Cross Modality Adaptive Feature Fusion for Multi-type and Multi-scale Impact Craters Identification on Mars

Chen Yang, *Senior Member, IEEE,* Minghao Zhao, Lorenzo Bruzzone, *Fellow, IEEE*, Renchu Guan, *Senior Member, IEEE* and Haishi Zhao, *Member, IEEE*

***Abstract*—Craters are the most typical geologic structures and landforms on the surface of Mars. Martian craters are widely distributed in a variety of morphology with multiple types and exhibit significant differences in scale. Many attempts have been made to automatic identification of Martian craters, yet existing methods do not satisfy the needs of large-scale identification. In this paper, we first integrate a Martian crater data set relate to the mid- and low-latitude regions, which contains different types and various scales of craters. Then, a dual CNN-Transformer-based cross modality adaptive feature fusion network (DCT-CMAFFNet) is proposed for an accurate identification of the multi-type and multi-scale craters at large scale on the Martian surface. The proposed network takes full advantage of the rich morphological features contained in Martian imagery and the topographical information reflected by the digital elevation model (DEM) data. It contains two modules: one is the dual CNN-Transformer part, which employs a hybrid architecture to extract the local detailed and global deep features of Martian craters from images and DEM; the other is CMAFF module, which exploits self-attention mechanism to learn the relationship between the images and DEM modalities and weight each position of the deep feature maps to ensure a comprehensive identification of multi-type and multi-scale Martian craters. By adaptively fusing the rich information from imagery and DEM, the proposed network identified 3166 new Martian impact craters larger than 1km, achieving 14%-24% improvement in accuracy compared to methods using either a single data source or data feature fusion modality.**

***Index Terms*—Mars, impact crater, multi-type, multi-scale, identification, cross modality, feature fusion, remote sensing.**

This work was supported in part by the National Natural Science Foundation of China under Grant 42272340 and 42302265, the Science-Technology Development Plan Project of Jilin Province of China grants 20230101311JC, (*Corresponding authors: Haishi Zhao*)

C. Yang, M. H. Zhao and H. S. Zhao are with the College of Earth Sciences, Jilin University, Changchun, 130061, China, and Key Laboratory of Lunar and Deep Space Exploration, National Astronomical Observatories, Chinese Academy of Sciences, Beijing, 100012, China (e-mail: yangc616@jlu.edu.cn, zhaomh20@mails.jlu.edu.cn, zhaohs@jlu.edu.cn).

L. Bruzzone is with the Department of Information Engineering and Computer Science, University of Trento, 38050 Trento, Italy (e-mail: lorenzo.bruzzone@unitn.it).

R. C. Guan is with the Key Laboratory for Symbol Computation and Knowledge Engineering of the Ministry of Education, College of Computer Science and Technology, Jilin University, Changchun, 130012, China (e-mail: guarenchu@jlu.edu.cn).

I. Introduction

MARS is currently the focus of many planetary science missions and research [1], [2]. Impact crater excavated on the surface of Mars is one of the prominent geologic structure and feature of Martian geomorphology. Impact craters have been a standard metric for the analysis of evolution on Mars, e.g., age-dating, subsurface minerals, geomorphologic and geologic mapping [3]-[10]. Compared to other typical landforms on Mars, the widely distributed craters exhibit circular or elliptical shapes. However, craters on Mars may be influenced and modified due to the Martian endogenic and exogenic dynamics [11]. Therefore, different morphological features of craters are presented on the Martian complex topography, including glacial, volcanic, aeolian, and fluvial landforms. Comprehensive identification of Martian craters is a necessary yet challenging task for revealing the evolutionary history of Mars.

The mapping of Martian craters and the establishment of databases are mainly based on long lasting manual counting by experts in planetary sciences. At present, the manually identified Martian craters are related to the works of Barlow et al. [12], [13], Robbins et al. [14] and Lagain et al. [15]. Among them, the most complete Martian crater database is given in Robbins et al. [14], contains 384,399 craters larger than 1 km in diameter. Lagain et al. [15] assessed and evaluated this crater database to correct misidentified entries. This process resulted in valuable classification information. Martian craters were classified into four types with their morphological characteristic, i.e., 288,155 standard craters, 8445 layered ejecta rampart sinuous (LERS) craters, 24,530 partially buried craters, and 55,309 secondary craters. The craters search used both the imagery, i.e., the Thermal Emission Imaging System (THEMIS) images [16], and the digital elevation model (DEM), i.e., MOLA topographic maps [17].

Another possible approach is to use automatic detection with target recognition algorithms that can efficiently identify Martian craters rapidly. Traditional automatic methods generally use crater shape-based interpolation or classifiers for recognizing craters by exploring radiometry, texture, and gradient information [18]-[21]. Salamunićcar et al. [19] established a catalogue (MA132843GT) that resulted in 132,844 Martian craters with a diameter larger than 1km. In the identification of Martian craters. Among existing DL automatic

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2025.3543861

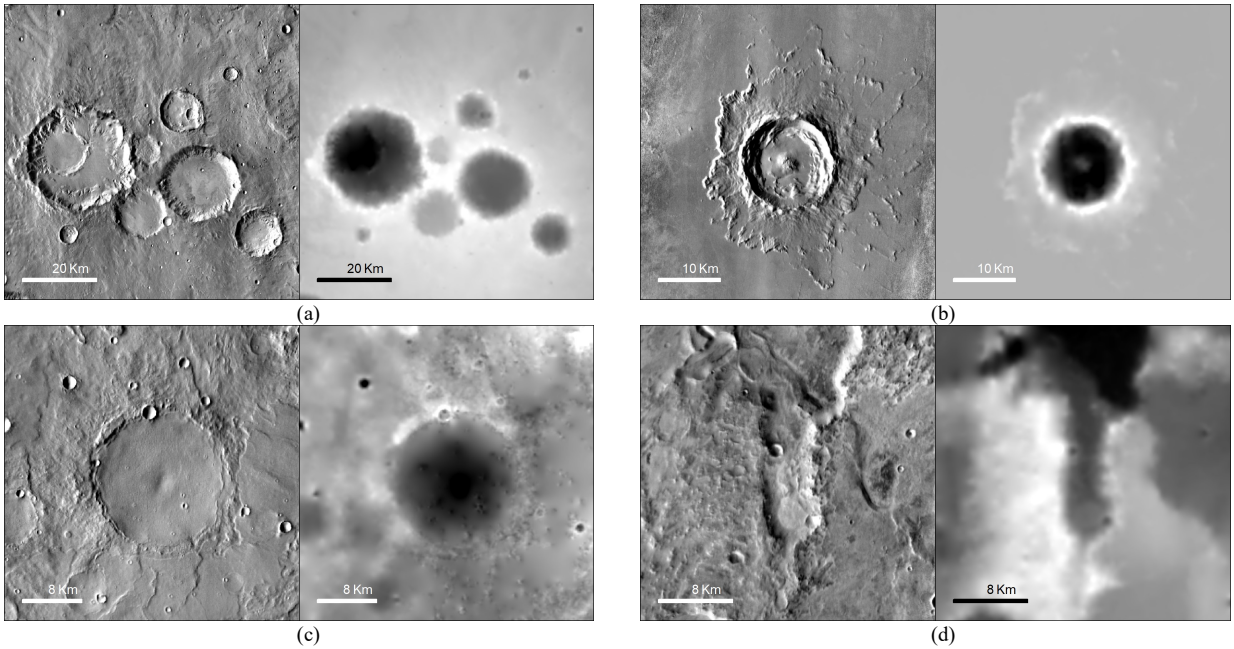

Fig. 1. Morphological features and topographical information of craters on Mar's surface (left: THEMIS; right: MOLA). (a) The standard craters are mainly observed on highlands and south of the dichotomy. There is a small amount of these craters in Northern lowlands, the Tharsis region, and terrains at latitudes higher than 45°N and 45°S due to the intensity of resurfacing processes. (b) The layered ejecta rampart sinuous (LERS) craters are filled in the low-middle latitudes, highland-plains and near the dichotomy boundary. (c) The partially buried craters are distributed at the southern highlands terrains. (d) A large number of secondary craters exists in the northern Isidis Planitia.

detection methods, convolutional neural networks (CNNs) -based models (i.e., LeNet, AlexNet, GoogleNet [22], CraterIDNet [23], U-Net [24],[25], Faster R-CNN[26], FPNet [27], and YOLO [28]), became the most widely used for the identification of craters. However, the above automatic methods have not identified Martian craters on a global scale. There are three factors for the limited range of automatic recognition. First, significant differences in Martian crater scales make it difficult to meet the needs of all scale-craters. As the number of impact structures increases exponentially with the small sizes, existing methods mostly focus on improving the prediction of small craters [27]-[30]. Second, the complex geological activities on Mars lead to area differences in the spatial distribution and to the Martian craters with different morphological and topographical features in the images and the DEM data (Fig.1). The automatic identification of craters in specific areas cannot fully consider the characters of the Martian craters. More importantly, existing models focus on data fusion between imagery and DEM data, making it difficult to obtain satisfactory results in the identification of dynamics influenced Martian craters.

According to the above analysis, in this paper we start from the manual counting database [14],[15], and also integrate the craters provided by the automatic recognition method [19] for obtaining a Martian craters dataset that contains 54,370 craters with different types and various scales (i.e., multi-type and multi-scale craters). Then, we convert the Martian craters identification problem into an object detection task and map craters using both the imagery and DEM data using a cross modality feature fusion Model. In greater detail, the main contributions of this paper are as follows.

1) For accurate and comprehensive automated identification of Martian craters, an integrated Martian crater data set distributed in the areas from 60°S to 60°N and from 180°W to 180 ° E on Mars (the mid- and low-latitude regions) is established. It contains large and medium-sized Martian craters of different types. Then, nearly 50% of the Martian craters remapped by glacial, volcanic, aeolian, and fluvial landforms were used for training.

2) A dual CNN-Transformer-based cross modality adaptive feature fusion network (DCT-CMAFFNet) is proposed. To fully capture the morphological features and the topographical information of multi-type and multi-scale craters, we designed a dual-branch architecture that can extract rich local features and long-range global context of Martian craters by using the CNN-Transformer hybrid model [31][32]. To comprehensively utilize the extracted morphological features and topographical information, we also established a cross-modality adaptive feature fusions module (CMAFF). The CMAFF model utilizes a self-attention mechanism to learn the relationship between the two modalities data, i.e., imagery and DEM, and weight each position of input feature maps to ensure effective identification of multi-type and multi-scale Martian craters. Compared with the methods with either only imagery or DEM and data fusion models, the proposed DCT-CMAFFNet has an 14%-24% improvement in detection accuracy.

3) 3166 Martian craters throughout low- to mid-latitudes on Mars are identified, which is more than one-tenth of the initial number of the annotated craters. Then, a cross modality adaptive feature fusion of the Martian crater database has been derived and made available, which is characterized by a low

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2025.3543861

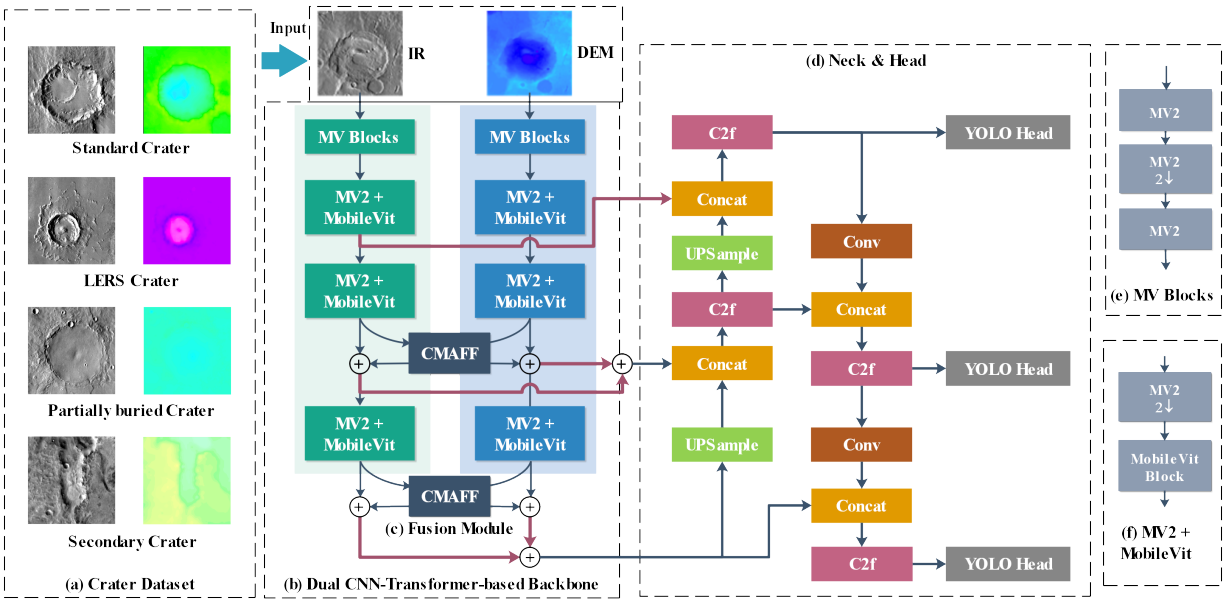

Fig. 2. Flowchart of the proposed dual CNN-transformer-based cross modality adaptive feature fusion network (DCT-CMAFFNet). (*a*) The crater dataset contains four types of craters, standard, LERS, partially buried, and secondary craters; (*b*) The dual CNN-transformer part consists of IR and DEM channels given as input to a MobileViT backbone, and contains one MV Blocks and three MV2+MobileViT; (*c*) The cross-modality adaptive feature fusions module (CMAFF) is adopted to fuse the imagery and DEM channels; (*d*) Neck and Head are the original of YOLOV8; they concatenate the feature maps and predict the bounding boxes, objectness scores, and classification scores.; (*e*) MV Blocks contains three MV2 modules, 2 ↓ indicates downsampling; (f) MV2+MobileViT contains a downsampling MV2 module and a MobileViT module.

amount of false positives, i.e., 12.9% for the multi-scale craters.

The rest of this paper is structured as follows. Section II introduces the related work on Martian crater identification based on current deep learning technology and the specific motivation of the proposed method. Section III describes the proposed DCT-CMAFFNet in detail. Experiments and analysis of results are given in Section IV. In Section V, we draw the conclusion of the proposed work.

II. Related Works and Motivation

A. Object Detection with CNN and Transformer

Object detection is a crucial computer vision task that makes the machine identify and localize objects in images [26]-[30]. Deep convolutional networks can learn robust and high-level feature representations of images, and have demonstrated greater accuracy than the other methods in object detection. These object-recognition methods have evolved into three main categories, i.e., the two-stage, one-stage, and ViT-based detectors. The two-stage detectors, such as Faster R-CNN [26], follow a coarse-to-fine processing paradigm. The coarse strives to improve recall ability, while the fine refines the localization based on the coarse detection and places more emphasis on discrimination ability. On the contrary, one-stage detectors, such as YOLO [28], can retrieve all objects in one-step inference. The one-stage detectors are faster than the two-stage ones, but their performance suffers noticeably when detecting dense and small objects.

Original CNN models can only capture local information, while lacking global context. Transformers were introduced in 2017 [33] and have been widely used in natural language processing. In 2020, the Vision Transformer was adapted for computer vision. ViT has demonstrated superior performance on large-scale datasets like ImageNet [34]. By adjusting the patch size, ViT can handle different image resolutions and sizes without a fixed input dimension. The self-attention mechanism allows ViTs to capture global relationships and dependencies across the entire image, potentially leading to better understanding and feature representation than CNN's local receptive fields [35]. Object detection with ViT adapts the core principles of the Vision Transformer architecture to identify and locate objects within an image. The ViT-based methods, such as the DETR [36], Deformable DETR [37], and Co-DETR [38], achieve extremely high accuracy on the COCO and large vocabulary instance segmentation datasets. However, these methods typically require large amounts of labeled high quality data to reach their full potential. The training ViT can be computationally expensive due to the quadratic complexity of the self-attention mechanism with respect to the input sequence length. On small datasets, they may underperform compared to the CNNs unless pre-trained on large datasets.

B. Identification Based on Feature Fusion

In the identification process of Martian craters, imagery and DEM data present different characteristics. Their effective fusion can further improve the performance of crater detection by utilizing complementary information [39-41]. For example, the THEMIS imagery captures longer thermal wavelengths and can characterize the morphological features of craters. It has a poor ability to identify buried craters. Compared with imagery, the DEM data provides three-dimensional information, which can be used to detect the craters by using the topographical

information. However, it is difficult to distinguish craters either in the areas with complex terrain or having small size.

With regard to the fusion strategy, existing fusion methods can be mainly classified into three families, i.e., data-level fusion, feature-level fusion and decision-level fusion methods [29]. The data-level fusion, also called pixel-level fusion [39], usually pre-processes the imagery and DEM data to form a three-channel image to provide as input to the network. However, the imagery and DEM data often contain inconsistent information on the characteristics of Martian craters. This simple fusion strategy cannot address the large dynamics of Martian craters. The feature-level fusion methods adopt multi-modal data that are fed into parallel branches [40]. Separate features of different modalities are extracted and then combined by some operations such as attention modules or simple concatenation [41]. The decision-level fusion methods integrate the detection results at the final stage and are more suitable for tasks with large differences in domains.

At present, existing Martian crater detection methods usually focus on single-source data or multi-source data-level fusion. In contrast, to make full use of imagery and DEM data, we establish a feature-level fusion module, i.e., the cross modality fusion module, by leveraging and balancing the morphological and topographical information presented in the imagery and DEM data to extract the features of multi-type and multi-scale Martian craters.

III. Methodology

In this section, the proposed DCT-CMAFFNet technique is described in detail. We designed a unique two-branch cross modality adaptive feature fusion network for Martian crater identification with CNN-Transformer, in which both imagery and DEM data can be collaboratively learned. The MobileVit network is reused as the backbone network for each branch and embeds the cross-modality adaptive feature fusions module (CMAFF) to facilitate modal fusion and interaction. Figure 1 illustrates the architecture of the proposed DCT-CMAFFNet. Considering the difference in the resolution of the imagery and DEM data, feature fusion occurs only at the outputs of the last two MobileVit modules, and the fused results are then fed into the neck network.

A. Dual CNN-Transformer Architecture

In the MobileViT backbone network, a hybrid architecture is defined by combining CNN and Transformer for leveraging the spatial inductive bias of CNNs. It can accelerate convergence and inference, while reducing the parameter count of the ViT by introducing spatial inductive bias parameters. In Fig. 3, MV2 represents the traditional lightweight CNN-based MobileViTv2 network. This module consists of two 1×1 convolutional blocks and a 3×3 depthwise convolutional block. The module marked with ↓ indicates a stride set to 2 (Fig 2. (e) and (f)), which is used for down sampling.

Fig. 3. The architecture of MV2 module.

The MobileViT network includes convolution, Transformer, and residual connection fusion modules, as illustrated in Fig. 4. MobileViT first extracts local features from imagery with an n×n convolution on the feature map, then projects vectors into a high-dimensional space through a 1×1 convolution that linearly combines them across the channels. Then the Transformer module employs an "unfold-transformer-fold" mechanism to extract global features. Subsequently, a 1×1 convolution projects the feature map back to a low-dimensional space. Finally, the feature map is concatenated with the original input along the channel dimension and fused with a 3×3 convolution to produce the final global feature output by using a shortcut connection.

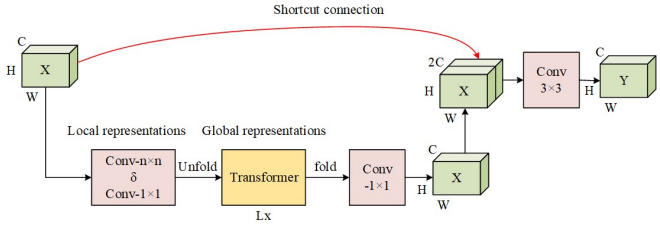

Fig. 4. The architecture of MobileViT.

B. Cross Modality Feature Fusion Module

The imagery feature map F_T and the DEM feature map F_D are obtained via the MobileViT block, and are given as input ito the CMAFF. Fig. 5 shows the proposed CMAFF architecture. The CMAFF first flattens each feature map and substitutes the matrices to obtain the vectors $I_T \in \mathbb{R}^{HW\times C}$ and $I_D \in \mathbb{R}^{HW\times C}$. Then, vectors of each modality are concatenated and a learnable positional embedding is added to obtain the input vector $I \in \mathbb{R}^{2HW\times C}$ of the Transformer. To extract the multiple complex relationships from different representational subspaces at different locations, we employ a multi-head attention mechanism [33], [37]. Next, by projecting the input sequence I onto the three weight matrices W^Q, W^K, and W^V (each with a size of $C\times D$) a set of queries, keys, and values is calculated as follows.

$$Q = IW^Q, \tag{1}$$

$$K = IW^K, \tag{2}$$

$$V = IW^V. \tag{3}$$

The self-attention layer uses the scaled dot product between Q and K to compute the weights, which are then multiplied by the value V to compute the output Z.

$$Z = \text{Attention}(Q,K,V) = \text{soft max}\left(\frac{QK^T}{\sqrt{D}}\right)V, \tag{4}$$

$$Z' = \text{MultiHead}(Q,K,V) = \text{Concat}(Z_1,...,Z_h)W^O, \tag{5}$$

where $1/\sqrt{D}$ is a scaling factor, Z' is the result of MultiHead, the subscript h denotes the number of heads of the multi-head attention mechanism, and $W^o \in \mathbb{R}^{h\cdot C\times C}$ denotes the projection matrix of Concat(Z_1, ..., Z_h). Finally, the output sequence I' is computed by the MLP, which has the same dimensions as the input sequence I.

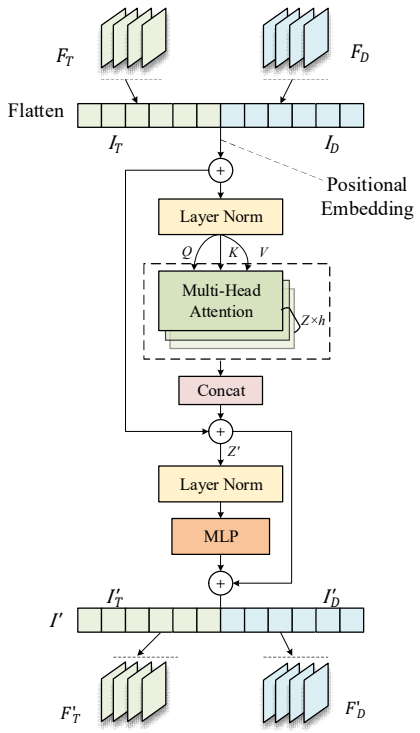

Fig. 5. The proposed cross-modality adaptive feature fusions module.

$$\begin{aligned} I' &= \mathrm{MLP}\left(Z'+I\right)+Z', \\ &= \mathrm{FC}\left(\mathrm{GELU}\left(\mathrm{FC}_1\left(Z'+I\right)\right)\right)+Z+I. \end{aligned} \tag{6}$$

As shown in Equation (6) and Fig. 5, it is generated by using a two-layer fully-connected network, with a Gaussian Error Linear Unit (GELU) [42] activation function. I' consists of I'_T and I'_D, which are converted into the recalibration results F'_T and F'_D.

The cross-modal fusion module utilizes a self-attention mechanism to learn the binary relationship between the IR and the DEM modalities, and uses a feature correlation matrix to weight each position of the input feature map. In Equation (7), $\alpha_{i,j}$ denotes the correlation between the *i*th position and the *j*th position on the feature map. $(1,...,HW)$ is the position index of IR channel and $(HW+1,...,2HW)$ is the position index of the DEM channel, so when calculating the correlation feature matrix α, the matrix can be partitioned into four matrix blocks, two blocks are intra-modality correlation matrix blocks, while the other two are inter-modality correlation matrix blocks.

The CMAFF module has eight repeating Transformer blocks. When the input image size is 640×640, it is down-sampled twice (160×160). Unlike the original CFT module [43], we use a global maximum pooling technique to down sample the feature map to a fixed low resolution (10×10) to reduce computation. The output is unsampled to the original resolution by bilinear interpolation before adding it to the original modal branch.

$$\alpha = \mathrm{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right) \tag{7}$$

$$= \begin{bmatrix} \alpha_{1,1} & \alpha_{1,2} & \cdots & \alpha_{1,HW} & \alpha_{1,HW+1} & \cdots & \alpha_{1,2HW} \\ \alpha_{2,1} & \alpha_{2,2} & \cdots & \alpha_{2,HW} & \alpha_{2,HW+1} & \cdots & \alpha_{2,2HW} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \alpha_{HW,1} & \alpha_{HW,2} & \cdots & \alpha_{HW,HW} & \alpha_{HW,HW+1} & \cdots & \alpha_{HW,2HW} \\ \alpha_{HW+1,1} & \alpha_{HW+1,2} & \cdots & \alpha_{HW+1,HW} & \alpha_{HW+1,HW+1} & \cdots & \alpha_{HW+1,2HW} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \alpha_{2HW,1} & \alpha_{2HW,2} & \cdots & \alpha_{2HW,HW} & \alpha_{2HW,HW+1} & \cdots & \alpha_{2HW,2HW} \end{bmatrix}$$

C. Neck and Head Networks

The proposed DCT-CMAFFNet can extract the features of multi-type and multi-scale Martian craters from MobileViT blocks within the CMAFF module. The neck of YOLO is responsible for aggregating and refining features at multiple scales, enabling the model to detect objects of different sizes effectively. Here, we also utilize a Path Aggregation Network (PANet)-like structure in the neck for multi-scale feature fusion. The UpSample modules in the Fig. 2 (d) employs bilinear upsampling to spatially align features at different levels, followed by concatenating them into a matrix. The C2f module combines low-level spatial details from earlier layers and high-level semantic information from deeper layers to enrich the feature representation. The top-down and bottom-up pathways of the neck network enhance detection capabilities for small, medium, and large objects by integrating features at different scales.

Finally, the head module of the proposed DCT-CMAFFNet provides the Martian crater detection results, including the coordinates of the bounding boxes to determine the location of the crater (i.e., the central coordinates of the detection boxes as well as their width and height values) and the confidence scores (ranging between 0 and 1) that indicate the model's confidence that the detected object is a crater. Moreover, by deploying head modules on different scale feature layers, the accurate detection of small, medium and large craters is achieved.

IV. Experimental Results and Analysis

This section first introduces the integrated Martian craters dataset, presents the training areas of craters on Mars and the evaluation criteria. Then, extensive experiments on THEMIS imagery and DEM data are carried out to verify the effectiveness of the proposed DCT-CMAFFNet. The results obtained in different types and various scales of the Martian craters are presented and discussed. Meanwhile, the newly identified craters are compared with the existing global Martian crater database. To further prove the robustness of the proposed model, ablation study is performed.

A. Description of Data Sets

THEMIS imagery – The Thermal Emission Imaging System (THEMIS) on board the Mars Odyssey spacecraft contains both visible/near-infrared and thermal infrared channels [16]. The THEMIS maps the entire Mars in day and night multi-spectral infrared images at 100m/pixel resolution. The morphological

feature of Martian craters was extracted using the THEMIS daytime infrared (IR) mosaic imagery which was released in 2014 by Arizona State University [44].

MOLA and HRSC DEM - The Mars Obiter Laser Altimeter (MOLA) is an instrument on the Mars Global Surveyor (MGS) [17]. MOLA obtains topography information on the surface of Mars by receiving reflections of laser pulses emitted to the planet. A blended digital elevation model (DEM) data [45] with a spatial resolution of 200m/pixel was derived from the MOLA and the High Resolution Stereo Camera (HRSC) of the European Space Agency's Mars Express (MEX) spacecraft. The MOLA and HRSC DEM data were used for learning the topographical features of Martian craters.

Integrated Martian crater data set – To ensure the reliability of Martian craters, we selected the craters with different scales and types in two manually identified Martian craters databases, i.e., Robbins et al. [14] and Lagain et al. [15]. Also, the automatic detection crater catalogue, i.e., Salamunićcar et al. [19] from 60°S to 60°N and from 180°W to 180°E on the Mars (the mid- and low-latitude regions) was considered. A Martian crater is regarded as correct when it belongs to Robbins et al. [14], Lagain et al. [15] and Salamunićcar et al. [19] simultaneously. The craters present only in one or two catalogues, are considered as disputed ones. The integrated Martian crater data set contains 54,370 craters with diameters between 3km and 220km. 26,649 Martian craters in 50% of the areas are in the training set and 27,721 craters in the remaining 50% are in the test set. In this data set, the types of Martian craters have been classified according to Lagain et al. [15], i.e., 36,692 standard craters, 6613 LERS craters, 10,754 partially buried craters, and 324 secondary craters, in which various morphological characteristics of Martian craters can be presented.

B. Experimental Settings

1) Training and Testing Areas

The distribution of Martian craters presents complexity and heterogeneity, i.e., different types and scales of Martian craters are scattered throughout Mars' surface. In this paper, a random partitioning strategy was utilized to ensure the identification of Martian craters comprehensively. The mid- and low-latitude regions of Mars are divided into training and testing areas with a certain interval, as shown in Fig 6. This partitioning strategy makes the number of Martian craters with different types and scales in training and testing areas relatively balanced. Table I lists the number of craters with the above four types in the integrated Martian crater data set for training and test sets. The data preprocessing methods used are as follows.

1. Data Matching: Down the THEMIS-IR images to a spatial resolution of 200 meters per pixel as the DEM data.

2. Coordinate System Alignment: Use Mercator projection to align the two images into a unified coordinate system.

3. Image Rendering: Render the DEM into RGB color using the ArcMap renderer.

4. Data Partitioning: Split the projected mosaic into 640×640 tiles with 50% overlap to improve the identification accuracy of Martian craters located at the boundary of the tiles. In the illustrative example of Fig. 7, the craters in the blue region can be divided into four tiles to ensure a relatively complete presentation.

Meanwhile, the tiles that do not contain Martian craters have been removed. Finally, the training set contains 20,510 tiles, and the test set contains 20,093 tiles.

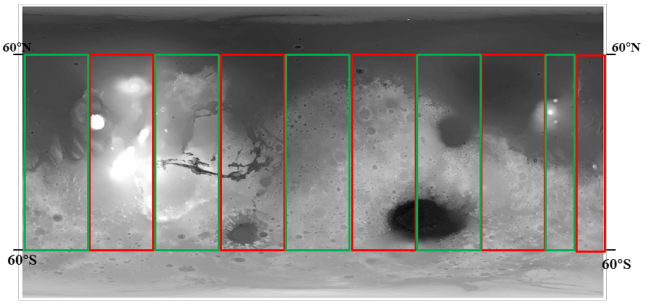

Fig. 6. Training and testing areas on the surface of Mars. The Martian coverage from 60°S to 60°N and from 180°W to 180°E is divided into training and testing areas with 40°, 40°, 40°, 40°,40°, 40°, 40°, 40°, 20°, and 20° interval. The green box regions represent the training areas and the red box regions represent the test areas.

TABLE I
THE NUMBER OF CRATERS WITH DIFFERENT TYPES IN THE INTEGRATED MARTIAN CRATERS DATA SET

Types	Training	Test	Total
Standard craters	17,857	18,844	36,692
LERS craters	3325	3288	6613
Partially buried craters	5305	5427	10,754
Secondary craters	162	162	324

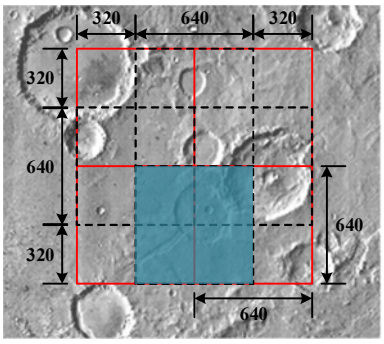

Fig. 7. Schematic diagram of the 50% overlap tile. The red solid boxes are the neighboring tiles, and the black dashed boxes are tiles with 50% overlap with the red box tiles.

We also split the mosaic after projection into 2560×2560 tiles to identify large Martian craters, which are down sampled to 640×640 for input into the model.

2) Implementation Details

All the experiments are conducted on an Ubuntu system with Python V3.9 and PyTorch 1.9.0, respectively. The CPU is an Intel Core1 i7-12700F and the GPU is a 24 GB RTX3090. We use SGD optimizer and the same hyperparameters during the training process in all experiments (see Table IV). Considering the characteristics of DEM data, the training process retains simple data augmentation methods, i.e., symmetry and rotation. In the inference step, we set the IoU threshold to 0.6 and the

confidence threshold to 0.2. In the post-processing step, we set the de-duplication IoU threshold to 0.2 to remove duplicate identified Martian craters. The confidence threshold is an evaluation of uncertainty for each identified crater. as a comparison, we use YOLOv8 middle structure as the base line.

TABLE II
RECALL AND PRECISION OF MARTIAN CRATERS IDENTIFICATION IN DIFFERENT CONFIDENCE LEVELS

Models	Data Source	conf=0.2		conf=0.25		conf=0.30		conf=0.35	
		R	P	R	P	R	P	R	P
YOLOV8	IR	0.792	0.916	0.790	0.918	0.787	0.919	0.784	0.921
YOLOV8	DEM	0.732	0.935	0.730	0.936	0.728	0.938	0.722	0.939
YOLOV8	Data-level IR-DEM	0.767	0.904	0.765	0.906	0.763	0.907	0.761	0.908
Proposed DCT- CMAFFNet	IR&DEM	**0.907**	0.888	**0.903**	0.893	**0.897**	0.898	**0.891**	0.903

TABLE III
RECALL OF THE MARTIAN CRATERS IDENTIFICATION IN DIFFERENT TYPES WHEN CONF=0.2

Models	Data Source	The Types of Martian Craters			
		Standard	LERS	buried	Secondary
YOLOV8	IR	0.907	0.974	0.649	0.599
YOLOV8	DEM	0.838	0.957	0.612	0.457
YOLOV8	Data-level IR-DEM	0.883	0.948	0.613	0.475
Proposed DCT- CMAFFNet	IR&DEM	**0.939**	**0.986**	**0.792**	**0.642**

TABLE IV
HYPERPARAMETRIC VALUES

Hyperparameters	Value
Initial earning Rate	0.01
Learning rate factor	0.1
Momentum	0.937
Weight Decay	0.0005
Epoch	150

3) Evaluation Criteria

In order to assess the performance of the proposed DCT-CMAFFNet, three quantitative evaluation indicators are adopted, including recall (R), precision (P), and false-positive rate ($FPRs$) of the newly identified multi-type and multi-scale Martian craters. All these quantitative indices range from 0 to 1. Meanwhile, the execution time of proposed DCT-CMAFFNet is reported to assess the computing complexity. The recall (R), precision (P), and false-positive rate ($FPRs$) are defined as:

$$R = \frac{TP}{TP + FN} \tag{9}$$

$$P = \frac{TP}{TP + FP} \tag{10}$$

$$FPR_s = \frac{FP}{FP + TN} \tag{11}$$

where true positive (TP) is the number of Martian craters that were identified correctly; false positive (FP) is the number of detected objects that were not a true Martian crater; false negative (FN) is the number of Martian craters was missed by the detectors; true negative (TN) is the number of objects that were correctly predicted as crater. Compared with other target recognition tasks, the crater identification task is more concerned with recall and precision, i.e., the rate at which the detector can find correct craters and new craters. At the same time, we selected all the newly identified Martian craters for false positive analysis.

C. Performance on Multi-type and Multi-scale Martian Craters

We conducted experiments considering separately the two single data sources (i.e., IR imagery and DEM data), a data fusion strategy (i.e., data-level fusion of IR-DEM), and the proposed DCT-CMAFFNet feature fusion strategy (i.e., feature-level fusion of IR&DEM). The data-level fusion strategy is based on a three-channel image consisting of IR imagery, DEM profile curvature, and DEM slope in degree. In feature fusion, i.e., IR&DEM, "&" indicates two-branch input. Table II shows the recall (R) and the precision (P) of crater identification results with confidence thresholds at 0.2, 0.25, 0.3, and 0.35. Table III lists the recall of the identification of Martian craters in different types. Bolded fonts indicate optimal results under the current evaluation criteria.

From Table II, one can observe that the recall decreases sequentially by increasing the confidence threshold, while precision increases by increasing the confidence threshold. The proposed DCT-CMAFFNet always exhibits the highest performance object detection accuracies, i.e., 0.907 in recall and 0.888 in precision when conf=0.2. The average identification time required for each image was 0.0131s. The recall value indicates that almost all of the recognized Martian craters in the test set can be detected with the proposed DCT-CMAFFNet; whereas from the precision, we can infer that the model identified many new Martian craters.

Table III show the identification accuracy of four types of Martian craters (i.e., standard, LERS, partially buried, and secondary craters) when different backbones and data sources were employed on the test set (conf=0.2). From the table III, one can see that the proposed DCT-CMAFFNet achieves the highest recall (i.e., 0.939 and 0.986 standard and LERS craters, respectively). For partially buried craters, the recalls decrease in all the models, especially in DEM and data-level IR-DEM that resulted in only 0.612, and 0.613, respectively. Aftercomplementing the advantages of IR and DEM data, the recall was improved to 0.792 with the proposed DCT-CMAFFNet. The secondary impact craters are limited by

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2025.3543861

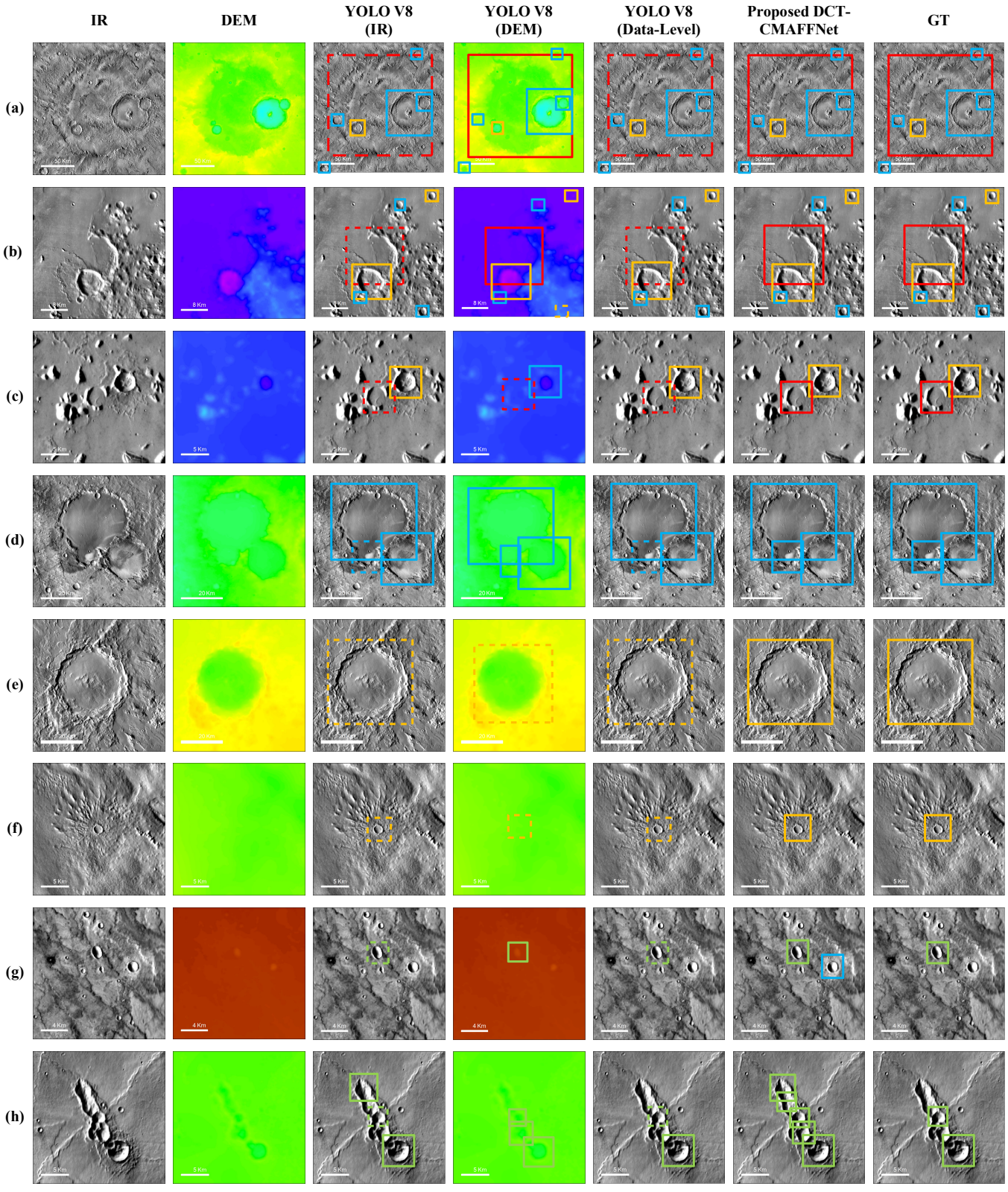

Fig. 8. Visual comparison of the identification results of different types of Martian craters obtained with the proposed DCT-CMAFFNet model. Different types of Martian craters, i.e., partially buried craters, standard craters, LERS craters, and secondary craters are labeled with red, blue, orange, and green squares, respectively. The dashed squares represent undetected craters. (a) A large crater, five standard craters and a LERS crater (Lat: 1.773°S, Long: 55.879°E); (b) one partially buried crater, two LERS craters and three standard craters (Lat: 21.696°N, Long: 172.703°E); (c) one partially buried crater and one LERS crater (Lat: 20.5°N, Long: 164.921°E); (d) three standard craters (Lat: 8.414°N, Long: 112.768°W); (e) one large LERS crater (Lat: 2.516°N, Long: 129.378°W); (f) one LERS crater (Lat: 2.516°N, Long: 129.378°W); (g) one secondary crater and one standard crater (Lat: 17.318°S, Long: 101.757°W); (h) several secondary craters (Lat: -17.643°S, Long: 109.607°E).

their small scale and their unique morphology characteristic, and all models have lower recall accuracies than the other types of Martian craters. The topographical information extracted from the DEM data is insufficient to reflect the feature of secondary impact craters. Therefore, fusing morphological features from imagery makes the recall significantly increased.

We selected typical results of the multi-type and multi-scale Martian craters of the above models for visual comparison from the test set. Fig. 8 (a), (b) and (c) show partially buried craters, standard craters, and LERS craters with different scales and morphologies. One can see that from the detection results of the IR imagery and data-level IR-DEM that it was difficult to identify the severely degraded partially buried craters by relying on imagery only or pixel-level fusion. Although the DEM data provides critical information for buried craters, they do not properly represent small-scale partially buried craters with unclear boundaries (see Fig. 8 (b) and (c)). Fig. 8 (d) shows three standard craters having large, middle and small scales, where the boundary of the small standard crater is unclear in the IR imagery.

On the contrary, due to the feature fusion in the proposed DCT-CMAFFNet, the multi-scale buried craters can be detected well compared to other methods. Fig. 8 (e) shows a large LERS crater that can be only detected in DEM and by the proposed DCT-CMAFFNet. Fig. 8 (f) contains one LERS crater with a small scale that is hard to identify for the two single data sources, and the data-level fusion strategy, whereas the proposed DCT-CMAFFNet could adaptively focus on the crater features. For the secondary craters in both Figs. 8 (g) and (h), the proposed model can distinguish the morphologic features and the topographic information of dense secondary craters comprehensively, detecting a large number of secondary craters. The IR imagery and the data-level IR-DEM did not recognize the secondary craters affected by terrain (see Fig 8 (g)) or craters overlapping with each other (see Fig 8 (h)).

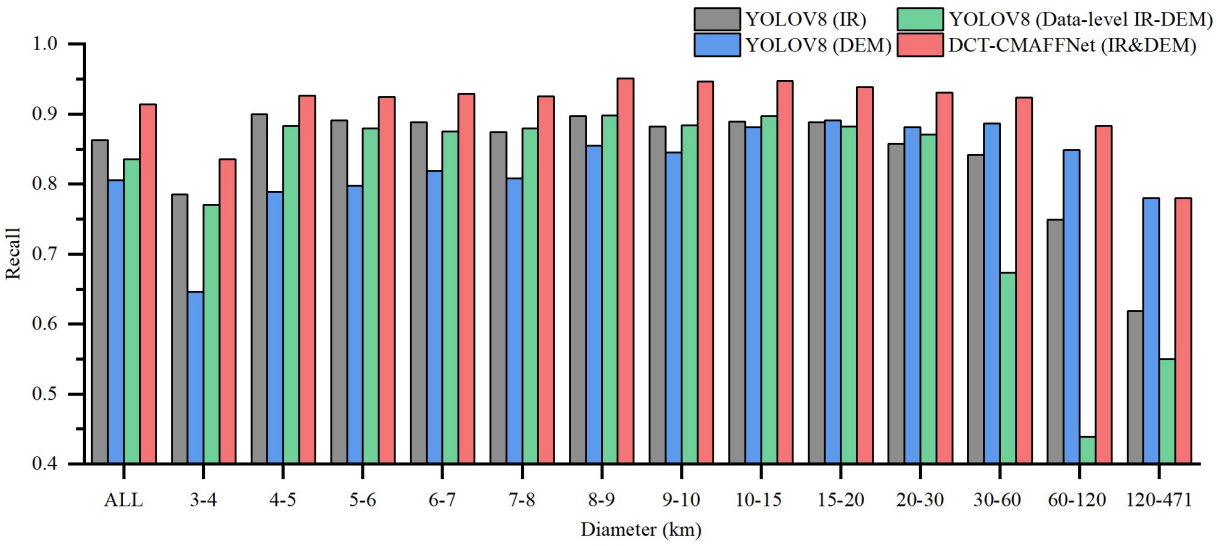

Fig. 9. Recall comparison of impact crater identification results at different scales.

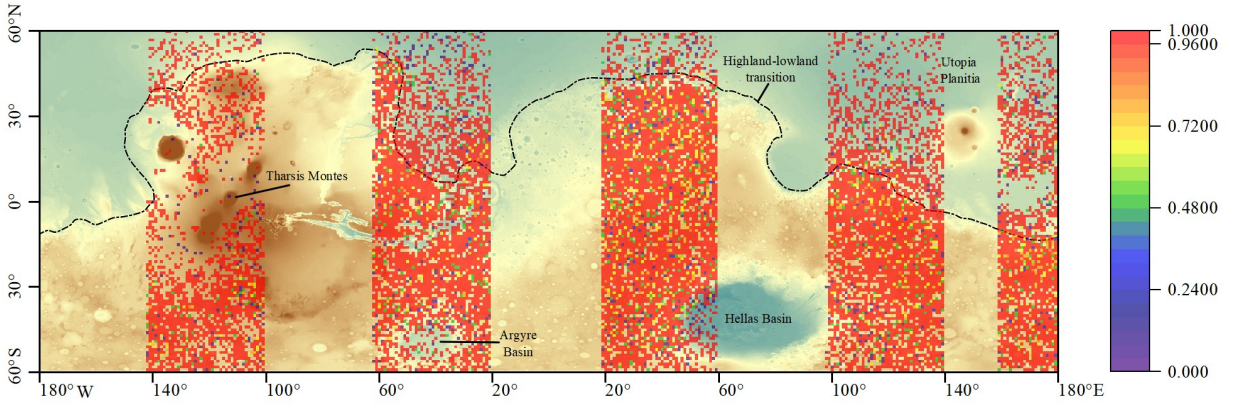

Fig. 10. Recall map in the testing area with the recall values shown by color bars. The gray region is the training area, the colorful region represents the test area, and the blank areas are those that do not contain >3km Martian craters.

We also evaluate the detection results versus the scales of Martian craters, as shown in Fig 9. The recalls of YOLOs V8 in DEM data are much lower than those in the IR imagery, the data-level IR-DEM and of the proposed DCT-CMAFFNet in the relatively small crater diameter sizes, i.e., from 3km to 10km; while the decline of the imagery recalls occurs in craters

larger than 30km. From a whole perspective, the proposed DCT-CMAFFNet performed the best in the recall of multi-scale Martian craters detection results. Especially when the diameter of the Martian crater is between 3km and 4km and larger than 30km, the recall of the proposed DCT-CMAFFNet is sharply higher than those of the IR imagery, the DEM data and the data-level IR-DEM. For small scale, i.e., between 3km and 4km, due to low resolutions the DEM data was not beneficial to extract features of craters, leading to a decrease of the recall of data-level IR-DEM. For larger scale, i.e., the diameter of craters between 30km and 471km, many partially buried and degraded craters exist. More than half of the craters between 60-471km in the test sets are partially buried and degraded craters. In this context, the IR imagery and the data-level IR-DEM cannot provide the information of partially buried and degraded craters, thus the recall drastically decline. On the contrary, the recall of the DEM data and the proposed DCT-CMAFFNet are slightly decreased in the identification of large-scale Martian craters. In summary, the proposed DCT-CMAFFNet can fully fuse the morphological features and the topographical information from the IR imagery and DEM data that achieves more than 90% recall at almost all scales of Martian craters.

Fig. 10 shows the recall map in the testing area. Each colored block of the testing area is 1° × 1°. In general, the Martian craters have a dense distribution and the proposed DCT-CMAFFNet achieved a high recall. However, some areas may experience low recall due to the influence of the altitude and terrain. Tharsis Montes, an area with three large shield volcanoes in the Tharsis region, contains relatively few craters. The recall in this region is low due to the high altitude of Tharsis Montes above the average Martian elevation that results in many craters exhibit unique morphological features and topographical information. Owing to the complex Martian endogenic and exogenic dynamics more blue and green color blocks appear at the left and right boundary of the Hellas Basin, the north area of the Argyre Basin, the Utopia Planitia, and the highland-lowland transition.

D. Comparison with the Global Martian Crater database

In the above section, the effectiveness of the proposed DCT-CMAFFNet has been proven. Then we identified Martian crater with the proposed DCT-CMAFFNet model throughout low- to mid-latitudes on Mars. Two complete manually identified Martian craters databases, i.e., Robbins et al. [14] and Lagain et al. [15], and one automatic detection crater catalogue, i.e., Salamunićcar et al [19], have been chosen for comparison. Fig. 11 shows the CSFD of identified Martian craters compared with the existing Martian crater databases. Table V lists the quantitative comparison results of identified Martian craters by using the proposed DCT-CMAFFNet with existing databases in different diameter scales. In the small-scale range, i.e., the diameter of the crater between 3km and 5km, the CSFD of Martian craters identified by the proposed DCT-CMAFFNet is lower than those of the manual count in Robbins et al. [14], Lagain et al. [15] and the automatic detection in Salamunićcar et al [19]. The identified small Martian craters have a certain mismatch with the manual databases, whereas the agreement increases as the diameter increases. This may be due to the fact that small craters are not detected sufficiently in our automated crater catalogue. On the contrary, the craters in the two automatically detected datasets have high consistency. For the crater diameter ranging between 5km and 50km, the CSFD of Martian craters in the four Martian crater datasets is basically comparable, and the matching assessment also increases accordingly. On a large scale with diameters between 50km and 471km, the CSFD of identified craters is higher than the existing Martian crater databases and catalogue. However, the matching assessment has decreased with the two manual databases. It might be due to the existence of disputed craters that presents degradation states. This was further analyzed in the subsection *E*.

TABLE V
THE NUMBER OF EXISTING MARTIAN CRATER DATABASES (NA) AND THE NUMBER OF MATCHED IDENTIFIED CRATERS (NM) WITH THE PERCENTAGE OF MATCHED CRATERS (MR) IN DIFFERENT DIAMETER SCALES.

Diameter (km)	Lagain's			Robbins'			Salamunićcar's		
	NA	NM	MR (%)	NA	NM	MR (%)	NA	NM	MR (%)
3-4	9405	5328	56.7	9613	5343	55.6	6037	4725	78.3
4-5	4985	3915	78.5	5143	3931	76.4	4074	3624	89.0
5-7	5733	4793	83.6	5879	4813	81.9	5461	4726	86.5
7-10	4611	4010	87.0	4775	4046	84.7	4317	3848	89.1
10-20	6002	5292	88.2	6210	5328	85.8	5831	5139	88.1
20-50	3903	3370	86.3	4088	3415	83.5	3712	3273	88.2
50-471	875	715	81.7	951	729	76.7	845	699	82.7
All	35,514	27,423	77.2	36659	27,605	75.3	30277	26,034	86.0

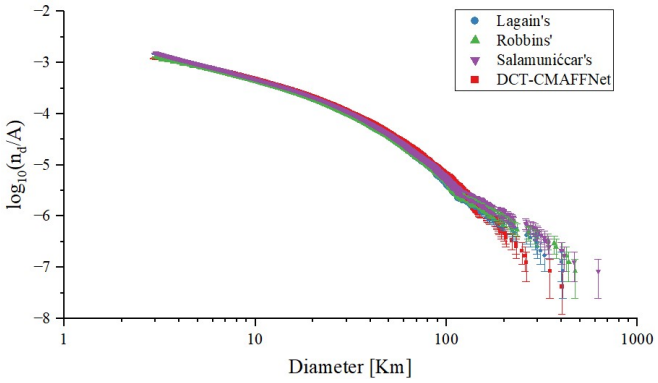

Fig. 11. The cumulative size-frequency distributions (CSFDs) of Martian craters with diameters between 3km and 471km in the existing databases and those identified in this paper.

E. False positive analysis

For manual assessment of the detection accuracy, the newly identified Martian craters (i.e., 3166 craters which are not included in the integrated Martian crater data set) are divided according to the scales of craters. The false-positive rates (FPRs) of the newly identified Martian craters are listed in Table VI. It should be noted that small craters with a diameter of less than 3km can be detected by the proposed DCT-CMAFFNet.

TABLE VI
FALSE-POSITIVE RATES OF NEWLY IDENTIFIED MARTIAN CRATERS

Diameter (km)	Num. newly identified	Num. false detection	FPR
1-3	1142	10	0.9%
3-4	965	35	3.6%
4-5	357	26	7.3%
5-7	307	61	19.9%
7-10	282	88	31.2%
10-20	347	148	42.7%
20-50	202	89	44.0%
50-249	33	12	36.4%
All	3635	469	12.9%

From Table VI, we can see that the FPRs of identified craters increase by increasing of the Martian craters scale. The proposed DCT-CMAFFNet detected a higher number of craters with a diameter of 1km-5km and obtained a relatively low false positive rate. However, a high false detection appeared at the medium-scale craters, i.e., 5km-50km. By analyzing the distribution of false detection craters, we found that some medium false detection craters are distributed in the transition zones and valleys, which exhibit large variations of Martian terrain and landforms. For large craters with diameters between 50km and 471km, the new detection is mainly related to the partially buried craters with serious degradation. Furthermore, we estimate the types of newly identified 3166 craters based on the criteria of Lagain et al. [15]. Craters without a complete boundary or those appearing only in DEM data are classified as degraded craters, comprising a total of 508 suspected degraded craters. Elliptical or chain-like craters surrounding large craters are classified as secondary craters. We identified 25 craters in this category. A total of 27 craters resembling those Fig. 8 (e) and Fig. 8 (f), with layered ejecta, are identified as LERS craters. The other craters that do not fall into the above categories are classified as standard craters.

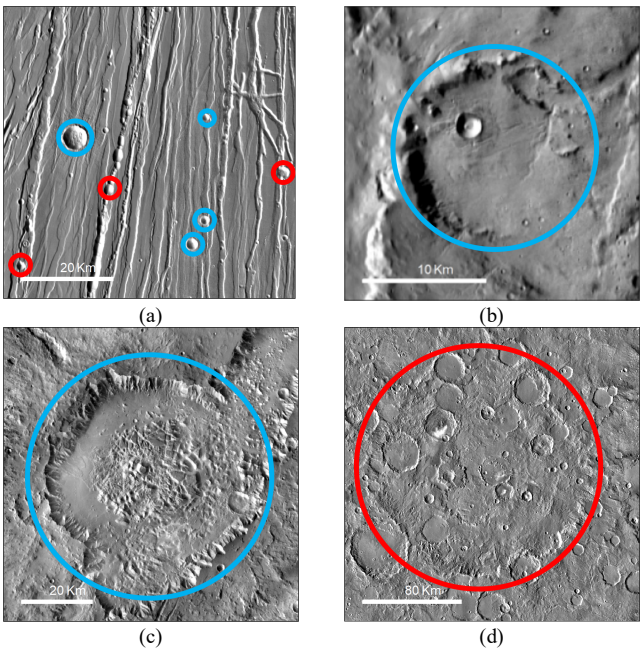

Fig. 12. Examples of newly identified Martian craters. (a) Lat: 25.711°N, Long: 103.373°W; (b) Lat: 26.933°S, Long: 26.853001°W; (c) Lat: 0.128°S, Long: 45.841°W; (d) Lat: 12.824°N, Long: 35.8281°E. Blue circles indicate confirmed Martian craters by experts, and red circles are misidentified or disputed craters.

Fig. 12 presents the examples of visualization results of newly identified Martian craters with small, medium, and large-scale. Fig. 12 (a) represents many newly identified small craters, all of which are less than 3km in diameter. Most craters in this area can be confirmed, even if some ones are still misidentified and disputed owing to the terrain. Fig. 12 (b) and (c) illustrate medium-scale craters, which are both partially buried craters and distributed in the Margaritifer Terra and Xanthe Terra, respectively. As shown in Fig. 12 (d), the DCT-CMAFFNet recognizes a large crater named Tikhonravov crater that is included in Robbins' database (ID 12-000001) [14] and Salamuniccar's catalogs [19]. However, some experts such as Lagain et al. argue that this is not a Martian crater. In general, the proposed DCT-CMAFFNet achieves high reliability and stability in the identification of Martian craters at different scales.

F. Ablation experiment

To further demonstrate the effectiveness of the proposed model, ablation experiments were conducted and are presented in this section. We evaluate the performance of component combinations. The results of ablation experiments are shown in Table VII. In Table VII, “CSPDarknet” is the original backbone of YOLO V8, which is a network structure composed entirely of CNNs. We define the first, second and third level fusions by adding a fusion module at the end of the corresponding MV2 +MobileViT block, i.e., at p3, p4 and p5 of the CSPDarknet backbone. “Add” is the addition of feature map of two channels.

The first five lines present the results of single source. The different levels of feature fusion are listed from lines 6 to 14. The last line reveals the results of the proposed DCT-CMAFFNet model.

TABLE VII
ABLATION STUDY FOR DIFFERENT COMBINATIONS OF THE PROPOSED MODEL COMPONENTS

#	Data Source			Backbone		Fusion module			Results	
	IR	DEM	Data-level	CSPDarknet	MobileVit	First level	Second level	Third level	R	P
1	√			√					0.792	0.916
2		√		√					0.732	0.935
3	√				√				0.902	0.744
4		√			√				0.807	0.783
5			√		√				0.767	0.888
6	√	√		√		None	CMAFF	CMAFF	0.852	0.911
7	√	√		√		None	ADD	ADD	0.849	0.918
8	√	√		√		None	CMAFF	None	0.840	0.915
9	√	√		√		None	None	CMAFF	0.812	0.912
10	√	√			√	None	ADD	ADD	0.902	0.755
11	√	√			√	CMAFF	CMAFF	CMAFF	0.907	0.787
12	√	√			√	None	CMAFF	None	0.895	0.828
13	√	√			√	None	None	CMAFF	0.892	0.811
14	√	√			√	None	CMAFF	CMAFF	**0.907**	0.888

As one can see from Table VII (#1, #6, #7, #9), when the CSPDarknet is adopted as a backbone network, the precision of Martian crater identification increases significantly but the recall decreases. By using the MobileViT as a backbone network, the recalls of all models were improved, with IR, DEM, and data-level fusion increasing to 0.904, 0.807, and 0.826, respectively. However, the precision shows a decline. When using "Add" with coarse feature fusion (#9), there is a slightly decrease in recall compared to YOLOV8-MobileViT with IR. Combining the MobileViT and different numbers of CMAFF with IR-DEM fusion, the proposed DCT- CMAFFNet achieves the same recall and better precision, i.e., 0.907 and 0.888 compared to the three CMAFF (#11). Only one CMAFF (#12, #13) lead to lower recall rates and higher precision rates compared to YOLOV8 with IR (#3). Due to the inconsistency of the spatial resolution of IR imagery and DEM data, the performance of crater detection cannot be improved by three CMAFF. The above results emphasized the crucial role played by the diverse components of the proposed DCT-CMAFFNet in obtaining the best detecion results for Martian craters.

The computational complexity and inference time of the YOLOV8, the YOLOV8-mobilevit and the proposed DCT-CMAFFNet are also evaluated (Table VIII).

TABLE VIII
MODEL'S COMPUTATIONAL COMPLEXITY, TRAINING AND INFERENCE TIMES

Models	Source	Parameters	Training times	Inference times
YOLOV8	Single	25.8M	10.2h	6.0ms
YOLOV8-mobilevit	Single	15.0M	13.16h	7.9ms
DCT-CMAFFNet	Double	17.5M	25.45h	13.8ms

As one can observe from Table VIII, the proposed DCT-CMAFFNet and YOLOV8m-mobilevit have fewer parameters compared to YOLOV8m, indicating lower model complexity. The YOLOV8-mobilevit has a longer inference time than YOLOV8 due to its network structure, but the increase in computational time is not substantial. The proposed DCT-CMAFFNet requires a longer training time compared to YOLOV8 and YOLOV8-mobilevit, reflecting its more complex dual- source network design.

V. CONCLUSION

The identification of Martian craters is an essential task in planetary science missions and research. Martian craters contain rich features and complex information, i.e., multiple types and multiple scales, and are usually modified by Martian endogenic and exogenic dynamics. On the global scale, it is difficult to identify Martian craters comprehensively and efficiently. In order to solve this problem, in this paper the DCT-CMAFFNet is proposed that exploits complementary information between different modal data, i.e., imagery and DEM for multi-type and multi-scale Martian craters identification.

First, a dual-branch architecture that combines CNN and Transformer networks is employed to capture both the local detailed features and the global context of craters. Then, to merge the morphological feature and topographical information from imagery and DEM data, we designed the CMAFF module in which each position of the deep feature maps can be weighted by using the self- attention mechanism. Through the experiments and comparison with the use of either only imagery or only DEM data, and also data-fusion networks, the following conclusions are drawn: 1) the proposed DCT-CMAFFNets can adaptively fuse imagery and DEM data with regard to the characteristics of Martian craters; 2) the proposed network makes adequate use of the morphological features and topographic information of multi-type and multi-scale craters on complex Martian surface, and provides acceptable identification of craters at low and middle latitudes. The craters recall is improved by 14.5%, 23.9% and 18.2% with respect to the use of only imagery, only DEM and data-fusion networks. As a result, 3166 new Martian impact craters larger than 1km are identified.

In future work, we will try to improve the smaller-scale (< 3km) craters identification with images at high spatial resolution, such as the Context Camera (CTX). Meanwhile, this model can be optimized to improve the identification accuracy in special terrains, e.g., basin margins and transition zones.

ACKNOWLEDGMENT

The authors would like to thank Stuart J. Robbins and A. Lagain for providing the D >1 km Martian craters databases, G. Salamunićcar for providing the automated crater catalogues.

REFERENCES

[1] D. Banfield *et al.*, "Summary of the Mars science goals, objectives, investigations, and priorities," *Bulletin of the American Astronomical Society,* vol. 53, no. 4, p. 142, 2021.

[2] J. J. Liu *et al.*, "Martian dunes indicative of wind regime shift in line with end of ice age,", *Nature,* vol. 620, no. 7973, pp. 303-309, Aug 10 2023, doi: 10.1038/s41586-023-06206-1.

[3] W. K. Hartmann and G. Neukum, "Cratering chronology and the evolution of Mars," (in English), *Space Sci Rev,* vol. 96, no. 1-4, pp. 165-194, Apr 2001, doi: 10.1023/A:1011945222010.

[4] A. Lagain, K. Servis, G. K. Benedix, C. Norman, S. Anderson, and P. A. Bland, "Model Age Derivation of Large Martian Impact Craters, Using Automatic Crater Counting Methods,", *Earth Space Sci,* vol. 8, no. 2, Feb 2021, doi: 10.1029/2020EA001598.

[5] Z. Y. Yue, K. C. Di, G. Michael, S. Gou, Y. T. Lin, and J. Z. Liu, "Martian surface dating model refinement based on Chang'E-5 updated lunar chronology function,", *Earth Planet Sc Lett,* vol. 595, Oct 1 2022, doi: 10.1016/j.epsl.2022.117765.

[6] V. Z. Sun and R. E. Milliken, "Ancient and recent clay formation on Mars as revealed from a global survey of hydrous minerals in crater central peaks,", *J Geophys Res-Planet,* vol. 120, no. 12, pp. 2293-2332, Dec 2015, doi: 10.1002/2015je004918.

[7] P. Singh, R. Sarkar, and A. Porwal, "Orbital remote sensing of impact-induced hydrothermal systems on Mars,", *Ore Geol Rev,* vol. 108, pp. 101-111, May 2019, doi: 10.1016/j.oregeorev.2017.12.024.

[8] S. Sharma *et al.*, "Diverse organic-mineral associations in Jezero crater, Mars,", *Nature,* vol. 619, no. 7971, pp. 724-732, Jul 27 2023, doi: 10.1038/s41586-023-06143-z.

[9] R. D. Capitan and M. J. Van de Wiel, "Regional morphometric and geomorphologic mapping of Martian landforms,", *Comput Geosci-Uk,* vol. 45, pp. 190-198, Aug 2012, doi: 10.1016/j.cageo.2011.11.030.

[10] K. L. Tanaka, S. J. Robbins, C. M. Fortezzo, J. A. Skinner, and T. M. Hare, "The digital global geologic map of Mars: Chronostratigraphic ages, topographic and crater morphologic characteristics, and updated resurfacing history,", *Planet Space Sci,* vol. 95, pp. 11-24, May 2014, doi: 10.1016/j.pss.2013.03.006.

[11] J. H. Roberts, "Chapter 17 - Endogenic origin of the Martian hemispheric dichotomy," in *Mars Geological Enigmas*: Elsevier, 2021, pp. 499-522.

[12] N. G. Barlow, "Crater size-frequency distributions and a revised Martian relative chronology," *Icarus,* vol. 75, no. 2, pp. 285-305, 1988.

[13] N. Barlow, "Revision of the``Catalog of Large Martian Impact Craters"," in *Sixth International Conference on Mars*, 2003, p. 3073.

[14] S. J. Robbins and B. M. Hynek, "A new global database of Mars impact craters ≥1 km: 1. Database creation, properties, and parameters," *Journal of Geophysical Research: Planets,* vol. 117, no. E5, 2012, doi: https://doi.org/10.1029/2011JE003966.

[15] A. Lagain *et al.*, "Mars Crater Database: A participative project for the classification of the morphological characteristics of large Martian craters," in *Large Meteorite Impacts and Planetary Evolution VI*, vol. 550, W. U. Reimold and C. Koeberl Eds.: Geological Society of America, 2021, p. 0.

[16] P. R. Christensen *et al.*, "The Thermal Emission Imaging System (THEMIS) for the Mars 2001 Odyssey Mission,", *Space Sci Rev,* vol. 110, no. 1-2, pp. 85-130, 2004, doi: Doi 10.1023/B:Spac.0000021008.16305.94.

[17] D. E. Smith *et al.*, "Mars Orbiter Laser Altimeter: Experiment summary after the first year of global mapping of Mars," *Journal of Geophysical Research: Planets,* vol. 106, no. E10, pp. 23689-23722, 2001, doi: 10.1029/2000JE001364.

[18] G. Salamuniccar, S. Loncaric, P. Pina, L. Bandeira, and J. Saraiva, "MA130301GT catalogue of Martian impact craters and advanced evaluation of crater detection algorithms using diverse topography and image datasets,", *Planet Space Sci,* vol. 59, no. 1, pp. 111-131, Jan 2011, doi: 10.1016/j.pss.2010.11.003.

[19] G. Salamuniccar, S. Loncaric, and E. Mazarico, "LU60645GT and MA132843GT catalogues of Lunar and Martian impact craters developed using a Crater Shape-based interpolation crater detection algorithm for topography data,", *Planet Space Sci,* vol. 60, no. 1, pp. 236-247, Jan 2012, doi: 10.1016/j.pss.2011.09.003.

[20] L. Bandeira, W. Ding, and T. F. Stepinski, "Detection of sub-kilometer craters in high resolution planetary images using shape and texture features," *Advances in Space Research,* vol. 49, no. 1, pp. 64-74, 2012.

[21] Y. R. Wang and B. Wu, "Active Machine Learning Approach for Crater Detection From Planetary Imagery and Digital Elevation Models,", *Ieee T Geosci Remote,* vol. 57, no. 8, pp. 5777-5789, Aug 2019, doi: 10.1109/Tgrs.2019.2902198.

[22] J. P. Cohen, *Automated crater detection using machine learning*. University of Massachusetts Boston, 2016.

[23] H. Wang, J. Jiang, and G. J. Zhang, "CraterIDNet: An End-to-End Fully Convolutional Neural Network for Crater Detection and Identification in Remotely Sensed Planetary Images,", *Remote Sens-Basel,* vol. 10, no. 7, Jul 2018, doi: 10.3390/rs10071067.

[24] D. M. DeLatte, S. T. Crites, N. Guttenberg, E. J. Tasker, and T. Yairi, "Segmentation convolutional neural networks for automatic crater detection on mars," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* vol. 12, no. 8, pp. 2944-2957, 2019.

[25] C. Lee, "Automated crater detection on Mars using deep learning," *Planet Space Sci,* vol. 170, pp. 16-28, 2019.

[26] C. Y. Hsu, W. W. Li, and S. Z. Wang, "Knowledge-Driven GeoAI: Integrating Spatial Knowledge into Multi-Scale Deep Learning for Mars Crater Detection,", *Remote Sens-Basel,* vol. 13, no. 11, Jun 2021, doi: 10.3390/rs13112116.

[27] S. Yang and Z. Cai, "High-resolution feature pyramid network for automatic Crater detection on Mars," *Ieee T Geosci Remote,* vol. 60, pp. 1-12, 2021.

[28] C. Y. Wang, A. Bochkovskiy, and H. Y. M. Liao, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors,", *Proc Cvpr Ieee,* pp. 7464-7475, 2023, doi: 10.1109/Cvpr52729.2023.00721.

[29] C.-Y. Hsu, W. Li, and S. Wang, "Knowledge-Driven GeoAI: Integrating Spatial Knowledge into Multi-Scale Deep Learning for Mars Crater Detection," *Remote Sens-Basel,* vol. 13, no. 11, p. 2116, 2021. [Online]. Available: https://www.mdpi.com/2072-4292/13/11/2116.

[30] M. Li, J. Liang, P. Yan, and X. Tian, "A new system for automatically generating a training dataset for Mars crater detection," in *2022 4th International Conference on Advances in Computer Technology, Information Science and Communications (CTISC)*, 2022: IEEE, pp. 1-5.

[31] X. Xie, D. Wu, M. Xie, and Z. Li, "GhostFormer: Efficiently amalgamated CNN-transformer architecture for object detection," *Pattern Recognition,* vol. 148, p. 110172, 2024.

[32] S. Mehta and M. Rastegari, "MobileViT: Light-weight, General-purpose, and Mobile-friendly Vision Transformer," doi: 10.48550/arXiv.2110.02178.

[33] A. Vaswani *et al.*, "Attention is all you need," *Advances in neural information processing systems,* vol. 30, 2017.

[34] Z. Liu *et al.*, "Swin Transformer: Hierarchical Vision Transformer using Shifted Windows,", *2021 Ieee/Cvf International Conference on Computer Vision (Iccv 2021),* pp. 9992-10002, 2021, doi: 10.1109/Iccv48922.2021.00986.

[35] M. Raghu, T. Unterthiner, S. Kornblith, C. Y. Zhang, and A. Dosovitskiy, "Do Vision Transformers See Like Convolutional Neural Networks?,", *Adv Neur In,* vol. 34, 2021. [Online]. Available: <Go to ISI>://WOS:000901616408029.

[36] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, and S. Zagoruyko, "End-to-End Object Detection with Transformers," Cham, 2020: Springer International Publishing, in Computer Vision – ECCV 2020, pp. 213-229.

[37] X. Zhu, W. Su, L. Lu, B. Li, X. Wang, and J. Dai, "Deformable DETR: Deformable Transformers for End-to-End Object Detection," doi: 10.48550/arXiv.2010.04159.

[38] Z. F. Zong, G. L. Song, and Y. Liu, "DETRs with Collaborative Hybrid Assignments Training,", *Ieee I Conf Comp Vis,* pp. 6725-6735, 2023, doi: 10.1109/Iccv51070.2023.00621.

[39] Q. C. Liu, L. Xiao, J. X. Yang, and Z. H. Wei, "CNN-Enhanced Graph Convolutional Network With Pixel- and Superpixel-Level Feature Fusion for Hyperspectral Image Classification,", *Ieee T Geosci Remote,* vol. 59, no. 10, pp. 8657-8671, Oct 2021, doi: 10.1109/Tgrs.2020.3037361.

[40] A. H. Gunatilaka and B. A. Baertlein, "Feature-level and decision-level fusion of noncoincidently sampled sensors for land mine detection,", *Ieee T Pattern Anal,* vol. 23, no. 6, pp. 577-589, Jun 2001, doi: Doi 10.1109/34.927459.

[41] S. C. Liu, H. Zhao, Q. Du, L. Bruzzone, A. Samat, and X. H. Tong, "Novel Cross-Resolution Feature-Level Fusion for Joint Classification of Multispectral and Panchromatic Remote Sensing Images,", *Ieee T Geosci Remote,* vol. 60, 2022, doi: 10.1109/Tgrs.2021.3127710.

[42] Hendrycks, Dan, and Kevin Gimpel. "Gaussian error linear units (gelus)." doi: 10.48550/arXiv. 1606.08415.

[43] F. Qingyun, H. Dapeng, and W. Zhaokui, "Cross-Modality Fusion Transformer for Multispectral Object Detection," doi: 10.48550/arXiv.2111.00273.

[44] C. S. Edwards, K. J. Nowicki, P. R. Christensen, J. Hill, N. Gorelick, and K. Murray, "Mosaicking of global planetary image datasets: 1. Techniques and data processing for Thermal Emission Imaging System (THEMIS) multi-spectral data," *Journal of Geophysical Research: Planets,* vol. 116, no. E10, 2011, doi: 10.1029/2010JE003755.

[45] R. Fergason, T. Hare, and J. Laura, "HRSC and MOLA blended digital elevation model at 200m v2. Astrogeology PDS Annex," *US Geological Survey,* vol. 400, 2018.