We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

5,500
Open access books available

136,000
International authors and editors

170M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Chapter

Application of Deep Learning Approaches for Enhancing Mastcam Images

Ying Qu, Hairong Qi and Chiman Kwan

Abstract

There are two mast cameras (Mastcam) onboard the Mars rover Curiosity. Both Mastcams are multispectral imagers with nine bands in each. The right Mastcam has three times higher resolution than the left. In this chapter, we apply some recently developed deep neural network models to enhance the left Mastcam images with help from the right Mastcam images. Actual Mastcam images were used to demonstrate the performance of the proposed algorithms.

Keywords: Mastcam, Curiosity rover, image fusion, pansharpening, deep learning, Dirichlet-net, U-net, transition learning

1. Introduction

The Curiosity rover (**Figure 1**) has several instruments that are used to characterize the Mars surface. For example, the Alpha Particle X-Ray Spectrometer (APXS) [1] can analyze rock samples collected from the robotic arm and extract compositions of rocks; the Laser Induced Breakdown Spectroscopy (LIBS) [2] can extract spectral features from the vaporized fumes and deduce the rock compositions at a distance of 7 m; and the Mastcam imagers [3] can perform surface characterization from 1 km away.

The two Mastcam multispectral imagers are separated by 24.2 cm [3]. As shown in **Figure 2**, the left Mastcam (34 mm focal length) has three times the field of view of the right Mastcam (100 mm focal length). In other words, the right imager has three times higher resolution than that of the left. To generate stereo image or construct a 12-band image cube by fusing bands from the multispectral imagers from the left and right Mastcams [4–6], a practical solution is to downsample the resolution of the right images to that of the left images, which would avoid the artifacts caused by Bayer pattern [7] or the JPEG compression loss [8]. Although this approach has practical merits, it may restrict the potential ability of Mastcams. First, downsampling the right images will throw away those high spatial resolution pixels in the right bands. Second, the lower resolution of the current stereo images may degrade the augmented reality or virtual reality experience of users. If one can apply some advanced pansharpening algorithms to the left bands, then one can have 12 bands of high-resolution image cube for the purpose of stereo vision and image fusion.

In the past two decades, there have been many papers discussing the fusion of a high resolution panchromatic (pan) image with a low-resolution multispectral image
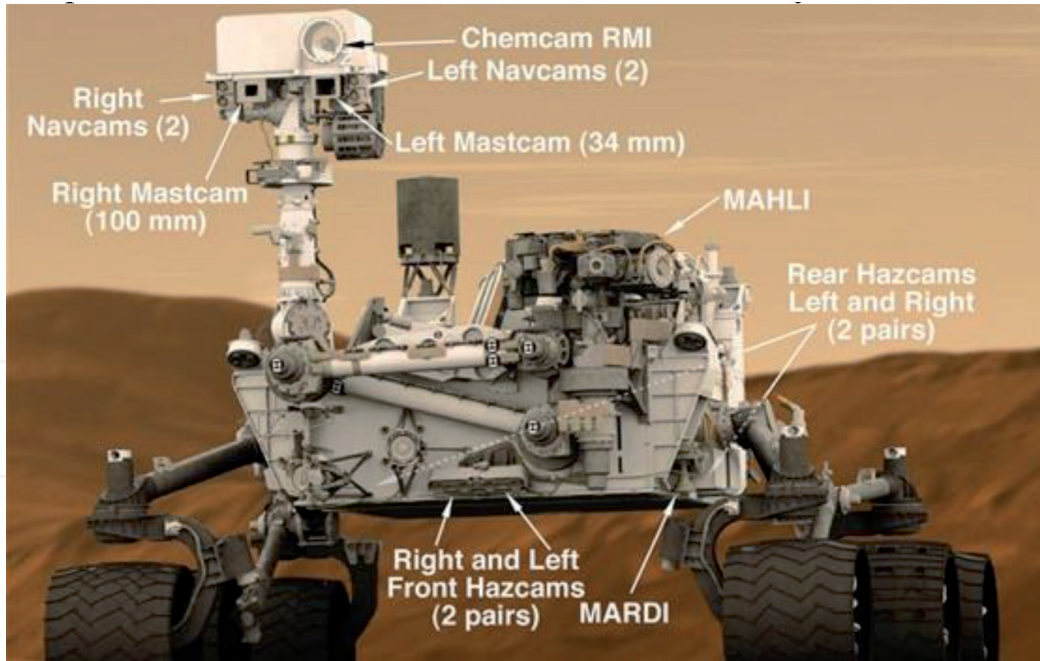
Figure 1.
Curiosity rover and its onboard instruments [7].
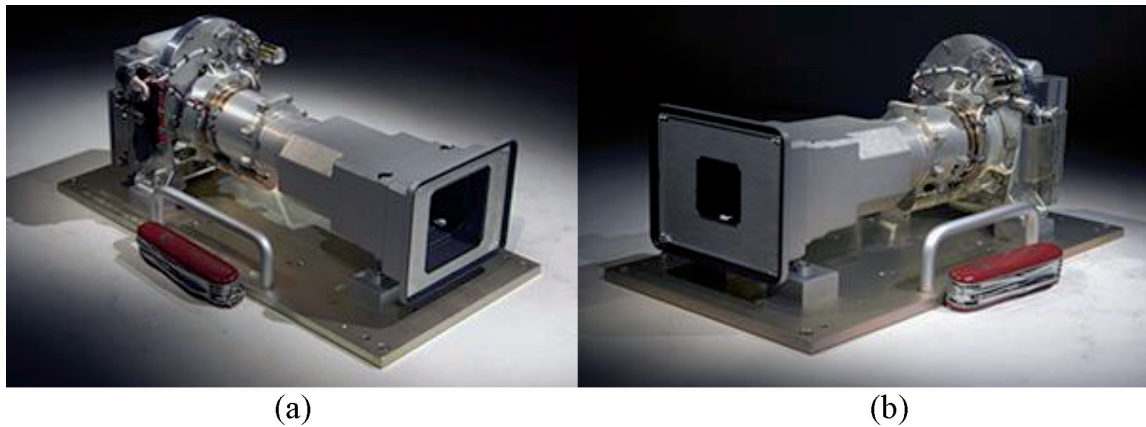
Figure 2.
The two Mastcam imagers [9]. (a) Left Mastcam (b) Right Mastcam.

(MSI) [10–14]. This is known as pansharpening. In our recent papers [15, 16], we proposed an unsupervised network structure to address the image fusion/super-resolution (SR) problem for hyperspectral image (HSI), referred to as HSI-SR, where a low-resolution (LR) HSI with high spectral resolution and a high-resolution (HR) MSI with low spectral resolution are fused to generate an HSI with high-resolution in both spatial and spectral dimensions. Similar to MSI, HSI has found extensive applications [17–21]. In this chapter, we adopt the innovative approaches designed in [15, 16], referred to as unsupervised sparse Dirichlet Network (uSDN), to enhance Mastcam images, where we treat the right Mastcam image as MSI with higher spatial resolution and the left Mastcam image as HSI with low spatial resolution.

In this chapter, we focus on the application of uSDN to enhance Mastcam images. In Section 2, we first introduce the problem of HSI-SR and then briefly summarize the key ideas of uSDN. In Section 3, we apply uSDN on actual Mastcam images. In Section 4, we include some further enhancements of uSDN and experiments. In Section 5, we introduce a transition learning concept, which is a natural extension of uSDN. Some preliminary results are also included. Finally, we conclude the chapter with some remarks.

2. The uSDN algorithm for HSI-SR

In this section, we describe the uSDN algorithm developed in [15, 16]. For more details, please refer to the reference. First of all, we will formulate the problem of HSI-SR to facilitate the discussion of Mastcam enhancement. **Table 1** summarizes the mathematical symbols used in this chapter.

The basic idea of uSDN is illustrated in **Figure 3**. First, the LR HSI, $\overline{Y_h} \in R^{m \times n \times L}$ with its width, height, and number of spectral bands denoted as m, n, and L, respectively, is unfolded into a 2D matrix, $Y_h \in R^{mn \times L}$. Similarly, the HR MSI, $\overline{Y_m} \in R^{M \times N \times l}$ with its width, height, and number of spectral bands denoted as M, N, and l, respectively, is unfolded into a 2D matrix $Y_m \in R^{MN \times l}$. And the SR HSI, $\overline{X} \in R^{M \times N \times L}$, is unfolded into a 2D matrix $X \in R^{MN \times L}$. Note that, generally, the spatial resolution of the MSI is much higher than that of the HSI, that is, $M \gg m$, $N \gg n$, and the spectral resolution of HSI is much higher than that of the MSI, that is, $L \gg l$. The objective is to reconstruct the high spatial and spectral resolution HSI, $\overline{X} \in R^{M \times N \times L}$, with LR HSI and HR MSI.

Due to the limitation of hardware, each pixel in an HSI or MSI may cover more than one constituent materials, leading to mixed pixels. These mixtures can be assumed to be a linear combination of a few basis vectors (or source signatures). Both LR HSI Y_h and HR MSI Y_m can be assumed to be a linear combination of c

HSI	Hyperspectral image
MSI	Multispectral image
HSI-SR	HSI super-resolution
HR	High-resolution
LR	Low-resolution
$\overline{Y_h} / Y_h$	3D/2D LR HSI
$\overline{Y_m} / Y_m$	3D/2D HR MSI
$\overline{X} / X$	3D/2D Reconstructed HR MSI
Φ_h	Spectral bases of HSI
Φ_m	Spectral bases of MSI
S_h	Coefficients/Representations of HSI
S_m	Coefficients/Representations of MSI
R	Transformation matrix
$\hat{Y}_h$	Reconstructed 2D HSI
W, b	Network weights and bias
$E_m(\theta_{he}) / E_m(\theta_{me})$	Encoder of the HIS/MSI
$D_h(\theta_{hd})$	Decoder of the HIS and MSI
$\theta_{he} / \theta_{me}$	Encoder weights of HIS/MSI
θ_{hd}	Decoder weights of HSI and MSI
s	Representations vector of a single pixel
v, u, β	Stick-breaking parameters
$H_p(s)$	Entropy function
$A(\tilde{S}_h, S_m)$	Angular difference

Table 1.
Symbols and abbreviations.
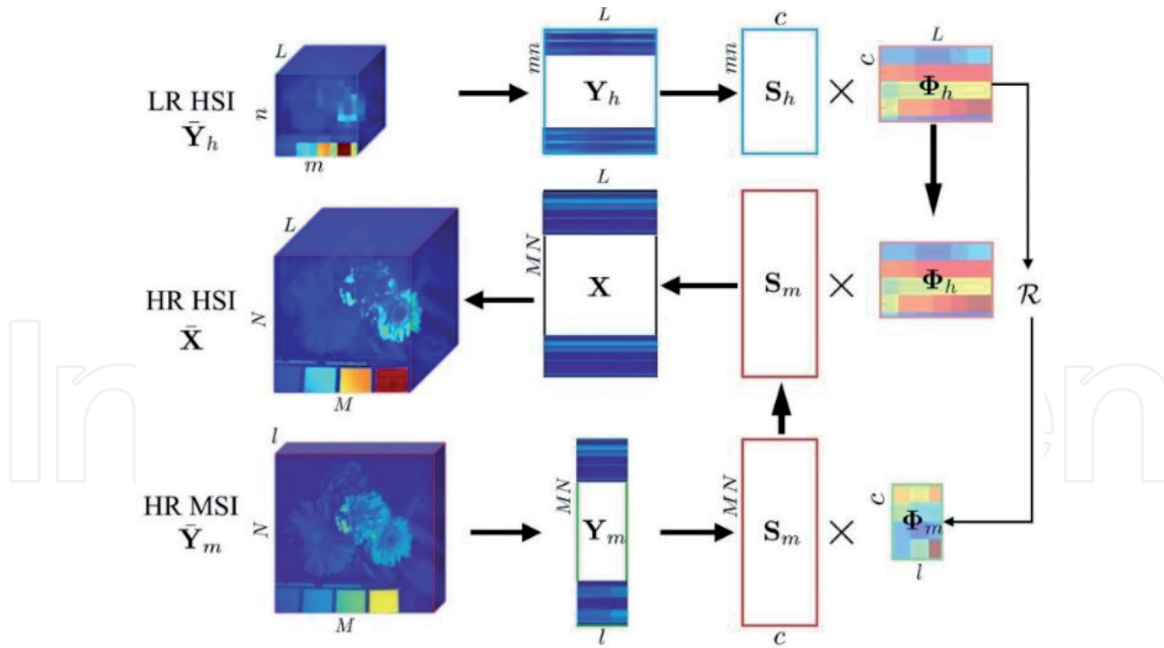
Figure 3.
General procedure of HSI-SR [15].

basis vectors with their corresponding proportional coefficients (referred to as representations in deep learning), as expressed in Eqs. (1) and (2), where $\Phi_h \in R^{c \times L}$ and $\Phi_m \in R^{c \times l}$ denote the spectral basis of Y_h and Y_m, respectively. They preserve the spectral information of the images. $S_h \in R^{mn \times c}$ and $S_m \in R^{MN \times c}$ are the proportional coefficients of Y_h and Y_m, respectively. Since the coefficients indicate how much each spectral basis has in constructing the mixed pixel at specific spatial locations, they preserve the spatial structure of HSI. The relationship between HSI and MSI bases can be expressed in the right part of Eq. (2), where $R \in R^{L \times l}$ is the transformation matrix given as a prior from the sensor [22–29].

$$Y_h = S_h \Phi_h, \tag{1}$$

$$Y_m = S_m \Phi_m, \quad \Phi_m = \Phi_h R, \tag{2}$$

$$X = S_m \Phi_h. \tag{3}$$

With $\Phi_h \in R^{c \times L}$ carrying the high spectral information and $S_m \in R^{MN \times c}$ carrying the high spatial information, the desired HR HSI, X, is generated by Eq. (3). See **Figure 3**. Since the ground truth X is not available, the problem has to be solved in an unsupervised fashion. In addition, the linear combination assumption enforces the representation vectors of HSI or MSI to be non-negative and sum-to-one, that is, $\sum_{j=1}^{c} s_{ij} = 1$, where s_i is the row vector of either S_m or S_h [24, 29].

The uSDN unsupervised architecture is shown in **Figure 4**. It has three unique structures. First, the network consists of two encoder-decoder networks, to extract the representations of the LR HSI and HR MSI, respectively. The two networks share the same decoder, such that both the spectral and spatial information from multi-modalities can be extracted with unsupervised settings. Second, the representations of both modalities, S_h and S_m, are enforced to follow a Dirichlet distribution where
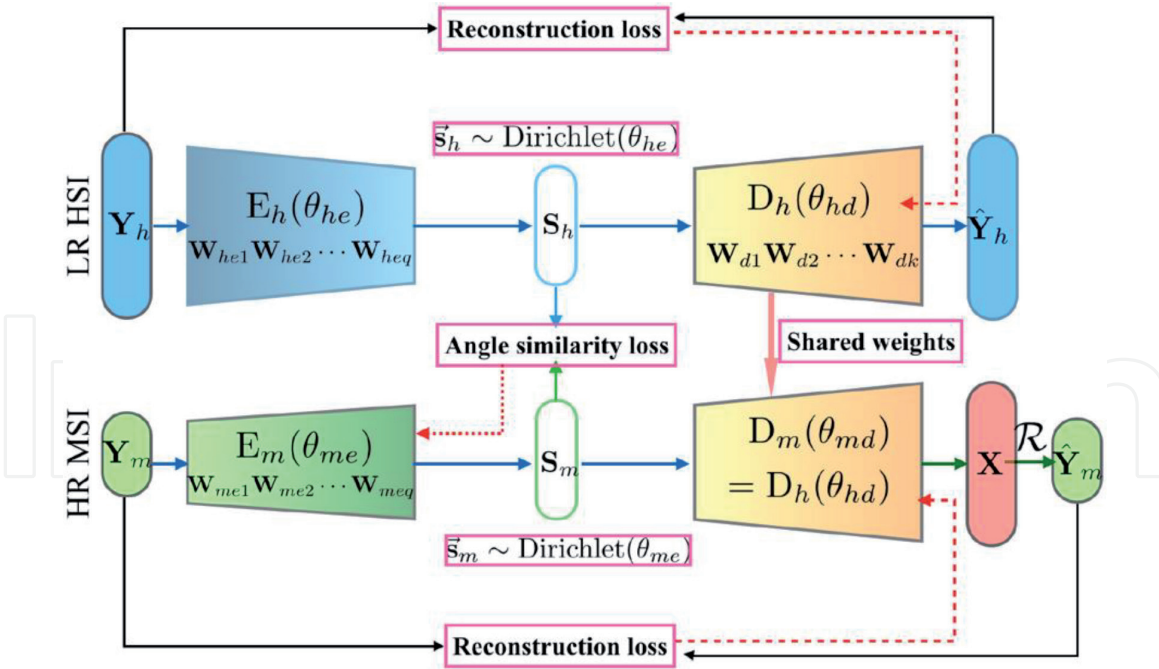
Figure 4.
Simplified architecture of uSDN [15].

the sum-to-one and non-negative properties are naturally incorporated into the network [30–34]. The solution space is further regularized with a sparsity constraint. Third, the angular difference of the representations from two modalities is minimized to preserve the spectral information of the reconstructed HR HSI.

3. Mastcam image enhancement using uSDN with improvements

### 3.1 Applying uSDN for Mastcam enhancement

uSDN has been thoroughly evaluated with two widely used benchmark datasets, CAVE [35] and Harvard [36]. Details can be found in [15, 16]. Here, we adopt uDSN to enhance the resolution of Mastcam images. As mentioned earlier, the right Mastcam has high resolution than the left. Hence, we treat the right Mastcam images are HR MSI and the left images as LR HSI. Although uSDN was introduced to deal with the general HSI super-resolution problem, we can treat the Mastcam image enhancement simply as a special case of HSI-SR.

For quantitative comparison, the root mean squared error (RMSE) and spectral angle mapper (SAM) are applied to evaluate the reconstruction error and the amount of spectral distortion, respectively.

The results are shown in **Figure 5**. The reconstructed image is very close to the ground truth. Most methods require that the size of high-resolution image should be equal to an integer multiplication of the size of low-resolution image. Thus, we only compare the method with CNMF [29] which works for arbitrary image size. The results are shown in **Table 2**. We observe that uSDN is able to outperform the CNMF.

3.2 Improvement based on uSDN

In this section, we summarize some further improvement of uDSN by fine-tuning the existing network structure in uSDN in order to further enhance the fusion performance.
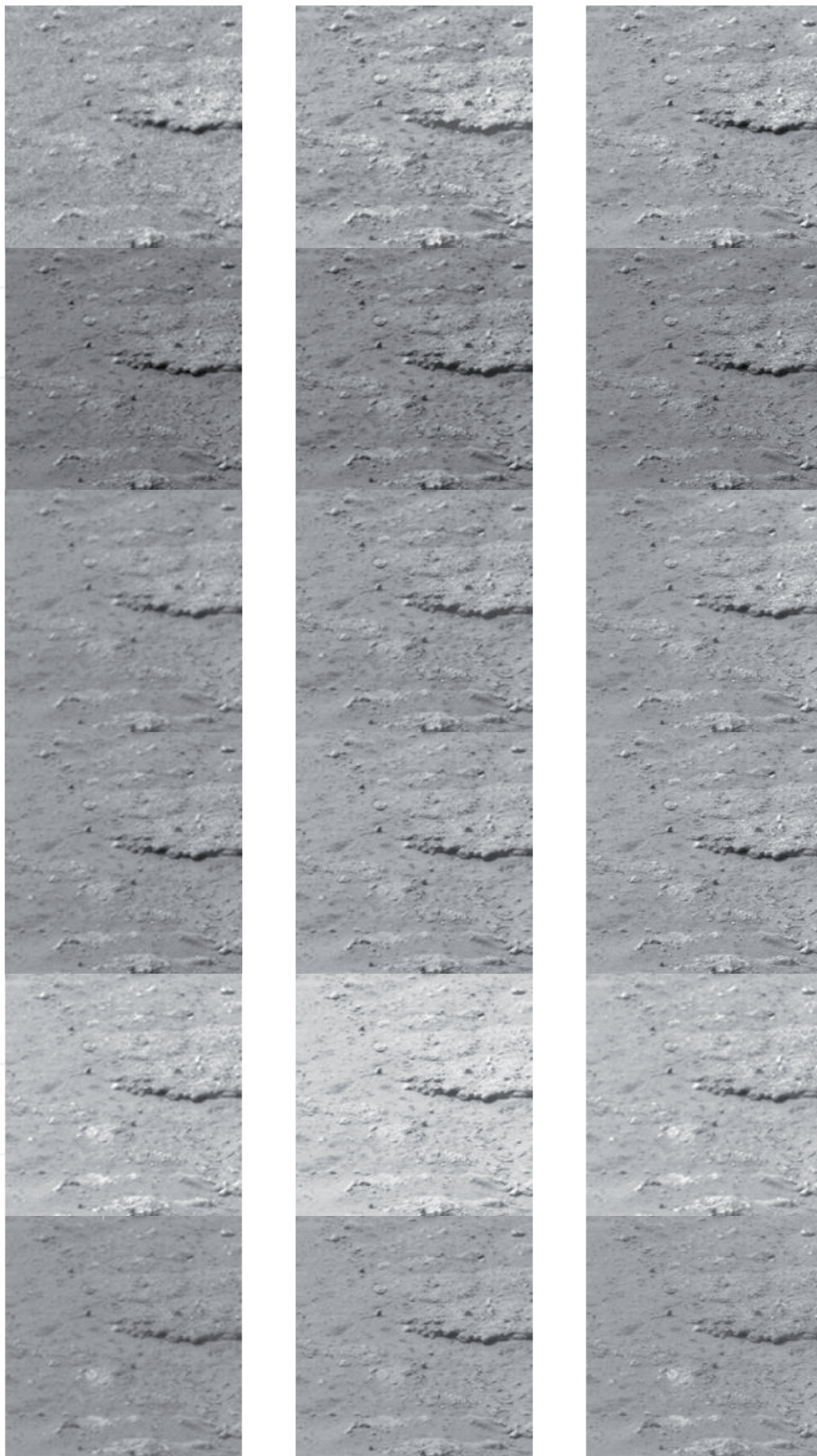
Figure 5.
Results of Mastcam image enhancement using uSDN. The left column shows the six bands from the left camera. The middle column shows the corresponding reconstructed results. The right column shows the six bands from the right camera.

Approaches	RMSE	SAM
CNMF	0.056	2.48
uSDN	0.033	2.09

Table 2.
Evaluations for image enhancement from Mastcam.

The existing structure of uDSN described in Section 3.1 is improved in two ways. First, in Section 3.1, the architecture consists of two deep networks, for the representation learning of the LR HSI and HR MSI, respectively. And only the decoders of the LR HSI and HR HSI networks are shared. The spectral information (i.e., the decoder of the LR HSI network) is extracted through the LR HSI network. Then the representation layer of the HR HSI is optimized by enforcing the spectral angle similarity. However, this introduces additional cost function, that is, angular difference minimization, and the optimization procedure is time consuming. In the improved uDSN, for the HR HSI network, most of the encoder weights are shared with the weights of the LR HSI encoder. Only a couple of encoder weights are updated during the HR HSI optimization. In this way, both the representations of the LR HSI and HR HSI networks are reinforced to follow Dirichlet distributions with parameters following the same trends. And the representations extracted from the LR HSI matches the patterns of that extracted from the HR HSI as shown in **Figure 6**.

Second, to further reduce the spectral distortion of the estimated HR HSI, instead of using l_2 loss, we adopt the l_{21} loss, which encourages the network to reduce the spectral loss of each pixel. Compared to the network with l_2 loss, the network with l_{21} loss is able to extract spectral information of images more accurately. The l_{21} loss can not only reduce the spectral distortion of the estimated HR HSI, but also improve the convergence speed of the network.

The result of the proposed method on individual HSI is visualized in **Figure 7**. When we optimize the network with l_{21} loss, we can observe that the difference between the estimated MSI and the ground truth MSI is very small, with RMSE of 1.7428 and SAM of 0.25615.
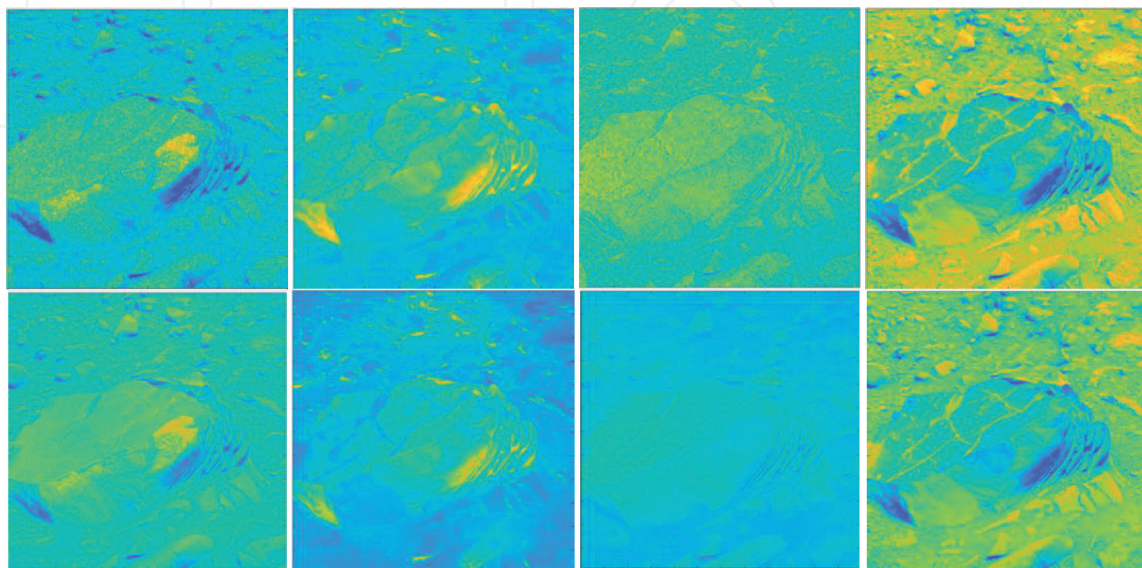
Figure 6.
Representations extracted from the LR HSI (top row) and the HR HSI (bottom row).
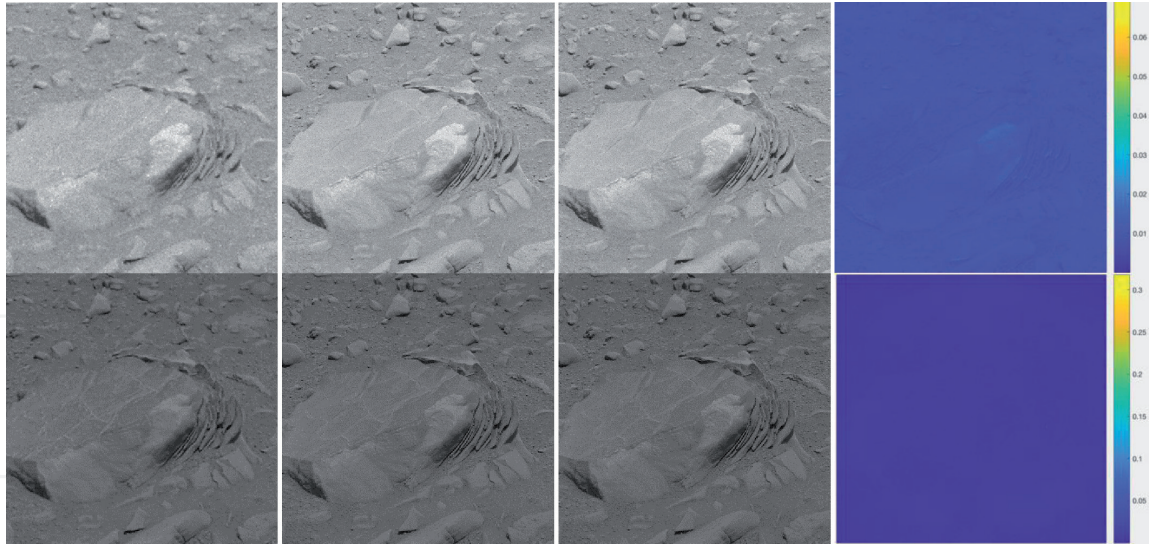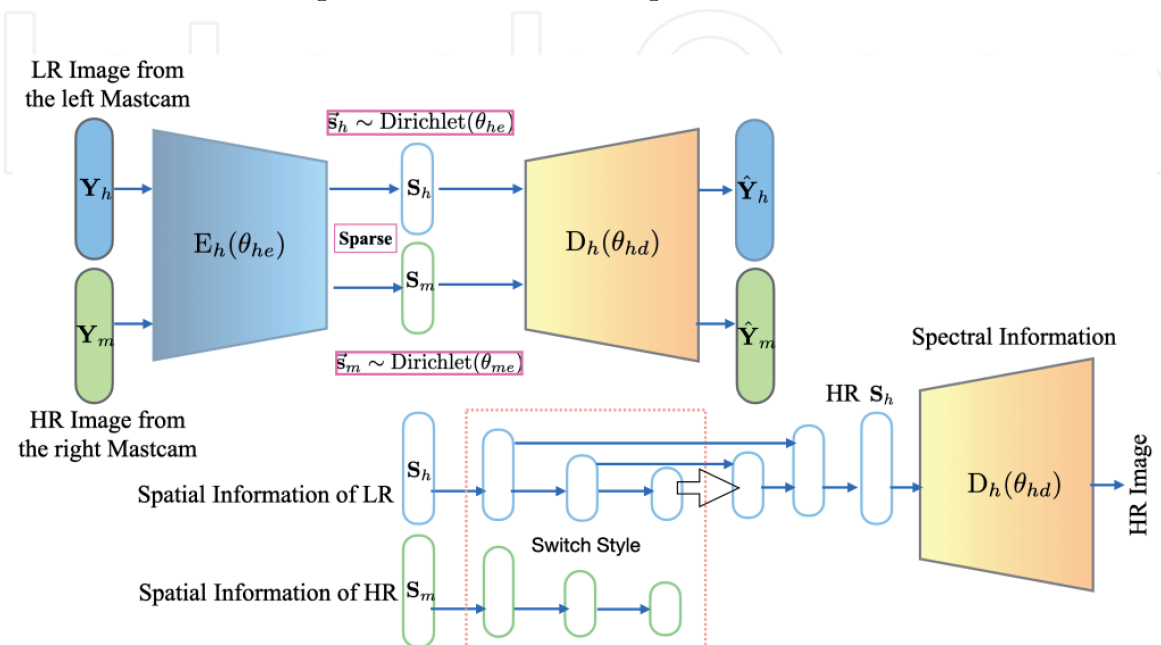
Figure 7.
The results using improved uSDN. The left column shows the first two bands from the left camera. The second column shows the corresponding reconstructed images from the improved uSDN. The third column shows the reference images from the right camera. The right column shows the absolute difference between the reconstructed images and the reference images.

4. Combination of Dirichlet-Net and U-Net

In this section, we propose to combine Dirichlet-Net with U-Net [37] to mitigate the mis-registration issue in the left and right Mastcam images.

Since in real scenarios, the images from the left and right cameras may not match each other perfectly even after registration, we propose a combination of Dirichlet-Net and U-Net to further improve the fusion performance using non-perfectly registered patches. We propose an unsupervised architecture as shown in **Figure 8**, which consists of two deep networks, an improved Dirichlet-Net for the representation learning of the MSI, and a U-Net for switching the low-resolution spatial information patches with high-resolution spatial information patches. Then the HR MSI of the left Mastcam image is generated by combining its spectral information with the spatial information of improved resolution.

Figure 8.
The architecture of the proposed approach that combines Dirichlet-net with U-Net.
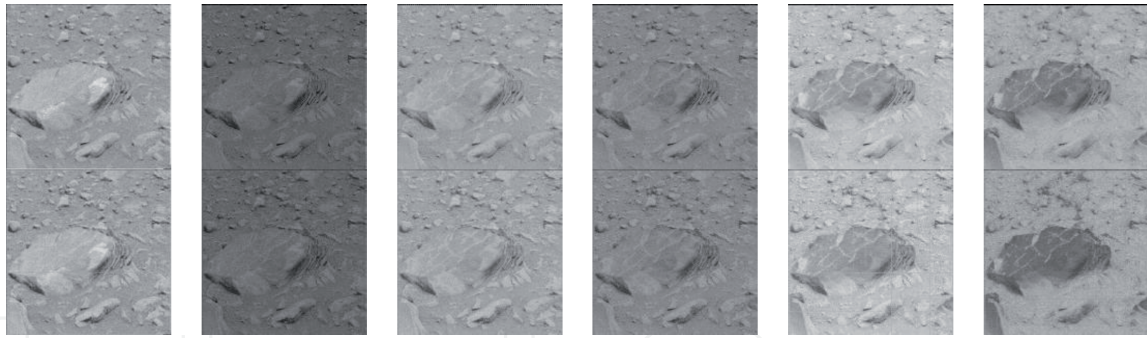
Figure 9.
The results of test image MSL_0002_0114_M1. The top row shows the six bands of raw images from the left camera. The bottom row shows the corresponding reconstructed images from the proposed method.
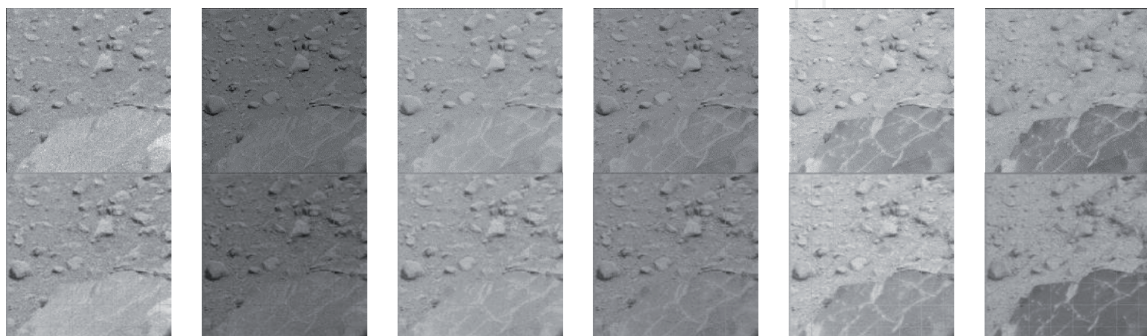
Figure 10.
The cropped results of test image MSL_0002_0114_M1. The top row shows the six bands of raw images from the left camera. The bottom row shows the corresponding reconstructed images from the proposed method.

From the last step in **Figure 8**, we are able to extract both the spectral and spatial information from LR MSI (left Mastcam) and HR MSI (right Mastcam). Although the scenes from the left and right camera are not the same, we assume they share the same group of spectral bases. And if we could improve the spatial information of the LR MSI using HR MSI, the quality of the LR MSI can be enhanced.

The architecture of the U-Net is illustrated in the lower part of **Figure 8**. We first learn a U-Net to recover the extracted spatial information, S_m, of HR MSI, Y_m, by convolution and deconvolution layers. The convolution layers extract HR spatial features from S_m, and the de-convolutional layers take these extracted features to rebuild the spatial information of S_m. Then we extract features from the spatial patches S_h of the LR HSI Y_h with the same convolution layers and switch these feature patches with their most similar feature patches in the HR spatial features [38]. Finally, the left Mastcam image with enhanced resolution, X, is generated by feeding the switching patches into de-convolutional layers of U-Net and the decoder of the Dirichlet-Net.

Here, we show experimental results from the proposed combination (Dirichlet-Net and U-Net) approach in **Figures 9** and **10**. We can observe that the reconstructed left Mastcam image is sharper than the raw MSI captured from the left camera directly and the spectral distortion of the recovered MSI is small, although only part of the high resolution MSI (right Mastcam image) is given from the right camera. Note that, due to the memory constraint, only a small patch can be recovered every time, thus there exist some disconnected parts in the results. This issue will be addressed in Section 5.

## 5. Spatial representation improvement with transition learning

High spatial resolution images have one natural property, that is, the transitions among pixel values are smooth. The patch-based method aims to replace the LR
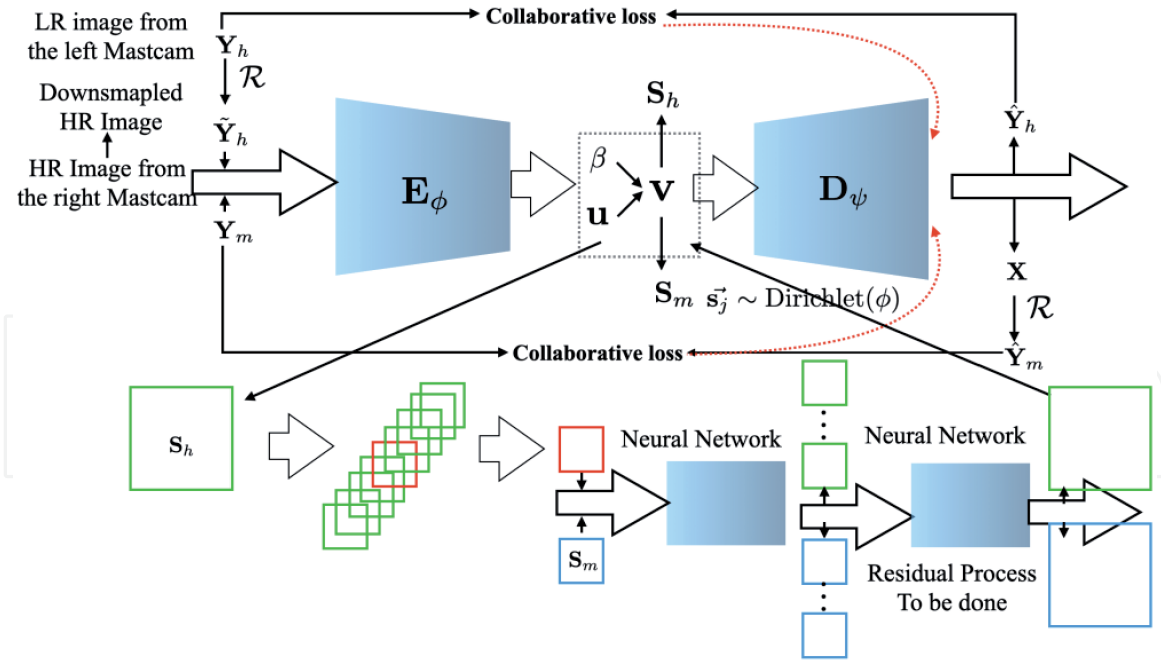
Figure 11.
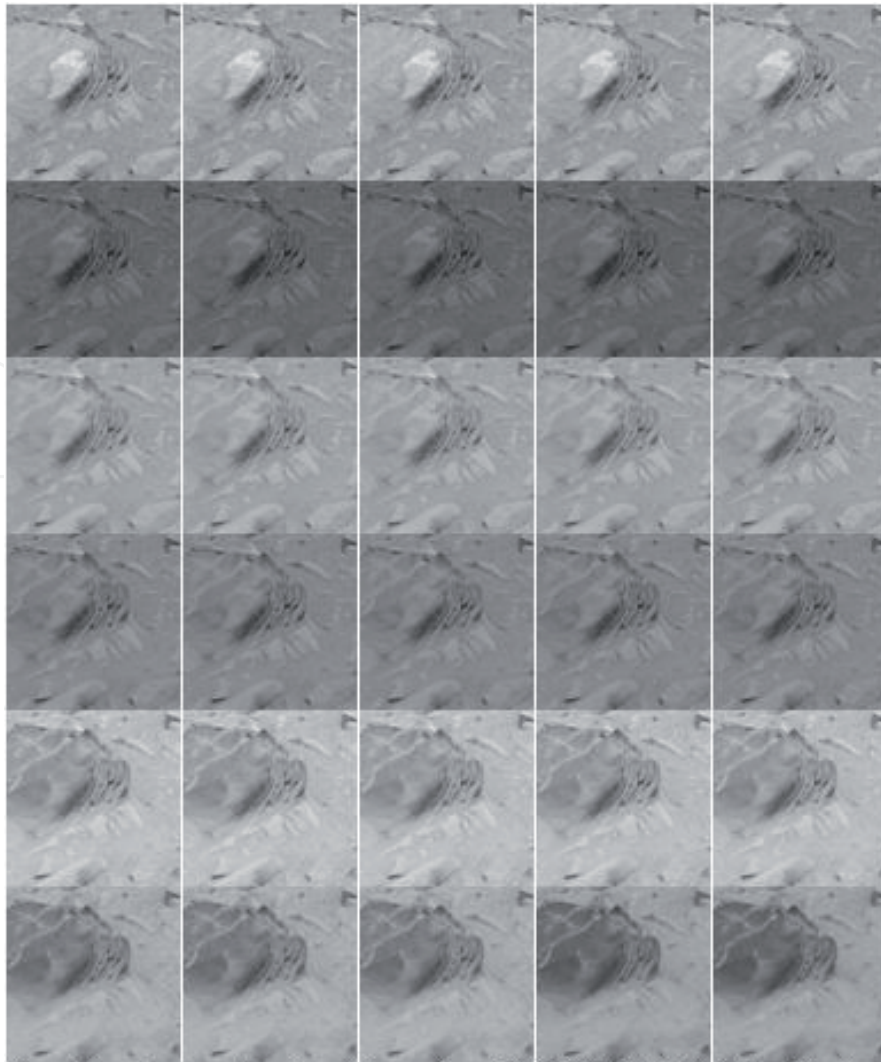*The architecture of the proposed transition learning approach.*
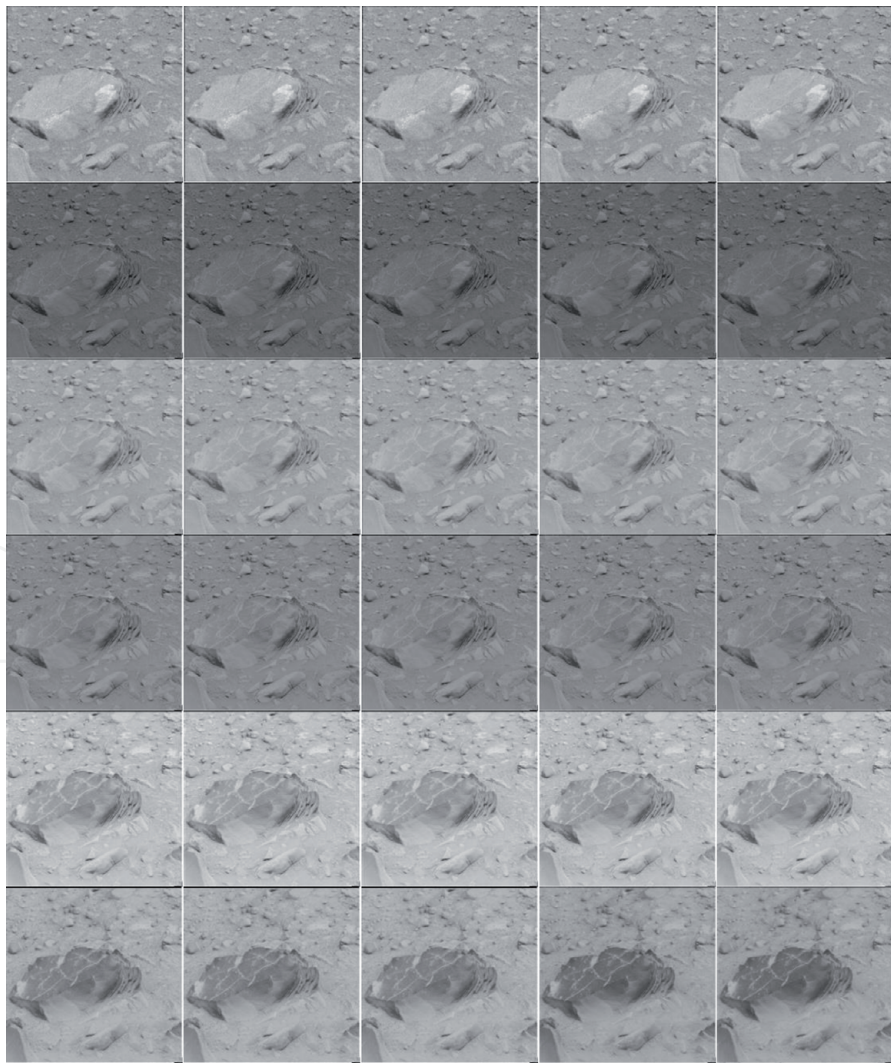
Figure 12.
The results of the test image MSL_0002_0114_M1. The left column shows the six bands of raw images from the left camera. The second, third, fourth and fifth columns show the corresponding reconstructed images from Bicubic, EnhanceNet, the proposed patch-based method and the residual-based transition-learning method, respectively.
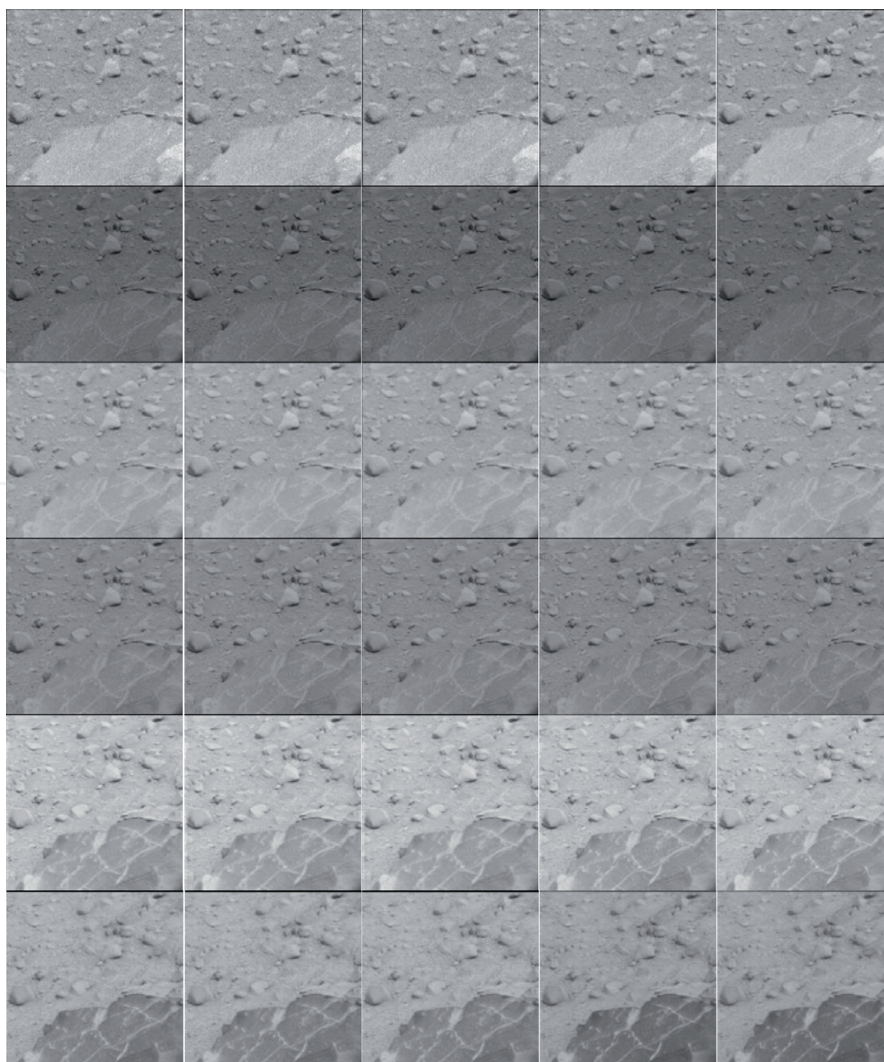
Figure 13.
The cropped results of the test image MSL_0002_0114_M1. The left column shows the six bands of raw images from the left camera. The second, third, fourth, and fifth columns show the corresponding reconstructed images from Bicubic, EnhanceNet, the proposed patch-based method and the residual-based transition-learning method, respectively.

patches from the LR MSI representations S_h with the most similar HR patches from the HR MSI representation, S_m. Since the LR MSI and HR MSI are unregistered and there is no ground truth of enhanced MSI, the patch-based improvement could not guarantee the smooth transitions in the reconstructed images, that is, the replaced patches may not match their neighbors. Therefore, in this section, we propose another structure based on transition-learning, to further improve the spatial resolution of LR HSI. The main structure is shown in **Figure 11**.

To learn smooth transitions between pixels, we first extract sub-images from the representations S_m, of HR MSI with stride 3, as shown in the lower part of **Figure 11**. For example, since the super-resolution factor is 3, we extract 9 sub-images from S_m. Then the network learns the transitions between the center sub-image with the other 8 sub-images. Since the LR MSI and HR MSI have similar statistic distributions, we assume that the transitions among pixels in both modalities are the same. Therefore, the representations S_h of LR MSI can be treated as the center sub-image of enhanced MSI and the other 8 sub-images of enhanced MSI can be estimated by feeding the representations S_h of LR MSI into the network trained by S_m. There are still residuals between the reconstructed and the ideal representations of S_m. This time, we adopt the principle described earlier to add high frequency residuals on the enhanced MSI.

Figure 14.
The cropped results of the test image MSL_0002_0114_M1. The left column shows the six bands of raw images from the left camera. The second, third, fourth, and fifth columns show the corresponding reconstructed images from Bicubic, EnhanceNet, the proposed patch-based method and the residual-based transition-learning method, respectively.

Here, the experimental results of the proposed approaches are compared with the results from Bicubic and the state-of-the-art single image super-resolution method EnhanceNet [39], as shown in **Figures 12**–**14**. Note that, since the EnhanceNet only offers the 4X pre-trained weights, we show its 4X reconstruction results for fair comparison, in case the down-sampling procedure reduces the quality of the reconstructed images. The Bicubic does not improve the resolution much. The EnhanceNet was trained on natural image dataset; thus it works poorly on remote sensing images. Compared to the bicubic or EnhanceNet methods, we can observe that the proposed methods can not only improve the spatial resolution of LR MSI, but also preserve the spectral information well, even though the images from the left and right camera are not registered. The transition-based approach works better than the patch-based one, because it learns the relationship between the reconstructed pixels.

6. Conclusions

In this chapter, we summarize the application of several deep learning-based image fusion algorithms to enhance Mastcam images from Mars rover. The first

algorithm termed as uDSN is based on the Dirichlet-Net, which incorporates the sum-to-one and sparsity constraints. Two improvements of the uDSN were then investigated. Finally, a transition learning-based approach was developed. Promising results using actual Mastcam images are presented. More research will be carried out in the future to continue the above investigations.

Acknowledgements

Author details

Ying Qu[1], Hairong Qi[1] and Chiman Kwan[2]*

1 University of Tennessee Knoxville, Knoxville, Tennessee, USA

2 Applied Research LLC, Rockville, Maryland, USA

*Address all correspondence to: chiman.kwan@arllc.net

IntechOpen

References

[1] Ayhan B, Kwan C, Vance S. On the Use of a Linear Spectral Unmixing Technique for Concentration Estimation of APXS Spectrum. Journal of Multidisciplinary Engineering Science and Technology. 2015;**2**(9):2469-2474

[2] Wang W, Li S, Qi H, Ayhan B, Kwan C, Vance S. Revisiting the preprocessing procedures for elemental concentration estimation based on CHEMCAM LIBS on MARS Rover. In: 6th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS). Lausanne, Switzerland; 2014

[3] Bell JF et al. The Mars Science Laboratory Curiosity Rover Mast Camera (Mastcam) Instruments: Pre-flight and in-flight calibration, validation, and data archiving. Earth and Space Science. July 2017;**4**(7): 396-452

[4] Ayhan B, Kwan C. Mastcam image resolution enhancement with application to disparity map generation for stereo images with different resolutions. Sensors. 2019;**19**:3526

[5] Kwan C, Chou B, Ayhan B. Enhancing stereo image formation and depth map estimation for Mastcam images. In: IEEE Ubiquitous Computing, Electronics & Mobile Communication Conference. New York City; 2018

[6] Ayhan B, Dao M, Kwan C, Chen H, Bell JF III, Kidd R. A novel utilization of image registration techniques to process Mastcam images in Mars rover with applications to image fusion, pixel clustering, and anomaly detection. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 2017;**10**(10):4553-4564

[7] Kwan C, Chou B, Bell JF III. Comparison of deep learning and conventional demosaicing algorithms for Mastcam images. Electronics. 2019;**8**:308

[8] Kwan C, Larkin J, Budavari B, Chou B. Compression algorithm selection for multispectral Mastcam images. Signal & Image Processing: An International Journal. 2019;**10**(1)

[9] Available from: https://msl-scicorner. jpl.nasa.gov/Instruments/Mastcam/

[10] Aiazzi B, Alparone L, Baronti S, Garzelli A, Selva M. Mtf-tailored multiscale fusion of high-resolution ms and pan imagery. Photogrammetric Engineering & Remote Sensing. 2006;**72**(5):591-596

[11] Aiazzi B, Baronti S, Selva M. Improving component substitution pansharpening through multivariate regression of ms+ pan data. IEEE Transactions on Geoscience and Remote Sensing. 2007;**45**(10)

[12] Akhtar N, Shafait F, Mian A. Bayesian sparse representation for hyperspectral image super resolution. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015. pp. 3631-3640

[13] Akhtar N, Shafait F, Mian A. Hierarchical beta process with Gaussian process prior for hyperspectral image super resolution. In: European Conference on Computer Vision. 2016. pp. 103-120

[14] Borengasser M, Hungate WS, Watkins R. Hyperspectral Remote Sensing: Principles and Applications. Boca Raton, Florida, USA: CRC Press; 2007

[15] Qu Y, Qi H, Kwan C. Unsupervised sparse Dirichlet-Net for hyperspectral image super-resolution. In: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2018. pp. 2511-2520

[16] Qu Y, Qi H, Kwan C. Unsupervised and Unregistered Hyperspectral Image Super-Resolution with Mutual Dirichlet-Net, arXiv preprint arXiv:1904.12175; 2019

[17] Zhou J, Kwan C, Ayhan B, Eismann MT. A novel cluster kernel RX algorithm for anomaly and change detection using hyperspectral images. IEEE Transactions on Geoscience and Remote Sensing. 2016;**54**(11):6497-6504

[18] Li S, Wang W, Qi H, Ayhan B, Kwan C, Vance S. Low-rank tensor decomposition based anomaly detection for hyperspectral imagery. In: IEEE International Conference on Image Processing (ICIP). 2015

[19] Kwan C, Budavari B, Bovik AC, Marchisio G. Blind quality assessment of fused WorldView-3 images by using the combinations of pansharpening and hypersharpening paradigms. IEEE Geoscience and Remote Sensing Letters. 2017;**14**(10):1835-1839

[20] Zhou J, Kwan C, Budavari B. Hyperspectral image super-resolution: A hybrid color mapping approach. Journal of Applied Remote Sensing. 2016;**10**(3):035024

[21] Kwan C, Ayhan B, Chen G, Wang J, Ji B, Chang C-I. A novel approach for spectral unmixing, classification, and concentration estimation of chemical and biological agents. IEEE Transactions on Geoscience and Remote Sensing. 2006;**44**(2):409-419

[22] Dian R, Fang L, Li S. Hyperspectral image super-resolution via non-local sparse tensor factorization. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017. pp. 5344-5353

[23] Kawakami R, Matsushita Y, Wright J, Ben-Ezra M, Tai Y-W, Ikeuchi K. High-resolution hyperspectral imaging via matrix factorization. In: 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2011. pp. 2329-2336

[24] Lanaras C, Baltsavias E, Schindler K. Hyperspectral super-resolution by coupled spectral unmixing. In: Proceedings of the IEEE International Conference on Computer Vision. 2015. pp. 3586-3594

[25] Loncan L, de Almeida LB, Bioucas-Dias JM, Briottet X, Chanussot J, Dobigeon N, et al. Hyperspectral pansharpening: A review. IEEE Geoscience and Remote Sensing Magazine. 2015;**3**(3):27-46

[26] Simoes M, Bioucas-Dias J, Almeida LB, Chanussot J. A convex formulation for hyperspectral image superresolution via subspace-based regularization. IEEE Transactions on Geoscience and Remote Sensing. 2015;**53**(6):3373-3388

[27] Vivone G, Alparone L, Chanussot J, Dalla Mura M, Garzelli A, Licciardi GA, et al. A critical comparison among pansharpening algorithms. IEEE Transactions on Geoscience and Remote Sensing. 2015;**53**(5):2565-2586

[28] Wei Q, Bioucas-Dias J, Dobigeon N, Tourneret J-Y. Hyperspectral and multispectral image fusion based on a sparse representation. IEEE Transactions on Geoscience and Remote Sensing. 2015;**53**(7):3658-3668

[29] Yokoya N, Yairi T, Iwasaki A. Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion. IEEE Transactions on Geoscience and Remote Sensing. 2012;**50**(2):528-537

[30] Wycoff E, Chan T-H, Jia K, Ma W-K, Ma Y. A non-negative sparse promoting algorithm for high resolution hyperspectral imaging. In: 2013 IEEE International Conference on Acoustics,

Speech and Signal Processing (ICASSP). 2013. pp. 1409-1413

[31] Sethuraman J. A constructive definition of Dirichlet priors. Statistica Sinica. 1994;**4**:639-650

[32] Ozkan S, Kaya B, Esen E, Akar G. EndNet: Sparse AutoEncoder Network for Endmember Extraction and Hyperspectral Unmixing, 08; arXiv:1708.01894; 2017

[33] Yokoya N, Grohnfeldt C, Chanussot J. Hyperspectral and multispectral data fusion: A comparative review of the recent literature. IEEE Geoscience and Remote Sensing Magazine. 2017;**5**(2):29-56

[34] Huang S, Tran TD. Sparse signal recovery via generalized entropy functions minimization. arXiv preprint arXiv:1703.10556; 2017

[35] Yasuma F, Mitsunaga T, Iso D, Nayar SK. Generalized assorted pixel camera: postcapture control of resolution, dynamic range, and spectrum. IEEE Transactions on Image Processing. 2010;**19**(9):2241-2253

[36] Chakrabarti A, Zickler T. Statistics of real-world hyperspectral images. In: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2011. pp. 193-200

[37] Isola P, Zhu J-Y, Zhou T, Efros AA. Image-to-image translation with conditional adversarial networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017. pp. 1125-1134

[38] Chen TQ, Schmidt M. Fast Patch-Based Style Transfer of Arbitrary Style. arXiv preprint arXiv:1612.04337; 2016

[39] Sajjadi M, Scholkopf B, Hirsch M. EnhanceNet: Single image super-resolution through automated texture synthesis. In: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017