DEPTH SEPARABLE ARCHITECTURE FOR SENTINEL-5P SUPER-RESOLUTION

Hyam Omar Ali[a,b], Romain Abraham[a], Bruno Galerne[a,c]

[a] Institut Denis Poisson, Université d'Orléans, Université de Tours, CNRS, France
[b] Faculty of Mathematical Sciences, University of Khartoum, Sudan
[c] Institut universitaire de France (IUF)

Abstract—Sentinel-5P (S5P) satellite provides atmospheric measurements for air quality and climate monitoring. While the S5P satellite offers rich spectral resolution, it inherits physical limitations that restricts its spatial resolution. Super-resolution (SR) techniques can overcome these limitations and enhance the spatial resolution of S5P data. In this work, we introduce a novel SR model specifically designed for S5P data that have eight spectral bands with around 500 channels for each band. Our proposed S5-DSCR model relies on Depth Separable Convolution (DSC) architecture to effectively perform spatial SR by exploiting cross-channel correlations. Quantitative evaluation demonstrates that our model outperforms existing methods for the majority of the spectral bands. This work highlights the potential of leveraging DSC architecture to address the challenges of hyperspectral SR. Our model allows for capturing fine details necessary for precise analysis and paves the way for advancements in air quality monitoring as well as remote sensing applications.

Index Terms—Sentinel-5P, Remote Sensing, Super-Resolution, Hyperspectral Images, Depth Separable Convolution

I. INTRODUCTION

In 2017, the Sentinel-5 Precursor (Sentinel-5P or S5P) satellite was launched by the European Space Agency (ESA) in collaboration with the European Commission aiming for enhanced Earth observation capabilities [1]. The satellite monitors the atmospheric species which impact global air quality, greenhouse gases, and ozone layer dynamics [2]. It provides data relevant to atmospheric and environmental observations and climate research. This global data supports policy-making and WHO efforts towards improving public health, prioritizing global air quality monitoring and reducing exposure to air pollution [3].

The S5P operates into 8 different spectral bands [4]. Through these bands, S5P measures radiance data across 3942 frequency spectral channels and captures hyperspectral (HS) images enabling the spectral analysis. The pixels of the HS images are generated through a combination of scans in the cross-track direction (perpendicular to the satellite's motion) and the ones in the along-track (parallel to the satellite's motion) [5]. The spatial resolution of the images is approximately $3.5 \times 5.5 \, km^2$ per pixel [1, 5].

While the S5P satellite offers rich spectral resolution, making it an invaluable tool for environmental monitoring, its spatial resolution is limited by physical constraints, such as the trade-off between pixel size and the amount of light detected, potentially affecting image quality [6]. These limitations restrict its potential for finer-scale analysis, necessitating ad-vanced image processing techniques such as super-resolution (SR) to enhance spatial resolution and improve the quality of S5P data for more precise environmental monitoring [7].

The unique spectral characteristics and inter-channel relationships of S5P HS data necessitate the development of more specific SR models. The S5Net model has recently been introduced as the first tailored SR approach for S5P data. Initially, this model was designed to handle individual spectral bands and a single channel at a time [8]. Later, S5Net was extended to process multiple spectral bands considering all the channels [9]. In this work, we present a novel approach, S5-DSCR, to enhance the spatial resolution of S5P HS images. Our contributions are as follows:

- We use Depth Separable Convolution (DSC) architecture that captures both the spatial and spectral relationships within and across channels. This lightweight architecture ensures capturing the interdependencies across all the spectral channels while reducing complexity.
- Considering the unique spectral characteristics of each band, we trained the same model separately for each specific band. This approach ensures the model is optimized and efficient for the unique features of each individual band.
- We compiled a comprehensive dataset comprising HS images from 15 orbits representing diverse spatial and spectral characteristics and providing maximum coverage of Earth's surface.

The plan of this paper is as follows: Section II reviews the related work. The methodology is described in Section III including problem formulation and the proposed method to resolve. Then, Section IV details the experimental setup. Section V presents and analyzes the results. Finally, Section VI concludes and draws potential perspectives

II. RELATED WORK

SR is an important branch of image processing that produces high-resolution (HR) images from low-resolution (LR) images. SR techniques are broadly classified into single-image super-resolution (SISR) and multi-image super-resolution (MISR) [10]. Since each S5P scan is unique, this work is focused on SISR approaches.

Over the years, numerous methods have been proposed to address SISR [10–14]. The classical and earliest approaches are interpolation-based techniques [15–19], such

as nearest neighbour and bicubic interpolation. Variational techniques [20–24], such as Total Variation (TV), iterative back-projection, and maximum a posteriori (MAP), leverage optimization and regularization techniques to enhance the SISR. However, with the rapid development of resources and techniques, deep learning approaches became dominant and achieved state-of-the-art performance in SR tasks. These approaches map the features of LR images to HR images. Although all the approaches basically employ the same essential ideas, they differ in network architecture and optimization. Designs such as residual networks [25–32], dense connections [33–36], attention mechanisms [30, 31, 37–39], GANs [33, 40–43], diffusion models [44], and group learning [29, 32, 45, 46], represent the diversity in approaches.

Unlike grey or RGB images, HS images have hundreds of spectral channels. This makes SR more challenging due to the need to map the features of a larger number of channels. The overall complexity of the model increases gradually as the number of channels increases making the process more challenging to consider the spectral dimension [14]. Considering reducing dimensions and applying convolution along the spectral dimension can achieve more improvement in SR for HS images [47–49]. However, as HS images exhibit a high correlation along the spectral dimensions, often these methods neglect to account for this correlation [50]. DSC addresses this limitation by representing standard convolution across the spectral channels with convolutions that incorporate the features across both the spectral channels and spatial dimension [51]. Several studies used DSC for SR [52–56].

Remote sensing imaging, particularly S5P images, differs from classical HS images due to its unique multi-band structure, spanning eight bands and almost 500 spectral channels with varying resolutions. For this reason, S5P images require the development of SR models tailored specifically to preserve spectral fidelity and handle its characteristics. The S5Net model introduced the first deep-learning approach designed specifically for S5P images with lightweight architecture [8]. Initially, this model was developed for the monochromic images (central spectral channel only), and S5net was later modified to process all the channels of the S5P images through the fine-tuning cascade procedure [9]. This approach begins at the central channel and iteratively expands to include all the spectral channels. However, it still processes each channel individually, missing the utilization of inter-channel information.

III. PROPOSED METHOD

A. Problem Formulation

The objective of this study is to enhance the spatial resolution of Sentinel-5P Level-1B radiance products using SISR techniques. For each band, the HS image consists of multiple spectral channels with relatively low spatial resolution. Following [9], to demonstrate the feasibility of SISR in S5P data, the original HS images were degraded to obtain the LR images while the original HS images acted as the HR ground truth.

(a) Architecture of S5-DSCR (L=5) and S5-DSCR-S (L=1) models
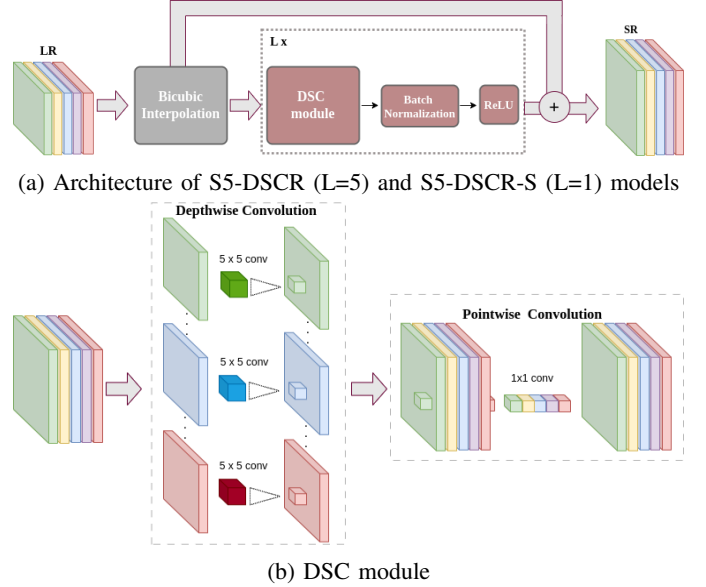

(b) DSC module

Fig. 1: Overview of our proposed models and their inner DSC module

The original degradation model [8] was used to simulate the LR images. This process involves simulating the real-world image acquisition process, such as blurring, to create realistic LR images that closely resemble the actual HS images. The degradation process involved applying a Point Spread Function (PSF) to model the blurring effect. The blurring kernel K is simulated by asymmetric Gaussian function with different standard deviations in the along-track and cross-track directions. For S5P data, the standard deviations vary for each detector [8]: 0.37 and 0.36 (UV), 0.44 and 0.74 (UVIS), 0.45 and 0.74 (NIR), and 0.15 and 0.20 (SWIR) in the across-track and along-track directions, respectively. A scaling factor of 4 was utilized to downsample the HR images.

B. Architecture

Our proposed model builds based on the DSC architecture. DSC is composed of two stages (Figure 1b): depthwise convolution (intra-channel) in which each input channel is independently convolved with a filter to extract features of this channel, followed by a pointwise convolution (intra-pixel) that combines the outputs of the depthwise stage with 1×1 convolution creating a new feature map. This representation significantly reduces the number of parameters and computational costs while preserving spectral fidelity [57, 58].

We introduce two versions of our model, S5-DSCR and its lightweight variant S5-DSCR-S. Figure 1 illustrates our proposed model. S5-DSCR incorporates multiple DSC modules with L=5, serving as five recursive layers, and integrates a residual connection to efficiently extract spatial and spectral features from HS images. S5-DSCR-S retains comparable performance with a single DSC module (L=1) making it suitable for resource-constrained environments.

The architecture adopts a cascaded design with a pre-upsampling approach that maintains effective feature extraction. Initially, the LR input image is upsampled using bicubic interpolation to generate an intermediate image that provides a baseline estimation of the final super-resolved (SR) image with the intended spatial dimensions. To extract the complex spatial and spectral features, the intermediate image is then given to the DSC module that processes jointly all the channels. This module restores missing high-frequency details to be added to the bicubic interpolation. Each depthwise convolution in the DSC module employs a 5×5 kernel, generating the intermediate feature maps. Each of these feature maps has the same size as the input image making the model memory efficient. Eventually, the residual connection combines the bicubic interpolation and the DSC output to obtain the final SR image.

C. Training Details

To consider the characteristics of each band, the same model was trained separately for each band. This ensures the model is optimized for each band's unique characteristics, such as the number of channels, spatial resolution, noise ratio, etc. Additionally, this approach allows for better adaptation to variations in spectral resolution and spatial complexity.

We trained both S5-DSCR and S5-DSCR-S. The training was performed on the HS images described below. The data was split into completely separate, non-overlapping training, validation, and testing with 65%, 20%, and 15% respectively, ensuring a balanced representation of spatial and spectral diversity. Each image was divided into overlapping patches of size 64×64 and 256×256 for LR and HR images, respectively. We employ Mean Square Error (MSE) loss to train the models and optimize them using the Adam optimizer with the default PyTorch parameters. The initial learning rate was set to 10^{-3} and reduced by a factor of 0.1 when the loss value did not improve for three epochs.

IV. EXPERIMENTS

A. Dataset

We used Sentinel-5P radiance data as the primary source of images for this study. These data are freely accessible on Copernicus' official website [59] as Level-1B radiance data. The dataset comprises data from 15 orbits covering distinct regions acquired on January 4, 2023, and September 7, 2023. The orbits cover data from diverse geographical regions, including Africa, the Gulf countries, America, Europe, Asia, and New Zealand. These regions were selected to represent various climatic and geographical conditions, ensuring that the dataset captures diverse spatial and spectral characteristics and provides maximum coverage of Earth's surface.

Each orbit contains radiance data of eight distinct bands as indicated in Table I. Band 1 was not included in this study because of its low signal-to-noise ratio [9]. Due to orbital and regional variations, the radiance data exhibits various spatial dimensions. The full radiance image spans a range of 4172 to 3735 along-track (scanlines) and 450 to 215 across-track

TABLE I: Sentinel-5P spectrometers characteristics

Spectrometer	# Channels	Name	Range (nm)
Ultraviolet (UV)	497	UV-1 / Band 1	270-300
		UV-2 / Band 2	300-320
Ultraviolet Visible (UVIS)	497	UVIS-1 / Band 3	320-405
		UVIS-2 / Band 4	405-500
Near-Infrared (NIR)	497	NIR-1 / Band 5	675-725
		NIR-2 / Band 6	725-775
Shortwave-Infrared (SWIR)	480	SWIR-1 / Band 7	2305-2345
		SWIR-2 / Band 8	2345-2385

(ground pixels), depending on the specific region and band. To ensure consistency and computational efficiency, we cropped each radiance image into multiple images of size 512×256 for all bands and 512×215 for SWIR bands. In contrast to [8], which focuses on a region corresponding to the central part of the swath, this cropping allows for the inclusion of more diverse regions. We excluded images exhibiting anomalies or extreme values due to inconsistent radiance measurements or sensor artefacts to further refine the dataset. This was performed using a combination of the interquartile range (IQR) to detect outliers beyond 1.5 times the IQR and 1% percentile clipping [60, 61].

B. Evaluation Metrics

The performance of the model was evaluated using standard metrics that assess image quality and spectral fidelity. These metrics provide an objective evaluation of the spatial and spectral resolution of each band. We used four measures [62–64]: Peak Signal-to-Noise Ratio (PSNR), Spatial Correlation Coefficient (SCC), Structural Similarity Index (SSIM), and Learned Perceptual Image Patch Similarity (LPIPS). PSNR quantifies pixel-level accuracy, and SCC measures spatial feature correlation. Considering the perceptual analysis, SSIM assesses perceptual similarity taking into account luminance, contrast, and structural information while LPIPS evaluates perceptual differences using deep features. These metrics collectively offer a global assessment focusing on both perceptual and structural fidelity.

V. RESULTS

The performance of our proposed models was quantitatively evaluated against the separate test dataset and compared against two methods: cubic interpolation and S5Net [9]. Since S5Net has three different weights, we used the validation dataset to decide the optimal weight for our data. S5-DSCR and S5-DSCR-S have 3.9M/3.6M and 0.25M/0.24M (bands 2-6/7-8) parameters, respectively. In contrast, S5Net has 57k parameters per channel, resulting in 28.5M/27.5M (bands 2-6/7-8) parameters when applied across all channels. Table II depicts the evaluation results.

Bicubic interpolation served as a baseline method showing how difficult the SR would be since it is not a learnable method. The heterogeneity in the performance of the bicubic interpolation confirms that each band need to be processed separately.

TABLE II: Performance metrics for SR models on our S5P testing dataset. Best results are in bold, second-best underlined.

Band	Method	PSNR ↑	SCC ↑	SSIM ↑	LPIPS ↓
BAND 2	BICUBIC	32.13	0.846	0.940	0.107
	S5NET	32.30	0.845	0.934	0.058
	S5-DSCR	**37.16**	0.866	**0.952**	**0.051**
	S5-DSCR-S	35.55	**0.875**	0.905	0.095
BAND 3	BICUBIC	27.10	0.887	0.875	0.194
	S5NET	28.98	0.846	0.906	0.135
	S5-DSCR	**33.80**	0.910	0.915	**0.063**
	S5-DSCR-S	32.74	**0.916**	**0.919**	0.132
BAND 4	BICUBIC	25.43	0.772	0.776	0.231
	S5NET	26.28	**0.8589**	0.7467	0.2944
	S5-DSCR	**26.96**	0.777	0.784	0.222
	S5-DSCR-S	26.83	0.788	**0.796**	**0.220**
BAND 5	BICUBIC	20.21	0.683	0.609	0.387
	S5NET	**26.79**	0.730	**0.745**	**0.269**
	S5-DSCR	24.71	0.726	0.692	0.292
	S5-DSCR-S	24.52	**0.743**	0.712	0.282
BAND 6	BICUBIC	22.62	0.757	0.710	0.273
	S5NET	**27.60**	**0.843**	**0.759**	0.269
	S5-DSCR	25.77	0.764	0.717	0.247
	S5-DSCR-S	24.45	0.775	0.725	**0.227**
BAND 7	BICUBIC	12.54	0.794	0.523	0.320
	S5NET	27.05	0.818	0.854	0.446
	S5-DSCR	**31.94**	**0.962**	**0.910**	**0.0541**
	S5-DSCR-S	30.02	0.867	0.811	0.195
BAND 8	BICUBIC	12.58	0.853	0.431	0.324
	S5NET	31.59	0.820	0.865	0.192
	S5-DSCR	**34.80**	**0.919**	**0.896**	**0.153**
	S5-DSCR-S	29.91	0.913	0.819	0.183

The S5-DSCR model consistently demonstrated superior performance across most of the bands. This is indicated by achieving the highest scores in PSNR, SCC, and SSIM while maintaining the lowest LPIPS values. Hence, S5-DSCR has a strong ability to recover HR images with both spatial and spectral fidelity, particularly its ability to maintain significant perceptual quality. On the other hand, S5-DSCR-S captures spectral correlations better than S5-DSCR while maintaining good perceptual quality that is comparable results to S5-DSCR. Notably, S5Net performed reasonably well overall in most of the metrics except for LPIPS, suggesting it has lower perceptual quality. Although bands 7 and 8 possess a low-resolution nature, S5-DSCR demonstrates significant improvements across all metrics. This showcases that S5-DSCR is robust for challenging bands such as these bands and can address various spatial complexities. On the other hand, the performance on bands 5 and 6 was not superior. We argue that this could be due to the higher variation across channels, especially considering that NIR spectrometers are vulnerable and sensitive to environmental factors such as light and weather conditions [65].

Additionally, we provide qualitative SR results of the proposed S5-DSCR model in Figure 2 (only three bands are displayed due to space constraints). For visualization, we applied Principal Component Analysis (PCA) to identify the three most significant components of the ground truth image.
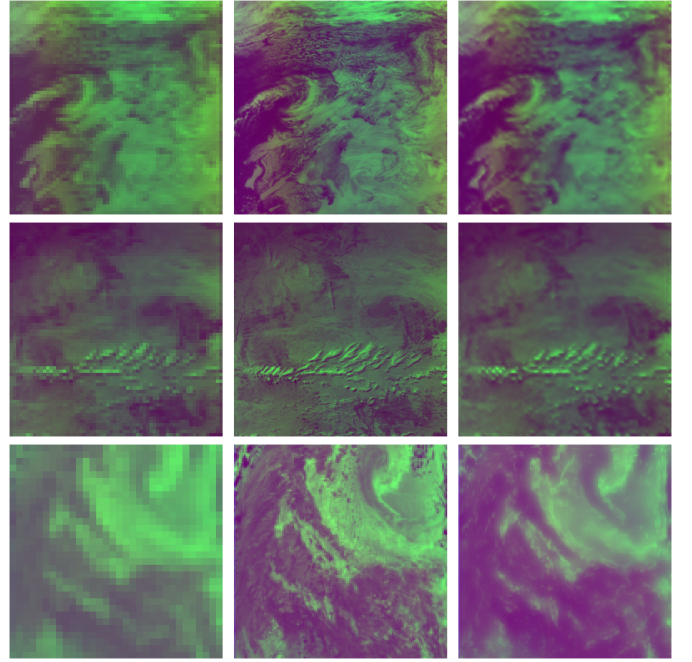

Fig. 2: SR results of S5-DSCR model for, from top to bottom, bands 3, 5 and 7, with, from left to right, LR, ground truth HR, and our result. For visualization, each image is displayed in the first three PCA components of the ground truth.

It can be observed that the reconstructed images contain high-frequency details consistent with the ground truth.

Overall, both quantitative and qualitative analyses demonstrate superior performance. Hence, the proposed models are robust and reliable models for S5P data.

VI. CONCLUSION

In this study, we introduced SR models designed specifically for S5P data, utilizing DSC architecture to capture inter-channel correlations while reducing computational complexity. We coupled DSC with residual connections to enhance feature extraction and stability. As a result, our models achieve superior performance across most of the spectral bands while being lightweight and memory efficient.

A limitation of this work is the reliance on simulated LR data which might not fully provide the degradation patterns. As a future research direction, we plan to explore self-supervised training for the original S5P data as well as testing the model on other satellites datasets to evaluate models' generalization ability.

REFERENCES

[1] Copernicus Program, "Sentinel-5P Mission Overview," 2024, accessed: 2024-12-23. [Online]. Available: https://sentiwiki.copernicus.eu/web/s5p-mission

[2] Copernicus Programm, "Sentinel-5P Applications Overview," 2024, accessed: 2024-12-23. [Online]. Available: https://sentiwiki.copernicus.eu/web/s5p-applications

[3] R. P. Velasco and D. Jarosińska, "Update of the who global air quality guidelines: Systematic reviews–an introduction," *Environment International*, vol. 170, p. 107556, 2022.

[4] Copernicus Program, "Sentinel-5P Products Overview," 2024, accessed: 2024-12-23. [Online]. Available: https://sentiwiki.copernicus.eu/web/s5p-products

[5] European Space Agency (ESA), *Input/Output Data Specification for the TROPOMI L01b Data Processor*, European Space Agency (ESA), 2018, accessed: 2024-12-21. [Online]. Available: https://sentinels.copernicus.eu/documents/247904/3119978/Sentinel-5P-Level-01B-input-output-data-specification

[6] A. Orych, "Review of methods for determining the spatial resolution of uav sensors," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 40, pp. 391–395, 2015.

[7] G. Camps-Valls, D. Tuia, L. Gómez-Chova, S. Jiménez, and J. Malo, "Remote sensing image processing," 2011.

[8] A. Carbone, R. Restaino, G. Vivone, and J. Chanussot, "Model-based super-resolution for sentinel-5p data," *IEEE Transactions on Geoscience and Remote Sensing*, 2024.

[9] A. Carbone, R. Restaino, and G. Vivone, "Efficient hyperspectral super-resolution of sentinel-5p data via dynamic multidirectional cascade fine-tuning," *IEEE Geoscience and Remote Sensing Letters*, 2024.

[10] H. Su, Y. Li, Y. Xu, X. Fu, and S. Liu, "A review of deep-learning-based super-resolution: From methods to applications," *Pattern Recognition*, p. 110935, 2024.

[11] W. Yang, X. Zhang, Y. Tian, W. Wang, J.-H. Xue, and Q. Liao, "Deep learning for single image super-resolution: A brief review," *IEEE Transactions on Multimedia*, vol. 21, no. 12, pp. 3106–3121, 2019.

[12] S. Ye, S. Zhao, Y. Hu, and C. Xie, "Single-image super-resolution challenges: a brief review," *Electronics*, vol. 12, no. 13, p. 2975, 2023.

[13] S. M. A. Bashir, Y. Wang, M. Khan, and Y. Niu, "A comprehensive review of deep learning-based single image super-resolution," *PeerJ Computer Science*, vol. 7, p. e621, 2021.

[14] N. Aburaed, M. Q. Alkhatib, S. Marshall, J. Zabalza, and H. Al Ahmad, "A review of spatial enhancement of hyperspectral remote sensing imaging techniques," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 16, pp. 2275–2300, 2023.

[15] R. Keys, "Cubic convolution interpolation for digital image processing," *IEEE transactions on acoustics, speech, and signal processing*, vol. 29, no. 6, pp. 1153–1160, 1981.

[16] B. Aiazzi, S. Baronti, M. Selva, and L. Alparone, "Bi-cubic interpolation for shift-free pan-sharpening," *ISPRS journal of photogrammetry and remote sensing*, vol. 86, pp. 65–76, 2013.

[17] X. Li and M. T. Orchard, "New edge-directed interpolation," *IEEE transactions on image processing*, vol. 10, no. 10, pp. 1521–1527, 2001.

[18] K.-W. Hung and W.-C. Siu, "Robust soft-decision interpolation using weighted least squares," *IEEE Transactions on Image Processing*, vol. 21, no. 3, pp. 1061–1069, 2011.

[19] Hung and Siu, "Fast image interpolation using the bilateral filter," *IET Image Processing*, vol. 6, no. 7, pp. 877–890, 2012.

[20] R. R. Schultz and R. L. Stevenson, "Extraction of high-resolution frames from video sequences," *IEEE transactions on image processing*, vol. 5, no. 6, pp. 996–1011, 1996.

[21] M. K. Ng, H. Shen, S. Chaudhuri, and A. C. Yau, "Zoom-based super-resolution reconstruction approach using prior total variation," *Optical Engineering*, vol. 46, no. 12, pp. 127 003–127 003, 2007.

[22] C. Fan, C. Wu, G. Li, and J. Ma, "Projections onto convex sets super-resolution reconstruction based on point spread function estimation of low-resolution remote sensing images," *Sensors*, vol. 17, no. 2, p. 362, 2017.

[23] M. K. Ng, H. Shen, E. Y. Lam, and L. Zhang, "A total variation regularization based super-resolution reconstruction algorithm for digital video," *EURASIP Journal on Advances in Signal Processing*, vol. 2007, pp. 1–16, 2007.

[24] J. Sun, Z. Xu, and H.-Y. Shum, "Image super-resolution using gradient profile prior," in *2008 IEEE conference on computer vision and pattern recognition*. IEEE, 2008, pp. 1–8.

[25] Z. Xiong, Z. Shi, H. Li, L. Wang, D. Liu, and F. Wu, "Hscnn: Cnn-based hyperspectral image recovery from spectrally undersampled projections," in *Proceedings of the IEEE international conference on computer vision workshops*, 2017, pp. 518–525.

[26] Y. B. Can and R. Timofte, "An efficient cnn for spectral reconstruction from rgb images," *arXiv preprint arXiv:1804.04647*, 2018.

[27] X.-H. Han, B. Shi, and Y. Zheng, "Residual hsrcnn: Residual hyper-spectral reconstruction cnn from an rgb image," in *2018 24th International Conference on Pattern Recognition (ICPR)*. IEEE, 2018, pp. 2664–2669.

[28] B. Kaya, Y. B. Can, and R. Timofte, "Towards spectral estimation from a single rgb image in the wild," in *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*. IEEE, 2019, pp. 3546–3555.

[29] R. Hang, Q. Liu, and Z. Li, "Spectral super-resolution network guided by intrinsic properties of hyperspectral imagery," *IEEE Transactions on Image Processing*, vol. 30, pp. 7256–7265, 2021.

[30] D. S. Nathan, K. Uma, D. S. Vinothini, B. S. Bama, and S. Roomi, "Light weight residual dense attention net for spectral reconstruction from rgb images," *arXiv preprint arXiv:2004.06930*, 2020.

[31] J. Li, S. Du, R. Song, C. Wu, Y. Li, and Q. Du, "Hasic-net: Hybrid attentional convolutional neural network with structure information consistency for spectral super-resolution of rgb images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–15, 2022.

[32] U. B. Gewali, S. T. Monteiro, and E. Saber, "Spectral super-resolution with optimized bands," *Remote Sensing*, vol. 11, no. 14, p. 1648, 2019.

[33] X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, Y. Qiao, and C. Change Loy, "Esrgan: Enhanced super-resolution generative adversarial networks," in *Proceedings of the European conference on computer vision (ECCV) workshops*, 2018, pp. 0–0.

[34] M. Haris, G. Shakhnarovich, and N. Ukita, "Deep back-projection networks for super-resolution," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 1664–1673.

[35] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu, "Residual dense network for image super-resolution," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 2472–2481.

[36] T. Tong, G. Li, X. Liu, and Q. Gao, "Image super-resolution using dense skip connections," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 4799–4807.

[37] J. Li, C. Wu, R. Song, Y. Li, and F. Liu, "Adaptive weighted attention network with camera spectral sensitivity prior for spectral reconstruction from rgb images," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2020, pp. 462–463.

[38] J. Li, C. Wu, R. Song, W. Xie, C. Ge, B. Li, and Y. Li, "Hybrid 2-d–3-d deep residual attentional network with structure tensor constraints for spectral super-resolution of rgb images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 3, pp. 2321–2335, 2020.

[39] X. Zheng, W. Chen, and X. Lu, "Spectral super-resolution of multispectral images using spatial–spectral residual attention network," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–14, 2021.

[40] Z. Zhu, H. Liu, J. Hou, H. Zeng, and Q. Zhang, "Semantic-

embedded unsupervised spectral reconstruction from single rgb images in the wild," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 2279–2288.

[41] X. Wang, L. Sun, A. Chehri, and Y. Song, "A review of gan-based super-resolution reconstruction for optical remote sensing images," *Remote Sensing*, vol. 15, no. 20, p. 5062, 2023.

[42] A. Bulat, J. Yang, and G. Tzimiropoulos, "To learn image super-resolution, use a gan to learn how to do image degradation first," in *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 185–200.

[43] M. Zhang and Q. Ling, "Supervised pixel-wise gan for face super-resolution," *IEEE Transactions on Multimedia*, vol. 23, pp. 1938–1950, 2020.

[44] C. Saharia, J. Ho, W. Chan, T. Salimans, D. J. Fleet, and M. Norouzi, "Image super-resolution via iterative refinement," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 4, pp. 4713–4726, 2023.

[45] X. Han, J. Yu, J.-H. Xue, and W. Sun, "Spectral super-resolution for rgb images using class-based bp neural networks," in *2018 Digital Image Computing: Techniques and Applications (DICTA)*. IEEE, 2018, pp. 1–7.

[46] S. Mei, Y. Geng, J. Hou, and Q. Du, "Learning hyperspectral images from rgb images via a coarse-to-fine cnn," *Science China Information Sciences*, vol. 65, pp. 1–14, 2022.

[47] G. Guarino, M. Ciotola, G. Vivone, and G. Scarpa, "Band-wise hyperspectral image pansharpening using cnn model propagation," *IEEE Transactions on Geoscience and Remote Sensing*, 2023.

[48] J. Qu, S. Hou, W. Dong, S. Xiao, Q. Du, and Y. Li, "A dual-branch detail extraction network for hyperspectral pansharpening," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–13, 2021.

[49] D. R. Hidalgo, B. B. Cort'es, and E. C. Bravo, "Dimensionality reduction of hyperspectral images of vegetation and crops based on self-organized maps," *Information Processing in Agriculture*, vol. 8, no. 2, pp. 310–327, 2021.

[50] J. He, Q. Yuan, J. Li, Y. Xiao, D. Liu, H. Shen, and L. Zhang, "Spectral super-resolution meets deep learning: Achievements and challenges," *Information Fusion*, vol. 97, p. 101812, 2023.

[51] L. Sifre, "Rigid-motion scattering for image classification," PhD thesis, Ecole Polytechnique - CMAP, 2014.

[52] K.-W. Hung, Z. Zhang, and J. Jiang, "Real-time image super-resolution using recursive depthwise separable convolution network," *IEEE Access*, vol. 7, pp. 99 804–99 816, 2019.

[53] S. Hussain and B. Lall, "Depth separable-cnn for improved spectral super-resolution," *IEEE Access*, vol. 11, pp. 23 063–23 072, 2023.

[54] W. Sun, R. Ke, Z. Liu, H. Lu, D. Li, F. Yang, and L. Zhang, "Enhanced feature refinement network based on depthwise separable convolution for lightweight image super-resolution," *Symmetry*, vol. 16, no. 11, p. 1406, 2024.

[55] W. Muhammad, S. Aramvith, and T. Onoye, "Multi-scale xception based depthwise separable convolution for single image super-resolution," *Plos one*, vol. 16, no. 8, p. e0249278, 2021.

[56] Z. Jiang, Y. Huang, and L. Hu, "Single image super-resolution: Depthwise separable convolution super-resolution generative adversarial network," *Applied Sciences*, vol. 10, no. 1, p. 375, 2020.

[57] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1251–1258.

[58] I. N. Junejo and N. Ahmed, "Depthwise separable convolutional neural networks for pedestrian attribute recognition," *SN Computer Science*, vol. 2, no. 2, p. 100, 2021.

[59] Copernicus Data Space, "Copernicus Data Space Ecosystem," 2024, accessed: 2024-12-23. [Online]. Available: https://dataspace.copernicus.eu/

[60] K. Merry, J. Ossorgin, and Z. Mekus, "Backbone architectures for space domain awareness," in *Proceedings of the¡ a href= https://www. amostech. com¿ Advanced Maui Optical and Space Surveillance (AMOS) Technologies Conference¡/a*, 2024, p. 16.

[61] K. Nakanishi, A. Kubo, Y. Yasui, and S. Ishii, "Robust off-policy reinforcement learning via soft constrained adversary," *arXiv preprint arXiv:2409.00418*, 2024.

[62] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.

[63] X. An, H. Gong, J. Yin, X. Wang, Y. Pan, X. Zhang, Y. Lu, Y. Yang, Z. Toth, I. Schiessl *et al.*, "Orientation-cue invariant population responses to contrast-modulated and phase-reversed contour stimuli in macaque v1 and v2," *PLoS One*, vol. 9, no. 9, p. e106753, 2014.

[64] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The unreasonable effectiveness of deep features as a perceptual metric," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 586–595.

[65] B. Stenberg, R. A. V. Rossel, A. M. Mouazen, and J. Wetterlind, "Visible and near infrared spectroscopy in soil science," *Advances in agronomy*, vol. 107, pp. 163–215, 2010.