Identification and Detection of Foliar Disease in Apple Tree Using Convolutional Neural Network

R Tanseer Ahmed[1], Dr. S.K Manju Bargavi[2]
[1]Dept of MCA
[2]Professor, Dept of MCA
[1, 2] School of CS& IT, Jain University,Bangalore, Karnataka, India

Abstract- *Apple trees are one of the most commonly planted crops in vast gardens and orchards.Simultaneously, apple plants are among the most affected diseases. Even the expert eye finds it difficult to detect disease at an early age and prevent it from spreading to other parts of the plant. As a result, a comprehensive program is needed to detect plant diseases at an early the level This paper shows the ability of Convolutional Neural Networks to detect and classify objects automatically. resolve problems. Pictures of Apple leaves from the Plath Pathology database, which includes a variety of diseases and healthy samplesare used to validate the findings To generate a broad range of train images and fine-tune the method, image filtering, compression, and production methods are used. With a total accuracy of, the qualified model achieves the best accuracy in all grades. 93.27 percent of all data, which was sampled and produced from 3642 labeled images.*

Keywords- Apple Tree ,Convolutional Neural Network, Deep Learning, Leaf Disease Detection, Machine Learning

I. INTRODUCTION

In the early stages, identifying plant diseases can play a very important role in plant protection. It will no doubt increase the value and quality of agricultural products. Monitoring plants for disease identification is a time-consuming, labor-intensive task that requires knowledge of plant diseases. However, recent advances in machine learning have empowered the task of diagnosing plant diseases on their own. There is no doubt that the discovery of machine-based information on the apples of apples helps to monitor large apple fields. An effective machine-based system can detect apple diseases at an early age, saving farmers from major losses / injuries. Monitoring plants for disease detection is a time-consuming, time-consuming process that requires knowledge of plant diseases. However, Recent advancements in machine learning have made it possible this plant-specific diagnostic work to be done on its own. Machine-induced apple disease undoubtedly helps to monitor large apples of apples. An effective machine-based system can detect apple diseases early, saving farmers money and time. To find, identify, and calculate different leaf parameters, Image acquisition processes such as capture, scanning, classification, and retrieval of image elements are used. Other common diseases of apple leaves include Apple Scab, Rust, Gray Spot, and Brown Spot. Apple tree leaf diseases can be effectively controlled, losses reduced, and healthy growth in the apple industry maintained early detection and accurate diagnosis. Neural convolutional deep networks work well in dual data processing, especially in image and video editing operations. Feature reuse feature in DenseNet dense block. Literature review by various methods of analysis of leaf parameters and techniques for diagnosing leaf diseases of various plants. This study demonstrates that The theoretical low-reading model offers a better approach for apple leaf disease control, with higher accuracy and a faster integration rate, and the imaging process described in this paper will increase the reliability of the convolutional neural network model.

II. RELATED WORK

Many researchers have contributed to this work on the analysis of leaf parameters and disease detection in the past. The following is a review of the text related to the study of disease in plant leaf.

In paper [1] 2970 photos of ATLD and healthy leaves were taken. There are Mosaic, Rust, Gray dot, Brown, Alternaria leaf spot, and healthy leaves are among the six diseases and healthy leaves in the archive. the proposed DCNN ATLD recognition model incorporates DenseNet and Xception, using standard global integration rather than fully connected layers. To improve overall network capacity and reduce overcrowding. To simulate variations in brightness, brightness, angle, and noise, data augmentation technology was used. while reconstructing images of apple leaves. The images are considered to mimic these changes by increasing and decreasing the brightness ratio by 30%, increasing the difference by 50% and decreasing the difference by 20%, increasing the sharpness by 100% and decreasing the sharpness by 70%, respectively. Make use of rotation degrees (90, 180, 270). The database is divided into three distinct subsets: a training database, a validation database, and a test database. A total of 60% of the storage is used as a training database, 20% as a verification database, and 20% as a test set, to ensure that each set includes a laboratory background and background plant background. The proposed convolutional neural network extracts the features, and then separates them using a vector support mechanism. Compared to other models, like the proposed one has a very high level of integration, a limited number of parameters, and high durability.

The authors in [2] have an apple leaf dataset that includes Good, General, Scab, Serious Apple Scab, Apple Gray Spot, General Cedar Rust, and Serious Cedar Apple Rust photos. The data collection included 2462 photographs. It is split into training and trial data sets at a ratio of 8:2. A Densenet-121 deep convolution network-based solution, three regression modes, multi-label split modes, and a focus loss feature are proposed for the diagnosis of apple leaf diseases such as good Apple, General Apple Scab, Serious Apple Scab, Apple Gray Spot, General Cedar Apple Rust, and -Serious Cedar Apple Rust. Using a Densenet-based network-based technique, distinguish apple leaf diseases such as good Apple, General Apple Scab, Serious Apple Scab, Apple Gray Spot, General Cedar Apple Rust, and Serious Cedar Apple Rust. It is proposed to use 121 deep convolution networks, three regression methods, multiple mark separation methods, and a focus loss function. Quantitative research has revealed that Densenet-121 deep convolution network approaches, three regression modes, and multi-label separation are efficient, as well as the concentration loss feature exceed the standard single multiclassification label and cross entropy loss function on the unbalanced data set.

In paper [3] To determine the region of interest, the process of classifying the neutrosophic logic-based segment form is used. The three membership elements distinguish a separate neutrosophic image: real, false, and central. The database contains there are four distinct groups images of both good and diseased basil leaves. The system proposed has been checked for 400 cases (200 healthy and 200 diseased). New features are released according to segment. These factors include the discriminatory force of the leaf's strength and texture to determine if the leaf is diseased or healthy. The purpose of predicting the delivery of encrypted data is a method of measuring a slightly restricted histogram to improve variability. After further processing, the image is converted into a background, which divides the image into three parts: real, false, and intermediate Feature omission: Create a new function pond focused on three distinct regions to distinguish between stable and sick leaves. A dataset-trained convolutional neural network is used in an in-depth learning process based on solving a previously described diagnostic problem. According to graphical reviews, RF surpasses certain models of machine learning with 98.4 percent accuracy. This analysis paper proposes a new method of segmentation and a new set of features. The accuracy of the proposed features is measured using nine separators.

In paper [4] The aim of this study is to develop a model. that will determine whether images of plant leaves are healthy or sick. If the leaf of a plant is tested for disease, the type of disease must also be diagnosed. The concept focuses only on defining and categorising Apple leaves infected with various apple diseases. The dataset is classified into four main categories. Three diseases, and one stage of the leaves are stable. 1000 leaf image samples and 1526 sick leaf photo samples This is a well-known set of Plant Village datasets. GoogLeNet Architecture is a CNN multi-layered 22-layer architecture. GoogLeNet has 5 million parameters, which is a small number compared to other standard formats with very high parameters, such as AlexNet. The use of "network-to-network" design in the form of startup modules is a key element in the development of GoogLeNet. These early modules use the same combination as the layers for more integration. Part of the CNN model is well-structured by adjusting parameters such as stop rate, batch size, and level of train separation. The test method will to be able to forecast type of disease the apple tree is suffering from. High accuracy statistics show that CNN cutting edge works well in the classification of plant diseasesApple Black Rot, Apple Cedar Apple Rust, Healthy Apple, and Apple Scab have an accuracy rate of 98.71 percent , 99.27 percent, 98.70 percent, and 97.3 percent respectively.

In paper [5] Proper design of an AlexNet-based To diagnose apple leaf diseases, a convolutional neural network is suggested. The suggested neural network has been learned to differentiate the four most common diseases of apple leaves. Four distinct diseases of apple leaves are represented in the database. A total of 1053 photographs were found with common symptoms of the disease, including 252 Mosaic, 319 Brown spot, 182 Rust, and 300 Alternaria leaf spots. A database of 13,689 diseased apple leaves has been used to improve the diagnostic properties of apple leaves Deep convolutional neural networks are used to power this system. In-depth education forms are exaggerated when a mathematical model produces sound or random errors rather than a basic relationship. Minor distortions are used for images in the test phase to reduce congestion in the training phase and improve the ability to withstand complex disruptions.The proposed diagnostic approach for network-based diagnostics achieves a complete 97.62 percent precision The proposed in-depth research model offers the best approach for apple leaf disease control, with high accuracy and fast concentration rate, and the imaging process can increase the robustness of the convolutional neural network model.

In paper [6] The detection model using deep CNNs is enhanced by an apple leaf database with laboratory images and complex images created using the expansion of data and image annotation technology using the Inception of GoogLeNet and Rainbow concatenation structure There are 2029 photos of sick apple leaves in the archive, which are classified into five categories: Alternaria leaf spot, Brown area, Mosaic, Gray area, and Rust. The model would learn as many inactive patterns as possible during the training process with further models after data magnification , preventing overheating and achieving high results. Additional functions such as rotation, horizontal and vertical horizontal, and vertical disturbances, including light distortion, sharpness, and contrast. using the Gaussian audio processing function. Each image creates 12 new images that are sick due to the above performance. The The proposed INAR-SSD (SSD with an initial module and a Rainbow concatenation model) is currently being prepared. The INAR-SSD model achieves a decent score of 78.80 mAP in the apple leaf database , according to test results. To process items of different sizes, the SSD model includes multiple maps of various features and resolutions. The SSD has a much higher recovery speed than the faster Raster CN, but the availability details for these two methods are almost identical. As a result, The SSD algorithm is used as the primary object detection algorithm and was created using a multi-angle feature combination. VGGNet is a standard migration learning model because it is very portable. VGGNet transcends traditional neural convolution networks in testing for apple leaf diseases. A high-performance approach for detecting apple leaf diseases early can detect these diseases in real time with greater accuracy and speed than previous methods.

In paper [7] The convolution neural network model is designed to detect disease in apples, and has three layers of convolution, three layers of integration, and two highly layers that are connected After playing with various numbers of convolution layers ranging from 2 to 6, it was discovered that three layers had the highest accuracy. The database mostly includes three diseases. Fungi include scabies, black rot, and apple cedar. Algorithms or traditional models such as SVM, Decision Tree, Logistic Regression, k-NN, LDA, Naive Bayes, and Random Forest were used to create images of cork 454, black rot 496, rust 220, and stable 1316. were used to test the proposed model. These algorithms were found to perform well in the same database collected after the installation used for the proposed CNN, the HSV histogram, and the Local Binary Pattern (LBP) The Gray Level Co-occurrence Matrix serves as the foundation for Haralick histogram texture functions (GLCM). Tests using the CNN model performed well in terms of precision, machine time, data, F1, and AVC-ROC curve and VGG16 results and InceptionV3 results show the performance of the proposed algorithm in addition to previously trained models and standard machine learning. Using the neural convolution network, an effective diagnostic model for Apple was created. When used on a previously trained model, the proposed model takes up just 20% of the room and integrates in less than one second, while pre-trained models take at least 30 seconds.

In paper [8] This paper offers an in-depth look at leaf parameter analysis, identifying stable, diseased, or injured leaf areas, and classifying leaf diseases using a number of plant approaches. A database of Pigeon Pea, Green Gram, and Black Gram plant leaves is used. In image processing, image retrieval is the act of retrieving an image. A digital camera is used to take pictures of different apple leaves. Photo and photo training set for testingIt's the first step in the process. Image Enhancement: Photo editing is used to make images more compatible. Image editing techniques include modification, image modification, sound removal, contrast enhancement, and unnecessary distractions. Image segmentation is used to determine the profit margin on image (ROI). The method of classifying an image into multiple categories is known as image classification. Often, image classification is used to find points and boundaries in images. Feature removal: Different features are removed using the debugging process in the feature removal function to define the regions using the selected presentation. Location is defined by its boundary, which is characterized by features such as texture and color A plant-specific algorithm would not work well with another leaf. To detect leaf diseases, special algorithms for the custard apple plant are needed in addition to the parameter leaf analyzer. Leaf area, weight, and width are all important factors in plant growth and photosynthesis study.

In paper [9] Various approaches, such as step models, image recognition, and deep learning models, are used to treat the disease.This document provides a variety of diagnostic tools for agricultural areas. Photography is a method of collecting images and converting them into a suitable output format. we can download a photo of the plant from any official website that contains a variety of leaf pictures or take one with a digital camera. The database contains images of plants, which can be taken with a digital camera and include various types of diseased leaf images.

Both the stable and the sick are found by hand using various names and numbers. Pre-image editing It is necessary to perform certain tasks such as sound and background removal. Both stable and sick images are stored in the RGB color format in the database and are manually accessed using different names and numbers. It provides research on a variety of schemes, as well as research, which can be used to recognise and classify plant leaf diseases, more than five articles SVM and Neural Network received more than 90% accuracy, competing with the best ML category models available for segmenting high-resolution data sets.

In paper [10] A low-cost, reliable, and highly accurate method for detecting apple leaf diseases. The MobileNet concept is used to do this. It is a low-cost application that can be quickly implemented on mobile devices as compared to other in-depth learning models. 334 data sets were used for training and evaluating the MobileNet model. The website lists two common forms of apple leaf damage: Alternaria leaf blotch and rust. Three versions were used: ResNet152, InceptionV3, and MobileNet. ResNet152 belongs to the standard CNN, as seen above. Either of the two basic versions is almost similar to the other. Through comparing in-depth sample models, the ResNet152AppleNet leaf diagnostic software will greatly reduce the burden of performance and accuracy. MobileNet is a CNN architecture designed specifically for mobile devices . the basic construction is constructed with the depth of the separated junction, which can be a type of set weight that includes the weight of the standard value has a depth of depth and the weight of 1 × 1 is called cognitive complexity. deep separation compliant with deep and clear layers with standard batch and ReLU. AppleNet leaf disease screening program can significantly reduce the burden of professional efficiency and accuracy by comparing in-depth study models. The accuracy of MobileNet is almost the same as for the more complex learning models now, and the low-end mobile app can easily be used.

III. EXISTING MODEL

The current way of diagnosing plant diseases is to simply inspect farmers' or experts' farms. This device requires a large number of people to track it. Plants should also be monitored continuously. When the land is too large, labor costs are too high. As mentioned earlier, visible farm monitoring takes time.it costs more than dishonesty to answer this question, Image processing is used to identify leaf diseases. However, there is no valid application for strategic separation. Leaf after collecting and describing its images features the distinction of plant leaves has completely changed.

IV. PROPOSED MODEL

The proposed project uses image processing to identify and identify leaf infections. Image processing techniques are used to process leaf images and extract important information for future research. When the disease is detected, the appropriate pesticide spray is sprayed on the infected leaf in the right amount. The CNN model is used to diagnose diseases. Accurately Distinguish between different diseases, usually several times on a single leaf. The accuracy and validity of the training data should be strong while the loss of attention is kept to a minimum.

V. SYSTEM ARCHITECTURE

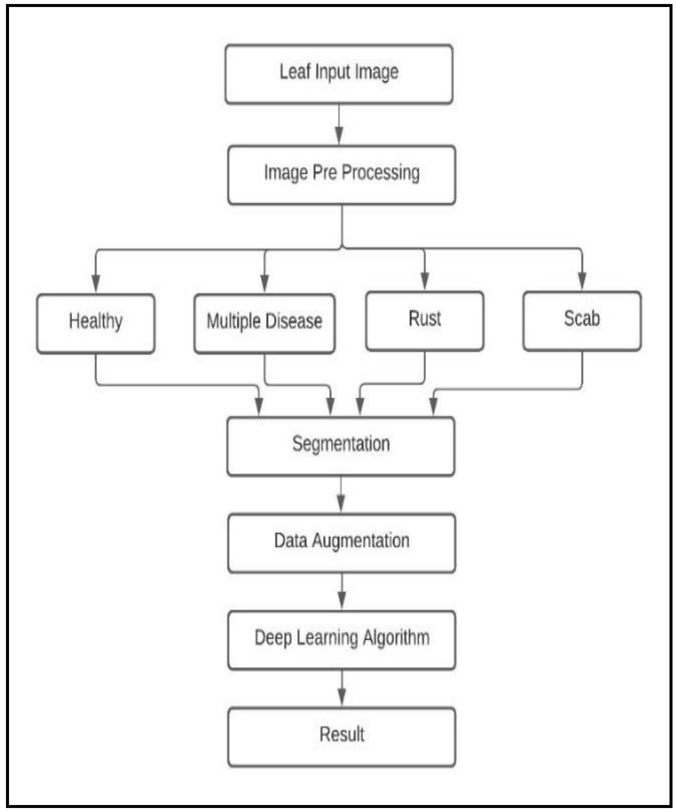


Fig-1: System Architecture

V. METHODOLOGY

The main purpose of the study was to create a model that could identify the leaves of healthy or diseased indoor plants. If a positive disease effect is found on a plant leaf, the type of disease should be determined. Research consists only in finding and separating Apple leaves infected with various apple diseases.

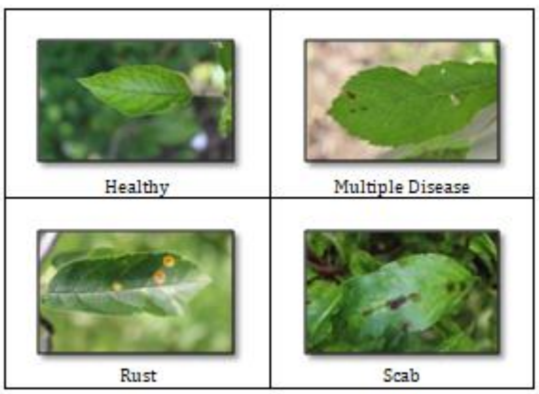
A. Architecture Used

GoogLeNet Architecture is a multi-layered CNN architecture with 22 layers. GoogLeNet has 5 million parameters, which is a small number compared to other standard properties with very high parameter values, such as AlexNet. The use of "network in network" structures in the form of startup modules is a key element of GoogLeNet's design. These are the first modules that use the same composites in combination with a multi-component coating, where similar variables refer to layers 1 size, 3 3, 5 5, and so on. This feature allows us to capture multiple features simultaneously.

B. Dataset Description

The database used for this project is divided into four different categories. Three of them are diseases, and one is a healthy leaf. There were 516 healthy leaf pictures and 3126 diseased leaf pictures. This is a subset of the well-known Plath Pathology database. The sorting process is required for a specific image of an apple leaf, which can separate the inclusion images in one of the classes shown in Table I below.

Table -1: Apple leaf Images with its disease



C. Dataset Description

The data augmentation implementation is made possible by the Keras library in Python 3. The dataset is initially loaded into the variable "train data," and the input images are resized to 224 x 224. Each image is further normalised by Taking the mean and dividing it by the standard deviation The images in the dataset are randomly rearranged, and the validation collection is made up of the first 730 images (roughly 20% of the initial dataset), while the remaining randomly ordered images are used to train the deep learning CNN algorithm. The function of the data generator at Cameras is used to generate new images by making simple adjustments such as rotation, rotation, height, and horizontal rotation to increase the size and capacity of the database to deal with all possible conditions while constructing limited data. As a result, to address the problem of having too few photographs in our database, the Camera library also includes photographic production. This generator creates new images in the database with rotation, horizontal rotation, vertical rotation, horizontal swipe, and randomly browse selected images.

D. Convolutional Neural Network

To complete this work, the Convolutional Neural Network (CNN), a neural network learning component widely used for image recognition and processing, is proposed. CNN is made up of neurons, much as most neural networks that can be trained to have readable weights and bias. Each neuron receives green image metrics as input, combines a weighted value, transmits them through activation feature that responds with output The network as a whole has a loss function that decreases with each iteration , and the network eventually reads a set of parameters that fit well with the partition function. Every class's corresponding vector of final probabilities is a vector of final probabilities. Unlike neural networks, which have a vector as input, CNN has an image with multiple channels as input. In the process of convolution, we take a filter of a certain size (say, 5x5x3), upload it to the whole image, and calculate the product of the dots between the filter and the base of the input image. The CNN model we propose follows a design pattern similar to that of the 1990s LeNet-5 (LeCun et al., 1989). Two convolutional layers are followed by a maxpool layer with a 2x2 filter and a double row, and two fully connected layers with sizes 500 and 4, respectively. CNN expects 60x60x3 input. The first convolutional layer contains 20 5x5x3 filters. Because "border mode" is set to "same," the length and width of the output do not change. A 2x2 maxpool layer with two steps is used. Output size is 60x30x10. The second convolutional layer contains 50 5x5x10 filters. This results in a size of 60x30x50. The maxpool layer expands to a size of 30x15x50, followed by an ANN layer of 500 neurons and a final layer of output of four neurons. In the first three phases, “ReLu” was used, followed by “softmax” in the final layer. To prevent the model from overheating, a 0.2 drop is used for each layer.

E. Algorithm And Implementation Details

Keras which operates beyond Tensorflow, is used to build and train the CNN model. All dependencies are imported using the Anaconda package manager. The following are the steps for implementation

- Load the train-test image dataset in the 80:20 proportion save them as numpy arrays. Train and test images are standardized using the Z-transformer, which removes the meanings of the images from each image and separates the effect by the standard deviation of the images.

- Load the train and test images' saved numpy arrays. Because we have a limited number of images for our work, we have used a data generator to generate new images from existing databases by making small changes to the current database using simple functions such as rotating, horizontal and vertical, horizontal and vertical shifts, and so on.
- The CNN model is similar to the standard LeNet architecture. The model has four basic layers. The first layer is a convolutional layer, which takes a colored image of size 60 x 60 as input. There is another layer of convolutional. Behind these two layers, there is the ReLu launch layer and the maxpool layer. The third and fourth layers are fully connected. The retaining layer of the output layer, which consists of four neurons that separate the insertion image into one of the four stages.
- After constructing the model, it is trained with the following parameters: optimizer = Adam, batch size = 12, epochs = 30. A model checkpoint is also included, which saves the best model so far

VI. EXPERIMENTAL RESULTS

The model's output was evaluated using an augmented dataset with modified images in the aircraft, as well as a repetitive operating model with fixed parameters but a small change in output, batch size, epochs, and a separate test variation in the confirmation size. After fine-tuning all aspects (drop-out, batch size, epochs, and subdivision test), it gets excellent results, making it the right choice for a stable, impartial model capable of achieving the highest levels of accuracy in almost all classes. As a result, our in-depth learning model gains an average accuracy of 93.47 percent.

Table -2: Deep Learning CNN Validations for different values of varying parameters

No	*Dropout*	*Epochs*	*Accuracy*	*Val.Loss*	*Val.Accuracy*
1	0.4158	15	86.06	36.20	87.78
2	0.3528	17	87.79	29.94	90.06
3	0.3121	19	90.29	31.36	90.62
4	0.2826	21	90.29	29.42	92.90
5	0.2420	23	91.61	31.60	90.34
6	0.2388	25	91.68	23.91	91.76
7	0.1988	27	92.65	26.30	90.34
8	0.1942	29	93.07	23.22	92.61

VII. CONCLUSION

This paper focuses on diagnosing diseases on apple trees using pictures of their leaves. Here, the proposed method an in-depth study approach based on diagnostics, using a database-trained convolutional neural network specified in the database. Now, whenever an image path is provided, all that is needed is to use the code, and the system will be able to predict the type of disease the apple tree is suffering from. Part of the CNN model is well-structured by adjusting parameters such as stop value, batch size, and train split rate. The best model accuracy in the total database is 93.47 percent output = 0.2.

VIII. FUTURE WORK

The decline in research is the scalability of work has yet to be assessed on other major crops. As a result, the addition of more plants and disease stages is a challenge ahead. One of the most important issues for this project is the database, because the images of the leaves here were collected and clicked under the recommended natural conditions, well arranged in a high quality, and with the same background. The real-world situation is very different from this set-up, and this is another challenge where the model has the potential to be greatly improved. Even for standard CPUs, improved training time and reduced test times to a fraction of a second is a contractual agreement and therefore very promising for taking mobile work. With the proliferation of advanced computer technology and the construction of dedicated Neural Processing Units, the opportunity to extend this work to hand-held technology seems promising. In addition, this method allows us to collect data on a timely basis and test it on various real-world samples to improve the model based on new inputs.

REFERENCES

[1] Chao, G. Sun, H. Zhao, M. Li and D. He, "Identification of Apple Tree Leaf Diseases Based on Deep Learning Models", *Symmetry*, vol. 12, no. 7, p. 1065, 2020.

[2] Y. Zhong and M. Zhao, "Research on deep learning in apple leaf disease recognition", *Computers and Electronics in Agriculture*, vol. 168, p. 105146, 2020.

[3] G. Dhingra, V. Kumar and H. Joshi, "A novel computer vision based neutrosophic approach for leaf disease identification and classification", *Measurement*, vol. 135, pp. 782-794, 2019.

[4] Baranwal, S. Khandelwal and A. Arora, "Deep Learning Convolutional Neural Network for Apple Leaves Disease Detection", *SSRN Electronic Journal*, 2019.

[5] B. Liu, Y. Zhang, D. He and Y. Li, "Identification of Apple Leaf Diseases Based on Deep Convolutional Neural Networks", *Symmetry*, vol. 10, no. 1, p. 11, 2017.
[6] P. Jiang, Y. Chen, B. Liu, D. He and C. Liang, "Real-Time Detection of Apple Leaf Diseases Using Deep Learning Approach Based on Improved Convolutional Neural Networks", *IEEE Access*, vol. 7, pp. 59069-59080, 2019.
[7] M. Agarwal, R. Kaliyar, G. Singal and S. Gupta, "FCNN-LDA: A Faster Convolution Neural Network model for Leaf Disease identification on Apple's leaf dataset", *2019 12th International Conference on Information & Communication Technology and System (ICTS)*, 2019.
[8] Gargade and S. Khandekar, "A Review: Custard Apple Leaf Parameter Analysis and Leaf Disease Detection using Digital Image Processing", *2019 3rd International Conference on Computing Methodologies and Communication (ICCMC)*, 2019.
[9] S. Swain, S. Nayak and S. Barik, "A Review on Plant Leaf Diseases Detection and Classification Based on Machine Learning Models", *Shabdbooks.com*, 2021.
[10] C. Bi, J. Wang, Y. Duan, B. Fu, J. Kang and Y. Shi, "MobileNet Based Apple Leaf Diseases Identification", *Mobile Networks and Applications*, 2020.