IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 6 June 2024, accepted 2 July 2024, date of publication 9 July 2024, date of current version 18 July 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3425654

RESEARCH ARTICLE

A Novel Attention-Based Early Fusion Multi-Modal CNN Approach to Identify Soil Erosion Based on Unmanned Aerial Vehicle

SHENG MIAO[1], YUFENG LIU[1], ZITONG LIU[1], XIANG SHEN[2], CHAO LIU[3], AND WEIJUN GAO[4]

[1]School of Information and Control Engineering, Qingdao University of Technology, Qingdao 266520, China
[2]Department of Statistics, The George Washington University, Washington, DC 20052, USA
[3]School of Environmental and Municipal Engineering, Qingdao University of Technology, Qingdao 266033, China
[4]Faculty of Environmental Engineering, The University of Kitakyushu, Kitakyushu 808-0135, Japan

Corresponding author: Chao Liu (liuchao@qut.edu.cn)

ABSTRACT Soil erosion poses significant ecological and economic challenges, necessitating precise and effective identification methods. Traditional models frequently overlook the intricate relationships between erosion factors and multispectral remote sensing data. To enhance these traditional methods, a novel dual-input gated fusion Convolutional Neural Network (CNN) has been developed, integrating channel and spatial attention mechanisms. This innovative model strengthens the connection between multispectral images and erosion factors, improving the accuracy and generalizability of erosion predictions. The model utilizes data collected from unmanned aerial vehicles (UAVs) equipped with high-precision multispectral sensors. By processing both spectral images and erosion factor data, the model effectively captures complex soil spatial distributions. The dual-input gated fusion mechanism allows the network to extract high-level semantics while suppressing redundant information, ensuring robust performance even in heterogeneous terrains. Experimental results indicate that this framework significantly enhances the performance of traditional models, providing superior predictions for small and medium-sized areas. Experimental results indicate that the presented framework can achieve better accuracy (96.92%) compared with other machine learning approaches, such as Random Forest (89.64%), VGGNET (91.52%), and RESNET (90.18%). Moreover, the proposed method can improve accuracy by 26.59% compared to the traditional RUSLE model. This improvement is critical for applications in ecological restoration and sustainable development. The integration of deep learning techniques with UAV-based data collection offers a powerful tool for environmental monitoring and management.

INDEX TERMS Convolutional neural network, gated fully fusion, convolutional block attention module, soil erosion, unmanned aerial vehicle.

I. INTRODUCTION

Soil erosion encompasses the deterioration, erosion, movement and deposition of the surface layer of the land due to natural elements such as water intrusion, wind intrusion and frost. Soil erosion poses a threat not only to soils, freshwater, oceans but also to food production and ecosystems [1]. Soils are intricately tied to human health and urban development [2]. Severe soil erosion can lead to direct damage on urban economies [3]. Studies have shown that the pace of soil erosion far surpasses the pace of natural soil formation [4]. Therefore, reducing the speed of soil erosion and improving the restoration and reconstruction of soil erosion hold great significance to ecological restoration and sustainable development [5]. The key step of solving soil erosion hazards is to learn the law of spatial differentiation of soil erosion

The associate editor coordinating the review of this manuscript and approving it for publication was Wenming Cao.

characteristics in the study area, accurately classify the erosion type of the erosion area, and promptly implement preventive measures to mitigate the erosion hazard.

Current research shows that remote sensing technology has made significant progress in the field of soil detection and research. Soil spectral images acquired by sensors contain important insights into the soil environment, the actual spatial distribution of the regional surface, etc., and are an important source of data for soil erosion research. The current mainstream erosion type classification schemes are mainly linear classification models based on erosion factors and image-based deep learning models. In complex soil environments, universal linear classification models usually need to be adapted to the local environmental equations in the study area [6]. For example, Di Stefano [7], Schurz [8], Gwapedza [9] and other scholars have adopted correction, improvement, calibration and other schemes to improve the applicability of soil erosion equations. In recent years, with the continuous development of deep learning, its special arithmetic characteristics make the model show obvious performance advantages in processing complex image semantic research. The deep learning model relies on the input data for adaptive parameterization, is insensitive to redundant data, and is able to receive more dimensional data for operation. Considering that this study area is a national coastal scenic area, its internal soil environment is complex and the remote sensing data contains more redundant information. Therefore, this study improved the data processing method by applying a deep learning model on the basis of a universal linear classification model to complete the classification of soil erosion degree.

The technical roadmap outlined in this paper, depicted in Fig. 1, the methodologies takes into account data relationships, enhances fitting capabilities, models more data using deep learning networks, employs algorithmic adjustments to boost the model's classification performance in intricate terrains, and establishes a more pragmatic soil erosion classification model. This paper introduces an improved convolutional neural network framework–a dual-input gated fusion convolutional neural network based on Convolutional Block Attention Module (CBAM) to jointly model erosion factors and multi-spectral images. Specifically, important features hidden in the two parts of the data are found in both dimensions, spatial and channel, and the dual-input gating mechanism is applied to simulate the interactions between the two sets of features to simulate a more realistic soil erosion scenario and improve the accuracy of the model. The contributions of this work can be summarized as follows:

1) An improved data processing method based on deep learning approach for soil erosion factors and multi-spectral remote sensing data.
2) Utilizing the Convolutional Block Attention Module (CBAM) mechanism enhances hidden features in soil erosion data at spatial and channel scales, while introducing an improved Dual Input Gate-controlled Full Fusion (GFF) module to simulate complex interactions between erosion factors and multispectral images.
3) In a specific case study, modeling was conducted on 2.9% of sample data within the study area, resulting in the overall identification of soil erosion levels in the research area.

The rest of the paper is organized as follows. Section II describes the related work. Section III presents the model applied in this research, including the attention mechanism and the gated fusion mechanism. Section IV evaluates the performance and efficiency of the model, and Section V summarizes the findings of this paper.

II. RELATED WORKS

In the initial stage of soil erosion research, soil researcher Wischmeier and Smith proposed an integrated approach to combine multiple erosion factor variable prediction models, establish the classical universal Soil Loss Equation (USLE) [10]. Renard developed the revised universal soil loss equation (RUSLE) [11], [12]. In addition, there is the Revised Wind Erosion Equation (RWEQ) [13].The equation took into account the relationship between the five major erosion factors, include rainfall erosivity [14], soil erodibility [15], sloped terrain [16], fractional vegetation cover [17] and soil and water conservation measures [18], which has been successfully applied to soil erosion studies in the eastern United State. Subsequently, the United States Department of Agriculture (USDA) introduced the official computational model known as Water Erosion Prediction Project (WEPP) [19], which incorporates more soil factors, further improving the accuracy of predictions. In subsequent studies, it was found that the resolution of remote sensing images directly affects the calculation accuracy of erosion factors, that is, on the spatial scale, the higher the resolution, the more soil details are included [20], and the corresponding accuracy and credibility of soil erosion prediction will be improved [21], [22], therefore, remote sensing technology has been widely used in soil research by many scholars. For example, M. Duan et al. combined remote sensing images and soil texture characteristics to produce a high-precision soil type map in the study area [23], and Zhang JR et al. calculated soil salinity content under different vegetation covers based on remote sensing image data [24]. Furthermore, Kurihara, Junichi et al. improved the accuracy of rice yield prediction in different environments based on remote sensing [25].

As machine learning continues to evolve, it is playing an extremely important role in many research directions in the field of soil science [26]. For example, Fang et al. used the Long Short-Term Memory (LSTM) network in deep learning to estimate the value of the Soil Moisture Active Passive (SMAP) [27]. Yamaguchi et al. built a rice leaf area index (LAI) prediction model based on deep learning [28]. Odebiri et al. applied deep learning frameworks to study remotely sensed Soil Organic Carbon (SOC) data [29]. Ren et al. apply deep learning models in the research on daily sea ice

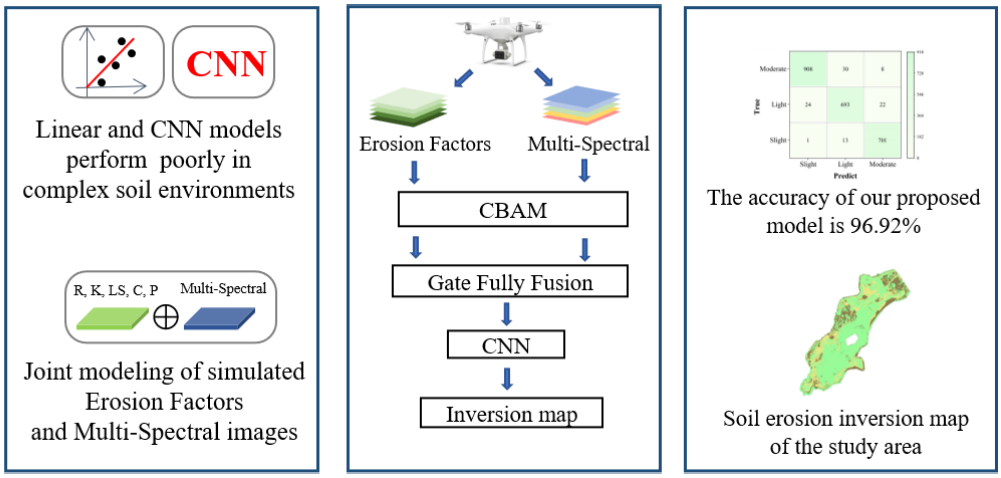


FIGURE 1. Technical roadmap.

concentration prediction in the Arctic region [30]. In recent years, Convolutional Neural Networks (CNN) have also made significant progress in the research of hyperspectral image classification [31]. The combination of CNNs and UAV remote sensing is often used to build high-precision models. Pottker et al. used CNN for classification to achieve plant community mapping [32]. Lanjewar et al. used deep learning CNN to classify soil types [33]. Zhao et al. completed high-precision vegetation type classification through a CNN model [34]. Whitehurst et al. employed CNN model to find traces of flood-induced erosion of the foundation from high-resolution drone images [35]. In practical applications, hyperspectral remote sensing data in space and channel, there are deeper features. Therefore, more and more scholars considered applying more complex convolutional neural networks. For example, Garajeh et al. proposed a new Deep Learning Convolutional Neural Network (DL-CNN) model to implement Soil Salinity (SSD) evaluation [36]. Xingmei et al. detected and counted of maize leaves using a two-stage deep learning model [37]. Han et al. proposed an Adaptive Multi-Source Data Fusion Network (AMSDFNet) based on deep learning to reduce labor costs in geological surveys and mapping [38]. In addition, Lee et al. proposed to a Contextual Deep CNN model to enhanced classification performance by extracting the Spatial-spectral feature of hyperspectral images [39]. Wang et al. Classification of individual cattle based on open pose mask R-CNN (OP-Mask R-CNN) [40]. Daanouni et al. Neural Structure Learning and a Multi-Head Attention (MHA) based cnn structure (NSL-MHA-CNN) for robust diabetic retinopathy prediction [41]. And Daoquan et al. traffic image classification using an Attention and Big-Step Convolutional Neural Networks (ABS-CNN) [42].

As shown in Table 1, the applicability of deep learning methods is broader, and the application scenarios are more diverse. Studies have shown that in universal linear classification models, not all spectral data have been fully utilized, and multiplication operations ignore the interrelationships between data dimensions. Additionally, the fixed parameters in the calculation formula need to be adjusted based on human knowledge, limiting the progress of actual research. To overcome the limitations in data calculation, this paper constructs a multimodal joint deep learning model. Building on the universal linear classification model data, it comprehensively demonstrates the impact of multi-spectral remote sensing data on soil erosion research, models the interrelationships between data, and adaptively improves parameters.

III. METHODOLOGY

In order to comprehensively demonstrate the impact of multi-spectral remote sensing data on soil erosion research, and to efficiently and dynamically allocate weight parameters to each spectral channel, as well as to investigate the interaction among different channels on the classification results, this paper proposes a DGCS-CNN model based on a dual-input convolutional neural network for jointly modeling multi-spectral images and soil erosion factors. In this model, the dual-input convolutional network receives the multi-spectral images and soil erosion factors, the gating fusion module is used to extract high-level semantics and suppress redundant information in the network, while the attention module is employed to compute the interaction between data, further enhancing the representation of feature information. To explore the value of the model in real-world applications, this paper applies the DGCS-CNN model in a coastal soil erosion prevention area. The scenic area contains an ecological protective wetland, several national marine parks and few many construction sites under planning and is connected to the main urban area. As the regional urbanization has been on the rise in recent years construction land has gradually expanded, and land use has tended to be intensive, which has changed the original soil environment,

TABLE 1. References collation.

Quantitative Models	Universal Soil Loss Equation Model [10-18]	Soil erosion studies in solution areas using numerical equations
	Water Erosion Prediction Project Model [19]	Incorporation of additional erosion factors further improves prediction accuracy
	Modified Revised Universal Soil Loss Equation Model [23] [24] [25]	Soil types mapped
		Calculated soil salinity content under different vegetation covers
		Improved the accuracy of rice yield prediction based on remote sensing
Deep Learning Models	Long Short-Term Memory (LSTM) Network [27]	Estimate the value of the Soil Moisture Active Passive (SMAP)
	Deep Learning Networks [28] [29] [30]	Leaf area index (LAI) prediction for rice
		Detection of remotely sensed soil organic carbon (SOC)
		Projections of daily sea ice concentrations in the Arctic
	Convolutional Neural Networks [32] [33] [34] [35]	Soil type image detection and recognition
		Mapping plant communities
		Classification of vegetation types to analyze the relationship between urban development and the natural environment
		Predicting traces of flood-induced foundation erosion
	A new Deep Learning Convolutional Neural Network model [36]	Soil Salinity evaluation
	Two-Stage Deep Learning [37]	Calculate the number of corn leaves
	Adaptive Multi-source Data Fusion Network [38]	Learning valuable spatial and spectral information from data sources using two branches
	Contextual Deep CNN [39]	Joint adjacent spatial-spectral relationships optimized to explore local contextual interactions
	Open pose mask R-CNN [40]	Classification of individual cattle
	Neural Structure Learning and a Multi-Head Attention CNN [41]	Retinopathy images prediction
	Attention and Big-Step Convolutional Neural Networks [42]	Traffic picture classification

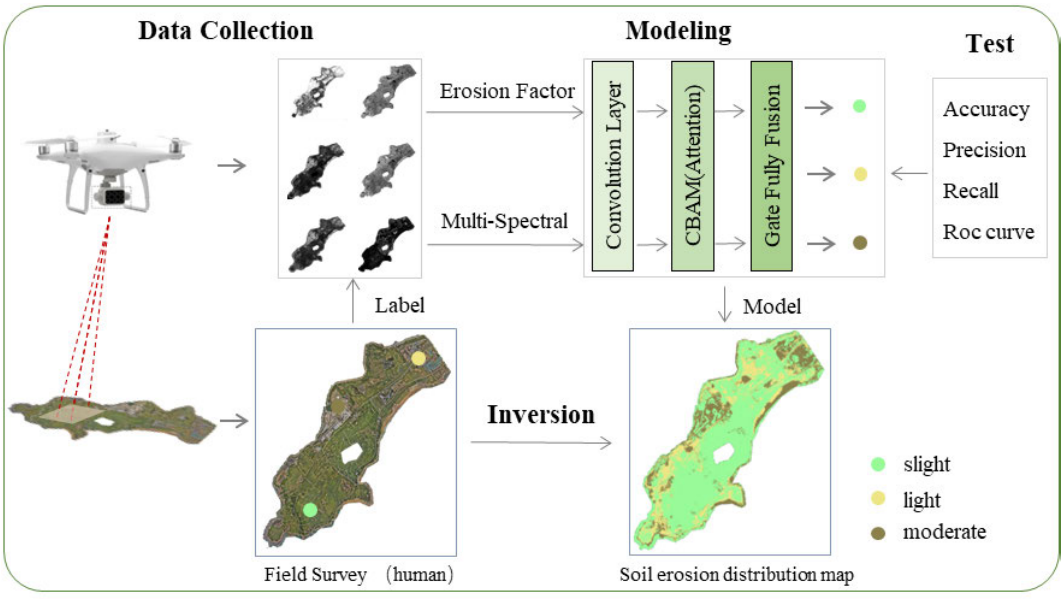


FIGURE 2. Research flowchart.

resulting in ecological risks such as soil pollution and soil erosion. Therefore, the study of soil erosion in the region holds exceptional importance.

The research process in this article is shown in Figure 2. First, the specific operation process of P4M UAV data collection in the research area, as well as the processing

process of multispectral data and erosion factor data, are described. According to the survey results of the field survey, the data were divided and marked, and then the processed two sets of data were input into the DGCS-CNN model. The parameter settings of each architecture of the model, as well as the attention mechanism and the gating full fusion mechanism, were introduced, and the applicability analysis of the soil erosion prediction model was introduced. Finally, the model is tested, the model with the best performance is selected, and the model is applied to complete the soil erosion inversion of the whole study area to obtain the erosion distribution map.

A. DATA COLLECTION

The collection of data in this study comes from UAV remote sensing. The drone selected for the flight test mission is a Phantom 4 Multi-spectral (P4M) UAV, which is equipped with a high-precision digital remote sensing sensor and a Real-Time Kinematics (RTK) module. The sensor boasts a visible lens and five narrowband spectral lenses, as outlined in the specification data displayed in Figure 3. In the experiment, the relative flight altitude of P4M was set to 100 meters to ensure the accuracy of the data. To ensure the integrity of the stitched images and the reliability of interpretation, the heading overlap rate is set to 80%, and the side overlap rate is set to 70%. In the actual mission, the total flight time was 20 hours, the cumulative flight distance was 286 miles, and 148,680 images were taken. In addition, to provide sample data for the identification model, after completing the UAV flight mission, field research work is carried out synchronously, sampling points are randomly arranged, and the actual soil erosion of sampling points and accurate latitude and longitude coordinates are recorded.

FIGURE 3. Sensor information for a P4M multispectral camera.

Based on the DEM data collected by P4M and the soil information data in the study area, five soil erosion factor data points of R, K, LS, C, and P were calculated in turn, the equations are shown in table 2. As shown in Equation (1), the R factor refers to the rainfall erosion coefficient, the parameter a in the equation is adaptively selected according to the latitude and longitude position of the region, P_n. represents the Nth month of rainfall in the study area. In Equation (2), the K factor refers to the erosion of the soil, where SAN is the sand content, SIL is the soil content, CLA is the clay content and C is the soil organic nitrogen content. In Equation (3), the LS factor is a composite factor of slope and slope length at the soil location. In the formula, L represents the slope length, and θ represents the slope. There are various ways to calculate L, and the reflux area method is used in this experiment. In Equation (4), the C factor was called sensitive factor in the prediction of soil erosion, where FCV represents the percentage of vegetation cover. In Equation (5), Soil and water conservation measures factor P, the value of P of each point in the area corresponds to the land use type of the point. The value range of the P value was between 0 and 1, where 0 means that there was no erosion at the location, and 1 means that no soil and water conservation measures have been taken at the location.

Considering the actual situation of the study area, this study determines the parameters of each factor calculation formula. In addition, according to the location and label of the sampling point recorded during the field survey, all the sample data are annotated, as shown in Figure 4, where the location of the sampling point is the latitude and longitude location information of the surveyed area, and the label is the true soil erosion degree of the sampling point which is used as classification model output. The classification and classification criteria of soil erosion in the study area were served as a reference, and multidimensional factors such as remote sensing data, DEM, and soil utilization type were extensively studied and analyzed. Finally, the soil erosion intensity in the study area was categorized into three levels; slight erosion, light erosion, and moderate erosion. In this paper, 432,063 photos were extracted and cut into 20 × 20 pixels images in the study area, of which 12000 photos were labeled as the dataset of the. The area of the dataset was approximately 2.77% of the total study area, with the proportions of the three erosion levels in the dataset being 28.9%, 39.6%, and 31.5%. In addition, to evaluate the performance of the model, the dataset is split into training dataset, validation dataset, and test dataset with a ratio of 7:1:2, where the test dataset does not participate in training and is only used to evaluate performance.

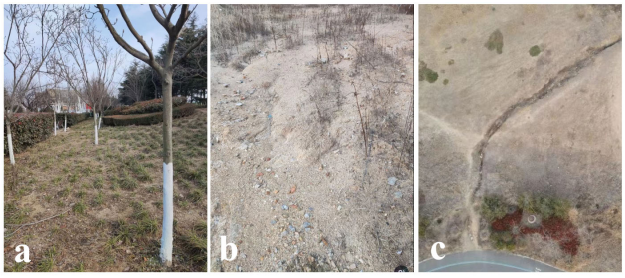


FIGURE 4. Sample data notes, (a) slight erosion, (b) light erosion and (c) moderate erosion.

B. MODELING

The CNN model is one of the most widely used algorithms of deep learning, and it is designed for efficiently processing grid-like data such as images with inductive bias. In this experiment, the spectral information and erosion factor of

TABLE 2. Soil erosion factor and equations [43].

Soil Erosion Factor	Equations
Rainfall erosion coefficient (R)	$R = \sum_{N=1}^{12} a \times P_N^{1.6295}$ (1)
Erosion of the Soil (K)	$\mathrm{K} = \frac{1}{7.59}\left\{0.2 + 0.3\mathrm{e}^{\left[-0.0256\mathrm{SAN}\left(1.0-\frac{\mathrm{SIL}}{100}\right)\right]}\right\} \times \left(\frac{\mathrm{SIL}}{\mathrm{CLA+SIL}}\right)^{0.3} \times \left[1.0 - \frac{0.25\mathrm{C}}{\mathrm{C}+\mathrm{e}^{(3.72-295\mathrm{C})}}\right] \times \left[1.0 - \frac{0.79(1-\mathrm{SAN}/100)}{1-\mathrm{SAN}/100+\mathrm{EXP}^{(-5.51+22.91-\mathrm{SAN}/100)}}\right]$ (2)
Composite factor of slope and slope length at the soil location (LS)	$\begin{cases} LS = L \cdot (10.8\sin\theta + 0.03 & \theta<5^\circ) \\ LS = L \cdot (16.8\sin\theta - 0.5 & 5^\circ \le \theta<10^\circ) \\ LS = L \cdot (21.9\sin\theta - 0.96 & \theta \ge 10^\circ) \end{cases}$ (3)
Sensitive factor in soil erosion (C)	$C = \begin{cases} 1 & \mathrm{FVC} \le 0.1 \\ 0.6508 - 0.34361g\,\mathrm{FVC} & 0.1 < \mathrm{FVC} \le 0.783 \\ 0 & \mathrm{FVC} > 0.783 \end{cases}$ (4)
Soil and water conservation measures factor (P)	$P \subset (0,1)^{(1\times R)}$ (5)

soil erosion are multi-channel image data, and CNN models can consumes the data as the input of the model and extract the characteristics of soil erosion through a convolution operation, while retaining the spatial invariance of erosion area. In addition, the CNN model employs the parameter sharing operation mechanism to reduce the number of parameters compared with fully connected neural network and reduce the complexity of feature extraction and classification. Additionally, there are rich resources of pretrained CNN backbones on large amount of images and the model can easily adapt to the new domain with limited labeled samples. In summary, the CNN model is selected to build the erosion classification model throughout the experiment. The structure of the proposed new architecture DGCS-CNN model is shown in Figure 5. First, the dual-input convolutional neural network was set up to receive multi-spectral data and erosion factors, the input embedding layer size of multi-spectral images was set to 20 × 20×8, and the size of the erosion factors was 20× 20 × 5. Three convolutional layers and three pooling layers were configured in the convolution module, two subsequent fully connected layers were designed to improve the classification accuracy, and the last fully connected layer was designed to output the prediction results of the DGCS-CNN model. Compared with conventional design of CNN models, we also added two modules: CBAM and GFF to improve the model performance for this specific task.

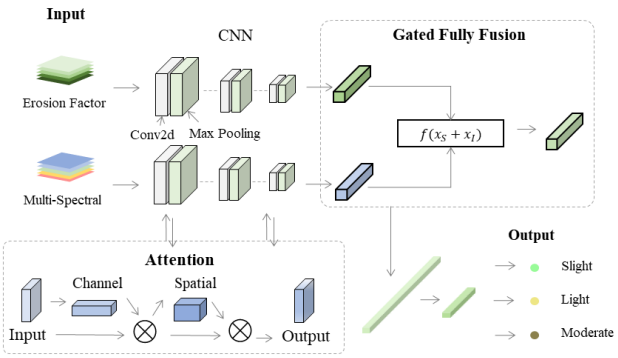


FIGURE 5. Model structure diagram.

1) CONVOLUTIONAL BLOCK ATTENTION MODULE

The Convolutional Block Attention Module (CBAM) is an improved attention-based module on CNN architecture. Previous work only focuses on the channel attention while CBAM introduced a combination of channel attention module and spatial attention module and achieved impressive performance on computer vision tasks. CBAM leverages a calculation process involving the feature attention matrix and the input feature map. These are sequentially multiplied along two independent dimensions: the channel level and the spatial level, thereby facilitating comprehensive feature optimization.

The architecture of CBAM is shown in Figure 6, which integrates the channel attention module and the spatial attention module.Specifically, in the channel attention module, as shown in Equation (6), the input multi-dimensional remote sensing images F, after passing through two parallel max pooling layers and average pooling layers, and two multilayer perceptron (MLP) layers with activation functions, and then the output feature $Out_c(F)$ and input F are summed by elementwise weighting, as shown in Equation (7), to obtain the output of the channel attention module F'.

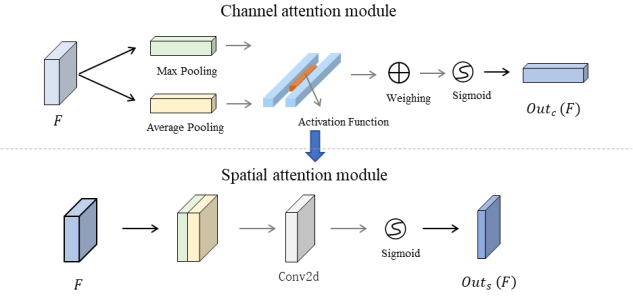


FIGURE 6. CBAM module.

In the spatial attention module, as shown in Equation (8), F' is taken as input, and then after the max pooling layers, the average pooling layers and convolution operations in turn, the spatial attention matrix $Out_s(F')$ is obtained, and the output F'' of CBAM is obtained by multiplying the input data again, as shown in Equation (9).

$$Out_c(F) = Sigmoid(MLP(AvgPooling(F) + MLP(MaxPooling(F)))) \quad (6)$$

$$F' = Out_c(F) \bigotimes F \quad (7)$$

$$Out_s(F) = Sigmoid(Conv2d([AvgPooling(F); MLP(MaxPooling(F))])) \quad (8)$$

$$F'' = Out_c(F') \bigotimes F \quad (9)$$

The data used in this experiment are multi-channel spectral data. At the channel level, for any location point in the study area, the weight distribution of the point in each spectrum should not be averaged and should be weighted based on the impact of each channel on the feature. In practical applications, it is desired that the network would pay more attention to the essential channels to improve the expression ability, and the CBAM module starts from the relationship between each channel, comprehensively considers the characteristics of each channel, finds the channel that is more meaningful for classification, and focuses its attention on the channel. At the spatial level, the spatial scale of UAV remote sensing images is at the centimeter-level, and the orthophoto of the study area contains all the detailed information of the area, such as a tree, and a very small amount of vegetation. Therefore, the CBAM module is used to extract important features from the spatial scale of the image, and further highlight the important In this study, the input data channel has high dimensions and rich information contained in the image, so the CBAM module is introduced to strengthen the channel with high weight allocation, weaken the channel with low weight allocation, and strengthen important features and interference features in complex remote sensing images, to assist the model to pay attention to key information.

2) GATED FULLY FUSION

In the DGCS-CNN model, the GFF module is employed to fuse spectral images and erosion factors on soil erosion prediction at early stages. CBAM focuses more on extracting high-level features in the data without considering the interactions between spectral and factors. In fact, the multi-spectral imaages reflect objective information at the spatial scale of the study area, while erosion factors focus more on actual soil erosion sensitivity, and all factors are closely related to soil characteristic information. Therefore, it is necessary to explicitly consider the influence of erosion factors on the prediction results during parameter transfer. The GFF module applies the feature fusion mechanism based on additive operation and two gates to filter the information before fusion The module efficiently aggregates the features and its architecture is shown in Figure 7.

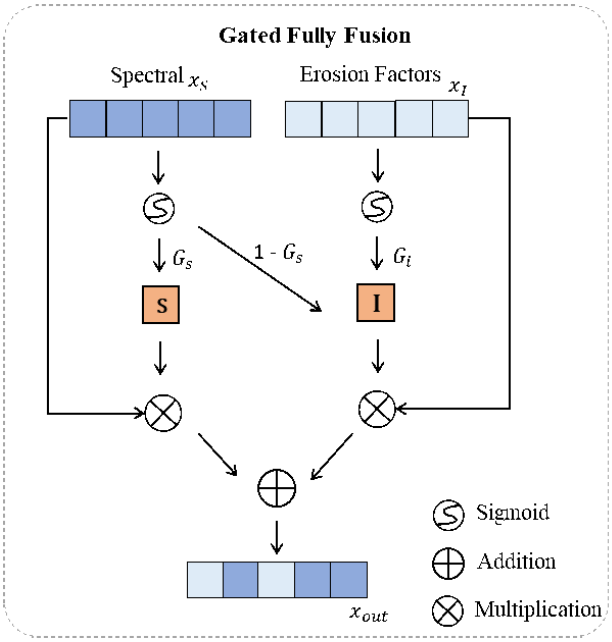


FIGURE 7. GFF module structure diagram.

Specifically, the input data of GFF x_S,x_I, come from the output of the upper layer of CNN, x_S the extracted features of the spectral data, and x_I the features extracted from the factor data. The inputs x_S and x_I were successively operated by $G_i = sigmoid(w_i \times x_i)$ to obtain G_s, G_i, where $\omega_i \subset (0, 1)^{(1\times R)}$ is the weight parameter, which is the same dimension as the input feature of the upper network, and the sigmoid function was designed for parameterizing the input data. The spectral gate G_S and factor gate G_I were set in the GFF to filter the information from both side before fusion. As shown in Equations (10) and (11),G_S passes both spectral and factor gates, while G_i only communicates with factor gates. a large value of the elements of G_I referes to higher weights of elements in x_I in the fusion.

$$G_s = (1 + G_s) \quad (10)$$

$$G_I = (1 - G_s)G_i \quad (11)$$

x_S and x_I were multiplied by the corresponding gate structure in turn to complete the feature fusion process, and the fusion calculation is shown in Equation (12):

$$x_{out} = f(G_S x_S + G_I x_I) \quad (12)$$

GFF inhibits the integration of invalid spectral information into the output, by measuring the importance of the information in the output, judging the impact of x_I on the output, in the fusion, reasonable control of the x_S and x_I ingress. In summary, GFF uses two gates to model the complex interactions of spectral images and factors on prediction, and further improves the performance of the model by measuring the complexity of the information and controlling the extraction of information, and effectively alleviates the data redundancy and noise information.

IV. EXPERIMENTAL RESULTS

To demonstrate the effectiveness of the DGCS-CNN model proposed in this paper, experiments were conducted in real-world data collected from soil erosion risk areas. In this section, first describe the configuration of the experiment and compare the common machine learning models with deep learning models. In addition, to further verify the performance of each module and the improvement effect of prediction accuracy, this paper conduct ablation experiments on CBAM and GFF modules. Throughout the study, implement the experiment under an computation environment with Python3.8 and TensorFlow2.6. All the models were trained on the server with two core CPUs(Intel(R) Xeon(R) Silver 4210R), one GPU(NVIDA GeForce RTX3090) and 256 RAM.

A. TRAINING AND VALIDATION

The experiment includes training dataset and validation dataset with 8400 and 1200 sampling respectively. The cross entropy loss function was selected during training. The essence of the cross function is to measure the distance of the probability distribution of ground truth and predicted values, and the goal of minimizing the cross-entropy loss provides reliable metrics for the update parameters of the whole network. The model achieves the best parameters of the whole network gradually during training process.

In the training process, this study employs the Adam optimizer, which offers a comprehensively approach by considering the first-order moment estimation and second-order moment estimation of the gradient. This enables each layer to independently adjust the learning rates from the two aspects of the mean and square of the gradient and improve the convergence speed of the model. In addition, the grid search method is applied to search the best combination of hyper-parameters including the number, size and step size of convolution kernels. The best configuration is determined according to the convergence degree and evaluation on validation dataset. Finally, the number of convolution kernels in each layer of the network are set to 64, 128, and 256, the sizes are 7, 5, and 3 respectively, and the step size is 1. Two fully connected layers for receiving the GFF module embedded in the size set to 800 and 3. The last fully connected layer outputs three erosion classification levels.

To monitor the evaluation of model training, 50 epochs of iterative training with 300 batches are employed. The accuracy and loss during training are shown in Figures 8a and 8b. During the training period, the loss of the validation set decreases rapidly and reaches a plateau state, and the accuracy gradually increases and tends to converge to a plateau state. In the last 10 epochs of training, the accuracy and loss value of the validation set remain stable, which indicates that the model converges and has good performance. Therefore, the model trained after 50 epochs was selected as the final model.

To further evaluate the performance of the model, a test dataset with 2400 samples are fed into the model in the experiment to evaluate the performance of the model. Because the samples in the test dataset do not participate in training, they can reflect the performance of model recognition. Finally, the accuracy of the model in the test set reached 96.9%. In addition, considering that the test dataset is composed of a random sample of the dataset, the sample size is usually unevenly distributed The confusion matrix of the model is calculated and visually analyzed in the form of a thermal map. As shown in Figure 9, it can be observed that the highlighted parts of each classification in the thermal map are distributed in the diagonal, so it can be judged that the model has a good classification ability at this time for each category.

B. OVERALL PERFORMANCE

In this experiment, Random Forest (RF) classification model is included as the baseline method. In the comparison experiment VGGNET, RESNET, CBAM-CNN, and the current mainstream method an Attention and Big Step Convolutional Neural Network (ABS-CNN) are set up. To ensure the fairness of the comparison, the training dataset and test dataset of all models are the same.

It is worth noting that, compared to the baseline method set in this paper, the accuracy of deep learning methods has significantly improved. This is attributed to the complexity of the study area's environment, the presence of considerable noise in the spectral data, and the erosion factors such as rainfall erosivity and vegetation coverage factors, which vary with climate change [44], climate change does not have a uniform dimension and is almost unpredictable, which renders the random forest algorithm unable to identify accurate feature relationships for soil erosion classification, thus affecting the classification accuracy. Therefore, the application of deep learning algorithms with stronger fitting ability and the capacity to suppress noise and redundant information leads to a more noticeable performance enhancement in soil erosion classification. In addition, the results of CBAM-CNN and DGCS-CNN show that the two key components of CBAM and GFF both improve the accuracy of the model to varying degrees. This paper proposed method, compared with the mainstream deep learning model ABS-CNN, has large performance improvement. To explore the comprehensive performance of each model, the accuracy, precision, recall, and F1-Score of the machine learning model are calculated in the test dataset, and the results are shown in Table 3. Precision and recall are used as evaluation indicators of the actual prediction ability of the model to measure the real prediction ability of each category. When the precision and recall of the model are high, it indicates that the real prediction ability of the model for each category and the retrieval ability for the real label are strong. In addition, this paper uses the F1-Score to represent the balance point between precision and recall; when the F1-Score is higher, the performance of the model is better.

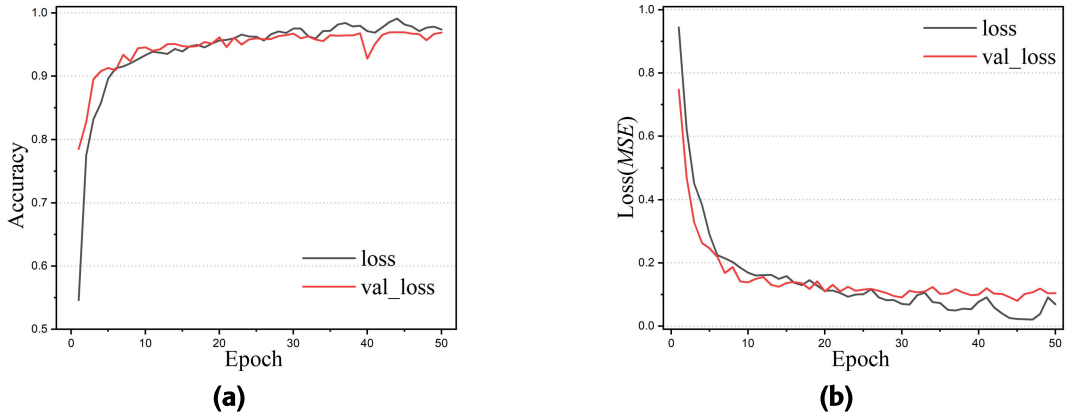


FIGURE 8. Training history of the accuracy and loss.

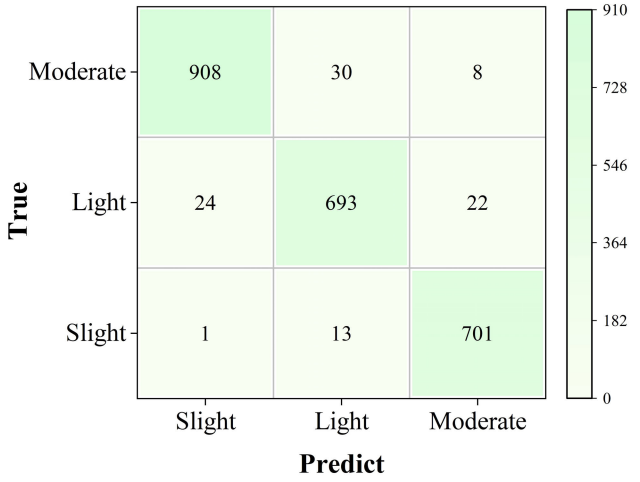


FIGURE 9. Confusion matrix of prediction model with GFF module.

To visualize the performance comparison results of the deep learning models, this study selected three representative deep learning models, and draw the ROC curves of these three deep learning models based on the test set, and calculate the area under curve (AUC) of slight, light, and moderate erosion in turn. As shown in Figure 10, the AUC of slight, light and moderate erosion in the DGCS-CNN model is significantly better the other models. This shows that DGCS-CNN achieved a better performance in this task.

The above results show that the DGCS-CNN model has higher accuracy and good classification performance. However, compared with other models, the DGCS-CNN model has the most complex network structure and requires more computational resources. To explore the computational feasibility of the model, this experiment calculates the time required for the three models to complete the training task and the inference task. As shown in Table 4, DGCS-CNN takes the longest time for training, and the training time is four times that of the vanilla CNN. However, in the area of soil erosion research, the most time consuming part is postprocessing the data so that the training time may not block the study in general and the training time is acceptable. Meanwhile, the inference time of the three models is almost the same, which indicates that the DGCS-CNN model has huge value in real-world application.

TABLE 3. Precision, recall, and F1-Score results.

Model	Soil erosion	Precision	Recall	F1-Score
Random Forest	Slight	89.06%	85.55%	87.31%
	Light	86.52%	82.24%	84.38%
Acc (Gain) = 89.64% (-)	Moderate	88.66%	83.35%	86.01%
VGGNET	Slight	95.95%	93.93%	94.94%
	Light	92.96%	89.08%	91.02%
Acc (Gain) = 91.52% (2.10%)	Moderate	92.08%	94.66%	93.37%
RESNET	Slight	96.24%	93.58%	94.91%
	Light	91.66%	93.56%	92.61%
Acc (Gain) = 90.18% (0.60%)	Moderate	94.58%	97.64%	96.11%
CBAM-CNN	Slight	97.26%	94.87%	96.05%
	Light	95.66%	95.99%	95.82%
Acc (Gain) = 93.2% (3.97%)	Moderate	95.28%	96.89%	96.08%
ABS-CNN	Slight	97.95%	95.93%	96.94%
	Light	96.09%	95.45%	95.77%
Acc (Gain) = 93.89% (4.74%)	Moderate	97.68%	98.06%	97.87%
Proposed	Slight	**99.56%**	**97.64%**	**98.59%**
	Light	**97%**	**97.77%**	**97.38%**
Acc (Gain) = 96.92% (8.12%)	Moderate	**97.78%**	**99.44%**	**98.60%**

TABLE 4. Training and inference time comparison.

Model	Train dataset time	Test dataset time
Random Forest	1000s	25s
VGGNET	1450s	26s
RESNET	1450s	26s
CBAM-CNN	1500s	26s
ABS-CNN	1550s	26s
DGCS-CNN	2150s	26s

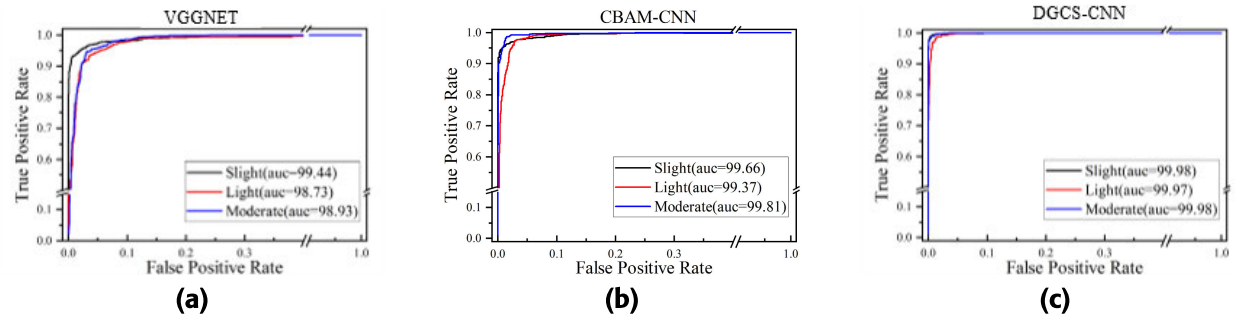


FIGURE 10. ROC curve and AUC of the model.

C. ABLATION STUDY

To further explore the performance of CBAM in capturing high-order feature information, this study randomly selects a subgraph in the training set and observes the changes in the feature map before and after CBAM application. The selected submap is labeled as Slight, which shows that there are many tall trees on the surface of the soil covered, although the soil consolidation effect of a single tree is limited, while multiple trees planted together can effectively protect the soil by preventing wind erosion, rain erosion, etc., so that the degree of erosion is low. In addition, we expect the model to pay attention to the characteristics of vegetation in this subplot and the information contained in the soil under the trees. As shown in Figure 11, by visualizing the input and output images of CBAM, it can be observed that CBAM captures the outline of the tree and the soil under the tree, successfully finds the area of interest and enhances the feature. The histogram in Figure 11 shows the sum of the weight parameters of the corresponding network of each channel of the spectral data calculated in the experiment, in which the two channels with the highest weight distribution NIR and B-V are selected in visible light and invisible light, respectively, and channel R with the lowest weight among all channels is selected for visual observation. In the figure, $NIR_1, B-V_1, R_1$ represents the raw data of the channel, NIR_2, $B-V_2$, R_2 represents the output after CBAM processing. By observing the two output channels NIR_2 and $B-V_2$, it can be found that the soil information in the output feature map is given higher weight, and the soil under the tree and the edge of the tree are more obvious, although the information in the input RE_1 channel is more blurred and does not map the more obvious features, and after multiplying with the feature attention matrix, the output R_2 channel still pays attention to the characteristics of the soil. In addition, the channel has less useful information in the spectrum, the feature distribution is not obvious, the impact on the classification results is minimal, and the network parameters of this layer are also the lowest. The results show that CBAM can efficiently capture the most important information in the image.

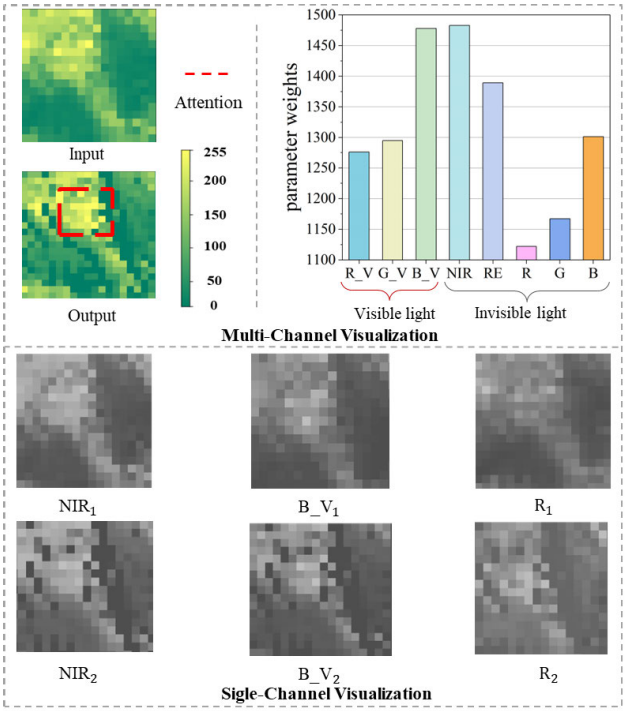


FIGURE 11. Visualization of the CBAM module.

D. RESULTS ANALYSIS

RUSLE is currently the most commonly used factorization model for soil erosion research. Table 5 provides a performance comparison between the RUSLE model and the DGCS-CNN model. In the case study discussed in this paper, the RUSLE model had a recall rate of zero in the mild and moderate categories, indicating its inability to accurately classify the soil erosion intensity in the study area. The DGCS-CNN model takes into account the internal relationship between erosion factors and multispectral data, offering a comprehensive understanding of the data dimensions. Deep learning models boast complex network structures, a large number of internal parameters, and enhanced fitting capabilities. Furthermore, while factorization model calculation parameters are fixed and require adjustment based on human knowledge in different environmental conditions to enhance classification accuracy, deep learning models can adapt and learn different feature rules for distinct regions based on environmental conditions. These models extract features

layer by layer to achieve data classification with strong adaptability and swift classification speeds.

TABLE 5. Precision, recall, and F1-Score results.

Model	Soil erosion	Precision	Recall	F1-Score
RUSLE	Slight	70.33%	0.00%	0.00%
	Light	**100.00%**	0.00%	0.00%
Acc = 70.33% (-)	Moderate	41.28%	0.00%	0.00%
Proposed	Slight	**99.56%**	**97.64%**	**98.59%**
	Light	97%	**97.77%**	**97.38%**
Acc = 96.92% (26.59%)	Moderate	**97.78%**	**99.44%**	**98.60%**

After a comprehensively analysis, the DGCS-CNN model demonstrates a strong ability to provide accurate predictions. Therefore, we proceeded to employ the DGCS-CNN model to the remaining 97.1% (420,063 data points) of the area to obtain the spatial distribution maps of the three soil erosion prediction results in the study area. To enhance the detailed observation of the distribution map, the image was processed with pseudo-color as shown in Figure 12, and the three erosion types of slight, light, and moderate were marked as green, yellow, and brown, respectively.

Soil erosion is attributed to a combination of natural factors and human factors. Natural factors lay the foundation of soil erosion while human factors significantly accelerate the process, including destruction of vegetation, overgrazing, road construction, engineering construction, etc. As construction sites expand, industrial production intensifies, and engineering projects advance, the disturbance caused to land and vegetation becomes more pronounced. This not only exacerbates pre-existing soil erosion but also exerts a severe toll on the ecological environment. The cumulative effects of these activities are substantial, affecting the stability of ecosystems and the overall environmental equilibrium.

As shown in Figure 12, the central area of the study area is covered by a large area of trees, and the degree of soil erosion is manifested as light. To further improve its water and soil conservation capacity, a small area of forest can be transformed into a multilayer structure of mixed forest, such as the establishment of coniferous forest, broad-leaved forest and shrub complex forest. Through the multilayer forest structure, litter is protected, enhancing the soil consolidation capacity and soil fertility of the area. In mountainside areas with large slopes and exposed surface vegetation, erosion is manifested as slight, and this area can effectively control soil erosion by increasing barren slope treatment and greening. Erosion in coastal areas manifests as light and moderate, so more coastal sheltered forests should be built and soil and water conservation forests should be constructed, including natural forests and plantations with high canopy density. Most of the moderate erosion in the study area occurred around the building land. Supervision and administrative departments shall make full use of remote sensing, unmanned aerial vehicles, information

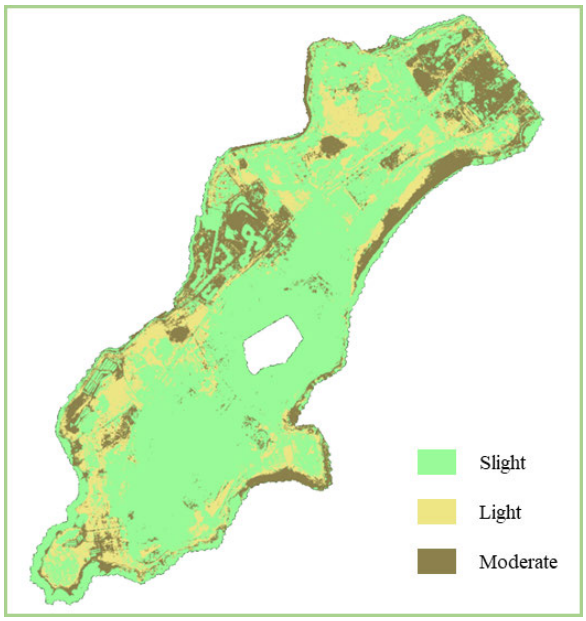


FIGURE 12. Inversion map of the DGCS-CNN model.

networks and other technologies to strengthen soil and water monitoring. To ensure the implementation of soil and water conservation measures in industrial production and construction activities, and to restrict or prohibit industrial production and construction activities that may cause soil erosion in important ecologically fragile and sensitive areas.

V. CONCLUSION

In this study, a dual-input gated attention convolutional neural network model is proposed for the accurate identification and prediction of soil erosion in small-and medium-sized loss areas. In the case study of a real-world soil and water loss area, the accuracy of the model achieves more than 96%. Furthermore, through the analysis of results from various algorithms, it was observed that deep learning models exhibit significantly higher performance in soil erosion prediction. This is attributed to the complex terrain environment leading to noise and a wealth of redundant information in the spectral data, as well as the intricate interactions among various erosion factors in soil erosion. Therefore, deep learning models are more suitable for the computation of soil data. For the classification of soil erosion at a medium and small scale research area, the DGCS-CNN achieves more accurate soil erosion localization and efficient, targeted management, which holds significant importance for ecological restoration and soil erosion prevention. The DGCS-CNN combines attention mechanism with convolutional neural networks to assign different weights to high spectral images in different bands. The dual input of the DGCS-CNN is designed for the soil erosion factors and remote sensing spectral images in the study area. In the DGCS-CNN, the dual-input gating fusion module aims to simulate the complex interactions between erosion factors and multi-spectral images. This study better

replicates the original appearance of the soil, resulting in more realistic and reliable predictive outcomes.

Using the CBAM mechanism to strengthen the channel weight ratio with greater impact on the spectral images could further improve the accuracy of the model, especially for the understory soil with a small amount of vegetation cover in the erosion. The gated fusion mechanism also demonstrates decent performance in the ablation study, and compared with the advanced mainstream models, it has obvious improvement. Overall, the fusion of different views provides a new angle to conduct studies in soil erosion and related area.

REFERENCES

[1] P. Borrelli, D. A. Robinson, P. Panagos, E. Lugato, J. E. Yang, C. Alewell, D. Wuepper, L. Montanarella, and C. Ballabio, "Land use and climate change impacts on global soil erosion by water (2015–2070)," *Proc. Nat. Acad. Sci. USA*, vol. 117, no. 36, pp. 21994–22001, Sep. 2020.

[2] P. M. Kopittke, N. W. Menzies, P. Wang, B. A. McKenna, and E. Lombi, "Soil and the intensification of agriculture for global food security," *Environ. Int.*, vol. 132, Nov. 2019, Art. no. 105078.

[3] H. R. Pourghasemi, F. Honarmandnejad, M. Rezaei, M. H. Tarazkar, and N. Sadhasivam, "Prioritization of water erosion–prone sub-watersheds using three ensemble methods in Qareaghaj catchment, southern Iran," *Environ. Sci. Pollut. Res.*, vol. 28, no. 28, pp. 37894–37917, Jul. 2021.

[4] D. Wuepper, P. Borrelli, and R. Finger, "Countries and the global rate of soil erosion," *Nature Sustainability*, vol. 3, no. 1, pp. 51–55, Dec. 2019.

[5] J. Ö. G. Jónsson, B. Davíosdóttir, E. M. Jónsdóttir, S. M. Kristinsdóttir, and K. V. Ragnarsdóttir, "Soil indicators for sustainable development: A transdisciplinary approach for indicator development using expert stakeholders," *Agricult., Ecosyst. Environ.*, vol. 232, pp. 179–189, Sep. 2016.

[6] R. Benavidez, B. Jackson, D. Maxwell, and K. Norton, "A review of the (revised) universal soil loss equation ((R)USLE): With a view to increasing its global applicability and improving soil loss estimates," *Hydrol. Earth Syst. Sci.*, vol. 22, no. 11, pp. 6059–6086, Nov. 2018.

[7] C. Di Stefano, V. Ferro, and V. Pampalone, "Applying the USLE family of models at the Sparacia (South Italy) experimental site," *Land Degradation Develop.*, vol. 28, no. 3, pp. 994–1004, Apr. 2017.

[8] C. Schuerz, B. Mehdi, J. Kiesel, K. Schulz, and M. Herrnegger, "A systematic assessment of uncertainties in large-scale soil loss estimation from different representations of USLE input factors—A case study for Kenya and Uganda," *Hydrol. Earth Syst. Sci.*, vol. 24, no. 9, pp. 4463–4489, Sep. 2020.

[9] D. Gwapedza, N. Nyamela, D. A. Hughes, A. R. Slaughter, S. K. Mantel, and B. van der Waal, "Prediction of sediment yield of the Inxu River catchment (South Africa) using the MUSLE," *Int. Soil Water Conservation Res.*, vol. 9, no. 1, pp. 37–48, Mar. 2021.

[10] W. H. Wischmeier and D. D. Smith, *Predicting Rainfall Erosion Losses: A Guide to Conservation Planning*. Washington, DC, USA: United States Department of Agriculture, 1978.

[11] M. K. Boitt, O. M. Albright, and H. K. Kipkulei, "Assessment of soil erosion and climate variability on Kerio Valley Basin, Kenya," *J. Geosci. Environ. Protection*, vol. 8, no. 6, pp. 97–114, 2020.

[12] J. Aswathi, K. S. Sajinkumar, A. Rajaneesh, T. Oommen, E. H. Bouali, R. B. B. Kumar, V. R. Rani, J. Thomas, K. P. Thrivikramji, R. S. Ajin, and M. Abioui, "Furthering the precision of RUSLE soil erosion with PSInSAR data: An innovative model," *Geocarto Int.*, vol. 37, no. 27, pp. 16108–16131, Dec. 2022.

[13] H. Du, X. Liu, X. Jia, S. Li, and Y. Fan, "Assessment of the effects of ecological restoration projects on soil wind erosion in northern China in the past two decades," *CATENA*, vol. 215, Aug. 2022, Art. no. 106360.

[14] D. Zhu, K. Xiong, and H. Xiao, "Multi-time scale variability of rainfall erosivity and erosivity density in the Karst region of southern China, 1960–2017," *CATENA*, vol. 197, Feb. 2021, Art. no. 104977.

[15] W. Ouyang, Y. Wu, Z. Hao, Q. Zhang, Q. Bu, and X. Gao, "Combined impacts of land use and soil property changes on soil erosion in a mollisol area under long-term agricultural development," *Sci. Total Environ.*, vols. 613–614, pp. 798–809, Feb. 2018.

[16] S. Peng, "Using GIS and RUSLE to study on the geographical distribution characteristics of soil erosion on the Dianchi watershed, China," *Desalination Water Treatment*, vol. 219, pp. 113–121, Apr. 2021.

[17] G. Hou, H. Bi, Y. Huo, X. Wei, Y. Zhu, X. Wang, and W. Liao, "Determining the optimal vegetation coverage for controlling soil erosion in *Cynodon dactylon* grassland in North China," *J. Cleaner Prod.*, vol. 244, Jan. 2020, Art. no. 118771.

[18] P. Tian, Z. Zhu, Q. Yue, Y. He, Z. Zhang, F. Hao, W. Guo, L. Chen, and M. Liu, "Soil erosion assessment by RUSLE with improved P factor and its validation: Case study on mountainous and hilly areas of Hubei province, China," *Int. Soil Water Conservation Res.*, vol. 9, no. 3, pp. 433–444, Sep. 2021.

[19] G. Biali, P. Cojocaru, and P. Schneider, "GIS-based water erosion modelling: The case of high slope water catchment area, Romania," *Environ. Eng. Manage. J.*, vol. 20, no. 4, pp. 613–624, 2021.

[20] A. Li, X. C. Zhang, and B. Liu, "Effects of DEM resolutions on soil erosion prediction using Chinese soil loss equation," *Geomorphology*, vol. 384, Jul. 2021, Art. no. 107706.

[21] C. von Hebel, S. Reynaert, K. Pauly, P. Janssens, I. Piccard, J. Vanderborght, J. van der Kruk, H. Vereecken, and S. Garré, "Toward high-resolution agronomic soil information and management zones delineated by ground-based electromagnetic induction and aerial drone data," *Vadose Zone J.*, vol. 20, no. 4, Jul. 2021, Art. no. e20099.

[22] J. Zhang, W. Guo, B. Zhou, and G. S. Okin, "Drone-based remote sensing for research on wind erosion in drylands: Possible applications," *Remote Sens.*, vol. 13, no. 2, p. 283, Jan. 2021.

[23] M. Duan, X. Song, X. Liu, D. Cui, and X. Zhang, "Mapping the soil types combining multi-temporal remote sensing data with texture features," *Comput. Electron. Agricult.*, vol. 200, Sep. 2022, Art. no. 107230.

[24] J. Zhang, Z. Zhang, J. Chen, H. Chen, J. Jin, J. Han, X. Wang, Z. Song, and G. Wei, "Estimating soil salinity with different fractional vegetation cover using remote sensing," *Land Degradation Develop.*, vol. 32, no. 2, pp. 597–612, Jan. 2021.

[25] J. Kurihara, T. Nagata, and H. Tomiyama, "Rice yield prediction in different growth environments using unmanned aerial vehicle-based hyperspectral imaging," *Remote Sens.*, vol. 15, no. 8, p. 2004, Apr. 2023.

[26] Q. Yuan, H. Shen, T. Li, Z. Li, S. Li, Y. Jiang, H. Xu, W. Tan, Q. Yang, J. Wang, J. Gao, and L. Zhang, "Deep learning in environmental remote sensing: Achievements and challenges," *Remote Sens. Environ.*, vol. 241, May 2020, Art. no. 111716.

[27] K. Fang, M. Pan, and C. Shen, "The value of SMAP for long-term soil moisture estimation with the help of deep learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 4, pp. 2221–2233, Apr. 2019.

[28] T. Yamaguchi, Y. Tanaka, Y. Imachi, M. Yamashita, and K. Katsura, "Feasibility of combining deep learning and RGB images obtained by unmanned aerial vehicle for leaf area index estimation in rice," *Remote Sens.*, vol. 13, no. 1, p. 84, Dec. 2020.

[29] O. Odebiri, O. Mutanga, J. Odindi, R. Naicker, C. Masemola, and M. Sibanda, "Deep learning approaches in remote sensing of soil organic carbon: A review of utility, challenges, and prospects," *Environ. Monit. Assessment*, vol. 193, no. 12, Dec. 2021, Art. no. 802.

[30] Y. Ren and X. Li, "Predicting the daily sea ice concentration on a subseasonal scale of the pan-arctic during the melting season by a deep learning model," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 4301315.

[31] W. Dong, Y. Yang, J. Qu, S. Xiao, and Y. Li, "Local information-enhanced graph-transformer for hyperspectral image change detection with limited training samples," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5509814.

[32] M. Pöttker, K. Kiehl, T. Jarmer, and D. Trautz, "Convolutional neural network maps plant communities in semi-natural grasslands using multispectral unmanned aerial vehicle imagery," *Remote Sens.*, vol. 15, no. 7, p. 1945, Apr. 2023.

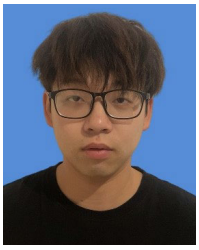
[33] M. G. Lanjewar and O. L. Gurav, "Convolutional neural networks based classifications of soil images," *Multimedia Tools Appl.*, vol. 81, no. 7, pp. 10313–10336, Mar. 2022.

[34] D. Zhao, X. Li, X. Wang, X. Shen, and W. Gao, "Applying digital twins to research the relationship between urban expansion and vegetation coverage: A case study of natural preserve," *Frontiers Plant Sci.*, vol. 13, Feb. 2022, Art. no. 840471.

[35] D. Whitehurst, B. Friedman, K. Kochersberger, V. Sridhar, and J. Weeks, "Drone-based community assessment, planning, and disaster risk management for sustainable development," *Remote Sens.*, vol. 13, no. 9, p. 1739, Apr. 2021.

[36] M. K. Garajeh, F. Malakyar, Q. Weng, B. Feizizadeh, T. Blaschke, and T. Lakes, "An automated deep learning convolutional neural network algorithm applied for soil salinity distribution mapping in Lake Urmia, Iran," *Sci. Total Environ.*, vol. 778, Jul. 2021, Art. no. 146253.
[37] X. Xu, L. Wang, M. Shu, X. Liang, A. Z. Ghafoor, Y. Liu, Y. Ma, and J. Zhu, "Detection and counting of maize leaves based on two-stage deep learning with UAV-based RGB image," *Remote Sens.*, vol. 14, no. 21, p. 5388, Oct. 2022.
[38] W. Han, J. Li, S. Wang, X. Zhang, Y. Dong, R. Fan, X. Zhang, and L. Wang, "Geological remote sensing interpretation using deep learning feature and an adaptive multisource data fusion network," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 4510314.
[39] H. Lee and H. Kwon, "Going deeper with contextual CNN for hyperspectral image classification," *IEEE Trans. Image Process.*, vol. 26, no. 10, pp. 4843–4855, Oct. 2017.
[40] J. Wang, X. Zhang, G. Gao, Y. Lv, Q. Li, Z. Li, C. Wang, and G. Chen, "Open pose mask R-CNN network for individual cattle recognition," *IEEE Access*, vol. 11, pp. 113752–113768, 2023.
[41] O. Daanouni, B. Cherradi, and A. Tmiri, "NSL-MHA-CNN: A novel CNN architecture for robust diabetic retinopathy prediction against adversarial attacks," *IEEE Access*, vol. 10, pp. 103987–103999, 2022.
[42] D. Li, X. Dong, J. Gao, and K. Hu, "Abnormal traffic detection based on attention and big step convolution," *IEEE Access*, vol. 11, pp. 64957–64967, 2023.
[43] K. G. Renard, *Predicting Soil Erosion by Water: A Guide to Conservation Planning With the Revised Universal Soil Loss Equation*. Washington, DC, USA: United States Department of Agriculture, 1997.
[44] J. Zhao, Z. Wang, Y. Dong, Z. Yang, and G. Govers, "How soil erosion and runoff are related to land use, topography and annual precipitation: Insights from a meta-analysis of erosion plots in China," *Sci. Total Environ.*, vol. 802, Jan. 2022, Art. no. 149665.

SHENG MIAO received the Ph.D. degree from Towson University, MD, USA, in 2017. He is currently an Associate Professor with the School of Information and Control Engineering, Qingdao University of Technology. He has published multiple high quality research papers in journals and conferences. His research interests include machine learning, human computation, smart healthcare, and intelligence systems.



YUFENG LIU was born in Mudanjiang, China, in 1999. He is currently pursuing the master's degree with the School of Information and Control Engineering, Qingdao University of Technology, China. His research interests include machine learning and artificial intelligence systems.

ZITONG LIU was born in Shandong, China, in 1999. She is currently pursuing the master's degree with the School of Information and Control Engineering, Qingdao University of Technology, China. Her research interests include machine learning and artificial intelligence systems.

XIANG SHEN received the B.S. degree in mathematics from Fudan University, in 2014, and the Ph.D. degree in statistics from George Washington University, in 2020. His main research interests include machine learning, deep learning, and natural language processing.

CHAO LIU received the Ph.D. degree from the School of Environmental and Municipal Engineering, Qingdao University of Technology, in 2018. She is currently an Associate Professor with the School of Information and Control Engineering, Qingdao University of Technology. Her current research interests include water environment system analysis and water pollution control.

WEIJUN GAO is currently a tenured Professor with The University of Kitakyushu, Japan, and an Academia Professor with Qingdao University of Technology, China. He has been a Visiting Professor at many universities, including China and USA, such as Xi'an Jiaotong University, Zhejiang University, and Lawrence Berkeley National Laboratory, USA. His research interests include the science, engineering, management, dissemination of city environment planning, distributed energy systems, building material recycle, health and environmental impacts of energy generation and use, geographic information systems, and climate change, especially in urban area and energy forecasting.