Results in Engineering 25 (2025) 103784

Contents lists available at ScienceDirect

Results in Engineering

journal homepage: www.sciencedirect.com/journal/results-in-engineering

Research paper

Tea leaf disease detection using segment anything model and deep convolutional neural networks

Ananthakrishnan Balasundaram [a,*], Prem Sundaresan [b], Aryan Bhavsar [b], Mishti Mattu [b], Muthu Subash Kavitha [c], Ayesha Shaik [a]

[a] Centre for Cyber Physical Systems, Vellore Institute of Technology, Chennai, 600127, India
[b] School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, 600127, India
[c] School of Information and Data Science, Nagasaki University, Nagasaki, 8528521, Japan

A R T I C L E I N F O

Keywords:
Tea leaf disease
Zero-shot segmentation
Segment anything model
CNN
MLP

A B S T R A C T

Tea is an important beverage across many cultures. Diseases affecting tea leaves can adversely impact the integrity, production and cause substantial economic losses. Hence, detecting these diseases efficiently and accurately at an early stage is extremely crucial. The dataset used in this work consists of 6 categories to be trained, namely: Algal Spot, Brown Blight, Gray Blight, Healthy, Helopeltis and Red Spot. In our proposed method, a convolutional neural network is used in conjunction with advanced image preprocessing techniques for detecting and segmenting the infected tea leaf region. OpenCV was employed to extract the Region of Interest (ROI) and image cropping was performed to focus only on the leaf. In the process of cropping, the leaf was identified in the image, a bounding box was drawn around it and then it was finally cropped to maximize the leaf in the image. Further, the Segment Anything Model's (SAM) zero-shot segmentation capabilities were tested to segment and extract the diseased regions of the leaf. Also, the images were fed into a custom Convolutional Neural Network (CNN) model to extract the relevant features. These features were subsequently assigned to various classifiers like MLP, SVM, and Decision Tree classifiers to classify the diseases. The performance of each model was analyzed and compared. An accuracy of 95.06 % was achieved demonstrating that the proposed model has relatively higher accuracy in identifying the tea leaf diseases than many of the existing models.

1. Introduction

Tea is the second most consumed beverage in the world, only behind water[1]. Derived from the leaves of plant *Camellia sinensis* almost 5000 years ago in Southeast Asia, the tea plant is now being cultivated in >30 countries. About three billion kilograms of tea is produced and consumed yearly[2]. The importance of tea in India cannot be overstated. The tea business in India is among the oldest and biggest in the world with >1600 tea estates and 500,000 small tea farmers. Over 3 million people are employed by this sector, which also significantly boosts the nation's foreign exchange earnings [5]. In fact, India is the fourth largest exporter of tea in the world, contributing about 10 % of total exports [6].

There are broadly three categories of tea – green tea, black tea, and oolong tea. Freshly harvested tea leaves must be processed to inactive enzymatic oxidation for green tea production, or to control the oxidation by the leaf enzymes to produce oolong and black teas. It can be grown in regions of various altitudes, from sea level to high mountains [3]. The most suitable conditions for tea cultivation include high humidity, with a temperature range of 13 °C to 30 °C, and acidic soils with a pH range of 4.5 to 5.5.

It is strongly believed that tea contains polyphenols and other components which may mitigate the risk of developing chronic diseases such as arthritis, diabetes, cancer and cardiovascular diseases. In fact, it is becoming clear that tea acts as a chemo-preventive agent against certain types of cancers[4]. Tea is also known to stimulate the cardiac function and central nervous system in humans. Tea leaves contain various minerals, including fluoride, manganese, chromium, selenium, calcium, magnesium, and zinc, in varying concentrations, which depend on factors such as the fermentation process, age, and size of the leaves [1].

The vast consumption of tea calls for substantial production, to which tea leaf diseases pose a major threat. Tea leaf diseases like gray blight, brown blight, red spot etc. cause damage to the tea plantations resulting in low quality tea and reduces the yield significantly. Early

* Corresponding author.
E-mail address: balasundaram.a@vit.ac.in (A. Balasundaram).

https://doi.org/10.1016/j.rineng.2024.103784
Received 6 November 2024; Received in revised form 3 December 2024; Accepted 18 December 2024
Available online 19 December 2024

detection of diseases will help the farmers to take preventive measures so that the disease does not spread rapidly in a larger area. Disease detection manually by experts is not only inaccurate but has turned out to be very expensive and since tea plants are grown in loamy soil, which is usually found in the hilly slopes, it is inconvenient to check for diseases.

Microscopic diagnosis and biological diagnosis are presently the prime operational methods being used for identifying tea leaf diseases. Microscopic methods, which concern subjective diagnosis, necessitate the need for skilled specialists. However, it is not very accurate, and even experts may make errors in disease detection. Alternatively, although biological detection methods have higher accuracy, they require relatively expensive equipment and extensive manpower, making them economically unfeasible. These conditions have led to the need for automated and cost-efficient methods for accurate detection of tea leaf diseases.

With advanced Machine Learning algorithms and Deep Learning techniques, leaf disease detection can be performed quickly and effectively. A residual-block based Res4net architecture [7] with a network interactive CBAM was proposed for accurate feature extraction in tea leaf disease diagnosis. A CNN model [8] to accurately detect 5 types of diseases in groundnut leaves, achieving an impressive accuracy of 96.50 %. The work in [22] used the hue value to identify the diseases portion of paddy leaves and proposed a CNN model to accurately detect paddy leaf diseases.

A lightweight federated learning approach [10] for leaf disease classification, ensuring data privacy while achieving high classification accuracy through experiments with both IID and non-IID datasets, demonstrating its effectiveness compared to traditional pre-trained models. Khullar et al. [32]. highlight that federated feature extraction techniques are less resource intensive than federated transfer learning for rice leaf disease classification. The work in [28] show that transfer learning models outperform machine learning classifiers in rice leaf disease detection, with InceptionResnetV2 achieving the highest validation accuracy.

About 70 % of India's population is dependent on agriculture, either directly or indirectly, and agribusiness provides around 17 % of the country's overall GDP [23]. It is imperative to recognize the need for robust leaf disease detection models that can detect leaf diseases early. Diseases like brown blight, gray blight etc. affect tea crop yield and hence cause economic losses to farmers. However, early disease detection and mitigation can reduce these losses. A fully automated means of detection is a sustainable solution to these problems. Deploying this model in the field can mitigate the losses of both small-scale and large-scale farmers. The costs involved in educating the farmers and deploying the technology can be offset by the amount of losses that this technology mitigates in the long run. Hence, adopting these technologies can be beneficial to the tea industry.

In this work, the focus is on early detection and accurate classification of tea leaf diseases. As part of this work, the preprocessing involves object detection, extracting Region of Interest (ROI) along with cropping the leaf images. We also explored Zero-shot segmentation capabilities of Segment Anything Model (SAM) in the form of diseased leaf segmentation. CNN, a very useful and widely used technique, was used for feature extraction and it has an excellent ability to handle large datasets, making it a preferable option for this work. The output of the CNN model will be fed to various classification models like Multilayer Perceptron (MLP), Support Vector Machine (SVM) and Decision Tree (DT) to classify the diseases, and their performance was compared. The key contributions of this research paper are listed below.

1. An image segmentation-based technique was developed and used to assess its impact while detecting diseased tea leaf region.
2. The capability of zero-shot segmentation provided by Segment Anything Model (SAM) was explored to extract and isolate the diseased portions of the tea leaves.
3. A custom CNN architecture was developed to generate highly informative feature maps that can aid in quick and accurate tea leaf disease identification
4. Three classifiers, namely MLP, SVM and DT were tried and the optimal classification approach that works best with the extracted features was determined.

The other sections of the paper include Section 2 discussing contemporary works and Section 3 discussing the proposed Method in detail. Section 4 discusses experimental results and interpretation. Finally, as part of Section 5, the conclusion section summarizes the results drawn from the research.

2. Related works

Leaf disease detection plays a pivotal role in the growth and yield of various crops. Recently developed deep learning and imaging techniques enable the detection of leaf diseases with great accuracy and efficiency. This section summarizes the contemporary works that focus on different methods and approaches of leaf diseases detection is set forth with a brief synopsis of their methodologies, approach, and the results. AX-RetinaNet [13] is used for detecting and identifying diseases on tea leaves. It uses a CNN for feature extraction and a region proposal network to generate bounding boxes. This approach achieved good accuracy in disease detection. The study in [14] introduces an advanced method for classification of tea leaf diseases using YOLOv7, a cutting-edge object detection model. Their findings highlight YOLOv7′s capability to effectively identify and classify tea leaf diseases, providing a valuable tool for enhancing disease management in tea cultivation.

RGB and hyperspectral images were used for developing a deep learning model [15] for rapid classification of tea coal disease. The model is trained on images that had been used for training ResNet18, VGG16 for RGB images, and LSTM and SVM for hyperspectral images. The best performance was attained using the CARS-LSTM for hyperspectral images, clearly showing higher accuracy compared to models that are CARS based. Heng, Yu, and Zhang [16] presented a novel method to perform tea leaf disease identification by employing the concept of hybrid pooling in their architecture. The work [17] introduced a model for classifying tea leaf diseases that integrates a developed CNN with feature fusion and classification techniques which demonstrates significant improvements in disease detection.

Tea leaf disease identification was performed using a CNN [18], demonstrating its effectiveness in enhancing disease detection. The work YOLO-Tea [19] presented an advanced model enhanced by YOLOv5, demonstrating significant improvements in detecting tea leaf diseases. A hybrid architecture based on CNN [20] used SVM and Convolutional Block Attention Module (CBAM), which can detect multiple diseases among multiple crop varieties. Deep Hashing with Integrated Autoencoders [21] was proposed for image retrieval in Tea Leaf images. In this method, autoencoders with skip connections enhanced the importance of key features in earlier tensors, creating a hybrid model for hashing and retrieving images from a tea leaf dataset. The work [9] introduced TRSRD, a database designed for research on risky substances in tea. They used NLP to syntactically analyze literature abstracts, classify them into 9 risk categories and 6 research categories, and extract chemical entities for constructing a Neo4j graph database of tea risk substances and their research relationships. Their work provides valuable resources for understanding and managing potential risks associated with tea consumption.

Kmeans clustering [11] was used for diseased portion extraction followed by transfer learning using EfficientNet to produce a lightweight model with high performance. The work [24] compared the performance of CNN-based architectures including VGG-16, Inception V4, DenseNet-121, and ResNet-50 after fine-tuning them. They found that DenseNet-121 was the best model when tested on the PlantVillage dataset. IterationVIT model [25] was introduced, combining convolutional and iterative transformer techniques to enhance tea disease classification and diagnosis, achieving a high accuracy of 98 % and an F1 score of 96.5 %. By incorporating attention mechanisms and bilinear interpolation, the model effectively addresses several challenges. A Vision Transformer (ViT) model [26] was proposed with hyperparameter tuning for Java plum leaf disease classification and achieved an impressive accuracy of 97.51 %. A combination of various convolutional neural networks coupled with GANs [27] was proposed for data generation to greatly improve small-sample tea disease identification. The usage of a lightweight Inception module [28] was explored to develop a generalized model that can identify diseases across crop varieties with good accuracy. Their generalized approach achieved 94.04 % on the widely acclaimed PlantVillage dataset. transfer learning on ResNet-50 and YOLO-V5 s [29] was utilized and the disease identification was split into three phases – crop identification, health classification and disease localization and identification. Their novel approach achieved accuracies of 95.83 %, 94.13 %, and 82.13 % in the three phases respectively. Vallabhajosyula et al. [30] proposed a novel hierarchical residual ViT(Vision Transformer) using improved Vision Transformer and ResNet9 models which can reduce the number of trainable parameters and extract meaningful features. A VARMAx-CNN-GAN Integration framework [31] was proposed wherein they use CNNs for early disease detection and GAN to synthesize images and improve generalization. Their proposed method combined deep learning with vector autoregressive moving average processes using eXogenous regressors (VARMAx processes) to quickly and accurately find tomato leaf diseases. The work [12] proposed a three-step method – image preprocessing to extract diseased portion of coffee leaves, followed by feature concatenation of GoogLeNet and ResNet based architectures, and finally classification by MLP. They achieved 99.08 % accuracy with their novel method

3. Proposed system

3.1. Dataset

The dataset used in this work was uploaded in Kaggle [34]. The images in this dataset consist of tea leaves from the tea estates of Unakoti district in the state of Tripura in India. The tea leaf disease identification was done under the consultation of tea leaf experts from the tea estates and by employing previously established techniques. The dataset used in this paper contains 5867 leaf images with 6 classes – Algal Spot, Brown Blight, Gray Blight, Healthy, Helopeltis and Red Spot as highlighted in Table 1. A 12MP mobile camera was used to photograph 2800 tea leaves in natural conditions. These photographs were captured at a resolution of 3000 × 3000 pixels by positioning the camera roughly 30 cm above the leaf [33]. As mentioned above, this dataset initially contained 2800 images which were later increased to 5867 images using data augmentation and resized to 256 × 256 pixels by the dataset author. This augmented dataset was uploaded to Kaggle [34] and same was used in this work. As illustrated in Table 1, this dataset contains 867 images of Brown Blight and 1000 images each for the other 5 classes, hence totaling 5867 images.

As illustrated in Table 1, 64 % of the images constituted the training dataset whereas 20 % constituted the testing dataset. Further, 16 % of the data was used to form the validation set during training to make sure that the model is not underfitting or overfitting. By monitoring the model's validation accuracy and validation loss, crucial hyperparameters such as number of epochs and the learning rate of the model were tuned.

Fig. 1

Table 1
Outlines the number of images used for training, validation and testing.

Disease (Class Label)	Number of images		
	Training Set	Validation Set	Test Set
Algal Spot	640	160	200
Brown Blight	554	139	174
Gray Blight	640	160	200
Healthy	640	160	200
Helopeltis	640	160	200
Red Spot	640	160	200

3.2. Image preprocessing

As part of image preprocessing automatic cropping of the leaf images was carried out so that the background details have a minimal effect on the model. This was achieved using the powerful OpenCV libraries wherein the image was first converted to the HSV format, and the Value channel of the image was used to segment the Region of Interest, which is the leaf. Then edge detection was performed using Canny followed by blurring of the image. Adaptive thresholding was then performed on this blurred image to generate a binary image of extracted edges. Using this image, the contours of the image were found. Finally, the contour with the maximum area was chosen to form a rectangular bounding box around the leaf. This bounding box was chosen as the reference to crop the unnecessary background to maximize the leaf's presence in the image. These steps have been illustrated in Fig. 4.

Canny edge detection is a well-known multistage algorithm used to detect edges in an image. Firstly, the image is smoothened to reduce noise. This smoothened image is then processed by a Sobel kernel to get the first derivative in the horizontal direction G_x and vertical direction G_y. With these derivatives, the edge gradient and direction for each pixel are determined in (1) below:

$$Edge\_gradient(G) = \sqrt{G_x^2 + G_y^2} \; Angle(\theta) = \left(\frac{G_y}{G_x}\right) \tag{1}$$

The gradient direction is normal to the edges and is rounded off to one of the four angles- vertical, horizontal, and the two diagonals.

Non-maximum Suppression is applied once the gradient magnitude and direction are found. This helps to suppress any pixels that may not constitute an edge. To achieve this, every pixel is examined to see if it is the local maximum in the direction of the gradient. For example, in Fig. 2, B lies on the horizontal edge and the gradient is perpendicular to it along the vertical direction. A and C lie along the gradient. Now, B is examined to determine if it is the local maximum. If it is not, it is suppressed to zero. In this manner, we obtain a binary image with thin edges. Finally, hysteresis thresholding is applied to pick out the "authentic edges" and drop the rest. This decision is taken through two threshold values "maxVal" and "minVal". If the Intensity Gradient of an edge falls below the minVal, it is discarded and if it exceeds the maxVal, the edge is retained as a "sure edge". However, if the intensity gradient falls between these two thresholds, then the edge is only retained if it connects to a sure edge, otherwise it is dropped.

In Fig. 3 [35], we can see that edge A exceeds the maxVal and is a sure edge. Edge B and edge C lie between maxVal and minVal, however, B is connected to a sure edge A and is retained, whereas C is discarded. In this manner, the strong edges are retained.

In this work, minVal was set to 0 and maxVal was set to 100 and an image with extracted edges was obtained. This image is then blurred before Adaptive Thresholding was applied on them to give a binary image with extracted edges. The contours of this image were found without using any chain approximation. The contour that best fits the leaf is chosen to draw a rectangular bounding box around the leaf which is ultimately used to crop most of the background out of the image.

The step-by-step process has been highlighted in Fig. 4. The first step for us was to choose a suitable channel of a suitable colour space on which we can perform thresholding to extract the RoI. We experimented with the channels of the RGB, LAB, and HSV colour spaces before settling on using the Value channel of the HSV colour space. We

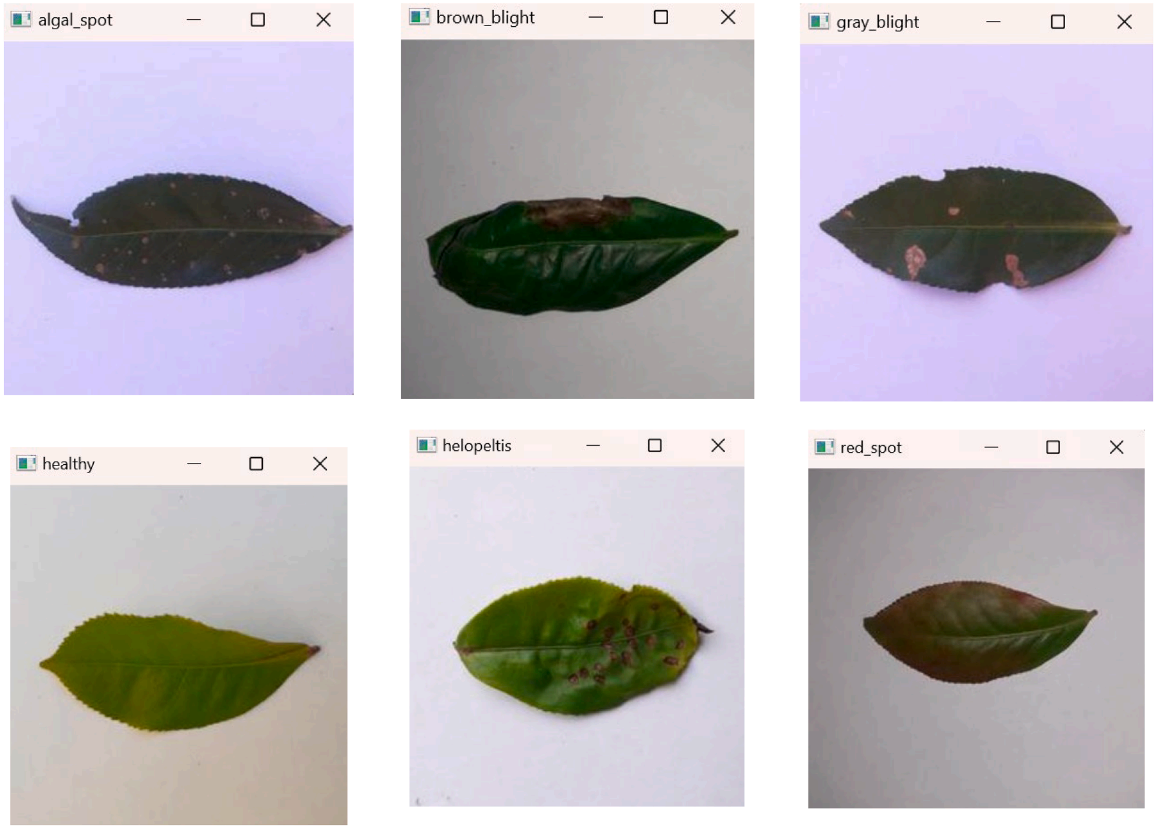


Fig. 1. Sample images of the six different classes in the dataset.

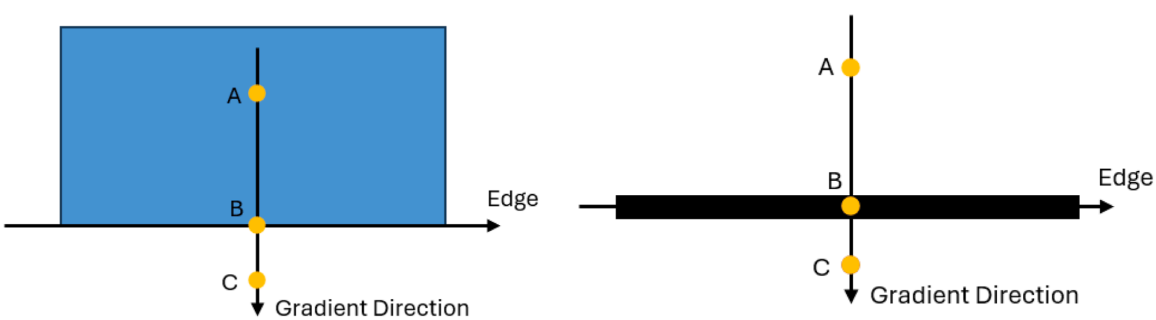


Fig. 2. Working of non-maximum suppression.

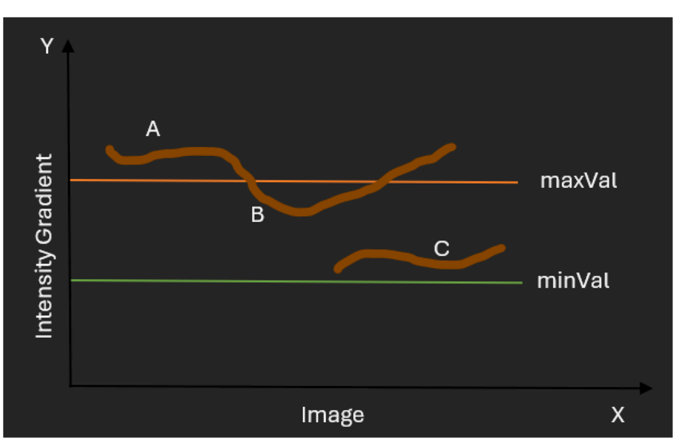


Fig. 3. Working of hysteresis thresholding.

performed Otsu thresholding to calculate the optimal threshold and used this value to perform binary thresholding. The result of this operation is seen in Fig. 4(b).

Canny edge detection was then performed on this RoI image to extract the edges. After blurring the extracted edges, we performed Adaptive thresholding on them to create a binary mask of extracted edges. Finally contours of this binary image were formed and the best contour was selected by selecting the one with the highest area as highlighted in Fig. 4(c). Finally, the original image is cropped according to the selected bounding box as shown in Fig. 4(d).

During this cropping process, we had to deal with varying contrasts and background colours, however, by choosing the Value channel of the HSV colour space and by choosing Otsu's method to find the right threshold, we were able to overcome these challenges and achieve accurate image cropping. This cropping was greatly helpful for us further down the line, especially for zero-shot segmentation with SAM as highlighted in Section 3.3. The SAM model was greatly sensitive to minute lighting changes in the background and by cropping the images

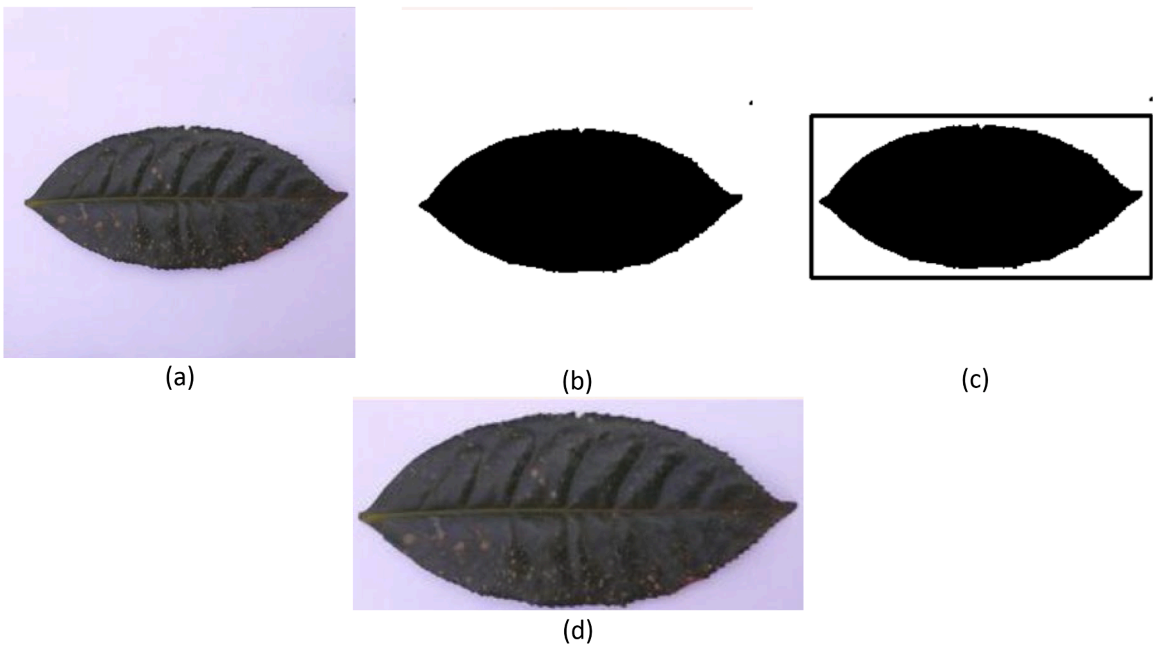


Fig. 4. a) Original leaf image. b) Extracted leaf Region of Interest c) Best fitting bounding box. d) The resulting cropped image.

and maximizing the leaf's presence, we were able to focus SAM's attention on the leaves leading to higher quality segmentation masks.

3.3. *Zero-shot segmentation with segment anything model (SAM)*

The Segment Anything Model (SAM) [36] developed by Meta AI is an advanced image segmentation model that excels in complex image segmentation tasks. Utilizing deep learning methods, SAM is capable of segmenting diverse objects, making it highly adaptable for various computer vision tasks. One of the features of SAM is its zero-shot segmentation capability. This means that SAM is so well generalized that it can perform segmentation on any unseen image without too many difficulties. In this work, the zero-shot segmentation capabilities of SAM has also been explored by using it to extract and isolate the diseased portions of tea leaves. The model checkpoint used in this paper is ViT-B.

The architecture of SAM is illustrated in Fig. 5. It consists of three main components. The image encoder generates high-dimensional feature maps. The prompt encoder encodes user inputs like text, boxes or points to guide SAM. The lightweight mask decoder combines the features maps and user inputs to produce the segmentation mask.

Fig. 6

In this work, we tested out the capabilities and limitations of SAM by using it purely in a zero-shot manner to isolate diseased portions of the leaves. Some of SAM's hyperparameters were tuned to obtain an optimal set of segmentation masks in the multistage process was devised for the zero-shot segmentation of diseased leaf images.

Firstly, the SAM model was initialized with some hyperparameters, including a Predicted IoU Threshold (pred_iou_thresh) of 0.9. This threshold is used to filter both the number and quality of output segmentation masks. In Zero-Shot Segmentation, SAM predicts the quality of each generated mask and compares that with the threshold. Only those masks which exceed the chosen threshold are output by SAM.

Apart from this, a stability score threshold of 0.92 was set. This score measures how consistently a mask maintains its shape when various internal thresholds are used by SAM to convert its soft mask predictions into binary masks. A higher stability score indicates that the mask largely retains its shape across various thresholds.

To handle multiple masks, a novel approach was used. Firstly, the Region of Interest (RoI) of the leaf was extracted in the form of a binary image with the leaf highlighted in white and the background in black. Then, a bitwise AND operation was performed between each segmentation mask and the extracted RoI. This bitwise AND operation ensures

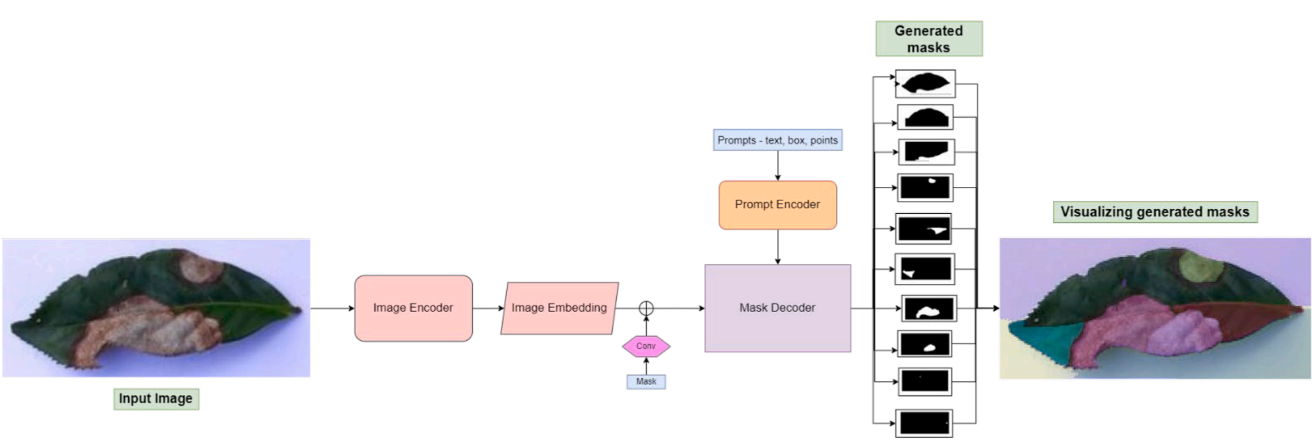


Fig. 5. Architecture of the Segment Anything Model (SAM).

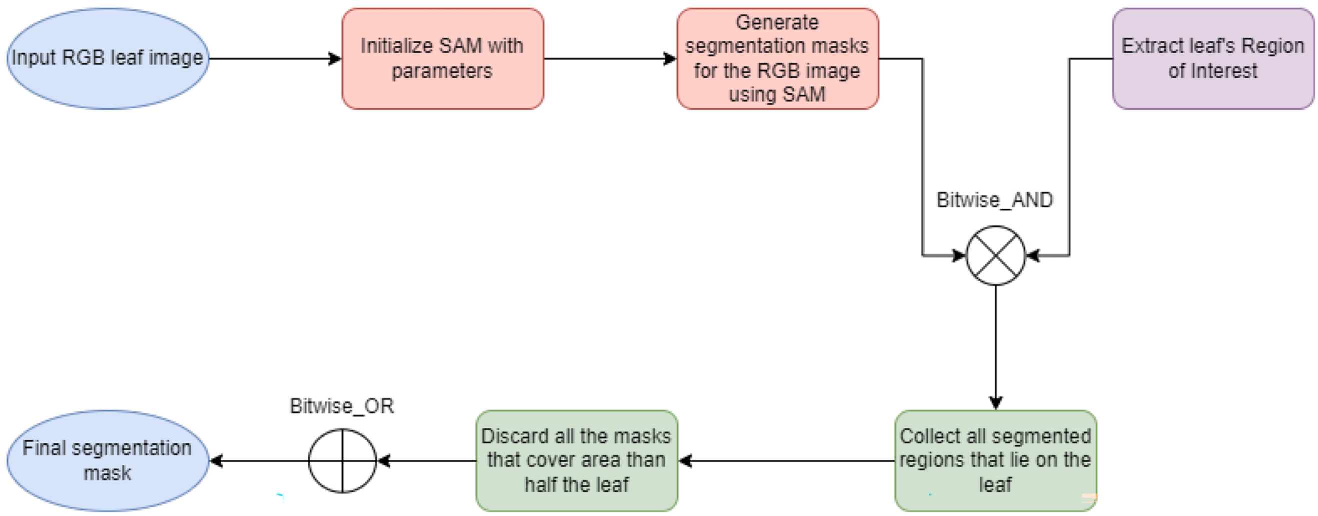


Fig. 6. Tea leaf segmentation using SAM.

that only those masks which contain regions lying on the leaf are retained. The other masks are discarded as they are irrelevant to the diseased leaf.

After obtaining the relevant masks, the size of the segmented region of each mask was compared with the size of the RoI extracted. Any mask which covered >50 % of the extracted Region of Interest's area was discarded. This was done because SAM also generates masks containing half of the leaf or more. Such segmentation masks were not useful to us as they almost cover the whole leaf as opposed to just the diseased part.

Finally, we performed bitwise OR on all the remaining segmentation masks. This was done to combine all the masks that represented some diseased portion on the leaf. Hence, the output of this operation is the culmination of all the diseased portions as recognized by the SAM. Sample outputs for the 6 classes have been illustrated in Fig. 7.

3.4. Proposed model

Convolution Neural Networks have more than tripled the performance of image classification tasks, making deep features extraction possible while efficiently training the models. In this work, a special type of CNN architecture is applied, implemented using the popular Python libraries of Tensor Flow and Keras.

Fig. 8

The key components of the proposed CNN model contribute to increasing the learning and generalizing capability of the network. The model takes an input RGB image with 128 × 128 dimensions. Hence, the input shape is represented as (128, 128, 3). The model has four blocks of convolution layers, each followed by a MaxPooling layer and Batch Normalization layer. All these layers are designed for progressively extracting spatial hierarchies of features from input images. The output from the convolution layers is then flattened into a one-dimensional vector to be used as a feature vector for the Dense Layers.

The proposed method uses MLP to perform disease classification. A dropout layer with a dropout rate of 0.5 to avoid over fitting is added, followed by batch normalization. Added to that is a Dropout layer with a rate of 0.5, followed by batch normalization. The final layer comprises of a fully connected layer with 6 units. Softmax activation is used to output class probabilities. The ReLU function is defined as $f(x) = \max(0, x)$. The ReLU introduces non-linearity in the model, and this activation function proves quite computationally efficient. It hinders the vanishing gradient problem to a certain extent, thus speeding up training. The batch normalization technique is used to normalize the input to every layer. This accelerates training and stabilizes the learning process. The inference mean is computed using (2), while the inference variance and scaling are computed using (3) and (4) respectively.

$$E_x = \frac{1}{m}\sum_{i=1}^{j}\mu_B^{(i)} \tag{2}$$

$$\mathrm{Var}_x = \left(\frac{m}{m-1}\right)\frac{1}{m}\sum_{i=1}^{j}\sigma_B^{2(i)} \tag{3}$$

$$y = \frac{\gamma}{\sqrt{\mathrm{Var}_x + \epsilon}}x + \left(\beta + \frac{\gamma E_x}{\sqrt{\mathrm{Var}_x + \epsilon}}\right) \tag{4}$$

The dropout technique of regularization prevents over fitting by randomly setting a fraction of the input units to zero while training a neural network. After closely monitoring the validation loss and validation accuracy across multiple epochs, we arrived at the dropout value of 0.5 used in our proposed system. ReduceLROnPlateau is the most striking feature of the proposed system. To boost the model's performance, we have applied this function to reduce the learning rate by a certain factor if the model's validation loss does not improve over a number of epochs as specified by the patience parameter. The inclusion of this function alone had boosted our accuracy by almost 3 %. The main motive behind using this function is to ensure that the model keeps improving steadily and does stagnate during the training process.

4. Experimental results and analysis

All model training and evaluations were conducted using the NVIDIA Tesla P100 GPU available on Kaggle. This GPU provides 16 GB of HBM2 memory, which facilitated efficient handling of the large dataset, and complex computations required for training the convolutional neural networks. The deep learning framework used was Tensorflow 2.15.0, and the experiments were executed within the Kaggle Notebooks environment. The parameters used for the learning rate scheduler are included in Table 2 whereas the hyperparameters are highlighted in Table 3.

Table 4, Table 5, Table 6

We trained the proposed CNN+MLP model on two types of training datasets, one containing only cropped leaf images, and another with cropped leaf images overlaid with SAM-generated masks (as shown in the examples on the third column of Table 7). While the model achieved a validation accuracy of >94 % on the cropped images dataset, it achieved a validation accuracy of 92 % on the dataset with SAM masks overlaid on them. While this highlights the limitations of SAM's zero-shot segmentation capabilities in the field of tea leaf disease detection,

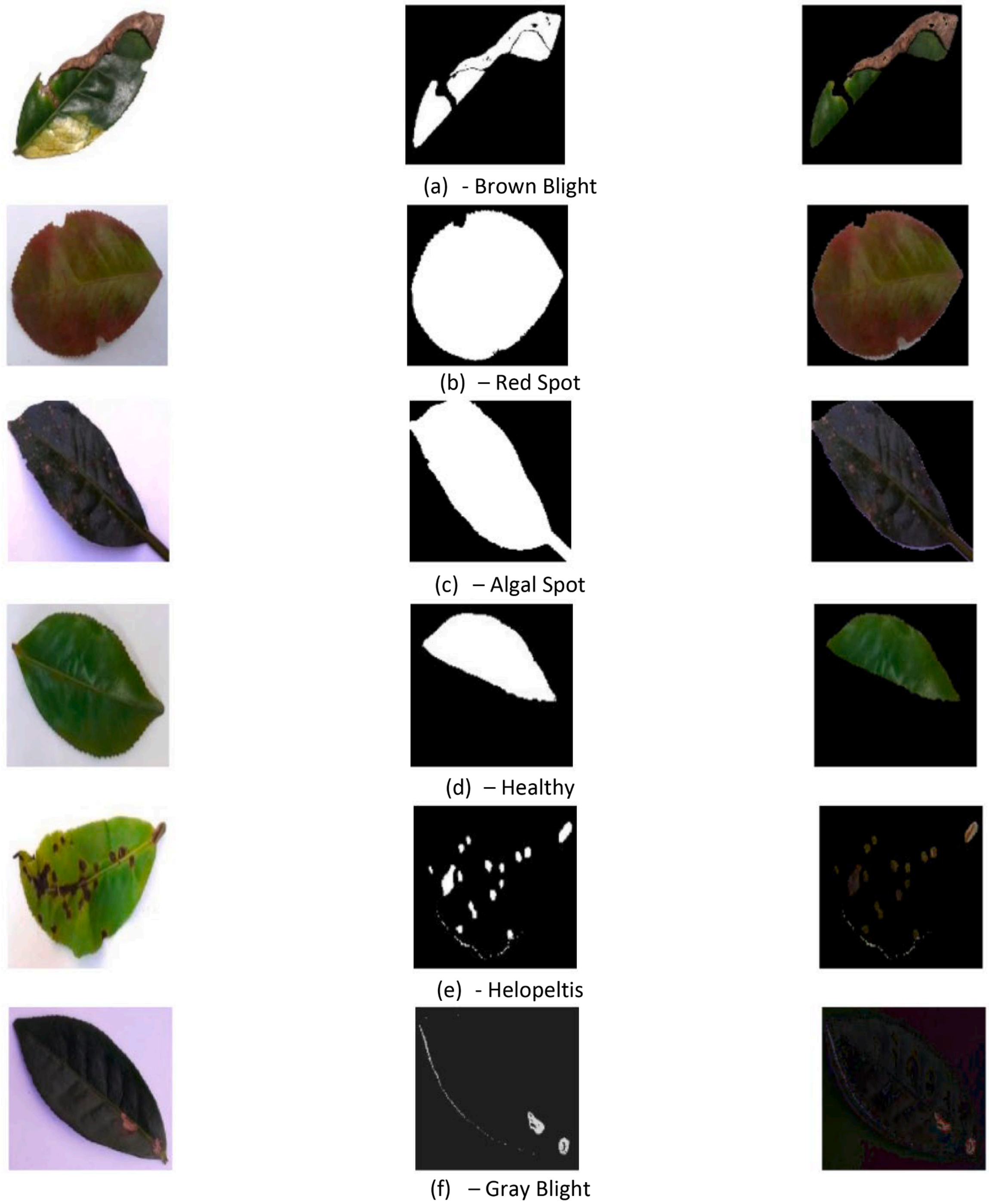


Fig. 7. (a-f) From left to right: The original image, the culmination of masks generated by SAM, the extracted image after overlaying the mask on the image.

we believe that future works can incorporate SAM into their experimentation and employ techniques like hyperparameter tuning to improve the generalization capabilities of their models. In this paper, we proceeded with the aforementioned former version of the model which was trained on the cropped image dataset since it achieved a better validation accuracy. Consequently, the results of this model were analyzed on the test dataset and are discussed below.

$$\mathrm{P} = \frac{TP + FP}{TP} \tag{5}$$

$$\mathrm{R} = \frac{TP + FN}{TP} \tag{6}$$

$$\mathrm{F1} = 2 \times \frac{\mathrm{P} \times \mathrm{R}}{\mathrm{P} + \mathrm{R}} \tag{7}$$

$$\mathrm{A} = \frac{TP + TN}{TP + TN + FP + FN} \tag{8}$$

The evaluation metrics used in this work include precision, recall, F1 score and accuracy which are calculated using the formulae highlighted in Eqs. (5), (6), (7) and (8) respectively. The accuracy of the proposed method was 95.06 %. The precision, recall and F1-score of each class are illustrated in Fig. 9. Fig. 10 shows the confusion matrix and classification report obtained with MLP classifier.

An analysis of the classification report illustrated in Fig. 10(a) shows

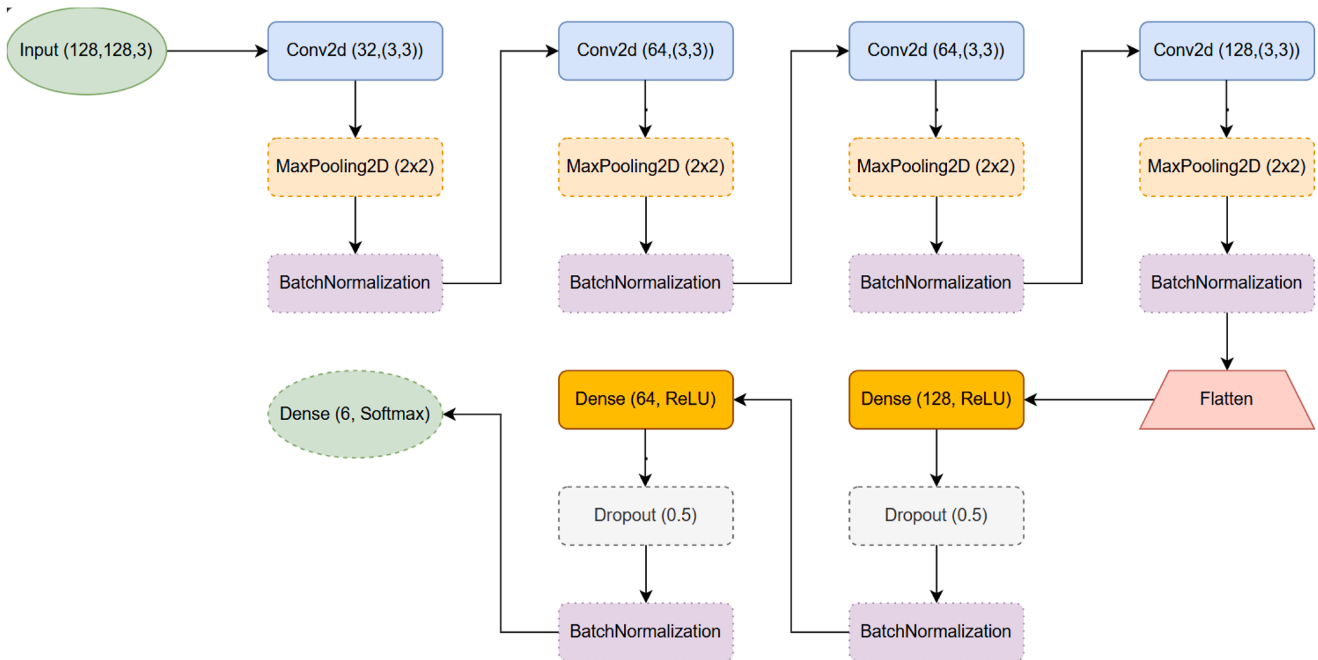


Fig. 8. Proposed CNN model architecture.

Table 2
Parameters of ReduceLROnPlateau function used in the proposed approach.

Parameters	Values
Initial learning rate	0.001
Patience	5
Factor	0.6
Minimum learning rate	0.0001

Table 3
Hyperparameters of the proposed CNN architecture.

Hyperparameters	Values
Batch Size	64
Epochs	80
Loss Function	Categorical Cross Entropy
Optimizer	Adam

Table 4
Shows the legend for the labels in the above classification report and confusion matrix.

Key	Class
0	Brown Blight
1	Red Spot
2	Algal Spot
3	Healthy
4	Helopeltis
5	Gray Blight

Table 5
Shows the class-wise test accuracy of our proposed CNN+MLP model.

Class	True Positives	Total samples	Test Accuracy
Brown Blight	160	174	91.95 %
Red Spot	190	200	95.00 %
Algal Spot	195	200	97.5 %
Healthy	197	200	98.5 %
Helopeltis	184	200	92.00 %
Gray Blight	190	200	95.00 %

Table 6
Hyperparameters for the Decision Tree using GridSearchCV.

Parameters	Values
max_depth	10
min_samples_leaf	2
min_samples_split	5

Table 7
Shows a comparison of our proposed method with existing solutions.

Methodology	Number of diseases	Test Accuracy
Proposed Method	6	95.06 %
Hairah et al. [18].	6	94.55 %
Hossain et al. [37]	3	93.33 %
Mukhopadhyay et al. [38]	5	83.00 %
Ihsan et al. [39]	8	77.48 %

that out of the 6 classes, our proposed model performs the best on Algal Spot and Red Spot in terms of F1 score, achieving scores of 0.97 and 0.96 respectively. On the other hand, our model achieved a marginally lesser F1 score of 0.93 on Brown Blight. This marginal difference shows the overall well-rounded performance of our model. We suspect that this marginal drop in F1 score may have arisen due to the presence of fewer samples of Brown Blight in our dataset and in our future work, we will experiment with utilizing data augmentation to solve the class imbalance in the Brown Blight class.

In terms of precision, our model was the most precise in identifying leaves suffering from Helopeltis, achieving a precision of 0.99. This is followed by a precision of 0.97 achieved on the Red Spot class. Our model achieved a relatively lower precision of 0.91 on the Healthy class. We hypothesize that this issue arises due to some of the diseased leaves having extremely minute discolourations. An analysis of the recall scores illustrated in Fig. 10(a) shows that our model achieved the highest recall of 0.98 on the Healthy class, whereas the lowest recall was 0.92 on the

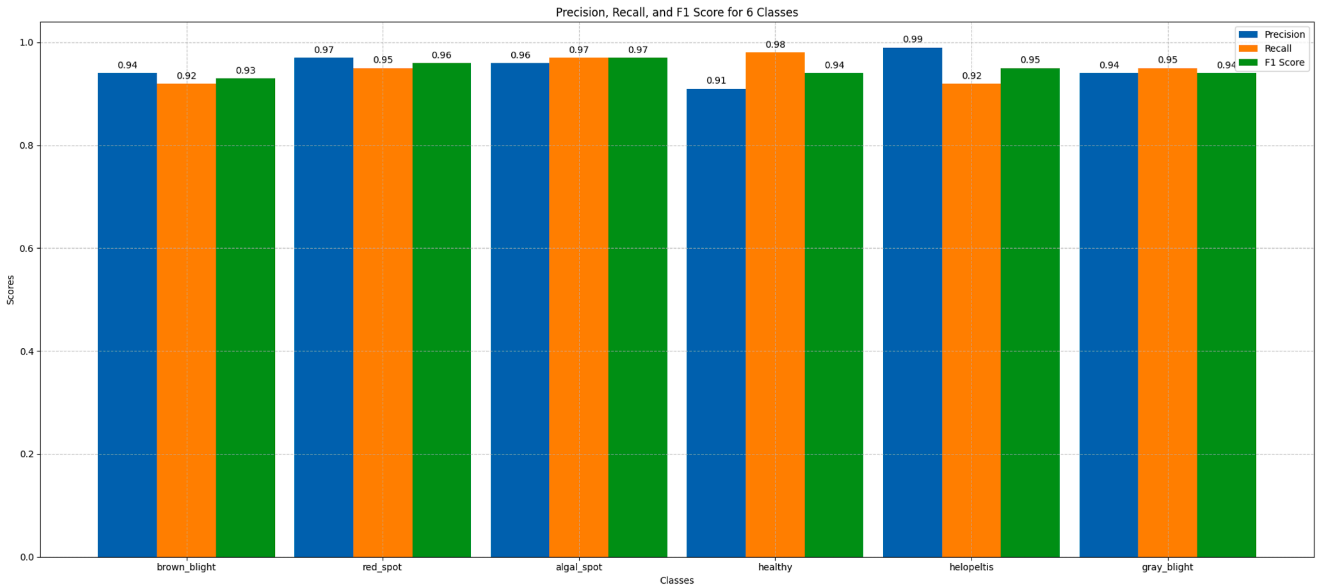
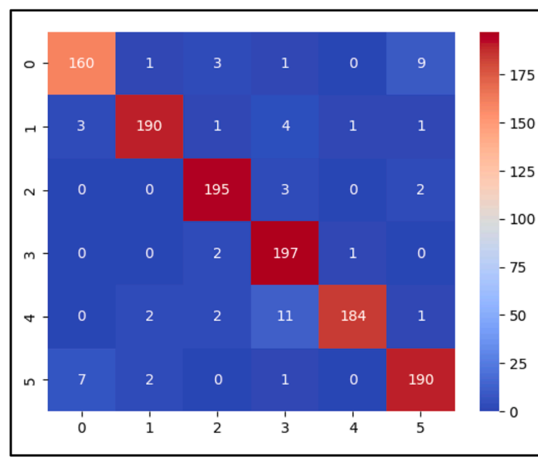


Fig. 9. Precision, recall and F1-score of each of the 6 classes.

	precision	recall	f1-score	support
0	0.94	0.92	0.93	174
1	0.97	0.95	0.96	200
2	0.96	0.97	0.97	200
3	0.91	0.98	0.94	200
4	0.99	0.92	0.95	200
5	0.94	0.95	0.94	200
accuracy			0.95	1174
macro avg	0.95	0.95	0.95	1174
weighted avg	0.95	0.95	0.95	1174

(a) Classification report-MLP



(b) Confusion matrix-MLP

Fig. 10. (a) Classification report-MLP (b) Confusion matrix-MLP.

Brown Blight and Helopeltis classes. This shows that our model is able to identify Healthy leaves well, which is especially crucial in disease detection since misclassifying healthy leaves may lead to unnecessary economic losses for farmers.

As seen in the confusion matrix, the biggest reason for error is the confusion between Gray Blight and Brown Blight. During our pre-processing experimentation, we found that KMeans Segmentation is a great way to differentiate between these two specific diseases. Hence, any future work that wants to perform binary classification on these two specific diseases can adopt KMeans in their method. However, in this paper, we found our approach to be the most effective and generalized one.

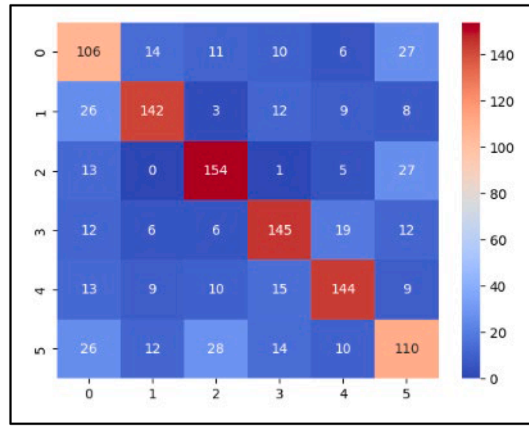
To further affirm the effectiveness of the proposed CNN+MLP model, an ablation study was conducted using Support Vector Machine (SVM) and Decision Tree (DT) classifiers alongside the Multi-Layer Perceptron (MLP). SVM was selected for its ability to handle high-dimensional feature spaces and provide robust classification boundaries, which contributes towards lowering misclassifications and enhancing accuracy. Decision Tree was included due to its simplicity and interpretability. These classifiers offer contrasting paradigms, enabling a comprehensive evaluation of the extracted features. By comparing the performance of these classifiers to MLP, the ablation study highlights the extent to which different machine learning models leverage CNN-extracted features for tea leaf disease classification.

It was observed that while SVM came close to the effectiveness of the proposed MLP, Decision Tree was significantly inaccurate when compared to our proposed system. The classification report as well as confusion matrix of Decision Tree and SVM have been illustrated in Fig. 11 and Fig. 12 respectively. The SVM was used with a Radial Basis Function Kernel. While SAM was successful in segmenting out the diseased portion of the leaves, it was still highly sensitive to the varying lighting effects and hence not completely reliable. SAM is unable to effectively capture the slight discoloration in the diseased leaves and hence the model trained on the dataset with cropped images overlaid with SAM masks underperformed slightly with respect to our proposed model trained on just cropped images.

In this ablation study, we tried to balance model complexity with interpretability. In our future work, we will experiment with other well-

	precision	recall	f1-score	support
0	0.54	0.61	0.57	174
1	0.78	0.71	0.74	200
2	0.73	0.77	0.75	200
3	0.74	0.72	0.73	200
4	0.75	0.72	0.73	200
5	0.57	0.55	0.56	200
accuracy			0.68	1174
macro avg	0.68	0.68	0.68	1174
weighted avg	0.69	0.68	0.68	1174

(a) Classification report - DT

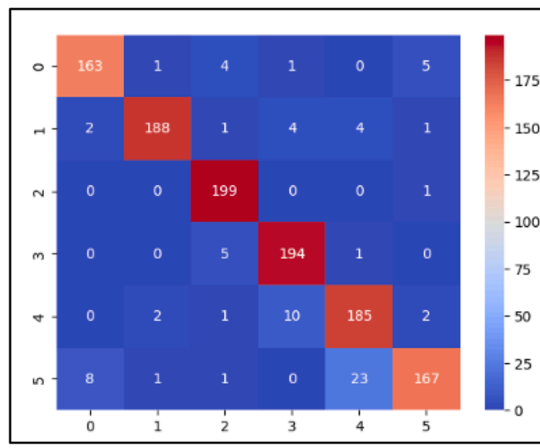


(b) Confusion matrix - DT

Fig. 11. (a) Classification report - DT (b) Confusion matrix - DT.

	precision	recall	f1-score	support
0	0.94	0.94	0.94	174
1	0.98	0.94	0.96	200
2	0.94	0.99	0.97	200
3	0.93	0.97	0.95	200
4	0.87	0.93	0.90	200
5	0.95	0.83	0.89	200
accuracy			0.93	1174
macro avg	0.94	0.93	0.93	1174
weighted avg	0.93	0.93	0.93	1174

(a) Classification report - SVM



(b) Confusion matrix - SVM

Fig. 12. (a) Classification report - SVM (b) Confusion matrix - SVM.

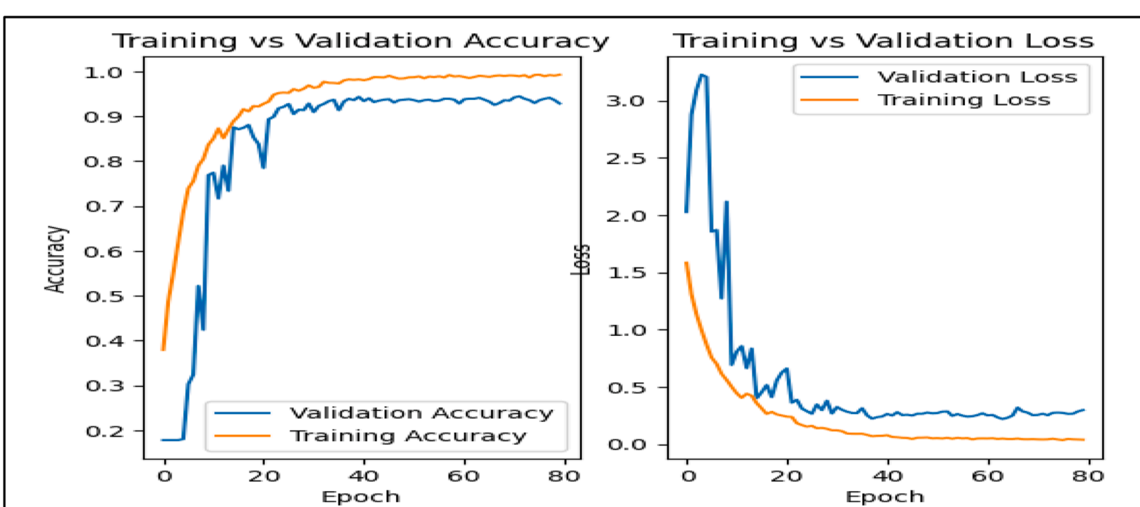


Fig. 13. Plot showing the accuracy and loss plot for the CNN-MLP model.

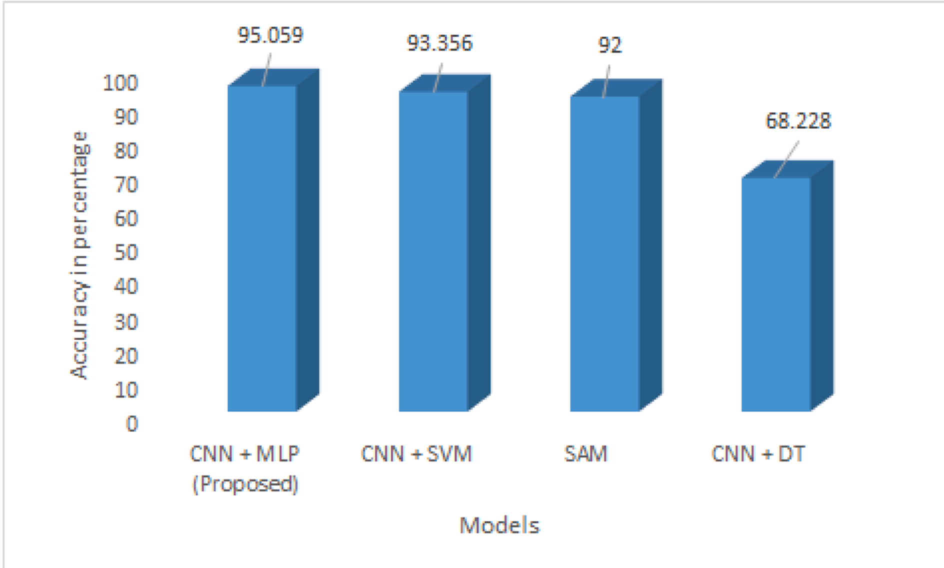


Fig. 14. Comparison of classification accuracy of different models.

known classifiers such as XGBoost and Random Forest among others.

Fig. 13 shows the accuracy and loss plot of the best performing model namely CNN-MLP. Fig. 14 shows the comparison of accuracies of different models. It could be observed that CNN+MLP provided the best accuracy of 95.06 %.

As highlighted in Table 7, our proposed method has performed better than existing solutions for tea leaf disease detection. Hairah et al. [18]. used a CNN with a MobileNetV2 backbone and achieved an accuracy of 94.55 %. Hossain et al. [37]. proposed an SVM classifier which achieved an accuracy of 93.33 %. Mukhopadhyay et al. [38]. utilized Non-dominated Sorting Genetic Algorithm (NSGA-II) based image clustering for detecting disease area and an SVM for classification and achieved an average accuracy of 83 %. Ihsan et al. [39]. compared the performance of multiple classifying algorithms, with Extra Tree Classifier achieving the highest accuracy of 77.48 %.

5. Conclusion and future work

This work was focused on tea leaf disease classification involving five prominent diseases namely Algal Spot, Brown Blight, Helopeltis, Red Spot, Gray Blight and a Healthy class. Having 5867 images, a train-test split of 80–20 was done and a novel approach for cropping the leaf images was devised to maximize the leaf's presence in the image and crop out most of the background. Then zero-shot segmentation capabilities of Meta's Segment Anything Model (SAM) was explored and found to be effective to an extent. However, the validation accuracy of the SAM model when fed with images containing the extracted diseased regions was marginally less than the best performing CNN+MLP model with an impressive 95.06 % test accuracy. As countless farmers depend on tea for their livelihood and millions of people consume tea every year, the proposed model will enable early detection and classification of diseases in tea plantation with greater accuracy, thereby boosting their economy and enhancing the quality of their yield. Future work will be focused on enhancing disease detection accuracy and scaling the model up towards detecting other major tea leaf diseases like Bacterial blight, Scab, Blister blight and Pink disease. This will be achieved through the incorporation of larger and more varied datasets.

CRediT authorship contribution statement

Ananthakrishnan Balasundaram: Writing – review & editing, Writing – original draft, Supervision, Software, Project administration, Methodology, Investigation, Formal analysis, Conceptualization. **Prem Sundaresan:** Writing – review & editing, Writing – original draft, Visualization, Methodology, Formal analysis. **Aryan Bhavsar:** Writing – review & editing, Writing – original draft, Methodology, Data curation. **Mishti Mattu:** Writing – review & editing, Writing – original draft, Visualization, Validation. **Muthu Subash Kavitha:** Writing – review & editing, Writing – original draft, Visualization, Investigation, Conceptualization. **Ayesha Shaik:** Writing – review & editing, Writing – original draft, Investigation, Formal analysis.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors wish to thank the management for VIT Chennai and the Research Centre for Cyber Physical Systems, VIT Chennai for their constant support and encouragement during the course of this work.

Data availability

The data used in this work is available in . https://www.kaggle.com/datasets/saikatdatta1994/tea-leaf-disease

References

[1] M.I. Prasanth, B.S. Sivamaruthi, C. Chaiyasut, T. Tencomnao, A review of the role of green tea (Camellia sinensis) in antiphotoaging, stress resistance, neuroprotection, and autophagy, Nutrients. 11 (2) (2019) 474, https://doi.org/10.3390/nu11020474.

[2] K. Hayat, H. Iqbal, U. Malik, U. Bilal, S. Mushtaq, Tea and its consumption: benefits and risks, Crit. Rev. Food Sci. Nutr. 55 (7) (2015) 939–954, https://doi.org/10.1080/10408398.2012.678949.

[3] C.J. Dufresne, E.R. Farnworth, A review of latest research findings on the health promotion properties of tea, J. Nutr. Biochem. 12 (7) (2001) 404–421, https://doi.org/10.1016/S0955-2863(01)00155-3.

[4] N. Khan, H. Mukhtar, Tea and health: studies in humans, Curr. Pharm. Des. 19 (34) (2013) 6141–6147, https://doi.org/10.2174/1381612811319340008.

[5] A.V. Nair, A. Damani, D. Khandelwal, H. Sachdev, S. Jain, Study on the tea market in India, ArXiv. (2023). https://arxiv.org/abs/2304.07851.

[6] S.V. Shinde, S. Lahade, Deep learning for tea leaf disease classification: challenges, study gaps, and emerging technologies. Applied Computer Vision and Soft

Computing With Interpretable AI, 1st ed., Chapman and Hall/CRC, 2023, p. 22, https://doi.org/10.1201/9781003359456.
[7] P. Bhuyan, P.K. Singh, S.K. Das, Res4net-CBAM: a deep CNN with convolution block attention module for tea leaf disease diagnosis, Multimed. Tools. Appl. 83 (10) (2024) 48925–48947, https://doi.org/10.1007/s11042-023-17472-6.
[8] Maheswaran, S., Indhumathi, N., Dhanalakshmi, S., Nandita, S., Shafiq, I.M., & Rithka, P. (2022). Identification and classification of groundnut leaf disease using convolutional neural network. In L. Kalinathan, R. P. S., & M. S. Kanmani (Eds.), Computational Intelligence in Data Science (Vol. 654, pp. 251–261). Springer. https://doi.org/10.1007/978-3-031-16364-7_19.
[9] Y. Wang, P. Wang, Y. Zhang, S. Yao, Z. Xu, Y. Zhang, TRSRD: a database for research on risky substances in tea using natural language processing and knowledge graph-based techniques, Database 2023 (baad031) (2023), https://doi.org/10.1093/database/baad031.
[10] M. Aggarwal, V. Khullar, N. Goyal, A. Alammari, M.A. Albahar, A. Singh, Lightweight federated learning for rice leaf disease classification using non independent and identically distributed images, Sustainability 15 (16) (2023) 12149, https://doi.org/10.3390/su151612149.
[11] G. Kumar, P. Sharma, A. Gupta, An integrated tea leaf diseases identification and retrieval model using machine learning and deep learning approach, SSRN (2024), https://doi.org/10.2139/ssrn.4757484. https://ssrn.com/abstract=4757484. or.
[12] B.M. Abuhayi, A.A. Mossa, Coffee disease classification using convolutional neural network based on feature concatenation, Inform. Med. Unlocked. 39 (2023) 101245, https://doi.org/10.1016/j.imu.2023.101245.
[13] W. Bao, T. Fan, G. Hu, D. Liang, H. Li, Detection and identification of tea leaf diseases based on AX-RetinaNet, Sci. Rep. 12 (2022) 415, https://doi.org/10.1038/s41598-022-06181-z.
[14] M.J.A. Soeb, M.F. Jubayer, T.A. Tarin, M.R. Al Mamun, F.M. Ruhad, A. Parven, et al., Tea leaf disease detection and identification based on YOLOv7 (YOLO-T), Sci. Rep. 13 (1) (2023) 6078.
[15] Y. Xu, Y. Mao, H. Li, A deep learning model for rapid classification of tea coal disease, Plant Methods 19 (2023) 98, https://doi.org/10.1186/s13007-023-01074-2.
[16] Q. Heng, S. Yu, Y. Zhang, A new AI-based approach for automatic identification of tea leaf disease using deep neural network based on hybrid pooling, Heliyon. 10 (5) (2024) e26465, https://doi.org/10.1016/j.heliyon.2024.e26465.
[17] N. Yücel, M. Yıldırım, Classification of tea leaves diseases by developed CNN, feature fusion, and classifier based model, Int. J. Appl. Math. Electronics Comput. 11 (1) (2023) 30–36, https://doi.org/10.18100/ijamec.1235611.
[18] U. Hairah, A. Septiarini, N. Puspitasari, A. Tejawati, H. Hamdani, S.E. Priyatna, Classification of tea leaf disease using convolutional neural network approach, Int. J. Electr. Comput. Eng. (IJECE) 14 (3) (2024) 3287–3294, https://doi.org/10.11591/ijece.v14i3.pp3287-3294.
[19] Z. Xue, R. Xu, D. Bai, H. Lin, YOLO-tea: a tea disease detection model improved by YOLOv5, Forests. 14 (2023) 415, https://doi.org/10.3390/f14020415.
[20] P. Kaur, A.M. Mishra, N. Goyal, S.K. Gupta, A. Shankar, W. Viriyasitavat, A novel hybrid CNN methodology for automated leaf disease detection and classification, Expert. Syst. (2024), https://doi.org/10.1111/exsy.13543.
[21] S.K. Jayapal, S. Poruran, Enhanced disease identification model for tea plant using deep learning, Intell. Autom. Soft Comput. (2023), https://doi.org/10.32604/iasc.2023.026564.
[22] Maheswaran, S., Sathesh, S., Rithika, P., Shafiq, I.M., Nandita, S., & Gomathi, R.D. (2022). Detection and classification of paddy leaf diseases using deep learning (CNN). In E. J. Neuhold, X. Fernando, J. Lu, S. Piramuthu, & A. Chandrabose (Eds.), Computer, communication, and Signal Processing (Vol. 651, pp. 51–62). Springer. https://doi.org/10.1007/978-3-031-11633-9_6.
[23] M. Aggarwal, V. Khullar, N. Goyal, Contemporary and futuristic intelligent technologies for rice leaf disease detection, in: Proceedings of the 10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), IEEE, 2022, pp. 1–6, https://doi.org/10.1109/ICRITO56286.2022.9965113.
[24] A. J, J. Eunice, D.E. Popescu, M.K. Chowdary, J Hemanth, Deep learning-based leaf disease detection in crops using images for agricultural applications, Agronomy 12 (10) (2022) 2395, https://doi.org/10.3390/agronomy12102395.
[25] B. Zhan, M. Li, W. Luo, P. Li, X. Li, H. Zhang, Study on the tea pest classification model using a convolutional and embedded iterative region of interest encoding transformer, Biology. (Basel) 12 (2023) 1017, https://doi.org/10.3390/biology12071017.
[26] A.C. Bhowmik, M.T. Ahad, Y.R. Emon, F. Ahmed, B. Song, Y. Li, A customised vision transformer for accurate detection and classification of Java plum leaf disease, Smart Agric. Technol. (2024) 100500, https://doi.org/10.1016/j.atech.2024.100500.
[27] G. Hu, M. Fang, Using a multi-convolutional neural network to automatically identify small-sample tea leaf diseases, Sustainable Comput. Informatics Syst. 35 (2022) 100696, https://doi.org/10.1016/j.suscom.2022.100696.
[28] I. Bouacida, B. Farou, L. Djakhdjakha, H. Seridi, M. Kurulay, Innovative deep learning approach for cross-crop plant disease detection: a generalized method for identifying unhealthy leaves, Inf. Process. Agric. (2024), https://doi.org/10.1016/j.inpa.2024.03.002.
[29] R. Abbasi, P. Martinez, R. Ahmad, Crop diagnostic system: a robust disease detection and management system for leafy green crops grown in an aquaponics facility, Artif. Intell. Agric. 10 (2023) 1–12, https://doi.org/10.1016/j.aiia.2023.09.001.
[30] S. Vallabhajosyula, V. Sistla, V.K.K. Kolli, A novel hierarchical framework for plant leaf disease detection using residual vision transformer, Heliyon. 10 (9) (2024) e29912, https://doi.org/10.1016/j.heliyon.2024.e29912.
[31] V. Cheemaladinne, S.R K, Tomato leaf disease detection and management using VARMAx-CNN-GAN integration, J. King Saud Univ. - Science 36 (8) (2024) 103340, https://doi.org/10.1016/j.jksus.2024.103340.
[32] M. Aggarwal, V. Khullar, N. Goyal, T.A. Prola, Resource-efficient federated learning over IoAT for rice leaf disease classification, Comput. Electron. Agric. 221 (2024) 109001, https://doi.org/10.1016/j.compag.2024.109001.
[33] S. Datta, N. Gupta, A novel approach for the detection of tea leaf disease using deep neural network, Procedia Comput. Sci. 218 (2023) 2273–2286, https://doi.org/10.1016/j.procs.2023.01.203.
[34] S. Datta, Tea Leaf Disease, Kaggle.com, 2022 [online]Available at: https://www.kaggle.com/datasets/saikatdatta1994/tea-leaf-disease.
[35] OpenCV. (n.d.). Canny edge detection. Retrieved August 1, 2024, from https://docs.opencv.org/4.x/da/d22/tutorial_py_canny.html.
[36] Facebook Research, *Segment Anything Model Architecture* [Image], GitHub, 2024. https://github.com/facebookresearch/segment-anything.
[37] S. Hossain, R.M. Mou, M.M. Hasan, S. Chakraborty, M.A. Razzak, Recognition and detection of tea leaf's diseases using support vector machine, in: 2018 IEEE 14th International Colloquium on Signal Processing & Its Applications (CSPA), IEEE, 2018, pp. 150–154.
[38] S. Mukhopadhyay, M. Paul, R. Pal, D.P. Mukherjee, Tea leaf disease detection using multi-objective image segmentation, Multimed. Tools. Appl. 80 (1) (2021) 753–771, https://doi.org/10.1007/s11042-020-09567-1.
[39] C.N. Ihsan, N. Agustina, M. Naseer, H. Gusdevi, J.F. Rusdi, A. Hadhiwibowo, F. Abdullah, Comparison of machine learning algorithms in detecting tea leaf diseases, Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi) 8 (1) (2024) 135–141, https://doi.org/10.29207/resti.v8i1.5587.