An Improved Double-Branch Network for Estimation of Crater Ages Based on Semisupervised Learning and Multi-Source Lunar Data

Zhonghua Hong , *Member, IEEE*, Linxuan Zhong , Xiaohua Tong , *Senior Member, IEEE*, Haiyan Pan , Ruyan Zhou , Yun Zhang , Yanling Han , Jing Wang , Shuhu Yang , and Haiyang He

Abstract—While various methods have been developed to estimate the age of impact craters, such as the crater size frequency distribution and morphology methods. Accurately and efficiently estimating the ages of lunar craters using traditional techniques is challenging due to their complex morphology and large number. As a result, the accuracy of age estimation algorithms for meteorite craters based on deep learning is restricted by factors such as a scarcity of age-labeled data and the complex morphology of these craters. To address these issues, this article presents an enhanced double-branch network for estimating crater ages via semisupervised learning and multisource lunar data. The algorithm consists of three steps: semisupervised training data augmentation, adaptive two-branch feature extraction, and a two-stage crater age classification process. The effectiveness of the improved approach was validated through ablation experiments, resulting in an overall accuracy of 83.7% on the test set of meteorite craters. This is 5.2% higher than the accuracy achieved by the previous deep learning method.

Index Terms—Crater-age classification, double-branch network, multisource lunar data, semisupervised learning.

I. INTRODUCTION

IMPACT craters dominate much of the lunar landscape and represent the most ubiquitous surface feature on the Moon. An accurate delineation of lunar crater chronology is paramount to determining the absolute age model of the lunar surface and the wider solar system. Furthermore, it holds significant implications for both lunar exploration and the prospect of future lunar resource utilization. In lunar chronology, one approach to determine the geological age of a crater is through the crater size frequency distribution (CSFD), which reveals the relationship between the number of impact craters per unit area and their respective diameters. To estimate the surface age of a given cell, the statistical frequency distribution of impact crater size is compared against the known impact crater generation function using either the cumulative distribution method or the relative distribution method [1], [2], [3]. However, the presence of secondary impact craters and degraded impact craters may lead to fluctuations in the number of impact craters and biased measurement results. The degradation state of the crater rim can be used as an alternative method for estimating the age of meteorite craters. Pohn and Offield [4] classified lunar craters into three categories based on the initial rim top's shape and established corresponding freshness categories for each category. Trang et al. [5] developed regression equations between the three levels of crater freshness and absolute model age. He calibrated the classification of the three freshness levels defined by Pohn and Offield [4] to the absolute model age and derived a transformation equation from crater freshness to the absolute model age. Although the impact crater morphology method has its merits, it is influenced by factors such as the uniformity of impact crater size distribution, the geological process of the lunar surface, and the solar incidence angle of the research base map, which results in considerable labor. Because of the complexity of the formation mechanisms of meteorite craters and the difficulty of obtaining lunar samples, traditional methods or approaches based on a single type of data to predict the age of these craters may not yield accurate and timely results.

In recent years, deep learning techniques have been widely applied in remote sensing image processing [6], [7], [8]. Convolutional neural networks (CNN) have gained widespread recognition as a preferred choice in deep learning networks for a range of applications [9], [10], [11]. The utilization of ResNet50 by Firat and Hanbay [12] has proved effective in the field of multispectral image classification. Similarly, Ye et al. [13] introduced a remote sensing image classification methodology that capitalizes on the lightweight VGG-16 model. In a related study, Li et al. [14] presented a novel approach to enhance multispectral remote sensing images affected by cloud pollution, making use of the MobileNet architecture. In the processing of sequence-based remote sensing data, recurrent neural networks are frequently utilized to extract time series

features [15]. In addition, graph convolutional networks can capture the spatial relationship among image pixels, thus the precision of node and graph classification can be improved [16]. Fully connected networks consist entirely of convolutional layers, which are able to preserve enough spatial information and are suitable for pixel-level classification tasks [17], [18]. Multimodal deep learning methods can effectively solve the problem of insufficient information in single-modal data, by utilizing multiple remote sensing data sources such as optical, spectral, and radar data, providing rich information pertaining to ground objects [15]. Traditional deep-learning methods require a substantial number of labeled training samples to extract sample features. However, transfer learning has emerged as a promising technique that reduces the cost of model training by leveraging a pretrained model to solve a different but related problem. Meanwhile, semisupervised learning has also gained attention as a viable approach for pattern recognition tasks, as it utilizes a significant amount of unlabeled data in conjunction with labeled data to improve overall accuracy (OA). Transfer learning and semisupervised learning have been successfully applied in the absence of labeled samples [19], [20], [21] and achieved high accuracy in predicting the age of lunar craters. Yang et al. [22] proposed an innovative method that combines transfer learning and deep learning for lunar crater age classification. First, they effectively transformed the age estimation problem into a classification problem. Second, a dual-channel classification network was constructed, in which one channel extracted features from the lunar digital orthophoto map (DOM) data and the other channel extracted features from other attribute data. The extracted features from both channels were then fused for classification. Third, to address the issue of limited labeled samples, the authors adopted a transfer learning strategy by utilizing the pretrained weights obtained from ImageNet as the initial weights to prevent random initialization. Moreover, Mean-Teacher semisupervised learning was employed to utilize a large amount of unlabeled data for preventing overfitting during training. The resulting classification model achieved an average accuracy of 85.44%. To the best of the authors' knowledge, this is the first and only study that has applied deep learning techniques to the task of predicting the age of meteorite craters. The achieved accuracy is considerable, marking a significant milestone in this research field. Despite the notable progress achieved, there remain several outstanding issues that warrant further investigation.

First, Mean-Teacher semisupervised approach, which introduces varying perturbations to identical unlabeled data during training to attain consistent predictions on the same sample with distinct perturbations, utilizing unlabeled data to curb overfitting during training and boost accuracy. With the advent of additional semisupervised strategies in recent times, the utilization of a significant quantity of unlabeled crater data has become a crucial focal point for the advancement of crater age estimation.

Second, as human exploration of the Moon continues, the volume of available lunar data is rapidly increasing. For example, the Chang'e-1 (CE-1) and Chang'e-2 (CE-2) mission, which launched in 2007 and 2010, utilized images from their CCD cameras to produce high-resolution DOM of the full Moon at 500, 120, 50, 20, and 7 m scales [23], [24]. The clementine probe's ultraviolet-visible (UVVIS) camera, equipped with three spectral filters, can be used to generate a color-scale mineral map of the Moon [25]. The U.S. Geological Survey (USGS) combined data from six Apollo missions to create a comprehensive geological map of the Moon [26]. These datasets offer significant insights into the lunar surface, including the ability to estimate the age of lunar craters, yet few studies have fully exploited their potential. How to effectively use the multisource data to improve the accuracy of lunar crater's age prediction still needs to be further investigated.

Therefore, this study aims to design a semisupervised learning method that can leverage a large amount of unlabeled data and a meteorite crater age classification model that can integrate multiple sources of data. The main contributions of this work are as follows.

1) We present an adaptive dual-channel classification network to extract features based on multisource data. In the feature layer, different from existing studies that fuse the features directly, an adaptive strategy is adopted. At the decision level, the crater classification process is initially partitioned into two stages, and then the Dempster-Shafer (D-S) theory is employed to combine the classification results of different networks.

2) This study first integrates multiple sources data, such as DOMs, multispectral images, as well as other morphological and geological feature data, for the purpose of estimating crater ages.

3) A labeled data augmentation algorithm based on a semisupervised learning strategy is presented. Unlabeled data with high confidence is selected to incorporate pseudolabels for training purposes. The confidence of the unlabeled data is determined by measuring the similarity between the unlabeled eigenvector and the eigenvector of the labeled data.

The rest of this article is organized as follows. In the Section II "Data," we present an overview of the experimental data that we utilized in our study. Section III "Methodology" outlines the method framework. In Section IV "Experiment and Analysis," we present the experimental framework. In Section V "Discussion," we compare and analyze the prediction results of individual samples. In Section VI "Conclusion," we summarize the proposed method and the key findings of our study and outlined potential avenues for future research.

II. DATA

This study utilized data from multiple sources, including DOMs generated by the Chang'E1 and Chang'E2 cameras, as well as five-band spectral image data obtained from the Clementine UVVIS camera. In addition, we obtained 1675 labeled lunar crater data and 6953 unlabeled lunar crater data from the Lunar and Planetary Institute's (LPI) 2015 data release [27]. The lunar geological data employed in this study were sourced from the USGS 1:5000000-scale geological map of the Moon published in 2013. For a more comprehensive description of the data sources used in this research, please refer to Table I.

TABLE I
DETAILED INFORMATION OF THE DATA

| Data sources	Data information	Download the website	function
CE-1	120m-DOM	http://moon.bao.ac.cn	DOM of craters with radius >7 km
CE-2	7m-DOM	http://moon.bao.ac.cn	DOM of craters with radius ≤7 km
Clementine	200m MS image	https://astrogeology.usgs.gov/	Crater MS information
LPI	Crater information	https:// www.lpi.usra.edu/lunar/surface/	Obtaining location, extent, and other morphological information of craters
Geological map	1:5M geologic map	https://astrogeology.usgs.gov/	Obtaining geological composition information of craters
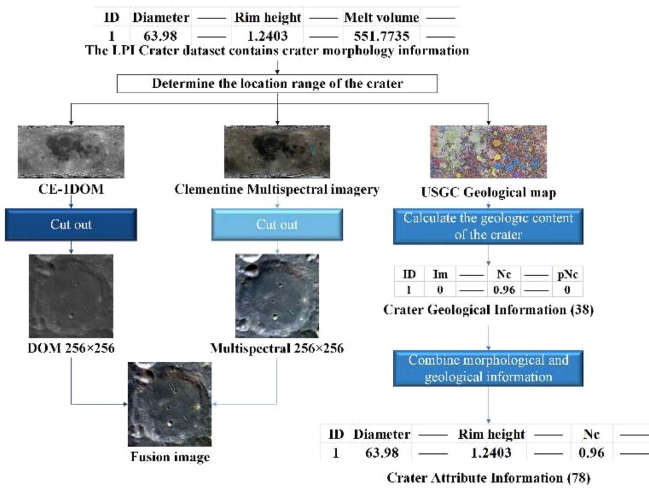
Fig. 1. Crater data processing process.

The data processing procedure is illustrated in Fig. 1. The longitude, latitude, and radius information of the craters in the crater database released by LPI is used to obtain the labeled and unlabeled crater locations and their ranges. Subsequently, the DOM and the multispectral image for each impact crater are separately extracted from the full Moon DOM and full Moon multispectral image based on the boundary range of each impact crater. The crater was fused with the multispectral images. Additionally, 40 crater-shape information attributes, including diameter, rim-to-interior depth, and interior volume, that cannot be extracted from the DOM were obtained from LPI. Furthermore, the content ratio of each lunar crater was derived from 1:500000 Lunar Geological Restoration map published by the USGS for 38 geological objects.

III. METHODOLOGY

Fig. 2 depicts the CNN-based algorithm for estimating the age of lunar craters using semisupervised learning. The algorithm consists of three parts: data augmentation via pseudolabeling of unlabeled data using nearest neighbor method in semisupervised learning, an adaptive dual-channel feature fusion module, and a two-stage crater classification algorithm.
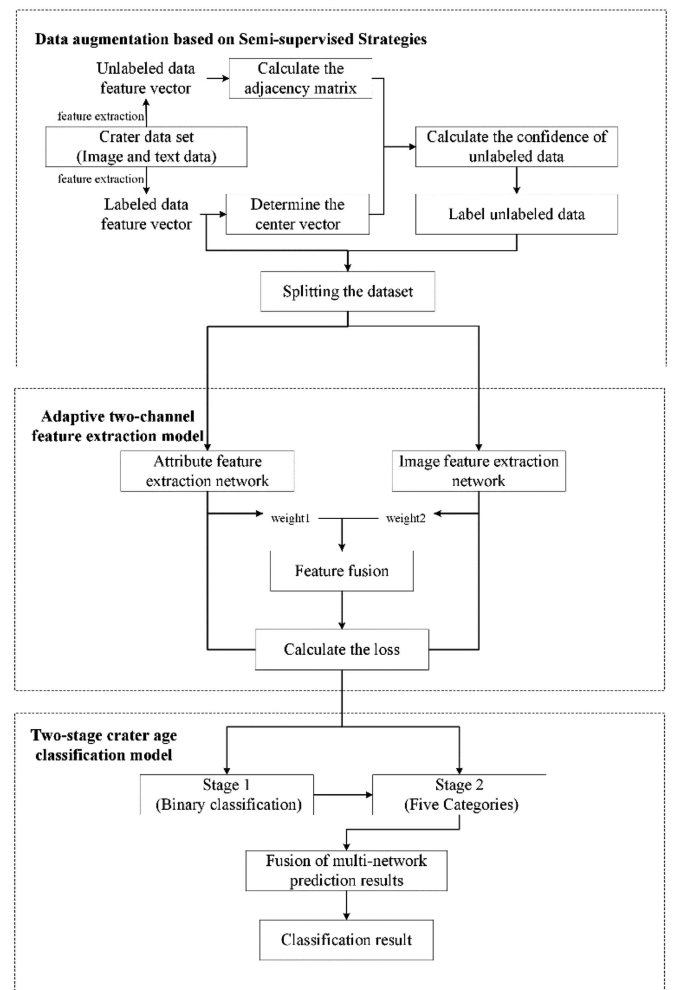

Fig. 2. Network structure diagram.

A. Data Augmentation Based on Semisupervised Strategies

The labeled data augmentation method, based on a semisupervised learning strategy, involves incorporating pseudolabeled data during training [28], [29], [30], [31], [32]. Craters sharing similar morphology and geology are typically indicative of the
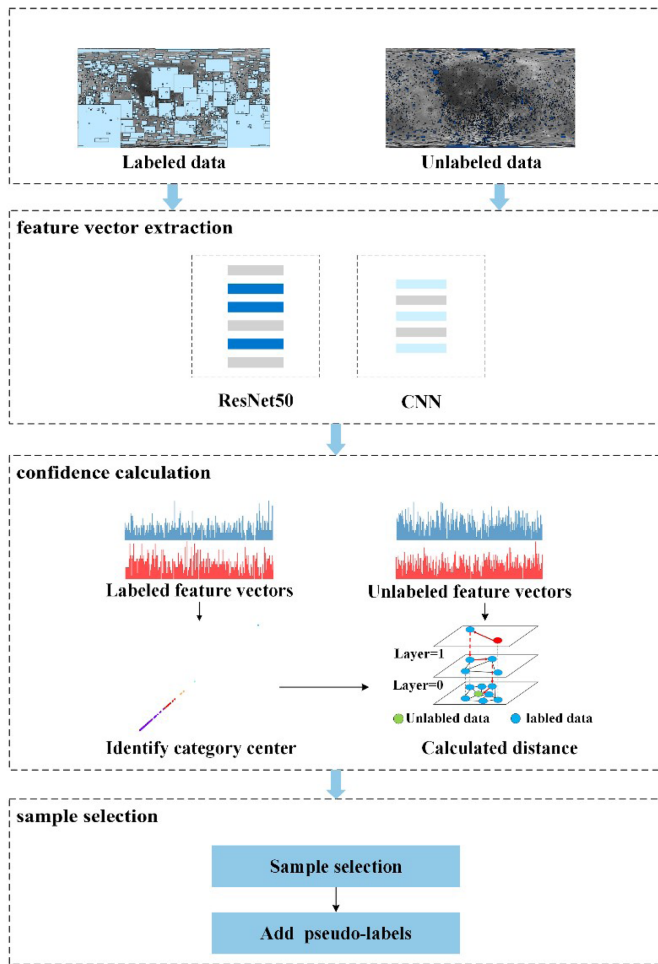
Fig. 3. Data expansion strategy.

same period. To assess the similarity in morphology and geological material content between unlabeled and labeled data, we define image feature confidence and attribute feature confidence for the unlabeled crater data. We select unlabeled data with high confidence and assign them pseudolabels. The similarity between samples is computed using the Hierarchical Navigable Small World Networks algorithm [33]. This algorithm constructs a hierarchical graph using skip lists composed of linked lists. The initial layer is created using a random subset of nodes, and subsequent layers establish connections with the nodes from the previous layer by incorporating a subset of nodes from the remaining pool until the final layer encompasses all nodes.

A total of 1337 lunar impact craters, including those with age labels, and 6953 craters without age labels, were selected from the Lunar Planetary Institute (LPI) crater database. Fig. 3 depicts the process of the label data augmentation technique. The proposed algorithm utilizes both lunar impact crater image data and attribute data. The image data consists of fused DOMs and multispectral images. The crater attribute data includes general attributes and geological information. The ResNet50 and the CNN network proposed by Zhang et al. [34] are utilized to extract image feature vectors and attribute feature vectors from image

data and attribute data, respectively. Feature centroid vectors for both image and attribute features are obtained based on labeled crater feature vectors. The feature vectors of the unlabeled data are then used to construct nearest neighbor graphs for both image and attribute features. In these graphs, the vertices represent the feature vectors of unlabeled samples, and the edges represent the distances between these vertices. The distances between the unlabeled image feature vectors and the image feature centroid vectors are computed to obtain image feature confidence scores. Similarly, the distances between the unlabeled attribute feature vectors and the attribute feature centroid vectors are computed to obtain attribute feature confidence scores. The fusion feature confidence score is obtained by combining the image feature confidence score and the attribute feature confidence score. Unlabeled samples with high fusion feature confidence scores are selected and assigned labels. To improve the reliability of the sample selection process, a dual-constraint approach is employed, giving priority to samples with high confidence scores for both image and attribute features.

B. Feature Extraction Network

This study employed two types of data: lunar crater image data and lunar crater attribute data. The CNN model proposed by Zhao et al. [36] was used to extract lunar crater attribute features. By leveraging end-to-end learning, CNN can automatically learn optimal feature representations, which is advantageous for handling complex textual data. In addition to the standard CNN architecture, dropout layers were incorporated to mitigate overfitting by temporarily discarding certain neural network units during training.

Given the diverse and intricate nature of lunar crater images, we aimed to capture more comprehensive feature representations by leveraging the strengths of different networks. Therefore, we selected ResNet50, MobileNet, and VGG16 to extract lunar crater image features. ResNet50 is a deep residual network that effectively addresses the problems of gradient vanishing and information loss, enabling better capture of fine-grained image details. MobileNet is a lightweight network that primarily focuses on learning low-level and fine-grained image features, demonstrating good perception capabilities for edges, textures, and colors. VGG16, on the other hand, primarily learns high-level and abstract image features, allowing for the capture of more complex image characteristics such as shape, texture, parts, and contextual information. By leveraging networks with distinct characteristics, our approach aims to capture features at different levels and perspectives, thereby enhancing the accuracy and robustness of classification. The network structures of some of the networks are presented in Fig. 4.

C. Adaptive Two-Channel Feature Extraction Model

The current study presents a novel self-adaptive dual-channel feature fusion model, which is developed to extract distinctive features from varied types of lunar crater data and subsequently integrate them. The first channel utilizes multiple CNNs to extract features from lunar crater image data, while the second channel uses the CNN to extract features from the lunar crater
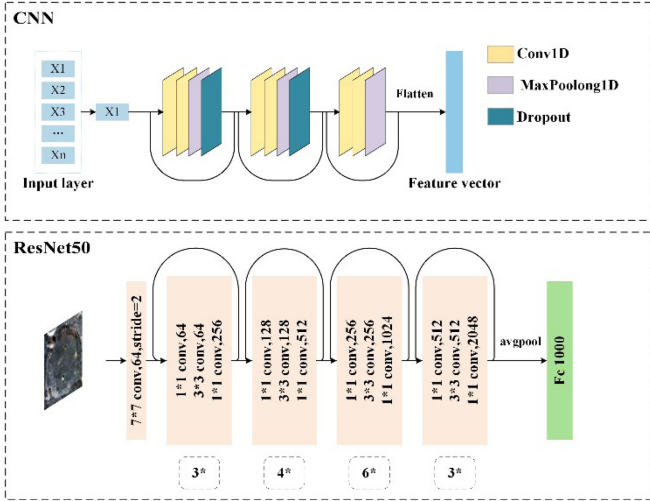
Fig. 4. Network structure diagram.

Fig. 5. Adaptive feature processing module.

attribute data. To improve the efficacy of different data in different classification tasks and extract more informative features for lunar crater age classification, Huang et al. [35] and Zhao et al. [36] incorporate an adaptive feature fusion weight for each of the two feature extraction channels,. The coefficients for adaptive feature fusion are adjusted based on the impact of the features from each channel on the network. Initially, the adaptive feature fusion weight for each channel is set to 1. The features from the two channels are then merged based on their respective adaptive feature fusion weights. The implementation of the adaptive dual-channel feature extraction module is illustrated in Fig. 5.

The adaptive feature fusion weights determine the relative importance of the features extracted from each channel and are updated during training to optimize the network's performance. Specifically, the adaptive feature fusion weight for the image channel is denoted by α_2 and the adaptive feature fusion weight for the attribute channel is denoted by α_1. The formula of feature fusion is as follows:

$$F_{\text{ff}} = \alpha_1 \times F_{\text{attr}} + \alpha_2 \times F_{\text{img}}$$

$$a_i = \frac{e^{w_i}}{\Sigma_j e^{w_j}} \ (i = 1, 2; j = 1, 2) \tag{1}$$

where F_{img} and F_{attr} are the feature vectors extracted from the image and attribute channels, respectively. The values of α_1 and α_2 are updated during training using an optimizer, such as stochastic gradient descent, to minimize the network's loss function. The weights are updated based on the gradients of the loss function with respect to the weights, and the learning rate determines the step size of the weight updates. By adaptively adjusting the feature fusion weights, the network can effectively utilize the information from both types of crater data to improve classification performance.

D. Two-Stage Crater Age Classification Model

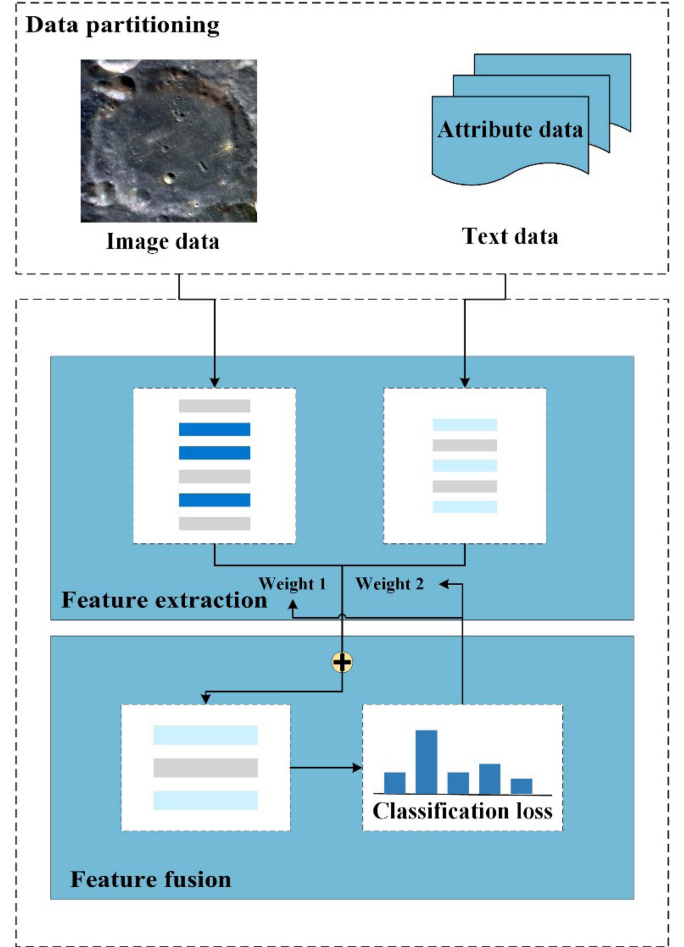The distinct features of pre-Imbrain and post-Imbrain crater image data are identified in this study. Furthermore, we observe that in the adaptive two-channel feature extraction stage, the image features hold greater weight when the crater category is bifurcated into pre-Imbrain and post-Imbrain periods. However, when the crater age is divided into five categories, the crater attribute features occupy greater feature weight. To optimize the performance of multisource data in the crater classification task, we bifurcate the crater age classification process into two stages. In the first stage, we initially classify the craters into two categories, and in the second stage, based on the results of the first stage, we further classify the craters into five categories.

Subsequently, a decision-level fusion was performed based on the confidence scores obtained from the predictions. The utilization of multiple networks for image feature extraction aimed to reduce overall bias and enhance the robustness of the classification process. In Fig. 6, the three distinct classification networks are represented by a, b, and c. The classification confidence for each class of the crater is calculated for every model (A, B, C, D, and E are the five classes of craters). The final prediction results are then fused using the D-S theory [37], [38].

Following D-S theory, the predicted category of a given crater is weighted by probability, with the category having the highest calculated value determined as the age category of the crater
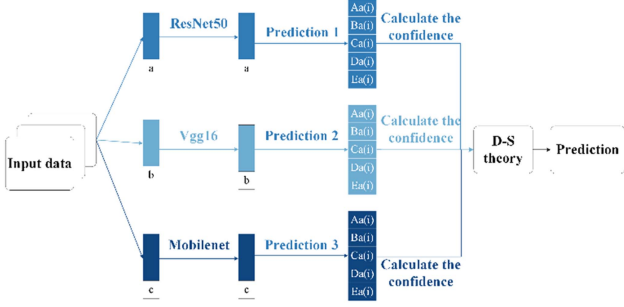
Fig. 6. Decision fusion model.

(where A, B, C, D, and E represent five categories of age periods for lunar craters, a, b, and c represent different classification networks, and i denotes the ith target). The fusion formula is expressed as follows: (2) shown at the bottom of this page.

IV. EXPERIMENT AND ANALYSIS

A. Experimental Environment and Model Training

All the experiments were performed on Intel(R) Core (TM) i9-10920X CPU at 3.50 GHz with 64 GB RAM and NVIDIA GeForce RTX2080Ti with 11 GB GDDR6X memory, CUDA 10.0 edition, Python 3.6, torch 1.2.

The training of the impact crater data is conducted in two stages, namely the freezing stage and the unfreezing stage. During the freezing stage, the feature extraction network is fine-tuned while preventing the weights from being destroyed in the initial phase of training to avoid the randomness of the backbone network's feature extraction effect. At the commencement of the training process, the learning rate is initialized to 0.001, while the weight decay coefficient and batch size are set to 0.0005 and 32, respectively, for a duration of 30 epochs. Upon entering the unfreezing stage, the learning rate is subsequently reduced to 0.0001, while the weight decay coefficient remains at 0.0005 and the batch size is reduced to 16, for an additional 30 epochs.

In this study, five experiments were conducted, where the available data was randomly divided into training, test, and validation sets in an 8:1:1 ratio for each experimental group. The training set data was utilized for label data augmentation. The distribution of the training, test, and validation sets for each experiment is presented in Table II.

The impact crater classification results were analyzed with OA. The confusion matrix is used to show the classification of different age periods. OA is defined as follows:

$$\text{OA} = \frac{1}{N} \sum_{i=1}^{C} x_{ii}. \tag{3}$$

TABLE II
EXPERIMENT DATA DISTRIBUTION

Crater-period	Training set	Test set	Validation set
Pre-Nectarian	353	45	44
Nectarian	588	74	74
Imbrian	228	26	26
Eratosthenian	104	14	13
Copernican	64	8	8

B. Result Analysis

This section presents a summary of the experimental methodology and findings presented in this article, which were carried out to assess the effectiveness of the improved techniques described in the previous section. The experiments are primarily divided into the following aspects.

1) Comparison of prediction results of different networks including image feature extraction network and attribute feature extraction network.
2) Multispectral image fusion and label data augmentation.
3) The improved adaptive dual-channel feature extraction and two-stage age classification network.
4) Our method's predictive performance for complex lunar craters.

To evaluate the effectiveness of the proposed method, the proposed method is compared with the state-of-the-art networks. Considering that there is only one network customized for crater's age prediction, three state-of-the-art commonly used networks and two classic machine learning algorithms are selected for comparison purpose, which are MobileNet, Vgg16, and ResNet50, support vector machine (SVM), and random forest.

The results presented in the table indicate that when utilizing fused image data of impact craters as input, the initial stage of binary classification based on crater age demonstrates comparable accuracy among the MobileNet, VGG16, and ResNet50 networks. However, when the craters are divided into five age categories, the ResNet50 network exhibits notably higher OA compared to the other networks. On the other hand, when attribute data of the impact craters are used as input, the initial stage of classification for two classes based on attribute data shows lower accuracy compared to the classification achieved using image data. Furthermore, the OA of the CNN surpasses that

$$
\begin{array}{c}
A_{a(i)} + \left(1 - A_{a(i)}\right) A_{b(i)} + \left(1 - \left(A_{a(i)} + \left(1 - A_{a(i)}\right) A_{b(i)}\right)\right) A_{c(i)} = A_{(i)} \\
B_{a(i)} + \left(1 - B_{a(i)}\right) B_{b(i)} + \left(1 - \left(B_{a(i)} + \left(1 - B_{a(i)}\right) B_{b(i)}\right)\right) B_{c(i)} = B_{(i)} \\
C_{a(i)} + \left(1 - C_{a(i)}\right) C_{b(i)} + \left(1 - \left(C_{a(i)} + \left(1 - C_{a(i)}\right) C_{b(i)}\right)\right) C_{c(i)} = C_{(i)} \\
D_{a(i)} + \left(1 - D_{a(i)}\right) D_{b(i)} + \left(1 - \left(D_{a(i)} + \left(1 - D_{a(i)}\right) D_{b(i)}\right)\right) D_{c(i)} = D_{(i)} \\
E_{a(i)} + \left(1 - E_{a(i)}\right) E_{b(i)} + \left(1 - \left(E_{a(i)} + \left(1 - E_{a(i)}\right) E_{b(i)}\right)\right) E_{c(i)} = E_{(i)}
\end{array}. \tag{2}
$$

TABLE III
RESULTS OF EACH NETWORK EXPERIMENT

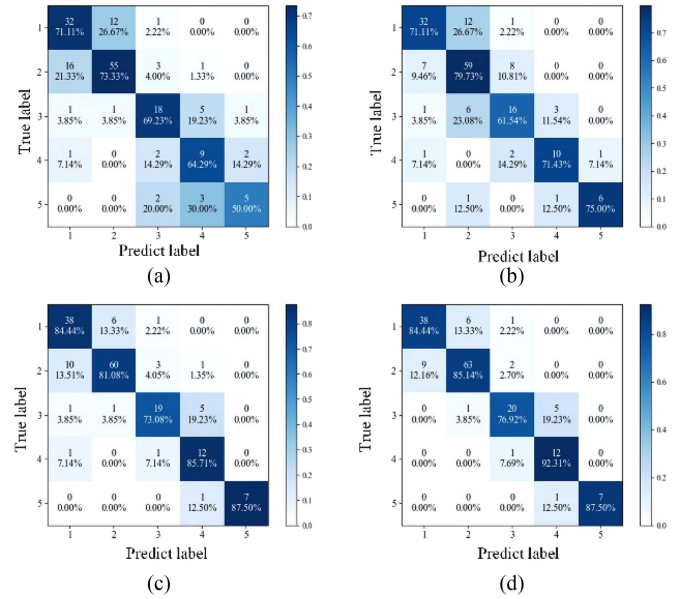| Network	Two-class accuracy	Overall accuracy	Time(S)
MobileNet	93.8%	62.7%	507
VGG16	94.1%	67.1%	979
ResNet50	95.2%	71.2%	667
SVM	90%	72%	6
Random forest	89%	71%	6
CNN	90.5%	73.6%	14
Our approach	95.6%	81.4%	692

Fig. 7. Confusion matrix for each experiment. (a) ResNet50 image data age classification. (b) CNN attribute data crater age classification. (c) Improving network crater age classification. (d) Tag augmented crater age classification.

of the SVM and Random Forest methods when the craters are divided into five age categories. This study proposes a two-stage adaptive dual-channel feature fusion method, achieving an OA of 81.4%, surpassing other networks. Additionally, the study compared the time cost of different methods. The results, as depicted in Table III, reveal that machine learning methods exhibit significantly higher time efficiency compared to deep learning methods. Among the five deep learning methods, CNN showcases the lowest time cost, requiring only 14 s. The time costs of MobileNet, ResNet50, and the proposed method are similar, all of which are faster than VGG16.

Table IV presents the results of five experiments: in experiment A, we exclusively utilized DOM data and employed the ResNet50 network for predicting the age of impact craters. Experiment B is based on Experiment A and adds the fusion of multispectral images. Experiment C employs CNN for age classification using the lunar crater attribute data. Experiment D refers to the semisupervised double-channel crater age classification (SSDC-LCC) method proposed by Yang et al. [22]. Experiment E is an adaptive double-channel lunar crater age classification method based on multisource data fusion, which is an extension of Experiments B and C. Experiment F is based on Experiment E and uses the idea of semisupervised learning by assigning labels to unlabeled data using labeled data and incorporating them into training.

The experiments were conducted in five trials for each experiment, and in each trial, the labeled experimental data were randomly divided into training set, testing set, and validation set with a ratio of 8:1:1. The average testing accuracy was calculated from the five trials for each experiment. The results of each experiment are presented in Table IV. Experiment A uses lunar crater DOM data and achieves an average testing accuracy of 67.3%. After fusing multispectral data in Experiment B, the accuracy improves to 71.2%. Experiment C uses only 78 lunar crater attributes and achieves an average classification accuracy of 73.6%. Moreover, Experiment E proposes an adaptive dual channel classification network based on multisource data fusion,

which achieves an average testing accuracy of 81.4%, a 5.2% improvement over the SSDC crater age classification method in Experiment D. To further improve model performance, Experiment F incorporates a semisupervised learning-based label expansion module into Experiment E. The average testing accuracy is improved by 1.8% to 83.2%. The effectiveness of the proposed method is demonstrated by the experimental results, and the confusion matrices for each experiment are shown in Fig. 7.

As depicted in Fig. 7, when using crater image data, the classification accuracy for distinguishing craters before and after the Imbrain period is notably high, with an average test accuracy of 95.2%. Conversely, the average test accuracy for age classification of craters, before or after the Imbrain period, using crater attribute features is merely 90.5%. Nonetheless, the average accuracy of the crater attributes was 82.7% for the Pre-Nectarian or Nectarian period, 82% for the post-Imbrain period, and only 75.4% for the image data. The results of the experiments agree with the proposed adaptive feature weight changes in different classification stages discussed in the method section. This indicates the significance of lunar crater image data in distinguishing the pre-Imbrian and post-Imbrian periods in the first stage, as well as the importance of lunar crater attribute data in the specific classification of the second stage. The results presented in this article add to the body of evidence supporting the effectiveness of the two-stage classification method.

In Fig. 8, we demonstrate the fusion effect of the lunar crater DOM and multispectral data. The left side is DOM crater, the middle is multispectral image, and the right side is the fusion image of orthophoto and multispectral image.

Fig. 9 illustrates the test results obtained after the label expansion. The results indicate that the impact of the label

TABLE IV
COMPARISON OF EXPERIMENTAL RESULTS

Exp	DOM	Fused image	Attribute data	Extended data	SSDC-LCC	Improved network	OA
A	☑	☐	☐	☐	☐	☐	67.3%
B	☑	☑	☐	☐	☐	☐	71.2%
C	☐	☐	☑	☐	☐	☐	73.6%
D	☑	☐	☑	☐	☑	☐	78.5%
E	☑	☑	☑	☐	☐	☑	81.4%
F	☑	☑	☑	☑	☐	☑	83.7%

Fig. 9. Tag expansion experiment.

Fig. 8. Image data of craters. (a) Pre-Nectarian period crater. (b) Nectarian period crater. (c) Imbrian period crater. (d) Eratosonian period crater. (e) Copernican period crater.
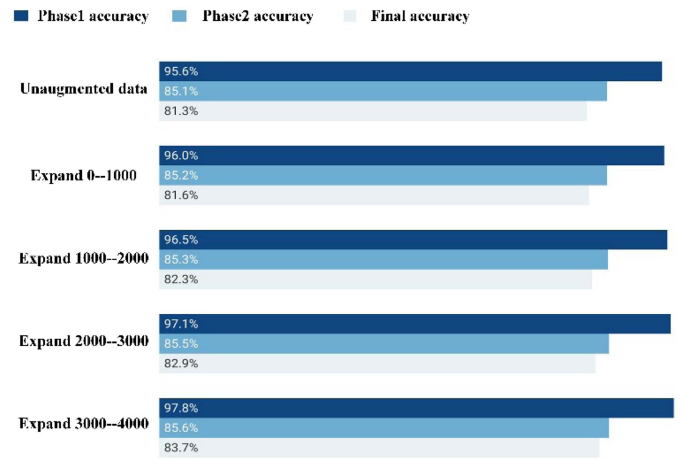
expansion on the experimental accuracy is mainly observed in the first stage, where the accuracy improved from 95.6% before expansion to 97.8% after expansion. Moreover, the final prediction accuracy increased from 81.3% before expansion to 83.7%. These experiments demonstrate the effectiveness of the semisupervised label expansion algorithm in improving the age classification accuracy, especially in the first stage as the number of labels increased. The results also highlight the efficient utilization of the unlabeled data in the process.

Complex craters exhibit intricate features, such as larger craters containing multiple smaller craters or adjacent craters. Accurately determining the age of these complex craters poses a significant challenge for deep learning methods. To assess the performance of the proposed network in estimating the age of complex craters, a carefully selected set of 196 complex craters was manually chosen. Fig. 10 presents the confusion matrix depicting the results. The findings demonstrate that when initially classifying the craters into pre-Imbrian and post-Imbrian periods during the first stage, the classification accuracy for complex craters reached 94.8%. When the craters were further divided into five categories, the OA achieved was 80.5%.

Notably, the accuracy of age classification for complex craters was primarily affected when initially dividing the craters into two categories, resulting in a 2.95% decrease in accuracy compared to the randomly selected test set. This decline in accuracy
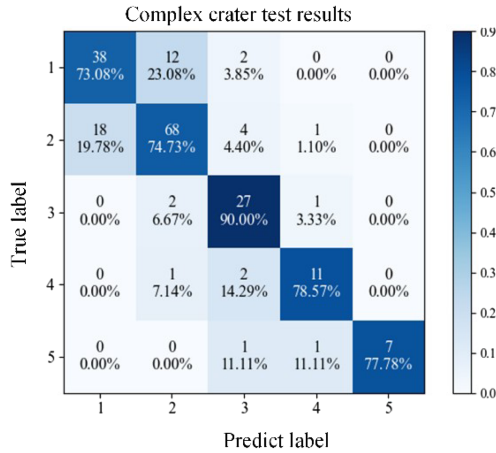
Fig. 10.   Complex crater test set confusion matrix.

TABLE V
LANDING AREA PREDICTED IMPACT CRATERS

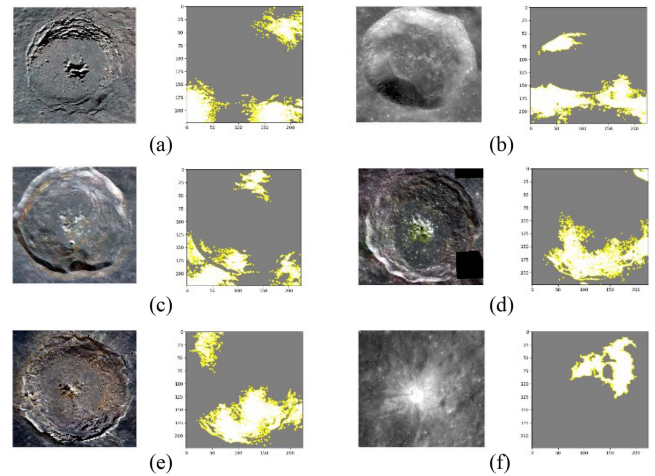| Crater	Detector	Existing results	Predicted results
Pythagoras	CE-5	Eratosthenian	Eratosthenian
Sharp B	CE-5	Copernican	Copernican
Harpalus	CE-5	Copernican	Copernican
Finsen	CE-4	Post-Imbrain	Copernican
Tycho	Apollo17	Copernican	Copernican
North Ray	Apollo17	Copernican	Copernican

Fig. 11.   LIME interpreter results. (a) Pythagoras. (b) Sharp B. (c) Harpalus. (d) Finsen. (e) Tycho. (f) NorthRay.

can be mainly attributed to the predominant influence of image features in the first stage of classification. Despite the highly challenging nature of predicting the age of complex craters, our approach has achieved satisfactory testing accuracy, further attesting to the robustness of our method.

V. DISCUSSION

The Chang'e-4 mission, comprising of the lander and rover, successfully landed within the Vonkam crater situated in the South Pole-Aitken, Basin. Finsen crater, located around 135 km northeast of the CE-4 landing site at coordinates 42.29° S, 177.72° W, and with a radius of 73 km, is a well-preserved complex that possesses a conspicuous central peak. According to Gou et al. [39], Finsen Crater was identified as an Imbrain crater through the crater Size-Frequency Distribution method. These findings agree with the results obtained by Wang and Wu [40].

Chang'e-5 successfully landed in the northeastern region of the Moon's Ocean of Storms at coordinates 43.06°N and 51.92°W. This region is mainly covered by Eratosonian basalts, and the lunar soil is predominantly formed by local basalts. Utilizing high-resolution lunar surface images acquired through remote sensing exploration, the planetary science team at Shandong University [41] identified small impact craters distributed on the bottom of each impact crater, sputtering blanket, and impact melting pool. The CSFD method was used to estimate the ages of Pythagoras, Sharp B, and Harpalus. Pythagoras is an Eratosonian impact crater, while Sharp B and Harpalus are relatively young Copernican impact craters.

Table V presents the results of predicting the age of Finsen crater using a deep learning method, indicating that it is a Copernican period crater, consistent with the LPI results but different from the previous analysis due to Finsen being a complex crater comprising multiple craters from different periods. Additionally, the results for Pythagoras, Sharp B, and Harpalus are consistent with Hou et al.'s [41] predictions, further validating the

reliability of the deep learning method for predicting crater ages.

Moreover, we utilized the LIME (Local Interpretable Model-Agnostic Explanations) image classification interpreter to scrutinize the lunar craters listed in Table V. The LIME interpreter annotated the parts of the image that contributed the most to the prediction, as shown in the right-hand side of Fig. 11. We annotated the areas of the classification outcomes that exhibited the highest benefit, as illustrated in Fig. 11. Our investigation disclosed that the annotated areas primarily appeared on the periphery of the craters, which is in accordance with our hypothesis of utilizing characteristics such as the extent of edge decay and brightness to differentiate the age of the craters. These discoveries provide a roadmap for future research that aims to improve the feature representation of the crater edges.

This article presents the prediction of the age of the crater situated in the vicinity of the potential landing zone for Artemis 3 mission, and the results of the predictions are presented in Table VI, which can be used as an additional factor for consideration in the selection of the landing site.

TABLE VI
ARTEMIS 3 CANDIDATE LANDING ZONE NEAR CRATER AGE

Crater name	Absolute age period	Age
Faustini	3.92–3.85 Ga	Nectarian System
De Gerlache	3.92–3.85 Ga	Nectarian System
Haworth	>3.92 Ga	pre-Nectarian System

VI. CONCLUSION

Accurately estimating the age of lunar craters is crucial for exploring lunar geological age. However, the limited availability of age-labeled crater data, the variable shape, and wide distribution of craters have resulted in a dearth of studies using deep learning methods for crater age estimation. To address these challenges, we propose an improved multibranch network model to estimate the age of lunar craters. To augment the training set, we propose a semisupervised learning strategy that leverages unlabeled data. Moreover, we integrate multisource data, including lunar DOM data, multispectral data, morphological data, and geological data, to extract the features of craters. The DOM data and multispectral data are fused using an image fusion algorithm. A two-branch network extracts the features of the crater data from different sources, and feature fusion is carried out based on the effects of the features on the network. A series of ablation and comparison experiments were performed to assess the proposed method. The experimental results indicate that the enhanced method leads to a significant improvement in the accuracy of crater age estimation. Specifically, we predict the age of craters near the landing areas of Chang'e-5, Chang'e-4, and Apollo 17, and our predictions are highly consistent with existing research results. Our conclusions are as follows.

1) Fusion of multispectral data with lunar DOM data improves the accuracy of the test set by 3.9% compared to using DOM data only, indicating the positive impact of multispectral data on lunar crater age estimation.

2) Significant improvement in accuracy was achieved through the semisupervised learning approach, which involves using unlabeled data to expand the training data. This suggests that the proposed algorithm is effective in enhancing the accuracy of the model.

3) The proposed model showcases accuracy comparable to state-of-the-art methods in experiments of lunar crater age estimation, thereby validating the effectiveness of deep learning techniques in this field. This finding highlights the potential of deep learning to contribute to improved accuracy and efficiency in the analysis of lunar data.

With the continuous advancement of lunar exploration technologies, increased deployment of lunar exploration satellites is anticipated in the future. The National Aeronautics and Space Administration of the United States has slated the Artemis 3 manned lunar mission for 2024. Artemis 3 will incorporate cutting-edge scientific instruments and sensors, enabling the acquisition of high-quality data encompassing DOM data, spectral data, and geological data. These diverse data sources will yield abundant information crucial for estimating crater ages. Consequently, multimodal deep learning emerges as a pivotal avenue for future development in this field.

REFERENCES

[1] H. Hiesinger, J. W. Head, U. Wolf, R. Jaumann, and G. Neukum, "Ages and stratigraphy of lunar mare basalts in Mare Frigoris and other nearside Maria based on crater size-frequency distribution measurements," *J. Geophys. Res.*, vol. 115, no. E3, Mar. 2010, doi: 10.1029/2009je003380.

[2] G. Neukum, B. A. Ivanov, and W. K. Hartmann, "Cratering records in the inner solar system in relation to the lunar reference system," *Space Sci. Rev.*, vol. 96, no. 1–4, pp. 55–86, Apr. 2001, doi: 10.1023/a:1011989004263.

[3] C. Riedel, G. Michael, T. Kneissl, C. Orgel, H. Hiesinger, and C. H. van der Bogert, "A new tool to account for crater obliteration effects in crater size-frequency distribution measurements," *Earth Space Sci.*, vol. 5, no. 6, pp. 258–267, Jun. 2018, doi: 10.1002/2018ea000383.

[4] R. Pohn and T. Offield, "Lunar crater morphology and relative age determination of lunar geologic units," *Open-File Rep.-USA Geol. Surv.*, pp. 153–162, 1970, doi: 10.3133/ofr69209.

[5] D. Trang, J. J. Gillis-Davis, and J. M. Boyce, "Absolute model ages from lunar crater morphology," *J. Geophys. Res., Planets*, vol. 120, no. 4, pp. 725–738, Apr. 2015, doi: 10.1002/2014je004639.

[6] S. Uhlmann and S. Kiranyaz, "Integrating color features in polarimetric SAR image classification," *IEEE Trans Geosci. Remote Sens.*, vol. 52, no. 4, pp. 2197–2216, Apr. 2014, doi: 10.1109/TGRS.2013.2258675.

[7] Z. Zhang, H. Wang, F. Xu, and Y.-Q. Jin, "Complex-valued convolutional neural network and its application in polarimetric SAR image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 12, pp. 7177–7188, Dec. 2017, doi: 10.1109/TGRS.2017.2743222.

[8] Z. Hong et al., "A high-precision recognition method of circular marks based on CMNet within complex scenes," *IEEE J Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 7431–7443, 2022, doi: 10.1109/JSTARS.2022.3194537.

[9] D. Marmanis, M. Datcu, T. Esch, and U. Stilla, "Deep learning earth observation classification using ImageNet pretrained networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 1, pp. 105–109, Jan. 2016, doi: 10.1109/lgrs.2015.2499239.

[10] F. Liu, L. Jiao, B. Hou, and S. Yang, "POL-SAR image classification based on Wishart DBN and local spatial," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 6, pp. 3292–3308, Jan. 2016, doi: 10.1109/TGRS.2016.2514504.

[11] A. K. Verma, S. Nagpal, A. Desai, and R. Sudha, "An efficient neural-network model for real-time fault detection in industrial machine," *Neural. Comput. Appl.*, vol. 33, no. 4, pp. 1297–1310, Jun. 2020, doi: 10.1007/s00521-020-05033-z.

[12] H. Firat and D. Hanbay, "Classification of hyperspectral images using 3D CNN based ResNet50," *IEEE Trans. Signal Process.*, vol. 14, no. 2, pp. 1–4, Jul. 2021, doi: 10.1109/SIU53274.2021.9477899.

[13] M. Ye et al., "A lightweight model of VGG-16 for remote sensing image classification," *IEEE J-STARS*, vol. 14, no. 99, pp. 6916–6922, Jun. 2021, doi: 10.1109/JSTARS.2021.3090085.

[14] X. Li, H. Ye, and S. Qiu, "Cloud contaminated multispectral remote sensing image enhancement algorithm based on MobileNet," *Remote Sens.*, vol. 14, no. 19, Aug. 2022, Art. no. 4815, doi: 10.3390/rs14194815.

[15] M. E. Paoletti, J. M. Haut, J. Plaza, and A. Plaza, "Scalable recurrent neural network for hyperspectral image classification," *J. Supercomputing*, vol. 76, no. 11, pp. 8866–8882, Feb. 2020, doi: 10.1007/s11227-020-03187-0.

[16] D. Hong, L. Gao, J. Yao, B. Zhang, A. Plaza, and J. Chanussot, "Graph convolutional networks for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 7, pp. 5966–5978, Jul. 2021, doi: 10.1109/TGRS.2020.3015157.

[17] F. Gang, C. Liu, Z. Rong, T. Sun, and Q. Zhang, "Classification for high resolution remote sensing imagery using a fully convolutional network," *Remote Sens.*, vol. 9, no. 6, Mar. 2017, Art.no. 498, doi: 10.3390/rs9050498.

[18] D. Pan, M. Zhang, and B. Zhang, "A generic FCN-based approach for the road-network extraction from VHR remote sensing images-using OpenStreetMap as benchmarks," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 2662–2673, 2021, doi: 10.1109/JSTARS.2021.3058347.

[19] T. Miyato, S.-I. Maeda, M. Koyama, and S. Ishii, "Virtual adversarial training: A regularization method for supervised and semi-supervised learning," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 8, pp. 1979–1993, Aug. 2019, doi: 10.1109/TPAMI.2018.2858821.

[20] R. A. R. Ashfaq, X.-Z. Wang, J. Z. Huang, H. Abbas, and Y.-L. He, "Fuzziness based semi-supervised learning approach for intrusion detection system," *Inf. Sci.*, vol. 378, no. 1, pp. 484–497, Feb. 2017, doi: 10.1016/j.ins.2016.04.019.

[21] J. Lu, V. Behbood, P. Hao, H. Zuo, S. Xue, and G. Zhang, "Transfer learning using computational intelligence: A survey," *Knowl. Based Syst.*, vol. 80, pp. 14–23, May 2015, doi: 10.1016/j.knosys.2015.01.010.

[22] C. Yang et al., "Lunar impact crater identification and age estimation with Chang'E data by deep and transfer learning," *Nature Commun.*, vol. 11, no. 1, Dec. 2017, Art. no. 6358, doi: 10.1038/s41467-020-20215-y.

[23] Z. Ouyang et al., "Primary scientific results of Chang'E-1 lunar mission," *Sci. China Earth Sci.*, vol. 53, pp. 1565–1581, Nov. 2010, doi: 10.1007/s11430-010-4056-2.

[24] W. Zuo, C. Li, and Z. Zhang, "Scientific data and their release of Chang'E-1 and Chang'E-2," *Chin. J. Geochem.*, vol. 33, pp. 24–44, Feb. 2014, doi: 10.1007/s11631-014-0657-3.

[25] J. K. Hillier, B. J. Buratti, and K. Hill, "Multispectral photometry of the Moon and absolute calibration of the Clementine UV/Vis camera," *Icarus*, vol. 141, no. 2, pp. 205–225, Oct. 1999, doi: 10.1006/icar.1999.6184.

[26] C. Fortezzo and T. Hare, "Completed digital renovation of the 1: 5,000,000 lunar geologic map series," in *Proc. 44th Annu. Lunar Planet. Sci. Conf.*, 2013, Art. no. 2114.

[27] T. Öhman, *Lunar and Planetary Institute*, Apr. 16, 2023. [Online]. Available: https://www.lpi.usra.edu/lunar/surface/

[28] Y. Du, F. Liu, J. Qiu, and M. Buss, "A semi-supervised learning approach for identification of piecewise affine systems," *IEEE Trans. Circuits Syst. I Reg. Papers*, vol. 67, no. 10, pp. 3521–3532, Oct. 2020, doi: 10.1109/tcsi.2020.2991645.

[29] M. Emadi, J. Tanha, M. E. Shiri, and M. H. Aghdam, "A selection metric for semi-supervised learning based on neighborhood construction," *Inf. Process. Manage.*, vol. 58, no. 2, Mar. 2021, Art. no. 102444, doi: 10.1016/j.ipm.2020.102444.

[30] Y. He, J. Wang, C. Liao, B. Shan, and X. Zhou, "ClassHyPer: ClassMix-based hybrid perturbations for deep semi-supervised semantic segmentation of remote sensing imagery," *Remote Sens.*, vol. 14, no. 4, Feb. 2022, Art. no. 879, doi: 10.3390/rs14040879.

[31] D. Hong, N. Yokoya, G. S. Xia, J. Chanussot, and X. X. Zhu, "X-ModalNet: A semi-supervised deep cross-modal network for classification of remote sensing data," *Int. Soc. Photogram. Remote Sens. J. Photogram. Remote Sens.*, vol. 167, pp. 12–23, Sep. 2020, doi: 10.1016/j.isprsjprs.2020.06.014.

[32] Y. Wang, L. Gao, Y. Gao, and X. Li, "A new graph-based semi-supervised method for surface defect classification," *Robot. Comput. Integr. Manuf.*, vol. 68, Apr. 2021, Art. no. 102083, doi: 10.1016/j.rcim.2020.102083.

[33] Y. A. Malkov and D. A. Yashunin, "Efficient and robust approximate nearest neighbor search using hierarchical navigable small world graphs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 4, pp. 824–836, Apr. 2020, doi: 10.1109/tpami.2018.2889473.

[34] Y. Zhang et al., "Wind direction retrieval from CYGNSS L1 level sea surface data based on machine learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5804413, doi: 10.1109/TGRS.2022.3221373.

[35] M. Huang, S. Liu, Z. Li, S. Feng, D. Y. Wu, and F. Shu, "Remote sensing image fusion algorithm based on two-stream fusion network and residual channel attention mechanism," *Wireless Commun. Mobile Comput.*, vol. 2022, Jan. 2022, Art. no. 8476000, doi: 10.1155/2022/8476000.

[36] W. Zhao, T. Mu, and D. Li, "Classification of hyperspectral images based on two-channel convolutional neural network combined with support vector machine algorithm," *J. Appl. Remote Sens.*, vol. 14, no. 2, May 2020, Art. no. 024514, doi: 10.1117/1.Jrs.14.024514.

[37] X. Deng, W. Jiang, and Z. Wang, "Zero-sum polymatrix games with link uncertainty: A Dempster-Shafer theory solution," *Appl. Math. Comput.*, vol. 340, no. 1, pp. 101–112, Jan. 2019, doi: 10.1016/j.amc.2018.08.032.

[38] T. Denoeux, "A k-nearest neighbor classification rule based on Dempster-Shafer theory," *IEEE Trans. Syst. Man Cybern.*, vol. 25, no. 5, pp. 804–813, May 1995, doi: 10.1109/21.376493.

[39] S. Gou, Z. Di, K. Cai, Z. Liu, and S. Li, "Absolute model age of lunar Finsen crater and geologic implications," *Icarus*, vol. 354, no. 1, Jan. 2021, Art. no. 114046, doi: 10.1016/j.icarus.2020.114046.

[40] Y. Wang and B. Wu, "Active machine learning approach for crater detection from planetary imagery and digital elevation models," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 8, pp. 5777–5789, Aug. 2019, doi: 10.1109/tgrs.2019.2902198.

[41] X. Hou et al., "Absolute model ages of three craters in the vicinity of the Chang'E-5 landing site and their geologic implications," *Icarus*, vol. 372, no. 15, Jan. 2022, Art. no. 114730, doi: 10.1016/j.icarus.2021.114730.

Zhonghua Hong (Member, IEEE) received the Ph.D. degree in geographic information system (GIS) from the Tongji University, Shanghai, China, in 2014.

He is currently a Professor with the College of Information Technology, Shanghai Ocean University, Shanghai, China, since 2022. His research interests including satellite/aerial photogrammetry, high-speed videogrammetric, planetary mapping, 3D emergency mapping, GNSS-R, deep learning, and processing of geospatial Big Data.

Linxuan Zhong received the B.S. degree in computer science and technology from the Jinggangshan University of Electronic and Information Engineering, Jiangxi, China, in 2021. She is currently working toward the M.S. degree in electronic information with the Shanghai Ocean University, Shanghai, China.

Her research interests include deep learning object detection and image classification.

Xiaohua Tong (Senior Member, IEEE) received the Ph.D. degree in geo information system (GIS) from the Tongji University, Shanghai, China, in 1999.

From 2001 to 2003, he was a Postdoctoral Researcher with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan, China. He was a Research Fellow with Hong Kong Polytechnic University, Hong Kong, in 2006. From 2008 to 2009, he was a visiting scholar with the University of California, Santa Barbara, CA, USA. His research interests include trust in spatial data, photogrammetry and remote sensing, and image processing for high-resolution satellite images.

Haiyan Pan received the Ph.D. degree in surveying and mapping from the Tongji University, Shanghai, China, in 2020.

She is currently a Lecture with the College of Information Technology, Shanghai Ocean University, Shanghai, China. Her research interests include multispectral/hyperspectral image classification, multitemporal remote sensing data analysis, and change detection.

Ruyan Zhou received the Ph.D. degree in agricultural bioenvironment and energy engineering from the Henan Agricultural University, Zhengzhou, China, in 2007.

She is currently an Associate Professor with the College of Information Technology, Shanghai Ocean University, Shanghai, China. Her research interests include photogrammetry and deep learning.

Jing Wang received the Ph.D. degree in biomedical engineering with the Department of Biomedical Engineering, Shanghai Jiaotong University, Shanghai, China, in 2014.

From 2015 to the present, she has been an Associate Professor with the College of Information Technology, Shanghai Ocean University, Shanghai, China. Her research interests include computer vision and medical image processing.

Yun Zhang received the Ph.D. degree in applied marine environmental studies from the Tokyo University of Maritime Science and Technology, Tokyo, Japan, in 2008.

From 2011 to the present, he has been a Professor with the College of Information and Technology, Shanghai Ocean University, Shanghai, China. His research interests include the study of navigation system reflection signal technique and its maritime application.
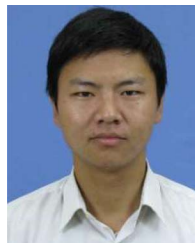
Shuhu Yang received the Ph.D. degrees in physics from the School of Physics, Nanjing University, Nanjing, China.

He is currently a Lecturer with the College of Information Technology, Shanghai Ocean University, Shanghai, China, since 2012. His research interests include hyperspectral remote sensing, evolution of the Antarctic ice sheet, and the use of navigational satellite reflections.

Yanling Han received the B.E. degree in mechanical design and manufacturing in 1996, and the M.E. degree in mechanical automation from the Sichuan University, Sichuan, China, in 1999, and the Ph.D. degree in engineering and control theory from the Shanghai University, Shanghai, China, in 2005.

She is a Professor and currently working with the Shanghai Ocean University, Shanghai, China. Her research interests include the study of ocean remote sensing, flexible system modeling, and deep learning.

Haiyang He received the M.Sc. degree in advanced control and system engineering from the University of Manchester, Manchester, U.K., in 2019.

He is currently working as a computer vision developing Engineer with Sense-Time Shanghai Ltd., Shanghai, China. His research includes computer vision, multiaxis-robot path planning, vision quality inspection, and control system optimization.