A comparative study of loss functions for hyperspectral SISR.
In 30th European Signal Processing Conference, EUSIPCO 2022 - Proceedings (pp. 484-487).
(European Signal Processing Conference; Vol. 2022-August)
. European Signal Processing Conference, EUSIPCO.

A Comparative Study of Loss Functions for Hyperspectral SISR

Nour Aburaed*, Mohammed Q. Alkhatib†, Stephen Marshall‡, Jaime Zabalza§, Hussain Al Ahmad¶
*†¶ College of Engineering and IT, University of Dubai, UAE
*‡§ Department of Electronic and Electrical Engineering, University of Strathclyde, UK
Email: *nour.aburaed,†mqalkhatib@ieee.org
*nour.aburaed, ‡j.zabalza, §stephen.marshall@strath.ac.uk, ¶halahmad@ud.ac.ae

***Abstract*—The spatial enhancement of Hyperspectral Imagery (HSI) is a popular research area among the community of image processing in general and remote sensing in particular. HSI contribute to a wide variety of industrial applications, such as Land Cover Land Use. The characterstic that distinguishes HSI from other type of images is the ability to uniquely describe objects with spectral signatures. This can be achieved due to the sensor's ability to capture reflectance in narrowly spaced wavelength bands, which yields an HSI cube with hundreds of bands. However, this ability compromises the spatial resolution of HSI, which must be improved for practicality and usability. There are several studies in the literature related to HSI Super Resolution (HSI-SR), especially using Convolutional Neural Networks (CNNs). Nonetheless, the investigation of the most suitable loss functions to train these networks is necessary and remains as an area to investigate. This paper conducts a comparative study of the most widely used loss functions and their effect on one of the state-of-the-art HSI-SR CNNs, mainly 3D-SRCNN. The paper also proposes a hybrid loss function based on the comparative results, and proves its superiority against other loss functions in terms of Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index Measurement (SSIM), and Spectral Angle Mapper (SAM).**

***Index Terms*—Hyperspectral, Super Resolution, CNN, 3D SR-CNN, Loss Function**

I. Introduction

Remote sensing applications are nowadays being integrated with various industrial fields, such as agriculture and urban planning. With the abundance and recent revolution of Image Processing (IP) and Artificial Intelligence (AI) tools, this integration has been occurring more rapidly and seamlessly. This is due to the fact that IP tasks, such as object detection and semantic segmentation, are now being achieved with minimal human intervention. Nonetheless, the effectiveness of IP and AI tool relies on the quality of the data. Remote sensing imagery can be categorized depending on its resolution into Hyperspectral Imagery (HSI), and Multispectral Imagery (MSI). This categorization was imposed by the natural trade-off in sensors, which limits their capability to capture either high spectral resolution, yielding HSI, or high spatial resolution, yielding MSI. Therefore, enhancing HSI spatially and MSI spectrally are topics of interest in the field of remote sensing. The former is the main theme of this paper.

Spatial enhancement of HSI can be categorized into Fusion and Single Image Super Resolution (SISR). Fusion methods require the availability of auxiliary information and assumes that MSI and HSI of the same scene are captured with precise georeferencing accuracy, which renders it impractical in scenarios were such requirement cannot be fulfilled. Therefore, this paper focuses on SISR, which does not require supplementary information. Bicubic interpolation [1] was one of the early methods utilized for image SISR generally, followed by enhanced versions of it and other interpolation methods. Then, machine learning methodologies were popularly used along with sparse representation and dictionary learning approaches with prior assumptions [2]. Convolutional Neural Networks (CNNs) revolutionized the field of IP generally and SISR particularly. Several state-of-the-art networks were developed that prevail over interpolation methods and traditional machine learning approaches [3]–[8]. It has been established by various studies that 2D CNNs achieve acceptable performance on MSI, but fall short when applied to HSI [9]. This is due to their inability to capture spectral signature; a powerful trait of HSI that eases the recognition and classification of objects. It is important to enhance HSI while preserving their spectral fidelity, which can be achieved by 3D CNNs due to their ability to successfully capture spectral context. There are many studies in the literature that utilize 3D CNNs for this purpose [9]–[14]. However, the studies do not investigate the effect of loss functions on the performance of the CNN. Deep Neural Networks (DCNNs), have known problems in training, such as reaching false (local) minimum and overfitting [15]. Thus, the choice of loss function is crucial for reaching optimum results. This paper deals with the analysis and comparisons of the most commonly used and state-of-the-art loss functions in SISR to gauge the effectiveness of each one. The ultimate goal of the paper is to establish the most suitable loss function for HSI-SR, and to lay the groundwork for the future work of proposing a new loss function. The studied loss functions are Mean Squared Error (MSE), Mean Absolute Error (MAE), Mean Squared Logarithmic Error (MSLE), Log Hyperbolic Cosine (LHC), Huber, Charbonier, Cosine Similarity (CS), and a proposed hybrid loss function. The network chosen to perform this study is 3D-SRCNN, which was previously proposed in [14] and has proven efficiency against other state-of-the-art approaches. The datasets used in this study are Pavia University [16] and Botswana [16], and the evaluation

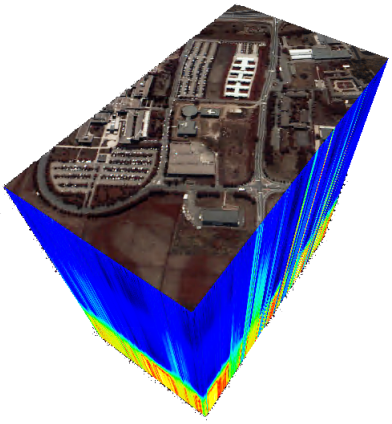

Fig. 1. Pavia University HSI cube.

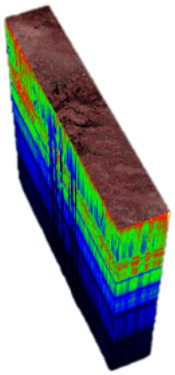

Fig. 2. Botswana HSI cube.

is performed based on Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM) [17], and Spectral Angle Mapper (SAM), which are the most frequently used evaluation metrics for HSI. PSNR and SSIM measure the spatial similarity between the Ground Truth (GT) image and the estimated (enhanced) one, while SAM measures the similarity between the spectra of the GT image and that of the estimated one.

The rest of the paper is organized as follows: Section II describes the loss functions to be studied and the experimental setup, Section III illustrates and analyzes the results obtained and draws conclusions based on that, and finally, Section IV summarizes the paper and states the future direction of this work.

II. Methodology

A. Datasets

The first dataset studied in this paper is Pavia University. It was captured by Reflective Optics System Imaging Spectrometer (ROSIS). The size of the image is 610×340 with 103 bands after removing the corrupted ones. The spatial resolution is 3.7m with spectral range of 430-960nm.

The second dataset studied in this paper is Botswana dataset, which was captured by NASA EO-1 satellite using Hyperion sensor. The size of the image is 1476×256 with 145 bands. The spatial resolution is 30m with a spectral range of 400-2500nm.

There are pre-processing steps performed on the datasets before experimentation. Since training neural networks requires a large amount of data, the datasets are divided into patches of 64×64 to increase the number of images for training and testing. Additionally, each patch is degraded using Gaussian blur and down-scaled by the required scale factor using nearest neighbor [18], [19]. In this study, all the experiments are performed on scale factor of 2. The resulting patch is considered as the Low Resolution HSI (LR-HSI), which will be enhanced and compared to the original GT HSI. The details about both datasets, including the number of training and testing patches, are summarized in Table I.

TABLE I
Description of the datasets used in this study.

Dataset	Sensor	Spatial Resolution	Bands	Size	Training	Testing
Pavia University	ROSIS	3.7m	103	610×340	41	4
Botswana	Hyperion	30m	145	1476×256	81	11

B. Loss Functions & Experimental Setup

For a GT HSI denoted $Y \in \mathbb{Z}^{M \times N \times B}$ of height M, width N, and a number of bands B, an estimated High Resolution HSI (HR-HSI) of the same size is denoted $\hat{Y} \in \mathbb{Z}^{M \times N \times B}$. A loss function's main objective is to calculate the residual error R between Y and $\hat{Y}$ for each band in order to mitigate this error during the training process of the 3D-CNN, which is in this case 3D-SRCNN. For more details about 3D-SRCNN, the reader is referred to this paper [14]. For HSI-SR, loss functions can be categorized into spatial and spectral. In this context, spatial loss refers to the pixel-wise comparison between the same band of GT HSI and the estimated HSI, while spectral loss refers to the comparison of the spectral signature vectors in a space with dimensionality equal to the number of bands.

1) Spatial Loss: The most straight forward and the most widely used loss function is MSE, which measures the sum of squared differences between every pixel at location (i, j) of band k in $Y_k(i, j)$ and $\hat{Y}_k(i, j)$. MSE is expressed by Equation 1. Measuring the absolute differences instead of square differences is the main distinction between MSE and MAE, also known as L1, which is expressed in Equation 2. MSLE is a variation of MSE that calculates the log error, as seen in Equation 3. MSLE is concerned with measuring the ratio between Y and $\hat{Y}$ and, unlike MSE, does not get affected by small differences.

$$R = \sum_{k=1}^{B} Y_k(i,j) - \hat{Y}_k(i,j)$$
$$L_{MSE} = \frac{1}{M \times N \times B} \sum_{i=1}^{M} \sum_{j=1}^{N} R(i,j)^2 \quad (1)$$

$$L_{MAE} = \frac{1}{M \times N \times B} \sum_{i=1}^{M} \sum_{j=1}^{N} |R(i,j)| \quad (2)$$

$$L_{MSLE} = \frac{1}{M \times N \times B} \sum_{k=1}^{B} \sum_{i=1}^{M} \sum_{j=1}^{N} \left(log\left(\frac{Y(i,j)+1}{\hat{Y}(i,j)+1} \right) \right)^2 \quad (3)$$

Huber [20] is a piece-wise loss function that was devised to be resilient to outliers. It is a hybrid combination of MSE and MAE, such that if the residual is less than or equal to a certain threshold δ, the loss is expressed by MSE, otherwise it is expressed by MAE, as seen in Equation 4. The challenge is to choose an optimal δ value, as it is dependent on the dataset inliers. The experiments in Section III will display different δ values for each dataset used.

$$L_{Huber} = \frac{1}{M \times N \times B} \sum_{i=1}^{M} \sum_{j=1}^{N} \begin{cases} \frac{1}{2} R(i,j)^2, & \text{if } |R(i,j)| \leq \delta \\ \delta \left(|R(i,j)| - \frac{1}{2}\delta \right), & \text{otherwise} \end{cases} \quad (4)$$

LHC loss was first proposed in [21], where the authors utilized this function to improve the performance of Variational Auto-Encoders. Even though LHC is not a piece-wise function, its performance is close to that of Huber. However, it lacks the adaptability of Huber, as δ is fixed in LHC. Additionally, LHC differs from MSE in the sense that it does not get affected by the occasional large errors.

$$L_{LHC} = \frac{1}{M \times N \times B} \sum_{i=1}^{M} \sum_{j=1}^{N} log(cosh(R(i,j))) \quad (5)$$

Charbonier loss function, originally proposed in [22], is a variant of MAE and it was adapted for HSI-SR in [23]. The authors argue that Charbonier loss is more resilient to outliers and provides more performance improvement over MAE. Charbonier loss is expressed in Equation 6. The standard value of ϵ is 10^-3.

$$L_{Charbonier} = \sum_{i=1}^{M} \sum_{j=1}^{N} \sqrt{R(i,j)^2 + \epsilon^2} \quad (6)$$

2) Spectral Loss: All the aforementioned loss functions measure the errors spatially. However, as explained in Section I, HSI differ from other types of images due to their high spectral resolution, which must be preserved while enhancing HSI. One loss function that takes spectral resolution into consideration is Cosine Similarity (CS), which is expressed in Equation 7. CS measures the similarity between the GT vector $\mathbf{y}$ of pixel values at position (i,j) and the estimated vector $\hat{\mathbf{y}}$ of pixel values at the same position across all bands.

$$L_{CS} = -\frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} \frac{\sum_{k=1}^{B} \mathbf{y_k}(\mathbf{i},\mathbf{j}) \hat{\mathbf{y}}_\mathbf{k}(\mathbf{i},\mathbf{j})}{\sqrt{\sum_{k=1}^{B} \mathbf{y}(\mathbf{i},\mathbf{j})^2} \sqrt{\sum_{k=1}^{B} \hat{\mathbf{y}}(\mathbf{i},\mathbf{j})^2}} \quad (7)$$

3) Hybrid Loss: Intuitively, one would argue that a hybrid loss function that combines spatial and spectral losses is the ideal solution to provide the best of both worlds. A hybrid loss function is proposed based on the spatial loss that yields the best performance combined with CS loss. The hybrid loss function is expressed in Equation 8:

$$L_{Hybrid} = \alpha L_s - \beta L_{CS}, \quad (8)$$

where $\alpha, \beta \in \mathbb{R}$. For this study, α and β will be set empirically depending on the dataset used. L_s will be chosen depending on the spatial loss function that scores the best results, which will be revealed in Section III.

TABLE II
PAVIA UNIVERSITY RESULTS SHOW THAT THE HYBRID FUNCTION ACHIEVES BETTER PSNR AND SSIM THAN ALL LOSS FUNCTIONS, AND BETTER SAM THAN SPATIAL LOSS FUNCTIONS.

Loss Function	PSNR (dB)	SSIM	SAM (°)
MSE	32.56	0.919	4.74
MAE	32.60	0.922	4.55
MSLE	31.45	0.917	5.75
Huber ($\delta = 1$)	32.64	0.921	4.69
LHC	32.51	0.920	4.85
Charbonier	32.66	0.921	4.50
CS	10.62	0.588	**4.34**
Hybrid ($\alpha = 1.5, \beta = 0.09$)	**32.73**	**0.923**	**4.45**

III. RESULTS AND ANALYSIS

After pre-processing the datasets as described in Section II-A, each dataset is then used for training and testing 3D-SRCNN with each described loss function from Section II-B. To ensure fairness of comparison, the training parameters are the same for every loss function trial. The training is performed of 200 epochs with Adam as an optimization function. Additionally, the same training and testing data are used for every loss function.

A. Pavia University

The results of Pavia University dataset are listed in Table II. From the spatial loss functions, the best results for Huber loss are obtained by setting $\delta = 0.5$, which was chosen with systematic testing. Charbonier scored the highest PSNR, and MAE scored the highest SSIM. The lowest scores of PSNR and SSIM were obtained by using MSLE. Furthermore, MSLE showed the most spectral distortions, as it resulted in the highest SAM value. The least spectral distortions were obtained from Charbonier, as it scored $4.34°$. The loss function with the least spectral distortions is CS, although it shows high spatial distortion as a trade-off, which makes it impractical as a loss function by itself. Based on these results, Charbonier is chosen as the best loss function. For the hybrid loss function, with reference to Equation 8, $L_s = Charbonier$ with $\alpha = 1.5$ and $\beta = 0.09$. Those parameters were chosen empirically based on the best observed results. The hybrid loss function yielded higher PSNR and SSIM than all loss functions. Even though its SAM is higher than that of CS, it is still lower than all the spatial loss functions, which means that the hybrid loss

TABLE III
BOTSWANA RESULTS SHOW THAT THE HYBRID FUNCTION ACHIEVES BETTER PSNR AND SSIM THAN ALL LOSS FUNCTIONS, AND BETTER SAM THAN SPATIAL LOSS FUNCTIONS.

Loss Function	PSNR (dB)	SSIM	SAM (°)
MSE	35.09	0.928	2.37
MAE	35.18	0.929	2.32
MSLE	35.07	0.928	2.35
Huber ($\delta = .5$)	35.14	0.928	2.33
LHC	35.05	0.928	2.37
Charbonier	35.25	0.930	2.28
CS	8.87	0.649	**2.20**
Hybrid ($\alpha = 1.3, \beta = 0.08$)	**35.36**	**0.931**	**2.24**

function successfully combines the best of both spatial and spectral loss functions.

B. Botswana

The results of Botswana dataset are listed in Table III. The best results for Huber loss function are obtained with parameters $\delta = 0.5$, similar to Pavia University. It is observed from the table that Charbonier loss function scored the highest PSNR and SSIM, in addition to the lowest SAM among the spatial loss functions. Similar to the case of Pavia University, CS showed the least spectral distortions while compromising spatial quality, which leads to the same conclusion of its impracticality as a loss function by itself. The hybrid loss function, set to $L_s = Charbonier$ with parameter values $\alpha = 1.3$ and $\beta = 0.08$, demonstrated the best PSNR and SSIM among all functions, in addition to the lowest SAM among the spatial loss functions.

IV. CONCLUSION

In this paper, a comparative analysis between some of the most widely used spatial and spectral loss functions have been discussed and compared in the context of HSI-SR. Based on this comparison, a hybrid loss function that combines the best of the spatial and spectral loss is proposed. The performance of the hybrid function is compared to the formerly analyzed loss functions. Quantitative evaluation in terms of PSNR, SSIM, and SAM prove that the hybrid loss function is more effective than the other loss functions. The future direction of this research involves devising a practical way to set the parameters α and β of the hybrid loss function rather than setting them empirically, as they are different depending on the dataset and finding the right parameter value can be a tedious process.

REFERENCES

[1] 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.

[2] A. Villa, J. Chanussot, J. A. Benediktsson, M. Ulfarsson, and C. Jutten, "Super-resolution: an efficient method to improve spatial resolution of hyperspectral images," in *2010 IEEE International Geoscience and Remote Sensing Symposium*, 2010, pp. 2003–2006.

[3] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *CoRR*, vol. abs/1501.00092, 2015.

[4] J. Kim, J. K. Lee, and K. M. Lee, "Accurate image super-resolution using very deep convolutional networks," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 1646–1654.

[5] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. P. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, "Photo-realistic single image super-resolution using a generative adversarial network," *CoRR*, vol. abs/1609.04802, 2016.

[6] J. Yu, Y. Fan, J. Yang, N. Xu, Z. Wang, X. Wang, and T. S. Huang, "Wide activation for efficient and accurate image super-resolution," *CoRR*, vol. abs/1808.08718, 2018.

[7] Weimin Chen, Yuqing Ma, Xianglong Liu, and Yi Yuan, "Hierarchical generative adversarial networks for single image super-resolution," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, January 2021, pp. 355–364.

[8] J. Andrew, T.S.R. Mhatesh, Robin D. Sebastin, K. Martin Sagayam, Jennifer Eunice, Marc Pomplun, and Hien Dang, "Super-resolution reconstruction of brain magnetic resonance images via lightweight autoencoder," *Informatics in Medicine Unlocked*, vol. 26, pp. 100713, 2021.

[9] S. Mei, X. Yuan, J. Ji, S. Wan, J. Hou, and Q. Du, "Hyperspectral image super-resolution via convolutional neural network," in *IEEE International Conference on Image Processing (ICIP)*, 2017, pp. 4297–4301.

[10] S. Mei, X. Yuan, J. Ji, Y. Zhang, S. Wan, and Q. Du, "Hyperspectral image spatial super-resolution via 3d full convolutional neural network," *Remote Sensing*, vol. 9, no. 11, 2017.

[11] W. Liu and J. Lee, "An efficient residual learning neural network for hyperspectral image superresolution," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 4, pp. 1240–1253, 2019.

[12] X. Dou, C. Li, Q. Shi, and M. Liu, "Super-resolution for hyperspectral remote sensing images based on the 3d attention-srgan network," *Remote Sensing*, vol. 12, no. 7, 2020.

[13] Y. Li, L. Zhang, C. Dingl, W. Wei, and Y. Zhang, "Single hyperspectral image super-resolution with grouped deep recursive residual network," in *2018 IEEE Fourth International Conference on Multimedia Big Data (BigMM)*, 2018, pp. 1–4.

[14] N. Aburaed, M. Q. Alkhatib, S. Marshall, J. Zabalza, and H. Al Ahmad, "3d expansion of srcnn for spatial enhancement of hyperspectral remote sensing images," in *2021 4th International Conference on Signal Processing and Information Security (ICSPIS)*, 2021, pp. 9–12.

[15] K. Kawaguchi, J. Huang, and L. P. Kaelbling, "Effect of depth and width on local minima in deep learning," *CoRR*, vol. abs/1811.08150, 2018.

[16] M Graña, MA Veganzons, and B Ayerdi, "Hyperspectral remote sensing scenes," *Grupo de Inteligencia Computacional (GIC)*.

[17] 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.

[18] Naoto Yokoya, Claas Grohnfeldt, and Jocelyn Chanussot, "Hyperspectral and multispectral data fusion: A comparative review of the recent literature," *IEEE Geoscience and Remote Sensing Magazine*, vol. 5, no. 2, pp. 29–56, 2017.

[19] Naoto Yokoya, Takehisa Yairi, and Akira Iwasaki, "Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 2, pp. 528–537, 2012.

[20] Jonathan T. Barron, "A more general robust loss function," *CoRR*, vol. abs/1701.03077, 2017.

[21] P. Chen, G. Chen, and S. Zhang, "Log hyperbolic cosine loss improves variational auto-encoder," 2019.

[22] C. Dong, C. C. Loy, K. He, and X. Tang, "Learning a deep convolutional network for image super-resolution," in *Computer Vision – ECCV 2014*, D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds., Cham, 2014, pp. 184–199, Springer International Publishing.

[23] L. Wang, T. Bi, and Y. Shi, "A frequency-separated 3d-cnn for hyperspectral image super-resolution," *IEEE Access*, vol. 8, pp. 86367–86379, 2020.