Identification of Cocoyam (Taro) Leaf Disease Using Deep Learning to Enhance Food Security

Nduanya U.I, Okoye Francis A, Ozor Godwin O.

Computer Engineering, Enugu State University of Science and Technology, Nigeria

***Abstract: -* For food security sustainability, the need for automatic identification and diagnosis of taro leaf diseases is highly desired. To improve the identification and classification accuracy of cocoyam leaf diseases and reducethe number of network parameters, the improved decision tree model of deep learning is proposed for leaf disease recognition in this paper. The deep learning model takes several images of the ravaged and normal leaves of different shapes and orientation. A high definition camera was used to capture the images to avoid blurred or overlapping images.With the techniques, there are significant improvements in identification and predictions of the presence of the taro leaf disease.**

***Keywords:* taro, deep learning, leaf disease.**

I. INTRODUCTION

Cocoyam or Taro (*Colocasia esculenta*) is a starchy edible tuber and bears broad greenish leaves. Cocoyam an important food. It is usually prepared in different pattern based on family choice. Another important aspect of the cocoyam was that is relatively cheap compared to other meals of its kind. However, in recent years, the number of species of cocoyam diseases and the degree of harm they cause have increased, mainly due to changes in cultivation systems, the variation of pathogen varieties, and inadequate of plant protection measures. In the year 2009 majority of farmers in the southern part of Nigeria experienced a serious loss as a result of taro leave diseases [1]. Epidemics of leaf blight(Phytophthoracolocasiae) may occur throughout the year during rainy, overcast weather when night temperatures are between 20-22°c and temperatures during the day are from 25-28°c. Entire fields may be blighted in five to seven days under these conditions [2].

Cocoyam leaf diseases have various symptoms. It may be more difficult for inexperienced farmers to diagnose diseases than for professional plant pathologists. As a verification system in disease diagnostics, an automatic system that is designed to identify plant diseases by the plant's appearance and visual symptoms could be of great help to farmers. Many efforts have been applied to the quick and accurate diagnosis of leaf diseases. By using digital image processing techniques, support vector machine (SVM), neural networks and other methods, we can detect and classify leaf diseases [4]_[7]. An SVM - based multi - classifier was proposed by Song *et al*. [8] and was applied to identify a variety of maize leaf diseases. The best recognition accuracy was 89.6%. The method of classification using SVM is only applicable to small samples, for a large number of samples, it cannot achieve high recognition accuracy. Chen and Wang [9] proposed a method for the identification of maize leaf diseases based on image processing technology and a probabilistic neural network (PNN). The best recognition accuracy of this method was 90.4%. However, for the PNN classifier, the identification accuracy and speed of this method decreases as the number of training samples increases.Amethod of leaf disease identification based on adaptive weighting multi-classifier fusion was proposed by Xu *et al.* [4]. Seven common types of leaf disease were tested by this method. The average recognition rate was 94.71%. Wang *et al*. [5] Qi *et al.* [6], and Zhang proposed different methods using digital image processing techniques based on Fisher discriminant, Retinex algorithm combined with principal component analysis (PCA) and SVM, and quantum neural network (QNN) and combination features for identification of leaf disease. The highest recognition accuracy of these studies was 95.3%, but fewer taro leaf diseases were involved in these methods. Different methods are used to identify leaf diseases [4] _ [7], the best recognition accuracy was 95.3%, which cannot meet the current requirements for high recognition accuracy. Therefore, in the follow - up study, we should focus on how to improve the identification accuracy. Machine learning has made tremendous advances in the past few years. It is now able to extract useful feature representations from a large number of input images. Machine learning provides an opportunity for detectors to identify crop diseases in a timely and accurate manner, which will not only improve the accuracy of plant protection but also expand the scope of computer vision in the field of precision agriculture.

II. METHODOLOGY

A. Required dataset

In object training, recognition and classification research, relevant dataset is a prerequisite for choosing appropriate algorithm or techniques and also for performance evaluation of the system. A totalof 725 images are collected from different sources, such as the smartphone cameras from different communities and Google websites, including different periods of occurrence of cocoyam leaf diseases, which are divided into 23 different categories based on the shapes and nature of the disease symptoms. There are additional 5 categories representing uninfected or healthy cocoyam leaves and a category representing healthy leaves. There are different species of cocoyam hence the leaves differ and the effect of the diseases based on the species were investigated. Fig 1 shows the samples of leaves collected for the research.

Figure 1: Cocoyam leaf with the both affected and unaffected

B. *Augmentation*

Training neural networks in deep learning requires substantial data. The more data the deep learning has to learn, the more features it can obtain. Since the original leaf image dataset collected in this study is not sufficient, it is necessary to expand the dataset by different methods to distinguish the different disease categories.After the original images are initialized, additional versionsare created by rotating the images in the various angles of 45o, 90o, 135o and 180o; bymirroring each rotated image; by cutting the center of the image by the same size; and by converting all processedimages to grayscale. The dataset is expanded by the abovemethods, which helps in reducing over - fitting during thetraining stage.

C. *Image Pre-Processing*

To improve feature extraction and increase consistency, the images in the dataset for the deep learning classifier are pre-processed before the model is trained. One of the most significant operations is the normalization of image size and format.

In this study, all images are resized to 224×224 pixels and which are automatically computed by MATLAB function. Alexnet was the model for the classification.

In the interest of confirming the accuracy of the classesin the dataset, agricultural experts examined leaf imagesgrouped by a keyword search and labelled all the images withthe appropriate disease acronym. It is well known that it isessential to use accurately classified images for the trainingand validation dataset. Only in that can may an appropriateand a reliable model be developed. In this stage, variousclasses of the dataset as well as the training set and the testingset are marked.In the proposed approach, the method adopted for extracting the feature set is called the *Color Co-occurrence Method* or CCM method in short. It is a method, in which both the color and texture of an image are taken into account, to arrive at unique features, which represent that image.

The image analysis technique selected for this study was the CCM method. The use of color image features in the visible light spectrum provides additional image characteristic features over the traditional gray-scale representation.

The CCM methodology established in this work consists of three major mathematical processes. First, the RGB images of leaves are converted into Hue Saturation Intensity (HSI) color space representation. Once this process is completed, each pixel map is used to generate a color co-occurrence matrix, resulting in three CCM matrices, one for each of the H, S and I pixel maps. (HSI) space is also a popular color space because it is based on human color perception. Electromagnetic radiation in the range of wavelengths of about 400 to 700 nanometers is called visible light because the human visual system is sensitive to this range. Hue is generally related to the wavelength of a light and intensity shows the amplitude of a light. Lastly, saturation is a component that measures the “colorfulness” in HSI space [14]. Color spaces can be transformed from one space to another easily. In these experiments, the Equations 1, 2 and 3 were used to transform the images components from RGB to HSI:

$$Hue(s) = \begin{cases} 2 - acos\left\{\frac{[(R-G)+(R-B)]}{\sqrt{(R-G)^2+(R-G)(G-B)}}\right\}, B > G \\ acos\left\{\frac{[(R-G)+(R-B)]}{\sqrt{(R-G)^2+(R-G)(G-B)}}\right\}, B \le G \end{cases} \tag{4}$$

$$Saturation(s) = 1 - \frac{3*\min(R,G,B)}{R+G+B} \tag{5}$$

$$Intensity(I) = \frac{R+G+B}{3} \quad (6)$$

The color co-occurrence texture analysis method was developed through the use of Spatial Gray-level Dependence Matrices (SGDM"s). The gray level co-occurrence methodology is a statistical way to describe shape by statistically sampling the way certain grey-levels occur in relation to other grey-levels.

These matrices measure the probability that a pixel at one particular gray level will occur at a distinct distance and orientation from any pixel given that pixel has a second particular gray level. For a position operator p, we can define a matrix *Pij* that counts the number of times a pixel with grey level *i* occurs at position *p* from a pixel with grey-level *j*. The SGDMs are represented by the function *P(i, j, d, Ɵ)* where *I* represents the gray level of the location *(x, y)* in the image *I(x,y)*, and *j* represents the gray level of the pixel at a distance *d* from location *(x, y)* at an orientation angle of *Ɵ*. The reference pixel at image position *(x, y)* is shown as an Asterix(***). All the neighbors from 1 to 8 are numbered in a clockwise direction. Neighbors 1 and 5 are located on the same plane at a distance of *1* and an orientation of *0* degrees. An example image matrix and its SGDM are already given in the three equations above.

In this research, a *one* pixel offset distance and a *zero-degree* orientation angle was used. After the transformation processes, we calculated the feature set for H and S, we dropped (I) since it does not give extra information. However, we use GLCM function in Matlab to create gray-level co-occurrence matrix; the number of gray levels is set to 8, and the symmetric value is set to "true", and finally, offset is given a" 0" value.

D. Convolution Neural Network

1. Convolution

Convolution is the most important operation in CNNs. The convolution calculation of the two - dimensional image can be mapped to the continuous sliding convolution window to obtain the corresponding convolution value. In CNNs, each feature map is convoluted by multiple input feature graphs. For an input *x* of the *i*th convolutional layer, it computes as (1),

$$h_{ic} = f(W_i * x) \quad (1)$$

where * represents the convolution operation, *Wi* represents the convolution kernels of the layer, and *f* represents the activation function. $W_i = [W_i^1, W_i^2, \dots, W_i^k]$, *K* is the number of convolution kernels of the layer. Each kernel $W_i^k$is an $M \times M \times N$weight matrix with *M* being the window sizeand *N* being the number of input channels [10].

2. Activate Function

The *ReLu* activation function is an unsaturated nonlinear function that can receive signals by simulating brain neurons. Saturated nonlinear function, such as *Sigmoid* and *Tanh*, have worse performance than unsaturated nonlinear functions when training a network.

In this test, the *ReLu*activation function is part of alexnet model, (it the last module before the classifier output), to prevent the problem of gradient dispersion while accelerating network training and to increase the identification accuracy.

3. Pooling

As the number of convolutional layers increases, the parameters of the network will increase exponentially. The pooling operation can effectively reduce the number of network parameters. To reduce the parameters in all regions, the pooling operation is performed by calculating the statistical characteristics of a region in order to represent the entire region's characteristics.

The effect of different pooling combinations on the identification accuracy of alexnet model will be explored in this study.

4. Dropout

Combining the predictions of many different models is a very successful way to reduce test errors[11, 12], but it appears to be too expensive for big neural networks that already take several daysto train. There is, however, a very efficient version of model combination that only costs about afactor of two during training. The recently-introduced technique, called "dropout" [13], consistsof setting to zero the output of each hidden neuron with probability 0.5. The neurons which are"dropped out" in this way do not contribute to the forward pass and do not participate in backpropagation.

So every time an input is presented, the neural network samples a different architecture, but all these architectures share weights. This technique reduces complex co-adaptations of neurons, since a neuron cannot rely on the presence of particular other neurons. It is, therefore, forced tolearn more robust features that are useful in conjunction with many different random subsets of theother neurons. At test time, we use all the neurons but multiply their outputs by 0.5, which is areasonable approximation to taking the geometric mean of the predictive distributions produced bythe exponentially-many dropout networks.

5. Loss Function

The loss function measures the discrepancy between the predictedresult and the label of the input, which is defined as in equation (2)

$$E(W) = -1/n \sum_{xi=1}^{n} \sum_{k=1}^{k} [y_{ik} logP(x_i = k) + (1 - y_{ik}) \log(1 - P(x_i = k))] \quad (2)$$

where W indicates the weighting matrixes of the convolutional and fully connected layers, n indicates the number of training samples, i is the index of training samples, and k is the index of classes. If the ith sample belongs to the kth class, $y_{ik} = 1$; else$y_{ik} = 0$. $P(x_i = k)$ is the probability of input xi belonging to the kth class that the model predicts, which is a function of the parameters W. Therefore, the loss function takes W as its parameters. Network training aims to find the value of W that minimizes the loss function E.

III. RESULTS AND DISCUSSION

Alexnet was much larger than previous CNNs used for computer vision tasks. The initial learning rate of the Alexnet model is 0.001, using the ``step" method to attenuate the learningrate. After 10000th iterations in a matter of few minutes, accuracy of 99.3% was recorded. Fig.2 shows the changesof test accuracy and Fig. 3 shows the curveof the system loss. We can see that the test identification accuracy graduallyconverges after 40000thiterations. The training time and the convergence time of theoriginal model are longer. The original model also has a largernumber of parameters.

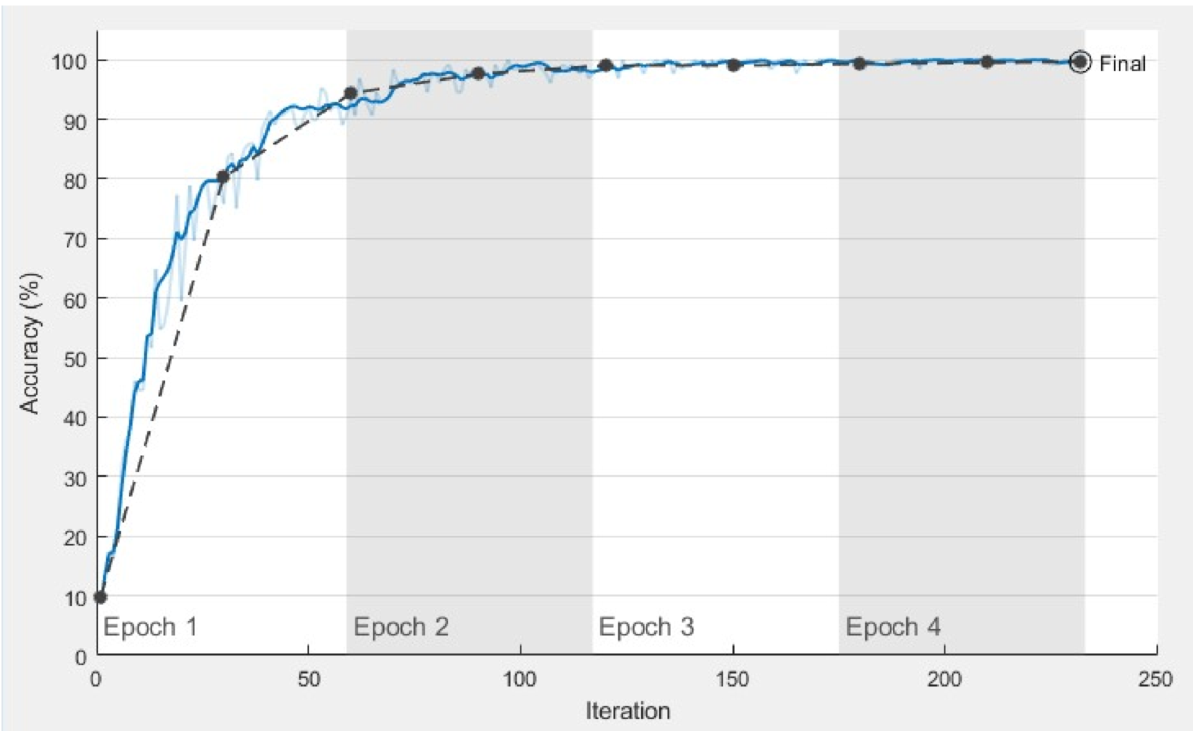

Fig 2: Changes in test accuracy

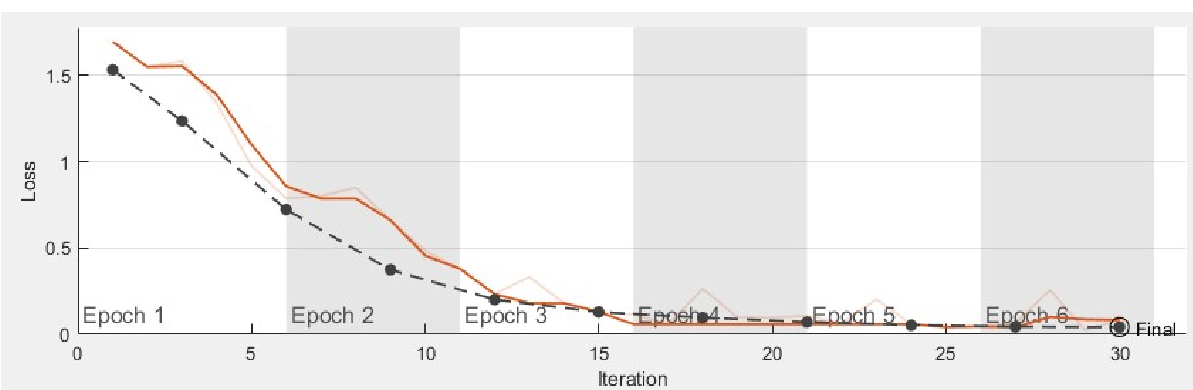

Fig 3: system loss gradient

Compared with the original unmodified model, the identification accuracy and system loss of the improved model arebetter than the original one. The improved model's identification accuracy is 0.7% higher than that of the originalone, the system loss is 10.8% less than the original one. The convergence time have been greatly improved,which can effectively improve the model training and recognitionefficiency.

IV. CONCLUSION

In this study, when identifying 5 types of cocoyam leaves, the improved deep learning models can achieve high identification accuracy, 99.3%. When the train - test set is 80 - 20 (80% of the whole dataset used for training, and 20% for testing), the classification algorithms used in this study allow the systems to acquire a diversity of sample conditions with strong robustness. Experiments show that it is possible to improve recognition accuracy by increasing the diversity of pooling operations, the reasonable addition of a *ReLu*function and dropout operations, and including multiple adjustments of the model parameters.

In future research, we will identify more types of cocoyam diseases and combine new algorithms and other deep learning structures for the training and testing of the model. Meanwhile, in order to enable agricultural producers to make quick and reasonable judgments about crop disease information. We will be working on the comprehensive data for artificial intelligent system.

REFERENCES

[1]. Bandyopadhyay R, Sharma K, Onyeka TJ, Aregbesola A, Kumar PL (2011) First Report of Taro (Colocasia esculenta) Leaf Blight Caused by Phytophthora colocasiae in Nigeria. APS Publication, https://doi.org/10.1094/PDIS-12-10-0890 , 3340 Pilot Knob Road, St. Paul, MN 55121 USA
[2]. Trujillo, E.E (1965). Effects of humidity and temperature on Phytophthora blight of taro. Phytopathology 55:183-188.
[3]. Onwueme I. (1999) Taro Cultivation in Asia and the Pacific.FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS. RAP Publication: 1999/16. Fao.org
[4]. L. F. Xu, X. B. Xu, and H. Min, ``Corn leaf disease identi_cation based on multiple classi_ers fusion,'' *Trans. Chin. Soc. Agricult. Eng.*, vol. 31, no. 14, pp. 194_201, 2015.
[5]. N. Wang, K. Wang, R. Xie, J. Lai, B. Ming, and S. Li, ``Maize leaf disease identi_cation based on _sher discrimination analysis,'' *Scientia Agricultura Sinica*, vol. 42, no. 11, pp. 3836_3842, 2009.
[6]. Z. Qi *et al.*, ``Identi_cation of maize leaf diseases based on image technology,'' *J. Anhui Agricult. Univ.*, vol. 43, no. 2, pp. 325_330, Feb. 2016.
[7]. F. Zhang, ``Recognition of corn leaf disease based on quantum neural network and combination characteristic parameter,'' *J. Southern Agriculture*, vol. 44, no. 8, pp. 1286_1290, 2013.
[8]. K. Song, X. Y. Sun, and J. W. Ji, ``Corn leaf disease recognition based on support vector machine method,'' *Trans. Chin. Soc. Agricult. Eng.*, vol. 23, no. 1, pp. 155-157, Jan. 2007.
[9]. L. Chen and L. Y. Wang, ``Research on application of probability neural network in maize leaf disease identi_cation,'' *J. Agricult. Mech. Res.*, vol. 33, no. 6, pp. 145_148, Jun. 2011.
[10]. G. Wang, Y. Sun, and J. X. Wang, (2017) ``Automatic image-based plant disease severity estimation using deep learning,'' in *Computational Intelligenceand Neuroscience*, pp. 1_8.
[11]. R.M. Bell and Y. Koren (2007) Lessons from the netflix prize challenge. ACM SIGKDD Explorations Newsletter, 9(2):75–79.
[12]. L. Breiman.(2001) Random forests. Machine learning, 45(1):5–32.
[13]. G.E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R.R. Salakhutdinov (2012) Improving neural networks by preventing co-adaptation of feature detectors. arXiv preprint arXiv:1207.0580.
[14]. Stone, M. C. (2001). "A Survey of Color for Computer Graphics". Course at SIGGRAPH 2001.