

# Hyperspectral Image Fusion Methods for the Mars Reconnaissance Orbiter

AJ Adkins, Sugandha Roy, and Chao Ren

**Abstract**—High resolution, hyperspectral images on Mars are essential to remote sensing and resource detection. However, hyperspectral images on Mars typically have low resolution or limited to some sub-regions. In contrast, multispectral or panchromatic images cover extended regions and have higher resolution. As a result, fusion of hyperspectral images with higher resolution multispectral or panchromatic images is a useful tool. A multitude of algorithms, both computational and machine learning-based, have been developed to achieve this task, but have been limited in application to terrestrial scenarios. In this paper, we apply two fusion techniques to imagery from the Mars Reconnaissance Orbiter (MRO), a dataset which has previously been unexplored with regards to image fusion. In particular, we employ approaches based on nonnegative matrix factorization and a guided deep decoder to perform fusion of hyperspectral images from the Compact Reconnaissance Imaging Spectrometer for Mars (CRISM) and panchromatic images from the Context Camera (CTX). We implement a deep-decoder network that is composed of an encoder-decoder network combined with attention mechanism that exploits multi-scale features of a guidance image, yielding a high resolution fused image rich in spectral information [1]. The results from both methods are individually evaluated based on established quantitative metrics and compared to each other to determine advantages and disadvantages between the techniques. High resolution hyperspectral images could lead to improved determination of water and mineral content on the Mars surface, informing landing sites for future manned missions to the planet.

**Index Terms**—Mars, Image Fusion, Hyperspectral, Non-negative matrix factorization, Deep learning model.

## I. INTRODUCTION

**H**YPERSPECTRAL imaging is an important tool in modern remote sensing, providing spectroscopic information over a two-dimensional area. This data is used to identify surface material composition, such as plant coverage or the presence of certain minerals. Outside of Earth-based observations, hyperspectral imaging is a crucial method to identify optimal landing positions for future missions to Mars. To achieve this, the Mars Reconnaissance Orbiter (MRO) is equipped with a handful of imaging instruments, including a hyperspectral system called the Compact Reconnaissance Imaging Spectrometer for Mars (CRISM). The multi-pixel spectrometer is able to cover the spectral range from 362 - 3920 nm with a spectral sampling of 6.55 nm/channel. Despite its ability to capture almost continuous spectral information, the spatial resolution is relatively low at 18.4 meters/pixel from a 300 kilometer altitude. The MRO is also equipped with a non-spectral Context Camera (CTX), which has over twice the spatial resolution at 6 meters/pixel at a 300 km altitude [2]. The ideal dataset would be one with the spatial resolution of the CTX camera, but containing all of the spectral

information of CRISM. In general, this problem has motivated the development of data fusion techniques for hyperspectral and multispectral/panchromatic images.

### A. Hyperspectral Image Fusion

Given a high spectral resolution, low spatial resolution image (hyperspectral) and a low spectral resolution, high spatial resolution image (multispectral), the idea of image fusion is to create a combined image that contains the high spatial spatial resolution of the multispectral image while retaining the full spectral information of the hyperspectral image. Mathematically, we wish to estimate a matrix  $\mathbf{Z} \in \mathbb{R}^{\lambda_h \times L_m}$  from two matrices  $\mathbf{X}_h \in \mathbb{R}^{\lambda_h \times L_h}$  and  $\mathbf{X}_m \in \mathbb{R}^{\lambda_m \times L_m}$ , where  $\lambda_h$  and  $\lambda_m$  are the number of spectral channels, and  $L_h$  and  $L_m$  are the number of pixels of the hyperspectral and multispectral images, each respectively. To date, many techniques have been developed to achieve this task. Computational methods include algorithms based on pansharpening [3], maximum a posteriori (MAP) estimation [4], and nonnegative matrix factorization (NMF) [5] [6]. Machine learning-based methods include both supervised and unsupervised deep learning [7] [8] [9]. Unlike many Earth-based remote sensing applications, there is no large scale ground truth for observations made from an orbiting satellite, which limits our selection to computational methods, or unsupervised deep learning methods.

## II. NONNEGATIVE MATRIX FACTORIZATION

A popular method to achieve fusion of hyperspectral and multispectral images is by nonnegative matrix factorization, which is closely related to the concept of spectral unmixing. The method is explored extensively, and can be reformulated in a wide variety of ways to accommodate different cost functions and constraints. We first introduce the general problem, then discuss specific implementations for image fusion.

### A. Linear Mixing Model

In the so called linear mixing model (LMM), each individual pixel of a hyperspectral image is composed of a weighted sum of spectral components, which may be representative of a combination of real materials present within the pixel. [10] These basis spectra are called *endmembers*, and their corresponding weights are called *abundances*. Given  $M$  endmembers, this can be expressed as

$$\mathbf{z} = \sum_{i=1}^M h_i \mathbf{w}_i = \mathbf{W}\mathbf{h}, \quad \mathbf{Z} = \mathbf{W}\mathbf{H} \quad (1)$$

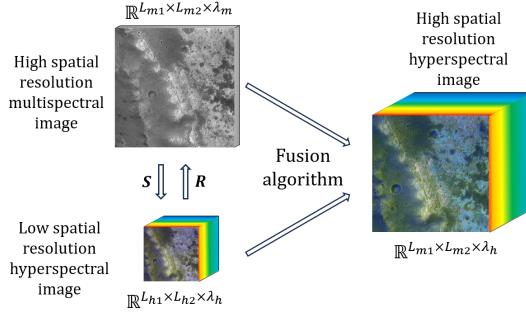


Fig. 1: Basic scheme for hyperspectral image fusion.

where  $\mathbf{Z}$  is the hyperspectral image,  $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_M]$  is the endmember matrix, and  $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_M]$  is the abundance matrix. Noise will also be present in this model, which we have omitted for illustrative simplicity.

The goal of nonnegative matrix factorization is then to factor  $\mathbf{Z}$  into  $\mathbf{W}$  and  $\mathbf{H}$  such that both matrices are nonnegative. The nonnegativity constraint is necessary to make the factorization physically meaningful as a real image would never contain negative reflectance values. After this process, the image can be considered spectrally unmixed. Qualitatively, the abundance matrix contains the spatial information of the image, equivalent to a grayscale representation, and the endmember matrix contains the spectral information.

The key idea applied to image fusion is that given a low resolution hyperspectral image  $\mathbf{X}_h$  and a high resolution multispectral image  $\mathbf{X}_m$ , they can each be factored into their abundance and endmember matrices

$$\mathbf{X}_h = \mathbf{W}_h \mathbf{H}_h \quad (2)$$

$$\mathbf{X}_m = \mathbf{W}_m \mathbf{H}_m \quad (3)$$

from which the fused high resolution hyperspectral image  $\mathbf{Z}$  can be constructed by multiplication of the high resolution abundance matrix with the high spectral resolution endmember matrix:

$$\mathbf{Z} = \mathbf{W}_h \mathbf{H}_m \quad (4)$$

### B. Coupling of Hyperspectral and Multispectral Images

A straightforward approach to solve the fusion problem is to individually unmix  $\mathbf{X}_h$  and  $\mathbf{X}_m$ , then form the fused image by (4), with the factors  $\mathbf{W}_m$  and  $\mathbf{H}_m$  left unused. However, the fact that both  $\mathbf{X}_h$  and  $\mathbf{X}_m$  are images of the same scene can be exploited to optimize the computation. In particular, the ideal fused image  $\mathbf{Z}$  can be transformed into the hyperspectral image  $\mathbf{X}_h$  by spatial downsampling, and into the multispectral image  $\mathbf{X}_m$  by spectral truncation. Mathematically,

$$\mathbf{X}_h = \mathbf{Z}\mathbf{S} \quad (5)$$

$$\mathbf{X}_m = \mathbf{R}\mathbf{Z} \quad (6)$$

where  $\mathbf{S}$  is the spatial degradation matrix, and  $\mathbf{R}$  is the spectral degradation matrix. Similarly, the hyperspectral abundance matrix can be considered the spatially degraded multispectral abundance matrix, and the multispectral endmember matrix is the spectrally degraded hyperspectral endmember matrix:

$$\mathbf{H}_h = \mathbf{H}_m \mathbf{S} = \tilde{\mathbf{H}}_m \quad (7)$$

$$\mathbf{W}_m = \mathbf{R}\mathbf{W}_h = \tilde{\mathbf{W}}_h \quad (8)$$

where we have defined  $\tilde{\mathbf{H}}_m$  and  $\tilde{\mathbf{W}}_h$  for notational simplicity. This coupled relationship between the two images allows us to construct a single cost function when performing the nonnegative matrix factorization.

### C. Minimization of Euclidean Distance

Given  $\mathbf{X}_h$  and  $\mathbf{X}_m$ , our task is to estimate  $\mathbf{W}_h$  and  $\mathbf{H}_m$ . We follow an approach similar to [11] by considering the following constrained least squares problem:

$$\arg \min_{\mathbf{W}_h, \mathbf{H}_m} \|\mathbf{X}_h - \mathbf{W}_h \tilde{\mathbf{H}}_m\|_F^2 + \|\mathbf{X}_m - \tilde{\mathbf{W}}_h \mathbf{H}_m\|_F^2 \quad (9a)$$

subject to

$$w_{h,ij} \geq 0, \forall i, j \quad (9b)$$

$$0 \leq h_{m,ij} \leq 1, \forall i, j \quad (9c)$$

$$\mathbf{1}^T \mathbf{H}_m = \mathbf{1}^T \quad (9d)$$

where constraint (9b) specifies the nonnegativity of the endmember matrix, constraint (9c) requires the abundance values (the basis coefficients) to be between 0 and 1, and constraint (9d) requires the abundance values to sum to one. An additional regularization term can be added to the cost function, which has been done in related work [12], but we have found that the fusion performs well without any regularization. Also notice that although this is a "double unmixing" problem, the minimization need only be done over the desired matrices due to the relationships (7) and (8).

Each term in the sum (9a) can be alternately minimized, with the degraded matrices  $\tilde{\mathbf{W}}_h$  and  $\tilde{\mathbf{H}}_m$  computed at each step by (7) and (8). Following a proximal gradient method, the update rules are

$$\mathbf{W}_h \leftarrow \text{prox}_{\mathbf{W}}[\mathbf{W}_h - \gamma_W (\mathbf{W}_h \tilde{\mathbf{H}}_m - \mathbf{X}_h) \tilde{\mathbf{H}}_m^T] \quad (10a)$$

$$\mathbf{H}_m \leftarrow \text{prox}_{\mathbf{H}}[\mathbf{H}_m - \gamma_H \tilde{\mathbf{W}}_h^T (\tilde{\mathbf{W}}_h \mathbf{H}_m - \mathbf{X}_m)] \quad (10b)$$

The step sizes can be computed as  $\gamma_W = 1/c \|\tilde{\mathbf{H}}_m \tilde{\mathbf{H}}_m^T\|_F$  and  $\gamma_H = 1/d \|\tilde{\mathbf{W}}_h \tilde{\mathbf{W}}_h^T\|_F$  where  $c$  and  $d$  are constants. The proximal operator is itself a sub-optimization problem that seeks to project the result of the gradient method step onto the feasible set defined by the constraints. Luckily, this projection is computationally very simple, amounting to setting all negative matrix values to zero for  $\text{prox}_{\mathbf{W}}$  and  $\text{prox}_{\mathbf{H}}$ , and setting all values greater than one for  $\text{prox}_{\mathbf{H}}$ . This is along the same lines as projection onto the simplex [13].

### III. IMAGE FUSION WITH DEEP-LEARNING BASED MODEL

We have employed a network architecture for fusing input and guidance images. Traditional methods apply task-specific handcrafted priors and lacked a unified approach. To overcome this, we have used a guided deep decoder (GDD) network as a general solution, which consists of an encoder-decoder network for exploiting multi-scale features of a guidance image and a deep decoder network for generating an output image.

#### A. Network Architecture

##### Encoder-Decoder Network with Skip Connections:

**Encoder:** The encoder part of the network takes an input image (which is the guidance image 'G' in this case) and processes it through several layers. Each layer typically consists of convolution operations (often with a stride of 2 for downscaling), followed by activation functions like Leaky ReLU, and normalization steps. The purpose of the encoder is to extract high-level semantic features from the input image. As we go deeper, the spatial resolution decreases while the feature dimensionality increases.

**Decoder:** The decoder takes the encoded features and reconstructs the input image back to its original size. It typically uses upsampling operations, such as bilinear upsampling, and convolution operations to increase the spatial resolution while reducing the feature dimensionality. The goal of the decoder is to reconstruct the input from its encoded features, ideally preserving all the important information.

**Skip Connections:** Skip connections are direct links from layers in the encoder to corresponding layers in the decoder. They help in preserving spatial information that might be lost during the encoding process and are particularly useful for tasks requiring precise localization (like segmentation).

##### Deep Decoder Network:

**Noise Map (Z):** This is a randomly initialized feature map that serves as the input to the deep decoder network. The noise map provides a starting point for the iterative reconstruction process.

**Deep Decoder Layers:** These layers consist of upsampling and convolution operations (similar to the decoder in the encoder-decoder network) but are designed to refine the noise map into a coherent output image. Each layer likely includes normalization and activation functions, with Leaky ReLU mentioned as the activation function in this architecture.

##### Final Output (X):

The deep decoder network outputs an image that is expected to be a reconstruction of the original input image (Y) but guided by the features of the guidance image (G). The entire process is guided by a loss function that measures the difference between the output image (X) and the input image (Y) while considering the semantic features from the guidance image (G). The loss function's role is to guide the training of the network so that the output image retains the semantic coherence and structure of the guidance image. In summary, this network architecture combines the high-level feature extraction capabilities of an encoder-decoder network with the refinement and reconstruction capabilities of a deep

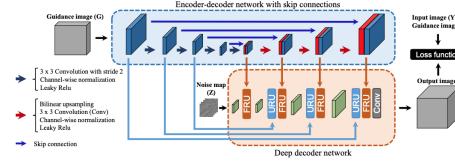


Fig. 2: Illustration of the GDD model architecture

decoder network, all while being guided by another input (the guidance image) to ensure that the output image aligns with the desired semantic features.

**Upsampling Refinement Unit (URU):** Upsampling: The purpose of upsampling is to increase the resolution of feature maps. In deep learning, this is often done using methods like bilinear interpolation, which can lead to smooth transitions but might miss finer details.

Refinement: After upsampling, the features may lack the necessary detail and sharpness. Refinement aims to recover the spatial details that are crucial for reconstructing high-quality images. The URU incorporates attention gates that apply weights to the upsampled features. These gates learn to focus on important features and help recover finer details that might have been lost during upsampling. Conditional Weights: The attention mechanism produces conditional weights that allow the network to emphasize features that are important for the current task, such as edges or textures.

**Feature Refinement Unit (FRU):** Feature Weighting: In FRU, the upsampled features from the deep decoder are combined with semantic features from the guidance image. High-Level Semantic Features: These are the features that encode more abstract representations of the content of the guidance image. They are crucial for ensuring that the output of the deep decoder matches the semantic content of the guidance image.

The FRU uses semantic features from the guidance image at multiple scales to refine the features at each level of the decoder. This helps the network align the output image semantically with the guidance image. The URU and FRU components of the GDD serve to refine the upsampled features through a combination of attention mechanisms and multi-scale semantic guidance from another image. This results in output that not only has high spatial resolution but also maintains semantic coherence with the guidance image, which is essential for tasks like super-resolution and image fusion.

#### B. Loss function

The proposed method effectively incorporates spatial details and semantic features from guidance images, making it a powerful regularizer.

Let  $X$ ,  $Y$ , and  $G$  be matrices in various forms, denoted as  $\hat{X} \in \mathbb{R}^{C \times W \times H}$ ,  $\hat{Y} \in \mathbb{R}^{C \times w \times h}$ , and  $\hat{G} \in \mathbb{R}^{C \times W \times H}$  respectively. Given the estimated HR-HS  $\hat{X}$ , the loss function can be defined as:

$$\mathcal{L}(X, \hat{Y}, G) = \mu \| \mathbf{X}S - \hat{Y} \|_F^2 + \| R\mathbf{X} - \hat{G} \|_F^2 \quad (11)$$

where  $\| \cdot \|_F$  is the Frobenius norm,  $S$  is the spatial downscaling with blurring, and  $R$  is the spectral response function

that integrates the spectra into  $R$ ,  $G$ ,  $B$  channels. The first term encourages the spectral similarity between the spatially downsampled  $\hat{X}$  and  $\hat{Y}$ . The second term encourages the spatial similarity between the spectrally downsampled  $\hat{X}$  and  $\hat{G}$ .  $\mu$  is a scalar controlling the balance between the two terms.

#### IV. DATASET

The dataset of imagery from the Mars Reconnaissance Orbiter was retrieved via the Nasa Planetary Data System. The database contains a wide variety of raw and processed imaging data. For the hyperspectral images, we used the CRISM Map-Projected Targeted Reduced Data Record (MTRDR), which contains over 10,000 processed hyperspectral images that are projected to account for distortions and registered to specific coordinates on Mars. We also accessed the CTX Experimental Data Record (EDR) for the high resolution panchromatic images.

The panchromatic images and the hyperspectral images cover the same locations on the martian surface. In general, the panchromatic images cover a much larger area than the hyperspectral images and have around 10x the spatial pixel resolution. For this reason, both images were registered and cropped so that they displayed the same exact content with different modalities. For simplicity, the images were cropped into squares with pixel values equal to a power of two. The registration was performed using coordinate values and pixel scales available in the metadata of the images. The process of registration and cropping is illustrated in Figure 3. After this procedure, the panchromatic image had 4x the spatial resolution as the hyperspectral image. The images are now prepared to perform fusion.

#### V. RESULTS

Figure 4 shows the visual comparison between the MS image (left) and the fused image (right) generated by the deep-decoder. The model was trained using MSE loss for 1000 iterations with learning rate = 0.01 and utilizing Adam optimizer. Figure 5 depicts the loss curve over 1000 iterations. For each of the 1024 pixels in the downsampled HS image ( $32 \times 32 \times 489$ ), a correlation is calculated between the spectral profile of that pixel and the average spectral profile of the corresponding  $32 \times 32$  block in the fused high-resolution image (Figure 6). The correlation computation for each of these 1024 points (each representing a pixel in the HS image) involves all 489 spectral bands. This means for each of the 1024 locations, we get a single correlation coefficient that summarizes the similarity across all 489 bands between the HS image pixel and its corresponding block in the fused image. In Figure 7, we see spectral intensities across different bands between HS image and fused image.

The results of NMF fusion are shown in Figure 9 (on the last page), which illustrates a few representative examples from different locations on Mars. Qualitatively, it can be seen that the fused images have the same spatial resolution as the panchromatic image while retaining the spectral information of the hyperspectral image, represented as the projection of the spectral bands with the highest variance onto RGB values.

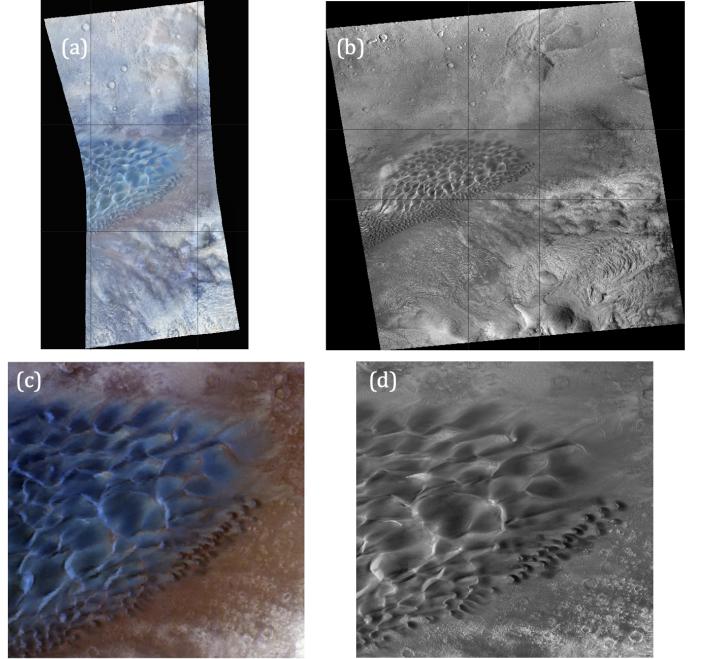


Fig. 3: (a) Uncropped hyperspectral image with cropping region highlighted, (b) Uncropped panchromatic image with cropping region highlighted, (c) Cropped hyperspectral image, (d) cropped panchromatic image

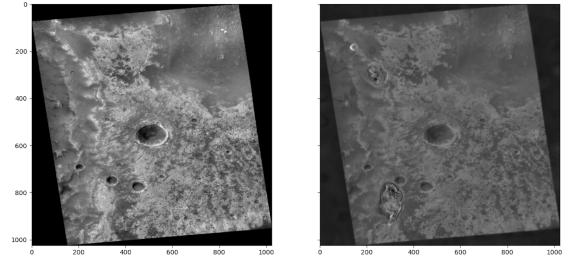


Fig. 4: Image fusion with deep-decoder

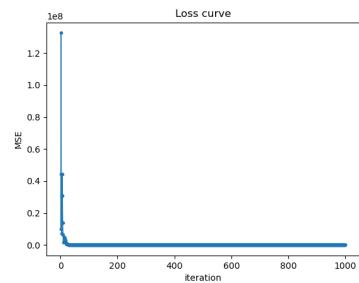


Fig. 5: MSE loss curve in deep-decoder

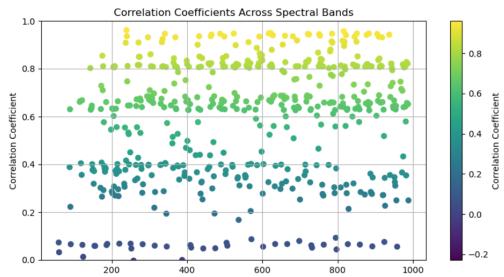


Fig. 6: Correlation across pixel locations

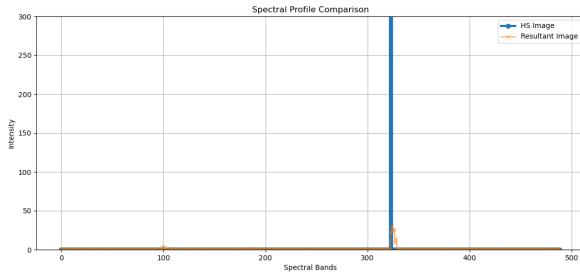


Fig. 7: Spectral profile alignment across all spectral bands

Some images exhibited artifacts after fusion, such as the band seen in the fused image in the last row of Figure 9. This can largely be attributed to the color projection algorithm.

## VI. DISCUSSION

The loss curve (Figure 5) shows a sharp decrease in the Mean Squared Error (MSE) value and stabilizes quickly within the first few iterations. The rapid convergence suggests that the model quickly finds a local minimum for the loss function. However, a fast convergence can also indicate that the model may be settling too quickly into a suboptimal solution without fully capturing the complex relationships or fine details in the data. In the correlation coefficients across spectral bands (Figure 6), the variability in correlation across different bands indicates inconsistency in the model's ability to preserve spectral fidelity across all bands. High correlation in some

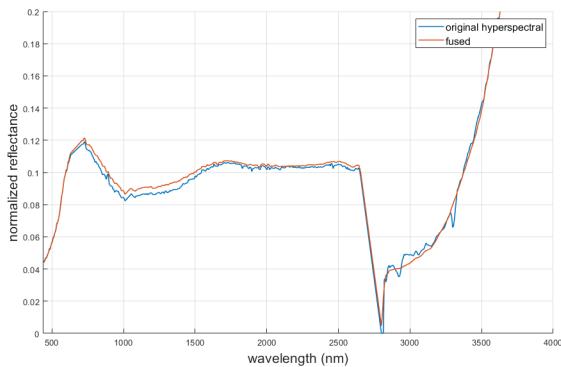


Fig. 8: Comparison of the spectrum at the same location of original hyperspectral image with NMF fused image

bands (yellow and green dots) suggests good spectral matching in these specific regions. However, many bands show moderate to low correlation (blue dots), indicating a loss of spectral integrity in these areas. This could be due to the model's inability to accurately reconstruct the spectral signature at those bands. The left image in the pair is the multispectral (MS) image (high spatial resolution but low spectral resolution), and the right image is the resultant fused image (Figure 4). Visual inspection suggests that the fused image lacks some clarity and detail compared to the MS image. This could be a result of the model's quick convergence, where it may not have sufficiently explored the parameter space to find a truly optimal solution that balances both spectral and spatial fidelity. High correlation areas might correspond to spectral bands where the characteristics of the original and fused images are similar, possibly due to the model effectively learning these specific band patterns. Low correlation areas may have occurred due to several reasons: The model might not have enough capacity to capture complex spectral signatures accurately across all bands. Rapid convergence might indicate overfitting to some spectral features while underfitting others. The optimization strategy (e.g., learning rate, iterations) might need tuning to allow the model more time to explore the parameter space thoroughly. In Figure 7 blue line in the HS image represents the spectral intensity values across different spectral bands for a specific point or region, i.e., location [10, 10] in the original HS image. Orange line in the resultant image shows the corresponding spectral intensities at the same point in the image that resulted from the fusion process. We noticed the significant peak around band 300 in both the HS and resultant images, indicating that the fusion process preserved key spectral features at this band. Across most bands, both images show very low intensities except for the sharp peak, suggesting that the fusion process primarily focused on major spectral features. For point [10, 10], we see that while there is a peak in both profiles at the same band, overall areas might not align perfectly, explaining the average correlation of 0.44.

Although the loss stabilized quickly, running for more iterations at a lower learning rate might help explore more subtle adjustments that could improve reconstruction. Incorporating regularization methods could prevent the model from converging too quickly to suboptimal solutions. Modifying the loss function to penalize spectral discrepancies more heavily might encourage better spectral alignment. Experimenting with different model architectures, learning rates, and other hyperparameters could lead to better overall performance.

Figure 8 shows a comparison of a selected band from the same location in the original hyperspectral image and the NMF fused image. The spectra were relatively similar, exhibiting a mean squared error of  $1.25 \times 10^{-5}$  and a PSNR of 49 dB across the bandwidth. It should be noted that the fused image was to some degree "smoothed out" compared to the original hyperspectral image. This is especially clear in the wavelengths above around 2800 nm, where fluctuations in this region are significantly dampened. However, data from these longer wavelengths is rarely used in practice due to its susceptibility to thermal variations, and the shorter wavelengths exhibit this effect to a lesser degree. This overall smoothing

effect can be attributed to the implicit interpolation property of the NMF process.

## VII. CONCLUSIONS

The implementation of nonnegative matrix factorization (NMF) and guided deep decoder (GDD) for hyperspectral image fusion using MRO data represents a significant step forward in Mars surface analysis capabilities. Our experiments have demonstrated that these methods can effectively combine spectral and spatial data. The guided deep decoder showed potential but requires further optimization for spectral fidelity and detail preservation. Challenges such as rapid model convergence indicate a need for better hyperparameter tuning and possibly more sophisticated optimization strategies to enhance both spectral and spatial reconstruction accuracy.

Additional work might include the implementation of a simple material segmentation task on both the hyperspectral image and its fused counterpart, highlighting some practical applications of these tools. The CRISM hyperspectral images cover only small segments of the Martian surface, whereas the panchromatic images cover nearly the entire planet. Therefore, more advanced directions could include a trained model capable of predicting the spectral information across all panchromatic images, effectively creating a high-resolution hyperspectral map of the entire planet.

## REFERENCES

- [1] T. Uezato, D. Hong, N. Yokoya, and W. He, "Guided deep decoder: Unsupervised image pair fusion," 2020.
- [2] CRISM web site. Publisher: JHU/APL. [Online]. Available: <http://crism.jhuapl.edu/instrument/design/overview.php>
- [3] Improving component substitution pansharpening through multivariate regression of MS -pan data | IEEE journals & magazine | IEEE xplore. [Online]. Available: <https://ieeexplore.ieee.org/document/4305344>
- [4] R. Nishii, S. Kusanobu, and S. Tanaka, "Enhancement of low spatial resolution image based on high resolution-bands," vol. 34, no. 5, pp. 1151–1158, conference Name: IEEE Transactions on Geoscience and Remote Sensing. [Online]. Available: <https://ieeexplore.ieee.org/document/536531>
- [5] N. Yokoya, "Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion," vol. 50, no. 2.
- [6] X. Liu, W. Xia, B. Wang, and L. Zhang, "An approach based on constrained nonnegative matrix factorization to unmix hyperspectral data," vol. 49, no. 2, pp. 757–772. [Online]. Available: <http://ieeexplore.ieee.org/document/5585746/>
- [7] Z. Wang, B. Chen, H. Zhang, and H. Liu, "Unsupervised hyperspectral and multispectral images fusion based on nonlinear variational probabilistic generative model," vol. 33, no. 2, pp. 721–735. [Online]. Available: <https://ieeexplore.ieee.org/document/9246732/>
- [8] S. Shi, L. Zhang, and J. Chen, "Hyperspectral and multispectral image fusion using the conditional denoising diffusion probabilistic model." [Online]. Available: <http://arxiv.org/abs/2307.03423>
- [9] J. Li, K. Zheng, J. Yao, L. Gao, and D. Hong, "Deep unsupervised blind hyperspectral and multispectral data fusion," vol. 19, pp. 1–5. [Online]. Available: <https://ieeexplore.ieee.org/document/9714360/>
- [10] N. Keshava and J. Mustard, "Spectral unmixing," vol. 19, no. 1, pp. 44–57. [Online]. Available: <http://ieeexplore.ieee.org/document/974727/>
- [11] C. Lanaras, E. Baltsavias, and K. Schindler, "Hyperspectral super-resolution by coupled spectral unmixing," in *2015 IEEE International Conference on Computer Vision (ICCV)*. IEEE, pp. 3586–3594. [Online]. Available: <http://ieeexplore.ieee.org/document/7410766/>
- [12] M. Simoes, J. Bioucas-Dias, L. B. Almeida, and J. Chanussot, "A convex formulation for hyperspectral image superresolution via subspace-based regularization," vol. 53, no. 6, pp. 3373–3388. [Online]. Available: <http://ieeexplore.ieee.org/document/7000523/>
- [13] L. Condat, "Fast projection onto the simplex and the  $l_1$  ball," vol. 158, no. 1, pp. 575–585. [Online]. Available: <https://doi.org/10.1007/s10107-015-0946-6>

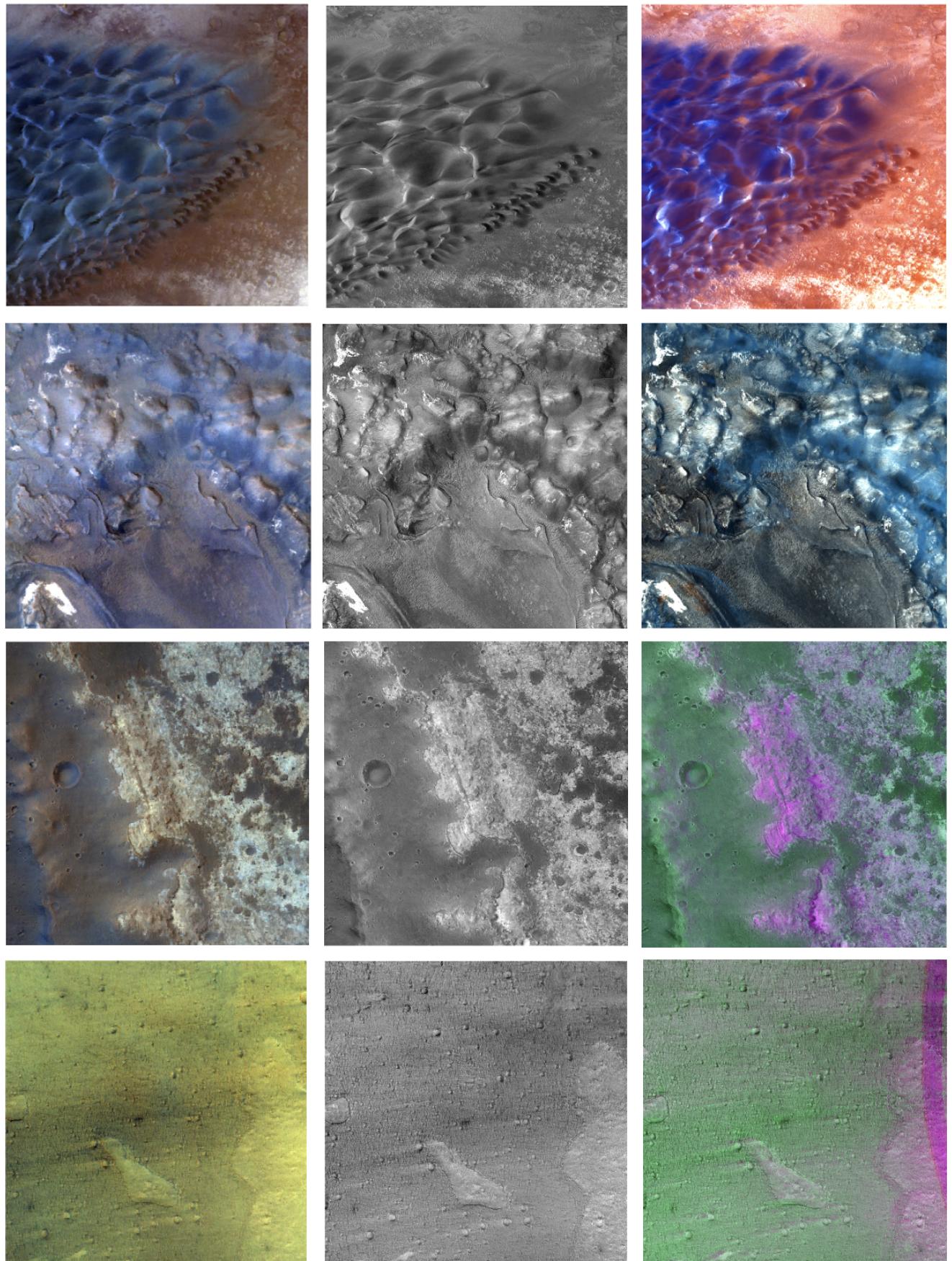


Fig. 9: Four examples of image fusion using NMF. The left column is the hyperspectral image, the middle column is the panchromatic image, and the right column is the fused image.