Potato Disease Classification Using Transfer Learning

Blessy Daffodil C Department of Data Science and Business Systems SRM Institute of Science & Technology Chennai, India Blessydaff25@gmail.com

Abstract—Food is a necessary and the key to survival of mankind and hence it is always is huge demand. As much as the need for the food is high, the supply to it contrasts with it. This might be because of the number of reasons among which one of the reasons is the failure of crops and the plants in the agricultural field. Thus, the necessity to save the plants from dving is increased and the way is to detect the defects on an early stage and thus the prevention could also be done in order. Latest technology in the field of Information Technology, has proved to be an outstanding benefit not only in our desired area of agriculture but also in applied in many other areas. In this article, we use transfer learning-based models to classify images of diseased Potato plant leaves into three types of plant diseases based on their defect using a Plant Village dataset. In this project, two pre-trained models ResNet50 and MobileNet are used and discovered that MobilNet performs better on training data with accuracy of 93%.

Keywords—Potato disease, potato leaves, transfer learning, deep learning, CNN, Artificial Intelligence, resnet, mobilenet

I. INTRODUCTION

Food is important to mankind and the demand for food is increased but the cultivation and production of food is being at risk by so many factors. Such factors include diseases in the crops and plants. Because of this, farmers face the consequences being at risk to losing their livelihood by trying to save the plants and their farming with their financial resources. Hence the technology's help is needed to detect the diseases in plants or people should consider to take food by importing than from the locals which will serve as a pain in both financial and health wise. The technology help is needed because some diseases in plants and the leaves are not recognizable to the human eye. Hence images of the diseased plants are considered as data for the mentioned problem statement and with the help of deep learning the diseases are recognized. Here the plant of interest is potato and diseases of potato are tried to be recognized.

A. Scope

Agriculture is one of the backbones of our country and it provides key benefits for one to lead one's life. It also provides many opportunities such as jobs to the population of the country. The major risk that are faced by farmers are the plant diseases because the diseased plant will also ruin the healthy plant by spreading. Hence the technology's help is needed to detect the diseases in plants or people should consider to take food by importing than from the locals which will serve as a pain in both financial and health wise. The technology help is needed because some diseases in plants and the leaves are not recognizable to the human eye. With all these problems one side, agriculture is not picked by many as personal choice as a decent employment and Srividhya. S Department of Data Science and Business Systems SRM Institute of Science & Technology Chennai, India Srividhs1@srmist.edu.in

many prefer the Desk jobs than the jobs on the field. Hence, the population in the field of Agriculture is less. The changes in the country now a days has led to decent paying for IT jobs and agriculture and the jobs in any field work is not greatly paying. Hence the technology has taken the upper hand and has decided to solved the agriculture issue. The main goal of the project is to identify not only diseased but what kind of diseases the plant might have so when it passed to the next phase or the next team, it will be easier to take necessary action. The reason why this needs the help of today's technology is because trying to identify a disease from a plant itself is tedious process. If the same must be done for an entire land or even more, the process would become time consuming. Not only that, but also some of the diseases in plants are not seen to the human eye and even if it does, sometimes it can go beyond rightly recognizable. One technology that can meet the earlier mentioned requirements today would be Computer vision. The computer vision helps agriculture to save plants by rightly recognizing the disease and it would save a lot of time and will lead to save farmers' lives and henceforth saving the livelihood of farmers and the population of the country. automated system which can help farmers to identify plant disease through computer vision would help farmers to save disease through computer vision, would help farmers to save time, plant, economy as well as all the efforts put by the farmers will not go in vain. Hence, a deep learning model is proposed in this project and few transfer learning models also to automate the process of recognizing the diseases in potato plant.

B. Methods

• Deep learning

Prior to the development of Convolutional Neural Network, the image recognition was being backed up on the traditional algorithms. But for the image processing, based on what the model is going to recognize the features must be recognized as per the requirement. But this process may be seen as a challenge due to the fact of defining features for different image types. A Model may study the feature and its representations on behalf of this which leads to deep learning representing the features for image on several levels of representation.

This project utilizes the neural network to recognize plant leaf images, and the pre-trained models are used to evaluate the performance based on the comparison of accuracy.

• Transfer learning

In deep learning, the popular one is the transfer learning due to the enormous number of resources requires to train these models. The TL is not exclusively related to problems

such as multi-tasking. When a model is trained on an experiment and the same model is re-assigned to different task is done using the transfer learning. The learning is improved in the new experiment based on the already acquired knowledge previously from the experiment i.e., it is transferred and hence the transfer learning. The former network is trained on a specific dataset for e.g., the imagenet and for the dataset related tasks. Later the same network is re-framed to be fed to the next similar experiment. The network cannot be applied to the next experiment if it does not find itself similar it. The TL works well with the image as the data also. For such kind of tasks, it is very usual to use the pre-trained model for huge image recognizing projects. Usually, the TL can be applied on tasks where the network needs to save time and do a decent performance at the same time. As mentioned earlier, it is ideal when a task that is relevant to the formed network with abundant data is found. Based on Transfer Learning for the precise diagnosis of plant diseases, CNN model was developed. The dataset here used is called the Plant Village and it consists of 2152 different photos and 66 other different images for validating and it is holding 3 unique directories of potatoes' leaf images. The focus is mainly here on resnet 50 and mobile net which are the pre-trained models.

II. LITERATURE STUDY

Numerous studies on image categorization and identification have been conducted. The author and the researcher in [1], used Maple and Hydrangea leaves with two different types of leaf disease were pre-processed, and their features were then extracted in order to be analysed. In order to segment the leaf into three parts—the infected, the leaf part, and the background part K-means clustering and ANN were utilized. The leaves were then categorized according to their disease.

The authors in [2], have concentrated on fine-tuning several convolutional neural network parameters for deep learning in order to classify normal images of cat and dog images. An ANN binary classifier is used for classification after the convolutional neural network is made to learn features. The network's performance is improved by using several levels of refinement, and this approach yields the best classification accuracy of 88.31%.

In the Reference [3], utilised a convolutional network of neural form to categorise pictures of food. Deep Learning is used to categorise 16643 food photos into different food categories. In the experiment, accuracy of 92.86% is attained.

M. Shaha et al. [4], done an image classification using pre-trained model by adjusting the VGG19 model that has already been trained. On two separate picture datasets, CalTech256 and GHIM10K, the model's performance is compared with two transfer learning models termed VGG16 and AlexNet as well as a hybrid Convolution Neural Network (CNN) model with an SVM as a classifier. According to the study's findings, VGG19 outperformed the other three models.

In the Reference [5], using a transfer learning model, the CNN model that is applying transfer learning to classify a collection of HEp-2 cell images. The photos in the dataset are divided into six categories primarily on their staining patterns only after model first applies feature selection to identify the key features that best represent them. They demonstrate at the end of the study that their CNN algorithm outperformed the other 4 algorithms created by earlier researchers.

Sindhuja et al. [6] described a fast, cost effective and reliable health monitoring sensor. To monitor plant health and diseases, they represented various technologies that have been used to detect the plant diseases.

Waldchen et al [7] released a review on the use of image processing techniques for plant disease diagnosis. Nearly 120 research articles were examined in their review, in addition to a detailed explanation of datasets.

Erika et al. [8] suggested a four-layer CNN model that incorporates seven different illnesses and healthy cucumber leaves. They noted the excellent and poor image quality and discovered an average accuracy of 82.3%.

Powara et al in [9] contrasted a few manually created feature descriptor strategies with CNN models. Their comparison includes HOG-BOW combined with SVM and MLP classifiers, as well as HOG-based characteristics mixed with KNN and kNearest Neighbors (HOG). Those models were contrasted with those created entirely, AlexNet and GoogleNet.

Piyush et al. [10] segmented a specific region of interest in the photographs of plants using color-based procedures. In this paper, illness spots were found using the YcbCr and CIELB colour models.

In the Reference [11], Some textural characteristics were discovered, including fatigue, uniformity, and coherence. They used photos to determine the mixture of dark and white level form, connected it to colour choice, and found illnesses of corn leaves.

Sachin D[12] used a neural network as a classifier with back propagation before using few segmentation techniques to gather features from images of leaves.

Authors in [13] set of leaf images taken as testbed and the critical elements make up the suggested framework, which is based on image processing. The K-Means technique is being used to segment the images at hand; after segmentation, the segmented images are then run through a neural network that has already been trained.

Melike et al. [14] CNN was designed as a tool for and extracting features automatically classifying. For study on plant leaf diseases, visual information is frequently used. In the approach, three channel components are subjected to the filters. The resulting feature vector from the convolutional component was input into the LVQ to train the network.

Authors in [15] applied the option adapted to determine the type of disease for categorizing grape leaves is the inception v3 architecture. Even while mobilenet provides accuracy comparable to that of inception v3, the model still overfit the provided dataset. The Adam optimizer was used here.

III. PROPOSED METHODOLOGY

A. System proposed

The work flow of the system as shown in Fig 1. The dataset is from Kaggle and is downloaded from there.



Fig 1. Proposed System Flowchart

The it is processed and removed redundant information. The train test split of the dataset is 80% and 20%.

The data which is clean after the process is given to the pretrained models, these models are already trained with some other dataset such as imagenet so that it can be used for transfer-learning. The deep learning technique is also used in here, which is a CNN model used for feature extraction for training-testing. The moels of pre-trained will be examined for performance with this data.

B. Dataset-taken

54,306 number 256x256 pixel images of Kaggle dataset are taken as the dataset in our experiment for testingtraining purpose. And in particularly potato leaf images contains 2152 images which of 1000 from potato-early blight and 1000 from potato-late-blight and 152 from potato-healthy as shown in Table1. Each image is RGB images and which means it is coloured and will have 3 channels such as red, green, blue.

TABLE 1. TOTAL NUMBER OF IMAGES IN DATASET

DISEASES	NO. OF IMAGES
Early Blight	1000
Late Blight	1000
Healthy	152

C. Dataset pre-processing

To decrease the unwanted data and noise from the data, we do pre-processing and this will make the accuracy increase and speed-up the executing. Data-pre-processing is done with Data-Augmentation, which is getting the different characteristics of the data images and combining it. Rotation, flip, zoom, fill, shear etc are the methods which is applied on every image for augmentation. Library which is used for augmentation is Keras.

• Augmentation of data

For getting big accuracy, the deep-neural network will need a lot of data. Sometimes the image sizes will not be enough for this. So, at this kind of places, we will use some data methods like flipping, rotating, zooming, shearing etc. to each image. These techniques will make new set of datasets which will be good for training purposes. It creates a new set of images from the already existing data and this process is called data-augmentation. The images needed for the training is not captured by ourselves. This is already available in public. The meaning of the term augmentation of data means that, different techniques like rotation is applied. Some of the data-augmentation techniques are as, Geometric - transformation, color - space - augmentation, filtering applied to kernel, picture blend, erasing randomly, feature-space-augmentations, adversarial training. generative-adversarial-network, neural - style - transfer, meta learning. Data augmentation strategies target overfitting at the training dataset, which is the source of the issue. This is done with the expectation that augmentations will allow for the extraction of more data from the original dataset. By data warping or oversampling, these augmentations artificially increase the size of the training dataset. Warping of data augmentation changes the already available images but it will remain the labels. In this technique, the methods are processes like erasing randomly, adversarial-training, colour changes such as Gray and geometric such as rotation, flip etc, and neural-styletransfer. Over samples are added to the training images to get the synthetic styles. The example of synthetic styles is blending of images, augmentation of feature-space, and generative-adversarial-networks. The data augmentation safety is ensured by retaining the labels and hence changing the data content will not affected by it. This process is safe for general image identification tasks such as identifying cat and dog, but this is not a good practice when comes to tasks like digit and signs. In that case, rotation and flip will create meaningless data. For forecasts which is un-certain, the nonlabelling technique will be efficient. Post-augmentation labelling is adjusting for this method to ensure this. The label as well as non-label preservation to the data will be giving better performance for the training as well as increase the accuracy in prediction-time.

• Deep-learning

This is a sub part of AI which calls artificial intelligence. Artificial-intelligence is the process of making the machine to learn itself and it will be capable of making decisions. Deep-learning is the technique which is under AI which is idea from brain of human and the neurons and learning process of it. There will have neurons like human and will act somewhat similar to them. We have used 2156 images for the deep learning purpose, which is learning the features of each image with neurons. There should be training as well as test data for this because if we use the same data which we have used for the training is taken for the testing as well will not give good accuracy. So that we will split the whole data into training-testing data. The

number of layers in CNN which is used in here is 3. And after that we add 2 maxpooling layer then two dense layers. 128,38 respectively at the first and last. To minimize overfitting of data, we can add dropouts of 25 percentage and 50 percentage. There must be activation functions to activate the neurons, for this purpose we use softmax as well as relu. The first function used between the layers and the second function used at the last. The total number of epoch used is 20 for training and testing and the batch size is 32.

• Transfer-learning

Transfer-learning is the technique which is used for prediction. In this case the model will be initially trained with a particular data and with that knowledge it will be trained to another set of data so that it can predict much efficient than the general convolutional networks. In CNN, the model will be trained from the scratch and it will be built for a particular problem. But in transfer learning, the model can be used for any problem and will be initially trained with some data. And we will mention this data training as weights when we load the pretrained model. Here the initially used dataset is ImageNet. The workflow is in Fig.2.



Fig.2. Transfer-Learning

The models for transfer learning is used general are vgg16, vgg19, resnet50, inceptionv2 etc. Pre-trained models which used here are resnet50 and mobilenet with weight imagenet.

IV. MODULES

A. Creation of models

The dataset contains Potato leaves images belonging to 3 different classes as shown in Table.2 with a total of 2156 images of 256x256 pixels were used for training and testing the model. Pre-trained models resnet50 and mobilenet are used as models here and they are pretrained with imagenet data and they are loaded here. Then the data features will be extracted and this will be fed to it.

CLASSES	TRAIN DATA	TEST DATA
Early Blight	805	195
Late Blight	805	195
Healthy	143	35
TOTAL	1669	417

B. CNN

CNN is a method used for machine to understand things like human brain and which is called convolutional neuralnetwork. It will have layers of neurons and the it will identify the features from the raw image. This information will be taken as input and there will be some output. There will be some loss after the output. It will be adjusted by back-propagation. three layers in CNN The are convolutional layer, fully connected layer, pooling layer. Making the model learn from the initial stage is a big task and it makes no more efficient than the pre-trained model. There are many pretrained models generally used, such as alexnet, mobilenet, vgg16, resnet. The data for implementation taken is the Potato images in Plant Village dataset which contains 2156 training images and 20% testing images.

Resnet 50

The common models in transfer learning are resnet50, alexnet, googlenet and vgg. These models behave different with different datasets. Such as the resnet can give 90 percentage or more accuracy with some dataset but will give less accuracy with some other dataset. But the pretrained model which is likely efficient in neural networks is resnet compared to vgg. Vgg have a greater number of layers and will take much time compared to resnet for training. So that we took resnet for the training purpose. But both of them are equally used for transfer learning. The problem arises with the deep neural networks are network-optimization, vanishing of gradient, and also the problem with degrade. The accuracy of detection will be higher with resnet because it uses new-methods.it reduces the problems comes with the deep neural network training, as well as the degrade, saturate of precision. The model resnet which is used in here have fifty layers of neurons.

The loops used here to notate the use of later layer in the current layer in the structure. The reduction and fill out and lowering of gradient of data is the difference between common neural networks and the resnet50. this is because the precision will increase first and the reduce there. The first layer of 64 and kernel of seven into seven, max-pooling layer is three into there. The grayscale layer shown first is 3 layers and they are same in all manner. The next phase has four layers of the same configuration and third one has four identical layers in same. The blue colour loops present in here are the connectors for two different types of layers. The classification is done by the last layer of 38 layers. The fully-connected layers are not used in the model which is used in here.

The images in our dataset will be transformed into 224 and 224 dimension to feed to the model. Data augmentation such as flip, rotate, etc will be done after the resize of images. The model which is used in here is have weights with the imagenet. SGD optimizer is also used in here. For activation softmax activation function is using. We can high the number of types of diseases in here, previously it will not be able to do it. The classifier is created with the help of keras and it will be added to the resnet model. And it will act as the feature extractor. We can change the imagenet

weight to some other one to test the precision of the output. Here we are experimenting which model is good for the dataset. Generally resnet and vgg show a good accuracy.We are comparing resnet and mobilenet here. The accuracy is good for mobilenet compared to resnet.It took lesser time compared to the other one as well as gives good result in training as well as test than the other.

• Mobilenet

The pretrained mobile net which is used in here have 2 levels. And they are before the pretraining and after the training with our data images. The first level is loading of pre-trained mobilenet model with imagenet weights. The top layers of the models makes untrainable then for the training.2D global average pooling layer of keras is used for feature transform and the layers of mobilenet is also used along with it. After that the feature vectors got from the images have given to it. The image augmentation output is taken as input and it is fed.

C. Fine tuning

We have trained the two models resnet and mobilenet with the plant village dataset of potato images. The training is done with the models with the top layer freeze. So that the top layers will not learn in training. Now the top layer will be unfrozen to make them learn. And, the bottom layers will be frozen. The model will have knowledge of the first training and with that, we will train the model again. Using this technique accuracy will increase compared to the conventional neural network training.

CNN or convolutional neural networks are layer of neurons in which they accept input and process output. They have different layers of neurons. They are convolutional layer, activation, pooling, fully connected, batch normalization, dropout. The convolutional layer sometimes includes the input layer also. The convolutional layer is the important part of the neural network. The input image will be divided into pixels and the input layer will take reach of the pixels and it will do scalar product with the weights assigned to the next layer. This will be fed to the next layer and so on.

Convolutional layer contains kernel or called them filters. These filters will make a feature map, which is nothing but another matrix like structure from the image. Filter size or kernel size is smaller than the image always. If we take a cat image, the nose, whiskers, eyes etc are some properties of that image which can be used to differentiate it from other images. This in the computer sense is called features.

Activation layer such as relu will decide which neuron must be activated. Before the activation, a bias value of each neuron will be added to the weights. So the neurons which are activated will transfer the output.

The pooling layer will conclude the features got from the convolutional layers as feature map. Which is, it will reduce the dimension of the feature map and it will be fed to the fully connected layer. The top layers which are frozen throughout the training are convolutional layer and dense layer. And the last dense layer will only be trainable. This layer will unfreeze then in the fine-tuning phase.

V. RESULTS

The results after train and test of the two models are show in the Table.3. The train and test has done with 10 epoch each. Resnet model gives more accuracy compared to mobilenet in the training time. But the test result is high for mobilenet. The loss rate is high for resnet compared to the other.

Model	Train Accuracy After 10 Epoch	Train Loss After 10 Epoch	Test Accuracy After 10 Epoch	Test Loss After 10 Epoch
ResNet50	0.45	0.41	0.58	0.68
Mobile Net	0.39	0.02	0.89	0.14

The resnet model have a higher size compared to mobilenet of 98mb as well as number of layers. Mobilenet have a depth of 88 layers. The train accuracy of resnet is slowly increasing with the epoch which is higher than the validation. Validation accuracy remains almost steady after the 3rd epoch.



Fig.4 Train and Validation Loss ResNet50

From Fig.3 and Fig.4, it is observed that Train and Validation Accuracy of resnet was in increasing manner from the start. And the final training accuracy was 45% and training loss was 40.1%. The training loss decreased gradually from the start but the validation loss remains steady till the end.



Fig.5 Training and validation accuracy MobileNet



Fig. 6 Training and validation loss MobileNet

Training accuracy and loss for MobileNet is 39% and 2% respectively. Training and validation loss was in gradual decrease as depicted from Fig.5 and Fig 6 together.



Fig.7 Training accuracy & loss of Resnet Fine-tuning

After finetuning, the accuracy of models has increased than the initial training and shows the results after finetuning. Accuracy of training of resnet has increased beyond 80 % from the 12th epoch. The training loss have decreased after 10th epoch and reduced to zero after 15th epoch is seen from the Fig 7.





Fig. 9 Loss of Mobilenet Fine-tuning

The Training accuracy of mobilenet increased from the tenth epoch to the end, that is 20th epoch. And the loss has the behaviour of decreasing from the beginning till the end is seen from the Fig 8 and Fog 9.

TABLE 4: RESULT AFTER FINE-TUNING

Model	Train Accuracy	Train Loss	Test Accuracy	Test Loss After
MobileNet	93%	0.09	0.63	0.05
ResNet50	82%	0.02	0.57	0.06

VI. CONCLUSION

The usage of AI is increasing day by day. The agriculture field have also need it the most. Finding the solutions to problems in agriculture such as loss of vegetation with the diseases and the soil degradation with the usage of improper fertilizers needs to end. This can be done with deep learning and machine learning techniques. The precision of the output won't go higher in deep training compared to transfer training. We got accuracy of almost 93 percentage with the mobilent model with the plant village dataset. And also, we compare the two pre-trained models resnet50 and mobilenet with the same dataset. For almost all the images, the models give correct prediction. The farmers who are new to the field will need the system the most.

REFERENCES

- Aakanksha Rastogi, Ritika Arora and Shanu Sharma," Leaf Disease Detection and Grading using Computer Vision Technology &Fuzzy Logic" 2nd International Conference on Signal Processing and Integrated Networks (SPIN), 2015.
- [2] Dhanabalan, S. S., Sitharthan, R., Madurakavi, K., Thirumurugan, A., Rajesh, M., Avaninathan, S. R., & Carrasco, M. F. (2022). Flexible compact system for wearable health monitoring applications.Computers and Electrical Engineering, 102, 108130.
- [3] M.T. Islam, B. M. N. Karim Siddique, S. Rahman, and T. Jabid, "Image Recognition with Deep Learning," 2018 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS), Bangkok, pp. 106-110, 2018, doi: 10.1109/ICIIBMS.2018.8550021.
- [4] M. Shaha, and M. Pawar, "Transfer Learning for Image Classification," 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, pp. 656-660, 2018, doi: 10.1109/ICECA.2018.8474802.
- [5] H. T. H. Phan, A. Kumar, J. Kim, and D. Feng, "Transfer learning of a convolutional neural network for HEp-2 cell image classification" 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), Prague, pp. 1208-1211, 2016, doi: 10.1109/ ISBI.2016.7493483.
- [6] S. Sankaran, A. Mishra, R. Ehsani, and C. Davis, "A review of advanced techniques for detecting plant diseases," Computers and Electronics in Agriculture, vol. 72, no. 1, pp. 1–13, 2010.
- [7] J. Wäldchen, and, P. Mäder, "Plant Species Identification Using Computer Vision Techniques: A Systematic Literature Review," Archives of Computational Methods in Engineering, vol. 25, issue 2, pp 507–543, April 2018.
- [8] E. Fujita, Y. Kawasaki, H. Uga, S. Kagiwada, and H. Iyatomi, "Basic investigation on a robust and practical plant diagnostic system," 15th IEEE International Conference on Machine Learning and Applications (ICMLA 2016), December 2016
- [9] P. Pawara, E. Okafor, O. Surinta, L. Schomaker, and M. Wiering, "Comparing Local Descriptors and Bags of Visual Words to Deep Convolutional Neural Networks for Plant Recognition," 6th International Conference on Pattern Recognition Applications and Methods (ICPRAM 2017), pp. 479-486, 2017.
- [10] P. Chaudhary, A. K. Chaudhari, A. N. Cheeran, and S. Godara, "Color transform based approach for disease spot detection on plant leaf," International Journal of Computer Science and Telecommunications, vol. 3, no. 6, pp. 65–69, 2012.
- [11] J. K. Patil, and R. Kumar, "Feature extraction of diseased leaf images," Journal of Signal & Image Processing, vol. 3, no. 1, p. 60, 2012.
- [12] S.D. Khirade, and A.B. Patil, "Plant Disease Detection Using Image Processing," 2015 International Conference on Computing Communication Control and Automation, Feb. 2015.
- [13] D. A. Bashish, M. Braik, and S. B. Ahmad, "A framework for detection and classification of plant leaf and stem diseases," International Conference on Signal and Image Processing, Dec. 2010
- [14] M. Sardogan, A. Tuncer, and Y. Ozen, "Plant Leaf Disease Detection and Classification Based on CNN with LVQ Algorithm," 3rd International Conference on Computer Science and Engineering (UBMK), Sept. 2018
- [15] Pazhani. A, A. J., Gunasekaran, P., Shanmuganathan, V., Lim, S., Madasamy, K., Manoharan, R., &Verma, A. (2022).Peer–Peer Communication Using Novel Slice Handover Algorithm for 5G Wireless Networks.Journal of Sensor and Actuator Networks, 11(4), 82.