Leaf Disease Detection using Machine Learning

G. Senthilvelan
Professor(s), Dept .of CSE, Dr.MGR Educational and Research Institute,
senthilvelan.cse@drmgrdu.ac.in

Dr. V. Rameshbabu
Professor(s), Dept .of CSE, Dr.MGR Educational and Research Institute,
drvramesh25@gmail.com

Dr. D. Usha
Professor(s), Dept .of CSE, Dr.MGR Educational and Research Institute,
usha.cse@drmgrdu.ac.in

Pravalika
Final Year Btech CSE, Dr.MGR Educational and Research Institute,
pulipatipravalika@gmail.com

Anuhya
Final Year Btech CSE, Dr.MGR Educational and Research Institute,
anuhyareddivari@gmail.com

Saileela
Final Year Btech CSE, Dr.MGR Educational and Research Institute,
roddasaileela524@gmail.com

Abstract

An essential component of identifying plants for tracking plant growth is plant phenotyping. In this research, an effective method for distinguishing between healthy and sick or infected leaves utilising machine learning algorithms and image processing techniques is presented. Different diseases impact the chlorophyll in leaves, causing brown or black markings on the leaf surface. The productivity of agriculture is largely influenced by the economy. When switching from one disease management strategy to another, farmers experience major difficulties. We are able to recognise or spot tomato leaf diseases, which is the typical method for detection for surveillance and monitoring professionals. Convolutional Neural Networks (CNN) are a type of machine learning technique that are employed in classification.

Keywords: *Plant leaf disease images, deep learning, Machine Learning, SVC, ANN, CNN, Resnet50.*

I. INTRODUCTION

Agriculture has advanced significantly with the automated diagnosis of plant diseases using plant leaves. Additionally, crop output and quality are improved by the prompt and early detection of plant diseases. Even agronomists and pathologists frequently struggle to recognise plant diseases by observing diseased leaves due to the widespread cultivation of a variety of crop items. However, the major method of disease detection in rural parts of poor nations is still ocular inspection. Researchers have come up with a number of methods to address the issues mentioned above. Machine learning can be used to classify plant diseases using a variety of feature sets. Traditional handcrafted and deep learning (DL)-based feature sets are the most well-liked feature sets among them. Before effectively extracting features, pre-processing such as picture enhancement, colour modification, and segmentation is necessary. Following feature extraction, various classifiers may be employed. Climate change and sustainable agriculture are both issues that are directly tied to the issue of effective plant disease protection. According to research findings, climatic change can affect the stages and rates of pathogen growth as well as host resistance, which has an impact on how physiologically hosts and pathogens interact. The fact that infections are spread throughout the globe more readily than ever before further complicates the matter. It's possible for new diseases to appear in regions where they haven't been previously recognised and where there isn't enough local knowledge to treat them. When viruses develop long-term resistance due to inexperienced pesticide use, their ability to defend themselves is drastically reduced. One of the cornerstones of precision farming is the prompt and precise identification of plant diseases. By addressing the issue of long-term pathogen resistance development and reducing the negative effects of climate change, it is critical to prevent needless waste of financial and other resources and achieve healthy output.

II. LITERATURE SURVEY

Crop diseases pose a significant threat to food security, but because the necessary infrastructure is absent in many places around the world, it is still difficult to quickly identify them. Impressive achievements have been obtained in the field of leaf-based image classification since the emergence of precise approaches. In order to distinguish between healthy and unhealthy leaves from the generated data sets, this article uses Random Forest. The phases of implementation included in our proposal include dataset construction, feature extraction, classifier training, and classification.The produced datasets of sick and healthy leaves are pooled and trained under Random Forest to categorise the images of sick and healthy leaves. Overall, utilising machine learning to train the vast publically accessible data sets gives us a clear technique to detect the disease existing in plants on a massive scale. The farmer in rural areas could believe it is difficult to distinguish the disease that can be present in their harvests. It is not reasonable for them to visit the agricultural office and find out what the infection might be. Our main goal is to identify the disease that is introduced in a plant by observing its shape using image processing and machine learning. The accuracy of the results and the recognition rate have been improved by using a contemporary strategy like machine learning and deep learning algorithms. Machine learning has been the subject of numerous studies aimed at identifying and diagnosing plant diseases.(Maniyath et al., 2018)

Crop diseases must be promptly identified and prevented if productivity is to be increased. Since CNNs have demonstrated outstanding achievements in the field of machine vision, deep convolutional neural network (CNN) models are used in this paper to recognise and diagnose plant illnesses from their leaves. Standard CNN models need a lot of parameters and are more expensive to compute. In this study, we switched from ordinary convolution to depth=separable convolution, which lowers the number of parameters and lowers the computational cost.

14 different plant species, 38 distinct categorical illness classes, and healthy plant leaves made up the open dataset used to train the models that were ultimately used. Various factors, including batch size, dropout rates, and different epoch counts, were taken into consideration to assess the models' performance. Using InceptionV3, InceptionResNetV2, MobileNetV2, and EfficientNetB0, respectively, the developed models outperformed more conventional handcrafted feature-based approaches in terms of disease-classification accuracy, achieving rates of 98.42%, 99.11%, 97.02%, and 99.56%. The implemented deep-learning model performed more accurately and with less training time than other deep-learning models. Additionally, the optimised parameter makes the MobileNetV2 design compatible with portable electronics. The deep CNN model is promising and can significantly influence the effective identification of the diseases. It may also have potential in the detection of diseases in real-time agricultural systems. (Hassan, S. M et al., 2021)

The majority of the world's food supply comes from plants. Plant diseases are a factor in productivity loss, but they can be controlled with ongoing observation. Monitoring plant diseases manually is time-consuming and prone to mistakes. Artificial intelligence (AI) and computer vision can be used to identify plant illnesses early on, reducing their negative consequences while also overcoming some of the limitations of constant human monitoringIn this study, we have thoroughly examined the performance of the several state-of-the-art convolutional neural networks (CNNs) classification network designs, namely ResNet18, MobileNet, DenseNet201, and InceptionV3, on 18,162 simple tomato leaf images to classify tomato illnesses. The comparative effectiveness of the models for the binary classification (healthy and unhealthy leaves), the six-class classification (healthy and diverse categories of diseased leaves), and the ten-class classification (healthy and various types of unhealthy leaves) are also given. With a 99.2% accuracy rate for binary classification of plain leaf pictures, InceptionV3 demonstrated outstanding performance. With a 97.99% accuracy rate for six-class classification, DenseNet201 also outperforms the competition. Finally, DenseNet201 attained a ten-class classification accuracy of 98.05%. It can be said that for the three tests, deep structures were more effective at classifying the disorders. Every experiment mentioned in this article beats previous research in terms of performance. (Muhammad E.H Chowdhury et al. 2021)

This research suggests an enhanced Faster RCNN to detect healthy tomato leaves and four diseases, including powdery mildew, blight, leaf mould fungus, and ToMV, in order to increase the recognition model accuracy of crop disease leaves and pinpointing sick leaves. In order to obtain more detailed disease features, we first swap out VGG16 with a depth residual network during image feature extraction. Following that, the bounding boxes are clustered using the k-means clustering technique. Using the findings of the clustering, we enhance the anchoring. The enhanced anchor frame leans toward the dataset's actual bounding box. Finally, using three distinct feature extraction networks, we conduct a k-means experiment. Compared to the original Faster RCNN, the improved approach for agricultural leaf disease detection had a faster detection speed and an identification accuracy that was 2.71% higher. The first step in preventing agricultural diseases and ensuring crop quality is crop disease detection. Traditional detection methods for crop disease mainly depend on manual observation and consequently lead to low detection efficiency and poor reliability. Farmers lack professional knowledge, and agricultural experts cannot serve the field at all times so that they miss the best time (Y. Zhang et al 2020)

III. MATERIAL AND METHODS

This model emphasizes an existing method that which is designed using the some of the algorithms of deep learning. Here the process is performed using the machine learning, which is one of the transfer learning methods, but this could not get the high accuracy. The suggested technique is useful for tracking vast fields of crops. There should be solutions for detecting and classifying the disease to get some knowledge which will later help in improving the quality of plants. So, patterns on the plants leaves will help in identifying what problem it has.

Disadvantages of Existing System:

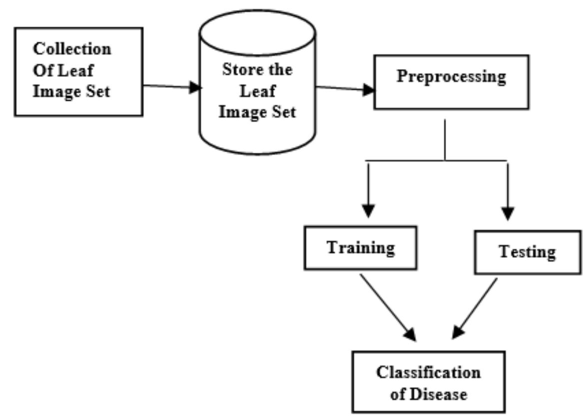
- Less feature compatibility.
- Low accuracy.
- The procedure is extremely slow.
- Consumption of time and space is also very high.
- Very few diseases have been covered. So, work needs to be extend to cover more diseases.

Convolution Neural Network (CNN) of deep learning together with machine learning techniques are being used in the purposed method to do the classification of either the Plant Leaf Disease detection. as techniques for detecting leaf disease based on image analysis. Therefore, accurate characterization of the Leaf illness is crucial, something that our suggested method will provide. Below is a block schematic of the suggested method. In order to categorise plant leaves as healthy or diseased, CNN, a machine learning technology, is utilised. If the plant leaf is diseased, CNN will identify the specific disease.

Advantages of Proposed System:

- Accurate Classification.
- Less Complexity
- High performance.
- Easy identification.
- Environment friendly of the identified.

Fig: 1 System Architecture



From fig 1, We have collected the data from the plant Village datasets. In plant Village datasets there are 54303 healthy and unhealthy leaf images divided into 38 categories by species and disease. In that categories we are collecting set of data such as potato and tomato leaves which are healthy and unhealthy.

The collected datasets are stored in the form of folder in our system.

Image preprocessing is used for the detection of plant diseases. Disease detection involves the steps like image acquisition, image preprocessing, image segmentation, feature extraction, and classification. Our paper discussed the methods used for the detection of plant diseases using their leaves images.

Training and Testing: In training, we will train the datasets of different leaf images and then we will go to testing part. In this testing we will use different algorithms to test the different leaf images to check whether the leaf is diseased or not.

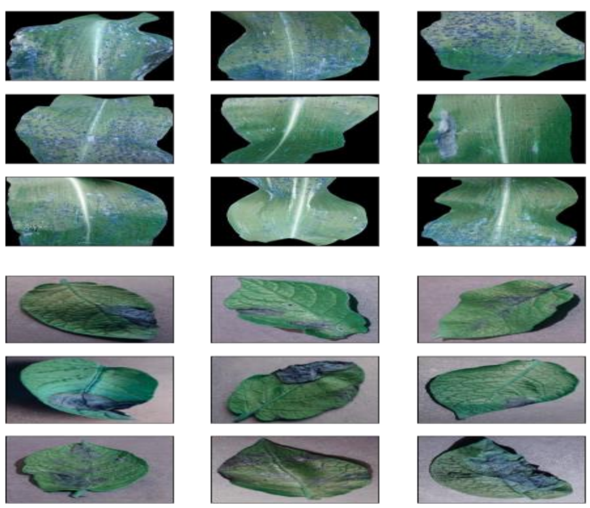
Classification: Finally, in classification by using CNN algorithm we will detect the disease which the leaf is effected and then it will provide the appropriate result.

IV. MODULE DESCRIPTION

Collection of Leaf Image Set:

The dataset comprising photos for classifying plant diseases is divided into training and testing datasets, with the test size being fixed at 30–20%. We attempt to forecast the crop-disease relationship using only a plant leaf image. The dataset for plant diseases can be obtained through self-gathering, network collection, and use of open data sets. Images from the dataset depict several plant diseases in a variety of different plants. We take into account some plants in this system, such as the tomato, potato, etc. Healthy leaves and leaves that are ill.

Fig: 2 Corn leaf And Potato leaf



They were all gathered for the plants mentioned above from various sources, such as photographs downloaded from the Internet. Datasets must be kept in the form of folders in a system after being created. It's a technique that gathers photos of plants and analyses them using machine vision tools to determine whether any of the images contain pests or illnesses.

Pre-processing:

In order to train our model, we resize and shape the photos. Image pre-processing is the process of processing digital photographs using algorithms that are implemented on computers. Utilizing a certain algorithm, we can identify the plant in the photograph. We employ a specific method and a comparable strategy for picture processing and detection. This technique depends heavily on the image quality; if the image isn't clear, we can't utilise the algorithm.

Training And Testing:

Use the pre-processed training dataset is used to train our model using CNN Deep learning and machine learning algorithms along with Resnet50 transfer learning methods. By using these algorithms we will first train the images that we have provided in the datasets and then we will test the trained datasets weather they are diseased or not diseased.

Convolutional Neural Network (CNN):

An input layer, hidden layers, and an output layer make up a convolutional neural network. Any middle layers in a feed-forward neural network are referred to as hidden layers since the activation function and final convolution hide their inputs and outputs.

Fig 3. Accuracy of model in CNN

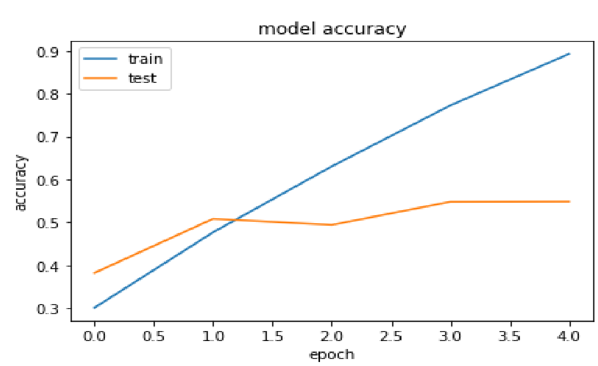
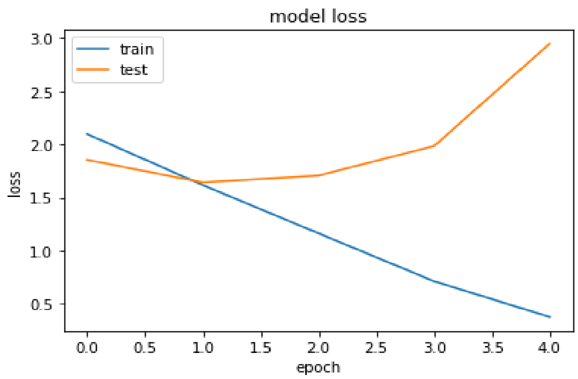


Fig 4. Loss of model in CNN



Artificial Neural Network(ANN):

The ANN learns datasets through a training method, which updates the neuron weights based on the error rate between the target and actual output. In general, ANN learns the datasets using the back propagation method as a training technique.

Fig 5. Accuracy of model in ANN

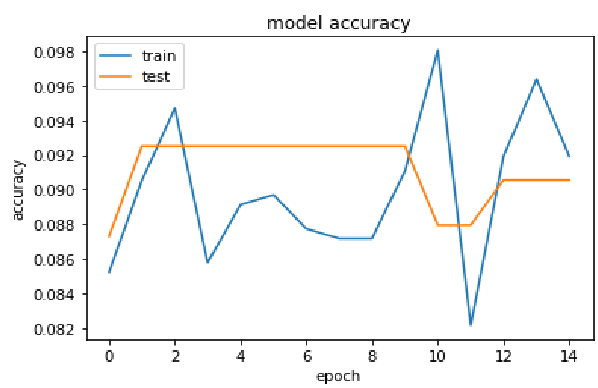
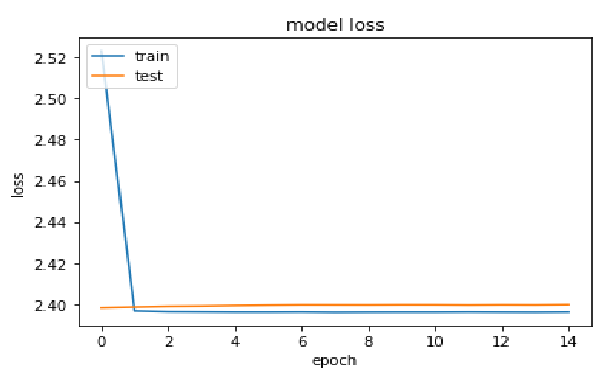


Fig 6. Loss of model in ANN



Support Vector Machine(SVM):

Although Support Vector Machines aretypically thought of as a classification tool, they may also be used to solve regression issues. Multiple categorical and continuous variables can be handled with ease. To divide several classes in multidimensional space, SVM creates a Hyper plane. SVM iteratively builds the ideal Hyper plane, which is used to reduce errors. The main goal of SVM is to identify the MMH, or maximum marginal hyperplane, that best separates the dataset into classes.

Fig 7. Accuracy of model in SVM

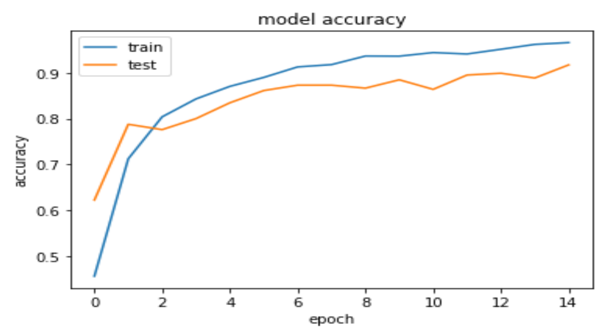
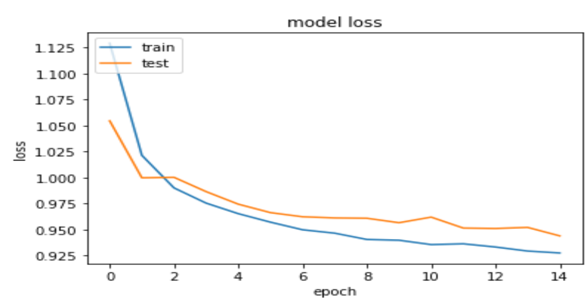


Fig 8. Loss of model in SVM



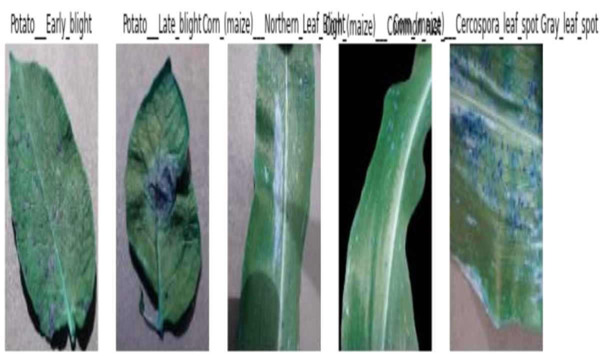
RESNET50:

A significant advancement in deep convolutional neural network training for computer vision problems was ResNet. While the original Resnet had 34 layers and employed 2-layer blocks, more sophisticated variations, such the Resnet50, used 3-layer bottleneck blocks to guarantee higher accuracy and shorter training times. Due of the ease with which models can be created with Keras, it is a well-known deep learning API. Anyone can use Keras for their research because it comes with a number of pre-trained models, including Resnet50. In light of this, creating a residual network in Keras for computer vision applications like picture categorization is rather straightforward.

Classification:

The results of our model are display of plant disease classification images are either with different labels with the use of different algorithms like CNN, ANN, SVM, and ResNet50. By using these algorithms, we will classify if the leaf is with disease or not. An also we it will give us the accuracy of the disease the leaf is having.

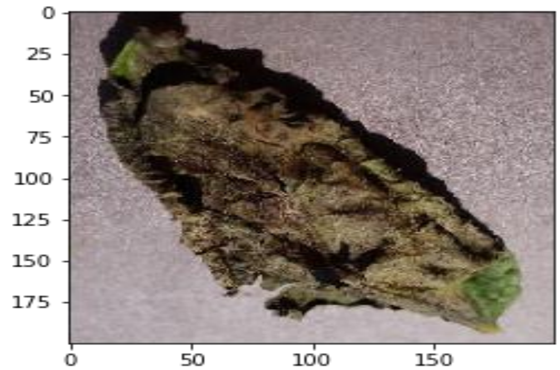
Fig. 9 Classification



V. RESULT

The models that we have suggested in this study makes use of the plant village dataset, which contains 152 photos of healthy plants and around 1000 images of leaves with early blight. The data for this model has been split into two sections, the training set and the testing set. The training set makes up 80% of the leaf image set that we collected, whereas the testing set makes up 20%. Based on its training and testing, the models that we given will show us the accuracy of the disease. In the below given leaf the accuracy of the disease it is having is 90%.

Fig: 10 Prediction of plant and its disease



VI. CONCLUSION

There are numerous methods for identifying plant diseases and offering treatments. Each offers advantages as well as drawbacks. On the one hand, visual analysis is the simplest and least expensive technique. With the help of the CNN, ANN, SVM, and ResNet50 algorithms, which we employed in our project to obtain the desired outcome, we have effectively finished the theory of modules, and got the result for the leaf image set that we tested and trained with the different algorithms.

Reference

[1] Shima Ramesh , Ramachandra Hebbar, Niveditha M, Pooja R, Prasad Bhat N, Shashank N, Vinod PV, " Plant Disease Detection using Machine Learning", published International Conference on Design Innovations for 3Cs Compute Communicate Control (ICDI3C), IEEE Xplore ,DOI: 10.1109/ICDI3C.2018.00017 , April 2018.

[2] Hassan, S. M, Maji, A. K., Jasiski, M., Leonowicz, Z., E. Jasiska, "Identification Of Plant-Leaf Diseases Using CNN And Transfer-Learning Approach", Electronics , 10(12), 1388; https://doi.org/10.3390/electronics10121388 , June 2021.

[3] Muhammad E.H Chowdhury, Tawsifur Rahman, Amith Khandakar, , Nabil Ibtehaz,, “Tomato Leaf Diseases Detection Using Deep Learning Technique, Technology in Agriculture, DOI: 10.5772/intechopen.97319, April 2021.

[4] Y. Zhang , Song, C., D. Zhang, “Deep Learning-Based Object Detection Improvement for Tomato Disease”, IEEE Access, OI: 10.1109/ACCESS.2020.2982456, March 2020.

[5] Sue Han Lee, Herve´ Goe¨au, Pierre Bonnet, Alexis Joly, New perspectives on plant disease characterization based on deep learning, Computers and Electronics in Agriculture, Volume 170, 2020, 105220, ISSN 0168-1699,

[6] Srdjan Sladojevic, Marko Arsenovic, Andras Anderla, Dubravko Culibrk, Darko Stefanovic, ”Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification”, Computational Intelligence and Neuroscience,2016

[7] Trivedi J., Shamnani Y., Gajjar R. (2020) Plant Leaf Disease Detection Using Machine Learning. In: Gupta S., Sarvaiya J. (eds) Emerging Tech- nology Trends in Electronics, Communication and Networking. ET2ECN 2020. Communications in Computer and Information Science, vol 1214. Springer, Singapore

[8] Jayme G.A. Barbedo, Factors influencing the use of deep learning for plant disease recognition, Biosystems Engineering,Volume 172,2018,Pages 84-91,ISSN 1537-5110,

[9] Bengio, Y. Lecun, Yann. (1997). Convolutional Networks for Images, Speech, and Time-Series.

[10] Kaiming He and Xiangyu Zhang and Shaoqing Ren and Jian Sun 2015 Deep Residual Learning for Image Recognition http://arxiv.org/abs/1512.03385

[11] Hussain M., Bird J.J., Faria D.R. (2019) A Study on CNN Transfer Learning for Image Classification. In: Lotfi A., Bouchachia H., Gegov A., Langensiepen C., McGinnity M. (eds) Advances in Computational Intelligence Systems. UKCI 2018. Advances in Intelligent Systems and Computing, vol 840. Springer, Cham.

[12] Mohanty Sharada P., Hughes David P., Salathe´ Marcel,”Using Deep Learning for Image-Based Plant Disease Detection”,Frontiers in Plant Science,vol.7, 2016

[13] AlexNet ,Krizhevsky, Alex, Ilya Sutskever, andGeoffrey E. Hin-ton,”ImageNet Classification with Deep ConvolutionalNeural Networks”, Advances in Neural Information Processing Systems 25 (NIPS 2012)

[14] Christian Szegedy and Wei Liu and Yangqing Jia and Pierre Sermanet and Scott Reed and Dragomir Anguelov and Dumitru Erhan and Vincent Vanhoucke and Andrew Rabinovich,”Going Deeper with Convolutions”, 2014,https://arxiv.org/abs/1409.4842

[15] J. Deng, W. Dong, R. Socher, L. Li, Kai Li and Li Fei-Fei, ”ImageNet: A large-scale hierarchical image database,” 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, 2009, pp. 248-255, doi: 10.1109/CVPR.2009.5206848.

[16] Ashish Vaswani and Noam Shazeer and Niki Parmar and Jakob Uszkor- eit and Llion Jones and Aidan N. Gomez and Lukasz Kaiser and Illia Polo- sukhin,

”Attention Is All You Need”, 2017,https://arxiv.org/abs/1706.03762

[17] Bichen Wu and Chenfeng Xu and Xiaoliang Dai and Alvin Wan and Peizhao Zhang and Zhicheng Yan and Masayoshi Tomizuka and Joseph Gonzalez and Kurt Keutzer and Peter Vajda,”Visual Transformers: Token-based Image Representation and Processing for Computer Vision”,2020

[18] Christian Szegedy and Vincent Vanhoucke and Sergey Ioffe and Jonathon Shlens and Zbigniew Wojna,”Rethinking the Inception Archi- tecture for Computer Vision”, 2015, https://arxiv.org/abs/1512.00567

[19] Konstantinos P. Ferentinos, ”Deep learning models for plant disease detection and diagnosis”,Computers and Electronics in Agriculture,Volume 145, 2018,Pages 311-318, ISSN 0168-1699,