# IMAGE-BASED PLANT DISEASE DETECTION USING DEEP LEARNING

**M. Gandhimathi**

Assistant Professor and Head, Department of BCA, G. Venakataswamy Naidu College (Autonomous), Kovilpatti

## ABSTRACT

Plant diseases pose a serious threat to food security, but the lack of basic infrastructure makes rapid diagnosis difficult in many sectors of industry. The combination of the worldwide popularity of mobile phones and recent advances in the idea of laptop computers enabled by extensive research has paved the way for mobile phone interference analysis. Using a public data set of 5,306 images of diseased and healthy plant leaves collected in a controlled environment, we developed a deep convolution neural network to recognize one plant species and 26 diseases. Have trained (or not recognized). The trained version achieves an accuracy of 99.35% of the extended control series demonstrating the feasibility of this technique. An important method for training deep learning models on increasingly large and publicly available photographic datasets is a direct approach to mobile phone-based analysis of plant diseases on a global scale.

**Keywords**: food security, rapid diagnosis, convolution, neural network, and datasets.

## Introduction

With the aid of modern technology, human society has been able to produce enough food to feed over 7 billion people. However, there are still a number of dangers to food security, including changing weather patterns (Tai et al., 2001), Odd, a plant disease, and the decline of pollinators (2016 Intergovernmental Technology and Policy Forum Full Report on Biodiversity Environment and Services Fourth Session Proceedings; Scott, 2005, among others). Even though plant diseases aren't the biggest threat to food security anywhere in the world, they can be devastating for smallholder farmers whose ability to produce healthy crops is essential to their survival. According to the UNEP (2013), smallholder farmers produce more than 80% of the rural goods in the expanding international region. It is also estimated that pests. In recent years, mobile phone-based devices have become commonplace, taking advantage of the unprecedented rapid spread of the mobile age across all sectors of industry (ITU,

2015). With high computing power, high performance, and a wealth of integrated accessories such as state-of-the-art HD cameras, smartphones offer state-of-the-art technology for disease detection. According to many forecasts, the number of smartphones in the world could increase from 5 billion to 6 billion by 2020. Towards the end of 2015, 69% of the business population was previously close to portable broadband protection, and 7% had access to multipurpose broadband in 2015, 12 times higher than in 2007 (ITU, 2015). A Image I combination of factors such as the proliferation of smartphones, high-definition cameras, and mobile devices with powerful processors will make fully computerized image-based disease analysis widely available, if technically feasible. There is a possibility. Using his 5,306 images of 26 sick (or healthy) plant species from the Plant Village project to demonstrate the technical feasibility of Deep His learning strategy. An example of each plant disease pair is shown in Image I.

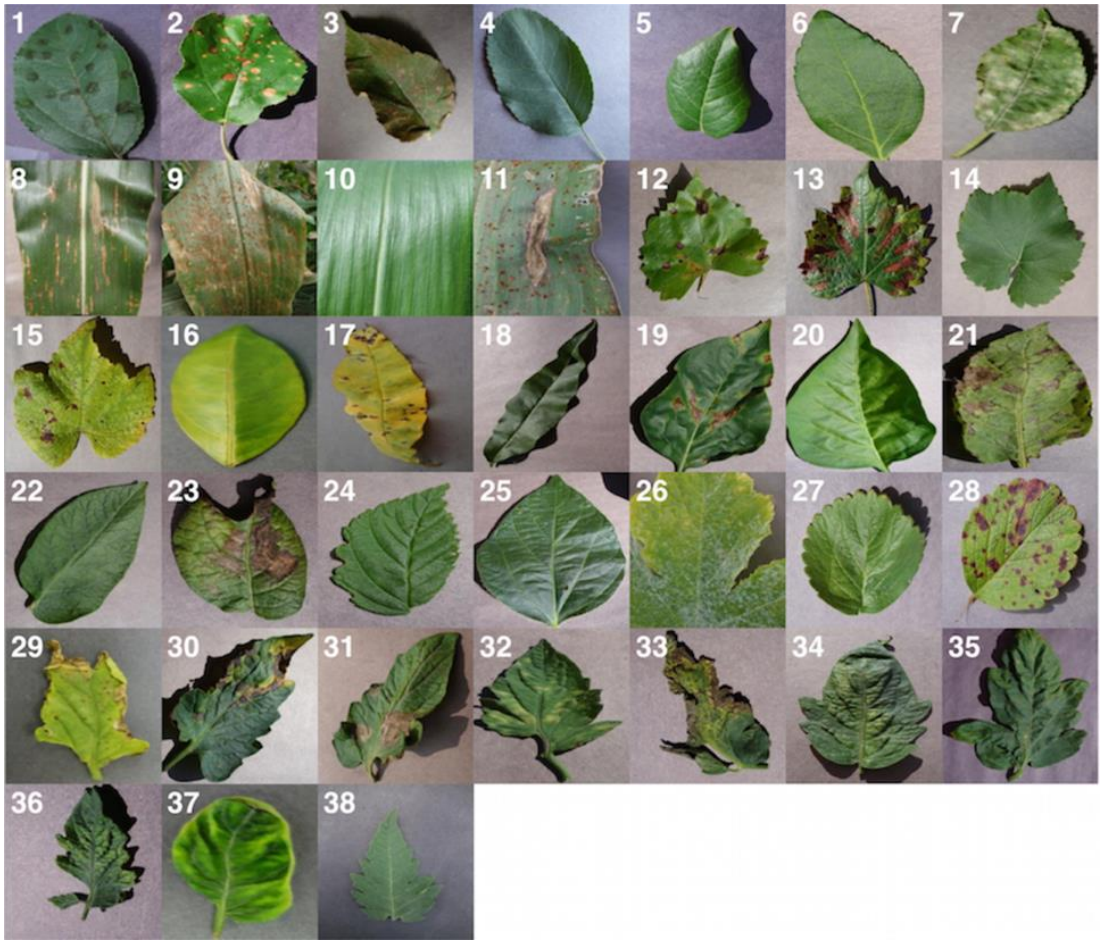



IMAGE 1. A Plant Village Dataset Example of Leaf Images That Represent Every Crop-Disease Pair Used.

In just a few years, computer ideas and object recognition in particular have come a long way. His PASCAL VOC task) and the larger and more comprehensive Large Image Credit Task (ILSVRC). As a reference point for many image-related problems in object classification and computer vision. In 2012, the vast and deep convolutional local regions of the brain showed a 16.% error in category snapshots for every 1000 categories used (Krizhevsky et al., 2012). His error rate decreased to 3.57% over the next three years thanks to numerous features in deep convolution neural networks Training large neural networks can be time-consuming, but skill mode can quickly classify images, so a smartphone client is also a good fit for that package. Deep brain networks are now used in various fields as examples of limited learning. For example, combine an image of a diseased plant with some plant disease input into the output of a neural network. Mathematical operations called knots in neural networks take numbers from edges and transform them into numbers in the output domain. A deep neural network connects input and output layers through a chain of stacked layers of nodes. Both network shape and features (nodes) and edge weights map inputs to outputs well, and either of these methods is necessary to build deep communities. Deep neural networks can set community parameters and thus improve mapping when incorporated into educational systems. This method is computationally difficult and has recently undergone significant progress thanks to a series of technical and conceptual revolutions (LeCun et al., 2015; 2015 (Schmidhuber)) including both healthy and diseased images. , required a huge, well-established dataset. We extend plants with correct image classification of their diagnostic features for plant diseases. Using 5,306 images and a convolutional neural community approach, this study recorded 26 disease classes for one crop species. bottom. The model's ability to identify ideal crop-disease combinations that provide 38 potential cues is used to assess their effectiveness. typical accuracy of ) indicates the technical feasibility of our strategy. Our findings represent the first step towards a smartphone-enabled system failure analysis device.

## Methods

### Dataset Description

5,306 photos of plant leaves were classified using 38 different tools. Based on photographs of plant leaves, our objective is to forecast plant disease pairs for each plant disease pair classifier. Images for each pair of plant diseases in the Plant City dataset are shown in Figure 1. All the methods presented in this article resize the images to 256 x 256 pixels and perform both model correction and prediction on these reduced

images. All experiments used three different versions of the entire Plant Village dataset. Let's start with the colour version of the Plant Village dataset. The Plant Village dataset is then tested in grayscale format, and finally, all tests are run on a fragmented version of the Plant Town dataset. In this way, we remove all irrelevant background information that may cause characteristic biases in the dataset. Traditional data collection for the Plant Village dataset. Segmentation was automated using a script configured to work well with a particular dataset. We chose a technique based on a series of masks created by analyzing the colour, brightness, and saturation components of different parts of the image in multiple colour spaces (Lab and HSB). One of the processing steps also facilitated the correction of very strong colour casts in some subsets of the dataset, eliminating another potential bias. The purpose of this series of experiments is to understand whether the neural network is actually learning a plant disease "concept" or just becoming aware of the dataset's biases. Image II displays various editions of the same magazine from a selection of publications.

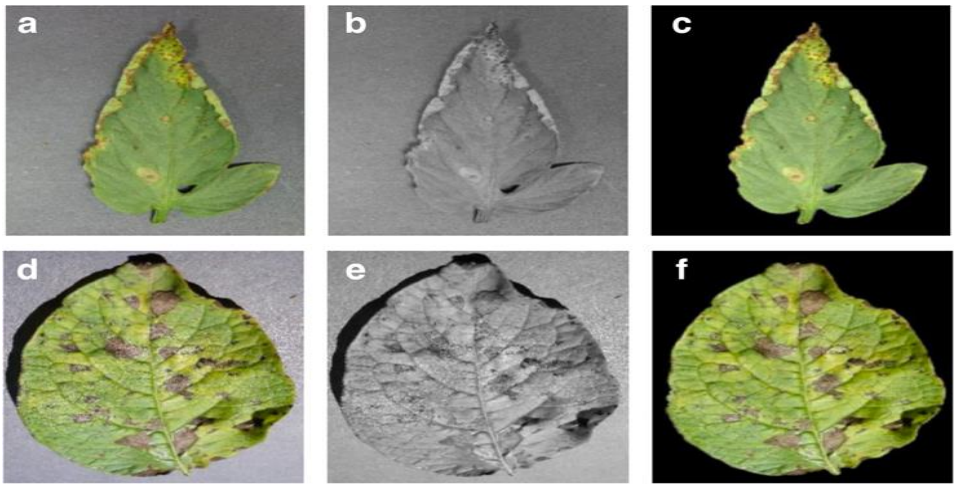



IMAGE II. Plant Village Dataset Example Images From The Three Different Versions Used In Various Experimental Configurations.

**Measurement of Performance**

Run all trials on various pieces of the 80-20 tensile test set to determine whether approaches are overfitting and how well they work with fresh, unheard-of data. 60-0 (60% of the data set is used for training and 0% is used for testing), 50-50 (50% of the data set is used for education and 50% is used for sign-out), 0-60 (percentage of zero data set used for training, 60 percent used for unenrollment), 20-80 (20 percent of data set used for training, 80 percent of data set used for testing) will be used). erase). Note that the Plant Village dataset often contains multiple

images of the same leaf taken in a single orientation, and we mapped these instances from his 5,306 photographs to 1,112 pixels. It is important to Also, many of these control trains splits ensure that all images for one magazine are sent to either check-in or training. Similarly, calculate the recommended accuracy for each test. This is half the average accuracy of the F1 classifier over the entire training period at frequent intervals (at the expense of each epoch). Use the final mean F1 score to estimate the impact of all exclusive experimental settings.

**Approach:**

Deep convolution neural networks' suitability for the aforementioned kinds of issues is contrasted. We focus on his two general models, specifically Alex Net and Google Net 2015 developed by the Large Visual Popularity Mission for the Image Net dataset 2009. The structure of Alex Net (see figure. Model of S2) The model is the same as the LeNet-Five model of the 1990s structure (LeCun et al., 1989). This layer is followed by a fixed stacked convolutional layer, in an architectural variation of LeNet-5.All community layers typically have a ReLu nonlinear actuator, and the convolutional layer precedes the actual aggregation layer. can optionally include an additional normalization layer at the end. Both the normalization and convolutional layers of the first two convolutional layers (conv1,2) are connected, followed by the convolutional layer of the last convolutional layer (conv5). The final fully connected layer (fc8) in our unique implementation of Alex Net, which serves as the input to the SoftMax layer, has 38 outputs, which equals the number of instructions in the dataset. The softmax layer exponentially normalizes the input (fc8) into a distribution of 38 value classes, each forming at least one. These values can be used to determine the network's confidence that the input image is correctly represented by the proper instructions. The ReLu nonlinear activation unit is connected to all 7 layers of the Alex network, and two identical layers (fc6,7) are connected to his final layer with a final ratio of 0.5. Although the Google Net structure has 22 layers and is instead 1 mile deep and wide, it still has a much smaller range of parameters in the network (5 million) than the Alex Net structure with 60 million parameters.

An essential component of Google Net's structure is the utility of the arrangement known as the "network of communities" (Lin et al., 2013) in terms of output modules. The seed module uses 1, 3, and 5 convolutions in parallel with maximum pooling layers to capture improvements simultaneously. Keeping in mind the practicality of the application and wanting to limit the amount of computation required, the 1:1 convolution is added

before (and after) the 3:3, and 5:5 circuits described above. above the max pooling layer) to reduce the dimensionality. The combined layer of the filter finally combines most of the outputs from the parallel layers. This org is one launcher, but the underlying Google Net model used in testing uses nine launchers. For a given overview of this structure, see reference (Szegedy et al., 2015). By first training, a new version on a single instance and then using a transfer study to modify the already trained model (mastering the ImageNet dataset), both architectures are fully integrated into the Plant Village dataset. Compare how well they work across the board. We reset the layer fc8 weights in Alex Net for the exchange study and reset the loss weights in Google Net for layers 1, 2, and the classifier. Second, contrary to what sometimes happens with knowledge acquisition, the training version does not limit the level of knowledge acquisition. The initial weight values of some layers that can already exchange a large amount of visual information are the main differences in knowledge acquisition between these two techniques (exchange and training from scratch). Acquisition of techniques illustrated by pretrained Alex Net and Google Net models derived from ImageNet (Russakovsky et al., 2015). than the Alex-Net structure with 5-60 million

We have 60 different experimental sets in total, with the following variables.:

1. Deep Choice Learning Architecture: Alex Net, Google Net.
2. Priority of educational mechanism: Transfer of studies, Education from scratch.
3. Database Type Request: Color, Grayscale, Segmented Sheets.

Distribution of training test series: Teaching: 80%, see: 20%, Practice: 60%, see: forty%, Teach: 50%, take a Look at 50%, Educate: forty%, test: 60%, Teach: 20%, check: 80%.

The usefulness of the structure known as the 'network of communities' (in terms of output modules is an important aspect of the structure of Google Net. Keeping in mind the practicality of the application and wanting to limit the amount of computation required, the 1:1 convolution is added before (and after) the 3:3, and 5:5 circuits described above. above the max pooling layer) to reduce the dimensionality. The combined layer of the filter finally combines most of the outputs from the parallel layers. This org is one launcher, but the underlying Google Net model used in testing uses nine launchers. For a given overview of this structure, see reference (Szegedy et al., 2015). By first training a new version on a single instance and then using a transfer study to modify the already trained model (mastering the

ImageNet dataset), both architectures are fully integrated into the Plant Village dataset. Compare how well they work across the board. We reset the layer fc8 weights in Alex Net for the exchange study and reset the loss weights in Google Net for layers 1, 2, and the classifier. Second, contrary to what sometimes happens with knowledge acquisition, the training version does not limit the level of knowledge acquisition. The initial weight values of some layers that can already exchange a large amount of visual information are the main differences in knowledge acquisition between these two techniques (exchange and training from scratch). Acquisition of techniques illustrated by pre-trained Alex Net and Google Net models derived from ImageNet (Russakovsky et al., 2015) than the Alex-Net structure with 5-60 million(Observation 3).

Figure 3

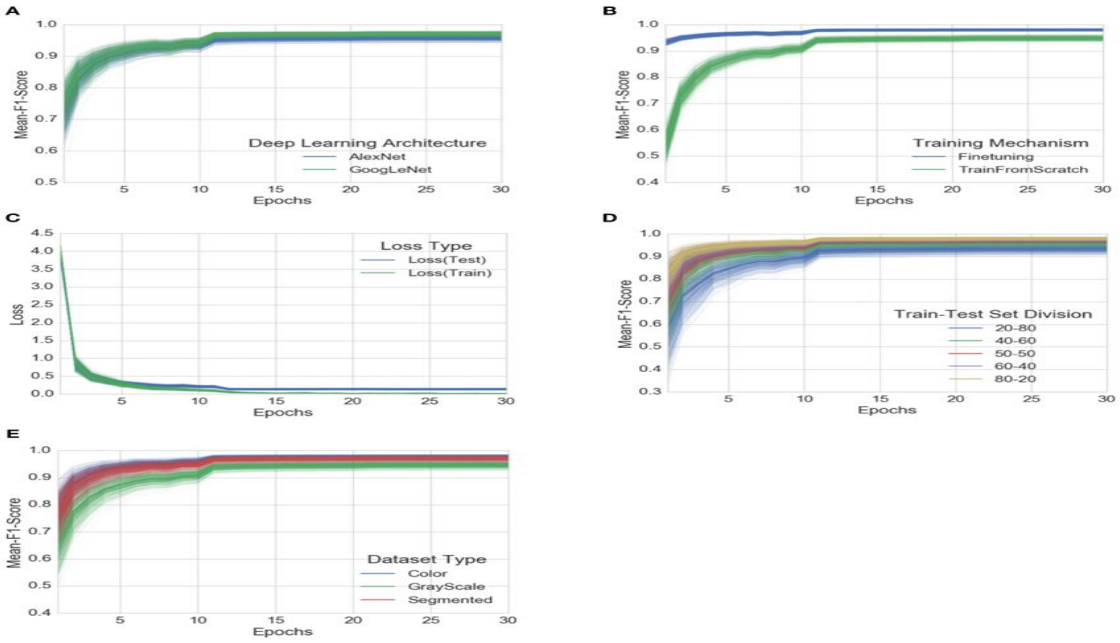



FIGURE 3. A Grouping by Experimental Configuration Parameters Mean F1 Score and Loss Progression over the 30 Epochs of Training across All Experiments.

In every experiment with the given configuration, the uncertainty of a given quantity is inversely proportional to the intensity of that quantity at any given point. A) comparison of the average F1 score development in all studies, grouped by training mechanisms; B) evaluation of training progress; and C) evaluation of the average F1 score development in all studies. (D) contrast of planned F1 classification succession in all trials, grouped by training–control set distributions; and (E) contrast of planned F1 classification progression in all trials, grouped by data type. In the supplementary material, it is possible to determine a plot of

all observations that are as similar as possible across all experimental setups. We also tried to standardize the hyperparameters for all experiments, and we used the following hyperparameters for all experiments to allow for a fair-haired comparison of the effects of all investigational settings:

- Type of solver: Stochastic Gradient Descent,
- The reward for the base study:0.0.5,
- Coverage for Learning Costs: Step (diminish ten times every thirty cycles or three), Momentum: zero. 9. Weightloss: Gamma: null.0005
- Stack Length: 0, 100 (for Alex Net) 1, 2 (for Google Net).

All the above experiments are used for their own fast and free deep learning framework. We can also reproduce key results, including overall accuracy, using well-known coffees. A selection of 30 epochs (as shown in the summary figure of all experiments).

Results: For all tests with the specified configuration, the elegance depth selected for each factor is proportional to the related uncertainty. (A) Evaluation of F1 rating evolution across all trials, clustered using the deep analysis structure, (B) Evaluation of F1 rating evolution across all trials, combined with the training mechanism(C) Assessment of test loss and learning loss progression at a certain point across all trials, (E) Progression estimate advises the F1 scores of all experiments grouped by dataset type usage. (D) succession estimate advises the F1 scores of all experiments grouped by train. In the supplementary tissue, a similar curve can be determined for all observations in all experimental settings, as far as it goes. To allow a fair comparison of results across experimental setups, we also attempted to standardize the hyperparameters for all experiments and used the following hyperparameters for all experiments:

- Solver Type: Stochastic Gradient Descent,
- Base Training Fee: Zero.0.5,
- Familiarity Fee Insurance: Step (sort of decreases by 10x every 30/three episodes),
- Momentum: Zero.9Lightweight:0.0005,
- Range:0.1,
- BatchLength (for Google Net), 100 (for Alex Net).

He used his branch of Caffe (Jia et al., 2001), a quick and open deep learning framework, to carry out all of the aforementioned tests. Taking regular coffee as an example, you can also understand the basic effect and general accuracy. Pathway to mobile phone-based crop disease analysis on a global scale.

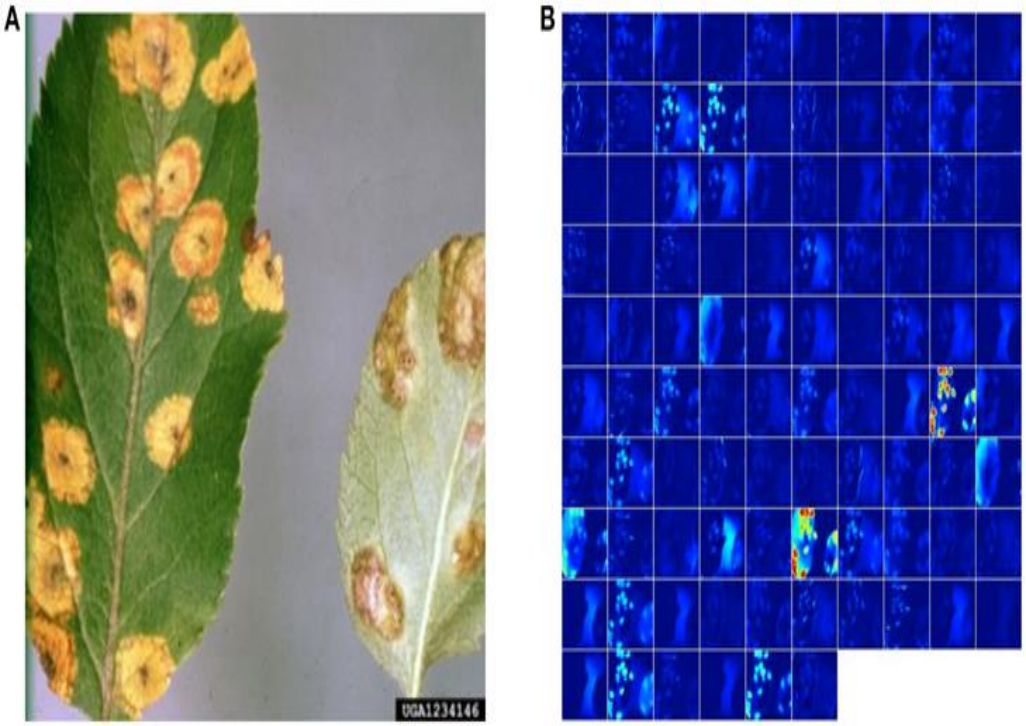
| | AlexNet | | GoogleNet | |
|---|---|---|---|---|
| | Transfer learning | Training from scratch | Transfer learning | Training from scratch |
| **TRAIN: 200%, TEST: 80%** | | | | |
| Color | 0.9736 (0.9742, 0.9737, 0.9738) | 0.9118 (0.9137, 0.9132, 0.9130) | **0.9820** (0.9824, 0.9821, 0.9821) | 0.9430 (0.9440, 0.9431, 0.9429) |
| Grayscale | 0.9361 (0.9368, 0.9369, 0.9371) | 0.8524 (0.8539, 0.8555, 0.8553) | **0.9563** (0.9570, 0.9564, 0.9564) | 0.8828 (0.8842, 0.6835, 0.8841) |
| Segmented | 0.9724 (0.9727, 0.9727, 0.9726) | 0.8945 (0.8956, 0.8963, 0.8969) | **0.9808** (0.9810, 0.9808, 0.9808) | 0.9377 (0.9388, 0.9380, 0.9380) |
| **TRAIN: 400%, TEST: 60%** | | | | |
| Color | 0.9860 (0.9861, 0.9861, 0.9860) | 0.9555 (0.9557, 0.9558, 0.9558) | **0.9914** (0.9914, 0.9914, 0.9914) | 0.9729 (0.9731, 0.9729, 0.9729) |
| Grayscale | 0.9584 (0.9588, 0.9589, 0.9588) | 0.9088 (0.9090, 0.9101, 0.9100) | **0.9714** (0.9717, 0.9716, 0.9716) | 0.9361 (0.9364, 0.9363, 0.9364) |
| Segmented | 0.9812 (0.9814, 0.9813, 0.9813) | 0.9404 (0.9409, 0.9408, 0.9408) | **0.9896** (0.9896, 0.9896, 0.9898) | 0.9643 (0.9647, 0.9642, 0.9642) |
| **TRAIN: 50%, TEST: 50%** | | | | |
| Color | 0.9896 (0.9897, 0.9896, 0.9897) | 0.9644 (0.9647, 0.9647, 0.9647) | **0.9916** (0.9916, 0.9916, 0.9916) | 0.9772 (0.9774, 0.9773, 0.9773) |
| Grayscale | 0.9661 (0.9663, 0.9663, 0.9663) | 0.9312 (0.9315, 0.9318, 0.9319) | **0.9788** (0.9789, 0.9788, 0.9788) | 0.9507 (0.9510, 0.9507, 0.9509) |
| Segmented | 0.9867 (0.9868, 0.9868, 0.9869) | 0.9551 (0.9552, 0.9555, 0.9556) | **0.9909** (0.9910, 0.9910, 0.9910) | 0.9720 (0.9721, 0.9721, 0.9722) |
| **TRAIN: 600%, TEST: 40%** | | | | |
| Color | 0.9907 (0.9908, 0.9908, 0.9907) | 0.9724 (0.9725, 0.9725, 0.9725) | **0.9924** (0.9924, 0.9924, 0.9924) | 0.9824 (0.9825, 0.9824, 0.9824) |
| Grayscale | 0.9686 (0.9689, 0.9688, 0.9688) | 0.9388 (0.9396, 0.9395, 0.9391) | **0.9785** (0.9789, 0.9786, 0.9787) | 0.9547 (0.9554, 0.9548, 0.9551) |
| Segmented | 0.9855 (0.9856, 0.9856, 0.9856) | 0.9595 (0.9597, 0.9597, 0.9596) | **0.9905** (0.9906, 0.9906, 0.9906) | 0.9740 (0.9743, 0.9740, 0.9745) |
| **TRAIN: 80%, TEST: 20%** | | | | |
| Color | **0.9927** (0.9928, 0.9927, 0.9928) | **0.9782** (0.9786, 0.9782, 0.9782) | **0.9934** (0.9935, 0.9935, 0.9935) | **0.9836** (0.9839, 0.9837, 0.9837) |
| Grayscale | **0.9726** (0.9728.0.9727, 0.9725) | **0.9449** (0.9451, 0.9454, 0.9452) | **0.9800** (0.9804, 0.9801, 0.9798) | **0.9621** (0.9624, 0.9621, 0.9621) |
| Segmented | **0.9891** (0.9893, 0.9891, 0.9892) | **0.9722** (0.9725, 0.9724, 0.9723) | **0.9925** (0.9925, 0.9925, 0.9924) | **0.9824** (0.9827, 0.9824, 0.9822) |

*Each cell in the table represents the mean $F_1$ score (mean precision, mean recall, overall accuracy) for the corresponding experimental configuration.*
*The bold values are the $F_1$ scores of the best performing models in the respective row/column.*

We discovered that the model was capable of resolving the overfitting issue after rigorous training with the best 20% of the data and testing the functioning model with the remaining 80%. We attain an overall accuracy of 98.21% (F1 score of 0.9820) for 20–80 cases using Google Net. Learning transfer: color. The general performance of Alex Network and Google Internet declines as we continue to increase the test-to-train-set ratio (see 3D configuration), as expected. but this drop always occurs when we build the expected model. Significant enough to be considered. necessarily. A lot of overlearning. Moreover, the parent 3C showed no difference in training and validation losses, indicating that the results of our experiments are not affected by excessive stretching. Figure 3A shows that the Google network consistently outperforms the Alex network between the Alex Internet and Google Internet architectures. Figure 3B also shows that mastering the switch always gives better results, regardless of the training method used. His three versions of the dataset (colour, grayscale, and segmented) exhibit normal variances in overall performance between studies, even though the other experimental parameters stayed the same model works well with the colour model of the dataset. We made sure the neural network could benefit the most from the hardware, recording technique, and illumination biases as we were designing

the experiment. When creating the experiment, we made sure that the neural network could benefit the most from the biases that the hardware, recording technique, and illumination had by nature. Hooray. As expected, the overall performance decreased compared to the dataset colour model experiments, but the average F1 score was still 0.852 (baseline accuracy of 85 or 53%). A segmented variant of the dataset was reorganized to examine the impact of background images on all performance. As given in the original version 3E, the version with the segmented image performs slightly worse than the model, but consistently outperforms the version with the grayscale image. Please use the colour version of the image.

It also evaluates the performance of versions of images from online materials that contain educational extensions. These strategies have important implications for the Plant Village dataset, which can now be collected in a controlled environment. Since these images shouldn't be too many, we performed a visual inspection step using Bing image search and computer download of IPM images to obtain a small managed dataset (Dataset 1) of 121 and 119 images. Got. Data set 2 (see accompanying material for detailed instructions). Using the ideal model for these datasets, we attained a normalized accuracy of 31. 0% of Dataset 1 and 31.69% of Dataset 2 correctly anticipated the correct class designation (such as yield and illness statistics) out of the 38 available class designations. Note that the overall accuracy of the arbitrary classifier is only 2.63%. The correct class was one of the top five predictions for all images 52.89% of the time in data set 1 and 65.61% of the time in data set 2. Google Net had the best model for two datasets. Segmentation: Transfer learning: 80-20 for Google Internet and Data Set 1: Colour Transfer learning: 80-20 for Data Set 2. The provided Image displays a representation of various data sets along with a visualization of their activations in the fundamental layers of the Alex Net architecture.




The provided Image displays a representation of various data sets along with a visualization of their activations in the fundamental layers of the Alex Net architecture. The credit for the image goes to Slide Collection, Bug Wood, USDA Cooperative Extension, and Clemson University. In (B), the identification of representations in the first convolutional layer (conv1) of the Alex Net building revealed that Alex Net employs: colour: Train of blank: 80-20, excluding the image in panel b. As the model must account for both crop species and disturbances, all effects have been reported so far. We can restrict the name to a particularly good state where yields are given based on the vegetation ID being predictable with the help of plant breeders. To assess the overall effectiveness of the model in this scenario, we restrict ourselves to vegetation where we have at least n > = 2 (to avoid a sparse class) or n > = 3 harvest rate lessons. If n is greater than 2, dataset 1 contains nine cuts and 33 orders. A random guess in this data set could be accurate to 0.225, however our model is accurate to 0.78. When n is more than three, the dataset comprises 25 wires spread over five different vegetation types. For such a data set, a random guess can have an accuracy of 0.179, while our version has an accuracy of zero. 11. Furthermore, the thirteen trials in data set 2 are assigned to plants when n is greater than 2. In such a data set, the random guess can be as high as 0.31, while the accuracy of our model is 0.5 5 The data set has eleven classes for each of the three crops if n is greater than 3. Random guessing's accuracy in each of these datasets is 0.288, while our

model's accuracy is 0.85. Discussion Recent advances have been made in the localization of images and convolution wave imaging of brain networks with objects.

Moreover, traditional machine learning-based disease classification methods typically focus on a limited number of categories within a single culture. For instance, a common extraction and classification technique classifies tomato powdery mildew from whole tomato leaves using words and stereopsis (Raza et al., 2015). Frost detection is performed using RGB photography in an uncontrolled environment. He uses RGBD images for apple skin detection Citrus yellow dragons are identified using fluorescence imaging spectroscopy (Wetterich et al., 2012), and citrus is detected using near-infrared spectral style-based sensors (Sankaran et al., 2011). Tomato Yellow Leaf Curl Virus is detected using traditional signature extraction procedures and directional vector machine pipelines among others. A recent overview of machine learning for plant phenotyping provides a comprehensive overview of the state of the field (Singh et al., 2015). Neural networks have previously been used to detect plant diseases (Huang, 2007) to classify and detect plant diseases of Phalaenopsis such as bacterial susceptible rot, bacterial brown spot, and phytophthora black rot. However, this method required careful image processing. Before the neural network categorized the texture features, I decided to generate a list of them. Based on Krizhevskij et al. (2012), our strategy found that using a deep convolutional neural ensemble structure and continuous supervised training is an effective way to outperform conventional methods even on a relatively wide range of image classification problems. For the first time, it is an alternative. Handcrafted properties are used with wider margins than common benchmarks. Functional design is a highly promising candidate for a rational and scalable method for computational verification of plant illnesses because it isn't arduous and the answer is generalizable. We developed a model for plant leaf photos using a sophisticated convolution neural network that can recognize crop varieties as well as the presence and kind of illness in previously unknown images. With 5,306 photos, 38 categories, 1 crop species, and 26 illnesses (or absence thereof), the Vegetable Village dataset has a 99.35% success rate. Consequently, the model proficiently classifies cropping and blending without requiring feature formation in 993 photos out of 1000 from 38 possible training sessions. Of utmost importance, The classification process itself can be speedy (less than 2d on a CPU), and it is simple to implement on a computer or smartphone, even though training on a

powerful GPU cluster computer could take several hours. This presents a clear pathway for the global analysis of plant diseases via smartphones.

Nevertheless, there are some limitations to peak phases that need to be addressed in future research. Specifically, the accuracy of the version decreased significantly to just over 31 percent when analyzing hard and fast footage taken outside of a school environment. Although this accuracy is considerably higher than that of 38 random training sessions (2.6%), it is necessary to diversify the training data to enhance accuracy. The present findings highlight the potential for significant improvement in accuracy by developing more diverse datasets with additional factual information. However, the challenge lies in the fact that we are currently restricted to individual journal categories that only cover a specific area of expertise. While these terms are useful, comprehensive software must be capable of classifying images of diseases in various institutions. It is worth noting that many diseases do not spread through the leaves, but rather through specific parts of the plant. As a result, fresh photo collections must take pictures from unusual perspectives and in as realistic a situation as they can. Farmers are also expected to have in-depth knowledge of the 38 species of plants they cultivate as well as the diseases that are known to affect each crop, which eventually reduces this issue to a more complex practical requirement. Restricting the class design might not have a discernible impact on the occurrence of diseases given the great precision of the Plant Village dataset.

Real datasets can, however, result in high precision. The technique is anticipated to improve significantly with more training data and generally performs astonishingly well for a variety of rare plant species and diseases. It is crucial to stress that our methods are intended to complement existing illness analysis techniques rather than to replace them. Long-term reliability of diagnoses based primarily on observable symptoms and signs is seldom guaranteed, and initial diagnosis by visual inspection alone is frequently difficult. Nevertheless, with over 5 billion smartphones predicted to be in use by 2020 (GSMA Intelligence, 2016), this technique can serve as an additional method to avoid lost sales. Moreover, you can enhance the accuracy by adding location and time records to your smartphone image in the future. Finally, it would be prudent to bear in mind the rapid advancement of mobile phone technology in recent years, which may continue to evolve. We are content with the results, so we acknowledge that an extremely precise diagnosis over the phone is only a matter of time.

**Resources**

The data and the code used in this paper are available at the following locations:

Data: https://github.com/salathegroup/plantvillage_deeplearning_paper_dataset

Code: https://github.com/salathegroup/plantvillage_deeplearning_paper_analysis

More image data can be found at https://www.plantvillage.org/en/plant_images

**References**

Bay, H., Ess, A., Tuytelaars, T. and Van Gool, L. (2008). Accelerated and continuous functions (surfing). Calculate. maybe Image Understand. 110, 3 6–359. Doi: 10.1016/j.cviu.2007.09.01

CrossRef Full Text | Google Scholar Chéné , Y. , Rousseau , D. , Lucidarme , P. , Bertheloot , J. , Caffier , V. , Morel , P. et al. (2012). On the use of a depth camera for 3D phenotyping of whole plants. Calculate. an electron Agriculture. 82, 122–127. Doi: 10.1016/j.compag.2011.12.007

CrossRef Full Text | Google Scholar Dalal, N., and Triggs, B. (2005). "Histograms of oriented gradients for human detection", julikussa Computer Vision and Pattern Recognition, 2005. CVPR 2005.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K. and Fei-Fei L. (2009). "Imagenet: a large-scale hierarchical image database", in Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on. (IEEE).

Ehler, L.E. (2006). Integrated pest management (IPM): definition, historical development and application, and other IPM. Pest Manager. Sci. Rev. 62, 787–7 Doi: 10.1002/ps.12 7

Everingham , M. , Van Gool , L. , Williams , C. K. , Winn , J. Kaj Zisserman , A . (2010). Pascal Visual Object Classes (VOC) challenge. Int. J. Komputo Vis. 88, 303–338. Doi: 10.1007/s11263-009-0275-

Garcia-Ruiz , F. , Sankaran , S. , Maja , J.M. , Lee , W.S. , Rasmussen , J. and Ehsani R . (2013). Comparison of two aerial imaging platforms for detection of citrus trees infected with huanglongbing. Calculate. an electron Agriculture. 91, 106–115. Doi: 10.1016/j.compag.2012.12.002

Harvey, C . A., Rakotobe, Z. L., Rao, N. S., Dave, R., Razafimahatratra, H., Rabarijohn, R. H. et al. (201). Extreme vulnerability of smallholders to agricultural risks and climate change in Madagascar. In Son. Trans. R. Soc. Londono. B

Biol. Sci. 369:20130089. Doi: 10.1098/rstb.2013.008 PubMed Resumo

He, K., Zhang, X., Ren, S. kaj Sun, J. (2015). Deep residual learning for image recognition. archiveXiv:1512.03385. PubMed abstracts

Hernandez-Rabadan, D.L. , Ramos-Quintana , F. and Guerrero Juk , J. . (201). Integrating somites and a Bayesian classifier for segmentation of diseased plants in uncontrolled environments. To know World J. 201: 21 67. doi: 10.1155/201 /21 67

Huang, K. Y. ( 2007 ). Application of an artificial neural network for disease detection of phalaenopsis seedlings using colour and texture features. Calculate. electron. Agriculture. 57, 3–11.
Doi: 10.1016/j.compag.2007.01.015

Hughes, DP and Salathe, M . (2015). An open-access archive of plant health-related images enabling the development of mobile disease diagnostics. arXiv:1511.08060 Google Scholar ITU (2015).

Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., et al. (2015). Caffe: a well-established architecture for rapid embedding of functions. archiveXiv:1 08.5093.

Krizhevskij , A. , Sutskever , I. and Hinton , G . E. (2012). "Imagenet classification with deep convolutional neural networks," in Advances in Neural Information Processing Systems, ed. F. Pereira, C.J. C. Burges, L. Bottou and K. Q. Weinberger (Curran Associates, Inc.), 1097-1105. LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W. et al. (1989). Backpropagation applied to handwritten zip code detection. Neurocomputing. 1, 5 1–551. Doi: 10.1162/neco.1989.1. .5 1

LeCun, Y., Bengio, Y. and Hinton, G. (2015). Deep learning. Nature 521, 36–. Doi: 10.1038/luonto1 539 PubMed Abstract

Mokhtar, U. , Ali , M . A., Hassanien, A.E. and Hefny, H. (2015). "Identify Two Tomato Leaflet Viruses using a Support Vector Machine," in Information Systems Design and Intelligent Applications, ed. J.K. Mandal, S. C. Satapathy, M. K. Sanyal, P. P. Sarkar, A. Mukhopadhyay (Springer), 771-782. Raza, S.-A., Princo, G., Clarkson, J. P., Rajpoot, N. M. et al. (2015). Automatic detection of diseased tomato plants using thermal and stereo photography. PLoS ONE 10:e0123262.

Doi: 10.1371/journal.pone.0123262. Available online at: https://wwws.plos.org/plosone/article?id=10.1371/journal.pone.0123262

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S. et al. (2015). ImageNet's large-scale visual recognition challenge. International J. Computing. maybe 115, 211–252. doi:10.1007/s11263-015-0816-y

Sanchez, PA and Swaminathan, MS (2005 ). Halve world hunger. Science 307, 357–359. doi: 10.1126/science.1109057

Sankaran, S., Mishra, A., Maja, J.M . and Ehsani, R. (2011). Visible infrared spectroscopy for detection of huanglongbing in citrus fruits. Calculate. an electron Agriculture. 77, 127–13. Doi: 10.1016/j.compag. 2011.03.00

Schmidhuber, J. (2015). Deep Learning in Neural Networks: An Overview. Neural Network 61, 85-117. Doi: 10.1016/j.neunet.201 .09.003

Simonyan, K. and Zisserman, A. (201). Ultra-deep convolutional networks for large-scale image recognition. Archive Xiv: 1 09.1556.

Singh, A., Ganapathy Subramanian, B., Singh, A. K., and Sarkar, S. (2015). Machine learning for high-throughput stress phenotyping in plants. Trends Plant Sci. 21, 110–12 doi: 10.1016/j.tplants.2015.10.015

Strange, R. N. and Scott, P.R. (2005). Plant diseases: a threat to global food security. Phytopathology 3, 83-116. doi: 10.11 6/annurev.phyto. 3.11300 .133839

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D. et al. (2015). "Going Deeper with Circuits" julikussa Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Google Scholar Tai, A. P., Martin, M . V. and Heald, C.L. (201). Ozone-induced climate change and air pollution threaten future global food security. Nat. Clima. Chang, 817-821. Doi: 10.1038/nclimate2317

UNEP (2013). Small farmers, food security, and the environment. Rome: International Fund for Agricultural Development (IFAD). Havebla rete at: https://www.ifad.org/documents/10180/666cac21b6 3c2876d9c2d1f01d5dd Wetterich, C. B., Kumar, R., Sankaran, S., Junior, J. B., Ehsani, S. L. (2011). A comparative study of the application of computer vision and fluorescence imaging spectroscopy