

# Medical, Aromatic, and Narcotic Plants Classification using an Artificial Neural Network

Margret E. Abdel Malek  
A Researcher in Faculty  
of Engineering -Electrical  
Department- Fayoum  
University  
[me1804@fayoum.edu.e](mailto:me1804@fayoum.edu.e)

Dr./ Rania A. Abuelsoud  
prof., Electrical  
Engineering Department,  
Fayoum University  
[raa00@fayoum.edu.eg](mailto:raa00@fayoum.edu.eg)

Dr./ Ahmed A. Nashat  
Associate prof., Electrical  
Engineering Department,  
Fayoum University  
[aan01@fayoum.edu.eg](mailto:aan01@fayoum.edu.eg)

## **Abstract**

Medical, Aromatic, and Narcotic plants are a natural treasure that grows in the desert without human being interference. They can be used in pharmaceutical industries (medicines), medical usage (medical anesthetic), perfumes industries, and cooking. Thus, they are very useful, available, and can be utilized for the sake of human beings. On the other hand, some of these plants are harmful to our bodies and must be strictly prohibited. So, it is necessary to design and implement an image processing system to detect these plants. This system can be applied by the Ministry of Agriculture and Armed Force. After surveying deserts and taking photos of plants by a small camera attached to a drone, they can be inserted into the system to detect the type of captured plant and take action. In this paper, an automatic computer vision system is proposed to identify six types of desert plants. First, a nine-class collected database is prepared. Second, an artificial neural network-based framework, which uses color, DWT, the ratio between the major and the minor axes of the plants, and Tamura statistical texture features, is employed to classify plants. Outcomes and the results of the suggested system have competed with several techniques such as the SVM, the Naive Bayes, the KNN, the decision tree, and discriminant analysis classifiers. Results reveal that the proposed system has the highest overall recognition rate, which is 94.3%, among other techniques.

## **Keywords**

Image segmentation, Features extraction, Medical Plants, Aromatic Plants, Narcotic Plants, DWT, Plants Classification, Computer vision, Artificial neural network.

## **1.INTRODUCTION:**

Medical, Aromatic, and Narcotic plants are very important because of their wide usage in pharmaceutical, medical, perfumes, and food industries. We prepared a database for this program. The database comprises important six types of plants. These are Calotropis, Cannabis, Chamomile, Marjoram, Poppy, and Rosemary. Then we implemented a MATLAB program to classify these plants. We established this program with this set of plants as a start but we can develop this set later. We made this program by extracting the plants' features which vary between texture, colour, oval shape axes ratio surrounded by plant body, and image decomposition by discrete wavelet analysis. This program is implemented by an artificial neural network. This system is done to detect, recognize desert plants, and give their names and it can be used to survey desert, and by those who work in agriculture, medical industries, and others.

## **2.RELATED RESEARCH:**

In the literature, a lot of methods, which depend on the shape, size, colour, and texture features, have been developed for plant classification. There are some works related to this field. Herdiyeni and Wahyuni (2012) used a fusion of fuzzy local binary pattern and fuzzy colour histogram and a probabilistic neural network (PNN) classifier on a dataset of 2448

foliage images (270 \*240 pixels) obtained from medicinal plants from the Indonesian forests to achieve a classification accuracy of 74.5% [1]. Prasvita and Herdiyeni (2013) improved an equivalent mobile application based on former research [2]. Sourcing the kernel descriptor (KDES) as a new characteristic for the extraction method, Le et al. (2014) developed a fully automated plant identification system [3]. The proposed technique was verified on a dataset of 55 medicinal plants from Vietnam and a very high correctness of 98.3% was achieved with a support vector machine (SVM) classifier. Moreover, their algorithm accomplished an exactness of 98.5% on the Flavia dataset, which is the best result printed so far on this dataset [3]. Using the discrete wavelet transform to extract translation-invariant features from a collection of 8 different ornamental plants in Indonesia, Arai et al. (2013) achieved an accuracy of 95.8% using a support vector machines (SVM) classifier [4]. The size of each image was 256\*256 pixels. Du et al. (2013) proposed an approach based on fractal dimension features based on leaf shape and vein patterns for the recognition and classification of plant leaves [5]. Manipulating a k-nearest neighbour classifier with 20 features, they were able to reach a high recognition rate of 87.1%. Using a volumetric fractal dimension methodology to produce a texture signature for a leaf and the Linear Discriminant Analysis (LDA) algorithm, Backes et al. (2009) were able to beat traditional approaches that were based on Gabor filters and Fourier analysis [6]. Using a k-nearest neighbour (KNN) classifier, Munisami et al. (2015) achieved an accuracy of 87.3% on a dataset of 640 leaves taken from 32 different plant species [7]. They used the shape and colour information only. The images were gotten using a smartphone camera with a resolution of 1980\*1024. Hernandez-Serna and Jimenez-Segura (2014) reached an accuracy level of 92.9% using the Flavia dataset [8]. Sixteen inputs (6 geometrical, 8 texture, and 2 morphological features) were fed to an artificial

neural network (ANN) with 60 nodes in the hidden layer and a learning rate of 0.1 over 50000 generations. Using the same dataset, Chaki et al. (2015) achieved an overall accuracy of 97.6% using a Neuro-Fuzzy classifier (NFC) with a 44-element texture vector and a 3-element shape vector [9]. Using shape features only on the Flavia dataset and Pattern Net (a flavour of neural network), Siravenha and Carvalho [10] reached a similar accuracy as Chaki et al. [9]. Their feed-forward neural network had two hidden layers with 26 neurons in each one and it was trained over 100 epochs. Interesting work was done by Carranza-Rojas and Mata-Montero (2016) in which they created two datasets: a clean one and a noisy one [11]. They applied the Histogram of Curvature over Scale (HCoS) algorithm to get contour information and the local binary pattern variance (LBPV) to extract texture information. In the best situation, the clean dataset outperformed the noisy dataset by only 7.3%. This suggests that images taken directly using a smartphone can produce satisfactory levels of accuracy compared with images that are manually processed in a lab and then classified. Earlier, Amin and Khan (2013) have used a distributed hierarchical graph neuron (DHGN) to take curvature information using 64 feature vectors and the k-nearest neighbour classifier with Canberra distance to obtain an accuracy of 71.5% [12].

In working with plant identification that is based on the leaf images, the most difficult part is to represent the leaf with the robust features. Wu et. Al. [13] focused on extracting 5 basic geometric features of a leaf which contain the following: diameter, length, width, area, and perimeter. Based on these geometric features, they also came up with 12 morphological features that can be used for leaf recognition. The trial result shows that his algorithm is effective with an accuracy greater than 90% on 32 kinds of plants. Hati et. G. [14] made plant leaf recognition using image processing on an artificial neural network. In describing the

structure and shape of the leaf, they make use of aspect ratio, circularity, Base angle, Apex ratio, Apex angle, Width ratio, and Moment ratio. They did a recognition accuracy of 92%. Kadir et. al. [15] in their research combined shape and vein, colour, and texture features to classify a leaf. A neural network called Probabilistic Neural Network (PNN) was used as the classifier and the result gives 93.75% accuracy.

In this paper, we propose a novel ANN adapted to fine-grained feature-based representation for classifying different species of desert plants that can be used in different fields. To show the power of the proposed classification system, we collected from the internet a database, comprising of 520 images of the six-selected -plants with different directions. The suggested pattern makes use of the Tamura’s contrast, coarseness, and the direction visual perceptual texture features to correct labelling the data [16],[17] & [18]. The rest of the paper is ordered as follows: Sec. 3 illustrates an overview of the proposed system for classifying the selected plants and their shapes. In Sec. 4, the algorithm steps begin with images acquisition. Then, these images go through preprocessing operations to be ready for analysis (crop and segmentation). In Sec. 5, extracting features which consist of measured values (for each shape of the plant) to be used in the feature vector F. In Sec. 6, an explanation for the medical, aromatic, and narcotic plants classification algorithm, how it works, and its block diagram. In Sec. 7, the proposed classification system is tested and compared against other supervised classification algorithms, such as SVM, Naive Bayes, KNN, decision tree, and discriminant analysis classifiers then analyzed. Finally, Sec. 8 offers a conclusion.

### 3. THE PROPOSED MEDICAL, AROMATIC, AND NARCOTIC PLANTS CLASSIFICATION SYSTEM:

we propose in this article, a classification system for Medical, Aromatic, and Narcotic Plants that perform the visual perception of the human operator in making decisions at a very fast rate. A computer vision system for Medical, Aromatic, and Narcotic Plants is realized generally based on several processes. The head part for computer vision machines is the image segmentation tracked by image classification. An artificial neural network is then used for plant classification decision, which depends upon colour, Tamura statistical texture features, DWT, and major/minor axes ratio. Fig 1 shows some of the used samples in the proposed system.



Fig. 1: Samples of our concerned plants

#### 4.IMAGES ACQUISITION, CROP, AND SEGMENTATION:

##### **4.1 Images Acquisition:**

A database of medical, aromatic, and narcotic plants is collected from different internet sites. The database is comprised of 520 images, with different shapes. the training dataset is 70% of the total dataset, while the testing dataset is around 30%.

##### **4.2 Image Crop:**

Our concerned plants consist of many forms. One of them is found separately as flower-like Chamomile. Another kind is Rosemary which consists of small sticks assembled in a bush. Calotropis has 3 forms (flower, fruit & leaf) assembled in a small tree and Poppy has 2 forms (fruit & flower). But Cannabis & Marjoram possess leaves of different shapes. In this stage, we concentrate on the plant itself by cropping it then segmentation by MATLAB.

##### **4.3 Image Segmentation:**

This module has two phases. First, we extract the plant from the background. The segmenting of the plant is achieved by identifying a polygonal mask created interactively and manually. MATLAB reads the image and displays it, along with directions to specify the initial contour location. We specify the contour for the plant manually then create the polygonal mask. Fig. 2 shows the mask steps that are used for the segmentation.



(a) (b) (c)

Fig. 2: The Segmentation Mask.

(a) represents the initial image.

(b)image after determining the manual contour.

(c)binary mask.

Second, a function is applied to re-establish the initial input plant image. This is done by

reading the created binary mask image pixel by pixel. A pixel is recognized as a background if its level value is zero. Otherwise, equate the level value of that pixel with the initial one.



Fig. 3: The Final Segmented Plant.

The training dataset is attached to this link: [https://drive.google.com/drive/folders/1mZI6GFeRbZcLpTNIVy3YjV\\_lxIFBz5zV](https://drive.google.com/drive/folders/1mZI6GFeRbZcLpTNIVy3YjV_lxIFBz5zV).

The dataset has 152 of Calotropis (flowers, fruits& leaves),76 of Cannabis,71 of Chamomile,72 of Marjoram,106 of Poppy (flowers & fruits),43 of Rosemary. They are all in RGB colour mode.

#### 5.EXTRACTING FEATURES:

##### **5.1 Colour:**

Intensity matrices for the R, G, and B colour system of all images in the dataset are calculated. Each image has three matrices, one for each colour component. The mean of each colour component for each image is calculated and is considered as a feature [AVGR2, AVGG2, AVGB2].

##### **5.2 Tamura Statistical Texture Feature:**

texture-based methods assist in surface, shape, and plant class determination. Statistical texture analysis is the most popular method used for image recognition. Tamura contrast, coarseness, and direction features are three quantitative second-order statistical texture measures for image classification. They are built on psychological analyses of human observation so they are 3 feature vectors in our research.

**5.2.1 Coarseness** mainly connects with the distance in grey levels of spatial variants, which is completely connected with the size of the original components making the texture. It has the exact association to measure and replication rates and the most essential texture

characteristic. An image will include repetitive textures model at several scales, coarseness objects to recognize the largest size at which a texture is existent, even where a smaller micro-texture is:

$$A_K(x, y) = \sum_{i=x-2^{k-1}}^{x+2^{k-1}-1} \sum_{j=y-2^{k-1}}^{y+2^{k-1}-1} \frac{f(i, j)}{2^{2k}} \quad (1)$$

Where  $2^k * 2^k$  size is the mean of neighbourhood

$$E_{k,h}(x, y) = |A_k(x + 2^{k-1}, y) - A_k(x - 2^{k-1}, y)| \quad (2)$$

This equation (2) computes the variance between pairs of averages equivalent to non-overlying neighbourhoods. [Fcoarseness]

### 5.2.2 Contrast

calculates the distribution of grey degrees that differs in an image and to what limit its pervasion is biased to black or white. The second-order and normalized fourth-order central moments of the grey degrees are used to state the contrast.

$$\text{Contrast} = \sigma / (\alpha_4)^n \quad (3)$$

$$\alpha_4 = \mu_4 / \sigma^4$$

where  $\mu_4$  is the 4th instant about the mean and  $\sigma^2$  is the variance  $n=1/4$  to give the nearest value according to Tamura.[16] [Fcontrast]

### 5.2.3 Directionality

is achieved by checking the sharp degree of a histogram which is composed of the gradient vectors of all the image pixels. It is given by the following equation:

$$\text{Directionality} = \frac{1 - r n_p}{\sum_{p=1}^{n_{peaks}} \sum_{a \in w_p} (a - a_p)^2 H_{\text{Directionality}}(a)} \quad (4)$$

where,  $n_p$ , number of peaks,  $a_p$ , is the point of the peak,  $w_p$ , is the range of the angles ascribed to the Pth peak,  $r$  refers to a normalizing factor linked to quantizing levels of the angles  $a$ , and  $a$  refers to quantized directional angle,

$H_{\text{Directionality}}$ , is the histogram of quantized direction values,  $a$  is formed by counting the number of the edge pixels with the equivalent directional angles [16] [Fdirection].

### 5.3 Discrete Wavelet Transform (DWT):

Using the 2-dimensional DWT (DWT2) to decompose an image will produce 4 sets of coefficients: one is a low-pass component of the image (called approximation coefficients(a)), and the other 3 sets are high-pass components (called details coefficients: horizontal(H), vertical (V), and diagonal (D)). Most information is in the high-pass components. Hence, we can decompose the image by selecting the HP coefficients. Depending upon the size of the image, the process of DWT decomposition can be repeated many times. For our database, it is eight times. Finally, 3 coefficients are gotten (horizontal(H), vertical (V), and diagonal (D)). Another function is implemented to take the average root square of the sum square of these coefficients (AV) which is considered as a feature in our proposed system [AV].

$$AV = \sqrt{(H^2 + V^2 + D^2)} \quad (5)$$

### 5.4 Major/Minor Axes Ratio:

value is given by the major axis length of the oval shape that surrounds the plant over the minor axis of the same shape. These lengths are measured by pixels. Figure 4 shows the Major axis and Minor axis. The ratio is considered as a feature vector [B].

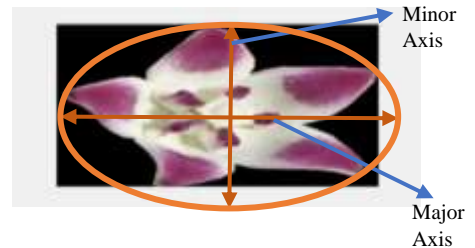


Fig. 4: The Major and Minor axes

## 6. MEDICAL, AROMATIC, AND NARCOTIC PLANTS CLASSIFICATION ALGORITHM:

Figure 5 shows a block diagram for the

proposed classifier. The Levenberg–Marquardt backpropagation optimization algorithm is used for training the recommended artificial neural network. The input layer of the pattern is made up of eight neurons. These are the quality indices of the vector  $F$ , defined by Eq.6 which consists of different measured values taken from each plant class. The model has ten hidden layers. The sigmoid function is worked as the activation function due to its easiness in derivative and its soft-switching ability. The number of neurons of the output layer is nine, which denotes the nine plants class. The output layer generates the real output of the neural network which is then subtracted from the required output (target) to generate the error. The mean squared error is used as a performance function and is directed back into the network hidden layers to renew the weighted sum of the input and bias of each neuron. In this system, the training dataset is 70% of the total dataset, while the testing dataset carried the remaining percentage. Therefore, out of 520 images as the real dataset, 107 (Calotropis-flowers, fruits& leaves), 53 (Cannabis), 49 (Chamomile), 49 (Marjoram), 74 (Poppy-flowers & fruits), and 30 (Rosemary) are used as the training group, while the residual 45 (Calotropis-flowers, fruits& leaves), 23 (Cannabis), 22 (Chamomile), 23 (Marjoram), 32 (Poppy-flowers & fruits) and 13 (Rosemary) are used as the testing set.

Calculate the texture feature vector,  $F$ , for each training plant image.

$$F = [AVGR2, AVGG2, AVGB2, Fcoarseness, Fcontrast, Fdirection, AV, B] \quad (6)$$

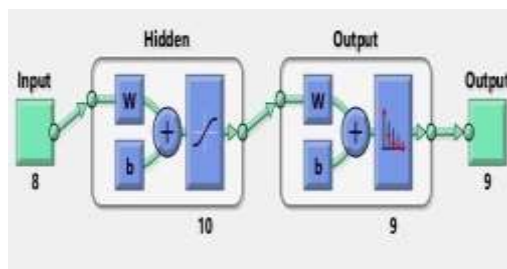


Fig.5 Architecture of the proposed Artificial Neural Network

## **7.EXPERIMENT RESULTS AND DISCUSSION:**

To test the performance of the plant classification system, 158 plants of assorted colours and shapes are selected for testing. Table 1 shows the results of the feature vector ( $F$ ), described by Eq. 6, for the 362 trained plant images. All these features combine to determine which class the image does belong to Table2. which shows the confusion matrix for the suggested artificial neural network for the Medical, Aromatic, and Narcotic Plants classification model. The model achieves a 94.3% overall correctness of the recognition rate. Tables 3, 4, 5, 6, and 7 show the confusion matrices for some other supervised classification algorithms. In this work, the performance of the SVM, the Naive Bayes, the k-nearest neighbors, the decision tree, and the discriminant analysis classifiers are compared with the ANN classifier. Simulation results show that the overall correctness of the model classification accuracy and the class sensitivity of the SVM, the Naive Bayes, the KNN, the discriminant analysis, and the decision tree classifiers are less than the proposed artificial neural network classification system.

Table1.shows the result of the vector (F) for each class

Plants Features	Calotropis Flower (1)	Calotropis Fruit (2)	Calotropis Leaf (3)	Cannabis (4)	Chamomile (5)	Marjoram (6)	Poppy-flower (7)	Poppy-fruit (8)	Rosemary (9)
	Color (Red)	55-122	48-127	33-122	20-111	83-165	47-117	76-164	63-124
Color (Green)	35-107	62-142	38-119	23-141	82-165	80-146	3-86	75-142	22-83
Color (Blue)	24-111	48-106	38-107	8-84	67-134	10-106	4-90	38-125	16-67
Coarseness	30-37	37-44	30-40	7-31	32-40	23-42	32-41	36-44	23-36
Contrast	42-94	37-87	33-70	10.2-79	81-112	36-105	21-78	52-87	30-73
Directivity	0.15-0.39	0.31-0.44	0.31-0.48	0.25-0.5	0.31-0.4	0.09-0.6	0.19-0.48	0.34-0.49	0.35-0.57
DWT (Average)	0-200	0-200	0-124	40-150	100-250	50-160	40-100	0-37	2-90
(Major/Minor)	1.01-1.4	1.1-1.6	1.06-2.2	3.5-7.2	1.04-1.9	1.1-2.6	1-2.6	1.05-1.49	1.2-6

Table 2. Confusion matrix for the suggested artificial neural network

Actual Class	Predicted class									Class Sensitivity%
	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8	Class 9	
Class 1	14	0	0	0	0	0	0	0	0	100
Class 2	0	15	0	0	0	1	0	0	0	93.8
Class 3	0	1	13	0	0	1	0	0	0	86.7
Class 4	0	0	0	23	0	0	0	0	0	100
Class 5	0	0	0	0	22	0	0	0	0	100
Class 6	1	0	0	1	0	21	0	0	0	91.3
Class 7	2	0	0	0	0	0	16	0	0	88.9
Class 8	1	0	0	0	0	0	0	13	0	92.9
Class 9	0	0	0	0	0	1	0	0	12	92.3
Class Precision %	77.8	93.8	100	95.8	100	87.5	100	100	100	Overall correctness= 94.3%

Table 3. Confusion matrix for the SVM classifier

Actual Class	Predicted class									Class Sensitivity%
	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8	Class 9	
Class 1	14	0	0	0	0	0	0	0	0	100
Class 2	0	16	0	0	0	0	0	0	0	100
Class 3	0	0	11	0	2	2	0	0	0	73.3
Class 4	0	0	0	23	0	0	0	0	0	100
Class 5	0	0	0	0	21	0	0	1	0	95.5
Class 6	0	0	1	1	0	20	0	1	0	87.0
Class 7	1	0	0	0	0	0	17	0	0	94.4
Class 8	0	0	0	0	0	0	0	14	0	100
Class 9	0	0	0	1	0	1	0	0	11	84.6
Class Precision %	93.3	100	91.7	92	91.3	87	100	87.5	100	Overall correctness = 93.0%

Table 4. Confusion matrix for the discriminant analysis classifier

Actual Class	Predicted class									Class Sensitivity%
	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8	Class 9	
Class 1	14	0	0	0	0	0	0	0	0	100
Class 2	0	10	2	0	0	4	0	0	0	62.5
Class 3	0	0	12	0	1	2	0	0	0	80.0
Class 4	0	0	0	23	0	0	0	0	0	100
Class 5	0	0	0	0	21	0	0	1	0	95.5
Class 6	0	0	0	0	1	21	0	1	0	91.3
Class 7	4	0	0	0	0	0	14	0	0	77.8
Class 8	0	0	2	0	0	0	0	12	0	85.7
Class 9	0	0	0	1	0	0	0	0	12	92.3
Class Precision %	77.8	100	75.0	95.8	91.3	77.8	100	85.7	100	Overall correctness = 88.0%



Table 5. Confusion matrix for the KNN classifier

Actual Class	Predicted class									Class Sensitivity%
	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8	Class 9	
Class 1	13	0	1	0	0	0	0	0	0	92.9
Class 2	0	9	2	1	0	4	0	0	0	56.3
Class 3	0	4	5	0	2	4	0	0	0	33.3
Class 4	0	0	0	23	0	0	0	0	0	100
Class 5	0	0	0	0	21	0	0	1	0	95.5
Class 6	0	2	1	4	0	16	0	0	0	69.6
Class 7	5	0	0	0	0	0	13	0	0	72.2
Class 8	0	2	1	0	0	0	0	11	0	78.6
Class 9	0	0	2	2	0	1	0	0	8	61.5
Class Precision %	72.2	52.9	41.7	76.7	91.3	64.0	100	91.7	100	Overall correctness = 75.3%

Table 6. Confusion matrix for the Naive Bayes classifier

Actual Class	Predicted class									Class Sensitivity%
	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8	Class 9	
Class 1	12	0	1	0	0	0	0	0	1	85.7
Class 2	0	7	3	0	0	6	0	0	0	43.8
Class 3	0	0	7	0	0	8	0	0	0	46.7
Class 4	0	0	0	23	0	0	0	0	0	100
Class 5	0	0	0	0	22	0	0	0	0	100
Class 6	0	1	3	0	0	18	0	0	1	78.3
Class 7	5	0	1	1	0	0	10	0	1	55.6
Class 8	0	0	5	0	0	0	0	9	0	64.3
Class 9	0	0	0	2	0	1	0	0	10	76.9
Class Precision %	70.6	87.5	35.0	88.5	100	54.5	100	100	76.9	Overall correctness = 74.7%

Table 7. Confusion matrix for the decision tree classifier

Actual Class	Predicted class									Class Sensitivity%
	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8	Class 9	
Class 1	11	0	2	0	0	0	0	0	1	78.6
Class 2	2	9	0	0	0	5	0	0	0	56.3
Class 3	2	0	4	0	0	9	0	0	0	26.7
Class 4	0	0	2	21	0	0	0	0	0	91.3
Class 5	2	0	0	0	20	0	0	0	0	90.9
Class 6	0	2	1	0	0	19	0	0	1	82.6
Class 7	4	0	0	0	0	0	14	0	0	72.8
Class 8	0	0	8	0	0	0	0	6	0	42.9
Class 9	0	0	6	2	0	1	0	0	4	30.8
Class Precision %	52.4	81.8	17.4	91.3	100	55.9	100	100	66.7	Overall correctness = 68.4%

Sensitivity and precision are two statistical performing methods for classification tests. Sensitivity is defined as the capability of the prediction pattern to pick the instance of a certain class from the dataset. It is the proportion of the actual positive classes which are correctly classified. On the other hand, precision is defined as the proportion of the predicted positive classes which are correctly classified. They are given by:

$$\text{Sensitivity} = \frac{TP}{(TP+FN)} \quad (7)$$

$$\text{Precision} = \frac{TP}{(TP+FP)} \quad (8)$$

Where TP, FP, and FN are the numbers of the true positive, the false positive, and the false-negative predictions for the studied class, in turn. Our suggested neural network has a sensitivity of 100% for classes 1, 4, and 5. Then the sensitivity is dropping gradually to reach 93.8% in 2nd class (Calotropis-fruit) as a test sample is misclassified as 6th class (Marjoram). It reaches 92.9% in the 8th class (Poppy-fruit). Also, it degrades to 92.3% in the 9th class (Rosemary). It is dropping to 91.3%

in the 6th class (Marjoram) and 88.9% in the 7th class (Poppy-flower). Finally, it reaches 86.7% in the 3rd class (Calotropis-leaf) that is the least sensitivity value in the suggested system. This is because a Calotropis-leaf sample is misclassified as Marjoram, and another sample of Calotropis-leaf is also misclassified as Calotropis- fruit.

Specificity and overall prediction accuracy for the class are other important statistical methods that clarify the execution of a classifier. They are used for binary classifiers. Specificity is described as the proportion of real negative classes (all except the positive class) that are correctly detected, while overall class prediction accuracy is defined as the proportion of the total number of correct predictions. Negative Predictive Value is defined as the proportion of negative classes that were correctly identified. They are given by:

$$\text{Specificity} = \frac{TN}{(TN+FP)} \quad (9)$$

$$\text{Overall Prediction Accuracy} = \frac{TP+TN}{(TP+TN+FP+FN)} \quad (10)$$

$$\text{Negative Predictive Value} = \frac{TN}{(TN + FN)} \quad (11)$$

To demonstrate the performance of the suggested ANN classification model, the (9×9) confusion matrices, shown in table 2., are transformed into four (2×2) confusion matrices. Tables 8 to 16 show the nine binary confusion matrices, each for our 9 plant classes, for the ANN, respectively. Tables 8 to 16, also, show that the overall class prediction accuracy for the proposed ANN. The classification of the plants is highly correct and true and the misclassification rate, which is the complement of the prediction accuracy, is minimum. Overall prediction accuracy for each class is 97.47%, 98.73%, 98.73%, 99.37%, 100%, 96.84%, 98.73%, 99.37%, and 99.37% respectively.

Tables 8 to 16 show the Binary confusion matrix for the proposed artificial neural network

Table8. Class (1): Calotropis-flower

Actual Class	Predicted Class		
	Class 1	Not Class 1	
Class 1	TP = 14	FN = 0	Sensitivity = 100%
Not Class 1	FP = 4	TN = 140	Specificity = 97.22%
	Precision = 77.8%	Negative predictive Value = 100%	Overall prediction accuracy = 97.47%

Table9. Class (2): Calotropis-fruit

Actual Class	Predicted Class		
	Class 2	Not Class 2	
Class 2	TP = 15	FN = 1	Sensitivity = 93.8%
Not Class 2	FP = 1	TN = 141	Specificity = 99.30%
	Precision = 93.8%	Negative predictive	Overall prediction

	Value = 99.30%	accuracy = 98.73%
--	----------------	-------------------

Table10. Class (3): Calotropis-leaf

Actual Class	Predicted Class		
	Class 3	Not Class 3	
Class 3	TP = 13	FN = 2	Sensitivity = 86.7%
Not Class 3	FP = 0	TN = 143	Specificity = 100%
	Precision = 100%	Negative predictive Value = 98.62%	Overall prediction accuracy = 98.73%

Table11. Class (4): Cannabis

Actual Class	Predicted Class		
	Class 4	Not Class 4	
Class 4	TP = 23	FN = 0	Sensitivity = 100%
Not Class 4	FP = 1	TN = 134	Specificity = 99.26%
	Precision = 95.8%	Negative predictive Value = 100%	Overall prediction accuracy = 99.37%

Table12. Class (5): Chamomile

Actual Class	Predicted Class		
	Class 5	Not Class5	
Class 5	TP = 22	FN = 0	Sensitivity = 100%
Not Class 5	FP = 0	TN = 136	Specificity = 100%
	Precision = 100%	Negative predictive Value = 100%	Overall prediction accuracy = 100%

Table13. Class (6): Marjoram

Actual Class	Predicted Class		
	Class 6	Not Class 6	
Class 6	TP = 21	FN = 2	Sensitivity = 91.3%
Not Class 6	FP = 3	TN = 132	Specificity = 97.78%
	Precision = 87.5%	Negative predictive	Overall prediction

	Value = 98.51%	accuracy = 96.84%
--	----------------	-------------------

Table14. Class (7): Poppy-flower

Actual Class	Predicted Class		
	Class 7	Not Class 7	
Class 7	TP = 16	FN = 2	Sensitivity = 88.9%
Not Class 7	FP = 0	TN = 140	Specificity = 100%
	Precision = 100%	Negative predictive Value = 98.59%	Overall prediction accuracy = 98.73%

Table15. Class (8): Poppy-fruit

Actual Class	Predicted Class		
	Class 8	Not Class 8	
Class 8	TP = 13	FN = 1	Sensitivity = 92.9%
Not Class 8	FP = 0	TN = 144	Specificity = 100%
	Precision = 100%	Negative predictive Value = 99.31%	Overall prediction accuracy = 99.37%

Table16. Class (9): Rosemary

Actual Class	Predicted Class		
	Class 9	Not Class 9	
Class 9	TP = 12	FN = 1	Sensitivity = 92.3%
Not Class 9	FP = 0	TN = 145	Specificity = 100%
	Precision = 100%	Negative predictive Value = 99.32%	Overall prediction accuracy = 99.37%

TP, TN, FN, and FP stand for true positive events, true negative events, false-negative events, and false-positive events.

Figure 6 shows another criterion to present the classification result which is ROC curves for the proposed ANN. the receiver operating characteristic (ROC) curve is a metric used to examine the quality of the classifier. It is the TP rate or the sensitivity against the FP rate or (1-specificity), which is the probability of false alarm, for the different possible cut-points of the test dataset. The proposed ANN classifier is

perfectly accurate for class 5 (olive) since the area under the ROC curve equals one. However, the area under the curve for the fourth class (lime) is 0.99, which is very close to one. Whereas class 6 (chocolate), is the least in specificity and sensitivity followed by class 3. Class 4&1 have a 100% sensitivity but in specificity 99.37% & 97.22% respectively. Class 8 ,9, 7&3 have a 100% specificity but sensitivity of 92.9%, 92.3%,88.9% & 86.7% respectively. Hence, the ANN model is better than all the former classifiers in differentiating between the right class and not the right class. The ANN model is the more perfect measure of separability between classes than the other supervised classification algorithms.

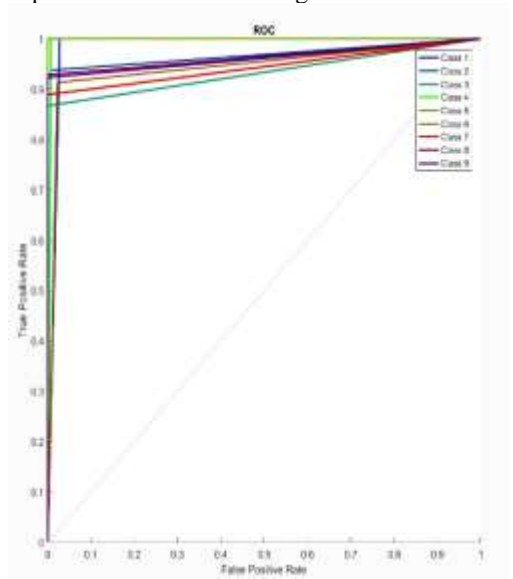




Fig.6 ROC of the proposed ANN classification model

Table17. color representation of classes

Classes	Color Name	Color
Class1(Calotropis-Flower)	blue	
Class2(Calotropis-Fruit)	teal	
Class3(Calotropis-Leaf)	green	
Class4(Cannabis)	lime	
Class5(Chamomile)	olive	
Class6(Marjoram)	chocolate	
Class7(Poppy-Flower)	red	

Class8(Poppy-Fruit)	red-violet	
Class9(Rosemary)	purple	

### **8. CONCLUSION:**

An artificial neural network-based system for the classification of Medical, Aromatic, and Narcotic Plants has been discussed. This research is unique as it is a mixture of features that varies among texture, color, axes length, and DWT. Also, the research includes a dataset that was collected randomly from the internet and cropped manually then segmented by a MATLAB function. All these tasks were performed by the research's authors. Otherwise, most researchers download a ready dataset and compose the algorithm or use a smartphone camera to take photos of samples to compose the dataset. It is considered that the ANN algorithm uses 10 neurons only in the hidden layer but achieves Overall correctness of 94.3%. similar researches concerned with leaf shape or leaves only. But this research combines different forms of plants like flowers, sticks, leaves, and fruits. Hati et. G. [14] made plant leaf recognition using image processing on an artificial neural network. In describing the structure and shape of the leaf, they make use of aspect ratio, circularity, Base angle, Apex ratio, Apex angle, Width ratio, and Moment ratio. They did a recognition accuracy of 92%. Kadir et. al. [15] in their research combined shape and vein, colour, and texture features to classify a leaf. A neural network called Probabilistic Neural Network (PNN) was used as the classifier and the result gives 93.75% accuracy. The excellency of this research appears in its ability to distinguish the plants at a high rate although it trained with 362 images which are considered a relatively small number. The proposed model uses Tamura's texture features, RGB color, DWT, and major/minor axes ratio properly to discriminate among nine classes of different shapes of Medical, Aromatic, and Narcotic Plants. The system has optimal performance as compared

with other supervised classification algorithms as the SVM, the Naive Bayes, the KNN, the decision tree, and the discriminant analysis classifiers. The overall class recognition accuracy of 100% is obtained by Chamomile, while it is 97.47% for Calotropis-flower, 98.73% for Calotropis-fruit, 98.73% for Calotropis-leaf, 99.37% for Cannabis, 96.84% for Marjoram, 98.73% for Poppy-flower, 99.37% for Poppy-fruit, 99.37% for Rosemary classes. The simplicity and the high recognition rate of the classification of 158 test samples, make it appropriate for implementing a productive and profitable computer vision machine for the pharmaceutical industry and agriculture. Future directions will focus on increasing the no. of training images and increasing the no. of plant species and using deep learning methods to supervised/semi-supervised machine learning algorithms. Increasing the no. of training images to enhance the classification rate and to minimize the misclassification of Medical, Aromatic, and Narcotic Plants application.

## References:

- [1] Y. Herdiyeni and N. K. S. Wahyuni, "Mobile application for Indonesian medicinal plants identification using Fuzzy Local Binary Pattern and Fuzzy Color Histogram," *2012 International Conference on Advanced Computer Science and Information Systems (ICACSIS)*, pp. 301-306, 2012, Depok, Indonesia.
- [2] Prasvita, Desta Sandya, and Yeni Herdiyeni. "Medleaf: mobile application for medicinal plant identification based on leaf image." *International Journal on Advanced Science, Engineering and Information Technology* 3.2, pp.5-8, 2013, Indonesia.
- [3] Le, Thi-Lan, Duc-Tuan Tran, and Van-Nam Hoang. "Fully automatic leaf-based plant identification, application for Vietnamese medicinal plant search." *Proceedings of the fifth symposium on information and communication technology*, pp. 146-154, 2014, Vietnam.
- [4] Arai, Kohei, Indra Nugraha Abdullah, and Hiroshi Okumura. "Identification of ornamental plant functioned as medicinal plant based on redundant discrete wavelet transformation." *International Journal of Advanced Research in Artificial Intelligence* 2.3, pp.60-64, 2013, Japan.
- [5] Du, Ji-xiang, Chuan-Min Zhai, and Qing-Ping Wang. "Recognition of plant leaf image based on fractal dimension features." *Neurocomputing* 116, pp.150-156, 2013, China.
- [6] Backes, Andre Ricardo, Dalcimar Casanova, and Odemir Martinez Bruno. "Plant leaf identification based on volumetric fractal dimension." *International Journal of Pattern Recognition and Artificial Intelligence* 23.06, pp.1145-1160, 2009, Brazil.
- [7] Munisami, Trishen, et al. "Plant leaf recognition using shape features and colour histogram with K-nearest neighbour classifiers." *Procedia Computer Science* 58, pp.740-747, 2015, Mauritius
- [8] Hernández-Serna, Andrés, and Luz Fernanda Jiménez-Segura. "Automatic identification of species with neural networks." *PeerJ* 2, e563, 2014, Colombia.
- [9] Chaki, Jyotismita, Ranjan Parekh, and Samar Bhattacharya. "Plant leaf recognition using texture and shape features with neural classifiers." *Pattern Recognition Letters* 58, pp. 61-68, 2015, India.
- [10] Siravenha, Ana Carolina Quintao, and Schubert R. Carvalho. "Exploring the use of leaf shape frequencies for plant classification." *2015 28th SIBGRAPI Conference on Graphics, Patterns and Images. IEEE*, pp. 297-304, 2015, Salvador.
- [11] Carranza-Rojas, Jose, and Erick Mata-Montero. "Combining leaf shape and texture for Costa Rican plant species identification." *CLEI Electronic journal* 19.1, pp.7-7, 2016, Costa Rica.
- [12] Amin, Anang Hudaya Muhamad, and Asad I. Khan. "One-shot classification of 2-D leaf shapes using distributed hierarchical graph neuron (DHGN) scheme with k-NN classifier." *Procedia Computer Science* 24, pp.84-96, 2013, Malaysia.

[13] S. G. Wu, F. S. Bao, E. Y. Xu, Y. Wang, Y. Chang and Q. Xiang, "A Leaf Recognition Algorithm for Plant Classification Using Probabilistic Neural Network," 2007 IEEE International Symposium on Signal Processing and Information Technology, pp. 11-16, 2007, Giza.

[14] Hati, Shayan, and G. Sajeevan. "Plant recognition from leaf image through artificial neural network." *International Journal of Computer Applications* 62.17, pp.8887, 2013, India.

[15] Kadir, Abdul, et al. "Leaf classification using shape, color, and texture features." arXiv preprint arXiv:1401.4447, 2013, Indonesia.

[16] Bagri, Neelima, and Punit Kumar Johari. "A comparative study on feature extraction using texture and shape for content based image retrieval." *International Journal of Advanced Science and Technology* 80.4, pp.41-52, 2015, *India*.

[17] Kebapci, Hanife, Berrin Yanikoglu, and Gozde Unal. , "Plant Image Retrieval Using Color, Shape and Texture Features," in *The Computer Journal*, vol. 54, no. 9, pp. 1475-1490, Sept. 2011 Turkey.

[18] Karmakar, Priyabrata, et al. "Improved tamura features for image classification using kernel based descriptors." 2017 International Conference on Digital Image Computing: Techniques and Applications (DICTA). IEEE, 2017, China

## تصنيف النباتات الطبية والعطرية والمخدرة باستخدام شبكة عصبية إصطناعية

النباتات الطبية والعطرية والمخدرة كنز طبيعي ينمو في الصحراء دون تدخل الإنسان فيمكن إستخدامها في الصناعات الدوائية (الأدوية) وأيضاً الإستخدام الطبي (التخدير الطبي) وصناعات العطور و الطبخ وبالتالي فهي مفيدة للغاية ومتاحة ويمكن إستخدامها من أجل البشر ومن ناحية أخرى ، فإن بعض هذه النباتات ضارة بأجسامنا ويجب منعها تماماً. لذلك من الضروري تصميم وتنفيذ نظام معالجة الصور لإكتشاف هذه النباتات حيث يمكن تطبيق هذا النظام من قبل وزارة الزراعة، فيعد مسح الصحاري والنقاط صور للنباتات بواسطة كاميرا صغيرة متصلة بطائرة بدون طيار يمكن إدخالها في نظام لاكتشاف نوع النبات الذي تم إنتقاطها وإتخاذ الإجراءات اللازمة. في هذا البحث ، تم اقتراح نظام رؤية حاسوبي آلي لتحديد ستة أنواع من النباتات الصحراوية. أولاً ، يتم إعداد قاعدة بيانات مجمعة من تسعة أنواع . ثانياً ، يتم استخدام إطار عمل قائم على الشبكة العصبية الإصطناعية ، والذي يستخدم اللون ، و (DWT(discrete wavelet transform) ، والنسبة بين المحور الرئيسي والثانوي للنباتات ، وخصائص نسيج تامورا الإحصائي ، لتصنيف النباتات ويتم تصنيفها باستخدام الشبكة العصبية الاصطناعية التي تستخدم خوارزمية Levenberg–Marquardt لتحسين الإنتشار العكسي في التدريب حيث تتكون طبقة المدخلات من ٨ خلايا عصبية وهي الخصائص التي تكون عبارة عن قيم مقياس مختلفة مأخوذة من كل فئة نباتية ويحتوي أيضا النموذج على ١٠ طبقات خفيه ويستخدم sigmoid function كدالة التنغيم، أما عن عدد الخلايا العصبية في الطبقة الناتجة (طبقة المخرجات) هو ٩ ، مما يدل على ٩ فئات (نبتة من ٣ أشكال مختلفة و نبتة من شكلين و ٤ نباتات تحمل شكل واحد لكل منها) حيث تولد طبقة المخرجات الناتج الحقيقي للشبكة العصبية والذي يتم طرحه بعد ذلك من الناتج المطلوب (الهدف) لتوليد الخطأ. تم تصميم هذا النظام ب ٥٢٠ صورة كمجموعة بيانات حقيقية (صور حقيقية مجمعة من شبكة الإنترنت)، تشكل مجموعة عينات التدريب ٧٠٪ من إجمالي مجموعهم ، بينما تحمل مجموعة عينات الإختبار النسبة المتبقية ويتم التأكد من كفاءة النظام المصمم وبعد ذلك تتم مقارنة النتائج وأداء النظام المصمم مع تقنيات مختلفة مثل :

(SVM (support vector machine) ، و Naive Bayes ، و KNN ، وشجرة القرار ، ومصنفات التحليل التمييزي. وتظهر النتائج أن النظام المقترح لديه أعلى معدل تمييز إجمالي ، وهو ٩٤,٣٪ ، من بين التقنيات أخرى.