Research Article

A System for Identifying Deficiency Elements in Apple Leaf Disease Based on Deep Learning

Linping Zheng1, Wanyi Li*, Nanjian Li1, Canming Huang1, Jinyi Cheng1, Yun Kuang1, Xiaoyun Guo1, Zou Ling2

1School of Computer Science, Guangdong University of Education, Guangzhou, Guangdong Province, China
2College of Information and Intelligence, Hunan Agricultural University, Changsha, Hunan Province, China

*Correspondence to: Wanyi Li, PhD, Lecturer, School of Computer Science, Guangdong University of Education, No. 30 Yingbin Avenue West, Huadu District, Guangzhou, 510303, Guangdong Province, China; E-mail: luther1212@163.com

Received: August 11, 2023 Revised: September 8, 2023 Accepted: October 18, 2023 Published: December 25, 2023

Abstract

Background: Plant diseases and pests were natural disasters that seriously threaten agricultural production. Leaf diseases posed a significant threat to the overall productivity and quality of apple orchards. Currently, the diagnosis of plant diseases in apple orchards mainly relied on manual labor, which was both time-consuming and costly. It took a considerable amount of time to train an expert capable of accurately identifying apple diseases and pests, and different crop diseases vary. However, with the use of AI technology, training could be completed in just a few hours to a day, transforming computers into experts in disease and pest identification. Although China became a major apple cultivation nation, the development data from 2021 revealed that traditional advantageous production areas were undergoing dynamic adjustments due to optimization of industrial structure, leading to a decrease in cultivated land. Conversely, the demand for apples in the Chinese market continued to rise, even as the country's population structure gradually trended towards an aging population. This situation underscored the urgent need to replace the laborious and high-cost task of identifying apple foliar diseases in the cultivation process with intelligent agricultural technological solutions. Taking into account the aforementioned factors of the natural environment and socio-cultural considerations and addressing the issue of reduced apple yields due to the high impact of foliar diseases, we initiated study on disease prevention and control during the cultivation process.

Methods: To overcome the objective impact of environmental factors, the system extracted images and performed operations such as enhancement and segmentation. Subsequently, a convolutional neural network was employed to process the enhanced images, extracting features such as color, shape, and texture of apple foliar lesions, which were then summarized into histograms using mathematical methods. These histograms served as the primary reference for distinguishing various diseases. Python was chosen as the programming language for the system, and the recognition system was operated by integrating frontend technologies, backend database design, and other relevant techniques.

Results: Based on a substantial number of experimental results, the apple leaf disease and nutrient
deficiency recognition system could accurately and with extremely high precision identifying major diseases commonly occurring in apple plants, including apple scab, leaf spot disease, rust disease, gray mold disease, and brown spot disease. Moreover, the system was capable of proposing remedies tailored to the specific pain points of each disease. It not only detected apple leaf diseases but also performs nutrient deficiency detection. This holds significant importance in enhancing both the yield and quality of apples.

**Conclusion:** For the commonly occurring foliar diseases in cultivation, the system exhibited the capability to autonomously recognize and provide decision feedback. In contrast to the less efficient and more costly traditional approach of consulting experts, the more accessible apple foliar disease recognition system was poised to replace it as a primary tool for prevention and control assistance. This system will serve as a guiding companion, safeguarding the robust growth of apple plants.

**Keywords:** apple foliar diseases, convolutional neural network; histogram, python

---

**1 INTRODUCTION**

As a major apple producer, China’s high-level development has become the foundation of the country’s apple-related industry chain. However, with a decrease in cultivation area, it is essential to increase the overall apple yield, which requires strict control over the improvement of apple yield per unit area. The most significant factor restricting this improvement is undoubtedly the damage caused by apple leaf diseases[3]. Behind the solution to apple foliar diseases lies the challenge of high-cost and low-efficiency manual identification, constrained by the threat of an aging population and traditional labor expenses. It also represents the determination of numerous regions with low productivity levels to embrace apple cultivation as a path to prosperity. According to statistics from the Food and Agriculture Organization of the United Nations, leaf diseases can lead to losses of approximately 10%-15% of the total apple production. Compared to the labor-intensive and cumbersome traditional methods of identification and control, an intelligent and data-driven apple leaf disease recognition system is undoubtedly more competent and suitable for the task. As shown in Figure 1, these were images of various apple foliar diseases on leaves.

In the current era of global digitization, technologies such as artificial intelligence, big data, and deep learning are thriving and gradually being widely applied in the agricultural sector. Building upon previous research, the functionalities of scanning, recognition, assessment, database enhancement, and assistive decision-making employed in the apple foliar disease recognition system have reached a considerable level of maturity. We intend to integrate these mature technologies with traditional recognition approaches, initiating a comprehensive series of studies aimed at further enhancing the system’s capabilities[2]. The apple leaf disease deficiency recognition system can perform real-time monitoring of the leaves of apple plants throughout their growth cycle. It accurately identifies the type of disease and nutrient deficiencies before they spread to other healthy plants. The system provides feedback to the operators with targeted preventive and control measures, fundamentally ensuring the improvement of apple yield and quality.

By leveraging this intelligent system, farmers can proactively detect and address potential issues, allowing for timely interventions and minimizing the impact of diseases and nutrient deficiencies on apple production. This technology-driven approach optimizes the management of apple orchards and enhances overall productivity and profitability in the apple industry.

Thanks to the widespread adoption of smartphones and the application and advancement of advanced image processing technologies in various electronic smart products, recognition systems have gained greater freedom and flexibility in their operation and use. This has lowered the entry barrier for grassroots growers. Such widespread adoption also has a significant role in expanding the market size after the system’s development. The elimination of additional labor costs for manual identification, along with a reasonable system development cost, aligns with the core purpose of our system’s development[3]. This progress facilitates the seamless integration of our system into traditional orchard operations, providing an intelligent and automated approach to usage. With this system, all they need to do is upload real-time images of diseased apple leaves, and the system will automatically handle image data processing, database operations, and knowledge graph visualization. The system can then provide corresponding
solutions for apple leaf diseases through an assistive decision-making system, essentially addressing the lack of expert resources at the grassroots level.

This technology empowers farmers with access to specialized knowledge and expertise without requiring them to be experts in the field. By utilizing the capabilities of the recognition system and leveraging the vast resources of image data and knowledge, farmers can make informed decisions and take appropriate measures to combat apple leaf diseases effectively. This not only improves their productivity but also contributes to the overall health and sustainability of the apple industry.

2 METHODS

2.1 Research Content

(1) Analyzing the characteristics of apple diseases and nutrient deficiency symptoms on leaves. Based on the predefined feature extraction approach, selecting and optimizing image preprocessing algorithms, and optimizing the best algorithms for each processing step to obtain images with strong contrast between disease spots and background grayscale. Researching image feature extraction algorithms based on color, shape, and texture of disease spots, and finally analyzed these feature parameters.

(2) Establishment of an apple leaf disease identification model. Using MATLAB software, many disease images such as apple rust, apple scab, and leaf spot disease were analyzed and processed. Relevant feature parameters of different diseases were extracted. The leaf segmentation was performed on the original images using this method[4]. The mean and standard deviation of these feature parameters were calculated. Then, by applying certain weights to the feature parameters, a feature discriminant function was established to effectively classify apple leaf diseases statistically.

(3) Identification of apple leaf diseases. Collecting the apple leaf images to be recognized, analyzed, and processed the images using MATLAB software to obtain corresponding feature parameters. Based on the established feature discriminant function, the disease type of the apple leaf can be effectively identified.

(4) Integration of disease recognition model and knowledge Q&A for assisted pest control decision making. After identifying the current disease type, the system provided detailed information about the disease and recommends appropriate pest control measures. This effectively prevented errors in judgment and delayed in
Innovation Forever Publishing Group

2.2 Research Plan

The apple leaf disease detection method comprised four main steps: The first step involved image enhancement; the second step involved image segmentation, where color image segmentation method was employed based on comparative studies; the third step involved feature extraction. For the segmented disease regions, features were extracted in terms of color, texture, and disease spot shape; the fourth step involved the design and implementation of the pattern recognition system algorithm.

2.3 Main Issues to be Addressed

(1) The issue of apple leaf variation and diverse image environments: The visual symptom variations of a single disease in different apple varieties or new varieties cultivated presented a major challenge for computer-based visual disease recognition. These variations arose from differences in natural and image capture environments, such as leaf color and shape, age of infected tissues, uneven image backgrounds, and varying lighting conditions during imaging.

(2) The issue of improving recognition system model accuracy: Addressing the challenge of ambiguous visual symptom identification for a single disease in different apple varieties or new cultivars, models like the Swin Transformer were employed to accurately classify specific leaf images from the test dataset into distinct disease categories. This approach allowed for the identification of individual disease symptoms from various disease conditions present in a single leaf image, thus enhancing the accuracy of the system’s recognition.

(3) The issue of establishing feature discriminant functions based on feature parameters: After constructing an image recognition system, further analysis and processing of the extracted image feature information was necessary to achieve effective discrimination. Therefore, by utilizing leaf feature parameters obtained from the recognition images and employing mathematical tools, a feature discriminant function was established for leaf features. This facilitated the capability to identify leaf diseases based on these features.

(4) The issue of disease image recognition and analysis on the server side: During the design of the recognition system, the process of apple image recognition needed to be analyzed, including image preprocessing, feature extraction, and disease spot recognition steps. Firstly, noise reduction was applied to disease images, followed by edge detection for disease images. Subsequently, effective classification feature parameters were selected and a feature library was constructed for disease images. This led to the selection of disease recognition classifiers, ultimately resulting in disease feature identification.

3 RESULTS AND DISCUSSION

3.1 System Technology

3.1.1 Open CV Open Source Computer Vision

Image processing was the process of using a computer to analyze the image to achieve the desired effect, which generally included image compression, enhancement and restoration, matching description, and recognition. The system mainly used the disease image data for data enhancement. The main expansion methods were as follows: inversion transformation, scale transformation, dislocation transformation and noise disturbance. As shown in Figure 2, these were images of the image processing results.

3.1.2 Feature Extraction Convolutional Neural Network

Convolutional neural network reduced many parameters due to its features of local connection of neurons and weight sharing. Meanwhile, the pooling layer utilized the principle of image local correlation to perform sub-sampling of images to reduce data processing while retaining useful information, which was the most important neural network model in image recognition at present. The process of BP neural network was mainly divided into two stages. The first stage was the forward broadcast of the signal, from the input layer to the hidden layer, and finally to the output layer; the second stage was the back propagation of the error, from the output layer to the hidden layer, and finally to the input layer, and the weight and bias of the hidden layer to the output layer are adjusted in turn. The neural network used the existing data to find the weight relationship between the input and the output (approximate), and then used the weight relationship to simulate. The current disease detection system had not yet achieved relatively accurate results, retaining many parameters. This system also used pooling layers to use the principle of local image correlation to subsample images, reducing the amount of data processing and improving its efficiency and accuracy.

Taken a single channel picture as an example, if the input picture size was 5*5 filter length and width was 3*3, then the data would first propagate forward, each filter would slide along the width and height of the input data, and the input data would be summed by chance, and finally generate a two-dimensional activation graph. The spatial position in the activation diagram represented the reaction of the raw data to this filter. For multi-channel images, the filters on the convolution layer of each channel would generate a two-dimensional activation diagram, and then superimpose in the depth direction according to the unused activation diagram to finally form the output of the convolution layer.
Firstly, four color features of the lesion image were extracted, and then seven features based on sum and difference histograms were extracted. Figure 3 were images recognized by leaf feature of anthracnose and leaf blight. The gray level and difference histogram of the lesion were calculated according to the sum and difference between pixels, in which the energy, entropy and consistency characteristics could reflect the probability distribution of the gray level of the image; the mean feature can reflect the average gray level of the image; the variance feature can reflect the clarity of the image; the contrast feature can reflect the similarity of the sum and difference histogram on the row and column [6]. The results showed that these 11 features reflected the lesion features of the diseased leaf image well and could be used for disease classification recognition.

3.1.3 Pooling Layer
After the convolution layer, the main function of the pooling layer was to reduce sampling, to reduce the dimension of features while retaining the main features, to reduce the amount of data processed by the next layer by compressing the input feature map, and to simplify the computational complexity of the network. Pooling operations were generally divided into two types: Max pooling and average pooling. Maximum pooling was used in many scenarios. As shown in Figure 4, these images were image segmentation results.

3.1.4 Swin Transformer Model
Swin Transformer was a brand-new Transformer model, which mainly used some concepts in CNN structural design (down sampling, dependency, etc.) to redesign Transformer. As shown in Figure 5, this was the structure diagram of the Swin Transformer model.

3.1.5 Activation Function
In practical applications, most of the data was distributed non-linearly. However, computations within neural networks were linear by nature. Therefore, to enable neural networks to effectively address non-linear problems, we incorporated activation functions within the neural network architecture. The role of activation functions was to introduce non-linearity into neural networks, thereby enhancing their expressive capabilities and enabling them to tackle a wider range of complex problems. In artificial neural networks, each neuron node corresponded to specific inputs and outputs. Figure 6 depicted the transmission process of neural data within a neuron in a neural network. The neuron’s node received the output from the previous layer’s neurons as input to its own node. This input was then processed by the activation function, and subsequently, the output value was passed on to the next layer. Different scenarios required the application of different activation functions, each with its own distinct characteristics [7]. Commonly used activation functions include the Sigmoid function and the ReLU function. The Sigmoid function restricted outputs to the range of (0, 1). If the input was a very large negative number, the output became 0. Similarly, if the input was a very large positive number, the output became 1. This activation function prevented data from diverging during propagation and was suitable for processing probability values. On the other hand, the ReLU (Rectified Linear Unit) function outputs 0 when the input was less than 0, and for positive input values, the output was equal to the input. Essentially, it acted as a function that selects the maximum value. This function was well-suited for training deep networks as it helped mitigate the vanishing gradient problem. Choosing the appropriate activation function for different scenarios could accelerate the convergence speed of model training and enhance the learning capability of the model more effectively.

3.1.6 Identify Disease Types
(1) Pattern recognition. There were methods based on decision theory and structured recognition. This topic used the recognition method based on decision theory. The extracted feature parameters were combined with the established feature discriminant function to identify the
(2) Using MATLAB software programming to design user-friendly GUI.

(3) At present, all the identification of disease types was not precise enough. The system segmented the images, then used algorithms to extract the disease spots from apple leaves. Finally, the extracted images were sent to a neural network for further analysis and processing, enabling real-time and convenient identification of apple leaf diseases.

3.2 System Function Design
3.2.1 Overall System Design

Apple leaf disease deficiency identification and classification system, the use of artificial intelligence professional technology, integrated machine learning, deep learning, image processing, data processing, web development, database technology and other high-tech, and brought together professional knowledge in apple disease related fields. The apple disease recognition model was built by neural network, which could identify the apple leaf disease accurately and quickly. In terms of the system providing prevention and control measures to assist decision-making, the system adopted intelligent and professional meant to help apple growers quickly and efficiently judge the current apple disease situation, to reduce the learning cost of growers, and to provide timely and effective treatment programs for growers as auxiliary decision-making, which could improve the application accuracy of apple growers to a certain extent and reduce unnecessary waste of pesticides. This, in turn, reduce the damage caused by diseases to fruit trees and
improve the yield and quality of apples.

The overall structure of apple leaf disease deficiency identification and classification system could be divided into two parts: front-end interaction system and back-end processing system. The front-end interaction system was used for human-computer interaction. After users entered the system, they could use the system functions by selecting corresponding functional modules. The system had two main functions, including disease deficiency identification and disease deficiency classification. When the user used different functions, the corresponding request is sent to the background processing system.

The background processing system took the constructed disease recognition model and disease knowledge map as the core, and took various databases as the data support. It was responsible for processing front-end requests, returning request results and displaying them on the front-end.

3.2.2 System Function Design

In terms of system function design, the apple disease drug application aid decision system primarily comprised two functional modules, which were integrated to provide apple growers with assistance in making decisions regarding drug application. Here was a general overview of each functional module:

The disease and deficiency recognition function utilized deep learning technology to construct an apple leaf disease recognition model. Upon uploading leaf images, the system employed the model to identify diseases, classify disease types and deficiency types, and provided recognition results. These results included detailed information such as the name, introduction, symptoms, occurrence conditions, as well as corresponding prevention and control measures for the identified disease or deficiency.

The disease deficiency classification function utilized the apple leaf disease recognition model to aggregate and categorize the recognition results. It also stored the historical identification information and offered disease cases for users to reference. Users accessing this module could view the overall profile of disease categories, as well as historical identification information which included the time of identification, disease name, description, symptoms, occurrence conditions, and corresponding prevention and treatment measures.

3.2.3 Design and Implementation of Apple Disease Deficiency Recognition Function

The user uploaded the disease image through the disease recognition function. After the system received the disease image input from the front-end page, it initiated the image preprocessing function. The preprocessed images were then fed into the AlexNet-F disease recognition model.
3.2.4 Design and Implementation of Apple Disease Deficiency Classification Function

Accumulate pertinent information regarding deficiencies in apple diseases, constructed a comprehensive dataset, utilized machine learning techniques to classify the data, and presented the outcomes to the users. Developed a robust database to store personal historical information, enabling the retrieval of past results based on the type of disease deficiency. This encompassed historical timestamps, names, descriptions, symptoms, occurrence conditions, as well as corresponding prevention and treatment measures, providing users with valuable references. The flow design of the apple disease classification function was exemplified in Figure 8.

3.2.5 System Design Issues and Development

(1) Limitations of system development: In the process of system development, there were also many problems. The first was the function design of the system disease deficiency recognition, the main core of which was the accuracy and speed of the recognition model. According to the requirements of model design, in the process of continuous model selection and training, we finally selected AlexNet-F model as the main model structure, which effectively improved the problem of inaccurate recognition results.

The second was the function design of disease deficiency classification of the system, the main core of which was the support of various data such as database. According to the functional requirements of systematic classification, we needed a lot of information basis such as disease deficiency type and drug application strategy. After collection and training, the data used could meet the needs of the system, but there were still problems such as the application strategy in the income bank, the insufficient information of disease deficiency and the insufficiency of discussion with traditional agricultural experts. Identifying problems with models and data still requires developers to refine and train collection.

(2) System development availability: After the preliminary design and development of the system, and through the testing department’s testing and self-testing, we opened the internal test for a small number of target users including large growers, orchard cooperatives, fruit farmers, agricultural technicians, and other apple industry related production industry personnel, in order to obtain the product in the function, process, interface and other existing problems and timely optimization and repair.

After a private beta, we received feedback from users. For the system interface, the user reflected that the interface beautification should correspond to the agricultural theme of the system, and the plate setting should be clear and clear. For the application function of the system, the user reflected the positive feedback of relatively accurate identification results and professionally correct identification and interpretation after use, which met the basic needs of users. It proved that the system has high usability.

(3) Scalability of system development: The system could meet the needs of most small-scale apple growing industries, but to put the system into larger scale use, it needed to expand the breadth and depth of functions. In terms of functional depth, to meet the needs of identification and classification of large tracts of farmland, the identification function needed to be more accurate and faster, and the classification function also needed to expand the data range to provide more types of disease information and drug application decisions. To this end, the system should expand its functional depth towards training more efficient models and collecting more extensive and effective information.

For the functional breadth, to meet the management needs of large farmland, new functions could be expanded in the direction of big data management. Added farmland real-time management function, making it easy for users to view farmland conditions, achieve disease identification, disease management, crop management integration, easily to use.

4 CONCLUSION

This paper mainly introduced an apple disease application assisted decision-making system based on deep learning technology, aiming to improve the level of apple yields and quality, and to solve the challenges faced by China’s apple industry, i.e., the reduction of planting area, the aging of the population, and the identification of apple foliar diseases. The system mainly consisted of two modules: disease deficiency recognition function and disease deficiency classification function. The disease deficiency identification function used image processing, convolutional neural network, mathematical methods and other technologies to construct an apple leaf disease identification model, which could monitor and identify the diseases of apple leaves in real time and could accurately identify a variety of apple diseases and provide targeted prevention and control suggestions. The disease deficiency classification function summarized and classified the recognition results, stored historical recognition information, provided users with disease cases for reference, and improve the efficiency of users’ decision-making.

The realization of the system mainly relied on the
python programming language and deep learning technology, through the front-end and back-end technology operation, it could accurately identify a variety of apple diseases and provide targeted prevention and treatment recommendations. It can effectively improve the accuracy and efficiency of disease identification. In addition, the system also had a friendly user interface and history record query function, which made it easier for users to obtain relevant information and make decisions.

5 PROSPECTIVE

This apple foliar disease deficiency identification system is a typical case of smart agriculture technology in practical application, which has a wide range of application prospects. With the continuous development of smartphones, image processing technology and artificial intelligence, the system will continue to be optimized in terms of ease of operation, accuracy, and cost-effectiveness. The following are specific future work and research directions related to system improvement and adaptation to the changing agricultural environment: Real-time monitoring and early warning, further development of the system to enable real-time monitoring of disease conditions on apple leaves. By combining it with sensor technology, the disease can be detected in time and early warning messages can be sent to farmers to help them take timely and effective control measures.

Big data analysis and decision support: by combining the system with big data analysis technology, the outbreak trend and propagation pattern of the disease can be predicted more accurately by collecting and analyzing a large amount of disease data. Based on these data and analysis results, farmers can be provided with more targeted and scientific prevention and control recommendations to help them make more informed decisions.

Automated Agricultural Management System: Combining the system with automated agricultural management technologies,
such as robots and automated spraying systems, enables comprehensive management of apple orchards. By collecting leaf image data through robots, the system can monitor the health of apple trees in real time and automate disease identification and application operations. Such an automated agricultural management system will improve the efficiency and sustainability of agricultural production.

Agroecological environment protection: While disease prevention and control, focusing on agro-ecological environment protection is also an important research direction. By studying the correlation between diseases and environmental factors, more environmentally friendly and sustainable disease control measures can be developed to reduce the negative impact on the ecological environment.

Acknowledgements
This work is supported by the Guangdong Province Undergraduate Innovation Training Project under Grant (No. S202314278042), the Collaborative Project for the Development of Guangzhou Philosophy and Social Science in 14th Five-year Plan (No. 2023GZGJ171), the Educational Science Planning Project of Guangdong Province (No. 2022GXJK073, No. 2023GXJK125), the Collaborative Project for the Development of Guangdong Province Philosophy and Social Science (No. 2023GD23XTY05), the National Undergraduate Innovation Training Project of China under Grant (No. 202314278014), the Natural Science Foundation of Hunan Province (No. 2022JJ30308), the Ideological and Political Special Research Project of Guangdong University of Education in 2022 (No. 2022SZZX13), the 2022 Party Building Research Project of the Guangdong Provincial Higher Education Party Building Research Association (No. 2022BK024), and the Special Support Program for Cultivating High-level Talents of Guangdong University of Education (2022 Outstanding Young Teacher Cultivation Object: Wanyi Li).

Conflicts of Interest
The authors declared that there is no conflict of interest regarding the publication of this paper.

Author Contribution
Zheng L wrote the manuscript and organized the experimental setup. Li W supervised the paper-writing process and provided the research platform for the team. Li N and Huang C designed the experiments and conducted testing for the project. Cheng J, Kuang Y, Guo X, and Ling Z performed data analysis and collected data. All authors contributed to the writing of this article, participated in the review process, and approved its submission.

References