Identification of philippine herbal medicine plant leaf using artificial neural network

Conference Paper · December 2017

DOI: 10.1109/HNICEM.2017.8269470

7 authors, including:

Renann Baldovino
De La Salle University

Anton Louise De Ocampo
De La Salle University

Ira Valenzuela
De La Salle University

Alvin B. Culaba
De La Salle University

Identification of Philippine Herbal Medicine Plant Leaf Using Artificial Neural Network

Robert G. de Luna, Renann G. Baldovino, Ezekiel A. Cotoco,
Anton Louise P. de Ocampo, Ira C. Valenzuela, Alvin B. Culaba, Elmer P. Dadios
Gokongwei College of Engineering, De La Salle University Manila, Philippines
robert_g_deluna@dlsu.edu.ph

***Abstract*—The study described in this paper consists of a system that involves image processing techniques to extract relevant features related to leaf in conjunction with using artificial neural network in order to detect and identify some Philippine herbal plants. Real samples of twelve different herbal medicine plant leaves are collected where each leaf are isolated in single image. Several features are extracted using techniques in image processing. With the artificial neural network acting as autonomous brain network, the system can identify the species of the herbal medicine plant leaves being tested. The system can also provide information about the diseases the herbal plant can cure.**

For the training, a features dataset of 600 images coming from 50 images per herbal plant are used. With the aid of Python, a neural network model with optimized parameters are established producing 98.16 % identification for the whole dataset. To evaluate the actual performance of the system, a separate 72 sample images of herbal plants are tested with the neural network model implemented in MATLAB. Experimental results demonstrate a 98.61 % accuracy of herbal plant identification.

***Index Terms*—feature extraction, image processing, artificial neural network, leaf identification**

I. INTRODUCTION

Plants have significant contributions towards human lives and play a predominant role in the well-being of the global population. They are one of the major sources of food, raw materials, medications and etc. Since then, different plants are well-known to be specifically used as remedy to some ailment or disease. Most of the people are aware of their importance and explore to gain more knowledge on how to use certain plant to cure certain illnesses. Until now, different plants including the herbal medicine plants pose a big impact to the health of every people around the world. According to the World Health Organization (WHO) in 2009, 80% of the people around the world still rely on botanical drugs or medicine. Recently, numerous promulgation of propagandas on the use of herbal medicine plants have been noted; that even health practitioners and other people in the field of science see their underlying values. And now, the challenge of exploring into deeper analysis of the herbal medicine plants and breakthroughs is posted with the advent of technologies being developed.

One of the problems arising related to herbal medicine plants are their applications and utilization. Though many are aware of the existence of different herbal medicine plants and their familiar applications; many are still unable to identify which are these herbal medicine plants from the vast diversity of plants in the environment; so as their noted and approved applications by health professionals. In addition, researchers in the field of botany, medicine, chemical structure analysis, and other related fields concerning plant studies are faced with the application of considerable effort on identifying plants. It is stated that plant identification demands extensive knowledge and uses complex terminology; even professional botanists need to take much time in the field to master plant identification [1]. In computer vision, despite many efforts [2–7], plant identification is still considered a challenging and unsolved problem.

The problem regarding the identification and familiarization of the plants, limited to the specified herbal medicine plant leaves, along with the seen necessity for the development of herbal medicine plant identification are addressed by the study; wherein the researchers came up with an idea of making a system that will aid the problem. The system can help the user to easily and correctively identify and know more about the particular herbal medicine plants together with their medical applications and accorded procedures to be done based on the existing literature released by authorized body of the government and health sectors.

In the system, the researchers provided a program that will serve as the tool for the user in order to distinguish specific herbal medicine plant leaves, with the aid of a camera installed along with the system to capture image of a leaf specimen. The captured image will be processed and analyzed using different techniques of image processing in order to obtain defined parameters. Various features consisting of more than 20 to 30 are used for recognition of plant [8-12]. For this study, several primary features will be used to obtain around 5 secondary features that will serve the purpose of identification. Many recent studies exist on plant classification and identification based on different plant features. However, to handle such features information, finding an efficient classification method has become an area of active study [13], [14-17]. The value of the parameters will act as the input in the artificial neural network. This is a computer program that has the ability to learn from examples and can thus also perform recognition of previously unseen patterns. A supervised multilayer perceptron (MLP) artificial neural network (ANN) will be used in this system. Training is carried out by presenting a succession of data records (the training set) to the network, each record containing data from a specimen or record of known identity.

978-1-5386-0912-5/17/$31.00 ©2017 IEEE

The resultant ability of the network to recognize previously unseen patterns is periodically tested using an independent "validation" dataset, also containing data records of known classes. For further information about the principles of ANNs, see [18] and [19]. The After identification, necessary display of result is presented. In this proposed work, experimental analysis was carried out with 12 herbal plant species approved by the Department of health (DOH) for medical applications which includes akapulko (acapulco), ampalaya (bittermelon), balbas-pusa (java tea) bayabas (guava), lagundi (five-leaved chase tree), malunggay (ben oil tree), niyog-niyogan (rangoon creeper), oregano, sambong (ngai camphor), tsaang-gubat (wild tea), ulasimang-bato (peperonia), and yerba buena (peppermint) [20]. Fifty leaves were taken from each plant species and clustered.

This study will not only address the problem arising from unawareness and/or misidentification of herbal medicine plant leaves, but more importantly, it will promote enhancement of identification schemes of herbal medicine plant leaves.

II. RELATED WORKS

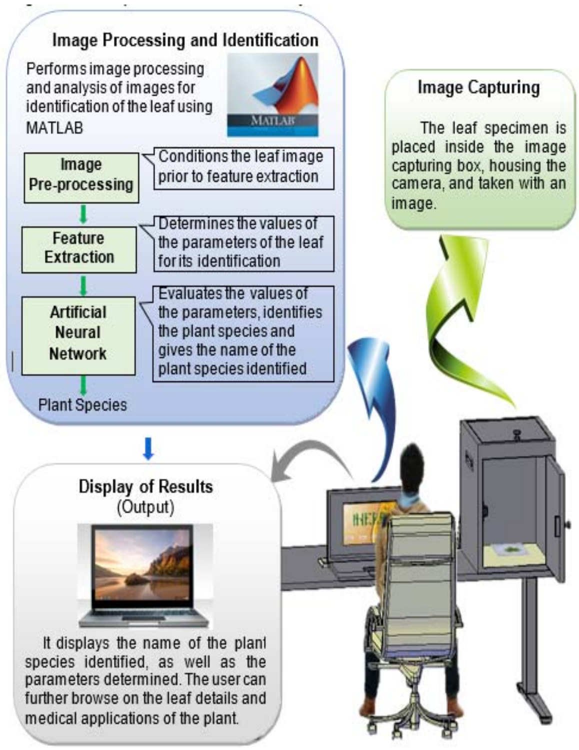
In working with plant identification that is based from the leaf images, the most difficult part is to represent the leaf with the robust features. Wu et. Al. [21] focused on extracting 5 basic geometric features of leaf which includes the following: diameter, length, width, area, and the perimeter. Based from these geometric features, they also came up with 12 morphological features that can be used for leaf recognition. Lin et. al. [22] computed basic descriptors like area, perimeter, length, width, and convex hull including some dimensionless shape factors like compactness, roundness, elongation, and roughness.

Hati et. al. [23] made a plant leaf recognition using a C++ -based image processing on artificial neural network. In describing the structure and shape of the leaf, they make use of aspect ratio, circularity, convexity, and solidity. Kadir et. al. [24] in their research incorporated shape and vein, color, and texture features to classify a leaf. A neural network called Probabilistic Neural Network (PNN) was used as the classifier. The use of neural networks for classification of medicinal plants was attempted in Basvaraj S.Anami et al. [25] and [26] in which they take the image of the entire plant and they consider the color, texture and edges. C. H. Arun et al. [27] classify medicinal plants based on texture features. Janani et al. [28] identify the medicinal plants based on leaf features using ANN. In this proposed work, experimental analysis was carried out to identify 12 Philippine herbal plant species using neural network in Python and MATLAB.

III. THE PROPOSED SYSTEM

The proposed system in the identification of herbal plant is composed of four fundamental building phases: **(1) Image acquisition phase**: This is the first phase of the system where digital image of the collected samples of herbal plant leaves are acquired. **(2) Pre-processing phase**: The quality of the acquired images should be improved to make the feature extraction more reliable. Several image processing techniques are applied. **(3) Feature extraction phase**: The 600 pre-processed images are subjected for feature extraction and represented in a database as vector values. **(4) Classification phase**: This is the last phase of the system and is implemented using the artificial neural network. The complete phase is shown in Figure 1.

Fig. 1. Conceptual Framework of the System for Identification of Philippine Herbal Medicine Plant Leaf



A. *Image Acquisition Phase*

An A4Tech web-camera with a resolution of 16M pixels was used for imaging. At a fixed distance of 37 cm, image was captured placing leaf on a white background.

B. *Image Preprocessing Phase*

Image pre-processing is necessary in order to enhance, smoothen, and remove some noise while capturing the images of the samples. This is done using MATLAB. Depending on how good is the pre-processing output, the efficiency of the classification and identification is affected by the generated features of the images.

C. *Feature Extraction Phase*

The extracted features are limited only to the parameters that had been considered in the study, including the length, width, area and perimeter of the original leaf, the area and perimeter of the convex hull image equivalent, and the leaf vein area which will dictate the morphological parameters including the aspect ratio, circularity, convexity, solidity and leaf vein density that defines the structure, shape and venation of the leaf.

D. Classification Phase

The classifier used in this study is the artificial neural network. The input to this phase are the training dataset feature vectors and their corresponding classes, whereas the outputs are the identification of what specific herbal plant. Using Python programming, datasets are analyzed and evaluated for different algorithms. With neural network as the main classifier, datasets are trained, implemented and tested establishing the optimized parameters to be used. It is then modelled in MATLAB where separated samples are tested.

IV. Experimental Results

The simulation experiments in this article are done using laptop with Python and MATLAB running on Windows 10. The datasets used for experiments are constructed based on real sample leaf images of 12 identified Philippine herbal medicine plants. These images were captured using a web-camera with 16M pixel resolution in jpg format. The dataset consisted of 12 categories pertaining to the specific herbal plant.

The camera's functionality is one of the most critical parts of the system since it will provide input images to the system for identification. The functionality of the camera depends on how is it placed in the capturing box, thus, adjustments on its height and/or position can be made in order to correct the error and prevent inaccuracies with the results. In testing the camera, as incorporated with the image acquisition, the researchers made use of the GUI to display the result both for background capturing and leaf image capturing. The captured images should be able to contain the whole leaf and send it directly to the computer with the aid of MATLAB GUI in displaying them. The GUI lay-out to be used in verifying the output of the image acquisition module is shown.

Fig. 2. GUI Layout for Camera Testing.

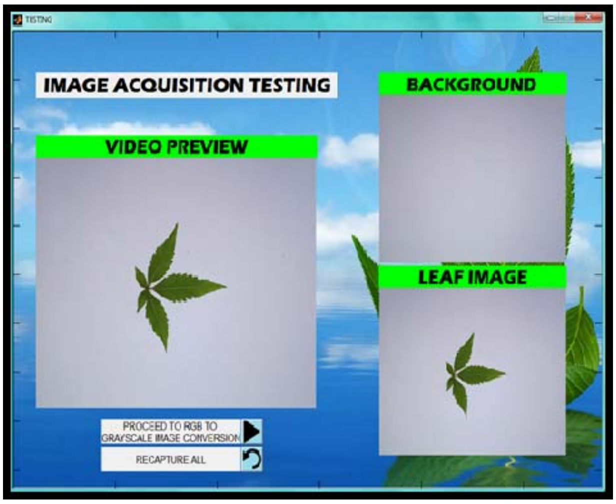


Figure 3 shows the sample leaf images of the identified herbal plants used in this study.

Fig. 3. Sample Images of the Herbal Plant Leaf

a. Akapulko b. Ampalaya c. Balbas-pusa

d. Bayabas e. Lagundi f. Malunggay

g. Niyog-niyogan h. Oregano i. Sambong

j. Tsaang-gubat k. Ulasimang-bato l. Yerba-buena

After the acquisition of leaf image, it will undergo the process of image preprocessing in order to condition the leaf image for feature extraction. It includes the RGB to grayscale conversion, thresholding, and image conversion to convex hull. In testing the functionality of the RGB to grayscale conversion module, the original leaf image and the converted image will be presented through the GUI. To verify the conversion process, the dimensions of the RGB image as well as the dimensions of the output grayscale image will be displayed. RGB images are three-dimensional images having three layers of matrix corresponding to three layers of its color components - red, green and blue. In contrast, the grayscale images are two-dimensional images having only length and width, defined also by the camera's resolution where the values of every component, are known as pixel. The GUI used is shown in Figure 4.

The grayscale image goes through thresholding making the image a component of black and white pixels. To verify the result of thresholding, as in Figure 5, every pixel will be monitored for the presence of only white and black pixels. The total number of accumulated black and white pixels will be compared to the total number of pixels of the image. If the total number of black and white pixels is not equal to total number of pixels of the grayscale image, the module poses errors as it indicated the presence of other pixels that are not

978-1-5386-0912-5/17/$31.00 ©2017 IEEE

converted to binaries 0 or 1; or some pixels are cut off due to coding schemes. If the two counts are equal, then the system achieves correct functionality.

Fig. 4. GUI Layout for RGB to Grayscale Image Conversion Testing

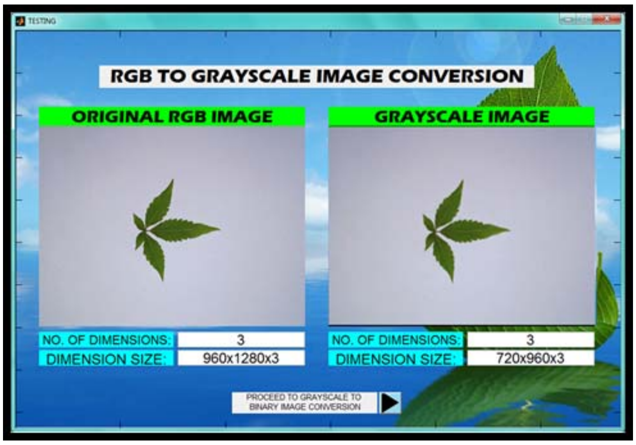
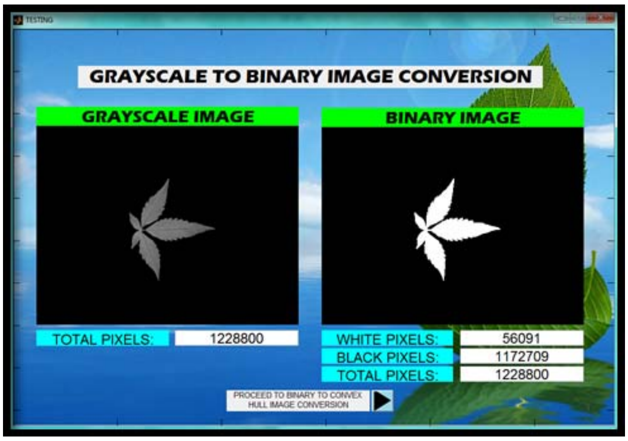


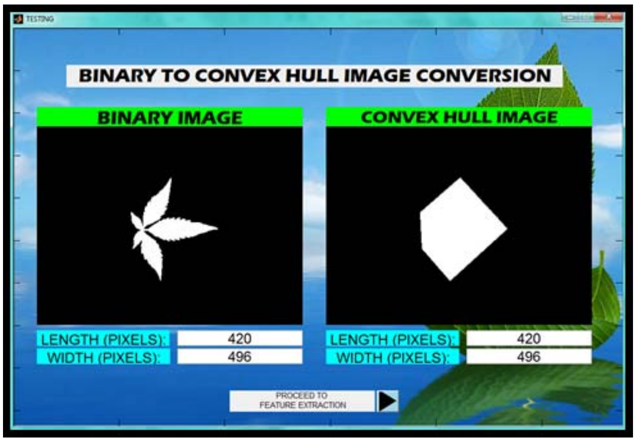
Fig. 5. GUI Layout for Binary to Convex Hull Image Conversion Testing



The convex hull conversion is tested by presenting the binary image and the convex hull image. To measure its functionality, the length and width of the binary image will be compared to the length and width of the convex hull image as presented in the GUI shown in Figure 6. If all these two parameters are equal, the system's convex hull conversion is functioning properly.

From the processed images, the feature extraction module determines the values of the parameters needed in identification of the plant leaf. This module consists of primary parameter extraction and secondary parameter determination. For the primary parameters, the binary image is extracted for the values of its length, width, perimeter and area while the convex hull image is extracted for the perimeter and area only. To do this, the length and width are measured in pixels by computing the difference of the maximum and minimum column having white pixel element and the difference of the maximum and minimum row having white pixel, respectively, as shown in Figure 6. The perimeter is determined by introducing another pre-processing; that is locating and setting the boundary of the white pixel area creating a border-outlined image; and then, counting the number of white pixels present. While the area is simply the total number of white pixels present in the binary image. The same process in getting the perimeter and area of the binary image is done in getting the perimeter and area of the convex hull image.

Fig. 6. GUI Layout for Binary to Convex Hull Image Conversion Testing



For the secondary parameters, these are defined as the following:

$$\text{Aspect Ratio}=\frac{\text{Width of Leaf Image}}{\text{Length of Leaf Image}} \tag{1}$$

$$\text{Circularity}=\frac{4\pi(\text{Area of Leaf Image})}{(\text{Perimeter of Leaf Image})^2} \tag{2}$$

$$\text{Convexity}=\frac{\text{Original Perimeter}}{\text{Perimeter of Convex Hull Image}} \tag{3}$$

$$\text{Solidity}=\frac{\text{Area of Leaf Image}}{\text{Area of Convex Hull Image}} \tag{4}$$

$$\text{Rectangularity}=\frac{\text{Length of Leaf Image} * \text{Width of Leaf Image}}{\text{Area of Leaf Image}} \tag{5}$$

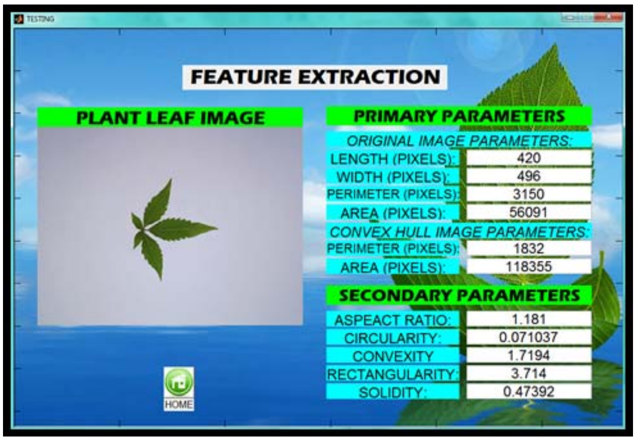
Figure 7 shows the GUI of the extracted primary and secondary parameters.

With all the features extracted, a dataset for 12 herbal plant leaf features are established. It is then subjected to some type of data analysis and performed using Python Version 3.5. Different algorithms are applied to the dataset as to analyze the classification learning rate. The algorithm includes Logistic Regression (LR), Linear Discriminant Analysis (LDA), K Nearest Neighbors (KNN), Classification and Regression Trees (CART), Naïve Bayes (NB), State Vector

978-1-5386-0912-5/17/$31.00 ©2017 IEEE

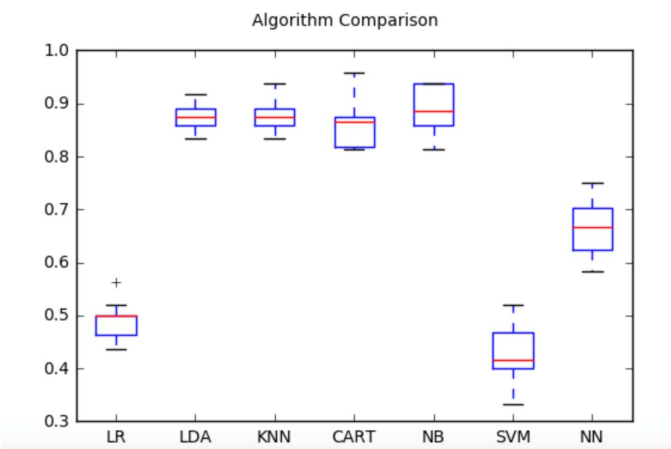
Machine (SVM), and the Neural Network (NN).

Fig. 7. GUI Layout for Feature Extraction Testing



Using the default parameters of different classifiers in Python, Figure 8 shows the result of mean accuracy and standard deviation of the assessment done. Naïve Bayes displayed high accuracy compared to all algorithms with State Vector Machine giving the least value. Since one of the objective of the study is to produce a learning mechanism using Artificial Neural Network, several analyses are made to improve the percent accuracy of the said algorithm.

Fig. 8. Accuracy and Standard Deviation Results of Different Algorithm



Dataset are subdivided into training set which comprised the 80 % of the dataset while the remaining 20 % is for the validation dataset. Using neural network as the classifier with the default value of its parameters, figures 9 to 10 depicted the accuracy score, confusion matrix, and classification report of the model as tested to the validation dataset and whole dataset respectively.

Fig. 9. ANN Result Using Default Parameters Tested in Validation Dataset

```
ACCURACY SCORE:
0.708333333333

CONFUSION MATRIX:
[[ 7  0  0  0  0  1  0  0  0  0  0  0]
 [ 0 15  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  4  0  0  0  0  0  7  0  0  0]
 [ 0  0  0  7  0  0  0  0  2  5  0  0]
 [ 0  1  0  0  8  0  0  0  0  0  0  0]
 [ 0  0  0  0  0 12  0  0  0  0  1  0]
 [ 3  0  1  0  0  0  2  0  1  1  0  0]
 [ 0  0  0  0  0  0  0  6  0  0  0  0]
 [ 0  0  2  0  0  0  0  0 10  2  0  0]
 [ 4  0  0  0  0  0  2  0  0  1  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  6  0]
 [ 0  0  0  0  0  0  0  0  0  0  2  7]]

CLASSIFICATION REPORT:
                precision    recall  f1-score   support

      Akapulko       0.50      0.88      0.64         8
      Ampalaya       0.94      1.00      0.97        15
   Balbas-pusa       0.57      0.36      0.44        11
       Bayabas       1.00      0.50      0.67        14
       Lagundi       1.00      0.89      0.94         9
     Malunggay       0.92      0.92      0.92        13
 Niyog-Niyogan       0.50      0.25      0.33         8
       Oregano       1.00      1.00      1.00         6
       Sambong       0.50      0.71      0.59        14
  Tsaang-Gubat       0.11      0.14      0.12         7
Ulasimang-Bato       0.67      1.00      0.80         6
   Yerba-Buena       1.00      0.78      0.88         9

   avg / total       0.75      0.71      0.71       120
```

Fig. 10. ANN Result Using Default Parameters Tested in Whole Dataset

```
ACCURACY SCORE:
0.721666666667

CONFUSION MATRIX:
[[43  0  0  2  0  4  1  0  0  0  0  0]
 [ 0 48  0  0  2  0  0  0  0  0  0  0]
 [ 0  0 17  0  0  0  0  0 33  0  0  0]
 [ 6  0  0 39  0  0  0  0  5  0  0  0]
 [ 0  3  0  0 47  0  0  0  0  0  0  0]
 [ 0  0  0  0  0 49  0  0  0  0  1  0]
 [12  0  2 13  0  0 20  0  1  2  0  0]
 [ 0  0  0  0  0  0  0 46  0  0  4  0]
 [ 0  0  9  4  0  0  0  0 37  0  0  0]
 [16  0  3 19  0  0  9  0  2  1  0  0]
 [ 0  0  0  0  0  0  0  0  0  0 50  0]
 [ 0  0  0  0  0  7  0  0  0  0  7 36]]

CLASSIFICATION REPORT:
                precision    recall  f1-score   support

      Akapulko       0.56      0.86      0.68        50
      Ampalaya       0.94      0.96      0.95        50
   Balbas-pusa       0.55      0.34      0.42        50
       Bayabas       0.51      0.78      0.61        50
       Lagundi       0.96      0.94      0.95        50
     Malunggay       0.82      0.98      0.89        50
 Niyog-Niyogan       0.67      0.40      0.50        50
       Oregano       1.00      0.92      0.96        50
       Sambong       0.47      0.74      0.58        50
  Tsaang-Gubat       0.33      0.02      0.04        50
Ulasimang-Bato       0.81      1.00      0.89        50
   Yerba-Buena       1.00      0.72      0.84        50

   avg / total       0.72      0.72      0.69       600
```

Without varying the default parameters of the neural network, accuracy score is around 70 to 72 % which is not good for machine learning. To improve the performance of the model, an optimized value for some parameters are experimentally done specifically for the following:

1. activation: Activation function for the hidden layer.
2. solver: The solver for weight optimization.
3. alpha: Penalty (regularization term) parameter
4. learning_rate: Learning rate schedule for weight updates.

After several simulations, the following optimized parameters are established and implemented in the neural network model in Python along with other default values: activation=tanh, solver=lbfgs, alpha=0.0001, and learning_rate=invscaling. The tanh is the hyperbolic tan function, returns f(x)=tanh(x), lbfgs is an optimizer in the family of quasi-Newton methods, and invscaling is gradually decreases the learning rate at each time step ‘t’ using an inverse scaling exponent of ‘power_t’. Using these optimized values, figures 11 to 12 depicted the accuracy score, confusion matrix, and classification report of the model as tested to the validation dataset and whole dataset respectively.

Fig. 11. ANN Result Using Optimized Parameters Tested in Validation Dataset

```
ACCURACY SCORE:
0.925

CONFUSION MATRIX:
[[ 8  0  0  0  0  0  0  0  0  0  0  0]
 [ 0 15  0  0  0  0  0  0  0  0  0  0]
 [ 0  0 11  0  0  0  0  0  0  0  0  0]
 [ 0  0  0 14  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  9  0  0  0  0  0  0  0]
 [ 0  0  0  0  0 12  0  0  0  0  1  0]
 [ 0  0  0  2  0  0  5  0  1  0  0  0]
 [ 0  0  0  0  0  0  0  6  0  0  0  0]
 [ 0  0  0  0  0  0  1  0 12  1  0  0]
 [ 0  0  0  1  0  0  1  0  0  5  0  0]
 [ 0  0  0  0  0  0  0  1  0  0  5  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  9]]

CLASSIFICATION REPORT:
                precision    recall  f1-score   support

      Akapulko       1.00      1.00      1.00         8
      Ampalaya       1.00      1.00      1.00        15
   Balbas-pusa       1.00      1.00      1.00        11
       Bayabas       0.82      1.00      0.90        14
       Lagundi       1.00      1.00      1.00         9
     Malunggay       1.00      0.92      0.96        13
 Niyog-Niyogan       0.71      0.62      0.67         8
       Oregano       0.86      1.00      0.92         6
       Sambong       0.92      0.86      0.89        14
  Tsaang-Gubat       0.83      0.71      0.77         7
Ulasimang-Bato       0.83      0.83      0.83         6
   Yerba-Buena       1.00      1.00      1.00         9

   avg / total       0.93      0.93      0.92       120
```

Fig. 12. ANN Result Using Optimized Parameters Tested in Whole Dataset

```
ACCURACY SCORE:
0.981666666667

CONFUSION MATRIX:
[[50  0  0  0  0  0  0  0  0  0  0  0]
 [ 0 50  0  0  0  0  0  0  0  0  0  0]
 [ 0  0 50  0  0  0  0  0  0  0  0  0]
 [ 0  0  0 49  0  0  1  0  0  0  0  0]
 [ 0  0  0  0 50  0  0  0  0  0  0  0]
 [ 0  0  0  0  0 50  0  0  0  0  0  0]
 [ 0  0  0  0  0  0 47  0  0  3  0  0]
 [ 0  0  0  0  0  0  0 50  0  0  0  0]
 [ 0  0  0  0  0  0  0  0 49  1  0  0]
 [ 0  0  0  0  0  0  6  0  0 44  0  0]
 [ 0  0  0  0  0  0  0  0  0  0 50  0]
 [ 0  0  0  0  0  0  0  0  0  0  0 50]]

CLASSIFICATION REPORT:
                precision    recall  f1-score   support

      Akapulko       1.00      1.00      1.00        50
      Ampalaya       1.00      1.00      1.00        50
   Balbas-pusa       1.00      1.00      1.00        50
       Bayabas       1.00      0.98      0.99        50
       Lagundi       1.00      1.00      1.00        50
     Malunggay       1.00      1.00      1.00        50
 Niyog-Niyogan       0.87      0.94      0.90        50
       Oregano       1.00      1.00      1.00        50
       Sambong       1.00      0.98      0.99        50
  Tsaang-Gubat       0.92      0.88      0.90        50
Ulasimang-Bato       1.00      1.00      1.00        50
   Yerba-Buena       1.00      1.00      1.00        50

   avg / total       0.98      0.98      0.98       600
```

Varying some default parameters of the neural network classifier, accuracy score is improved for validation dataset from 70.83 % to 92.50 % and from 72.16 % to 98.16 % for the whole dataset which is good for machine learning.

The artificial neural network is then developed using the MATLAB’s NEURAL NETWORK Toolbox™. The input layer of the network that the researchers had created corresponds to the five parameters that are derived from the parameters that had been extracted by the image processing module. The input had been encoded in MATLAB in matrix form, which is a 5x600 matrix, since there are five parameters which correspond to the Aspect Ratio, Circularity, Convexity, Solidity, and Rectangularity, respectively and there are 600 samples of leaves. The 600 samples came from the 50 leaf samples per species of the 12 herbal medicine plants; wherein the 50 leaf samples are collected from only one variety or sub-species of each of the herbal medicine plants. The 50 samples of every herbal medicine plant are consisting of leaf samples of variable sizes, to make the system learn proper recognition of them in any size. The 600 samples used for the training are all in good shape, or in other words, have no deformities.

In contrast, the output layer corresponds to the target output of the system based on the given set of inputs. Since there are 12 species of herbal plants to be identified and there are 600 samples of leaf images, the target output is a 12x600 matrix. Both the input and output matrices were retrieved from an Excel file that holds the different features that had been extracted by the feature extraction module during the Feature Extraction Phase.

After establishing the learning of the module in MATLAB, the actual testing of ANN is done with the random set of herbal leaf samples. Samples are placed in the image capturing box of the system, undergo the process of image preprocessing, classified by the neural network model and displayed the results. The sample output GUI are presented in Figure 13.

Fig. 13. Sample Test Output for the ANN Module

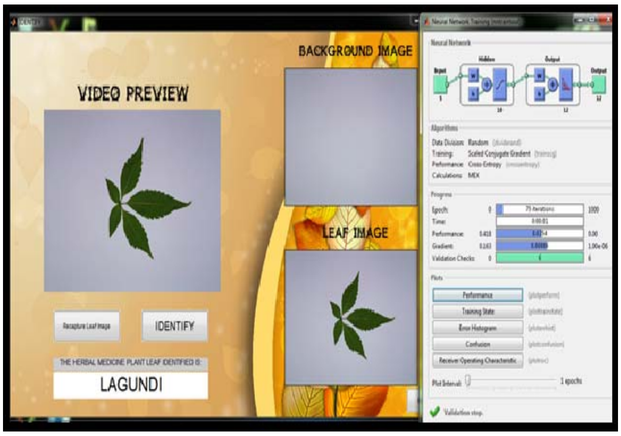


Figure 14 is the sample GUI of the use of herbal plant which can be accessed after the identification.

Fig.14. Sample GUI of Herbal Plant Usage

To finally test the complete system, 72 new leaf samples 6 per herbal plant are collected and used in the testing. The system was able to give a total of 71 successful identifications from the 72 trials done. The percentage of accuracy of the results can be computed by the ratio of the number of correct identification over the total number of samples multiplied by 100 thus giving a result of 98.61 %.

V. CONCLUSIONS

Researchers had successfully developed the herbal medicine plant leaf identification system. Python was used in the analysis of the dataset in establishing optimized parameters for the neural network model. It was then implemented in MATLAB and integrated in the whole system. Datasets are derived from 12 types of herbal plants with 50 samples per type giving 600 samples where the primary and secondary features are extracted. The 12 herbal medicine plants used in the study are akapulko, amplaya, balbas-pusa, bayabas, lagundi, malunggay, niyog-niyogan, oregano, sambong, tsaang-gubat, ulasimang-bato, and yerba-buena. Classification rate of the system are increased from 72.16 % to 98.16 % for the whole dataset by varying the alpha, solver, learning rate and activation parameter of the neural network model to its optimized value.

With regards to the evaluation of the functionality and accuracy of the system, the researchers had conducted actual testing of the different parts of the system, including the hardware part and the software components. Based on the results obtained from the 72 new leaf samples, the complete system provided herbal leaf identification of 98.61% accuracy.

VI. RECOMMENDATIONS

The following are suggestions that can be considered in the improvement of the system created:

1. future researcher can include additional number of Philippine herbal medicine plant leaf with increased number of samples to attain higher accuracy of identification.

2. future researcher should have considered adding parameters or characteristic definitions of the plants aside from the different morphological parameters like defining the location or place of the herbal medicine plant commonly found, in order to provide more distinct characterization of the plants.

3. the concept of the study can also be used for the identification of different underutilized and indigenous fruit plants based also in its leaf samples and actual fruit samples, considering the shape and structure of the fruit and other possible characteristics of the fruit itself.

ACKNOWLEDGMENT

The authors would like to acknowledge the Engineering Research and Development for Technology (ERDT) of the Department of Science and Technology (DOST) for funding this research. The authors also recognize the assistance and technical expertise of members of the Intelligent Systems Learning (ISL) Laboratory of De La Salle University (DLSU).

REFERENCES

[1] Rademaker C. A. 2000. The classification of plants in the United States Patent Classification system. World Patent Information 22: 301–307

[2] Neeraj Kumar, Peter N Belhumeur, Arijit Biswas, David W Jacobs, W John Kress, Ida C Lopez, and João VB Soares, “Leafsnap: A computer vision system for automatic plant species identification,” in ECCV, pp. 502–516. Springer, 2012.

[3] Cem Kalyoncu and Önsen Toygar, “Geometric leaf classification,” Computer Vision and Image Understanding, vol. 133, pp. 102–109, 2014.

[4] Abdul Kadir, Lukito Edi Nugroho, Adhi Susanto, and Paulus Insap Santosa, “Leaf classification using shape, color, and texture features,” arXiv preprint arXiv:1401.4447, 2013.

[5] Thibaut Beghin, James S Cope, Paolo Remagnino, and SarahBarman, “Shape and texture based plant leaf classification,” in Advanced Concepts for Intelligent Vision Systems, 2010, pp. 345–353.

[6] James Charters, Zhiyong Wang, Zheru Chi, Ah Chung Tsoi, and David Dagan Feng, “Eagle: A novel descriptor for identifying plant species using leaf lamina vascular features,” in ICME-Workshop, 2014, pp. 1–6.

[7] James S Cope, Paolo Remagnino, Sarah Barman, and Paul Wilkin, “The extraction of venation from leaf images by evolved vein classifiers and ant colony algorithms,” in Advanced Concepts for Intelligent Vision Systems, 2010, pp. 135–144.

[8] Xiao-Feng Wang, Ji-Xiang Du, and Guo-Jun Zhang "Recognition of Leaf Image based on Shape Feature Using Hyper sphere Classifier".

[9] Wang, Z. Chi, Z. Feng, D. "Shape based leaf image retrieval" publication year 2003, pages 34-43.

[10] Xun Liu Keyi Hou Liliang Wang Ping Liu "Index and recognition for the shape counter of plant leaves" Publication year 2009, pages 436-440.

[11] Takeshi Saitosh, Kimiya Aoki, and Toyohisa Kaneko "Automatic Recognition of Blooming Flowers" ICPR'04 10514651/04 IEEE.

[12] Anderson Rocha, Daniel C.Hauagge, Jacques Wainer, Siome Goldenstein "Automatic fruit and vegetable classification from images" Comput.Electron.Agric. (2009) doi: 1 0.1 016/j.compag.2009.

[13] X. Wang, D. Huang, J. Dua, H. Xu and L. Heutte, Classification of plant leaf images with complicated background, Appl. Math. Comput., vol.205, pp.916-926, 2008.

[14] S. G. F. Wu, S. Bao, E. Y. Y. Xu, X. Y. Wang, F. Chang and Q. L. Xiang, A Leaf Recognition Algorithm for Plant Classification using Probabilistic Neural Network, The 7th IEEE International Symposium on Signal Processing and Information Technology, 15-18 Dec. 2007, Egypt, pp. 11-16, 2007.

[15] K. B. Lee and K. S. Hong, Advanced Leaf Recognition based on Leaf Contour and Centroid for Plant Classification, The 2012 International Conference on Information Science and Technology, Shanghai, China, vol. 3, pp. 133-135, 2012.

[16] J. H. Kim, R. G. Huang, S. H. Jin and K. S. Hong, Mobile-based flower recognition system, 3rd International Symposium on Intelligent Information Technology Application, Nan Chang, China, 21-22 Nov. 2009, pp.580-583, 2009.

[17] V. Satti , A. Satya , S. Sharma, An Automatic Leaf Recognition System For Plant Identification Using Machine Vision Technology International Journal of Engineering Science and Technology, vol.5, pp.874-879, 2013.

[18] J.A. Freeman and D.M. Skapura, Neural networks: algorithms, applications, and programming techniques, Reading, Massachusetts, USA: Addison-Wesley, 1992.

[19] S. Haykin, Neural networks - a comprehensive foundation. New York, USA: Macmillan College Publishing Company, Inc.,1994.

[20] Philippine Herbal Medicine.2014. Retrieved date: September 12, 2015.Retrieved from:http://www.philippineherbalmedicine.org/.

[21] Wu, S. G., et. al., A Leaf Recognition Algorithm for Plant Classification Using Probabilistic Neural Network, in Signal Processing and Information Technology, 2007 IEEE International Symposium 2007: Giza p. 11-16.

[22] Lin, T.-T., Y.-T. Chi, and W.-C Liao. Leaf Boundary Extraction and Geometric Modeling of Vegetable Seedlings in 2000 ASAE Annual International Meeting. 2000. Milwaukee, Wisconsin, USA.

[23] Hati, Shayan & G, Sajeevan. (2013). Plant Recognition from Leaf Image through Artificial Neural Network. International Journal of Computer Applications. 62. 15-18. 10.5120/10172-4897.

[24] Kadir, Abdul & Edi Nugroho, Lukito & Susanto, Adhi & Santosa, Paulus. (2011). Leaf Classification Using Shape, Color, and Texture Features. International Journal of Computer Trends and Technology. 1. 225-230.

[25] Manisha Kaushal, Arjan Singh, Baljit Singh, "Adaptive thresholding for Edge Detection in Gray Scale Images", International Journal of Engineering Science and Technology", Vol. 2(6), pp. 2077-2082, Jan. 2010.

[26] Raman Mani, Dr. Himanshu Aggarwal, "Study and Comparison of Various Image Edge Detection Techniques", international Journal of Image Processing (lJIP), Vol.3 (I), pp. 1- 12, Jan. 20 II.

[27] C.H Arun, W.R Sam Emmanuel, D. Christopher Dhurairaj, "Texture Feature Extraction for Identification of Medicinal Plants and Comparison of Different Classifiers", International Journal of Computer Applications (0975 - 8887), Volume 62(12), Jan. 2013.

[28] Janani R, Gopal A, "Identification of Selected Medicinal Plant Leaves Using Image Features and ANN", IEEE International Conference on Advanced Electronic Systems (ISAES), pp. 238- 242, Sept. 2013.

978-1-5386-0912-5/17/$31.00 ©2017 IEEE