IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Vol. 12, Issue 4, April 2023

DOI: 10.17148/IJARCCE.2023.12490

Identification and diagnosis of fruit diseases through image processing techniques

Nilesh Khambalkar[1], Snehal Belkhode[2], Nikhita Tarare[3], Chetan Raut[4],Dr.Rahul Burange[5]

Department of Electronics Engineering, K.D.K. College of Engineering, Nagpur,India[1-5]

Abstract: In the past, detecting fruit diseases relied on human visual inspection, which was often unreliable due to subjective judgment and limitations in detecting microorganisms. This approach was time-consuming, costly, and less accurate. However, using MATLAB-based approaches for quick and accurate diagnosis is a better choice compared to outdated methods. Symptoms of infection or disease can manifest on fruits, leaves, and lesions of plants, and this project aims to accurately diagnose the condition based on submitted images through image segmentation, preprocessing, feature extraction, and labeling. Various factors such as insect transmission, weather, and environmental conditions can cause infectious diseases in fruits, caused by viruses, fungi, or bacteria. The project will focus on identifying the cause of contamination in fruits to determine the type of infection, by extracting major and minor axes of fruit characteristics from images for effective classification.

Keywords: K-Means Clustering, Local Binary Pattern, Multi-class Support Vector Machine, Texture Classification

I. INTRODUCTION

Computer vision research strives to develop recognition systems that can match the reliability of human perception. In the agricultural industry, images play a crucial role in collecting and analyzing scientific data, with photography being the preferred technology for accurate data reproduction and reporting. However, processing and quantifying photographic data using mathematical methods can be challenging. To overcome these challenges, the integration of computers, microelectronics, and conventional photography has led to the development of digital image analysis and processing technologies, which enable researchers to improve and study their data across various magnification levels, from microscopic to telescopic.

Effective monitoring of fruit and tree health is vital for sustainable agriculture, as early detection and treatment of diseases are critical. However, there are currently no widely available sensors for real-time monitoring of tree health, and the most common method is labor-intensive scouting, which can be time-consuming and costly. Molecular methods, such as polymerase chain reaction, are often used for identifying fruit diseases, but they require extensive sampling and processing, adding to the cost and time involved.

Fruit infections can have a significant impact on harvest success, resulting in reduced yields and removal of affected varieties from production. Early identification of diseases and monitoring of crop health are crucial for managing disease vectors, applying appropriate measures, and maximizing production. Traditional methods of visual inspection by experts have been relied upon for detecting and identifying fruit diseases, but in some developing countries, accessing on-site specialists can be time-consuming and expensive.

Fruit diseases, such as soybean rust in soybeans, can cause substantial economic losses in production and quality during harvest. However, even partial eradication, as little as 20% of the infection, can result in significant profits of nearly $11 million for farmers.

It's also important to note that diseases appearing in fruits can impact the foliage and structural components of trees. Early identification of fruit issues can help reduce losses and prevent the spread of diseases.

II. LITERATURE REVIEW

The method proposed by S. Malathy et al. [1] aims to detect fruit diseases and identify specific types of diseases that affect fruits through a comparative analysis. To achieve this, Convolutional Neural Networks (CNN), a type of deep learning algorithm commonly used for visual image evaluation, is employed. CNN takes input images and distinguishes

DOI: 10.17148/IJARCCE.2023.12490

them based on various features and parameters extracted from the images. This method is expected to greatly benefit farmers in increasing agricultural yields in the near future. Further research on this approach will be carried out using the Python programming language. When implemented, this method has shown a success rate of 97%.

The research conducted by R. Ramya et al. [2] highlights the importance of early diagnosis of fruit diseases in the agriculture sector. In this study, emphasis is placed on utilizing Cloud computing to identify and analyze diseases in fruits in specific plant regions, as well as storing and retrieving data related to agricultural fields and farmer characteristics. Various factors such as insects, soil quality, and weather conditions can contribute to the occurrence of fruit diseases. Image processing techniques are employed to assess and record relevant information about the plants and their environment.

The research conducted by M. Senthamil Selvi et al. [3] underscores the significance of agricultural output to the country's economy, with vegetable and fruit yields being greatly impacted by plant diseases. Despite efforts to address this issue, annual losses due to pests and diseases in India are estimated at 50,000 crore rupees, leading to food scarcity for millions of people. Accurate diagnosis of plant diseases is therefore crucial for reducing food loss and improving agricultural product quality and quantity. While manual analysis of leaf patterns and disease identification are effective methods for detecting plant diseases, they have limitations, such as the need for human intervention and the time-consuming nature of large-scale monitoring in big farms. In recent years, various agribusiness establishments and technological advancements have emerged to enhance agricultural productivity. This study proposes a groundbreaking approach that utilizes Image Processing (IP) technology to accurately identify plant diseases, reducing the need for physical observation and manual detection. The proposed method involves image capture, pre-processing, segmentation, feature extraction, and classification, with the histogram of oriented gradient used for image feature extraction. Through the analysis of acquired images, the affected area of a leaf can be easily pinpointed, aiding in timely disease detection and management.

The authors of this study, Yan Qi et al. [4], addressed the significant concern of fruit diseases in the fruit-growing industry by developing a new plant disease diagnosis model based on deep learning. By identifying fruit leaves, the model was able to effectively manage fruit diseases in complex environmental conditions, leading to increased fruit output and improved quality. The model employed various image processing techniques, including image normalization processing and the MSRCR defogging algorithm, to enhance image quality. Additionally, the Canny SLIC algorithm based on gradient was used for data set fragmentation to obtain leaf blades exhibiting characteristics of disease spots. The identification process was concluded by feeding the fruit disease photos into an upgraded version of the DenseNet algorithm, which accurately detected and categorized illness characteristics within the images. Notably, the model outperformed the gold standard CNN convolutional architecture model with an impressive average accuracy of 98.98% when tested on data from three different types of fruit diseases: Grape spot anthracnose, Grapevine white rot, and Grapevine anthracnose. This novel model significantly improved the clarity and reliability of fruit disease image recognition in complex environments, and has the potential to be used for automated detection and recognition of fruit diseases.

III. METHODOLOGY

The decline in fruit production can be attributed to various factors, including infectious diseases that affect the fruit. Identifying the specific disease affecting the crop is a challenging task that requires innovative approaches for accurate and timely categorization. Manual inspection of fruits by farmers is labor-intensive and time-consuming, which is why image processing and machine learning algorithms offer a faster and more accurate solution for disease detection and pesticide recommendation. The image processing pipeline typically involves several steps, including grayscale conversion, noise reduction, smoothing, and other enhancements. Feature extraction is then performed by identifying significant differences in pixel values and locating edges in the image. Segmentation is used to separate the region of interest from the rest of the image, and finally, images are classified based on their characteristics using machine learning algorithms.

One popular deep learning method used for computer vision tasks, such as citrus disease classification, is the convolutional neural network (CNN). The CNN model consists of various building blocks, such as convolutional layers, pooling layers, and fully connected layers, which are used to create a powerful neural network. The CNN model is trained using techniques like backpropagation, allowing it to learn complex spatial feature patterns from the input images. In the CNN model, the input image is first decomposed into individual pixels, forming a three-by-three matrix of red, blue, and green colors. The subsequent layers of the CNN perform convolutional operations between the input matrix and filter grid to extract features. The convolved feature map is then sent to the Maxpooling layer, which reduces the dimensions

DOI: 10.17148/IJARCCE.2023.12490

of the feature map by applying filtering operations to extract high-level features. As the CNN model progresses through layers, it focuses on the most essential details.

The flattened feature vector produced by the Maxpooling layer is then passed to the Activation Functions module, where the citrus fruit images are grouped and characterized based on the information obtained from the previous layers. The SoftMax activation method is commonly used to calculate the probabilities of each citrus crop disease that can be identified. The classification of the input image is determined based on the highest probability value.

IV. PREPROCESSING

Image pre-processing is a common practice to improve the quality of image data and prepare it for further analysis and processing. This is typically done at the lowest level of abstraction and involves various techniques that exploit the redundancy present in images without increasing the amount of information they contain. For example, neighboring pixels that represent the same physical object often have similar or identical brightness values. As a result, if a pixel is corrupted, its original value can be approximated by taking the average of its surrounding pixels. This helps to eliminate unwanted distortions and enhance essential features in the image. It is recommended to apply different image pre-processing algorithms before saving the acquired picture to an image database.

V. INPUT IMAGE

The images provided in the context show maize fruits that have been impacted by bacterial and fungal diseases, including foliar fruit spot and apple fruit spot. These images were captured randomly from a maize field, and the lighting conditions during photography were not consistent. To commence the segmentation process, the images were resized to a standardized resolution of 256 by 256 pixels.

VI. SEGMENTATION

Segmentation is a crucial technique used to isolate and extract important regions within an image for further analysis. To achieve this, we utilize the user-friendly k-means clustering algorithm for image segmentation. This technique is particularly useful for separating and retrieving visual objects, even when their boundaries are indistinct. In order to generate a large number of clusters that can reliably isolate different picture objects, we employ a color space transformation that encompasses all color components. This color space includes a luminosity layer as well as two chromaticity levels. Euclidean distance matrices are utilized to measure the dissimilarity between hues, and the k-means distance is computed to determine the class for each image cluster based on the pixel positions within the image.

VII. K-MEAN CLUSTERING

K-means clustering is a widely used technique for comprehending and analyzing photos. Clustering involves dividing a set of data into smaller subsets based on common characteristics using a predefined distance metric. In image clustering, various features present in the image, such as shapes or textures, can be utilized. K-means clustering partitions the data into a predetermined number of groups, with initially arbitrary cluster centers. The next step involves determining the centroid of the data set and associating each point in the data set with it. Each pixel in the image is then assigned to a cluster based on its proximity to the centroid of that cluster, which is calculated using the Euclidean distance metric.

VIII. BLOCK DIAGRAM

In the process of analyzing photos, k-means clustering is utilized to group data based on common characteristics using a predetermined distance metric. For example, in fruit image analysis, the RGB fruit picture is transformed using color space conversion, and pre-processing methods such as cropping, smoothing filter, and histogram equalization may be applied to clean up the image. Segmentation techniques like RGB to HIS conversion and boundary/spots detection can be used to divide the image into sections with similar characteristics. Thresholding algorithms can transform grayscale images into binary images by determining cutoff values based on pixel intensities. Support Vector Machine (SVM) is a supervised learning model that can be used for classification and regression tasks in image analysis. In agriculture, image processing methods are used to detect leaf diseases, but there is a need for accurate categorization of leaves after feature recognition. Various methods such as fuzzy logic, principal component analysis, and K-Nearest Neighbor Classifier can be employed for leaf disease categorization. Different plant species like apples, grapes, potatoes, and tomatoes can be labeled based on their health condition. The dataset used for analysis may consist of thousands of photos of different crops, resized to a standard size, and divided into training and testing datasets for CNN model training.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⅔ Impact Factor 8.102 ⅔ Vol. 12, Issue 4, April 2023

DOI: 10.17148/IJARCCE.2023.12490

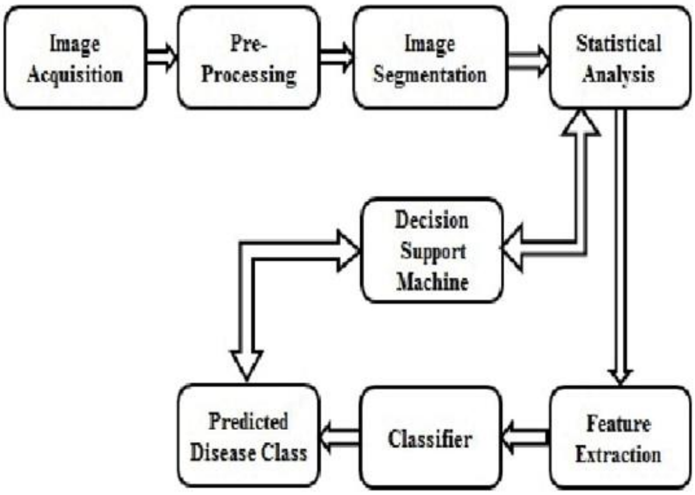

Fig1. Overall Diagram of System Architecture

IX. CLASSIFICATION

Support vector machines (SVMs) are a type of supervised learning models that utilize specialized algorithms for data analysis, specifically for tasks like classification and regression. In this study, SVM algorithm was employed to classify images and its effectiveness in classification was investigated. SVMs are particularly powerful for binary classification tasks as they can provide a fast classifier function after training. There are various methods to apply SVMs for situations with three or more classes. SVMs are widely used supervised learning models in machine learning, and their associated learning algorithms analyze data for classification and regression tasks. Dual-class classification serves as the foundation of SVMs, and it is the traditional approach for multiclass classification. The classifier evaluation involves labeling output values that exceed the threshold as "true" and those that do not as "false". The SVM classifier is utilized for binary image classification.

X. CONCLUSION

The world is rapidly moving towards a future where technology plays a crucial role. Farmers often express frustration as they invest significant resources in fertilizers only to see their crops devastated by viruses. However, there is a shortage of experts in this field, and opinions of experts and non-experts can differ. Thus, seeking expert advice before taking action is advisable. To improve the accuracy of disease diagnosis, it was discovered that increasing the number of training samples and fine-tuning SVM parameters were beneficial. This approach provides a framework for identifying and categorizing fruit diseases. K-means segmentation is used to segment the infected area, followed by extraction of texture characteristics using GLCM. Finally, SVM is employed for classification, enhancing the overall accuracy of disease identification.

REFERENCES

1.Tian Youwen and Li Chenghua, "Color Image Segmentation Method of Plant Disease Basing on Statistics Model Identification [J]", Journal of Jilin University (Gongxue Version), no. 02, pp. 291-293, 2004.
2.Shang Yijun, Zhang Shanwen and Zhang Yunlong, "Plant Disease Detection Method Basing on Image of the Plant Leaf [J]", Jiang su Agricultural Science, vol. 42, no. 04, pp. 340-342, 2014.
3.Chao Xiaofei, Research on Disease Identification and Disease Spot Segmentation Methods of Normal Apple Leaves Basing on Deep Learning [D], Northwest A&F University, 2021.
4.Guo Xingang, Wang Jia, Qu Nuoxi and Cheng Chao, "Canny SLIC Image Segmentation Algorithm Basing on Gradient Direction[J]", Computer Simulation, vol. 38, no. 09, pp. 465-469+500, 2021.
5.Lan Luo, Yalan Ye and Zehui Qu, "A Mini Fish tailed Lion: The Intelligent Fishbone Based on Golden Fish", the 2012 International Conference on Systems and Informatics, 5. 2012.
6.Cheng Yanyan, "Method Simulation of Self Adaptive Fast Removal of Unaligned Distributed Redundant Data[J]", Computer Simulation, vol. 36, no. 09, pp. 389-392, 2019.
7.Wang Xiaofang, Fang Dengjie, He Hairui and Zou Qianying, "MSRCR Image Defogging Algorithm Basing on Multi-scale Detail Optimization[J]", Experiment Technology and Management, vol. 37, no. 09, pp. 92-97, 2020.
8.Retinex Enhancing Algorithm of Image Processing (SSR MSR MSRCR), [online] Available: https://www.bbsmax.com.

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Vol. 12, Issue 4, April 2023
DOI: 10.17148/IJARCCE.2023.12490

9.R Achanta, A Shaji, K Smith et al., Slic super-pixels[R], 2010.
10.S Hochreiter, "The vanishing gradient problem during learning recurrent neural nets and problem solutions[J]", Int J Unc Fuzz Knowl Based Syst, vol. 6, no. 2, pp. 107-116, 1998.
11.Liu Huasong, Research on Segmentation Method of Toxic Strain Embryo Egg Image Basing on In-depth Learning[D], Tianjin:Tiangong University, 2019.
12.Goyal P Linty, R Girshick et al., "Focal loss for dense object detection[J]", IEEE Transactions on Pattern Analysis & Machine Intelligence, vol. PP, no. 99, pp. 2999-3007, 2017.
13.Yang Qirui, "Deep Network Fire Disaster Image Identification Algorithm Basing on improved DenseNet[J]", Computer Application and Software, vol. 36, no. 2, pp. 258-263, 2019.
14.Wu Yunzhi, Liu Aoyu, Zhu Xiaoning, Liu Chenxi, Fan Guohua, Le Yi, et al., "FI-DenseNet: Convolution Network Applied to Plant Disease Image Identification[J/OL]", Journal of Anhui Agricultural University, pp. 1-7, 11 2021.
15.Lan Luo, QiongHai Dai, ChunXiang Xu and ShaoQuan Jiang, "An Applicaiton Study to the Ciphers Weigh in Faithful Transmission", Applied Mechanics and Materials Switzerland, 2012.
16. Zulkifli Bin Husin, Abdul Hallis Bin Abdul Aziz and Ali Yeon Bin Md Shakaff Rohani Binti S Mohamed Farook, "Feasibilit y Study on Plant Chili Disease Detection Using Image Processing Techniques", 2012 Third International Conference on Intelligent Systems Modelling and Simulation.
17. R. Dhaya, "Flawless Identification of Fusarium Oxysporum in Tomato Plant Leaves by Machine Learning Algorithm", Journal of Innovative Image Processing (JIIP), vol. 2, no. 04, pp. 194-201, 2020
18. J. Samuel Manoharan, "Flawless Detection of Herbal Plant Leaf by Machine Learning Classifier Through Two Stage Authentication Procedure", Journal of Artificial Intelligence and Capsule Networks, vol. 3, no. 2, pp. 125-139, 2021.
19. Godliver Owomugisha, John A. Quinn, Ernest Mwebaze and James Lwasa, "Automated Vision-Based Diagnosis of Banana Bacterial Wilt Disease and Black Sigatoka Disease", Preceding of the 1'st international conference on the use of mobile ICT in Africa, 2014.
20 M. Sharif, M. A. Khan, Z. Iqbal, M. F. Azam, M. I. U. Lali, and M. Y. Javed, ''Detection and classification of citrus diseases in agriculture based on optimized weighted segmentation and feature selection,'' Comput. Electron. Agricult., vol. 150, pp. 220–234, Jul. 2018.
21 Manavalan, ''Automatic identification of diseases in grains crops through computational approaches: A review,'' Comput. Electron. Agricult., vol. 178, Nov. 2020, Art. no. 105802.
22 W. Pan, J. Qin, X. Xiang, Y. Wu, Y. Tan, and L. Xiang, ''A smart mobile diagnosis system for citrus diseases based on densely connected convolutional networks,'' IEEE Access, vol. 7, pp. 87534–87542, 2019.
23 G. Wang, Y. Sun, and J. Wang, ''Automatic image-based plant disease severity estimation using deep learning,'' Comput. Intell. Neurosci., vol. 2017, Jul. 2017, Art. no. 2917536.
24 U. P. Singh, S. S. Chouhan, S. Jain, and S. Jain, ''Multilayer convolution neural network for the classification of mango leaves infected by anthracnose disease,'' IEEE Access, vol. 7, pp. 43721–43729, 2019..