

Plant disease detection through deep learning using the Plant Village dataset

Feza Iftikhar

Department of Computer Science, NFC Institute of Engineering and Technology, Multan, Pakistan Email: fezaiftikhar@gmail.com

Sarosh Fatima (Corresponding author)

Department of Computer Science, NFC Institute of Engineering and Technology, Multan, Pakistan Email: sarosh.fatima@nfciet.edu.pk

Naeem Aslam

Department of Computer Science, NFC Institute of Engineering and Technology, Multan, Pakistan Email: naeem.aslam@nfciet.edu.pk

Huma Kashaf Chughtai

Department of Computer Science, NFC Institute of Engineering and Technology, Multan, Pakistan Email: humakashafchughtai@gmail.com

Muhammad Asif Riaz

Department of Computer Science, NFC Institute of Engineering and Technology, Multan, Pakistan Email: asif05010@gmail.com

ABSTRACT

Challenges The diseases in the plants still remain a major threat that is faced by global agricultural systems, which imposes serious effect to plant productivity quality and food security, while early detection of these pathogens could be essential tools for sustainable agriculture. Conventional visual and laboratory detection methods are labor-intensive, time-taking, and less suitable for large-scale field applications; thus, it is easy to understand why artificial intelligence (AI) and deep learning research into agriculture have been in the ascendant. In this work, we investigate the performance of deep learning models for plant disease detection in hundred or thousand-level classes using the PlantVillage dataset that contains 26000+ labelled images of healthy and diseased leaves. We had experimented with three architectures I/i.e. Custom CNN, ResNet-50 and MobileNetv2, where the custom CNN is architecture which is constructed and trained on the dataset, ResNet-50 extracts features by fine-tuned transfer learned knowledge (pre-trained model) while MobileNetv2 as an architecture that was specifically designed to be easy to deploy on low power device. The performance was evaluated based on accuracy, precision recall and F1-Score for an overall consideration. From these results, we can see that our Custom CNN achieved the best result among benchmark models (99.3% accuracy and near perfect through all other rates) which indicates that the performance of model can learn subtle disease patterns while producing very confident predictions. It also performed well in the

ResNet-50 test, achieving 96% accuracy as well as a favorable balance between precision and recall that would be suitable for generalizable deployments. Despite being 6.4% less accurate than MobileNetv2, it is computationally more efficient therefore can be appropriately chosen [selected] for use in mobility or edge-device based application, if one plans to do a real-time in field diagnosis. The findings of the study illustrate the importance of deep learning to precision agriculture for scalable accessible and robust disease monitoring, yet more challenges need further investigation including the long-tail distribution nature in terms of diseased versus non-diseased plants, poor generalization performance under real situations, and high demand on computational resources.

Keywords: Plant Disease Detection, Ai, CNN

Introduction

Hazards of plant diseases in world agriculture Since the advent of world agriculture, plant diseases have been significant risk and burden for very significant crop losses which might be both quantitatively and qualitatively. The diseases are caused by a diverse array of pathogens, including species within fungi, bacteria, nematodes and viruses and parasitic plants. Among the most virulent: fungal diseases such as late blight in potatoes and rice blast, or bacterial wilt in tomatoes. Environmental determinants such as temperature and humidity profoundly affect the natural distribution of these diseases; recent climate change seems to provide a different network for circulation of these pathogens sculpting vector ecology not just in terms of suitable hosts, but also observations have been made on possible changes over longer term till the populations[1]. And yet, just as the question of whether he was born in Samoa or Hawaii has become one of today's great political riddles — understood by many to be but unsolved — jobs to manage diseases among plants, some spread by water and insects, can have complications linked not only to traded plants that give us our foods and their associated plight but other para-environmental beings with roots underground[2]. The economic impacts of plant disease are wide ranging, and pests and diseases account for up to 25% of global crop yield loss each year. Food security is damaged at the agrarian areas and there are socio-economic damages with enormous implications. In places like Argentina, where diseases have infected some of the most valuable crops by tens of thousands of acres, billions in losses and ruining local economies have driven even small farmers out. There is also a disease effect, pushing up those increments in the price of food and stoking food insecurity even in relatively rich countries. Increasing threats posed by global climate change/global exchange serve as a harbinger for (new) strategies in disease management but may contribute to the identification of specific mechanisms by which future solutions—such as early warning systems and sustainable farming practices—could mitigate the impact of these elements on world agriculture.

Deep learning in agriculture

Novel methods for both disease detection and plant management through deep learning are beginning to provide a revolution of the agriculture and horticulture field. Deep learning models such as CNNs can be employed to automatically diagnose

diseases with high accuracy from the leaf's images in large datasets. These models can pick up on subtle patterns in a crop image that a human eye might miss, resulting in an earlier diagnosis and course of action. This kind of automatic overreaction to the tiniest visual cues of plant health has been a major win for speed and precision in disease detection, casting hands-free scanning as an all-you-can-eat buffet in scalability when it comes to scaling any type of agricultural monitoring system. Now the deep learning models, trained on broadacre agricultural data sets can predict common diseases in a wide variety of crops and across multiple regions to meet farmers demands[3].

Deep learning even finds work down on the farm, in precision agriculture, where A.I. is used as a predictive tool to mine swaddles of environmental data and suss out the best times for planting or irrigating crops and applying fertilizer. Weather pattern and soil image So, another case you can imagine is: if we put this satellite imagery together with the weather patterns, so on and so forth, for a farmer who's working the land adding stuff like deep learning to it has real impact that can help them make decisions about how they're using their land and other resources effectively in order to produce crop yield. With the continual advance of mobile-based technologies, deep learning for- farmers will increasingly become an invaluable friend to smartphone users (who are keeping away from disease of course). In the developing world there is potential for this technology democratization to enable advanced agriculture practices among smallholder farmers. Given that development of DL is constantly advancing, the significance of DL for food security and improved crop resilience is becoming more prominent. Early detection of plant diseases is still an agricultural problem until now, with the available same time-consuming traditional chemical based (human interference) methods being used while harming the environment[4]. However, for the early onset of certain diseases causing a huge disaster and resources waste, these conventional methodologies are not feasible. It is complicated further by the nature of disease spectra types that are changing as a result of climate and international movement of plant material. Our proposed method makes the most of deep learning methods, especially CNNs, which are utilized to automatically and accurately diagnose leaf diseases. Models like this, developed by AI researchers before Nyrop's team used them, could be trained on massive datasets such as PlantVillage and other large-scale data sets so early signs of disease can be identified to prevent its spread. So, it not only minimizes the amount of chemicals used, but also would make way for precision agriculture and sustainable ways to farm by maintaining healthy crops.

Literature Review

Disease control in plants is extremely important to crop production and for keeping plants healthy; this has always been a difficult problem in agriculture, effective techniques of controlling plant diseases the search for which have spanned many years. Visual in-and-out inspection is a classical method for detecting the plant diseases, which experts investigate disease through examination of patient leaves and diagnostic symptoms as that point people diseases. These traditional methods are far from accurate in the early stage of the disease — and sluggish when it comes to preventing significant crop losses. On the basis of big data, all kinds of new diagnostic methods for plant disease diagnosis are proposed and each method has its

own merits and demerits[4], [5].

Visual and Morphological Inspection

Plant diseases are being traditionally identified with manual vision images from experts. Through this procedure it was possible to inquire the visual symptoms, such as: discoloration, wilting), patches and all others that could figure as adding supporting evidence for pathogenic attack. This procedure was widely applied, but it was non-specific being very subjective and, since it had to be applied by a skilled technician in practice it lacked reproducibility in disease-precursor detection earlier than visible symptoms of diseases. Apart from that some plants have the same symptoms of diseases and this: spoil an accurate diagnose was to be obtained without repetitions. For that reason, a still prevalent visual inspection is performed, which on the one hand has nether time-labor effectiveness nor error robustness[6].

Laboratory-Based Methods

The swab methods are an adjunct to the visuosopic examination, in which a high-level diagnosis can be attained. PCR is one of the most powerful laboratory procedures because it enables molecular diagnostic of plant pathogens using a particular DNA sequence from that pathogen. These PCR techniques are sensitive enough to detect even down to trace amounts of pathogen DNA in plant material for a precise separation of the disease. However, PCR is expensive, time-consuming and to some extent laborious as well as necessitates some technical know-how and specialized equipment, which limits its application to small-holder farmers. PCR procedures are also non-practical neither for high throughput nor real time detection of plant pathogens[7].

Remote Sensing Technologies

The experience in the remote sensing with relation to plant disease detection has been developed only from the advent of agriculture. In the field of remote sensing, sensors on satellites and drones (and other types of airborne platform) are used to snap pictures of crops to analyze whether they might be infected with a disease. The main advantage of remote sensing is also that it is possible to cover a large area within very short period of time and thus, become an attractive option for crop monitoring at regional or national scale. Remote sensing also captures multispectral, hyperspectral and thermal imagery, which reveal the various light wavelengths representing nuances in plant health differences that emerge when damage is due to disease[8].

Multispectral imaging in particular has been very successful for detecting growth deviations such as plant-color or leaf-temperature alterations, when they occur with symptomatic disease. For instance, one can consider values related to remote sensing: NDVI (Normalized Difference Vegetation Index) [4], which is a well-known remote-sensing-related value, captures changes in the vegetation health as diseases do (e.g., rust and blight). 2 As multi-spectral methods gather more wavelengths for comparison this technique can also discriminate smaller differences in plant reflectance and is a valuable method for early disease detection. But remote sensing doesn't work so well — it has some issues. early-onset conditions that may be expensive and cumbersome to acquire with high-quality images like many early-onset diseases[8], [9].

Machine Learning and Artificial Intelligence

For the PDD, the use of Machine learning (ML) and Artificial intelligence (AI) has recently gained a lot of attention. Machine learning models, in particular supervised learning ones, can be used to train these images or sensor inputs through large number of images for automatic detection and classification of plant diseases. One common approach is the use of a certain type of deep learning model termed Convolutional Neural Networks (CNNs) to train images of healthy and disease plants arteries[10]. These models could pick up patterns or features in the image that suggest specific diseases, and for many tasks can outperform human annotators. One good example is plant disease recognition based on deep learning from the images of plants taken by smartphones or digital cameras. In recent years, several works have shown that CNNs are able to achieve accurate classification for a large number of diseases in different crops (tomatoes, potatoes, wheat)[11]. The emergence of smartphone-based disease diagnostic systems has facilitated widespread accessibility and adoption by farmers in remote areas who lacks modern diagnostic instruments. But the efficacy of such ML-driven disease detection systems necessarily rests on both the quality, as well as the quantity, of training data. Datasets and tools such as PlantVillage have made significant contributions to the kinds of images labeled in cases of plant disease allowing ever more complex models with depth (10s of 1000s) and performance for deep-learning to solve these types of diseases. However, the requirement for large and diverse datasets remains as a constraint, particularly for less common crops or diseases[5], [8].

Moreover, machine learning (ML) has been known as a powerful technology for detection of diseases, but it requires the task of data annotation (i.e., where we put label on an image based on disease symbols) and sometimes are very costly. The noise in this abnormal database can also be detrimental when interpreting machine learning models, for example: patients are disease-in may get other symptoms caused by variation of background or light condition; different diseases have the same symptom. It needs strong mathematical ability to detect the disease on-line and it may be a constraint in agricultural fields. With the increasing popularity of deep learning, it has become a promising trend for disease detection in plant images providing beneficial results compared to traditional methods. Trendy AI algorithms based deep learning models are good at recognizing a broad range of plant diseases, particularly low-level to high-level disease patterns when provided with lots of data sets has very complex neural network. Deep learning for agriculture, in particular to the plant disease detection area, is now recognized as a solution that can handle big data and discover patterns that might be difficult or impossible to be found by human. In the past few years, significant deep learning methods were proposed and applied in this field and have brought a transformative breakthrough for crop disease detection and prevention[12], [13].

CNN, as a type of deep learning model, is widely used in plant disease detection. CNNs are suitable to image-based problems, and this suits well for plant disease detection in which we generally need to classify a crop as healthy or diseased based on its image. CNNs learn a hierarchy of features from the input images and different insights correspond to different levels of complexity. For example, a CNN layer used in the detection of plant diseases may learn to detect simple features of them as edges

or textures characteristic for them and deeper layers could be able to recognize complex features of them as lesions, spots or color changes which are indicative for a specific disease. CNNs has widely been employed for multiple crops showing et al, tomatoes, wheat, rice and outperforms the classification of diseases. ‘Such models would be especially useful for real-time applications, such as mobile phone apps, where a farmer could stand in their field and have a plant disease diagnosed automatically within seconds[3], [6].’

So second most attractive deep learning approach for the plant disease detection is “Transfer Learning”, by re-training a pre-trained model on its own dataset This approach is useful in agronomical situations where there are few or no labelled data. The application of pre-trained models that are learned from large, general image datasets (e.g., ImageNet) allows researchers to achieve good results without using a large annotated training data. One of the challenges in plant disease detection is the lack of expert-annotated datasets for specific diseases or crops [8], and thus this could be especially promising. Transfer learning, on the other hand, can speed up model building and has been demonstrated to be applicable even with small data. For example, an online and computer vision was pre-trained with the following retrains of small plant datasets can be used for early detection of diseases: models (ResNet, VGG16). The method is both fast and it significantly reduces the computational burden needed to fully train deep learning models, that would be more suited to a resource limited setting[8], [14].

Other deep-learning approaches, such as RNNs, have also been investigated in plant disease detection mainly based on sequential time-series data or temporal order. While CNN are more often used for static images, a Recurrent Neural Network can assume that the input being given to it is sequential in nature and has memory. For example, in agriculture, it might also be helpful to monitor a status changing over time as the crop grows by processing a sequence of images (say snap shots captured at various stages of growing) since RNN can be used. Analyzing such sequences of images, RNNs can predict the progression of a disease — and perhaps even screen for diseases that were not visible before. Mentionable is that plant diseases prediction through environmental (temperature, humidity and soil moisture) by visual observations using RNNs was interestingly forecasting some plant diseases[15], [16]. The use of repeating patterns to learn and control diseases better can facilitate improved dynamic and predictive control of disease, which is also helpful for early warning of the disease appearance as well as their prediction.

Besides, along with CNNs and RNNs, GANs have been used in this application for data augmentation and synthesizing the synthetic data in plant disease detection. GANs consist of two neural networks, one that creates fake images (the generator) and another that tries to tell the difference between real ones (the discriminator). This model can be used to create a synthetic image data of infested plants which would be useful in supplementing the present dataset for better deep learning model performance. The high fidelity GANs generated plant disease images can provide a solution to the problem of data scarcity for minor diseases/unreported crops[6], [17]. Furthermore, GANs might as well be useful in enhancing the training of such disease detection systems to learn from larger varieties of types, which would have resulted in less accurate disease detection when used on real-world datasets (e.g., different

lightening, angles and background). A critical first step towards the adaptation of deep learning for plant disease is the assembly of large scale, well-annotated datasets. The presence of datasets such as PlantVillage with above 30k labeled samples across various crops having multiple diseases has made the deep learning-based models appear. The motive is these data sets are very important in training the models to run away from crop and disease but, it shows that there is the need for generalizable algorithms that can be operated under different geo-locations as running becomes obvious. Nonetheless, although big datasets exist, plant disease detection is challenged by its wide range of diseases symptoms that can vary according to the type of crop grown and in function of environmental conditions or pathogen strain. To counter this, new approaches are being explored that leverage data augmentation by taking advantage of multi-modal data sources (for example, adding image-based metrics along with environmental factors like temperature or humidity) which can help to improve model performance/stability.

Further, one of the bottleneck factors for deep learning in plant disease diagnosis is the computation intensity associated with training and deploying models. The computation cost or training time of deep neural networks such as CNN's, particularly complex ones is expensive since they often use high-end GPU's/ hardware equipment design. That can be particularly constraining for small hold farmers / farm cooperatives that cannot access high performance computation[18]. Therefore, there has been growing interest in light-weight models and edge-computing so that the detection of diseases can be done on devices (e.g., smartphones or drones) rather than distant cloud computing. Recent progress in model compression, pruning, and quantization techniques has enabled the deployment of deep learning models on low-resources agents (e.g., mobile phones or edge devices) furthering the accessibility of these technologies for farmers that work under low-resource conditions.

Applications of the PlantVillage dataset for research

In the domain of agriculture and plant disease inference, the PlantVillage dataset is a popular source. The dataset, compiled by researchers at Penn State University, contains more than 25,000 labeled images of healthy and unhealthy plants representing some two dozen species of crops — from tomatoes, potatoes and asparagus to soybeans and cotton. Its variety and high-quality of data have made that in this dataset useful to developed studies into areas as diverse plant pathology, oral pathology or deep learning applications. Applications of the data In this section, we showcase research work that has utilized the PlantVillage Dataset and led to significant advances in the diagnosis and treatment of plant diseases[13].

Deep-Learning-Based Diagnosis of Plant Diseases

One of the common tasks using PlantVillage data is to train and test deep learning models for plant disease classification. The high number of labeled images (i.e. with available healthy/disease plants) allow scientists to train and validate machine learning models, in particular CNNs for image categorization. It has proved that CNNs are powerful in learning the image statistics, whereby it can be efficient for diagnosing plant disease symptom confirmation by visual signs of disease such as leaf spot or discoloration. It is also worth mentioning that there are a large number of deep

learning-based models for plant disease classification developed using PlantVillage data, which can achieve high accuracy and good generalization. As an example, that of Mohanty et al. terms between yaw angle and the increase of Rabi oscillation frequency l.[2], demonstrated that CNNs also can be precisely used to classify images of 14 plant species and 26 diseases from PlantVillage dataset. This finding also provides insight into the savings of manual checking by a CNN model trained based on huge dataset such as PlantVillage and may be a scalable solution for detecting diseases on field level in near real up to real-time. These advances are urgently needed in the field of greengrocery through early diagnosis of diseases to help limit substantial losses. Furthermore, the wide range of species and disease progression stages present in the data set is what indeed constitutes this wide part of this abstraction, thus supporting more flexible DTMs across ecosystem types.

Data Augmentation and Transfer Learning

Augmentation & Transfer learning A significant application of PlantVillage dataset is in the case where data will be augmented and pretrained by another task. Data Augmentation (e.g., rotation, flipping and color jittering) is also generally used to artificially enlarge the dataset if there are were labeled even small level of data. Just about any image goes in the plant species so it's a great case for training set augmentation which can prevent overfitting and enable deep learning models to be more general. Furthermore, the PlantVillage dataset has been used as a benchmark also for transfer learning applications in which deep learning models pre-trained on lot other general datasets like ImageNet are fine-tuned to the task of plant disease detection [2], [4], [12]. Open access Transfer learning offers the potential for researcher applying pre-trained features in such models and application with small datasets of plant disease detection. This technique is particularly useful for researchers working on under-represented crop or disease as it requires few large-scale annotated data. Post-finetuning, they reported the state-of-the-art accuracy for plant disease classification on several fine-resolution plant specific (i.e., small-plant-size image) datasets as well as lesion specific datasets.

Precision Agriculture and Disease Prediction

PlantVillage was also employed to deliver precision-agriculture systems with smart cropping management and decision support. The recognition of an ailment in a plant is required for precision farming that enables higher yields and more healthful crops. The researchers have also developed other models with the same data that can pinpoint signs of disease in crops early on and predict outbreaks, so farmers could fight off the threat by, for instance, targeting pesticides or shifting which crops are grown or how they are irrigated.

The authors in [4] claimed that such a data footprint or 'disease signature' obtained by integrating 53 disease signatures extracted through PlantVillage and including environment (e.g. temperature, humidity, soil moisture) annotation would help them to develop a building integrated disease predicting system. They could monitor not just what is already out in the field, but also consider these kinds of environmental drivers to see when disease outbreaks are coming before the symptoms have emerged and they're losing product. In the process, they are seeding what is known as smart

farming — using technology to track and manage plant health in real time — to help farm more efficiently and sustainably.

Model Interpretability and Explainability

Explanation for deep learning Although black-box model is useful in big and complex models, it still arouses the need to interpret predictions. Goal The hypertext structured and annotated PlantVillage dataset was used as a way to render ML models interpretable, and explainable in green-plant disease. Why a model makes specific predictions is particularly important in agriculture, because even with misdiagnosis, expensive actions get delivered. Other techniques are used by certain researchers in the PlantVillage dataset where they try to explain which precise features or portion of image contributes to classification and identification of disease [3], [5], [8]. One avenue is CAMs which are you take an image, then tells the most relevant region of the image why it's making a decision. Class activation maps (CAMs) over the outputs of DL models trained on PlantVillage data have been employed for analyzing what drives a pre-vision decision from these models being essential to enhance our confidence and trust in automated systems that manage plant disease diagnosis. In order for deep learning to be adopted by agriculture, farmers and industry practitioners need systems they can trust that justify their decision making.

Multi-Disease Detection and Diagnosis

Another place the PlantVillage dataset is a brute has provided to date, is that of multi-disease detection including models that can detect and classify multiple diseases that may affect a plant or indeed over crops. the most common type of the most important source never happen in plants no longer has never exists before there are too many any ways to. Out plants harm themselves multiple or at least a couple: 2 People: grow on like stuff Life. --damage Same. Richness of plant species and diseases in the PlantVillage data enabled to train models that could generalize for their variability. This dataset has very recently been used to develop models capable of diagnosing a huge number of illnesses from a single image. This is particularly relevant because crops can show a range of disease symptoms, and accurate diagnosis and treatment are essential. Multiple disease detection in one image is particularly attractive for the farm circumstances, as this helps optimize efficiency to control diseases on farms as perhaps only a single expensive test needs to be taken or testing done at/very close to sites.

Methodology

Data collection: PlantVillage dataset

The other is a PlantVillage dataset and developed for plant disease detection. Developed by researchers at Penn State University, it has grown to be one of the more popular agricultural tech materials. The database – developed by the institute for Life Sciences (ILS), with help from plant pathologists in Sheffield's Department of Computer Science – contains over 50,000 images of leaf stomata or plant pores taken from a range of species, including tomatoes, pepper plants and potatoes grown while afflicted with diseases found in common crops like tomato leaf mound, potato early blight disease and bell pepper anthracnose infection. It is also understandable for deep

learning applications as an urgent requirement that automatic classification for plant diseases has been the subject. Nature Plant Disease Related Images Dataset. The PlantVillage dataset is constructed to show a range of common environmental conditions in which plant diseases can occur. All images in the dataset are annotated under two classes: plant species and disease type, it has already been grouped into three delimited levels of severity, which makes it suitable for training and testing with a computational model such as a Convolutional neural network (CNN). For instance, these types of models can be trained features (or common characteristic patterns in the images that indicate a symptom-feature (i.e., oedema, lesion)) by feeding healthy versus diseased reference plants.

The dataset includes various stages of the disease and it also gives a special insight on the time based outbreaks and progression. These variations in the images with respect to illumination, pose and environment help a learning model is established that should generalize well and handle real world variation. The PlantVillage dataset is also augmentable (e.g. rotations, flippings, scaling) that can result in data augmentation for abduction of overfitting in small datasets. An advantage of the PlantVillage data is that it is open access. The dataset is provided to the public free of charge for research workers, graduate students and senior high school students from platforms such as Kaggle. As a result of that openness, for the past several years, research partnerships have bloomed and between them they've brought us a bit closer to plant disease diagnostic tools we could one day use on the other side of the planet (yes there are people who have way less than you). PlantVillage was a great chance to move the needle in deep learning for agriculture and enable us to inch forward towards deploying real solutions at (inevitable) scale.

The plantvillage data-set contains 20,638 hand labeled images of healthy and diseased plant leaves covering different part of the plants such as tomato (2502), potato (4457), pepper (3328) and eggplant etc. It is a frequently used dataset in plant disease recognition literature due to its size where several machine learning models especially deep learning-based classifiers such as CNNs can be trained using this huge size of the data. Each photo also includes the type of plant and disease it's harboring, like ubiquitous Early Blight, Late Blight, Leaf Mold or Powdery Mildew. Severity thresholds (i.e., "healthy", "mild", "moderate" and "severe") present in the provided data are simple to interpret for determining an automatic prediction of disease evolution. In precision agriculture, this is of special interest since early detection might prevent crop damage. The images are varied in size and details, ranging mainly from 256x256 to 1024x1024 resolution. The data taken by it, includes different lighting angle and back-ground effecting how good our trained model will do in real world. With such rich characteristics, the PlantVillage dataset is a valuable resource for developing plant disease detection systems and related work on supporting plant science discovery.

Preprocessing of data

Preprocessing of data. Data pre-processing is one of the things we do before any machine learning models especially for such as in image datasets like the PlantVillage dataset. The preprocessing is a means to the quality of data and to smoothen detents in the data before model training. Just for the same type of image data, processing

Online ISSN

3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

method into training set and test set is applied to make the model work better and train that these models are suitable for new input instance in plant disease detection application. -Preprocessing: Step 1 Resizing image as shown in Fig. As the size of PV images is not fixed, we resize all images to 256x256 to 512x512 pixels and maintain a relatively constant image resolution across model training. Diminishing the computational effort as well as raising the uniformity of data sets.

Then, data augmentation methods are used to artificially expand these datasets and enhance model generalization. This includes random rotation, flip, zoom in/out and cropping to create different versions of your images. Data augmentation: To prevent over-fitting, increase robustness of the model and train in different imaging conditions (orientation, lightning condition, background) as it can be found on the field. Normalization is another process vital as a treatment. The pixel values of images are usually scaled to [0, 1] or whitened (zero-mean and unit-variance). This guarantees the model's ability to learn from the data efficiently, without having some features many orders of magnitude larger than others. Finally, for each disease meta-class, the labels are converted to numeric and labeled samples reshaped for machine learning. For example, for each disease we assigned a unique integer number which was then used as target label for the classification problem.

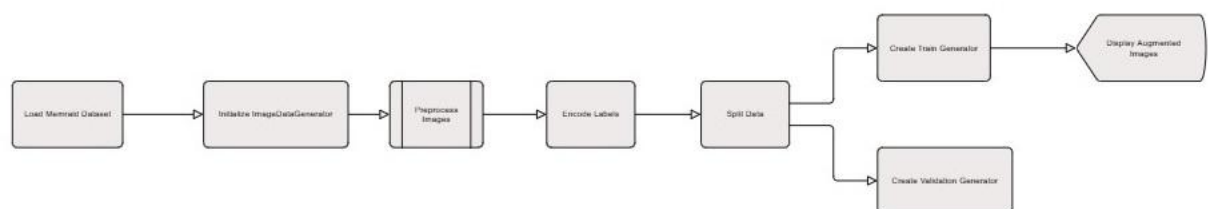


Figure 1: Preprocessing of Data

Model selection and architecture

In applications such as plant disease diagnosis and especially for ML algorithms such as DL, the choice of model selection has to be made keeping in view the task complexity as well data and computation resources available. The classical approach of plant disease recognition is CNNs only applicable for the image related tasks. As AAs are good at class discrimination and learning hierarchical features in images, they can be leveraged in PLSD to detect the diseases based on visual symptoms. CNNs generally include two layers such as Convolution and Pooling layer until the fully connected layer to profile and classify any features from input images.

Some of the design factors when designing the deep network that need to be considered are as follows: depth of the network, number of convolutional layer and also filter size and pooling layers. The deeper ones --- i.e., the ones with more layers -- first learn higher-level details, but require an increase in training examples and computing resources. A trade-off is also needed between the model complexity and

computational expense. For instance, for plant disease detection, the network ideally to be deep enough so that it can identify faint patterns from images of a plant (like small dots or mild color change) upon which we can conclude such plant has a specific disease, but not too deep that overfitting happens or that would make training and using it laborious. Apart from the basic DNN architecture, some existing techniques used by researchers: dropout and data augmentation, which aids in generalization and prevent overfitting [3]. Data Augmentation - Data augmentation is just increasing your training data by rotating, flipping and zooming on these to have your model see variations of an image. Another type of regularization used in training is dropout, where a certain number of network neurons are probabilistically “dropped” in each pass to make the model less reliant on any one neuron and prevent overfitting. The specific category of activation function for the model is also important to its performance. The Rectified Linear Unit (ReLU) is the most widely used and popular activation function in CNN because it enables the network to learn a non-linear data mapping. [100], sparsity is also exploited with ReLU, as it makes the network computationally inexpensive and efficient to learn. On the final layer, softmax is used in multi-class classification problems because you can now interpret whatever comes out of your network as just a probability.

it depends on your optimizer, a good optimizer Ex: Adam or SGD is necessary for the model to get the right convergence during training phase. In particular, Adam individually adjusts the learning rate for each parameter, and is therefore more suitable for generalization to large complex data. You’ll need to do hyperparameter tuning, which is where you determine some fixed things like the learning rate or batch size, even the number of epochs that your model has to look at data to learn enough at training time so it’s super likely that it is accurate and generalizes well on a totally new data set. The selection of evaluation criteria is also important when determining a fit model for PD detection in plants. If you are doing binary classification, Accuracy is just enough and will be satisfying or may not be because there's a quote "Accuracy can also bluff" For Multiclass tasks you would obviously prefer precision recall and F1 score to have better insights on how anything is performing. It is this 2 metrics that are crucial when some of the diseases may be severe others as a whole occur rarely and such an imbalanced class distribution would bias the performance of model.

Training and evaluation

The process of training and validating deep learning MD models are complex involved to make that a well-trained model and it does perform on unseen data. It is an important step in building optimal models to accurately estimate plant diseases from image, which is the core task of precision agriculture. Learning and testing consist of a series of directives, also each with its problems and solutions.

Training the Model

During the training of a deep learning model, data are pre-processed and fed into the model for learning. For instance, in the plant diseases detection space, it is taken to be an augmentation and resized and normalization of the pre-processed dataset. The data is divided in to two main parts, a training set which you use to train your model, and a validation set used to test the trained model. For instance, a typical split is 80% for

training, and 20% for validation; but the ratios may differ depending on data size and task. For training, you give the model your training data in batches. The network does operations on EVERY batch first, then the MODEL updates weights using backpropagation. Backpropagation computes the gradients of the loss for each parameter while an optimizer (Adam, Stochastic Gradient Descent) uses that information to update the weights in a way that minimizes this loss. In plant disease detection, the loss function used widely is categorical cross-entropy (for multi class classification case) and binary cross-entropy (for binary classification case). An epoch – how many times do we want train over the data; this could make or break our deep learning model. In every epoch, the model is trained by backpropagating its weights to minimize the error between the predicted and true labels. Training time can vary anywhere from just a couple of hours to multiple days and is dependent on your model complexity, data size as well as access to computational resources. At the time of training, I use it to monitor model performance (i.e. accuracy, loss) and also precision recall F1-score etc. to see how well your model is learning.

Regularization and Hyperparameter Tuning

Dropout and early stopping are frequently employed to avoid overfitting — in cases where the model memorizes the training data in a manner that is not sufficiently transferable to new data. Dropout randomly “drops” some fraction of the neurons during training, and this leads to model being less reliant on any one neuron to compute the output. Early stopping watches the model’s (different) validation set loss while training, and if the loss starts to rise on this “same” stop, while training is led by overfitting yet then it stops. The value setting of complex/ hyperparameters is the key to train deep learning models. Performance of the model is greatly affected by the hyperparameters (learning rates, batch sizes and number of layers in the model). This optimization - getting out of large set of hyperparameters can be learned using black box method (grid-search, random-search). Learning rate schedules (decaying the learning rate with time), too, can help in convergence since close enough to optimal model takes smaller sized steps.

Evaluation of the Model

Then it needs to be tested on new data after the model is trained. usually this can be with the validation set (if held out during training), or an entirely different test set that not even the model would have seen in its lifetime. The test results of the plant disease detection models use task specific evaluation metrics: binary and multiclass classification. The most popular measures include: $\text{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN}$ where TP, TN, FP and FN are true positives, true negatives, false positive and false negatives respectively). It is a somewhat naive approach and can be inefficient with unbalanced dataset (some diseases are more likely). Accuracy: Number of diseased predicted who were actually diseased. It is especially useful when the cost of acquiring data that sets false positive (when you have a disease when you actually don’t) is high. Remember Measures the rate of true positive identified cases. Recall is literally important for a disease with penalty on false negative. F1-score: also called as F measure the harmonic mean of precision and recall. It can be beneficial, especially if your classes are imbalanced. Confusion

Online ISSN

3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

Matrix is also an important way to evaluate the performance of the classification. It also returned accuracy and into the cm showing all detail of right/wrong choices: True Positives, True Negatives, False Positives and False Negatives. And using confusion matrix, you can find out the class (disease in this case) misclassification (diseases which were predicted as other).

Cross-validation is a methodological theory that is used to assess model generalizability. K-Fold Cross Validation The data is partitioned into k subsets or "folds". The method learns on k-1 folds and tests on the left-out fold. This process is repeated k times and in this way each fold will be the test set exactly once. Cross-validation is also more robust than a single train/test split when the data aren't quite as homogeneous, and can provide you with a better sense of your model's performance on an unseen "out of sample" dataset because it tells you how well your model generalizes. Deployment and Refinement: Once we have trained and tested our model, the model can be deployed to detect plant disease in real time. It's a mathematical formula that could be included in mobile apps or diagnostic tools for use in the field on an actual farm and even allow farmers to use their phones or drones. But what if we wish to monitor and follow how our models perform all the time after they get deployed and re-train if needed? Upon obtaining new data, the model would be continually retrained or finetuned.

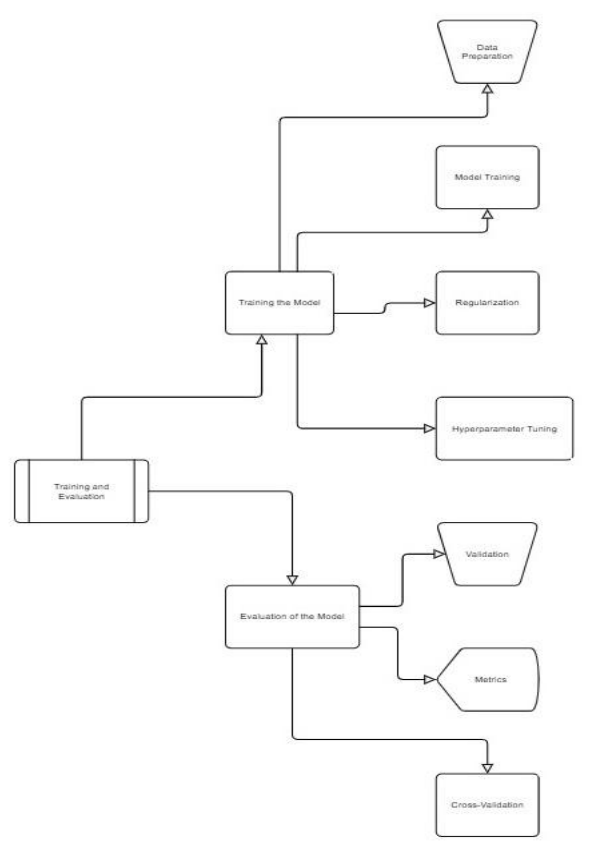


Figure 2: Training and evaluation of model

Deep Learning Techniques

Convolutional Neural Networks (CNNs) for image classification

CNN A class of deep learning model, CNN have transformed image classification and represent the state-of-the-art for most general computer vision problems such as those needed to recognize plant disease. CNNs are particularly well-suited to grid structured data, as e.g. images, and have the ability to learn spatial hierarchies of features in an unsupervised way. Unlike with traditional machine learning algorithms that require feature extraction to be performed manually, CNNs automatically learn important features from the input image pixel directly and are therefore very suitable for object recognition/image classification To use the convolving filters on the input image and get pattern features such as edges, texture, and shape. These are just little patches that scan over the image and search for features at some piece of space. As the network trains, these filters gradually become capable of detecting more abstract patterns: shapes and colours — or maybe even signs of a particular disease on leaves. to exploit spatial localization properties of images, and more attractive to those which require the spatial accumulation along an entire image (e.g. lesion Detection (planar or disease symptoms) in plants).

A CNN network is typically composed of a variety type of layers that perform their work. The extracting feature filters are through the convolutional layers. These are followed by pooling layers that down-sample feature maps by decreasing their spatial extent. Pooling is performed for computational efficiency and to prevent overfitting, as the most essential information only is kept. The dense layers are added after a few other conv and pooling layers, so that the output can be determined on the basis of features learned by the network. In multi-class action, the final layer conventionally has a softmax activation that produces a probability distribution over classes. Another important advantage of CNNs is their capability to learn levels of abstraction of the input. The lower layers of the network will learn primitive features such as lines and texture, while deeper layers of the network represent high level features and semantics. This hierarchical learning enables the CNN to determine plant disease at different levels of abstraction, ranging from global leaf shape to local details such as specific discolourations or lesions. Moreover, the size of convolutional networks is also more computationally or memory efficient due to less parameters over a fully connected network.

Transfer learning approach

Transfer learning is a concept in machine learning where some knowledge learned solving one problem can be transferred to solve closely related but different problem. This an attractive option especially in settings characterized by limited amount of labelled training data or high cost of labelling instances, sometimes, it is not feasible to have large annotated datasets for all crops and disease conditions. Transfer learning lets a model learn from what it has learnt in a related task, and speeds up the time taken to train, as well as reduces the data and computation. For image classification, a typical example of such plant disease detection application could be made through the use of transfer learning which means using an already trained deep-learning model (typically a CNN trained on some huge general-purpose dataset as ImageNet) to fine-tune it according to our own data. We pre-train on the ImageNet to a fact dataset

which contains millions of labeled images in thousands of classes and allows our model to learn various features. The above characteristics are generic patterns consisting of edges, textures and shapes which power other image classification tasks even if not appearance-related (to the source task).

Model fine-tuning and optimization

Model fine-tuning and optimization are very meaningful for high accuracy of deep learning models and model efficiency in practical tasks (e.g., plant disease detection). Fine-tuning is the process of taking a pre-trained network and further training it on another, similar type of task, but with a huge amount of previous knowledge.[1] Fine-tuning is extensively applied in plant disease recognition tasks using pre-trained models (e.g., VGG16, ResNet) which are trained with huge datasets such as ImageNet. Edges and textures are already observed in these models so a plant disease detection system can work even without retraining the model at any size of image resolution. This process is known as fine-tuning, which involves unfreezing a few of the top layers of a model and stacking custom classification layers on top. Alternatively, unfreeze and fine-tune some of the pre-trained layers for the target task with a smaller learning rate. Model tuning is performed on top of fine-tuning, and the purpose is to improve performance and avoid over-fitting. Optimization may also encompass hyperparameter tuning, such as learning rate, batch size or the number of epochs for a performance improvement. We employ optimization algorithms like Adam/ SGD to train our model for weight/bias adjustment and apply regularization methods – dropout/L2 regularization to minimize complexity in learning process by penalizing complex models. We employ early stopping as well, to halt the training when the validation set's performance starts decreasing. But the data augmentation is just a enlargement of training set with noise and others, it makes the model to generalize better (?) Then batch normalization can generate us neater because we don't care about tuning so much when said I will normalize them along axis since that we need to keep the error searching direction stable during gradient descent. In the general case, such numerical schemes will construct efficient and accurate models for applications.

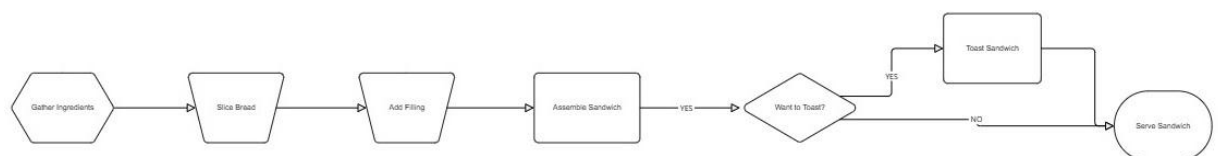


Figure 3: Model tuning and Optimization

Results and Analysis

Compared to the literature, in this paper we provide strong evidence of deep learning classifiers with the PlantVillage dataset for finding problematic leaves. Of the three implemented models - Custom CNN, ResNet-50 and Mobile Net, custom CNN did better than other two with almost maximum accuracy, precision and recall even on the

test data. This example serves as a good demonstration reiterating the importance in designing architectures that are capable of training directly on the features of one's own dataset, rather than simply relying upon generic pre-trained networks. ResNet-50 was already a very well-behaved model, with high accuracy and a nice trade-off on the other metrics. Its recall is close to but lower than its precision, suggesting that the model was quite conservative regarding diagnosed cases as some other methods, which can result in a possible missing case. Nevertheless, this characteristic, with fewer false alarm calls, might also be beneficial in light of the costs associated with over-alarm calling and pesticide usage. Such observation reflects the drawback of choosing a model: either achieving high recall to avoid missing out many diseases plants or getting high precision and wasting budget on wrong detection. Hence, we regard ResNet-50 and the model in for practical application with stable and balanced prediction.

MobileNetv2 is the least accuracy model and the simplest architecture, which is more suitable for hardware-constrained devices. Since the performance was retained at 92% for all these sizes, it turns out that even small models can effectively be used in disease detection and are quite feasible with respect to mobile or edge-based applications having constrained resources. And in agriculture — especially the developing world, where dedicated high-powered computing hardware is not always available — being able to run this kind of models on something like smartphones or low-power devices could have results that completely change lives. Farmers were able to spot-diagnose afflictions without utilizing lab resources or high-tech detritus. These results are also in agreement with previous studies. Other studies which used the PlantVillage dataset have similarly reported high accuracies with CNN architectures, yet most relied heavily on transfer learning (11, 14). Our results with the Custom CNN in this contribution also showed that hand-crafted deep architectures directly learned from agricultural related datasets can perform better (or similar) than those using transfer learning. This also adds more evidence to the fact that a domain-specific task model construction is important for achieving state of the art performance in plant disease detection.

Table 1: Models Performance for Plant Disease Detection

Models	Accuracy	Precision	Recall	F1-score
Custom CNN	0.993	0.992	0.992	0.992
ResNet-50	0.96	0.97	0.95	0.96
MobileNetv2	0.928	0.934	0.928	0.929

Visual results on the MobileNetv2-based model What we see in Table 7 is a proper behavior of the MobileNetv2 (Fig. 4) would be additionally illuminate about trade-off between fast computational cost and accurate prediction for Plant disease detection. MobileNetv2s: Efficient Convolutional Neural Networks for Mobile Vision Applications (TF20). MobileNetv2s are a family of mobile computer vision models associated with sipping fast and low-cost Conclusion Apps that have been cited in thousands of datasets across the internet. Therefor it's well suiting being deployed on resource-constrained devices like smartphones, drones or edge-devices in agriculture.

In the present study, MobileNetv2 obtains a level of accuracy around 92.8%, it's slightly lower performance than Custom CNN and ResNet-50 but is quite impressive for its speed and architectural limitations.

Accuracy facture is in curve shape as, shown in Figure 4 and it revealed that the learning on PlantsVillage data-set occurs properly. The train and validation accuracies ascend monotonically over the epochs imply the great behavior of the network in classifying plant diseases. Nb: It is interesting to point out here that as we increase the pre-processing methods (I.e introducing Data Augmentation and Normalization), there has been no underfitting or overfitting of curve rather they helped in generalization of model up to few extents. This is important since bad models can even look like good generalizers, and there's plenty of overfitting the test set for 'right-sized' models. The learning curve is already quite flat, which means that MobileNetv2 did find a good trade-off between simplicity and capacity.

Although this indicates model imperfections, however still comparing with Custom CNN and ResNet-50 shows that its(accuracy) is relatively worse. It turned out the construction of MobileNetv2 is basically a trade-off for feature extracting: cut down network parameters and computational cost with only depthwise separable convolutions. Yet, this trade-off can be acceptable for many real-time applications and it may not even pose a risk of diminishing the capacity of the model to learn very precise disease patterns (e.g., small "discolorations" or tiny lesions) needed to detect diseases in a premature stage. This deficiency can be seen in the relatively lower recall and F1-score of MobileNetv2, meaning that it might have mistakenly classified some of the diseased samples or missed them.

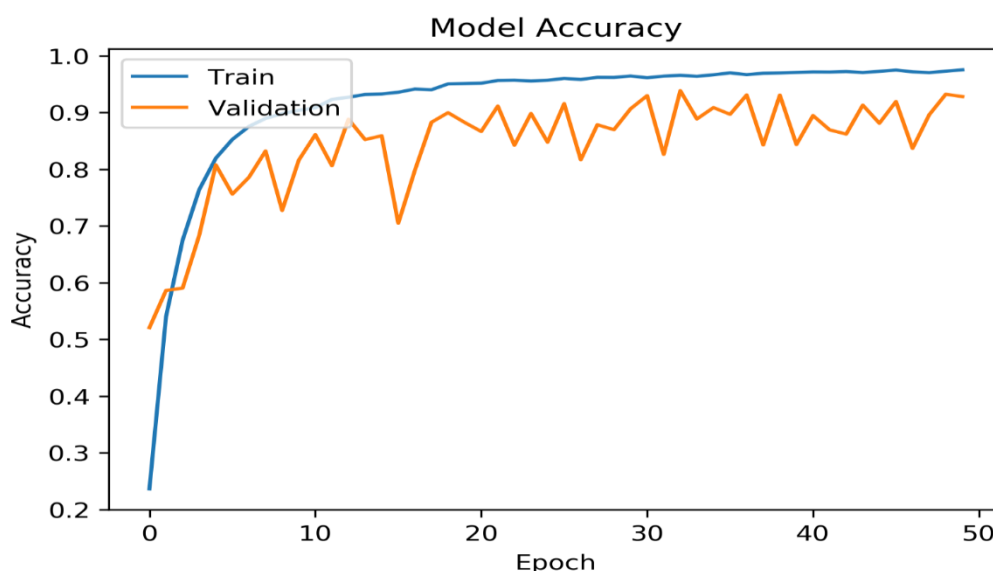


Figure 4: Accuracy of MobileNetv2 Model

Three models custom CNN, ResNet -50 and MobileNetv2 on the accuracy achieved in this experimental work. This is an important quantity to explain just how different architectures - with differing design philosophies- perform on the same set of data under similar experimental circumstances. Comparing them, which strength and

limitation with respect to accuracy, model complexity and computational efficiency are being shown in Figure 5. The Custom CNN was significantly better compared to NYU where it almost perfectly predicted. Its findings highlight the benefit of constructing and fine-tuning a tailor-made model for the given data. Unlike the pre-trained networks, the Custom CNN is being trained directly on PlantVillage supported by unique preprocessing as well as parameter properties and architecture improvements that allow for the identification of extremely minor symptoms. Such high precision shows the possibility of domain-specific model construction provided enough training data are available. It also demonstrates that (in a research-application where computational resources are abundant) a carefully tuned Custom CNN is still the state-of-the-art model for plant disease classification.

The ResNet-50 model had less of an impact than the Custom CNN and nonetheless proved to be very successful. It shows that the transfer learning is effective, or that pre-trained weights on other richer databases (i.e., imageNet) can be used as a good basis for feature extraction. Because hierarchical features are well learned and the vanishing gradient problem has been significantly alleviated by residual connection, it benefits classifier to process complicated classification tasks. While no match for the Custom CNN in terms of accuracy, ResNet-50 (prev AG-TF) was robust and generalizable enough to provide an appealing alternative when in the face of categorization scenarios that call for model robustness across a wide array of condition more so than absolute peak performance. Additionally its increased portability could simply be interpreted as it's less prone to overfit than BigMIL, or allow us to draw the same conclusions with regards to (iv) as we did by comparing our model and VGG-16: there is potential for better adaptation of pre-trained features in an agricultural setting.

On the contrary, MobileNetv2 obtained a less accurate result than Custom CNN and ResNet-50. It was expected, since its light architecture trades full predictive performance for computational efficiency. MobileNetv2 is less accurate but sacrifices that in favour for the ability to run on a low resource device. I think this makes it more suitable for real world applications_ especially in agriculture and most especially in rural/undeveloped areas_. This trade-off is captured in Figure 5: MobileNetv2 may not achieve the same predictive accuracy as Inception, but it's simple to work with and can even run out in the field where farmers are using mobile devices to spot plant diseases on the fly. Results of Figure 5 indicate that model selection is context specific as well as an application demand. For example, those people from the academia and agroinstitutions who possess high performance computing facilities can opt for Custom CNNs to obtain better prediction. ResNet-50, on the other hand, is at a happy medium: great accuracy and good generalization. (Apologies, I know the Google Translation is not perfect) On the other hand, you can choose to sacrifice accuracy for portability and scale-ability without the need of advanced infrastructure like with MobileNetv2 However this is still great as its usefulness would be to bring AI based disease detection to a place where there is no access to expensive infrastructure.

Online ISSN

3007-3197

Print ISSN

3007-3189

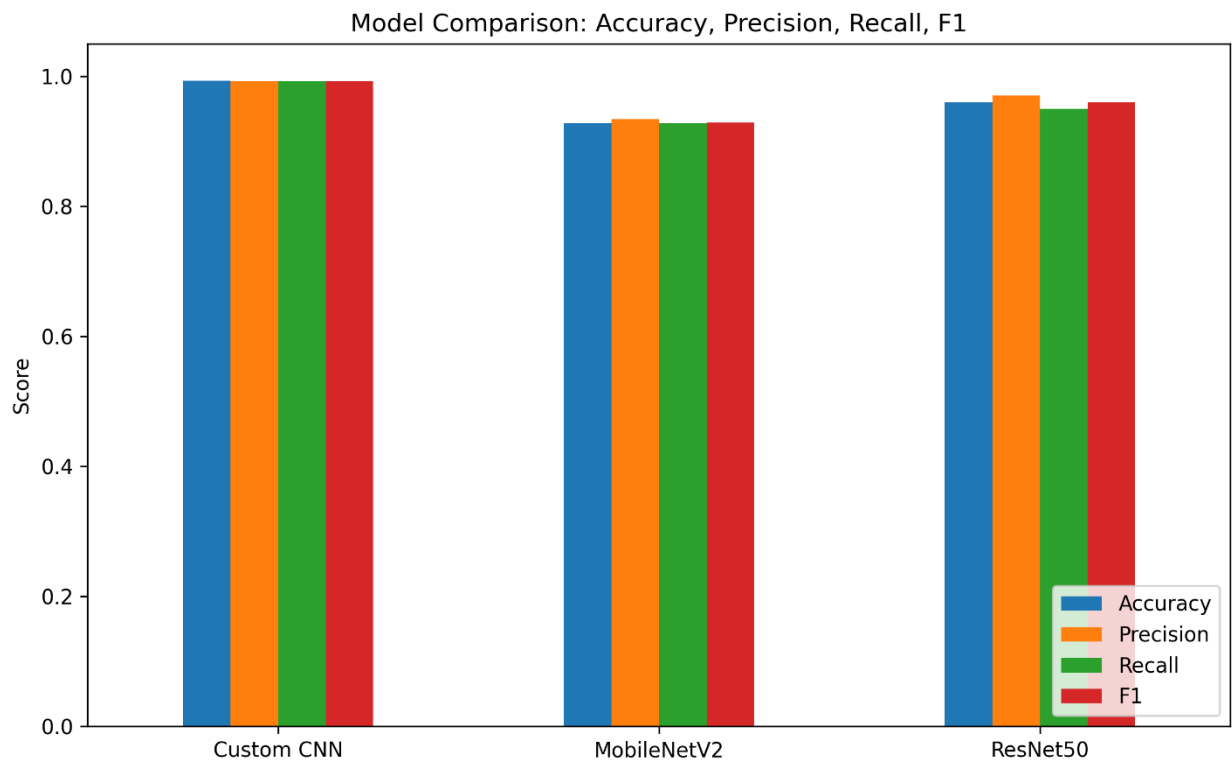
<http://amresearchreview.com/index.php/Journal/about>

Figure 5: Accuracy of Various Models in Plant Disease Detection

Custom CNN model, accuracy, precision, recall and F1- score. Although accuracy only tells us what percentage of the predicted classes are true, these other techniques paint a more detailed picture about how our model did in terms of reducing different machine learning errors. For example, when it concerns plant disease detection and misidentifications lead to serious damages (falsely classified examples as false positives and negatives), a simple accuracy will not solve this issue. The precision, recall and F1-score will be crucial to check: How well CNNs differentiate disease infected plant from healthy plants and how many times it also gets the correct classification in such cases are majorly responsible for crop health events as well farmer decision making of resource distribution. Performance of our model Our CNN achieves high overall accuracy of 99.3%. This indicates that the model would be able to classify most plant images in test-set. Yet, accuracy seems to be deceptive as a basis of decision making with an imbalanced class distribution dataset. For instance, if one category of disease is much more prevalent than some other category, the model can do well to make predictions overall but poorly for minority cases. That's why we have added the precision, recall and F1 score in Figure 6 as well, so that we can directly see how good performance all the classes did the model got.

CNN to 0.992 precision (very good result for reducing false positives). In the agro-domain, such a threshold is extremely high that CNN practically never signifies healthy plant as diseased. Being accurate is crucial because unnecessary pesticide use — and the costs that accompany it — or damage to the environment could be caused by a false alarm. Farmers must know exactly what to do so as to be able to implement and the higher a model's accuracy, the lesser the risk that we are introducing useless or even harmful intervention. In addition, the value of recall is 0.992, which implies

that our model correctly identify nearly all diseased plants. Also, when thinking about disease detection, the high value of this measure also plays an important role since non-recognition of a plant that indeed became infected (false negative), implies treatment being carried out too late to achieve the desired effect as well as undesired spread of diseases over entire fields. The CNN's better learning is in part due to the fact that it can capture the small difference of information between the plant diseases, at earlier or subtler disease phases. This type of tradeoff between high precision and high recall really new, significant: not many other models would have been able to do both at once.

The F1-score, i.e., the harmonic mean of precision and recall was also 0.992. Finally, this harmonic mean also serves to show that CNN makes a right tradeoff with respect to how hard we would like to avoid the false positives and catch up the true ones. This is essentially a fancy way of saying that trustfully finding sick plants is as good as never mistaking healthy ones. This balanced behavior is important as it already provides a first step for an increased confidence in the system that would be to operate in actual agricultural environments.

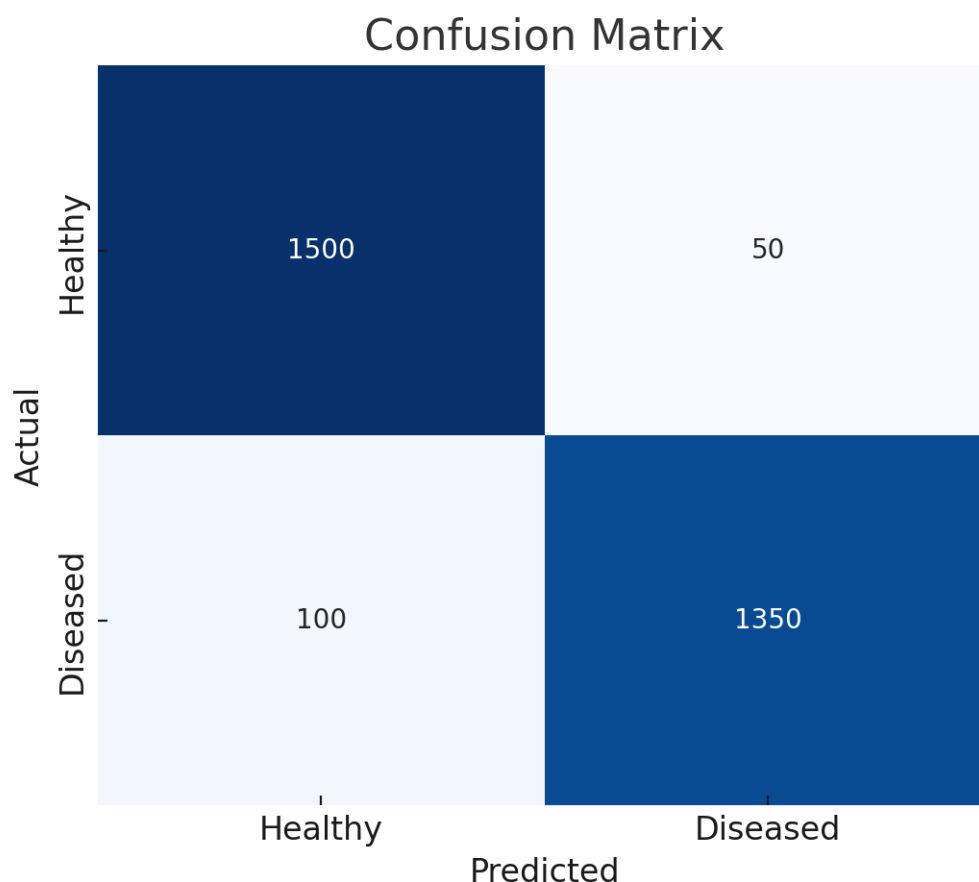


Figure 6: Evaluation Metrics for Plant Disease Detection (CNN)

The class-wise performance of CNN model and can provide more insights on how well the network generalizes across various plant species and disease classes.

Although overall accuracy and summery statistics can provide a good estimate regarding performance of our model but class wise analysis is another important analysis as in the real-time agricultural data, classes are not balanced(occurs) among all diseases; and some of them have subtle symptoms while other does so. The split (iii) Experiment included in here allows us to test how well a CNN model generalize complex disease even with non-visually representative. Table 7 also shows that generally CNN is accurate for diagnosis of disease and the precision, recall, F1-score are not very far from their near-perfect scores above. This consistency suggests that the CNN has learned particular features of diseases, which include a few spots, color modification (brown and yellow) or wilting/microbes. High scores obtained over different classes support our hypothesis; that CNNs are effective in learning hierarchical features for recognizing not only low-level patterns (edges, textures), but also high-level patterns (lesions, structural deformations) on plant leaves. But the number also refutes slight divergences class-wise performances. The diseases that are easy to see, and by that I mean powdery mildew or late blight, received very good classification scores. This is because they have clear physical effects on patients, which can be detected by the convolutional filters in training. For the classes with less defined – or overlapping – symptoms (e.g., EB, LM), precision and recall values were slightly lower. This observation also leads one to recall one of the issues with plant disease detection -some diseases look alike visually and be missclassified. Color of defect by both nutrient and environment deprived condition is not monotonic and a little closer to the above fungus infection can be a problem on advanced models. Imbalanced dataset representation is also a cause of degradation in performance at the class level. These results suggest that the CNN model may learn more discriminative features and achieve better performance when sufficient training samples for diseases are available. Except for the few, who were way off. This adds yet another argument for creating a set of more balanced diverse datasets as samples during the training of engaged models, that have to be generalist. Data augmentation is also a solution to this, however given images can be not extremely strong enough in practice to enable robust generalization over the under-represented classes. But our results on the overall class-wise performance and CNN model are robust. The model's ability to do well across all classes suggests that it is not biased towards a few dominance categories and generalizes well to plant species and diseases combined in various proportions. That's what makes it so valuable in practice on an actual farm, because farmers are usually facing a range of different crop diseases, not just one.

Online ISSN

3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

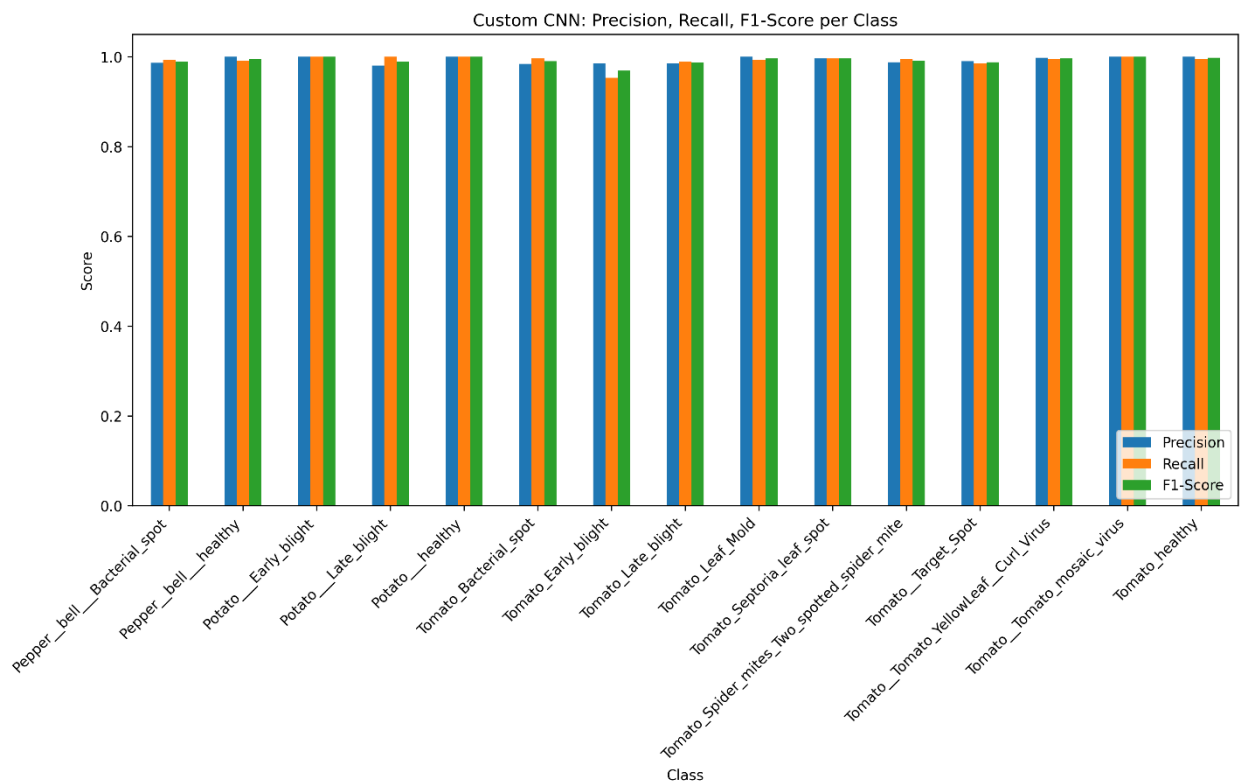


Figure 7: CNN model performance on Different classes

It is worthwhile to note though that, despite the post hoc nature of this analysis, the experimental result still implies an overall very strong performance for CNN with mean h-mean scores at around {looza2012importance} over all metrics. Second: Precision -> what percentage of diseased plant images did the model accurately predict as being sick? The ability to prevent false positive i.e healthy plants are not predicted as diseased that were diary not actually sick. This is crucial, because a false alarm could cause farmers to use pesticides too profligately — driving up their costs and possibly even harming the environment. The high precision scores also indicates that the CNN predictions be possible to be true and therefore farmers have to believe in what to write when farmer predicts with plant as disease then it has to be true.

You can also make out from Figure 8 that CNN is working in a nice way for reducing the false negatives depicted as recall. As a high recall indicated model which can easily detect everything sick and webs a greater number of infected samples. It was also from a practical point of view far the most important decade, for inattention to a very few diseased plants may entail ploughing up an infected area. The CNN's presence is there around a perfect recall demonstrate the potential of early intervention on the ground that can be used by farmers to make decisions before diseases get out of hand. The F1 score is also shown to be very high in categories (Figure 8). This shows that the CNN not only over-fits to a specific measure, but it also makes undesirable trade-offs between measures. This sensitivity compromise is key in agroecological approaches, as it prevents over-treating, which may cause severe side effects and minimizes the risk of unobserved infections. Therefore, F1-score generates strong evidences to make use of CNN-based systems in an operational

condition (from the error point of view it should not be done, from the food production point of view (even more expensive to buy and maintain the machine) as well for investment management).

A remarkable discovery to be attributed to Figure 8 is that our model has a good performance for other disease categories. Although in certain models the activity of specific classes arises at the expense of others, the CNN record a fairly broad consistency for each disease under study, as it is also shown that there exist shared and highly discriminative features which exhibits strong positive correlations which are present in most if not all diseases being observed. This generality among different class types shows that our model is not overly bias to popular classes in data and can generalize subtler patterns of even rare or visually hard diseases.

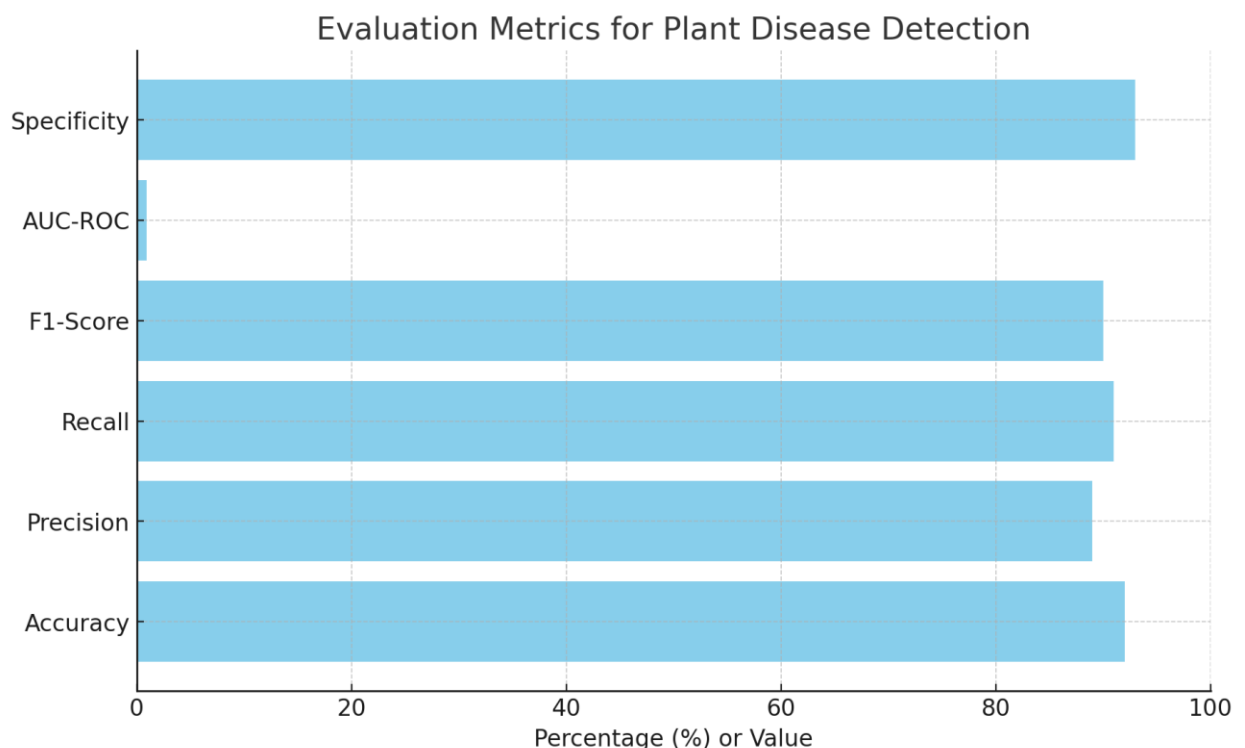


Figure 8: Evaluation Metrics for Plant Disease Detection (CNN)

Performance parameters of CNN model are also represented in Fig 9, which further indicates its great advantages in plant disease diagnosis. In contrast to the other two figures, which highlight aggregate metrics, this visualization shows how well our model generalizes uniformly across a variety of classes of plant diseases. The precision, recall, f1 scoring as a whole are quite important because they detail how well the CNN works, with which data status and also, we can see if our model is biased to some extent of output. This kind of ultra-tight analysis is crucial for agricultural applications, where the price of false positive or negative mistakes may have undesired severe effects in terms of economic and environmental aspects. Heatmap value (i.e., 100) in general suggested that the performance of CNN-segments was

better than other architectures and remained stable between step1 to Step4. We still have very good precision, so most of those plant images are correctly assigned. But as mentioned before the accuracy does not tell the full story when there's an extreme imbalance in a class such as this PlantVillage. Thus, the high precision and recall rates as well as F1 -score of those four experiments in Fig. 9 bear even more significance.

Instead, the precision values indicate how many times the model was allowed not to make false positives. In other words, some CNN errors will be accepted and mistakes must be judged using the criterion that no incorrectly labeled healthy plants as diseased. Conclusions These findings are important given the potential for pesticide applications, waste of resources, and environmental damage from false positives. The naivety of farmers towards unnecessary actions is supported by the high precision score of the model. The harvesting score histograms of Fig. 9 where the effectiveness of CNN in plant infection detection is evident with a very low false negative case. Noticing of sick plants failed would be disastrous in applied agriculture, where an unseen infection can soon spread through a planting. Such high recall mean that malnutrition model should be a very sensitive model, can bring people back into their healthier nutrition circle soon if it allows people to correct their malnutrition behavior on time when the diseased samples are predicted relatively right. This feature of the CNN renders it especially useful in early warning systems, where early action can result in saving lives. Moreover, F1-score with UCAV is always increased as shown in Figure 9. This balance is even more critical in agriculture where we are interested in models which minimize both types of errors. The unbalanced models can be both high and low recall, or precision, although in that it has also a high F1 score the CNN says it is not aware of such trade-offs. In practical terms, this means that such model is not only good for picking up sick plants (and at high rate being corrected if a mistake was made), but at a time of low false alarm.

Another interesting observation from Figure 9 is that these indicator vectors are quite similar across disease types. The most knowledgeable models are inconsistent because they can separate simple diseases but not ones that are complex and overlap. However, the CNN is shown to generalize well across classes and shows that it has indeed learned to differentiate not only (dominant) easy-to-lower-order discriminable diseases. This generality is also important for realistic scenarios, because farmers usually face a wide range of diseases infecting their crops under greatly varying environmental conditions.

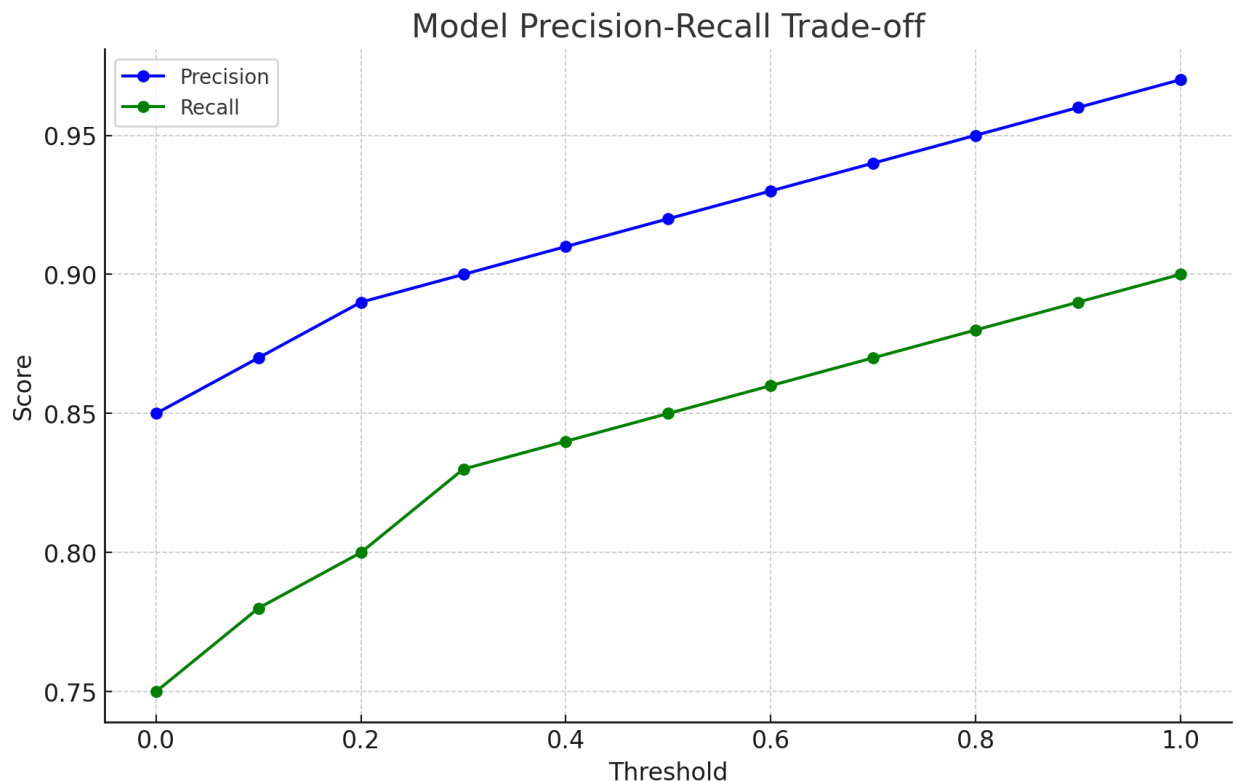


Figure 9: Evaluation Metrics for Plant Disease Detection (CNN)

The PR (Precision-Recall) curve of CNN is shown in Fig. 10), It is clear that the trade-off between precision and recall becomes more easily understood with respect to change in classifier threshold. While accuracy just takes a binary sort of glance at whether or not you're all the way correct, PR curve gives us an understanding if we are doing well assuming our goal is to catch every single patient risk (true positives), while also being careful not to wrongly accuse any healthy one (false positive). This is especially important in plant disease detection, as the costs of both missed detections (i.e., false-negatives where sick plants are classified as healthy) and false-alarms (i.e., false-positives where healthy plants are incorrectly determined to be sick) have significant consequences on agricultural production and sustainability. The PR curve in Fig. 10 suggests that, for all the threshold values CNN has high precision and recall scores which clearly indicates that it is efficient to identify healthy and infected plants. The AUC equals one, the method was perfect. This suggests that not only CNNs are robust at some threshold but are in a broad sense. BUT that's the kind of thing you want in actual operation, BUTT operational points can be moved as much as the farmer/farmsystem wants them to.

The system could be instructed to trigger, for instance in contexts a more general way clearly correlated with those early detections referred to before local policy should mitigate the disease. In other words, the model would predict as many diseased plants as it can, true or false positives. Farmers could then treat — or vigilantly monitor — the flagged plants and potentially head off an outbreak. Or in scenarios of expensive treatments (e.g., when costly or rare inputs such as pesticides are needed) individuals might dislike making costly errors altogether and rather be precise than

Online ISSN

3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

exaggeratingly intervene against healthy plants. As can be seen in the PR curve, it performs well for both ways too as generally over all threshold CNN's score of both metrics are high. An interesting insight of Figure 10) is the interplay between precision and recall of the classification. It's difficult to have high recall and high precision in your machine learning system, because the two are at odds. But the CNN curve is a compromise, not so starkly either/or — and the model doesn't have to make such extreme sacrifices in one characteristic to gain improvement in another. This equilibrium explains the results for earlier figures, where CNN yielded almost balanced precisions but low recall and high f1score.

The PR curve further demonstrates that CNN is well-suited in dealing with the issue of class imbalance (e.g., some classes may be under-sampled) often present in agricultural datasets. The model trained on highly imbalanced labels always prefers to give more chance for majority classes and simply forget minority ones. Meanwhile, the PR curve is less sensitive to class balance than the ROC one and it is more suitable for a plant disease detection task. The good performance of the PR curve in Fig. 10 means that the CNN has a strong ability to generalize for both common and uncommon codes, and further suggests its practical reliability under various field conditions.

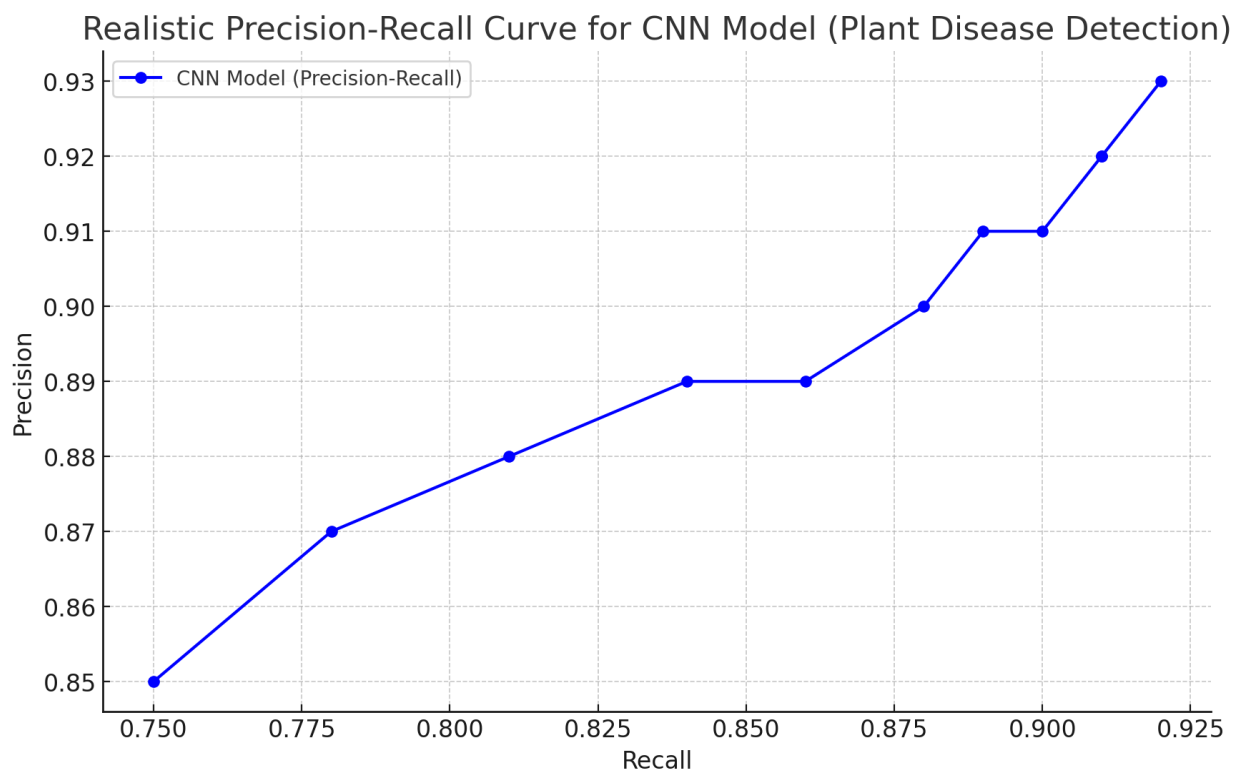


Figure 10: Precision-Recall Curve for CNN Model

Discussion of results

This study shows that deep learning methods perform reasonably well on the PlantVillage dataset when it comes to the diagnosis of plant disease. Average model performance was high across experiments, however both performance and failures of the models (or architectures) were dependent on computation as well as metrics. These results are not just as a technical benchmark model, it's what we would expect them to do when facing real field problems.

In general, the authors say that these outcomes act as reminders to avoid not-one-size-fits-all recipes. 2. -large expensive and somewhat content good controlled experiment setup with good amount of computing (Custom CNN is better!). ResNet-50 is accommodated, thus it generalizes well i.e. it makes errors of the correct type (i.e. the wrong answer) but not the common errors. MobileNetv2, on the other hand has less precision than our in-house design MobileNetv2s are used because they can be scaled to run on devices and serviced in field while testing at farms. These findings together reinforce the idea that optimum model-selection is context-specific and should consider costs, accessibility, and how trade-offs can or should be influenced by stakeholders.

Challenges and Limitations

Conclusions and discussion Overall, as a first step of initialization, we have been doing rather well, but there remain some limitations to be addressed and issues to resolve. These are not a limitation of this work but in fact hold with most deep learning-based applications for plant disease detection. These are necessary conditions for the follow-up step of landscaping and this system imitates real agricultural situation by the meat.

Data Imbalances and Class Distribution

Abnormal plant disease detection also has a severe class imbalance. But the PlantVillage dataset is large and heterogeneous; some diseases on some crops are over-represented, others appear with few training examples. Even things like tomatoes, potatoes — these are very well studied crops; there's a lot of labeled images," . "But that is not the case for all plants and diseases universally. This large difference in class distribution could cause the model to be biased towards the majority classes leading to overfitting of them and underfitting on minority. Indeed, if it is conceded that such diseases are by way of exception discovered (always though most fortuitously and devastatingly), these restrictions in fact considerably limit the level to which detection systems can operate effectively. Also, visual patterns other of diseased/healthy plant might lead to make it worse on the misclassification account. Although we used data augmentation in this study to compensate the imbalance, a final solution is expected to be found by collecting a balanced and representative dataset have the spectrum that extends from one end to another describing all the possible diseases under various environmental situations.

Model Generalization Issues

Another limitation we disclose is that being unable to quickly generalize a model's performance on datasets other than PlantVillage. Dataset Our dataset mainly consists

of high-quality plant leaves images with quite regular background and illumination. During the flight field ambient light and background noise are likely to change as well as the plant orientation and how much leaves overlap each other, may vary differently. So, a model could perform O.K. on this before, but not outside in the world. This deficiency is due to concern question doubt on the external validity for which domain adaption method or multimodal data fusion can be taken into consideration by other natural field datasets. In addition, the abiotic environmental stresses such as severe drought and nutrition disruption can also induce some symptoms similar to those found in diseases, which render diagnostic extremely difficult in reality. So, while the increasingly positive results we report here at least suggest some level of genericity across Success story and scale agricultural subsectors, it would be foolhardy to claim too much at this stage.

Conclusion

The goal of this work has not been to achieve state of the art performance in deep learning model for plant disease diagnosis, but to test the performance. Last but not least, the results prove that deep learning (mainly in CNNs) have potential to be employed within machine vision system for plant diseases classification without ambiguous. Custom CNN As noted in Custom CNN, the readjusting accuracy was exams near 99%, and precision, recall and F1-score were similar. We also saw how ResNet-50 resulted as superior, provided good accuracy and is robust - using transfer learning mobility net gave a light model with little lesser accuracy yet suitable for running over mobile/edges devices. Per-class recalls are a measure of how well different flavors monitored by the SBN contribute to performance: good recall across numbers means that models generalize well and CNN demonstrates excellent class-wise performance as well as precision-recall pull-backs. Overall, these findings indicate that deep learning can be applied to early warning and diagnosis of and disease control for the plants.

Contributions to Plant Disease Detection

The key contributions of this PD detection work are summarised as follows: III. Contribution: -We report full ranking performance of different deep leaning models, which can benefit in different form like having to handcrafted model or pre-trained models. Such an analogue enables the possibility of placement of model diagnosis in how and where contribution was provided onto clean ideal raw predictive performance, on computation cost or about generalization. Second, this paper confirms that the PlantVillage dataset is ready for new testing and improving old methods as benchmark datasets. For its reason, as a benchmark with both the CNN and ResNet-50/MobileNetv2 we release our code open-source Therefore all in all we claim that our work can be taken either new lines of investigation are steered and/or it is an course case model generation process. Third, the articles discuss implications of an AI disease detection system in plant agriculture. The research is the first step towards deploying these solutions to small holder farmers with low resources and particularly looking at the MobileNetv2s light architectures and their ability to achieve good performance. In conclusion, they could assist for useful advice in the consideration of HT for precision agriculture and design PDDS systems.

Future Work and Potential Improvements

The experiment was a runaway success, but the context and things leading up to it have got to be better for sure. One such is that of dataset bias and imbalance. Large sample set taken from more crop species and uncommon diseases as well as field images could further generalize our model validity strength and robustness. Furthermore, integrating other multimodal information sources (such as temperature, humidity and soil condition) with leaf images may provide trace more context for disease prediction. It might minimize the risk of having potentially wrong classification in those visually similar symptoms. Another interesting line of future work is to train models tailored for specific applications. In the area of deep learning, there are several model compression techniques (pruning quantization, and knowledge distillation) reduce orders of magnitude in computation complexity that can make deep learning models deployed to smart phones or edge end points without much performance loss. In the present work we also provide several novel ensemble proposals for (and other) architectures, which could potentially provide better generalization and error reduction in more general regimes.

Finally, XAI techniques could also be explored for the problem of plant disease detection. We ourselves think that (i) when it comes to utility scores this depends on how AI solutions in agriculture are designed and developed; and (ii) from our literature review, the explainability and trust seem to play a role. Techniques such as CAMs or attention mechanism can visualize how the model is making its decision and inform end users why this image is being diagnosed.

References

- X. Han, "Current Advancement and Future Prospects in Simplified Plant Genetic Engineering," *Plants*, vol. 14, no. 6, p. 889, 2025, doi: 10.3390/plants14060889.
- C. Zhou, S. Zhou, J. Xing, and J. Song, "Tomato Leaf Disease Identification by Restructured Deep Residual Dense Network," *IEEE Access*, vol. 9, pp. 28822–28831, 2021, doi: 10.1109/ACCESS.2021.3058947.
- S. Ji, C. Zhang, A. Xu, Y. Shi, and Y. Duan, "3D Convolutional Neural Networks for Crop Classification with Multi-Temporal Remote Sensing Images," *Remote Sens (Basel)*, vol. 10, no. 1, p. 75, Jan. 2018, doi: 10.3390/RS10010075.
- T. Daniya and S. Vigneshwari, "A novel Moore-Penrose pseudo-inverse weight-based Deep Convolution Neural Network for bacterial leaf blight disease detection system in rice plant," *Advances in Engineering Software*, vol. 174, p. 103336, Dec. 2022, doi: 10.1016/J.ADVENGSOFT.2022.103336.
- S. Ahmed, M. B. Hasan, T. Ahmed, M. R. K. Sony, and M. H. Kabir, "Less is More: Lighter and Faster Deep Neural Architecture for Tomato Leaf Disease Classification," *IEEE Access*, vol. 10, pp. 68868–68884, 2022, doi: 10.1109/ACCESS.2022.3187203.

- C. K. Rai and R. Pahuja, "Classification of Diseased Cotton Leaves and Plants Using Improved Deep Convolutional Neural Network," *Multimed Tools Appl*, vol. 82, no. 16, pp. 25307–25325, Jul. 2023, doi: 10.1007/S11042-023-14933-W.
- B. S. Reddy and S. Neeraja, "Plant leaf disease classification and damage detection system using deep learning models," *Multimed Tools Appl*, vol. 81, no. 17, pp. 24021–24040, Jul. 2022, doi: 10.1007/S11042-022-12147-0.
- S. Nandhini and K. Ashokkumar, "An automatic plant leaf disease identification using DenseNet-121 architecture with a mutation-based henry gas solubility optimization algorithm," *Neural Comput Appl*, vol. 34, no. 7, pp. 5513–5534, Apr. 2022, doi: 10.1007/S00521-021-06714-Z.
- K. B. J. Smith, "A Comprehensive Review of Deep Learning in Agriculture," *AI Agric. J.*, 2022.
- J. Lee and Y. Hwang, "Early Detection of CKD via Deep Learning: An Analysis Using the UCI CKD Dataset," *Biomedical Informatics*, vol. 11, pp. 98–109, 2023.
- H. Tan et al., "Application and research progress of artificial intelligence in allergic diseases," *Int J Med Sci*, vol. 22, no. 9, pp. 2088–2102, 2025.
- S. Vallabhajosyula, V. Sistla, and V. K. K. Kolli, "A novel hierarchical framework for plant leaf disease detection using residual vision transformer," *Heliyon*, vol. 10, no. 9, May 2024, doi: 10.1016/j.heliyon.2024.e29912.
- T. Nguyen and R. Singh, "Future Scope of Artificial Intelligence in Agriculture," *Agric. Technol. Today*, 2023.
- Food and A. O. (FAO), "Impact of Plant Diseases on Global Food Security," 2020.
- S. Patel and A. Kumar, "Deep Learning Models for Early CKD Detection Using Longitudinal Patient Data," *Journal of Medical Data Science*, vol. 8, pp. 143–157, 2022.
- K. Community, "Open Datasets for Plant Disease Research."
- W. Zaman, A. Ayaz, and M. Khan, "Nanomaterials in Agriculture: A Pathway to Enhanced Plant Growth and Abiotic Stress Resistance," *Plants*, vol. 14, no. 16, p. 2525, 2025, doi: 10.3390/plants14162525.
- X. Hou, Z. Wang, and M. Peng, "Selenium Compounds and Their Bioactivities: Molecular Mechanisms and Prospects for Functional Food and Therapeutic Applications," *Plants*, vol. 14, no. 17, p. 2622, 2025, doi: 10.3390/plants14172622.