IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium | 978-1-7281-6374-1/20/$31.00 ©2020 IEEE | DOI: 10.1109/IGARSS39084.2020.9324488

DATA-DRIVEN AND MODEL-DRIVEN SPECTRAL SUPERRESOLUTION ALGORITHMS: COMBINATION, ANALYSIS AND APPLICATION FOR CLASSIFICATION

Jiang He[1], *Jie Li*[1], *Member, IEEE, Qiangqiang Yuan*[1], *Member, IEEE,*

[1] School of Geodesy and Geomatics, Wuhan University, Wuhan, China

ABSTRACT

In this paper, five spectral superresolution (SSR) algorithms are compared to verify the availability of SSR results as input data in classification. To enhance the spectral resolution, SSR algorithms are proposed to increase the channel number of multispectral images, which can be divided into model-driven and data-driven methods. To combine the advantage of these two types of algorithms, we proposed an optimization-inspired convolutional neural network (OCNN) by unfolding a traditional variational model. The proposed method combines data-driven training with model-driven optimization together to enhance the spectral resolution of high-resolution (HR) multispectral images (MSIs) to obtain HR hyperspectral images (HSIs). Experiments in both SSR and classification are made to show the proposed method is of efficiency and superiority.

Index Terms— Spectral superresolution, Classification, Spectral information.

1. INTRODUCTION

Hyperspectral images (HSIs) are famous for the rich spectral information because of the finer spectral resolution with tens or hundreds of spectral channels and have attracted increasing attention in many researches, such as object recognition [1], spectral unmixing [2], and classification[3], among others.

But the fine spectral resolution of HSIs is at the cost of spatial resolution because of the energy separation of the hyperspectral sensors, which makes HSIs unavailable for those applications where the high spatial resolution is required. To obtain HR HSIs, many researchers have proposed several methods including hyperspectral superresolution, fusion-based methods, and SSR. SSR means acquiring HR HSIs through spectral enhancement of HR MSIs. The existing SSR algorithms can be divided into two groups, model-driven methods and data-driven methods.

For the model-driven algorithms, Arad and Ben-Shahar [4] proposed to computed the dictionary representation of each RGB pixel by the Orthogonal Match Pursuit (OMP) algorithm. By introducing a novel shallow method based on A+ of Timofte et al. [5] from super-resolution, Wu et al. [6] substantially improves over Arad's method. These algorithms can enhance the spectral resolution of MSIs but with spectral distortion in some object.

Besides, with the help of data-driven methods, i.e. deep learning, Galliani et al. [7] proposed a DenseUnet with 56 convolutional layers which get a good SSR performance. And also Can et al. [8] proposed a moderately deep CNN model with residual blocks to reach better SSR effect. Those methods can certainly enhance the spectral resolution but show a not good spatial fidelity.

To improve the spectral resolution as well as maintain good spatial details, combining the model-driven and data-driven algorithms, this paper presents an optimization-inspired CNN by unfolding a variational model, which shows good performance in spectral enhancement. Unlike the methods alternately running CNN and total variational model, the proposed OCNN is with a total end-to-end manner, which can be trained as data-driven models.

To make better use of the image data, classification has attracted many researchers. And to increase the information to help classification, many researchers select HSIs as input data at the cost of spatial resolution. With the help of SSR, we can enhance the spectral channels of MSIs. So this paper also discussed whether the SSR algorithms can help classification very well.

The remaining paper is organized as follows. Section 2 shows two model-driven methods, two data-driven SSR methods, and the proposed OCNN. Experimental results are shown in section 3, and both SSR quantitative evaluation and classification results are displayed. Conclusions are given in the last section.

2. SPECTRAL SUPERRESOLUTION METHODS

2.1. Model Formulation

To present the relationship between HSIs and MSIs clearly, a spectral degradation model is proposed. Let $\boldsymbol{X}_h \in R^{W\times H\times C}$ means the observed hyperspectral image, where C is the number of the spectral channels and W, H is the width and height, respectively. and $\boldsymbol{X}_m \in R^{W\times H\times c}$ presents the observed multispectral image, where $c < C$ is the number of the multispectral bands. And the spectral degradation can be modeled with a transform matrix $\boldsymbol{\Phi} \in R^{c\times C}$ as follow

$$\boldsymbol{X}_m = \boldsymbol{\Phi}\boldsymbol{X}_h \tag{1}$$

2.2. Arad

Employing sparse dictionary learning, Arad et al. proposed a sparse recovery method. In this method, the multispectral and hyperspectral images are regarded as a dictionary $\boldsymbol{D_m}$ or $\boldsymbol{D}_h$ multiplied with a weight matrix $\boldsymbol{w}$ as

$$\boldsymbol{X}_m = \boldsymbol{D_m w} \quad (2)$$

$$\boldsymbol{X}_h = \boldsymbol{D}_h \boldsymbol{w} \quad (3)$$

And Arad proposed that the multispectral dictionary $\boldsymbol{D}_m$ can be transformed from the hyperspectral dictionary $\boldsymbol{D}_h$ with a matrix $\boldsymbol{R}$ as

$$\boldsymbol{D_m} = \boldsymbol{R}\boldsymbol{D}_h \quad (4)$$

So Arad firstly learns a rich hyperspectral dictionary $\boldsymbol{D}_h$ from a large number of samples, and use the transformation matrix $\boldsymbol{R}$ to generate $\boldsymbol{D_m}$ by (4). Then employing Orthogonal Match Pursuit (OMP) algorithm in the MSIs, the weight matrix $\boldsymbol{w}$ can be computed and used to reconstruct HSIs by (3).

2.3. A+

For the A+ proposed by Wu et al., they use a pretrained overcomplete dictionary, not as proposed by Arad for a superposition but as anchor points to perform a nearest neighbor search. A projection matrix is computed for each anchor, using a collection of neighboring samples from the complete training set to approximate a local mapping from RGB to hyperspectral values as shown in Fig. 1.

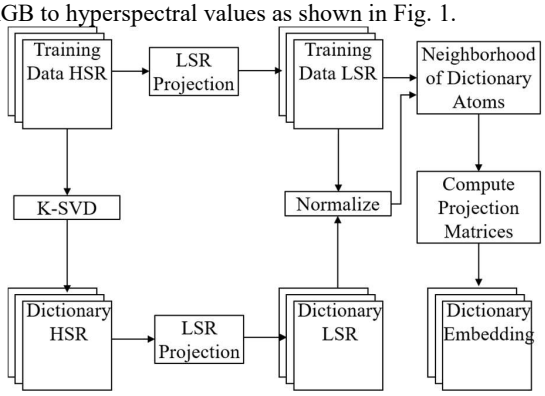

Fig. 1 the processing of the A+

2.4. DenseUnet

Galliani et al. proposed a very deep Unet with Densenet block, which is called DenseUnet in this paper. The network is with a total of 56 layers and each convolution has size 3× 3. The image gets down-scaled 5 times by a factor of 2 with a 1×1 convolution followed by max-pooling. In its own terminology, each Densenet block has a growth rate of 4 with 16 layers, which means 4 convolutional layers per block, each with 16 filters, seeing Fig. 2.

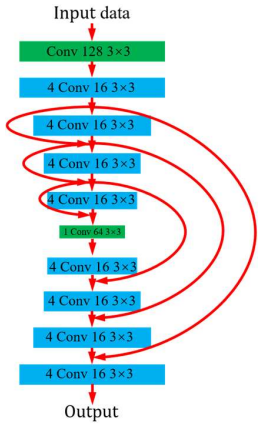

Fig. 2 the architecture of the DenseUnet

2.5. CanNet

Rather than a very deep convolutional neural network (CNN) proposed by Galliani et al., Can et al. proposed a shallow CNN with residual blocks to achieve the spectral enhancement, which is shown in Fig. 3. And we call this CNN as CanNet. It's noted that all the activation functions are PReLU.

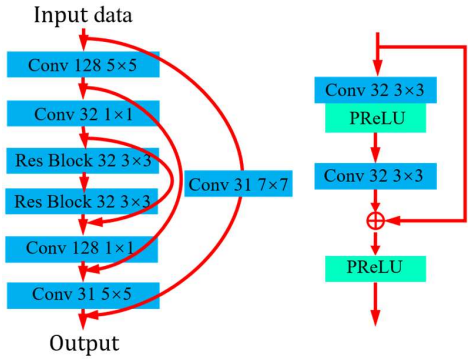

Fig. 3 the architecture of the CanNet

2.6. Proposed OCNN

2.6.1. Networks

With the spectral degradation model, we can easily give the formulation of the SSR problem with some prior knowledge as regularization term as

$$\widehat{\boldsymbol{X}}_h = arg\min_{\boldsymbol{X}_h} \|\boldsymbol{X}_m - \boldsymbol{\Phi}\boldsymbol{X}_h\|^2 + \gamma\mathcal{R}(\boldsymbol{X}_h) \quad (5)$$

where γ is a trade-off parameter, and $\mathcal{R}(\cdot)$ is the regularization term. To better solve this minimization problem, the variable splitting technique can be employed to separate the two terms. In this paper, Equation (5) can be solved by solving two subproblems iteratively as

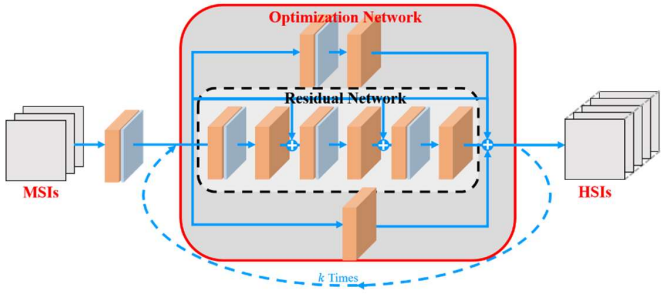

Fig. 4 the processing of the OCNN

$$\widehat{\boldsymbol{X}}_h^{k+1} = arg\min_{\boldsymbol{X}_h} \|\boldsymbol{X}_m - \boldsymbol{\Phi}\boldsymbol{X}_h\|^2 + \mu\|\boldsymbol{G}^k - \boldsymbol{X}_h\|^2 \quad (6)$$

$$\widehat{\boldsymbol{G}}^{k+1} = arg\min_{\boldsymbol{X}_h} \|\boldsymbol{G} - \boldsymbol{X}_h^{k+1}\|^2 + \lambda\mathcal{R}(\boldsymbol{G}) \quad (7)$$

where $\lambda = \frac{\gamma}{\mu}$ is another penalty parameter related to μ and γ. Inspired by [9], we employ a residual network to learn the prior $\boldsymbol{G}^k$ implicitly. And we also employed two different convolutional blocks to update $\widehat{\boldsymbol{X}}_h^k$ for the end-to end manner as shown in Fig. 4.

2.6.2. Multi-level Learning

As we know, CNN is famous for its data-driven training, which is always affected by the loss function. To make sure the optimization stages in the OCNN operate normally, multi-level constraints to stages are employed. and the loss functions in each level are all Mean Absolute Error (MAE) loss function, i.e. L1 loss.

$$Loss = \sum_{j=1}^{K} \alpha_j |\widehat{\boldsymbol{X}}_h^k - \boldsymbol{X}_h| \quad (8)$$

where α_j is the trade-off parameter in the jth stage. $\widehat{\boldsymbol{X}}_h^k$ means the intermediate result reconstructed by the jth stage.

Introducing multi-level loss functions, the proposed OCNN can learn more accurate mapping functions at different optimization stages and achieve better spectral enhancement.

3. EXPERIMENTS AND DISCUSSION

We evaluated five SSR methods using CAVE dataset and remote sensing (RS) datasets. And four image quality metrics are utilized, including correlation coefficient (CC), peak signal-to-noise ratio (PSNR), structural similarity (SSIM), and spectral angle mapper (SAM).

The CAVE dataset is a popular hyperspectral image dataset in hyperspectral image processing, which consists of 32 scenes with spatial size of 512×512. All the HSIs in CAVE cover spectral range from 400nm to 700nm with 10m spectral resolution, which contain 31 bands. Besides, the RGB images covering the same scene as HSI data are available. 3 images from the CAVE dataset are selected as testing images and others are used to train models.

As for the RS datasets, a Chinese hyperspectral dataset from Orbita hyperspectral satellites (OHS) with 10m spatial resolution and 32 channels cover spectral range from 400nm to 1000nm are selected as HSIs. Because of the same spatial resolution between OHS data and Sentinel-2 data, we simulated Sentienl-2 MSIs from OHS HSIs using spectral downsampling with the spectral response functions. In this manner, the errors caused by geographic registration and the inconsistency of imaging time between Sentinel-2 and OHS data can be avoided, which helps us to build the RS SSR dataset. We also selected 3 images as testing data and 750 images as training samples.

TABLE I
NUMERICAL COMPARISON OF FOUR IMAGE QUALITY METRICS BETWEEN RESULTS IN CAVE DATASET AND RS DATASET

Method	CAVE dataset				RS dataset			
	CC	PSNR	SSIM	SAM	CC	PSNR	SSIM	SAM
Arad	0.9444	26.3057	0.8303	20.3662	0.8319	23.6114	0.5979	12.2771
A+	0.9860	35.2462	0.9387	22.5772	0.9023	25.8905	0.7434	10.7717
DenseUnet	0.9914	35.0648	0.9695	8.5961	0.9431	26.0308	0.8914	9.8650
CanNet	0.9915	34.7093	0.9642	9.9054	0.9586	26.6592	0.8895	9.8871
OCNN	**0.9941**	**37.3844**	**0.9818**	**8.3442**	**0.9623**	**28.5653**	**0.9210**	**9.6879**

The SSR performance of five methods are shown in TABLE I. The proposed OCNN get the best SSR performance in both CAVE dataset and RS dataset. And CanNet get good performance in spatial fidelity but with a worse spectral maintaining than DenseUnet. A+ shows a surprising improvement in PSNR but with high spectral distortion.

Then, we verified the classification results by support vector machine (SVM) only in RS dataset. Due to lack of ground truth of classification in the proposed RS dataset, we

simulated a HS-MS image couple from the Indian Pines dataset with the same spectral range as RS dataset. The quantitative results are shown in Table II.

TABLE II
THE OVERALL ACCURACY AND KAPPA COEFFICIENT OF CLASSIFICATIONS

Method	Kappa	OA(%)
Original MSI	0.5760	64.0103
Arad	0.5343	60.6296
A+	0.6074	66.5320
DenseUnet	0.5960	65.4602
CanNet	0.6006	65.6350
OCNN	**0.6113**	**67.2160**

Comparing with original MSI, except Arad, the images enhanced by the SSR methods get better classification results, which shows that the SSR methods can provide more spectral information for original multispectral images to help classification.

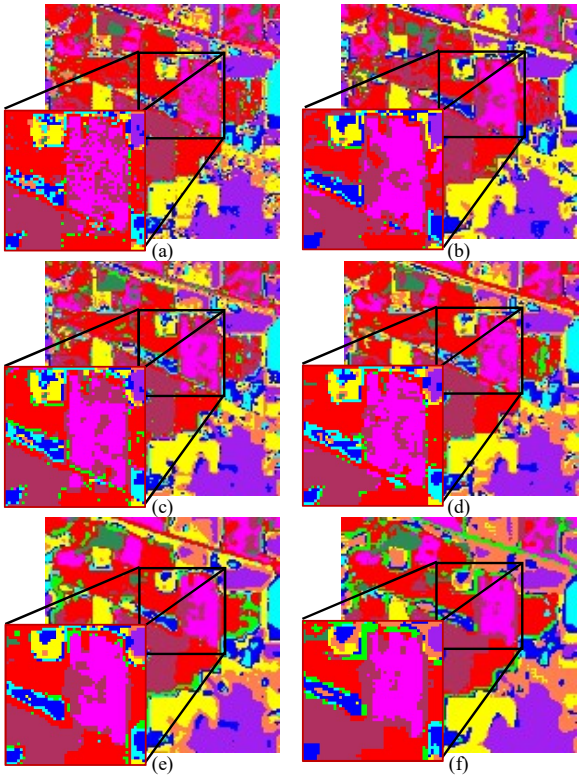

Fig. 5 the classification results of different input. (a) original MSI with 4 bands. (b) the SSR result of Arad. (c) the SSR result of Arad. (d) the SSR result of DenseUnet. (e) the SSR result of CanNet. (f) the SSR result of OCNN.

The classification results are shown in Fig. 5. The SSR algorithms can improve more information to help SVM better group the same object as shown in the area zoomed in. And the classification results using the SRR data by the proposed OCNN shows better integrality.

4. CONCLUTION

To enhance spectral resolution of MSI, this paper proposed an OCNN from a physical model with deep learning manner, and compared with four famous SSR algorithms based on dictionary learning and deep learning. Verified in natural image dataset and RS image dataset, the proposed OCNN get the best SSR effect. Besides, a series of classification experiments are also presented to show the SSR algorithms can truly help classifier better distinguish different objects. And compared with four mentioned SSR methods, the proposed OCNN improves the accuracy of classification the most, which proves the effectiveness and superiority of the proposed method.

5. ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China under Grant 41701400 and 61671334.

6. REFERENCES

[1] J. Jung and M. M. Crawford, “Extraction of features from LIDAR waveform data for characterizing forest structure,” *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 3, pp. 492–496, May 2012.

[2] J. M. Bioucas-Dias *et al.*, “Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches,” *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 354–379, Apr. 2012.

[3] C.-I Chang, *Hyperspectral Imaging: Techniques for Spectral Detection and Classification*, vol. 1. New York, NY, USA: Plenum, 2003.

[4] B. Arad, and O. Ben-Shahar, “Sparse Recovery of Hyperspectral Signal from Natural RGB Images,” In *the Proceedings of the European Conference on Computer Vision (ECCV)*, Amsterdam, The Netherlands, October 2016.

[5] R. Timofte, V. De Smet, and L. Van Gool, “A+: Adjusted anchored neighborhood regression for fast super-resolution,” in *Asian Conference on Computer Vision (ACCV)*, Springer, 2014, pp. 111–126.

[6] J. Wu, J. Aeschbacher, and R. Timofte, "In Defense of Shallow Learned Spectral Reconstruction from RGB Images," *2017 IEEE International Conference on Computer Vision Workshops (ICCVW)*, Venice, 2017, pp. 471-479.

[7] S. Galliani, C. Lanaras, D. Marmanis, E. Baltsavias, and K. Schindler, “Learned Spectral Super-Resolution,” *arXiv*, 1703. 09470 (2017).

[8] Y. Can, and R. Timofte, “An efficient CNN for spectral reconstruction from RGB images,” *arXiv*, 1804. 04647 (2018).

[9] W. Dong, P. Wang, W. Yin, G. Shi, F. Wu and X. Lu, "Denoising Prior Driven Deep Neural Network for Image Restoration," *in IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 10, pp. 2305-2318, 1 Oct. 2019.