INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

RESEARCH ARTICLE

Check for updates

OPEN ACCESS

**Received:** 18-05-2023

**Accepted:** 21-05-2023

**Published:** 07-06-2023

**Citation:** Donu JV, Santhi K (2023) Real Time Classification and Detection of Apple Leaf Diseases at Early Stage by Employing Enhanced Paddy Field Pattern Search Method (EPF-PSM) with 2LC. Indian Journal of Science and Technology 16(22): 1682-1693. https://doi.org/10.17485/IJST/v16i22.1206

[*] **Corresponding author**.

donuscholar@gmail.com

**Funding:** None

**Competing Interests:** None



Published By Indian Society for Education and Environment (iSee)

**ISSN**
Print: 0974-6846
Electronic: 0974-5645

# Real Time Classification and Detection of Apple Leaf Diseases at Early Stage by Employing Enhanced Paddy Field Pattern Search Method (EPF-PSM) with 2LC

**Jose V Donu[1]*, K Santhi[2]**

**1** Department of Computer Science, Dr.N.G.P Arts and Science College, Tamilnadu, 641048, Coimbatore, India
**2** Department of Information Technology, Dr.N.G.P Arts and Science College, Tamilnadu, 641048, Coimbatore, India

## Abstract

**Objectives:** To propose a suitable machine learning approach to detect and classify the leaf diseases in apple trees at the pre-mature state to reduce plant degradation. In order to boost the accuracy level, EPF-PSM (Enhanced Paddy Field-Based Pattern Search Method) is employed along with 2LC (Two Level Classifier) technique. **Methods:** Preprocessing images, which include FS and FE, is done by employing EPF-PSM to detect the apple leaf disease. 2LC classifier is utilized to classify the type of disease based on the available features like gradient, pixel, edges, etc. The decision tree algorithm is employed to search for patterns in images and compare them for accurate prediction. With MC-Apriori (Multi Label Classification Apriori), the relationship between infected and non-infected leaves in search space is explored, and RGB scheme (RGBCS) is employed to discover the colour depth ratio of infected leaves. EPF-PSM extracts attributes from the ALDD-NW dataset (collected at the Apple Experiment Station of Northwest A&F University, China) and selects the suitable function to perform ALD classification, pre-mature detection, and accuracy. MATLAB software is utilized for implementation and assessment of the novel ML approach. The findings are compared with contemporary models such as the IFPA-GA with SVM-SVI, the FR-CNN, the R-SSD, and the INARSSD. **Findings:** The suggested model gives 98.01% accuracy rate, 97.06% detection speed in 29 seconds, 98.04% sensitivity, 95.92% specificity, 98.01% precision, 96.17% recall, 98.4% TPR, 99.03% TNR, and 96.07% F-Score, which is comparatively higher than the existing methods in terms of detection and classification of ALD. **Novelty:** The results shows that the suggested machine learning (ML) technique EPF-PSM with 2L-C has the ability to produce accurate detection and classification of ALD at the pre-mature stage, which helps the farmers to treat apple plants properly. The evident results outperform the prevailing methods IFPA-GA with SVM-SVI, FR-CNN, R-SSD, and INARSSD.

**Keywords:** Apple Leaf Disease Detection; Classification Techniques; Machine Learning Algorithm; Pattern Search Method; Paddy Field Method; Image Processing

## 1 Introduction

Apple trees are affected by a wide range of diseases, many of which cause significant damage to their leaves. These diseases can be caused by environmental conditions and also by fungi, bacteria, and viruses. Detection and diagnosis of ALD at a pre-mature stage tend to be very difficult with the use of prevailing image recognition methods. Researchers use various methods to predict and classify the disease, but the accuracy is not impressive. Real-time detection using image processing and advanced ML methods to detect ALD and classify the type of disease at its earliest occurrence has been taken as a key problem for this study. Alternaria leaf spot, brown spot, mosaic spot, grey spot, and rust spot are the 5 types of ALD taken as samples for this study. ALD can have a significant impact on the health and productivity of apple trees, which can in turn have an impact on the availability and price of apples in the marketplace. In order to overcome the drawbacks of conventional methods, the new bio-inspired technique is used in combination with 2LC, PSM, MC-Apriori, and RGBCS for accurate classification and detection of ALD, which will help farmers treat the disease at the correct time and increase the productivity of apples. As a part of background study, many DL methods were proposed to detect and classify the ALD in a robust manner. But all the works have reported with less accuracy in disease prediction and has few drawbacks in the system in terms of root scanning, image depth ratio calculation, colour segmentation etc. hence, the present work will be suitable for classification and prediction at early stage with high accuracy.

Various ML techniques were used to detect and boost the level of classification. DCNN with an attention mechanism was proposed to predict ALD at an early stage, where the attention wedge was integrated into the system to get the spatial location of the image. To quantify the process quickly and easily, the swish activation function was utilized. The competitive performance was achieved in detection but failed to classify the type of ALD in advance[(1)]. Scientometric analysis was done by the authors with a sample of 1897 apple leaf images to perform image processing where the CNN method was employed. IR-Image Restoration, IF-Image Filtration, and IE-Image Enhancement were used during the preprocessing stage, and the EDS method was employed along with CNN to spot the colour variation in the leaves to identify the disease in a rapid manner. The accuracy achieved is 89%, which is comparatively low in terms of classifying the ALD type[(2)]. IFPA-MCSVM with SVI genetically inspired techniques was proposed to detect the disease with the help of SGDM and RGA methods. The colour segments of the leaves are taken for comparison to identify the infected and non-infected areas. Biotic and abiotic methods were used along with SGDM to spot the ALD. 90% accuracy is achieved, and the system has drawbacks in spotting the ALD at an early occurrence[(3)]. The E-IWDARP method was proposed to boost accuracy, as the error detection rate of this algorithm is remarkable. Rust grading is done using AM, where mobile devices are used to scan the images in real-time. During the process, the system will grade the rust formation in the leaves, which will help the farmers treat and prevent it at an early stage. The ALDD-FP dataset was used to test and train the algorithm[(4,5)]. Virus detection in GF leaves is done using an intelligent technique called FCM and K-Fold validation. Unhealthy parts in the grape leaves are marked using FCM and validated using K-Fold. 500 iterations were carried out, which resulted in virus detection and classification of the GF leaf. The average accuracy reported by the system is 87%. RVSRP-IR is employed to improve the sensing of images in the search space during real-time scanning[(6,7)].

The DCNN method was introduced to maximize the speed of ALD classification. The ensemble of pre-trained images was utilized, and various image augmentation processes were carried out. The Canny edge-based detection method was used along with DCNN and produced an accuracy result of 89.06%. Though the speed is nominally acceptable, the system failed to detect ALD during the real-time process [8]. The PCM (Parallel Convolutional Model) was introduced to identify the ALD in its growing stage. The 3x3 matrix method is integrated with PCM for feature selection and extraction. Further, AM is added to scan the complexity of the images uploaded in the dataset. Finally, the difference in the leaves is spotted, and detection is done with an accuracy of 86.78%. Here, the mathematical GBM method is also employed to multiply the matrix values [9]. The colour segmentation method along with DWTM was introduced by the author to spot the gradient and rust colours in the leaves to classify the type of disease. Colour histogram features and vector-based feature selection were carried out during the preprocessing stage, and the fusion method was deployed to detect the ALD. The plant village dataset was used. 91% detection accuracy is attained with a 95% hue ratio [10]. The RCABCRP technique was developed to accelerate deep sensing for classification and detection in a robust manner and achieve 91% accuracy. The performance of the ALD detection method was analyzed using ML and DL models using VGG-19, and significant results were produced by this model, which overcomes the drawbacks of versions like CGG-17 and NETRes-16 [11,12]. LICM (Leaf Image Classification Model) was introduced by the authors to prove that technological advancements in smart farming are necessary in order to avoid spraying pesticides on apple trees that affect human health. XGBoost and Gaussian-Blur methods were also employed, along with LICM, where Kaggle VGA datasets were used with 9876 images on 5 different types [13]. An improvised CNN-based detection system was proposed for detecting the guava leaf diseases at an early stage, and the RGBT method was deployed to mark the edges and measure the search space to compare the values and get the optimal value. The process is repeated for 500 iterations to reach an accuracy level of 89.09%. The ALDD-MA dataset was used to test and train the model [14]. The attention-based CNNM method attempted by the researchers aims to spot the disease in the complicated search space and low-clarity images. Multi-scale JB was employed to identify the similarities in apple leaves and mark them to train the algorithm. Asymmetric convolutional techniques are also employed along with MM to offer dynamic detection of ALD. But the drawback of MM-ABCNNM is that the type of disease is not classified accurately [15]. The new detection method called canny edge detection is employed for automatic disease segmentation in humid climates. During real-time scanning, the system will initially scan the boundaries of the leaves and then go for root scanning to spot the colour difference in the leaves. Once the colour difference is spotted, the type of disease is identified with the help of similarity. The accuracy of this model was 92%, and the speed of detection was 47 seconds [16]. The E-SSD model was projected as the best approach to efficiently identify the ALD in any environmental condition using the mathematical wheel selection method for feature selection and extraction. This approach is well used in terms of real-time scanning using mobile devices [17]. FRCNN was proposed to identify the disease in real time by utilizing RGN. It is one of the pattern analysis methods in machine intelligence that aims to use the system in all climatic conditions. Loss function and cross function are deployed for preprocessing. PASCAL-VOC datasets were used. The bounding box regression technique is also employed to calculate the depth ratio in leaves [18]. A new method called INARSSD was developed to measure the pixel values and compute the marked proportion rate. The similarities are initially spotted, and we randomly record the significant values to get an optimal result. The accuracy in detection was achieved up to 87%, whereas the classification is done, which is not impressive [19]. Other ML and DL models were introduced to detect multiple diseases in plants, which showed a progressive increase in terms of detection and classification accuracy [20–23]. Though all prevailing approaches have not shown the remarkable performance, the new method is proposed in this research for efficient detection and classification of ALD. To carry out the uninterrupted operations in detection at pre-mature stage, few steps to be followed such as,

- **Image Transformation:** Geometric, colour, and frequency domain transformations, which include scaling, rotating, translating, etc., will be used for image recognition. Adjustment ratios called brightness, contrast, and saturation are also processed in image transformation.
- **Restoration :** IR is the process of correcting the damaged or non-clarity images with the help of denoising, inpainting, super-resolution techniques to achieved desire quality.
- **Filtrating & Enhancement:** F&E is the process of removing certain images and modifying the appearance of the images in ALDD-NW dataset.
- **Segmentation :** Converting images into segments or regions for easy identification of all the parts of image for robust scanning.

The purpose of this work is to propose a novel method to (1) classify the type of ALD; (2) detect and classify at a pre-mature level; (iii) increase the accuracy level of detection and classification by employing the bio-inspired machine learning method EPF-PSM with 2LC; and (iv) categorize the infected and non-infected leaves with the help of the RGBCS scheme.

## 2 Methodology

The suggested approach focused mainly on the classification and detection of ALD at a pre-mature level as well as categorizing the infected and non-infected leaves using the RGBCS scheme. EPF-PSM works with a unique pattern search method in combination with 2LC to boost the accuracy level. The dataset is loaded, and pre-processing is done by ML methods. The suggested model uses a bioinspired paddy field-based pattern search method for FS and FE. MC-Apriori is employed to identify the difference in leaves (infected or non-infected), and RGBCS is used to mark the edges and colour segmentation. When compared to prevailing methods such as IFPA-GA with SVM-SVI, FR-CNN, R-SSD, and INARSSD, the new approach provides high accuracy in classifying the disease type and speedy detection of ALD, which will help farmers treat the apple leaves in a robust manner.

### 2.1 Proposed Methodology

The suggested method aims to identify and classify the kind of ALD at an early stage, classify the type of leaves (healthy and non-healthy), and maximize true positive instances to enhance performance. EPF-PSM with 2LC has introduced a new pattern search approach in which the image's intensity ratio is determined first. It is one of the hybrid approaches where the boundary and image subject are calculated from the trained data. The image intensity level (IIL) is identified and the threshold level for segmentation is chosen based on IIL. The ALD-spotted region is identified using RGBCS, and pixels are noted. High sensitivity and specificity are recorded when MC-Apriori is employed and multi-class variants are used to plot the infected region in the leaf images. Leaf patterns are identified and compared with trained data to spot the disease in an effective manner. PSM have capability of visually inspecting the leaves and search for patterns in apple leaves, records the location and severity and then compare the same with the test data. PSM boosts the detection speed where the scanners are used in real time scenario. The dataset is converted to TF (Transactional Format) and clean the irrelevant process and apply multi class label to improvise the class vectors. Assume that the number of loaded ALLD-NW Images $N$, $I$ represents the preprocessing parameter. Set the threshold value as $Th$ and set the shape of the image as , $W$ , where $H$ represents height and $W$ represents width of the image. Assign the value $P$ for pattern marking and measure the region and mark as $HL$ and $UHL$. Derive the threshold value by using the following equation,

$$Th\,Value\; = \left(\frac{N_{th-n}}{sum\,(H\;and\;W)}\right]^{\hat{}}(P/(I-1)\;\mid match\,UHL\;and\;HL = PR) \tag{1}$$

where,$N_{th-n}$ is the image threshold value and $sum\,(H\;and\;W)$ is to calculate the total boundaries of leaf in the region. After calculating the position, apply PSM model, where $PR$ denotes the pattern recognition and $P/(I-1)$ calculates the marking value of whole image. Calculate the updated region by using the below equation,

$$Region\;R_{image} = \left(\frac{Range\;\;(N)}{Pattern\;\;(I)}\right]^{1}\,x\,\left(\frac{2}{N-1}\right)range\;value) \tag{2}$$

where, $Region\;R_{image}$ is the pattern ratio of the healthy and unhealthy leaves. PSM is deployed to mark the boundaries and patterns at the initial level and pixel values are recorded to compare the patterns using test data.

### 2.2 Data attainment, acquisition and pre-processing

This study makes use of 26,377 images from the ALDD-NW (apple leaf disease dataset, China) which has five different types of disease spots. Machine learning techniques are used to pre-process collected images. The data was gathered and all of the images are from different environmental conditions. Table 1 shows the amount of train and test images, and Figure 1 shows the different types of disease images.

**Table 1.** Apple Leaf Disease Dataset - ALDD-NW

| S.No | Disease | Train/Test Images | Total Number |
|---|---|---|---|
| 1. | ALS-Alternaria Leaf Spot | 4121/1222 | 5343 |
| 2. | BP-Brown Spot | 4316/1339 | 5655 |
| 3. | MP-Mosaic Spot | 3770/1105 | 4875 |
| 4. | GS-Grey Spot | 3328/1482 | 4810 |
| 5. | RP-Rust Spot | 4238/1456 | 5694 |

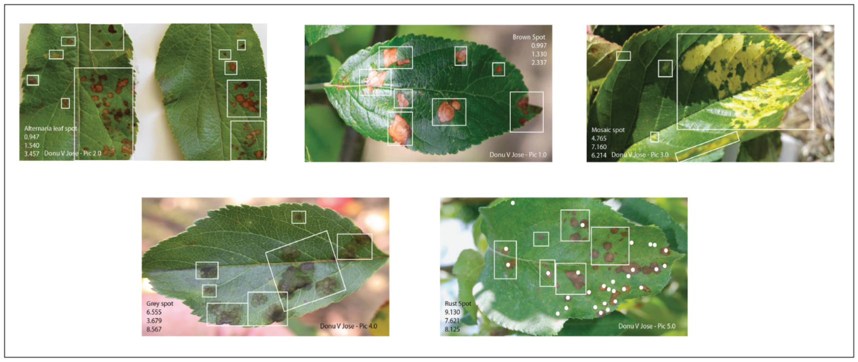


**Fig 1.** Different Types of Common Apple Leaf Diseases (ALS-Alternia Spot, BS-Brown Spot, MP-Mosaic Spot, GS-Grey Spot and RP-Rust Spot)

The data distribution is done for training and testing the effectiveness of the algorithm. EPF-PSM with 2LC is used for FS and FE and evaluates the performance to achieve the accuracy level. The steps of FS and FE, which include

- Raw data collection & pre-processing to remove noisy datas
- Resize the ALDD-NW images for best fit
- Data filtration and restoration task is carried out
- Convert the leaf images into grayscale which can reduce the computational complexity by utilizing colour function vector
- Augment the data (flipping, rotating, shifting, increasing brightness, contrast and hue)
- Splitting the datas – Train, Test and Validate
- Label encoding using L-Encoder

## 2.3 Enhanced Paddy Field based Pattern Recognition ML Method

It is one of the nature inspired techniques to find the best out of given population. The core concept behind this EPF-PSA method is the seed pollination process of a paddy crop (PC). Pattern search uses matrix model to recognize the pattern and find the optimal value. When a PC grows in a region, it determines the prospect generations of that crop based on the fitness owing to pollination. Some seeds are dispersed in adjacent locations, when their possibility and fitness are at their peak. When a seed is less fit, it will generate inferior seeds. This process is repeated until the best solution is found. The population of paddy seeds has grown. EPF-PSM method combines and gives the optimal value. PSM follows 7 steps for image recognition such as,

**Input:** ALDD-NW pre-processed image to be searched and a pattern to search for.

**Output:** The global positions of the pattern in the ALDD-NW image.

- Convert the image and the pattern into grayscale version
- Compute the dimensions of the pre-processed image and the pattern
- Rows and Columns iteration can be done to spot pixel values
- For each pixel in the image, compare the pattern against the first pixel
- If the pixel values are equivalent, execute a supplementary comparison of the current pattern and the corresponding section of the image to determine if the pattern is present
- If the pattern is present, record the global positions of the pattern in the ALDD-NW image
- Return the position of the pattern in the image.

Let us take that $P(a\ \&\ b)$ as position of the pattern image, where $I$ is the number of segmented images and $O$ is the optimal value. The Equation for the same is written as,

$$Optimal\ Value\ =\ P(a,b) = ((a_x,\ b_y)_x\ €\ R_{n,}\ J_y\ €\ \{0,1\}\}\ O^{i=1} \tag{3}$$

where, $P_x$ can be written as, $P_x = (R_{1,}\ \ R_{2,}\ \ R_{3,}\ \ R_n\ \ \}$ . Here, R denotes Number of Rounds/Iterations). The identified pattern sets are compared against the test data and derived by the following equation to identify the supersets of images.

$$Superset =\ P\ (a.b)\,match^{\,i+1} = IR\left(P^{features\ st}, \eta\right)\ \ - \left(a^{i+1}.B^{finite+h}\right) \tag{4}$$

where, $P\ (a.b)\,match^{\,i+1}$ denotes the finite pattern matching and number of ALDD-NW images are selected for patter recognition.

**EPF-PSM with 2LC Algorithm**

1. **Input:** The Matlab Settings with ALDD-NW Images
2. **Begin**:
3. Initialize the datasets with a random population of seeds$PS$
4. Calculate the fitness of each seed
5. ***If*** fitness value is best $FS_x$
6. **Then** set new value $OV_x$
7. Sort the paddy seeds, in order of fitness $FS_x$
8. Produce seeds from the best individuals, where the best fits produce more seeds
9. Pollinate $IP_n$seeds in the neighboring space
10. Disperse the plants, According to Gaussian distribution, the next generation of seeds produced by each plant is scattered within the parameter space; the positions of the seeds are measured $Xi+1 = NXi+PS$
11. Reach termination if the termination conditions are met $OV_x = PS$ .
12. Update the best position $Nx$ position
13. ***End IF***
14. **Output:** Optimal Solution
15. End

## 2.4 2LC (Two Level Classifier) for classification

2LC is one of the most effective approaches for determining the specific type of ALD. To determine the optimal value, two major levels are employed. The 1st level seeks to classify the ALDD-NW images into classes, while the follow up level focuses on particular disease subcategories. Let's assume that the trained set of data and test data input are as follows, $(A_{Train}\ ,\ A1_{Train})$ and$(A_{Test})$. $(X_{Train}\ )$ and $(\ B_{predict})$ is the prediction value of the 2LC classifier. The following are the steps to be followed in classification:

- For each ALD category identified in the 1st level classification, create exclusive datasets with masked images for that specific disease
- Train the separate classifiers for each ALD category to further classify the images within the specific disease category
- Fine-tune the classifiers using the disease-specific datasets
- Evaluate the performance of each disease-specific classifier

The optimal solution $ALD\ Solution$ can be calculated with the help of below equation,

$$Optimal\ V2\ ALD\ Solution\ (\ B_{predict})\ = (A_{Train}\ ,\ A1_{Train})|\ (A_{Test})\ (It_{1-n}) \tag{5}$$

where, $(It_{1-n})$ shows different iterations to perform 2LC for ALD classification operation.

## 2.5 RGBCS and ERFE technique for colour segmentation & pattern recognition

Enhanced Recursive Feature Eradication is one of the regions filling and extraction models used to detect the leaf edges from a boundary. RGBCS is the colour encoding scheme used to isolate the colour ranges in the leaves, which helps the user differentiate between healthy and unhealthy leaves. Initially Load the images to run the RGBCS scheme and convert them into colour space. Train the model and predict the disease based on the number of iterations. 2-LC combines both L1-linear and L2-nonlinear classifiers to enhance accuracy in spotting specific types of images. The enhanced RGBCS performance is shown in Equation 6.

$$IRGBCS =\ \ (A_{Train}\ ,\ A1_{Train})|\ (A_{Test})\ (It_{1-n})\ \ *\textstyle\sum_{i=0}^{n}\left(\frac{i}{j}\right)x^{i} \tag{6}$$

where, the term $IR$ denotes the fitness value and with the help of loaded train and test data $(A_{Train}\ ,\ A1_{Train})$ the most prediction state is obtained.

### 2.6 Accuracy and Effectiveness of EPF-PSM with 2LC

The following are the various states of performance and accuracy when employing EPF-PSM with 2LC:

- The detection of ALD in the pre-mature state is obtained with the help of the new paddy crop method.

All five types of diseases are detected at the initial stage as Class A and Class B sets for further extraction, and the findings are used to assess the performance level of EPF-PSM with 2LC.

### 2.7 About MATLAB R2020a Implementation Process

The proposed EPF-PSM with 2LC model against the baseline ML models to detect the ALD at pre-mature stage is evaluated using the MATLAB tool. MATLAB gives users a clear pictorial analysis where the image processing technique can be easily demonstrated. Wide range of toolboxes is available for IP, ML and DL. IP toolbox, CV toolbox, DL toolbox, ML toolbox were used for easy access. MATLAB is currently used by many researchers, engineering applications, IT calculations, prediction analysis graphs etc. The application permits the user to explore raw data to graphical representations and compares the existing versions to current version in the form of plotting graph with the help of various parameters.

### 2.8 Performance Analysis Metrics of EPF-PSM with 2LC

The recommended ML method EPF-PSM with 2LC was compared against the prevailing approaches IPFA-GA with SVM-SVI [3], R-SSD [17], FR-CNN [18] and INARSSD [19] which were selected as the existing methods in the preceding section. To assess the performance of EPF-PSM the following formula were used.

$$\text{Accuracy} = \frac{(TPE+TNE)}{(TPE+TNE+FPE+FNE)} \times 100 \quad (7)$$

$$\text{Sensitivity} = \frac{TPE}{(TPE+FNE)} \times 100 \quad (8)$$

$$\text{Specificity} = \frac{TNE}{(TNE+FPE)} \times 100 \quad (9)$$

$$\text{Precision} = \frac{TPE}{(TPE+FPE)} \times 100 \quad (10)$$

$$CC = \frac{T_1}{\sqrt{T_2 \times T_3 \times T_4 \times T_5}} \times 100 \quad (11)$$

where, the PEM equation is derived using,
$T_1 = (TPE \times TNE - FPE \times FNE)$, $T_2 = (TPE+FPE)$, $T_3 = (TPE+FNE)$, $T_4 = (TNE+FPE)$, and $T_5 = (TNE+FNE)$.

- **Sensitivity and Specificity** - A statistical method is used to evaluate the screening test and measure how well the test spots leaves with disease.
- **Accuracyand F score** - Accuracy is the proportion of exact predictions of disease in apple leaves made by the system out of all predictions where Fscore is used to balance the P&R in classifier.
- **True Positive Rate and True Negative** - TPR shows the actual positive cases correctly spotted by the system where TNR shows the actual negative cased spotted among all leaf images.
- **Detection Speed** - Evaluates the time and speed of the ALD detection by the system and records the values in all iterations.
- **Precision and Recall** - Precision measures the amount of TP among all positive predictions, whereas recall calculates the amount of TP among all true positive cases.

## 3 Results and Discussion

This section clearly demonstrates the implementation results of the unique ML and DL technique EPF-PSM with 2LC in comparison to the baseline versions IPFA-GA with SVM-SVI [3], R-SSD [17], FR-CNN [18] and INARSSD [19]. The proposed model shows evident results and overcomes the drawbacks in terms of accurate classification and ALD prediction in the pre-mature state. FS, FE, and object detection are done using enhanced paddy fields and pattern search methods with non-linear constraints. The MC Apriori approach is used to determine the association between infected and non-infected leaves. Figures 2, 3, 4, 5, 6 and 7 show the observed findings, with the X axis plotted with performance efficiency metrics and the Y axis plotted with final output values.

### 3.1 Sensitivity and Specificity Performance Analysis

Figure 2 depicts the sensitivity and specificity performance analysis of EPF-PSM with 2LC against IPFA-GA with SVM-SVI [3], R-SSD [17], FR-CNN [18] and INARSSD [19]. It has been observed that EPF-PSM with 2LC provides remarkable and improved results than existing models. The MC-AP approach differentiates between infected and non-infected leaves, allowing the system to detect disease at an early stage with high accuracy rate. The suggested DL technique performs well in terms of searching for leaf patterns, separating them, and generating optimal value due to thorough classification and data extraction. 98% and 95% are achieved in EPF-PSM, where all other models show less performance.

**Table 2.** Performance Values of Sensitivity and Specificity

| Metrics / Schemes | R-SSD [17] | FR-CNN [18] | INARSSD [19] | IFPA [3] | EPF-PSM with 2-LC |
|---|---|---|---|---|---|
| **Sensitivity** | 82.08 | 80.07 | 83.64 | 94.01 | 98.04 |
| **Specificity** | 80.17 | 79.13 | 83.76 | 93.38 | 95.91 |

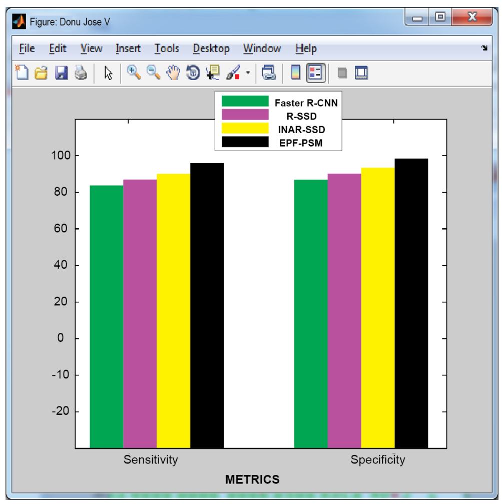


**Fig 2.** Sensitivity and Specificity Performance Analysis

### 3.2 Accuracy in detection Analysis

Figure 3 clearly portrays the accuracy level of detecting ALD at a pre-mature stage of the proposed EPF-PSM with 2LC against IPFA-GA with SVM-SVI [3], R-SSD [17], FR-CNN [18] and INARSSD [19]. As EPF-PSM with 2LC utilized both the global and local pattern search methods with the RGB colour scheme for image scanning and 2LC for classification, it resulted in high accuracy in identifying the infected leaves at an initial level in all environmental conditions. Also, the accuracy level was met up to 96% and goes up to 98%, where it outperforms the baseline versions.

**Table 3.** Performance Values of Accuracy

| Metrics / Schemes | R-SSD [17] | FR-CNN [18] | INARSSD [19] | IFPA [3] | EPF-PSM with 2-LC |
|---|---|---|---|---|---|
| **Accuracy IT-1** | 81.02 | 79.01 | 85.07 | 94.01 | 96.09 |
| **Accuracy IT-N** | 82.05 | 76.01 | 87.04 | 94.09 | 98.01 |

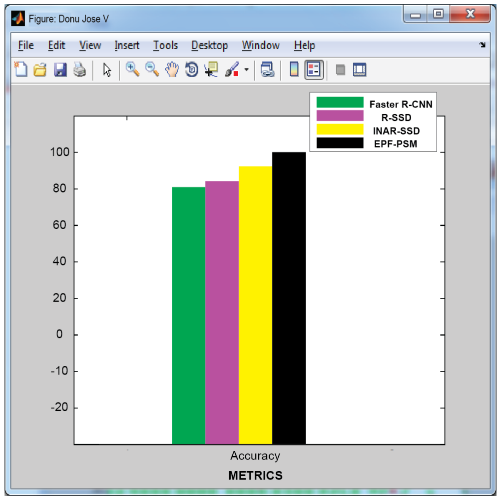


**Fig 3.** Performance ofAccuracy in detection & classification

### 3.3 True Positive and True Negative Rate Performance Analysis

Figure 4 reveals the true positive and true negative performance analysis of the proposed ML approach EPF-PSM with 2LC against IPFA-GA with SVM-SVI [3], R-SSD [17], FR-CNN [18] and INARSSD [19]. On chosen ALDD-NW datasets with specified parameters like as colour, spots, pixel ratio, and so on, the Paddy field technique outperforms the previous methods. The suggested ML approach has high TP & TN rate in disease classification and early prediction of ALD due to its extensive selection and extraction procedures. Based on the dataset values, the confusion matrix is utilized to determine TPR and TNR. 98.4% and 99.03% are achieved in EPF-PSM with the 2LC method, whereas other algorithms have not shown superior performance.

**Table 4.** Performance Values of TP and TN

| Metrics / Schemes | R-SSD [17] | FR-CNN [18] | INARSSD [19] | IFPA [3] | EPF-PSM with 2-LC |
|---|---|---|---|---|---|
| **True Positive** | 76 | 82 | 84 | 90 | 98.4 |
| **True Negative** | 74 | 79 | 81 | 87 | 99.03 |

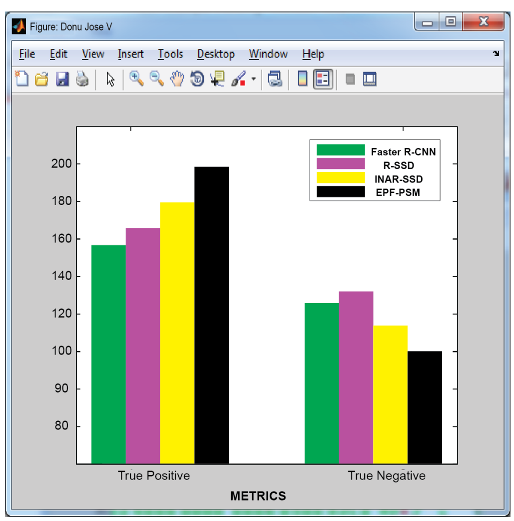


**Fig 4.** Performance of TP and TN

### 3.4 Speed Detection Analysis

Figure 5 demonstrates the speed and time of ALD detection of the proposed technique EPF-PSM with 2LC against IPFA-GA with SVM-SVI (3), R-SSD (17), FR-CNN (18) and INARSSD (19). When compared to existing models, the new ML approach, EPF-PSM with 2LC, provides greater performance with proven results. The suggested technique calculates the process of 2LC and boundary scanning speed to increase detection speed and reduce time. The DT technique is used to search for patterns in the leaves using training and test data. In 29 seconds, 97% of the speed is attained.

**Table 5.** Performance Values of Speed Detection

| Metrics / Schemes | R-SSD (17) | FR-CNN (18) | INARSSD (19) | IFPA (3) | EPF-PSM with 2-LC |
|---|---|---|---|---|---|
| **Speed** | 71.02 | 62.03 | 79.14 | 93.07 | 97.06 |
| **Time (in seconds)** | 79 | 98 | 76 | 51 | 29 |

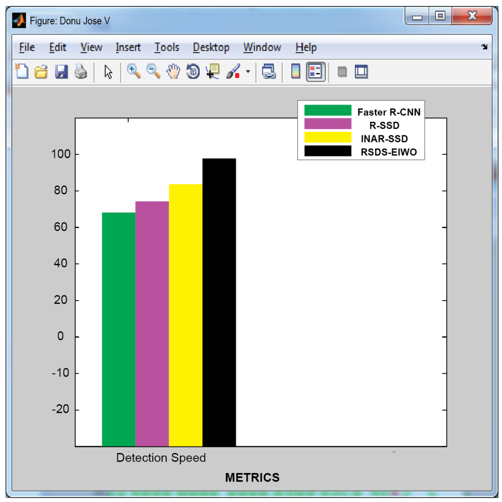


**Fig 5.** Performance of Speed & Time

### 3.5 Precision and Recall Performance Analysis

Precision and recall analysis is projected in Figure 6 of the proposed novel DL method EPF-PSM with 2LC and compared against IPFA-GA with SVM-SVI (3), R-SSD (17), FR-CNN (18) and INARSSD (19). In this new method, the image boundaries are scanned and the disease is detected based on the colour variance of the images. Because of its high scanning depth ratio, the EPF-PSM with 2LC performs efficiently and consistently in instances where other ML models fail to achieve.

**Table 6.** Performance Values of Precision and Recall

| Metrics / Schemes | R-SSD (17) | FR-CNN (18) | INARSSD (19) | IFPA (3) | EPF-PSM with 2-LC |
|---|---|---|---|---|---|
| **Precision** | 87.03 | 84.08 | 90.01 | 95.01 | 98.7 |
| **Recall** | 83.08 | 80.02 | 89.07 | 93.17 | 96.3 |

### 3.6 F-Score Analysis

Figure 7 illustrates the comparative analysis of F-Score of the EPF-PSM with 2LC and compared against IPFA-GA with SVM-SVI (3), R-SSD (17), FR-CNN (18) and INARSSD (19). On the chosen images in the ALDD-NW datasets, the EPF-PSM with 2LC outperforms the existing methods. The proposed model performs well in segregating the leaf images into infected and non-infected by spotting the colour difference using the ERFE method. The F-Score results achieved up to 95% for n iterations. It is noteworthy that the proposed hybrid approach has achieved proven results.

**Table 7.** Performance Values of F-Score analysis

| Metrics / Schemes | R-SSD [17] | FR-CNN [18] | INARSSD [19] | IFPA [3] | EPF-PSM with 2-LC |
|---|---|---|---|---|---|
| **F-Score (It 1)** | 81.02 | 79.08 | 88.08 | 92.07 | 95 |
| **F-Score (It N)** | 80.01 | 77.11 | 87.04 | 91.09 | 94 |

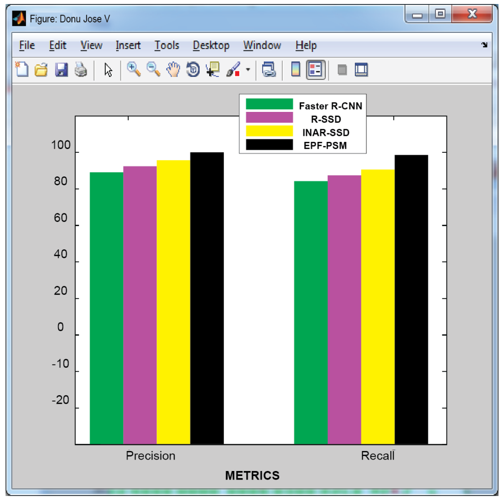


**Fig 6.** Performance ofPrecision and Recall

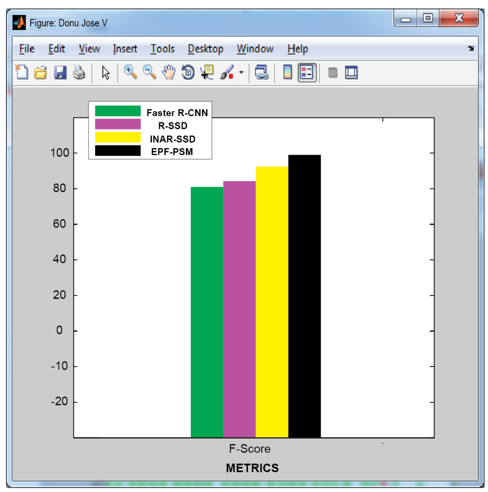


**Fig 7.** Performance of F-Score

## 4 Conclusion

The proposed hybrid approach has a unique ALD prediction and classification method called EPF-PSM (Enhanced Paddy Field Pattern Search Method) with 2LC to improve disease prediction and classification at the pre-mature stage. The accuracy is achieved up to 98% with an ALD detection speed rate of 97% in 29 seconds, whereas existing algorithms only attain up to 91% with a speed of 51 seconds. EPF-PSM uses an apple leaf dataset (ALDD-NW) collected from the Apple Experiment Station of Northwest A&F University in China for testing and implementation. EPF-PSM has global search capability, where it searches the image and marks the boundaries with the help of the RGB colour scheme. Here leaf patterns are compared to identify infected and non-infected leaves with the help of trained and test data. 2LC is employed to boost the classification accuracy and identify the type of disease in the leaf. MC-AP is utilized to identify the relationship between the infected and non-infected leaves. The ERFE technique is applied to recognize the significant features in apple leaf images to boost accuracy at a higher level. The results show 98% and 95% sensitivity and specificity, 96% F-score, 98% and 96% precision and recall, and 98% and 94% TPR and TNR. A few limitations in EPF-PSM with 2LC are noted during the study, where the EPF-PSM method will work only if non-linear and linear classifiers mark the boundary and edges of the images during root scan process. But, in the case of changing leaf images like guava, areca, etc., the EPF-PSM will slow down in accuracy. The algorithm may be enhanced in the future to spot diseases in all types of leaves, which will be a significant contribution for agriculture and horticulture.

## References


1) Wang P, Niu T, Mao Y, Zhang Z, Liu B, He D. Identification of Apple Leaf Diseases by Improved Deep Convolutional Neural Networks With an Attention Mechanism. *Frontiers in Plant Science*. 2021;12. Available from: https://doi.org/10.3389/fpls.2021.723294.
2) K S, P VR, P R, M PK, S P. Apple Leaf Disease Detection using Deep Learning. In: 6th International Conference on Computing Methodologies and Communication (ICCMC). 2022;p. 1063–1067. Available from: https://doi.org/10.1109/ICCMC53470.2022.9753985.
3) Donu V, Jose, Santhi K. 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*. 2022;15(29):1440–1450. Available from: https://doi.org/10.17485/IJST/v15i29.1235.
4) Nithyanandh S, Jaiganesh V. Quality of service enabled intelligent water drop algorithm based routing protocol for dynamic link failure detection in wireless sensor network. *Indian Journal of Science and Technology*. 2020;13(16):1641–1647. Available from: https://doi.org/10.17485/IJST/v13i16.19.
5) Mi Z, Zhang X, Su J, Han D, Su B. Wheat Stripe Rust Grading by Deep Learning With Attention Mechanism and Images From Mobile Devices. *Frontiers in Plant Science*. 2020;11. Available from: https://doi.org/10.3389/fpls.2020.558126.
6) Mohammadpoor M, Nooghabi MG, Ahmedi Z. An Intelligent Technique for Grape Fanleaf Virus Detection. *International Journal of Interactive Multimedia and Artificial Intelligence*. 2020;6(1):62–62. Available from: https://doi.org/10.9781/ijimai.2020.02.001.
7) Nithyanandh S, Jaiganesh V. Dynamic Link Failure Detection using Robust Virus Swarm Routing Protocol in Wireless Sensor Network. *International Journal of Recent Technology and Engineering*. 2020;8(2):1574–1578. Available from: https://doi.org/10.35940/ijrte.b2271.078219.
8) Bansal P, Kumar R, Kumar S. Disease Detection in Apple Leaves Using Deep Convolutional Neural Network. *Agriculture*. 2021;11(7):617. Available from: https://doi.org/10.3390/agriculture11070617.
9) Fu L, Li S, Sun Y, Mu Y, Hu T, Gong H. Lightweight-Convolutional Neural Network for Apple Leaf Disease Identification. *Frontiers in Plant Science*. 2022;13:831219. Available from: https://doi.org/10.3389/fpls.2022.831219.
10) Hasan S, Jahan S, Islam MI. Disease detection of apple leaf with combination of color segmentation and modified DWT. *Journal of King Saud University - Computer and Information Sciences*. 2022;34(9):7212–7224. Available from: https://doi.org/10.1016/j.jksuci.2022.07.004.
11) Nithyanandh S, Jaiganesh V. Reconnaissance Artificial Bee Colony Routing Protocol to Detect Dynamic Link Failure in Wireless Sensor Network. *International Journal of Scientific & Technology Research*. 2020;10(10):3244–3251. Available from: https://doi.org/10.35940/ijstr.b2271.0986231.
12) Sujatha R, Chatterjee JM, Jhanjhi NZ, Brohi SN. Performance of deep learning vs machine learning in plant leaf disease detection. *Microprocessors and Microsystems*. 2021;80:103615. Available from: https://doi.org/10.1016/j.micpro.2020.103615.
13) Kalwad PD, Kanakaraddi SG, Preeti T, Ichalakaranji S, Salimath S, Nayak S. Apple Leaf Disease Detection and Analysis Using Deep Learning Technique, Information and Communication Technology for Competitive Strategies (ICTCS 2020). 2022. Available from: https://doi.org/10.1007/978-981-16-0739-4_76.
14) Di J, Li Q. A method of detecting apple leaf diseases based on improved convolutional neural network. *PLOS ONE*. 2022;17(2):e0262629. Available from: https://doi.org/10.1371/journal.pone.0262629.
15) Perveen K, Kumar S, Kansal S, Soni M, Alshaikh NA, Batool S, et al. Multidimensional Attention-Based CNN Model for Identifying Apple Leaf Disease. *Journal of Food Quality*. 2023;2023:1–12. Available from: https://doi.org/10.1155/2023/9504186.
16) Srinidhi VV, Sahay A, Deeba K. Plant Pathology Disease Detection in Apple Leaves Using Deep Convolutional Neural Networks : Apple Leaves Disease Detection using EfficientNet and DenseNet. *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)*. 2021;p. 1119–1127. Available from: https://doi.org/10.1109/ICCMC51019.2021.9418268.
17) Jeong J, Park H, Kwak N. Enhancement of SSD by concatenating feature maps for object detection. *Procedings of the British Machine Vision Conference 2017*. 2017;p. 1–20. Available from: https://arxiv.org/abs/1705.0958720.
18) Ren S, He K, Girshick R, Sun J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2017;39(6):1137–1149. Available from: https://doi.org/10.1109/TPAMI.2016.2577031.
19) Jiang P, Chen Y, Liu B, He D, Liang C. Real-Time Detection of Apple Leaf Diseases Using Deep Learning Approach Based on Improved Convolutional Neural Networks. *IEEE Access*. 2019;7:59069–59080. Available from: https://doi.org/10.1109/ACCESS.2019.2914929.
20) Luo Y, Sun J, Shen J, Wu X, Wang L, Zhu W. Apple Leaf Disease Recognition and Sub-Class Categorization Based on Improved Multi-Scale Feature Fusion Network. *IEEE Access*. 2021;9:95517–95527. Available from: https://doi.org/10.1109/ACCESS.2021.3094802.
21) Singh S, Gupta S, Tanta A, Gupta R. Extraction of Multiple Diseases in Apple Leaf Using Machine Learning. *International Journal of Image and Graphics*. 2022;22(03). Available from: https://doi.org/10.1142/S021946782140009X.
22) Eldho KJ. Impact of Unbalanced Classification on the Performance of Software Defect Prediction Models. *Indian Journal of Science and Technology*. 2022;15(6):237–242. Available from: https://doi.org/10.17485/IJST/v15i6.2193.
23) Bonkra A, Bhatt PK, Rosak-Szyrocka J, Muduli K, Pilař L, Kaur AL, et al. Apple Leave Disease Detection Using Collaborative ML/DL and Artificial Intelligence Methods: Scientometric Analysis. *International Journal of Environmental Research and Public Health*. 2023;20(4):3222. Available from: https://doi.org/10.3390/ijerph20043222.