PREDICTION OF DISEASE IN TOBACCO LEAF USING DEEP BELIEF NETWORK

D. Sindhu
Department of Computer Science and Engineering, Anna University, India

Abstract

The quality of crop yield declines as far as leaf diseases in agriculture are concerned. Leaf diseases may therefore be recognised immediately, in order to increase plant yields. Most of the device is therefore impaired by the absence of different patterns of the leaf disease that impair the detection accuracy. In this paper we build an IT model that helps shape the context of collecting images, extracting features and classifying images in real time. The classifiers indicate the outcome, whether or not the leaf is ill. We use the Deep Belief Network (DBN) in this paper to categorise images in real-time. Tobacco plant experimental findings indicate that a higher sorting rate has been suggested than other current approaches.

Keywords:

Real-Time Image Acquisition, Artificial Neural Network, Leaf Disease Detection, Tobacco Plant

1. INTRODUCTION

Sustainable agricultural activity and climate change are closely linked to effective defence against plant diseases [1]. Research indicates that climate change can change pathogenicity and concentrations, change tolerance to hosting and change the physiological interactions between host and pathogen [2]. With viruses circulating across the globe more exponentially than ever before, the situation is complex. There may be new pathogens where they are not known previously and local knowledge cannot be used to fight them [3].

Tobacco is a major economic crop. At the moment, disease detection is based on the conventional approach based on the experience of tobacco growers, which often causes economic damage due to incorrect and premature diagnostics. Agricultural disease image perception attracts a wide scientific focus in the advancement of computer vision technologies. Studies on tobacco images are however minimal.

Tabacco Mosaic Virus (TMV) is a crippling worldwide epidemic - can quickly impact the consistency and yield of tobacco. In order to reduce global loss of sustainable agriculture, it is important that the TMV inoculated tobacco disease is correctly detected at an early stage of infection for plant disease prevention and control. The identification of diseases on plant leaves is historically focused on direct and indirect methods. Visual analyses, serology and molecular techniques are primarily direct methods. Indirect approaches are based on detection technology based on biomarkers. Sadly, these plant disease detection approaches are timely, ineffective and harmful. An advanced and timely approach is also important for early knowledge in diverse areas about disease diagnosis, which will make disease control simpler and increase efficiency by proper management strategies.

Long-term pathogens can develop resistance and significantly lower battle capabilities by accidental use of pesticides. The prompt and accurate detection of plant diseases [4] is the cornerstone of precision agriculture. Financial and other services cannot be disproportionately expended and best controlled by addressing the problem of developing long-term pathogen tolerance and reducing climate change adverse reactions.

In this changing world there has been never more adequate, prompt monitoring of illnesses, including early avoidance. Pathologies of plants can be observed in various ways. Some diseases have no apparent impact, or the affect is too hard to operate and in these situations a thorough analysis is required. However, certain pathogens induce a certain kind of manifestation of the visual spectrum, such that the main technique for plant identification is advanced research.

In order to get accurate diagnosis of plant diseases [6], a plant pathologist should be outstanding analysts in identifying characteristic signs. Variations of the symptoms of diseased plants can lead to improper diagnosis as amateurs and hobbies are more difficult to determine than qualified pathologists. An electronic computer that recognises the conditions and visual signs of the plant as a verifier of disease diagnosis may provide amateurs in the garden sector and trained practitioners with substantial benefits.

Computer vision improvements provide opportunities [10] - [16] to broaden and consolidate effective plant safety practises as well as to boost demand for specifically implemented agricultural computer vision. The detection and identification of plant diseases was achieved by common digital imaging techniques, such as colour analyses and thresholds [5].

Deep learning is a recent trend in deep learning which offers cutting-edge results in many fields of science, including computational vision and bioinformatics. A variety of methods are currently available for deeper learning on the identification of plant diseases and the most commonly used are the Convolutional Neural Network (CNN). Deep learning [8] [9] gains from the opportunity to use data directly without man-made materials [7].

In this paper we try to introduce a deep method of learning for classification of plant diseases, focusing mainly on the diseases found in the leaf pictures. This paper establishes a computer vision framework by framing a model consisting of acquiring images, removing features and classifying images. For the analysis of images in real time, a deep research classifier known as Deep Belief Network (DBN). The research results for tobacco plants suggest that the rates of classification have improved compared to other current methods by the proposed process. The test results demonstrate whether the leaf is sick or not.

2. RELATED WORK

Related Work Manual [1] [2] or semi-automatic [3] methods for the study of the fixed disease categories are frequently used for directions modelling features. However, manual methods must solve environmental problems [4]. The possibility of automatic extraction of features is provided by deep learning technology [5] [6]. While images are generally collected in controlled conditions in public plant datasets.

Study [7] have shown that when using the model trained on the PlantVillage data set in the real environment, precision is rapidly decreasing. Duo learning was used to address the absence of statistics because of the difficulty of field collection [8]. Furthermore, for weakly supervised finely grained classes, the visual care mechanism is effective. In order to locate the distinguishing regions in the image the attention proposal network (APN) [9] was put forward. The squeeze/excitation (SE) [10] networks were designed to explore the dependencies between the channels.

The study [11] developed the Residual Care Network composed of a number of Residual Care Modules with residual learning-based constraints. The classification of the fine grain on a small dataset remains challenging, as it is difficult for CNNs to extract high levels from limited data. Moreover, environment factors can easily affect the quality of the images collected in the field. In this article, the method proposed is to resolve the problems described above.

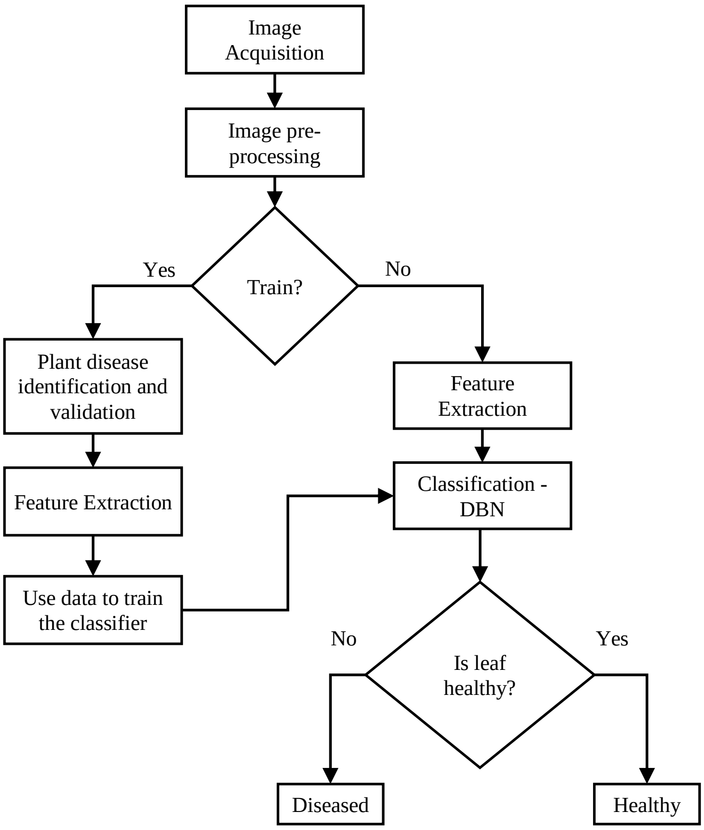

Fig.1. Prediction architecture using DBN model

3. METHODS

For the classification of tobacco plants, the DBN classification system outlined in Fig.1 is used. The classification system consists of: image collection, pre-processing of images, elimination of functions and classification. The classification scheme is designed to assign the leaf to a normal or malignant tobacco leaf.

- *Image Acquisition*: These images are captured using a different orientation, different format, backdrop, illumination and positions digital camera.
- *Pre-processing*: The system can find in the inventory details of local as well as global repositories a vast volume of safe and sick leave images. Three RGB channels are available in each frame. Each image in our data set will be converted into a grid of 256×256 pixels in the preprocessing step. When a pixel is chosen for cleaning purposes, noise is extracted from image samples taken.

3.1 FEATURE EXTRACTION

Gray Level Co-occurrence Matrix is used for the extraction of features such that the matrix location is counted with the transformation of the Scale Invariant function Transform (SIFT).

3.2 CLASSIFICATION USING DBN

DBN requires structured data and additional outputs are provided by understanding these structured data with additional data sets and human interference is required if desired output is not obtained. Therefore, no human interference is required. One of Deep Learning's benefits is the potential to train and capture high-level functionality in a gradual order that reduces the need for field experience.

The deep learning strategies gain prominence because of their skillful structure. DBN is a change in typical neural networks that decay exponentially when rotating through recurrent networks with the input and the output results of the hidden layers. This is modified once more by modifying the hidden neuron structure in the long-term memory of DBN.

The binary variant of the factor analysis can be treated as restricted Boltzmann machine (RBMs). Thus, a binary variable dictates the network output, instead of having several variables. The RBNs allow the generative weight of the cached units to be trained more effectively. These secret units are learned to catch the associations of higher order of data contained in visible units. The generative weights are obtained by an unmonitored layer-by-layer process, allowing a difference in contrasts (Hinton, 2002). The RBN training method known as the sampling of Gibbs begins with the introduction of a vector, which transfers values to the hidden units. The observable unit inputs are inverted to recreate the initial input stochastically. Then, these new apparent neuron activations are forwarded such that secret unit activations can be accomplished by one-step reconstruction.

3.3 LONG SHORT TERM MEMORY MODELS (LSTM)

LSTM is emerging in a recurrent neural network (RNN). Last layer performance is supported for standard RNN modules over a special Tanh function while LSTMs are used for feedback loops and gates. The LSTMs consist of four complex layers of NNs in each node, a module, an input path, an exit door and a forgotten gate where the cell maintains values over autonomous interval times. LSTM can add or remove data from module status (which is the main flow chain) using gates from sigmoid.

LSTM model is a supervised deep learning form which is exceptionally effective in the production of time series predictions. In this situation, information is transferred to certain cell states by means of processing.

This step is used to train powerful architectures on a broad data set such as ImageNet. The goal is to enable the weight of the network for the next step. The network is improved from the first step. We also have the current performance stage that replaces the pretrained network output layer for two kinds of tobacco diseases.

4. RESULTS AND DISCUSSIONS

The evaluation of the proposed classification model in a high-end computing system with the Matlab Framework is discussed in this section. Initially, DBN classification output is measured on tobacco leaves, which are 1500 images of good and diseased leaves.

We used data improvement techniques, such as cropping, random rotation, random tossing, and Gaussian noise to address the issue of inadequate data in the dataset, in order to extend the training collection. Provided that the color characteristic is a distinctive feature of lesions, there are no improvement approaches for color-dithering. In comparison, the quantities of data in various groups in the initial dataset vary considerably. And by monitoring the volume of updated data currently used, we retain data balance.

The DBN classifier is tested against accuracy, sensitivity, specificity and f-measure. The DBN framework is compared against back propagation neural network (BPNN), feed forward neural network (FFNN), Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) to validate the accuracy of the classifier.

Table.1. Performance of 70% of training and 30% testing data

Statistical Parameter	FFNN	BPNN	RNN	CNN	DBN
Accuracy	56.90	57.20	59.55	60.91	81.70
F-measure	39.62	41.72	53.11	55.48	84.88
G-mean	73.76	73.99	75.53	75.94	86.80
MAPE	29.55	26.61	22.63	22.05	17.35
Sensitivity	62.97	66.48	86.77	87.43	97.47
Specificity	75.40	75.59	79.12	80.50	81.34

Table.2. Performance of 80% of training and 20% testing data

Statistical Parameter	FFNN	BPNN	RNN	CNN	DBN
Accuracy	97.13	97.15	97.25	97.27	97.33
F-measure	78.60	78.73	80.34	81.02	81.31
G-mean	80.66	80.90	82.17	82.49	82.68
MAPE	32.37	32.01	29.88	29.38	29.06
Sensitivity	65.68	66.04	68.17	68.67	68.99
Specificity	95.94	96.00	97.27	97.68	98.03

Table.3. Performance of 90% of training and 10% testing data

Statistical Parameter	FFNN	BPNN	RNN	CNN	DBN
Accuracy	98.66	98.66	98.74	98.76	98.81
F-measure	87.17	87.29	89.26	90.62	90.63
G-mean	95.32	95.32	95.76	96.09	96.13
MAPE	72.07	71.89	62.66	55.30	54.62
Sensitivity	91.82	91.82	92.76	93.50	93.57
Specificity	98.75	98.76	98.85	98.94	98.94

5. CONCLUSIONS

DBN develops a system which uses a number of frameworks, including the acquisition of images, the extraction of the corresponding characteristic and the classification of tobacco plant for the classification of the leaf disease. The deep classification of DBN classifies the photos of the tobacco leaf. DBN increased performance (86.18%) over existing models is verified by the performance. It offers more detailed results than other approaches for classifications of the leaf samples.

REFERENCES

[1] S. Selvakani Kandeeban and R.S. Rajesh, "A Genetic Algorithm Based Elucidation for Improving Intrusion Detection through Condensed Feature set by KDD 99 Data Set", *Information and Knowledge Management*, Vol. 1, No. 1, pp. 1-9, 2011.

[2] K.A. Garrett, S.P. Dendy and E.E. Frank, "Climate Change Effects on Plant Disease: Genomes to Ecosystems", *Annual Review of Phytopathology*, Vol. 44, pp. 489-509, 2006.

[3] J.G.A. Barbedo, "Digital Image Processing Techniques for Detecting, Quantifying and Classifying Plant Diseases", *Springer Plus*, Vol. 2, No. 1, pp. 660-673, 2013.

[4] H. Al-Hiary, S. Bani-Ahmad and M. Reyalat, "Fast and Accurate Detection and Classification of Plant Diseases", *International Journal of Computer Applications*, Vol. 17, No. 1, pp. 31-38, 2011.

[5] Y. Lu, S. Yi, N. Zeng and Y. Liu, "Identification of Rice Diseases using Deep Convolutional Neural Networks", *Neurocomputing*, Vol. 267, pp. 378-384, 2017.

[6] A. Krishnakumar and A. Narayanan, "A System for Plant Disease Classification and Severity Estimation using Deep learning Techniques", *Proceedings of International Conference on ISMAC in Computational Vision and Bio-Engineering*, pp. 447-457, 2018.

[7] S. Chakraborty and A.V. Tiedemann, "Climate Change: Potential Impact on Plant Diseases", *Environmental Pollution*, Vol. 108, No. 3, pp. 317-326, 2000.

[8] Sabri Arik, Tingwen Huang, Weng Kin Lai and Qingshan Liu, "Soil Property Prediction: An Extreme Learning Machine Approach", *Proceedings of International Conference on Neural Information Processing*, pp. 666-680, 2015.

[9] X. Yang and T. Guo, "Machine Learning in Plant Disease Research", *European Journal of Biomedical Research*, Vol. 3, No. 1, pp. 6-9, 2017.

[10] C. Hu, X. Wang and F. Wu, "Motioncast: On the Capacity and Delay Tradeoffs", *Proceedings of ACM International Symposium on Mobile Ad Hoc Networking and Computing*, pp. 18-21, 2009.

[11] Mostaque Md. Morshedur Hassan, "Current Studies on Intrusion Detection System, Genetic Algorithm and Fuzzy Logic", *International Journal of Distributed and Parallel Systems*, Vol. 4, No. 2, pp. 35-47, 2013.

[12] T. Amalraj Victoire and M. Sakthivel, "A Refined Differential Evolution Algorithm Based Fuzzy Classifier for Intrusion Detection", *European Journal of Scientific Research*, Vol. 65, No. 2, pp. 246-259, 2011.

[13] R. Ghadge, J. Kulkarni, M. Pooja, N. Sachee and R.L. Priya, "Prediction of Crop Yield using Machine Learning", *International Research Journal of Engineering and Technology*, Vol. 5, No. 2, pp. 31-37, 2018.

[14] Fabrizio Balducci, Donato Impedovo and Giuseppe.Pirlo, "Machine Learning Applications on Agricultural Datasets for Smart Farm Enhancement", *Machines*, Vol. 6, No. 3, pp. 21-38, 2018.

[15] M.A. Adejumobi, H.A. Hussain and O.R. Mudi, "PhysioChemical Properties of Soil and Its Influence on Crop Yield of Oke-Oyi Irrigation Scheme, Nigeria", *International Research Journal of Engineering and Technology*, Vol. 6, No. 4, pp. 1-8, 2019.

[16] T. Truong, A. Dinh and K. Wahid, "An IoT Environmental Data Collection System for Fungal Detection in Cropfields", *Proceedings of 30th International Conference on Electrical and Computer Engineering*, pp. 1-4, 2017.