PLANT LEAF DISEASE CLASSIFICATION USING CNN

Priyanka V[1], Bhavani R[2]

[1]Student, Dept. of Computer science and Engineering, Government College of Technology, Coimbatore, India
[2]Assistant Professor, Dept. of Computer science and Engineering, Government College of Technology, Coimbatore, India

---***---

Abstract -*Plants are the significant wellspring of nourishment for a wide range of living creatures. With the increment in population, it is currently more essential to keep this supply proceed. To cop-up with this, it is necessary to protect the plants from different sorts of infections. So, there is need for identifying those diseases or infections at the early stage. Most recent advances in Deep Learning and Image Processing techniques would help the farmers in timely identification and classification of leaf diseases. In this paper, a convolutional neural network (CNN) based model is proposed to classify the leaf diseases. The customized CNN model was built using keras library in Google colab. The model was trained for 50 epochs with a batch size of 64. The proposed CNN model achieved an accuracy of 90% on training and 89% on testing.*

***Key Words*: CNN, Maxpooling Layer, Fully Connected Layer, Leaf Disease**

1. INTRODUCTION

Agriculture is a sector which has a huge impact on life and economic stature of humans. Agriculture is a primary source of income for approximately 58 percent of India's population. In terms of farm yields, India is ranked second in the world Agriculture, it was stated in 2018, provided employment to more than 50 percent of the workforce, adding 18–20 percent to the country's GDP. As a result, India has established itself as one of the leading countries in terms of agricultural yield and productivity. Given that agriculture employs the majority of the people, it is critical to understand the issues that this sector faces. There are a numerous of problems that the agriculture field faces such as inefficient farming strategies and techniques, Inadequate use of compost, manures, and fertilizers, insufficient water supplies, and different plant diseases, to name a few. Diseases are extremely destructive to plants health, which has an impact on their growth. The attack of these many forms of diseases on plants causes a significant reduction in yield performance, both in terms of quality and quantity. Disease-affected plants account for roughly 20-30% of all crop losses. The proposed system is used to predict the leaf diseases and categorize leaf illnesses, using deep learning. Deep Learning has recently achieved outstanding results in disciplines such as image recognition, speech recognition, and natural language processing. The convolution Neural Network has produced excellent results in the problem of Plant Disease Detection. Convolutional Neural Network (CNN) is a type of Artificial Neural Network (ANN) with a large number of hidden layers between the input and output layers and these hidden layers are called as Convolution, Max Pooling, Dropout, etc. The convolutional neural network finds the correct mathematical operation to turn the input into the output, whether it's a linear correlation or a non-linear relationship. By training a huge number of images CNN, a sort of machine learning can achieve the goal of accurate detection. In the realm of generic identification, CNN does not rely on specific features and has a good detection implication. CNN does not depend on selected features, and has a good detection implication in the field of generalized identification. Deep Learning have fully dominated the field of image classification for the final few years; thus, the proposed Convolution Neural Network is used to detect plant leaf disease classification using CNN diseases and categorize them into 38 classes.

2. RELATED WORKS

Many researches have done in the recent years for the plant leaf diseases diagnosis by using deep learning algorithms.
In paper [2], the authors build a classifier based on the extraction of morphological features using a Multilayer Perceptron with Ada boosting techniques. Certain pre-processing approaches are used to prepare a leaf image before starting the feature extraction phase. The authors used various morphological features named as centroid, major axis length, minor axis length, solidity, perimeter, and orientation are mined from the images of various classes of leaves. k-NN, Decision Tree and Multi Layer Perceptron are applied in the Flavia dataset to test the accuracy of the model. The proposed machine learning classifier outperformed state-of-the-art algorithms, achieving a precision rate of 95.42 percent. In paper [18], Smart farming, which employs the appropriate infrastructure, is an innovative technique that aids in the improvement of the quality and quantity of agricultural production in the country including tomato production. Diseases cannot be prevented because tomato plant farming takes into account a variety of factors such as the atmosphere, soil, and amount of sunlight. The current cutting-edge computer system innovation enabled by deep learning has paved the path for camera-detected tomato leaf disease. This research resulted in the creation of a unique approach for disease detection in tomato plants. To identify and distinguish leaf diseases, a

motor-controlled picture capturing box was built to capture four sides of each tomato plant. The test subject was a specific tomato variety known as Diamante Max. Phroma Rot, Leaf Miner, and Target Spot were among the illnesses identified by the method. Diseased and healthy plant leaves are collected in the dataset leaves. Then, to identify three diseases train a deep convolutional neural network. The system used a Convolution Neural Network to determine whether tomato illnesses were present on the plants being monitored. The F-RCNN trained anomaly detection model has an accuracy of 80 percent, whereas the Transfer Learning illness recognition model has a 95.75 percent accuracy. The automated image capture system was tested in the field and found to be 91.67 accurate in detecting tomato plant leaf diseases. In paper [14], The authors proposed approach contains three phases such as pre- processing, feature extraction and classification. The pre-processing stage includes a classic image processing steps such as converting to grayscale and boundary enhancement. The feature extraction stage deduces the common DMF from five fundamental features. The authors are using the Support Vector Machine (SVM) for the classification of leaf recognition. Leaf features which are mined and orthogonalized from the leaf images into 5 principal variables that are given as input vector to the SVM. Model tested with Flavia dataset and a real leaf images and compared with k-NN classifier. The proposed model yields very high accuracy and takes very less execution time. In paper [4], A method for detecting image segmentation based on region, followed by texture feature extraction. For classification, SVM is utilized. Pictures of roses with bacterial disease, beans leave with bacterial disease, lemon leaves with Sun burn disease, banana leaves with early scorch disease, and fungal disease in beans leaves are included in the data collection. The SVM model has a 95 percent accuracy rate. In paper [15], Gathering learning thinks about a solitary classifier as well as a bunch of classifiers. The joined expectation of order dependent on casting a ballot is finished. Tomato leaf illness order was completed utilizing surface highlights removed from GLCM. Numerous classifiers viz SVM. Multi-facet perceptron and Random Forest were utilized for arrangement. A delicate democratic classifier predicts the result dependent on most elevated likelihood of picked class as result. The singular arrangement precision was 88.74%, 89.84% and 92.86% for RF, MLP and SVM classifier. Delicate democratic based group learning yielded a precision of 93.13%. In paper [16], convolutional neural organization is utilized to characterize soybean plant illness. The data set incorporates pictures taken from normal assets. The framework execution of 99 .32% is done inside the kind of turmoil. The investigation moreover portrays that insights expansion at the instruction set works on the general exhibition of the organization. The proposed CNN form of this paper incorporates three convolutional layers each layer is seen by utilizing a maximum pooling layer. The last layer is totally connected to MLP.ReLu initiation work is completed to the result of each convolutional layer. The result layer is given to the softmax layer which offers the likelihood conveyance of the 4-result preparing. The records set has been separated into schooling (70%), validation (10%) and attempting out (20%). ReLu enactment trademark is utilized in light of the fact that it impacts in quicker instruction. From the impacts of type, it very well might be seen that the precision of tinge pictures is higher than the grayscale pix. Henceforth, hue pix are better for work extraction. The aftereffect of the adaptation additionally demonstrates that the model is overfitting. Overfitting happens when the form suits the instruction set effectively. It then, at that point, will become to sum up new models that have been currently not inside the preparation set. In paper [19], The five types of apple leaf disease discussed in this study are aria leaf spot, brown spot, mosaic, grey spot, and rust. In apple, this is a problem. Deep learning techniques were employed in this study to improve convolution neural networks (CNNs) for disease detection in apple leaves. The apple leaf disease dataset (ALDD) is utilized in this paper to generate a new apple leaf disease detection model that leverages deep- CNNs by employing Rainbow concatenation and Google Net Inception structure. The suggested INAR- model was trained on a dataset of 26,377 photos of apple leaf disease and then utilized to detect five common apple leaf diseases. The INAR- SSD model achieves 78.80 detection performance in the experiments with a high-detection speed of 23.13 FPS. The findings show that the innovative INAR-SSD model provides a high-performance solution for the early detection of apple leaf illnesses that can identify these diseases in real time with greater accuracy and speed than earlier methods. In paper [17], For leaf disease identification, the advised device in this article has a particular deep learning version which is built on a specific convolutional neural network. The images in the facts set were taken with a variety of cameras and resources. Pests and disorder, according to research, are not a problem in agriculture since healthy and developing plant life in rich soil is able to withstand pest/ailment. They employed Faster Region-Based Convolutional Neural Network (Faster R-CNN), Region-Based Fully Convolutional Network (R-FCN), and Single Shot Detector as detectors (SSD). To execute the experiment the dataset had been divided into three units validation, education, and trial sets for example. The first assessment is carried out at the validation set after that education of the neural community is carried out at the education set and very last evaluation is executed at the testing set. They use education and validation units to execute the training system and trying out set for evaluating effects on uncooked statistics. In paper [11], Plant diseases are a major factor in crop productivity posing a threat to food security and reducing farmer profits. Identifying plant illnesses and taking adequate feeding methods to cure them early is the key to preventing losses and a reduction in productivity. The scientists employed two approaches to identify and classify healthy and sick tomato leaves in their study. Using the k-nearest neighbor strategy the tomato leaf is evaluated as healthy or sick in the first procedure. They then use a probabilistic neural network and the k-nearest neighbor approach to classify the sick tomato leaf in the second technique. For categorization reasons characteristics such as GLCM, Gabor, and color are used. Experimentation was carried out on the authors own dataset. There are 600 healthy and sick leaves in all. The results show that using a fusion strategy with a PNN classifier beats other methods.

3. METHODOLOGY

Deep learning is a trendy technique for image processing and data analysis that provides correct results. Deep learning has been successfully applied in numerous domains and it recently entered into agriculture too. Therefore, the proposed system applies deep learning to make a formula for machine-controlled detection of plant leaf diseases. The leaf disease classification consists of five pipelined procedures: dataset collection, resizing, reshaping and rescaling the image, model building and classifying diseased leaf image using CNN. This approach identifies the type of leaf disease in less time with more accuracy.

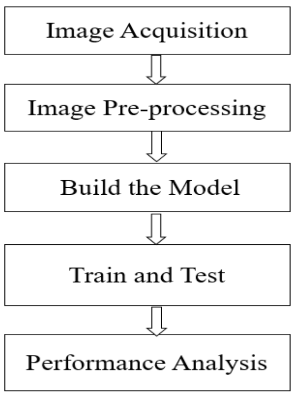


Fig.1. Image Classification Steps

3.1 Image Acquisition

Plant leaf images are downloaded from plant village dataset [13]. The dataset consists of 5148 images grouped into 38 classes. The class names are listed below;

['Apple___Apple_scab',
'Apple___Black_rot',
'Apple___Cedar_apple_rust',
'Apple___healthy',
'Blueberry___healthy',
'Cherry___Powdery_mildew',
'Cherry___healthy',
'Corn___Cercospora_leaf_spot Gray_leaf_spot',
'Corn___Common_rust',
'Corn___Northern_Leaf_Blight',
'Corn___healthy',
'Grape___Black_rot',
'Grape___Esca_(Black_Measles)',
'Grape___Leaf_blight_(Isariopsis_Leaf_Spot)',
'Grape___healthy',
'Orange___Haunglongbing_(Citrus_greening)',
'Peach___Bacterial_spot',
'Peach___healthy',
'Pepper,_bell___Bacterial_spot',
'Pepper,_bell___healthy',
'Potato___Early_blight',
'Potato___Late_blight',
'Potato___healthy',
'Raspberry___healthy',
'Soybean___healthy',
'Squash___Powdery_mildew',
'Strawberry___Leaf_scorch',
'Strawberry___healthy',
'Tomato___Bacterial_spot',
'Tomato___Early_blight',
'Tomato___Late_blight',
'Tomato___Leaf_Mold',
'Tomato___Septoria_leaf_spot',
'Tomato___Spider_mites Two-spotted_spider_mite',
'Tomato___Target_Spot',
'Tomato___Tomato_Yellow_Leaf_Curl_Virus',
'Tomato___Tomato_mosaic_virus',
'Tomato___healthy']

Fig.2.sample diseased leaf image

3.2 Image Preprocessing:

The Plant Village dataset's image sizes are found to be 256 x 256 pixels. Then using the Keras deep-learning framework, data processing and image augmentation are carried out. After that, a training/testing split was applied to the data applying 80% of the images for training and 20% for testing.

3.3 Convolution Layer

Convolutional layers are the core part of CNN where lot of computation occurs. A filter moves across the receptive fields of the image checking whether the feature is present. Next the filter shifts by a stride, repeating the process until the filter has slid across the entire image. The final output is the feature map which is the dot product of input image pixels and the filter. Fig.2. illustrates the convolution operation of input image size 5*5 and filter size 3*3 and stride 1.

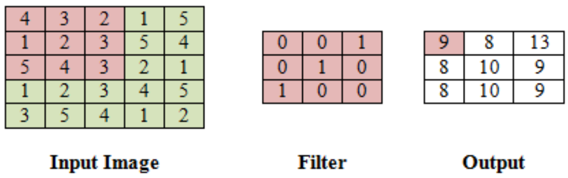


Fig.3. Convolution Process

3.4 Max Pooling Layer

Pooling layer perform down sampling by reducing the number of parameters in the inputs. In the pooling layer the filter does not have any weights. The kernel applies an aggregation function to the values within the receptive field. Fig.3. illustrates the maxpooling with filter 2*2 and stride 1.

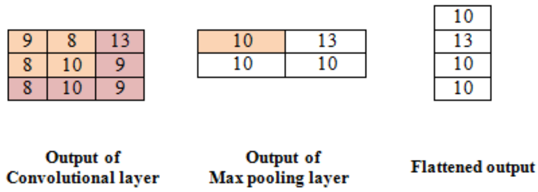


Fig.4.Maxpooling Process

3.5 Flatten Layer

Once the pooled featured map is obtained, the next step is to flatten it. Flattening comprises converting the entire pooled feature map matrix into a single column matrix which is then fed to the neural network for processing.

3.6 Fully Connected Layer

As previously stated, the output layer in a CNN is a densely linked layer that flattens and provides information from the other layers in order to turn the output into the appropriate number of output classes or labels. The word "Fully Connected" in the fully connected layer point out that every neuron in the next layer is connected to every single neuron in the previous layer. This stage is complete up of the input layer, the fully connected layer, and the output layer. This layer is described as a Multilayer Perceptron which consumes a softmax activation function present in the output layer. The input representation is flattened into a feature vector and sent through a network of neurons to predict the output probabilities in the fully-connected operation of a neural network. The pooling and convolutional layers outputs explain the high-resolution properties of the input picture. The fully connected layer on the foundation of the training dataset consumes these features for classifying input images into different classes.

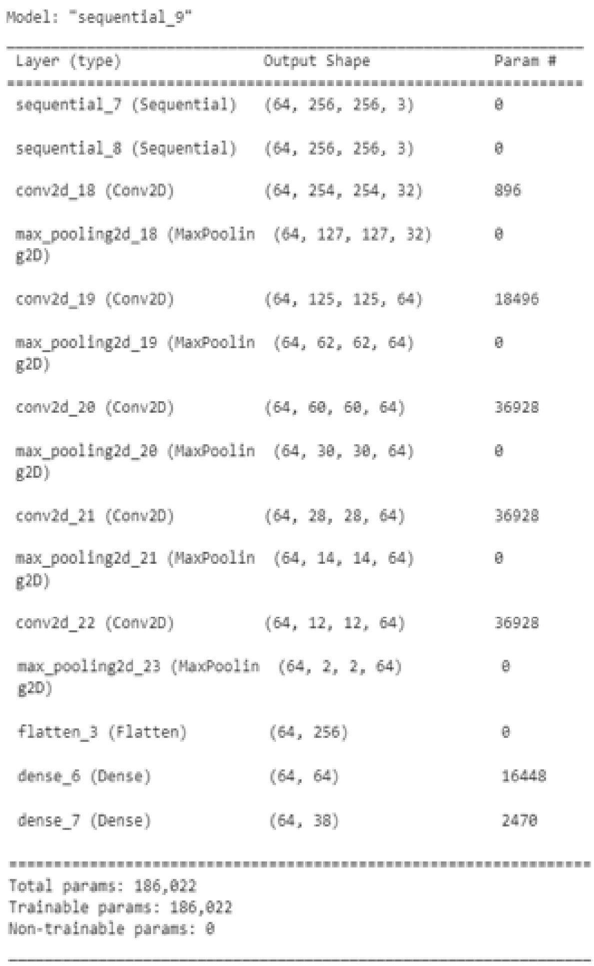
3.7 Softmax Layer

The soft-max layer creates a probability distribution with one output value.

- Softmax is implemented by a completely linked network layer at the end output.
- There must be the same number of neurons in the Softmax layer as there are in the output classes.
- Like a sigmoid function, the Softmax function compresses the outputs of each component to be between 0 and 1. It does, however, split each result so that the entire sum of the outputs equals one.

3.8 Model Summary

It is used to view every parameter and form in every layer of our models. It is observed that the total number of parameters is 186,002 and the total number of trainable parameters is 186, 002. The non-trainable parameter is zero.



```
Model: "sequential_9"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 sequential_7 (Sequential)   (64, 256, 256, 3)         0

 sequential_8 (Sequential)   (64, 256, 256, 3)         0

 conv2d_18 (Conv2D)          (64, 254, 254, 32)        896

 max_pooling2d_18 (MaxPoolin  (64, 127, 127, 32)       0
 g2D)

 conv2d_19 (Conv2D)          (64, 125, 125, 64)        18496

 max_pooling2d_19 (MaxPoolin  (64, 62, 62, 64)         0
 g2D)

 conv2d_20 (Conv2D)          (64, 60, 60, 64)          36928

 max_pooling2d_20 (MaxPoolin  (64, 30, 30, 64)         0
 g2D)

 conv2d_21 (Conv2D)          (64, 28, 28, 64)          36928

 max_pooling2d_21 (MaxPoolin  (64, 14, 14, 64)         0
 g2D)

 conv2d_22 (Conv2D)          (64, 12, 12, 64)          36928

 max_pooling2d_23 (MaxPoolin  (64, 2, 2, 64)           0
 g2D)

 flatten_3 (Flatten)         (64, 256)                 0

 dense_6 (Dense)             (64, 64)                  16448

 dense_7 (Dense)             (64, 38)                  2470

=================================================================
Total params: 186,022
Trainable params: 186,022
Non-trainable params: 0
_________________________________________________________________
```

Fig .5. Model Summary

4. Results and Discussion

The performance of the proposed CNN model is analyzed using accuracy. Accuracy is defined as following equation,

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \qquad (1)$$

where TP refers to True Positive
TN refers to True Negative
FP refers to False Positive
FN refers to False Negative

In fig.5. the x-axis represents number of epochs and y-axis represent the metrics via accuracy and loss. from the graph it is observed that the model performs well for 50 epochs and reaches an accuracy of 90.62%.

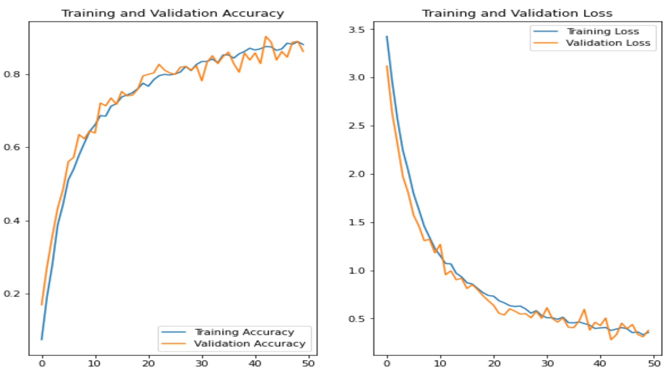


Fig.6.Performance analysis

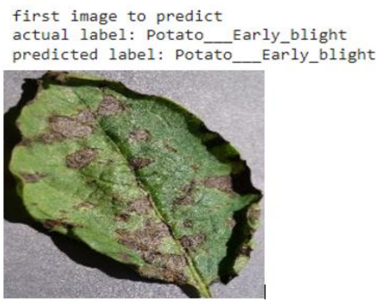


Fig.7.illustrates the prediction of the trained model

4. CONCLUSION

There are many advanced methods for identifying and categorizing plant diseases using image processing and deep learning. In this paper, a customized CNN is employed to identify plant leaf disease using images of healthy and diseased plant leaves. The CNN model was trained for various epochs. The model performs well for 50 epochs with same training and validation accuracy. The model yields the accuracy of 90.62%. Isn future, the hyperparameters of the model can be tuned for yielding the best accuracy results.

REFERENCES

[1] C. A. Priya, T. Balasaravanan, and A. S. Thanamani,(2012) “An efficient leaf recognition algorithm for plant classification using support vector machine,” in Proc. Int. Conf. Pattern Recognit., Inform. Med. Eng. (PRIME), pp. 428-432.

[2] Munish Kumar,Surbhi Gupta,Xiao-Zhi Gao,Amitoj Singh,(2019) “Plant Species Recognition Using Morphological Features and Adaptive Boosting Methodology”, IEEE Access,vol. 7, pp. 163-170.

[3] Sandika, B., Avil, S., Sanat, S. and Srinivasu, P., 2016, November. Random forest based classification of diseases in grapes from images captured in uncontrolled environments. In 2016 IEEE 13th International Conference on Signal Processing (ICSP) (pp. 1775-1780). IEEE.

[4] Singh, V. and Misra, A.K., 2017. Detection of plant leaf diseases using image segmentation and soft computing techniques. Information processing in Agriculture, 4(1), pp.41-49.

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

[6] Ramcharan, A., Baranowski, K., McCloskey, P., Ahmed, B., Legg, J. and Hughes, D.P., 2017. Deep learning for image-based cassava disease detection. Frontiers in plant science, 8, p.1852.

[7] Shrivastava, V.K., Pradhan, M.K., Minz, S. and Thakur, M.P., 2019. Rice plant disease classification using transfer learning of deep convolution neural network. International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences.

[8] Hassan, S.M., Maji, A.K., Jasiński, M., Leonowicz, Z. and Jasińska, E., 2021. Identification of Plant-Leaf Diseases Using CNN and Transfer- Learning Approach. Electronics, 10(12), p.1388.

[9] Robert G. de Luna, Elmer P. Dadios, Argel A. Bandala, “Automated Image Capturing System for Deep Learning-based Tomato Plant Leaf Disease Detection and Recognition,” International Conference on Advances in Big Data, Computing and Data Communication Systems (icABCD) 2019.

[10] Peng Jiang, Yuehan Chen, Bin Liu, Dongjian He, Chunquan Liang, “Real-Time Detection of Apple Leaf Diseases Using Deep Learning Approach Based on Improved Convolution Neural Networks,” IEEE ACCESS 2019.

[11] Balakrishna K Mahesh Rao, "Tomato Plant Leaves Disease Classification Using KNN and PNN," International Journal of Computer Vision and Image Processing 2019.

[12] Prajwala TM, Alla Pranathi, Kandiraju Sai Ashritha, Nagaratna B. Chittaragi, Shashidhar G. Koolagudi, "Tomato Leaf Disease Detection using Convolutional Neural Networks," Proceedings of 2018 Eleventh International Conference on Contemporary Computing (IC3), 2018.

[13] https://data.mendeley.com/datasets/tyw btsjrjv/1

[14] Priya, C.A., Balasaravanan, T. and Thanamani, A.S., 2012, March. An efficient leaf recognition algorithm for plant classification using support vector machine. In International conference on pattern recognition, informatics and medical engineering (PRIME-2012) (pp. 428-432). IEEE.

[15] Jeyalakshmi, S. and Radha, R., 2021. Classification of Tomato Diseases using Ensemble Learning. ICTACT Journal on Soft Computing, 11(4), pp.2408-2415.

[16] Wallelign, S., Polceanu, M. and Buche, C., 2018, May. Soybean plant disease identification using convolutional neural network. In The thirty-first international flairs conference.

[17] Akila, M. and Deepan, P., 2018. Detection and classification of plant leaf diseases by using deep learning algorithm. International Journal of Engineering Research & Technology (IJERT), 6(7), pp.1-5.

[18] De Luna, R.G., Dadios, E.P. and Bandala, A.A., 2018, October. Automated image capturing system for deep learning-based tomato plant leaf disease detection and recognition. In TENCON 2018-2018 IEEE Region 10 Conference (pp. 1414-1419). IEEE.

[19] Jiang, P., Chen, Y., Liu, B., He, D. and Liang, C., 2019. Real-time detection of apple leaf diseases using deep learning approach based on improved convolutional neural networks. IEEE Access, 7, pp.59069-59080.