Data fidelity-oriented spatial-spectral fusion of CRISM and CTX images

Qunming Wang [a, *], Wenjing Ma [a], Sicong Liu [a], Xiaohua Tong [a], Peter M. Atkinson [b, c]

[a] College of Surveying and Geo-Informatics, Tongji University, 1239 Siping Road, Shanghai 200092, China

[b] Faculty of Science and Technology, Lancaster University, Lancaster LA1 4YR, UK

[c] Geography and Environment, University of Southampton, Highfield, Southampton SO17 1BJ, UK

*Corresponding author. E-mail: wqm11111@126.com.

Abstract: The Compact Reconnaissance Imaging Spectrometer for Mars (CRISM) is a Mars-dedicated compact reconnaissance imaging spectrometer that captures remote sensing data with very fine spectral resolution. However, the spatial resolution of CRISM data is relatively coarse (18 m), limiting its application to regional scales. The Context Camera (CTX) is a digital camera equipped with a wide-angle lens, providing a finer spatial resolution (6 m) and larger field-of-view, but CTX provides only a single panchromatic band. To produce CRISM hyperspectral data with finer spatial resolution (e.g., 6 m of CTX images), this research investigated spatial-spectral fusion of 18 m CRISM images with 6 m CTX panchromatic images. In spatial-spectral fusion, to address the long-standing issue of incomplete data fidelity to the original hyperspectral data in existing methods, a new paradigm called Data Fidelity-oriented Spatial-Spectral Fusion (DF-SSF) was proposed. The effectiveness of DF-SSF was validated through experiments on data from six areas on Mars. The results indicate that the fusion of CRISM and CTX can increase the spatial resolution of CRISM hyperspectral data effectively. Moreover, DF-SSF can increase the fusion accuracy noticeably while maintaining perfect data fidelity to the original hyperspectral data. In addition, DF-SSF is theoretically applicable to any existing spatial-spectral fusion methods. The 6 m CRISM hyperspectral data inherit the advantages of the original 18 m data in spectral resolution, and provide richer spatial texture information on the Martian surface, with broad application potential.

Keywords: Compact Reconnaissance Imaging Spectrometer for Mars (CRISM), Context Camera (CTX), downscaling, spatial-spectral fusion, data fidelity, area-to-point kriging (ATPK).

1. Introduction

As a neighboring planet in the solar system, Mars has always been a focus of human fascination and, more recently, an emerging target for human exploration. Its known, abundant resources, based on our gradually deepening understanding of its geology, landform and atmosphere, may be important for the future development of human society. With advances in space technology, Mars exploration is increasingly seen as an important pathway for space resources development and scientific technological innovation.

As an advanced imaging technology, hyperspectral remote sensing provides data with a very fine spectral resolution for observing the surface of Mars, offering unique opportunities for a deeper understanding of the geology and environment of Mars. The main advantage of hyperspectral images over multispectral images lies in the richer spectral information, which is crucial for studying the mineral composition and formation mechanisms on the surface of Mars. The Mars Reconnaissance Orbiter (MRO) was launched on August 12, 2005 (Zurek and Smrekar, 2007) to investigate the geology and climate of Mars. Its scientific objectives include observing the current climate of Mars, searching for water activity, mapping surface features on Mars, and studying potential future landing sites. The MRO carries several instruments, including the Compact Reconnaissance Imaging Spectrometer for Mars (CRISM) (Murchie et al., 2007) and the Context Camera (CTX) (Malin et al., 2007). These two sensors capture information about the same area with different spectral and spatial resolutions simultaneously. Specifically, CRISM is operated in hyperspectral mode, acquiring hyperspectral images covering more than four hundred spectral bands from the visible to near-infrared wavelengths. However, its spatial resolution is 18 m, which is relatively coarse for observing detailed spatial texture information in local areas. The primary function of CTX is to provide background information for other MRO instruments through simultaneous observations. The coverage of the images captured by CTX is larger than that of CRISM and the spatial resolution is about 6 m, which is three times finer than for CRISM (Malin et al., 2007). However, these images are single-band and do not provide spectral information. This study proposed to downscale the 18 m CRISM hyperspectral images to 6 m, with the aid of the 6 m fine spatial resolution CTX images. The 6 m hyperspectral images can potentially provide more spatial details about the surface of Mars while preserving the fine spectral resolution of CRISM.

Spatial-spectral fusion (also known as pan-sharpening) aims to fuse images with fine spectral resolution, but coarse spatial resolution with images with fine spatial resolution, but coarse spectral resolution (e.g., panchromatic image) in the same region to create images with both fine spatial and spectral resolutions (Zhang and Shen, 2016). It can resolve the trade-off that occurs between spatial and spectral resolution when designing a single sensor. The existing spatial-spectral fusion methods include component substitution (CS)-based, multi-resolution analysis (MRA)-based, variational optimization (VO)-based, matrix factorization, learning-based, and geostatistical approaches (Ghamisi et al., 2019; Loncan et al., 2015; Thomas et al., 2008; Vivone et al., 2014; Yang et al., 2022). The main idea of the CS methods is to transform the multi/hyperspectral images into another space, and utilize the panchromatic image to substitute the transformed coarse spatial resolution component (Aiazzi et al., 2007; Thomas et al., 2008). Representative methods include principal component analysis (PCA) (Shah et al., 2008; Shettigara, 1992), Gram-Schmidt transformation (GS) (Laben and Brower, 2000), Gram-Schmidt adaptive (GSA) (Aiazzi et al., 2007). The core idea of MRA is to extract fine spatial resolution details from the panchromatic image and inject them into the coarse spatial resolution multi/hyperspectral images (Chavez et al., 1991). Representative methods include smoothing filter-based intensity modulation (SFIM) (Liu, 2000), generalized Laplacian pyramid (GLP) with modulation transfer function (MTF)-matched filter (MTF-GLP) (Aiazzi et al., 2006), and GLP with MTF-matched filter and multiplicative injection model (MTF-GLP-HPM) (Lee and Lee, 2009). The VO-based methods construct a variational optimization model to take full advantage of the spatial information of panchromatic image and spectral information of coarse spatial resolution multispectral image. For example, the variational approach developed by Fang et al. (2013) consists of three terms, which aim to minimize the difference in spatial gradients between the panchromatic image and fused image, the difference between original coarse multispectral image and (degraded) fused image, and the difference in spectral gradients between original coarse multispectral image and fused image. The matrix factorization methods were proposed from the perspective of spectral unmixing. A representative method in this category is coupled nonnegative matrix factorization (CNMF) (Berne et al., 2010; Yokoya et al., 2012), which extracts endmembers from coarse spatial, but fine spectral resolution image and proportions from fine spatial, but coarse spectral resolution image. The CNMF prediction is the linear combination of the proportions and endmembers. The learning-based methods focus on establishing a nonlinear mapping relationship between the fine spatial resolution panchromatic images and coarse spatial resolution multi/hyperspectral images or learning the intrinsic structure of the data observed. Deep learning has become a common choice for spatial-spectral fusion due to its strong fitting ability. Early attempts are mainly supervised methods, which require multi/hyperspectral images at the target fine spatial resolution. Examples for this type of methods include the pan-sharpening neural network (PNN) (Masi et al., 2016) and a deep network architecture for pansharpening (PanNet) (Song et al., 2018; Yang et al., 2017, 2018). Recently, more advanced versions were developed, such as domain transform model driven by deep learning (Sun et al., 2024) and progressive multi-iteration registration-fusion co-optimization network (Qu et al., 2024). For supervised methods, however, the applicability in reality can be compromised, as it can be difficult to collect required fine spatial resolution multi/hyperspectral images for training. Alternatively, unsupervised deep learning strategies have been developed for spatial-spectral fusion in recent years (Ma et al., 2020; Qu et al., 2023; Sun et al., 2023), which do not need multi/hyperspectral images at the target fine spatial resolution for training, but learn the intrinsic structure from the input data. Specifically, the input panchromatic image and coarse multispectral image are used in the construction of loss factions to constrain the predictions of the networks. Spatially, it is assumed that the spatial information presented by the (spectrally degraded) fused image should be similar to that of the panchromatic image. Spectrally, it is assumed that the (spatially degraded) fused image should be the same as the input multispectral image. The geostatistical approaches can take into account the changes in the spatial support of the data and the point spread function (PSF) effect of sensors (Atkinson et al., 2008; Wang et al., 2014). The most significant advantage of geostatistical methods is that the downscaling results are consistent with the original coarse spatial resolution data. Representative methods include area-to-point regression kriging (ATPRK) (Wang et al., 2016), downscaling cokriging (DSCK) (Pardo-Igúzquiza et al., 2006), and kriging with an external drift (KED) (Sales et al., 2012).

Significant progress has been made in research on spatial-spectral fusion based on various Earth observation datasets. However, research on fusion of datasets on other planets is relatively limited. To the best of our knowledge, there has been no study on the fusion of CTX and CRISM images for Mars. The surface of Mars is generally covered by rocks and minerals with various terrain and landform features, which is substantially different from that for the Earth surface. The main advantage of hyperspectral image over multispectral image is that the former can provide more detailed spectral information, which significantly enhances the ability to distinguish between different types of rocks and minerals. In this case, the effectiveness of existing spatial-spectral fusion methods needs to be validated. Most current spatial-spectral fusion methods suffer from a long-standing issue: they cannot achieve data fidelity of the original coarse spatial resolution images. That is, when the spatial-spectral fusion results are degraded to the original coarse spatial resolution, they are not consistent with the original coarse image. To realize data fidelity is crucial for hyperspectral sharpening, as the reliability of spectra plays key role in downstream applications such as rock and mineral identification.

Geostatistics provides a new solution for realization of spatial-spectral fusion with perfect data fidelity to the original coarse spatial resolution images. Specifically, various models based on area-to-point kriging (ATPK) (Atkinson, 2013; Kyriakidis, 2004; Kyriakidis and Yoo, 2005; Wang et al., 2015) including DSCK, KED, and ATPRK, have significant advantages in preserving the original coarse image. That is, when the spatial-spectral fusion result is degraded to the original spatial resolution, it is completely consistent with the original image. This characteristic of complete data fidelity is also referred to complete data coherence in geostatistics.

This paper proposed a completely new paradigm of spatial-spectral fusion called DF-SSF (Data Fidelity-oriented Spatial-Spectral Fusion) for fusion of CTX and CRISM images, where the CTX image is treated as the panchromatic image. It aims to preserve perfectly the original CRISM data. DF-SSF achieves this by utilizing ATPK to downscale the difference (i.e., coarse residual image) between the existing fusion result (i.e., produced by any of the existing spatial-spectral fusion methods) and the original coarse spatial resolution image. The produced fine spatial resolution residual image is then added to the existing fusion result to obtain the final fusion result with complete data fidelity. The contributions of this paper mainly lie in two aspects.

1) It is the first study on the fusion of CTX and CRISM images on Mars to enhance the spatial resolution of CRISM hyperspectral images. Although spatial-spectral fusion has been developed over decades, to the best of our knowledge, there is no research on Mars remote sensing data dominated by rocks and minerals, especially for fusion of CTX and CRISM data, where the preservation of original spectra is crucial for downstream applications.
2) A completely new paradigm DF-SSF, is proposed for spatial-spectral fusion with complete data fidelity to the original coarse spatial resolution data. DF-SSF is theoretically applicable to any existing spatial-spectral fusion method.

The remainder of this paper is divided into five sections. Section 2 provides a brief introduction to the study area, research data, and the corresponding processing. Section 3 explains the principles of the proposed DF-SSF method. The experimental results for validation of the method are presented in Section 4. Section 5 discusses issues related to the proposed method and outlines future directions. Finally, Section 6 concludes the paper.

2. Study area and data

2.1. Study area

The study area of this paper consists of six regions, as shown in Table 1 (Bennett et al., 2014). Among them, Eberswalde Crater, Mawrth Vallis, and Holden Crater (Poulet et al., 2014) were once considered as candidate landing sites for Mars missions. The Melas Chasma area is a canyon on the Martian surface, while the Gale Crater (Peulvast and Masson, 1993; Poulet et al., 2014) was the location where the Curiosity Rover landed successfully in 2012 and is currently considered for scientific exploration. The Jezero Crater (Goudge et al., 2015) was the landing site of the Perseverance Rover, which landed successfully in 2021. These six regions have become the focus of scientific research due to their unique geological and geomorphological characteristics. Utilizing the hyperspectral remote sensing data covering these areas, it is possible to study Martian rocks, minerals and many other topics. The abundance, distribution, and properties of these rocks and minerals can provide insights into the composition and evolutionary processes of the Martian surface. In-depth study of these regions can help us understand the geological history and hydrological processes of Mars more comprehensively.

2.2. CRISM images

All CRISM images used in this paper are Map-Projected Targeted Reduced Data Record (MTRDR) (Murchie et al., 2007) products obtained at the latest, publicly available calibration level. The MTRDR includes map-projected Targeted Empirical Record corrected (TER) calibrated I/F (the ratio of the sensor's spectral irradiance to the solar spectral irradiance) spectral information and excludes spectral channels with questionable radiance measurements ("bad bands"). The hyperspectral image cubes in the MTRDR product suite are stored as 32-bit real number units. As a hyperspectral sensor, CRISM covers the spectral range from 362 nm to 3920 nm with a fine spectral resolution of 6.55 nm per channel. This range includes the visible, near-infrared and shortwave infrared wavelengths. Through these channels, CRISM obtains spectral information from the Martian surface, which can be used for studying Martian mineralogy, geology and geomorphology. The CRISM data used in this study are MTRDR products with a spatial resolution of about 18 m in this hyperspectral mode.

2.3. CTX images

The CTX images used in this paper are Experiment Data Record (EDR) (Malin et al., 2007) products containing raw CTX images along with their associated metadata information (e.g., observation time, exposure time, camera parameters, etc.). The raw EDR data are convenient for further processing, analysis and utilization. Unlike multispectral images, CTX images contain only a single-band, with a spectral range from 500 nm to 700 nm. On the MRO's near-circular, near-polar mapping orbit, the spatial resolution of CTX is about 6 m. This relatively fine spatial resolution enables CTX to capture more detailed surface features on Mars, including impact craters, canyons, dunes, and some other intricate information.

2.4. Data processing

The CRISM MTRDR images were used as originally provided, while CTX images were processed using the ISIS3 (Sucharski et al., 2020) pipeline. The raw CTX image files were first converted to ISIS3 image format and then subjected to image map projection (Equidistant cylindrical). Since precise alignment between the input hyperspectral and panchromatic images is essential for spatial-spectral fusion, image registration was also performed. Geometric registration between the CRISM and CTX images was achieved by using rasterio's virtual warping to reproject the CRISM images into the coordinate reference system of the corresponding CTX images. However, even within the same coordinate system, the CRISM and CTX images may not be perfectly aligned. Therefore, subsequent registration is necessary. In this paper, the software ENVI was employed for local registration through the Harris corner detection algorithm. The final 6 m CTX and 18 m CRISM images of the study area are shown in Fig. 1. The spatial sizes of CTX and CRISM images used in each region are 900×900 and 300×300 pixels, respectively.

Table 1. Locations of the six study regions.

Regions	Latitude (°)		Longitude (°)	
	Min	Max	Min	Max
Eberswalde crater	-24.5	-23.4	-34.0	-32.7
Mawrth Vallis	23.3	24.6	-19.6	-18.4
Holden crater	-27.6	-25.9	-36.0	-34.0
Melas Chasma	-10.5	-8.7	-78.0	-75.0
Jezero crater	18.0	18.8	77.2	78.4
Gale crater	-0.9	-3.8	135.9	139.9

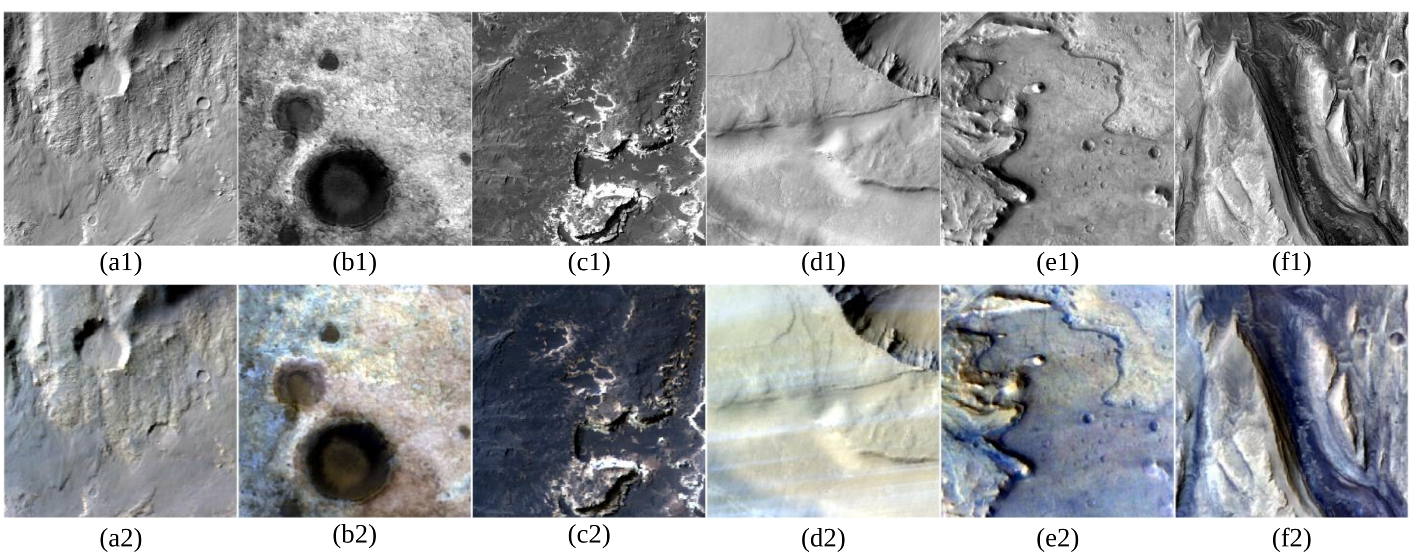


Fig. 1. The CTX (6 m) and CRISM (18 m) images of the six regions (bands 37, 25 and 12 as RGB). The spatial sizes of the CTX and CRISM images are 900×900 and 300×300, respectively. (a) Eberswalde Crater. (b) Mawrth Vallis. (c) Holden Crater. (d) Melas Chasma. (e) Jezero Crater. (f) Gale Crater.

3. Methods

3.1. Overview of DF-SSF

The main requirement of the Wald protocol I (Wald et al., 1997) in spatial-spectral fusion is that there should be no deviation between the spatial-spectral fusion result and the original coarse image. That is, when the fused image is degraded to the original coarse spatial resolution, it should be exactly the same as the observed coarse image. However, the reality is that existing spatial-spectral fusion models struggle to achieve data fidelity in the true mathematical sense, meaning that the residuals (i.e., the difference between the fusion result and the original coarse data) are commonly non-zero. Achieving data fidelity to the original coarse image plays a crucial role in enhancing the reliability of the fused data. Therefore, this paper proposes a completely new spatial-spectral fusion paradigm (DF-SSF) that can realize complete data fidelity to the original coarse image:

$$\widehat{\mathbf{HS}}_k = \widehat{\mathbf{HS}}_k^{'} + \Delta\mathbf{R}_k,\ \ k = 1,2,\ldots,N \tag{1}$$

where k denotes the result of the k-th band (k=1,2,…,N, where N is the total number of bands), $\widehat{\mathbf{HS}}_k$ is the spatial-spectral fusion result that enables complete data fidelity, $\widehat{\mathbf{HS}}_k^{'}$ is the prediction of any existing spatial-spectral fusion model, and $\Delta\mathbf{R}_k$ represents the fine spatial resolution residuals present in the existing spatial-spectral fusion model. Details for calculation of $\widehat{\mathbf{HS}}_k^{'}$ and $\Delta\mathbf{R}_k$ are introduced in the following Sections 3.2 and 3.3, respectively. The whole flowchart of the proposed DF-SSF method is sketched in Fig. 2.

3.2. Estimation based on existing SSF methods

The widely used spatial-spectral fusion algorithms generally follow the basic principles of CS or MRA. Therefore these two categories of methods are the focus of this paper. Their main principles are briefly described as follows.

For the CS-based methods, the spatial-spectral fusion result is defined as:

$$\widehat{\mathbf{HS}}_k^{'} = \widetilde{\mathbf{HS}}_k + g_k^{\mathrm{C}}(\mathbf{P} - \mathbf{I}_L),\ \ k = 1,2,\ldots,N \tag{2}$$

in which $\widehat{\mathbf{HS}}_k^{'}$ denotes the fusion result for the k-th band, $\widetilde{\mathbf{HS}}_k$ denotes the multi/hyperspectral image interpolated to the spatial size of the panchromatic image, g_k^{C} is the weight of the fine spatial resolution gain injected into the k-th band (where C represents the CS-based methods), and $\mathbf{P}$ denotes the panchromatic image. $\mathbf{I}_L$ is defined as:

$$\mathbf{I}_L = \sum_{i=1}^{N} \omega_i \widetilde{\mathbf{HS}}_i \tag{3}$$

in which ω_i represents the fitting weight for the i-th (i=1,2,…,N) band of the multi/hyperspectral image.

For the MRA-based method, the spatial-spectral fusion result is defined as:

$$\widehat{\mathbf{HS}}_k' = \widetilde{\mathbf{HS}}_k + g_k^{\mathrm{M}}(\mathbf{P} - \mathbf{P}_L),\ \ k = 1,2,\ldots,N \tag{4}$$

where $\mathbf{P}_L$ represents the coarse spatial resolution version of the image $\mathbf{P}$, and g_k^{M} is the weight of the fine spatial resolution gain for the k-th band (where M represents the MRA-based methods).

3.3. Estimation of the residuals at fine spatial resolution

For existing spatial-spectral fusion models, it is inevitable that there are coarse spatial resolution residuals (denoted as $\Delta\mathbf{R}_k^C$) in their predictions, as defined as follow:

$$\Delta\mathbf{R}_k^C = \mathbf{HS}_k - \widehat{\mathbf{HS}}_k'\uparrow \tag{5}$$

where $\mathbf{HS}_k$ represents the observed coarse spatial resolution image of the k-th band, $\uparrow$ denotes the degradation operation, and $\widehat{\mathbf{HS}}_k'\uparrow$ indicates the result of degrading the spatial-spectral fusion result of the k-th band to the original coarse spatial resolution.

In this paper, we utilized ATPK to estimate the spatial resolution fine residuals $\Delta\mathbf{R}_k$ in Eq. (1). Specifically, for the residual of the fine spatial resolution pixel at spatial position $\mathbf{x}$ in the k-th band (denoted as $\Delta\mathbf{R}_k(\mathbf{x})$), its value can be predicted through a linear combination of L spatially adjacent coarse residuals in image $\Delta\mathbf{R}_k^C$:

$$\Delta\mathbf{R}_k(\mathbf{x}) = \sum_{i=1}^{L} \beta_i \Delta\mathbf{R}_k^C(\mathbf{x}_i),\ \ s.t. \sum_{i=1}^{L} \beta_i = 1 \tag{6}$$

where $\mathbf{x}_i$ denotes the spatial location of the i-th neighborhood, β_i denotes the weight of its corresponding coarse residual, and L is the number of neighboring coarse pixels used in the prediction. The L weights (i.e., β_1, β_2,..., β_L) in Eq. (6) are calculated by the kriging equation as follows:

$$\begin{bmatrix} \gamma_{CC}^{k}(\mathbf{x}_1,\mathbf{x}_1) & \cdots & \gamma_{CC}^{k}(\mathbf{x}_1,\mathbf{x}_L) & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \gamma_{CC}^{k}(\mathbf{x}_L,\mathbf{x}_1) & \cdots & \gamma_{CC}^{k}(\mathbf{x}_L,\mathbf{x}_L) & 1 \\ 1 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_L \\ \theta \end{bmatrix} = \begin{bmatrix} \gamma_{FC}^{k}(\mathbf{x},\mathbf{x}_1) \\ \vdots \\ \gamma_{FC}^{k}(\mathbf{x},\mathbf{x}_L) \\ 1 \end{bmatrix} \tag{7}$$

where the term $\gamma_{CC}^{k}(\mathbf{x}_m,\mathbf{x}_n)$ (m, n=1,2,…,L) is the coarse-to-coarse residual semivariogram between coarse pixels centered at $\mathbf{x}_m$ and $\mathbf{x}_n$ in the k-th band, $\gamma_{FC}^{k}(\mathbf{x},\mathbf{x}_m)$ is the fine-to-coarse residual semivariogram between fine and coarse pixels centered at $\mathbf{x}$ and $\mathbf{x}_m$ in the k-th band, and θ is the Lagrange multiplier. Further details on calculation of the semivariograms can be found in Wang et al. (2015, 2016).

3.4. Perfect data fidelity of DF-SSF

An important advantage of ATPK is the perfect coherence of the prediction with the input coarse image. That is, once the ATPK prediction is degraded to the original coarse spatial resolution, it is exactly identical to the original coarse data. Based on the perfect coherence of ATPK, the coarse residuals $\Delta\mathbf{R}_k^C$ can be accurately reproduced when the ATPK predictions $\Delta\mathbf{R}_k$ are degraded to the coarse spatial resolution as follows:

$$\Delta\mathbf{R}_k\uparrow = \Delta\mathbf{R}_k^C \tag{8}$$

As described in Eq. (1), the prediction of DF-SSF is composed of $\widehat{\mathbf{HS}}_k^{'}$ predicted by any of the existing spatial-spectral fusion methods and $\Delta\mathbf{R}_k$ predicted by ATPK. Combining Eqs. (1), (5) and (8), we can derive the following:

$$\begin{aligned} \widehat{\mathbf{HS}}_k\uparrow &= \left(\widehat{\mathbf{HS}}_k^{'} + \Delta\mathbf{R}_k\right)\uparrow \\ &= \widehat{\mathbf{HS}}_k^{'}\uparrow + \Delta\mathbf{R}_k\uparrow \\ &= \widehat{\mathbf{HS}}_k^{'}\uparrow + \Delta\mathbf{R}_k^C \\ &= \widehat{\mathbf{HS}}_k^{'}\uparrow + \mathbf{HS}_k - \widehat{\mathbf{HS}}_k^{'}\uparrow \\ &= \mathbf{HS}_k \end{aligned} \tag{9}$$

Eq. (9) means that once the prediction of DF-SSF (i.e., $\widehat{\mathbf{HS}}_k$) is degraded to the original coarse spatial resolution, it is exactly identical to the original coarse image (i.e., $\mathbf{HS}_k$), thereby, achieving complete data fidelity to the original data.

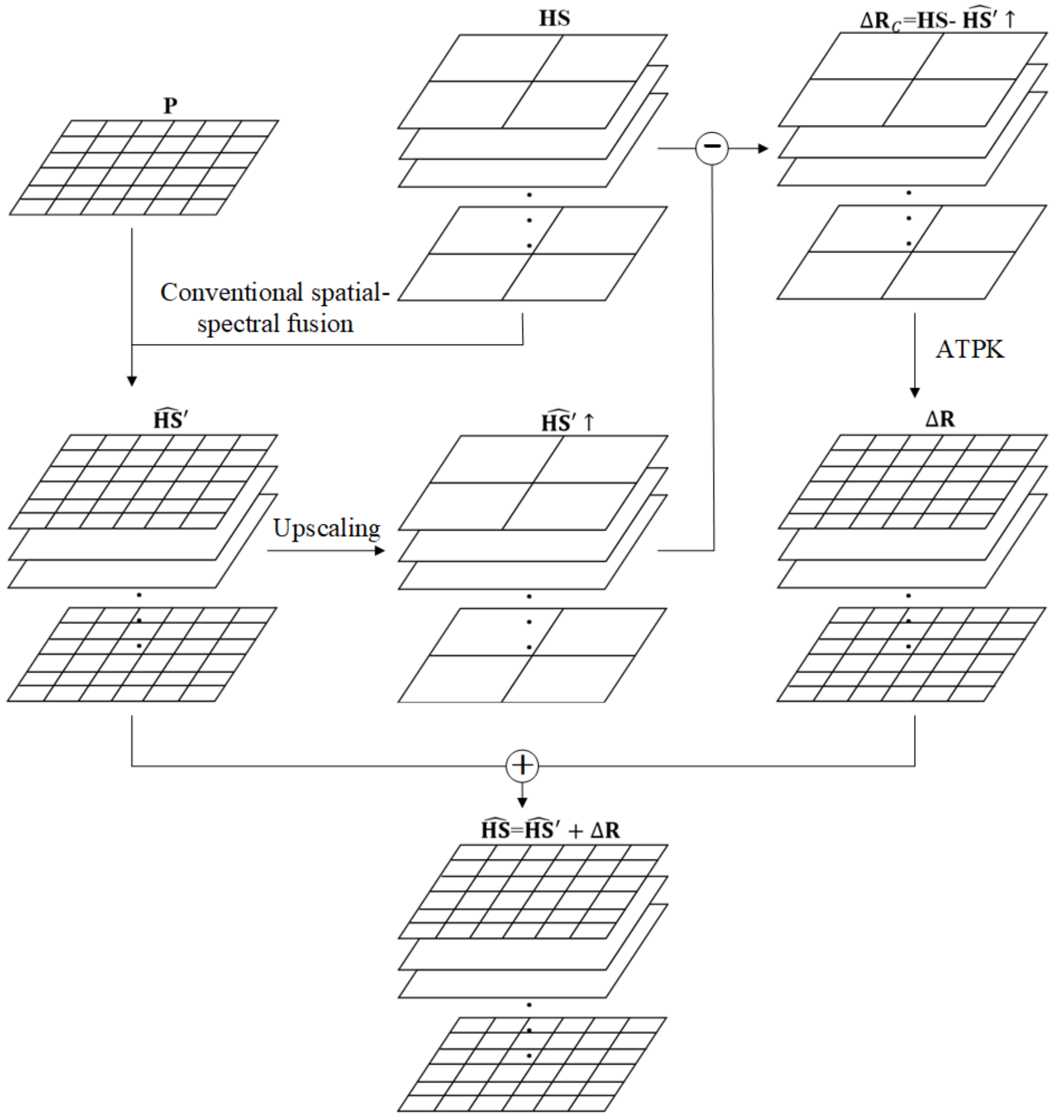


Fig. 2. Flowchart of the proposed DF-SSF.

4. Experiments

4.1. Experimental setup

In this paper, two sets of experiments were designed to validate the feasibility of the spatial-spectral fusion of CRISM and CTX data, as well as the effectiveness of our proposed DF-SSF model (e.g., its data fidelity ability). The six regions introduced in Section 2 were selected. For the CRISM hyperspectral data, after removing the noisy bands, the first 70 bands covering the spectral range similar to that of the CTX panchromatic image were selected in the experiments.

In the first set of experiments, the effectiveness of DF-SSF was validated utilizing simulated data. Specifically, due to the absence of 6 m CRISM images, there are no reference data for objective evaluation of the 6 m results produced by fusion of the 18 m CRISM data and the 6 m CTX data. Therefore, to ensure the existence of reliable hyperspectral reference images at the target fine spatial resolution, a commonly used strategy was adopted: the 18 m CRISM hyperspectral image and the 6 m CTX panchromatic image were degraded to 54 m and 18 m, respectively. Then the two degraded images were fused to obtain the 18 m hyperspectral image by spatial-spectral fusion. The original 18 m CRISM hyperspectral image was used as the reference image to evaluate the accuracy of the 18 m fusion result. During the degradation process, the Gaussian PSF (with a convolution kernel parameter of 0.5) was used in the experiments. This paper employs five evaluation metrics for quantitative assessment, including correlation coefficient (CC), spectral angle mapper (SAM), root mean square error (RMSE), relative global-dimensional synthesis error (ERGAS) (Ranchin and Wald, 2000) and universal image quality index (UIQI) (Wang and Bovik, 2002). For CC, RMSE, and UIQI, the values were computed band-by-band, and then averaged across all bands. For SAM, it was first calculated pixel-by-pixel and finally averaged across all pixels. To evaluate the data fidelity capability, we also evaluated the metric of coherence, which involves degrading the spatial-spectral fusion image to the original coarse spatial resolution and to calculate the CC with the original input coarse resolution image.

In the second set of experiments, the actual 18 m CRISM hyperspectral data were fused with 6 m CTX panchromatic data to obtain CRISM hyperspectral data at 6 m spatial resolution. The spatial-spectral fusion performance was evaluated mainly based on visual inspection and the metric of coherence.

CS and MRA are two types of the most widely used spatial-spectral fusion methods. Therefore, we applied DF-SSF to seven methods within the two categories: GS, GSA, PCA, SFIM, MTF-GLP, MTF-GLP-HPM, and guided filter PCA (GFPCA) (Liao et al., 2015). In addition, we also examined the applicability of DF-SSF to the unsupervised deep learning-based method, and a typical method of this type, that is, pansharpening based on a generative adversarial network (Pan-GAN) (Ma et al., 2020), was considered.

4.2. 18 m fusion results

1) Visual evaluation: Fig. 3 shows the spatial-spectral fusion results of different CS and MRA methods in the six regions. Note that the sub-areas marked in red are zoomed in Fig. 4 for clearer comparison. Visually, it is evident that all methods produce results closer to the reference images after considering data fidelity by

DS-SSF. Specifically, using the DF-SSF method, the spectral distortion present in the GS and PCA methods (such as in the alluvial fan and channel areas of the Eberswalde Crater region in Fig. 3) is significantly reduced. The fusion results of the GFPCA and SFIM methods based on DS-SSF reproduce more spatial structures (as seen in the Mawrth Vallis region where surface features are depicted as relatively small and dense layers of sediment). For the MTF-GLP and MTF-GLP-HPM methods, the spectral and spatial distortions in the results are relatively minor, but when considering data fidelity, the results are closer to the reference images in the hue.

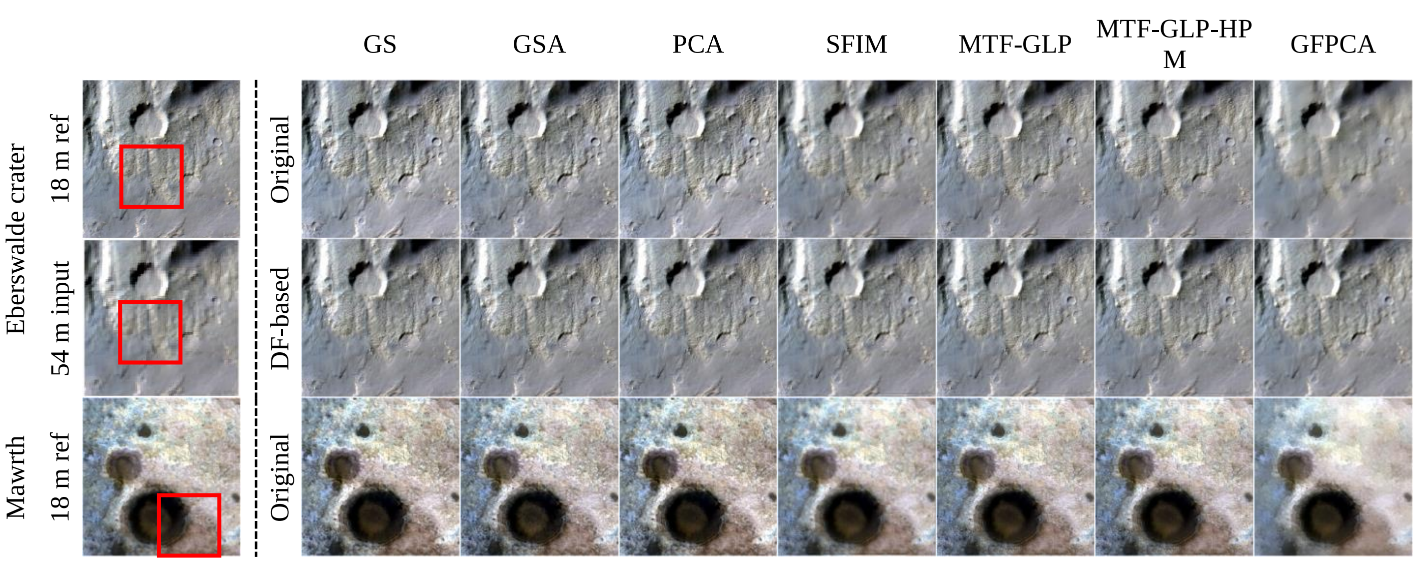
2) Accuracy evaluation: Firstly, for a clearer comparison of the results from different methods, we selected two bands from the fusion results in Fig. 3 for analysis, and produced the error maps in Fig. 5. It can be observed clearly that, for all seven methods, the errors are significantly reduced when using the DF-SSF method, particularly in areas such as smooth river channels, weathered regions, and impact craters. Hence, the fusion results obtained by DF-SSF exhibit smaller errors compared to those of the original methods. Secondly, scatterplots representing the results of the original and DF-SSF methods are given in Fig. 6. From the scatterplots, it is apparent that the results of the original spatial-spectral fusion methods are relatively scattered along the axes, while the fusion results of DF-SSF are more concentrated around the y=x line, indicating that its results are closer to the reference.

Table 2 provides quantitative evaluation results for the various methods in the six regions. It can be seen that the accuracy of the DF-SSF results is obviously greater than that of the original methods. For example, the CC and UIQI values of all six regions are increased by over 0.0120, the ERGAS values are all decreased by more than 0.0500, and the SAM values are all decreased by more than 0.0003. It is worth noting that even in some regions where several methods present lower accuracies (e.g., GS, GSA, and PCA in the Jezero Crater area), the accuracy is increased noticeably by using the DF-SSF method.

3) The effect of band correlation: Generally, spatial-spectral fusion methods perform better on bands that have larger correlation with the panchromatic image. Fig. 7 presents the fusion accuracy (using CC as an example) of various methods for all bands in the six regions (left) and also the relationship (in terms of CC) between the original CTX data and CRISM hyperspectral bands (right). As shown in Fig. 7, the trend of all curves in the left column is similar to that of the corresponding curves in the right column. Taking the Eberswalde crater region as an example, in the right plot of Fig. 7(a), the first 20 bands present a smaller correlation with the panchromatic image compared to other bands. Correspondingly, the fusion accuracy of all methods for the first 20 bands in the left plot of Fig. 7(a) is relatively lower. The physical reason for this phenomenon is that the first 20 bands are not completely covered by the panchromatic image. For the remaining 50 bands in the visible range, the panchromatic image provides more correlated information for the fusion process, leading to more accurate fusion results.

4) Coherence: Coherence is an important indicator for evaluating the quality of the spatial-spectral fusion results in terms of data fidelity. We degraded the 18 m spatial-spectral fusion results of various methods to 54 m, and compared them with the input 54 m CRISM hyperspectral data. Table 3 shows the overall coherence (in terms of CC) assessment for all six regions. It can be seen clearly that the proposed DF-SSF model achieves perfect coherence for all original methods, indicating its ability to preserve the original coarse data completely. This advantage of DS-SSF lies in the perfect coherence property of ATPK, as demonstrated mathematically in Section 3.4.

5) Unsupervised deep learning-based fusion results: We apply DF-SSF to Pan-GAN to validate its applicability to the unsupervised deep learning method. Since the performance of deep learning-based methods relies on the number of training data, different numbers of pairs (including 53, 28 and 14 pairs) of CRISM and CTX images were considered for training, and DF-SSF was examined in all these cases. The accuracy evaluation results are shown in Table 4. It can be seen that the spatial-spectral fusion accuracy of the proposed DF-SSF method is consistently greater than that of the original methods in all cases. For example, in the case of 53 pairs of training data, both CC and UIQI are increased by more than 0.0200, and ERGAS is decreased by more than 0.1500. Moreover, when the number of training data decreases, the accuracy of original Pan-GAN decreases correspondingly, but the advantage of DF-Pan-GAN is more obvious. For example, for the Eberswalde Crater data, the CC gain of DF-Pan-GAN over Pan-GAN is around 0.07 in the case of 53 pairs, and the gain further increases to 0.10 in the case of 14 pairs. Therefore, the proposed DF-SSF method is also effective for the unsupervised deep learning approach, even under various numbers of training data.



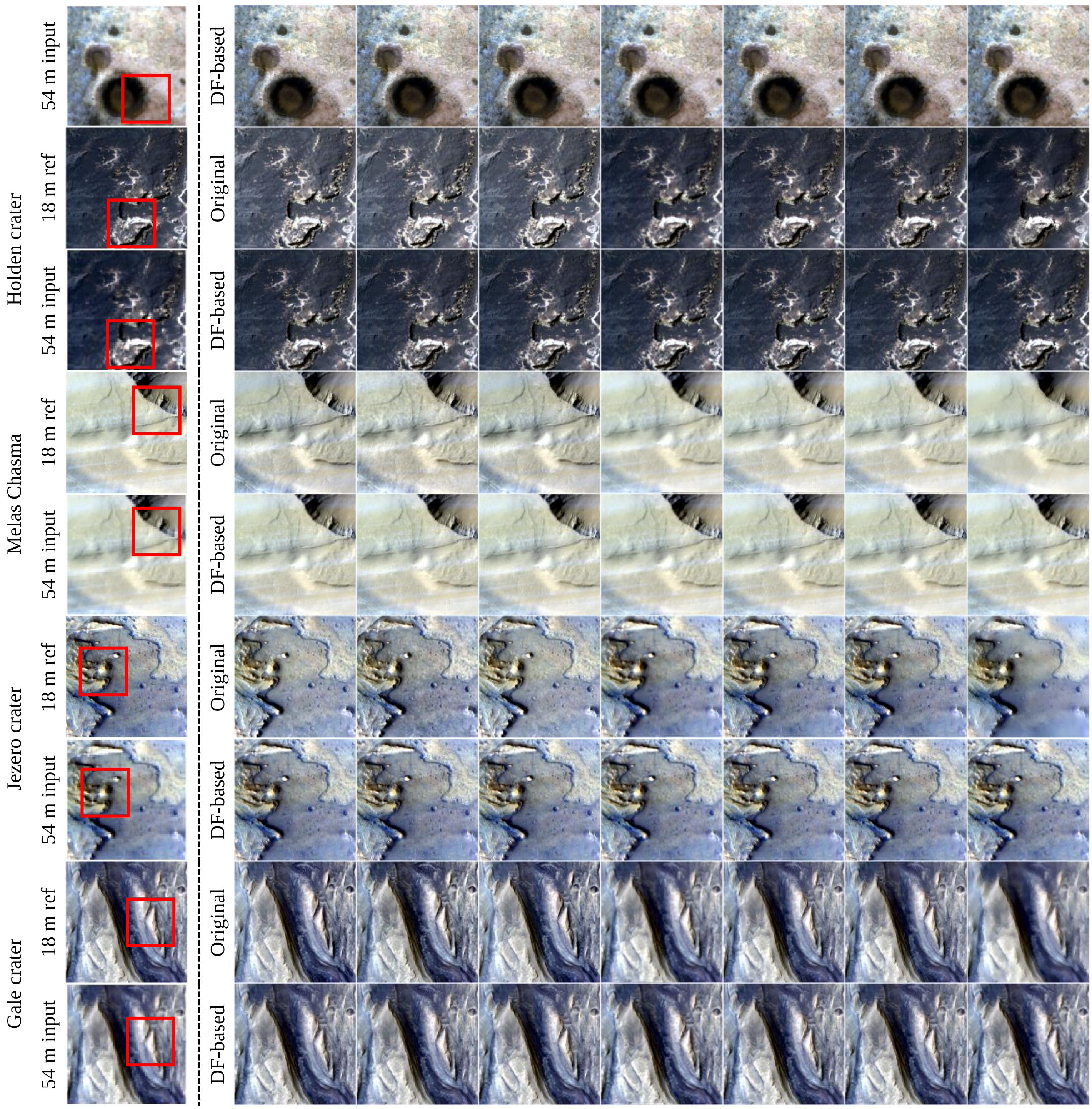
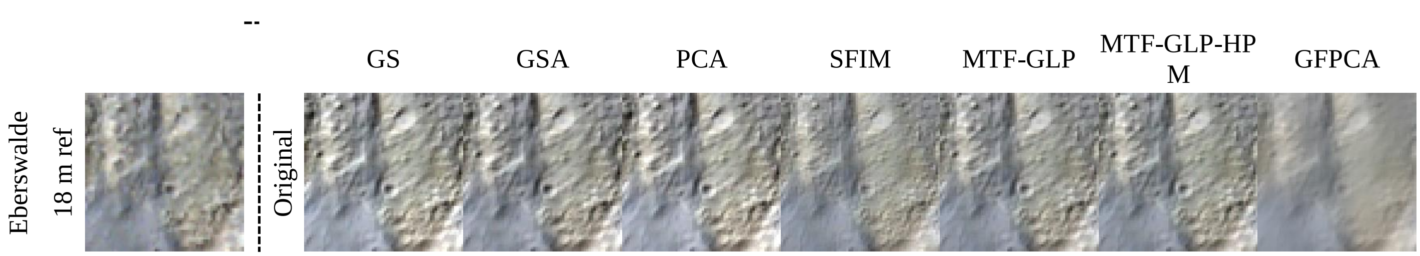


Fig. 3. 18 m spatial-spectral fusion results (300×300 pixels) of the six regions (bands 37, 25 and 12 as RGB). The first and second rows of each region are the results of the original SSF and DF-SSF, respectively. The sub-areas marked in red are zoomed in Fig. 4.



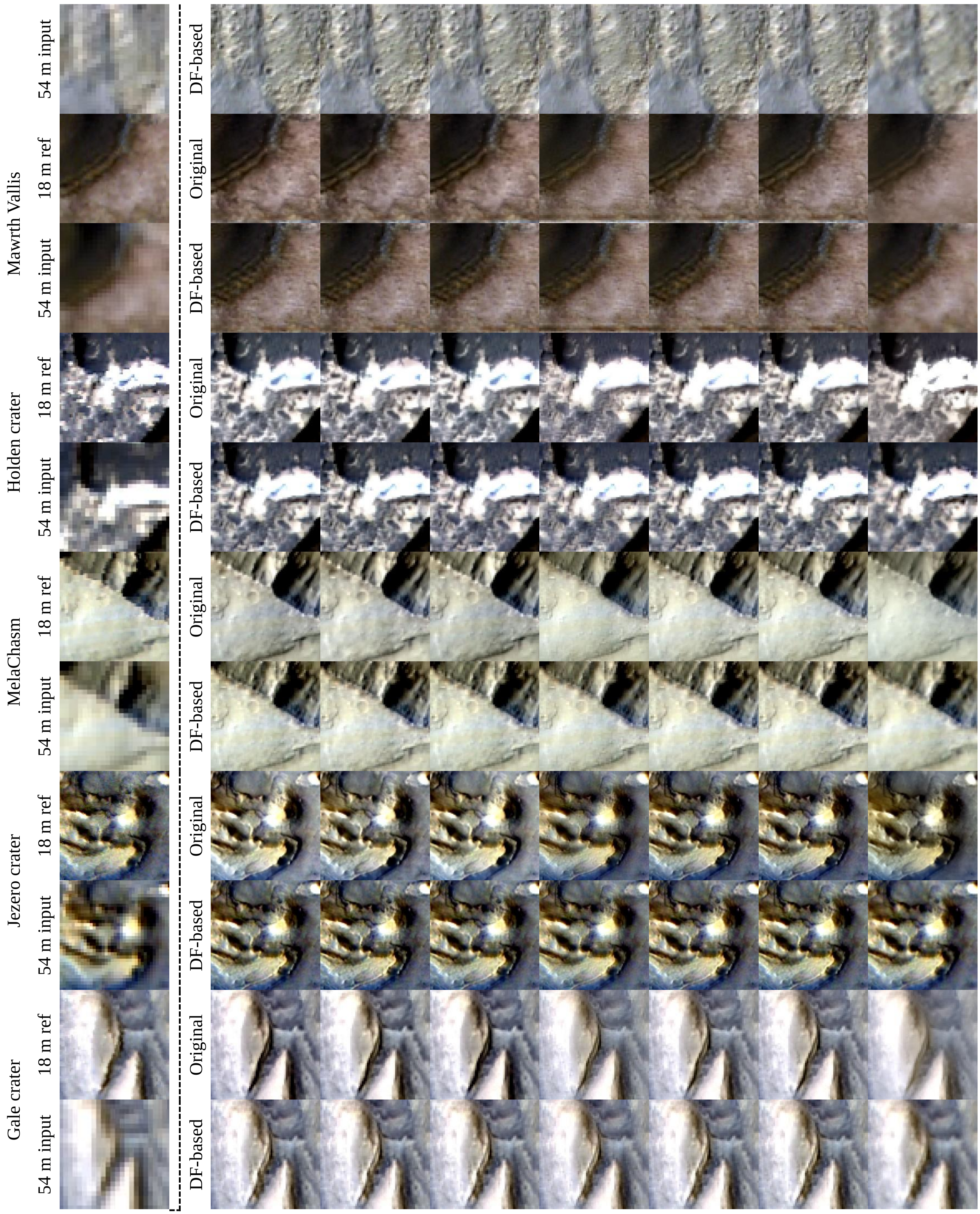


Fig. 4. The sub-areas of 18 m spatial-spectral fusion results in Fig. 3.

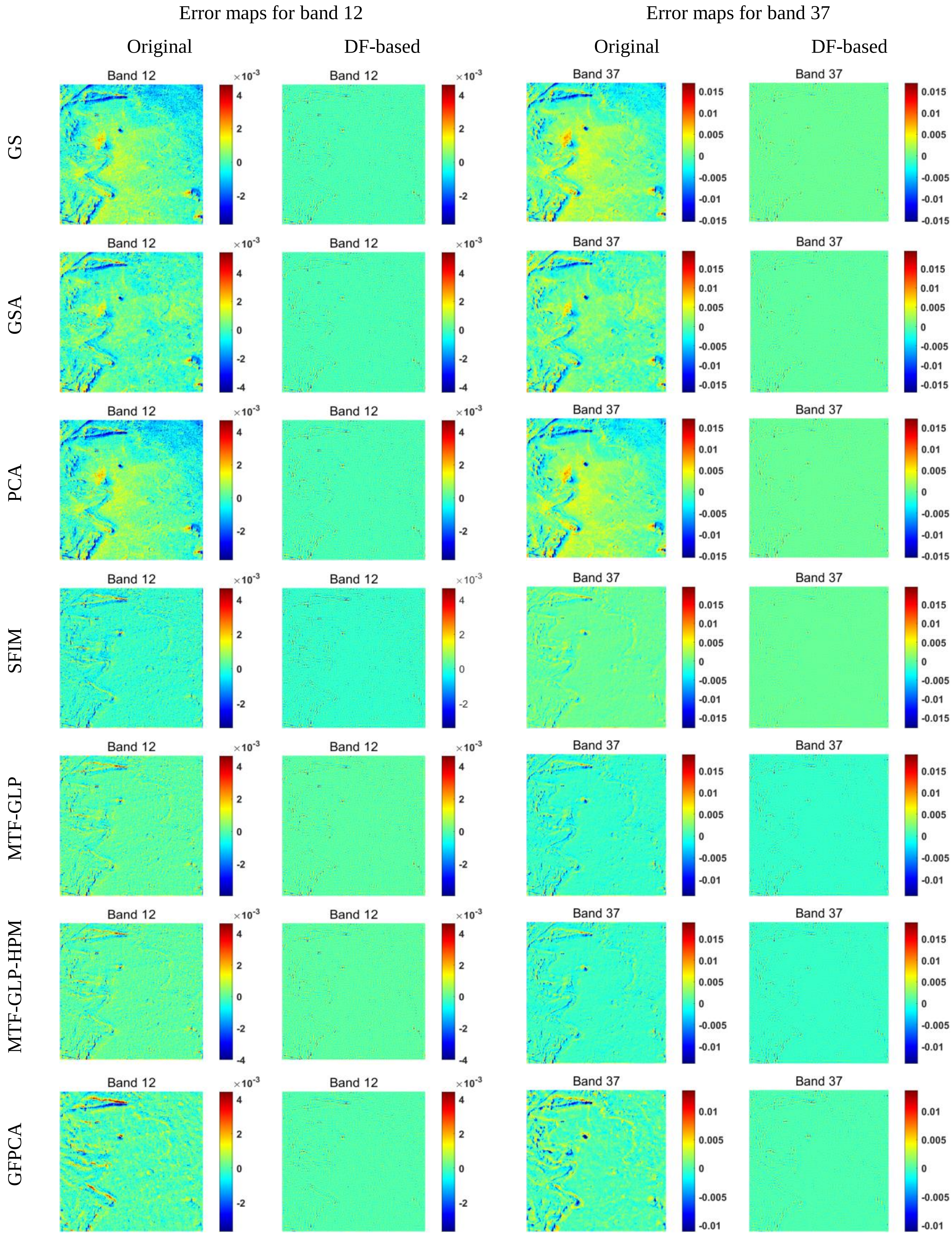
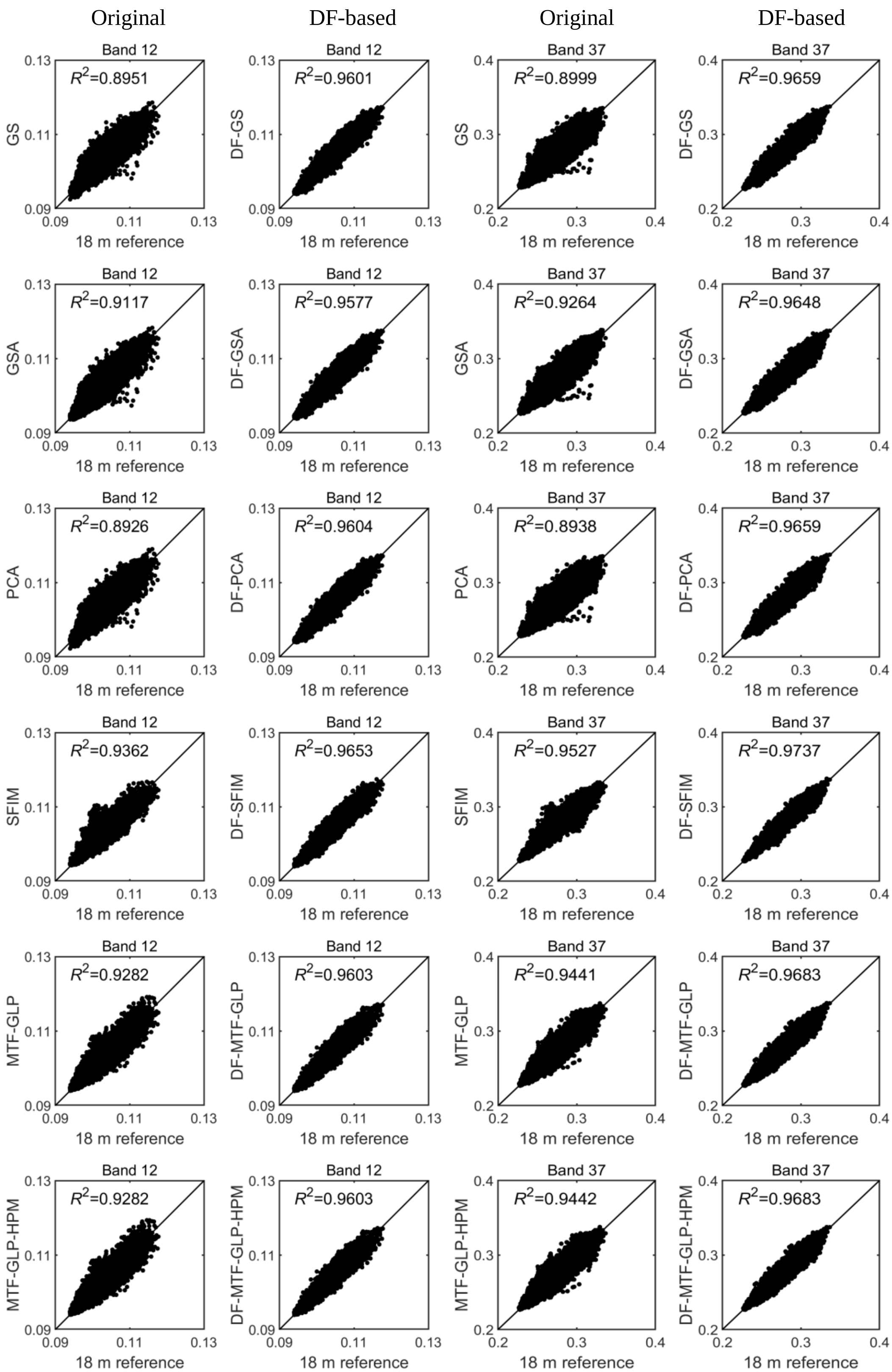


Fig. 5. Error maps for selected bands 12 and 37 in the Jezero crater region.

Table 2. Quantitative assessment of different spatial-spectral fusion methods for the six regions.

	Methods	CC		SAM		RMSE		ERGAS		UIQI	
		Original	DF-based	Original	DF-based	Original	DF-based	Original	DF-based	Original	DF-based
Eberswalde Crater	GS	0.8937	0.9654	0.0065	0.0051	0.0044	0.0025	1.3041	0.7638	0.8912	0.9654
	GSA	0.9264	0.9620	0.0062	0.0051	0.0038	0.0027	1.1300	0.8041	0.9260	0.9619
	PCA	0.8937	0.9655	0.0065	0.0051	0.0044	0.0025	1.3054	0.7638	0.8912	0.9654
	SFIM	0.9318	0.9682	0.0059	0.0050	0.0035	0.0024	1.0521	0.7252	0.9310	0.9682
	MTF-GLP	0.9317	0.9660	0.0060	0.0050	0.0036	0.0025	1.0685	0.7548	0.9317	0.9659
	MTF-GLP-HPM	0.9312	0.9656	0.0060	0.0051	0.0036	0.0025	1.0723	0.7603	0.9312	0.9655
	GFPCA	0.9561	0.9796	0.0063	0.0049	0.0029	0.0019	0.8602	0.5755	0.9505	0.9705
Mawrth Vallis	GS	0.9477	0.9817	0.0039	0.0027	0.0049	0.0029	0.7137	0.4203	0.9474	0.9816
	GSA	0.9598	0.9809	0.0035	0.0027	0.0043	0.0030	0.6266	0.4280	0.9597	0.9809
	PCA	0.9451	0.9817	0.0040	0.0027	0.0051	0.0029	0.7355	0.4204	0.9449	0.9817
	SFIM	0.9725	0.9851	0.0032	0.0025	0.0034	0.0026	0.4996	0.3721	0.9727	0.9851
	MTF-GLP	0.9686	0.9824	0.0033	0.0026	0.0038	0.0028	0.5543	0.4075	0.9686	0.9824
	MTF-GLP-HPM	0.9687	0.9825	0.0033	0.0026	0.0038	0.0028	0.5439	0.4072	0.9686	0.9824
	GFPCA	0.9734	0.9900	0.0037	0.0024	0.0033	0.0020	0.4861	0.2962	0.9722	0.9900
Holden Crater	GS	0.8701	0.9576	0.0127	0.0095	0.0105	0.0061	2.9637	1.7348	0.8632	0.9576
	GSA	0.8971	0.9540	0.0121	0.0097	0.0096	0.0064	2.7250	1.8149	0.8971	0.9540
	PCA	0.8702	0.9577	0.0129	0.0095	0.0105	0.0061	2.9659	1.7339	0.8635	0.9577
	SFIM	0.9104	0.9583	0.0113	0.0095	0.0088	0.0061	2.4776	1.7155	0.9067	0.9579
	MTF-GLP	0.9134	0.9576	0.0113	0.0095	0.0086	0.0061	2.4446	1.7303	0.9127	0.9575
	MTF-GLP-HPM	0.9131	0.9574	0.0113	0.0095	0.0087	0.0061	2.4520	1.7378	0.9123	0.9573
	GFPCA	0.9121	0.9656	0.0121	0.0092	0.0088	0.0055	2.4991	1.5544	0.8934	0.9650
Melas Chasma	GS	0.8938	0.9814	0.0036	0.0026	0.0059	0.0025	1.0204	0.4311	0.8930	0.9814
	GSA	0.9244	0.9803	0.0035	0.0027	0.0052	0.0026	0.9023	0.4455	0.9234	0.9803
	PCA	0.8974	0.9815	0.0036	0.0026	0.0058	0.0025	1.0036	0.4305	0.8967	0.9815
	SFIM	0.9636	0.9836	0.0031	0.0026	0.0034	0.0023	0.5984	0.4030	0.9635	0.9836
	MTF-GLP	0.9613	0.9820	0.0032	0.0026	0.0036	0.0024	0.6212	0.4227	0.9613	0.9820
	MTF-GLP-HPM	0.9615	0.9821	0.0032	0.0026	0.0035	0.0024	0.6191	0.4212	0.9618	0.9821
	GFPCA	0.9623	0.9897	0.0036	0.0026	0.0034	0.0018	0.6045	0.3176	0.9599	0.9897
Jezero Crater	GS	0.8077	0.9607	0.0029	0.0021	0.0023	0.0010	0.3687	0.1640	0.8052	0.9606
	GSA	0.8509	0.9543	0.0027	0.0021	0.0021	0.0011	0.3477	0.1783	0.8483	0.9541
	PCA	0.8068	0.9609	0.0029	0.0021	0.0023	0.0010	0.3719	0.1639	0.8048	0.9608
	SFIM	0.9377	0.9670	0.0023	0.0020	0.0012	0.0009	0.2030	0.1461	0.9362	0.9669
	MTF-GLP	0.9282	0.9618	0.0024	0.0020	0.0013	0.0010	0.2213	0.1593	0.9280	0.9617
	MTF-GLP-HPM	0.9281	0.9617	0.0024	0.0020	0.0013	0.0010	0.2216	0.1596	0.9278	0.9616
	GFPCA	0.9144	0.9744	0.0027	0.0020	0.0014	0.0007	0.2381	0.1253	0.9011	0.9742
Gale Crater	GS	0.9306	0.9793	0.0052	0.0033	0.0062	0.0034	1.0395	0.5700	0.9297	0.9793
	GSA	0.9471	0.9782	0.0045	0.0033	0.0054	0.0035	0.9186	0.5846	0.9470	0.9781
	PCA	0.9280	0.9794	0.0053	0.0033	0.0063	0.0034	1.0619	0.5702	0.9273	0.9793
	SFIM	0.9617	0.9819	0.0040	0.0031	0.0045	0.0031	0.7651	0.5258	0.9612	0.9819
	MTF-GLP	0.9590	0.9801	0.0042	0.0032	0.0047	0.0033	0.7987	0.5535	0.9590	0.9801
	MTF-GLP-HPM	0.9588	0.9800	0.0042	0.0032	0.0047	0.0033	0.8013	0.5560	0.9588	0.9800
	GFPCA	0.9697	0.9873	0.0042	0.0029	0.0041	0.0026	0.6934	0.4349	0.9665	0.9872



Original
DF-based
Original
DF-based
Band 12
Band 37
R^2=0.8951
R^2=0.9601
R^2=0.8999
R^2=0.9659
GS
DF-GS
R^2=0.9117
R^2=0.9577
R^2=0.9264
R^2=0.9648
GSA
DF-GSA
R^2=0.8926
R^2=0.9604
R^2=0.8938
R^2=0.9659
PCA
DF-PCA
R^2=0.9362
R^2=0.9653
R^2=0.9527
R^2=0.9737
SFIM
DF-SFIM
R^2=0.9282
R^2=0.9603
R^2=0.9441
R^2=0.9683
MTF-GLP
DF-MTF-GLP
R^2=0.9282
R^2=0.9603
R^2=0.9442
R^2=0.9683
MTF-GLP-HPM
DF-MTF-GLP-HPM
18 m reference

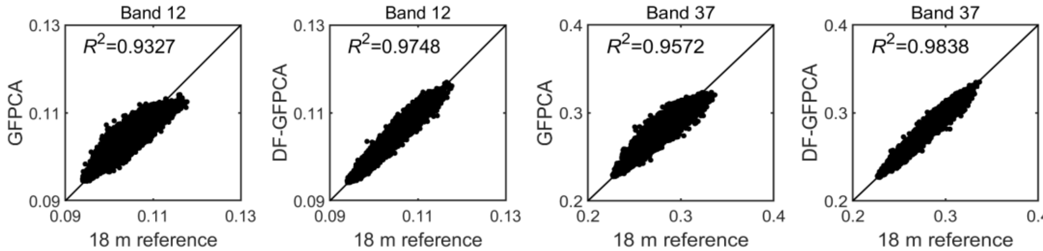
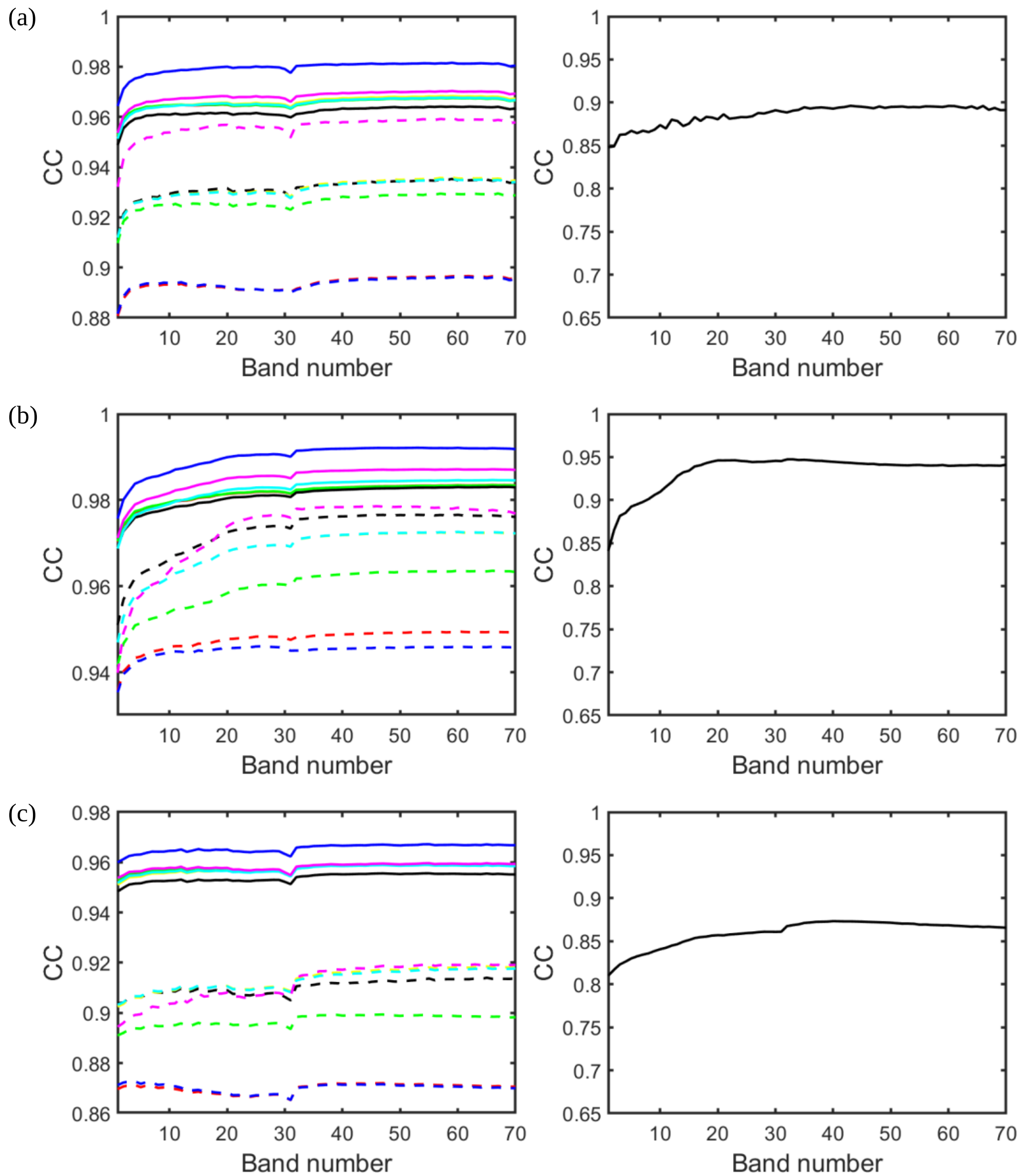


Fig. 6. Scatterplots of predicted against actual coarse pixel values (18 m) in the Mawrth Vallis region.



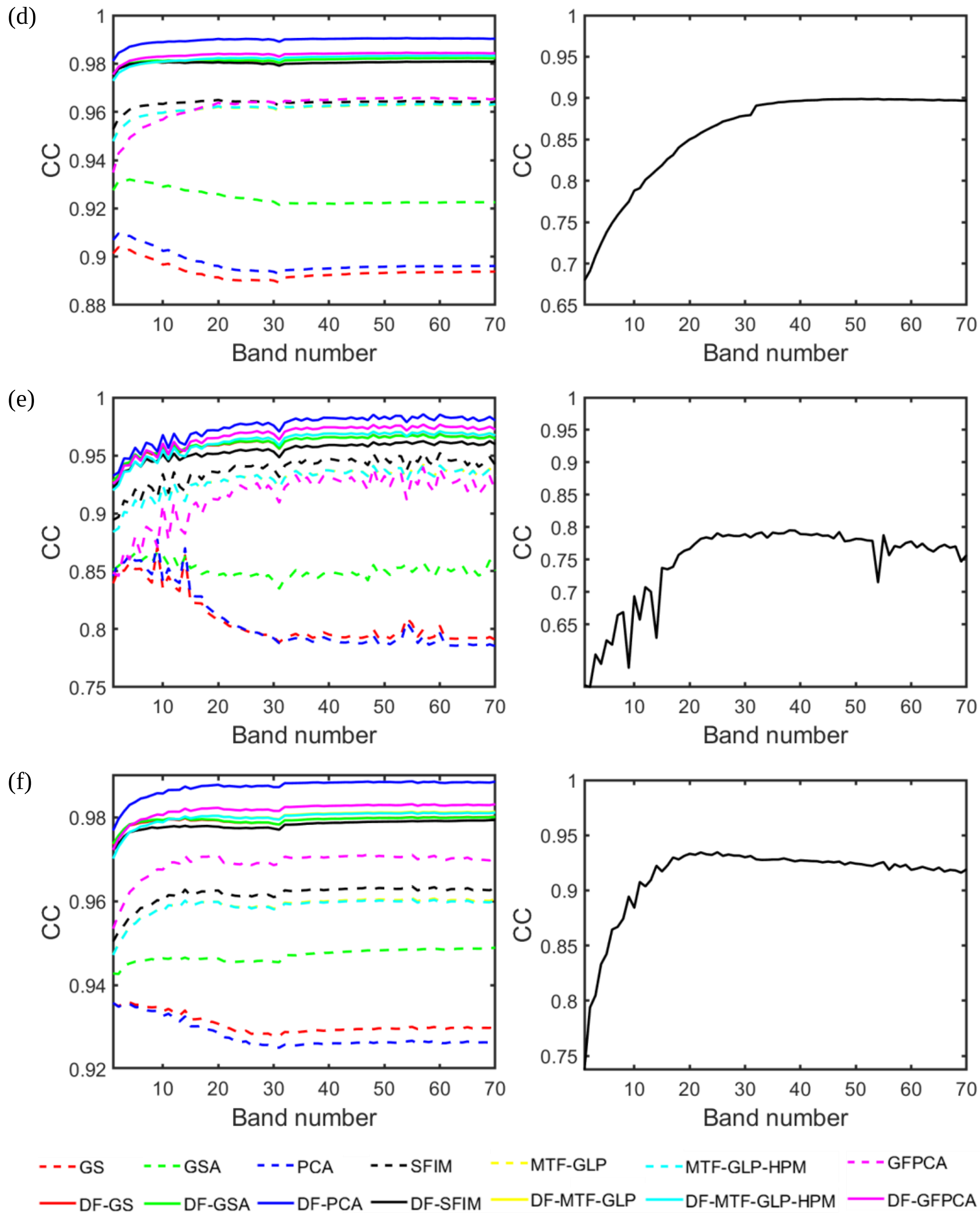


Fig. 7. The spatial-spectral fusion accuracy (in terms of CC) of each band (left) and relation (in terms of CC) between the panchromatic (degraded) and hyperspectral bands (right) for the six regions. (a) Eberswalde crater. (b) Mawrth Vallis. (c) Holden crater. (d) Melas Chasma. (e) Jezero crater. (f) Gale crater.

Table 3. Evaluation (in terms of CC) of the data fidelity ability of the 18 m spatial-spectral fusion results for the six regions.

		Eberswalde crater	Mawrth Vallis	Holden crater	Melas Chasma	Jezero crater	Gale crater
GS	Original	0.9397	0.9760	0.9319	0.9285	0.8611	0.9634
	DF-based	0.9998	0.9999	0.9997	0.9998	0.9997	0.9999
GSA	Original	0.9763	0.9885	0.9616	0.9584	0.9144	0.9808
	DF-based	0.9998	0.9999	0.9997	0.9998	0.9996	0.9999
PCA	Original	0.9395	0.9732	0.9311	0.9320	0.8591	0.9591
	DF-based	0.9998	0.9999	0.9997	0.9998	0.9997	0.9999
SFIM	Original	0.9771	0.9934	0.9714	0.9878	0.9833	0.9884

	DF-based	0.9998	0.9999	0.9997	0.9999	0.9999	0.9999
MTF-GLP	Original	0.9794	0.9934	0.9738	0.9886	0.9818	0.9887
	DF-based	0.9998	0.9999	0.9998	0.9999	0.9998	0.9999
MTF-GLP-HPM	Original	0.9795	0.9934	0.9738	0.9886	0.9818	0.9887
	DF-based	0.9998	0.9999	0.9997	0.9999	0.9998	0.9999
GFPCA	Original	0.9884	0.9909	0.9663	0.9833	0.9597	0.9914
	DF-based	0.9999	0.9999	0.9997	0.9999	0.9998	0.9999

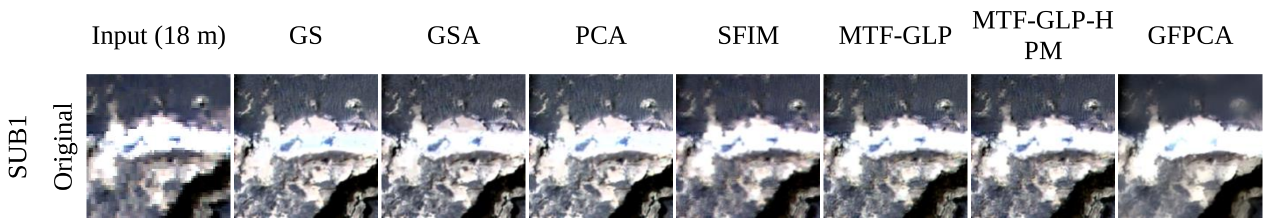
Table 4. Quantitative assessment of the unsupervised deep learning method (Pan-GAN) for the six regions.

	Regions	CC		SAM		RMSE		ERGAS		UIQI	
		Original	DF-based	Original	DF-based	Original	DF-based	Original	DF-based	Original	DF-based
53 pairs of training data	Eberswalde Crater	0.9057	0.9756	0.0093	0.0044	0.0049	0.0022	1.4642	0.6674	0.8981	0.9753
	Mawrth Vallis	0.9716	0.9934	0.0029	0.0012	0.0037	0.0016	0.5531	0.2396	0.9689	0.9933
	Holden Crater	0.9424	0.9843	0.0150	0.0087	0.0082	0.0038	2.3353	1.0794	0.9361	0.9840
	Melas Chasma	0.9596	0.9919	0.0034	0.0013	0.0040	0.0016	0.6958	0.2867	0.9570	0.9919
	Jezero Crater	0.9463	0.9874	0.0227	0.0007	0.0016	0.0006	0.2773	0.0979	0.9318	0.9871
	Gale Crater	0.9468	0.9880	0.0042	0.0019	0.0056	0.0023	0.9637	0.3952	0.9389	0.9878
28 pairs of training data	Eberswalde Crater	0.8986	0.9736	0.0094	0.0048	0.0054	0.0024	1.6128	0.7143	0.8845	0.9729
	Mawrth Vallis	0.9609	0.9899	0.0036	0.0015	0.0045	0.0020	0.6515	0.2955	0.9578	0.9897
	Holden Crater	0.9363	0.9806	0.0156	0.0088	0.0090	0.0044	2.3837	1.7723	0.9315	0.9802
	Melas Chasma	0.9524	0.9881	0.0036	0.0015	0.0046	0.0018	0.7861	0.3039	0.9517	0.9880
	Jezero Crater	0.9393	0.9800	0.0236	0.0009	0.0018	0.0008	0.3008	0.1203	0.9177	0.9795
	Gale Crater	0.9449	0.9864	0.0047	0.0022	0.0057	0.0025	0.9720	0.4233	0.9360	0.9861
14 pairs of training data	Eberswalde Crater	0.8609	0.9632	0.0117	0.0061	0.0063	0.0028	1.8772	0.8310	0.8450	0.9625
	Mawrth Vallis	0.9491	0.9852	0.0044	0.0022	0.0049	0.0023	0.7235	0.3486	0.9449	0.9848
	Holden Crater	0.9349	0.9763	0.0169	0.0100	0.0091	0.0046	2.4019	2.0302	0.9279	0.9740
	Melas Chasma	0.9510	0.9874	0.0044	0.0021	0.0054	0.0021	0.7879	0.3616	0.9509	0.9872
	Jezero Crater	0.9297	0.9795	0.0216	0.0013	0.0023	0.0015	0.3196	0.1439	0.9269	0.9732
	Gale Crater	0.9439	0.9847	0.0047	0.0028	0.0058	0.0025	1.0006	0.4401	0.9354	0.9841

4.3. 6 m fusion results

In this set of experiments, the original fusion and DF-SSF methods were used for spatial-spectral fusion of the original 18 m CRISM hyperspectral data and 6 m CTX panchromatic data, to create 6 m CRISM hyperspectral images. The Gaussian PSF with a convolution kernel parameter of 0.5 was used. The 6 m fusion results for the Holden crater are shown in Figs. 8 and 9, while the results for the Gale crater regions are shown in Figs. 10 and 11. Note that Figs. 9 and 11 show the zoom images of three sub-regions in Figs. 8 and 10, respectively. It is evident that the 6 m CRISM images are visually more pleasant, which present more spatial details. Furthermore, the 6 m DF-SSF results are more similar in color to the original 18 m images, indicating greater data fidelity. Table 5 provides the coherence evaluation results for the Holden Crater and Gale Crater regions. It is apparent that the data fidelity of the seven methods is increased using the proposed model (e.g., the coherence is increased by at least 0.0080).

Fig. 8. 6 m spatial-spectral fusion results for the Holden crater region (bands 37, 25 and 12 as RGB). The sizes of the panchromatic and hyperspectral images are 900×900 and 300×300, respectively. (a1) 18 m CRISM. (a2) 6 m CTX. (b1)-(h1) show the results of the original SSF, while (b2)-(h2) show the corresponding results of DF-SSF. (b) GS. (c) GSA. (d) PCA. (e) SFIM. (f) MTF-GLP. (g) MTF-GLP-HPM. (h) GFPCA.



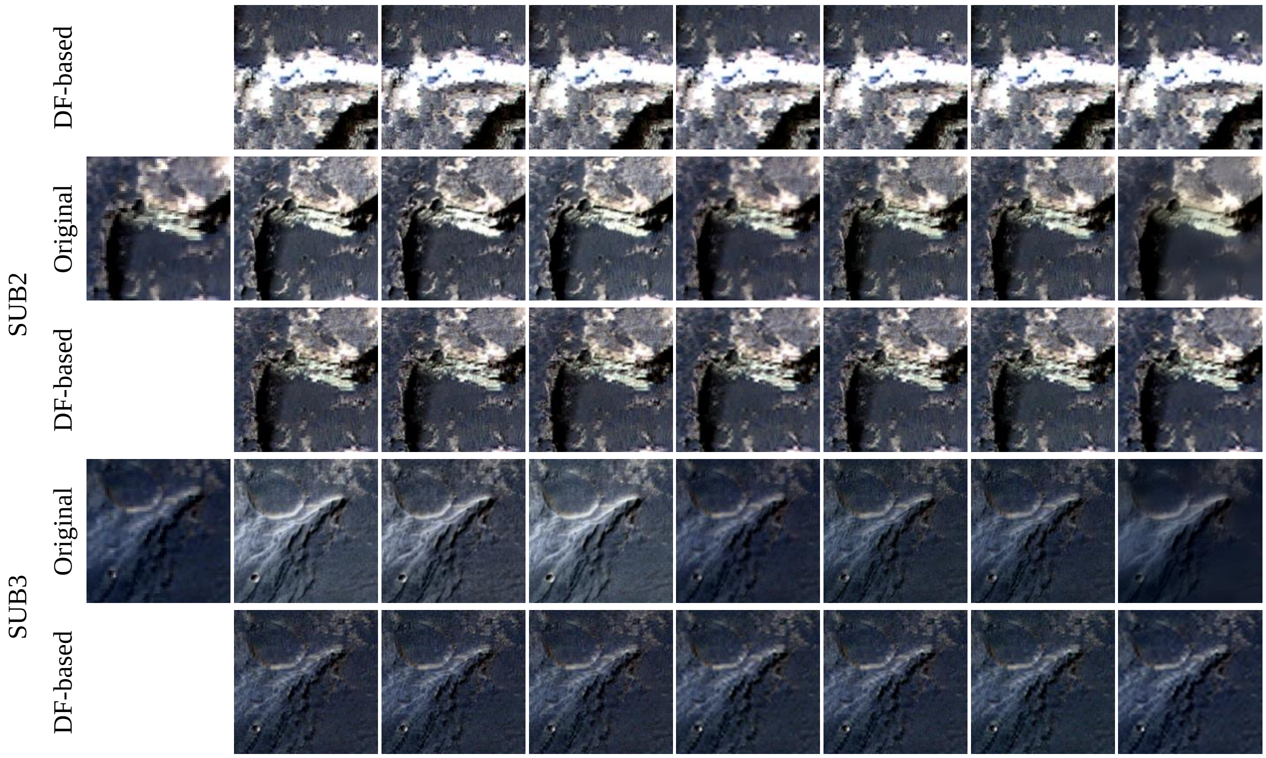
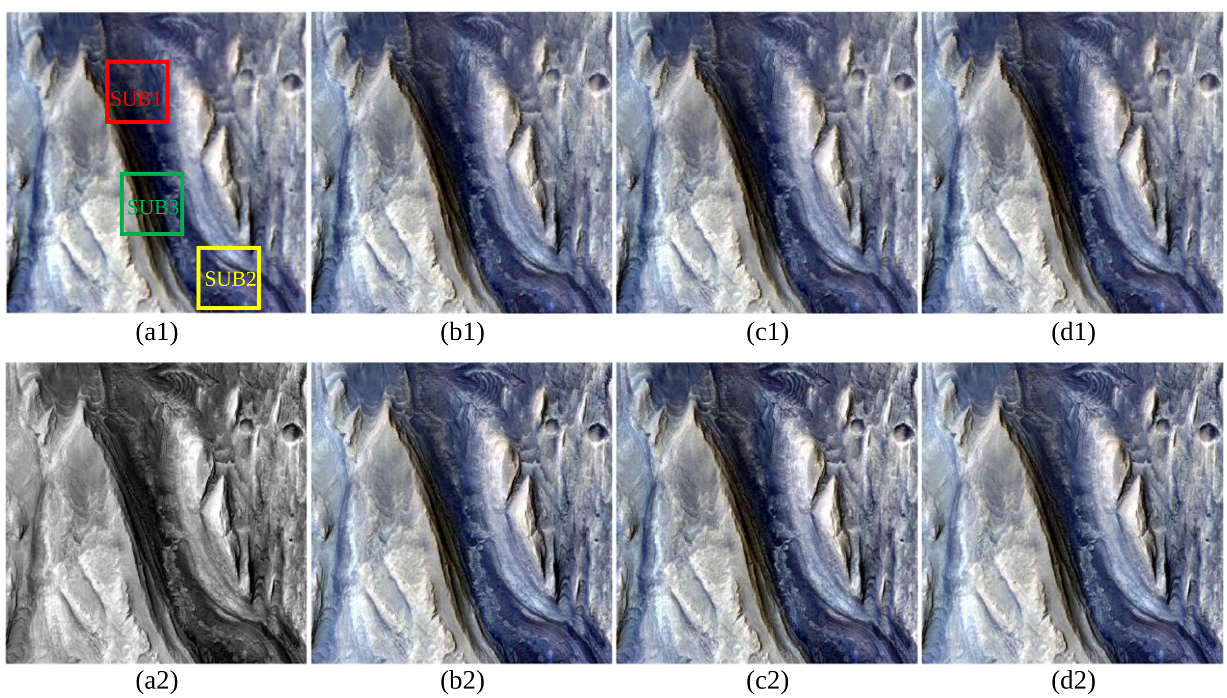


Fig. 9. 6 m spatial-spectral fusion results of the three sub-areas in Fig. 8. The first row of each sub-area shows the results of original SSF and the second row shows the results of DF-SSF.



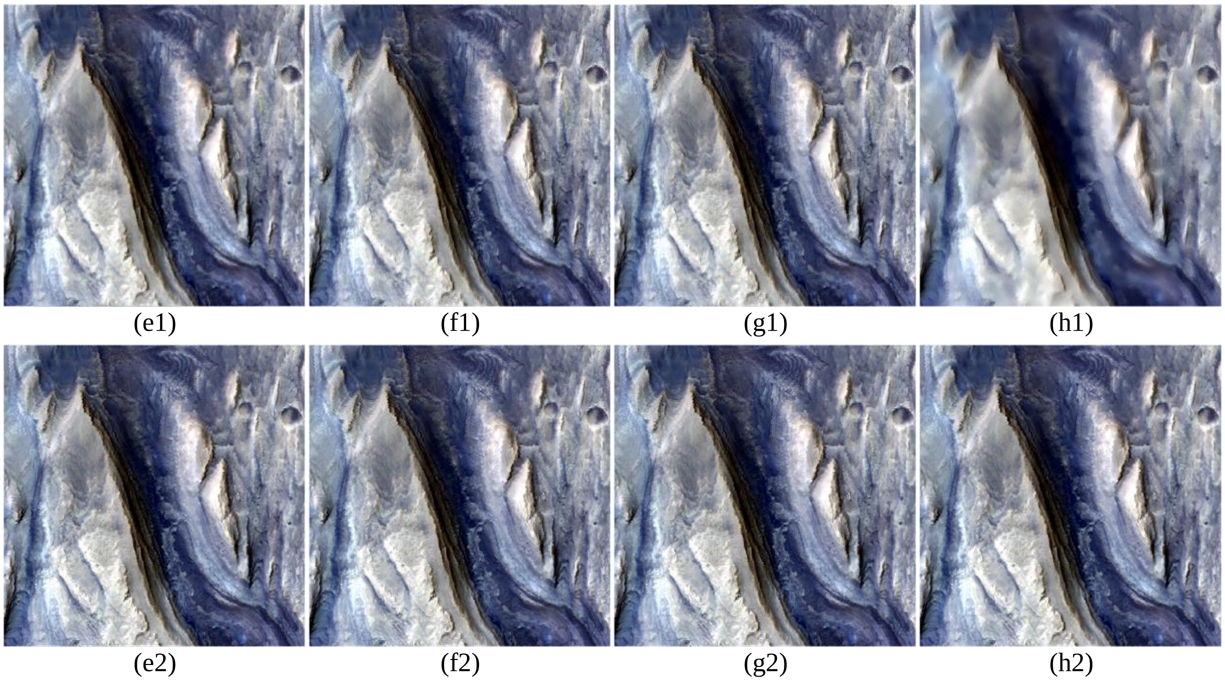
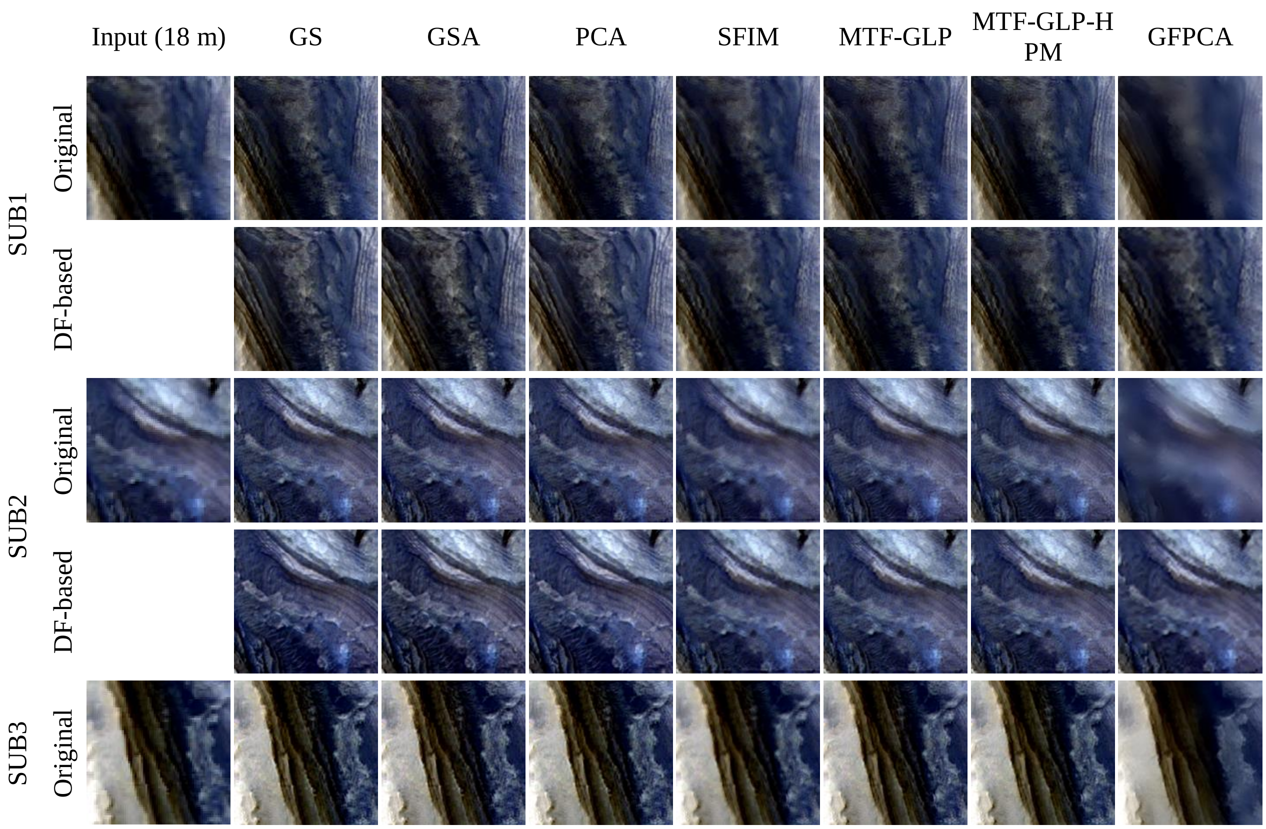


Fig. 10. 6 m spatial-spectral fusion results for the Gale crater region (bands 37, 25 and 12 as RGB). The sizes of the panchromatic and hyperspectral images are 900×900 and 300×300, respectively. (a1) 18 m CRISM. (a2) 6 m CTX. (b1)-(h1) show the results of the original SSF, while (b2)-(h2) show the corresponding results of DF-SSF. (b) GS. (c) GSA. (d) PCA. (e) SFIM. (f) MTF-GLP. (g) MTF-GLP-HPM. (h) GFPCA.



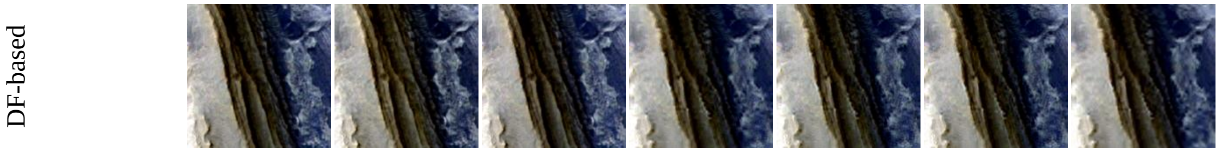


Fig. 11. 6 m spatial-spectral fusion results of the three sub-areas in Fig. 10. The first row of each sub-area shows the results of original SSF and the second row shows the results of DF-SSF.

Table 5. Evaluation (in terms of CC) of the data fidelity ability of the 6 m spatial-spectral fusion results for the Holden crater region and the Gale crater region.

	Holden crater		Gale crater	
	Original	DF-based	Original	DF-based
GS	0.8739	0.9993	0.9325	0.9997
GSA	0.8960	0.9992	0.9475	0.9997
PCA	0.8745	0.9993	0.9300	0.9997
SFIM	0.9796	0.9997	0.9929	0.9999
MTF-GLP	0.9789	0.9996	0.9917	0.9999
MTF-GLP-HPM	0.9786	0.9996	0.9917	0.9999
GFPCA	0.9521	0.9995	0.9808	0.9998

5. Discussion

5.1. Uncertainty in spatial-spectral fusion of CRISM and CTX images

While the experiments validated the effectiveness of fusing CTX and CRISM hyperspectral data, there still exists uncertainty in the fusion process.

Spectrally, uncertainty may be introduced if the spectral ranges of CTX image and CRISM hyperspectral images do not match, especially for hyperspectral bands that are not covered by the spectral range of the CTX panchromatic image (as shown in the fusion accuracy of the first 20 bands in Fig. 7). Therefore, the spectral overlap between the CTX image and the CRISM hyperspectral image should be considered before spatial-spectral fusion. Since the CTX image contains only one visible band, while the CRISM image contains both visible and infrared bands, to mitigate the uncertainty in the experiments, only the first 70 bands of CRISM that roughly match with the CTX spectral coverage were considered. For further analysis of the uncertainty, we evaluated the fusion accuracy for the 71st to the 110th bands of the CRISM hyperspectral data in the Mawrth Vallis region. The results are shown in Fig. 12. It is seen clearly that when there is no spectral overlap between CTX and CRISM hyperspectral images (e.g., for the 71st to the 110th bands), the fusion accuracy decreases obviously, especially for the original version without the consideration of DF.

Temporally, it is important to ensure that the CTX and CRISM data were observed at the same time. In practice, due to many factors (e.g., shooting time, camera operating mode, etc.), it can sometimes be challenging to obtain CTX and CRISM data for the same area at the same time. When there is a significant time gap, there may be certain changes in Martian surface between the two datasets (e.g., the formation of new impact craters on the Martian surface due to recent impact events), which may lead to uncertainties in the fusion process.

Spatially, accurate geometric registration is a crucial prerequisite for spatial-spectral fusion. When fusing CTX images with CRISM hyperspectral images, it is essential to consider fully the geometric registration accuracy between the two datasets. Although reliable registration was performed between the CTX and CRISM data in this study, it is worth noting that the registration accuracy may not be perfect, which can also introduce some uncertainties.

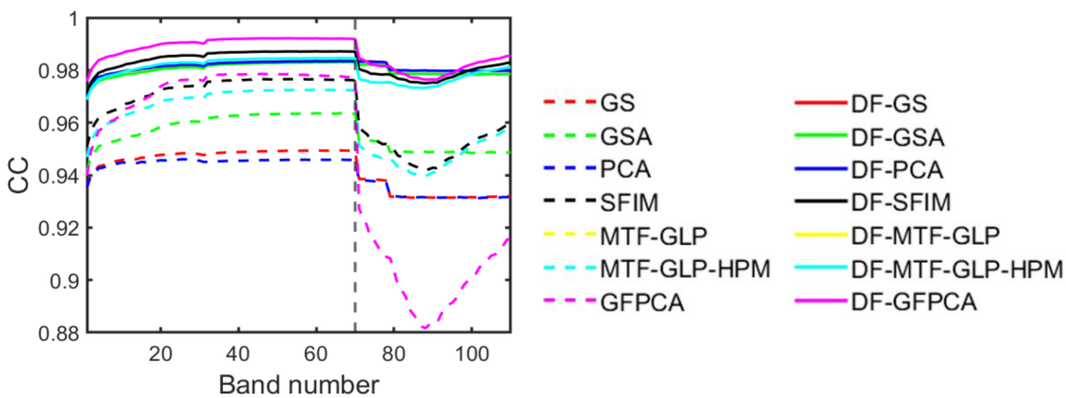


Fig.12. Quantitative assessment (in terms of CC) of the spatial-spectral fusion methods for bands 1 to 110 (the Mawrth Vallis region as an example).

5.2. Advantages of using ATPK in DF-SSF

The use of ATPK in DF-SSF lies in the appealing advantage of maintaining completely the original coarse data. To analyze the advantages of ATPK in residual downscaling in DF-SSF, we selected the bicubic interpolation (BI) method to downscale the residuals in existing spatial-spectral fusion methods and compared it with the proposed method. BI offers a good balance between accuracy and computational complexity and has been used widely in various studies, but it cannot preserve the original data perfectly. The experimental results are shown in Table 6. It is evident that the proposed method produces more accurate fusion results in all cases. This illustrates directly the benefit of achieving complete data fidelity in spatial-spectral fusion. Note that in existing studies, some methods were also designed with the objective to achieve data fidelity, such as the VO-based methods. Mathematically, however, they can only approach the original coarse images gradually to achieve approximate fidelity. This is different from ATPK, which can achieve data fidelity in the true mathematical sense.

Table 6. Comparison between the use of bicubic interpolation (BI) and ATPK (i.e., the proposed method) for residual downscaling in spatial-spectral fusion (the Eberswalde crater region as an example).

	CC		SAM		RMSE		ERGAS		Coherence	
	BI-based	DF-based	BI-based	DF-based	BI-based	DF-based	BI-based	DF-based	BI-based	DF-based
GS	0.9592	0.9654	0.0053	0.0051	0.0028	0.0025	0.8304	0.7638	0.9983	0.9998
GSA	0.9554	0.9620	0.0054	0.0051	0.0029	0.0027	0.8767	0.8041	0.9980	0.9998
PCA	0.9592	0.9655	0.0053	0.0051	0.0028	0.0025	0.8304	0.7638	0.9983	0.9998
SFIM	0.9625	0.9682	0.0053	0.0050	0.0026	0.0024	0.7832	0.7252	0.9981	0.9998
MTF-GLP	0.9590	0.9660	0.0053	0.0050	0.0028	0.0025	0.8258	0.7548	0.9981	0.9998
MTF-GLP-HPM	0.9585	0.9656	0.0053	0.0051	0.0028	0.0025	0.9312	0.7603	0.9981	0.9998
GFPCA	0.9745	0.9796	0.0052	0.0049	0.0021	0.0019	0.6452	0.5755	0.9984	0.9999

5.3. DF-SSF vs ATPRK

In ATPRK, the fusion process is achieved by two steps: regression modeling and ATPK-based residual downscaling. The regression part links the coarse multi/hyperspectral image and fine panchromatic image through a linear fitting process, and the ATPK part downscales the coarse residuals in the regression process to the target fine spatial resolution. However, the relation between multi/hyperspectral and panchromatic images is sometimes complicated, which may not be characterized accurately by a simple linear model. As a result, the residuals of the regression model may be large, and the uncertainty introduced into the predictions of post-ATPK may be correspondingly large. As an alternative, the DF-SSF method proposed in this paper uses any of the existing spatial-spectral fusion method as the first step (i.e., instead of the regression step in ATPRK). To analyze the influence of magnitude of the residuals in the methods, Table 7 presents the quantification of coarse residuals for various spatial-spectral fusion methods and the regression part of ATPRK (in terms of the RMSE), while Table 8 compares the accuracy between DF-SSF and ATPRK. It can be observed that the RMSE of the regression part of ATPRK is larger than that of the existing spatial-spectral fusion methods and the accuracy is generally smaller than the DF-SSF-based methods. This suggests that the DF-SSF method can take full advantage of existing spatial-spectral fusion methods by using them as the primary step to reduce the coarse residuals, which is more competitive than the regression step in ATPRK.

Table 7. RMSE of the seven spatial-spectral fusion methods and regression modeling of ATPRK (R-ATPRK) for the Gale crater region.

	GS	GSA	PCA	SFIM	MTF-GLP	MTF-GLP -HPM	GFPCA	R-ATPRK
RMSE	0.0062	0.0054	0.0063	0.0045	0.0047	0.0047	0.0041	0.1864

Table 8. Accuracy evaluation of DF-SSF and ATPRK for the Gale crater region.

Methods	CC	SAM	RMSE	ERGAS	UIQI
DF-GS	0.9793	0.0033	0.0034	0.5700	0.9793
DF-GSA	0.9782	0.0033	0.0035	0.5846	0.9781
DF-PCA	0.9794	0.0033	0.0034	0.5702	0.9793
DF-SFIM	0.9819	0.0031	0.0031	0.5258	0.9819
DF-MTF-GLP	0.9801	0.0032	0.0033	0.5535	0.9801
DF-MTF-GLP-HPM	0.9800	0.0032	0.0033	0.5560	0.9800
DF-GFPCA	0.9873	0.0029	0.0026	0.4349	0.9872
ATPRK	0.9785	0.0033	0.0034	0.5792	0.9785

5.4. Generalization ability of DF-SSF

In this paper, DF-SSF was proposed for Mars remote sensing data with complex surface configurations and examined on seven classical spatial-spectral fusion methods of CS and MRA and also an unsupervised deep learning-based method. By applying the DF-SSF model to these methods, the fusion accuracy is increased noticeably, demonstrating the generalization ability of DF-SSF for existing methods. In this section, we also further examined the generalization ability of DF-SSF from three aspects: 1) extension to spatial-spectral fusion of Earth observation data, 2) extension to the more challenging fusion of hyperspectral and multispectral images and 3) application to more fusion methods. Accordingly, a GF-5 hyperspectral dataset covering an urban area in Shanghai, China was used. The spatial resolution is 30 m. The spatial size is 300×300 pixels, and the number of VNIR bands used is 150. A four-band 30 m multispectral image was synthesized by averaging 37 (or 38) consecutive bands of the 150 bands. Moreover, a 120 m coarse hyperspectral image was simulated by spatially degrading the 30 m hyperspectral image with a zoom factor of four. The three images of the study area are shown in Fig. 13. The task is to fuse the 120 m hyperspectral image with the four-band, 30 m multispectral image to reconstruct the 30 m hyperspectral image. For CS and MRA-based methods, two types of schemes (i.e., the selected band and synthesized band schemes) were considered for using multiple bands of the 30 m multispectral image (Butera et al., 2015). In addition, the matrix factorization (i.e., CNMF) method was also examined. The accuracy evaluation results are exhibited in Table 9. It is clear that the accuracy is increased by using the DF-SSF method. This indicates that the method proposed in this paper is also applicable to the three cases listed above.

In future, DF-SSF can be applied to more up-to-date spatial-spectral fusion methods. For example, supervised learning-based spatial-spectral fusion methods have received increasing attention in recent years (Sun et al., 2021; Ren et al., 2022). Such methods require a large amount of training data (particularly for deep learning-based versions), which is harder to obtain for the Mars remote sensing data studied in this paper, thus, preventing validation of these methods. In future research, the generalization ability of DF-SSF for supervised learning-based spatial-spectral fusion methods can be investigated based on Earth observation data.

(a) (b) (c)

Fig. 13. The GF-5 hyperspectral image used for test. (a) 30 m GF-5 hyperspectral image (300×300 pixels; bands 150, 39 and 3 as RGB). (b) 30 m multispectral image (300×300 pixels; bands 4, 2 and 1 as RGB) simulated by degrading (a) spectrally. (c) 120 m hyperspectral images (75×75 pixels; bands 150, 39 and 3 as RGB) simulated by degrading (a) spatially.

Table 9. Quantitative assessment of the spatial-spectral fusion methods for the GF-5 data set.

	Methods	CC		SAM		RMSE		ERGAS		UIQI	
		Original	DF-based	Original	DF-based	Original	DF-based	Original	DF-based	Original	DF-based
Selected band	GS	0.8992	0.9447	0.0468	0.0352	0.0115	0.0072	2.5827	1.7943	0.8578	0.9300
	GSA	0.9957	0.9965	0.0176	0.0160	0.0020	0.0018	0.4954	0.4451	0.9955	0.9964
	PCA	0.9000	0.9251	0.0476	0.0387	0.0111	0.0075	2.5747	1.9875	0.8470	0.9091
	SFIM	0.9612	0.9675	0.0377	0.0314	0.0086	0.0068	1.8736	1.5024	0.9314	0.9590
	MTF-GLP	0.9749	0.9782	0.0310	0.0266	0.0067	0.0055	1.4832	1.2454	0.9571	0.9708
	MTF-GLP-HPM	0.9756	0.9790	0.0315	0.0269	0.0065	0.0053	1.4553	1.2188	0.9584	0.9717
	GFPCA	0.9071	0.9226	0.0612	0.0421	0.0147	0.0101	3.1901	2.2263	0.7558	0.9067
Synthesized band	GS	0.9010	0.9459	0.0447	0.0337	0.0114	0.0071	2.5675	1.7746	0.8602	0.9313
	GSA	0.9981	0.9984	0.0087	0.0081	0.0013	0.0012	0.3115	0.2854	0.9981	0.9984
	PCA	0.9004	0.9262	0.0460	0.0373	0.0110	0.0073	3.4169	1.9711	0.8486	0.9103
	SFIM	0.9628	0.9687	0.0357	0.0294	0.0085	0.0067	1.8513	1.4786	0.9331	0.9602
	MTF-GLP	0.9768	0.9796	0.0282	0.0241	0.0065	0.0054	1.4499	1.2135	0.9591	0.9722
	MTF-GLP-HPM	0.9774	0.9803	0.0288	0.0245	0.0063	0.0052	1.4216	1.1863	0.9603	0.9731
	GFPCA	0.9047	0.9217	0.0596	0.0411	0.0146	0.0101	3.1860	2.2267	0.7536	0.9057
	CNMF	0.9864	0.9915	0.0214	0.0142	0.0043	0.0026	0.9973	0.6317	0.9787	0.9905

5.5. Applicability of the 6 m CRISM data

Compared to 18 m CRISM data, 6 m CRISM data present significant advantages. First, from the spatial aspect, 6 m CRISM data can provide more detailed texture information. Second, from the spectral aspect, 6 m CRISM data fully inherit the fine spectral resolution (i.e., 6.55 nm) of the original 18 m CRISM data, which is much finer than that of multispectral sensors. The hyperspectral data can capture more detailed spectral characteristics of the Martian surface, offering broader application prospects. For example, when spacecraft select suitable landing sites on the Martian surface in the future, if some small impact craters, canyons and ditches are smaller than 18 m, they are often represented as mixed pixels in 18 m CRISM data, making it difficult to determine the fine topography and geomorphology. However, in 6 m spatial resolution CRISM data, these small-sized features can be more effectively identified. Additionally, the Martian surface is covered by complex geological and topographical characteristics. At 18 m spatial resolution, the texture details of mineral-rich areas cannot be observed clearly. For example, the boundaries of impact craters (e.g., in the Eberswalde crater area) and canyons (e.g., the Gale crater area) appear blurred, and weathering layers and river sediments are difficult to identify. In the 6 m CRISM data, the textures of these features can be recognized more effectively, with great potential to enhance the accuracy of mineral identification.

6. Conclusion

The CRISM can capture hyperspectral images spanning multiple spectral channels from the visible to the near infrared, offering unique advantages for studying minerals, geology, and surface features on Mars. However, the spatial resolution of CRISM data is 18 m, which may be relatively coarse for observing surface texture details in local areas. Here, by spatial-spectral fusion, 18 m CRISM hyperspectral data were downscaled to 6 m, using 6 m CTX images as the panchromatic image. To address the challenge of data fidelity to the original coarse hyperspectral data in spatial-spectral fusion, a novel paradigm called DF-SSF was proposed. It utilizes ATPK to downscale the difference (i.e., coarse residual image) between the spatial-spectral fusion result of any existing method and the original hyperspectral images. The produced fine spatial resolution residual image is then added to the spatial-spectral fusion result of the existing method to yield a spatial-spectral fusion result with complete data fidelity to the original hyperspectral image. The experimental results in six regions show that the fusion of CRISM and CTX can result in finer spatial resolution CRISM hyperspectral images with satisfactory accuracy (based on DS-SSF, the CC is above 0.9600 in all cases). Furthermore, by applying the DF-SSF model to existing spatial-spectral fusion methods, complete data fidelity to the original CRISM hyperspectral original data can be achieved, increasing the fusion accuracy of the existing methods. DF-SSF is theoretically applicable to any existing spatial-spectral fusion methods. Compared with the 18 m CRISM data, the 6 m CRISM data not only provide more detailed texture information, but also inherit the fine spectral resolution, offering broader application potential.

Acknowledgements

This research was supported by the National Natural Science Foundation of China under Grants 42221002, 42222108 and 42171345.

References

Aiazzi, B., Alparone, L., Baronti, S., Garzelli, A., Selva, M., 2006. MTF-tailored multiscale fusion of high-resolution MS and Pan imagery. Photogrammetric Engineering & Remote Sensing 72, 591-596.

Aiazzi, B., Baronti, S., Selva, M., 2007. Improving component substitution pansharpening through multivariate regression of MS + Pan data. IEEE Transactions on Geoscience and Remote Sensing 45, 3230-3239.

Atkinson, P.M., 2013. Downscaling in remote sensing. International journal of applied earth observation and geoinformation 22, 106-114.

Atkinson, P.M., Pardo-Iguzquiza, E., Chica-Olmo, M., 2008. Downscaling cokriging for super-resolution mapping of continua in remotely sensed images. IEEE Transactions on Geoscience and Remote Sensing 46, 573-580.

Bennett, K.J., Wang, J., Scholes, D., 2014. Accessing PDS Data in Pipeline Processing and Websites Through PDS Geosciences Orbital Data Explorer's Web-Based API (REST) Interface, in: 45th Annual Lunar and Planetary Science Conference. pp. 1026.

Berne, O., Helens, A., Pilleri, P., Joblin, C., 2010. Non-negative matrix factorization pansharpening of hyperspectral data: an application to mid-infrared astronomy, in: workshop on hyperspectral image & signal processing: Evolution in Remote Sensing. pp. 1-4.

Butera, F., Selva, M., Chiarantini, L., Aiazzi, B., Baronti, S., 2017. Hyper-sharpening: a first approach on SIM-GA data. IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing 8, 3008-3024.

Chavez, P., Sides, S.C., Anderson, J.A., 1991. Comparison of three different methods to merge multiresolution and multispectral data Landsat TM and SPOT panchromatic. Photogrammetric Engineering and Remote Sensing 57, 295-303.

Fang, F., Li, F., Shen, C., Zhang, G., 2013. A variational approach for pansharpening. IEEE Transactions on Image Processing 22, 2822-2834.

Ghamisi, P., Rasti, B., Yokoya, N., Wang, Q., Hofle, B., Bruzzone, L., Bovolo, F., Chi, M., Anders, K., Gloaguen, R., 2019. Multisource and multitemporal data fusion in remote sensing: a comprehensive review of the state of the art. IEEE Geoscience and Remote Sensing Magazine 7, 6-39.

Goudge, T.A., Mustard, J.F., Head, J.W., Fassett, C.I., Wiseman, S.M., 2015. Assessing the mineralogy of the watershed and fan deposits of the Jezero crater paleolake system, Mars. Journal of Geophysical Research: Planets 120, 775-808.

Kitanidis, P.K., 1993. Generalized covariance functions in estimation. Mathematical Geology 25, 525-540.

Kyriakidis, P.C., 2004. A geostatistical framework for area-to-point spatial interpolation. Geographical Analysis 36, 259-289.

Kyriakidis, P.C., Yoo, E.-H., 2005. Geostatistical prediction and simulation of point values from areal data. Geographical Analysis 37, 124-151.

Laben, C.A., Brower, B.V., 2000. Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening. Google Patents. U.S. Patent 6011875.

Lee, J., Lee, C., 2009. Fast and efficient panchromatic sharpening. IEEE Transactions on Geoscience and Remote Sensing 48, 155-163.

Liao, W., Huang, X., Van Coillie, F., Gautama, S., Pižurica, A., Philips, W., Liu, H., Zhu, T., Shimoni, M., Moser, G., 2015. Processing of multiresolution thermal hyperspectral and digital color data: outcome of the 2014 IEEE GRSS data fusion contest. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 8, 2984-2996.

Liu, J.G., 2000. Smoothing filter-based intensity modulation: A spectral preserve image fusion technique for improving spatial details. International Journal of Remote Sensing 21, 3461-3472.

Loncan, L., De Almeida, L.B., Bioucas-Dias, J.M., Briottet, X., Chanussot, J., Dobigeon, N., Fabre, S., Liao, W., Licciardi, G.A., Simoes, M., 2015. Hyperspectral pansharpening: A review. IEEE Geoscience and remote sensing magazine 3, 27-46.

Ma, J., Yu, W., Chen, C., Liang, P., Guo, X., Jiang, J., 2020. Pan-GAN: An unsupervised pan-sharpening method for remote sensing image fusion. Information Fusion 62, 110-120.

Malin, M.C., Bell III, J.F., Cantor, B.A., Caplinger, M.A., Calvin, W.M., Clancy, R.T., Edgett, K.S., Edwards, L., Haberle, R.M., James, P.B., 2007. Context camera investigation on board the Mars Reconnaissance Orbiter. Journal of Geophysical Research: Planets 112. E5 (2007):E05S04.

Masi, G., Cozzolino, D., Verdoliva, L., Scarpa, G., 2016. Pansharpening by convolutional neural networks. Remote Sensing 8, 594.

Murchie, S., Arvidson, R., Bedini, P., Beisser, K., Bibring, J.-P., Bishop, J., Boldt, J., Cavender, P., Choo, T., Clancy, R.T., 2007. Compact reconnaissance imaging spectrometer for Mars (CRISM) on Mars reconnaissance orbiter (MRO). Journal of Geophysical Research: Planets 112 (E5).

Pardo-Igúzquiza, E., Chica-Olmo, M., Atkinson, P.M., 2006. Downscaling cokriging for image sharpening. Remote Sensing of Environment 102, 86-98.

Peulvast, J.P., Masson, P.L., 1993. Melas Chasma: Morphology and tectonic patterns. Earth, Moon, and Planets 61, 219-248.

Poulet, F., Carter, J., Bishop, J.L., Loizeau, D., Murchie, S.M., 2014. Mineral abundances at the final four curiosity study sites and implications for their formation. Icarus 231, 65-76.

Qu, J., Dong, W., Li, Y., Hou, S., Du, Q., 2023. An interpretable unsupervised unrolling network for hyperspectral pansharpening. IEEE Transactions on Cybernetics 53, 7943-7956.

Qu, J., Liu, X., Dong, W., Liu, Y., Zhang, T., Xu, Y., Li, Y., 2024. Progressive multi-iteration registration-fusion co-optimization network for unregistered hyperspectral image super-resolution. IEEE Transactions on Geoscience and Remote Sensing 62, 1-14.

Ranchin, T., Wald, L., 2000. Fusion of high spatial and spectral resolution images: The ARSIS concept and its implementation. Photogrammetric Engineering and Remote Sensing 66, 49-61.

Ren, K., Sun, W., Meng, X., Yang, G., Peng, J., Huang, J., 2022. A locally optimized model for hyperspectral and multispectral images fusion. IEEE Transactions on Geoscience and Remote Sensing 60, 1-15.

Sales, M.H.R., Souza, C.M., Kyriakidis, P.C., 2012. Fusion of MODIS images using kriging with external drift. ieee Transactions on Geoscience and Remote Sensing 51, 2250-2259.

Shah, V.P., Younan, N.H., King, R.L., 2008. An efficient pan-sharpening method via a combined adaptive PCA approach and contourlets. IEEE Transactions on Geoscience and Remote Sensing 46, 1323-1335.

Shettigara, V.K., 1992. A generalized component substitution technique for spatial enhancement of multispectral images using a higher resolution data set. Photogrammetric Engineering and Remote Sensing 58, 561-567.

Song, H., Liu, Q., Wang, G., Hang, R., Huang, B., 2018. Spatiotemporal satellite image fusion using deep convolutional neural networks. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 11, 821-829.

Sucharski, T., Mapel, J., jcwbacker, Kristin, Lee, K., AgoinsUSGS, Shepherd, M., Jr, C.R.C., dcookastro, Stapleton, S., Becker, K.J., Rodriguez, K., Sides, S., Cole, jusflag, Wilson, T., acpaquette, Williams, K., jlaura, kledmundson, paarongiroux, Milazzo, M., isis3mgr, Beyer, R., Adoram-Kershner, L., ahowington, jshinaman, AustinSanders, Rsaleh57, lwellerastro, 2020. USGS-Astrogeology/ISIS3: ISIS4.1.0 Public Release.

Sun, W., Ren, K., Meng, X., Xiao, C., Peng, J., 2021. A band divide-and-conquer multispectral and

hyperspectral image fusion method. IEEE Transactions on Geoscience and Remote Sensing 60, 1-13.
Sun, W., Ren, K., Meng, X., Yang, G., Li, C., Huang, K., Li, J., Huang, J., 2024. Domain transform model driven by deep learning for anti-noise hyperspectral and multispectral image fusion. IEEE Transactions on Geoscience and Remote Sensing 62, 1-17.
Sun, W., Ren, K., Meng, X., Yang, G., Peng, J., Li, J., 2023. Unsupervised 3-D tensor subspace decomposition network for spatial-temporal-spectral fusion of hyperspectral and multispectral images. IEEE Transactions on Geoscience and Remote Sensing 61, 1-17.
Thomas, C., Ranchin, T., Wald, L., Chanussot, J., 2008. Synthesis of multispectral images to high spatial resolution: a critical review of fusion methods based on remote sensing physics. IEEE Transactions on Geoscience and Remote Sensing 46, 1301-1312.
Vivone, G., Alparone, L., Chanussot, J., Dalla Mura, M., Garzelli, A., Licciardi, G.A., Restaino, R., Wald, L., 2014. A critical comparison among pansharpening algorithms. IEEE Transactions on Geoscience and Remote Sensing 53, 2565-2586.
Wald, L., Ranchin, T., Mangolini, M., 1997. Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images. Photogrammetric Engineering and Remote Sensing 63, 691-699.
Wang, Q., Shi, W., Atkinson, P.M., 2016. Area-to-point regression kriging for pan-sharpening. ISPRS journal of Photogrammetry and Remote Sensing 114, 151-165.
Wang, Q., Shi, W., Atkinson, P.M., Zhao, Y., 2015. Downscaling MODIS images with area-to-point regression kriging. Remote Sensing of Environment 166, 191-204.
Wang, Q., Shi, W., Zhang, H., 2014. Class allocation for soft-then-hard subpixel mapping algorithms with adaptive visiting order of classes. IEEE Geoscience and Remote Sensing Letters 11, 1494-1498.
Wang, Z., Bovik, A.C., 2002. A universal image quality index. IEEE Signal Processing Letters 9, 81-84.
Yang, J., Fu, X., Hu, Y., Huang, Y., Ding, X., Paisley, J., 2017. PanNet: a deep network architecture for pan-sharpening, in: proceedings of the IEEE international conference on computer vision. pp. 1753-1761.
Yang, J., Zhao, Y.-Q., Chan, J.C.-W., 2018. Hyperspectral and multispectral image fusion via deep two-branches convolutional neural network. Remote Sensing 10, 800.
Yang, Y., Su, Z., Huang, S., Wan, W., Tu, W., Lu, H., 2022. Survey of deep-learning approaches for pixel-level pansharpening. National Remote Sensing Bulletin 26(12）, 2411-2432.
Yokoya, N., Member, S., IEEE, Yairi, T., Iwasaki, A., 2012. Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion. IEEE Transactions on Geoscience and Remote

Sensing 50, 528-537.
Zhang, L.P., Shen, H.F., 2016. Progress and future of remote sensing data fusion. J. Remote Sensing 20, 1050-1061.
Zurek, R.W., Smrekar, S.E., 2007. An overview of the Mars Reconnaissance Orbiter (MRO) science mission. Journal of Geophysical Research: Planets 112 (E5).