Bulletin of Electrical Engineering and Informatics
Vol. 12, No. 1, February 2023, pp. 75~85
ISSN: 2302-9285, DOI: 10.11591/eei.v12i1.4388

Cluster-based segmentation for tobacco plant detection and classification

Thirthe Gowda Mallinathapura Thimmegowda[1], Chandrika Jayaramaiah[2]
[1]Department of Computer Science and Engineering, Government Engineering College, Karnataka, India
[2]Department of Computer Science and Engineering, Malanad College of Engineering, Karnataka, India

Article Info

Article history:

Received Jul 9, 2022
Revised Sep 5, 2022
Accepted Sep 30, 2022

Keywords:

Classification
CNN
Machine learning
Motion trajectory
Spatial patterns
Tobacco plant detection
Unmanned aerial vehicle

ABSTRACT

Tobacco is one of the major economical crops in the agriculture sector. It is essential to detect tobacco plants using unmanned aerial vehicle (UAV) images for improved crop yield and plays an important role in the early treatment of tobacco plants. The proposed research work is carried out in three phases: In the first phase, we collect images from UAV's and apply the French Commision Internationale de l'eclairage (CIE) L*a*b colour space model as pre-processing operations and segmentation. And then two prominent motion descriptors namely histogram of flow (HOF) and motion boundary histogram (MBH) are combined with the optimal histogram of oriented gradients (HOG) descriptor for exploring optimal motion trajectory and spatial measurements. And finally, the spatial variations with respect to the scale and illumination changes are incorporated using the optimal HOG descriptor. Here both dense motion patterns and HOG are refined using hierarchical feature selection using principal component analysis (PCA). The proposed model is trained and evaluated on different tobacco UAV image datasets and done a comparative analysis of different machine learning (ML) algorithms. The proposed model achieves good performance with 95% accuracy and 92% of sensitivity.

Corresponding Author:

Thirthe Gowda Mallinathapura Thimmegowda
Department of Computer Science and Engineering, Government Engineering College
Hassan, Karnataka, 573202, India
Email: mailto:thirthegowda.gech@gmail.com

1. INTRODUCTION

Now a day's many commercial crops are grown and cultivated worldwide and also selling their cultivated crops in national and international markets. In any crop management, with the increase in the area under cultivation, the diversity and type of weeds will also increase. Tobacco is one of the most commercially significant crops grown in and cultivated majorly in many countries like China, India, America, Brazil, and Russia. Because of its high economic value [1], [2]. India is the second largest producer and exporter after china and Brazil. In India, different breeds of tobacco plants are grown, flue-cured verginia (FCV) is one of them. Using of appropriate technology, crop management plays an important role in production, marketing, and exporting. The quality of the cured FCV tobacco leaves depends on classification, evaluation of the tobacco leaves [3], and also the growth of the tobacco plant. Many of the parameters are used to evaluate and estimate the growth of the tobacco plant. The evaluation and estimation of the growth of the tobacco plant play an important role in the management of tobacco crop planting better i.e. crop yield estimation, and disease detection, analysis [4], [5]. So these need to modernize and evolve new technology to improve production practices and crop management. Currently, the traditional yield estimation [6] is based on manual inspection on-site and which is time-consuming. Several studies have been made on yield estimation. In recent years several vision-based applications use both images and videos because these provide visual data for analysis and different kinds of decision-making systems [7].

There are many platforms that exist today for collecting and acquiring images quickly and cheaply such as remote sensing very high-resolution satellites [8], world view-2, GeoEye-1 [9], and unmanned aerial vehicle (UAV). The remote sensing images and videos are used for many applications for object detection and tracking, identifying and detecting buildings, offices, and also shapes of the building [10]. In the past few years due to cost-effectiveness, portability, mobility, and safety, flexibility in using unmanned aerial vehicles becomes more popular [11]. It has the flexibility to find solutions to many problems. The UAV's can be used to find solutions to many agricultural problems like vegetation monitoring [11], detection and counting of plants [12], detection of crop plants for early treatment [13], crop yield estimation [14], and commercial plant detection and analysis [15], [16]. In computer vision, object detection and recognition [17] are major fields for many applications. Processing of images is computationally expensive due to the vast amount of pixels that need to be processed along with the certain limitation of image processing. Most of the detection process focuses on images with improved spatial conditions and the vision system provides limited real-time detection properties.

Many studies have been made on machine vision methods for detection of the tobacco plants [18], [19]. With the advancements in image processing techniques, few attempts have been made on grading FCV tobacco leaves. The availability of a broad range of features to interpret local data of an object of interest gives way to several improvements in the object detection process. Xie *et al*. [20] proposed an algorithm for recognizing and counting a tobacco plant. The French Commision Internationale de l'eclairage (CIE) color models and morphological reconstruction have been used for the recognition of plant regions and support vector machines for classification. In this case, the use of the K nearest neighbor (KNN) method for identifying the local region of each plant is inefficient hence KNN requires the number of K values of classifying the number of pixels belonging to a single plant, and values of K must be known. Wang *et al.* [21] proposed an automatic method for the detection of an individual oil palm tree in UAV images. Here the system uses the histogram of oriented gradients (HOG) feature descriptor for the extraction of features for the machine learning (ML) algorithm. The computed HOG descriptor may vary if dense grids are not uniformly spaced and it requires normalizing overlapped local contract. Fan *et al*. [22] developed a deep neural network-based system for the automatic detection of the tobacco plant in UAV images. Here the proposed system uses the morphological operations and watershed algorithm for segmentation. The properties which are considered for identifying a central region of plants are not adequate if the plant orientation is varied, illumination, overlapped leaves, and sometimes the brightness of the leaves is also the same as the central region of the plant. Chen *et al*. [23] proposed citrus tree segmentation based on the monocolor machine vision technique and support vector machine (SVM). This system addresses the issues of illumination and orientation of the plants. Neupane *et al*. [24] invented a model for banana plant detection and counting using deep learning. It is incorporated with three image processing techniques such as contrast stretching, synthetic color transformation, and orthodontic methods for plant region extraction. This method is very expensive and time-consuming. Sun *et al*. [25] proposed an algorithm for tobacco plant detection using UAV red, green, and blue (RGB) images.

The region of interest is computed by the grouping of pixels. The grouping can be done on the searching algorithm and clustering method. Normally many of the searching algorithms like breadth first search (BFS) and depth first search (DFS) are used for grouping and in this work density-based spatial clustering algorithm with noise (DBSCAN) clustering methods have been integrated for the grouping of pixels. However, one of the drawbacks of the DBSCAN clustering technique is not suitable for high dimension data, neck type of dataset, and varying density cluster. Aitelkadi *et al.* [26] proposed a methodology based on the deep learning approach for the detection of fruit trees and counting. In this proposed statistical structural data are extracted and used for model learning. In form last decade, some of the research work is carried out for detection systems by making use of existing neural network models like convolutional neural networks (CNN), region-based CNN (R-CNN) [27], faster R-CNN (FR-CNN) [28], and Yolo for detection and classification purposes [29]. In some case HOG [30] are extensively used for extraction of object location information and discriminatively trained template models are created using deformable HOG parts. Zhu *et al.* [31] used a cascade of histograms of oriented gradients from a dense set of small-scale blocks over the entire detection window to mitigate the detection problems over extremely complicated backgrounds and dramatic illumination changes with the limitation of increased computation cost.

The efficiency of local descriptors is extended to represent the 3D volumes from each interest point, such as histogram of optical flow (HOF) [32], and 3D HOG [33] which can explore object and motion information's effectively for improved classification rate. Duta *et al.* [34] investigated the histograms of motion gradients (HMG) descriptor to model both temporal and spatial derivation between two succeeded frames. and shape difference [35] driven aggregated descriptors are used to encode the unique shape information about actions of various classes. Though it offers considerable performance metrics still it has limitations over large-scale video indexing due to its computational complexity. Kong and Fu [36] developed the heritage at risk (HAR) system by extracting spatio-temporal features which are consistent with intrinsic characteristics of motion. Scale invariant feature transform (SIFT)-driven compound matching is carried out in [37] over consecutive frames for event classification. However, these algorithms have limitations in terms of diminished precision levels relating to the location of the human body. Liu *et al.* [38] introduced a novel multimodal and multi-view concurrent video analysis to overcome challenges involved with intra-class variance. This cross-domain and multitask ML model significantly reduce the intra-class variations due to view changes.

Following the art of review, we aim to use edge-based segmentation and cluster-based segmentation for the efficient detection of tobacco plants. In this paper, a hybrid clustering-based approach is proposed for the detection of tobacco plants in UAV images. Instead of employing the traditional method, we proposed to use the CIE Lab color space model for segmenting the image primarily from the dynamic background. Edge-based segmentation is incorporated to detect and extract edge features and edges are traced using corresponding edge features. This method helps identify certain specific centric features of plants for clustering. Hierarchical clustering is employed initially for groping all selected feature points and the K-means clustering algorithm is applied to obtain the centric of each cluster for each plant. Furthermore, an improved watershed algorithm is applied to segment connected plant regions based on cluster points. Finally, perform extraction of plant regions that are inputted to the ML or deep learning for classification of the tobacco plant and weed plants.

The major contributions to this research work are as follows: i) edges of occluded plant regions are detected by using an edge segmentation algorithm, ii) all bright pixel points present inside the plant region are detected using template matching, which helps to identify the center portion of each plant, iii) hybrid clustering algorithm for grouping and combining many clusters to find the centroid of each plant, and iv) each plants are segmented based on certriod using watershed algorithm. The organization of this research paper is as follows: section 3 presents materials and proposed methods which include edge detection and feature extraction for clustering. Experimental results and detailed analysis are presented in section 4. Finally, section 5 presents the conclusion of the research work.

Material and methods: dataset: in machine vision, collecting and acquiring an input image for various algorithms is a critical job. Our RGB color images are acquired by using UAVs equipped with the digital camera. The images are captured in a tobacco plantation region with different altitudes, and different illumination conditions in the village Ramanathapuram (nearer to the central tobacco board), Arakalagud, Hassan, Karnataka, India. We captured the same tobacco plantation agriculture regions three times in different duration i.e. 20 days gap duration for each capture for crop yield estimation. The dataset has 500 color RGB aerial images with 4k horizontal resolution with 3,840 pixels in width and 2,160 pixels in height. RGB aerial image is the 24-bit color image and each pixel is digitized with 8 bits per channel. The dataset increased to 7,715 individual plant region images by augmentation and the dataset were divided into 75% of the images as training data and 25% of the images as validation and testing dataset.

2. PROPOSED METHOD OF TOBOCCA PLANT DETECTION

The proposed tobacco plant detection system comprises four stages. In the first stage of pre-processing, connected plant region segmentation is in the second stage, from the third stage is individual plant regions segmentation, and the fourth stage is plant detection. Pre-processing and plant region segmentation: For accurate segmentation of the plant region it requires pretreating the captured UAV RGB color image. Normally the pixel values as an image have different values in all the channels of an RGB image and coordinates in an RGB image have tristimulus values ranging from 0 to 1.

$$X = 0.490R + 0.310G + 0.200B \tag{1}$$

$$Y = 0.177R + 0.831G + 0.010B \tag{2}$$

$$Z = 0.000R + 0.010G + 0.990B \tag{3}$$

The RGB values are normalized by defining.

$$\begin{aligned} r &= R/(R+G+B), \\ g &= G/(R+G+B), \\ b &= B/(R+G+B) \end{aligned} \tag{4}$$

To segment the plant region and soil region, it needs to binarize the RGB image by preserving green pixels. Further, the obtained green channel values are changed to 1 and the remaining values to zero on some thresholding values. The corresponding equations to compute this are as (5).

$$ROI(x,y) = Ig(x,y) - Gray(x,y) \tag{5}$$

Where ROI (x,y) refers to the preserved green pixel value pointed by coordinates x and y. Ig (x,y) is the original green channel its values are higher than the other two channels in RGB image as per tristimulus. Gray(x,y) is the gray value of the RGB, and pixel values are the average of the RGB channels. It represents the contribution of lightness or luminance by all three channels. The main objective of this computation is to reduce high computational costs. As per the review study, many researchers use different non-linear computational methods to preserve green channels. To speed up luminance calculations we employed the linear approximation method. The conversion equation for RGB to gray is as (8).

$$Gray(x,y) = 0.299R + 0.587G + 0.114B \tag{8}$$

Where Gray has 256 different values ranging from 0 to 255. In order to remove salt and pepper noise, smoothing and blurring images median blurring is applied. The resultant image ROI(x,y) is convolved with 2D median blurring kernel blr(x,y) with dimensions 3×3.

$$ROI\ (x,y) = ROI(x,y) * blr(x,y) \tag{7}$$

ROI (x, y) is the resultant grayscale denoised image, the pixels belonging to ROI on this grayscale are selected by binarizing using morphological filtering shape analysis by assigning 0 and 1to respectively. The resultant binary image is then post-processed by filling background pixels that can't be reached and it. can be filled from the image edge. Finally, the region of interest is labeled that contains objects found in the binary image. The process of preprocessing and candidate region segmentation is shown in Figure 1. It has certain regions marked with a red circle indicating regions that are not connected to the background. These regions are filled by the fill-hole method with color space driven plant region detection and binarized edge detection as shown in Figures 2(a) and 2(b) respectively. It also shows labeled each blob with 8 connectivity regions and each region is displayed with a different color. This resultant image, clearly shows that the segmented region in an image has more than one plant due to occlusion, scale invariance, dynamic background, and different illumination condition.

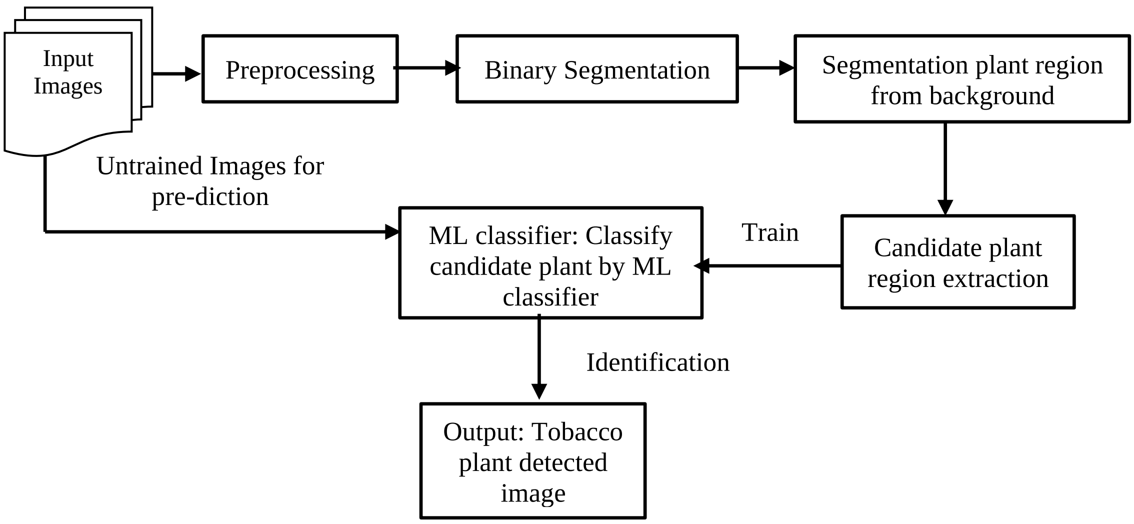
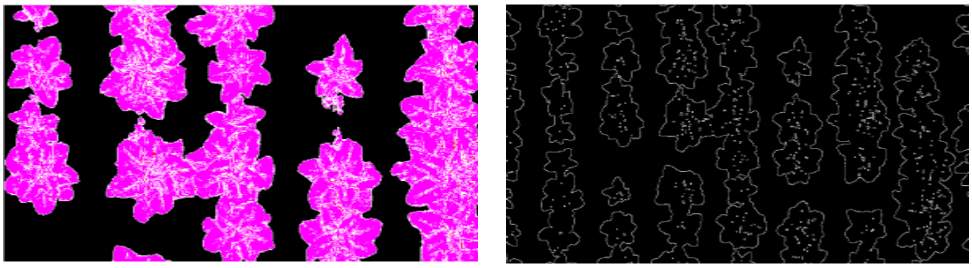


Figure 1. General architecture of plant detection system

2.1. Plant region detection and extraction

As we discussed in the previous section binary segmentation was employed to segment the candidate plant region and soil region. The preserved plant regions on the UAV RGB color image are segmented using grabcut algorithm to identify an individual plant from the occluded region. At the vigorous stage and maturing stage, the tobacco plant leaves are overlapped and occluded, all plant regions are connected, leaves portion is also has brighter region and plant region may have more number of brighter region. Hence, It is very difficult to identify the central region of the plant by using a brighter region because of the existence of many bright regions in a single plant region and it is inadequate for classification algorithms. Finding brighter pixels in a very large dataset using KNN is computationally expensive. To address these issues we employ a series of segmentation methods.



(a) (b)

Figure 2. Process of plant region segmentation using grabcut and CIE lab (a) segmented plant region after CIE lab in binary and in color space and (b) edge detected in plant region using edge segmentation

2.1.1. Cluster-based segmentation

To improve the tobacco plant detection rate, an approximate canny edge detection algorithm is applied to find the boundary of the entire region from the binary segmented region. The main objective of this algorithm is, to with the plant margin shape of all bright regions inside the plants detected. Optimized brighter regions have been extracted from plant regions using a defined template and the centroid of each region is used as a feature point for clustering to detect the center point of each plant. The following sub-matrix is used as a template for feature matching with zero padding.

$$\begin{matrix} [1,1,1,0 \\ 1,0,1,0 \\ 1,1,1,0 \\ 0,0,0,0] \end{matrix} \quad (8)$$

There after distance between every pair of feature point pixel is calculated using Euclidian distance. The equation for calculating distance is as (9):

$$\text{Dist }(x, y) = \sqrt[n]{\sum_{i=1} (y_i - p_i)^2} \quad (9)$$

Where dist(x,y) is a distance matrix between pair of two points. Thereafter establish a linkage between every pair by creating using a tree with the specific method which describes the measuring distance between clusters. Further, we employ an agglomerative hierarchical cluster from the linkage. Through this, it achieves a grouping of a selected feature point which helps in identifying a region from the plants. The hierarchical cluster returns a cluster a point assigned to each observation as bounded region as shown in Figure 3(a). Figures 3(b) and 3(c) respresents the dendrogram of clusters and clustering of selected tree respectively. Finally, the classification results shown in Figure 3(d) clearly shows that there are many clusters in single individual plants. To make a grouping of many cluster points to one unique single point for each plant a K-means cluster algorithm to apply. Based on the single plant pixel location, the boundaries of the individual plant are segmented using the watershed algorithm. After segmenting the plant region, an area of the individual plant regions pointed with coordinates by a bounding box is extracted from the original UAV image. It is very important to detect and classify tobacco plants and weed plants in a UAV image. In this step, various ML algorithms have been used for the detection and classification of tobacco plants.

2.1.2. Image labeling

Before training a ML algorithm with data samples, all extracted individual plants must be labeled first. Labeling an image is a human task that requires labeling cropped images. In this step, all cropped plant region images from the previous stage are labeled with tobacco, weed, and gross labels manually.

2.1.3. Image augmentation and resizing

Annotated images from the previous stage have a different size because images are annotated based on coordinates of bounding boxes and they may vary from each other. Before training these images to the model it needs to be resized because all the classification model accepts only the fixed size of the input. In our work, all the annotated images have been resized a shape of 64×64×3. One of the drawbacks of the ML algorithm is overfitting. If overfitting is high, the trained model can't able to classify the data correctly. To reduce overfitting, it needs to increase the training sample size, tune the hyperparameter, reduce the complexity of the model and system regularizatio. To increase and provide an adequate number of training samples to ML algorithms a data augmentation technique has been used in our work. The image augmentation in image processing consists of geometrical transformations like scaling, rotation, cropping, flipping, translation, and intensity transformation. In our case, rotation with θ (90^o, 180^o, 270^o, and 360^o), translation, resizing, shear, median blur, Gaussian blur, and average blur have been used.

2.1.4. ML algorithm

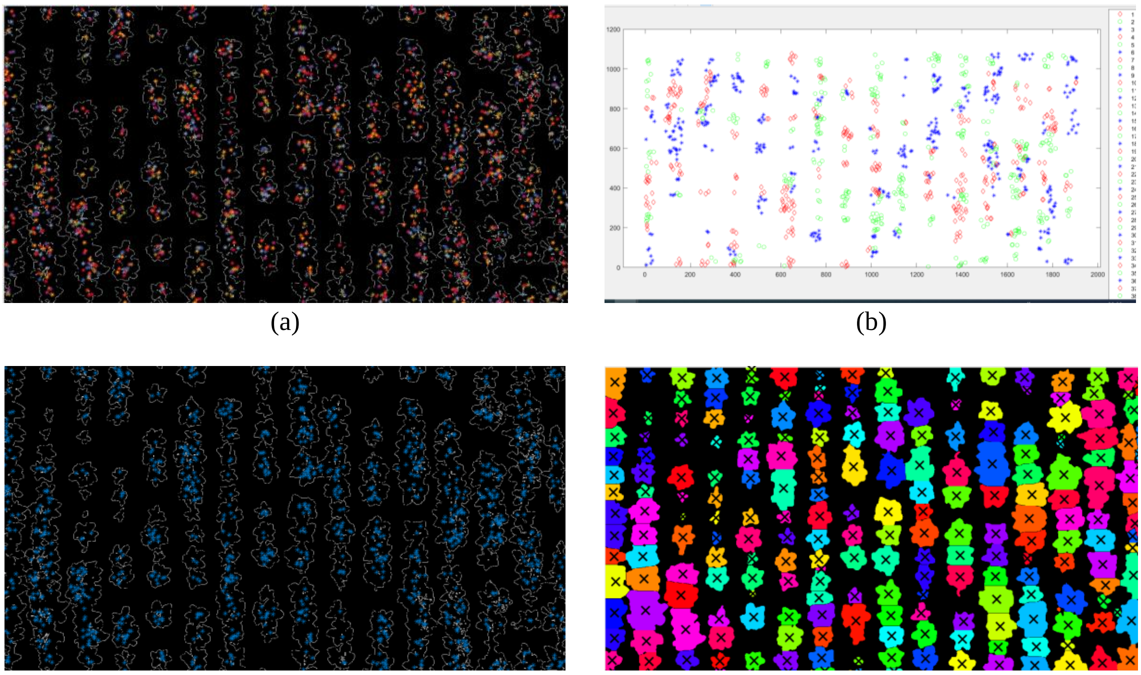
ML algorithm plays an important role in many applications. Some ML classification algorithms are used for object detection, recognition, feature extraction, and classification of the object. For our work, 2 different ML classifiers have been used for the detection and classification of tobacco plants in an image.

2.1.5. K-nearest neighbors

KNN is a non-parametric method for classification and regression. It uses a k nearest training samples in a dataset. Here the output of a KNN is a class member and an object is classified based on data points nearest neighbors and assigned to a class on the score of its neighbors. In our experiment, extracted individual plant regions are trained and classified using KNN. To increase the performance and the overfitting problem is solved by tuning the values of n neighbor's parameter.

2.1.6. Random forest algorithm

Random forest is a supervised ML algorithm. It is based on the concept of ensemble learning by combining multiple classifiers to increase performance and solve complex problems. It integrates multiple decision trees to predict the data of the dataset belonging to the class. Due to its advantages of small learning time, maintaining accuracy even if data missing and high prediction accuracy, we employ a random forest algorithm for the classification of the tobacco plant.



(a) (b) (c) (d)

Figure 3. Process of cluster base segmentation (a) and (b) set of features set selected for clustering, (c) clustering of selected points based on tree, and (d) individual plant region detected classification

3. RESULTS AND DISCUSSION

The performance analysis of tobacco plant object detection using a segmentation that is based on cluster-based along with a watershed is carried out with the help of the ML classifier.

3.1. Experimental setup

In this research work, we experimented with our tobacco plant datasets. As we explained in the previous section, the dataset size is 500 images still it is very small. To increase prediction performance and avoid overfitting, we employ data transformation and intensity transformation techniques as we explained in the previous section. After data augmentation and resizing, the dataset size is increased to 7,715 individual plant region images. The dataset of 7,715 images was divided into 75% of the images as training data, 15% of the images as validation data, and the remaining 10% image dataset used as testing. In our experiment, the proposed system has been trained, validated, and tested with CPU Intel core I7 2.4 GHz processor and 16 GB Memory. The image processing methods are employed for pre-processing, segmentation of plant region and ML models have been used for classification using MATLAB 2017a. When compared to the conventional method cluster-based segmentation shows enhanced performance in region segmentation on occluded regions, KNN, and random forest ML yield the best results on classification.

Metrics: to evaluate the performance of detection and classification, the implemented algorithms' performance can be evaluated using the following general metrics: sensitivity, specificity, and accuracy. These can be described in terms of true positive (TP), true negative (TN), false positive (FP), and false negative (FN). These performance results associated to given models are evaluated as illustrated in Table 1 for validation.

Table 1. Terms used to define metrics

Outcome of the test	Condition		Total
Positive	TP	FP	TP+FP
Negative	FN	TN	FN+TN
Total	TP+FN	FP+TN	=TP+TN+FP+FN

$$Sensitivity = TP/(TP + FN) \tag{10}$$

$$Specificity = TN/(TN + FP) \tag{11}$$

$$Precision = TP/(TP + FP) \tag{12}$$

$$Accuracy = (TP + TN)/TP + TN + FP + FN) \tag{13}$$

Where TP is a true positive indicates samples correctly classified as tobacco plant, TN is a true negative indicating correctly classified as weed plant, FP is a false positive indicating incorrectly identified as tobacco plant and FN is a false negative. Hyperparameter selection: to improve model performance and avoid overfitting during training and testing, many hyperparameters have been used. The proposed model has been trained and tested with KNN and random forest classifiers. The hyperparameter used for experimentation are listed in Table 2.

Table 2. Hyperparameter values

Algorithm	Hyperparameter	Tested value
KNN	n_neighbors	30 -3
	Leaf_size	30-50
	P	3-4
RF	N_estimators	10-100
	Max_features	Auto, sqrt, log2
	Random_state	30

3.2. Experimental results

From the experimental results the efficiency of the proposed method is proved to be an optimal one for automatic detection of tobacco plants in UAV images and this cluster-based approach achieves better accuracy than all other conventional methods. The following are attributes considered for classification. Samples from the real dataset are selected randomly for training and testing. The number of individual tobacco plant samples used for training, validation, and testing is given in the Table 3. Before training and testing a model, some of the features are extracted from segmented samples. Three main features such as color, texture, and shape have been considered and extracted before training it into classification algorithms as shown in Table 4. These features are the main features for classification. The accuracy of the KNN and RF algorithms obtained with all the above-mentioned features as given in Table 5 gives minimum accuracy, maximum accuracy, and average accuracy of KNN and RF respectively.

Table 3. Attribute considered for image plant classification

Attributes	Features
Image size	64×64×3
Image type and format	RGB, *.jpg
Number of images for training	5785
Number of images for validation	1160
Number of images for testing	770
Image category	Set1: highly occluded plants Set2: Partially occluded plant Set3: poor illumination condition

Table 4. Classification results using KNN and Table 5 shows training and testing scores of KNN

Features	Train score		Test score		Average Accuracy (%)
	Min. acc (%)	Max. acc (%)	Min. acc (%)	Max. acc (%)	
Color	97.68	99.46	97.51	99.01	98.41
texture	93.17	98.73	93.26	97.92	95.77
shape	90.05	95.91	89.99	91.70	91.91
No-feature	73.35	86.66	72.99	73.66	76.65

Table 5. Performance metrics scores on testing

Features	Classification metrics performance on testing %			
	Sensitivity	Specificity	Precision	Accuracy
KNN	95.98	98.61	96.20	96.93
RF	98.33	99.7	99	99.01
	Classification metrics performance on validation in %			
KNN	94.01	98.12	95.10	95.44
RF	97.50	99.24	98.99	98.67

Table 5 performs the KNN and RF algorithms on testing and prediction. The KNN achieves an average of 95.98% of sensitivity, 98.61% of specificity, and 96.10% of precision on color, texture, and shape features respectively. The performance of the random forest is higher than the KNN on both testing and prediction. The table, clearly shows that the overfitting is very high when the k-is large and degrades the learning rate. The overfitting can be reduced by tuning hyperparameters such as k-size, leaf size, and P. As we increase the k size, the model becomes so generalized but it drops its performance. Similarly, in the case of RF, the overfitting can be avoided by fine-tuning a hyperparameter such as n_estimator, Max_features, and random_state.

3.3. Confusion matrices analyzes

To validate the performance of the proposed tobacco classification system, the results have been compared with all different state-of-the-art methods as shown in Table 6 and Figure 4 with confusion metrics. In this method, both spatial and temporal descriptions are extracted through frequency detail analyzes. This method encodes the transformed frequency domain coefficients trajectories based on the frequency of occurrence and geometrical changes in various bands to model the motion characteristics which gives superior end results as shown in Figure 5. It was observed that due dependency between accuracy and the number of transform domain features tradeoff between accuracy and codebook’s size causes significant performance degradation. To improve computational efficiency PCA transform is applied on wavelet descriptor to stress the discrimination between various classes.

Table 6. Accuracy performance measure for RHOF HOG and other methods

Measures / Video Sets	Accuracy (%)		
	MBH	HOF	Proposed RHOF
Initial stage	79.16	76.42	84.55
Intermediate stage	66.10	69.33	71.23
Final stage	64.25	63.44	74.50

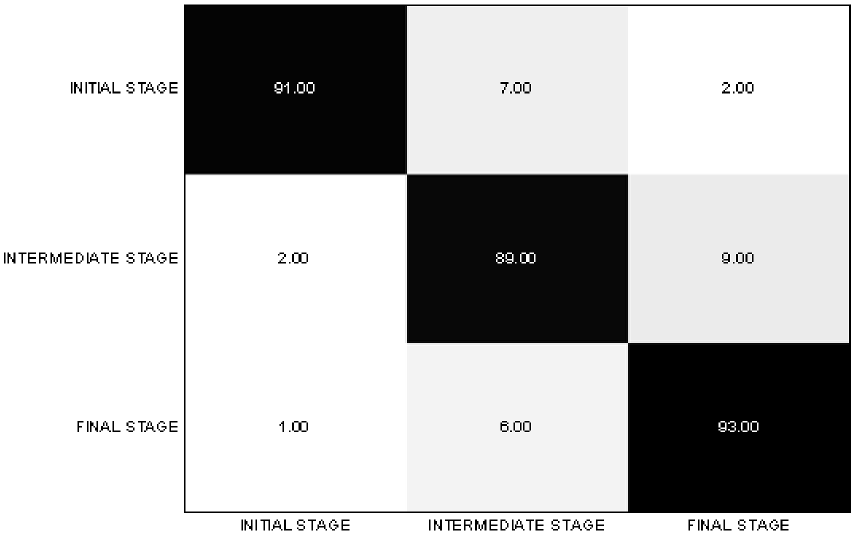


Figure 4. Confusion matrices of tobacco video classification system with improved feature subset

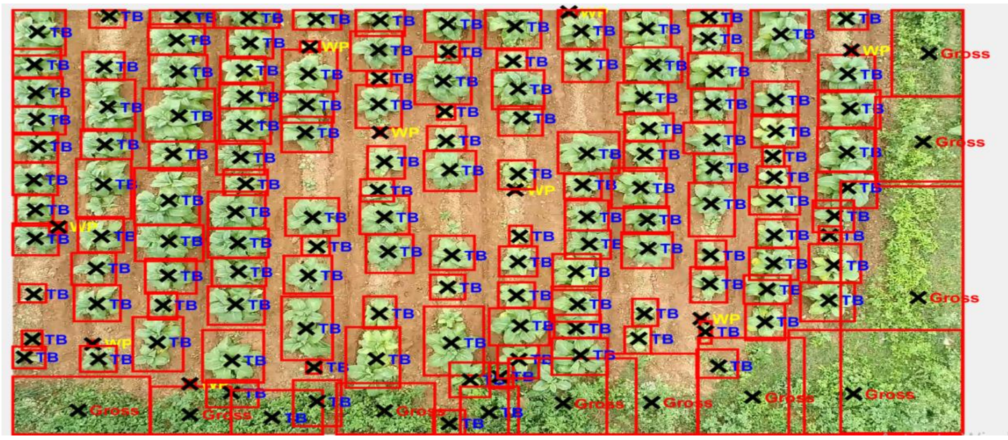


Figure 5. Visualization of final detection results

4. CONCLUSION

The cluster-based segmentation is introduced in this research. The performance analysis of tobacco plant object detection using a segmentation that is cluster-based along with a watershed technique is carried out with the help of the ML classifier. The study on ML classifier based on color, texture, and shape features is presented. The proposed model achieved a good average classification accuracy of 98% and 95% in RF and KNN models respectively. It offers improved recognition rates, and the performance measures showed the superiority of the proposed plant detection system as compared to the HOF+HOG-based deformable Model and PCA-driven histogram of flow method. The efficiency of the proposed method for automatic detection of tobacco plants in UAV images using a cluster-based approach achieves better accuracy than any other conventional method. This research work can be extended to improve the detection of unhealthy plants by considering certain important features of plants.

REFERENCES

[1] D. S. Guru, P. B. Mallikarjuna, S. Manjunath, and M. M. Shenoi, “Min-max representation of features for grading cured tobacco leaves,” *Statistics and Applications*, vol. 9, no. 1–2, pp. 15–29, 2011.

[2] D. S. Guru, P. B. Mallikarjuna, S. Manjunath, and M. M. Shenoi, “Machine vision based classification of tobacco leaves for automatic harvesting,” *Intelligent Automation and Soft Computing*, vol. 18, no. 5, pp. 581–590, Jan. 2012, doi: 10.1080/10798587.2012.10643267.

[3] F. Zhang and X. Zhang, “Classification and quality evaluation of tobacco leaves based on image processing and fuzzy comprehensive evaluation,” *Sensors*, vol. 11, no. 3, pp. 2369–2384, Feb. 2011, doi: 10.3390/s110302369.

[4] K. Yamamoto, W. Guo, Y. Yoshioka, and S. Ninomiya, “On plant detection of intact tomato fruits using image analysis and machine learning methods,” *Sensors*, vol. 14, no. 7, pp. 12191–12206, Jul. 2014, doi: 10.3390/s140712191.

[5] A. Fuentes, S. Yoon, S. Kim, and D. Park, “A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition,” *Sensors*, vol. 17, no. 9, pp. 1–21, Sep. 2017, doi: 10.3390/s17092022.

[6] W. Li, H. Fu, L. Yu, and A. Cracknell, “Deep learning based oil palm tree detection and counting for high-resolution remote sensing images,” *Remote Sensing*, vol. 9, no. 1, pp. 1–13, Dec. 2016, doi: 10.3390/rs9010022.

[7] J. P. Vasconez, J. Delpiano, S. Vougioukas, and F. A. Cheein, “Comparison of convolutional neural networks in fruit detection and counting: a comprehensive evaluation,” *Computers and Electronics in Agriculture*, vol. 173, pp. 1–12, Jun. 2020, doi: 10.1016/j.compag.2020.105348.

[8] F. M. -Noguer, P. N. Belhumeur, and S. K. Nayar, “Active refocusing of images and videos,” *ACM Transactions on Graphics*, vol. 26, no. 3, pp. 1–9, Jul. 2007, doi: 10.1145/1276377.1276461.

[9] C. S. Marzan and N. Marcos, “Towards tobacco leaf detection using haar cascade classifier and image processing techniques,” in *Proceedings of the 2nd International Conference on Graphics and Signal Processing - ICGSP’18*, 2018, pp. 63–68, doi: 10.1145/3282286.3282292.

[10] G. Montavon, W. Samek, and K. -R. Müller, “Methods for interpreting and understanding deep neural networks,” *Digital Signal Processing*, vol. 73, pp. 1–15, Feb. 2018, doi: 10.1016/j.dsp.2017.10.011.

[11] H. G. -Alvarez, J. E. Herrick, M. Mattocks, D. Toledo, and J. V. Zee, “Comparison of three vegetation monitoring methods: their relative utility for ecological assessment and monitoring,” *Ecological Indicators*, vol. 9, no. 5, pp. 1001–1008, Sep. 2009, doi: 10.1016/j.ecolind.2008.11.011.

[12] L. Chaerle, S. Lenk, D. Hagenbeek, C. Buschmann, and D. V. D. Straeten, “Multicolor fluorescence imaging for early detection of the hypersensitive reaction to tobacco mosaic virus,” *Journal of Plant Physiology*, vol. 164, no. 3, pp. 253–262, Mar. 2007, doi: 10.1016/j.jplph.2006.01.011.

[13] M. A. Aguilar, M. M. Saldaña, and F. J. Aguilar, “GeoEye-1 and WorldView-2 pan-sharpened imagery for object-based classification in urban environments,” *International Journal of Remote Sensing*, vol. 34, no. 7, pp. 2583-2606, 2013, doi: 10.1080/01431161.2012.747018.

[14] Q. Wang, S. Nuske, M. Bergerman, and S. Singh, “Automated crop yield estimation for apple orchards,” in *Experimental Robotics*, Heidelberg: Springer Tracts in Advanced Robotics, 2013, pp. 745–758.

[15] S. Li, Z. Tao, K. Li, and Y. Fu, “Visual to text: survey of image and video captioning,” *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 3, no. 4, pp. 297–312, Aug. 2019, doi: 10.1109/TETCI.2019.2892755.

[16] F. Martinelli *et al.* “Advanced methods of plant disease detection. A review, ” *Agronomy for Sustainable Development*, vol. 35, no. 1, pp. 1-25, 2015, doi: 10.1007/s13593-014-0246-1.

[17] Y. N. Wu, Z. Si, H. Gong, and S. -C. Zhu, “Learning active basis model for object detection and recognition,” *International Journal of Computer Vision*, vol. 90, no. 2, pp. 198–235, Nov. 2010, doi: 10.1007/s11263-009-0287-0.

[18] C. Liu, S. Li, F. Chang, and Y. Wang, “Machine vision based traffic sign detection methods: review, analyses and perspectives,” *IEEE Access*, vol. 7, pp. 1–19, 2019, doi: 10.1109/ACCESS.2019.2924947.

[19] C. Koch, K. Georgieva, V. Kasireddy, B. Akinci, and P. Fieguth, “A review on computer vision based defect detection and condition assessment of concrete and asphalt civil infrastructure,” *Advanced Engineering Informatics*, vol. 29, no. 2, pp. 196–210, Apr. 2015, doi: 10.1016/j.aei.2015.01.008.

[20] H. Xie, Z. Fan, W. Li, Y. Rong, Y. Xiao, and L. Zhao, “Tobacco plant recognizing and counting based on SVM,” in *2016 International Conference on Industrial Informatics - Computing Technology, Intelligent Technology, Industrial Information Integration (ICIICII)*, Dec. 2016, pp. 109–113, doi: 10.1109/ICIICII.2016.0037.

[21] Y. Wang, X. Zhu, and B. Wu, “Automatic detection of individual oil palm trees from UAV images using HOG features and an SVM classifier,” *International Journal of Remote Sensing*, vol. 40, no. 19, pp. 1–15, Oct. 2018, doi: 10.1080/01431161.2018.1513669.

[22] Z. Fan, J. Lu, M. Gong, H. Xie, and E. D. Goodman, “Automatic tobacco plant detection in uav images via deep neural networks,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 3, pp. 1–12, Mar. 2018, doi: 10.1109/JSTARS.2018.2793849.

[23] Y. Chen *et al.*, “Citrus tree segmentation from UAV images based on monocular machine vision in a natural orchard environment,” *Sensors*, vol. 19, no. 24, p. 5558, Dec. 2019, doi: 10.3390/s19245558.

[24] B. Neupane, T. Horanont, and N. D. Hung, “Deep learning based banana plant detection and counting using high-resolution red-green-blue (RGB) images collected from unmanned aerial vehicle (UAV),” *PLOS ONE*, vol. 14, no. 10, pp. 1–22, Oct. 2019, doi: 10.1371/journal.pone.0223906.

[25] X. Sun, J. Peng, Y. Shen, and H. Kang, “Tobacco plant detection in RGB aerial images,” *Agriculture*, vol. 10, no. 3, pp. 1–15, Feb. 2020, doi: 10.3390/agriculture10030057.

[26] K. Aitelkadi, H. Outmghoust, S. Laarab, K. Moumayiz, and I. Sebari, “Detection and counting of fruit trees from RGB UAV images by convolutional neural networks approach,” *Advances in Science, Technology and Engineering Systems Journal*, vol. 6, no. 2, pp. 887–893, Apr. 2021, doi: 10.25046/aj0602101.

[27] F. Gao *et al.*, “Multi-class fruit-on-plant detection for apple in SNAP system using faster R-CNN,” *Computers and Electronics in Agriculture*, vol. 176, pp. 1–10, Sep. 2020, doi: 10.1016/j.compag.2020.105634.

[28] S. K. Behera, A. K. Rath, and P. K. Sethy, “Fruits yield estimation using Faster R-CNN with MIoU,” *Multimedia Tools and Applications*, vol. 80, no. 12, pp. 19043–19056, May 2021, doi: 10.1007/s11042-021-10704-7.

[29] J. Liu and X. Wang, “Tomato diseases and pests detection based on improved yolo V3 convolutional neural network,” *Frontiers in Plant Science*, vol. 11, pp. 1–12, Jun. 2020, doi: 10.3389/fpls.2020.00898.

[30] S. Abbasi and F. Tajeripour, “Detection of brain tumor in 3D MRI images using local binary patterns and histogram orientation gradient,” *Neurocomputing*, vol. 219, pp. 526–535, Jan. 2017, doi: 10.1016/j.neucom.2016.09.051.

[31] Q. Zhu, M. -C. Yeh, K. T. Cheng, and S. Avidan, “Fast human detection using a cascade of histograms of oriented gradients,” in *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume 2 (CVPR’06)*, vol. 2, pp. 1491–1498, doi: 10.1109/CVPR.2006.119.

[32] R. V. H. M. Colque, C. Caetano, M. T. L. de Andrade, and W. R. Schwartz, “Histograms of optical flow orientation and magnitude and entropy to detect anomalous events in videos,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 27, no. 3, pp. 1–10, Mar. 2017, doi: 10.1109/TCSVT.2016.2637778.

[33] A. Kläser, M. Marszałek, and C. Schmid, “A spatio-temporal descriptor based on 3D-gradients,” *BMVC 2008 - Proceedings of the British Machine Vision Conference 2008*, 2008, doi: 10.5244/C.22.99.

[34] I. C. Duta, J. R. R. Uijlings, B. Ionescu, K. Aizawa, A. G. Hauptmann, and N. Sebe, "Efficient human action recognition using histograms of motion gradients and VLAD with descriptor shape information," *Multimedia Tools and Applications*, vol. 76, no. 21, pp. 22445–22472, Nov. 2017, doi: 10.1007/s11042-017-4795-6.

[35] J. Sun, X. Dai, Q. Wang, M. C. M. van Loosdrecht, and B. -J. Ni, "Microplastics in wastewater treatment plants: Detection, occurrence and removal," *Water Research*, vol. 152, pp. 21–37, Apr. 2019, doi: 10.1016/j.watres.2018.12.050.

[36] Y. Kong and Y. Fu, "Human action recognition and prediction: a survey," *International Journal of Computer Vision*, vol. 130, no. 5, pp. 1–37, 2022, doi: 10.1007/s11263-022-01594-9.

[37] F. Rançon, L. Bombrun, B. Keresztes, and C. Germain, "Comparison of SIFT encoded and deep learning features for the classification and detection of esca disease in bordeaux vineyards," *Remote Sensing*, vol. 11, no. 1, pp. 1–26, Dec. 2018, doi: 10.3390/rs11010001.

[38] A. -A. Liu, N. Xu, W. -Z. Nie, Y. -T. Su, Y. Wong, and M. Kankanhalli, "Benchmarking a multimodal and multiview and interactive dataset for human action recognition," *IEEE Transactions on Cybernetics*, vol. 47, no. 7, pp. 1–14, Jul. 2017, doi: 10.1109/TCYB.2016.2582918.

BIOGRAPHIES OF AUTHORS

Thirthe Gowda Mallinathapura Thimmegowda received B.E degree (Branch: IS & E, College: BTLITM, Bangalore, Karnataka, India), M.Tech degree (Specialization: Software Engineering, College: SJCE, Mysore, Karnataka, India in the years 2004 and 2006 respectively from Visvesvaraya Technological University (VTU), Belagavi, Karnataka, India. He has 18 years of teaching experience in reputed organizations at graduate and post graduate level. He is currently working as Assistant Professor, Department of Computer Science and Engineering, Government Engineering College, Hassan, Hassan district, Karnataka, India. He has published papers in various national and international journals and conferences. His areas of interest include image and video processing, data mining, natural language processing, machine learning and data analytics. He can be contacted at email: thirthegowda.gech@gmail.com.

Chandrika Jayaramaiah received B.E degree (Branch: CS &E, College: MCE, Hassan, University: Mysore, Karnataka, India), M.Tech degree (Specialization: Software Engineering, College: SJCE, Mysore, University: Mysore, Karnataka, India) and Ph.D. degree (University: VTU, Belagavi, Karnataka, India) in the years 1991, 1997 and 2014 respectively. She has more than 30 years of teaching experience in reputed organizations at graduate and post graduate level. She is currently working as Professor, Department of Computer Science and Engineering, Malnad College of Engineering, Hassan, Karnataka, India. She has published papers in various national and international journals and conferences. Her current research interests include image processing, data mining, miming stream data, designed novel algorithms for classification and clustering of stream data. She can be contacted at email: jc@mcehassan.ac.in.