Cassava Disease Classification using Deep Learning Algorithms

Maman Mythien S
mamanmythien.cs17@bitsathy.ac.in
Bannari Amman Institute of Technology
www.bitsathy.ac.in
Sathyamangalam

Prasanth K S
prasanthks.cs17@bitsathy.ac.in
Bannari Amman Institute of Technology
www.bitsathy.ac.in
Sathyamangalam

***Overview* – Cassava is one of the key food security crops cultivated by Africa's small-holder farmers since it can withstand harsh conditions. It is the 2nd largest carbohydrates provider in Africa. At least 80% of small-holder farmers residing in Sub-Saharan Africa cultivate cassava but they all suffered due to viral diseases affecting cassava plants which result in poor yielding. In this project, we are going classify the plants affected by various diseases such as Cassava Bacterial Blight virus (CBB), Cassava Mosaic viral Disease (CMD), Cassava Brown Streak viral Disease (CBSD), Cassava Green Mottle Virus (CGM) along with the healthy ones based on the images.**

For this purpose, we used a deep learning technique known as Convolutional Neural Network (CNN) which was implemented with the help of TensorFlow's inception model. More than 8000 images were used for training and testing purpose. We stored and retrieved images with the help of NoSQL database implemented through CouchDB. By implementing this project in cassava cultivation will definitely reduce the poor yielding and loss of farmers.

Keywords—convolutional neural network, tensorflow, inception model, NoSQL, CouchDB

I. INTRODUCTION

Cassava is the largely cultivated food crop in Africa due to its richness in carbohydrates and proteins and it can also grow in a wide range of conditions. Due to that, in Africa, more than 500 million people depend on the root of cassava as their main staple and an estimated 800 million people worldwide consume cassava. For many small-holder farmers in Africa, cassava is the primary source of food and income. They suffer a lot when their crops were affected by viral diseases such as affect Cassava Bacterial Blight virus (CBB), Cassava Mosaic viral Disease (CMD), Cassava Brown Streak viral Disease (CBSD), Cassava Green Mottle Virus (CGM). These are the most common diseases that cassava cultivation in a deadly way. The major problem for them is to identify which type of disease attacks their plant. For this problem, we created a model which can classify plants affected by the above viral diseases depend on the images. A dataset of 5 fine-grained cassava leaf disease categories with 9,436 labeled images was used as a training image for our model. We used TensorFlow's inception v3 model which composes of Convolutional Neural Network (CNN) to classify those images. The images are getting stored and retrieved along with classified results with the aid of NoSQL through CouchDB. Classifying diseased plants will help the farmers to understand the situation easier. It helps them to analyze which type of medicines and nutrients would help their cultivation efficiently.

II. METHODOLOGIES IMPLEMENTED

A. *Convolutional Neural Network (CNN)*

Convolutional Neural Network is a different type of feed-forward artificial neural network which is inspired by visual cortex. The visual cortex is nothing but a small region presents in our brain that is sensitive to specific regions of the visual field which enables us to classify images seen by us. In Convolutional Neural Network, the neuron in a layer will only be connected to a small region of the layer before it, instead of all of the neurons in a fully-connected manner which is used in Fully Connected Networks. CNN consists of the following layers; convolution layer, pooling layer, ReLu layer, and Fully Connected layer. CNN compares the piece of the image by piece. The pieces that it looks for are called features.

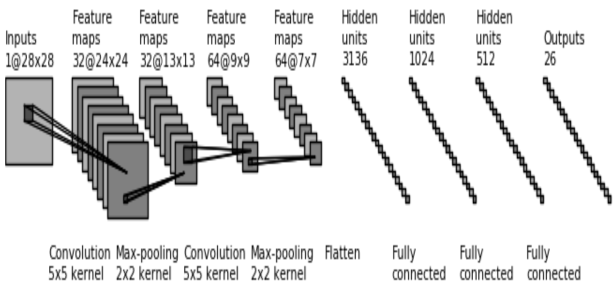


Fig 1: Architecture of Convolutional Neural Network

For each feature, a matrix of numbers will be created which gets reduced by walking through the above layers one by one. Finally, an array will be getting as an output for the Fully Connected layer. Similarly, an array will be getting from the test image. To see similarity between two images, CNN matches the rough features of one image with the other, in roughly the same positions. By doing so, the image classification of CNN gets accurate.

B. *Inception V3 Model*

Inception model consists of two important parts; convolutional network's fully extraction part and fully connected network's classification part. In the first part, the general features of the images were extracted from the input while in the rest the images are classified based on those features. The inception v3 model is a pre-trained deep learning model which attains the state-of-the-art precision in identifying general objects. It contains many layers and many networks and, in each layer, a feature gets extracted and saved for classification

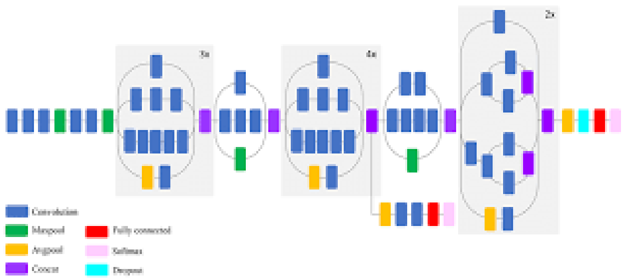


Fig 2: Inception V3 Model Structure

III. PoC Architecture

In our model, more than 8000 images were given for training purpose. Our model extracts many features for each and image using a convolutional neural network. For each feature, a matrix of vector values was assigned. While flowing through the model, the matrix values get reduced and in the last an array of values will remain. Our model saved those array values for all images along with their labels. When an unknown image is given, the same process is applied to extract features and then the extracted features are compared with the saved features of training images using fully connected networks. Based on the accuracy, the result will be displayed. The highest accuracy of our model so far is 95.3%

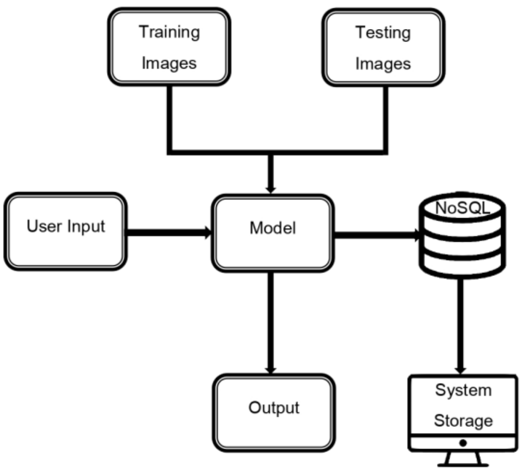


Fig 3: Cassava Disease Classification System

The user's input image is being fed into our model. After classification, the image along with their results is being stored in a NoSQL database with the help of CouchDB. The data present in the NoSQL database gets converted into a CSV file and stored in the local system for better visualization.

IV. PoC Design

For the more user-friendly purpose, we have included Graphical User Interface (GUI) for including images in our model. Our model can run for a single image as well as for a bunch of images so that user can clearly verify the disease for a single plant with lots of images of that particular plant in different angles.

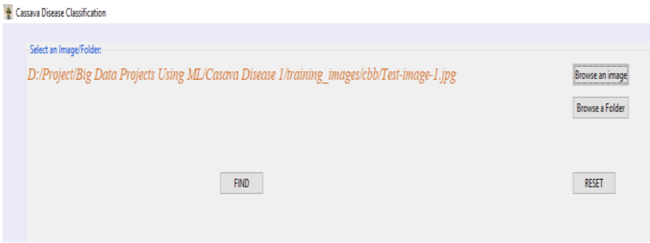
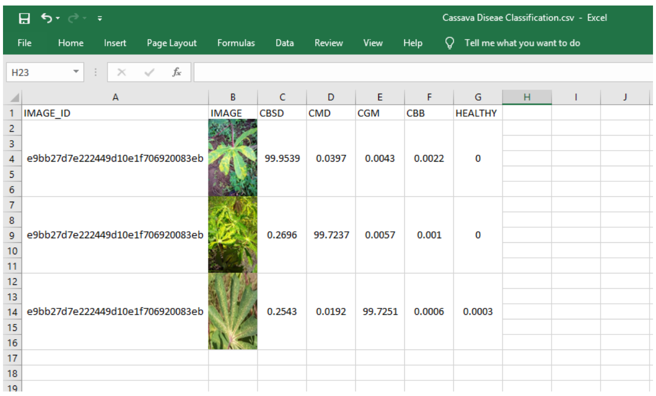


Fig 4: Cassava Disease Classification System GUI

The images with the classification result will be displayed in a separate window for more efficiency.

Fig 5: Result Window with Classification

The results are being stored in a CSV format so that user can easily track a record of which plants affected by which diseases along with the percentage of accuracy.



IMAGE_ID	IMAGE	CBSD	CMD	CGM	CBB	HEALTHY
e9bb27d7e222449d10e1f706920083eb		99.9539	0.0397	0.0043	0.0022	0
e9bb27d7e222449d10e1f706920083eb		0.2696	99.7237	0.0057	0.001	0
e9bb27d7e222449d10e1f706920083eb		0.2543	0.0192	99.7251	0.0006	0.0003

Fig 6: Results are Exported in the CSV Format

With the implementation of this model, users can easily identify and take the necessary steps in case any disease affects their plant.

V. Innovations Implemented

Our model can store a large volume of images since it uses a NoSQL database. We developed our model in a way that that large volume of images doesn't affect the model's speed and efficiency. We have implemented sequential processing where images are processed one by one and each time when an image is processed the following image in the sequence gets processed in a background format. The first processed image gets visualized in a window manner and the next image is only be visualized when the first window gets closed. The most interesting part is that each time an image is analyzed, it gets stored in the NoSQL database in the background.

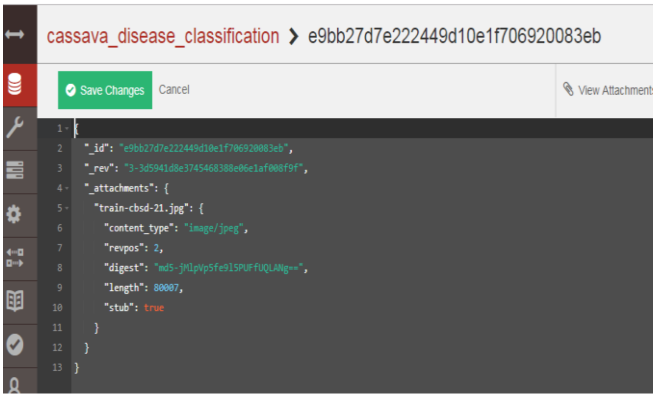


Fig 7: Results Stored in a NoSQL Database

Our model can enhance the life of small-holder farmers since it helps in yielding lossless cultivation. It also helps botanists for their research purposes where the relation between one and other cassava diseases are known.

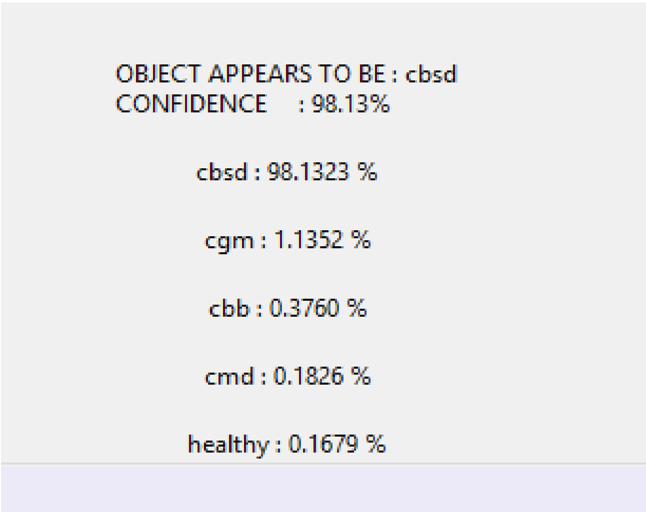


Fig 8: Output for a Particular Image

VI. Complications Faced

The most difficult part while creating our model is to process a group of images in a sequential manner. In my model when one image gets processed, the other needs to be processed too for time efficiency. While implementing that we struggle a lot. Then the crucial part arrives when we need to connect our model to the NoSQL database so that the results can be stored into it along with the images. We handled that with the help of the CouchDB server and python programming interface. Then we need to store and display the data in a CSV format which makes it easy for researchers and farmers to handle a large dataset or field efficiently. We had also overcome that problem with the help of NumPy and Pandas packages.

VII. Conclusion

With the help of our model along with 95.3% accuracy in prediction, we can easily detect and classify diseases which results in high yielding of cassava. This helps those small-holder farmers for leading their life without the fear of slight yielding. We can develop our model to classify not only the selected diseases but all the diseases that could possibly affect cassava plants. We can even enhance our project to classify all possible diseases for all vital food and economic crops in the entire world by implementing cloud computing. Implementing our model will provide a better life for farmers.

Project Demo : https://youtu.be/PKUFQEhX_XI

References

[1] H. Park, J. S. Eun, S. H. Kim, "Image-based disease diagnosing and predicting of the crops through the deep learning mechanism", In Information and Communication Technology Convergence (ICTC) IEEE 2017 International Conference on, pp. 129-131, 2017.C.

[2] K. Elangovan, S. Nalini, "Plant disease classification using image segmentation and SVM techniques", International Journal of Computational Intelligence Research, vol. 13, no. 7, pp. 1821-1828, 2017L.

[3] S. H. Lee, C. S. Chan, S. J. Mayo, P. Remagnino, "How deep learning extracts and learns leaf features for plant classification", Pattern Recognition, vol. 71, pp. 1-13, 2017.

[4] K. P. Ferentinos, "Deep learning models for plant disease detection and diagnosis", Computers and Electronics in Agriculture, vol. 145, pp. 311-318, 2018.

[5] H. Durmuş, E. O. Güneş, M. Kırcı, "Disease detection on the leaves of the tomato plants by using deep learning", Agro-Geoinformatics IEEE 6th International Conference on, pp. 1-5, 2017.

[6] H. A. Atabay, "Deep residual learning for tomato plant leaf disease identification", Journal of Theoretical & Applied Information Technology, vol. 95, no. 24, 2017.

[7] V. Singh, A. Misra, "Detection of plant leaf diseases using image segmentation and soft computing techniques", Information Processing in Agriculture, 2016.

[8] S. Zhang, W. Huang, C. Zhang, "Tree-channel convolutional neural networks for vegetable leaf disease recognition", Cognitive Systems Research, 2018.

[9] D. P. Hughes, M. Salathé, "An open access repository of images on plant health to enable the development of mobile disease diagnostics", 2015.

[10] Y. LeCun, Y. Bengio, G. Hinton, "Deep Learning", Nature, vol. 521, pp. 436-444, 2015.

[11] V. Tümen, Ö. F. Söylemez, B. Ergen, "Facial emotion recognition on a dataset using convolutional neural network", 2017 International Artificial Intelligence and Data Processing Symposium (IDAP), November 2017.

[12] B. Mokbel, B. Paassen, F. M. Schleif, B. Hammer, "Metric learning for sequences in relational LVQ", Neurocomputing, vol. 169, pp. 306-322, 2015.

[13] X. Li, Y. Zhang, "Digital image edge detection based on LVQ neural network", 2016 IEEE 11th Conference on Industrial Electronics and Applications (ICIEA), pp. 1251-1255, 2016.

[14] Savita N. Ghaiwat, Parul Arora, "Detection and Classification of Plant Leaf Diseases Using Image processing Techniques: A Review", International Journal of Recent Advances in Engineering & Technology, vol. 2, no. 3, 2014.

[15] Prof Sanjay B. Dhaygude, Mr. Nitin P. Kumbhar, "Agricultural plant Leaf Disease Detection Using Image Processing", International Journal of Advanced Research in Electrical Electronics and Instrumentation Engineering, vol. 2, no. 1, January 2013.

[16] S. J. G. A. Barbedo, "A new automatic method for disease symptom segmentation in digital photographs of plant leaves", European Journal of Plant Pathology, vol. 147, no. 2, pp. 349-364, 2016.

[17] J. G. A. Barbedo, "A novel algorithm for semi-automatic segmentation of plant leaf disease symptoms using digital image processing", Tropical Plant Pathology, vol. 41, no. 4, pp. 210-224, 2016.

[18] J. G. A. Barbedo, L. V. Koenigkan, T. T. Santos, "Identifying multiple plant diseases using digital image processing", Biosystems Engineering, vol. 147, pp. 104-116, 2016

[19] Vijay Jumb, Mandar Sohani, Avinash Shrivas, "Color Image Segmentation using K-means Clustering and Otsu's Adaptive Thresholding", International Journal of Innovative Technology and Exploring Engineering (IJITEE), vol. 3, no. 9, February 2014.