Pak. J. Bot., 58(8) August 2026.

ARTICLE | DOI: http://dx.doi.org/10.30848/PJB2026-8(15)

ESTIMATION OF LARGE-SCALE TOBACCO (*NICOTIANA TABACUM* L.) PLANTING AREA BASED ON IMPROVED UNET NETWORK AND SPATIAL SAMPLING

FENGHUA HUANG[1,2,3*], XIANG HU[1,2,3], CHANGPING YANG[3], QIANYU ZHAO[4], GUOSHENG CHI[5], YANG ZHOU[5] AND ZONGPING ZHANG[5]

[1]Fujian Key Laboratory of Spatial Information Perception and Intelligent Processing, Yango University, Fuzhou 350015, China
[2]Fujian University Engineering Research Center of Spatial Data Mining and Application, Yango University, Fuzhou 350015, China
[3]College of Artificial Intelligence, Yango University, Fuzhou 350015, China
[4]Academy of Digital China (Fujian), Fuzhou University, Fuzhou 350108, China
[5]Guangze Branch of Nanping Tobacco Company, Nanping 354100, China

*Corresponding author's email: fhhuang@ygu.edu.cn

Abstract

Due to the implementation of a relatively strict tobacco monopoly policy in China, every local government controls the tobacco planting area according to the allocated indicators. In the process of large-scale tobacco cultivation, how to quickly and cost-effectively predict the area of pre-planted tobacco by local tobacco farmers every year before seedling transplanting, timely detect phenomena of over-planting or under-planting, and taking corresponding remedial measures, is an important topic that needs to be studied. This work intends to comprehensively utilize spatial sampling, remote sensing interpretation, and deep learning techniques for estimating the large-scale tobacco planting area. Firstly, it combines Support Vector Machines (SVM) and medium-resolution multispectral satellite remote sensing images to estimate the area of film-mulched farmlands across the whole experimental regions, and calculates the proportion of film-mulched farmland through the existing farmland layer. Secondly, the whole experimental regions are divided into groups based on township administrative boundaries, and the most representative townships are selected for further analysis through cluster sampling. Then, within the representative township areas, spline sampling was conducted, and existing tobacco cultivation monitoring grids were selected as the spline sampling areas. An extraction method for film-mulched tobacco fields (FMTF) based on an improved Unet network (EMFMTFIUN) is proposed. The EMFMTFIUN network and high-resolution satellite images were used to extract FMTF in the spline sampling areas, distinguishing between FMTF and non- FMTF, and then calculating the proportion of FMTF in representative townships. Finally, based on the total area of farmlands, the proportion of film-mulched farmlands in the whole experimental regions, and the proportion of FMTF in representative townships, the area of FMTF and the total tobacco planting area in the whole experimental regions were calculated, and a corresponding SaaS (Software as a Service) cloud service platform for tobacco planting area estimation is constructed. Taking Guangze County in Fujian Province as an example, experimental results compared with other 7 representative algorithm models (DeeplabV3+, Hrnet, Pspnet, SegFormer, Attention-Unet, ResUnet and Unet++), using the EMFMTFIUN model to extract FMTF in the spline sampling regions (monitoring grids) in the county from 2022 to 2025 achieves higher accuracy (with an average of over 97.0%). The average estimation accuracy of the total tobacco cultivation area in the county from 2022 to 2025 reaches 96.91%, with an average estimation time consumption of 118.8 seconds. This effectively improves the accuracy and efficiency of large-scale tobacco planting area estimation, meets the relevant requirements of local tobacco management departments, and has good promotion and application value.

Key words: Deep learning; Spatial sampling; Film-mulched tobacco field extraction; Improved unet network

Introduction

Tobacco (*Nicotiana tabacum* L.) is an important economic crop and one of the major agricultural products available for export in China, and the tobacco industry is one of the significant sources of tax revenue for the country. China carries out a strict tobacco monopoly policy, through which specialized management agencies are established in various regions to coordinate and plan the processes of tobacco planting, processing and acquisition, including allocating tobacco planting areas and other indicators to different regions. The tobacco planting process generally includes greenhouse seedling cultivation, tobacco field ridging, tobacco field mulching, tobacco seedling transplanting, fertilization and maintenance, tobacco leaves picking, and tobacco leaves curing. Tobacco field mulching can reflect the actual willingness of tobacco farmers to transplant and plant tobacco. Estimating the area of film-mulched tobacco fields (FMTF) before tobacco seedling transplanting can accurately reflect the actual area of tobacco planted. In the process of large-scale tobacco cultivation, how to quickly and cost-effectively achieve high-accuracy extraction of local FMTF every year before tobacco seedling transplanting, estimate the area of pre-planted tobacco by local tobacco farmers, timely detect the phenomena of over-planting or under-planting, and take corresponding remedial measures, is an important topic that needs to be studied.

Traditional methods for estimating the area of large-scale crops planting generally involve manually conducted

Received: 25-11-2025 | Revised: 20-02-2026 | Accepted: 24-02-2026 | Online: 14-04-2026

regional surveys and hierarchical statistical aggregation, which are not only time-consuming and labor-intensive but also prone to omissions, misreporting and falsification due to human interference (Bei *et al*., 2025; Yang *et al*., 2026). Satellite remote sensing imagery, characterized by its macroscopic, timely and objective features, has been widely applied in agricultural fields such as agricultural statistics and surveys, agricultural zoning, farmland change detection, crop growth monitoring, and yield estimation. In particular, high-resolution satellite remote sensing imagery has become one of the primary data sources for estimating the areas of various crops planting. However, relying solely on expensive high-resolution satellite imagery for estimating the area of large-scale crops planting often encounters issues such as excessively high costs of data acquisition, scarcity of eligible satellite data, and excessive data volume affecting estimation efficiency (Hu & Zhou, 2022; Pan *et al*., 2025). In recent years, some scholars have comprehensively employed spatial sampling techniques and high-resolution satellite imagery interpretation to estimate the areas of large-scale food crops (such as wheat, rice and corn) planting and achieve good results. However, this technology has rarely been applied to the field of estimating the area of large-scale tobacco planting (Ren *et al*., 2021; Xun *et al*., 2021; Zhang *et al*., 2023; Liu *et al*., 2024; Tian *et al*., 2024; Zhang *et al*., 2025).

Furthermore, although the combination of traditional machine learning algorithms (such as support vector machines, random forests, etc.) with high-resolution satellite imagery has become prevalent for estimating the planting area of crops on a large scale, shallow neural networks are always difficult to achieve high learning accuracy due to their limited learning capacity in complex scenarios such as the extraction of FMTF. In recent years, the emerging deep learning technology has unique automatic feature learning capabilities, as well as robust nonlinear complex function representation and fitting abilities. They can generate more abstract high-level feature representations through layer-by-layer learning, and effectively enhance the utilization rate of crop low-level features and optimize organizational structures, thereby improving the accuracy of crop classification and tobacco field semantic segmentation. The representative researches include: Some scholars created a small-sample tobacco dataset using Gaofen-2 satellite imagery and then extracted tobacco planting regions using the Unet network (Zhang *et al*., 2021). Compared to other machine learning methods, they achieved over 3% improvements in both accuracy and recall rate; Gaofen-1 satellite panchromatic multispectral fused imagery was used as the data source and employed three deep learning models, namely Unet, PSPNet and DeepLabv3+, for crop classification research (Tian *et al*., 2022). This provided a reference for accurately obtaining the categories, areas and spatial distribution information of the crops with complex planting structures; a relatively accurate tobacco planting area estimation method based on drone remote sensing imagery and the DeeplabV3+ deep semantic segmentation model were proposed (Fu & Huang *et al*., 2022). The original Atrous Conv structure was replaced with four classic lightweight backbone networks, resulting in a general improvement in the accuracy of tobacco field semantic segmentation. Currently, the accuracy and efficiency of these various models applied to tobacco field remote sensing image semantic segmentation still need to be further improved, and the researches directly applying deep learning to FMTF extraction and planting area estimation are still rare both domestically and internationally. Although crop classification or planting area estimation methods based on deep learning and high-resolution remote sensing imagery exhibit higher accuracy compared to traditional machine learning methods, deep learning models are complex, its training processes are time-consuming and require high computational capabilities. Achieving a balance between the accuracy and efficiency of tobacco planting area estimation is an important issue that remains to be studied.

In summary, the comprehensive adoption of spatial sampling, remote sensing interpretation and deep learning techniques is an effective approach to solve the challenge of estimating the large-scale tobacco planting area. Taking Guangze County in Nanping City as an example, this work first extracted the total area of film-mulched farmlands by combining support vector machines and medium-resolution multispectral satellite remote sensing images, and calculated the proportion of total film-mulched farmlands through the existing total farmland layer. Secondly, a combination of cluster sampling and spline sampling techniques is used to select representative townships and spline sampling regions, and an extraction method for FMTF based on an improved Unet network (EMFMTFIUN) is proposed. The EMFMTFIUN network and high-resolution satellite images are employed to extract FMTF in the spline sampling regions, thereby estimating the proportion of FMTF in representative townships. Finally, based on the total farmland area of the experimental regions, the proportion of total film-mulched farmlands, and the proportion of FMTF in representative townships, the total area of FMTF and the total tobacco planting area are calculated. A corresponding SaaS (Software as a Service) cloud service platform for tobacco planting area estimation is constructed, effectively improving the accuracy, efficiency and automation level of estimating the area of large-scale tobacco planting, and significantly reducing economic costs. The aim is to demonstrate and apply this method on a large scale throughout Fujian Province and even nationwide.

Material and Methods

Overview of experimental regions: Guangze County is under the jurisdiction of Nanping City, Fujian Province. Located in the northwest of Fujian Province and the northwest section of the Wuyi Mountain Range, the county covers a total area of approximately 2,240 square kilometers and currently administers 8 townships. The terrain within Guangze County is high in the surrounding areas and low in the center, with the highest point in the northeast. The terrain is mainly composed of hills, river valleys and plains, with an altitude ranging from 300 to 1,800 meters, and suitable space for farming is relatively small. Guangze County belongs to the mid-subtropical zone, with four relatively uniform, warm and humid

seasons. The annual average temperature is 17.6°C, and the highest temperature reaches about 39.7°C. Guangze County has the highest proportion of primary industry in its economy. The county has an existing farmland area of approximately 210,000 mu, including permanent tobacco fields covering about 147,000 mu and key protected tobacco fields covering about 60,000 mu. It is one of the important tobacco planting and production bases in Nanping City. The administrative division and distribution of permanent tobacco fields in Guangze County are shown in Fig. 1. According to investigations, the local tobacco planting cycle generally spans from December to August every year and the main category of tobacco planting is K326. From December to early February every year, it is the greenhouse seedling cultivation period, and from January to February, it is also the tobacco field film-mulching period (mulching work generally ends in mid-February every year). From late February to mid-March, it is generally the transplanting period, and from March to May, it is the fertilization and growth period. From June to August, it is the harvesting and baking stage. According to research, the local relevant departments have a significant demand for automatic estimation of the total area of tobacco planting, and it is recommended that the best time window for remote sensing monitoring of local tobacco planting conditions is between the end of film-mulching and the beginning of tobacco seedling transplanting.

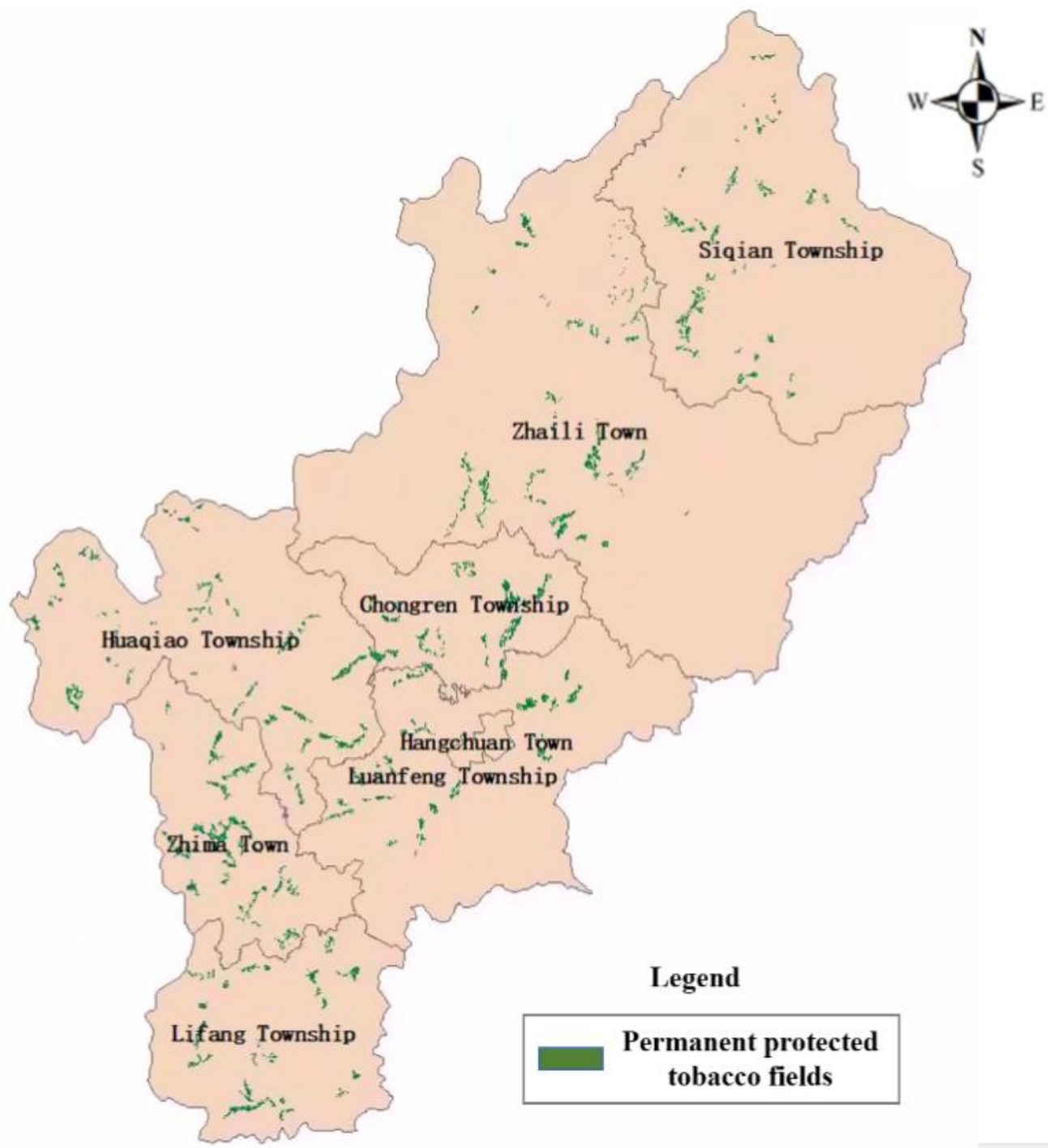

Fig. 1. Administrative divisions of Guangze County and distribution of permanent protected tobacco fields.

Data acquisition and preprocessing

The data obtained in this work mainly included the following parts:

(1) The high-resolution satellite remote sensing imagery covering the entire territory of Siqian Township and Zhima Township in Guangze County (source: Jilin-1 wide-swath 01B satellite), including one 0.5-meter resolution panchromatic band and four 2-meter resolution multispectral bands (blue, green, red and near-infrared), totaling four scenes (one scene per year from 2022 to 2025, with imaging time ranging from mid-February to early March).
(2) The multi-spectral satellite imagery of the entire Guangze County (source: Gaofen-6 satellite), encompassing one 2-meter panchromatic band and four 8-meter multi-spectral bands (blue, green, red, and near-infrared), with a total of four scenes (one scene per year from 2022 to 2025, with imaging occurring from mid-February to early March).
(3) The vector maps of permanent tobacco fields in Guangze County (including key monitoring grids, source: Guangze Branch of Nanping Tobacco Company).
(4) The vector maps of administrative divisions in Guangze County (including township and village layers, source: Tianditu website).
(5) The vector maps of land use in Guangze County (Source: Natural Resources Bureau of Guangze County, Nanping City, China).

This work first preprocessed the satellite images in the above data (including radiometric correction, topographic correction and atmospheric correction), followed by fusion of the higher-resolution panchromatic band and the lower-resolution multispectral bands using the Gram-Schmidt algorithm as needed, to generate a multispectral fused image with higher resolution.

Overall technical scheme: This work comprehensively utilizes spatial sampling, remote sensing interpretation and deep learning techniques to estimate the large-scale tobacco planting area. The research technical route is shown in Fig. 2. Taking Guangze County in Nanping City as an example, firstly, a field investigation was conducted in the experimental regions to obtain the overall crop planting structure information and the relevant medium and high spatial resolution satellite remote sensing images, and the relevant data was preprocessed (as detailed in Section 2.2 above). Secondly, combining support vector machines (SVM) and "Gaofen-6" multispectral satellite remote sensing images, an area estimation model of film-mulched farmlands based on SVM (AEMFMFSVM) was built and optimized. The total area of film-mulched farmlands was estimated through AEMFMFSVM, and then combined with the existing farmland layer of the whole experimental regions to calculate the proportion of total film-mulched farmlands. At the same time, the whole experimental regions were divided according to township boundaries, and representative groups were selected through cluster sampling. Within these representative groups, spline sampling is conducted, and the existing tobacco planting monitoring grids were selected as the spline sampling survey regions. An extraction model of FMTF based on improved Unet (EMFMTFIUN) was built and optimized, and the EMFMTFIUN network and "Jilin-1" high-resolution satellite images were used to extract FMTF in the spline survey regions, so as to distinguish between FMTF and non-FMTF and then calculated the proportion of FMTF in representative groups. Furthermore, based on the total farmland area of the experimental regions, the proportion of total film-mulched farmlands and the proportion of the FMTF in representative groups, the total area of FMTF was calculated, and the total tobacco planting area was estimated by multiplying correction factor h. which is mathematically defined as formula (1):

$$h = S_1/(S_1 - S_2) \quad \text{..... (1)}$$

where S_1 represents the total area of the tobacco fields, and S_2 represents the total area of the field ridges and the gaps between the film-mulched blocks. The values of S_1 and S_2 depend on the internal structure and field measured average value of the tobacco fields in the experimental area. Finally, the analysis of the total tobacco planting area estimation results was conducted, and a large-scale tobacco planting area estimation SaaS cloud service platform was constructed to promote project outcomes.

AEMFMFSVM model building: Firstly, feature extraction was conducted on Gaofen-6 fused imagery of the whole experimental regions, calculating the features of the Soil Adjusted Vegetation Index (SAVI), Normalized Vegetation Index (NDVI) and texture attributes (“Energy”, “Contrast”, “Correlation” and “Homogeneity”) respectively. Then a combined feature vector containing 10 attributes was formed by combining the above features and the original four reflectivity feature bands. Among the combined vector, “Energy” feature was primarily used to measure the uniformity of the textures in the film-mulched farmland images, “Contrast” feature was primarily used to measure the clarity or depth of the textures in the film-mulched farmland images, “Correlation” feature was primarily used to measure the linear dependency of pixel pairs in the film-mulched farmland image textures, and “Homogeneity” feature was primarily used to describe the degree of local variation in the film-mulched farmland image textures. The aforementioned texture features can all be calculated through the Gray-Level Co-occurrence Matrix (GLCM). Secondly, based on the aforementioned combined feature vector and a small number of labeled non-film-mulched farmland samples, an SVM-based non-film-mulched farmland recognition model was trained, and the parameters of the model were fine-tuned and optimized using the ant colony algorithm. Finally, the optimized model was used to extract non-film-mulched farmland across the whole experimental regions, and combined with the available farmland layer for masking operations to obtain the film-mulched farmland layer, and estimated the area of film-mulched farmlands in the whole experimental regions. The area of film-mulched farmlands was divided by the total area of farmland in the whole experimental regions yields the proportion of film-mulched farmland. The construction process of the AEMFMFSVM model is shown in Fig. 3.

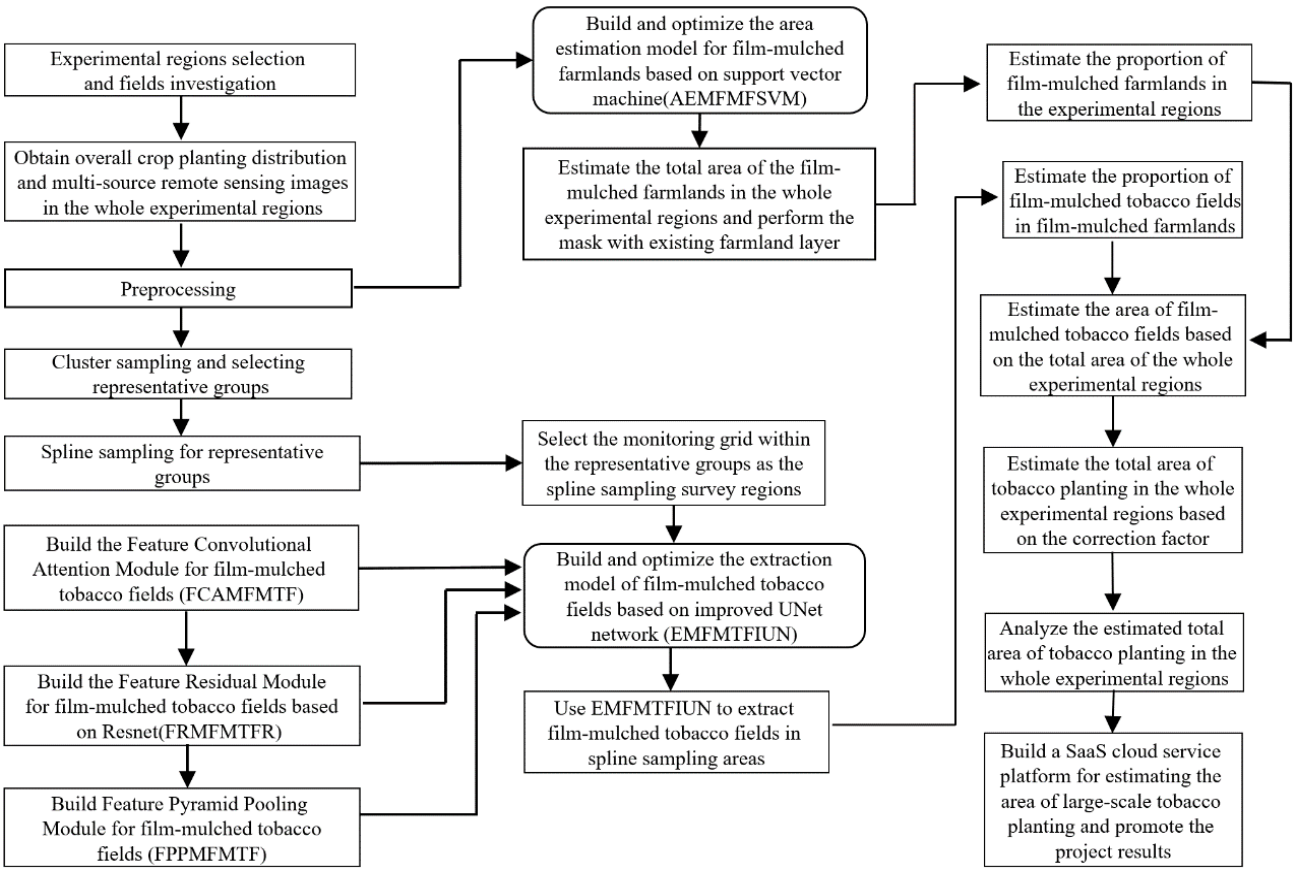

Fig. 2. Overall technical flowchart of this work.

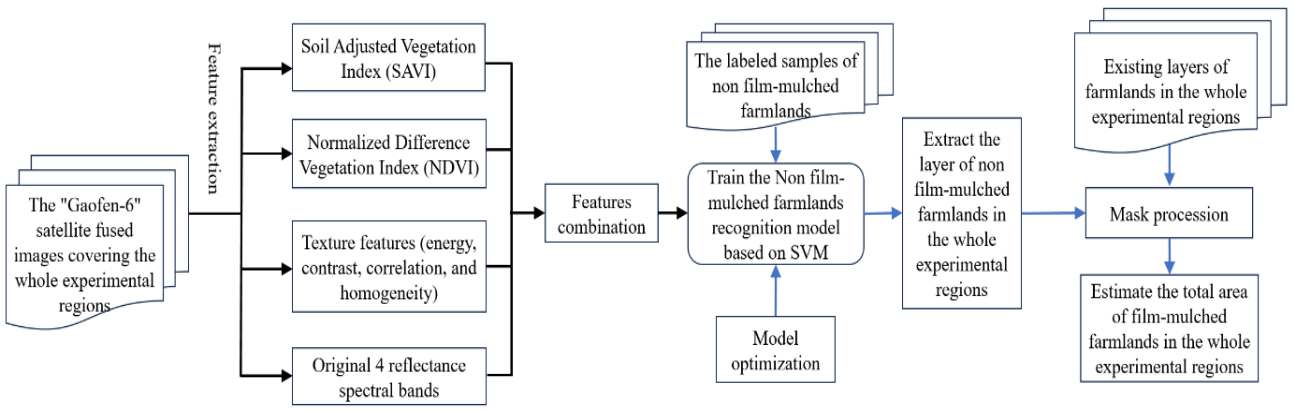

Fig. 3. Construction process of the AEMFMFSVM model.

It is well known that support vector machine (SVM) and random forest (RF) are both excellent algorithms for extracting film mulched tobacco fields based on high-resolution remote sensing images, but they also have their own different characteristics. Compared with RF, SVM is the better choice when the number of labeled samples is small, the dimension of classification features is relatively high (such as text and image) and the generalization ability requirement of the model is good. In this work, the number of labeled sample images of film mulched tobacco fields is limited, the classification feature dimension is relatively higher and the model is expected to have good generalization ability. Obviously, it is more appropriate to use SVM to estimate the area of film-mulched farmlands based on the "Gaofen-6" multispectral satellite remote sensing images.

Design of spatial sampling scheme: This work adopts a combination of cluster sampling and spline sampling for spatial sampling, aiming to effectively reduce the cost of estimating the large-scale tobacco planting area while maintaining high accuracy.

1. Cluster sampling scheme: Cluster sampling refers to a sampling method that first divides the whole samples into several non-overlapping and non-duplicative clusters, then randomly selects several representative clusters in the above clusters, and finally conducts a comprehensive survey on all units within the selected clusters. The smaller the differences between clusters and the greater the differences within clusters, the higher the sampling efficiency. In this work, based on the administrative division structure, the whole

experimental regions were divided into seven clusters according to township boundaries (Hangchuan Town has no tobacco planting records). The differences in topography and tobacco field distribution structure between clusters were relatively small, while the aforementioned differences within each cluster were significant. In addition, as for the average proportion of permanent tobacco fields in the past three years (2022-2024), Siqian Township in the northeast corner and Zhima Town in the southwest corner on the diagonal were the most stable and closest to the whole experimental regions. Therefore, Siqian Township and Zhima Township were selected as the selected clusters for cluster sampling in this work.

2. Spline sampling scheme: Through the aforementioned cluster sampling, the scope of the required in-depth investigation has been significantly narrowed. However, the spatial areas of the two selected clusters are still relatively large, with the total areas of Zhima Townshhip and Siqian Township reaching 160.01 and 418.94 square kilometers, respectively. Moreover, the tobacco planting areas were concentrated around roads and in the intermountain plains, and the historical total area proportion was generally less than 2% of the total area of the two places. Therefore, in order to improve survey efficiency and reduce information omissions, it was necessary to perform another sampling on the two selected clusters from the cluster sampling. Spline sampling was a typical stratified two-stage sampling method. To improve survey accuracy and reduce costs, the existing tobacco planting monitoring grids (key monitored polygonal areas designated by the local tobacco management department) in the two selected clusters were used as a framework in the first stage of sampling. Taking Siqian Township, one of the selected clusters, as an example, the distribution of tobacco planting monitoring grids in 2024 is shown in Fig. 4. In the second stage, the traditional field survey method was changed, and high-resolution satellite imagery covering the two selected clusters and the EMFMTFIUN model were used to extract FMTF from all monitoring grid areas, in order to calculate the area of FMTF in the selected clusters from the cluster sampling. Compared with the area of the relevant film-mulched farmlands, the proportion of FMTF in the two selected clusters can be calculated.

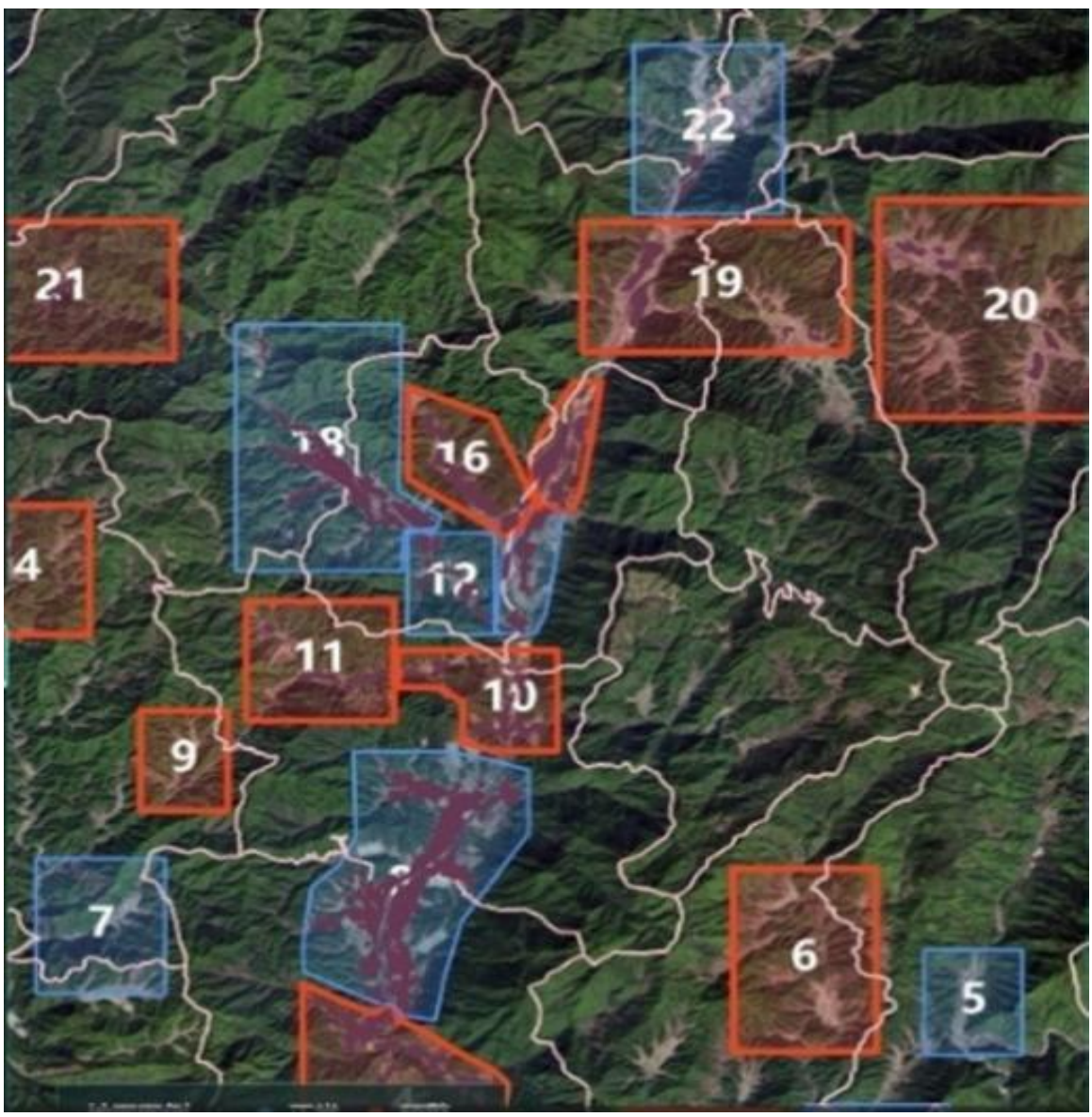

Fig. 4. Distribution of tobacco planting monitoring grids in Siqian Township in 2024.

According to the feature combination structure of EMFMTFIUN model and relevant experimental results, the extraction accuracy of EMFMTFIUN model is not affected by the grid shape. Because in the feature extraction process of the EMFMTFIUN model, the size of a convolution window is much smaller than that of a grid itself, and the remote sensing images within each grid can be segmented into a large number of sub images for parallel processing. Except for a small number of sub images at the edges of the grids, other sub images inside the grids are not affected by the grid shape during the convolution operation. Obviously, the grid shape has no significant impact on the training results of the EMFMTFIUN model.

EMFMTFIUN model building: Existing research has demonstrated that Unet and its improved versions possess advantages such as simpler image semantics, relatively fixed structure, fast convergence rate and high accuracy (Gu *et al*., 2020; Shen *et al*., 2020; Zhang *et al*., 2021; Jian *et al*., 2023).They are more suitable for intelligent semantic segmentation of large-scale remote sensing images and are currently a type of deep learning algorithm model that is more applicable to the identification of FMTF based on high-resolution remote sensing images. This work proposes an extraction model for FMTF based on an improved Unet network (EMFMTFIUN). It incorporates the Feature Convolutional Attention Module for FMTF (FCAMFMTF), the Feature Residual Module for FMTF based on Resnet (FRMFMTFR) and the Feature Pyramid Pooling Module for FMTF (FPPMFMTF) on the basis of the traditional Unet model, and combines the above modules to generate a new improved Unet network model. EMFMTFIUN simultaneously focuses on semantic information and detailed information, which can compensate for the deficiency of Unet in detecting small targets and minimize the influence of the morphology of FMTF themselves and other vegetation. The structure of the EMFMTFIUN network model is shown in Fig. 5.

As can be seen from Fig. 5, EMFMTFIUN, similar to the traditional Unet network, adopts a symmetric encoding and decoding network structure, performing four upsampling and downsampling operations in the left and right halves, respectively. In the left half of EMFMTFIUN, each encoder mainly consists of a FRMFMTFR residual module, a pooling layer and a downsampling layer. Downsampling reduces the spatial dimension of the data, allowing for the acquisition of high-level semantic features from shallow to deep layers. The feature map is gradually downsampled from 256×256×3 to 16×16×512, filtering out noise and unimportant high-frequency information. At the final stage of the encoding path, the FMTF feature pyramid pooling module (FPPMFMTF) is connected. The FPPMFMTF module acts as a bridge in the network, expanding the filter's field view. The right half of the EMFMTFIUN network constitutes the decoding path, mainly composed of bottom-up deconvolution layers that upsample the extracted deep-level features to the desired size using nearest neighbors. Each decoder in the right half is mainly composed of a FCAMFMTF spatial attention module, a FRMFMTFR residual module and an upsampling layer. Its input includes two parts: the deep-level features obtained from the previous deconvolution layer and the shallow-level features extracted from the corresponding layer in the left half of the network and processed by the FCAMFMTF channel attention module. The features extracted from the corresponding layers in the left and right halves are fused using a residual skip path, gradually restoring the detailed features and spatial dimensions of the FMTF.

Compared with other Unet network variants (such as Attention-Unet, ResUnet and Unet++), EMFMTFIUN absorbs their advantages and combines FCAMFMTF, FRMFMTFR and FPPMFMTF to generate a new improved Unet network model. In addition, the FCAMFMTF channel attention module is used as the skip connection module to prevent over-fitting and improve the accuracy of EMFMTFIUN model. FCAMFMTF enables EMFMTFIUN model to focus on the region of interest adaptively, and suppress irrelevant regions such as noise and background; FRMFMTFR uses the idea of ResNet for reference to solve the problems of gradient disappearance and information loss in deep network; FPPMFMTF can generate multi-scale feature information of FMTF and improve the feature extraction effect. Therefore, EMFMTFIUN has better comprehensive performance than Attention-Unet,ResUnet,Unet++ for the FMTF extraction.

FCAMFMTF module construction: The primary task of constructing FCAMFMTF involves designing the structures of the Channel Attention and Spatial Attention modules for FMTF features and assigning different weights to them. The general structures of the two modules are illustrated in Figs. 6 and 7 below.

The structure and working process of the channel attention module are illustrated in Fig. 6. Given a feature map F, the channel attention module first compresses the feature map F using global max pooling and global average pooling. The compressed features are then input into a multi-layer perceptron (MLP) for dimensionality expansion and reduction operations. Subsequently, the two one-dimensional vectors output by the MLP are summed, and the channel attention coefficient Mc is obtained through the sigmoid function.

The FCAMFMTF channel attention module multiplies the input feature map F by the channel attention coefficient Mc to obtain a new feature F'. Then, F' is input into the spatial attention module, where average pooling and max pooling are used to obtain two different two-dimensional vectors. These two feature descriptions are then merged, and a convolution operation is applied to generate the spatial attention coefficient Ms. The structure and working process of the FCAMFMTF spatial attention module are illustrated in Fig. 7.

FRMFMTFR module construction: FRMFMTFR is a combination of a convolutional layer, Batch Normalization (BN) and a Rectified Linear Unit (ReLU) activation function. Utilizing the BN layer for normalization can accelerate model convergence and enhance the model's generalization performance. Employing the ReLU activation function can eliminate the vanishing gradient problem in deep networks and prevent underfitting. The structure and working process of the FRMFMTFR module are illustrated in Fig. 8.

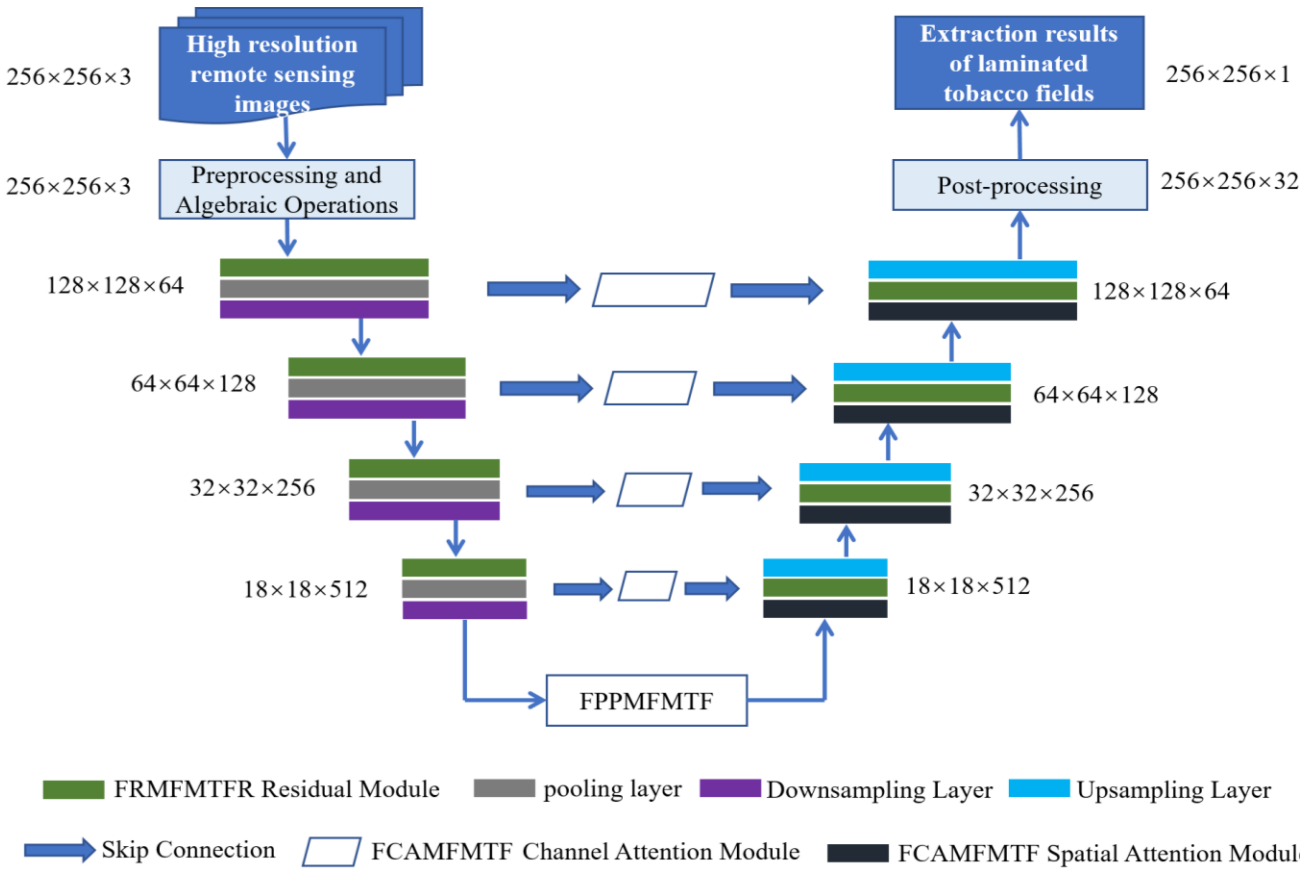

Fig. 5. EMFMTFIUN network model structure.

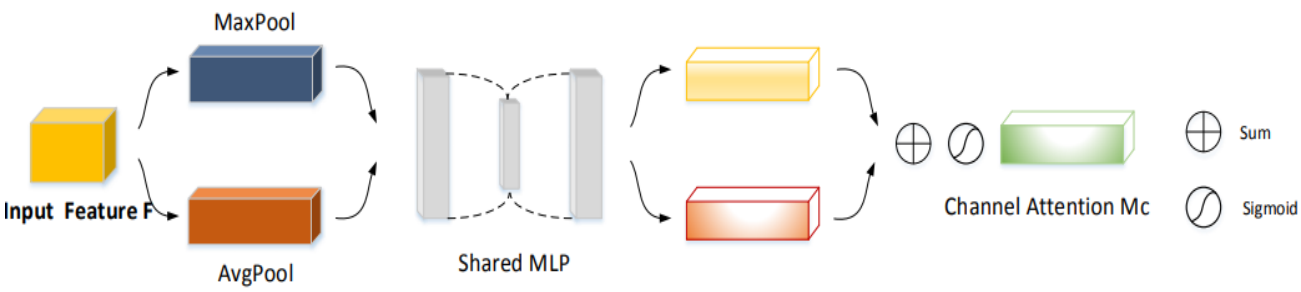

Fig. 6. Structure and working process of the FCAMFMTF channel attention module.

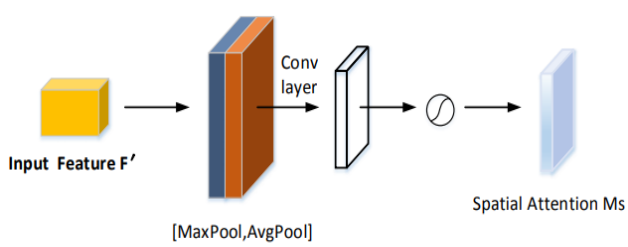

Fig. 7. Structure and working process of the spatial attention module in FCAMFMTF.

FPPMFMTF module construction: The FPPMFMTF proposed in this work consisted of a set of pooling modules with different scales, which could perform multiple pooling operations on the input feature maps at different scales to generate multiple feature maps of different scales. Subsequently, a 1×1 convolution operation can be utilized to reduce the number of channels in the feature maps, and the pooling results at different scales can be upsampled to the size of the input image. Finally, the features on different channels are concatenated to generate multi-scale FMTF feature information. The structure and working process of the FPPMFMTF module are illustrated in Fig. 9.

Estimation and analysis of large-scale FMTF area: The total farmland area was calculated based on the existing farmland layer of the whole experimental regions, and then multiplied by the proportion of film-mulched farmlands within the whole experimental regions, as well as the proportion of FMTF within the whole film-mulched farmlands obtained above, so as to calculate the total area of FMTF in the whole experimental regions. Finally, a correction coefficient was applied for fine-tuning to estimate the total tobacco planting area. Since the FMTF patches extracted by the aforementioned EMFMTFIUN model included parts of edge field ridges or small field paths, according to relevant experience and estimation results from the local agricultural department, the area proportion of these ridges and paths is approximately 3%. Therefore, the value of the correction coefficient in this work is set at 0.97. After estimating the tobacco planting area in the whole

experimental regions, it was necessary to analyze the estimation results, including changes in tobacco planting regions in the current year, the main reasons for these changes, and the areas affected, so as to provide decision support for the local tobacco management departments.

Construction of SaaS cloud service for estimating the area of large-scale FMTF: To better disseminate the relevant findings of this work and serve more tobacco farmers and relevant management departments, this work has constructed a large-scale tobacco planting area estimation SaaS cloud service platform, named "Guangze Tobacco Planting Area Smart Verification Platform," specifically for Guangze County. Users can upload high-resolution satellite images of the sampled groups in Guangze County from a specific year. The platform performs tobacco planting area estimation using the method proposed in this work on the cloud and returns the estimation results. Due to the adoption of a deep learning network structure in the extraction of mulched tobacco fields in this work, the platform requires high computing capability during model training, validation and testing. This work employs a multi-GPU parallel working mode, effectively enhancing the efficiency of the SaaS cloud platform for large-scale tobacco planting area estimation.

This work used the multi-tenant model to build the above SaaS platform, which can support multiple users to use the planting area estimation function in the cloud in parallel. In this work, the SaaS platform adopted four-layer architecture composed of presentation layer, scheduling layer, business layer and data layer, having good scalability. The presentation layer uses web browser to realize human-computer interaction and graphical display of tobacco planting area estimation results; The scheduling layer is first responsible for identifying and authenticating each user request, and then scheduling reasonably according to the load and service characteristics of the service processor; The business layer is responsible for receiving the requests forwarded by the scheduling layer, and calling the tobacco planting area estimation model (such as EMFMTFIUN) to process the relevant requests and return the processing result (the processing objects are the satellite images uploaded to the data layer). The tobacco planting area estimation model in the business layer is separately deployed on a multi-GPU server to ensure sufficient computing power to support multiple users to use the estimation model in parallel to process related services; the data layer uses distributed server clusters to store transaction data, user operation logs and various satellite images uploaded by users.

Performance evaluation metrics: The performance evaluation of the tobacco planting area estimation method proposed in this work included two parts: the overall performance evaluation of tobacco planting area estimation and the performance evaluation of FMTF identification, as detailed below:

(1) Overall performance evaluation of tobacco planting area estimation.
Using two evaluation metrics, namely overall estimation accuracy (the deviation ratio between estimated area and actual area) and overall time consumption, the overall performance of the tobacco planting area estimation method proposed in this work was comprehensively evaluated.

(2) Evaluation of identification performance in FMTF

In terms of the evaluation of identification accuracy for FMTF, this work utilizes a confusion matrix (as shown in Table 1) to evaluate the identification effectiveness of FMTF, where the positive class represents FMTF and the negative class represents the background.

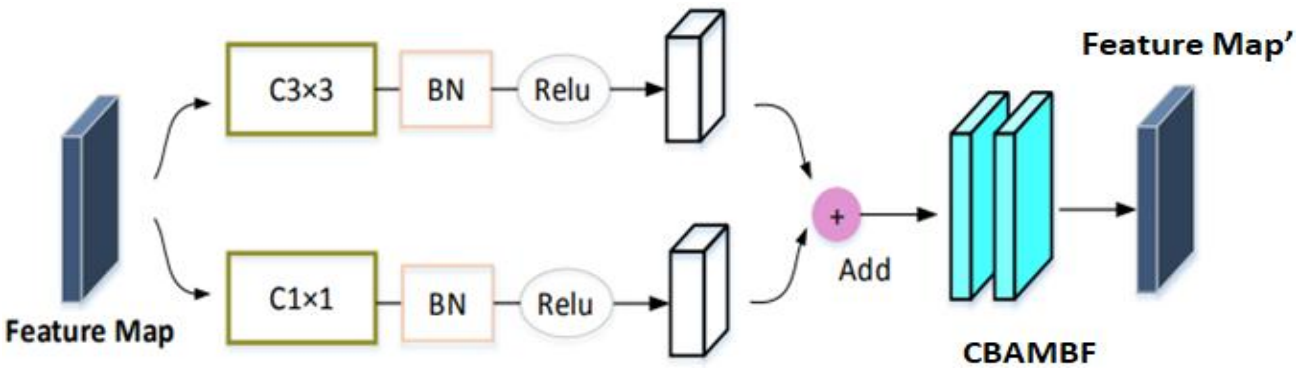

Fig. 8. Structure and working process of the FRMFMTFR module.

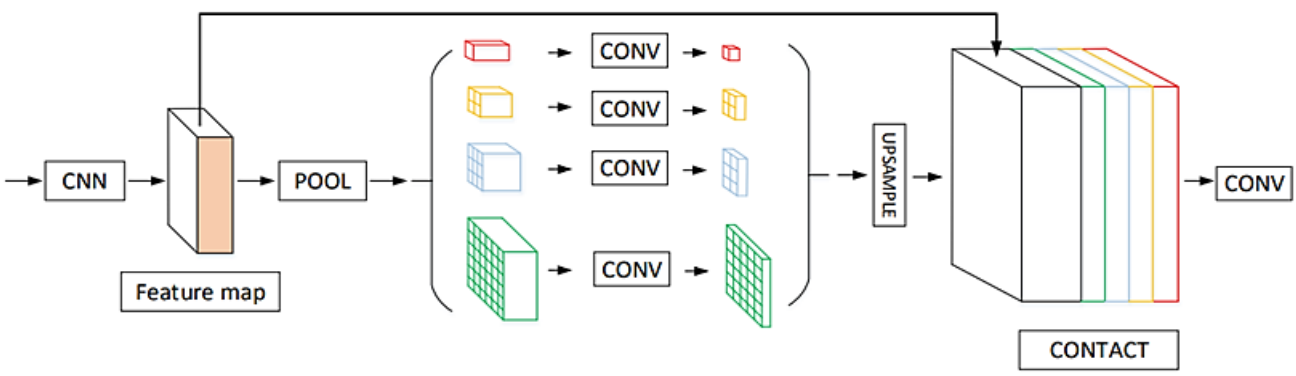

Fig. 9. Structure and working process of FPPMFMTF module.

In this work, six metrics, namely Precision(*P*), Recall(*R*), Overall Accuracy (*OA*), Intersection over Union (*IoU*), mean Intersection over Union (*mIoU*), and F1 score (*F*1), were primarily employed to evaluate the extraction accuracy of FMTF in practical FMTF extraction. Additionally, the average time consumption (*Ts*) was used to measure the efficiency of FMTF identification. The specific definitions of the aforementioned accuracy evaluation metrics are as follows:

Precision(*P*): It represents the ratio of the number of samples correctly classified as positive to the total number of samples classified as positive, that is, the proportion of actual positive samples among the predicted positive samples. Its calculation procedure is shown in formula (2).

$$P = \frac{TP}{TP + FP} \quad \text{(2)}$$

Recall(*R*): It represents the ratio of the number of samples classified as positive to the actual number of positive samples in the test set, that is, the proportion of correctly classified as positive samples among the all actual positive samples in the test set. Its calculation procedure is as shown in formula (3).

$$R = \frac{TP}{TP + FN} \quad \text{(3)}$$

Overall accuracy (*OA*): It represents the proportion of all correctly classified samples (i.e., the sum of elements on the diagonal of the confusion matrix) to the total number of samples (i.e., the sum of all elements in the confusion matrix). It is one of the core indicators for evaluating the overall performance of a classification model. In this work, it specifically refers to the proportion of the total number of samples correctly classified as FMTF and non-FMTF to the total number of samples. Its calculation procedure is shown in formula (4).

$$OA = \frac{TP + TN}{TP + TN + FP + FN} \quad \text{(4)}$$

Intersection-over-union (*IoU*): It represents the ratio of the intersection to the union of the actual category samples and the predicted category samples. Specifically, it is the ratio of the number of correctly classified positive samples to the sum of three terms: the number of correctly classified positive samples, the number of positive samples misclassified as negative, and the number of negative samples misclassified as positive. Its calculation is shown in formula (5).

$$IoU = \frac{TP}{TP + FP + FN} \quad \text{(5)}$$

Mean intersection-over-union (*mIoU*): It represents the average IoU value of all classes. The calculation procedure is as shown in formula (6), where *n* denotes the number of classes.

$$mIoU = \frac{\sum_i^n IoU}{n} \quad \text{(6)}$$

F1-Score (*F*1): It is the harmonic mean of Precision and Recall, as well as an indicator that comprehensively evaluates classification accuracy. It is particularly suitable for finding the balance point between the two metrics of *P* and *R*. Its calculation procedure is shown in Formula (7). Generally, the F1-Score is high only when both *P* and *R* are high.

$$F1 = \frac{2 \times P \times R}{P + R} \quad \text{(7)}$$

Results and Discussion

EMFMTFIUN training and FMTF extraction results: This work takes two selected clusters (Siqian Township and Zhima Township) from cluster sampling as examples, with all the monitoring grid coverage areas within them serving as two independent experimental regions (region01 and region02), respectively. During the training process of the EMFMTFIUN model, it is necessary to first create a training sample set. This work utilizes "Jilin-1" wide-band 01B satellite remote sensing image data as the data source. During field investigations, handheld GPS was used to manually locate some points in the typical FMTF, and a batch of image blocks with a size of 256×256 pixels was cut from the aforementioned satellite images centered on the above points using ArcGIS 10.5. Each image block corresponds to a labeled image (where pixels are labeled as 0 for non-field-mulched tobacco fields and 1 otherwise), thus forming a labeled sample set. Since the extraction of FMTF using deep learning requires a large number of labeled samples, and the number of manually created original labeled samples is limited, sample augmentation is necessary to increase the number of labeled samples. In this work, sample augmentation can be achieved by rotating the sample image blocks. Through horizontal flipping, vertical flipping and diagonal mirroring, the number of labeled sample image blocks can be increased to four times the original number. The augmented sample set is divided into a training set, a validation set and a test set, and the sample numbers of the above three sets corresponding to region01 are 8832, 1263, and 2523, respectively, while the sample numbers of the above three sets corresponding to region02 are 5350, 840, and 1562, respectively.

To evaluate the performance of the EMFMTFIUN model proposed in this work, comparative experiments were conducted using representative alternative models, namely DeeplabV3+ (Fu & Huang, 2022), Hrnet (Sun *et al*., 2019), Pspnet (Zhang *et al*., 2022) and SegFormer (Xie *et al*., 2021; Tian *et al*., 2023), Attention-Unit (John & Zhang, 2022; Huang *et al*., 2024), ResUnet (Zhao *et al*., 2023) and Unet++ (Aaseegha & Venkataramana, 2026),along with EMFMTFIUN for the extraction of FMTF. For the whole regions of Region01 and Region02, the aforementioned 8 different network models were employed for the extraction of FMTF, and the comparison of extraction results (average values from 2022 to 2025) is presented in Tables 2 and 3. It should be emphasized that the proposed model (EMFMTFIUN) and the compared baseline models (DeeplabV3+, Hrnet, Pspnet, SegFormer，Attention-Unet、ResUnet、Unet++) had the same training conditions, including the same training samples, validation samples and test samples, as well as the same environments of software and hardware.

Table 1. Confusion matrix.

Confusion matrix		Actual values	
		Number of pixels belonging to actual mulched tobacco fields	Number of pixels belonging to actual non-FMTF
Predicted values	Number of pixels predicted as FMTF	*TP*	*FN*
	Number of pixels predicted as non-FMTF	*FP*	*TN*

Table 2. Comparison of extraction results of FMTF in Region 1 based on different network models (average values from 2022 to 2025).

	DeeplabV3+	Hrnet	Pspnet	SegFormer	Attention-Unet	ResUnet	Unet++	EMFMTFIUN
P	0.84	0.82	0.72	0.86	0.80	0.82	0.81	0.86
R	0.82	0.81	0.83	0.90	0.79	0.80	0.78	0.91
OA	0.95	0.95	0.96	0.97	0.93	0.95	0.93	0.97
IOU	0.74	0.72	0.63	0.82	0.73	0.78	0.73	0.83
mIOU	0.83	0.85	0.80	0.89	0.80	0.81	0.80	0.91
*F*1	0.84	0.87	0.82	0.90	0.81	0.83	0.83	0.90
Ts/hours	16.93	19.77	18.42	24.34	13.24	17.39	18.31	10.33

Table 3. Comparison of extraction results of Region02 FMTF based on different network models (average values from 2022 to 2025).

	DeeplabV3+	Hrnet	Pspnet	SegFormer	Attention-Unet	ResUnet	Unet++	EMFMTFIUN
P	0.84	0.82	0.78	0.89	0.81	0.83	0.82	0.90
R	0.83	0.82	0.87	0.93	0.81	0.83	0.80	0.93
OA	**0.97**	0.96	0.96	0.97	0.95	0.96	0.96	0.98
IOU	0.78	0.77	0.69	0.83	0.76	0.77	0.77	0.85
mIOU	0.88	0.87	0.85	0.89	0.85	0.88	0.86	0.92
*F*1	0.84	0.83	0.83	0.91	0.83	0.85	0.81	0.92
Ts/hours	7.33	9.07	8.47	9.21	6.13	8.35	8.77	5.37

Table 4. Analysis of estimated tobacco planting area in Guangze County(2022-2025).

Year	Actual area (hectare)	Predicted area (hectare)	Overall estimation accuracy /%	Estimated time (second)
2022	2314.49	2264.60	97.85%	123.3
2023	2307.95	2254.73	97.69%	109.4
2024	2314.49	2216.51	95.77%	110.7
2025	2367.85	2280.54	96.31%	131.9
Average	2326.23	2254.13	96.91%	118.8

From Tables 2 and 3, it can be seen that compared to the seven other representative models, the EMFMTFIUN model proposed in this work achieves higher accuracy and efficiency in extracting FMTF, and its overall accuracies of FMTF extraction in Region01 and Region02 reach 0.97 and 0.98, respectively, with time consumptions of 10.33 hours and 5.37 hours, respectively. In addition, the overall accuracy of the SegFormer model is equal to or slightly lower than that of EMFMTFIUN, while the performance of the DeeplabV3+ model is second only to SegFormer. The extraction accuracy and efficiency of Pspnet and SegFormer models are relatively lower. In terms of extraction effects, the seven models of DeeplabV3+, Hrnet, Pspnet ,SegFormer, Attention-Unet, ResUnet and Unet++ have their own advantages and disadvantages in terms of missed and incorrect extractions, and the extracted FMTF all exhibit varying degrees of boundary missing or adhesion. Compared with the above 7 models, EMFMTFIUN has the best overall effects of the FMTF boundaries extraction, followed by SegFormer, DeeplabV3+, Hrnet, ResUnet,Unet++, Attention-Unet in order of extraction accuracy from high to low, while the Pspnet model has most missed extractions, especially in the extraction of small tobacco fields. The boundaries of the FMTF extracted by the EMFMTFIUN model are clearer and more complete, and in areas with a higher distribution of fragmented small fields, the efficiency and accuracy of extracting small FMTF are higher compared to the other 7 models. Taking Siqian Township, one of the two selected cluster sampling townships , as an example, the extraction results of large-scale FMTF based on EMFMTFIUN are shown in Fig. 10.

Estimation results of the total tobacco planting area in the whole experimental regions: Using the large-scale tobacco planting area estimation method based on the improved Unet and spatial sampling proposed in this work, the actual tobacco planting area of the entire Guangze County in Nanping City from 2022 to 2025 was estimated, and the results are presented in Table 4.

As indicated in Table 4, the average estimation accuracy of tobacco planting area in Guangze County, Nanping City from 2022 to 2025 reached 96.91%, with an average estimation time consumption of 118.8 seconds. Both the accuracy and efficiency of large-scale tobacco planting area estimation meet the relevant requirements of the local tobacco management department, indicating promising market application prospects.

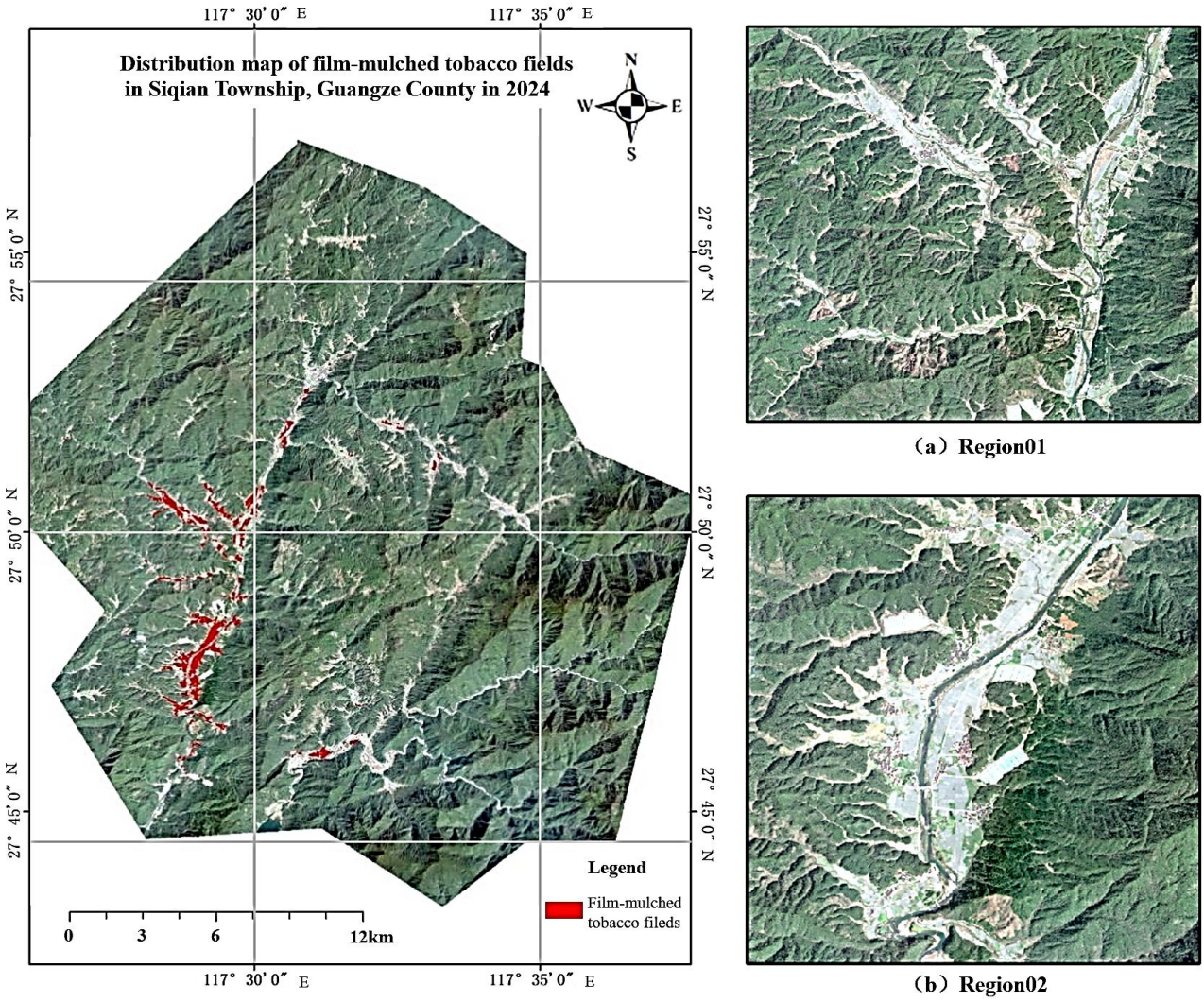

Fig. 10. Distribution map of extraction results of FMTF in Siqian Township using EMFMTFIUN (2024).

Smart verification platform for tobacco planting area of Guangze County: The "Smart Verification Platform for Tobacco Planting Area of Guangze County" constructed in this work is a hybrid platform combining Web and SaaS. The core functional modules of the system (such as tobacco planting area estimation and tobacco plant count statistics) can all be implemented on the cloud. Users can upload high-resolution satellite images of relevant monitoring grid regions, extract FMTF within these grid regions through the cloud-based EMFMTFIUN model, and combine the spatial sampling rules mentioned earlier to estimate the total tobacco planting area in the whole experimental regions. This provides free or paid services for tobacco farmers and tobacco management departments.

The main novelty of this work is employing the new methods of FMTF extraction and spatial sampling, aiming to effectively reduce the cost of estimating the large-scale tobacco planting area while maintaining high accuracy.

In terms of FMTF extraction, the EMFMTFIUN model proposed in this work exhibits notable advantages in both accuracy and efficiency compared to other 7 compared models of DeeplabV3+, Hrnet, Pspnet ,SegFormer, Attention-Unet, ResUnet and Unet++. Furthermore, when compared to the some Unet variants (such as SegFormer) with complex network structures, the EMFMTFIUN model, utilizing an improved Unet network structure, proves to be more accurate and efficient. Although the EMFMTFIUN model's advantage in extraction accuracy is not particularly pronounced, it demonstrates a clear superiority in extraction efficiency compared to the other seven models, making it more suitable for scenarios requiring large-scale remote sensing image datasets for learning. EMFMTFIUN is only built and used in Goungze County, but it can also been used for extracting FMTF and estimating tobacco planting areas on a larger scale with the similar terrain, climate and tobacco planting category in Fuzhou provinces, China, because the larger images can been divided into some sub images which can be processed in parallel. According to the experiments of this work, a county is not the most suitable scale for the estimation experiments

As for spatial sampling, the combination of cluster sampling and spline sampling is employed to improve the efficiency of large-scale FMTF extraction in this work. In cluster sampling, instead of randomly selecting several clusters as the sampled clusters, we divided clusters based on existing administrative divisions after analyzing historical records. We then selected two or more clusters (townships) with the closest permanent tobacco field planting proportions over the years to the whole experimental region as the sampled clusters. This approach

not only avoids unpredictable errors that may arise from random sampling but also demonstrates good stability. The aforementioned cluster sampling method is more suitable for situations where there is relatively little overall difference between clusters within the groups but significant internal differences within each cluster. Additionally, in spline sampling, we use the existing tobacco planting monitoring grids (key monitoring regions designated by the local tobacco management department) in the sampled clusters as the scope for further extracting mulched tobacco fields, effectively reducing survey workload and information omissions.

Of course, there are also some shortcomings in this work. First, EMFMTFIUN model proposed in this work is mainly used to estimate the planting area of tobacco with category of K358, and have not been used to estimate the planting area of other crops, so its generalization ability for other crops remains to be verified. Secondly, the EMFMTFIUN model can be used to other regions except Guangze County in Fujian province, China because of the similar terrain, climate and tobacco planting category, but its generalization ability in other province of China have not verified. Third, county-scale is not necessarily the most suitable scale for the spatial sampling method proposed in this work, because the different counties may have quite different size. How to determine the optimal scale of spatial sampling is an important issue to be solved in the follow-up research.

Conclusion

This work comprehensively utilizes techniques such as spatial sampling, remote sensing image interpretation and deep learning, as well as existing vector data from the experimental regions, to perform estimating the large-scale tobacco planting area. By combining cluster sampling and spline sampling techniques, the survey scope is relatively narrowed down, and the proposed EMFMTFIUN model is adopted to effectively improve the accuracy and efficiency of extracting FMTF. Finally, based on the total farmland area, the proportion of film-mulched farmland in the whole experimental regions, and the proportion of FMTF in the sampling regions, the total area of FMTF and the total area of tobacco planting are calculated, and a corresponding SaaS cloud service platform for tobacco planting area estimation is constructed. Taking Guangze County in Nanping City as an example, the average estimation accuracy of tobacco planting area from 2022 to 2025 reached 96.91%, with an average estimation time consumption of 118.8 seconds. The accuracy of FMTF extraction using the EMFMTFIUN model reached over 0.97, effectively improving the accuracy and efficiency of large-scale tobacco planting area estimation, reducing estimation costs, meeting the relevant requirements of local tobacco management departments, and demonstrating promising market application prospects.

Conflicts of Interest: The authors declare no conflict of interest.

Author Contributions: Conceptualization, F.H.; methodology, F.H. and C.Y.; validation, F.H. and Q.Z.; formal analysis, F.H. and X.H.; investigation, Q.Z. and Y.Z.; data curation, Q.Z. and Z.Z.; writing—original draft preparation, F.H. and C.Y.; writing—review and editing, F.H. and Q.Z.; supervision, F.H and G.C.; funding acquisition, F.H. and Y.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Fujian Province Science and Technology Plan Project (Guided Project, 2023N0021), Fuzhou science and technology plan project (2025-SG-012), Fujian education and research projects for young and middle-aged teachers (JZ240084).

Acknowledgments

The authors would like to thank Changping Yang, Xiang Hu in College of Artificial Intelligence (China) for their assistance, suggestions, and discussions.

References

Aaseegha, M.D. and B. Venkataramana. 2026. A deep learning framework for cervical cancer detection: Unet++ Segmentation and DenseNet-121 Classification. *Ind. J. Gynecol. Oncol*., 24(21): 1-14.

Bei, Y.T., X.Y. Zhang, K.H. Wang, J.X. Gao and X.B. Sun. 2025. Research progress of crop production based on remote sensing technology. *Chin. Agric. Sci. Bull.*, 41(27): 149-156.

Fu, B.H. and L. Huang. 2022. Extraction of tobacco planting area from UAV images based on deep semantic segmentation. *Commun. Technol.*, 55(02): 181-186.

Gu, L., S.Q. Xu and L.Q. Zhu. 2020. Detection of building changes in remote sensing images via FlowS-Unet. *Acta Autom. Sin.*, 46(06): 1291-1300.

Hu, Y.S. and C.Y. Zhou. 2022. Research status and development trends of agricultural monitoring technology based on remote sensing and spatial sampling methods. *Agric. Technol.*, 42(13): 60-63.

Huang, L.B., Y.Y. Lin, P. Cao, X. Zou, Q. Qin, Z.Y. Lin, F.T. Liang and Z.Y. Li. 2024. Automated detection and segmentation of pleural effusion on ultrasound images using an attention U-net. *J. Appl. Clin. Med. Phys*., 25(1): 14231.

Jian, R.B., Z.Y. Cai, Z.S. Yang, W.Z. Wang, Y. Liu, J.F. Chen and M.L. Wang. 2023. Research on image segmentation method of field wheat harvest boundary based on U-Net. *J. Henan Agric. Univ.*, 57(03): 444-450.

John, D. and C. Zhang. 2022. An attention-based U-Net for detecting deforestation within satellite sensor imagery. *Int. J. Appl. Earth Obs. Geoinf.*, 107: 102685.

Liu, T.J., S.B. Duan, N.T. Liu, B.A. Wei, J.T. Yang, J.K. Chen and L. Zhang. 2024. Estimation of crop leaf area index based on Sentinel-2 images and PROSAIL-Transformer coupling model. *Comput. Electron. Agric.*, 227(2): 109663.

Pan, Y.Z., Q.Z. Li, J.S. Zhang, J.T. Shi, X.F. Zhu, X.H. Gu, W.Q. Zhu, J.W. Yue, K. Jia, X. Du, F.M. Zhang, Y. Zhang, H.Y. Wang, W.N. Wang and L. Li. 2025. Summary and review of research progress on multi-dimensional and multi-scale stereoscopic statistical remote sensing survey technology for major crop areas. *J. Beijing Norm. Univ. (Nat. Sci.)*, 61(01): 118-125.

Ren, J., Y. Shao, H. Wan, Y.H. Xie and A. Campos. 2021. A two-step mapping of irrigated corn with multi-temporal MODIS and landsat analysis ready data. *J. Photogramm. Remote Sens.*, 176: 69-82.

Shen, X.D., X.L. Wu and Y.D. Lei. 2020. A building change detection method based on improved Unet. *Pract. Electron.*, (1): 15, 30-31.

Sun, K., B. Xiao, D. Liu and J.D. Wang. 2019. Deep high-resolution representation learning for human pose estimation. *Proc. 2019 IEEE/CVF Conf. Comput. Vis. Pattern Recognit*. (*CVPR2019*), Long Beach, CA, USA, 5686-5696.

Tian, H.R., P.X. Wang, K. Tansey, J. Wang, W.T. Quan and J.M. Liu. 2024. Attention mechanism-based deep learning approach for wheat yield estimation and uncertainty analysis from remotely sensed variables. *Agric. For. Meteorol.*, 356: 110183.

Tian, T., D. Wang, Z. Wang and H.B. Li. 2022. Precise classification of crops in complex planting structure area based on deep learning model. *Chin. J. Agric. Resour. Reg. Plan.*, 43(12): 147-158.

Tian, X.W., J.L. Wang, M. Chen and S.Q. Du. 2023. Improved SegFormer network based method for semantic segmentation of remote sensing images. *Comput. Eng. Appl.*, 59(08): 217-226.

Xie, E.Z., W.H. Wang, Z.D. Yu, A. Anandkumar, J.M. Alvarez and P. Luo. 2021. SegFormer: Simple and efficient design for semantic segmentation with transformers. *Adv. Neural Inform. Process. Syst.*, 34: 12077-12090.

Xun, L., J.H. Zhang, D. Cao, J.W. Wang, S. Zhang and F.M. Yao. 2021. Mapping cotton cultivated area combining remote sensing with a fused representation-based classification algorithm. *Comput. Electron. Agric.*, 181: 105940.

Yang, J., C.B. Zhang, S.M. Chen, Z.C. Li and Q.P. Chen. 2026. Application of remote sensing image technology in accurate extraction of tobacco planting information. *Agric. Mach. Agron.*, 57(01): 61-64.

Zhang, L., C.H. Liu, L.F. Shi and Y. Zhang. 2021. Tobacco planting information extraction based on U-net neural network. *Agric. Technol.*, 41(22): 44-47.

Zhang, R., J. Chen, L. Feng, S. Li, W. Yang and D. Guo. 2022. A refined pyramid scene parsing network for polarimetric SAR image semantic segmentation in agricultural areas. *IEEE Geosci. Remote Sens. Lett.*, 19: 1-5.

Zhang, X., Z.L. Liu, X. Li, H. Bao, N.N. Zhang and T.C. Bai. 2025. High-accuracy cotton field mapping and spatiotemporal evolution analysis of continuous cropping using multi-source remote sensing feature fusion and advanced deep learning. *Agriculture*, 15(17): 1814.

Zhang, N.N., X. Zhang, T.C. Bai, P. Shang, W.H. Wang and L. Li. 2023. Identification method of cotton leaf pests and diseases in natural environment based on CBAM-YOLO v7. *Trans. Chin. Soc. Agric. Mach.*, 54: 239-244.

Zhao, Z.J., F. Zhang, Q.Wu, Z.Q. Li, X. Tong, J.W. Li and W. Han. 2023. Cloud identification and properties retrieval of the fengyun-4A satellite using a ResUnet model. *IEEE Trans. Geosci. Remote Sens*., 61: 1-18.