ISSN No:-2456-2165

Tobacco Disease Detection and Classification for Grading System Using Convolutional Neural Network

Pempho Jimu
Student: Department of Computer Science
DMI-St. Eugene University, Zambia
Lilongwe, Malawi

Dr. G. Glorindal Selvam
Project Supervisor: Department of Computer Science
DMI-St. John the Baptist University, DMI-SJBU
Lilongwe, Malawi

Abstract:- Tobacco being one of the perennial crops that are the major source of forex in Malawi. But due to the decrease in agricultural extension workers specializing in Tobacco this paper proposes a remedy for tobacco disease detection and grading using a convolution neural network (CNN). This paper proposes the use of CNN as a machine learning algorithm for the training a system to detect tobacco diseases while grading the tobacco. This paper also compares CNN algorithm with other neural network algorithms in terms of sensitivity, specificity and other parameters such that the CNN proves to be efficient and effective hence offering higher performance of the algorithms. This research work will be designing and developing a learning model, which will classify, grade and detect the disease in the tobacco plant leaf by using the convolutional neural network based on the type of the tobacco Leaf. As these convolutional networks work with a large number of images for the training of the model, real-time images will be collected from the reliable source such as Agricultural Research and Extension Trust (ARET) and Tobacco Auction Flow Point respectively.

Keywords:- *Image Processing, Convolutional Neural Network (CNN), Leaf Detection, Agricultural Research and Extension Trust (ARET), Tobacco Classification, Tobacco Grading*

I. INTRODUCTION

The effective preservation of plant diseases is closely related to bearable agriculture and climate variation [8]. Digital image processing has been applied in a variety of fields especially in medicine and agriculture. Digital image processing is used in a numerous-fields and it helps in automation of various processes requiring human being visual perception which is effectively side stepped by the inception of digital image processing. Digital image processing is very important because it involves the use of machines which are accurate every time they are running since they do not get tired despite the long time spent running. Another importance of digital image processing is the capability of using a high computation power obtained from advanced computers and computing concepts such quantum computing which can help in performing highly complex tasks within a small amount of time. Not only that, but also, application in a wide variety of areas which makes it very flexible. Although digital image processing acquisition is much of attention in technological developments, it is also very incomplete especially when it comes to interoperability of a given system from one area of application to another. Usually this is the case because a digital image processing system is hard coded to only be effective and efficient in only one area of application and might not perform well in another field. This problem is corrected by adding highly advance digital image processing system with artificial intelligence, specifically neural networks due to their capability of learning new tasks by specifying urgings in the algorithm. Digital image processing has a lot of algorithms which use different techniques but share common steps to achieve a given objective whether it is image compression or image recognition. The fundamental steps in digital image processing image acquisition, image enhancement, image restoration, color image processing, wavelets and multiresolution processing, compression, morphological processing, segmentation, representation and description, and object recognition. This paper involves the usage of digital image processing together with neural networks in the recognition and identification of tobacco diseases [15].

II. RELATED WORKS

Agricultural equipment is evolving specifically precise prediction of diseases in plants [1]. This has facilitated very much in the objective of attaining food security in a lot of countries and making sure that farmers have the maximum quality of produce from the farm. The capability to detect diseases and to grade tobacco reduces subjective interference of human beings who are fallible. Expert systems were one of the necessary steps toward automation of detection of tobacco diseases and grading, but they were not as accurate and fast as those of neural networks [2]. Expert systems are rule-based where by the algorithm is based on conditions defined by a programmer or system engineer. A field as complex as medicine it is almost impossible for a system developer to anticipate all possible conditions without missing out. This gives enough of a reason why machine learning using neural networks fits better for problems such as detection and prediction of diseases. In the paper [3] convoluted neural networks (CNN) are used as the basis for training a system. Deep neural networks (DNN) are the most basic but they have a disadvantage with vanishing gradients. In order to overcome the drawbacks of such traditional method, new techniques are needed to predict the grading process. This has been proven by the previous studies [4,9], where they applied image processing and machine learning techniques to develop an automatic grading system for flue-cured tobacco leaves. They projected algorithms to extract the leaves' color, shape, quality, injury and waste tolerance, stalk point, and maturity.

Relative to classification, machine learning algorithms like Fuzzy [4], Support Vector Machine [7], K-Nearest Neighbor [6], Generalized Regression Neural Network [9], and Backpropagation Neural Network [8] were used to classify tobacco leaves. Because the classification features were selected based on human experience, these approaches enhanced the accuracy of tobacco grading. However, the sorting accuracy is still not high enough and is vulnerable to artificial feature selection. Developed in recent years, Convolutional Neural Networks (CNNs) have become popular especially in computer vision because they achieved remarkable performances on different tasks such as image classifications in the field of agriculture [9,11]. CNN is an end-to-end pipeline that can automatically discover the useful features for classification, whose advantage lie in local 3-D coherence in the input images, which allow them to have fewer bulks as some constraints are shared. Inspired by the development of CNN in image-based acknowledgment and not much studies on the application of CNN in tobacco leaf grading, this study attempts to use CNN in grading and classification of tobacco leaves and compare with the existing models [10]. This paper endorses convolution neural networks (CNN) which do not jump certain layers without suppressing the reduction endangered gradients which destructively impact the overall training of neural networks as they are being used. CNN is also more significant because the training is relatively faster than the DNN and is more novel in finding a lot of connections within the neural network.

III. METHODOLOGY

Before you begin to format your paper, first write and save the FORWARD PROPAGATION:

A neural network given the weight of W^ (l-1, l) for connection from layer l-1 to layer l and the weight W^(l-21, l) for connection from l-2 to l. This makes the forward propagation to be as shown below:

$$\boldsymbol{a^l} \coloneqq \boldsymbol{g}(\boldsymbol{W^{l-1,l}}.\boldsymbol{a^{l-1}} + \boldsymbol{b^l} + \boldsymbol{W^{l-2,l}}.\boldsymbol{a^{l-2}})$$
$$\boldsymbol{g}(\boldsymbol{Z^l} + \boldsymbol{W^{l-2,l}}.\boldsymbol{a^{l-2}})$$

Whereby the variables are:
a^l the activations of neurons in a layer l
g the activation function of layer l
$\boldsymbol{Z^l} = \boldsymbol{W^{l-1,l}}.\boldsymbol{a^{l-1}} + \boldsymbol{b^l}$.

This is the forward propagation of the system of a convolution neural network. A general case of layers more than three, the equation for the neuron activation is represented as follows:

$$\boldsymbol{a^l} \coloneqq \boldsymbol{g}(\sum_{\boldsymbol{k-2}}^{\boldsymbol{k}} \boldsymbol{W^{l-k,l}}.\boldsymbol{a^{l-k}})$$

Whereby k is the number of layers within a given neural network.

Below is a diagram that shows a neural network in forward propagation with three layers:

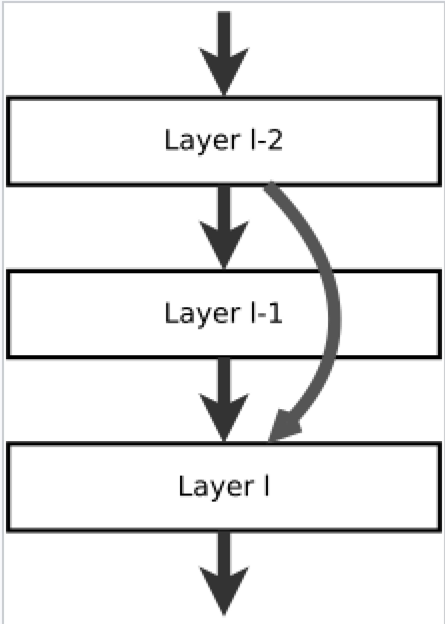

Fig 1:- neural network with three layers

➢ *Image Acquisition*
Digital images are acquired manually by users using either the user interface during the process of detecting tobacco disease and grading tobacco within an image and images are also acquired in a bulk in the form of a dataset as required by the code in MATLAB by a specified path. Acquisition conveys images to the image enhancement.

➢ *Image Enhencement*
Image enhancement converts the scale of an image or images to a uniform scale so as to minimize errors during classification and to avoid feeding a lot of errors to a given neural network [7]. The process is not complicated because the acceptable scale of both horizontal and vertical axis is the defined initially in the code.

➢ *Feature Extraction*
In as much as accelerative propagation is concerned, feature extraction is solely a result of the neural network's forward propagation. The system uses activation of neurons whereby each and every layer has its own feature to extract. Since this system uses RESNET-50 which has 50 layers and the algorithm does not jump other layers making it to be functionally a convolution neural network [4,5].

➢ *Classification*
Classification is nothing more than the assigning of a meaningful label to a given image by mapping a selected image feature weighted probabilities and comparing them to each of the feature weighted probabilities by the datasets. If there is congruency within the values a corresponding label is assigned to the selected image during classification.

Below is a diagram that shows how the modules explained above are interconnected together.

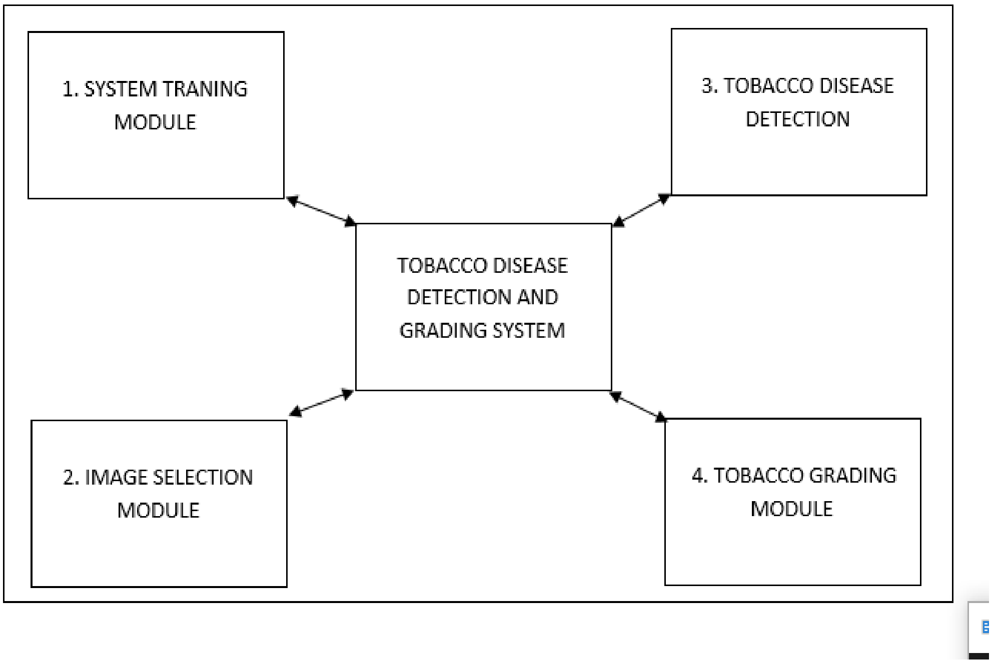

Fig 2:- module description figure

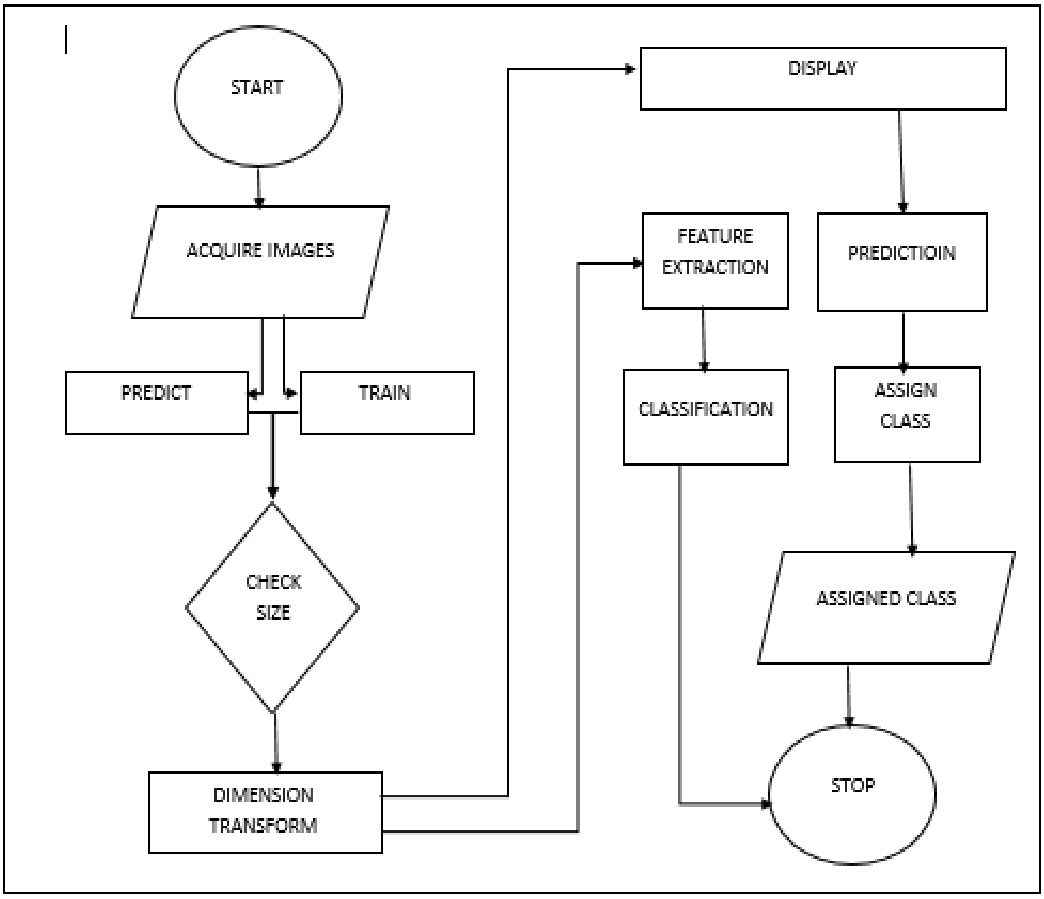

Fig 3:- tobacco disease prediction flow diagram

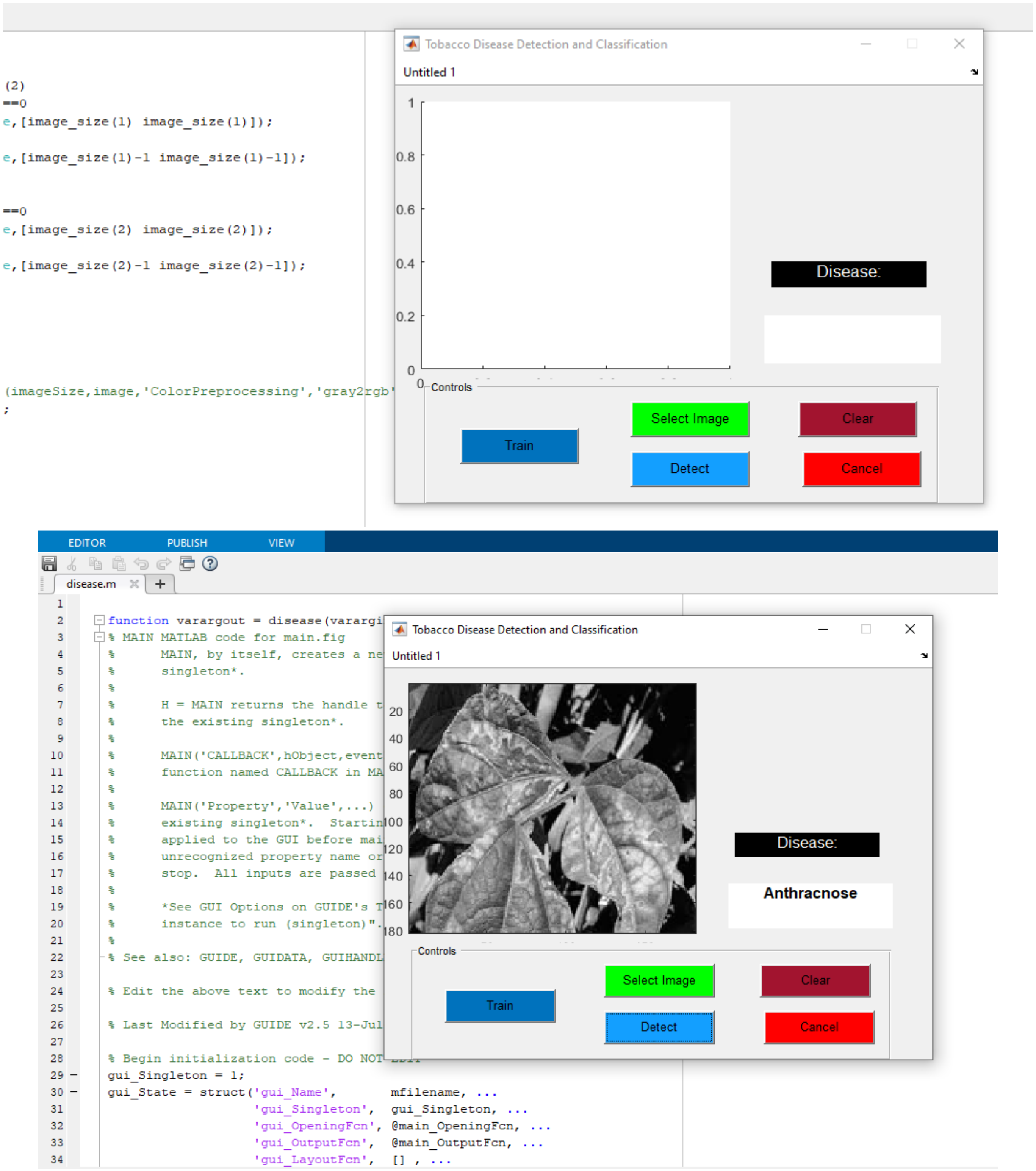

Fig 4:- tobacco disease prediction implementation

IV. RESULTS AND DISCUSSION

After the text edit has been completed, the paper is ready Using MATLAB quantitative results were collected with the help of high-end performance computers so as to have a detailed account of how the residual neural network algorithm works and performs as compared to other types of neural network algorithms.

Figure 5 shows that CNN is the best algorithm in terms of accuracy as compared to both deep neural network (DNN) due to the missing or disappearing gradients because in the DNN measurements are taken frequently hence little to no change is observed.

Accuracy is given by the formula:

$$\text{Accuracy}=\frac{TN+TP}{TN+TP+FP+FN}$$

Whereby, TN means true negative
TP means true positive,
FP means false positive,
and FN means false negative.

ISSN No:-2456-2165

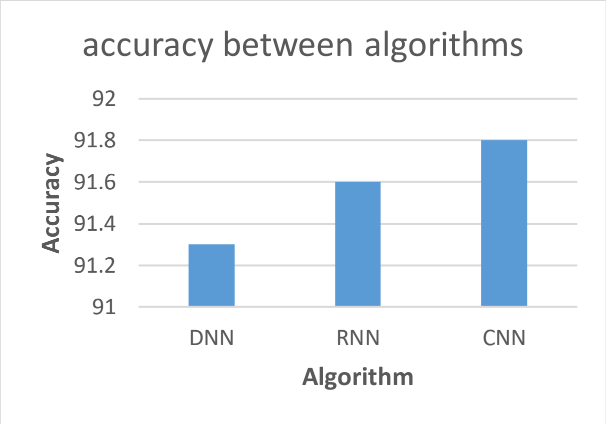

Fig 5:- accuracy between algorithms

From figure 6 sensitivity of the CNN algorithm is exponentially greater than that of both RNN and DNN because the CNN does not jump layers like the RNN and it takes into consideration small difference between objects or features.

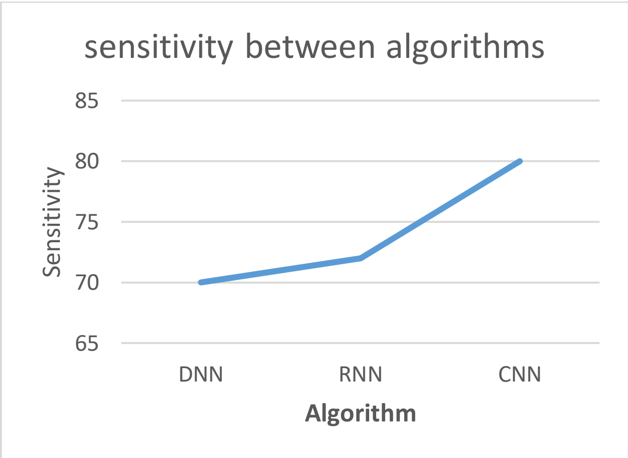

Fig 6:- sensitivity between algorithms

Figure 7 shows specificity of CNN during disease detection and grading of tobacco is greater than that of RNN and DNN because the CNN is a general algorithm that is very flexible to learn almost everything including image classification.

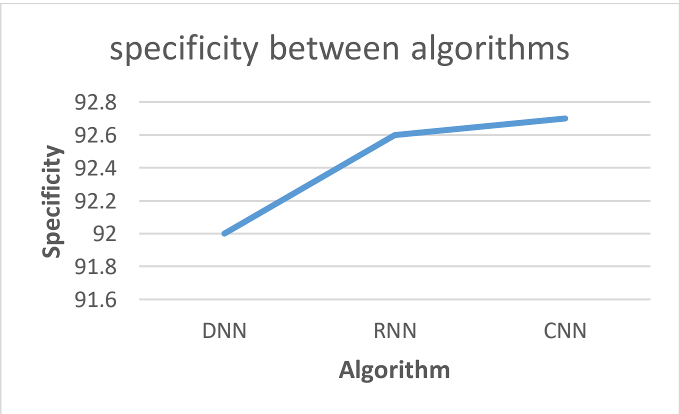

Fig 7:- specificity between algorithms

Figure 8 portrays the F-measure between DNN, RNN and CNN. From the graph it shows that both DNN and RNN have a similar F-measure while the CNN has a different F-measure which is of a higher value than both RNN and DNN.

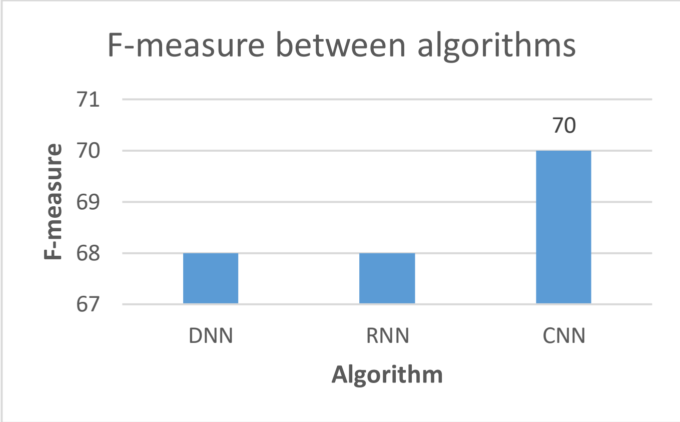

Fig 8:- F-measure between algorithms

V. CONCLUSION

This paper uses CNN with the help of RNN in such a way that the traditional structure and partition memory works as feedback loops. RNN (residual neural network) was designed and developed using MATLAB Resnet-50 and it successfully predicted tobacco diseases upon training it with a given dataset. The RNN algorithm is proven to be better than the DNN algorithm and it is a little less than the CNN algorithm. The findings show that RNN is better than current deep and machine-learning models. The results for the detecting and grading of leaf samples are also more accurate than other approaches. The use of CNN reveals that segmented and non-segmented experimental results of images. The classification of images can effectively be as high as 97.23% for grading tobacco and it took 6.78ms to classify a single leaf for a tobacco. This method and approach makes easier for grading and detecting the disease effectively.

REFERENCES

[1]. Abu-Naser, S. S., Kashkash, K. A., & Fayyad, M. (2010). Developing an expert system for plant disease diagnosis. Journal of Artificial Intelligence; Scialert,3(4), 269-276.
[2]. Jiyu Peng, Kunlin Song "Fast detection of tobacco mosaic virus infected tobacco using laser-induced breakdown spectroscopy," March 2017
[3]. Abu-Nasser, B. S., & Abu-Naser, S. S. (2018). Cognitive System for Helping Farmers in Diagnosing Watermelon Diseases. International Journal of Academic Information Systems Research (IJAISR), 2(7), 1-7.
[4]. Abu Naser, S. S., & Shaath, M. Z. (2016). Expert system urination problems diagnosis. World Wide Journal of Multidisciplinary Research and Development, 2(5), 9-19.
[5]. J. Liu, J. Shen, Z. Shen, and R. Liu, "Grading tobacco leaves based on image processing and generalized regression neural network," in Proc. International Conference on Intelligent Control, Automatic Detection and High-End Equipment, 2012, pp. 89-93.

[6]. Y. Sari and R. A. Pramunendar, "Classification quality of tobacco leaves as cigarette raw material based on artificial neural networks," Int. J. Comput. Trens Technol., vol. 50, no. 3, pp. 147-150, August 2017.
[7]. A. Santhosh, K. Saravanakumar, P. Saravanakumar, and S. Vimalraj, R. Saranya "DETECTION OF LEAF DISEASES AND CLASSIFYING THEM USING MULTICLASS SVM, " March, 2019.
[8]. S. K. Dasari, K. R. Chintada, and M. Patruni, "Flue-cured tobacco leaves classification: A generalized approach using deep convolutional neural networks," Cognitive Science and Artificial Intelligence Advances and Application, pp. 13-21, 2018.
[9]. Yi, S. C., Hajaligol, M. R., & Jeong, S. H. (2005). The prediction of the effects of tobacco type on smoke composition from the pyrolysis modeling of tobacco shreds. Journal of analytical and applied pyrolysis, 74(1-2), 181-192
[10]. W. S. Jeon and S. Y. Rhee, "Plant leaf recognition using a convolutional neural network," International Journal of Fuzzy Logic and Intelligent Systems, vol. 17, no. 1, pp. 26-34, March 2017
[11]. S. Pan, M. Kudo, and J. Toyama, "Edge detection of tobacco leaf images based on fuzzy mathematical morphology," in Proc. International Conference on Information Science and Engineering, 2009, pp. 1219-1222.
[12]. V. Premalatha, M. Sumithra, S. Deepak, and P. Rajeswari, "Implementation of spatial fem for leaf image segmentation in pest detection," Int. J. Adv. Res. Comput. Sci. Softw. Eng., vol. 4, no. 10, pp. 471-477, October 2014.
[13]. P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in Proc. the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, 2001, p. 1.
[14]. Mohanad H. Al-Qadi, Mohammed F. El-Habibi "Developing an Expert System to Diagnose Tomato Diseases," Pages: 34-40, May 2022
[15]. Abu Naser, S. S., & Zaqout, I. S. (2016). Knowledgebased systems that determine the appropriate students major: In the faculty of engineering and information technology. World Wide Journal of Multidisciplinary Research and Development, 2(10), 26-34.