



# Medicinal plants recognition using heterogeneous leaf features: an intelligent approach

Manoj Sharma<sup>1</sup> · Naresh Kumar<sup>2</sup> · Shallu Sharma<sup>3</sup> · Sumit Kumar<sup>4</sup> ·  
Sukhjinder Singh<sup>1</sup> · Seema Mehandia<sup>5</sup>

Received: 4 May 2023 / Revised: 27 September 2023 / Accepted: 29 October 2023 /

Published online: 13 November 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

## Abstract

In this research, an intelligent approach for the identification of Indian medicinal plant leaves is presented. Herein, heterogeneous features extracted from leaves of Indian medicinal herbs serve as input attributes for classifier models. A bilateral approach is employed to bring out heterogeneous features. In the first approach, a framework for extracting hand-crafted feature descriptors characterized by edge histograms, oriented gradients and binary patterns is proposed. In the second approach, deep features are extracted by Convolutional Neural Networks (CNN) using a transfer learning approach. Z score normalization is applied for the normalization of extracted features. Normalized features are fused using intermediate level fusion techniques to bring out heterogeneous features. ‘Uniform Manifold Approximation and Projection (UMAP)’ is employed to tackle the challenge offered by the high dimensional heterogeneous feature vector dataset. Finally, the reduced heterogeneous feature vectors are fed into different classifiers for the identification and classification of medicinal herbs. We compare the models’ performances in terms of accuracy, precision, sensitivity, Probability of False Alarm (PFA), Matthews Correlation Coefficient Score (MCC), Jaccard Score and Receiver Operating Characteristics Curve (ROC). The results reveal that the Random Forest classifier with hybrid feature vector attains accuracy, precision and sensitivity of more than 99% and attains a very low PFA of 0.02%. Hence the proposed models with heterogeneous feature framework not only improve the identification and classification performance but at the same time reduce the probability of false alarm to substantial amounts. Our model presents remarkable results and marks up to 03% performance improvement in contrast to other available models and it demonstrates significant improvement in the identification and classification accuracy of Indian medicinal plant leaves.

**Keywords** Medicinal plants · Machine learning · Handcrafted features · Deep learning features · Classifiers

## 1 Introduction

In the fast moving lifestyles due to the techno-savvy world, we are losing our connection with nature. We need to re-draw our focus on nature because we are part of it & can't afford to escape from nature. Natural products which include herbs are eco-friendly, and comparatively safe as most of the synthetic drugs are unsafe not only to humans but to the environment also. Traditionally medicinal plants like Aloe, Tulsi, Neem, Turmeric and Ginger are used in India & other parts of globe for several common ailments in the different seasons and these are safe and free from side effects too. The primary source of traditional Indian medicines are medicinal plants that give basic protection for human health. Traditional medicine offers the basic safeguards for human health which primarily draws its ingredients from medicinal herbs. Medical plant leaves are used in medicine, diet, fragrance, flavours etc. in India and other countries. According to the Botanical Survey of India, around 8,000 different herbal species are found in India, and around 95% of them are obtained from forest areas around India through destructive/ deforestation means in the region. In general, across the nation, there are 41% herbs (including grasses), 26% trees, 17% shrubs and 16% climbers. However, the extinction scenario poses a threat to the plant's genetic resources due to the growing population and the expansion of metropolitan areas [1]. India is a country with a history of healing systems using medicinal plants and a past study reveals that the use of medicinal plants has significant effects on traditional medicine resource preservation, authenticity, and identification instruction [2]. Medicinal plants are a vital source of herbal medicines, teas and other products. For several sectors, accurate plant identification is crucial. Over the past two decades, there has been a tremendous increase in the use of herbal medicine; however, there is still a significant lack of research data in this field. Hence, it's important to study medicinal plants, their identification & classification for better use. Present work is focussed on the identification of medicinal plants using their leaves. Researchers are increasingly using image processing techniques to identify plants from images of their leaves [3]. Machine learning and artificial intelligence can play a vital role in classifying and identifying medicinal plant leaves. The changing leaf colour and the differences in leaf form with ageing make it difficult to determine the type of plant. One of the most significant jobs in our ecology is played by plants. However, due to rapidly declining plant diversity, it was found that some of the plant species are on the verge of extinction and the problem needs our immediate attention and efforts for their conservation [4, 5]. It is necessary and yet valuable to develop a computerized automated system for various medicinal leaf species identification and classification to help the researchers and the public to identify them easily [6, 7]. On one hand to manage and stop the malpractices in the crude drug sector where medicinal leaves are vital ingredient in drug manufacturing, and on the other hand lack of expertise in discriminating medical plant leaves on their physical shapes, geometric features and venation architectures, there is an urgent requirement of automated and reliable recognition and stratification methodology required for the distinction of medicinal plant leaves. Hence, the classification of plant leaves is a very difficult and crucial problem to resolve. Herein, to address the challenge of identifying and classifying the medicinal plant leaves, a computer-aided model based on medicinal plant leaves hybrid features is proposed and medicinal plant leaves considered are given in Table S1. The distinguishing and identification capabilities of proposed models are evaluated on performance matrices such as Accuracy, Precision, Sensitivity, Probability of False Alarm (PFA), Matthews Correlation Coefficient (MCC) Score, Jaccard Score and Receiver Operating Characteristics (ROC) Curve. The structure

of the manuscript is sorted as: Section 2 critically reviews the work and the proposed methodology is explained in Section 3. Evaluation matrices and result analysis are presented in Sections 4 and 5 respectively. Sections 6 and 7 presents the discussion and conclusion respectively.

## 2 Related work: critical review

Several methods for the identification and classification of medicinal leaves are presented by various researchers. Gopal et al. [6] reported a system for medical plant leaf classification using various image-processing techniques. The authors evaluated the performance of their system by considering 100 leaf images for the training of the model and 50 leaf images to test the model. They extracted boundary-based, moment-based and colour-based features etc. from medical plant leaf images for further classification and their proposed model attained an accuracy of 92%. T. Sathwik et al. [4] developed a system for the identification and classification of medicinal plants based on extracted texture features (statistical, structural and modelling) from the leaf images using Gray Level Co-occurrence Matrix (GLCM). They explored amalgamations of texture features and obtained an accuracy of 94%. Malaysia-based tropical medicinal plants classification and identification framework using various angle features were proposed by Sainin et al. [8]. They employed genetic algorithms for feature selection and thereafter, used Direct Ensemble-Classifier for Imbalanced Multiclass Learning (DECIML) for the classification of medicinal plants. Their proposed model DECIMLFS. WIG achieved an accuracy of 85% with 181 selected features. Turkoglu et al. [9] proposed plant leaves recognition using local binary patterns. In addition, they evaluated the robustness of the proposed methodology against salt & pepper and Gaussian noise. With extracted features, the leaves were classified and tested using the Extreme Learning Machine (ELM) method and their proposed model obtained excellent classification results on Flavia, Swedish, ICL, and Foliage datasets. Prajwala Tm. et al. [10] proposed a method for tomato leaves disease detection using the CNN model named as LeNet. For their proposed model they considered 13,360 images for the training of the model and 4800 images for testing the model with 10-classes of tomato leaves and their proposed model achieved an accuracy of 94–95%. Another approach for classifying and detecting healthy and unhealthy leaves was developed by Yadav et al. [11]. Authors extracted deep features from plant leaf images using deep CNN (Convolutional Neural Network) followed by a feature selection method namely Particle Swarm Optimization (PSO). They used a dataset of plant leaves with 23-classes and achieved an accuracy of 97.39% for their approach.

Classification of medicinal plant leaves using Whale Optimization Algorithm (WOA) as an optimizer for feature selection and Random Forest (RF) as a classifier was presented by Pankaja et al. [12]. Swedish dataset and Flavia leaf datasets were used to test the performance of their proposed model. Suitable pre-processing techniques to remove noise and enhance the quality of leaf images were employed before the classification and identification of leaves. An accuracy of 97.58% was achieved with their proposed method. Tan et al. [13] proposed a CNN-based technique called D-Leaf was proposed for plant species classification. Three distinct pre-trained CNN models were used to extract the deep features. Five distinct machine-learning techniques were employed to classify plant leaves based on extracted deep features. They used a traditional morphometric approach to compute the morphological measures based on the Sobel segmented veins. The authors achieved

maximum accuracy of 95.54% using fine-tuned AlexNet. Naeem et al. [14] suggested a method for classifying medicinal plant leaves based on machine-learning approaches. The authors collected the 6-different types of medicinal plant leaves as a dataset for their experimental work. Sobel filter for intensity-based edge/line identification was employed on greyscale formatted captured images. The extracted texture-based, run-length matrix-based and multi-spectral-based features are used to prepare the feature vectors. Finally, they explored various classifiers for the classification of medicinal plant leaves based on extracted features and their model obtained excellent classification results.

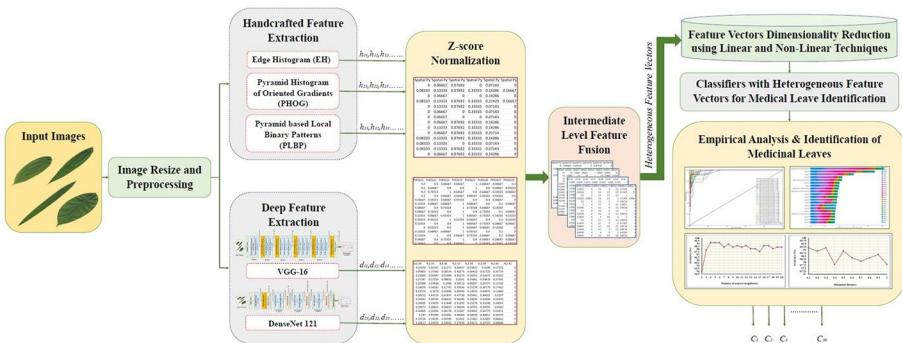
Dudi et al. [7] proposed a system for the classification of plant leaves using a deep-learning approach. For improving the image quality of leaves histogram equalization followed by median filtering was applied. The activation function was optimized using a hybrid Shark Smell-based Whale Optimization Algorithm (SS-WOA) to achieve maximum accuracy. Their proposed model achieved excellent classification and identification results. Shreyan et al. [3] developed a model named as BLeafNet for the classification of plant leaves. In the proposed methodology, the authors combined Bonferroni fusion with deep learning to achieve improved classification accuracy. They used Malayakew, Leafsnap and Flavia datasets to evaluate the performance of their proposed system. Their findings support the model's supremacy as their model showed better performance than many other state-of-the-art models.

Among the presented survey works, the majority of proposed models employed hand-crafted or deep features extracted from CNN networks were employed for medicinal leaves classification and identification. We can observe that an approach that attempts to fuse or amalgamate hand-crafted and deep features of medicinal plants is unnoticeable and hence can be employed for the identification and classification of Indian medicinal plant leaves. So, a novel methodology that considers the fusion of relevant information extracted by hand-crafted features and deep learning features can be employed for improving classification performance. Here, a model based on fused heterogeneous hybrid features of leaf is presented for identifying and classifying Indian medicinal plant leaves. The main contribution of the presented study is:

- i) Novel heterogeneous features generated by the intermediate-level fusion of hand-crafted and deep features for the classification of Indian medicinal plant leaves.
- ii) Employment of intermediate-level fusion techniques for medicinal plant classification makes our proposed methodology differ from available state of art models [13–16].
- iii) In addition to the classification performance analysis with the heterogeneous fused features, an in-depth comparative analysis of medicinal plant leaves classification with hand-crafted and deep features is also presented.
- iv) In addition to enhancing the classification performance, the proposed methodology with fused hybrid features also reduces the probability of false alarm to a substantial amount.
- v) The interpretability and explainability of ML “Black Boxes” are described to demonstrate the impact of fused feature variables on ML models.

### 3 Proposed methodology

The proposed methodology is focused on computer-aided automated classification and identification of medicinal plant leaves and is shown in Fig. 1. The methodology is divided into: - (a) Dataset description and pre-processing of medicinal leaves; (b)



**Fig. 1** An illustration of medicinal plant leaves classification model

extraction of handcrafted and deep features; (c) intermediate level feature fusion and dimensionality reduction and; (d) identification of medicinal herbs.

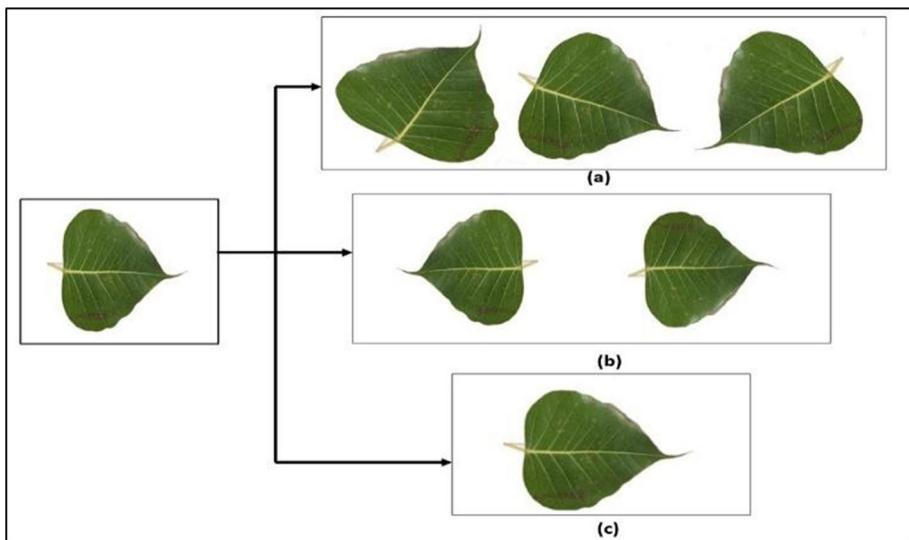
### 3.1 Data source and pre-processing

The dataset [17] undertaken in study consists of 1835 images of healthy Indian medicinal plant leaves collected from 30 species such as *Artocarpus Heterophyllus* (Jackfruit), *Azadirachta Indica* (Neem), *Ficus Religiosa* (Peepal Tree), *Ocimum Tenuiflorum* (Tulsi), *Trigonella Foenum-graecum* (Fenugreek), and much more. The medicinal leaves are collected from distinct locations of South Indian state Karnataka. There may exist inter and intra class similarities in leaves which is primarily due to color, illumination, shape or texture of the leaves. The medicinal plant leaves are picked from distinct plants of the same species. Approximately 60 to 100 images of every single species are collected using Samsung S9+ camera. Utter care has been taken to pluck healthy and mature leaves from distinct plants of identical species. Figure 2 shows the distinct species of medicinal plant leaves. 30 medicinal plant leaves under investigation with their pharmaceutical properties and respective number of samples considered are provided in Table S1.

The original images are resized to resolution of  $1600 \times 1200$ . After resizing of images, the image background is removed by segmentation technique using Python (OpenCV). The original image background is replaced with white colour. Haze noise removal technique was utilised to enhance the quality of images with parametric values: (i) amount of smoothing ( $\text{sigma\_s}$ ) = 25 and; (ii) value of dissimilar colours within neighbourhood region ( $\text{sigma\_r}$ ) = 0.25. As shown in Fig. 3, the images are rotated in both x and y directions ( $0^\circ$ - $360^\circ$ ); flipped horizontally and vertically and additionally, as the input layer size of VGG-16 and DenseNet-121 is  $224 \times 224 \times 3$ , the images are resized to  $224 \times 224$  before applying to CNN networks. The pre-processing plays a vital role in improving the training performances of the models by removing the noise from images (overfitting) and avoids under-fitting by image augmentation.



**Fig. 2** Distinct species of medicinal plant leaves



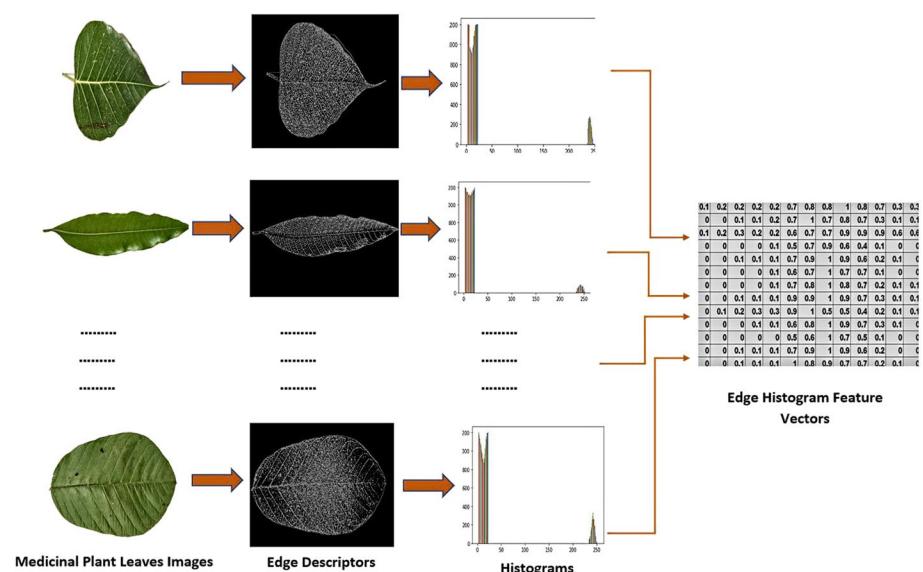
**Fig. 3** **a** rotation; **b** flipping and; **c** image resizing (224×224) of medicinal plant leaves during pre-processing stage

### 3.2 Feature extraction

This phase is applied to extract: (a) handcrafted features and; (b) deep features. The hand-crafted feature vectors are presented in terms of Edge Histogram (EH), Pyramid based Local Binary Patterns (PLBP) and Pyramid Histogram of Oriented Gradients (PHOG) descriptors. Whereas, pre-trained VGG-16 and DenseNet-121 CNN networks are employed for deep feature vector extraction.

**Handcrafted feature extraction approach** Herein, we applied EH, PHOG and PLBP for extracting handcrafted features. The handcrafted approaches are implemented by considering (i) the feature size and (ii) the classification accuracy [18]. The feature size is based on the feature extraction depending upon the internal changes in the parameters of the leaves. Classification accuracy is based on how much information is extracted which can be used for maximal differentiation with different classes and minimal differences within a class.

In the edge histogram feature extraction approach, the descriptors are expressed with unique features described in terms of brightness change repetition and directionality. Whereas in conventional edge detection technique the original images are represented by pixels that belong to the edges [19, 20] of the image. Edge in image is an important low-level feature. It can describe both shape and texture features, which are essential elements for content-based image analysis. Its importance makes the edge to be one of the most frequently used image features for the content-based image analysis, demanding a standardized means for its description. Herein, 80 edge histogram features are extracted to represent the original medicinal plant leaves. In the present work, the images are segmented into  $4 \times 4$  sub images to detect the edges and type of edges and its density. The edge histogram descriptors for each sub image are represented in terms of vertical local edges, horizontal local edges,  $45^\circ$  local edges,  $135^\circ$  local edges and non-directional local edges i.e., 5 local edge descriptors. The maximum value attained by five edge descriptors is compared with a defined threshold value to evaluate the dominant local edge descriptor. In our case a total of 80 histograms ( $5 \times 16 = 80$ ) values are normalized and quantized for sub images to represent the original medicinal plant leaves [19, 20]. The bin values are dependent on the frequency of local edges. Figure 4 represents edge histograms of the medicinal plant leaves with their histograms.

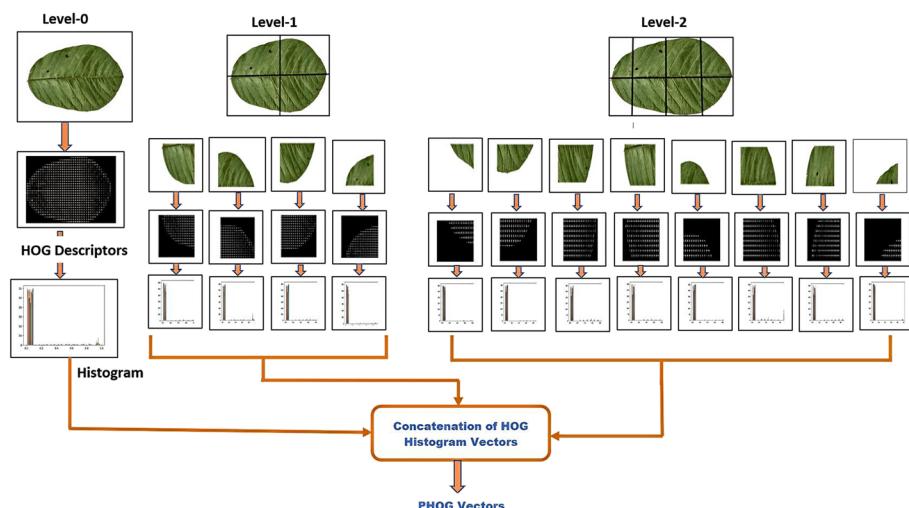


**Fig. 4** Representation of edge histogram feature descriptors

In the PHOG feature extraction approach, the descriptors are based on spatial pyramid matching [21]. The PHOG descriptor is capable of detecting the object at all positions and scales. The PHOG descriptor is an add-on to HOG descriptor in which the descriptors are evaluated in terms of orientation of edges. In addition, the HOG based description is invariant to noise and intensity changes in addition to the local position and scale changes. Because of these properties, the PHOG descriptor is capable of having invariant gradient information at various scales and locations. These characteristics make PHOG description as a strong candidate from various conventional, static, and handcrafted image feature extraction techniques. The PHOG descriptors exhibit more robustness against illumination, geometric variations and pose and find its applications in various domains, especially in pattern recognition [22]. Here, as shown in Fig. 5, to estimate the PHOG descriptors, the medicinal plant leaf image is divided into blocks considered as levels. The histograms of gradient for respective blocks are evaluated. Finally, concatenation of all the histograms is performed to attain PHOG descriptor vectors. As shown in Fig. 5, at level 0, the complete image is considered to be a single region of interest and the histograms of edge orientations are evaluated for the region.

For level 1, the entire medicinal plant leaf is segmented into four equal regions divided into height and width wise. Now, histograms of edge orientations are evaluated for five segments (one complete segment of level 0 and 4 segments of level 1). Similarly, for level 2, the entire image segmented into eight equal regions divided into height and width wise. Now histograms of edge orientations are evaluated for thirteen regions (eight segments at level 2 and five segments from previous levels). Hence the dimensions of feature descriptors can be increased by increasing the number of levels. In our case, we employed level 0, 1 & 2 and 30 bins of histograms to extract 630 PHOG descriptors from medicinal plant leave images (no. of PHOG descriptors = no. of bins of histohrams  $\times \sum_{l=0}^2 4^l = 630$ ; where  $l$  is the number of levels)

Lastly, PLBP is employed for extracting handcrafted features. Pyramid transform is an effective multi-resolution analysis approach. In contrast to Locally Binary Patterns (LBP) wherein the value of the threshold of a neighbouring pixel depends on the current pixel

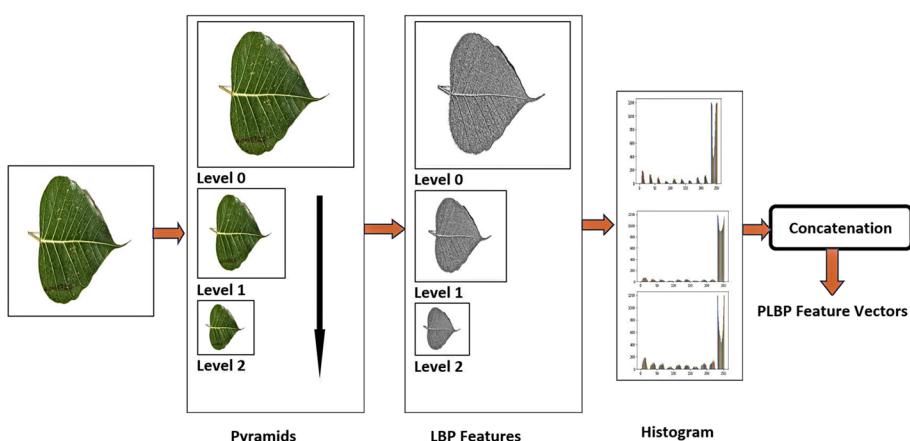


**Fig. 5** Representation of PHOG feature descriptors

value [23, 24], in PLBP the histogram is mapped on bins that were created for each uniform pattern and the variable patterns are allocated to a single bin. The PLBP represents LBP in the spatial pyramid domain and extracts significant feature vectors from images with markable texture variations. The lower-level spatial pyramid pixels can be achieved by down sampling the neighbours low-pass filtered high-resolution image for each beneath pyramidal levels. In our case, the medicinal plant leaves are down sampled by factor with three generated pyramid levels named as Level 0, Level 1 and Level 2. It is noted that with each pyramidal transformation the size of the image is half the size of the previous level image i.e., the size of  $k^{\text{th}}$  level image is half the size of  $(k-1)^{\text{th}}$  level image. In our case, as observed from Fig. 6, the size of image of level 2 is half the size of image at level 1 and so on. The LBP for the original sized ( $1600 \times 1200$ ) medicinal plant leaves images at level 0 are evaluated. As mentioned, the images at level 0 are high resolution images and hence they contain detailed features of plant leaves. At level 1 i.e., first pyramidal level, the images with size half of level 0 ( $800 \times 600$ ), the LBP patterns are generated. The level is implemented to deal with uniformity in texture rotation and same procedure is applied for level 2.

In the end, all the histograms generated by each level are concatenated to achieve the PLBP feature vector. In our case, a total of 756 PLBP descriptors are extracted from each medicinal plant leaf image.

**Transfer learning approach for deep feature extraction using pre-trained CNN** The CNNs are classified as subsets of deep neural networks and mainly employed for computer vision and pattern recognition applications. They are similar to convention neural networks and formulated using neurons, their respective weights, bias and activation functions. Typical structure of CNN comprises of input layers, convolutional layers, polling layers and dense layer. Images with size, where  $w$  is the dimesion of image and  $d$  represents the amount of depth/no. of channels of image (normally, its value is 1 for grey scale image and 3 for RGB image) are exposed to conventional layers followed by input layers. The convolutional layer employed learnable filters/kernels of size  $(w \times d)$ , to perform convolution operation and activates the image feature maps. The output received from convolutional layers are submitted to activation function. For effective mapping,

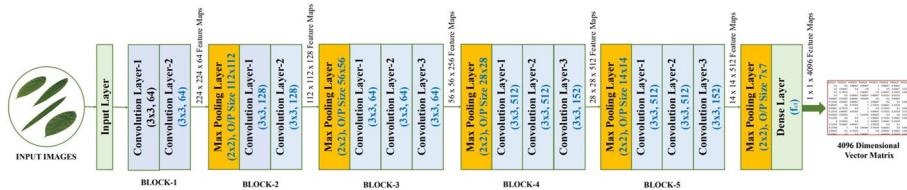


**Fig. 6** Representation of PLBP feature descriptors

rectified linear unit (ReLU) is adopted as activation function that maps the negative values to zero and retain the positive values and the activated features are forwarded to next layers. The pooling layer performs the function of down-sampling or image sub sampling and significantly lowers the learning parameters. Normally, activation, convolution and pooling operations are repeated across number of layers and every layer corresponds with distinct features. The fully connected dense layer which is “Multilayer Neural Network” is employed to give rise a “D” dimension vector, where “D” is the number of classes/target that the CNN can predict. The last layers (Softmax or Sigmoid layers) are engaged for binary or multi class classification task. The CNNs after supervised training can perform the task for classification or can be used as feature extractor using transfer learning approach.

Initially, these CNNs were trained with large image datasets and at later stage transfer learning approach [22] is employed to realize the characteristics acquired from large image datasets to similar classification assignment (with lesser number of images dataset) for progression of training process. Herein, we engaged the transfer learning approach due to unavailability large medicinal leaves database required to train the CNN model. Moreover, it is very difficult and time-consuming process to train the deep CNN models from scratch with large image database. So, by adjustments in the pre-trained CNN, with previously achieved weights and parametric values, the networks attain significant results towards the newly assigned task. The pre-trained CNN networks are mainly attracted as: (i) feature extractor and; (ii) fine-tunning networks on new task. The conventional architecture of pre-trained network, without amending their pre-defined weights, is utilized for feature extraction of images. The dense layers are substituted with classifiers and the features extracted by pre-trained network layers are fed as input to the classifiers for classification task. So, the pre-trained CNN models can be utilized as means to bring out the significant features from image dataset. Whereas, in case of fine tuning, the entire CNN model or subset of higher layers are fine-tuned with new dataset/task. Herein, the pre-trained CNN model weights will serve as initial weights and subsequently the weights are updated during the training process. The lower layers capture low level features such as texture, edges etc. that are comprehensive and the higher layers extracts target specific features from the images.

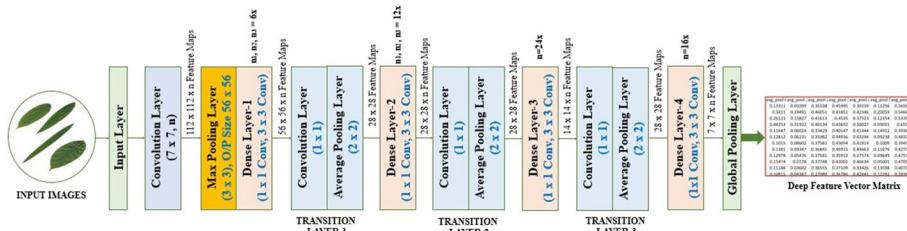
In the proposed work VGG-16 and Densenet-121 pre-trained CNN networks are implemented for deep feature extraction. The VGG16 feature extraction model has the capability of extracting a huge amount of data and results in good accuracy. It is one of the most popular techniques of image feature extraction, which performs better if any DL model is applied for classification tasks. Hence, the VGG16 model has been chosen for feature extraction in the proposed work. The use of max-pooling layers in VGG16 helps to reduce the spatial dimensionality of the feature maps, which can improve generalization performance and reduce overfitting. VGG-16 CNN network architecture [25] with  $3 \times 3$  convolution filters as feature extractor is employed and presented in Fig. 7. The resized ( $224 \times 244 \times 3$ ) images of medicinal plant leaves are brought to convolutional layer followed by input layers. The VGGNet-16 splits the feature extractor model into 5 segments/blocks consisting of multiple convolutional layer networks connected in series. Convolutional layers having  $3 \times 3$  convolutional filters are supplemented to extend the depth of the network. The convolutional layers denoted by  $a \times a \times r$  shows  $a \times a$  filter size and  $r$  number of channels. As shown in Fig. 7, for segment 1,



**Fig. 7** VGG-16 CNN Model as a feature extractor

$224 \times 224 \times 64$  low level feature maps are activated at the convolutional layer. Pooling layer (max) of stride  $2 \times 2$  are introduced for dimensionality reduction, and size of  $112 \times 112$  is obtained at first pooling layer. Correspondingly for each block, the deep feature maps excited by convolutional layers and dimensionality reduction is performed by pooling layers. An output of  $14 \times 14 \times 512$  excited feature map is obtained at convolutional layer of block five and the dimensions are further reduced to  $7 \times 7$  by pooling layer. The output of pooling layer is feed to dense layer (fc1) to extract further deep features. Finally,  $1 \times 1 \times 4096$  feature maps are extracted at output of fc1 layer.

In the second approach, the DenseNet-121 CNN network [26] was utilized for deep feature extraction. DenseNet has achieved excellent performance, while also utilizing less memory and processing power than other state-of-the-art techniques. DenseNet strengthens the features and gradients of each layer by using the top classifier to supervise all layers through feature connection. Figure 8 shows DenseNet CNN with convolutional layers, pooling layers and dense blocks. The feed-forward approach is used in DenseNet CNN to connect the layers. The dense layer mainly performs two operations:  $1 \times 1$  Conv (convolutional operation) for fetching the features and  $3 \times 3$  Conv (convolutional operation) for reducing the feature depth/number of channels. Concatenation of features is normally employed in DenseNet architecture. The variables  $n_1$ ,  $n_2$  and  $n_3$  shown in Fig. 8 characterize the repetitions of deep blocks for different architecture of DenseNet CNN networks. As illustrated, the resized ( $224 \times 244 \times 3$ ) images of medicinal plant leaves are feed to  $7 \times 7$  convolutional layer. A deep feature map of size  $112 \times 112 \times n$  is obtained at output of convolutional layer. The pooling layer followed by convolutional layer are used for dimensionality reduction. A reduced output is obtained at pooling layer (max) with  $56 \times 56$  size. Dense blocks are employed to further extract the deep features from images. Transition layers are sandwiched between two dense layers to further scale down the number of channels and down sample the features. The transition layer consists of  $1 \times 1$  convolution operation to lower the channel count to half and



**Fig. 8** DenseNet CNN as feature extractor

$2 \times 2$  pooling layer (average) is implemented to further down sample the feature size in terms of width and height. The output of the Dense Block-4 having  $7 \times 7 \times n$ -dimensional feature map array is feed to global pooling layer. This array is converted to a  $1 \times 1 \times n$ -dimensional vector by using a global polling layer (max-pool layer) feature vectors are extracted from pooling layer. Finally,  $1 \times 1 \times 1024$  feature maps are extracted at output layer.

Herein, the weight parameters of VGG-16 and DenseNet-121 CNN network pre-trained on ImageNet's dataset were used to bring down execution time and to attain improved training results. The medicinal plant leave dataset splits to 8:2 ratio for training and testing of the model. For employing VGG-16 as feature selector, the model is trained in batch size of 28, learning rate of  $1e^{-4}$  with “ReLU” as activation function and “Admas” as optimizer. The model is tuned in 60 epochs.

### 3.3 Feature fusion and dimensionality reduction

Feature fusion can impart a vital role in enhancing the performance of the classifiers. In this, two or more incongruent features are amalgamated into a single feature vector. For feature fusion, typically three strategies: low-level, middle level or intermediate level and high-level fusion [27] are employed. In low-level or data fusion, diverse data sources are amalgamated to bring out new raw data and are assumed to contain [Supplementary Information](#). In middle-level or intermediate-level fusion, various features are fused to obtain prominent feature information from all. In high-level fusion, different classifier outputs are combined to deliver improved results. Amongst all, middle-level or intermediate-level feature fusion is considered to be much more efficient because the fused feature vector comprises added information in contrast to low-level or decision-level fusion [28].

In the present study, an intermediate-level fusion of handcrafted and deep features to obtain a heterogeneous feature dataset is employed. The feature fusion technique can be implemented using methods such as pooling, gate units or concatenation [27]. In our case, the feature fusion is performed by concatenating feature vectors extracted from handcrafted and deep CNN pre-trained networks into one large feature vector. However, two primary issues: feature compatibility or feature scaling and; b) high dimensionality are encountered in the task of feature fusion. For feature compatibility, we employed z-score normalization to transform the features. The z-score normalization normalizes each value in the dataset such that the mean value of all the values is zero and the standard deviation is 1. During z-score normalization, the distribution shape remains the same and the z-score has the advantage of robustness against outliers. The z-score normalization is taking the advantage of considering both the mean value and variability in raw scores. The z-score can be given as

$$z = \frac{x - \mu}{\sigma} \quad (1)$$

where  $z$  is the z-score;  $x$  is the original feature vector;  $\mu$  is the mean and  $\sigma$  is the standard deviation.

The second primary issue is the dimensionality reduction of the fused feature vector dataset. The dimensionality reduction techniques can be commonly classified as linear and non-linear dimensionality reduction techniques [29–31]. The linear dimensionality

reduction techniques establish a linear amalgamation of the input attributes/features to construct new variables. Linear Discriminant Analysis (LDA), PCA, Factor Analysis (FA) and time-structure Independent Component Analysis (tICA) are a few linear dimensionality reduction techniques. Whereas, the non-linear dimensionality reduction techniques construct new features/attributes by mapping the input attributes to nonlinear functions. t-SNE, Isometric Mapping (Isomap) and UMAP are a few such nonlinear dimensionality reduction techniques. The nonlinear dimensionality reduction techniques have the advantage of better analysis and visualization of complex data [32]. Herein, we employed UMAP, developed by McInnes et al. [33] to encounter the high dimensionality issue. UMAP is based on a fuzzy topology to bring down the dimensionality and establish probability distribution in a high-dimension manifold. The UMAP illustrates data points by high dimensional weighted graphs to formulate a fuzzy topological structure. The edge weight represents the likelihood of connecting two points. Dimensionality reduction using UMAP can be evaluated as follows:

- i) for a given dataset  $X = \{x_n\}, n = 1, 2, 3, \dots, N$  with data points in high dimensional space  $S'$ , evaluate the sets  $v_n$  comprising of  $k$  neighbor points for every data point in the dataset  $x_n$ .
- ii) for the respective  $n^{th}$  data point, evaluate the nearest neighbor and the distance as

$$\rho_n = \min(d(x_n, x_q) \mid x_q \in v_n, d(x_n, x_q) > 0) \quad (2)$$

and the value of  $\sigma_n$  such that

$$\sum_{x_q \in v_n} \exp\left(\frac{-\max(0, d(x_n, x_q) - \rho_n)}{\sigma_n}\right) = \log_2 k \quad (3)$$

where,  $d(x_n, x_q)$  is the distance between  $n^{th}$  and  $q^{th}$  data point and  $\rho$  is distance between  $n^{th}$  data point and its first nearest neighbor.

- iii) using adjacency matrix establish UMAP graph as unidirectional weighted graph

$$B = A + A^T - A^T A \quad (4)$$

and the elements of A are stated by weights in the corresponding directed graph:

$$w(x_n, x_q) = \exp\left(\frac{-\max(0, d(x_n, x_q) - \rho_n)}{\sigma_n}\right) \quad (5)$$

- iv) the elements  $y_n, n = 1, 2, 3, \dots, N$  in low dimensional space are evaluated by the force-directed placement of the graph using attractive and repulsive forces  $F^a$  and  $F^r$  respectively between vertices  $n$  and  $q$  can be given as

$$F_{n,q}^a = \frac{-2bc \|y_n - y_q\|_2^{2(c-1)}}{1 + \|y_n - y_q\|_2^2} w(x_n, x_q) (y_n - y_q) \quad (6)$$

$$F_{n,q}^r = \frac{b}{(\mu + \|y_n - y_q\|_2^2)(1 + \|y_n - y_q\|_2^2)} (1 - w(x_n, x_q)) (y_n - y_q) \quad (7)$$

where  $b, c$  and  $\mu$  are constants.

Binary cross entropy is used as a cost function in UMAP and its derivative is employed for updating the coordinates of the low dimensional data points. The coordinates are updated to optimize the projection space till convergence. The parameters number of neighbours and minimum distance are fine-tuned to attain their optimal values. The accuracy of local and global structures of data is balanced by the parameter number of nearest neighbours and the minimum distance between data can be optimized with the parameter minimum distance [33].

## 4 Performance evaluation matrices and experimental setup

### 4.1 Performance evaluation matrices

Table 1 presents the performance evaluation matrices for our proposed model. The matrices are evaluated using TP, FP, TN and FN.

Where, TP, FP, TN and FN have their usual meanings defined in the confusion matrix [34]. The term *TNR* (True Negative Rate) defines the specificity of the classifier. Additionally, ROC-AUC curves are plotted (between TPR and FPR) to further estimate the classifier performance. The curves at the top left signify good classification, whereas, the lower right corner displays poor classification. The area under the curve is evaluated and a high score of AUC ( $\approx 1$ ) represents outstanding classification results.

### 4.2 Experimental setup

The training and testing of medicinal plant leaves identification and classification models were simulated in Google Colaboratory. For all the experiments, the data source is arbitrarily split into training and test datasets. The training data source (80%) is employed for building and training of the model and the test dataset (20%) is engaged for testing the model. Three classifiers namely: Random Forest (RF), Extreme Gradient Boosting (XGBoost) and Multi-Layer Perceptron Neural Network (MLP-NN) are employed for classification. The simulation parameters of classifiers, CNN feature extraction models and dimensionality reducers are presented in Table 2.

## 5 Results analysis

Detailed empirical analyses are presented in this section. Section 5.1 presents the empirical analysis of models with handcrafted features; Section 5.2 presents the analysis with deep features, and Section 5.3 presents the empirical analysis with heterogeneous fused features.

### 5.1 Medicinal plant identification utilizing handcrafted features

For the present study, EH, PLBP and PHOG algorithms are employed for extracting the edge, appearance, shape and texture relied on features of medicinal plant leaves. Table 3 presents the empirical analysis results based on handcrafted features.

It is observed that a very low identification performance with edge histogram features is derived by the classifiers. It is noticeable that the PHOG extracts a lesser number of

**Table 1** Performance measurement matrices for classification model of Indian medicinal plant leaves

Parameter	Mathematical representation	Significance
Accuracy	$\frac{TP+TN}{TP+FP+TN+FN}$	Provides correct predicted events among all the events examined
Precision	$\frac{TP}{TP+FP}$	Also defined as Positive Predicted Values (PPV) and defines the proportion of positive predictions that are actually positive.
Sensitivity	$\frac{TP}{TP+FN}$	The ability of the classifier to correctly detect the positive events from total positive events. However, it does not provide information about the events that are wrongly classified as positive from other classes of events.
Probability of false alarm	$1 - TNR$	Presents the probability of retrieving irrelevant events amongst all irrelevant events. It is evaluated as the ratio of negative events wrongly classified as positive events and total negative events.
MCC	$\frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$	Evaluates the correlation coefficient (Pearson product-moment) between the true value and forecasted value.
Jaccard Score	$\frac{TP}{TP+FP+FN}$	Commonly known as the Jaccard similarity coefficient, and is a statistical tool to find the similarity and dissimilarity of test and predicted datasets. A close value of 1 indicates the highest similarity between the two datasets.

**Table 2** Simulation parameters

Simulation parameters	
RF	criterion for quality of split = Gini, number of features for splitting = square root of no. of features, stopping threshold = 0, number of samples required to be leaf node = 1, node splitting samples = 2, trees in forest = 100
XGBoost	Criteria used for splitting = 'gain', interaction constraints = none, learning rate = none, maximum depth of tree = none, n of estimators = 100.
MLP-NN	Activation function = ReLU, number of neurons in i <sup>th</sup> hidden layer = 100, weight optimization solver = Adams, alpha = 0.0001, number of iterations = 200, learning rate = constant, beta-1 for early stopping = 0.9, beta-2 for early stopping = 0.999, number of epochs = 10
CNN Models	learning rate = 1e-4; epochs = 60; optimizer = Adams; batch size = 28, activation function = ReLU
UMAP	nearest number of neighbours = 5, minimum distance = 0.3, number of components = 3
t-SNE	Dimensional space = 2, perplexity = 30, learning rate = 200, total iterations employed for optimization = 1000

features in contrast to PLBP still, it provides higher classification and identification accuracy for Indian medicinal plant leaves. With PHOG features, using spatial layout and local shape seized by edge orientations and tilting the medicinal leaves images at several resolutions, improved accuracy is achieved. As observed, the performance is improved with the amalgamation of handcrafted features. Maximum classification performed is attained by the MLP-NN classifier with amalgamated handcrafted features. An accuracy and

**Table 3** Empirical analysis with handcrafted features

Handcrafted features	Extracted features	Classifier	Accuracy (%)	Precision (%)	Sensitivity (%)	PFA (%)
EH	80	RF	69.20	75.54	69.20	1.04
		XGB	69.48	71.65	69.48	1.05
		MLP-NN	70.57	72.32	70.57	1.00
PLBP	756	RF	77.11	78.74	77.11	0.85
		XGB	77.65	79.07	77.65	0.77
		MLP-NN	82.83	84.07	82.83	0.65
PHOG	630	RF	80.38	83.19	80.38	0.64
		XGB	77.38	79.12	77.38	0.78
		MLP-NN	81.47	82.83	81.47	0.65
EH + PLBP	836	RF	82.88	83.99	82.88	0.61
		XGB	82.56	83.09	82.56	0.60
		MLP-NN	89.10	89.61	89.10	0.37
EH + PHOG	710	RF	83.65	84.71	83.65	0.56
		XGB	82.28	83.71	82.28	0.61
		MLP-NN	89.10	90.33	89.10	0.37
PLBP + PHOG	1386	RF	84.74	86.36	84.74	0.52
		XGB	85.28	86.05	85.28	0.50
		MLP-NN	88.82	90.38	88.82	0.38
EH + PLBP + PHOG	1466	RF	87.19	88.26	87.19	0.44
		XGB	85.28	85.51	85.28	0.50
		MLP-NN	90.46	91.47	90.46	0.32

sensitivity of 90.46%, and precision of 91.47% with a low percentage (0.32%) of false alarm is obtained with amalgamated handcrafted features. A low value of the probability of false alarm signifies a low probability of retrieving irrelevant events amongst all irrelevant events.

## 5.2 Medicinal plant identification utilizing CNN extracted features

VGG-16 and DenseNet-121 pre-trained convolutional neural networks have been utilized for feature extraction from medicinal plant leaves and the model's performance with deep features is tabulated in Table 4.

It is noted that with VGG-16 and DenseNet-121 extracted features, an accuracy of more than 90% has been achieved by the classifiers. In contrast to the DenseNet-121 feature extractor, the VGG-16 extracts a four-fold number of deep features. The MLP-NN Classifiers with VGG-16 extracted features attained maximum accuracy & sensitivity of 93.73% and precision of 94.26%. An improvement in classification performance has been observed with deep features extracted by the DenseNet-121 CNN network. An enhancement of almost 5% in classification accuracy has been observed with DenseNet-121 extracted features and the MLP-NN classifier attains maximum accuracy & sensitivity of 98.63% and precision of 98.69%. A significant reduction of 0.18% in the probability of false alarm was also noticed with MLP-NN classifiers utilizing DenseNet-121 extracted features. Though with a lesser number of dense layers, the DenseNet-121 extracts significant features and demonstrates excellent results as compared to VGG-16.

## 5.3 Medicinal plant identification utilizing fused heterogeneous features

The fused heterogeneous dataset is proposed to complement the feature set vector. The model's performance analysis with fused heterogeneous features is presented in Table 5. With a heterogeneous feature dataset, the RF classifier shows significant performance enhancement having a low probability of false alarm. Approximately the same performance has been marked by the models with CNN + PLBP + PHOG fused features and CNN + EH + PLBP + PHOG fused features. Hence, revealing that EH extracts descriptors hold the least significant information. The classifiers with DN-121 + PLBP + PHOG fused heterogeneous features (2410 features) and DN-121 + EH + PLBP + PHOG fused features (2490 features) deliver the same classification performance. However, augmenting the EH feature does not mark a significant improvement in classification performance but a decrease in the probability of false alarm has been noted signifying that EH descriptors

**Table 4** Empirical analysis with CNN extracted feature

	Extracted features	Classifier	Accuracy (%)	Precision (%)	Sensitivity (%)	PFA (%)
VGG-16	4096	RF	93.46	94.11	93.46	0.19
		XGB	90.73	91.04	90.73	0.32
		MLP-NN	93.73	94.26	93.73	0.22
DenseNet-121	1024	RF	97.82	97.92	97.82	0.08
		XGB	93.46	94.10	93.46	0.22
		MLP-NN	98.63	98.69	98.63	0.04

**Table 5** Empirical analysis with fused heterogeneous features

	Features	Classifier	Accuracy (%)	Precision (%)	Sensitivity (%)	PFA (%)
VGG-16						
VGG-16 + PLBP + PHOG	5482	RF	97.00	97.26	97.00	0.10
		XGB	93.18	93.49	93.18	0.23
		MLP-NN	98.09	98.24	98.09	0.06
VGG-16 + EH + PLBP + PHOG	5562	RF	96.18	96.57	96.18	0.13
		XGB	93.18	93.58	93.58	0.23
		MLP-NN	97.54	97.71	97.54	0.08
DenseNet-121 (DN-121)						
DN-121 + PLBP + PHOG	2410	RF	98.09	98.26	98.09	0.06
		XGB	94.27	94.76	94.27	0.19
		MLP-NN	97.82	97.96	97.82	0.07
DN-121 + EH + PLBP + PHOG	2490	RF	98.91	98.97	98.91	0.03
		XGB	94.55	94.04	94.55	0.18
		MLP-NN	97.28	97.17	97.28	0.07

aid in decreasing the occurrence of false events. The Random Forest classifier with DN-121 + EH + PLBP + PHOG fused heterogeneous features attains maximum classification performance with 98.91% accuracy and sensitivity, 98.97% precision and 0.03% probability of false alarm.

It is needed to be noted here that in Table 5 the heterogeneous fused dataset suffers from high dimensionality problems. Usually, feature selection or feature extraction is employed for dimensionality reduction [35]. We prefer feature extraction in contrast to feature selection as in feature selection information can be lost during the process of irrelevant feature elimination [35]. We employed PCA as a linear dimensionality reducer and; UMAP and t-SNE as non-linear dimensionality reducers. Several experiments were techniques to observe the classification performance utilizing heterogeneous dataset and Table 6 presents the comparative analysis of RF and classification performance with UMAP, Principal Component Analysis (PCA) and t-distributed Stochastic Neighbor Embedding (t-SNE) dimensionality reduction techniques. Amongst presented dimensionality reduction techniques, the classifier shows the least classification performance with the PCA-reduced fused dataset. For the present study, the PCA dimensionality reduction technique marks its best performance on the handcrafted and DenseNet-121 fused dataset with an accuracy and sensitivity of 97.54%, precision of 97.68% and PFA of 0.08%. An improvement in classification performance with t-SNE reduced fused feature dataset has been noted in comparison to PCA. The RF classifier with handcrafted and DenseNet-121 fused feature and dimensionality reduced with t-SNE technique attains accuracy, precision and sensitivity of greater than 98% and a probability of false alarm of 0.04%. However, with UMAP reduction technique the classifier attains classification performance results with more than 99% accuracy, precision and sensitivity and reduces the probability of false alarm up to 0.02% in equivalence to t-SNE and 0.06% in comparison to PCA.

The prediction results for misidentified medicinal plant leaves by RF classifier with heterogeneous leaves features for the test dataset are presented in Table 7. Five medicinal plant species are misidentified and the remaining twenty-five medicinal leaves are 100% correctly identified during the testing phase of the model. The misclassified medicinal plant species are: *Ficus Auriculata* (Roxburgh fig), *Moringa Oleifera* (Drumstick), *Santalum*

**Table 6** Performance metrics of RF classifier with dimensionality reduced fused feature set (D1: VGG-16+PLBP+PHOG; D2: VGG-16+EH+PLBP+PHOG; D3: DN-121+PLBP+PHOG; D4: DN-121+EH+PLBP+PHOG)

Accuracy (%)			Precision (%)			Sensitivity (%)			PFA (%)			
UMAP	PCA	t-SNE	UMAP	PCA	t-SNE	UMAP	PCA	t-SNE	UMAP	PCA	t-SNE	
<b>VGG-16</b>												
D1	97.00	95.91	97.54	97.36	96.39	97.82	97.00	95.91	97.54	0.10	0.14	0.08
D2	96.45	96.18	97.00	96.69	96.54	97.27	96.45	96.18	97.00	0.12	0.13	0.10
DenseNet-121 (DN-121)												
D3	98.36	97.54	98.63	98.47	97.73	98.75	98.36	97.54	98.63	0.05	0.08	0.04
D4	99.18	97.54	98.63	99.20	97.68	98.72	99.18	97.54	98.63	0.02	0.08	0.04

*Album* (Sandalwood), *Syzygium Jambos* (Rose Apple) and *Tabernaemontana Divaricata* (Crape Jasmine). From test results it was found that 80% of images of *Ficus Auriculata* (Roxburgh fig) are correctly classified whereas 20% images of *Ficus Auriculata* (Roxburgh fig) are misidentified as *Basella Alba* (Basale) (10%) and *Jasminum* (Jasmine) (10%). The *Moringa Oleifera* (Drumstick) medicinal plant leaves are 95% correctly identified and 5% are misclassified as *Carissa Carandas* (Karanda). The *Santalum Album* (Sandalwood) species are 89% correctly identified by the model and 11% of the species are misidentified as *Murraya Koenigii* (Curry). Another medicinal plant species *Syzygium Jambos* (Rose Apple) are 91% correctly categorized and 9% of them are misclassified as *Nerium Oleander* (Oleander). Similarly, the *Tabernaemontana Divaricata* (Crape Jasmine) species are correctly identified upto 83% and 17% are misclassified as *Moringa Oleifera* (Drumstick). The misidentifications are mainly observed due to their similar structure, texture, color and shape of their edges.

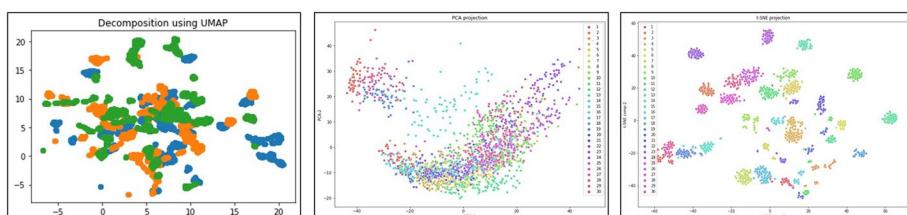
**Table 7** Analysis of misidentified medicinal leaves during testing phase

Medicinal Plant leave Species	Correctly Identified (%)	Misidentified (%)	Misidentified Medicinal Plant Species
<i>Ficus Auriculata</i> (Roxburgh fig)	80 %	20 %	<i>Basella Alba</i> (Basale) (10%) <i>Jasminum</i> (Jasmine) (10%)
<i>Moringa Oleifera</i> (Drumstick)	95%	05%	<i>Carissa Carandas</i> (Karanda)
<i>Santalum Album</i> (Sandalwood)	89%	11%	<i>Murraya Koenigii</i> (Curry)
<i>Syzygium Jambos</i> (Rose Apple)	91%	09%	<i>Nerium Oleander</i> (Oleander)
<i>Tabernaemontana Divaricata</i> (Crape Jasmine)	83%	17%	<i>Moringa Oleifera</i> (Drumstick)

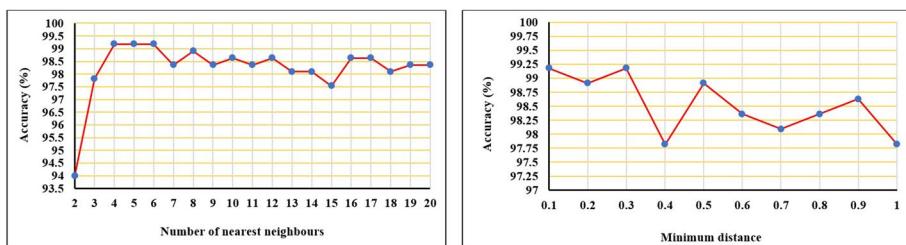
Figure 9 presents the decomposition of the fused heterogeneous dataset using UMAP, PCA and t-SNE dimensionality reduction technique. Figure 9b presents PCA decomposition with eigenvectors with their eigenvalues. Herein, the Eigenvectors with the lowest eigenvalues are dropped out as they support minimal information about the distribution of the data. In the PCA decomposition plot, the classes are represented by different colours and it is noted that the classes are not well separated from each other and marked variance from  $-40$  to  $40$  for component1 and  $-20$  to  $40$  for component 2 and hence showed inferior classification results with reduced dimensionality dataset as compared to t-SNE and UMAP. In t-SNE student-t distribution is employed for computing similarity between two points in lower dimensional space. As seen from Fig. 9c, in contrast to PCA, the t-SNE decomposition has well-separated classes after decomposition and t-SNE component 1 (x-axis) marked variance from  $-60$  to  $60$  and  $-40$  to  $60$  for t-SNE component 1 (y-axis) and hence shows better classification results. Figure 9a presents the UMAP dimensionality-reduced dataset. The UMAP has the advantage that it can maintain a more global structure than t-SNE. Herein, the UMAP reduced dataset provides excellent classification results in comparison to t-SNE and PCA as the clusters of classes in UMAP are well defined and show a variance from  $-5$  to  $20$  for both the components (x- and y-axis).

Variation in the classifier's performance in terms of accuracy with UMAP tuning parameters is presented in Fig. 10. As observed maximum accuracy has been realized with 4–6 number of nearest neighbours along with a minimum distance of 0.1 or 0.3. Except for 2 nearest neighbours, an accuracy of greater than 97.5% is achieved by the classifier with another number of nearest neighbours.

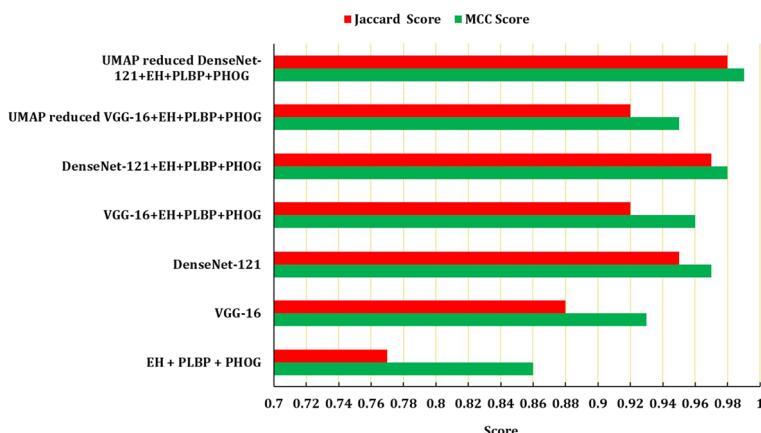
The execution of proposed classification models with fused heterogeneous features is further validated in terms of MCC and Jaccard Score and presented in Fig. 11. The classifier with fused handcrafted features demonstrates the least MCC and Jaccard scores as 0.86 and 0.77 respectively. With CNN extracted features an improvement in terms of MCC and Jaccard score is noted. In contrast to fused handcrafted features, an improvement of 0.11 and 0.15 has been observed in MCC and Jaccard scores respectively in classifiers with DN-121 extracted features. With the proposed hybrid handcrafted and DenseNet-121 extracted fused feature dataset, an appreciative performance with MCC and Jaccard score of greater than 0.97 has been achieved. Maximal MCC and Jaccard score of 0.99 and 0.98 respectively has been attained by the classifier with UMAP-reduced handcrafted and DN-121 extracted fuse dataset.



**Fig. 9** Decomposition of the fused heterogeneous dataset using (a) UMAP; b PCA and; c t-SNE



**Fig. 10** Accuracy v/s UMAP tuning parameter plots



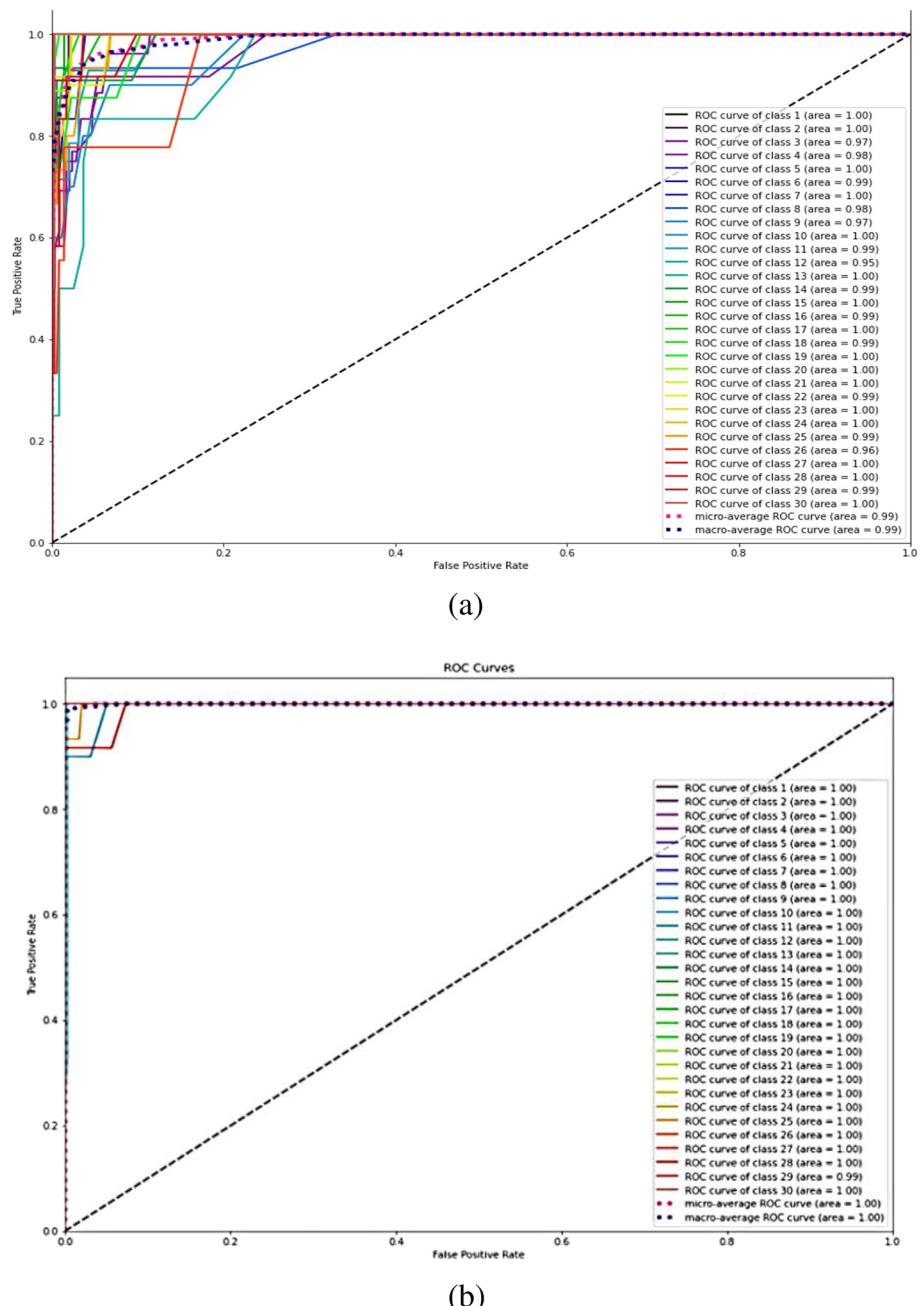
**Fig. 11** Performance matrices in terms of MCC and Jaccard Score

## 6 Discussion

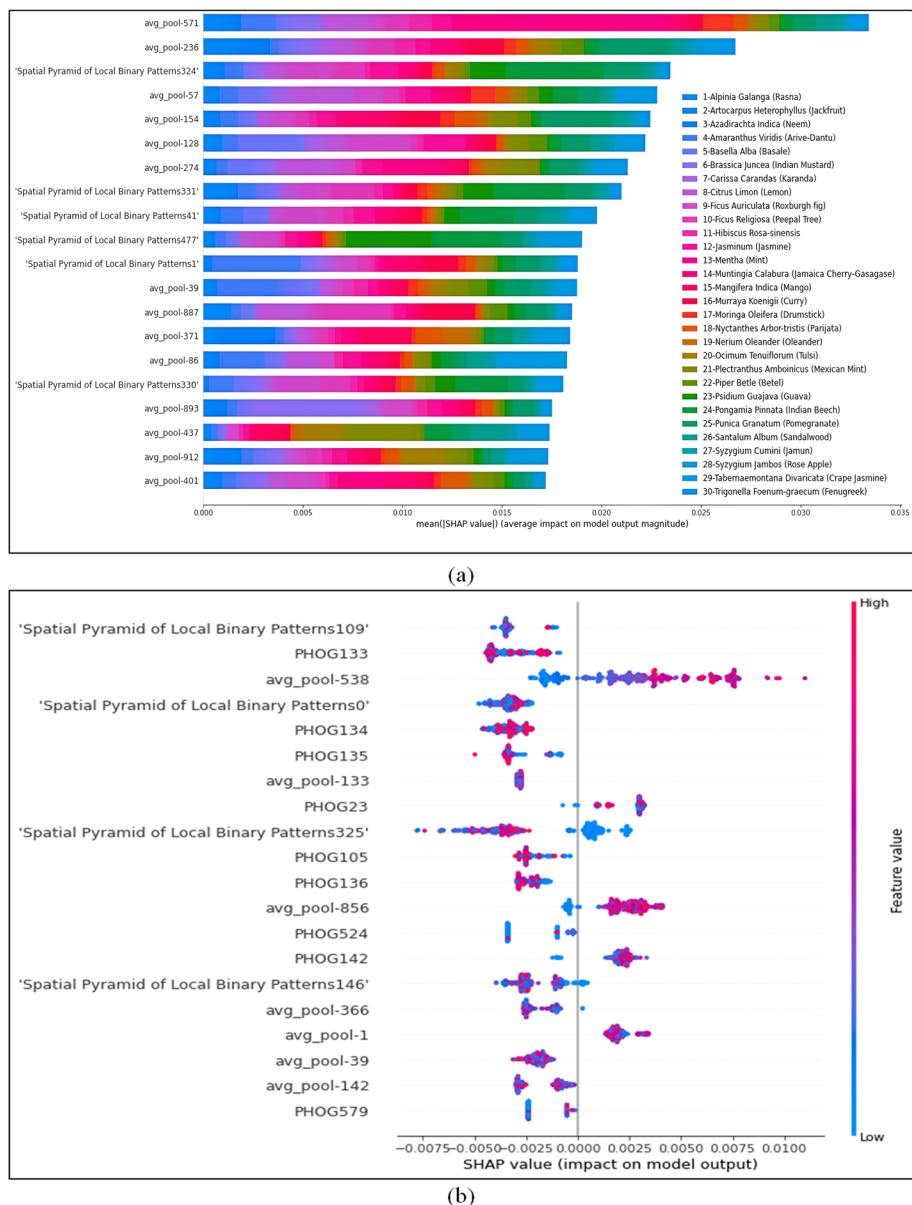
Here's an idea for Indian medicinal plant leaf classification based on fused heterogeneous leaf features that were investigated and presented. The classifiers with handcrafted features evidence unsatisfactory classification performance. However, the performance improvement was noted with the amalgamation of handcrafted features. With amalgamated handcrafted features an accuracy and sensitivity of 90.46%, precision of 91.47% and 0.32% of false alarm were achieved by MLP-NN classifier. With DN-121 extracted features, the MLP-NN classifier marked an improved performance with more than 98% accuracy, precision and sensitivity with significant limiting in the probability of false alarm. The classification performance was further enhanced with an intermediate-level fusion of handcrafted and CNN-extracted deep features. The evaluation matrices (at the single operating point) presented in Table 1 are too specific and reflect the results for single or average test instances and the matrices do not deliver performance information where the performance may change rapidly [36, 37]. Herein, in addition to these single point quantification matrices, ROC-AUC curves are demonstrated as the analysis derived from ROC-AUC is not influenced by decision criteria [38, 39]. The ROC plots the True Positive Rate (TPR) against the False Positive Rate (FPR) for several threshold values and the area underneath signifies the competency of the classifier to distinguish various classes. The

higher the ROC-AUC score, the better the classifier's performance in distinguishing the classes. For an ideal classifier, the graph would score 100% (value of 1) for TPR (Y-axis) with 0% (value of 0) for FPR (X-axis) and a value of AUC as 1 for the excellent classification model and 0 for the poor classification model. For the proposed work, the classification competency of the RF classifier with handcrafted and fused features in terms of ROC-AUC performance description is presented in Fig. 12. Figure 12a-b shows the ROC-AUC curve of the RF classifier with hand-crafted and hybrid fused features. As noted from Fig. 12 (a)-(b), the classifier attains an average ROC curve area greater than or equal to 0.99 but the classifier attains dissimilar values for TRP and FPR. As stated, for excellent classification results the ROC-AUC curve should have a flat upper left corner indicating the value of 1 at the y-axis and 0 at the x-axis. Figure 12a-b, shows substantial variation in increase and decrease in TPR and FPR values. From Fig. 12a, the classifier with handcrafted features exhibits approximately TPR values of 0.03 to 1 and FPR values between 0 and 0.38 for more than 20 different classes of medicinal plant leaves and a diminished flat left top corner of the ROC-AUC curve with an average ROC curve area value of 0.99. However, as noted in Fig. 12b, a significant improvement in values of TPR and FPR has been noted with fused features. The classifier exhibits the value of 1 and 0 for TPR and FPR for maximum classes of medicinal plant leaves and marginally demonstrates approximately value of 0.9 to 1 as TPR and 0.08 to 0 as FPR values for three classes of medicinal plant leaves and layout a considerable flat left top corner of ROC-AUC curve with an average ROC curve area value of 1 and validates excellent classification results.

From results, it confers that machine learning models present remarkable medicinal plant leaves classification results; however, the models struggle with the issue of interpretability and explainability [40, 41] and offer a demanding task to understand and interpret the results presented by these “Black Boxes” [42]. Many Researchers presented methods for the interpretability of these “Black Boxes” and hence invoked the conceptualization of eXplainable AI (XAI) to make the model more understandable. Herein, the “SHapley Additive explanation (SHAP)” proposed by Lundberg and Lee [43] has been implemented for the interpretability and explainability of the proposed model. The SHAP framework analyses the influence of independent feature variables on the predicted output. The impact is interpreted using SHAP values that account on: - local accuracy; missingness and; consistency [43]. Figure 13a illustrates the influence of fused heterogeneous feature variables on the prediction of the proposed model in terms of SHAP values. The heterogeneous feature variables are represented along the Y-axis and their impact on prediction performance (defined in terms of mean SHAP value) is represented along the x-axis. For instance, the fused features “avg\_pool571” and “avg\_pool401” exhibit the highest and lowest impact on the prediction performance among the presented feature variables. Figure 13b presents the impact of feature variables on prediction performance for a typical class of medicinal plant leaves. Herein, the y-axis represents the feature variables and their impact (defined in terms of SHAP values) along the x-axis. Positive and negative Shaply values show the favourable and unfavourable influence of feature variables represented along the x-axis. Feature significance from high to low is pronounced with a change in colour from red to blue. For instance, the feature variable “avg\_pool538” shows both positive and negative impacts on the prediction performance for the typical class of medical plant leaves. The feature indicates a low negative impact and a strong positive impact on the classification performance. It is also observed that the features: Spatial Pyramid Local Binary Pattern 109, PHOG133, Spatial Pyramid Local Binary Pattern 0, PHOG134, PHOG135, avg\_pool 133, avg\_pool39, avg\_pool142 and PHOG 79 marks negative impact on prediction performance i.e., with increasing feature value the predicted values decrease.



**Fig. 12** **a** ROC-AUC Curve of RF Classifier with handcrafted features; **b** ROC-AUC Curve of RF Classifier with hybrid fused deep and handcrafted features



**Fig. 13** **a** Feature importance defined in SHAP value and; **b** influence of heterogeneous features on a typical class of medicinal plant leaves

Our proposed methodology presents excellent classification results and a comparative analysis of our methodology with other state of art methodologies is presented in Table 8. As marked out in Table 8, the proposed model for the classification of medicinal plant leaves demonstrates excellent classification results amongst the other state of art models. In [44], the authors presented a transfer learning approach for deep feature extraction using VGG-16,

VGG-19, Inception V3 and Xception pre-trained CNN networks. Artificial Neural Network, Support Vector Machine (SVM) and SVM+Bayesian Optimization (BO) were employed for the classification of medicinal plant leaves. The proposed Xception/ANN model achieved an accuracy of 97.50% with precision and sensitivity of 98%. Though their model achieved excellent results, even though the model is not suited for classifying the images in complex backgrounds. Dileep M.R et al. [16] proposed Alexnet pre-trained CNN network for feature extraction of medicinal plant leaves. They employed an SVM classifier for classifying medicinal plant leaves. Their model attains an accuracy of 96.76% with 5 cross-validations. A CNN-LSTM model was projected by Haryono, K. et al. [45] The CNNLSTM architecture is employed for feature extraction and classification of herbal plant leaves and their CNN-LSTM model obtained 94.96% accuracy. Studies presented in [16,45] were evaluated in terms of single performance parameter “accuracy” only and no methodology was proposed by authors to reduce the high dimensionality of the deep features dataset. Our proposed RF classifier with heterogeneous fused features attains remarkable results in terms of all the considered performance evaluation matrices. The dimensionality of the fused dataset is reduced by UMAP and the model attains an accuracy and sensitivity of 99.18% and precision of 99.20%.

## 7 Conclusion

This research intends to identify and classify the Indian medicinal plant leaves using machine learning techniques which would help us to achieve better detection and classification results. The proposed methodology was based on intermediate-level fusion of features, dimensionality reduction of heterogeneous fused feature vectors, and classification using machine learning classifiers. The handcrafted features are extracted using EH, PLBP and PHOG techniques and the deep features are extracted by VGG-16 and DenseNet-121 CNN networks. Improved classifier performance was observed with deep features with a notable drop in the probability of false alarm. The empirical analysis revealed that the classification performance can be further improved with the fusion of features. However, the primary disadvantage of a fused feature vector is its dimensionality. Linear and non-linear dimensionality techniques were employed for dimensionality reduction Techniques. It was found that with UMAP an improvement of approximately 1.30% was observed in the classifier’s performance in terms of accuracy, precision and sensitivity and the probability of false alarm dips to the value of 0.02%. This improvement is primarily due to intermediate-level feature fusion and dimensionality reduction. The proposed fusion-based approach with UMAP dimensionality reduced feature vector obtained accuracy, precision and sensitivity of greater than 99%. So, traditional feature, colour,

**Table 8** Proposed methodology vs. other state of art models

Authors	Methodology	Accuracy (%)	Precision (%)	Sensitivity (%)
S. Roopashree et al. [44]	Xception/ANN	97.50	98	98
Dileep M. R. et al. [16]	AlexNet/SVM	96.76	---	---
Haryono, K. et al. [45]	CNN-LSTM	94.96	---	---
<b>Proposed methodology</b>	RF Classifier with input attribute as handcrafted and deep fused features with subsequent UMAP for feature-set dimensionality reduction.	99.18	99.20	99.18

edge etc. based identification methods can be incorporated with transfer learning methods in the domain of medicinal plants for better and more accurate identification and classification. These computer-aided identification techniques provide significant results without the need for human expertise with less effort. The future scope of the work can be extended by expanding the dataset with herbal plant leaves collected from different regions and different climatic conditions. The authors will also investigate to further decrease the fused feature complexity.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s11042-023-17639-1>.

**Funding** None.

**Data availability** The dataset is publicly available at <https://data.mendeley.com/datasets/nnytj2v3n5> (accessed on 29 June, 2022).

## Declarations

**Ethical approval** This article does not contain any studies with human participants or animals performed by any of the authors.

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

1. Vilasini M, Ramamoorthy P (2020) Machine learning approaches for classification of Indian leaf species using smartphone images. A Thesis submitted to Faculty of Information and Communication Engineering, Anna University. <http://hdl.handle.net/10603/333970>. Accessed 15 Nov 2022
2. Kan HX, Jin L, Zhou FL (2017) Classification of medicinal plant leaf image based on multi-feature extraction. *Pattern Recognit Image Anal* 27:581–587. <https://doi.org/10.1134/S105466181703018X>
3. Ganguly S, Bhowal P, Oliva D, Sarkar R (2022) BLeafNet: a bonferroni mean operator based fusion of CNN models for plant identification using leaf image classification. *Ecol Inf* 69:101585. <https://doi.org/10.1016/j.ecoinf.2022.101585>. (ISSN 1574–9541)
4. Sathwik T, Yasaswini R, Venkatesh R, Gopal A (2013) Classification of selected medicinal plant leaves using texture analysis. 2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT), 2013, pp. 1–6. <https://doi.org/10.1109/ICCCNT.2013.6726793>
5. Figueiredo MSL, Grelle CEV (2009) Predicting global abundance of a threatened species from its occurrence: implications for conservation planning. *Divers Distrib* 15:117–121. <https://doi.org/10.1111/j.1472-4642.2008.00525.x>
6. Barré P, Stöver BC, Müller KF, Steinhage V (2017) LeafNet: a computer vision system for automatic plant species identification. *Ecol Inf* 40:50–56. <https://doi.org/10.1016/j.ecoinf.2017.05.005>
7. Gopal A, Prudhveeswar Reddy S, Gayatri V (2012) Classification of selected medicinal plants leaf using image processing. 2012 International Conference on Machine Vision and Image Processing (MVIP), pp. 5–8, <https://doi.org/10.1109/MVIP.2012.6428747>
8. Sainin MS, Alfred R (2014) Feature selection for Malaysian medicinal plant leaf shape identification and classification. 2014 International Conference on Computational Science and Technology (ICCST), pp. 1–6, <https://doi.org/10.1109/ICCST.2014.7045183>
9. Turkoglu M, Hanbay D (2019) Leaf-based plant species recognition based on improved local binary pattern and extreme learning machine. *Phys A: Stat Mech Appl* 527:121297. <https://doi.org/10.1016/j.physa.2019.121297>
10. Tm P, Pranathi A, SaiAshritha K, Chittaragi NB, Koolagudi SG (2018) Tomato leaf disease detection using convolutional neural networks. 2018 Eleventh International Conference on Contemporary Computing (IC3), pp. 1–5, <https://doi.org/10.1109/IC3.2018.8530532>
11. Yadav R, Kumar Rana Y, Nagpal S (2019) Plant Leaf Disease detection and classification using particle swarm optimization. In: Renault É, Mühlethaler P, Boumerdassi S (eds) *Machine learning for networking*. MLN 2018. Lecture Notes in Computer Science, vol 11407. Springer, Cham. [https://doi.org/10.1007/978-3-030-19945-6\\_21](https://doi.org/10.1007/978-3-030-19945-6_21)

12. Pankaja K, Suma V (2020) Plant leaf recognition and classification based on the Whale Optimization Algorithm (WOA) and Random Forest (RF). *J Inst Eng India Ser B* 101:597–607. <https://doi.org/10.1007/s40031-020-00470-9>
13. Tan JW, Chang SW, Abdul-Kareem S, Yap HJ, Yong KT (2020) Deep learning for plant species classification using Leaf Vein morphometric. In: IEEE/ACM transactions on Computational Biology and Bioinformatics. 17(1):82–90. <https://doi.org/10.1109/TCBB.2018.2848653>
14. Naeem S, Ali A, Chesneau C, Tahir MH, Jamal F, Sherwani RAK, Ul Hassan M (2021) Ul Hassan, M. The classification of medicinal plant leaves based on multispectral and texture feature using machine learning approach. *Agronomy* 11:263. <https://doi.org/10.3390/agronomy11020263>
15. Ibrahim Z, Sabri N, Mangshor NNA (2018) Leaf recognition using texture features for herbal plant identification. *Indonesian J Electr Eng Comput Sci* 9(1):152–156
16. Dileep MR, Pournami PN (2019) AyurLeaf: A deep learning approach for classification of medicinal plants. In: Proc. IEEE Region Conf. (TEN-CON), Oct. pp. 321\_325
17. Roopashree S, Anitha J (2020) Medicinal Leaf Dataset. Mendeley Data V1. <https://doi.org/10.17632/nnytj2v3n5.1>
18. Tsalera E, Papadakis A, Samarakou M (2022) I. feature extraction with handcrafted methods and convolutional neural networks for facial emotion recognition. *Appl Sci* 12:8455. <https://doi.org/10.3390/app12178455>
19. Won CS, Park DK, Park S-J (2002) Efficient Use of MPEG-7 Edge Histogram Descriptor. *Journal ETRI* 24(1 February):23–30
20. Won CS (2004) Feature extraction and evaluation using Edge Histogram Descriptor in MPEG-7. In: Aizawa K, Nakamura Y, Satoh S (eds) Advances in Multimedia Information Processing - PCM 2004. PCM 2004. Lecture Notes in Computer Science, vol 3333. Springer, Berlin, Heidelberg. doi: [https://doi.org/10.1007/978-3-540-30543-9\\_73](https://doi.org/10.1007/978-3-540-30543-9_73)
21. Bosch A, Zisserman A, Munoz X (2007) Representing shape with a spatial pyramid kernel. *Image Process* 5:401–408
22. Kumar N, Sharma M, Singh VP, Madan C, Mehandia S (2022) An empirical study of handcrafted and dense feature extraction techniques for lung and colon cancer classification from histopathological images. *Biomed Signal Process Control* 75:103596. <https://doi.org/10.1016/j.bspc.2022.103596>
23. Ojala T, Pietikainen M, Maenpaa T (2002) Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans Pattern Anal Mach Intell* 24(7):971–987
24. Qian X, Hua X-S, Chen P, Ke L (2011) An effective local binary patterns texture descriptor with pyramid representation. *Pattern Recogn* 44:10–11
25. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556
26. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ (2017) Densely connected convolutional networks. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA. pp. 2261–2269. <https://doi.org/10.1109/CVPR.2017.243>
27. Huang SC, Pareek A, Seyyedi S, Banerjee I, Lungren MP (2020) Fusion of medical imaging and electronic health records using deep learning: a systematic review and implementation guidelines. *NPJ Digit Med* 3(1):1–9
28. Hu A, Zhang R, Yin D, Zhan Y (2014) Image quality assessment using a SVD-based structural projection. *Signal Process Image Commun* 29:293–302
29. Roweis ST, Saul LK (2000) Nonlinear dimensionality reduction by locally linear embedding. *Science* 290(5500):2323–2326
30. Cunningham JP, Ghahramani Z (2015) Linear dimensionality reduction: survey, insights, and generalizations. *J Mach Learn Res* 16(89):2859–2900
31. Sugiyama M (2016) Nonlinear dimensionality reduction. In Introduction to Statistical Machine Learning; Elsevier, pp 429–446. <https://doi.org/10.1016/b978-0-12-802121-7.00047-9>
32. Mishra S, Panda M (2020) 2 an analysis on non-linear dimension reduction techniques. *Mach Learn Appl* 19–42. <https://doi.org/10.1515/9783110610987-004>
33. McInnes L, Healy J, Saul N, Großberger LUMAP (2018) Uniform manifold approximation and projection. *J Open Source Softw* 3(29):861
34. Sharma M, Monika, Kumar N et al (2021) Badminton match outcome prediction model using Naïve Bayes and feature weighting technique. *J Ambient Human Comput* 12:8441–8455. <https://doi.org/10.1007/s12652-020-02578-8>
35. Velliangiri S, Alagumuthukrishnan S, Iwin Thankumar joseph S (2019) A review of dimensionality reduction techniques for efficient computation. *Procedia Comput Sci* 165:104–111. <https://doi.org/10.1016/j.procs.2020.01.079>

36. McClish DK (2012) Evaluation of the Accuracy of medical tests in a region around the optimal point. *Acad Radiol* 19(12):1484–1490. <https://doi.org/10.1016/j.acra.2012.09.004>. ([Online]. Available)
37. Bradley AP (1997) The use of the area under the {ROC} curve in the evaluation of machine learning algorithms. *Pattern Recogn* 30:1145–1159
38. Hajian-Tilaki K (2013) Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. *Casp J Intern Med* 4(2):627–635
39. Fawcett T (2006) An introduction to ROC analysis. *Pattern Recogn Lett* 27(8):861–874. <https://doi.org/10.1016/j.patrec.2005.10.010>
40. Rudin C (2019) Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat Mach Intell* 1:206–215. <https://doi.org/10.1038/s42256-019-0048-x>
41. Toja E, Guan C (2021) A survey on Explainable Artificial Intelligence (XAI): toward Medical XAI. *IEEE Trans Neural Netw Learn Syst* 32(11):4793–4813
42. Linardatos P, Papastefanopoulos V, Kotsiantis S, Explainable AI (2021) A review of machine learning interpretability methods. *Entropy* 23:18. <https://doi.org/10.3390/e23010018>
43. Lundberg SM, Lee SI (2017) A unified approach to interpreting model predictions. *Adv Neural Inf Process Syst* 30:4765–4774. <https://arxiv.org/abs/1705.07874>
44. Roopashree S, Anitha J (2021) DeepHerb: a vision based system for medicinal plants using Xception features. In: *IEEE Access* 9:135927–135941. <https://doi.org/10.1109/ACCESS.2021.3116207>
45. Haryono K, Anam, Saleh A (2020) Herbal leaf identification and authentication using deep learning neural network. A Novel International Conference on Computer Engineering, Network, and Intelligent Multimedia (CENIM), pp 338–342. <https://doi.org/10.1109/CENIM51130.2020.9297952>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

## Authors and Affiliations

Manoj Sharma<sup>1</sup> · Naresh Kumar<sup>2</sup>  · Shaloo Sharma<sup>3</sup> · Sumit Kumar<sup>4</sup> ·  
Sukhjinder Singh<sup>1</sup> · Seema Mehandia<sup>5</sup>

- ✉ Manoj Sharma  
[neelmanoj@gmail.com](mailto:neelmanoj@gmail.com)
- ✉ Naresh Kumar  
[naresh\\_uiet@yahoo.com](mailto:naresh_uiet@yahoo.com)
- ✉ Sumit Kumar  
[kumarsumit8@gmail.com](mailto:kumarsumit8@gmail.com)
- Sukhjinder Singh  
[er.ssrankhra@gmail.com](mailto:er.ssrankhra@gmail.com)

<sup>1</sup> Department of ECE, Giani Zail Singh Campus College of Engineering & Technology, MRSPTU, Bathinda, India

<sup>2</sup> Department of ECE, UIET, Panjab University, Chandigarh, India

<sup>3</sup> School of Computer Science and Engineering Technology, Bennett University, Greater Noida, U.P., India

<sup>4</sup> SEEE, Centre for Space Research, Division of R&D, Lovely Professional University, Phagwara, Punjab, India

<sup>5</sup> Department of Biotechnology, UIET, Panjab University, Chandigarh, India