GCFC: Graph Convolutional Fusion CNN Network for Cross-Domain Zero-Shot Extraction of Winter Wheat Map

Chunyang Wang, Peipei Zhou, Yudong Zhang, *Senior Member, IEEE*, Junding Sun, Bibo Lu, Zhaozhao Xu, Baishun Su, and Xingwang Li, *Senior Member, IEEE*

***Abstract*—Accurate extraction of winter wheat and its planted area holds significance for agricultural research and government real-time food monitoring. Traditional machine learning methods often demand extensive data and corresponding labels for training in cross-domain classification problems. The heterogeneity of land cover types causes an uneven distribution of samples, leading to unsatisfactory results when these methods are applied directly to other regions. This article introduces a two-branch Graph Convolutional Fusion CNN network incorporating dynamic weighted stratified loss to address these challenges. To reduce the weight of losses generated by easily classified pixels, this loss function adds task masks and category dynamic weights to the cross-entropy loss. Dual branching merges global insights from graph convolutional networks with local emphasis from convolutional neural networks. It enhances the handling of cross-domain classification problems. The first branch introduces an adaptive mechanism and applies it to the graph's adjacency matrix to enhance the model's adaptability to different domain graph structures. The second branch alleviates the oversmoothing problem of edge clustering caused by graph convolution and handles multiscale and spectral information more efficiently. The experimental results showed that the proposed method achieved 99.98% accuracy and good classification results on the Zhoukou dataset. The zero-shot cross-domain prediction on the Suixian dataset achieved 96.11% accuracy. Ultimately, the entire winter wheat planting area of Shangqiu City was extracted with an accuracy of 91.92%. Numerous experiments and practical applications confirm that the proposed method is feasible and effective for winter wheat cross-domain extraction.**

***Index Terms*—Cross-domain, graph convolution, oversmoothing, winter wheat.**

Manuscript received 5 February 2024; revised 2 April 2024; accepted 16 April 2024. Date of publication 22 April 2024; date of current version 1 May 2024. This work was supported in part by the Chunhui Program Cooperative Research Project of Chinese Ministry of Education under Grant HZKY20220279, in part by the Henan Provincial Science and Technology Research Project under Grant 232102211019 and Grant 222102210131, in part by the Key Research Project Fund of Institution of Higher Education in Henan Province under Grant 23A520029, in part by the National Natural Science Foundation of China under Grant 62303167, and in part by the Nationally Funded Postdoctoral Researcher Program of China under Grant GZC20230707. *(Corresponding authors: Chunyang Wang; Yudong Zhang.)*

Chunyang Wang, Peipei Zhou, Junding Sun, Bibo Lu, Zhaozhao Xu, and Baishun Su are with the School of Computer Science and Technology, Henan Polytechnic University, Jiaozuo 454000, China (e-mail: wcy@hpu.edu.cn; 212209020056@home.hpu.edu.cn; sunjd@hpu.edu.cn; lubibo@hpu.edu.cn; xzz@hpu.edu.cn; subaishun@hpu.edu.cn).

Yudong Zhang is with the School of Computer Science and Technology, Henan Polytechnic University, Jiaozuo 454000, China, and also with the Department of Information Technology, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia (e-mail: yudongzhang@ieee.org).

Xingwang Li is with the School of Physics and Electronic Information Engineering, Henan Polytechnic University, Jiaozuo 454000, China (e-mail: lixingwangbupt@gmail.com).

Digital Object Identifier 10.1109/JSTARS.2024.3392448

I. Introduction

FOOD constitutes the fundamental foundation for human survival and development. Winter wheat occupies a prominent position within the spectrum of major grains in China [1], [2]. Winter wheat plays a crucial role in China's agricultural production, with its output intricately connected to the nation's import and export trade policies [3]. Furthermore, the precise delineation of winter wheat planting areas is of immense significance for China's food security and economic prosperity [4], [5]. Consequently, meticulously examining methods for accurately extracting winter wheat acreage becomes imperative.

Conventional crop harvesting methods rely on sampling, counting, and reporting, necessitating extensive field visits by managers [6]. However, these methods exhibit inherent drawbacks, including high subjectivity, substantial errors, time-intensive procedures, slow update rates, and a lack of spatial distribution [7]. To overcome these limitations, integrating remote sensing techniques into the extraction of feature areas has emerged as a viable solution for delineating crop cropping structures [8], [9].

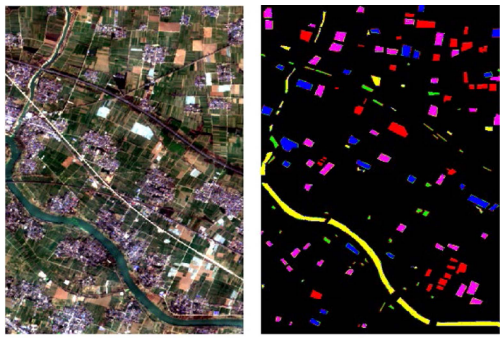
Due to the early limitations of computer hardware and software, machine learning-based methods have been used for remote sensing image classification for finite region extraction applications. Commonly employed approaches included neural networks [10], [11], support vector machines (SVMs) [12], [13], decision trees [14], [15], and random forests [16], [17]. Moustakidis et al. [18] introduced a fuzzy decision tree that utilizes a binary SVM for node discrimination. This approach demonstrates superior classification performance compared to traditional SVM and other methods. Liu et al. [19] innovatively extracted maize drought areas by constructing a vegetation index time series curve, achieving an extraction with 90.03% accuracy. Huang et al. [20] introduced a rule-set-based remote sensing classification method, attaining commendable results in extracting and estimating date palm plantation areas in the study region. Zhao et al. [21] employed an enhanced composite kernel to augment the SVM method, enhancing not only the kernel's

generalization ability and learning efficacy but also achieving high accuracy in remote sensing image classification. While these methods effectively extract features, it is essential to note that the extracted features are single-pixel features, and several drawbacks and shortcomings persist. Challenges include high data requirements, difficulty in feature extraction, algorithmic complexity, limited model generalization ability, and sensitivity to parameter tuning.

With the advancement of deep learning, there is a growing trend in the semantic segmentation of remote-sensing images using deep learning techniques [22], [23], [24]. Unlike traditional machine learning algorithms, deep learning semantic networks provide a complete pipeline for feature extraction. They use convolutional operations to maximize the utilization of spatial information, capturing local structures and relationships in images. Chen et al. [25] applied high-resolution remote sensing images for end-to-end segmentation using full convolutional networks, enabling efficient pixel-level prediction and more precise segmentation results. In a notable contribution, Zhang et al. [26] introduced an enhanced segmentation network that integrates attention with U-Net, achieving superior results in distinguishing between Amazonian rainforest and nonrainforest areas. Le et al. [27] used the SegNet architecture for the semantic segmentation of multichannel airborne images. In addition, Li et al. [28] proposed reducing the complexity of the dot product attention mechanism to $O(N)$ core attention. A multiscale attention network is devised to conduct semantic segmentation of high-resolution remote sensing images, consistently delivering optimal classification performance with the highest accuracy. Liu et al. [29] presented multitask pixels end-to-end CNN RoadNet capable of simultaneously predicting road surface, edges, and centerline. This work introduces a user interaction operation, effectively addressing shadows and occlusions in the road region. Zhao et al. [30] proposed a new spatial attention pyramid convergence module. Combining the attention mechanism and the multiscale module improves the accuracy of high-resolution aerial image labeling. Despite their robust feature representation capabilities, there is a noted underutilization and under-recognition of features that impact their full potential. The demanding requirement for a substantial amount of labeled data for training poses a challenge, considering the time-consuming and expensive nature of data labeling and sample acquisition. This challenge is particularly pronounced in the context of winter wheat sample acquisition, which necessitates availability within critical cycles. Further investigations are needed to explore methods that involve reduced data labeling and more suitable modeling approaches.

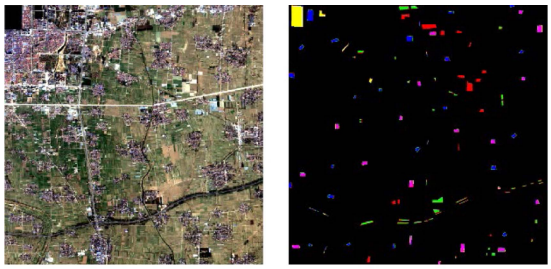
Bengio et al. [31] employed a pretraining model using a task-specific source database to tackle the abovementioned issue. This approach links an extensive set of labeled source data with a limited quantity of unlabeled target data, thereby advancing research in computer vision in visual transfer adaptive learning. You et al. [32] proposed a generalized domain adaptive approach that tackles the disparity between the training and testing fields. Tian et al. [33] introduced a multisource subdomain adaptive transfer learning method to transfer multisource knowledge to cross-domain knowledge. The method uses a multibranch network structure to match the feature space distribution of each source and target domain. It employs the local maximum mean difference for a detailed local comparison of subdomain distributions within different domains of the same category. Zhu et al. [34] successfully solved the problem that there may be too little labeled data to construct a good classifier by utilizing a large amount of unlabeled data and a small amount of labeled data. Marfurt et al. [35] introduced the application of deep transfer learning in remote sensing scene classification. It discussed the advantages and challenges of deep transfer learning methods in addressing high data demand and insufficient model generalization capability. Mehmood et al. [36] summarized how transfer learning is applied in remote sensing image classification. It discussed the advantages and limitations of transfer learning methods in addressing issues related to sample scarcity and feature extraction difficulties. Zhang et al. [37] proposed a novel meta-transfer learning approach called MetraSAR for recognizing synthetic aperture radar (SAR) targets across tasks and domains. This method enhances the effectiveness of transferring knowledge across different tasks and domains. Knowledge transfer is usually applied between similar tasks. In remote sensing data, due to climatic reasons, the remote sensing images of the same feature over different periods are more different, and there are also phenomena of the same feature. While transfer learning has made strides in remote sensing image classification, misclassification remains common when classifying across geographic regions due to variations in spectral reflectance and differences in similar features between different geographic areas.

Compared to the mentioned structures, a graph convolutional network has proven more effective in establishing correlations among various land covers [38], [39]. This capability facilitates the transfer of learned knowledge for application in diverse domains. First proposed in 2005, it was initially intended to solve some strictly graph-theoretic problems. The most widely used GCN, initially proposed by Bruna in 2013 (a student of LeCun), relies on two classification methods: one in the frequency domain (Spectral domain) and another in the spatial domain (Spatial domain) [40]. Zhang et al. [41] proposed a superpixel-based framework for graphical model categorization. The framework adopts the superpixel as the fundamental unit of the graphical model, allowing it to capture contextual information and spatial dependencies among the superpixel. Ma et al. [42] introduced a new attention graph convolutional network for effective superpixel segmentation in extensive SAR image datasets. This model enhances segmentation algorithm efficiency and graph convolutional network performance by incorporating an attention layer. Qi et al. [43] proposed a 3-D graphical neural network designed for constructing k-nearest neighbor graphs on 3-D point clouds. The network demonstrates proficiency in capturing long-distance dependencies in images, which poses a challenge for traditional methods to model. Tian et al. [44] proposed a novel EGAN-LSTM framework for multiclass geospatial object detection in HSRI. This framework involves parameterizing and embedding graph-structured data extracted from image data into the network for joint learning and model optimization. Liu et al. [45] introduced CNN-enhanced GCNs for heterogeneous deep



(a) (b)

Fig. 1. Zhoukou dataset. (a) RGB-color map. (b) Ground-truth map.



(a) (b)

Fig. 2. Suixian dataset. (a) RGB-color map. (b) Ground-truth map.

networks. Integrating graph encoding into the network addresses the structural incompatibility between Euclidean data-oriented CNNs and non-Euclidean data-oriented GCNs. In addition, this integration enhances the learning of node features, leading to improved adaptability of graphs to HSI content. Yu et al. [46] proposed the ConGCN algorithm for HSI classification, incorporating an adaptive graph enhancement technique to improve contrast learning performance by integrating spectral-spatial prior. While the graph structure enables more accurate feature extraction, challenges like poor edge effects persist. The method proposed in the study aims to address issues like pixel ratio imbalance and the unsatisfactory direct application of the model across regions in the winter wheat classification task. The primary contributions of the research are as follows.

1) Aiming at the challenge of information transfer in heterogeneous graphs within the cross-domain winter wheat extraction task, the study proposed the adaptive graph convolutional network (AGCN), innovatively enhancing the traditional graph convolutional network. Before constructing the adjacency matrix, an adaptive mechanism is incorporated to weight the edges in the graph dynamically. This enhancement allows the AGCN network to adjust edge weights based on the input data's features and context, enhancing the model's flexibility in accommodating heterogeneous graph structures. Consequently, the adaptive mechanism significantly improves the model's generalization ability.
2) The study proposed that the PCSNet network tackle the edge clustering issue arising from graph convolution techniques. Various image regions receive dynamic attention for feature processing by integrating a spatial channel attention module (CAM). A spatial spectrum separation convolution (SSON) [47] module is also introduced to perceptively analyze features in different frequency bands. The synergy of the two modules enables the PCSNet network to capture the multiscale and spectral information of winter wheat in the image, effectively improving the grasp of edge information.
3) To further enhance the performance of the cross-domain extraction task for winter wheat images, the study optimized the PCSNet network by introducing an encoder-decoder structure to better learn and express the key information in the images. The optimization boosts performance in the cross-domain extraction task and strengthens the network's ability to extract features.
4) The study proposed the dynamic weighted stratified loss (DWSL) algorithm to address sample category imbalance and enhance segmentation ability for challenging samples in winter wheat extraction. The algorithm, refined using the cross-entropy loss function, incorporates task masks and category dynamic weights to focus the model on crucial categories and difficult samples. The approach effectively resolves sample imbalance issues.

II. Materials

A. Dataset

The datasets used in the study were obtained through Sentinel-2 and retrieved from the Google Earth Engine platform. The Zhoukou dataset was constructed by selecting the Zhoukou region at a size of 700 pixels $\times$ 500 pixels. The Suixian dataset was constructed by selecting the Suixian County area at a size of 1000 pixels by 1000 pixels. Figs. 1 and 2 display the RGB color map and the ground truth of the images. Each time series dataset contains only one image with 28 spectral bands and a spatial resolution of 10 m. The data captures a wealth of information about the surface in a multiband format and includes images from a total of four periods. Images for each period consisted of red, green, blue, near infrared, normalized difference vegetation index, normalized difference built-up index, and normalized difference water index. Real color images use a combination of red, green, and blue bands to simulate the colors the human eye sees, creating realistic ground colors. Different combinations of red and near-infrared bands are used to reflect the state of the vegetation more accurately when assessing vegetation health and growth.

The datasets were categorized into five distinct feature types: winter wheat, forest, buildings, water, and other land, presented in a specified order. About 6% of the Zhoukou dataset (350 000 pixels in total, with 32 392 labeled pixels) was annotated. Approximately 2% of the Suixian dataset (1 000 000 pixels in total, with 24 997 labeled pixels) was annotated. The training and test samples of the datasets are distributed in a 7:3 ratio, with the validation set including 1% randomly selected samples from the

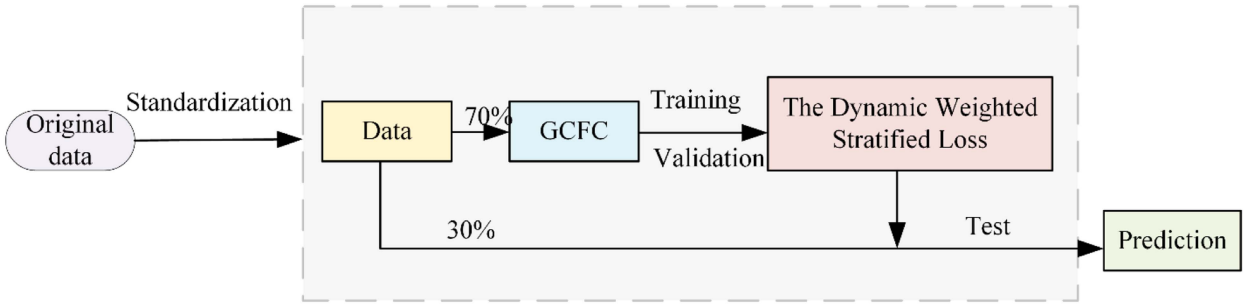


Fig. 3. Flowchart of winter wheat extraction.

TABLE I
NUMBER OF PIXELS IN THE TRAINING, VALIDATION, AND TEST SETS FOR EACH CATEGORY ON THE ZHOUKOU DATASET

Class Name	Training	Validation	Test
Winter Wheat	6273	268	2688
Forest	1629	69	698
Buildings	3901	167	1672
Water	5941	254	2546
Other Land	4938	211	2112
Total	22 682	969	9716

TABLE II
NUMBER OF PIXELS IN THE TRAINING, VALIDATION, AND TEST SETS FOR EACH CATEGORY ON THE SUIXIAN DATASET

Class Name	Training	Validation	Test
Winter Wheat	3508	150	1503
Forest	3539	152	1517
Buildings	3714	159	1592
Water	3526	151	1511
Other Land	3211	138	1376
Total	17 498	750	7499

overall test set for validation purposes. Tables I and II show the detailed categorization of the dataset partitions.

III. METHODS

Fig. 3 shows the flowchart of the winter wheat extraction method. First, the image is normalized. Next, the model is trained and validated using 70% of the image data and labels. Finally, the remaining 30% of the image data and labels are fed into the trained network to obtain the prediction map. DWSL is presented in Section III-A. The proposed Graph Convolutional Fusion CNN (GCFC) network is given in Section III-B. The forward propagation process of GCFC is given in Algorithm 1. The training process of GCFC is given in Algorithm 2.

A. Dynamic Weighted Stratified Loss

There is an imbalance in the ratio of each category within the dataset in the study, with the distribution of categories in the Zhoukou dataset used for training approximately being address 6:2:4:6:5. In addressing these issues, the study proposed the

Algorithm 1: GCFC Algorithm Forward Propagation Process.

Input: Input feature $X \epsilon R^{H \times W \times B}$; $SLIC()$ is the superpixel segmentation algorithm with a segmentation scale of d; $Encoder - Decoder()$ consists of multiple convolutional layers, normalization layers, batch operations; σ is a nonlinear function.

Output: The prediction map F

$M \leftarrow SLIC(X)$
$\hat{M}$ is obtained by normalizing M
The undirected graph node V is obtained from (7)
The reduced network feature X^* is obtained from (8)
$F_1 \leftarrow \sigma(X^*)$
X_1 is obtained from (12)
X_2 is obtained from (13)
X_3 is obtained from (14)
$X_4 \leftarrow Encoder - Decoder(X_3)$
$F_2 \leftarrow \sigma(X_4)$
$F \leftarrow \sigma(Concat(F_1, F_2))$

DWSL method. DWSL adopts a comprehensive strategy to sample imbalances in remote sensing images. The model adapts weights dynamically in real time by incorporating cross-entropy loss, sample weights, and dynamic weight adjustment. This strategic adjustment ensures heightened attention to categories characterized by fewer samples, significantly enhancing the model's overall performance The cross-entropy loss denoted as $L_{CE_{ij}}$ served as a metric to quantify the disparity between the model output and the actual label. To ensure numerical stability, incorporate a minor constant (10^{-15}) to prevent the potential emergence of infinite values during logarithmic computations. The loss utilizes the true label mask to filter out areas not considered

$$L_{CE_{ij}} = \left(- y_{ij} \times \log\left(p_{ij} + 10^{-15}\right)\right) \times m_{ij} \quad (1)$$

where i represents the sample index and j represents the category index. y_{ij} signifies whether sample i belongs to category j, displaying the information labeling one-hot coding vector. p_{ij} denotes the model's predicted probability for sample i belonging to category j, and m_{ij} represents a mask whether the sample i considered in the loss calculation.

Algorithm 2: GCFC Algorithm Training Process.

Input: Dataset D; ground-truth gt; Optimizer Adam, the initial learning rate is 0.0005; Dynamic Weighted Stratified Loss (DWSL); epoch K, k, $\forall k \in \{0, \ldots, K\}$; best loss.

Output: Weights file

```
The gt is divided and processed to get training_gt_onehot, training_gt_mask, val_gt_onehot, and val_gt_mask
Initialize model parameters and the best loss
X is obtained by data preprocessing from D
for k = 0 ... K do
    Using the Adam gradient to zero
    Forward propagation (Algorithm 1's input X gets the predicted map F)
    Loss value ← DWSL (F, training_gt_onehot, training_gt_mask)
    Backward propagation computes the gradient
    Adam updates the model parameters
    validation loss ← DWSL (F, val_gt_onehot, val_gt_mask)
    if validation loss < best loss
        best loss = validation loss
        Save weights file
    end if
end for
```

Address the sample imbalance by introducing sample weights w_{ij}. Sample weights are computed for each category in the training set by taking the inverse of the sample count. This ensures greater weight in the loss calculation for samples from the scarce category

$$w_{ij} = \frac{1}{1 + \sum_i \sum_j y_{ij}} \times m_{ij}. \tag{2}$$

The weighted loss $L_{w_{ij}}$ is the product of the sample weights w_{ij} and the cross-entropy loss $L_{CE_{ij}}$. This helps the model focus more on those categories with scarce samples by equalizing the contributions of different categories

$$L_{w_{ij}} = L_{CE_{ij}} \times w_{ij}. \tag{3}$$

Dynamic weights δ_{ij} refer to learnable parameters achieved through a Softmax operation applied to sample weights. They ultimately result in the computation of dynamic weighted stratified loss L_{DWSL}. This loss function can effectively deal with the sample imbalance problem, improving the model's focus on low-frequency categories by fully utilizing the dynamic weighting mechanism

$$\delta_{ij} = Softmax\left(w_{ij}\right) \tag{4}$$

$$L_{DWSL} = \sum_i \sum_j \left(\delta_{ij} \times L_{w_{ij}}\right). \tag{5}$$

Traditional multitask learning methods commonly employ a simple weighted average strategy, overlooking the distinctions in characteristics among various tasks. In contrast, the loss function proposed in the study incorporates a cross-entropy loss with a masking mechanism to assess the disparity between predicted and actual labels. Simultaneously, sample weights are assigned to different categories. The cross-entropy loss of the introduced masking mechanism is weighted for each category by multiplying it with the sample weights. The Softmax operation on the sample weights allows the model to adaptively weight each category instead of just using predefined fixed weights. Ultimately, DWSL is applied to the weighted loss function for the dynamic weights. The proposed loss function empowers the model to focus more on challenging samples during training, enhancing the overall segmentation performance.

B. Methodology for GCFC

To enhance winter wheat cross-domain extraction, the article proposed a fusion network called GCFC. Fig. 4 shows the overall GCFC network architecture. It uses a two-branch structure to exploit graph convolution and convolutional neural networks fully. The graph convolution branch has an adaptive mechanism and a graph convolution layer based on superpixels. The CNN branch comprises multiple hybrid modules from the PCSNet network. Softmax realizes the information fusion to output accurate segmentation results.

Specifically, the article will be for a 700 × 500 × 28 image input network. Initially, to enhance spectral resolution, it involves minimizing redundant information within the image through a 3 × 3 convolution operation. Subsequently, a batch normalization layer is applied to each feature map to ensure a stable distribution of features. To augment the model's nonlinear capabilities, LeakReLU [48] is chosen as the activation function instead of the commonly used ReLU [49]. Following this, the processed image is bifurcated into two branches leading into the AGCN and PCSNet networks. The AGCN network outputs a feature map F_1 (700 × 500 × 14) with dimensionality. The PCSNet network generates a feature map F_2 (700 × 500 × 14) with dimensionality. The two branches extract feature information from the image using distinct perspectives. Ultimately, the fine results are obtained by fusing the features of F_1 and F_2 with the Softmax operation. The innovations of the network are described in detail in the rest of the section. AGCN and PCSNet are introduced in the following two sections.

1) AGCN: In cross-domain scenarios, considering interdomain transfer learning is necessary due to distributional differences between domains. The model must exhibit sufficient generalization for efficient feature extraction in the target domain. This section introduces a novel graph convolutional network named AGCN. The model utilizes the superpixel method [50] for processing the input data. To better align with the data characteristics of different domains, the model's computational complexity can be varied by flexibly adjusting the size of the superpixel scale (d). AGCN incorporates an adaptive mechanism (SA) [51] applied to the adjacency matrix. This mechanism gives the model greater flexibility to capture the dependencies between nodes more accurately by assigning different weights to each

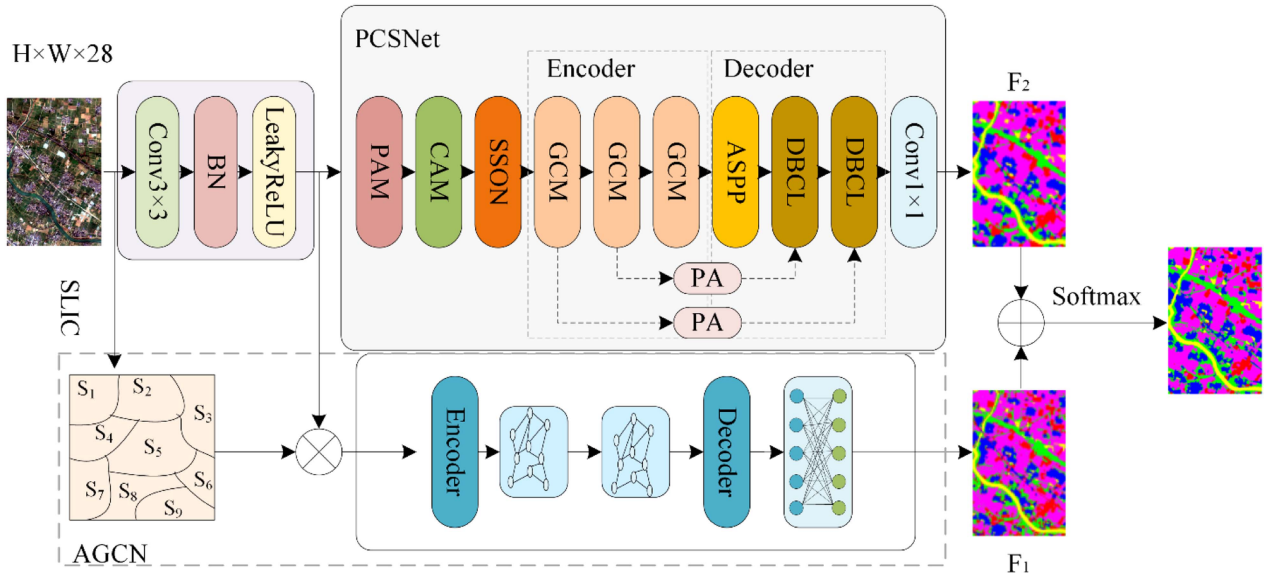


Fig. 4. Structure of GCFC.

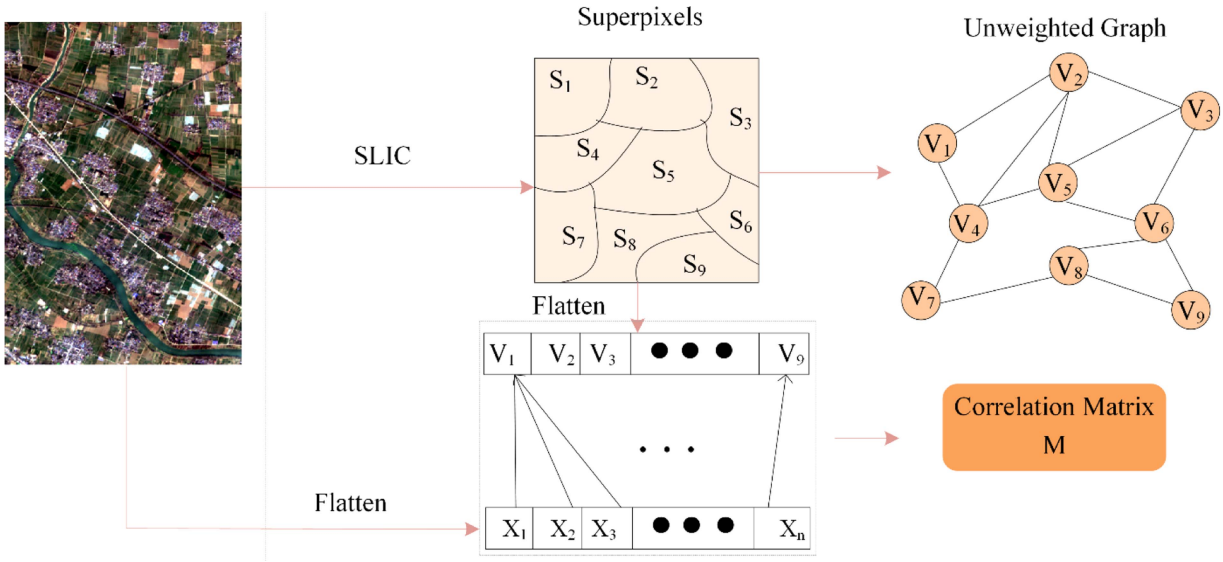


Fig. 5. Conversion process of pixel-to-superpixel data, where X_i means the ith pixel in the flattened image data and V_j denotes the average radiance of the pixels contained in the superpixel S_j.

node. It enables the model to better adapt to complex information transfer in heterogeneous graphs and improve overall performance.

Fig. 5 shows the conversion process of pixel-to-superpixel data. Initially, the image undergoes segmentation into multiple superpixels, emphasizing spatial connectivity and spectral similarity using the linear iterative clustering technique. Subsequently, the adjacency between superpixels is established, transforming the image into an undirected graph denoted as $G = (V, E)$, where V represents a node and E represents an edge. In this case, the node's features are the superpixel's average pixel features. In the representation, the pixel points in the hyperpixel set is made to correspond to the nodes in the graph

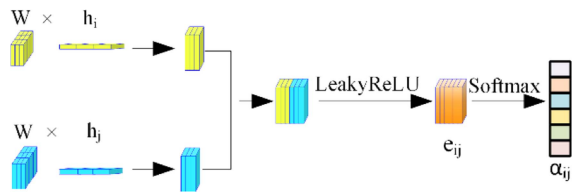


Fig. 6. Structure of SA.

using a hyperpixel-based approach. Specifically, the image is segmented into N superpixels through clustering, where N is calculated as $\lceil (H \times W)/d \rceil$. H and W represent the image's

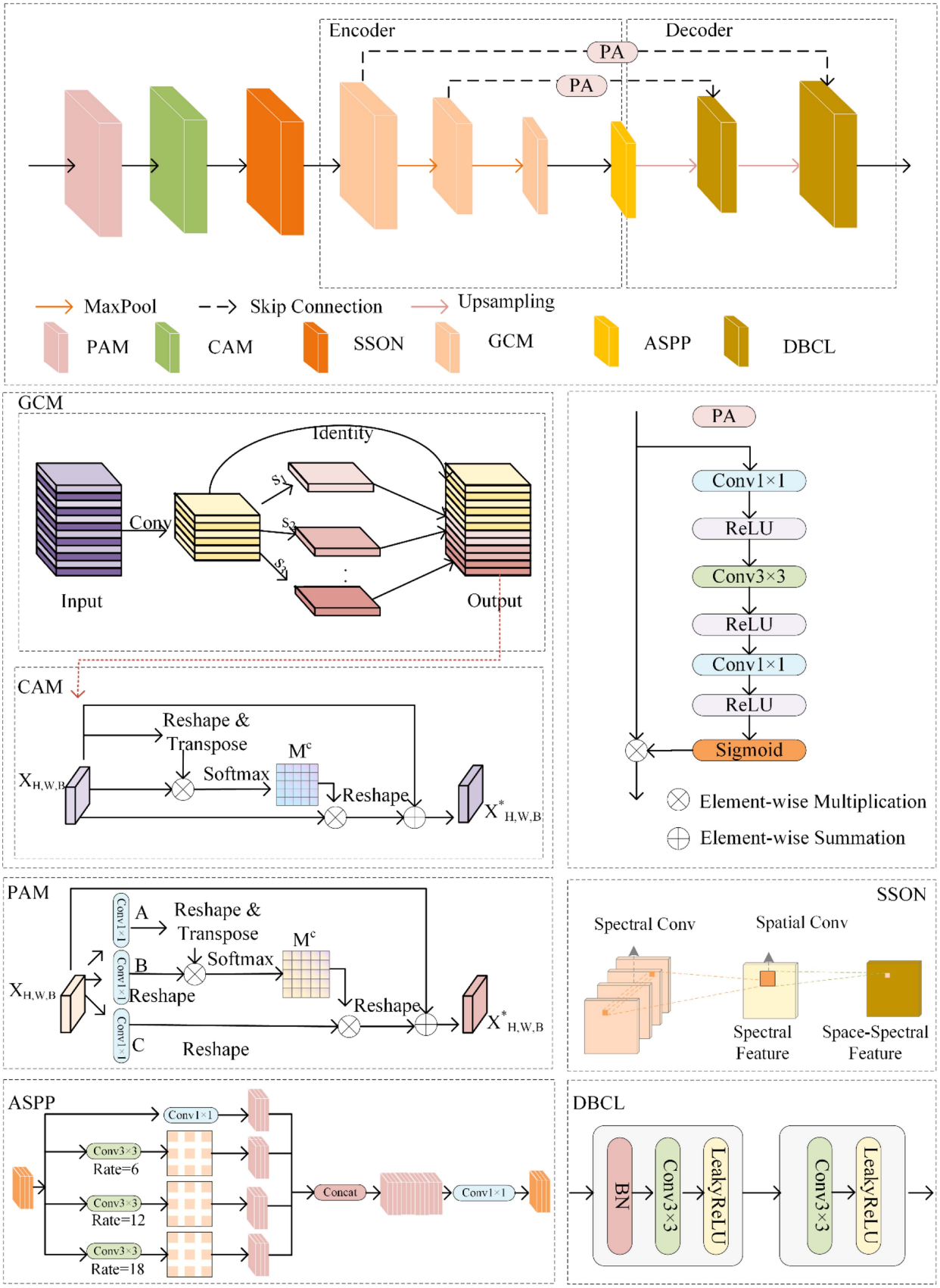


Fig. 7. Structure of PCSNet.

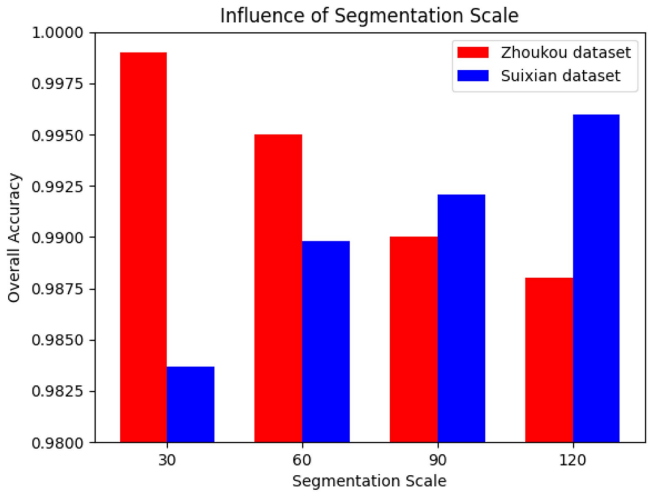


Fig. 8. Classification accuracy of GCFC at different segmentation scales.

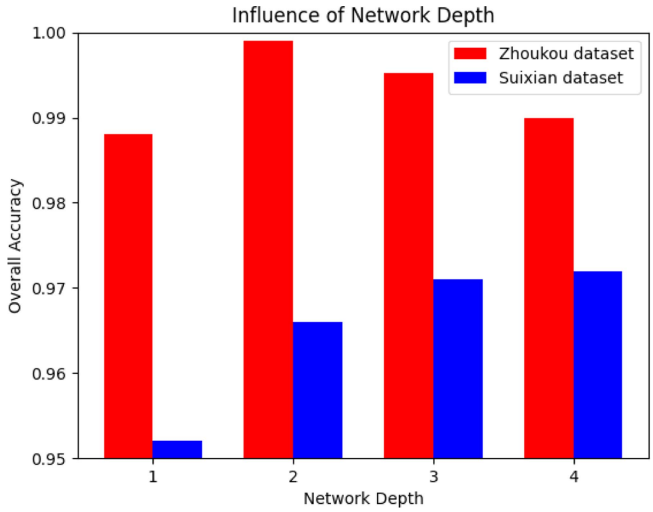


Fig. 9. Classification accuracy of GCFC at different network depths.

length and width, respectively. The parameter d $(1 \leq d)$ serves as the segmentation scale, controlling the sensory field size, range of information propagation in the graphical convolution operation, and superpixel size. Denoting $M \in R^{H \times W \times N}$, it represents the correlation matrix between pixels and superpixels

$$M_{i,j} = \begin{cases} 1, if\ \bar{X}_i \in S_j \\ 0, if\ \bar{X}_i \notin S_j \end{cases} \quad \bar{X} = Flatten\,(X) \tag{6}$$

where $Flatten(\cdot)$ denotes flattening image data along the spatial dimension, $M_{i,j}$ denotes the value of M at position (i, j), and $\bar{X}_i$ denotes the ith pixel in $\bar{X}$. The node V in the undirected graph is then reconstructed and the feature transformation is implemented

$$V = Encoder\ (X; M) = \hat{M}^T\ Flatten\,(X) \tag{7}$$

$$X^* = Decoder\ (V; M) = Reshape\,(M \times V) \tag{8}$$

where $\hat{M}$ represents the normalization of M, the $Encoder(;)$ operation encodes the graph data into the nodes V of the undirected graph based on the correlation matrix M. $Reshape(\cdot)$ indicates reshaping along the spatial dimension of the reduced data. Revert to network features X^* using the $Decoder(;)$ operation.

In contrast to other methods that compute edge weights from the original image, the article utilizes an SA for computing attentional weights on the graph edges. More precisely, the network can dynamically emphasize or attenuate connections between nodes during the learning process by calculating an attention score applied to a weighted adjacency matrix. Fig. 6 shows the SA network structure.

Assigning weights to each node in the adaptive mechanism. For each node v_i and v_j, the attention score between them, i.e., the similarity matrix e_{ij} is computed. Use the inner product form to express

$$e_{ij} = LeakyReLU\,(Concat\,(W \times h_i, W \times h_j)) \tag{9}$$

where $LeakyReLU$ is the activation function and W is the learnable weight matrix. $Concat$ denotes the splicing operation, and h_i is the feature representation of node v_i.

Subsequently, the attention scores were converted into attention weights using the Softmax function

$$\alpha_{ij} = \frac{exp\,(e_{ij})}{\sum_{k \epsilon N_i} exp\,(e_{ik})} \tag{10}$$

where N_i is the set of neighboring nodes of node v_i. Thus, the node characterization can be expressed as

$$z_{ij} = ReLU\left(\sum_{j \epsilon N_i} (\alpha_{ij} \times (W \times h_j))\right). \tag{11}$$

The adaptive mechanism calculates the similarity between nodes using the inner product form. It uses the attention mechanism to dynamically adjust the weights to get the weighted internode features z_{ij}. This approach enhances flexibility in capturing relationships between nodes, improving the overall characterization of the model.

In summary, AGCN networks get rid of their reliance on a fixed, predefined adjacency matrix by learning the weights of the relationships between each node. Adaptive weight learning enhances the model's flexibility. It enables the model to better adapt to the characteristics of diverse tasks and data. This further improves its modeling capabilities when handling graph-structured data. The model can dynamically adjust the strength of the relationship between nodes according to specific input samples, presenting higher flexibility and adaptability.

2) PCSNet: To address the edge clustering issue in AGCNs, the section introduces the PCSNet network. Its purpose is to enhance edges and minimize misclassification in winter wheat detection. The PCSNet utilizes the positional and CAMs to capture spatial and channelwise dependencies in the input feature map [52]. It also incorporates the SSON module for spectral feature extraction, considering both depth and spatial information. To preserve high-level semantic information and detailed features,

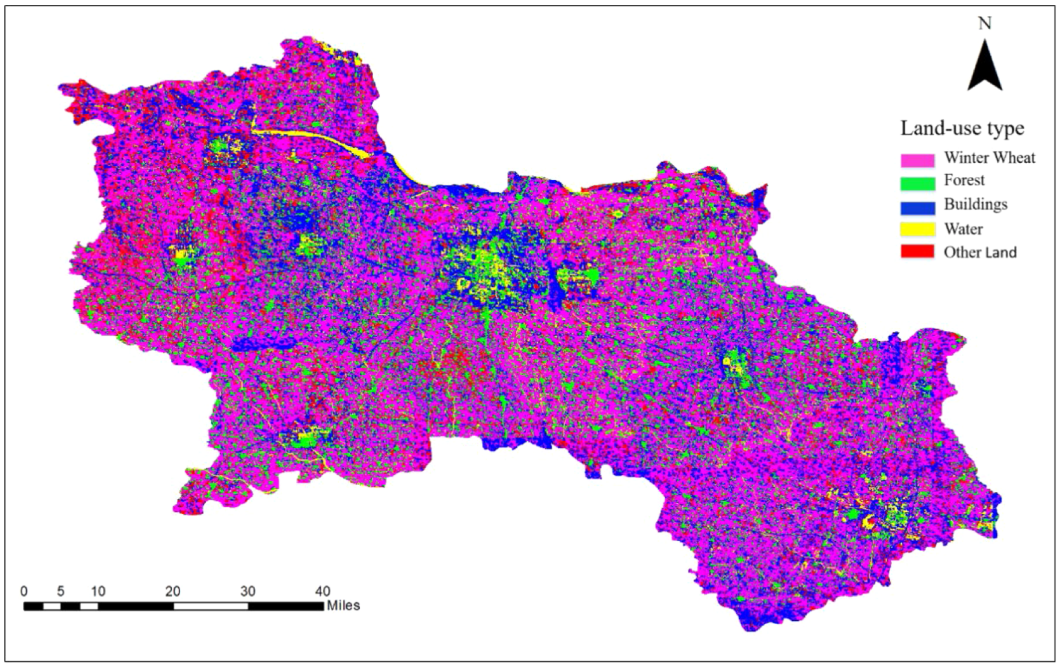


Fig. 10. Plot plan of winter wheat planting pattern in Shangqiu.

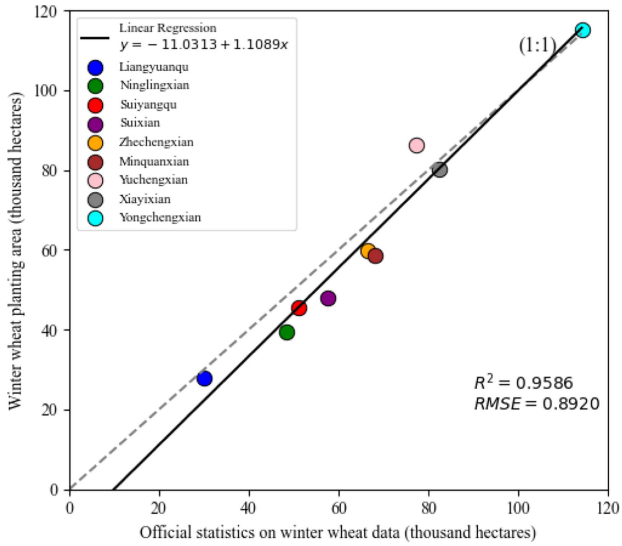


Fig. 11. Comparison results of the winter wheat statistical area and the extracted area in Shangqiu counties and districts.

the model employs a coding–decoding structure, enhancing its ability to effectively handle edge information. Fig. 7 illustrates the structure of PCSNet.

The positional attention module (PAM) is used to capture the correlation between different positions in an image. The PAM module helps increase the network's sensitivity to spatial information by modeling relationships in image space. Let the input feature map be $X \in R^{H\times W\times B}$, where B is the number of channels, H is the height, and W is the width. First, three 1×1 convolutional layers are used to reduce the number of channels in the input feature map to one eighth of the original. This process produces representations for the query and key while preserving the original number of channels for the value. The query and key representations are derived by convolving the input, and the attention score is computed through tensor multiplication. Second, the attention scores are Softmax normalized to obtain weights at each position that sum to one. The values are then weighted and summed using normalized attention mapping to output the results. Finally, residual crosstalk fuses the output with the original input and introduces a learnable parameter γ to tune the fusion. This process ultimately yields the output of the module $X_1 \in R^{H\times W\times B}$

$$PAM\ (X) = X+\gamma\times\left(\left(Softmax\left(\frac{X_{query}^{T}\times X_{key}}{\sqrt{B_{query}}}\right)\right)^{T}\times X_{value}\right) \tag{12}$$

where $X_{query} = Conv_{1\times1}\ (X)$ is the query tensor and $X_{key} = Conv_{1\times1}\ (X)$ is the key tensor. $X_{value} = Conv_{1\times1}\ (X)$ is the value tensor, γ is the learned scaling factor, and B_{query} is the dimensionality of the query tensor.

The CAM enhances the network's capability of discerning various channels. Allowing it to effectively prioritize mission-critical information by dynamically modeling the significance of each channel. Initially, the number of channels in the input feature map X_1 is adjusted by the adaptive parameter γ, and the Softmax operation is initiated. The attention map is obtained by computing the product between the channels of the input feature map. Subsequently, the attention scores are normalized via Softmax to derive weights for each channel. The output is

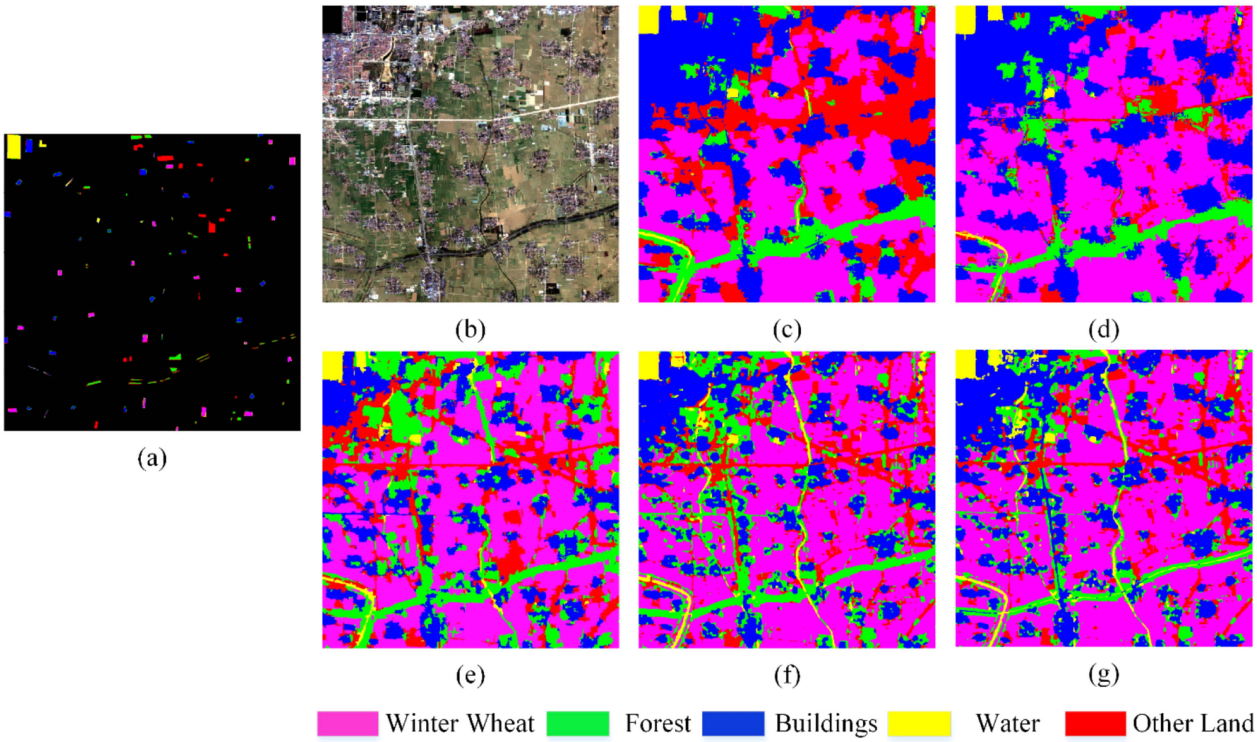


Fig. 12. Visualization results of GCFC model ablation on the Suixian dataset. (a) Ground-truth map. (b) RGB-color map. (c) GCN. (d) GCN+SA. (e) GCN+PCSNet. (f) GCN+SA+PCSNet. (g) GCFC.

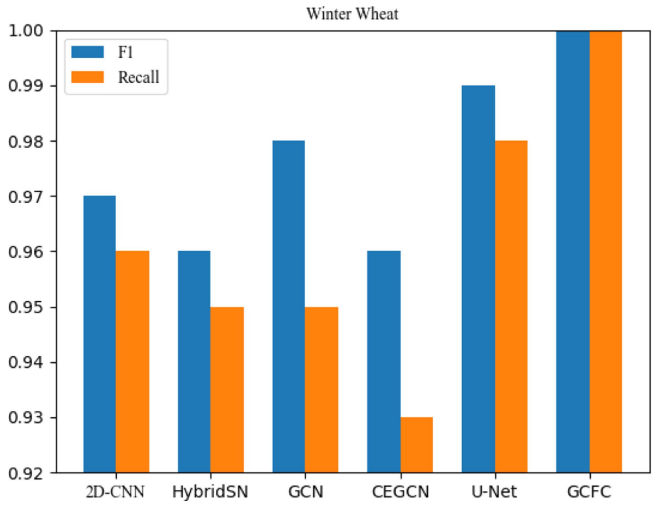


Fig. 13. Histograms of different models F1 and Recall metrics on the Zhoukou dataset.

obtained by summing up features for each channel with normalized attention weights. Ultimately, a residual concatenation akin to PAM is executed to produce the module output, denoted as $X_2 \in R^{H\times W\times B}$

$$CAM\ (X_1) = X_1\ + \gamma \times \left(Softmax\left({X_1}^T \times X_1\right) \times X_1\right). \tag{13}$$

The SSON module is primarily employed to process structural and textural information in images. It enhances the network's understanding of image structure by separating image features in spectral and spatial dimensions to extract finer-grained features. Specifically, the input feature map X_2 is normalized to expedite convergence. Subsequently, the spatial features of the data are captured. This is achieved by linearly combining the input features between channels. The operation utilizes a convolution operation with a convolution kernel size of 1×1. Nonlinear mapping using the LeakyReLU activation function. Deep convolutional operations are used to obtain independent learning for each channel's convolutional kernel and to capture unique features in the channel dimension. This operation used a convolution kernel size of 5×5, and the number of groupings was configured to match the size of the channel count. Finally, the LeakyReLU activation function is applied once more to nonlinearly map the output of the deep convolution, resulting in the output $X_3 \in R^{H\times W\times B}$

$$SSON\,(X_2) = LeakyReLU(Conv_{5\times5}(\\ LeakyReLU\,(Conv_{1\times1}\,(BN\,(X_2))))) \tag{14}$$

where $Conv_{5\times5}$ denotes a deep convolution operation with a convolution kernel size of 5×5. $Conv_{1\times1}$ denotes a point-by-point convolution operation with a convolution kernel size of 1×1.

Compared to the traditional coding layer structure, the article introduces a ghost channel in regular convolution instead of employing the typical double convolution operation. In addition, a CAM module is introduced to extract information from key channels, contributing to the formation of a GCM module. The

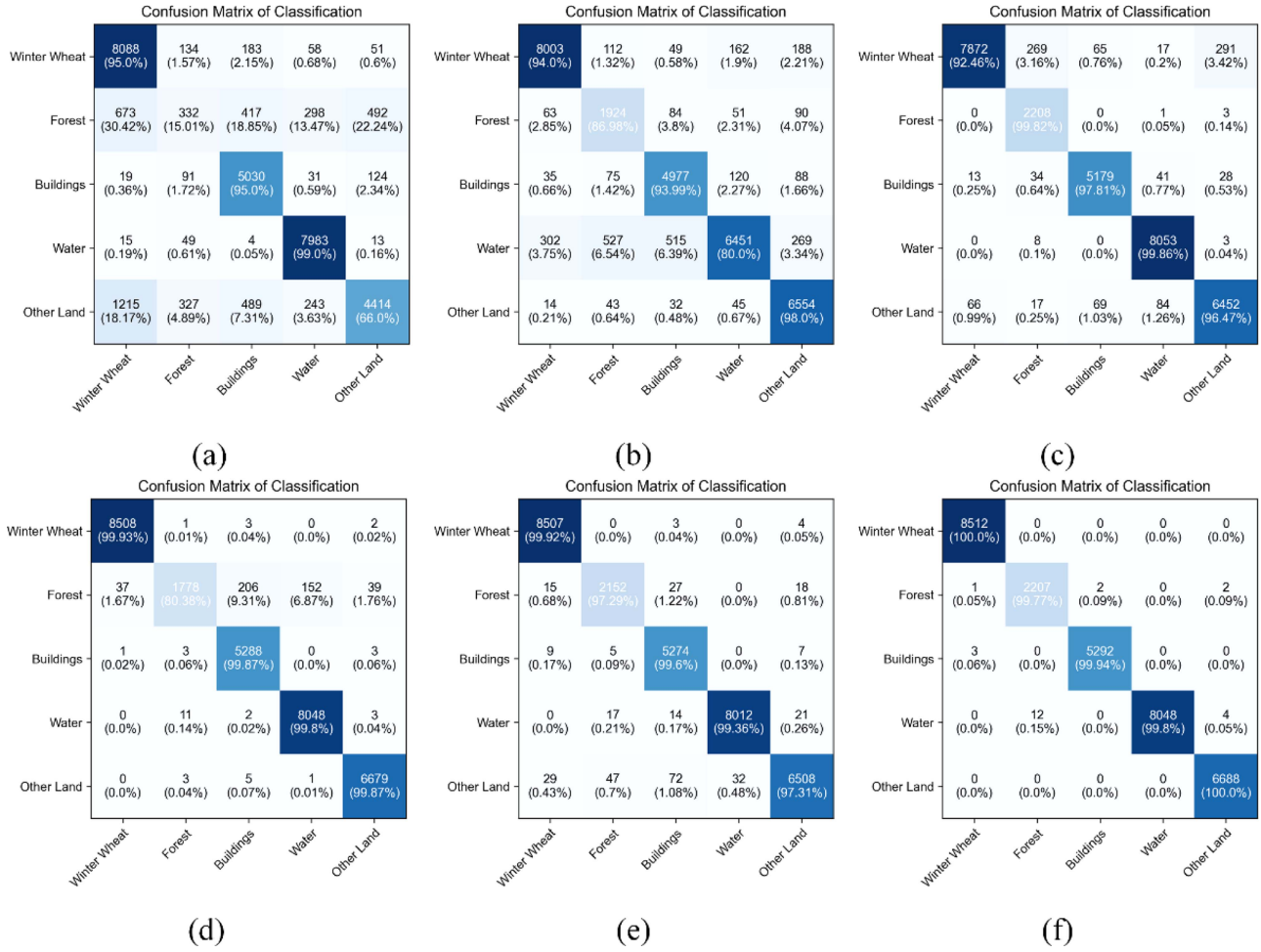


Fig. 14. Confusion matrix plots of different models on the Zhoukou dataset. (a) 2D-CNN. (b) HybridSN. (c) GCN. (d) CEGCN. (e) U-Net. (f) GCFC.

atrous spatial pyramid pooling module is introduced in the third GCM module to obtain additional feature information from various levels, following the introduction of two similar GCM modules for feature extraction. The decoding layer conducts an upsampling operation to restore the feature map size to a higher resolution. Pixel attention (PA) module and jump connection techniques are introduced to further improve the interaction between high-level and low-level features. Features are extracted through a double convolutional operation. Following similar operations, the features are then mapped to the output channel using a 1×1 convolutional layer. Finally, the final binary segmentation result is obtained by applying the Sigmoid function for activation.

Specifically, the decoding layer's PA attention uses three convolutional layers to calculate the query (Q), key (K), and value (V). As illustrated in Fig. 7, consider the input feature map denoted as $Z \in R^{H \times W \times B}$. Initially, the channel dimension of Z transforms using a 1×1 convolutional layer to reduce the number of channels to one eighth of the original. Subsequently, the second convolutional layer employs a 3×3 convolutional kernel and introduces the ReLU [49] activation function for learning more complex feature representations. Lastly, the features are further adjusted in the channel dimension using a 1×1 convolutional layer to restore the number of channels to the original size. The PA calculation can be expressed as follows:

$$PA(Z) = Conv_{1\times1}\left(ReLU\left(Conv_{3\times3}\left(ReLU\left(Conv_{1\times1}(Z)\right)\right)\right)\right) \tag{15}$$

where $Conv_{1\times1}$ denotes a convolution operation with 1×1 convolution kernel size and $Conv_{3\times3}$ denotes a convolution operation with a 3×3 convolution kernel size.

In summary, the proposed PCSNet incorporates the SSON module, the spatial CAM, and the coding–decoding structure. They effectively extract and integrate image features from different aspects to provide more accurate and rich information.

C. Experimental Setup

The hardware configuration environment for the study is as follows: Intel (R) Core (TM) i3-12100F, 16 GB memory, NVIDIA GeForce GTX1650, 64-b Windows operating system. The PyTorch [53] version is 1.10.2. After conducting several experiments, the hyperparameters were finalized as follows: the initial learning rate is 0.0005, the epoch number is 300, the number of rounds is 10, and the optimizer is Adam [54]. The

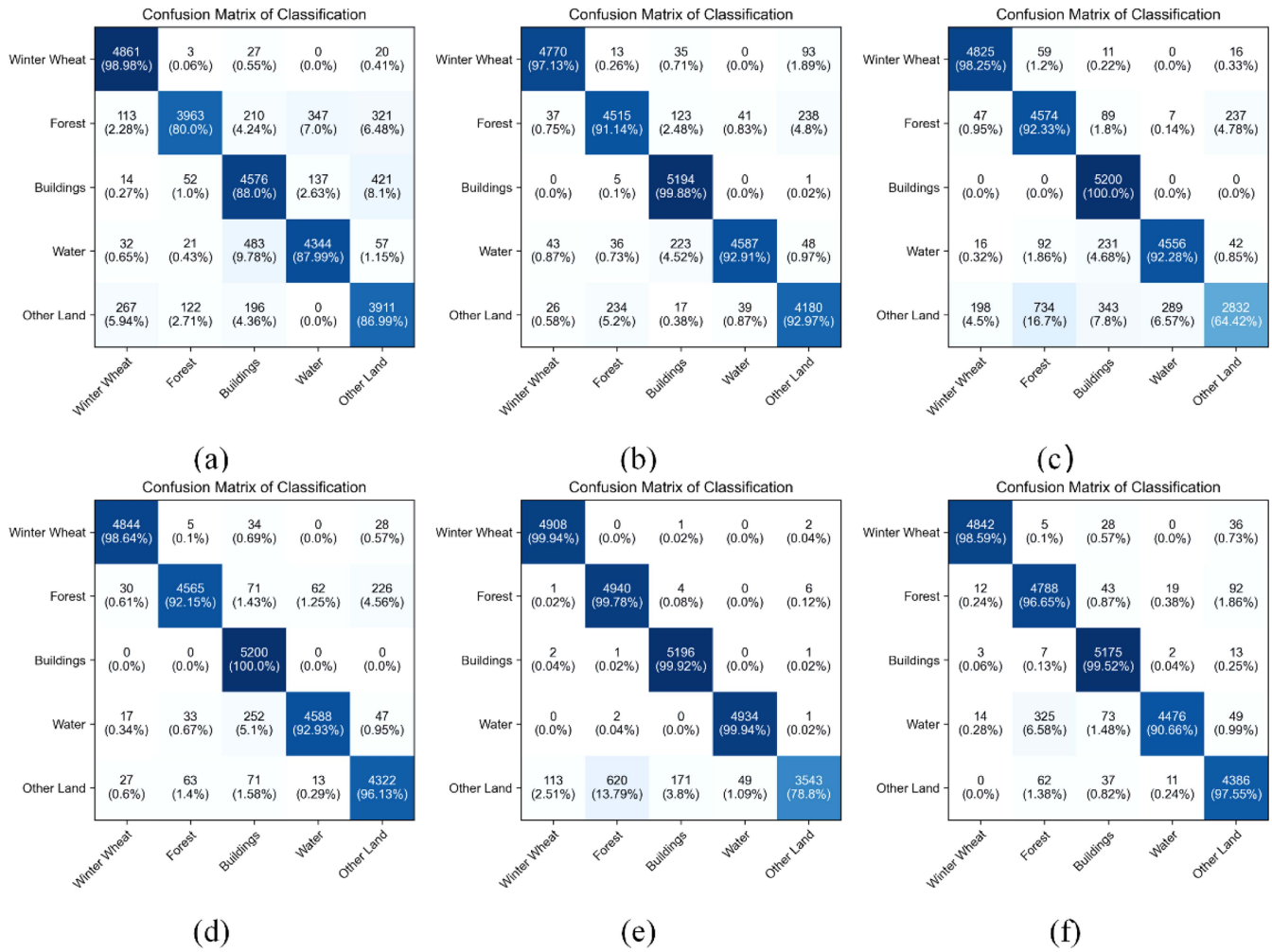


Fig. 15. Confusion matrix plots of different models on the Suixian dataset. (a) 2D-CNN. (b) HybridSN. (c) GCN. (d) CEGCN. (e) U-Net. (f) GCFC.

superpixel segmentation scale (d) is set to 30, and the network depth is set to 2. During training, the model did not use pretrained weights. To ensure that the model was adequately tuned and adapted to the dataset, we trained the model from scratch.

D. Evaluation Indicators

Extracting winter wheat from remote sensing images can be viewed as a semantic segmentation challenge. To rigorously assess the model's segmentation performance in both quantitative and qualitative aspects, we employ metrics including class accuracy (CA), overall accuracy (OA), average accuracy (AA), Kappa statistic (KPP), F1 score, and Recall

$$CA_i = \frac{c_i}{n_i} \tag{16}$$

$$OA = \frac{C}{N} = \frac{\sum_{i=1}^{k} c_i}{\sum_{i=1}^{k} n_i} \tag{17}$$

$$AA = \frac{1}{k} \sum_{i=1}^{k} \frac{c_i}{n_i} \tag{18}$$

$$KPP = \frac{p_0 - p_e}{1 - p_e} = \frac{OA - \sum_{i=1}^{k} a_i * b_i}{1 - \frac{1}{N^2} \sum_{i=1}^{k} a_i * b_i} \tag{19}$$

$$F1_i = \frac{TP_i}{TP_i + \frac{TN_i + FP_i}{2}} = \frac{2TP_i}{2TP_i + TN_i + FP_i} \tag{20}$$

$$Recall_i = \frac{TP_i}{TP_i + FN_i} \tag{21}$$

where c_i represents the number of correctly categorized samples for each category ($i = 1, 2 \ldots, k$), n_i denotes the total number of samples in each category, and k is the total number of categories, set to 5 in the article. a_i and b_i signify the actual and predicted quantities per category, respectively. N indicates the total number of samples. TP_i represents the count of positive samples accurately predicted by the model to belong to class i. FN_i signifies the count of samples from the true category i that the model erroneously predicts as belonging to other categories. FP_i denotes the count of samples genuinely belonging to category i that the model incorrectly classifies as other categories. TN_i is the count of samples where the model accurately predicts samples not belonging to class i as belonging to other categories.

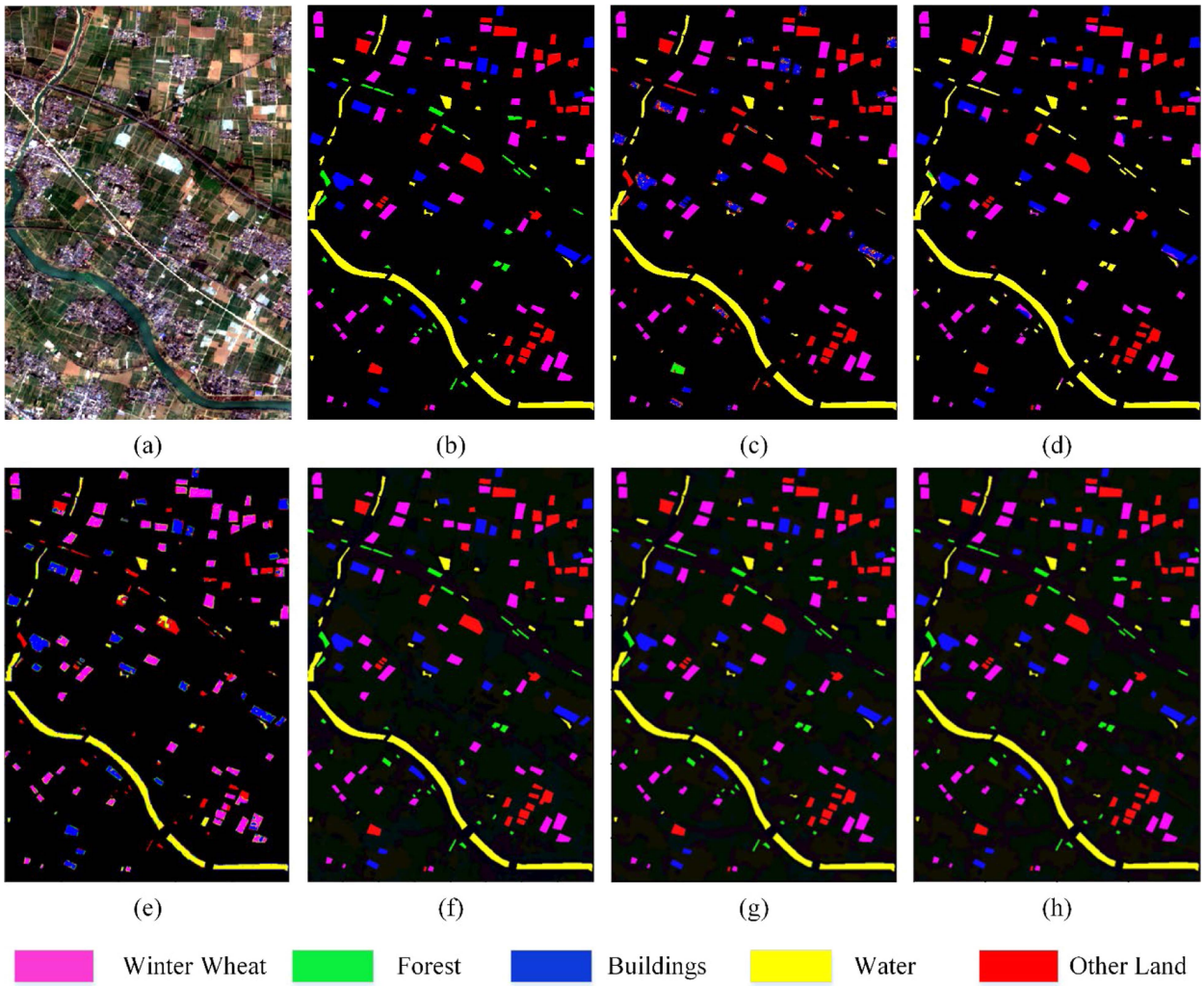


Fig. 16. Labeled regression results for different models on the Zhoukou dataset. (a) RGB-color map. (b) Ground-truth map. (c) 2D-CNN. (d) HybridSN. (e) GCN. (f) CEGCN. (g) U-Net. (h) GCFC.

To measure the performance of the model in cross-domain extraction of winter wheat area, we used the metrics WheatAcc, root mean square error (RMSE), and coefficient of determination (R^2)

$$\mathrm{WheatAcc}_l = 1 - \frac{|y_l - \hat{y}_l|}{y_l} \tag{22}$$

$$RMSE = \sqrt{\frac{\sum_{l=1}^{m} (y_l - \hat{y}_l)^2}{m}} \tag{23}$$

$$R^2 = 1 - \frac{\sum_{l=1}^{m} (y_l - \hat{y}_l)^2}{\sum_{l=1}^{m} (y_l - \bar{y}_l)^2} \tag{24}$$

where l stands for the area's index ($l = 1, 2 \ldots, m$), and m represents the number of districts and counties in Shangqiu City. y_l means the official statistics of winter wheat area in each district and county of Shangqiu City and $\hat{y}_l$ indicates the winter wheat area in each district and county of Shangqiu City extracted by the model across domains. $\bar{y}_l$ denotes the mean value of the official statistics of winter wheat area in each district and county of Shangqiu City.

IV. Results and Discussion

A. Influence of Segmentation Scale

The segmentation scale affects the GCFC segmentation results. To comprehensively investigate the impact of the segmentation scale on our model, we systematically configured scales at 30, 60, 90, and 120. Subsequently, we evaluated the segmentation results on both the Zhoukou dataset and the Suixian dataset. Fig. 8 depicts the variations in the OA index.

The choice of the segmentation scale significantly influences the size of the constructed graph. A smaller plot produced by a larger scale preserves larger objects and reduces noise. Conversely, a smaller scale results in a larger graphic with more noise and smaller objects retained. Fig. 8 shows that the model's accuracy on the Zhoukou dataset decreases as the scale increases. This indicates a prevalence of smaller objects, highlighting the need for a smaller scale to retain detailed information. In

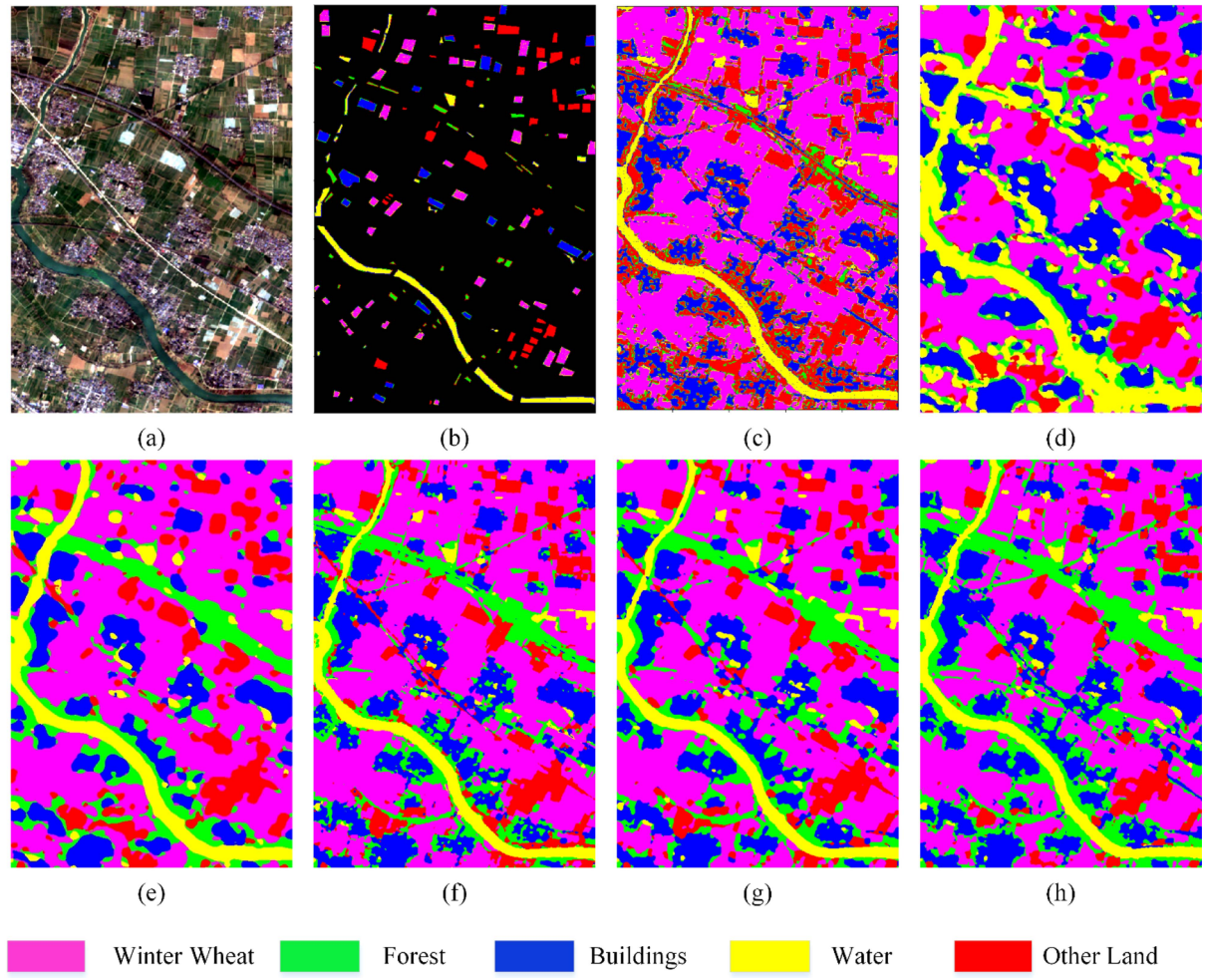


Fig. 17. Classification results of different models on the Zhoukou dataset. (a) RGB-color map. (b) Ground-truth map. (c) 2D-CNN. (d) HybridSN. (e) GCN. (f) CEGCN. (g) U-Net. (h) GCFC.

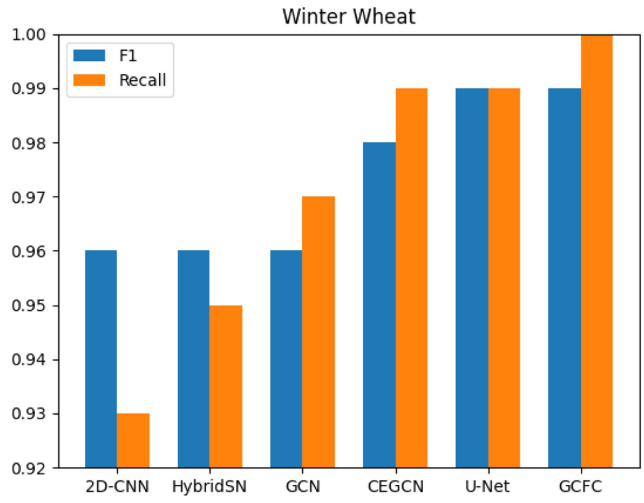


Fig. 18. Histograms of different models F1 and Recall metrics on the Suixian dataset.

contrast, the accuracy on the Suixian dataset improves with a larger scale, indicating the presence of more large-scale objects. We decided to keep things balanced and avoid generating excessively smooth classification maps by consistently setting the scale at 30 for all experiments. This aims to achieve optimal segmentation performance across diverse datasets.

B. *Influence of Network Depth*

Network depth serves as an important hyperparameter that affects neural network performance. In general, increasing the depth of the network does not always significantly improve performance due to the gradient vanishing problem. Therefore, choosing an appropriate network depth is crucial to improving the stability and performance of the model. In the section, we vary the network depths from 1 to 4 and evaluate the model's performance at each depth. Each experiment is replicated ten times, employing OA as an evaluation metric.

In Fig. 9, the model performance gradually decreases as the network depth increases on the Zhoukou dataset. This phenomenon could be attributed to the inherent smoothing traits of graph convolutional networks. As the network depth increases, these networks tend to produce more sophisticated classification maps. The Suixian dataset is predominantly composed of

TABLE III
CLASSIFICATION PERFORMANCE OF GCFC WITH DIFFERENT KERNEL SIZES ON DIFFERENT DATASETS

	(3,3)	(3,5)	(5,3)	(5,5)
Region I	95.18	96.23	97.94	**99.90**
Region II	93.41	94.37	95.07	**96.61**

(5, 5) denotes that the SSON module deep convolutional operations convolutional kernel is 5× 5 (bold means the best performance).

extensive land cover. Moderate smoothing of their classification maps helps to suppress more noise and outliers. This aligns with the findings presented in Section IV-A. However, the smoothing property of graph convolutional networks might be less effective for remote sensing images containing smaller-scale objects. Consequently, an excessively deep network model could lead to a degradation in performance. To address this, we maintained a fixed network depth of 2 in all other experiments. This decision ensures optimal segmentation performance across diverse datasets. The network depth was chosen not only to prioritize model performance but also to achieve a balance suitable for different datasets. In addition, the uniform network depth across multiple datasets guarantees the generalizability of the results.

C. Influence of Kernel Size

Considering that the kernel size is directly related to the computational cost, we set the kernel sizes of deep convolutional operations volumes in the SSON module to (3, 3), (3, 5), (5, 3), and (5, 5), respectively. As shown in Table III, the effect of kernel size on the results inspired us to conduct further experiments using (5, 5) kernel sizes. The two regions correspond to the Zhoukou dataset and the Suixian dataset, respectively. This selection achieves a balance between performance and computational cost, as the (5, 5) kernel size effectively captures both local and global features and performs well in handling both Region I and Region II. This decision not only enhances the model's performance but also minimizes unnecessary computational overhead, thereby enhancing the efficiency and scalability of the model.

D. Cross-Domain Extraction of Winter Wheat Area

The model cross-domain extracts the winter wheat area in nine districts and counties throughout the Shangqiu region through training on the Zhoukou dataset. Fig. 10 illustrates the classification results.

In the study, we focused on winter wheat sowing areas across all counties and districts of Shangqiu. A comparative analysis with official statistics was conducted to assess accuracy, as presented in Table IV. The model's accuracy was evaluated at both the Shangqiu as a whole, highlighting its practical viability. Fig. 11 illustrates the comparison between statistically reported winter wheat areas and those extracted by the model in each county and district of Shangqiu. The results affirm the method's reliability. Data points closely align along the 1:1 line, resulting in an R^2 value of 0.9586. This provides strong evidence for the feasibility and reliability of the proposed approach.

TABLE IV
COMPARISON OF WINTER WHEAT PLANTING AREAS (THOUSAND HECTARES) IN VARIOUS DISTRICTS AND COUNTIES IN SHANGQIU

Area	$\hat{y}$	y	WheatAcc(%)
Suiyangqu	45.63	51.00	89.47
Suixian	47.87	57.65	83.04
Ninglingxian	39.36	48.40	81.32
Zhechengxian	59.97	66.54	90.13
Liangyuanqu	28.00	30.13	92.93
Minquanxian	58.47	68.08	85.88
Yuchengxian	86.21	77.27	88.43
Xiayixian	80.36	82.32	97.62
Yongchengxian	115.35	114.24	99.03
Total	556.67	605.58	91.92

TABLE V
RESULTS OF ABLATION EXPERIMENTS CONDUCTED ON THE SUIXIAN DATASET

Test	Model	OA(%)	AA(%)	KPP(%)
1	GCN	87.26	87.81	83.31
2	GCN+SA	94.95	94.78	93.69
3	GCN+PCSNet	90.35	90.19	87.94
4	GCN+SA+PCSNet	96.41	96.36	95.51
5	GCFC (Ours)	**96.61**	**96.60**	**95.76**

(Bold means the best performance).

E. Ablation Experiments

In this section, we conduct ablation experiments on the GCFC model structure. All tests were carried out under identical experimental conditions. In this study, the network was trained using 70% of the Zhoukou dataset, and the model's evaluation indexes were obtained from the Suixian dataset. The evaluation metrics, including OA, AA, and KPP, are employed to assess both the segmentation performance and cross-domain capability of various models. The results are further supported through visualizations for cross-domain winter wheat extraction.

The section designs five sets of experiments to verify the performance of the proposed SA, PCSNet, and DWSL. Specifically, Test 1 uses GCN as the baseline model. Test 2 and Test 3 used Test 1 as a baseline and incorporated SA and PCSNet, respectively. In addition, Test 4 introduces both SA and PCSNet based on Test 1. Test 5 introduces SA and PCSNet simultaneously, replacing the loss function with the proposed DWSL. Table V presents the evaluation results of the various test models, and Fig. 12 shows the visualization results.

In Table V, comparing Test 1 and Test 2, the model introducing SA outperforms the baseline model in winter wheat cross-domain extraction. On the Suixian dataset, the winter wheat cross-domain extraction's OA, AA, and KPP improved by 7.69%, 6.97%, and 10.38%, respectively. With PCSNet added, the model's segmentation accuracy in winter wheat cross-domain extraction was much higher than it was in Test 1. The improvements were 3.09%, 2.38%, and 4.63% for OA, AA, and KPP, respectively. The results indicate that SA helps the

TABLE VI
COMPARATIVE EXPERIMENTAL RESULTS OF DIFFERENT MODELS ON THE ZHOUKOU DATASET

Method	CA(%)					OA	AA	KPP
	Winter Wheat	Forest	Buildings	Water	Other Land	(%)	(%)	(%)
2D-CNN [55]	95.00	15.01	95.00	99.00	66.00	85.67	72.90	81.22
HybridSN [56]	94.00	86.98	93.99	80.00	98.00	90.45	78.31	87.43
GCN [57]	92.46	**99.82**	97.81	**99.86**	96.47	97.56	97.28	96.79
CEGCN [45]	99.93	80.38	99.87	99.80	99.87	98.46	95.96	98.00
U-Net [58]	99.92	97.29	99.60	99.36	97.31	98.94	98.69	98.64
GCFC(Ours)	**100**	99.77	**99.94**	99.80	**100**	**99.92**	**99.90**	**99.90**

(Bold means the best performance).

graph structure perform adaptive learning by adapting different datasets for graph construction and enhancing the model's cross-domain capabilities. PCSNet can optimize edge information, enhancing the distinction between winter wheat and other vegetation, so it can also improve the model's cross-domain capability. Comparing Test 1, Test 2, and Test 3 to the baseline model in Test 4, the combination of SA and PCSNet performed best in terms of winter wheat segmentation and cross-domain. This suggests that the method aims to maximize both segmentation accuracy and cross-domain capability, leveraging SA's adaptability to diverse datasets for graph construction and PCSNet's optimization of edge features. In addition, Test 5's cross-domain performance is better than Test 4. Cross-domain extraction resulted in improvements of 0.2%, 0.24%, and 0.25% for OA, AA, and KPP, respectively. The findings highlight that the proposed DWSL enhances the model's ability to extract winter wheat across domains by addressing sample imbalance issues.

To facilitate a more visual comparison of cross-domain performance among different experiment groups, Fig. 12 illustrates the cross-domain classification results for each experimental set. In Fig. 12(c) and (d), Test 1's cross-domain performance is noticeably worse, and the image has multiple missegmentations in addition to being overly smooth. Compared with Test 1, Test 2 can get a more accurate classification map during feature extraction, but the effect map is still too smooth, and elements such as roads, farmland, and woodland cannot be well distinguished. Fig. 12(e) classifies a river woodland with good edge detail, but the display appears noisy, with broken, blurred, or distorted elements such as roads and woodland, hindering Test 3's ability to precisely identify and extract complete, continuous features. Test 3 exhibits limitations in feature extraction, failing to capture the full content and contextual information in the image.

Fig. 12(f) shows the winter wheat cross-domain segmentation map for Test 4. Compared to Fig. 12(g), which showed a misallocation of buildings, other cultivated land, and woodland, this contrast underscores the efficacy of the designed DWSL. In summary, the proposed method demonstrates robust segmentation and cross-domain performance in the task of winter wheat cross-domain extraction.

F. Comparison Experiments

The section evaluates the performance of the GCFC model in accurate winter wheat extraction and cross-domain applications. To assess its effectiveness, a comparison is made with five other models using the metrics of CA, OA, AA, and KPP. The models selected for comparison include 2D-CNN [55], HybridSN [56], GCN [57], CEGCN [45], and U-Net [58]. The initial comparison utilizes the Zhoukou dataset, with a 7:3 ratio for training and testing samples. Table VI presents the evaluation metrics for the various methods.

Table VI shows that the method proposed in this article performs best in terms of accurate segmentation performance in winter wheat. Notably, 2D-CNN is relatively ineffective due to underutilized training samples. HybirdSN has a low accuracy rate even with its strong expressive power. While GCN excels in global feature learning, our model demonstrates a slight underperformance in individual categories but outperforms GCN by 2.36%, 2.26%, and 3.11% in terms of OA, AA, and KPP, respectively. CEGCN enhances graph convolution by integrating CNN with feature learning for both small-scale regular regions and large-scale irregular regions. U-Net uses encoding and decoding structures to capture the details and contextual information of an image to improve segmentation accuracy. In comparison with CEGCN and U-Net, the proposed approach in the article exhibits enhanced expressive ability and effectively improves model accuracy. Compared with CEGCN and U-Net, our method improves 1.46% and 0.98% in OA, 3.94% and 1.21% in AA, and 1.90% and 1.62% in KPP, respectively.

Fig. 13 vividly illustrates our model's extraction performance for winter wheat on the Zhoukou dataset. In comparison to other models, our model outperforms them both in terms of Recall and F1 metrics. This suggests that our model has a significant advantage in the winter wheat extraction task.

Figs. 14 and 15 present confusion matrix plots detailing the segmentation results generated by the model proposed on the Zhoukou dataset and the Suixian dataset. The observed superior performance in winter wheat segmentation results can be attributed to the extensive cultivation of winter wheat in the Henan region. Winter wheat possesses simple spatial characteristics and distinctive spectral features, making it less reliant on spatial information compared to other land covers. But other land covers, like cropland and built-up areas, with similar spectra but complex spatial characteristics require more spatial information to be supported. Consequently, the accuracy of segmentation for these areas is relatively lower. This emphasizes the importance of enhancing models to effectively capture spatial information, particularly when dealing with

TABLE VII
COMPARATIVE EXPERIMENTAL RESULTS OF DIFFERENT MODELS ON THE SUIXIAN DATASET

Method	CA(%)					OA	AA	KPP
	Winter Wheat	Forest	Buildings	Water	Other Land	(%)	(%)	(%)
2D-CNN [55]	98.98	80.00	88.00	87.99	86.99	89.00	88.40	88.43
HybridSN [56]	97.13	91.14	99.88	92.91	92.97	94.89	94.81	93.61
GCN [57]	98.25	92.33	**100**	92.28	64.42	89.80	89.29	87.22
CEGCN [45]	98.64	92.15	**100**	92.93	96.13	96.11	96.13	95.14
U-Net [58]	**99.94**	**99.78**	99.92	**99.94**	78.80	96.04	95.69	95.05
GCFC(Ours)	98.59	96.65	99.52	90.66	**97.55**	**96.61**	**96.60**	**95.76**

(Bold means the best performance).

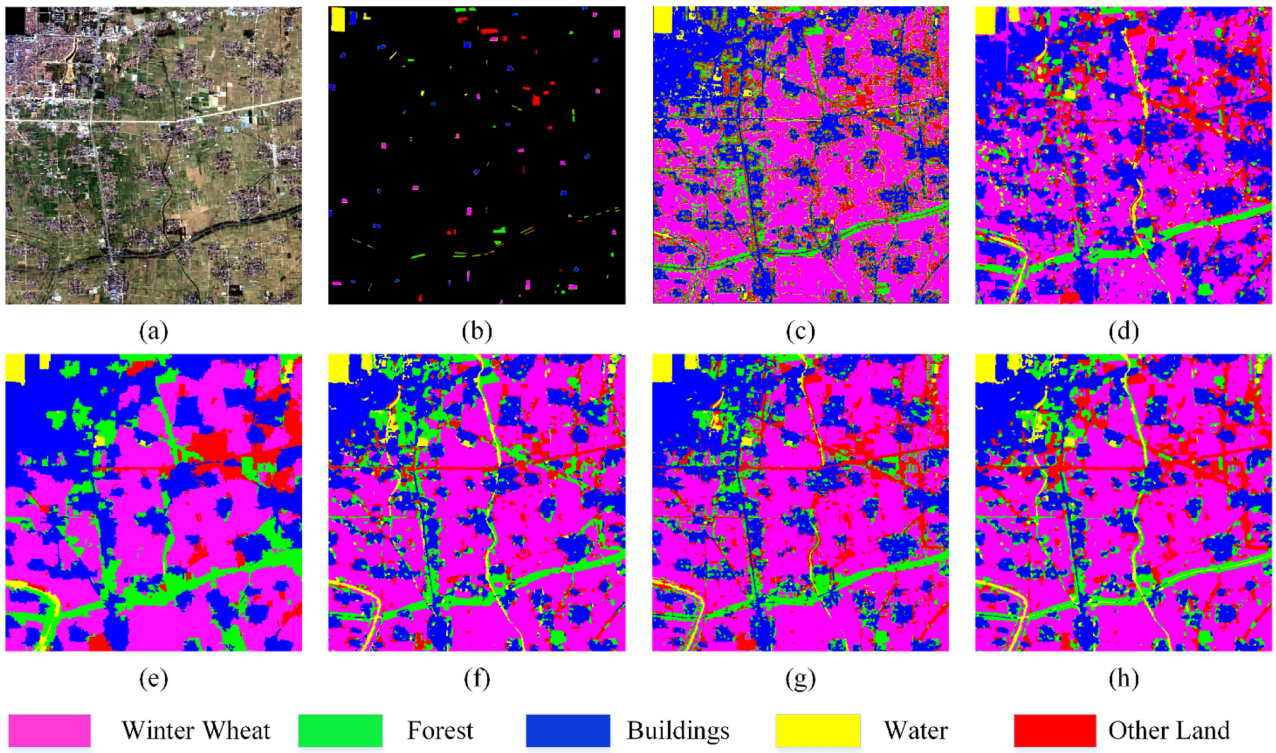


Fig. 19. Classification results of different models on the Suixian dataset. (a) RGB-color map. (b) Ground-truth map. (c) 2D-CNN. (d) HybridSN. (e) GCN. (f) CEGCN. (g) U-Net. (h) GCFC.

land cover types characterized by complexity in their spatial features.

For qualitative analysis, Fig. 16 shows the regression effect plots of six different networks on the labeled graphs on the Zhoukou dataset. The proposed GCFC model has a high regression rate. The categories Buildings and Forest appear to be characterized by missegmentation and noise from other methods. However, the proposed method yields a smoother regression effect plot while ensuring correct categorization. Fig. 17 shows a full map of the predicted ground for each of the different models in a dataset with a more complex feature distribution. The proposed model demonstrates a high degree of agreement with ground truth in terms of accuracy and edge detail. This emphasizes the effectiveness of the proposed GCFC model in capturing intricate features and maintaining OA in the segmentation results.

In summary, the approach proposed in the article exhibits exceptional precision and accuracy in classifying winter wheat and other categories. This is consistent with the findings reflected in the objective indicators detailed in Table VI.

To evaluate the cross-domain capability of the proposed method in the article, we subject the other five comparative models to classification training on the Suixian dataset. In contrast, the proposed method in the article undergoes training on the Zhoukou dataset, and zero-sample cross-domain classification is applied to the Suixian dataset. Table VII shows that the method in the article outperforms the highest CEGCN by 0.5%, 0.47%, and 0.62% in OA, AA, and KPP, respectively. Fig. 18 shows that the GCFC model performs better on both F1 and Recall metrics compared to the model trained using samples. Fig. 19 displays the GCFC model's outstanding cross-domain performance. Facing the distributed complexity of the features,

TABLE VIII
COMPARATIVE EXPERIMENTAL RESULTS OF DIFFERENT MODELS FOR CROSS-DOMAIN CLASSIFICATION ON THE SUIXIAN DATASET

Method	CA(%)					OA	AA	KPP
	Winter Wheat	Forest	Buildings	Water	Other Land	(%)	(%)	(%)
2D-CNN [55]	84.32	90.37	**99.68**	92.49	20.76	85.83	77.52	81.02
HybridSN [56]	66.21	68.76	47.79	85.74	83.95	75.36	70.49	67.60
GCN [57]	83.39	79.62	99.53	92.02	77.08	87.26	86.32	83.31
CEGCN [45]	94.49	95.34	98.93	**99.41**	76.13	95.09	92.86	93.85
U-Net [58]	93.43	87.01	99.21	92.30	72.07	92.28	88.75	89.68
GCFC(Ours)	**98.59**	**96.65**	99.52	90.66	**97.55**	**96.61**	**96.60**	**95.76**

(Bold means the best performance).

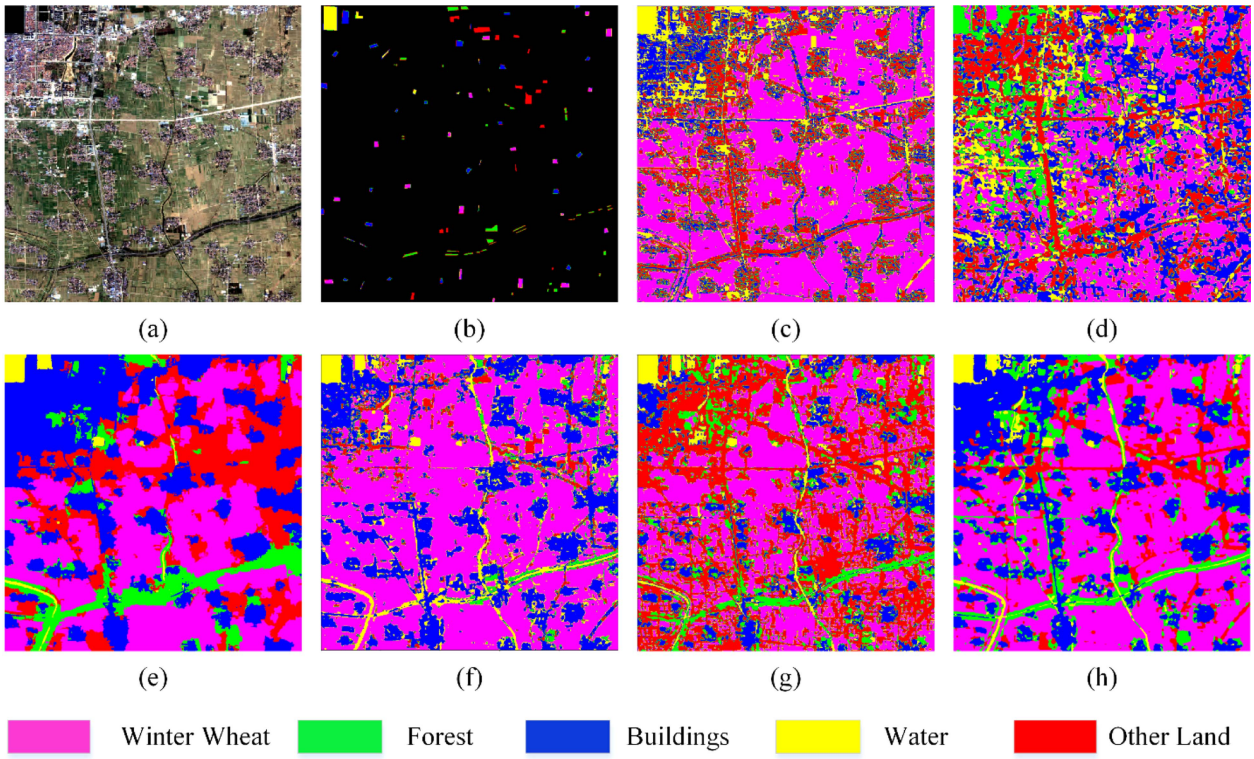


Fig. 20. Cross-domain classification results of different models on the Suixian dataset. (a) RGB-color map. (b) Ground-truth map. (c) 2D-CNN. (d) HybridSN. (e) GCN. (f) CEGCN. (g) U-Net. (h) GCFC.

it is possible to extract small areas of farmland planted with other crops from large areas of winter wheat. There is also good recognition of fine features that should be separable but are misclassified. Compared to training with samples, it even outperforms it and is closer to the distribution of real surface features.

Both the five comparison models and the model proposed in this article undergo training on the Zhoukou dataset, followed by zero-sample cross-domain validation using the Suixian dataset. This experimental design aids in providing a more comprehensive demonstration of the excellent performance exhibited by the proposed GCFC model in cross-domain feature extraction. As depicted in Table VIII, these five comparison models, although trained on one dataset, still exhibit good performance on another dataset. However, OA, AA, and KPP were lower compared to the GCFC method. Fig. 20 clearly illustrates that the resultant maps generated by the classification method proposed in this study closely resemble the real ground conditions. The GCFC model shows adaptive learning capabilities. It integrates graph structure and CNN convolution for edge feature extraction. These capabilities enable it to handle domain bias and inconsistent data distribution challenges. Such capabilities aid in addressing challenges encountered across diverse geographies and datasets, consequently enhancing the accuracy and reliability of the winter wheat extraction task.

Based on the classification results and accompanying data analyses provided in the preceding tables and figures, we draw the following conclusions: The methods presented in this article show outstanding ability to generalize, performing well in both sample-based training and zero-sample cross-domain validation, highlighting the exceptional performance of the model. This substantiates the feasibility and reliability of the proposed approach.

G. Discussion

The GCFC model proposed in this article is a two-branch graphical convolutional neural network model specifically designed for winter wheat extraction. Comparative analysis shows that the GCFC model performs best regarding segmentation accuracy, effectiveness, and cross-domain capability. In addition, the segmentation performance of the GCFC model is verified by ablation experiments with different branches, modules, and loss functions. Comparing GCFC with the other five control models shows that GCFC performs better in classification, both on real ground truth maps and 0-sample cross-domain result maps. This comparison further highlights that GCFC can significantly reduce the burden of sample data collection. In practical applications, rapid extraction of winter wheat area can provide important help for subsequent work such as yield estimation.

This article provides insight into the feasibility of applying GCFC to winter wheat extraction. The methodology of GCFC incorporates the advantages of graph convolutional networks and CNN convolutional neural networks to meet the demand of extracting winter wheat areas across domains. In addition, GCFC has a good generalization ability and strong feature characterization ability to capture the spectral and spatial features of winter wheat and other land features. The performance of GCFC was validated solely in this study for winter wheat area extraction across different domains within one municipality. GCFC exhibits strong performance in training with small samples, suggesting its potential for extraction across larger areas. In our next work, we will further increase the sample size and enrich the diversity of the dataset to extend the applicability of the model to larger regions.

V. Conclusion

The study focused on overcoming challenges encountered in traditional one-to-one training for remote sensing image classification, particularly in addressing issues related to sample distribution imbalance and feature inconsistency across various domains. We introduced a novel two-branch fusion network GCFC to overcome these obstacles, which incorporates DWSL. The loss function improves the model's focus on important categories and difficult samples by using task masks and dynamic category weights. This helps address the problem of imbalanced samples. The proposed AGCN network offers a flexible solution for the scarcity of labeled samples in practical applications. AGCN facilitates winter wheat prediction in cross-domains with no labeled data by transferring knowledge from labeled samples and adapting to diverse graph structures. Furthermore, the introduction of the PCSNet network for edge enhancement significantly improves the accuracy of cross-domain feature extraction. Experimental results support the effectiveness of our proposed method. Without labeled samples in the Shangqiu area, the accuracy of this method for extracting winter wheat was 91.92%. The comprehensive set of experiments demonstrates the robust performance of our approach in the challenging task of cross-domain extraction of winter wheat. Moreover, the method can extract area information of winter wheat across domains, providing a crucial reference for subsequent crop management and decision-making.

Future research will focus on further optimizing the cross-domain network to improve the accuracy of cross-domain extraction in winter wheat. Our goal is to extend the model's applicability over a larger region, thereby improving its generalization capabilities and enabling it to be applied more effectively to a wide range of winter wheat extraction tasks. This includes improving the feature learning capability of the model, optimizing the network structure to adapt to different geographical variations, and effectively fusing multisource remote sensing data to enhance the overall performance. Through these efforts, we expect to contribute more accurate and reliable solutions to developing the field of remote sensing extraction of winter wheat.

References

[1] W. K. J. Guofeng, "Early estimation of winter wheat planting area in Qingyang city by decision tree and pixel unmixing methods based on GF-1 satellite data," *Remote Sens. Technol. Appl.*, vol. 33, no. 1, pp. 158–167, Mar. 2018, doi: 10.11873/j.issn.1004-0323.2018.1.0158.

[2] B. Franch et al., "Improving the timeliness of winter wheat production forecast in the United States of America, Ukraine and China using MODIS data and NCAR growing degree day information," *Remote Sens. Environ.*, vol. 161, pp. 131–148, May 2015, doi: 10.1016/j.rse.2015.02.014.

[3] W. Li, H. Zhang, W. Li, and T. Ma, "Extraction of winter wheat planting area based on multi-scale fusion," *Remote Sens.*, vol. 15, no. 1, Jan. 2023, Art. no. 1, doi: 10.3390/rs15010164.

[4] C. Atzberger, "Advances in remote sensing of agriculture: Context description, existing operational monitoring systems and major information needs," *Remote Sens.*, vol. 5, no. 2, Feb. 2013, Art. no. 2, doi: 10.3390/rs5020949.

[5] J. Zhang, L. Feng, and F. Yao, "Improved maize cultivated area estimation over a large scale combining MODIS–EVI time series data and crop phenological information," *ISPRS J. Photogramm. Remote Sens.*, vol. 94, pp. 102–113, Aug. 2014, doi: 10.1016/j.isprsjprs.2014.04.023.

[6] N. Wang, X. Fan, J. Fan, and C. Yan, "Random forest winter wheat extraction algorithm based on spatial features of neighborhood samples," *Mathematics*, vol. 10, no. 13, Jan. 2022, Art. no. 13, doi: 10.3390/math10132206.

[7] C. Wei, J. Huang, L. R. Mansaray, Z. Li, W. Liu, and J. Han, "Estimation and mapping of winter oilseed rape LAI from high spatial resolution satellite data based on a hybrid method," *Remote Sens.*, vol. 9, no. 5, May 2017, Art. no. 5, doi: 10.3390/rs9050488.

[8] C. Xiuwan, "Using remote sensing and GIS to analyse land cover change and its impacts on regional sustainable development," *Int. J. Remote Sens.*, vol. 23, no. 1, pp. 107–124, Jan. 2002, doi: 10.1080/01431160010007051.

[9] Y. X. Mu, M. Q. Wu, Z. Niu, W. J. Huang, and J. Yang, "Method of remote sensing extraction of cultivated land area under complex conditions in southern region," *Remote Sens. Technol. Appl.*, vol. 35, no. 5, pp. 1127–1135, Nov. 2020, doi: 10.11873/j.issn.1004-0323.2020.5.1127.

[10] J. Gao, K. Wang, X. Y. Tian, and J. Chen, "A BP-NN based cloud detection method for FY-4 remote sensing images," *J. Infrared, Millimeter, THz Waves*, vol. 37, no. 4, pp. 477–485, Aug. 2018, doi: 10.11972/j.issn.1001-9014.2018.04.016.

[11] S. Kim, W.-J. Song, and S.-H. Kim, "Double weight-based SAR and infrared sensor fusion for automatic ground target recognition with deep learning," *Remote Sens.*, vol. 10, no. 1, Jan. 2018, Art. no. 1, doi: 10.3390/rs10010072.

[12] T. He, Y.-J. Sun, J.-D. Xu, X.-J. Wang, and C.-R. Hu, "Enhanced land use/cover classification using support vector machines and fuzzy k-means clustering algorithms," *J. Appl. Remote Sens.*, vol. 8, no. 1, May 2014, Art. no. 083636, doi: 10.1117/1.JRS.8.083636.

[13] X. Li, X. Lyu, Y. Tong, S. Li, and D. Liu, "An object-based river extraction method via optimized transductive support vector machine for multi-spectral remote-sensing images," *IEEE Access*, vol. 7, pp. 46165–46175, 2019.

[14] X. Sang et al., "Intensity and stationarity analysis of land use change based on CART algorithm," *Sci. Rep.*, vol. 9, no. 1, Aug. 2019, Art. no. 1, doi: 10.1038/s41598-019-48586-3.

[15] K. Zhang and B. Hu, "Individual urban tree species classification using very high spatial resolution airborne multi-spectral imagery using longitudinal profiles," *Remote Sens.*, vol. 4, no. 6, Jun. 2012, Art. no. 6, doi: 10.3390/rs4061741.

[16] L. F. Santos Pereira, S. Barbon, N. A. Valous, and D. F. Barbin, "Predicting the ripening of papaya fruit with digital imaging and random forests," *Comput. Electron. Agriculture*, vol. 145, pp. 76–82, Feb. 2018, doi: 10.1016/j.compag.2017.12.029.

[17] N. Wang, Q. Li, X. Du, Y. Zhang, L. Zhao, and H. Wang, "Identification of main crops based on the univariate feature selection in Subei," *Yaogan Xuebao/J. Remote Sens.*, vol. 21, no. 4, pp. 519–530, 2017, doi: 10.11834/jrs.20176373.

[18] S. Moustakidis, G. Mallinis, N. Koutsias, J. B. Theocharis, and V. Petridis, "SVM-based fuzzy decision trees for classification of high spatial resolution remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 1, pp. 149–169, Jan. 2012.

[19] S. Liu, J. Tian, S. Wang, D. Wang, T. Chi, and Y. Zhang, "Crop drought area extraction based on remote sensing time series spatial-temporal fusion vegetation index," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2019, pp. 6271–6274.

[20] W.-J. Huang et al., "Research on extraction of Zizyphus jujuba planting area in Jia county of Shaanxi," *Zhongguo Zhong Yao Za Zhi Zhongguo Zhongyao Zazhi China J. Chin. Mater. Medica*, vol. 44, no. 19, pp. 4116–4120, Oct. 2019, doi: 10.19540/j.cnki.cjcmm.20190731.103.

[21] J. Zhao, W. Gao, Z. Liu, G. Mou, L. Lu, and L. Yu, "A classification of remote sensing image based on improved compound kernels of SVM," in *Computer and Computing Technologies in Agriculture III* (IFIP Advances in Information and Communication Technology). D. Li and C. Zhao Eds., Berlin, Germany: Springer, 2010, pp. 15–20, doi: 10.1007/978-3-642-12220-0_3.

[22] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 4, pp. 834–848, Apr. 2018.

[23] F. Yu, V. Koltun, and T. Funkhouser, "Dilated residual networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 636–644.

[24] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 3431–3440.

[25] L. Chen, X. Dou, J. Peng, W. Li, B. Sun, and H. Li, "EFCNet: Ensemble full convolutional network for semantic segmentation of high-resolution remote sensing images," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, Jan. 2022, Art. no. 8011705, doi: 10.1109/LGRS.2021.3076093.

[26] D. John and C. Zhang, "An attention-based U-Net for detecting deforestation within satellite sensor imagery," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 107, Mar. 2022, Art. no. 102685, doi: 10.1016/j.jag.2022.102685.

[27] N. Audebert, B. Le Saux, and S. Lefèvre, "Semantic segmentation of earth observation data using multimodal and multi-scale deep networks," in *Proc. Asian Conf. Comput. Vis.*, 2017, pp. 180–196, doi: 10.1007/978-3-319-54181-5_12.

[28] R. Li et al., "Multiattention network for semantic segmentation of fine-resolution remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, Jan. 2022, Art. no. 5607713, doi: 10.1109/TGRS.2021.3093977.

[29] Y. Liu, J. Yao, X. Lu, M. Xia, X. Wang, and Y. Liu, "RoadNet: Learning to comprehensively analyze road networks in complex urban scenes from high-resolution remotely sensed images," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 4, pp. 2043–2056, Apr. 2019.

[30] Q. Zhao, J. Liu, Y. Li, and H. Zhang, "Semantic segmentation with attention mechanism for remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, Jan. 2022, Art. no. 5403913, doi: 10.1109/TGRS.2021.3085889.

[31] Y. Bengio, "Deep learning of representations for unsupervised and transfer learning," in *Proc. ICML Workshop Unsupervised Transfer Learn., JMLR Workshop Conf. Proc.*, 2012, pp. 17–36, Accessed: Jan. 05, 2024. [Online]. Available: https://proceedings.mlr.press/v27/bengio12a.html

[32] K. You, M. Long, Z. Cao, J. Wang, and M. I. Jordan, "Universal domain adaptation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 2715–2724.

[33] J. Tian, D. Han, M. Li, and P. Shi, "A multi-source information transfer learning method with subdomain adaptation for cross-domain fault diagnosis," *Knowl.-Based Syst.*, vol. 243, May 2022, Art. no. 108466, doi: 10.1016/j.knosys.2022.108466.

[34] X. Zhu, "Semi-supervised learning literature survey," Dept. Comput. Sci., Univ. Wisconsin-Madison, Madison, WI, USA, Tech. Rep. 1530, 2005, Accessed: Jan. 25, 2024. [Online]. Available: https://minds.wisconsin.edu/handle/1793/60444

[35] R. Pires de Lima and K. Marfurt, "Convolutional neural network for remote-sensing scene classification: Transfer learning analysis," *Remote Sens.*, vol. 12, no. 1, Jan. 2020, Art. no. 1, doi: 10.3390/rs12010086.

[36] M. Mehmood, A. Shahzad, B. Zafar, A. Shabbir, and N. Ali, "Remote sensing image classification: A comprehensive review and applications," *Math. Problems Eng.*, vol. 2022, Aug. 2022, Art. no. e5880959, doi: 10.1155/2022/5880959.

[37] Y. Zhang, X. Guo, H. Leung, and L. Li, "Cross-task and cross-domain SAR target recognition: A meta-transfer learning approach," *Pattern Recognit.*, vol. 138, Jun. 2023, Art. no. 109402, doi: 10.1016/j.patcog.2023.109402.

[38] S. Zhang, H. Tong, J. Xu, and R. Maciejewski, "Graph convolutional networks: A comprehensive review," *Comput. Soc. Netw.*, vol. 6, no. 1, Nov. 2019, Art. no. 11, doi: 10.1186/s40649-019-0069-y.

[39] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," 2016, *arXiv:1609.02907*.

[40] M. Henaff, J. Bruna, and Y. LeCun, "Deep convolutional networks on graph-structured data," 2015, *arXiv:1506.05163*.

[41] G. Zhang, X. Jia, and J. Hu, "Superpixel-based graphical model for remote sensing image mapping," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 11, pp. 5861–5871, Nov. 2015.

[42] F. Ma, F. Gao, J. Sun, H. Zhou, and A. Hussain, "Attention graph convolution network for image segmentation in big SAR imagery data," *Remote Sens.*, vol. 11, no. 21, Jan. 2019, Art. no. 21, doi: 10.3390/rs11212586.

[43] X. Qi, R. Liao, J. Jia, S. Fidler, and R. Urtasun, "3D graph neural networks for RGBD semantic segmentation," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2017, pp. 5209–5218.

[44] S. Tian, L. Kang, X. Xing, J. Tian, C. Fan, and Y. Zhang, "A relation-augmented embedded graph attention network for remote sensing object detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, Dec. 2021, Art. no. 1000718, doi: 10.1109/TGRS.2021.3073269.

[45] Q. Liu, L. Xiao, J. Yang, and Z. Wei, "CNN-enhanced graph convolutional network with pixel- and superpixel-level feature fusion for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 10, pp. 8657–8671, Oct. 2021.

[46] W. Yu, S. Wan, G. Li, J. Yang, and C. Gong, "Hyperspectral image classification with contrastive graph convolutional network," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, Jan. 2023, Art. no. 5503015, doi: 10.1109/TGRS.2023.3240721.

[47] W. Wang, S. Dou, Z. Jiang, and L. Sun, "A fast dense spectral–spatial convolution network framework for hyperspectral images classification," *Remote Sens.*, vol. 10, no. 7, Jul. 2018, Art. no. 7, doi: 10.3390/rs10071068.

[48] A. L. Maas, "Rectifier nonlinearities improve neural network acoustic models," 2013, Accessed: Feb. 02, 2024. [Online]. Available: https://www.semanticscholar.org/paper/Rectifier-Nonlinearities-Improve-Neural-Network-Maas/367f2c63a6f6a10b3b64b8729d601e69337ee3cc

[49] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks," in *Proc. 14th Int. Conf. Artif. Intell. Statist., JMLR Workshop Conf. Proc.*, 2011, pp. 315–323, Accessed: Feb. 02, 2024. [Online]. Available: https://proceedings.mlr.press/v15/glorot11a.html

[50] B. Jiang, P. Zhang, and L. Huang, "Visual object tracking by segmentation with graph convolutional network," 2020, *arXiv:2009.02523*.

[51] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Liò, and Y. Bengio, "Graph attention networks," *Statistics*, vol. 1050, pp. 10–48550, Feb. 2018.

[52] J. Fu et al., "Dual attention network for scene segmentation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Apr. 2019, pp. 3146–3154.

[53] A. Paszke et al., "PyTorch: An imperative style, high-performance deep learning library," *Adv. Neural Inform. Process. Syst.*, vol. 32, Dec. 2019.

[54] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *3rd Int. Conf. Learn. Representations*, San Diego, 2015, Jan. 2017, doi: 10.48550/arXiv.1412.6980.

[55] S. Ji, C. Zhang, A. Xu, Y. Shi, and Y. Duan, "3D convolutional neural networks for crop classification with multi-temporal remote sensing images," *Remote Sens.*, vol. 10, no. 1, Jan. 2018, Art. no. 1, doi: 10.3390/rs10010075.

[56] S. K. Roy, G. Krishna, S. R. Dubey, and B. B. Chaudhuri, "HybridSN: Exploring 3-D–2-D CNN feature hierarchy for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 2, pp. 277–281, Feb. 2020.

[57] M. Chen, Z. Wei, Z. Huang, B. Ding, and Y. Li, "Simple and deep graph convolutional networks," *PMLR*, vol. 119,pp. 1725–1735, Jul. 2020. [Online]. Available: https://proceedings.mlr.press/v119/chen20v.html

[58] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," *MICCAI*, vol. 9351, Nov. 2015, doi: 10.1007/978-3-319-24574-4_28.

Chunyang Wang received the Ph.D. degree in surveying and mapping science and technology from the School of Surveying and Land Information Engineering, Henan Polytechnic University, Jiaozuo, China, in 2015.

From 2018 to 2019, he was a Visiting Scholar with Chiba University, Chiba, Japan. He is currently an Associate Professor with the School of Computer Science and Technology, Henan Polytechnic University. His research interests include image processing, artificial intelligence and deep learning remote sensing applications, and spatial and temporal Big Data remote sensing applications.

Peipei Zhou received the B.S. degree in software engineering from Shangqiu Normal University, Shangqiu, China, in 2022. She is currently working toward the master's degree in software engineering with Henan Polytechnic University, Jiaozuo, China.

Her research interests include deep learning and remote sensing image processing.

Yudong Zhang (Senior Member, IEEE) received the Ph.D. degree in information science and engineering from Southeast University, Nanjing, China, in 2010.

He was a Postdoctoral Researcher with Columbia University, New York, NY, USA, from 2010 to 2012, and an Assistant Research Scientist with the Research Foundation for Mental Hygiene, New York, from 2012 to 2013. He was a Full Professor with Nanjing Normal University, Nanjing, from 2013 to 2017. He is currently a Professor with the University of Leicester, Leicester, U.K., and is also an External Professor with the School of Computer Science and Technology, Henan University of Technology, Zhengzhou, China. His research interests include deep learning and medical image analysis.

Junding Sun received the B.S. degree in computer applications and M.S. degree in control theory and control engineering from Henan Polytechnic University, Jiaozuo, China, in 1998 and 2001, respectively, and the Ph.D. degree in computer applications from Xidian University, Xi'an, China, in 2005.

His research interests include image processing, image retrieval, and pattern recognition.

Bibo Lu received the B.S. and M.S. degrees in computational math and Ph.D. degree in applied math from Jilin University, Changchun, China, in 2002, 2005, and 2008, respectively.

He is currently a Professor with the School of Computer Science and Technology, Henan Polytechnic University, Jiaozuo, China. His research interests include image processing, video processing, and deep learning.

Zhaozhao Xu received the M.E. degree in software engineering technology, in 2015, and the Ph.D. degree in computer science and engineering, in 2018, both from the National Pilot School of Software, Yunnan University, Kunming, China.

His research interests include data mining and feature selection.

Baishun Su received the Ph.D. degree in mechanical design and theory from China University of Mining and Technology, Beijing, China, in 2012.

His research interests include fiber grating sensing technology, wireless sensor networks, and embedded systems.

Xingwang Li (Senior Member, IEEE) received the M.Sc. degree from the University of Electronic Science and Technology of China, Chengdu, China, in 2010, and the Ph.D. degree from Beijing University of Posts and Telecommunications, Beijing, China, and 2015, both in communication and information systems.

From 2010 to 2012, he was with the Comba Telecom Ltd., Guangzhou, China, as an Engineer. From 2017 to 2018, he was a Visiting Scholar with Queen's University Belfast, Belfast, U.K. He is currently an Associate Professor with the School of Physics and Electronic Information Engineering, Henan Polytechnic University, Jiaozuo, China. His research interests include wireless communications, intelligent transport systems, artificial intelligence, and the Internet of Things.

Dr. Li was the recipient of Exemplary Reviewer for IEEE TRANSACTIONS ON COMMUNICATIONS and the *Journal of Electronics and Information Technology* in 2022. He is on the editorial board of IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, IEEE SYSTEMS JOURNAL, IEEE SENSORS JOURNAL, and *Physical Communication*. He was a Guest Editor for the Special Issue on "Integrated Sensing and Communications for 6G IoE" of the IEEE INTERNET OF THINGS JOURNAL, "Computational Intelligence and Advanced Learning for Next-Generation Industrial IoT" of IEEE TRANSACTIONS ON NETWORK SCIENCE AND ENGINEERING, and "AI-Driven Internet of Medical Things for Smart Healthcare Applications: Challenges, and Future Trends" of the IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS. He was a TPC member for the IEEE International Conference on Communications and IEEE Global Communications Conference.