MarsMapNet: A Novel Superpixel-Guided Multi-view Feature Fusion Network for Efficient Martian Landform Mapping

Hui Zhao, Sicong Liu, *Senior Member, IEEE*, Xiaohua Tong, *Senior Member*, *IEEE*, Qian Du, *Fellow*, *IEEE*, Lorenzo Bruzzone, *Fellow, IEEE,* Kecheng Du, Jie Zhang, Xuanning Lu

***Abstract*—Landform classification and mapping of the Martian surface using Mars orbiter images can provide an important reference for landing site selection and rovers' traversability evaluation in Mars exploration. Moreover, specific Martian landforms are closely associated with the evidences of water-related activities and Martian life, thus have crucial research importance. This article proposes a novel superpixel-guided multi-view feature fusion network (MarsMapNet) for efficient mapping of the Martian landforms. In particular, the proposed MarsMapNet first generates the superpixel-level segments from Mars orbiter images by considering local morphological homogeneity of landforms. Then a multi-view feature extraction and fusion (MVF) network is developed, where abstract convolutional features are extracted based on scene-level patches, and multi-textures are extracted based on local landform from shallow-to-deep feature learning. After the network being trained on scene-level samples and guided by the superpixel segmentation, Martian landforms can be correctly classified in an efficient way, whose mapping time cost sharply decreased when compared to the reference methods. The proposed MarsMapNet has been validated on three real landing sites from several Mars missions (*i.e.*, the Jezero Crater, the Southern Utopia Planitia, the Oxia Planum) by using the Mars Reconnaissance Orbiter's Context Camera (CTX) images. Qualitative and quantitative analysis on the obtained experimental results confirm the effectiveness and efficiency of the proposed MarsMapNet when compared with the state-of-the-art methods, demonstrating its potential for supporting a Martian global landform mapping in the future.**

***Index Terms*—Martian landform mapping, Mars exploration, CTX images, superpixel segmentation, shallow and deep features, feature fusion.**

This work was supported in part by the National Natural Science Foundation of China under Grants 42221002, 42241130, 42071324, and by the Shanghai Rising-Star Program (21QA1409100) (*Corresponding author: Sicong Liu*)

Hui Zhao, Sicong Liu, Xiaohua Tong, Kecheng Du, Jie Zhang and Xuanning Lu are with the College of Surveying and Geoinformatics, Tongji University, Shanghai, 200092, China (e-mail: sicong.liu@tongji.edu.cn; zhaohui@tongji.edu.cn; xhtong@tongji.edu.cn; kecheng_du@tongji.edu.cn; zhangjie22@tongji.edu.cn; lu0707@tongji.edu.cn).

Qian Du is with the Department of Electrical and Computer Engineering, Mississippi State University, Starkville, MS 39762, USA (e-mail: du@ece.msstate.edu).

Lorenzo Bruzzone is with the Department of Information Engineering and Computer Science, University of Trento, I-38123, Italy (e-mail: lorenzo.bruzzone@unitn.it).

I. Introduction

Martian surface exhibits various morphological features. Such surface landforms were caused by different types and degrees of geological processes in different historical periods of Mars, and shaped by water, winds, lava flow and seasonal phenomena [1-3]. Study of the Martian landform features can help us gain insight into the geological background, evolutionary history, global climate change, water ice, water availability and potentially habitable environment [4-9]. It is noteworthy that mapping the Martian surface's landform is an essential and preliminary step for investigating the scientific goals, engineering constraints in Mars exploration landing site selection [10]. The Martian surface features, such as slope, roughness and surface load-bearing, are highly related to the landing safety of Mars rovers, and the terrain type is essential for evaluating the traversability for rovers [11]. On the other hand, the spatial distribution of water/ice, mineral and biosignature preservation related landforms can make further contributions for scientific constraints on the candidate landing site selection or sample collection [10].

Many studies have been conducted in order to identify a specific type of Martian landform, such as *impact craters* [12], *dunes* [13], *dark slope streaks* [14], *volcanic rootless cones* and *transverse aeolian ridges* [15, 16], *CO_2 jet deposits* [17] and *chaos terrains* [18]. The morphology and distribution, together with material composition analysis, of these landforms can be used to reveal their causes. In addition, there are works focusing on the multi-class landform classification problem. In the earlier studies, the digital elevation model (DEM) constructed by the Mars Orbiter Laser Altimeter (MOLA) data was the main data source used for extracting terrain attribute features and then implementing an unsupervised Martian terrain classification [19-21].

The topography maps generated by DEM support the analysis of engineering requirements for Mars exploration, but they mainly reflect features such as slope, roughness, and undulations. Due to the coarse spatial resolution of DEM products, fine-scale landform types cannot be properly represented and distinguished. The high spatial resolution images acquired by Mars orbiter cameras offer a great opportunity to identify a variety of Mars landforms at a finer scale. Studies have been conducted by using different Mars Mission orbiter images, such as the Mars Reconnaissance Orbiter' Context camera (CTX) [22] and the High Resolution

Imaging Science Experiment (HiRISE) camera [23], the Mars Express high-resolution stereo camera (HRSC) [24], the High Resolution Imaging Camera (HiRIC) in China's First Mars Exploration Tianwen-1 Mission [25], and the Mars Global Surveyor's Mars Orbiter Camera (MOC) [26].

Most of the current landform classification methods for Martian orbiter images were built upon supervised classification schemes. Such methods can be divided into three main categories, including manual visual interpretation, traditional machine learning, and deep learning. Planet Four: Terrains (P4T) is an online citizen science project, where the general public was enlisted to identify and manually label seasonally sculpted terrains in the South Polar region of Mars using CTX images. The terrains include *araneiform*, *swiss cheese terrain* and *craters* [27]. In [28], firstly three classification systems were created: four categories based on terrain features, two categories based on regions exhibiting aeolian deflation, and six categories based on geologic classification distilled from USGS maps. Then multi-source data including HRSC and MOLA DEM were segmented by using the standard normalized cuts algorithm, where multiple distinguishing features for each superpixel were extracted to form a feature vector. Finally, a traditional classification method, such as the Naïve Bayes classifier or the boosting classifier, was used to classify the extracted features.

In recent years, a growing number of advanced deep learning methods have been applied to Martian landform classification. To help users to automatically discover Mars images of interest from the Planetary Data System (PDS) image atlas, the AlexNet trained on Earth images was fine-tuned and used for classifying twenty-four classes of Mars rover images and six classes (*i.e.*, *craters*, *bright sand dunes*, *dark sand dunes*, *dark slope streaks*, *other* and *edge*) of HiRISE orbital images [29]. Based on the same method developed in [29], three new classes of interest, which are *impact ejecta*, *spider*, and *swiss cheese* were added in [30]. A novel software with the capability of Soil Property and Object Classification (SPOC) was proposed for classifying HiRISE orbital images and the Curiosity's Navigation Camera images. The terrain classification model of SPOC was built based on the fully-convolutional neural networks - DeepLab, and 17 terrain classes were identified for eight candidate landing sites of the Mars 2020 Rover (M2020) mission [31]. Then the obtained terrain classification result was used to analyze the traversability of eight candidate landing sites [32]. In [33], a specialized terrain classification system named "Novelty or Anomaly Hunter - HiRISE" (NOAH-H) with DeepLab model was used to classify HiRISE images of two ExoMars candidate landing sites, *i.e.*, Oxia Planum and Mawrth Vallis, and 14 descriptive classes were identified for the traversability analysis. In our previous work, we proposed a VGG (Visual Geometry Group) - Like network for topography mapping of the China's First Mars Mission Tianwen-1 landing area by using the high resolution HiRISE and HiRIC orbiter images, where three landforms (*i.e.*, *smooth regolith*, *rough surface* and *dunes*) were automatically identified along the Zhurong rover's routing path [34]. In [35], a DoMars16k data set that contains 15 landforms commonly found on the Martian surface was constructed by using CTX images, and a pre-trained Densenet161 was used to map Martian landforms based on a pixel-by-pixel sliding-window approach. Finally, the classification map was smoothed by using the Markov Random fields (MRFs) method.

Note that the classification systems for Mars landforms usually differ according to the specific application purpose and the considered region at a given spatial resolution of the orbiter image. A relatively comprehensive class system can be found in [31, 33], but the sample annotation is pixel-based. Pixel-level annotation means a polygon is used to outline the landform along its distinct boundary, and it is labor intensive and also requires domain expert knowledge. In [35], a DoMars16k data set with 15 landforms and 16,150 CTX sample images was built. The sample annotation in DoMars16k data set is a square scene-level patch (e.g., 200×200 pixels) that labeled as a given landform class. Since it is a scene-level data set, its annotation is much easier than pixel-based samples as in [31, 33]. In this work, we use the publicly available scene-level DoMars16k samples to validate the proposed novel approach.

On the basis of the aforementioned analysis, some open issues observed in the current literature need to be further addressed. **First**, natural images usually contain rich color or spectral information that is suitable for identifying different objects. The popular deep learning networks were initially proposed for three-band (*i.e.*, RGB) natural images. However, the high-resolution Martian orbiter camera data (*e.g.*, CTX, HiRISE, HiRIC) are all single-band and gray-scale panchromatic images. Thus, directly using the pre-trained networks based on natural smooth images (*e.g.*, ImageNet) on the single-band Martian images as in the existing works may reduce the effectiveness of classification performance. Moreover, the unique textural information of various landform types has not been adequately exploited in the current developed works. **Second**, classification methods developed in the literature used spatial windows [34, 35]. The overlapped patches between adjacent pixels inevitably increases the data processing burden and the computational cost, and may produce salt-and-pepper noises in the mapping results. Despite that some post-processing techniques such as MRFs can be used to smooth the mapping results [35], it is critical to develop an efficient automatic classification method, especially when dealing with a large Martian area. **Third**, in practical applications we require a classification map with detailed landform boundaries. Therefore, the fine pixel-based samples were labeled and used by semantic segmentation model (*i.e.*, DeepLab) to obtain a classification map with detailed landform boundaries [31, 33]. However, the annotation of pixel-based samples are more time-consuming compared to scene-level samples (*e.g.*, those in [35]). How to map Martian landforms with boundaries between them by using scene-level samples becomes a critical task.

In order to address the aforementioned open issues and challenges, in this work a superpixel-guided multi-view feature fusion network (denoted as the MarsMapNet) is

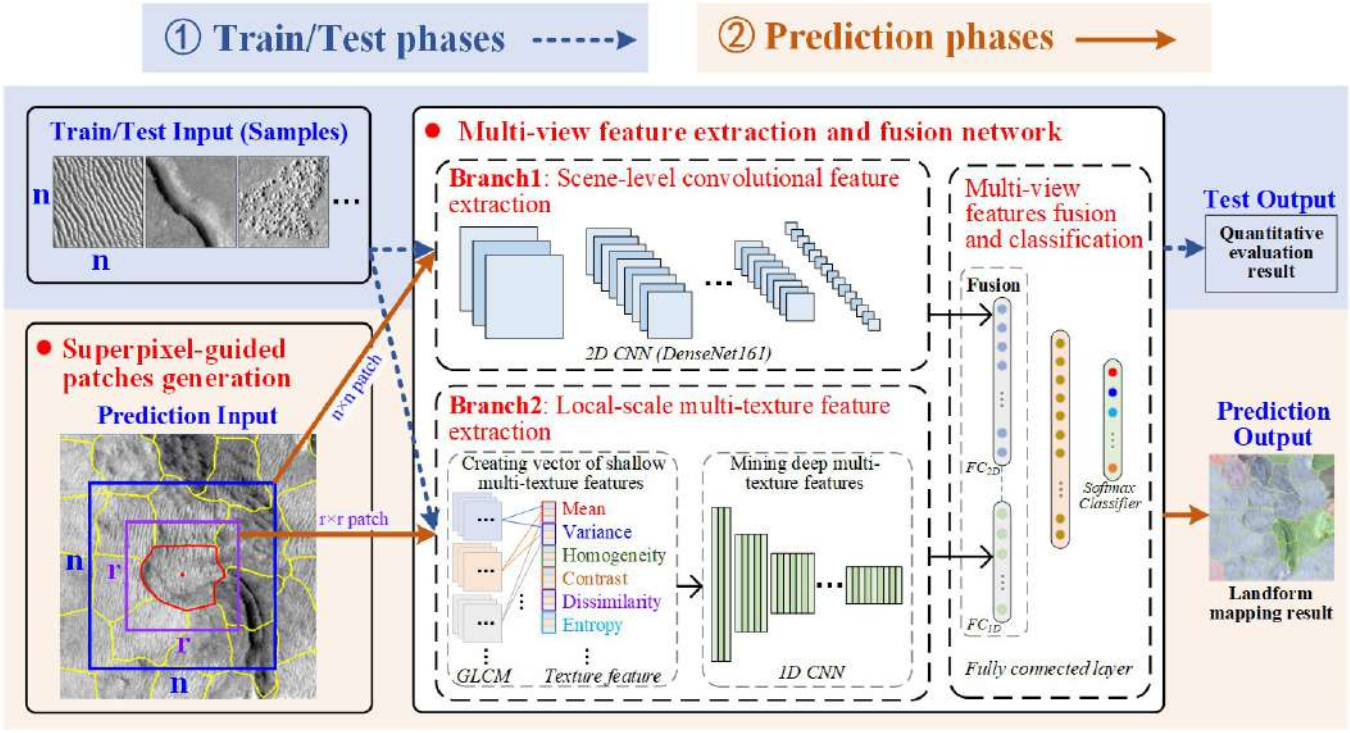

Fig. 1 Block diagram of the proposed MarsMapNet approach.

proposed for efficient Martian landform mapping. The DoMars16k samples [35] were used for network training and accuracy testing of the proposed approach, and experimental results obtained on three Martian landing areas including the Jezero Crater (*i.e.*, the Mars Perseverance rover landing region), the Southern Utopia Planitia (*i.e.*, the Tianwen-1 Zhurong rover landing region), and the Oxia Planum (*i.e.*, the prospective ExoMars landing region) confirm the effectiveness of the proposed approach. The main contributions of this paper are highlighted as follows.

1) A novel multi-view feature extraction and fusion (MVF) network is developed. The proposed MVF is designed to fully extract the multi-texture features of Martian landform at local scales and convolutional features at a scene-level. Then the extracted multi-view features are fused to offer a more comprehensive representation of Martian landforms even on the single-band and gray-level orbiter images. This allows a more robust exploitation of shallow-to-deep and abstract features for complex Martian landforms, and finally leads to more accurate mapping results. To the best of our knowledge, there is no similar work in literature to solve the same problem.

2) A new superpixel-guided patch generation process is designed. In the prediction process, patches are generated according to the guidance of superpixels, and then are incorporated with the proposed MVF network. This alleviates the salt-and-pepper noise problem [34, 35], and also reduces the computational cost (by about 124 times, 99 times and 112 times on three datasets, respectively) when compared to the state-of-the-art (SOTA) pixel-wise classification method. Obtained results exhibit good local homogeneity of landforms, as well as accurate representation of their discriminable boundaries. Note that no additional post-processing is required to smooth the results as in the SOTA methods.

3) The proposed MarsMapNet is easy to be used with a good sample augmentation capability. Since it allows the classification by using the more accessible scene-level samples when compared with pixel-based samples. This greatly reduces the annotation difficulty when dealing with the complex Martian scenarios (*i.e.*, low-texture, various illumination condition regions), and makes sample augmentation easier in practical applications.

The rest of this paper is organized as follows. The proposed approach is described in detail in Section II. Experimental results and analysis are presented in Section III. Finally, Section IV draws the conclusions.

II. Proposed MarsMapNet

The proposed MarsMapNet aims to automatically classify the Martian landforms and generate efficiently a mapping result over a large geographical area. Fig. 1 shows the block diagram of the proposed approach, which mainly consists of two parts: 1) Superpixel-guided patches generation; 2) Multi-view feature extraction and fusion network.

As shown in Fig. 1, there are two phases in the whole mapping process: ① the Train/Test phase; ② the Prediction phase. In the train/test phase, sample maps of $n \times n$ size are independently inputted into Branch1 and Branch2 of MVF network to train the network and calculate the quantitative evaluation results. In the prediction phase, $n \times n$ and $r \times r$ patches created by the guidance of superpixels are independently inputted into Branch1 and Branch2 of the trained network to generate the landform mapping result. Note

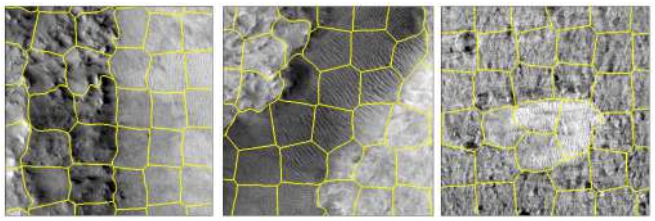

Fig. 2 Examples of superpixel-level segmentation results.

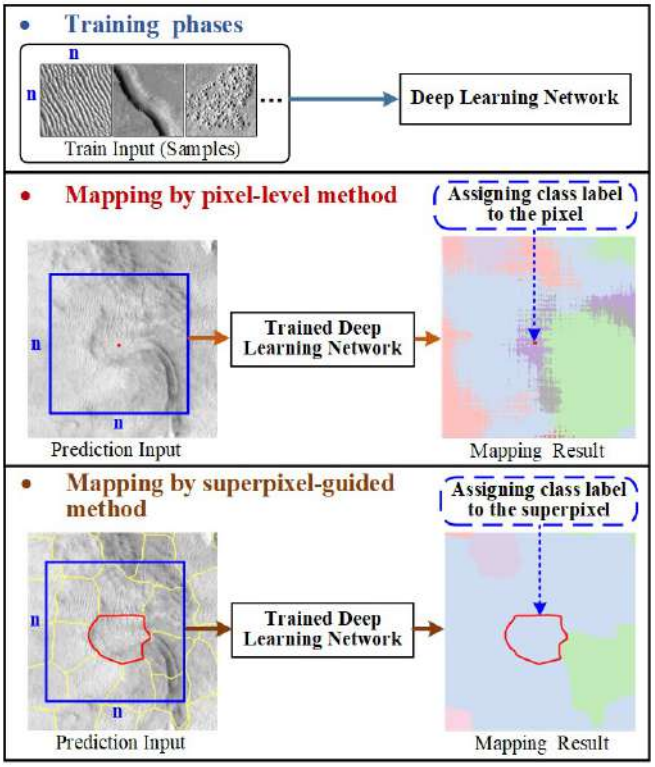

Fig. 3 The comparison of pixel-level and superpixel-guided mapping methods.

that the superpixel segmentation is only used in the prediction phase to guide the patch generation, which is not used in the train/test phase.

A. Superpixel-guided patches generation

The input data for the training process are the manually annotated scene-level samples, where each patch of a sample image represents a given landform category. Those sample images contain homogeneous and regular texture features (*e.g.*, *aeolian straight*, *rough terrain*), or some remarkable features (*e.g.*, *channel*, *crater*) of various landforms. However, we aim to obtain a more detailed landform mapping result with boundaries between different landforms, rather than a scene-level classification result. To this end, superpixel segmentation is performed, and superpixel-level segments with local homogeneity are used as the basic units for classification.

The simple linear iterative clustering (SLIC) [36] is applied to obtain the superpixel-level segments. It adopts the *k*-means clustering approach to generate superpixels efficiently, exhibiting better boundary adhesion performance. In addition, the compactness and number of the superpixels are flexible. SLIC superpixels correspond to clusters in the *lxy* gray-scale image space, where *l* represents the gray-scale value of the

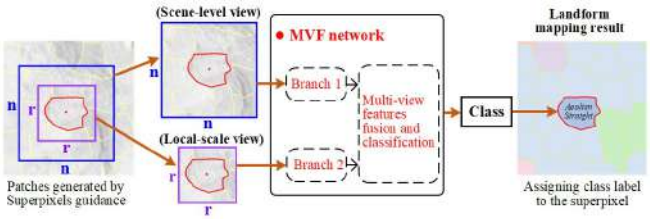

Fig. 4 The flowchart of the label prediction process for a given superpixel.

gray-scale image, and (x,y) are the pixel's position. Each pixel j within a superpixel has the smallest distance D to the clustering center i of that superpixel. This can be expressed as:

$$\begin{aligned} D &= \sqrt{\left(\frac{d_c}{m}\right)^2 + \left(\frac{d_s}{S}\right)^2}, \\ d_c &= \sqrt{(l_j - l_i)^2}, \\ d_s &= \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2}, \end{aligned} \tag{1}$$

where $S = \sqrt{N/K}$, N is the number of pixels, and K is independent of the number of superpixels. The compactness m allows us to weigh the relative importance between color similarity d_c and spatial proximity d_s. The larger the m, the more compact and regular the resulting superpixels, and vice versa. In this work, K is defined according to the size of input image and the size of landforms of interest, which can be empirically formalized as:

$$K = (H \times W)/(b \times b) \tag{2}$$

where H and W are the height and weight of the input image, respectively, and $H \times W = N$. The size of $b \times b$ (65 in the experiment) is pre-defined close to the size of the interested landform in the sample patches. It is obvious that by using the superpixel segmentation, the processing burden can be largely reduced nearly $b \times b$ times compared to the pixel-by-pixel sliding-window approach. Fig. 2 showed some examples of superpixel segments with parameters $m = 0.2$ and $b = 65$. As one can see, the superpixel-level segmentation resulted in the excellent boundary adhesion performance. Nevertheless, there are still some factors affecting the accuracy of segmentation boundaries, such as multi-scale complex landforms and image contrast quality. However, superpixel-level segmentation can achieve the tradeoff between accuracy and efficiency, which is suitable in most of practical application cases.

The superpixel-level segments are generated to guide the classification process. Fig. 3 compares the pixel-level and superpixel-guided mapping methods. It can be seen that the pixel-level mapping result exhibits salt-and-pepper noises, and overlapped patches between adjacent pixels inevitably increase the data processing burden. However, the superpixel-guided classification obtains excellent boundary adhesion performance with less processing burden. The label prediction process for a given superpixel in the proposed approach is shown in Fig. 4. During the prediction phase, when superpixel blocks are generated, patches of size $n \times n$ (see the blue box in Fig. 4) and $r \times r$ (see the purple box in Fig. 4) centered on a given superpixel centroids are created. The size of $n \times n$ (200 × 200) is the same as the sample, and an $n \times n$ patch contains

the superpixel and its context at a scene-level. r = 120 is defined as a tradeoff between b and n to effectively extract local multi-texture of superpixel ($b \times b$) and also to avoid the influence of other landforms in a scene-level patch ($n \times n$). Then, patches with two sizes are input into two branches of the MVF network, respectively, to predict the category value of landform. Finally, the superpixel is given the predicted label.

B. Multi-view feature extraction and fusion network for landform classification

1) Scene-level convolutional feature extraction

Popular CNN architectures, such as AlexNet [37], VGG [38], GoogLeNet [39], Inception [40], ResNet [41], EfficientNet [42] and DenseNet [43] are widely used for scene-level image classification. They can automatically learn abstract and robust convolutional features from the scene-level view. Experimental comparison was conducted in Section III.B to evaluate the performance of several SOTA architectures, demonstrating that DenseNet161 exhibited better classification performance than others. Therefore, the DenseNet161 is used as the basic 2D-CNN deep feature extraction architecture.

In DenseNet161, for each layer, feature maps of all preceding layers are used as inputs, and its own feature maps are given as input to all subsequent layers. Thus, the vanishing-gradient problem is alleviated, while the feature propagation and reuse are enhanced. DenseNet161 contains convolution, pooling, dense block, transition and classification layers. For more details, readers can refer to [43].

2) Local-scale multi-texture feature extraction

The Martian orbiter CTX images are single-band gray-level images, thus the effective mining of their spatial, textural and brightness features becomes an essential prerequisite for a successful classification task. Accordingly, the distinctive textural features exhibited by landforms can be used to distinguish their specific types. To this end, the simple yet effective Gray Level Co-occurrence Matrix (GLCM) [44] is used to extract the shallow multi-texture features for landforms at a local scale. As a popular statistical method for image texture analysis, GLCM defines the probability of gray-level α occurring in the neighborhood of another gray-level β at an offset (Δ_x, Δ_y). Given a training patch I, of size $n \times n$, these probabilities create a co-occurrence matrix G as follows:

$$G(\alpha,\beta)=\sum_{x=1}^{n}\sum_{y=1}^{n}\begin{cases}1, \text{ if } I(x,y) = \alpha \text{ and } I(x+\Delta_x, y+\Delta_y)=\beta \\ 0, \text{ otherwise}\end{cases} \quad (3)$$

where the offset (Δ_x, Δ_y) is expressed by the combination of a distance d and a direction θ [44].

In this paper, nine statistical values including mean, variance, homogeneity, contrast, dissimilarity, entropy, angular second moment, correlation and auto correlation are extracted from the GLCM. The maximum gray-level of the considered CTX image is 256, and we set both α and β as 32 in the experiment, which balances the effectiveness and a proper statistical representation with the computational effort. Four different distances d (*i.e.*, 2, 4, 6, 8) and four different directions θ (*i.e.*, 0°, 45°, 90°, 135°) are used to generate textural features. Finally, feature vector (m_0×1) can be obtained for each input patch image, and $m_0 = 4\times4\times9$.

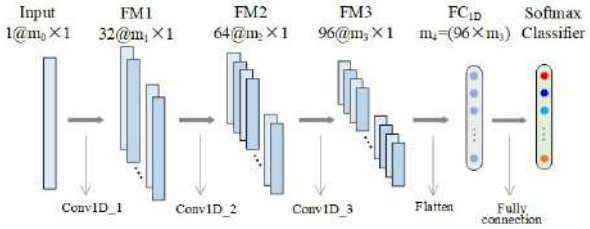

Fig. 5 Architecture of the developed 1D-CNN.

1D-CNN was effective for feature extraction of 1D signals in various application fields [45-47]. In our work, 1D-CNN is used to process 1D shallow artificial features. The architecture of the developed 1D-CNN is illustrated in Fig. 5, which contains three 1D convolution layers (Conv1D) and one fully connected layer. The m_0-dimensional feature vector extracted from each patch sample is inputted to the 1D-CNN for discriminable feature mining. The number of channels for feature maps 1, 2 and 3 (denoted as FM1, FM2 and FM3) are 32, 64 and 96, respectively. Let m_i be a length of vector sequence after the convolution operation, which can be expressed as follows:

$$m_i = \frac{m_{i-1} - KernelSize + 1}{Stride} + 1, \quad i = 1,2,3. \quad (4)$$

where *KernelSize* is the size of the convolution kernel, *Stride* is the step size of the convolution. After the flatten operation, the final number of fully connected layer nodes is $m_4 = 96 \times m_3$.

3) Multi-view features fusion and classification

Fusion of shallow artificial features and deep features has shown its effectiveness in multi-source image classification [48, 49]. In order to offer a more comprehensive representation for Martian landforms from the single-band and gray-scale images, the extracted abstract convolutional features from scene-level view and the multi-texture features from local landform view are fused. The fusion training is divided into two steps. One is to independently train the 2D-CNN in Branch1 and 1D-CNN in Branch2. The other is to fuse the Branch1 and Branch2 that trained in the previous step, and then re-train the fusion model to fine-tune the weights with the idea of transfer learning. The procedure of fine-tune fusion plays a crucial role in achieving better classification performance and constructing a more robust network, which has been validated in Refs. [47] and [50].

In the first step, the 2D-CNN (*e.g.*, DenseNet161) and 1D-CNN architectures are trained independently. Note that for the 2D-CNN branch, the final fully connected layer of DenseNet161 is removed, as it was pre-trained on 1.2 million Earth images from 1,000 classes in the ImageNet data set [51]. A new fully connected layer with the number of Martian landform classes is added, then all parameters are re-trained based on the pretrained parameters. 1D-CNN is trained on the 1D shallow feature vector from scratch.

In the second step, the final classification layer of the trained 2D-CNN and 1D-CNN architecture is removed. Then the FC_{2D} and FC_{1D} are defined as the final fully connected layer of 2D-CNN and 1D-CNN architectures, respectively. As shown in Fig. 1, they are concatenated to generate the fusion layer F_{fusion} as:

$$F_{fusion} = F_{2D} \,||\, F_{1D} \tag{5}$$

where || is the concatenating operation. F_{fusion} is followed by two fully connected layers for classification. Finally, the constructed MVF network is re-trained with the last two fully connected layers being randomly initialized and other layers being initialized using the pre-trained parameters.

III. Experimental Results and Analysis

To evaluate the effectiveness of the proposed MarsMapNet for Martian landform mapping, quantitative and qualitative results obtained on three landing sites (*i.e.*, Jezero Crater, Southern Utopia Planitia, Oxia Planum) were analyzed in detail.

A. Experimental Setup and Parameter Setting

SLIC: Based on multiple trials, the compactness m is set to 0.2, and b is set to 65, thus K can be automatically calculated according to Equation (2).

GLCM: Gray-level α and β are set to 32, four distance values d (2, 4, 6, 8) and four directions θ (0°, 45°, 90°, 135°) are selected to extract the GLCM feature vectors for each image patch.

Model training: The Stochastic Gradient Descent (SGD) optimization is used, where the learning rate is 0.01, the momentum is 0.9, and the batch size is set to 64, the training epoch of 1D-CNN, 2D-CNN and the proposed MVF network are 2000, 50 and 50, respectively.

Landform prediction: The training samples size n is equal to 200 [35]. The patch size r for calculating 1D features is defined as 120.

Algorithms were implemented by using Python, where the deep learning networks were built by using Pytorch. Experiments were carried out on the Ubuntu 18.04.5, with Intel Xeon Gold 6130 CPUs at 2.10 GHz, 159 GB RAM, and GPU of NVIDIA GRID P40-24Q, 22GB.

B. Quantitative Results Analysis on the Public Data Set

The available public DoMars16k data set is used to quantitatively evaluate the classification performance of the proposed approach. The DoMars16k data set is built based on the Mars Reconnaissance Orbiter's CTX images. In the DoMars16k data set, common Martian landforms are divided into five thematic groups, including *Aeolian Bedforms*, *Topographic Landforms*, *Slope Feature Landforms*, *Impact Landforms*, and *Basic Terrain Landforms*. Each group contains some specific classes, and in total there are 15 classes of landforms in the data set, including *Aeolian Curved*, *Aeolian Straight*, *Cliff*, *Ridge*, *Channel*, *Mounds*, *Gullies*, *Slope Streaks*, *Mass Wasting*, *Crater*, *Crater Field*, *Mixed Terrain*, *Rough Terrain*, *Smooth Terrain* and *Textured Terrain*.

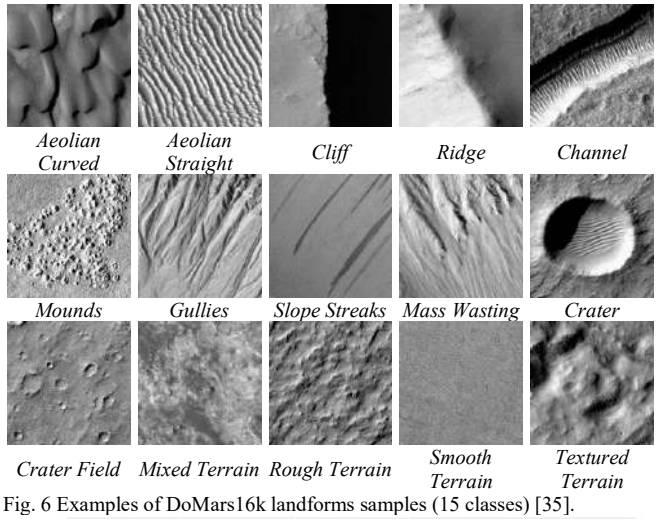

Fig. 6 Examples of DoMars16k landforms samples (15 classes) [35].

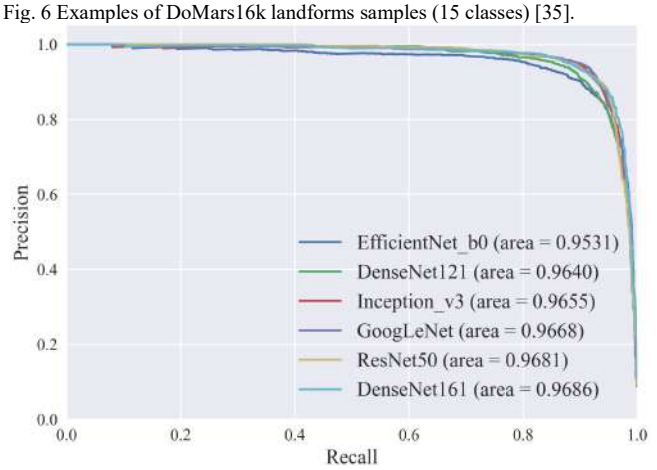

Fig. 7 Micro-average *P-R* curves obtained by different CNN architectures.

Fig. 6 illustrates examples of the considered 15 classes of landforms. The data set consists of 16,150 sample images, which were randomly subdivided into training (70%), validation (20%), and test (10%) sets, respectively. The size of the sample image is 200 × 200 pixels (*i.e.*, 1.2 km × 1.2 km), which is a trade-off between the annotation cost and details of the final mapping result.

To qualitatively analyze the effectiveness of the proposed approach, we compared multiple metrics: Precision (*P*), Recall (*R*), *P-R* curve and *F1-scores* [52]. Detailed formulas of those metrics are listed below:

$$P = \frac{TP}{TP + FP} \tag{6}$$

$$R = \frac{TP}{TP + FN} \tag{7}$$

where *TP* is the counts of True Positives, *FP* is the counts of False Positives, *FN* is number of False Negatives. For binary classification, *F1-score* is defined as the harmonic mean of *P* and *R*:

$$F1\text{-}Score = \frac{2 \times P \times R}{P + R} \tag{8}$$

For the multi-class case, Micro-*F1* equals *OA* (overall accuracy) that is a direct measure when the number of different classes is balanced. Macro-*F1* is more reliable when

Table 1 Comparison of the multiple metrics (%) obtained by different 2D-CNN networks in Branch 1 of the proposed MVF network.

Landform Classes	Branch 1 (2D-CNN)					
	GoogLeNet	EfficientNet_b01	ResNet50	DenseNet121	Inception_v3	DenseNet161
Aeolian Curved	98.56	97.58	99.53	98.58	98.56	98.58
Aeolian Straight	98.04	97.06	98.04	96.12	96.59	97.09
Cliff	89.66	88.89	86.41	88.56	90.29	91.00
Ridge	97.05	96.64	98.29	98.28	98.72	87.13
Channel	98.15	97.67	98.15	97.67	98.62	94.96
Mounds	93.72	95.40	97.85	96.14	96.14	98.02
Gullies	94.53	93.94	96.97	95.05	92.46	92.86
Slope Streaks	90.48	92.09	91.59	91.59	90.48	98.17
Mass Wasting	91.18	91.18	89.62	89.86	92.38	93.40
Crater	83.58	86.54	82.65	85.00	86.73	98.72
Crater Field	92.82	94.58	91.35	93.60	95.61	96.63
Mixed Terrain	97.00	95.00	97.51	96.04	95.57	90.73
Rough Terrain	95.91	94.70	97.36	95.52	96.27	95.10
Smooth Terrain	96.64	95.73	95.00	98.31	95.76	96.23
Textured Terrain	72.34	77.61	72.63	77.78	74.49	77.72
Macro-*P*	92.70	93.00	92.91	93.19	93.25	**93.79**
Macro-*R*	92.77	93.00	93.01	93.28	93.36	**93.83**
Micro-*F1*	92.94	93.12	93.18	93.43	93.49	**93.99**
Macro-*F1*	92.64	92.97	92.86	93.20	93.24	**93.76**

Table 2 Accuracy metrics (%) of the ablation architectures of the proposed MVF network.

Landform Classes	Branch 1 (2D-CNN: DenseNet161)	Branch 2 (GLCM&1D-CNN)	Proposed MVF network
Aeolian Curved	98.58	83.11	99.06
Aeolian Straight	97.09	89.22	97.06
Cliff	91.00	71.58	92.93
Ridge	87.13	71.92	88.56
Channel	94.96	74.46	95.83
Mounds	98.02	87.80	97.54
Gullies	92.86	72.63	95.52
Slope Streaks	98.17	85.98	98.15
Mass Wasting	93.40	61.19	91.94
Crater	98.72	74.26	98.31
Crater Field	96.63	75.40	96.63
Mixed Terrain	90.73	81.48	94.06
Rough Terrain	95.10	89.72	94.63
Smooth Terrain	96.23	85.71	97.05
Textured Terrain	77.72	53.97	81.03
Macro-*P*	93.79	77.20	**94.59**
Macro-*R*	93.83	77.58	**94.61**
Micro-*F1*	93.99	77.45	**94.73**
Macro-*F1*	93.76	77.23	**94.55**

the number of different classes is unbalanced. Micro-*P* and Micro-*R* is calculated based on *TP*, *FP* and *FN* for all classes, and then the Micro-*F1* is calculated based on Micro-*P* and Micro-*R*:

$$\text{Micro-}P = \frac{TP_1 + TP_2 + \cdots + TP_{nc}}{TP_1 + TP_2 + \cdots + TP_{\text{nc}} + FP_1 + FP_2 + \cdots + FP_{nc}} \quad (9)$$

$$\text{Micro-}R = \frac{TP_1 + TP_2 + \ldots + TP_{nc}}{TP_1 + TP_2 + \ldots TP_{nc} + FN_1 + FN_2 + \ldots + FN_{nc}} \quad (10)$$

$$\text{Micro-}F1 = \frac{2 \times \text{Micro-}P \times \text{Micro-}R}{\text{Micro-}P + \text{Micro-}R} \quad (11)$$

Macro-*P* and Macro-*R* are calculated by averaging the *P* and *R* for each class. Macro-*F1* is calculated based on Macro-*P* and Macro-*R*:

$$\text{Macro-}P = \frac{1}{nc}\sum_{i=1}^{nc} P_i \quad (12)$$

$$\text{Macro-}R = \frac{1}{nc}\sum_{i=1}^{nc} R_i \quad (13)$$

$$\text{Macro-F1} = \frac{2 \times \text{Macro-}P \times \text{Macro-R}}{\text{Macro-}P + \text{Macro-R}} \quad (14)$$

where *nc* is the number of classes.

Table 1 provides the obtained *F1-Scores* for each landform class, as well as the Macro-*P*, Macro-*R*, Micro-*F1* and Macro-*F1*. The results were obtained on the DoMars16k test set by different 2D-CNN architectures (*i.e.*, GoogLeNet, EfficientNet_b01, ResNet50, DenseNet121, Inception_v3, DenseNet161) in Branch 1 of the proposed MVF network. Note that the highest F1-measure was selected after five running of experiments in order to obtain the optimal classification model. To further compare the effect of each 2D-CNN based networks, we plot the Micro-average *P-R* curves, and provide the numeric results of the area values under the curves of different CNN architectures in the legend of Fig. 7. Note that a large area under the *P-R* curve represents both high recall and high precision. From Table 1 and Fig. 7, one can see clearly that DenseNet161 achieved the best performance among all compared 2D-CNN models.

To evaluate the effectiveness of the proposed MVF network, we compared multiple metrics according to the ablation analysis. Obtained results are provided in Table 2. Due to the insufficient spatial features, Branch 2 exhibits the worst performance with the lowest Macro-*P*, Macro-*R*, Micro-*F1* and Macro-*F1* values are equal to 77.20%, 77.58%, 77.45% and 77.23%, respectively. The Branch 1 models performed better than Branch 2 with higher accuracies. It should be noted that, by taking advantage of both branches, the proposed MVF network achieves the highest Macro-*P* (*i.e.*, 94.59%), Macro-*R* (*i.e.*, 94.61%), Micro-*F1* (*i.e.*, 94.73%) and Macro-*F1* (*i.e.*, 95.55%), as well as the highest class-accuracy values on most of landforms. This indicates the superiority of multi-view features fusion.

C. Results and Analysis on Real CTX Data

1) Description of the Study Areas and Used Data

Three study areas on Mars were selected in our experiments, including the NASA's Perseverance landing region - Jezero Crater (18.38°N, 77.58°E), the China's Tianwen-1 Zhurong landing region - Southern Utopia Planitia (25.066°N, 109.925°E) and the prospective ESA's ExoMars candidate landing region - Oxia Planum (18.28°N, 335.37°E). Fig. 8 shows the location of the three study areas on the color-coded MOLA DEM map. The used Martian orbiter CTX images are shown in Fig. 9 (a1) - (c1). In particular, the Jezero Crater is located in the Martian Isidis Planitia region, it is supposed that an ancient river flowed into the Jezero Crater, and formed a delta at its west-side [see in Fig. 9 (a1)]. Utopia Planitia is the largest recognized circular impact basin in the northern plains of Mars [53], and it may be an area that was covered by a huge ancient ocean in the northern hemisphere of Mars [54, 55]. Oxia Planum is a 200 km wide low-relief terrain characterized

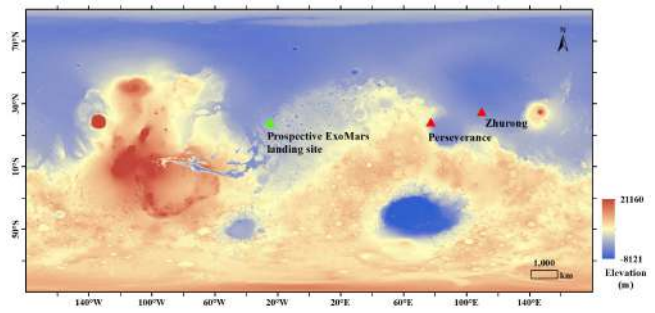

Fig. 8 Three considered study areas on the color-coded MOLA DEM.

by hydrous clay-bearing bedrock units located at the southwest margin of Arabia Terra [56]. This region exhibits Noachian- aged terrains. It contains abundant exposures of phyllosilicate- bearing bedrock and a fan-shaped sedimentary structure at the terminus of an ancient fluvial channel system [57]. Fig. 9 (a2) – (c2) show the reference maps of three study areas, which were made according to careful manual image interpretation.

2) Qualitative Analysis of the Landform Mapping Results

Landform mapping results obtained by the proposed MarsMapNet are shown in Fig. 10. Note that we provided two versions of mapping results: 1) obtained by using multi-view features, which is denoted as MarsMapNet (multi-view); 2) obtained by using only abstract convolutional features from scene-level (*i.e.*, 2D-CNN), which is denoted as MarsMapNet (single-view).

In Fig. 10, one can see that the two versions of the proposed method provided results that are macroscopically similar. However, differences are present on some landforms at local areas. In particular, the main landforms in the Jezero Crater region can be observed in Fig. 10 (a1) and (a2) are *Channel*, *Crater*, *Aeolian Straight*, *Ridge*, *Slope Streaks*, *Mixed Terrain*, *Textured Terrain* and *Crater Field*. Landscape features indicative of liquid water include erosional forms (like channels) and depositional forms (like deltas and alluvial fans) [58]. In the northwest corner of the map, the supposed ancient river [highlighted with a yellow line in Fig. 10 (a2)] was divided into *Channel* at both ends with *Aeolian Straight* in the middle. This mapping result reflects the real situation due to the fact that the bottom of river's central section is now covered by *linear dunes*. Such identified landforms provide important evidence for ancient aqueous processes and aeolian processes. In addition, the *Aeolian Straight* [highlighted in white polygon in Fig. 10 (a2)] in the southwest corner of the map and *Ridge* [highlighted in blue rectangles marker in Fig. 10 (a2)] in the west of the map were correctly classified. *Slope Streaks* [highlighted in yellow oval in Fig. 10 (a2)] where found around *Ridge*, which is confirmed on the high-resolution HiRISE image. However, some errors still exist in the mapping results. For example, since the *Mounds* class contains only *mound* clusters at a scene-level rather than individual *mound*, the larger-size individual *mound* [highlighted in red squares in Fig. 10 (a2)] may be mistakenly identified as *crater*, as they have very similar geometric properties.

The geomorphic features that appear in the vicinity of the Zhurong rover's landing site include *pitted cones*, *craters*, *mass wasting*, *troughs*, *ridges*, and *transverse aeolian ridges*, where the formation of *rampart crater*, *cones* and *troughs* are highly related to the scientific target water/ice [59-61]. From the landform mapping result of Southern Utopia Planitia area

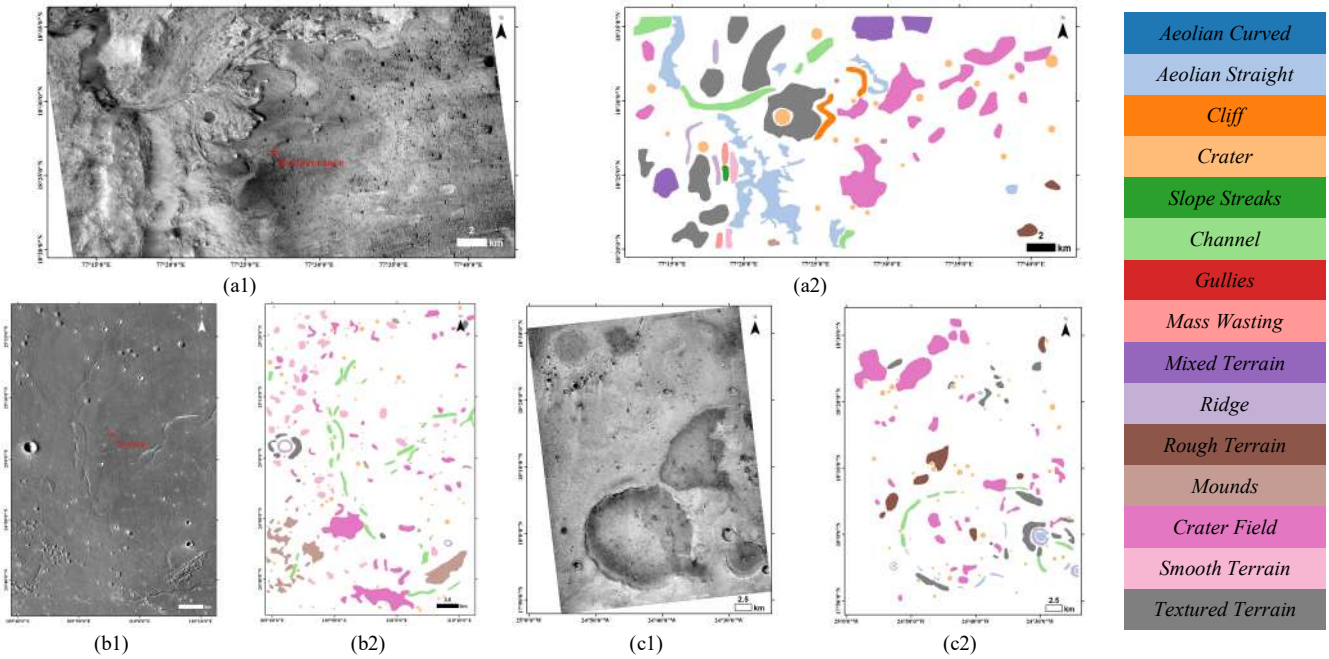

Fig. 9 The CTX images used for landform classification and their corresponding reference maps on three study areas: (a1) Perseverance landing region: Jezero Crater (the subset of the CTX image: D14_032794_1989_XN_18N282W, image size: 5600 × 3000 pixels); (b1) Tianwen-1 Zhurong landing region: Southern Utopia Planitia (subset of CTX mosaic image of F04_037553_2068_XN_26N250W and D22_035786_2060_XN_26N250W, image size: 5843 × 8911 pixels); and (c1) Prospective ExoMars candidate landing region: Oxia Planum (CTX image: F13_040921_1983_XN_18N024W, image size: 6361 × 8012 pixels). (a2), (b2) and (c2) are reference maps to (a1), (b1) and (c1), respectively.

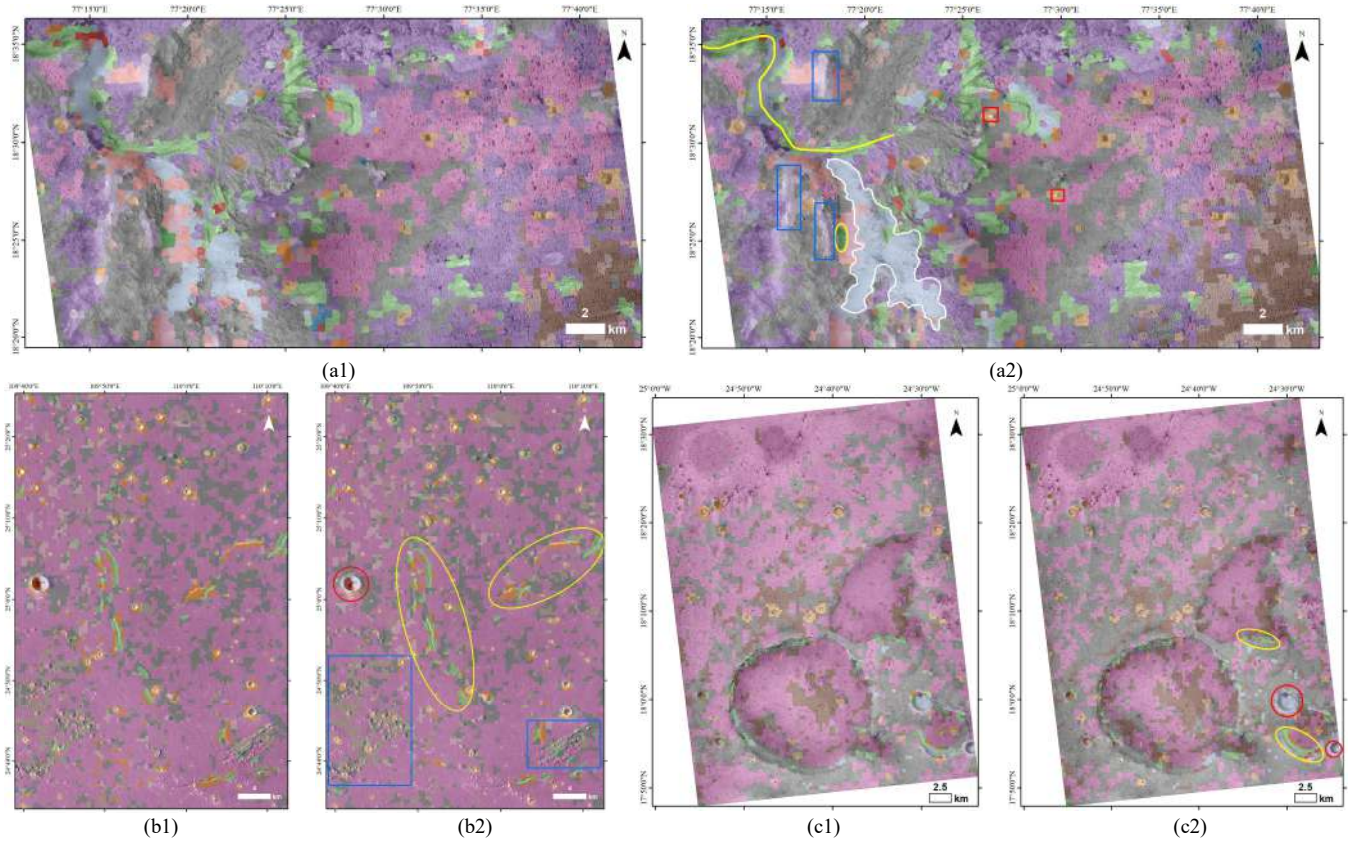

Fig. 10 Landform mapping results obtained by the proposed MarsMapNet (single-view) (a1, b1, c1) and MarsMapNet (multi-view) (a2, b2, c2) on three study areas: (a) Jezero Crater, (b) Southern Utopia Planitia, and (c) Oxia Planum.

[see in Fig. 10 (b)], one can see that there are seven classes of landforms identified, including *Crater Field*, *Crater*, *Mounds*, *Channel*, *Ridge*, *Textured Terrain* and *Smooth Terrain*. The *Channel* [highlighted in yellow oval marker in Fig. 10 (b2)] corresponding to *troughs* in [59-61] were correctly identified, and *steep slopes* on the *Channel* boundary were divided into *Cliff*. The *troughs* in Utopia Planitia were considered to be related to volatile activity, and it was suggested that they were formed by the volumetric compaction of a fine-grained sedimentary material covering an uneven buried surface [59, 62]. The *Mounds* located in the region's southwest and southeast [highlighted in blue rectangles in Fig. 10 (b2)] correspond to *cones* as stated in [59-61] and they were consistently identified. The Zhurong rover planned to continue its journey south to explore these cones and investigate their origins. In addition, there is a large *rampart crater* with a diameter of about 2.3 km [highlighted in red circle in Fig. 10 (b2)], which is much larger than the fixed window size. So, this *crater* was divided into several parts in the mapping results, *i.e.*, *Ridge* and *Cliff* on the rim, *Mass Wasting* on the slope, and *Gullies* on the bottom.

As shown in Fig. 10 (c), the identified landforms in the Oxia Planum region consist of *Crater*, *Crater Field*, *Channel*, *Aeolian Straight*, *Ridge*, *Textured Terrain* and *Rough Terrain*. *Craters* fitting the window size were correctly identified, whereas smaller *craters* were divided into the *Crater Field*. *Channels* [highlighted in yellow oval in Fig. 10 (c2)) distributed along the rim of *craters* were correctly identified. In addition, large areas were grouped into *Textured Terrain*, which means the bedrock of the region is covered by loose material, but the underlying bedrock is not completely buried [35]. It is important to note that *channels* and bedrock indicate long-term low-intensity erosion processes in the Oxia Planum region, and this region was chosen as a candidate landing site due to this widespread bedrock exposure, which may contain physical and chemical biomarkers for seeking out possible past life and aqueous activity on Mars [57]. However, there are still remaining issues due to the window-size limitations. *Craters* [highlighted in a red circle in Fig. 10 (c2)] larger than the window size were divided into *Ridge* and *Aeolian Straight* parts, which are actually the crater rim and the crater floor.

In order to further analyze the mapping results qualitatively, we calculated the Gradient-weighted Class Activation Maps (Grad-CAM) [63] at local regions [see in Fig. 11 (b2) and (c2), and Fig. 12 (b2) and (c2)]. For a particular category, a CAM indicates the discriminative image regions used by the CNN to identify that category [64]. Grad-CAM is a generalization of CAM and is applicable to a significantly broader range of CNN model families. In Grad-CAM, discriminative regions of the image that are associated with the landforms is localized and highlighted.

In Fig. 11, both mapping results obtained by the proposed MarsMapNet (single-view) and MarsMapNet (multi-view) are correct. Fig. 11 (a) shows some examples of correctly classified superpixel landforms. The number of superpixel is illustrated in red (*e.g.*, 3239 is the center superpixel of *Crater*

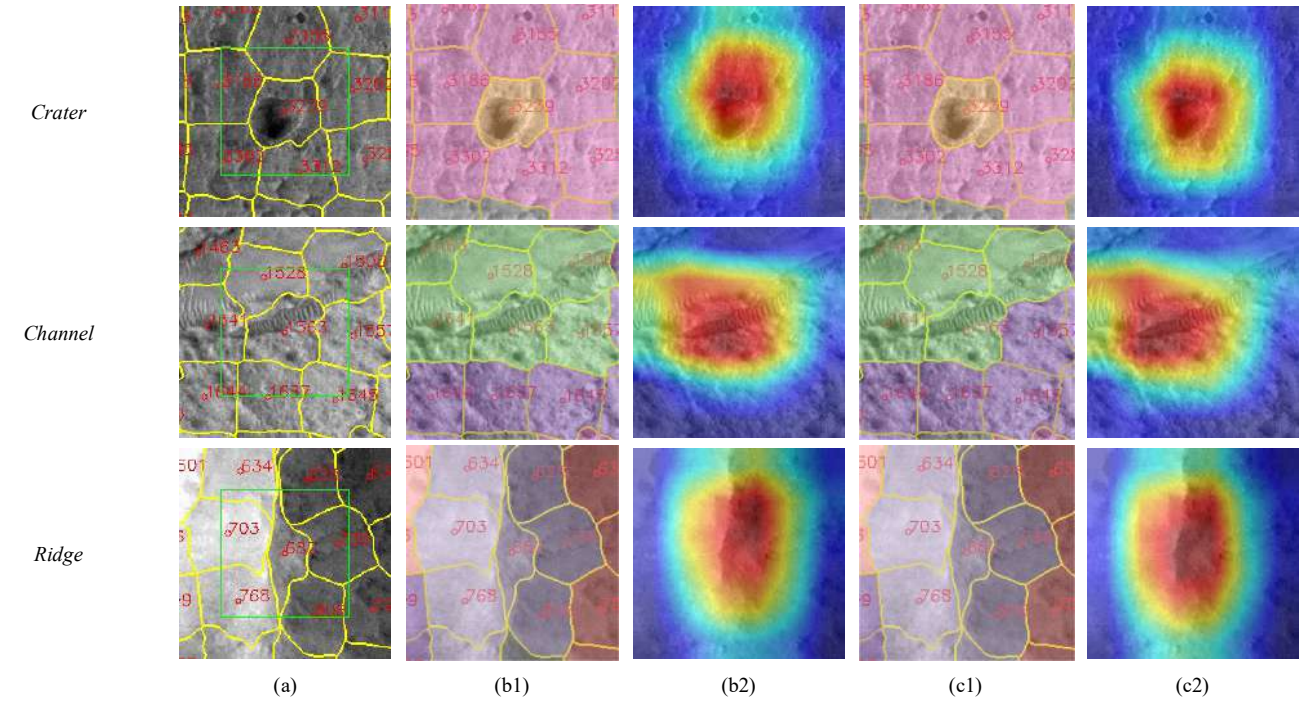

Fig. 11 Examples of correct mapping results obtained by the proposed MarsMapNet approaches and the corresponding Grad-CAM at a local scale. (a) Superpixel-corresponding patch within the CTX image patch. The whole patch size is 200×200 (inputted into 2D-CNN), the local patch size in the green box is 120×120 (used to calculate 1D features and inputted into 1D-CNN); (b1) and (b2): mapping results obtained by the proposed MarsMapNet (single-view) and its Grad-CAM, respectively; (c1) and (c2): mapping results obtained by the proposed MarsMapNet (multi-view) and its Grad-CAM, respectively.

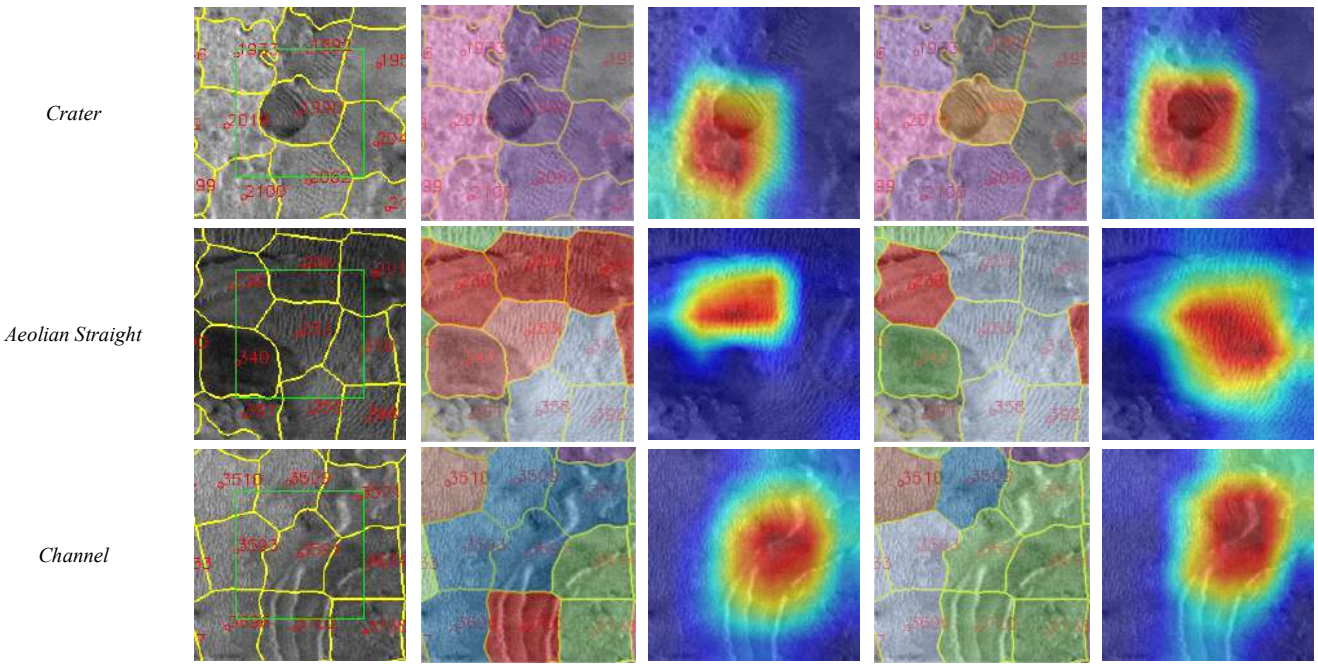

Fig. 12 Examples of wrong (b1: MarsMapNet (single-view)) and correct (c1: MarsMapNet (multi-view)) mapping results obtained by the proposed MarsMapNet approaches and their corresponding Grad-CAMs at a local scale. (a) Superpixel-corresponding patch within the CTX image patch. The whole patch size is 200×200 (inputted into 2D-CNN), the local patch size in the green box is 120×120 (used to calculate 1D features and inputted into 1D-CNN); (b1) and (b2): mapping results obtained by the proposed MarsMapNet (single-view) and its Grad-CAM, respectively; (c1) and (c2): mapping results obtained by the proposed MarsMapNet (multi-view) and its Grad-CAM, respectively.

in the first row). Fig. 11 (b1) and (c1) show the mapping results. One can see that the highlighted regions obtained by two proposed methods are similar, corresponding exactly to the landform type (*i.e.*, *Crater*, *Channel*, *Ridge*) of the center superpixels [see Fig. 11 (b2) and (c2)].

Fig. 12 shows wrong and correct examples of mapping results where the proposed MarsMapNet (multi-view) outperformed the proposed MarsMapNet (single-view). In the first and second rows, classes of the center superpixel 1990 and 263 are *Crater* and *Aeolian Straight*, respectively, which

Table 3 The Micro-*F1* (%) of mapping results obtained by different methods.

Study areas	Reference methods		Proposed MarsMapNet (multi-view)	
	without MRFs	with MRFs	without MRFs	with MRFs
Jezero Crater	82.65	84.25	86.03	86.41
Southern Utopia Planitia	72.93	73.54	78.32	78.55
Oxia Planum	84.04	84.35	85.51	85.70

were mis-classified into *Mixed Terrain* and *Mass Wasting* by the proposed MarsMapNet (single-view), and the highlighted regions were off-center the center superpixels. However, they were correctly predicted in the proposed MarsMapNet (multi-view), and the highlighted regions exactly matched the center superpixels. There are some examples (*e.g.*, third row in Fig. 12) where the proposed MarsMapNet (single-view) and MarsMapNet (multi-view) both acquire the highlighted region correctly, but the MVF allows MarsMapNet (multi-view) to achieve better recognition performance.

3) Comparison with other methods

To further demonstrate the effectiveness of the proposed MarsMapNet, landform mapping results obtained by the MarsMapNet (multi-view) were compared with those obtained by the state-of-the-art methods with or without post processing [35]. For a fair comparison, the trained Densenet161 that achieved the highest testing accuracy was also used the reference method. The MRFs was used to smooth the pixel-wise classification noise by a sliding window. Table 3 lists the Micro-*F1* scores of mapping results on three study areas. The Micro-*F1* scores were calculated based on the manually labeled reference maps (see in Fig. 9 (b2), (c2) and (d2)). The accuracy of reference method was improved by using MRFs. Since the mapping result of the proposed MarsMapNet exhibited good boundary adhesion performance and local homogeneity, the high accuracy can be obtained without MRFs. It is worth noting that the accuracy of the proposed MarsMapNet without using MRFs is also higher than that of the reference method using MRFs.

Fig. 13 provides a comparison of the obtained classification maps on the Jezero Crater area data set. The two compared methods produced in general similar results at a global scale. However, some major differences are highlighted with yellow boxes in Fig. 13, which are enlarged and further compared in Fig. 14. We can observe the salt-and-pepper noise in the mapping results obtained by the reference method without using MRFs filtering [see in Fig. 14 (c1)]. After MRFs filtering, noise was removed, resulting in a smoother landform classification map [see in Fig. 14 (c2)]. The proposed MarsMapNet achieved better mapping results [Fig. 14 (d1)] than the reference method without using MRFs [Fig. 14 (c1)] and achieved with more regular and smooth boundaries. However, homogeneity can be also observed within the given classes. Despite the use of MRFs in the proposed approach, there is no significant improvement in the classification result [see Fig. 14 (d2)]. This indicates that the proposed MarsMapNet can obtain comparable mapping results without post-processing operations to the reference method that relies on the MRFs filtering. In addition, classification results at a local scale are compared in more details [see the red box highlighted in Fig. 14 (a) and enlarged in Fig. 14 (b)]. There is *Aeolian Straight* in the first-row image, a *Crater* in the second-row image, and *Aeolian Straight* (left part) and *Channel* (right part) in the third-row image. They were all correctly identified by the proposed MarsMapNet [see in Fig. 14 (d1)], but were misclassified by the reference method [see Fig. 14 (c2)].

4) Comparison of computing time

Table 4 shows the computing time taken by different methods. Especially, time cost in the reference methods contains two parts: the model prediction and the MRFs filtering steps. The time cost in the proposed MarsMapNet (single-view) includes two parts: the superpixel segmentation and model prediction, and the proposed MarsMapNet (multi - view) contains three parts: the superpixel segmentation, multi-texture features generation and model prediction. Three 2D-CNN models (*i.e.*, DenseNet161, Inception_v3 and DenseNet121) were also compared according to their high performance in the quantitative evaluation as shown in Table 1. We can see that due to the deeper model and more parameters, DenseNet161 took the most time, followed by DenseNet121 and Inception-v3.

Table 4 Computational time taken by alternate methods (Seconds).

Study areas	Methods		Classification unit	Superpixel segmentation	Multi-texture features generation	Model prediction	Post-classification processing (MRFs)	Total
Jezero Crater	Reference methods	DenseNet161	Pixel	/	/	78,413	859	79,272
		Inception_v3		/	/	46,500	879	47,379
		DenseNet121		/	/	42,660	861	43,521
	Proposed MarsMapNet (single-view)		Superpixel	14	/	34	/	48
	Proposed MarsMapNet (multi-view)			14	596	29	/	639
Southern Utopia Planitia	Reference methods	DenseNet161	Pixel	/	/	265,351	2,803	268,154
		Inception_v3		/	/	155,160	2,810	157,970
		DenseNet121		/	/	140,280	2,807	143,087
	Proposed MarsMapNet (single-view)		Superpixel	57	/	121	/	178
	Proposed MarsMapNet (multi-view)			59	2,555	93	/	2,707
Oxia Planum	Reference methods	DenseNet161	Pixel	/	/	257,484	2,796	260,280
		Inception_v3		/	/	150,840	2,781	153,621
		DenseNet121		/	/	134,460	2,795	137,255
	Proposed MarsMapNet (single-view)		Superpixel	47	/	100	/	147
	Proposed MarsMapNet (multi-view)			48	2,203	81	/	2,332

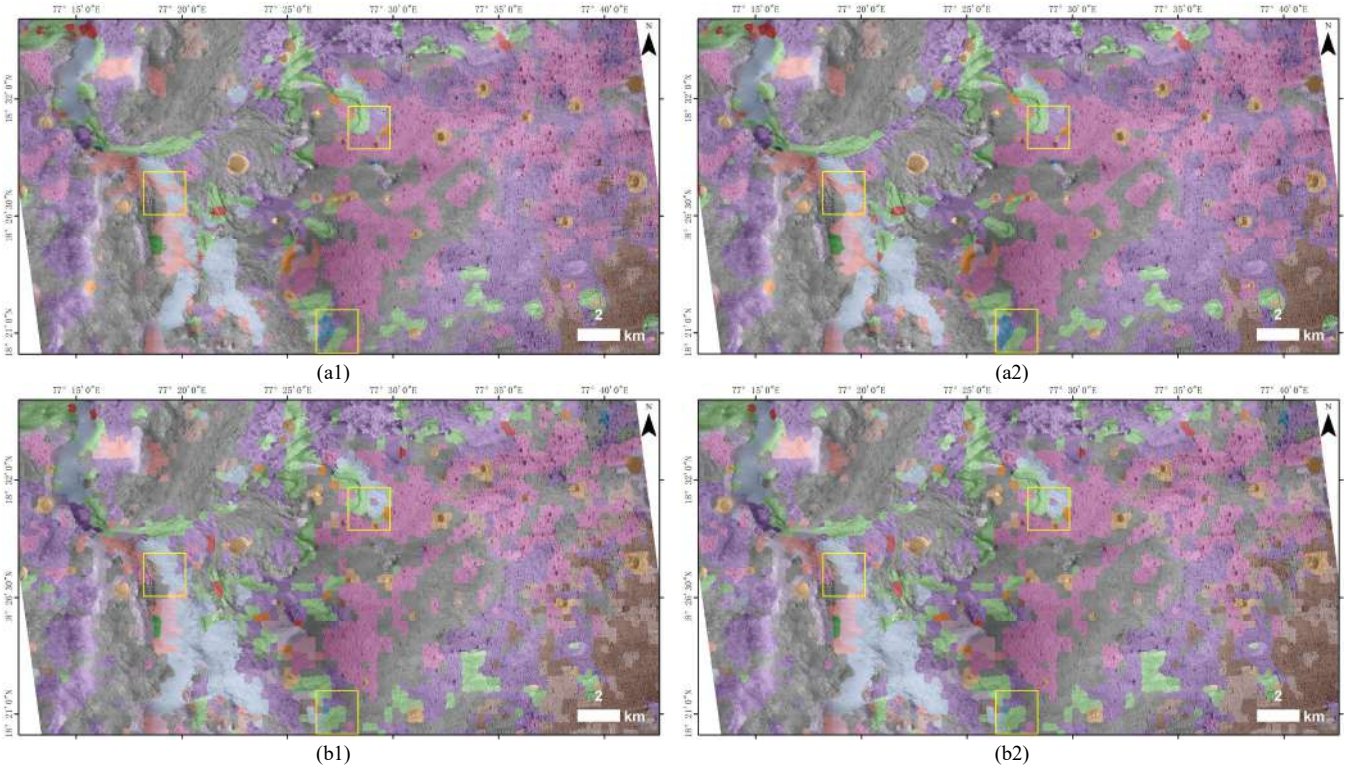

Fig. 13 Landforms mapping results at a global scale obtained by different methods (Jezero Crater area): reference method [35] without (a1) and with MRFs (a2); the proposed MarsMapNet without (b1) and with MRFs (b2).

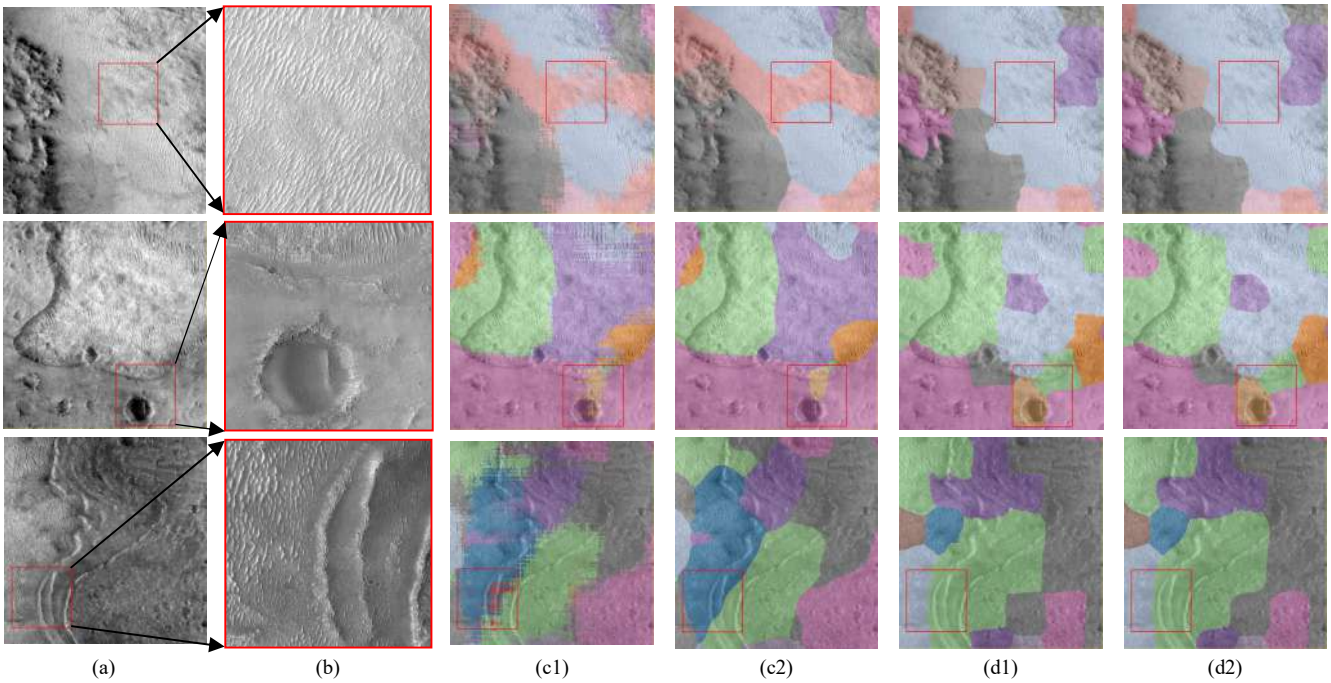

Fig. 14 Landforms mapping results at a local scale obtained by different methods (Jezero Crater area). (a) CTX images; (b) enlarged view (HiRISE image) of the red box in (a); reference method [35] without (c1) and with MRFs (c2); the proposed MarsMapNet without (d1) and with MRFs (d2).

The total time consumption of the proposed MarsMapNet approach is significantly smaller than that of the reference method (DenseNet161). In particular, the MarsMapNet (single-view) decreased by 1,652 times, 1,506 times and 1,771 times of time consumption on three data sets, respectively. The MarsMapNet (multi-view) decreased by 124 times, 99 times and 112 times on three data sets, respectively. Even when compared with the reference methods (more light-weight Inception-v3 and DenseNet121), two versions of the proposed MarsMapNet also exhibited a significant advantage on computational efficiency. A sharp time reduction can be seen mainly in the model prediction process. With regard to

the two versions of the proposed MarsMapNet techniques, the main difference in computing cost is due to shallow-feature generation step. Although the MarsMapNet (multi-view) took more time than the MarsMapNet (single-view), we consider this an acceptable tradeoff with respect to its improved classification accuracy. Furthermore, the time cost analysis also indicates that the proposed MarsMapNet (single-view) exhibits great advantages for a rapid landform mapping, which offers the potential for efficient Martian landform mapping over a large area.

IV. Conclusion

In this article, a novel superpixel-guided multi-view feature fusion network (MarsMapNet) is proposed for efficient mapping of the Martian landforms. To fully exploit texture information on single-band and gray-scale Martian orbiter images, the specially designed multi-texture features are extracted from a local view and inputted into 1D-CNN for further feature mining. Then they are fused with abstract convolutional features that are extracted by 2D-CNN from a scene-level view. When performing Martian landforms mapping, superpixel segmentation is incorporated into the trained network to both reduce salt-and-pepper noise in pixel-wise classification and allows a better local homogeneity and more accurate landform boundaries. Moreover, superpixel-level classification results can be obtained by using more accessible scene-level samples.

Quantitative analysis on the considered data set and qualitative mapping analysis on three Mars missions' landing (or candidate landing) regions confirmed the effectiveness of the proposed method in terms of better classification accuracy and high computational efficiency when compared with state-of-the-art methods. The proposed approach has the potential for discovering Martian surface scientific targets of interest, analyzing their pattern and spatial distribution, especially in the application of landing site selection. For future works, we plan to investigate the potential of the proposed MarsMapNet in producing a Martian global landform product. Meanwhile, we will explore more advanced techniques to improve the Martian landform mapping results at local scale.

Acknowledgments

The authors would like to thank PDS Geosciences Node Mars Orbital Data Explorer (ODE) for providing CTX images (https://ode.rsl.wustl.edu/mars/productsearch), authors of the paper [35] for providing the DoMars16k public data set (http://dx.doi.org/10.5281/zenodo.4291940), and MOLA team of Goddard Space Flight Center for providing the MOLA DEM product (https://astrogeology.usgs.gov/search/map/Mars/GlobalSurveyor/MOLA/Mars_MGS_MOLA_DEM_mosaic_global_463m).

References

[1] J. S. Kargel, V. R. Baker, J. E. Beget, J. F. Lockwood, T. L. Pewe, J. S. Shaw, and R. G. Strom, "Evidence of Ancient Continental-Glaciation in the Martian Northern Plains," *Journal of Geophysical Research: Planets*, vol. 100, no. E3, pp. 5351-5368, Mar. 1995.

[2] J. W. Head, M. A. Kreslavsky, and S. Pratt, "Northern lowlands of Mars: Evidence for widespread volcanic flooding and tectonic deformation in the Hesperian Period," *Journal of Geophysical Research: Planets*, vol. 107, no. E1, Jan. 2002.

[3] K. L. Tanaka, J. A. Skinner Jr., T. M. Hare, T. Joyal, and A. Wenker, "Resurfacing history of the northern plains of Mars based on geologic mapping of Mars Global Surveyor data," *Journal of Geophysical Research: Planets*, vol. 108, no. E4, Apr. 2003.

[4] Tanaka, K.L., Skinner, J.A., Jr., Dohm, J.M., Irwin, R.P., III, Kolb, E.J., Fortezzo, C.M., Platz, T., Michael, G.G., and Hare, T.M., 2014, Geologic map of Mars: U.S. Geological Survey Scientific Investigations Map 3292, scale 1:20,000,000, pamphlet 43 p., https://dx.doi.org/10.3133/sim3292.

[5] T. A. Goudge, A. M. Morgan, G. Stucky de Quay, and C. I. Fassett, "The importance of lake breach floods for valley incision on early Mars," *Nature*, vol. 597, no. 7878, pp. 645-649, Sep. 2021.

[6] J. W. Head, J. F. Mustard, M. A. Kreslavsky, R. E. Milliken, and D. R. Marchant, "Recent ice ages on Mars," *Nature*, vol. 426, no. 6968, pp. 797-802, Dec. 2003.

[7] M. A. Ivanov, H. Hiesinger, G. Erkeling, D. Reiss, "Mud volcanism and morphology of impact craters in Utopia Planitia on Mars: Evidence for the ancient ocean," *Icarus*, vol. 228, pp. 121-140, 2014.

[8] M. Pajola, R. Pozzobon, S. Silvestro, F. Salese, S. Rossato, L. Pompilio, G. Munaretto, L. Teodoro, A. Kling, E. Simioni, A. Lucchetti, L. L. Tornabene, L. Marinangeli, A. C. Tangari, J. Wilson, G. Cremonese, M. Massironi, and N. Thomas, "Geology, in-situ resource-identification and engineering analysis of the Vernal crater area (Arabia Terra): A suitable Mars human landing site candidate," *Planetary and Space Science*, vol. 213, 2022.

[9] A. J. Brown, S. Byrne, L. L. Tornabene, T. Roush,"Louth crater: Evolution of a layered water ice mound," *Icarus*, vol. 196, no. 2, pp. 433-445, 2008.

[10] J. A. Grant, M. P. Golombek, S. A. Wilson, K. A. Farley, K. H. Williford, and A. Chen, "The science process for selecting the landing site for the 2020 Mars rover," *Planetary and Space Science*, vol. 164, pp. 106-126, 2018.

[11] M. Ono, B. Rothrock, E. Almeida, A. Ansar, R. Otero, A. Huertas, and M. Heverly, "Data-Driven Surface Traversability Analysis for Mars 2020 Landing Site Selection," *2016 IEEE Aerospace Conference*, pp. 1-12, 2016.

[12] D. M. DeLatte, S. T. Crites, N. Guttenberg, and T. Yairi, "Automated crater detection algorithms from a machine learning perspective in the convolutional neural network era," *Advances in Space Research*, vol. 64, no. 8, pp. 1615-1628, 2019.

[13] L. Bandeira, J. S. Marques, J. Saraiva, and P. Pina, "Automated Detection of Martian Dune Fields," *IEEE Geoscience and Remote Sensing Letters*, vol. 8, no. 4, pp. 626-630, 2011.

[14] Y. Wang, K. Di, X. Xin, and W. Wan, "Automatic detection of Martian dark slope streaks by machine learning using HiRISE images," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 129, pp. 12-20, 2017.

[15] L. F. Palafox, C. W. Hamilton, S. P. Scheidt, and A. M. Alvarez, "Automated detection of geological landforms on Mars using Convolutional Neural Networks," *Computers & geosciences*, vol. 101, pp. 48-56, Apr. 2017.

[16] T. Nagle-McNaughton, T. McClanahan, and L. Scuderi, "PlaNet: A Neural Network for Detecting Transverse Aeolian Ridges on Mars," *Remote Sensing*, vol. 12, no. 21, 2020.

[17] K. M. Aye, M. E. Schwamb, G. Portyankina, C. J. Hansen, A. McMaster, G. R. M. Miller, B. Carstensen, C. Snyder, M. Parrish, S. Lynn, C. Mai, D. Miller, R. J. Simpson, and A. M. Smith, "Planet Four: Probing springtime winds on Mars by mapping the southern polar CO2 jet deposits," *Icarus*, vol. 319, pp. 558-598, 2019.

[18] H. Shozaki, Y. Sekine, N. Guttenberg, and G. Komatsu, "Recognition and Classification of Martian Chaos Terrains Using Imagery Machine Learning: A Global Distribution of Chaos Linked to Groundwater Circulation, Catastrophic Flooding, and Magmatism on Mars," *Remote Sensing*, vol. 14, no. 16, 2022.

[19] T. Stepinski and R. Vilalta, "Digital Topography Models for Martian Surfaces," *IEEE Geoscience and Remote Sensing Letters*, vol. 2, no. 3, pp. 260-264, 2005.

[20] B. D. Bue and T. F. Stepinski, "Automated classification of landforms on Mars," *Computers & Geosciences*, vol. 32, no. 5, pp. 604-614, 2006.

[21] T. F. Stepinski and C. Bagaria, "Segmentation-Based Unsupervised Terrain Classification for Generation of Physiographic Maps," *IEEE*

Geoscience and Remote Sensing Letters, vol. 6, no. 4, pp. 733-737, 2009.

[22] M. C. Malin, J. F. Bell, B. A. Cantor, M. A. Caplinger, W. M. Calvin, R. T. Clancy, K. S. Edgett, L. Edwards, R. M. Haberle, P. B. James, S. W. Lee, M. A. Ravine, P. C. Thomas, and M. J. Wolff, "Context Camera Investigation on board the Mars Reconnaissance Orbiter," *Journal of Geophysical Research: Planets*, vol. 112, no. E5, 2007.

[23] A. S. McEwen, E. M. Eliason, J. W. Bergstrom, N. T. Bridges, C. J. Hansen, W. A. Delamere, J. A. Grant, V. C. Gulick, K. E. Herkenhoff, L. Keszthelyi, R. L. Kirk, M. T. Mellon, S. W. Squyres, N. Thomas, and C. M. Weitz, "Mars Reconnaissance Orbiter's High Resolution Imaging Science Experiment (HiRISE)," *Journal of Geophysical Research: Planets*, vol. 112, no. E5, 2007.

[24] R. Jaumann, G. Neukum, T. Behnke, T. C. Duxbury, K. Eichentopf, J. Flohrer, S. v. Gasselt, B. Giese, K. Gwinner, E. Hauber, H. Hoffmann, A. Hoffmeister, U. Köhler, K. D. Matz, T. B. McCord, V. Mertens, J. Oberst, R. Pischel, D. Reiss, E. Ress, T. Roatsch, P. Saiger, F. Scholten, G. Schwarz, K. Stephan, and M. Wählisch, "The high-resolution stereo camera (HRSC) experiment on Mars Express: Instrument aspects and experiment conduct from interplanetary cruise through the nominal mission," *Planetary and Space Science*, vol. 55, no. 7-8, pp. 928-952, 2007.

[25] Q. Meng, D. Wang, X. Wang, W. Li, X. Yang, D. Yan, Y. Li, Z. Cao, Q. Ji, T. Sun, W. Yan, K. Wang, X. Li, J. Huang, Z. Wang, W. Zhao, Y. Wang, Y. He, X. Hao, W. Liu, B. Zhang, P. Zhou, Y. Li, H. Zhao, L. Lu, H. Guan, D. Zhou, F. Wu, F. Zhang, S. Zhu, and J. Dong, "High resolution imaging camera (HiRIC) on China's first mars exploration Tianwen-1 mission," *Space Science Reviews*, vol. 217, no. 3, 2021.

[26] M. C. Malin and K. S. Edgett, "Mars Global Surveyor Mars Orbiter Camera: Interplanetary cruise through primary mission," *Journal of Geophysical Research: Planets*, vol. 106, no. E10, pp. 23429-23570, 2001.

[27] M. E. Schwamb, K.-M. Aye, G. Portyankina, C. J. Hansen, C. Allen, S. Allen, F. J. Calef, S. Duca, A. McMaster, and G. R. M. Miller, "Planet Four: Terrains - Discovery of araneiforms outside of the South Polar layered deposits," *Icarus*, vol. 308, pp. 148-187, 2018.

[28] A. Kraut and D. Wettergreen, "Classification of mars terrain using multiple data sources," *CIDU*, pp. 54-68, 2010.

[29] K. L. Wagstaff, Y. Lu, A. Stanboli, K. Grimes, T. Gowda, and J. Padams, "Deep Mars: CNN Classification of Mars Imagery for the PDS Imaging Atlas," *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 7867-7872, 2018.

[30] K. Wagstaff, S. Lu, E. Dunkel, K. Grimes, B. Zhao, J. Cai, S. B. Cole, G. Doran, R. Francis, J. Lee, and L. Mandrake, "Mars Image Content Classification: Three Years of NASA Deployment and Recent Advances," *Proceedings of the AAAI Conference on Artificial Intelligenc*, vol. 35, pp. 15204-15213, 2021

[31] B. Rothrock, R. Kennedy, C. Cunningham, J. Papon, M. Heverly, and M. Ono, "SPOC: Deep Learning-based Terrain Classification for Mars Rover Missions," *AIAA SPACE 2016*, 2016.

[32] M. Ono, M. Heverly, B. Rothrock, E. Almeida, F. Calef, T. Soliman, N. Williams, H. Gengl, T. Ishimatsu, A. Nicholas, E. Stilley, K. Otsu, R. Lange, and S. M. Milkovich, "Mars 2020 Site-Specific Mission Performance Analysis: Part 2. Surface Traversability," *2018 AIAA SPACE and Astronautics Forum and Exposition*, 2018.

[33] A. M. Barrett, M. R. Balme, M. Woods, S. Karachalios, D. Petrocelli, L. Joudrier, and E. Sefton-Nash, "NOAH-H, a deep-learning, terrain classification system for Mars: Results for the ExoMars Rover candidate landing sites," *Icarus*, vol. 371, 2022.

[34] S. Liu, X. Tong, S. Liu, H. Xie, H. Zhao, and D. Liu, "Topography Modeling, Mapping and Analysis of China's First Mars Mission Tianwen-1 Landing Area from Remote Sensing Images," *Journal of Deep Space Exploration*, vol. 9, no. 3, pp. 338-347, 2022.

[35] T. Wilhelm, M. Geis, J. Püttschneider, T. Sievernich, T. Weber, K. Wohlfarth, and C. Wöhler, "DoMars16k: A Diverse Dataset for Weakly Supervised Geomorphologic Analysis on Mars," *Remote Sensing*, vol. 12, no. 23, 2020.

[36] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 11, pp. 2274-82, Nov. 2012.

[37] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84-90, 2017.

[38] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition." In *Proc. International Conference on Learning Representations*, 2014, pp.1-14. [Online]. Available: http://arxiv.org/abs/1409.1556.

[39] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, "Going Deeper with Convolutions," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1-9, 2015.

[40] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2818-2826, 2016.

[41] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770-778, 2016.

[42] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," 2019. *DOI:10.48550/arXiv.1905.11946.*

[43] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4700-4708, 2017.

[44] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural Features for Image Classification," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 3, no. 6, pp. 610-621, 1973.

[45] J. Zhao, X. Mao, and L. Chen, "Speech emotion recognition using deep 1D & 2D CNN LSTM networks," *Biomedical Signal Processing and Control*, vol. 47, pp. 312-323, 2019.

[46] S. Kiranyaz, T. Ince, and M. Gabbouj, "Real-time patient-specific ECG classification by 1-D convolutional neural networks," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 3, pp. 664-75, Mar. 2016.

[47] X. Xu, W. Li, Q. Ran, Q. Du, L. Gao, and B. Zhang, "Multisource Remote Sensing Data Classification Based on Convolutional Neural Network," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 2, pp. 937-949, 2018.

[48] S. Liu, H. Zhao, Q. Du, L. Bruzzone, A. Samat, and X. Tong, "Novel Cross-Resolution Feature-Level Fusion for Joint Classification of Multispectral and Panchromatic Remote Sensing Images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-14, 2021.

[49] S. Liu, Y. Zheng, Q. Du, L. Bruzzone, A. Samat, X. Tong, Y. Jin, and C. Wang, "A Shallow-to-Deep Feature Fusion Network for VHR Remote Sensing Image Classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-13, 2022.

[50] M. Zhang, W. Li, Q. Du, L. Gao, and B. Zhang, "Feature Extraction for Classification of Hyperspectral and LiDAR Data Using Patch-to-Patch CNN," *IEEE Transactions on Cybernetics*, vol. 50, no. 1, pp. 100-111, Jan 2020.

[51] J. Deng, W. Dong, R. Socher, L. J. Li, K. Li, and F. F. Li, "ImageNet: A Large-Scale Hierarchical Image Database," *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248-255, 2009.

[52] S. Raschka, V. Mirjalili, "Python Machine Learning," *Packt Publishing Ltd.: Birmingham*, UK, 2017.

[53] M. M. Mills, A. S. McEwen, and C. H. Okubo, "A Preliminary Regional Geomorphologic Map in Utopia Planitia of the Tianwen-1 Zhurong Landing Region," Geophysical Research Letters, vol. 48, no. 18, 2021.

[54] G. Di Achille and B. M. Hynek, "Ancient ocean on Mars supported by global distribution of deltas and valleys," *Nature Geoscience*, vol. 3, no. 7, pp. 459-463, 2010.

[55] B. M. Hynek, M. Beach, and M. R. T. Hoke, "Updated global map of Martian valley networks and implications for climate and hydrologic processes," *Journal of Geophysical Research*, vol. 115, no. E9, 2010.

[56] C. Quantin-Nataf, J. Carter, L. Mandon, P. Thollot, M. Balme, M. Volat, L. Pan, D. Loizeau, C. Millot, S. Breton, E. Dehouck, P. Fawdon, S. Gupta, J. Davis, P. M. Grindrod, A. Pacifici, B. Bultel, P. Allemand, A. Ody, L. Lozach, and J. Broyer, "Oxia Planum: The Landing Site for the ExoMars "Rosalind Franklin" Rover Mission: Geological Context and Prelanding Interpretation," *Astrobiology*, vol. 21, no. 3, pp. 345-366, Mar 2021.

[57] A. M. Barrett, J. Wright, E. Favaro, P. Fawdon, M. R. Balme, M. J. Woods, S. Karachalios, E. Bohachek, E. Sefton-Nash, and L. Joudrier,

"Oxia Planum, Mars, classified using the NOAH-H deep-learning terrain classification system," *Journal of Maps*, pp. 1-14, 2022.

[58] Mustard, J.F., M. Adler, A. Allwood, D.S. Bass, D.W. Beaty, J.F. Bell III, W.B. Brinckerhoff, M. Carr, D.J. Des Marais, B. Drake, K.S. Edgett, J. Eigenbrode, L.T. Elkins-Tanton, J.A. Grant, S. M. Milkovich, D. Ming, C. Moore, S. Murchie, T.C. Onstott, S.W. Ruff, M.A. Sephton, A. Steele, A. Treiman (2013): Report of the Mars 2020 Science Definition Team, 154 pp., posted July, 2013, by the Mars Exploration Program Analysis Group (MEPAG) at http://mepag.jpl.nasa.gov/reports/MEP/Mars_2020_SDT_Report_Final .pdf.

[59] J. Zhao, Z. Xiao, J. Huang, J. W. Head, J. Wang, Y. Shi, B. Wu, and L. Wang, "Geological characteristics and targets of high scientific interest in the Zhurong landing region on Mars," *Geophysical Research Letters*, vol. 48, no. 20, 2021.

[60] Y. Liu, X. Wu, Y. S. Zhao, L. Pan, C. Wang, J. Liu, Z. Zhao, X. Zhou, C. Zhang, Y. Wu, W. Wan, and Y. Zou, "Zhurong reveals recent aqueous activities in Utopia Planitia, Mars," *Science Advances*, vol. 8, no. 19, May 2022.

[61] J. Liu, C. Li, R. Zhang, W. Rao, X. Cui, Y. Geng, Y. Jia, H. Huang, X. Ren, W. Yan, X. Zeng, W. Wen, X. Wang, X. Gao, Q. Fu, Y. Zhu, J. Dong, H. Li, X. Wang, W. Zuo, Y. Su, D. Kong, and H. Zhang, "Geomorphic contexts and science focus of the Zhurong landing site on Mars," *Nature Astronomy*, vol. 6, no. 1, pp. 65-71, 2021.

[62] D. L. Buczkowski, K. D. Seelos, and M. L. Cooke, "Giant polygons and circular graben in western Utopia basin, Mars: Exploring possible formation mechanisms," *Journal of Geophysical Research: Planets*, vol. 117, no. E8, 2012.

[63] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization," *International Journal of Computer Vision*, vol. 128, no. 2, pp. 336-359, 2019.

[64] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, "Learning Deep Features for Discriminative Localization," *Proceeding of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2921-2929, 2016.