Dec,2023 A computer and a world map

Description automatically generated with medium confidence

Mango Leaf Diseases Classification using CNNs.

Moustafa Omar Mohammed, Abdel-Rahman Ramadan Abuela, and Abdullah Farghl Abdullah

Student, Department of Artificial Intelligence, Cairo University, Egypt.

Email: [imoustafaomar.fci@gmail.com](mailto:imoustafaomar.fci@gmail.com)

Student, Department of Artificial Intelligence, Cairo University, Egypt.

Email: [bodya9936@gmail.com](mailto:bodya9936@gmail.com)

Student, Department of Artificial Intelligence, Cairo University, Egypt.

Email:  [abdallakamal365@gmail.com](mailto:manhas.jatinder@gmail.com)

**Abstract:**

**Agriculture faces many plant diseases that affect their quality, one of these plants that face a lot of diseases challenge is mango.**

**Traditional methods of disease detection using the naked eye or manual inspections are often unreliable and time-consuming, making it imperative to develop more accurate and efficient techniques for early disease detection, advancements in computer vision and machine learning have paved the way for the development of automated disease detection systems for agricultural crops. This project focuses on the application of Convolutional Neural Networks (CNNs), a deep learning architecture specialized in image analysis, to address the problem of mango leaf disease detection. CNNs have demonstrated remarkable success in image recognition tasks, making them a suitable choice for this project.**

**Keywords: Convolution Neural Network (CNN), Crop, Deep learning, Image classification, Mango.**

# **Introduction**

Mango (Mangifera indica), a tropical product of huge monetary significance and inescapable development, is famous for its delightful taste as well as for its bunch wellbeing benefits. Regardless of its conspicuousness, the mango business defies a relentless danger from different sicknesses focusing on the leaves of mango trees. If uncontrolled, these illnesses can possibly significantly reduce both the quality and yield of mangoes. Customary ways to deal with sickness recognizable proof, depending on visual perception or manual assessments, frequently end up being inconsistent and tedious. Thusly, there is a convincing need to foster more exact and productive methods for the early discovery of illnesses.  
This model was trained on dataset that has 8 classes, each class represents one of mango leaf disease and one class represents healthy leaf, [Anthracnose, Bacterial Canker, Cutting Weevil, Die Back, Gall Midge, Powdery Mildew, Sooty Mould], with applying suitable data preprocessing methods(e.g. resize images data, image generator), feature extraction with VGG-16 model, feature selection with optimizer technique (e.g. Lasso and Ridge regularization ), classification algorithm with KNN algorithm. By utilizing innovation, one can recognize infections with an enormous scope. On account of mango leaves, there are different kinds of illnesses present, like fine buildup, anthracnose, red rust, and so on. In the current work, a profound learning (DL)-based model has been proposed for the grouping of different mango leaf illnesses (fine buildup, anthracnose, and red rust) at the underlying stages. Exactness, review, accuracy, and F-score have been utilized to assess the model.

1. **EXPERIMENTAL RESULT**

**Dataset and Preprocessing**

the dataset sourced from Kaggle , which collected from Four mango orchards of Bangladesh, namely Sher-e-Bangla Agricultural this dataset consists from 4000 images around 1800 are of distinct leaves, and the rest are prepared by zooming and rotating where deemed necessary which sized as 240\* 320 images were categorized into seven diseases

A close up of a plant

Description automatically generated A green leaf with brown spots

Description automatically generated A green leaf on a white background

Description automatically generated

namely Anthracnose, Bacterial Canker, Cutting Weevil ,Die Back, Gall Midge, Powdery Mildew, and Sooty Mould.

Number of classes: Eight (including the healthy category). Each of the eight categories contains 500 images preprocessing.

All images are resized as 224\*224\*3 pixels and was divided to train data about 2560 samples and test data about 800 samples and validation data about 640 samples.

One hot encoding was used to encode categorical labels .

A collage of a leaf

Description automatically generated A collage of a leaf

Description automatically generated

**Evaluation metrics**

**Recall**

Known as sensitivity , it means the ability of model to find positive samples , recall score means that the model correctly identified all the TP samples, with no FN

**F1 score**

It is combination of recall and precision score The F1 score is the harmonic mean of the precision and recall

**Precision**

precision answers the question: Out of all the instances predicted as positive, how many were actually positive?, high precision score means that the model is effective with positive predictions

A screenshot of a computer

Description automatically generated

**Confusion matrix**

It is a table used to evaluate the performance of the algorithm by comparing the predictions of the algorithm and actual values, it is square table that has 4 categories

**TP**: correctly predicted as positive(actual and predictions both are positive)

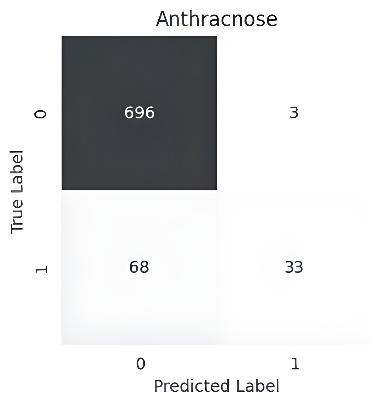
**TN**: correctly predicted as negative(actual and predictions both are negative)

**FP**: incorrectly predicted as positive (predicted as positive and actual is negative)

**FN**: incorrectly predicted as negative (predicted as negative and actual is positive)

A list of positive results

Description automatically generated



Figure

A black and white chart

Description automatically generated

Figure

**PROPOSED MODEL**

Model consists of 3 mainly stages.

**-pretrained model**

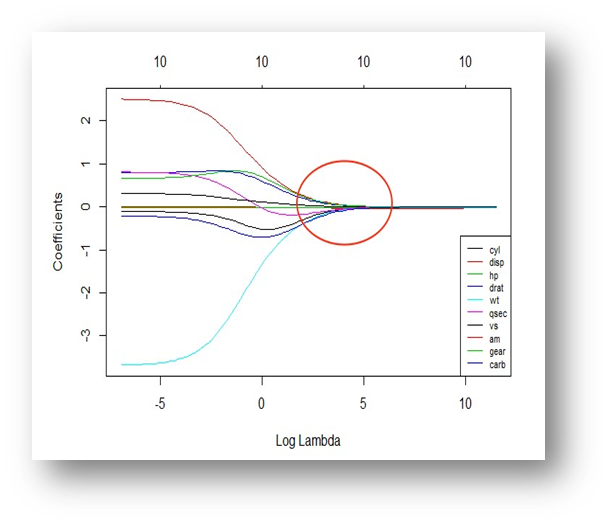
We used VGG-16 to train this model for feature extraction, this architecture mainly was applied with ImageNet dataset , it is one of the most used CNN architecture , The convolutional layers use small 3x3 filters with a stride of 1, Max pooling is applied after some of the convolutional layers, and Relu was used as activation function after each convolution layer and finally our data was used as shape 224\*224\*3 to be trained with this architecture

A diagram of a rectangular object

Description automatically generated

**-feature selection optimizers**

lasso and Ridge were used as optimizers for feature selection , this is effective stage to raise model accuracy and performance by choosing the most important features to use in our classifier , Lasso and Ridge can help with feature selection by shrinking the coefficients of less important features towards zero., This leads to automatic feature selection as features with zero coefficients are effectively excluded from the model. but Ridge tends to shrink them towards zero without making them exactly zero, Lasso helped us as model has many features that needed to be reduced but Ridge helped to get rid of highly correlated features that prevent any overfitting

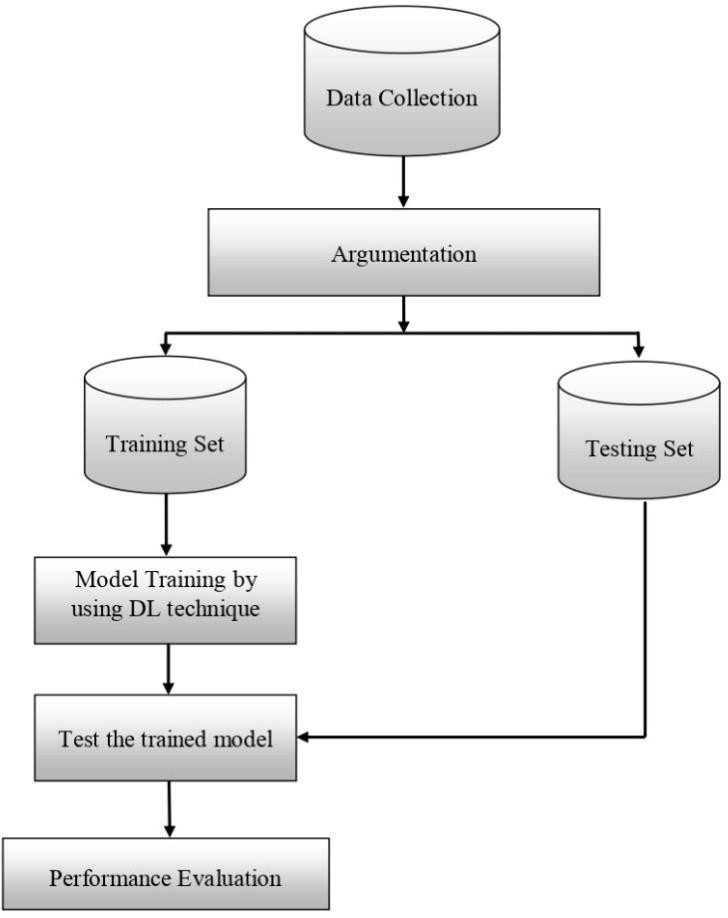


**- KNN algorithm**

KNN classifies or predicts its target value based on the majority class or average of its k nearest neighbors in the feature space, Euclidean distance is commonly used as the distance metric to measure the similarity between data points

A diagram of a training course

Description automatically generated



**Result and CONCLUSION**

In conclusion, we improved deep learning model to detect mango leaf diseases, by training model with VGG-16 architecture as CNN layers with max pooling and Relu activation function to extract features from data ,then by Lasso and Ridge selected the most important features and KNN algorithm classified data with high accuracy 86%

A graph with numbers and points

Description automatically generated

A graph with blue lines and numbers

Description automatically generated

# **Related work**

Leaf disease detection researches have a big history, they try many techniques using pattern recognition and machine learning algorithms to have better classification for the diseases they used Convolutional Neural Network [1] to perform feature extraction and classification together itself, Artificial Neural Network [2], Back Propagation Neural Network [3],Support Vector Machine [4] and other image processing methods [5,6]. Color Co-occurrence matrix [7], Angle Code Histogram [3], Zooming algorithm [8], Canny edge detector [5] and other algorithms for feature extraction and some of them don't need for preprocessing of images.

This graph show the difference between results of those papers:

Ma in [9] used the Deep Convolutional Neural Network (DCNN) to classify the diseases from real-time images.

Ferentinos in [10] used a VGG convolutional neural network to identify and classify the plant leaves from healthy and diseased images. The dataset was large and this help the deep learning to achieve high accuracy.

Too et al.in [11] used four different deep convolutional network architectures (VGG 16, Inception V4, ResNet and DenseNets) to classify the diseases from plantVillage dataset which consists of 54000 images have 52 classes (38 diseased- 4 healthy).

Barbedo in [12] presented many diseases and plants which affect the performance of the network.

In [13] Picon et al. used DCNN to classify three fungal diseases affect the wheat plant.

Zhang et al. in [14] he used GoogLeNet and Cifar10 network with real-time images called the maize leaf contains images have 9 different disease which was collected for two locations for three years.

In [2] Lu et al. proposed the Deep Convolutional Neural Network (DCNN) to classify ten kinds of rice leave diseases from dataset of about 500 images containing the infected and healthy images. Authors also used the 10-fold cross validation to achieve better results.

Gandhi et al [15] used Generative Adversarial Networks (GANs) and CNN to identify diseases from the leaf images using a mobile application. he used AlexNet and SqueezeNet deep learning network

Durmus et al. in [16] classify plant leaf diseases. from the images which taken from the plantVillage database (tomato leaf images) with 10 classes.

Jain et al. in [17] proposed Convolutional Neural Network (CNN) for the real-time classification of the diseases from the plant leave images. on a cloud environment.

1. **References**

[1] Y. Lu, S. Yi, N. Zeng, Y. Liu, and Y. Zhang, ‘‘Identification of rice diseases using deep convolutional neural networks,’’ Neurocomputing, vol. 267, pp. 378–384, Dec. 2017. doi: 10.1016/j.neucom.2017.06.023. [https://www.sciencedirect.com/science/article/pii/S0925231217311384]

[2] Elham Omrani, Benyamin Khoshnevisan, Shahaboddin Shamshirband, HadiShaboohi, Nor Badrul Anuar, MohdHairul nizam md nasir, (2014), Potential of redial basis function- based support vector regression for apple disease detection, Measurement, vol.55,pp:. 512-519. [https://www.semanticscholar.org/paper/Potential-of-radial-basis-function-based-support-Omrani-Khoshnevisan/06533c1dd5b8ed522662c67624bd09c2343a09c3]

[3] Kshitij Fulsoundar , Tushar Kadlag , Sanman Bhadale , PratikBharvirkar, (2014), Detection and classification of plant leaf diseases, International Journal of Engineering Research and General Science Volume 2, Issue 6. [https://www.semanticscholar.org/paper/DETECTION-AND-CLASSIFICATION-OF-PLANT-LEAF-DISEASES-Fulsoundar-Kadlag/d36eef135e3cd2867ef8d49980726e112502fee2]

[4] Shivaputra S.Panchal, Rutuja Sonar, (2016). Pomegranate Leaf Disease Detection Using Support Vector Machine, International Journal Of Engineering And Computer Science ISSN: 2319-7242 ,Volume 5 Issues 6. [https://www.technoarete.org/common\_abstract/pdf/IJERECE/v4/i8/Ext\_68231.pdf]

[5] R.Preethi, S.Priyanka, U.Priyanka , A.Sheela, (2015), Efficient knowledge based system for leaf disease detection and classifi- cation, International Journal of Advance Research In Science And Engineering, Vol. No.4, Special Issue (01) https://www.semanticscholar.org/paper/EFFICIENT-KNOWLEDGE-BASED-SYSTEM-FOR-LEAF-DISEASE-Preethi-Priyanka/6ffed9a8da19c12b2eac3af4101f4b4a4c40d9f7]

[6] P.Revathi , M.Hema Latha, (2012), Classification Of Cotton Leaf Spot Diseases Using Image Processing Edge Detection Techniques International Journal of Applied Research ISBN, 169-173. https://ieeexplore.ieee.org/document/6513900]

[7] S. Arivazhagan, R. Newlin Shebiah , S. Ananthi, S. Vishnu Varthini, (2013), Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features, CIGR Journal, Vol. 15, No.1 211. https://www.researchgate.net/publication/287577015\_Detection\_of\_unhealthy\_region\_of\_plant\_leaves\_and\_classification\_of\_plant\_leaf\_diseases\_using\_texture\_features]

[8] Santanu Phadikar, Jaya Sil, (2008), Rice Disease Identification Using Pattern Recognition Techniques, Proceedings Of 11th International Conference On Computer And Information Technology, 25-27. https://ieeexplore.ieee.org/document/4803079]

[9] J. Ma, K. Du, F. Zheng, L. Zhang, Z. Gong, and Z. Sun, ‘‘A recognition method for cucumber diseases using leaf symptom images based on deep convolutional neural network,’’ Comput. Electron. Agricult., vol. 154, pp. 18–24, Nov. 2018. doi: 10.1016/j.compag.2018.08.048. https://www.researchgate.net/publication/328664988\_A\_recognition\_method\_for\_cucumber\_diseases\_using\_leaf\_symptom\_images\_based\_on\_deep\_convolutional\_neural\_network]

[10] K. P. Ferentinos, ‘‘Deep learning models for plant disease detection and diagnosis,’’ Comput. Electron. Agricult., vol. 145, pp. 311–318, Feb. 2018. doi: 10.1016/j.compag.2018.01.009. https://www.sciencedirect.com/science/article/pii/S0168169917311742]

[11] E. C. Too, L. Yujian, S. Njuki, and L. Yingchun, ‘‘A comparative study of fine-tuning deep learning models for plant disease identification,’’ Comput. Electron. Agricult., to be published. doi: 10.1016/j.compag.2018.03.032 https://www.sciencedirect.com/science/article/pii/S0168169917313303]

[12] J. G. A. Barbedo, ‘‘Factors influencing the use of deep learning for plant disease recognition,’’ Biosyst. Eng., vol. 172, pp. 84–91, Aug. 2018. doi: 10.1016/j.biosystemseng.2018.05.013[https://www.sciencedirect.com/science/article/pii/S1537511018303027]

[13] A. Picon, A. Alvarez-Gila, M. Seitz, A. Ortiz-Barredo, J. Echazarra, and A. Johannes, ‘‘Deep convolutional neural networks for mobile capture device-based crop disease classification in the wild,’’ Comput. Electron. Agricult., to be published. doi: 10.1016/j.compag.2018.04.002. [https://www.sciencedirect.com/science/article/abs/pii/S0168169917312619]

[14] X. Zhang, Y. Qiao, F. Meng, C. Fan, and M. Zhang, ‘‘Identification of maize leaf diseases using improved deep convolutional neural networks,’’ IEEE Access, vol. 6, pp. 30370–30377, 2018. doi: 10.1109/ACCESS.2018.2844405. - https://www.researchgate.net/publication/325612493\_Identification\_of\_Maize\_Leaf\_Diseases\_Using\_Improved\_Deep\_Convolutional\_Neural\_Networks

[15] R. Gandhi, S. Nimbalkar, N. Yelamanchili, and S. Ponkshe, ‘‘Plant disease detection using CNNs and GANs as an augmentative approach,’’ in Proc. IEEE Int. Conf. Innov. Res. Develop. (ICIRD), Bangkok, Thailand, May 2018, pp. 1–5. doi: 10.1109/ICIRD.2018.8376321. https://www.researchgate.net/publication/325911752\_Plant\_disease\_detection\_using\_CNNs\_and\_GANs\_as\_an\_augmentative\_approach

[16] H. Durmu‡, E. O. Güne‡, and M. Kırcı, ‘‘Disease detection on the leaves of the tomato plants by using deep learning,’’ in Proc. 6th Int.Conf. Agro-Geoinformatics, Fairfax, VA, USA, Aug. 2017, pp. 1–5. doi: 10.1109/Agro-Geoinformatics.2017.8047016. https://www.researchgate.net/publication/319971856\_Disease\_detection\_on\_the\_leaves\_of\_the\_tomato\_plants\_by\_using\_deep\_learning

L. Jain, M. A. H. Vardhan, M. L. Nishanth, and S. S. Shylaja, ‘‘Cloud-based system for supervised classification of plant diseases using convolutional neural networks,’’ in Proc. IEEE Int. Conf. Cloud Comput. Emerg. Markets, Nov. 2017, pp. 63–68. doi: 10.1109/CCEM.2017.22. [https://www.researchgate.net/publication/324381375\_Cloud-Based\_System\_for\_Supervised\_Classification\_of\_Plant\_Diseases\_Using\_Convolutional\_Neural\_Networks]