IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 13, No. 1, March 2024, pp. 876~887
ISSN: 2252-8938, DOI: 10.11591/ijai.v13.i1.pp876-887

Optimized deep learning-based dual segmentation framework for diagnosing health of apple farming with the internet of things

Harsha Raju, Veena Kalludi Narasimhaiah
School of Electronics and Communication Engineering, Reva University, Bengaluru, India

Article Info

Article history:

Received Jul 8, 2023
Revised Oct 17, 2023
Accepted Nov 15, 2023

Keywords:

Convolution neural network
Genetic algorithm
Internet-of-things
Otsu technique
Supervised learning

ABSTRACT

The high disease prevalence in apple farms results in decreased yield and income. This research addresses these issues by integrating internet of things (IoT) applications and deep neural networks to automate disease detection. Existing methods often suffer from high false positives and lack global image similarity. This study proposes a conceptual framework using IoT visual sensors to mitigate apple diseases' severity and presents an intelligent disease detection system. The system employs the augmented Otsu technique for region-aware segmentation and a colour-conversion algorithm for generating feature maps. These maps are input into U-net models, optimized using a genetic algorithm, which results in the generation of suitable masks for all input leaf images. The obtained masks are then used as feature maps to train the convolution neural network (CNN) model for detecting and classifying leaf diseases. Experimental outcomes and comparative assessments demonstrate the proposed scheme's practical utility, yielding high accuracy and low false-positive results in multiclass disease detection tasks.

Corresponding Author:

Harsha Raju
School of Electronics and Communication Engineering, Reva University
Bengaluru, India
Email: harsh4no1@gmail.com

1. INTRODUCTION

Agriculture is a significant contributor to the economies of many countries. Exporting and importing fruit varieties is also common, especially apple fruit, which is in high demand worldwide [1]. Many diseases attack apple trees in crop fields due to their unavoidable presence in agriculture. However, apple rust, cider rust, and scab are the most common fungal diseases that heavily impact apple crop yields and economic values [2], [3]. According to estimates, diseases to plants result in a loss of 35-40% of crop yields in agriculture [4]. Due to a lack of facilities and awareness, most farmers have no idea what type of disease they are dealing with. In these situations, they often use fungicides and insecticides to prevent yield loss due to fungi and pests [5]. This practice raises concerns due to excessive chemical use, posing risks to health, the environment, and environment-friendly insects [6]. Early disease detection is crucial for planning sustainable crop maintenance and reducing chemical usage in apple plant care [7]. An affordable and automated disease identification method is essential. Many research works are conducted to automate the disease detection and classification process. However, each previous approach has different drawbacks, such as high false positive results in classification tasks due to imprecise preprocessing and feature modelling. Few researchers have achieved good detection accuracy but at the cost of computational complexity.

The prime motivation of the proposed study is drawn from the findings discussed by Bonkra *et al.* [8], where different machine and deep learning-based schemes have been used. This study has been explicitly carried out concerning apple leaves towards the detection of disease. However, no exclusive study model is witnessed for diagnosing the health of apple farming considering the autonomous monitoring in the internet of things (IoT). Moreover, convolutional neural networks (CNN) have been the most famous disease detection model in previous studies based on the Plant-Village dataset. But its effectiveness depends on the extensive training samples. A comprehensive analysis of various disease detection studies using plant leaf images is discussed in [9]. Several approaches based on image processing and learning algorithms have been explored to improve plant disease recognition. The learning techniques include CNN [10], artificial neural networks (ANN) [11], backpropagation [12], support vector machines (SVM) [13], and various other image processing techniques [14], [15]. CNN is robust for image processing tasks because it can perform feature extraction and classification. It is shown [16] how CNN can be used to classify apple flowers effectively and automatically semantically. A comparative analysis validates the performance of the presented work, where the introduced mechanism effectively counts total flowers. In [17], the researchers introduced a multiclass apple disease classification system using optimized segmentation and feature extraction techniques. Gaussian and median filtering enhances disease portions, followed by correlated pixel-based segmentation and genetic algorithm-based feature extraction. In [18], the researchers focus on improving feature extraction with histogram-based segmentation, genetic algorithm, and correlation-based attribute selection using a custom dataset for SVM classification. The work by [19] uses K-means clustering with colour-based segmentation, gray-level co-occurrence matrices (GLCMs) for feature extraction, and SVM for soybean leaf disease identification. Further, the study of [20] presented graph cut extraction and texture-based feature selection, using k-nearest neighbor (KNN) and SVM for classification, although limited by feature maps and background segmentation.

In [21], the researchers discuss Gaussian mixture models for image segmentation, swarm optimization, fuzzy c-means (FCM) for disease portion segmentation, and SVM for binary leaf classification, assuming uniform pixel intensity variations. The authors combined Delta-E and expectation maximization to generate segmented images. ANNs detect diseases in brinjal plants based on colour, texture, and structural features but lack k-fold cross-validation [22]. The researchers in [23] improved grape leaf images using a haze-removing algorithm, segmented enhanced images into multiple channels and used mathematical morphology after selecting the best sub-channels with a weight function. It utilized SVM for classification but did not focus on disease portion segmentation and considered only a single channel intensity image. Nesterov *et al.* [24] presented a hybrid contrast stretching method that enhanced diseased portions in Apple leaves, using Mask R-CNN for disease segmentation, ResNet-50 for feature learning, and SVM for disease classification. However, there was no ground truth for Mask R-CNN. In [25], MobileNetV2 was used for automated feature extraction and classification of cassava plant leaf diseases with data augmentation, although artificial images with unnatural colours could impact classification. Gaikwad *et al.* [26] addressed multi-crop leaf disease detection, using colour, black and white, and grayscale images for Apple, Custard apple, and Guava plants, concluding that colour is a significant factor. In [27], the authors have suggested a custom 197-layer deep CNN model for multiple crop disease detection with high accuracy (99.58%) but at a high computational cost. In [28], the authors applied capability transformers to CNN models to enhance generalization for apple disease identification, achieving 96.85% accuracy with synthetic data, but computational complexity and potential underfitting issues remain challenges. The review of the literature also reveals that there are significant challenges in disease identification for apple leaf, viz. symptom variation [29], overlapping symptoms [30], multiple infections [31], environmental factors [32], issues in early disease detection [33], leaf variation [34], limited visual clues [35]. While existing studies have tried to tackle these issues, they tend to be symptomatic and lack a comprehensive solution. The identified research gaps include the need for effective segmentation in the presence of similar symptoms in background regions, attention to factors affecting model performance, a solution for segmentation without ground truth, and room for improvement in existing disease detection models. The proposed scheme aims to address these problems:

- Leaf symptom segmentation is challenging because symptoms resemble background regions and must be detectable from the background. Existing studies do not address noise or effectively preprocess leaf images with multiple diseases and similar background conditions.
- Extracting significant features is essential for disease classification, but previous work does not consider factors affecting model performance, such as lighting variations, multiple diseases, and different lesion sizes. Irrelevant features reduce accuracy and slow system execution, so feature selection is crucial.
- Existing image segmentation methods often require ground truth data, which is not always available. An efficient technique is needed to segment images without ground truth.

This paper proposes a plant disease detection system using colour conversion-based preprocessing and learning-driven segmentation schemes. The system includes two segmentation schemes: primary and secondary. The primary scheme segments the disease region from the original image using Otsu's algorithm and U-Net. The secondary scheme optimizes U-Net using a genetic algorithm to detect multiclass plant diseases. The system also explores pre-trained classification models, ResNet and MobileNet, for benchmarking. The goal is to increase the practical utility of disease prediction algorithms in IoT-based smart apple farming. The proposed scheme contributes as follows:

- It uses apple leaf images to conceptualise a smart apple farm network for large-scale health diagnosis.
- It introduces a novel exploratory-based preprocessing method to enhance colour intensity, improving leaf health diagnosis.
- It presents a dual segmentation approach. Primary segmentation uses a U-Net learning model and Otsu-based thresholding, while secondary segmentation integrates a U-Net model with a Genetic algorithm optimizer.
- The study demonstrates through extensive analysis that a CNN-based classifier provides superior detection and classification performance compared to commonly used ResNet and MobileNet models.

2. METHOD

The prime motive of the proposed study is to present a simplified yet learning-based computational framework capable of identifying and classifying the severity of crop disease considering leaf images. The case study considered is the leaf images of apples towards segmentation, identification, and classification of diseases. Disease detection and early classification are critical challenges in agriculture, but technology can automate plant disease detection and diagnosis, making it accurate and cost-effective. The IoT will enable farmers to make informed decisions and improve crop productivity and quality. Figure 1 shows a conceptual architecture of smart IoT applications for sustainable apple farming.

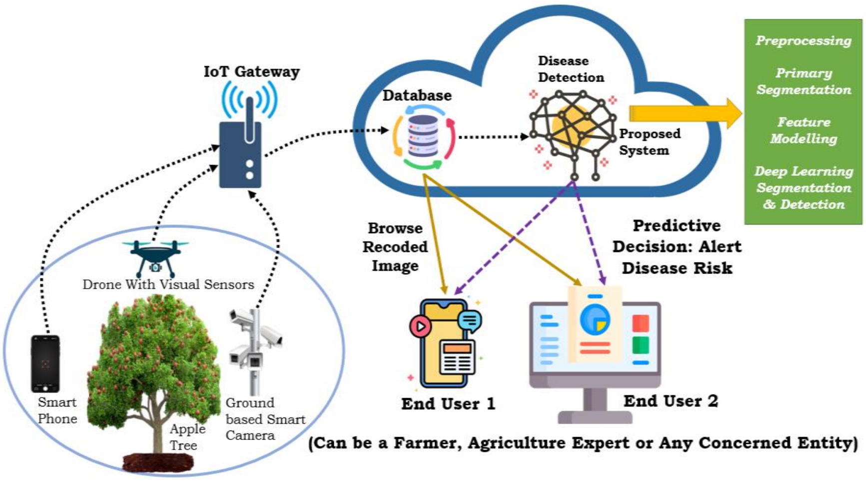


Figure 1. Illustrates a conceptual framework of a smart apple farm based on an IoT application

The conceptual framework of the proposed scheme entails smart visual sensors, an IoT gateway, a cloud database, cloud analytics for disease detection, and an end-user interface with cloud apps for real-time reporting and predictive analytics. This system comprises various visual sensors that periodically capture plant leaf images, including ground-based smart cameras, drones, and smartphones. The IoT gateway connects these sensors to the cloud, storing the data in a cloud database. An intelligent algorithm analyzes the images for disease detection. If a disease is detected, an alert is sent to the end user, enabling timely action in an end-user interface with cloud apps for real-time reporting and predictive analytics. This defines the practicality and significance of implementing the proposed learning model for leaf disease detection and classification from an application perspective.

The primary consideration in the proposed system is that the leaf disease can be better identified if its characteristics differ from background portions and other similar disease symptoms. Therefore, region-aware segmentation is carried out where the input leaf image is separated into three parts viz: i) background region, ii) foreground region, and iii) region of disease spot. Harnessing this set of information, the learning model generalizes the underlying pattern of data samples more effectively. The schematic architecture of the proposed system is illustrated in Figure 2.

The scheme proposed in this paper adopts various computational intelligence algorithms to devise an effective leaf disease detection system. The first operation is an effective preprocessing operation over an input apple leaf image followed by exploratory analysis. Based on the exploratory analysis, a contrast enhancement is done to correct the visual characteristics of the colour leaf image. It should be noted that the preprocessing operation does not convert colour to grayscale representation because colour represents an essential feature aspect, meaning the greener the leaf, the healthier it is. The proposed preprocessing algorithm also involves a colour conversion scheme applicable after primary image segmentation. The proposed scheme induces two forms of the learning-based segmentation process, i.e., primary and secondary.

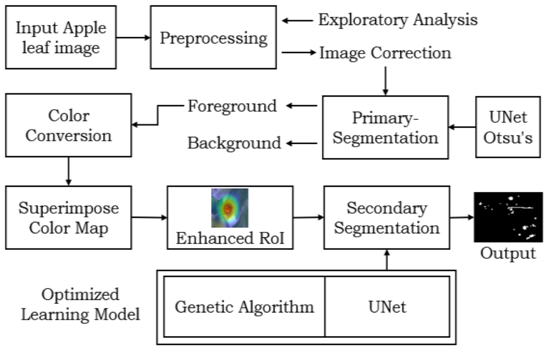


Figure 2. Schematic architecture of disease detection system

The primary segmentation uses the U-Net learning model and Otsu's thresholding with an objective of foreground and background separation. As a result, the distortion is reduced in the learning process by preserving the original structure of the image. Considering the foreground part contains disease regions, it is subjected to a colour conversion algorithm, where the disease spot is first transformed in the hue saturation value (HSV) colour map and then superimposed to the original image in the red, green, blue (RGB) colour map. As a result, enhanced disease region is obtained as a significant feature map, which serves as optimal inputs to the secondary segmentation for generating binary masks and acts as training data. The unique contribution made in this phase is that optimization is done in the U-Net network with the genetic algorithm, which will optimally adjust the learning parameters, i.e., weight and bias. It is to be noted that adopting a genetic algorithm acts as an optimizer, where a customized fitness function is developed that serves the role of the loss function in the binary mask generation. Hence, deep learning classifiers are adopted with obtained binary masks to perform disease classification. The successful implementation and execution of the proposed system can serve as a cost-effective and quick decision-making process in the agriculture sector for disease diagnosis in real-time scenarios.

2.1. Leaf image preprocessing and primary segmentation

In this section, the implementation of the preprocessing and segmentation methodology adopted in the proposed system is discussed. The exploratory analysis is conducted as the first step in the preprocessing phase to get an insight into the apple leaf image dataset adopted in the proposed work [36]. The exploratory findings reveal the following significant observations:

- The dataset consists of leaf images subjected to the most common diseases, as shown in Figure 3.
- There is a total of 7,771 images of apple leaves in the dataset. The total number of images is sufficient for the deep learning model; hence, no data augmentation procedure must be applied.
- The dataset is nearly uniform and does not associate with the class imbalance problem.
- The outer shape is similar, and a greener colour indicates a healthier leaf.
- The stem is always present in the centre of the leaf, and the edge of the leaf is similar.
- To perform disease classification, preprocessing and effective segmentation are required to extract the region of interest (ROI).
- In preprocessing, the images undergo contrast enhancement and resizing to accommodate the substantial dataset size. Subsequently, the dataset is divided into two parts, one for training and one for testing, using an 80:20 ratio, where 80% of the total 7,771 images are assigned to the training set, and the remaining 1,555 images are allocated to the testing set. This ensures a balanced and representative distribution for training and evaluating the model's performance.

Post preprocessing operation, the enhanced leaf images are subjected to a primary segmentation operation where the foreground region of the colour leaf image is separated from the background portion. The study proposes an augmented Otsu's thresholding technique, a joint mechanism of conventional Otsu's segmentation approach, and the U-Net learning model. The reason is that the conventional Otsu technique is computationally efficient to generate binarized images, but it it cannot be directly applied to colour images as it is susceptible to colour variation. Since the colour plays an essential role in the detection, the leaf images are not converted to grayscale representation. Therefore, in the original Otsu's algorithm, augmentation is done by integrating with the U-Net learning model to perform region-aware segmentation without losing vital aspects of the colour leaf image. The U-Net is a computational architecture based on the principle of fully convolutional networks that capture contextual features and localization [37].

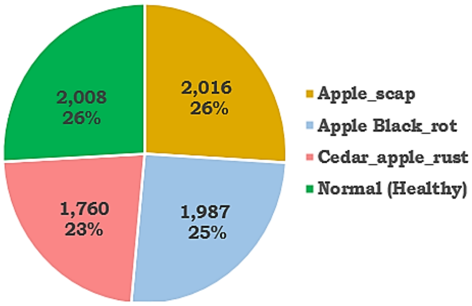


Figure 3. Image distribution statistics of apple leaf image dataset

2.2. Feature representation

The obtained segmented colour leaf images are in the RGB colour space, which may not provide a better feature set or represent colour attributes significant to different diseases in apple leaf because multiple diseases may occur on the same leaf and share similar colours. Therefore, the proposed feature representation scheme adopts a colour transformation where the RGB image is converted into an HSV colour map to comprehend better the colour variation among the diseases in the leaf image. The HSV stands for Hue Saturation and value, where the Hue component plays a vital role in colour detection, and the value determines the darkness of the leaf to detect diseases like scabs. The procedure for RGB to HSV is discussed in algorithm 1.

Algorithm 1: Color space conversion from RGB to HSV
Input: segmented foreground leaf image (I_{FS})
Output: HSV Image
Start
For $\forall$ $i_{x,y}$ in I_{FS}
Compute $C_{max} = f_{max}(R, G, B)$
Compute $C_{min} = f_{min}(R, G, B)$
Compute Diff = $C_{max} - C_{min}$
if $C_{max} == C_{min}$
Value of $H = 0$
else if $C_{max} == R$
Value of $H = \left(60 \times \left(\frac{G-B}{Diff}\right) + 360\%360\right)$
else if $C_{max} == G$
Value of $H = \left(60 \times \left(\frac{B-R}{Diff}\right) + 120 \% 360\right)$
else if $C_{max} == B$
Value of $H = \left(60 \times \left(\frac{R-G}{Diff}\right) + 240 \% 360\right)$
end
if $C_{max} == C_{min}$
Value of $S = 0$
else
Value of $S = \left(\frac{Diff}{C_{max}}\right) \times 100$
end
Value of $V = C_{max} \times 100$
end
Return HSV color map
End

Algorithm 1 takes the input of the segmented foreground image IFS obtained from region segmentation operation using augmented Otsu's technique. After the successful execution of all the steps, it returns the image in the HSV colour map from the RGB colour space. The algorithm initially computes the maximum followed by Cmax and minimum Cmin, the colour value of the RGB components for pixels ix,y in the input image. Further, it computes a chroma (strength) by computing the difference (Diff) of Cmax and Cmin. It is known that the chroma C value is a direct representation of colour purity. Therefore, when the maximum and minimum value of chroma is found to be the same (Cmax=Cmin), it will represent the instant when the considered image possesses a balanced chroma score and is ideal for performing further computation of other entities of HSV colour space. Afterward, the algorithm executes an operation of conditional check to compute the individual value of H, S, and V. Here, the RGB components, saturation (S), and value (V), lie in the range (0,1). At the same time, the hue (H) should have values in the range (0,360). A modulo operation is performed in each step of computing the HVS value.

Figure 4 illustrates the output of algorithm 1 as a feature map highlighting the diseased region on the leaf image. Figure 4(a) illustrates the segmented leaf image (obtained using the proposed UNet-Otsu's approach) is transformed into the HSV colour map, highlighting the colour information. Figure 4(b) shows an RGB colour map of the original leaf image, while Figure 4(c) demonstrates the segmented disease region, which is superimposed over the input image to emphasize the ROI, which corresponds to the disease spot.

(a) (b) (c)

Figure 4. Presents a visual depiction of the feature map: (a) HSV colour map, (b) shows the original RGB colour map, and (c) visualizes a feature that shows enhanced disease region

2.3. Secondary segmentation

The agenda of secondary segmentation is generating a binary mask by utilizing the capability of the U-Net segmentation model that takes its input as a feature map, an output generated from algorithm-1 towards localizing the disease in the leaf image. However, the ground truth (GT) mask is not available for the obtained enhanced disease region or feature, so training the U-Net learning model for disease spot (instance) segmentation without GT is not feasible, and its output cannot effectively be evaluated against the actual problem. Therefore, the U-Net is integrated with a genetic algorithm, and a specialized fitness function is developed to train the model appropriately. Here, the genetic algorithm is an optimizer, and the proposed fitness function acts like a loss function aware of all input leaf image types and corresponding disease classes. Both optimizer and loss function are essential parts of any neural network or deep learning model.

The optimizer is used to adjust the attributes of the learning model, such as weights and learning rate, to reduce the losses. Similarly, the genetic algorithm serves as an improved optimizer function to adjust the learning parameters of the U-Net model in the training phase. On the other hand, the loss function is an algorithm for assessing the suitability of U-Net learning network models in the training data. The proposed scheme offers a flexible scheme of computing fitness functions based on different variants of leaf classes, which directly contributes towards the accurate detection of variable leaf cases. The schematic illustration of the proposed optimized learning model for the secondary segmentation process is shown in Figure 5.

In order to build the fitness function, the algorithm selects two threshold values for each type of image I_{FS} and SI. These thresholds are determined from training images, where specific samples are chosen, and their colour values in the HSV colour space are averaged out. The fitness function evaluates the output binary mask based on these threshold values. When the class of the object is 0, the fitness is determined by counting the total number of unmasked pixels in the segmented image. All pixels will be masked for a normal image without any region of interest, leading to a higher fitness value when more pixels are masked. For class type 1, which corresponds to a reddish region of interest in a leaf, the algorithm measures the hue values in the masked image. If the hue values of the region of interest fall under the specified threshold, it contributes to a higher fitness value. The fitness algorithm computes the fitness value for all generated segmented images of the type mask produced by the U-Net learning model. These fitness values are then utilized by the genetic algorithm as optimizers during the training process to adjust the weights and biases, thus guiding the model to accurately improve its performance in segmenting leaf images. Figure 6, provides a visual representation of the secondary segmentation process, showing the successful outcome of returning a segmented mask of an input apple leaf image.

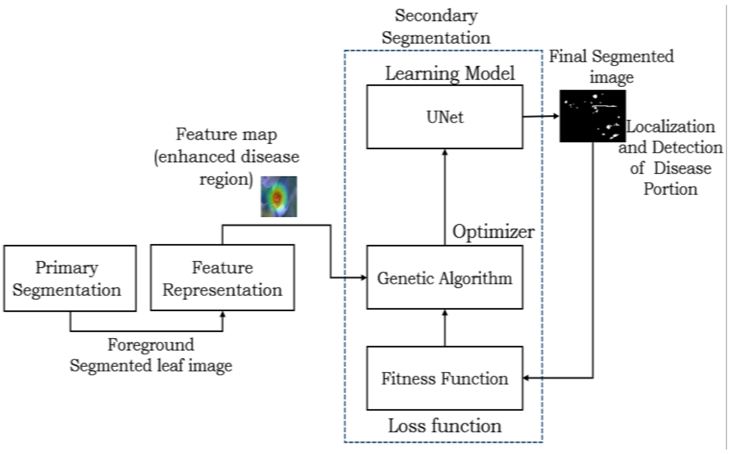
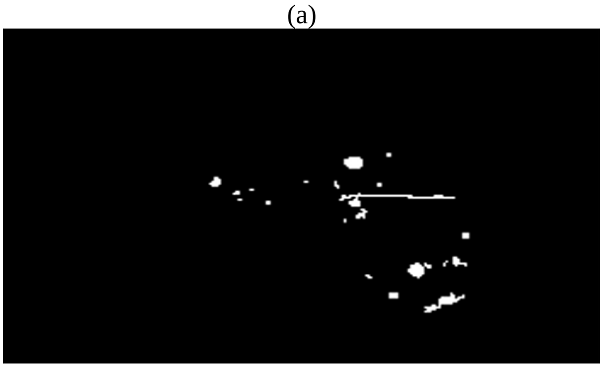


Figure 5. The procedure of secondary segmentation using U-Net and genetic algorithm

Algorithm 2: Fitness function for genetic algorithm for U-Net optimization
Input: segmented foreground leaf image (I_{FS}), segmented image or mask (SI), leaf class (Y)
Output: fitness (*f*)
Start
Initialize $f = 0$
if $Y == 0$
compute $f = 100 - \left(\frac{\sum_{i=1}^{n}\sum_{j=1}^{m} SI_{i,j}}{m \times n}\right) \times 100$
end of if
if $Y == 1$
I_{FS} = HSV (I_{FS})
$X = I_{FS} \times SI$
S = $\sum_{i=1}^{n}\sum_{j=1}^{m} SI_{i,j}$
$compute\ f = 100 * \left(\frac{s-\sum_{i-1}^{n}\sum_{j=1}^{m} T1 > Hue(X_{i,j}) > T2}{s}\right)$
end of if
if $Y == 2$
I_{FS} = HSV (I_{FS})
$X = I_{FS} \times SI$
S = $\sum_{i=1}^{n}\sum_{j=1}^{m} SI_{i,j}$
compute $f = 100 * \left(\frac{s-\sum_{i-1}^{n}\sum_{j=1}^{m} D1 > Volume(X_{i,j}) > D2}{s}\right)$
end of if
if $Y == 3$
I_{FS} = HSV (I_{FS})
$X = I_{FS} \times SI$
S = $\sum_{i=1}^{n}\sum_{j=1}^{m} SI_{i,j}$
compute $f = 100 * \left(\frac{s-\sum_{i-1}^{n}\sum_{j=1}^{m} M1 > Hue(X_{i,j}) > M2}{s}\right)$
else
$Y == 4$
End of if
Return output f
End

Figure 6 showcases the outcome of the U-Net-based secondary segmentation for binary mask generation, where Figure 6(a) shows the original input image and Figure 6(b) shows a binary mask, i.e., segmented disease (cedar apple rust) part. This visual analysis significantly reveals how the genetic algorithm, incorporating the fitness function, effectively optimizes the U-Net model for leaf image segmentation, improving disease region localization. In order to identify and classify diseases in apple leaf images, the generated masks of the original colour image will be used to train a deep classification model. The next section shows the outcome analysis and performance discussion based on the experimental analysis using the CNN model trained with generated masks for training and testing the colour image dataset.

(a)



(b)

Figure 6. Outcome of secondary segmentation: (a) input leaf image and (b) segmentation disease part

3. RESULTS AND DISCUSSION

The proposed system is developed using Python programming language executed in the Anaconda distribution, installed on the Windows 10 OS system. This section presents the outcome is discussed based on the experimental analysis, where the obtained segmented masks used to train supervised classification models to detect disease. The study trained the CNN model for disease classification with the obtained mask and ground truth in the training set. The overall result and performance discussion are presented concerning the visual and numerical outcomes concerning the accuracy, precision-recall rate, and F1 score. The statistical information about the testing dataset and class label is given in Table 1.

Table 1. Labels of disease class

Disease class	No. of samples	Label
Apple scab	431	0
Black Rot	391	1
Cider Apple Rust	367	2
Healthy	365	3

3.1. Model analysis

Figure 7 presents the confusion plot for the classified disease obtained while testing the proposed learning model. Since the proposed system is a multiclass classification system, the outcome is shown for each disease class, including healthy leaf. Further, numerical outcomes of classification for i) apple scab is shown in Table 2, ii) black rot disease is shown in Table 3, iii) cider apple rust is shown in Table 4, and iv) healthy leaf is shown in Table 5. The complete classification outcomes are shown as correct and false predictions for better accuracy analysis.

A closer examination of Table 2 reveals that out of 431 Apple Scab leaves, the proposed system accurately predicted 415 samples corresponding to Scab disease, but it misclassified 16 samples. The table provides a breakdown of the outcome statistics for the classification of Apple Scab, showing the total count, correct predictions, and false predictions for each class, including Black Rot, Cider Apple Rust, and Healthy. Table 3 presents outcome statistics for classifying Black Rot, including the total count, correct predictions, and false predictions for each class, such as Apple Scab, Cider Apple Rust, and Healthy. The analysis of the numerical outcomes derived from Table 3 reveals a total of 391 black rot leaves, of which the proposed system correctly predicted 373 samples, while 18 were misclassified.

Table 4 presents outcome statistics for Cider Apple Rust disease classification. The model's performance is evident, with 367 cider apple rust leaves, of which the proposed system accurately predicted 349 samples, while 18 samples were misclassified. This evaluation shows the model's effectiveness in identifying the cider apple rust disease. Table 5 presents comprehensive outcome statistics for the Healthy leaf classification, correct predictions, and false predictions for Black Rot, Cider Apple Rust, and Black Rot. The analysis of healthy leaves reveals a total of 365 samples, out of which the proposed system accurately predicted 355 samples, while ten samples were misclassified. This evaluation provides valuable insights into the model's performance in distinguishing healthy leaves and can aid in further refining the classification process for improved accuracy and reliability. Based on the results in Tables 2 to 5, the proposed classification model demonstrates promising performance in identifying different apple leaf diseases. The high number of correct predictions for Apple Scab, Black Rot, and Cider Apple Rust indicates the model's effectiveness in disease classification.

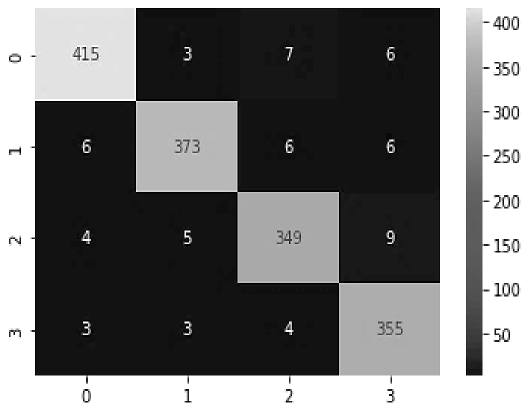


Figure 7. Confusion plot for classification model CNN

Table 2 Outcome statistics for classification of Apple Scab

Total Apple Scab	Correct prediction	False prediction		
	Apple Scab	Black Rot	Cider Apple Rust	Healthy
431	415	3	7	6

Table 3. Outcome statistics for classification of Black Rot

Total Black Rot	Correct prediction	False prediction		
	Black Rot	Apple Scab	Cider Apple Rust	Healthy
391	373	6	6	6

Table 4. Outcome statistics for classification of Cider Apple Rust

Total Cider Apple Rust	Correct prediction	False prediction		
	Cider Apple Rust	Apple Scab	Black Rot	Healthy
367	349	4	5	9

Table 5. Outcome statistics for classification of Healthy leaf

Total Healthy	Correct prediction	False prediction		
	Healthy	Black Rot	Cider Apple Rust	Black Rot
365	355	3	3	4

3.2. Comparative analysis

In this section, a comparative analysis of various approaches for leaf disease classification is presented. Table 6 highlights the comparison with the state of art methods which shows that proposed scheme offers significant edge of advantage towards addressing the identified research problem. The outcome in Table 6 also showcases that these existing studies with varied approaches introduces both beneficial aspect as well as limiting factors.

Table 6. Comparison towards state-of-art method

Approaches	Advantage	Limitation
CNN [16], [26], [28]	Higher accuracy	Higher resource dependencies
Genetic algorithm, correlation [17], [18], SVM [18]	Search-based optimization simplified and faster	Fitness function cannot adapt to potential change when test dataset is increased
K-Means clustering, color segmentation [19]	Simplified approach, effective segmentation	Narrowed analysis to offer higher precision
SVM, KNN [20]	Supports instance-based learning	Biased prediction not addressed; outcomes have outliers too
Swarm Optimization, FCM [21]	Applicable for varied forms of images even with noisy	Highly sensitive during initialization & parameters
Wavelet, ANN [22]	- Simplified feature learning for leaf disease	- Demands large number of parameters, leading to overfitting
SVM [23]	- Capable of processing high-dimensional problem space - Less overfitting issues	- Computationally intensive operation when exposed to multiple classes of leaf disease
MobileNet [25]	- Faster inference - Effective transfer learning and finetuning	- Reduced model capacity - Sensitive to input quality - Lower accuracy
CNN with ResNet [27]	Optimal resource dependencies	- Higher computational burden - Higher accuracy
Proposed (CNN, Genetic Algorithm, U-Net)	- Lower computational burden - Higher accuracy - Highly optimized performance - Supports generalization - Highly adaptable and flexible	- Not assessed over high resolution images

4. CONCLUSION

This paper has presented an enhanced decision support system to benefit the agriculture system cost-effectively. The modelling of the proposed system adopts a systematic implementation procedure that includes basic preprocessing, learning-based augmented segmentation, a colour space transformation algorithm for feature representation, and an optimized U-Net for mask generation. The proposed study has optimized the learning process of U-Net for mask generation, where a new fitness function is developed as a loss function, and a genetic algorithm is used as an optimizer for training. Using this scheme, a region-aware segmentation is achieved, with better training performance. This region-aware segmentation model generates an adequate mask of input leaf images, which further served as a basis for developing a supervised learning-based classification system. The proposed study implements a CNN model for multiclass disease classification. The experimental results verified the proposed system's scope: higher precision, recall rate, and F1 score. The result analysis demonstrates its scope and can be applied to real-world applications to benefit farmers through an effective decision-making process. Using the proposed system, the disease can be detected early, and suitable treatment can be applied. Hence, this leads to quality-aware fruits, higher production, and provide economic benefits to the producer. In future work, the study considers building a prototype of the suggested conceptual architecture of smart farms with more modification and optimization.

REFERENCES

[1] J. Liu and X. Wang, "Plant diseases and pests detection based on deep learning: a review," *Plant Methods*, vol. 17, no. 1, p. 22, Dec. 2021, doi: 10.1186/s13007-021-00722-9.
[2] H. Orchi, M. Sadik, and M. Khaldoun, "On using artificial intelligence and the internet of things for crop disease detection: a contemporary survey," *Agriculture (Switzerland)*, vol. 12, no. 1, p. 9, Dec. 2022, doi: 10.3390/agriculture12010009.

[3] N. Ganatra and A. Patel, "A survey on diseases detection and classification of agriculture products using image processing and machine learning," *International Journal of Computer Applications*, vol. 180, no. 13, pp. 7–12, Jan. 2018, doi: 10.5120/ijca2018916249.
[4] D. M. Rizzo, M. Lichtveld, J. A. K. Mazet, E. Togami, and S. A. Miller, "Plant health and its effects on food safety and security in a One Health framework: four case studies," *One Health Outlook*, vol. 3, no. 1, p. 6, Dec. 2021, doi: 10.1186/s42522-021-00038-7.
[5] A. Alengebawy, S. T. Abdelkhalek, Su. R. Qureshi, and M. Q. Wang, "Heavy metals and pesticides toxicity in agricultural soil and plants: ecological risks and human health implications," *Toxics*, vol. 9, no. 3, p. 42, 2021.
[6] F. Jacquet et al., "Pesticide-free agriculture as a new paradigm for research," *Agronomy for Sustainable Development*, vol. 42, no. 1, p. 8, Feb. 2022, doi: 10.1007/s13593-021-00742-8.
[7] P. Vasavi, A. Punitha, and T. V. Narayana Rao, "Crop leaf disease detection and classification using machine learning and deep learning algorithms by visual symptoms: a review," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 12, no. 2, p. 2079, Apr. 2022, doi: 10.11591/ijece.v12i2.pp2079-2086.
[8] A. Bonkra et al., "Apple leave disease detection using collaborative ML/DL and artificial intelligence methods: scientometric analysis," *International Journal of Environmental Research and Public Health*, vol. 20, no. 4, p. 3222, Feb. 2023, doi: 10.3390/ijerph20043222.
[9] H. Raju and V. K. Narasimhaiah, "Insights on assessing image processing approaches towards health status of plant leaf using machine learning," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 12, no. 2, p. 884, Jun. 2023, doi: 10.11591/ijai.v12.i2.pp884-891.
[10] G. K. V. L. Udayananda, C. Shyalika, and P. P. N. V. Kumara, "Rice plant disease diagnosing using machine learning techniques: a comprehensive review," *SN Applied Sciences*, vol. 4, no. 11, p. 311, Nov. 2022, doi: 10.1007/s42452-022-05194-7.
[11] C. Jackulin and S. Murugavalli, "A comprehensive review on detection of plant disease using machine learning and deep learning approaches," *Measurement: Sensors*, vol. 24, no. 100441, p. 100441, Dec. 2022, doi: 10.1016/j.measen.2022.100441.
[12] B. V. Nikith, N. K. S. Keerthan, M. S. Praneeth, and D. T. Amrita, "Leaf disease detection and classification," *Procedia Computer Science*, vol. 218, pp. 291–300, 2023, doi: 10.1016/j.procs.2023.01.011.
[13] P. Sindhu and G. Indirani, "Internet of things enabled pomegranate leaf disease detection and classification using cuckoo search with sparse auto encoder," *Mathematical Statistician and Engineering Applications*, vol. 71, no. 3, pp. 904–917, 2022.
[14] R. Mishra et al., "I-LDD: an interpretable leaf disease detector," *Soft Computing*, pp. 1–17, May 2023, doi: 10.1007/s00500-023-08512-2.
[15] X. Liang, "Few-shot cotton leaf spots disease classification based on metric learning," *Plant Methods*, vol. 17, no. 1, p. 114, Dec. 2021, doi: 10.1186/s13007-021-00813-7.
[16] P. A. Dias, A. Tabb, and H. Medeiros, "Multispecies fruit flower detection using a refined semantic segmentation network," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 3003–3010, Oct. 2018, doi: 10.1109/LRA.2018.2849498.
[17] M. A. Khan et al., "An optimized method for segmentation and classification of apple diseases based on strong correlation and genetic algorithm based feature selection selection," *IEEE Access*, vol. 7, pp. 46261–46277, 2019, doi: 10.1109/ACCESS.2019.2908040.
[18] Z. Chuanlei, Z. Shanwen, Y. Jucheng, S. Yancui, and C. Jia, "Apple leaf disease identification using genetic algorithm and correlation-based feature selection method," *International Journal of Agricultural and Biological Engineering*, vol. 10, no. 2, pp. 74–83, 2017.
[19] S. B. Jadhav, V. R. Udup, and S. B. Patil, "Soybean leaf disease detection and severity measurement using multiclass SVM and KNN classifier," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 9, no. 5, p. 4077, Oct. 2019, doi: 10.11591/ijece.v9i5.pp4077-4091.
[20] A. A. Bracino, R. S. Concepcion, R. A. R. Bedruz, E. P. Dadios, and R. R. P. Vicerra, "Development of a hybrid machine learning model for apple (malus domestica) health detection and disease classification," in 2020 *IEEE 12th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM)*, Dec. 2020, pp. 1–6, doi: 10.1109/HNICEM51456.2020.9400139.
[21] P. Kumar, M. 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 and Computation: Practice and Experience*, vol. 33, no. 3, p. e5312, Feb. 2021, doi: 10.1002/cpe.5312.
[22] R. Anand, S. Veni, and J. Aravinth, "An application of image processing techniques for detection of diseases on brinjal leaves using k-means clustering method," in 2016 *International Conference on Recent Trends in Information Technology (ICRTIT)*, Apr. 2016, pp. 1–6, doi: 10.1109/ICRTIT.2016.7569531.
[23] A. Adeel et al., "Diagnosis and recognition of grape leaf diseases: an automated system based on a novel saliency approach and canonical correlation analysis based multiple features fusion," *Sustainable Computing: Informatics and Systems*, vol. 24, no. 100349, p. 100349, Dec. 2019, doi: 10.1016/j.suscom.2019.08.002.
[24] D. A. Nesterov, B. M. Shurygin, A. E. Solovchenko, A. S. Krylov, and D. V. Sorokin, "A CNN-based method for fruit detection in apple tree images," *Computational Mathematics and Modeling*, vol. 33, no. 3, pp. 354–364, Jul. 2022, doi: 10.1007/s10598-023-09577-2.
[25] O. O. Abayomi-Alli, R. Damaševičius, S. Misra, and R. Maskeliūnas, "Cassava disease recognition from low-quality images using enhanced data augmentation model and deep learning," *Expert Systems*, vol. 38, no. 7, Nov. 2021, doi: 10.1111/exsy.12746.
[26] S. S. Gaikwad, S. S. Rumma, and M. Hangarge, "Fungi affected fruit leaf disease classification using deep CNN architecture," *International Journal of Information Technology*, vol. 14, no. 7, pp. 3815–3824, Dec. 2022, doi: 10.1007/s41870-022-00860-w.
[27] A. Pandian J., K. K., N. R. Rajalakshmi, and G. Arulkumaran, "An improved deep residual convolutional neural network for plant leaf disease detection," *Computational Intelligence and Neuroscience*, vol. 2022, no. 5102290, pp. 1–9, Sep. 2022, doi: 10.1155/2022/5102290.
[28] X. Li and S. Li, "Transformer help CNN see better: a lightweight hybrid apple disease identification model based on transformers," *Agriculture*, vol. 12, no. 6, p. 884, Jun. 2022, doi: 10.3390/agriculture12060884.
[29] S. Hasan, S. Jahan, and M. I. Islam, "Disease detection of apple leaf with combination of color segmentation and modified DWT," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 9, pp. 7212–7224, Oct. 2022, doi: 10.1016/j.jksuci.2022.07.004.
[30] Z. ur Rehman et al., "Recognizing apple leaf diseases using a novel parallel real-time processing framework based on MASK RCNN and transfer learning: an application for smart agriculture," *IET Image Processing*, vol. 15, no. 10, pp. 2157–2168, Aug. 2021, doi: 10.1049/ipr2.12183.
[31] P. Bansal, R. Kumar, and S. Kumar, "Disease detection in apple leaves using deep convolutional neural network," *Agriculture*, vol. 11, no. 7, p. 617, Jun. 2021, doi: 10.3390/agriculture11070617.

[32] M. Jung et al., "Genetic architecture and genomic predictive ability of apple quantitative traits across environments," *Horticulture Research*, vol. 9, Jan. 2022, doi: 10.1093/hr/uhac028.
[33] D. Jose and K. Santhi, "Early detection and classification of apple leaf diseases by utilizing IFPA genetic algorithm with MC-SVM, SVI and deep learning methods," *Indian Journal Of Science And Technology*, vol. 15, no. 29, pp. 1440–1450, Aug. 2022, doi: 10.17485/IJST/v15i29.1235.
[34] A. Jain, S. Sarsaiya, Q. Wu, Y. Lu, and J. Shi, "A review of plant leaf fungal diseases and its environment speciation," *Bioengineered*, vol. 10, no. 1, pp. 409–424, Jan. 2019, doi: 10.1080/21655979.2019.1649520.
[35] H. U. Rehman, S. Anwar, and M. Tufail, "Machine vision based plant disease classification through leaf imagining," *Ingénierie des systèmes d information*, vol. 25, no. 4, pp. 437–444, Sep. 2020, doi: 10.18280/isi.250405.
[36] "Apple leaves images," 2017. https://www.kaggle.com/datasets/abdallahalidev/plantvillage-dataset.
[37] X.-X. Yin, L. Sun, Y. Fu, R. Lu, and Y. Zhang, "U-Net-based medical image segmentation," *Journal of Healthcare Engineering*, vol. 2022, Apr. 2022, doi: 10.1155/2022/4189781.

BIOGRAPHIES OF AUTHORS

Harsha Raju received his master degree in VLSI and embedded systems. Currently persuing his doctor of philosophy in the school of electroniccs and communinaction engineering, REVA university. Harsha is currently working as an Assistant Professor in Dr Ambedkar Institute of Technology Bengaluru. He has had many national and international patents under his belt, out of which few patents are granted. He has published many international journals which are indexed in scopus, google scholor and web of science REVA University I would thank REVA university for providing me the oppourtunity for persuing my Doctrate of philosophy. The infrastructure in the campus is just excellent which makes any scholor to put in effort for persuing his education. He can be contacted at email: harsh4no1@gmail.com.

Dr. Veena Kalludi Narasimhaiah received her Ph.D. degree in wireless sensor networs which was awarded by Kuvempu University under the guidance of Dr. B.P. Vijaya Kumar, Professor and Head, Department of Information Science and Engg., MSRIT, Bangalore. She is working as an Associate Professor in the school of Electronics and Communication Engineering, Reva university. She has had around 50 international journals which are indexed in scopus, google scholor and web of science. She was Qualified for GATE All India level exam in 1999 and joined NITK Surathkal for M. Tech in Digital Electronics and Advanced Communication course. She has had many national and international patents under his belt, out of which few patents are granted. She can be contacted at email: veenakn@reva.edu.in.