)}, y_{\text{pred}}^{(i)} \rangle}{\|y_{\text{true}}^{(i)}\| \|y_{\text{pred}}^{(i)}\|} \right)$ measures the angular discrepancy between predicted and ground-truth spectral vectors, $\mathcal{L}_{\ell_2} = \frac{1}{N} \sum_{i=1}^{N} \|y_{\text{true}}^{(i)} - y_{\text{pred}}^{(i)}\|_2^2$ reinforces global spectral consistency. Here, λ_1 and λ_2 are weighting coefficients that balance the influence of SAM and ℓ_2 loss, respectively. In our experiments, we set $\lambda_1 = 0.5$ and $\lambda_2 = 0.03$.

III. EXPERIMENTAL SETUP

A. Datasets and Implementation

Two publicly available hyperspectral datasets, PaviaC and PaviaU, containing 102 and 103 spectral bands, respectively, are used in our study. We used a patch size of 144×144 , following the protocol of previous work [20], for training and testing the datasets. Specifically, the test set patch is extracted from the bottom center of the PaviaC dataset and the top left of the PaviaU dataset, while the remaining image regions are used for training [20].

To generate low-resolution images, we employ area-based sampling and downscale them by factors of $2\times$, $4\times$, and $8\times$. During training, the Adam optimizer was employed with a batch size of 4. Additionally, an early stopping function was used to prevent fixed epochs and avoid overfitting. Both datasets were divided into band groups of size 32, with one-fourth

 TABLE II

 EVALUATION ON DATASETS (PAVIAC, PAVIAU) IN DIFFERENT SCALING SETUPS. THE COMPARISON RESULTS ARE REPORTED FROM [20].

	N 11	PaviaC		PaviaU			
Scale Factor	Nidel	MPSNR ↑	MSSIM↑	SAM↓	MPSNR ↑	MSSIM ↑	SAM↓
2×	VDSR [32]	34.879	0.9501	3.689	34.038	0.9524	3.258
	MCNet [31]	34.626	0.9455	3.865	33.743	0.9502	3.359
	EDSR [33]	34.580	0.9452	3.898	33.985	0.9511	3.334
	MSDformer [12]	35.028	0.9493	3.691	34.159	0.9553	3.211
	MSFMNet [34]	35.200	0.9506	3.656	34.980	0.9582	3.160
	AS3 ITransUNet [35]	35.221	0.9511	3.612	35.163	0.9591	3.149
	PDENet [11]	35.244	0.9519	3.595	35.275	0.9594	3.142
	CSSFENet [20]	35.522	0.9544	3.542	35.924	0.9625	3.038
	DSDCN (Ours)	36.434	0.9578	3.538	35.941	0.9442	3.703
4×	EDSR [33]	28.591	0.7782	6.573	29.894	0.7791	5.074
	VDSR [32]	28.317	0.7707	6.514	29.904	0.7753	4.997
	MCNet [31]	28.756	0.7826	6.385	29.993	0.7835	4.917
	MSDformer [12]	28.810	0.7833	5.897	30.098	0.7905	4.885
	MSFMNet [34]	28.873	0.7863	6.300	30.283	0.7948	4.861
	AS3 ITransUNet [35]	28.874	0.7893	5.972	30.289	0.7940	4.859
	PDENet [11]	28.951	0.7900	5.876	30.295	0.7944	4.853
	CSSFENet [20]	29.054	0.7961	5.816	30.689	0.8107	4.839
	DSDCN (Ours)	29.665	0.8152	4.826	30.524	0.7958	4.807
8×	VDSR [32]	24.804	0.4944	7.588	27.028	0.5962	7.133
	EDSR [33]	25.067	0.5282	7.507	27.467	0.6302	6.678
	MCNet [31]	25.096	0.5391	7.429	27.483	0.6254	6.683
	MSDformer [12]	25.215	0.5462	7.427	27.323	0.6341	6.668
	MSFMNet [34]	25.257	0.5464	7.449	27.586	0.6356	6.615
	AS3 ITransUNet [35]	25.258	0.5435	7.417	27.689	0.6413	6.574
	PDENet [11]	25.288	0.5436	7.402	27.738	0.6457	6.531
	CSSFENet [20]	25.359	0.5493	7.306	27.825	0.6569	6.505
	DSDCN (Ours)	25.463	0.5553	6.345	27.827	0.6235	6.487

overlap between them. We adopted several widely used metrics to evaluate the quality of reconstructed images quantitatively, including mean peak signal-to-noise ratio (MPSNR), mean structural similarity index (MSSIM), and Spectral Angle Mapper (SAM) [36].

B. Ablation Study on Band Grouping

Table I presents an ablation study evaluating the effects of band grouping size and the custom loss function on the performance of the proposed DSDCN model. The baseline without band grouping yields an MSSIM of 0.9482 and an MPSNR of 35.017 dB. It can be observed that using a band grouping size of 16 improves performance, while a grouping size of 32 achieves the best results (MSSIM: 0.9578, MPSNR: 36.434 dB). However, increasing the group size to 48 slightly reduces accuracy, suggesting that excessive grouping may weaken the spatial-spectral feature alignment. Performance is also reported using only the MSE loss (without the custom loss), with the group size fixed at 32. This results in a performance drop (MSSIM: 0.9507, MPSNR: 36.127 dB), highlighting the benefits of incorporating spectral angle and L2 constraints into the loss function.

C. Comparison with State-of-the-Art Methods

As shown in Table I, we compare the parameter count of our proposed model with several state-of-the-art super-resolution methods, including ERCSR (1.59M) [30], MCNet (2.17M) [31], PDENet (2.30M) [11], and CSSFENet (1.61M) [20]. DSDCN contains only 0.96 million parameters, substantially fewer than those of the other methods.

Table II presents a detailed quantitative comparison under multiple scaling factors (2×, 4×, and 8×) on both the PaviaC and PaviaU datasets. The comparison includes various stateof-the-art methods such as VDSR [32], EDSR [33], MCNet [31], MSDformer [12], MSFMNet [34], AS3 ITransUNet [35], PDENet [11], and CSSFENet [20]. Despite its lightweight nature, our model ranks first in the PaviaC dataset and achieves competitive performance on the PaviaU dataset in terms of PSNR and SAM. These results emphasize the strength of our architecture in capturing both spatial and spectral information for hyperspectral image super-resolution.

IV. CONCLUSION

In this study, we introduced a lightweight depthwise separable dilated convolutional network (DSDCN) for hyperspectral super-resolution. The model leverages depthwise separable convolutional blocks with a dilated convolution fusion mechanism to improve spatial and spectral feature extraction. Additionally, we proposed a custom loss function that integrates MSE, L2 regularization, and a spectral angle-based loss to ensure highfidelity reconstruction. Experimental results demonstrate that our approach achieves competitive performance on two publicly available hyperspectral datasets while maintaining a relatively compact model size. This makes our method a practical and efficient solution for real-world hyperspectral image superresolution tasks.

ACKNOWLEDGMENT

This project has been funded by the European Union's NextGenerationEU instrument and the Research Council of

Finland under grant N_{\circ} 348153, as part of the project Artificial Intelligence for Twinning the Diversity, Productivity and Spectral Signature of Forests (ARTISDIG). We also gratefully acknowledge CSC for providing access to the LUMI supercomputer, operated by the EuroHPC Joint Undertaking.

REFERENCES

- P. Singh, P. C. Pandey, G. P. Petropoulos, A. Pavlides, P. K. Srivastava, N. Koutsias, K. A. K. Deng, and Y. Bao, "Hyperspectral remote sensing in precision agriculture: Present status, challenges, and future trends," in *Hyperspectral remote sensing*. Elsevier, 2020, pp. 121–146.
- [2] L. Loncan, L. B. de Almeida, J. M. Bioucas-Dias, X. Briottet, J. Chanussot, N. Dobigeon, S. Fabre, W. Liao, G. A. Licciardi, M. Simões, J.-Y. Tourneret, M. A. Veganzones, G. Vivone, Q. Wei, and N. Yokoya, "Hyperspectral pansharpening: A review," *IEEE Geoscience and Remote Sensing Magazine*, vol. 3, no. 3, pp. 27–46, 2015.
- [3] Y. Altmann, A. Halimi, N. Dobigeon, and J.-Y. Tourneret, "Supervised nonlinear spectral unmixing using a postnonlinear mixing model for hyperspectral imagery," *IEEE Transactions on Image Processing*, vol. 21, no. 6, pp. 3017–3025, 2012.
- [4] R. A. Borsoi, T. Imbiriba, J. C. M. Bermudez, C. Richard, J. Chanussot, L. Drumetz, J.-Y. Tourneret, A. Zare, and C. Jutten, "Spectral variability in hyperspectral data unmixing: A comprehensive review," *IEEE Geoscience and Remote Sensing Magazine*, vol. 9, no. 4, pp. 223–270, 2021.
- [5] U. Muhammad, W. Wang, S. P. Chattha, and S. Ali, "Pre-trained vggnet architecture for remote-sensing image scene classification," in 2018 24th International Conference on Pattern Recognition (ICPR). IEEE, 2018, pp. 1622–1627.
- [6] U. Muhammad, W. Wang, and A. Hadid, "Feature fusion with deep supervision for remote-sensing image scene classification," in 2018 IEEE 30th international conference on tools with artificial intelligence (ICTAI). IEEE, 2018, pp. 249–253.
- [7] U. Muhammad, W. Wang, A. Hadid, and S. Pervez, "Bag of words kaze (bowk) with two-step classification for high-resolution remote sensing images," *IET Computer Vision*, vol. 13, no. 4, pp. 395–403, 2019.
- [8] U. Muhammad, M. Z. Hoque, W. Wang, and M. Oussalah, "Patch-based discriminative learning for remote sensing scene classification," *Remote Sensing*, vol. 14, no. 23, p. 5913, 2022.
- [9] C. Peng, P. Shu, X. Huang, Z. Fu, and X. Li, "Lcrca: image superresolution using lightweight concatenated residual channel attention networks," *Applied Intelligence*, pp. 1–15, 2022.
- [10] L. Bu, X. Xiao, Z. Zhang, and D. Mingjun, "Sbhsr: single-band superresolution method for hyperspectral images based on blind degradation and fusion of auxiliary band," *Geocarto International*, vol. 39, no. 1, p. 2294901, 2024.
- [11] J. Hou, Z. Zhu, J. Hou, H. Zeng, J. Wu, and J. Zhou, "Deep posterior distribution-based embedding for hyperspectral image super-resolution," *IEEE Transactions on Image Processing*, vol. 31, pp. 5720–5732, 2022.
- [12] S. Chen, L. Zhang, and L. Zhang, "Msdformer: Multi-scale deformable transformer for hyperspectral image super-resolution," *IEEE Transactions* on Geoscience and Remote Sensing, 2023.
- [13] Y. Guo, J. Wang, Q. Chen, J. Cao, Z. Deng, Y. Xu, J. Chen, and M. Tan, "Towards lightweight super-resolution with dual regression learning," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024.
- [14] W. Lee, J. Lee, D. Kim, and B. Ham, "Learning with privileged information for efficient image super-resolution," in *Computer Vision–ECCV* 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXIV 16. Springer, 2020, pp. 465–482.
- [15] Q. Gao, Y. Zhao, G. Li, and T. Tong, "Image super-resolution using knowledge distillation," in *Asian Conference on Computer Vision*. Springer, 2018, pp. 527–541.
- [16] Z. Zhan, Y. Gong, P. Zhao, G. Yuan, W. Niu, Y. Wu, T. Zhang, M. Jayaweera, D. Kaeli, B. Ren *et al.*, "Achieving on-mobile realtime super-resolution with neural architecture and pruning search," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 4821–4831.
- [17] R. Wu, Q. Li, X. Fu, and W.-K. Ma, "A convex low-rank regularization method for hyperspectral super-resolution," in 2018 IEEE Statistical Signal Processing Workshop (SSP). IEEE, 2018, pp. 383–387.

- [18] J. Xin, N. Wang, X. Jiang, J. Li, H. Huang, and X. Gao, "Binarized neural network for single image super resolution," in *Computer Vision–ECCV* 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part IV 16. Springer, 2020, pp. 91–107.
- [19] Y. Ma, H. Xiong, Z. Hu, and L. Ma, "Efficient super resolution using binarized neural network," in *Proceedings of the IEEE/CVF conference* on computer vision and pattern recognition workshops, 2019, pp. 0–0.
- [20] J. Zhang, R. Zheng, Z. Wan, R. Geng, Y. Wang, Y. Yang, X. Zhang, and Y. Li, "Hyperspectral image super-resolution based on feature diversity extraction," *Remote Sensing*, vol. 16, no. 3, p. 436, 2024.
- [21] A. G. Howard, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," arXiv preprint arXiv:1704.04861, 2017.
- [22] N. Ahn, B. Kang, and K.-A. Sohn, "Fast, accurate, and lightweight super-resolution with cascading residual network," in *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 252–268.
- [23] J. Kim, J. K. Lee, and K. M. Lee, "Deeply-recursive convolutional network for image super-resolution," in *Proceedings of the IEEE conference* on computer vision and pattern recognition, 2016, pp. 1637–1645.
- [24] H. Hu, C. Yu, Q. Zhou, Q. Guan, and H. Feng, "Hdconv: Heterogeneous kernel-based dilated convolutions," *Neural Networks*, vol. 179, p. 106568, 2024.
- [25] W. Sun, X. Zhou, X. Zhang, and X. He, "A lightweight neural network combining dilated convolution and depthwise separable convolution," in *Cloud Computing, Smart Grid and Innovative Frontiers in Telecommunications: 9th EAI International Conference, CloudComp 2019, and 4th EAI International Conference, SmartGIFT 2019, Beijing, China, December 4-5, 2019, and December 21-22, 2019 9.* Springer, 2020, pp. 210–220.
- [26] F. Yu and V. Koltun, "Multi-scale context aggregation by dilated convolutions," arXiv preprint arXiv:1511.07122, 2015.
- [27] Z. Wang, D. Li, M. Zhang, H. Luo, and M. Gong, "Enhancing hyperspectral images via diffusion model and group-autoencoder superresolution network," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 38, no. 6, 2024, pp. 5794–5804.
- [28] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in *Proceedings* of the IEEE conference on computer vision and pattern recognition, 2018, pp. 4510–4520.
- [29] C. Dong, C. C. Loy, and X. Tang, "Accelerating the super-resolution convolutional neural network," in *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14,* 2016, Proceedings, Part II 14. Springer, 2016, pp. 391–407.
- [30] Q. Li, Q. Wang, and X. Li, "Exploring the relationship between 2d/3d convolution for hyperspectral image super-resolution," *IEEE Transactions* on Geoscience and Remote Sensing, vol. 59, no. 10, pp. 8693–8703, 2021.
- [31] Q. W. Qiang Li and X. Li, "Mixed 2d/3d convolutional network for hyperspectral image super-resolution," *Remote sensing*, vol. 12, no. 10, p. 1660, 2020.
- [32] J. Kim, J. K. Lee, and K. M. Lee, "Accurate image super-resolution using very deep convolutional networks," in *Proceedings of the IEEE conference* on computer vision and pattern recognition, 2016, pp. 1646–1654.
- [33] B. Lim, S. Son, H. Kim, S. Nah, and K. Mu Lee, "Enhanced deep residual networks for single image super-resolution," in *Proceedings of the IEEE* conference on computer vision and pattern recognition workshops, 2017, pp. 136–144.
- [34] J. Zhang, M. Shao, Z. Wan, and Y. Li, "Multi-scale feature mapping network for hyperspectral image super-resolution," *Remote Sensing*, vol. 13, no. 20, p. 4180, 2021.
- [35] Q. Xu, S. Liu, J. Wang, B. Jiang, and J. Tang, "As 3 itransunet: Spatial-spectral interactive transformer u-net with alternating sampling for hyperspectral image super-resolution," *IEEE Transactions on Geoscience* and Remote Sensing, 2023.
- [36] Y. Chudasama, U. Muhammad, V. Mayra, F. Guiotte, and J. Laaksonen, "A comparison of hyperspectral super-resolution techniques for boreal forest imagery," in *IGARSS 2024-2024 IEEE International Geoscience* and Remote Sensing Symposium. IEEE, 2024, pp. 1226–1230.