IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 24 December 2024, accepted 17 January 2025, date of publication 24 January 2025, date of current version 30 January 2025.
Digital Object Identifier 10.1109/ACCESS.2025.3534024

METHODS

Tea Disease Recognition Based on Image Segmentation and Data Augmentation

JI LI[1] **AND CHENYI LIAO**[2]
[1]College of Bioscience and Biotechnology, Hunan Agricultural University, Changsha 410128, China
[2]School of Environmental and Chemical Engineering, Foshan University, Foshan 528000, China
Corresponding author: Ji Li (jili@stu.hunau.edu.cn)

ABSTRACT Accurate identification of tea leaf diseases is crucial for intelligent tea cultivation and monitoring. However, the complex environment of tea plantations—affected by weather variations and uneven lighting—poses significant challenges for building effective disease recognition models using raw field-captured images. To address this, we propose a method that combines two-stage image segmentation with an improved conditional generative adversarial network (IC-GAN). The two-stage segmentation approach, integrating graph cuts and support vector machines (SVM), effectively isolates disease regions from complex backgrounds. The IC-GAN augments the dataset by generating high-quality synthetic disease images for model training. Finally, an Inception Embedded Pooling Convolutional Neural Network (IDCNN) is developed for disease recognition. Experimental results demonstrate that the segmentation method improves recognition accuracy from 53.36% to 75.63%, while the IC-GAN increases the dataset size. The IDCNN achieves 97.66% accuracy, 97.36% recall, and a 96.98% F1 score across three types of tea diseases. Comparative evaluations on two additional datasets further confirm the method's robustness and accuracy, offering a practical solution to reduce tea production losses and improve quality.

INDEX TERMS Conditional generative adversarial network, disease recognition, deep learning, image generation.

I. INTRODUCTION

China is a leading producer and consumer of tea [1]. In 2022, the country's tea production reached 3.18 million tons, generating an output value of 31.8 billion yuan [2], [3]. Cultivating tea trees and producing high-quality tea are crucial pathways for tea farmers to achieve prosperity. However, the prevalence of approximately 130 types of tea diseases in China poses significant threats, leading to reduced tea quality and yield. Consequently, accurately identifying tea diseases and implementing timely preventive measures are vital for minimizing yield losses, enhancing tea quality, and boosting the income of tea farmers.

The traditional method of identifying tea diseases involves observing symptoms and spots with the naked eye, then diagnosing the disease type based on experience [4], [5] [6]. However, due to the limitations of human perception, manual identification is inefficient, costly, and impractical for large-scale tea plantation disease monitoring. As a result, automatic tea disease detection and recognition has become a prominent research area in machine learning and computer vision [7]. In recent years, many researchers have conducted extensive studies on automatic plant leaf disease recognition methods based on image processing. Several classification approaches have been explored, including self-supervised learning techniques to improve feature representations for disease classification. Some studies [3] have attempted to use color spaces and traditional machine learning methods for better classification. However, existing methods are often hindered by complex background interference and poor robustness due to the similarity in color between leaf diseases, resulting in weak generalization and difficulty in practical applications. This work proposes a two-stage segmentation framework to better extract complex background features of leaf diseases, facilitating improved classification performance.

The associate editor coordinating the review of this manuscript and approving it for publication was Mario Donato Marino.

Conversely, numerous studies have focused on utilizing YOLO-based convolutional neural networks (CNNs) to identify tea diseases [8], [9]. These approaches often involve integrating modules such as MobileNetV2 or other components into the network structure to mitigate background interference [10] or enhancing small-sample tea disease recognition by incorporating attention mechanisms [11]. While these methods have demonstrated notable improvements in recognition accuracy, they also present several limitations. The high complexity of CNN models and their substantial demand for computational resources can restrict their applicability in resource-constrained environments. Additionally, the adaptability of these methods to diverse tea diseases and varying environmental conditions remains a significant challenge.

In summary, the field still faces the following challenges:

(i) Small sample problem. Due to geographical, temporal, and weather-related limitations in collecting disease samples, the existing dataset is small, making it difficult to meet the training requirements of deep learning models. Additionally, during actual data collection, the number of samples for certain diseases may be much larger than for others, leading to an imbalanced dataset and negatively affecting the model's generalization ability.

(ii) Complex background interference. Plant disease images obtained in real cultivation environments often contain various background elements such as leaves, trunks, stems, roots, soil, weeds, straw, plastic film, fallen leaves, stones, and standing water. These complex backgrounds make it difficult to detect disease lesions, directly impacting the accuracy of disease region segmentation. Moreover, the color and shape of disease spots may closely resemble those of other objects in the background, further decreasing classification accuracy.

It can be seen from various studies [12], [13], [14] that the process of identifying plant leaf diseases generally involves three stages: (1) Image segmentation: isolating the disease-affected area from the leaf image; (2) Feature extraction: extracting relevant features from the segmented disease image; and (3) Disease recognition: classifying the extracted features to identify the type of leaf disease. Researchers have proposed various methods for each of these stages to improve disease recognition performance, particularly under small sample conditions. This study was conducted on three datasets and compared with nine other methods. The results demonstrate that the proposed method effectively improves the accuracy of the classification of tea leaf disease, showing superior performance in key metrics such as accuracy, recall, and F1 score when compared to previous methods.

The main contributions of this paper are as follows: (1) A novel two-stage image segmentation framework is proposed, integrating the Graph-Cut algorithm with support vector machine (SVM) technology to achieve high-precision segmentation of tea disease regions in complex backgrounds, providing high-quality input data for subsequent disease recognition tasks and enhancing the accuracy of the recognition process. (2) An improved conditional generative adversarial network (IC-GAN) is designed and optimized by dividing the data manifold into overlapping neighborhoods and modeling the distribution of complex datasets as a mixture of local conditional distributions. This approach significantly improves the model's generation stability and data augmentation capabilities, effectively addressing small-sample learning challenges. (3) An Inception module-embedded deep neural network (IDCNN) is developed, combining dilated convolution techniques to achieve a substantial improvement in tea disease image classification accuracy, with an average accuracy of 96.17%. (4) Key challenges in tea disease recognition, such as small sample sizes, complex backgrounds, and uneven sample distributions, are addressed through effective solutions, offering new technical pathways for plant disease recognition and contributing valuable insights to the advancement of related research fields.

The rest of this paper is organized as follows. The second part introduces related work. The third part provides a detailed description of the proposed method. The fourth part presents and discusses the experimental results. The final section is the conclusion of this paper.

II. RELATED WORK

Research on plant disease recognition primarily focuses on traditional machine learning methods, deep learning methods, and recent improvements in techniques for complex backgrounds and small sample problems. This paper summarizes relevant studies and analyzes the differences between this work and existing research.

A. IMAGE SEGMENTATION FRAMEWORK

Using image segmentation to separate the foreground from the background allows for the accurate identification of disease-affected areas on tea leaves. Kumar et al. [15] applied a Gaussian mixture model (GMM) to separate the foreground and background of tea leaf disease images, combining it with a particle swarm optimization-based fuzzy C-means algorithm for disease region segmentation. Khan et al. [16] used an expectation maximization algorithm to optimize strongly correlated pixels and extracted disease features from apple leaf disease images. Shin et al. [17] trained a strawberry leaf grayscale histogram, extracted features for normalization, and then used SVM and other methods to classify and identify strawberry pests and diseases. Zhang et al. [18] employed superpixel segmentation to process complex background tomato disease leaf images, combined it with local binary pattern (LBP) for feature extraction, and used SVM for classification. Mokhtar et al. [19] applied Gabor wavelet transform to extract texture features from tomato leaf images and combined it with an SVM classifier for disease detection. While these methods have shown effectiveness under specific conditions, they are typically reliant on manually designed features and struggle to adapt to complex natural backgrounds or a wide variety of plant

TABLE 1. Comparison of recognition accuracy of IDCNN models based on lesion images and original tea disease images.

Author(s)	Technique	Methodology (Backbone)	Limitations
Pravin Kumar et al.	Machine Learning	Gaussian Mixture Model + SVM	Low accuracy, requires clear background
Zhao WL et al.	Deep Transfer Learning	SE-DenseNet-FL	Uneven distribution of disease images
Ramesh et al.	Deep Learning	DNN-JOA+Jaya	Fewer identified categories
Usman Afzaal et al.	Deep Feature Extraction+ Segmentation	Mask R-CNN	Requires large-scale annotated data

diseases. In contrast, we use a two-stage segmentation model to accurately segment the disease-affected areas.


(a) (b) (c) (d)

FIGURE 1. Example of a dataset of tea disease images captured in the field. (a) tea circular red spot disease; (b) tea red leaf spot disease; (c) tea cloudy leaf blight; (d) healthy tea leaves.

B. IMPROVED TECHNIQUES FOR SMALL SAMPLE AND COMPLEX BACKGROUND PROBLEMS

To address the challenges of small sample size and complex backgrounds, researchers have proposed various improved methods, including data augmentation, transfer learning, and generative adversarial networks (GAN). Khirade and Patil [20] used an OTSU threshold-based method to segment the disease area, which was then classified using an artificial neural network. Onyango and Marchant [21] separated the cotton plant from the soil using RGB color features and achieved fast segmentation extraction by combining morphological processing techniques with color segmentation methods. Quan [22] accurately separated the melon fruit and surface defects from the background using a parallel watershed algorithm and deployed the segmentation on an FPGA platform to enhance real-time performance. Chen et al. [23] developed a novel plant disease recognition model based on transfer learning and deep learning to achieve efficient disease detection. Although these methods alleviate the issue of small sample size to some extent, the enhancement of augmented data diversity remains a key challenge that needs further attention.

C. A DEEP LEARNING-BASED METHOD FOR TEA DISEASE IDENTIFICATION

Deep learning technology has overcome the dependence on manually designed features in traditional methods by automatically extracting image features and has been widely applied in disease identification in recent years. Rai and Pahuja [24]proposed a disease leaf symptom recognition technique based on convolutional neural networks to improve the accuracy of leaf disease recognition. Li et al. [25] employed a SE-DenseNet-FL model based on transfer learning to perform tea disease recognition. Shao et al. [26] used the OTSU method to segment and extract spotted disease areas from tobacco leaves, followed by multi-feature selection based on a BP neural network for disease feature extraction. Ramesh and Vydeki [27] proposed optimizing the weight of a deep neural network using the Jaya optimization algorithm for rice leaf disease recognition, followed by K-Means clustering to remove background features and isolate the disease areas. Krisnandi et al. [28] completed disease classification through contrast enhancement and K-means segmentation of disease regions, combined with the SVM algorithm. Afzaal et al. [29] achieved automatic detection and segmentation of tomato leaf diseases using Mask R-CNN, significantly improving the efficiency of disease recognition in complex scenes. However, deep learning methods typically require a large labeled dataset for support and may not perform well in small sample conditions. Additionally, the robustness of these methods in complex backgrounds and varying lighting conditions still requires improvement. Table 1 illustrates the main methods for plant disease recognition in recent years and their characteristics.

Research on plant leaf disease using deep learning methods has primarily focused on fruits and food crops, with relatively few studies dedicated to tea diseases. Tea diseases have a significant impact on both the yield and quality of tea. To enable timely prevention and control of tea diseases and reduce economic losses for tea farmers, this paper proposes a method for small-sample recognition of tea leaf diseases based on a two-stage image segmentation approach and IC-GAN for image enhancement. The goal is to achieve accurate recognition of tea diseases under small sample conditions and field environments. The method is demonstrated using three common tea diseases—Tea Red Spot Disease, Tea Leaf Blight, and Tea Cloud Spot Disease—as examples. The superiority of the proposed method is validated by first segmenting the original disease images of tea leaves using the two-stage image segmentation approach, then enhancing the segmented diseased images with IC-GAN, and finally employing a dilated convolutional neural network embedded with Inception for the recognition of the augmented diseased images.

III. MATERIALS AND METHODS

A. COLLECTION OF TEA DISEASE IMAGES

Anshun City in Guizhou Province is a thousand-year-old tea-producing region, where a small portion of tea cultivation supports the economic development of several counties and districts. Tea resources are abundant in Anshun, and various common diseases of tea crops are prevalent here. Therefore, the tea disease images used in this study were collected from the tea garden in Damyang Town, Ziping County, Anshun City.

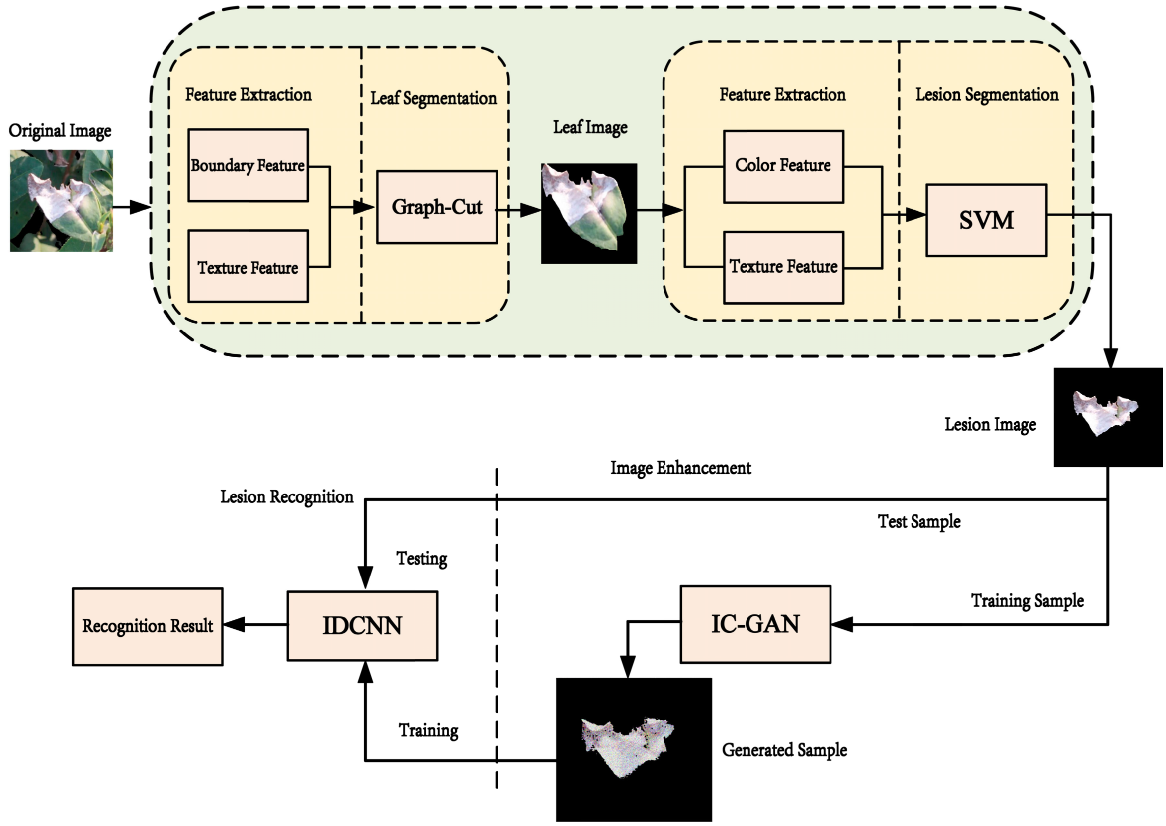


FIGURE 2. **A framework for small sample recognition of tea diseases based on two-stage image segmentation and IC-GAN image enhancement.**

To facilitate accurate labeling of the disease images by experts, a Mi 10s smartphone was used to photograph both diseased and healthy tea leaves, ensuring that each image contains only one type of tea disease. A total of 400 images were collected, including three common tea diseases: tea circular red spots, tea red leaf spots, and tea cloudy vein necrosis, as well as healthy tea leaves. For each type of tea, 100 images were collected, with a resolution of 2000 × 1824 pixels. Fig. 1 shows examples of typical images from each category selected from the original dataset.

B. SMALL SAMPLE RECOGNITION METHOD FOR TEA DISEASES

The small-sample tea disease recognition method proposed by this research institute consists of three parts: The first part involves segmenting disease spots from the entire image captured under field conditions to obtain the disease image. The second part enhances the disease image to expand the dataset and address the small sample problem. The third part focuses on disease image recognition to accurately diagnose the type of tea disease. Fig. 2 illustrates the flowchart of the small-sample tea disease recognition method.

C. A TWO-STAGE IMAGE SEGMENTATION METHOD FOR PATHOLOGICAL CHANGES

Lesion image segmentation refers to the process of extracting the diseased regions from complete tea leaf images, forming the basis for accurately and efficiently capturing lesion image features. In this study, a two-stage segmentation method combining Graph-Cut [30] and SVM [31] is proposed for segmenting tea leaf lesion images. As shown in Fig. 2, in the first stage, Graph-Cut leverages texture and boundary information to segment tea leaves from the original image. In the second stage, SVM uses texture and color features to segment the lesion areas from the tea leaf images.

Specifically, the color and texture feature vectors of lesion images and healthy tea leaves are used as training samples for SVM. The color features are extracted from color histograms in the RGB space, while the texture features are obtained using a Gabor filter.

The two-stage segmentation method based on Graph-Cut and SVM offers significant advantages, which justify its combination and application to tea disease recognition. While it may seem like a straightforward combination of well-known methods, the merit lies in the synergy between Graph-Cut's ability to handle complex background and SVM's robust classification capabilities. Graph-Cut, by leveraging graph-based models, is particularly effective in segmenting regions of interest, such as disease spots on tea leaves, from cluttered backgrounds. This method excels at accurately delineating the boundaries between diseased areas and the background, even under uneven illumination conditions. However, Graph-Cut alone can sometimes suffer from noise and incomplete segmentation, which is where SVM plays a crucial role. This two-stage approach significantly improves segmentation performance, as evidenced by our experimental results, and highlights the potential of combining traditional methods in an innovative way.

Fig. 3 illustrates the two-stage segmentation process and results for a typical diseased tea leaf image. As seen in Fig. 3, Graph-Cut effectively overcomes noise, successfully segments the tea leaf from the background, and lays a foundation for subsequent lesion image segmentation. Moreover, the segmentation method proposed in this study is unsupervised, automatically selecting the regions of interest without requiring human intervention.

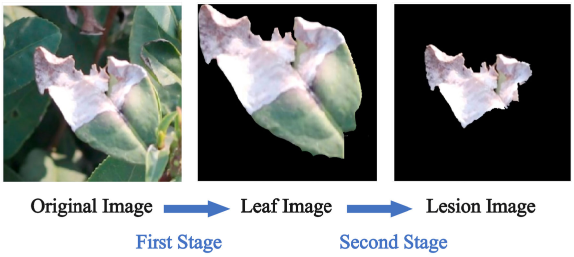


FIGURE 3. Schematic diagram of two-stage lesion image segmentation method.

1) GRAPH-CUT ALGORITHM

In previous image processing algorithms, previous research [32] was accustomed to viewing an image as a matrix of pixel points. However, in addition to this matrix representation, graph-cut-based algorithms understand images from the perspective of graph theory, where the image is abstracted as a graph structure. For example, if the processed image is a 3 $\times$ 3 pixel matrix.The algorithm abstracts an image as a graph structure, where the set of nodes V contains all pixel points, along with two terminal nodes S (source) and T (sink). The set of edges E includes all n-links (edges between adjacent pixels) and t-links (edges between pixel points and terminal nodes). By finding the minimum cut, the image pixels can be divided into foreground (connected to source node S) and background (connected to sink node T).

In the segmentation process, the region term $R(A)$ measures the cost of cutting t-links, which penalizes the mismatch between the feature of each pixel and the foreground or background models. It can be expressed as:

$$R(A) = \sum_{p \in P} R_p(A_p) \tag{1}$$

where $R_p(A_p)$ represents the penalty for pixel point p being classified as a certain category (foreground or background), which can be obtained by calculating the negative log of the posterior probability of the pixel value:

$$R_p\left(A_p = I\right) = -\ln \mathbb{P}\left(I_p|\mathcal{F}\right) \tag{2}$$

$$R_p\left(A_p = 0\right) = -\ln \mathbb{P}\left(I_p|\mathcal{B}\right) \tag{3}$$

I_p, F, and B refer to the pixel values, foreground region, and background region at pixel p, respectively. It is worth noting that the operation of the Graph Cut algorithm depends on the information input by the user. In the initial stage of the processing process, this study will use the user-annotated foreground and background regions to construct the corresponding probability models $\mathbb{P}(\cdot|\mathcal{F})$ and $\mathbb{P}(\cdot|\mathcal{B})$.

The boundary term $B(A)$ measures the cost of cutting n-links, which penalizes the discontinuity between neighboring pixel pairs. The boundary term can be expressed as:

$$B(A) = \sum_{(p,q) \in N} B(p, q)\delta(A_p, A_q) \tag{4}$$

where $B(p, q)$ is a function that measures the difference between pixel points p and q in the spatial and color domains. The greater the difference, the more likely these two points are not of the same category locally.

The penalty function or loss function $E(A)$, which evaluates the segmentation quality, is the sum of the region term and the boundary term. It is expressed as:

$$E(A) = \lambda R(A) + B(A) \quad (\lambda > 0) \tag{5}$$

where λ is a weight coefficient that adjusts the relative influence of the region term and the boundary term. The advantage of the graph-cut method lies in its fast execution speed and the robustness of the numerical solution. However, its disadvantage is that the loss function significantly influences the segmentation result, requiring user input for the segmentation process. Fig. 4 shows the process of how Graph-Cut segments images.

2) SVM SEGMENTATION

Support Vector Machines (SVM) are a classic binary classification technique. Its learning strategy is to maximize the margin, which can be formalized as a convex quadratic programming problem, or equivalently, the minimization of the regularized hinge loss function. The learning algorithm for SVM is an optimization algorithm for solving the convex quadratic programming problem.For a given training dataset $T = \{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\}$. Where $x_i \in R^n$ is the feature vector. $y_i \in \{-1, +1\}$ is the category label, and the optimization goal of SVM is to minimize the objective function under the constraint to obtain the optimal solution. By using the Lagrange multiplier method, we can obtain the normal vector of the hyperplane $w$and the bias b, and finally construct the classification decision function:

$$f(x) = sign(w_i \cdot x + b) \tag{6}$$

To solve the nonlinear classification problem, SVM introduces the kernel function $K(x, z)$ to perform feature mapping, mapping the data from the low-dimensional space to the high-dimensional feature space, so that it becomes linearly separable in the high-dimensional space. The objective function of the nonlinear SVM is represented by the kernel function $K(x_i, x_j)$, and the optimal Lagrange multiplier α_i is solved through quadratic programming (QP). The final

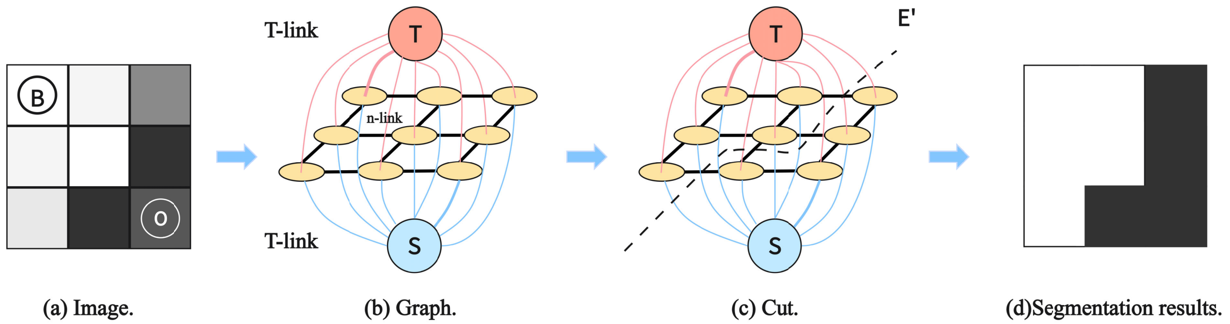


FIGURE 4. Graph-cut Structural Image.

classification decision function is:

$$f(x) = sign(\sum_{i=1}^{N} \alpha_i^* y_i K(x_i, x) + b^*) \tag{7}$$

Among them, α_i^* is the optimal solution, and b^* is the bias term. SVM solves linear and nonlinear classification problems effectively by maximizing the margin strategy and kernel function techniques. Its optimization process depends on the solution of quadratic programming, and it achieves the linear separability of data in high-dimensional feature space through the kernel function, thus showing high classification performance in many real-world applications. Fig. 5 shows the schematic diagram of how SVM performs segmentation.

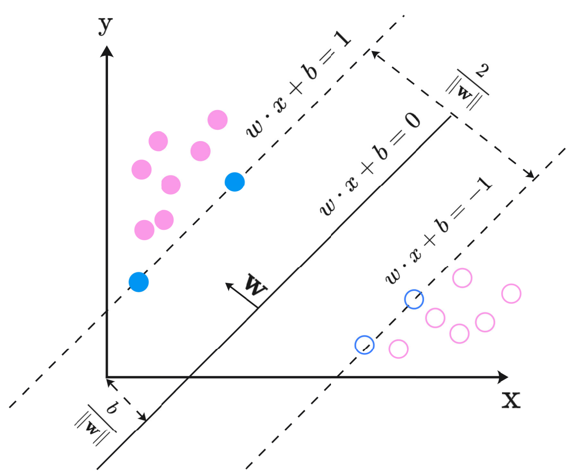


FIGURE 5. Diagram of SVM segmentation.

D. DISEASE DATA AUGMENTATION AND DATA SET CREATION

In this study, the collected tea leaf disease image dataset contains a limited number of images for each type of disease, making it a typical small-sample image recognition task. However, deep learning-based recognition methods require a sufficient number of samples to learn the features of leaf disease images in order to achieve high recognition accuracy. Therefore, this paper proposes an improved Conditional Generative Adversarial Network (IC-GAN), which improves the stability of the network by partitioning the data manifold into overlapping neighborhoods. The IC-GAN network can generate supervised tea leaf images consistent with given conditions, avoiding the generation of single-type tea leaf disease images and effectively expanding the tea leaf disease image dataset. A Generative Adversarial Network (GAN) consists of a generator G and a discriminator D. The generator is used to maximize the simulation of the original target data distribution and generate target samples, with the goal of minimizing the discriminator's ability to correctly classify generated data, i.e., the generator wants the discriminator to misclassify its generated data as real data. This can be expressed by the following function:

$$\min_{G} \mathbb{E}_{n\sim p_z(n)} \left[\log\left(l - D\left(G(n)\right)\right)\right] \tag{8}$$

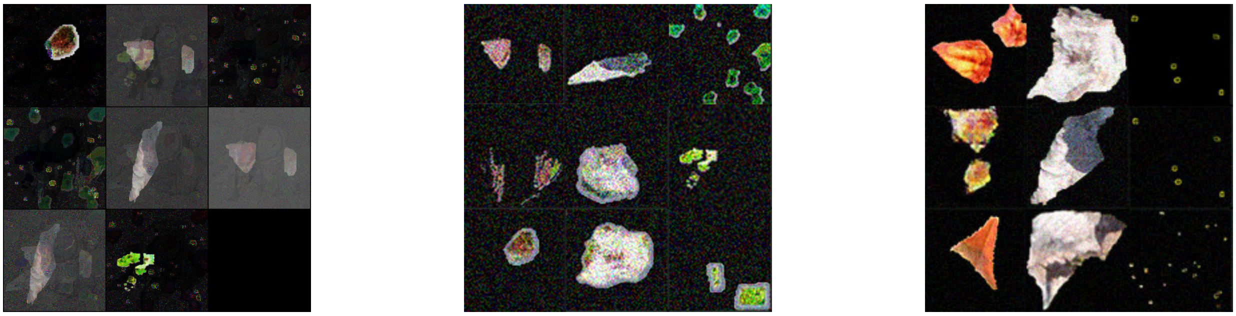
$G(n)$ is the data generated by the generator based on random noise, and $\mathbb{E}_n \sim p_n(n)\left[\log\left(1 - D(G(n))\right)\right]$ represents the probability that the discriminator will classify the generated data as true. That is, the generator's goal is to maximize the discriminator's error rate.

The discriminator is used to distinguish between generated samples and real samples. Its goal is to maximize its ability to classify real data and generated data. This can be expressed by the following function:

$$\max_{D} \mathbb{E}_{i\sim p_{\text{data}}(\text{i})}[\log D(i)] + \mathbb{E}_{m\sim p_z(n)} \left[\log\left(l - D\left(G(n)\right)\right)\right] \tag{9}$$

D(i) is the probability that the discriminator classifies the real data i as true, and $D(G$(n)) is the probability that the discriminator classifies the generated data G(n) as true.The first term $\mathbb{E}_{\text{i}\sim p_{\text{data}}(x)}[\log D(i)]$ represents the probability that the discriminator classifies a real sample as true, and the second term $\mathbb{E}_{n\sim p_z(n)}\left[\log\left(l - D\left(G(n)\right)\right)\right]$ represents the probability that the discriminator classifies a generated sample as false. The overall goal function of GAN is the combination of the generator and discriminator goal functions, forming a Min-Max Problem. The goal function f_{GAN} is as follows:

$$\begin{aligned} &f_{GAN}\left(D, G\right) \\ &\quad = E_{i-P_{Set}(i)}\left[\log D\left(i\right)\right] + E_{n-P_{max}(n)}\left[\log\left(1 - D\left(G\left(n\right)\right)\right)\right] \end{aligned} \tag{10}$$



(a) GAN (b) CGAN (c) IC-GAN

FIGURE 6. Comparison of tea disease images generated by different generative adversarial networks.

During the training process, these two networks will compete with each other and eventually reach a balanced state. However, GAN cannot generate images of tea leaves of specified categories during data augmentation. The Conditional Generative Adversarial Network (C-GAN) is a network that inputs specified condition information (such as the category of tea disease) into the generator and discriminator of GAN to help the generator generate high-quality data that corresponds to the specified condition. Compared to GAN, C-GAN can generate more targeted and controllable data. Let the condition information be the category of tea disease image to be generated, represented by C. This means that the data generated by the generator and the data judged by the discriminator are both dependent on the conditional information. The goal of the discriminator D is to maximize its ability to correctly classify real data and generated data while considering the conditional information.

The goal of the generator G is to minimize the ability of the discriminator to correctly classify the generated data, meaning that the generator wants the discriminator to misclassify its generated data as real data. This is achieved by maximizing $\mathbb{E}_{\mathrm{n}\sim p_{\mathrm{noise}}(\mathrm{n})}\left[\log\left(l - D\left(G(\mathrm{c} \mid \mathrm{C})\right)\right)\right]$, where the generator wants the discriminator to classify the generated data and its conditioning information as true with a high probability. In this way, C-GAN can generate data that is consistent with additional conditioning information during generation, effectively expanding the dataset of four tea image categories studied in this research. The objective function f_{C-GAN} is shown in equation (12), at the bottom of the next page.

Due to the need for additional conditioning information, the training and generation complexity of C-GAN will also increase accordingly, making it prone to training instability problems such as gradient vanishing. Therefore, to solve this problem, this paper uses the conditional adversarial generation network (IC-GAN), which draws on kernel density estimation techniques by dividing the data manifold into overlapping neighborhoods and modeling the distribution of a complex data set as a mixture of local condition distributions, learning the distribution of each data point around it [33]. Specifically, the data set is divided into overlapping neighborhoods consisting of each data point and its nearest neighbors. The goal of the model is to approximate the entire data distribution $p(x)$ as a mixture of these local condition distributions $\mathrm{p}(\mathrm{x}|h_{\mathrm{i}})$:

$$p(x) \approx \frac{1}{M}\sum_{i=l}^{M} p\left(x|h_i\right) \tag{11}$$

$\mathbf{h}_i$ is the feature vector of the ith data point, and M is the total number of data points. To extract instance features, IC-GAN uses a parameterized feature extraction function $\mathbf{f}_\phi$, which maps each data point $\mathbf{x}$ to a feature vector $\mathbf{h_i} = \mathbf{f}_\phi(\mathbf{x})$. These feature vectors are used to define the neighborhood of each data point and serve as additional inputs for the generator and discriminator. The generator G and discriminator D are optimized through adversarial training. During training, the generator tries to generate samples that are similar to real data points, while the discriminator tries to distinguish between real samples and generated samples. The training objective is defined by the following Min-Max Game Problem, as shown in equation (13), at the bottom of the next page.

The expected log probability of the discriminator on the true data is $\mathbb{E}_{x\sim p(x),x_n\sim U(A_i)}[\log D(x_n, f_\phi(x))]$. The expected log probability $\mathbb{E}_{x\sim p(x),z\sim p(z)}[\log(1 - D\left(G(z, f_\phi(x)), f_\phi(x)\right))]$ of the discriminator on the generated fake data by the generator G is the log probability of the discriminator on the generated data. The generator G tries to minimize $V(D, G)$, i.e. reduce the discriminator's ability to correctly classify real and generated data. Meanwhile, the discriminator D tries to maximize $V(D, G)$, i.e. improve its ability to distinguish real and generated data. In this way, the generator and the discriminator compete with each other during training, ultimately reaching a Nash equilibrium, at which neither party can improve its utility by unilaterally changing its strategy.

Fig. 6 shows a comparison of different GAN used to generate defect images, and it can be seen that the images generated by IC-GAN have relatively high quality and are suitable for training models. We used IC-GAN to generate 1,900 new images for each type of tea image as described earlier, and combined the generated images with the real images collected to obtain an enhanced tea image dataset. There are 2,000 images of each type of tea, for a total of 8,000 images. Using IC-GAN for image enhancement solves the problem of small sample recognition and lays a data foundation for accurate image recognition of tea diseases in the subsequent process. At the same time, the enhanced data for each type of tea image was divided into a training set (6,400 images) and a test set (1,600 images) at a ratio of 8:2.

E. METHODS FOR EXTRACTING AND RECOGNIZING PATHOLOGICAL IMAGE FEATURES

After enhancing the data volume of segmented disease image data through IC-GAN, deep learning methods can be used to identify the type of disease. Traditional neural network methods have low recognition accuracy, while classification methods such as SVM and Random Forest (Random Forest, RF) [34] require manual feature extraction from disease image texture, color or shape, and modeling takes a long time and feature extraction is difficult. Deep learning methods have good results in recognition tasks and do not require manual feature extraction. Therefore, this study uses deep learning methods to identify tea disease types. As the task becomes more complex, the performance of many previously widely used deep learning models [35] (such as AlexNet, VGG, and ResNet) has reached a bottleneck. Increasing depth or width is usually the simplest way to obtain high-quality deep network models, but this design approach often leads to overfitting and gradient disappearance problems. The emergence of the Inception model made it possible to maintain the sparsity of the network structure and take advantage of the high computational performance of dense matrices [28]. Compared with other deep learning networks, Inception V3 has two significant improvements. First, Inception V3 uses parallel small filters to decompose the large filters in convolution, which saves more computing resources and improves training speed. And to reduce the model's excessive confidence in the training labels and enhance the model's generalization ability, Inception V3 adopts the label smoothing regularization (Label Smoothing Regularization, LSR) technology, which modifies the loss function of the model by the following equation (14):

$$q'(k) = (I - \epsilon)\delta_{k,y} + \frac{\epsilon}{K} \tag{14}$$

$q'(k)$ is the true label distribution, $\delta_{k,y}$ is a Dirac δ function, which is 1 when $k = y$ and 0 otherwise, $u(k)$ is the prior uniform distribution of labels, and ϵ is a smoothing parameter between 0 and 1. Additionally, Inception V3 employs multi-scale convolutional kernels to extract multi-scale features from the input image. By fusing these features, the model's recognition accuracy and robustness are improved [36].

This study designs a Dilated Convolutional Neural Network (IDCNN) embedded in Inception V3 for tea disease recognition. The dilated convolution introduces a dilation rate parameter, which allows the same-sized convolution kernel to have a larger receptive field. Consequently, dilated convolution requires fewer parameters compared to standard convolution with the same receptive field size. Fig. 8 illustrates the structure of the IDCNN network, which consists of 52 layers, including 6 convolution layers (Conv1 to Conv6), 3 pooling layers (Pooling1 to Pooling3), and 1 Softmax classifier layer. The remaining 42 layers are composed of 3 Inception blocks. Block 1 contains 3 modules with 9 layers; Block 2 contains 5 modules with 23 layers; and Block 3 contains 3 modules with 10 layers.

Unlike traditional Inception networks, IDCNN replaces the first convolution layer (Conv1) with a dilated convolution layer (Dilated Conv1), uses a convolution without padding for Conv3 instead of the original convolution, and replaces the fully connected layer for regression with a convolution layer (Conv6). All convolution layers use ReLU activation functions and batch normalization, which help mitigate overfitting and improve training efficiency. The number of convolution kernels for Conv1, Conv2, Conv3, Conv4, Conv5, and Conv6 are 32, 32, 64, 80, 192, and 1000, respectively. The kernel size for Dilated Conv1, Conv2, Conv3, and Convs is 3×3 dpi, while the kernel size for Conv4 and Conv6 is 1×1. The pooling size for Pooling1 and Pooling2 is 3×3 dpi, and the pooling window size for Pooling3 is 7×7.

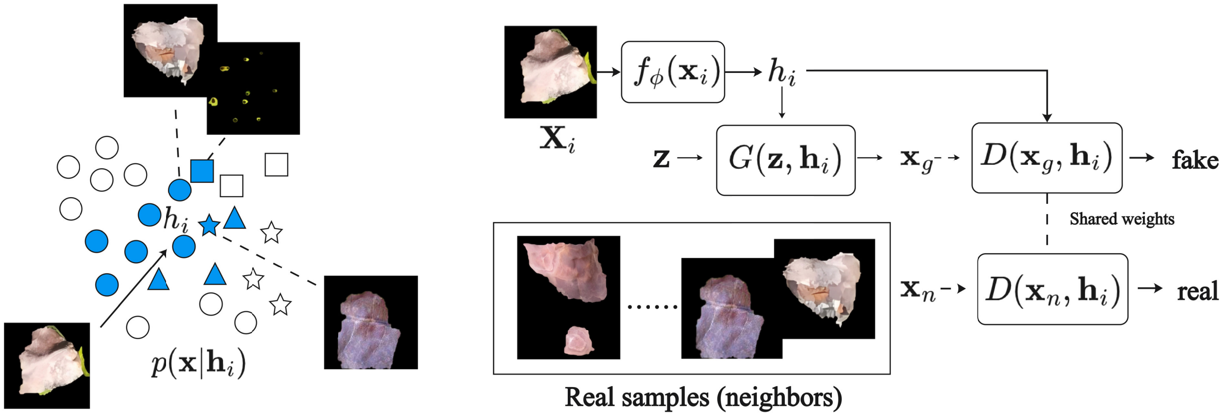
IV. RESULTS

The IC-GAN and IDCNN deep learning models were trained using the TensorFlow framework on two Nvidia Tesla P100 (16 GB memory) graphics cards. The training strategies and hyperparameters of these deep learning models were determined based on preliminary experiments. Specifically, the batch gradient descent algorithm was used to optimize the network parameters. The batch size of the IC-GAN and IDCNN models was set to 8 and 24, respectively. The initial learning rate was set to 0.001, which was reduced by 0.1 every 20 epochs. The maximum training epoch was set to 200.

This study conducted two experiments to verify the performance of the proposed tea leaf disease image segmentation and recognition method. Specific indicators used include accuracy (A), recall rate (R), and F1 score (F). The Equations for calculating these indicators are shown as equations (15)-(17). To further evaluate the efficiency of the proposed methods, the computational complexity in terms of parameters (Millions) and floating-point operations (FLOPs in Billions) was also analyzed. These metrics provide insights into the computational resources required by the models, highlighting their suitability for practical applications with

$$f_{C-GAN}(D, G) = E_{i-P_{\text{Sat}}(i)}\left[\log D(i \mid C)\right] + E_{n-P_{\text{nocia}}(n)}\left[\log\left(1 - D\left(G\left(n \mid C\right)\right)\right)\right] \tag{12}$$

$$\min_{G}\max \mathbb{E}_{x\sim p(x),x_n\sim U(A_i)}[\log D(x_n, f_\phi(x))] + \mathbb{E}_{x\sim p(x),z\sim p(z)}[\log(1 - D\left(G(z, f_\phi(x)), f_\phi(x)\right))] \tag{13}$$



(a) Neighborhood A_i of instance h_i (b) Schematic illustration of the IC-GAN workflow

FIGURE 7. Overview of IC-GAN for Tea Leaf Disease Detection (a)The generator aims to produce realistic images that resemble the neighbors of instance features $\mathbf{h}_i$, determined in the embedding space using cosine similarity. The figure illustrates several neighbors. Note that images within the same neighborhood may represent different classes, indicated by varying shapes. (b) Conditioned on instance features $\mathbf{h}_i$ and noise $\mathbf{z}$, the generator G generates synthetic samples $\mathbf{x}_g$. These generated samples and real samples (neighboring instances $\mathbf{x}_n$ are input into the discriminator D, which is also conditioned on $\mathbf{h}_i$. The discriminator distinguishes between generated "fake" samples and real samples while sharing weights.

limited computational power.

$$\mathrm{A} = \frac{TP}{TP + FP} \tag{15}$$

$$R = \frac{TP}{TP + FN} \tag{16}$$

$$F = \frac{2PR}{P + R} \tag{17}$$

A. LESION IMAGE SEGMENTATION EXPERIMENT

For the segmentation of tea disease images under uneven illumination and field environment, this study proposed a two-stage disease image segmentation method and compared it with the Segment Anything Model (SAM) [37], the K-means algorithm, and SVM to verify the superiority of the proposed segmentation method. The segmentation results of the four segmentation methods on the typical sample images of the three types of tea diseases studied in this paper are shown in Fig. 9. From Fig. 9, it can be seen that, due to uneven local illumination on the leaves, SAM failed to separate the boundary between the shadowed and normal areas, leading to excessive segmentation of both the healthy areas around the disease spots and the background. In the K-means segmentation results, large background areas were not removed because the K-means algorithm performs poorly under uneven illumination. Compared to SAM and K-means, SVM handles uneven lighting and cluttered backgrounds better, but it suffers from some missed disease spots. In contrast, the two-stage segmentation method based on Graph-Cut and SVM proposed in this study successfully segments all disease regions, with only minor segmentation errors in a few pixels, demonstrating its ability to provide high-quality tea disease images for further precise disease recognition.

Using the four segmentation methods mentioned above to segment the disease images and the original tea disease images as the training samples of the IDCNN model, we indirectly illustrate the necessity of image segmentation for tea disease images. Based on this, the recognition accuracy rate of the IDCNN model for the disease images and the original disease images in the test set of three types of tea disease is shown in Table 2. It can be seen that separating the disease images from the original tea disease images using different graph segmentation methods can greatly improve the recognition accuracy rate of the IDCNN model. Among them, the IDCNN model with the two-stage segmentation method based on Graph-Cut and SVM has the highest recognition accuracy rate for the three disease images, and has a significant improvement over the other three segmentation methods, which also proves the superiority of the two-stage segmentation method proposed in this paper for tea disease image segmentation task. The IDCNN model trained with the disease images separated by the two-stage segmentation method has an average recognition accuracy rate of 22.27% higher than the IDCNN model trained with the original tea disease images. Therefore, it is necessary to segment the disease images from the original tea disease images before image enhancement and image recognition.

B. LESION IMAGE RECOGNITION EXPERIMENT

1) IDCNN MODEL RECOGNITION RESULTS BASED ON DIFFERENT DATA AUGMENTATION METHODS

To address the small sample problem in accurate identification of tea diseases, this paper proposes an image enhancement method for disease images using IC-GAN to expand the data volume of partitioned tea disease images. To verify the superior performance of IC-GAN in expanding tea disease image data, it is compared with classic data augmentation methods such as rotation and translation, GAN network, and C-GAN network. The training sets of tea disease image datasets expanded by these four data

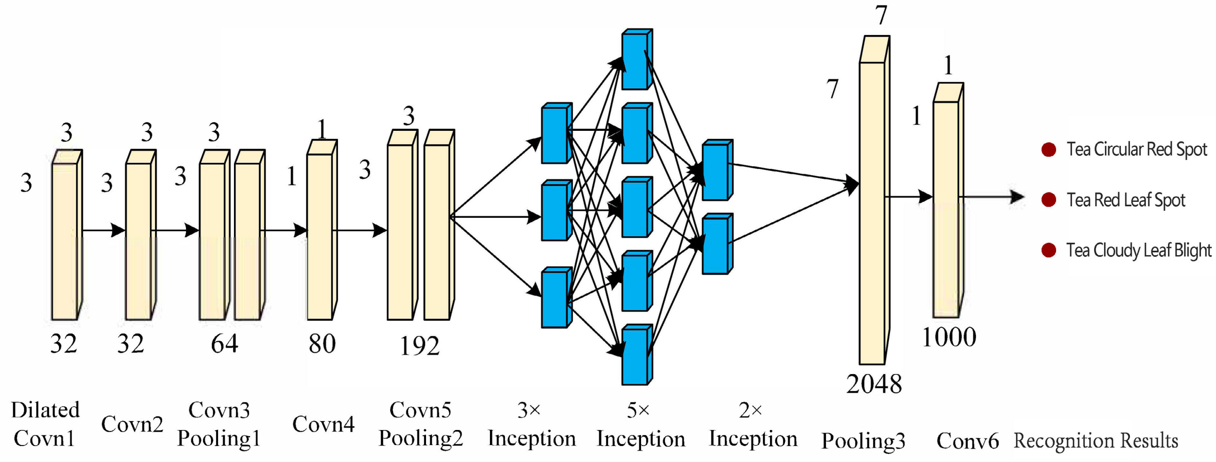


FIGURE 8. **IDCNN Model Structure.**

augmentation methods are input into the IDCNN recognition model proposed in this paper for training. The recognition accuracy of IDCNN model on the test set samples is used as an evaluation index of the performance of the four data augmentation methods. The recognition results of the IDCNN model are shown in Fig. 10.

From Fig.10 it can be seen that the IDCNN model trained with the original disease image dataset without data augmentation has the lowest accuracy in identifying three tea diseases. This shows that in the context of small sample size, expanding the scale of tea disease image dataset through data augmentation libraries is indeed helpful in improving disease recognition accuracy. Compared with non-data augmentation, the use of rotation and translation for image enhancement can slightly improve the accuracy of recognition. After image enhancement with GAN and C-GAN, the average accuracy of IDCNN model in identifying three tea diseases is 86.52% and 88.79%, which is significantly better than the rotation and translation method, because GAN can increase the diversity of the expanded samples compared with rotation and translation. With IC-GAN, the introduction of gradient penalty in the discriminator has reconstructed the loss function, so the average accuracy of IDCNN model in identifying tea circular red spots disease, tea red leaf spot disease, and tea cloudy leaf spot disease is 97.33%, 95.42%, and 95.75%, all of which are greater than 95%. IC-GAN can not only generate tea disease images of specified categories, but also generate high-quality images from a limited number of original disease images. Therefore, the IC-GAN proposed in this paper is an effective method for enhancing tea disease images.

2) ABLATION STUDY AND DIFFERENT DATASETS

To comprehensively validate the role of the IC-GAN module in the tea disease recognition method, we designed the following ablation experiments on our self-built tea disease dataset to evaluate the image generation quality and classification task performance. In this configuration, we directly use the basic IDCNN model for classification, testing the fundamental performance of the classification network as a benchmark. Additionally, to validate the effect of IC-GAN on mitigating the small sample problem, we compared the model's performance under different data scales and conducted a fine-grained analysis, such as the classification performance of small disease spot areas. To further validate the contribution of the IC-GAN module and other core components to the overall method's performance, we designed various experimental configurations to compare the performance under different module combinations. Table 3 summarizes the classification results under each experimental configuration.

Furthermore, we included apple leaf disease data (4 classes) from the Plant Village dataset and tea disease data (8 classes) from the Tea Sickness Dataset to broaden the scope and reliability of the evaluation. The PlantVillage is a plant disease image database commonly used as a base dataset for research on crop and plant diseases. The images in this database were captured in a laboratory, and currently it contains 54,305 plant disease leaf images from 13 plants with 26 disease categories. The data on apple leaf disease used in this study consists of 3,171 images from 4 categories. The Tea Sickness Dataset contains tea samples of 7 common tea diseases, with a total of 885 images collected from the Johnstone Boiyon Farm in the Koiwa region, Bomet County.

The study designed four model configuration schemes, including the basic IDCNN model, an IDCNN model combined with two-stage segmentation technology, an IDCNN model enhanced with IC-GAN technology, and our proposed IDCNN model combining both two-stage segmentation and IC-GAN enhancement. The results show that on a self-built dataset, the configuration combining two-stage segmentation and IC-GAN enhancement significantly improved recognition performance, with accuracy increasing from 75.63% to 96.17%, recall rate rising from 76.56% to 96.45%, and F1 score improving from 74.98% to 96.38%. However, on the Plant Village dataset, which has a larger scale with 3,171 images, the effect of IC-GAN enhancement on

TABLE 2. Comparison of recognition accuracy of IDCNN models based on lesion images and original tea disease images.

IDCNN Model Training Samples	tea circular red spot	tea red leaf spot	tea cloudy leaf blight
Original Tea Disease Image	59.42%	47.78%	52.89%
SAM	67.57%	65.42%	64.37%
K-means Algorithm	69.81%	67.7%	68.69%
SVM	27.80%	59.29%	37.85%
Two-Stage Segmentation Method	78.33%	72.89%	75.67%

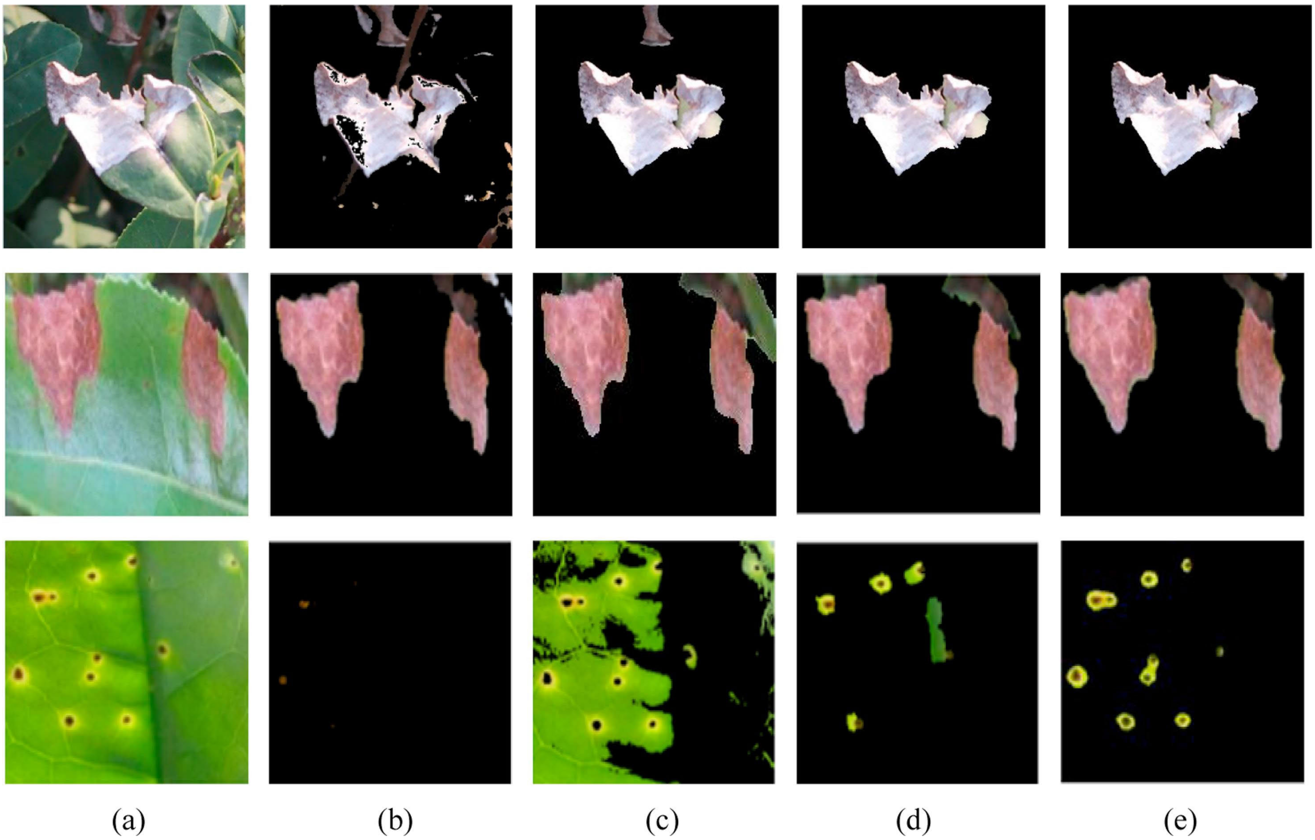


FIGURE 9. Tea disease image segmentation results. (a) is the original tea disease image; (b) is the SAM segmentation result; (c) is the K-means segmentation result; (d) is the SVM segmentation result; (e) is the two-stage segmentation method based on Graph-Cut and SVM segmentation result.

model performance was limited. The basic IDCNN model achieved an accuracy of 89.33%, a recall rate of 88.45%, and an F1 score of 89.32%, while the best configuration only improved accuracy to 94.53%, recall to 95.53%, and F1 score to 93.01%. This phenomenon may be due to the relatively high sample size of the Plant Village dataset, which limits the contribution of IC-GAN enhancement in further improving feature discrimination capability. On the Tea Sickness dataset, the configuration combining two-stage segmentation and IC-GAN enhancement achieved the best performance, with accuracy, recall, and F1 score reaching 97.66%, 97.36%, and 96.98%, respectively, fully demonstrating the effectiveness of these techniques in improving the model's recognition accuracy. In conclusion, this study shows that two-stage segmentation and IC-GAN enhancement technologies perform outstandingly in improving recognition performance on small-scale datasets, but their contribution is limited on large-scale datasets, possibly due to the sample diversity inherent in the dataset itself.

C. COMPARISON OF RECOGNITION EFFECTS OF DIFFERENT RECOGNITION METHODS

From Fig. 11, it can be observed that the accuracy rapidly increases and the loss rapidly decreases in the early stages of training. As the training progresses, the accuracy tends to stabilize, and the loss curve also tends to stabilize, indicating that the model may have converged. Overall, this figure shows that the IDCNN model performs well during the training process of two-stage segmentation, with improved accuracy and reduced loss, and maintains a lower loss on the test set, indicating that the model has good generalization ability.Fig. 12 shows a comparison of traditional machine learning methods (such as RF, SVM) and deep learning methods (IDCNN, VGG16) in the accuracy of tea disease image recognition. From Fig. 12, it can be seen that the IDCNN and VGG16 models have significantly higher accuracy in recognizing three types of tea disease than RF and SVM. Furthermore, compared with VGG16, the proposed IDCNN model has an average improvement of 10%

TABLE 3. Performance metrics for different model configurations across various datasets.

Dataset	IDCNN	Two-Stage	IC-GAN	Accuracy	Recall	F1
Self-Built Dataset	✓			75.63	76.56	74.98
	✓	✓		80.54	81.63	80.23
	✓		✓	90.63	89.56	90.87
	✓	✓	✓	**96.17**	**96.45**	**96.38**
Tea Sickness Dataset	✓			82.56	82.68	81.55
	✓	✓		86.37	86.77	88.21
	✓		✓	88.35	88.46	89.98
	✓	✓	✓	**97.66**	**97.36**	**96.98**
Plant Village	✓			89.33	88.45	89.32
	✓	✓		91.36	91.36	90.69
	✓		✓	92.65	92.10	91.46
	✓	✓	✓	**94.53**	**95.53**	**93.01**

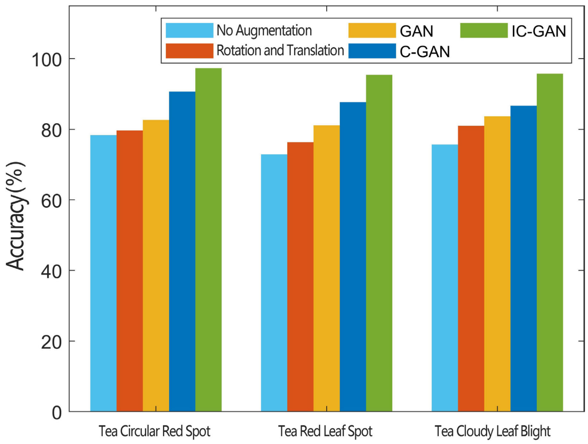


FIGURE 10. Comparison of recognition accuracy of IDCNN models based on different image enhancement methods.

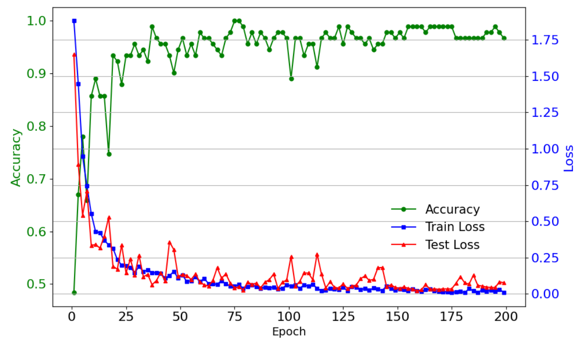


FIGURE 11. The accuracy and loss curves of IDCNN model in the training process after two-stage segmentation.

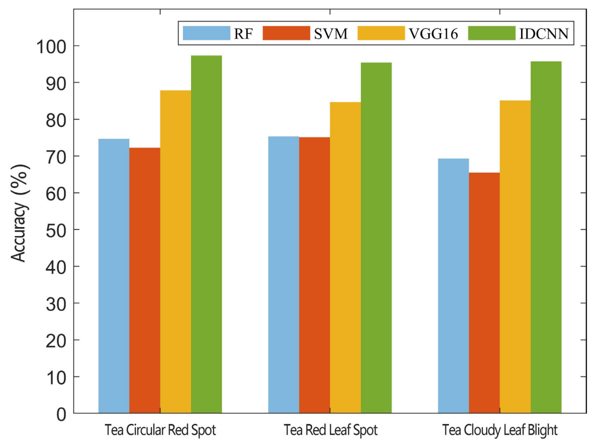


FIGURE 12. Comparison of recognition accuracy of different recognition methods for three types of tea diseases.

in tea disease recognition accuracy. In addition, during the experiment, IDCNN required fewer iterations and shorter training time than VGG16, while VGG16 required more computing resources and convergence time to update its large number of weight parameters. This is because IDCNN's first convolutional layer uses a dilated convolution, which expands the local receptive field and enhances the feature extraction ability of the convolutional layer, reducing the loss of spatial information and internal data structure in the pooling process. At the same time, IDCNN uses convolutional layers to replace the fully connected layer in the logical regression part of the Inception V3, greatly reducing the number of network parameters and saving computing resources.

In order to objectively evaluate the recognition performance of different recognition methods, more detailed evaluation indicators are presented as shown in Table 2. As shown in Table 4, the IDCNN model outperforms almost other methods in terms of recognition performance, with an average accuracy of 96.17%, recall rate of 96.45%, and F1 score of 96.38% for image recognition of the type of tea disease. This demonstrates its superior ability to correctly identify tea diseases. Although Swin-Transformer achieves a slightly higher accuracy of 97.02%, it comes with significantly higher computational costs, requiring 50 million parameters and 8.7 billion FLOPs. In contrast, IDCNN has only 19.0 million parameters and 6.2 billion FLOPs, making it much more computationally efficient while maintaining competitive performance. Furthermore, YOLOv11s also shows a high accuracy of 93.85%, but with 20.1 million parameters and 21.7 billion FLOPs, its computational load is considerably higher than that of IDCNN. The reduction in both parameters and FLOPs for IDCNN translates into faster training times, lower memory usage, and faster inference

TABLE 4. Average recognition performance of different recognition methods for tea diseases.

Model	Accuracy/%	Recall/%	F1-Score/%	Parameters (Millions)	FLOPs (Billions)
YOLOv5s	86.68	85.87	86.60	7.2	16.5
VGG16	90.45	90.80	90.80	138.4	15.3
Mask-R CNN	92.24	92.65	92.10	44.0	17.0
SE-DenseNet-FL	93.56	93.33	93.69	27.2	16.0
YOLOv11s	93.85	93.34	94.17	9.4	21.7
ViT (Vision Transformer)	96.10	96.31	96.10	86.0	50.0
Swin-Transformer	97.02	97.49	96.88	50	8.7
IDCNN (Ours)	**96.17**	**96.45**	**96.38**	**19.0**	**6.2**

speeds, which makes it particularly suitable for resource-constrained environments, such as mobile or edge devices. Additionally, experiments confirm that IDCNN can achieve near-identical recognition performance even with fewer parameters, indicating its robustness and efficiency.

V. CONCLUSION

This study aims to achieve accurate recognition of images of tea plant disease captured in field environments. The diseases considered for recognition include Tea Leaf Curl Disease, Tea Red Leaf Spot Disease, and Tea Clouded Leaf Wilt Disease. The study proposes a small-sample recognition method for tea plant diseases based on image segmentation, image enhancement, and image recognition.

To address the challenges of complex backgrounds and uneven lighting in field-captured tea plant images, a two-stage segmentation method based on Graph-Cut and SVM is introduced. This method has strong discrimination ability. Under the experimental conditions, the recognition accuracy of the IDCNN model after segmentation improved by 18.91%, 25.11%, and 22.78% for the three types of tea diseases compared to before segmentation. This demonstrates that the proposed method can accurately segment the diseased areas on tea leaves with almost no manual intervention, showing superior performance.

Next, to address the problem of limited labeled tea disease image samples, the study proposes using the IC-GAN network to enhance the segmented diseased images and generate high-quality images of tea plant diseases. Finally, the IDCNN model is used to recognize the disease types of the enhanced tea disease images.

The experimental results show that the proposed small-sample tea disease recognition method can accurately identify the disease types in tea images taken in field scenarios. The recognition accuracy for Tea Leaf Curl Disease, Tea Red Leaf Spot Disease, and Tea Clouded Leaf Wilt Disease is 97.33%, 95.42%, and 95.75%, respectively, which is significantly higher than other recognition methods. This is due to: (1) the two-stage segmentation method removes background interference and enhances the diseased area features; (2) the IC-GAN network expands the tea disease image sample size and increases the diseased image features; (3) the IDCNN model extracts deep, effective key features from the processed diseased areas using Inception and dilated convolution methods.

This study proposes an effective method suitable for timely plant leaf disease recognition in agricultural fields, which has significant practical value. Although the proposed small-sample tea disease recognition algorithm can significantly improve the recognition accuracy of tea disease images captured in complex field environments, it still has limitations. Tea disease recognition is still influenced by many factors. There are still many challenges in terms of dataset, lightweight models, and practical applications that need to be overcome and improved.

Future work will focus on extending the proposed method to recognize other crop diseases, broadening the applicability of the model to various agricultural contexts. Additionally, there are plans to deploy the IDCNN model in real-time monitoring systems, providing farmers with intelligent on-site disease recognition services. This will not only enhance the precision of pest and disease management but also contribute to the sustainability of crop production. Furthermore, the method will be explored for integration with mobile software that can offer efficient disease detection on-site, assisting farmers in early diagnosis and rapid intervention. Multi-spectral and multi-temporal data could also be incorporated in future research to enhance the robustness of the model under various environmental conditions, further improving disease detection performance.

REFERENCES

[1] G. Wei, J. Huang, and J. Yang, "The impacts of food safety standards on China's tea exports," *China Econ. Rev.*, vol. 23, no. 2, pp. 253–264, Jun. 2012.

[2] J. Li, S. Han, R. Xu, X. Zhang, J. Liang, M. Wang, and B. Han, "Insight into the effect of geographic location and intercropping on contamination characteristics and exposure risk of phthalate esters (PAEs) in tea plantation soils," *J. Integrative Agricult.*, vol. 23, no. 11, pp. 3896–3911, Nov. 2024.

[3] Y. Peng et al., "Where is tea grown in the world: A robust mapping framework for agroforestry crop with knowledge graph and sentinels images," *Remote Sens. Environ.*, vol. 303, Mar. 2024, Art. no. 114016.

[4] S. Xie, C. Wang, C. Wang, Y. Lin, and X. Dong, "Online identification method of tea diseases in complex natural environments," *IEEE Open J. Comput. Soc.*, vol. 4, pp. 62–71, 2023.

[5] W. Li, W. Zhan, W. Zhou, T. Han, P. Wang, H. Liu, M. Xiong, and Y. Sun, "Research and application of lightweight YOLOv7-tsa network in tea disease detection and identification," *J. Henan Agricult. Sci.*, vol. 52, no. 5, p. 162, 2023.

[6] G. Hu, H. Wang, Y. Zhang, and M. Wan, "Detection and severity analysis of tea leaf blight based on deep learning," *Comput. Electr. Eng.*, vol. 90, Mar. 2021, Art. no. 107023.

[7] S. Lee, S. Lin, and S. Chen, "Identification of tea foliar diseases and pest damage under practical field conditions using a convolutional neural network," *Plant Pathol.*, vol. 69, no. 9, pp. 1731–1739, Dec. 2020.

[8] Y. Chen, X. Wu, Z. Zhang, J. Yan, F. Zhang, and L. Yu, "Method for identifying tea diseases in natural environment using improved YOLOv5s," *Trans. Chin. Soc. Agricult. Eng.*, vol. 39, no. 24, pp. 185–194, 2023.

[9] L. Shuai, J. Mu, X. Jiang, P. Chen, B. Zhang, H. Li, Y. Wang, and Z. Li, "An improved YOLOv5-based method for multi-species tea shoot detection and picking point location in complex backgrounds," *Biosystems Eng.*, vol. 231, pp. 117–132, Jul. 2023.

[10] Z. Li, Y. Li, C. Yan, P. Yan, X. Li, M. Yu, T. Wen, and B. Xie, "Enhancing tea disease identification with lightweight mobileNetv2," College Big Data Inf. Eng., Guizhou Univ., Guiyang, China, Tech. Rep., 2023, doi: 10.32604/cmc.2024.051526.

[11] X. Deng and C. Photong, "Evaluation of tea leaf disease identification based on convolutional neural networks VGG16, ResNet50, and DenseNet169 image recognitions," in *Proc. 12th Int. Electr. Eng. Congr. (iEECON)*, Mar. 2024, pp. 01–04.

[12] K. Indhumathi and K. S. Kumar, "Efficient prediction of seasonal infectious diseases using hybrid machine learning algorithms with feature selection techniques," *Int. J. Artif. Intell. Tools*, vol. 32, no. 7, Nov. 2023, Art. no. 2350044.

[13] S. Mukhopadhyay, M. Paul, R. Pal, and D. De, "Tea leaf disease detection using multi-objective image segmentation," *Multimedia Tools Appl.*, vol. 80, no. 1, pp. 753–771, Jan. 2021.

[14] M. Xin and Y. Wang, "Image recognition of crop diseases and insect pests based on deep learning," *Wireless Commun. Mobile Comput.*, vol. 2021, no. 1, Jan. 2021, Art. no. 5511676.

[15] S. K. P. Kumar, M. G. Sumithra, and N. Saranya, "Particle swarm optimization (PSO) with fuzzy c means (PSO–FCM)–based segmentation and machine learning classifier for leaf diseases prediction," *Concurrency Computation: Pract. Exper.*, vol. 33, no. 3, p. 5312, Feb. 2021.

[16] M. A. Khan, M. I. U. Lali, M. Sharif, K. Javed, K. Aurangzeb, S. I. Haider, A. S. Altamrah, and T. Akram, "An optimized method for segmentation and classification of apple diseases based on strong correlation and genetic algorithm based feature selection," *IEEE Access*, vol. 7, pp. 46261–46277, 2019.

[17] J. Shin, Y. K. Chang, B. Heung, T. Nguyen-Quang, G. W. Price, and A. Al-Mallahi, "Effect of directional augmentation using supervised machine learning technologies: A case study of strawberry powdery mildew detection," *Biosyst. Eng.*, vol. 194, pp. 49–60, Jun. 2020.

[18] S. Zhang, Y. Zhu, Z. You, and X. Wu, "Fusion of superpixel, expectation maximization and PHOG for recognizing cucumber diseases," *Comput. Electron. Agricult.*, vol. 140, pp. 338–347, Aug. 2017.

[19] U. Mokhtar, M. A. S. Ali, A. E. Hassenian, and H. Hefny, "Tomato leaves diseases detection approach based on support vector machines," in *Proc. 11th Int. Comput. Eng. Conf. (ICENCO)*, Dec. 2015, pp. 246–250.

[20] S. D. Khirade and A. B. Patil, "Plant disease detection using image processing," in *Proc. Int. Conf. Comput. Commun. Control Autom.*, Feb. 2015, pp. 768–771.

[21] C. M. Onyango and J. A. Marchant, "Segmentation of row crop plants from weeds using colour and morphology," *Comput. Electron. Agricult.*, vol. 39, no. 3, pp. 141–155, Aug. 2003.

[22] F. Quan, "Hardware accelerated watershed based musk melon image segmentation," *J. Shanxi Agricult. Sci.*, vol. 2011, no. 7, pp. 726–728, 2011.

[23] J. Chen, J. Chen, D. Zhang, Y. Sun, and Y. A. Nanehkaran, "Using deep transfer learning for image-based plant disease identification," *Comput. Electron. Agricult.*, vol. 173, Jun. 2020, Art. no. 105393.

[24] C. K. Rai and R. Pahuja, "Classification of diseased cotton leaves and plants using improved deep convolutional neural network," *Multimedia Tools Appl.*, vol. 82, no. 16, pp. 25307–25325, Jul. 2023.

[25] Z. Li, J. Sun, Y. Shen, Y. Yang, X. Wang, X. Wang, P. Tian, and Y. Qian, "Deep migration learning-based recognition of diseases and insect pests in Yunnan tea under complex environments," *Plant Methods*, vol. 20, no. 1, p. 101, Jul. 2024.

[26] Y. Shao, G. Xuan, Y. Zhu, Y. Zhang, H. Peng, Z. Liu, and J. Hou, "Research on automatic identification system of tobacco diseases," *Imag. Sci. J.*, vol. 65, no. 4, pp. 252–259, May 2017.

[27] S. Ramesh and D. Vydeki, "Recognition and classification of paddy leaf diseases using optimized deep neural network with Jaya algorithm," *Inf. Process. Agricult.*, vol. 7, no. 2, pp. 249–260, Jun. 2020.

[28] D. Krisnandi, H. F. Pardede, R. S. Yuwana, V. Zilvan, A. Heryana, F. Fauziah, and V. P. Rahadi, "Diseases classification for tea plant using concatenated convolution neural network," *COMMIT Commun. Inf. Technol. J.*, vol. 13, no. 2, pp. 67–77, Oct. 2019.

[29] U. Afzaal, B. Bhattarai, Y. R. Pandeya, and J. Lee, "An instance segmentation model for strawberry diseases based on mask R-CNN," *Sensors*, vol. 21, no. 19, p. 6565, Sep. 2021.

[30] K. He, D. Wang, M. Tong, and X. Zhang, "Interactive image segmentation on multiscale appearances," *IEEE Access*, vol. 6, pp. 67732–67741, 2018.

[31] G. Hu, H. Wu, Y. Zhang, and M. Wan, "A low shot learning method for tea leaf's disease identification," *Comput. Electron. Agricult.*, vol. 163, Aug. 2019, Art. no. 104852.

[32] Y. Wang, Y. Wei, X. Qian, L. Zhu, and Y. Yang, "AINet: Association implantation for superpixel segmentation," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2021, pp. 7078–7087.

[33] A. Casanova, M. Careil, J. Verbeek, M. Drozdzal, and A. R. Soriano, "Instance-conditioned gan," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 34, 2021, pp. 27517–27529.

[34] M. Bhagat and D. Kumar, "Performance enhancement of kernelized SVM with deep learning features for tea leaf disease prediction," *Multimedia Tools Appl.*, vol. 83, no. 13, pp. 39117–39134, Oct. 2024.

[35] Y. Wang, R. Xu, D. Bai, and H. Lin, "Integrated learning-based pest and disease detection method for tea leaves," *Forests*, vol. 14, no. 5, p. 1012, May 2023.

[36] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2818–2826.

[37] M. Henila and P. Chithra, "Segmentation using fuzzy cluster-based thresholding method for apple fruit sorting," *IET Image Process.*, vol. 14, no. 16, pp. 4178–4187, Dec. 2020.

JI LI is currently pursuing the master's degree with the College of Biological Science and Technology, Hunan Agricultural University. His research interests include image processing, deep learning, and plant protection.

CHENYI LIAO received the bachelor's degree in management from the School of Environmental and Chemical Engineering, Foshan University. Currently, he is working with Seed Breeding Company. His main interests include tea industry research, tea production management, and tea planting area division.