INTRODUCTION

Advanced research in agriculture has led to the concept of smart farming and development of novel techniques to improve the crop yield and provide assistance to the farmers, but pests and plant diseases are still a dominant threat to the overall plant health and food production. With the current population growth rate, it is imperative that researchers investigate every potential solution that would facilitate the protection of crop yield. For sustainable development, there is a need to minimize the economic and production losses, along with the degrading impact of fertilizers and pesticides on the environment. Several researchers have consecrated their valuable efforts in developing new and innovative practices for disease diagnosis in plants. Previously, this domain has remained highly dependent on the conventional and long-established approaches. Traditional methods encompass visual inspections by experienced individuals, be it the farmer or a botanist. Detecting plant diseases in a lab is one other option. It comprises of bringing back samples from the field and conducting microscopic evaluation and diagnostic experiments, such as ELISA, PCR etc. While this is an accurate procedure, it is time-consuming, not to mention costly, requiring laboratory equipment set-up and highly labor-intensive. Given the restricted access to resources and limited expertise in plant pathology, there is a dire need for automated processes. Farmers are struggling throughout the world to protect their crop from the onslaught of several harmful microorganisms or pathogens such as virus, bacteria, fungus, nematodes, protozoa (Shankar, Harsha, & Bhandary, 2014) as well as feeding of the insects. A susceptible host and favorable environmental conditions are all these pathogens require, in order to infect the plants and eventually degrade their growth, sometimes resulting in high mortality. They are ubiquitous and stay inactive in the soil, air and/or water, even in the crop debris for multiple seasons, until they encounter suitable conditions to infest the plant. In the event of infestation of pests in the crop, it undergoes biological and chemical stress that initially fails to provide any visual cues. It is only after the disease has manifested to a greater severity level, that visual symptoms appear and demand immediate action to control the outbreak. The visual symptoms might encompass changes in color, shape and/or size of the parts of the plant. Also, to detect any such symptoms requires constant monitoring of the crop, which can prove to be tiresome task. For Agronomists and researchers, plant disease diagnosis has become a greater concern over time and has led them to devise a multitude of techniques for early detection and prediction of diseases that acutely affect the yield and quality of the grain. Recent research in the sensor based techniques have led to the identification and development of imaging technologies that are extensively being utilized for detection as well as identification of plant diseases, in addition to being non-destructive in nature. Images capture any anomalies in the occurrence of a plant, which might indicate the attack of a pathogen. Optical properties of a plant are exploited to detect plant stress levels and disease severity. Lately, the emphasis has been towards early disease detection, which is possible due to the recent advances in the remote sensing and imaging technologies (Mahlein, 2016) such as multispectral imaging, hyperspectral imaging, thermal imaging etc. As in humans, early diagnosis in plants can help eliminate and cure the disease, with minimum cost, effort and use of a pesticide, hence preserving the ecosystem from its damaging effects. For more than last four decades, researchers have contributed significantly towards achieving this goal by utilizing the computational power of advanced technologies such as machine learning, deep learning, digital image processing, computer vision, genetic algorithms, big data, internet of things etc. These areas of computing have led to various suitable solutions in agriculture, from crop protection and monitoring to crop quality and its management. With their combined application, the process of data collection and analysis can be largely automated, leading to the timely prediction of most probable diseases in crops. It would also eliminate the dependency on manual tasks and subject matter experts.

Table 1 gives the nomenclature of the general terms used throughout the text. Following sections in this chapter highlights numerous image processing, machine learning and deep learning based plant disease diagnosis techniques devised in the last few decades, along with the implementation of five CNN models and a mobile application to predict tomato plant diseases.

Table 1. Nomenclature

ANN	Artificial Neural Networks
CNN	Convolutional Neural Networks
REG	Regression
SVM	Support Vector Machines
BPNN	Back Propagation Neural Networks
GRNN	Generalized Regression Neural Networks
MAE	Mean Absolute Error
VI	Vegetation Indices
MDC	Minimum Distance Criterion
FFBPNN	Feed Forward Back Propagation Neural Networks
GA	Genetic Algorithms
ROC	Receiver Operator Characteristic
PCA	Principal Components Analysis
RBF	Radial Basis Function
MLP	Multi-Layer Perceptron
FCM	Fuzzy C Means
LVQ	Learning Vector Quantization
DPLS	Discriminant Partial Least Squares
MD	Mahalanobis Distance
MSE	Mean Squared Error

BACKGROUND

Pathogen outbreak is highly dependent on environmental conditions. Early prediction approaches employing forecasting technologies, based on weather, can prove to be highly beneficial in minimizing the pesticide usage for disease control in plants; thereby reducing the subsequent cost and environmental damage. Such systems can act as a warning system and heed in the timely control of the disease. Kaundal et al. (2006) have designed cross-year and cross-location models for duration of five consecutive years from 2000-2004 and 5 distant locations, for rice blast. The coefficient correlation, coefficient of determination and %MAE was computed and validated by implementing and comparing prediction accuracies of SVM, REG, GRNN and BPNN. Six environmental predictor variables were recorded: temperature (min/max), relative humidity (min/max), rainy days in a week and rainfall, where rainfall was found to be the most effective among the other variables. In response to the attack of several micro-organisms, plants have developed certain vital self defense mechanisms viz. genes for resisting diseases, also known as R genes. Pal et al. (2016) have investigated a SVM-based machine learning tool was trained for fast and accurate identification of R Proteins (112 in total), with 91.11% accuracy. These R Proteins can be exploited for the breeding of novel plant species that are inherently disease resistant. SVM has been the most efficient machine learning technique employed in various researches. Pattern recognition techniques are also majorly applied for disease identification in plants, whereby numerous features pertaining to color, shape and texture are extracted (Camargo & Smith, 2009). Not all features are helpful for training the machine learning tool, they may provide superficial, redundant or no information at all. Hence, such features can be discarded after thorough analysis. Petrellis (2015) has proposed a cost effective, low complexity approach for a mobile application that focuses on visible lesions, their coverage and count, rather than extracting features, which as mentioned, is a common practice in approaches based on Image Processing.

For majority of Image processing based techniques, plant leaf is the most utilized part of a plant for disease identification, as it shows clear visual symptoms of changes in color, shape or size. Recent advancements and research in sensor-based techniques have resulted in various non-destructive approaches, applicable to disease diagnosis and plant pathology. Measurements of disease severity level and stress in a plant can be captured by observing its optical properties. RGB imaging, multi- and hyperspectral imaging, thermal imaging, fluorescence spectroscopy are amongst the few non-invasive methods for prediction and prevention of plant diseases (Mahlein, 2016). As mentioned previously, visual symptoms indicate the manifestation of the disease to a greater severity level, hence modifying the appearance of the plant. In such cases, RGB images of the diseased plant are captured via a digital camera deployed in the field. These images are preprocessed and transformed into some suitable color model for feature extraction and finally, computed features are used for training a machine learning based tool for classification of the disease. Al-Bashish et al. (2011) have converted the RGB images into HSI color model, employed k-means clustering for segmentation, co-occurrence method for eleven texture-based feature extraction and trained FFBPNN with ten hidden layers for classification that resulted in 93% accuracy. Similarly, Sabrol and Kumar (2016) have suggested extracting ten statistical features such as mean, skewness, standard deviation and correlation, and exercising a decision tree with 80 nodes for classification of tomato diseases. These four steps (color space conversion, segmentation, feature extraction and classification) have been the most commonly used approach in Image Processing field concerning RGB imaging for disease identification in plants. With minor modifications, researchers have investigated the variations of this process to better the results such as using mean, normalized mean and green ratios of R, G, B values, for computing features and accuracy (Liu, Zhang, Shu, & Jin, 2013). Another promising imaging technique is Hyperspectral imaging that has emerged as an effective method for early prediction, without the necessity of appearance of the visual symptoms. It captures spatial as well as spectral data, conforming to the subtle modifications in the biological and chemical properties of a plant in the form of reflectance values in various wavelengths. These reflectance values are then employed to compute predefined VIs such as Normalized difference vegetation index (NDVI), Simple ratio index (SRI), Photochemical reflectance index (PRI), Plant senescence reflectance index (PSRI) etc., whose values indicate the biotic and abiotic stress in the plants such as water content, chlorophyll content, nutrients deficiency etc., hence, initiating the necessary course of action. Multispectral Imaging is similar to hyperspectral, as in both capture spectral data, but multispectral imaging records data for 3-10 broad bands while hyperspectral imaging captures data in thousands of narrow bands, hence proves to be more accurate, while posing difficulty in identifying significant bands for data analysis (Zhao, Li, Yu, Cheng, & He, 2016), not to mention expensive and sophisticated hardware setup. Increased temperature of the plant surface can also indicate infestation of pathogens, where different levels of infection have distinct heat signatures. For such cases, thermal imaging can be used to capture the reflected infrared radiations for disease diagnosis.

Table 2 and Table 3 summarizes the numerous machine learning based statistical methods and imaging techniques successfully applied for identification, quantification and classification of plant diseases. Genetic Algorithms based segmentation of colored images of cotton crop has been proposed by Singh and Misra (2016), whereby after preprocessing and image enhancement, green colored pixels are masked followed by the segmentation. Features such as energy, contrast, shade etc. are computed via the color co-occurrence method, and SVM and MDC are trained for classification. SVM outperforms MDC and gives a prediction accuracy of 95.71%. The combination of GA and SVM has been successfully applied for examining the quality of rice seeds at an early stage to differentiate between healthy and diseased food grains, infected with a seed-borne disease called Bakanae disease (Chung et al., 2016).

Table 2. Summary of statistical approaches for Plant Disease Identification and Classification

Publication and Year	Statistical Approach	Crop	Disease
Ferentinos (2018)	AlexNetOWTBn, VGG	25 Plant Species	58 different plant-disease pair

Wang, Sun, & Wang (2017)	Deep Learning (VGG16, VGG19, Inception-v3, ResNet50)	Apple	Apple Black Rot
Brahimi, Boukhalfa, & Moussaoui (2017)	Deep Learning: AlexNet, GoogleNet	Tomato	Spider Mites, Septoria Spot, Late blight, Early blight, Bacterial Spot, Tomato Yellow Leaf Curl Virus, Leaf Mold, Tomato Mosaic Virus, Target Spot.
Durmus, Günes, & Kirci (2017)	Deep Learning: AlexNet, SqueezeNet	Tomato	Spider Mites, Septoria Spot, Late blight, Early blight, Bacterial Spot, Tomato Yellow Leaf Curl Virus, Leaf Mold, Tomato Mosaic Virus, Target Spot.
Lu, Zhou, Gao, & Jiang (2017)	ROC curve analysis	Tomato	Yellow Leaf Curl Disease
Mohanty, Hughes, & Salathé (2016)	Deep Learning Models (GoogleNet and AlexNet)	14 crops	26 diseases
Sabrol & Kumar (2016)	Decision Trees	Tomato	Late Blight, Bacterial Leaf Spot, Septoria Leaf, Leaf Curl, Bacterial Canker
Zhao et al. (2016)	Partial Least Square Regression (PLSR)	Cucumber	Angular Leaf Spot (ALS) disease
Singh & Misra (2016)	SVM classifier + Minimum Distance Criterion+ K-means	Rose, beans leaf , lemon leaf, banana leaf	Banana leaf: Early scorch disease; Lemon leaf: Sun burn disease; Beans & Rose: Bacterial disease; Beans leaf: Fungal disease.
Chung et al. (2016)	SVM	Wheat	Bakanae disease
Sabrol & Kumar (2016)	Fuzzy Inference System (FIS), Adaptive neuro-fuzzy inference system, Multi-layer FFBPNN	Tomato	Late Blight, Bacterial Leaf Spot, Tomato Leaf Curl, Septoria Leaf Spot, Bacterial Canker.
Rupanagudi, Ranjani, Nagaraj, Bhat, & Thippeswamy (2015)	K-means Clustering and PCA	Tomato	Borer Insects

Raza, Prince, Clarkson, & Rajpoot (2015)	SVM classifier with linear kernel	Tomato	Powdery Mildew
Xie, Shao, Li, & He (2015)	Extreme Learning Machine (ELM) classifier model	Tomato	Early blight, Late blight
Liu, Zhang, Shu, & Jin (2013)	RBF SVM	Wheat	Stripe rust, powdery mildew, leaf rust, leaf blight
Arivazhagan, Shebiah, Ananthi, & Varthini (2013)	MDC, SVM	Tomato	Bacterial leaf spot, Leaf spot
Sankaran, Mishra, Ehsani, & Davis (2011)	LCA, QDA, K-nearest neighbor, SIMCA	Citrus Orchards	Huanglongbing (HLB)
Al-Bashish, Braik, & Bani-Ahmad (2011)	K-means clustering and FFBPNN	General leaf images	Late Scorch, Tiny Whiteness, Cottony Mold, Ashen Mold, Early Scorch.
Rumpf et al. (2010)	Decision Trees, ANN, SVM+RBF	Sugar Beet	Powdery mildew, sugar Beet rust, cercospora leaf spot,
Singh, Jayasa, & Paliwala (2010)	Linear, Quadratic, Mahalanobis and a BPNN classifier	Wheat Kernels	Grain Borer, Rusty Grain Beetle, Rice Weevil, Red Flour Beetle.
Ghaffari et al. (2010)	**Clustering:** K-Means clustering, PCA and FCM **For classification:** MLP, LVQ and RBF based ANNs	Tomato	Spider mite infected plants and Powdery mildew (Oidium lycopersicum)
Camargo & Smith (2009)	SVM	Cotton	Southern Green Stink Bug, Bacterial Angular, Ascochyta Blight
Wang , Zhang , Zhu, & Geng (2008)	BPNN with gradient descent	Tomato	Late Blight
Kuo-Yi Huang (2007)	BPNN	Phalaenopsis	Phalaenopsis seedling disease
Xu, Zhu, Ying, & Jiang (2006)	DPLS, MD.	Tomato	Tomato mosaic virus

Table 3. Imaging Techniques successfully investigated for several crops

Imaging Technique	Crop	Publication and Year

RGB Imaging	Apple, Tomato, PlantVillage Dataset (Blueberry, Potato, Raspberry, Strawberry, Cherry, Corn, Grape, Orange, Peach, Pepper, Soybean), Rose, beans leaf , lemon leaf, banana leaf, Wheat, Cotton, Phalaenopsis, Wheat Kernels, Watermelon, Pumpkin, Bell, Peach.	Ferentinos (2018), Wang, Sun, & Wang (2017), Brahimi, Boukhalfa, & Moussaoui (2017), Durmus, Günes, & Kirci (2017), Mohanty, Hughes, & Salathé (2016), Sabrol & Kumar (2016), Singh & Misra (2016), Chung et al. (2016), Raza, Prince, Clarkson, & Rajpoot (2015), Liu, Zhang, Shu, & Jin (2013), Al-Bashish, Braik, & Bani-Ahmad (2011), Camargo & Smith (2009), Kuo-Yi Huang (2007), Singh, Jayasa, & Paliwala (2010)
Hyperspectral Reflectance	Cucumber, Tomato, Sugar Beet, Wheat Kernels	Zhao et al. (2016), Xie, Shao, Li, & He (2015), Rumpf et al. (2010), Singh, Jayasa, & Paliwala (2010), Wang , Zhang , Zhu, & Geng (2008), Zhang, Qin, & Liu (2005), Zhang, Qin, Liu, & Ustin (2003)
Thermal & Stereo Visible Light Imaging	Tomato	Raza, Prince, Clarkson, & Rajpoot (2015)
Visible-near Infrared Spectroscopy	Citrus Orchards	Sankaran, Mishra, Ehsani, & Davis (2011), Xu, Ying, Fu, & Zhu (2007), Xu, Zhu, Ying**, &** Jiang (2006)

DEEP LEARNING: APPLICATIONS IN AGRICULTURE

Deep learning has been extensively applied to various computational fields viz. computer vision, speech recognition, natural language processing, image classification, but firstly it was utilized for detecting handwritten digits in documents. While machine learning requires pre-computed features to be fed to the classifier for training, deep learning streamlines the process and elicits these features directly from the data; hence, it encapsulates the feature extraction as well as classification processes. Its successful application in classifying images was exploited to distinguish among healthy and diseased plants, based on leaf images. But its performance is greatly dependent on two factors: size of the data and diversity of the data, both must be extensive for the deep learning model to present exceptional results. There are essentially four elementary operations performed by the CNN architectures, Convolution, that utilizes kernels or filters of varied sizes to generate multiple and distinct feature maps; Pooling (most commonly max pooling); ReLu (Rectified Linear Units) for introducing non-linearity; and Fully Connected Layer. Research shows that deep learning architectures outperform other machine learning tools significantly. In recent years, deep learning architectures have been largely evaluated with digital/RGB images and hyperspectral data as well. GoogleNet and AlexNet have been previously implemented on all plant images across each crop & disease pair of PlantVillage dataset (Mohanty, Hughes, & Salathé, 2016). This dataset is an open access depository of 54,306 images covering 14 crops and corresponding 26 diseases. This dataset has been utilized by many authors for implementing deep learning architectures. Another implementation includes evaluation of AlexNet and SqueezeNet models on PlantVillage dataset (Durmus, Günes, & Kirci, 2017). In contrast to other image processing techniques, where few hundred images suffice for image classification, deep learning requires thousands of images for training, as less number of training images might lead to overfitting. Apart from disease classification, deep learning has also been effectively applied to determine the severity levels of a disease ranging from healthy, early, to middle, and finally, end stage (Wang, Sun, & Wang, 2017). Authors trained the deep learning architectures on Apple black rot images from the PlantVillage dataset and achieved a test accuracy of 90.4%. Ferentinos (2018) has implemented AlexNetOWTBn and VGG on 87,848 images of 25 distinct plant species with 58 plant & disease pair, with 99.53% accuracy. These experiments have been performed on images collected/downloaded from the Internet, not digital images captured on the field in real environmental conditions. This can be considered as a shortcoming; hence, these applications must reach the farmers for actual evaluation of the technique.

SOLUTIONS AND RECOMMENDATIONS

Tomato plant was selected as the crop under consideration, with India being the second largest producer of tomatoes worldwide, as per the Food and Agriculture Organization of the United Nations (FAOSTAT). The authors have analyzed 18,160 images of tomato plant leaves, spread across 10 class labels assigned to them. Each class label is a predefined crop & disease pair. With the image of the plant leaf as input, the authors attempt to predict the correct plant-disease pair. Table 4 lists a few common tomato plant diseases along with their scientific names and category of pathogens. This application can be effectively utilized by farmers to meet their needs and protect the tomato yield by timely application of control measures. Through this chapter, the authors promote the practice of smart agriculture via automation. Our application aims to make the existing methods reach the growers, as well as, be user-friendly and accurate, with minimal human intervention. Once the disease has been diagnosed, treatment can be suggested through any communication medium.

Table 4. Common Tomato Plant Diseases

Disease Name	**Pathogen Scientific Name**	**Type**
Early Blight	Alternaria solani	Fungal
Late Blight	Phytophthora infestans	Oomycete
Buck Eye Rot	Phytophthora parasitica	Oomycete
Fusarium Wilt	Fusarium oxysporum f. sp. lycopersici	Hyphomycete
Septoria Leaf Spot	Septoria lycopersici	Fungal
Powdery Mildew	Leveillula taurica	Fungal
Bacterial Wilt	Pseudomonas solanacearum	Bacterial
Bacterial Leaf Spot	Xanthomonas campestris pv. Vesicatoria	Bacterial
Bacterial Canker	Clavibacter michiganensis pv. michiganensis	Bacterial
Tomato Mosaic	(ToMV)	Viral

Virus		
Tomato yellow leaf curl	Tomato yellow leaf curl virus	Viral
Anthracnose	Colletotrichum gloeosporioides	Fungal
Cercospora leaf mold	Pseudocercospora fuligena	Fungal

Source: *(Shankar, Harsha, & Bhandary, 2014)*

Figure 1 depicts images of healthy as well as diseased tomato leaves from the PlantVillage dataset. In every approach described in this chapter, the authors have resized the given images to 256 × 256 pixels, and performed model optimization along with predictions, on all downscaled images. Figure 2 demonstrates the resultant images after intermediate convolution, pooling and activation operations performed on the input images.

Figure 1. Tomato healthy and diseased leaf images form PlantVillage Dataset (Top row left to right) (a) Healthy (b) Tomato Mosaic Virus (c) Early Blight (d) Bacterial Spot (Bottom row left to right) (e) Late Blight (f) Septoria Leaf Spot (g) Yellow Leaf Curl (h) Leaf Mold

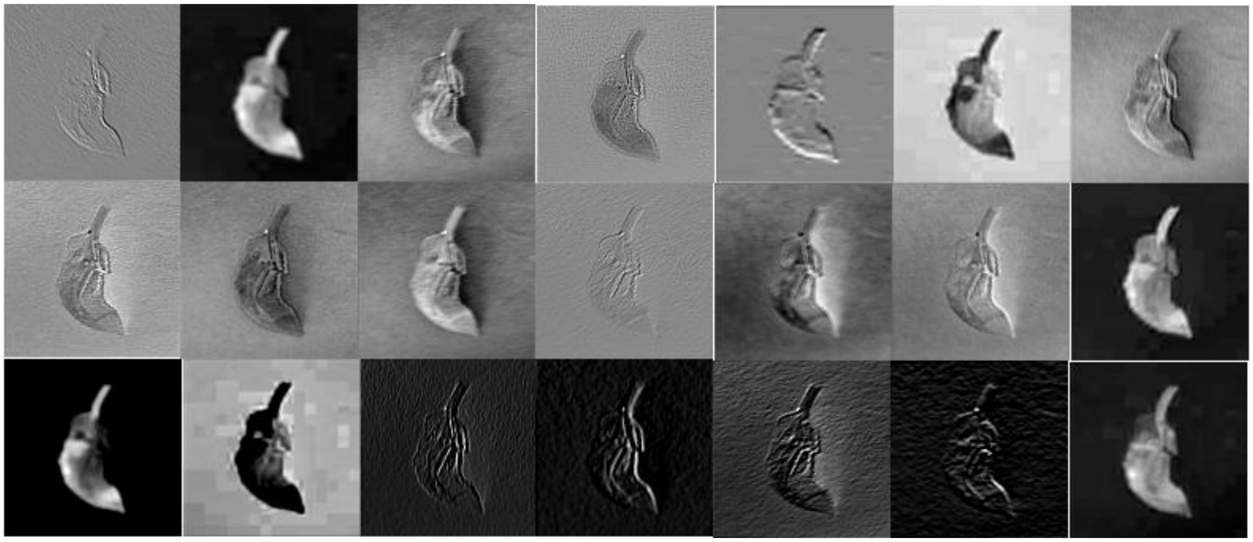

Figure 2. Intermediate resultant images after image processing by CNN

Due to hardware limitations, the authors decided to only use bottlenecking. The original approach was to take a pre-trained network, slice off the end (the part that classifies inputs), stick a new classification network on to replace it, freeze the convolutional layers, and then train the new part of the network. Easier said than done, also, much more expensive to do. By creating bottlenecks, one can simply record the feature map outputs as they are before they would have been fed into the classification layers that were removed. In other words, bottlenecks require zero training. They are produced by the model predicting on your new data set. It is exponentially quicker to use a model to predict than it is to train it (For eg. the computation that goes into stochastic/mini-batch gradient descent). Additionally, since bottlenecking does not require parallel computing, it can be done on a CPU. The authors have a train-test split of 70%-30%. And the performance measures include accuracy, loss, mean precision, mean recall, confusion matrix, and mean F1-micro score (as the proportion of images per class is not balanced). The authors assess the relevance of deep convolutional neural networks for the classification problem mentioned above. The authors target the five popular architectures, namely Inception_ResNet_V2, ResNet50, VGG16, VGG19, and Xception for image classification. Table 5 provides some differentiating details of all five models.

Table 5. Few details of CNN Architectures implemented

Model	Size (in MB)	Parameters	Depth
Xception	88	22,910,480	126
VGG16	528	138,357,544	23
VGG19	549	143,667,240	26
ResNet50	99	25,636,712	168
InceptionResNetV2	215	55,873,736	572

Source: https://github.com/GeorgeSeif/Transfer-Learning-Suite

Inception_ResNet_V2

Inception-ResNet-v2 is a modified version of Inception V3 model, roughly based on Microsoft's ResNet. It is a CNN that delivered better accuracy on the ILSVRC image classification benchmark. With residual connections, it leads to decreased training times.

- Basic structure of Inception-ResNet-v2 (layers, dimensions) are as follows:

    - Image -> Stem -> 5x Module A -> Reduction-A -> 10x Module B -> Reduction B -> 5x Module C -> AveragePooling -> Droput 20% -> Linear, Softmax
    - 299x299x3 -> 35x35x256 -> 35x35x256 -> 17x17x896 -> 17x17x896 -> 8x8x1792 -> 8x8x1792 -> 1792 -> 1792 -> 1000
- Modules A, B and C are similar.
- They contain 2 (B, C) or 3 (A) branches.
- Each branch starts with a 1x1 convolution on the input.
- All branches merge into one 1x1 convolution (which is then added to the original input, as usually in residual architectures).
- Module A uses 3x3 convolutions, B 7x1 and 1x7, C 3x1 and 1x3.
- The reduction modules also contain multiple branches. One has max pooling (3x3 stride 2), the other branches end in convolutions with stride 2.

VGG16 and VGG19

Simonyan & Zisserman introduced VGG network in 2014. VGG is individualized by its uncomplicated structure, utilizing *3×3* convolutional layers, each placed on top of the other in increasing depth. Max Pooling takes care of the reduced volume size. Two fully-connected layers with each having 4,096 nodes, is followed by a softmax classifier. The numbers '16' and '19' that are included in the names, represent the number of weight layers in the network. Deployment of VGG networks is tedious due to inherently slow training speeds and large architecture weights. It can be observed in Table 5 that VGG takes up more space as compared to other networks.

ResNet50

It was introduced in 2015. Unlike traditional architectures such as AlexNet, VGG & OverFeat, ResNet, with 50 weight layers, is build on micro-architecture modules also known as network-in-network architectures. Micro-architecture can also be viewed the set of "building blocks" utilised for construction of a network.
Updating the residual module for using identity mappings can help in obtaining better accuracy. ResNet being deeper than VGG16 & VGG19, the model size is remarkably smaller, as fully-connected layers have been replaced by the utilization of global average pooling— this significantly reduces the network size to 99MB for ResNet50.

Xception

Xception was first proposed by François Chollet, the creator of the Keras library. Xception is the extended version of the original Inception architecture, where the standard Inception modules are replaced with depth-wise separable convolutions. Xception exercises the smallest weight serialization at only 91MB. Xception, when trained on ImageNet dataset for image classification significantly outperformed Inception V3, due to the efficient usage of model parameters and incorporation of depth-wise separable convolution operation instead of Inception modules.

The authors exercised the following hyper-parameters in all of the experiments:

- Solver type: Stochastic Gradient Descent,
- Base learning rate: 0.001
- Decay: 1e-6
- Momentum: 0.9
- Gamma: 0.1

• Batch size: 128 (in case of Inception ResNet V2, ResNet50, Xception), 100 (in case of VGG16, VGG19).

ResNet50 had the highest train-test accuracies (99.12% - 96.14%), the highest F1 score (0.9612) and also the best *Confusion matrix*. This can be observed in Figures 3-8 depicting the resultant graphs of all five CNN architectures for performance measures accuracy, loss, and F1 mean score.

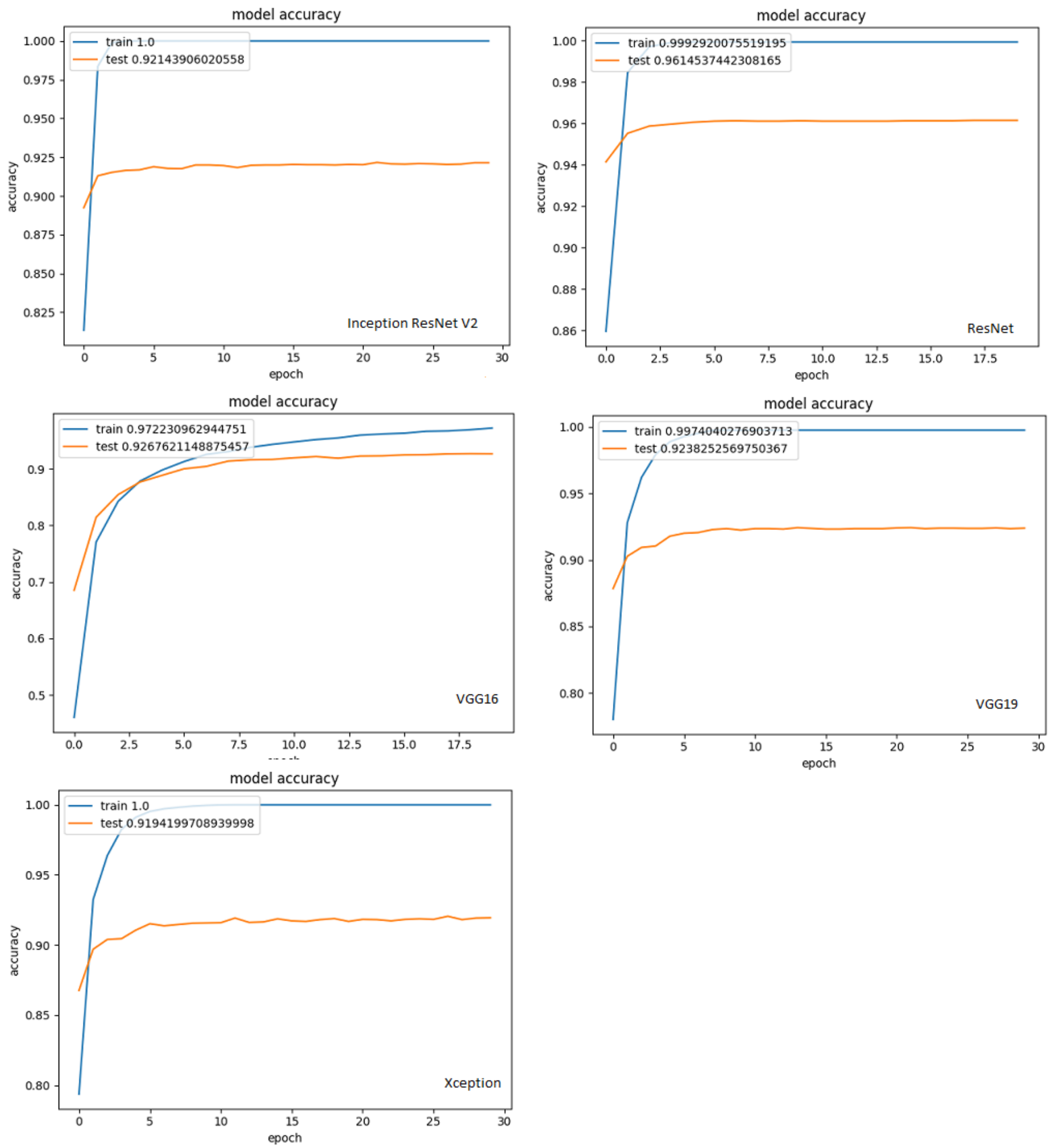

Figure 3. Model accuracies against epoch of all five CNN architectures

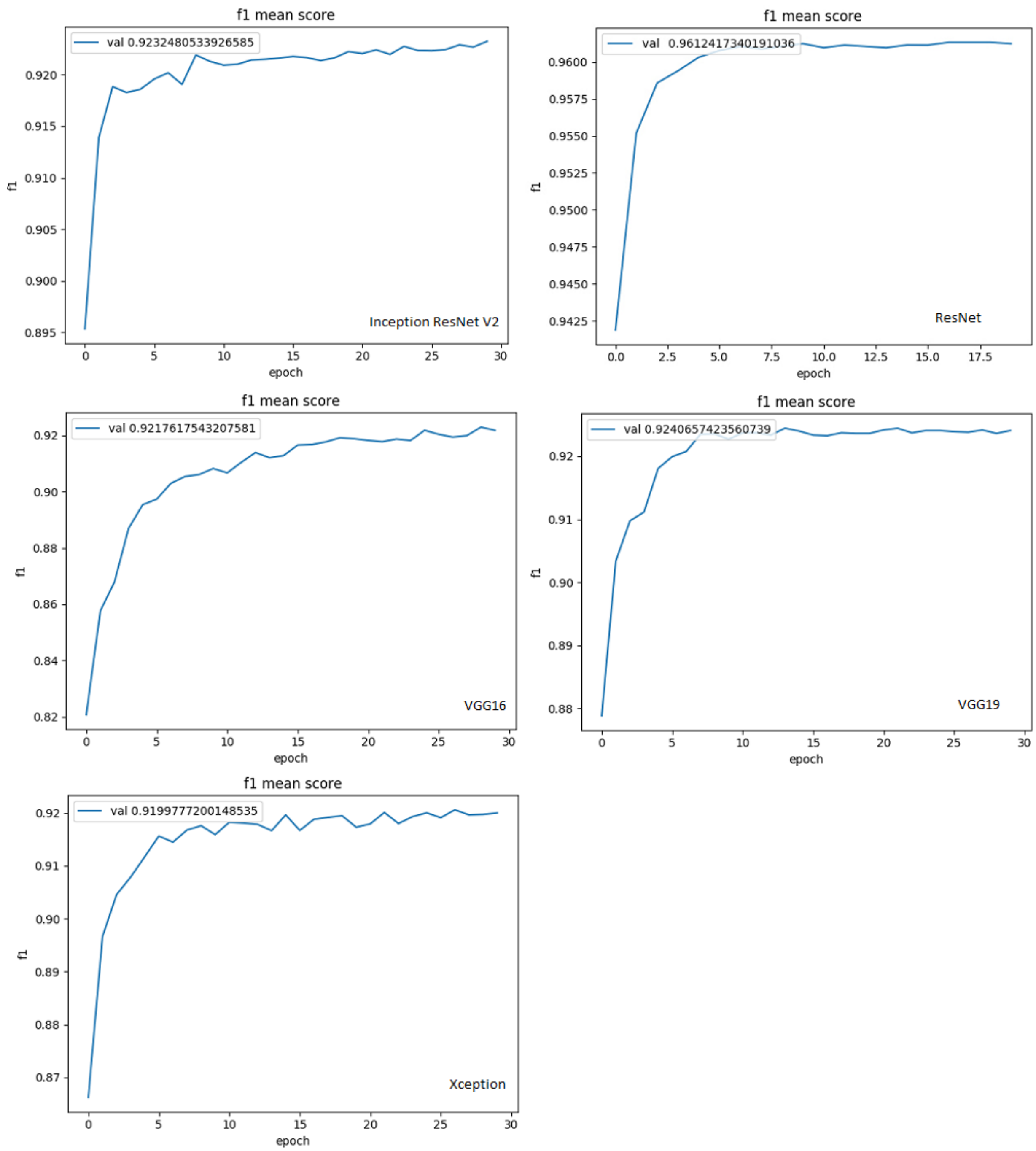

Figure 4. F1 Mean Score against epoch of all five CNN architectures

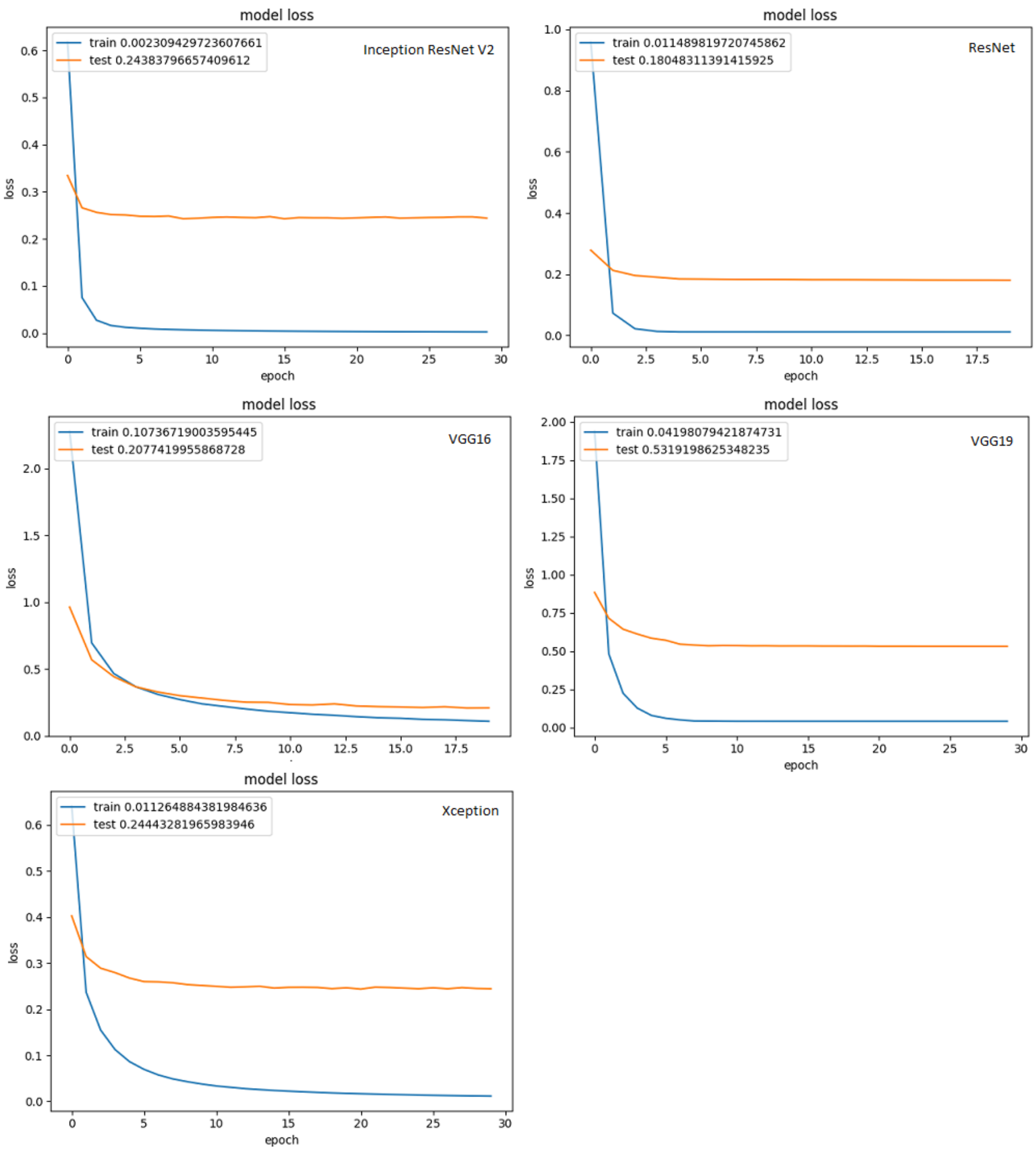

Figure 5. Model Loss against epoch of all five CNN architectures

Inception ResNet V2	Target_Spot	Late_blight	Tomato_mosaic_virus	Leaf_Mold	Bacterial_spot	Early_blight	Healthy	Tomato_Yellow_Leaf_Curl_Virus	Two-spotted_spider_mite	Septoria_leaf_spot
Target_Spot	**361**	0	1	2	4	6	5	1	22	8
Late_blight	12	**523**	0	9	3	13	1	5	3	7
Tomato_mosaic_virus	5	0	**81**	3	0	1	0	2	2	6
Leaf_Mold	14	2	2	**234**	2	3	0	1	4	6
Bacterial_spot	8	3	0	0	**598**	4	0	11	0	3
Early_blight	26	28	1	6	13	**230**	2	7	5	21
Healthy	15	1	0	0	0	1	**445**	0	0	2
Tomato_Yellow_Leaf_Curl_Virus	8	1	0	2	3	3	0	**1596**	8	0
Two-spotted_spider_mite	26	2	0	2	1	2	5	9	**473**	1
Septoria_leaf_spot	25	5	5	11	20	3	0	1	1	**451**

ResNet-50	Target_Spot	Late_blight	Tomato_mosaic_virus	Leaf_Mold	Bacterial_spot	Early_blight	Healthy	Tomato_Yellow_Leaf_Curl_Virus	Two-spotted_spider_mite	Septoria_leaf_spot
Target_Spot	**377**	3	1	2	0	4	5	1	13	4
Late_blight	3	**549**	0	7	1	10	1	3	0	2
Tomato_mosaic_virus	1	0	**92**	1	0	0	0	2	1	3
Leaf_Mold	3	2	0	**250**	1	2	0	1	3	6
Bacterial_spot	1	3	0	0	**617**	1	0	4	0	1
Early_blight	13	25	2	2	7	**277**	0	6	3	4
Healthy	5	1	0	0	0	0	**457**	0	1	0
Tomato_Yellow_Leaf_Curl_Virus	3	2	1	1	4	0	0	**1605**	4	1
Two-spotted_spider_mite	22	0	2	0	1	1	1	6	**487**	1
Septoria_leaf_spot	7	2	3	1	4	4	0	1	1	**499**

VGG-16	Target_Spot	Late_blight	Tomato_mosaic_virus	Leaf_Mold	Bacterial_spot	Early_blight	Healthy	Tomato_Yellow_Leaf_Curl_Virus	Two-spotted_spider_mite	Septoria_leaf_spot
Target_Spot	**376**	2	1	1	1	1	4	0	20	4
Late_blight	21	**521**	0	3	1	20	1	0	2	7
Tomato_mosaic_virus	4	0	**93**	0	0	0	0	0	1	2
Leaf_Mold	20	7	4	**212**	0	4	0	2	6	13
Bacterial_spot	14	5	0	0	**585**	8	0	7	1	7
Early_blight	31	31	0	0	9	**252**	0	2	2	12
Healthy	21	2	1	0	0	0	**438**	0	2	0
Tomato_Yellow_Leaf_Curl_Virus	13	0	1	1	15	4	0	**1573**	14	0
Two-spotted_spider_mite	44	3	2	1	0	2	2	8	**458**	1
Septoria_leaf_spot	28	4	3	0	4	3	2	0	7	**471**

VGG-19	Target_Spot	Late_blight	Tomato_mosaic_virus	Leaf_Mold	Bacterial_spot	Early_blight	Healthy	Tomato_Yellow_Leaf_Curl_Virus	Two-spotted_spider_mite	Septoria_leaf_spot
Target_Spot	**346**	3	1	3	3	11	9	0	25	9
Late_blight	7	**535**	0	11	0	14	0	0	1	8
Tomato_mosaic_virus	1	0	**91**	3	0	0	1	1	0	3
Leaf_Mold	2	8	2	**234**	1	5	0	2	10	4
Bacterial_spot	1	8	0	0	**595**	7	0	10	1	5
Early_blight	9	31	1	6	12	**257**	0	7	3	13
Healthy	11	2	1	0	0	1	**445**	0	4	0
Tomato_Yellow_Leaf_Curl_Virus	0	4	3	4	8	3	0	**1581**	18	0
Two-spotted_spider_mite	21	2	1	2	1	8	3	7	**472**	4
Septoria_leaf_spot	5	10	3	11	3	6	2	2	2	**478**

Xception	Target_Spot	Late_blight	Tomato_mosaic_virus	Leaf_Mold	Bacterial_spot	Early_blight	Healthy	Tomato_Yellow_Leaf_Curl_Virus	Two-spotted_spider_mite	Septoria_leaf_spot
Target_Spot	**367**	2	0	0	1	5	7	2	17	9
Late_blight	20	**518**	0	9	1	13	0	7	2	6
Tomato_mosaic_virus	7	0	**81**	4	0	1	0	0	4	3
Leaf_Mold	11	9	1	**224**	2	4	0	2	8	7
Bacterial_spot	7	5	0	0	**602**	3	0	4	0	6
Early_blight	30	35	0	4	10	**237**	2	2	5	14
Healthy	14	3	0	0	0	2	**442**	0	2	1
Tomato_Yellow_Leaf_Curl_Virus	6	4	0	1	7	4	0	**1585**	11	3
Two-spotted_spider_mite	42	1	0	1	2	0	5	5	**462**	3
Septoria_leaf_spot	23	4	5	5	8	11	2	4	3	**457**

Figure 6. Comparison Matrices of all five CNN architectures

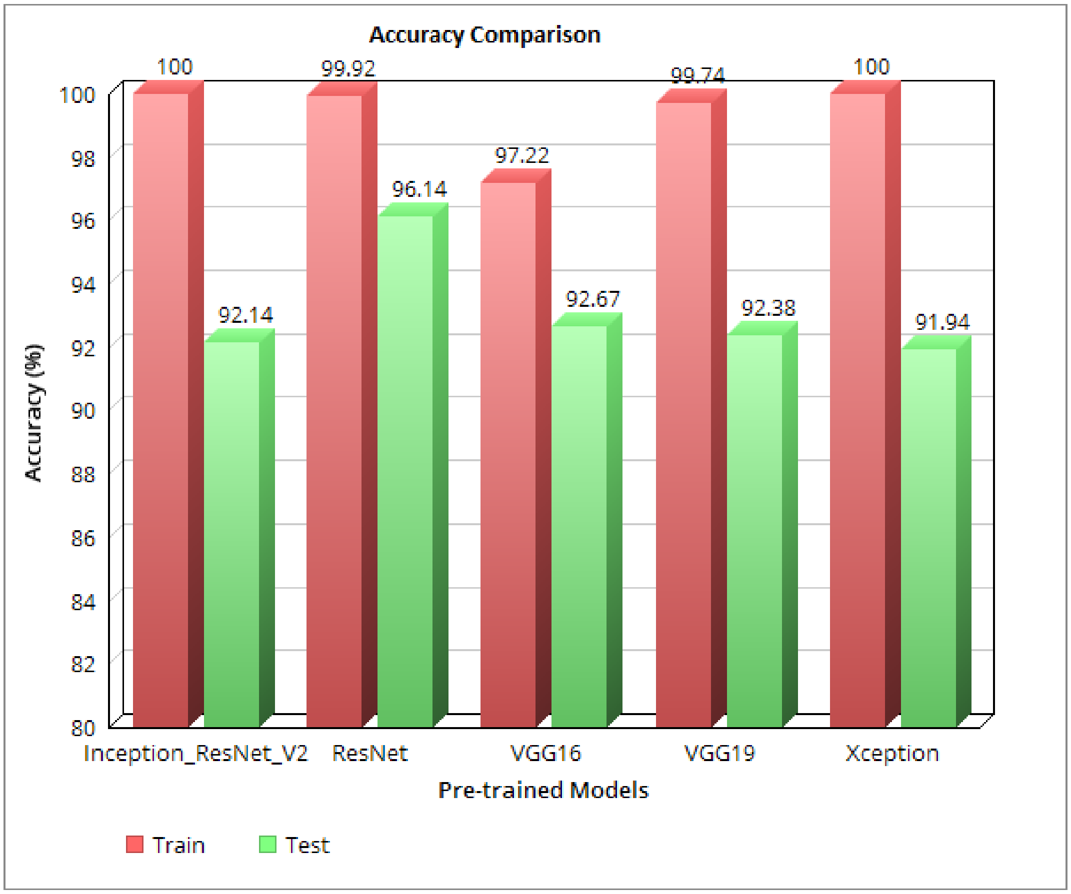

Figure 7. Train-Test dataset Comparison of all five CNN architectures

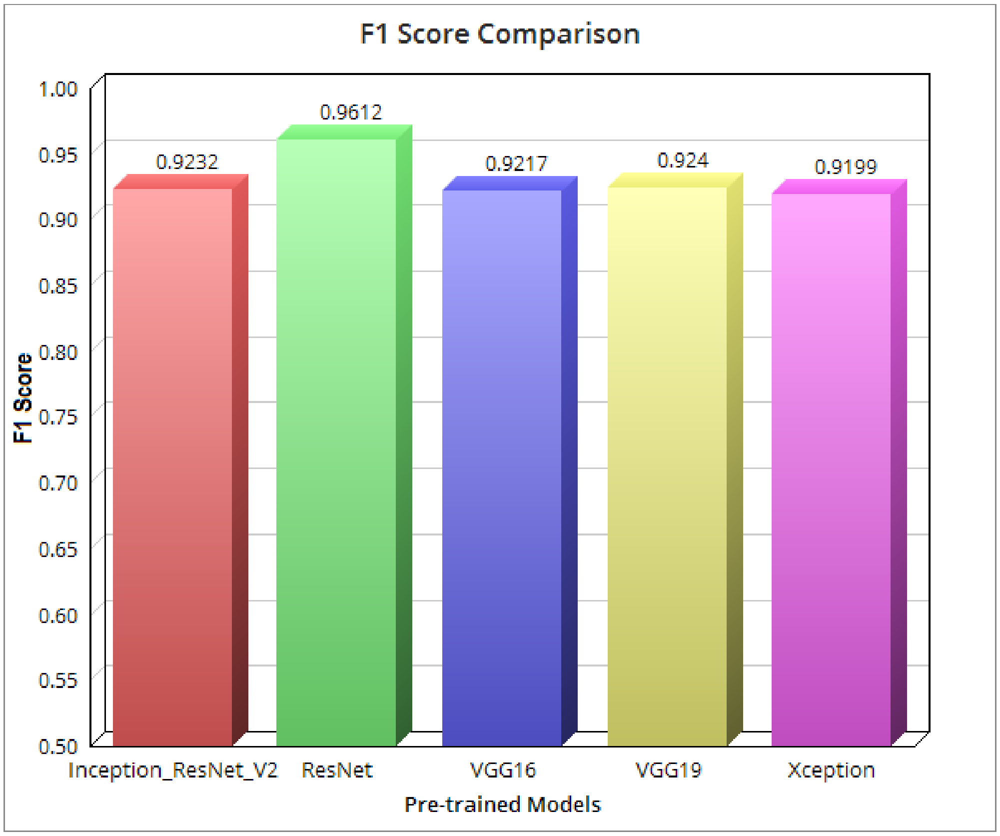

Figure 8. Performance Comparison of all five CNN architectures

ResNet50 was utilized to create a mobile application for identification and classification of Tomato plant diseases. The Tomato disease detection app is an android based application. The Purpose of this application is to help the growers detect the diseases in the tomato plant using the Image Based Detection techniques. In the current scenario, android phones are widely used by majority of the people in India because of its simplicity and user-friendly interface. Moreover, due to a drop in prices of android phones due to a highly competitive market, it has become easy for all sections of society to own an android phone and reap its benefits along with a lot of applications present on this platform. This app can be utilized in the agricultural sector to a great extent for disease detection. It will prove to be a great assistance particularly for farmers as it would eliminate the dependency on expensive machinery, high-end setup and experts, for diagnosing diseases in the plant. This app is easy to use and can be easily habituated by anyone because of its simple design and highly accurate results. After installing the application, it will appear as shown in Figure 9. The figure demonstrates the complete flow of the application.

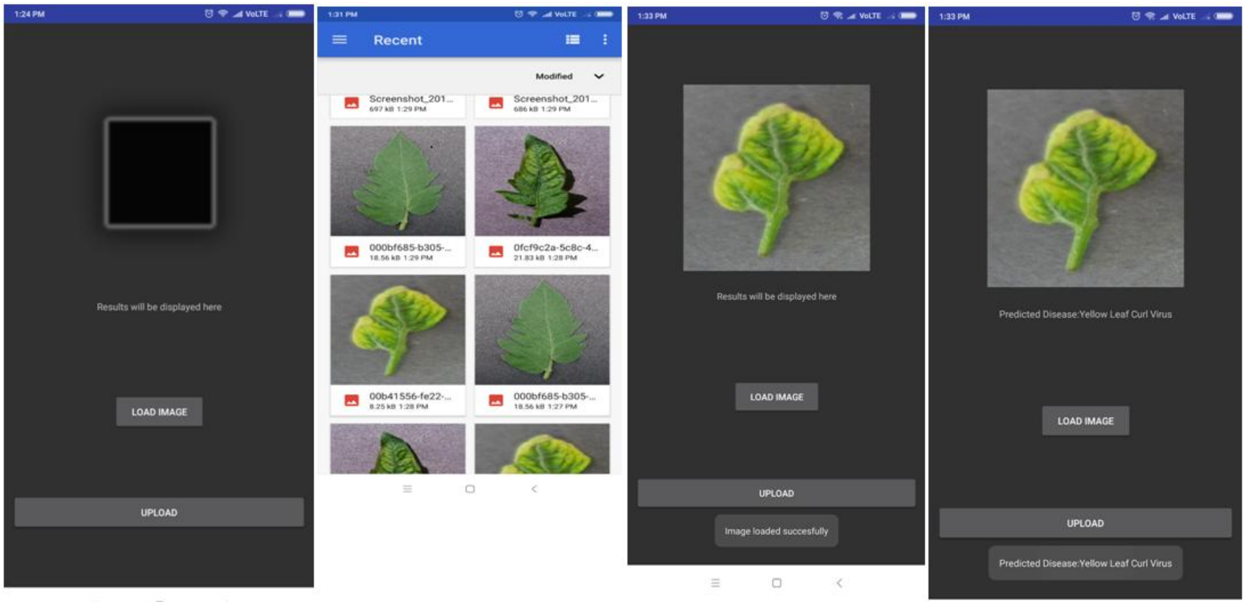

Figure 9. Complete view and flow of the Mobile Application

It has two buttons and an imageview where images are displayed.
Load Image - This button is used to load an image from gallery. This image is then displayed on imageview. Choose an image from gallery, where you select an image for which you want to test for a disease.
Upload - This button sends the loaded image to the Flask Based server and returns the results and displays them in result space. Click *Upload* to send the selected image to server for further analysis and classification results. Figure 10 depicts snapshots of classification results for healthy as well as Tomato mosaic virus infected leaf.

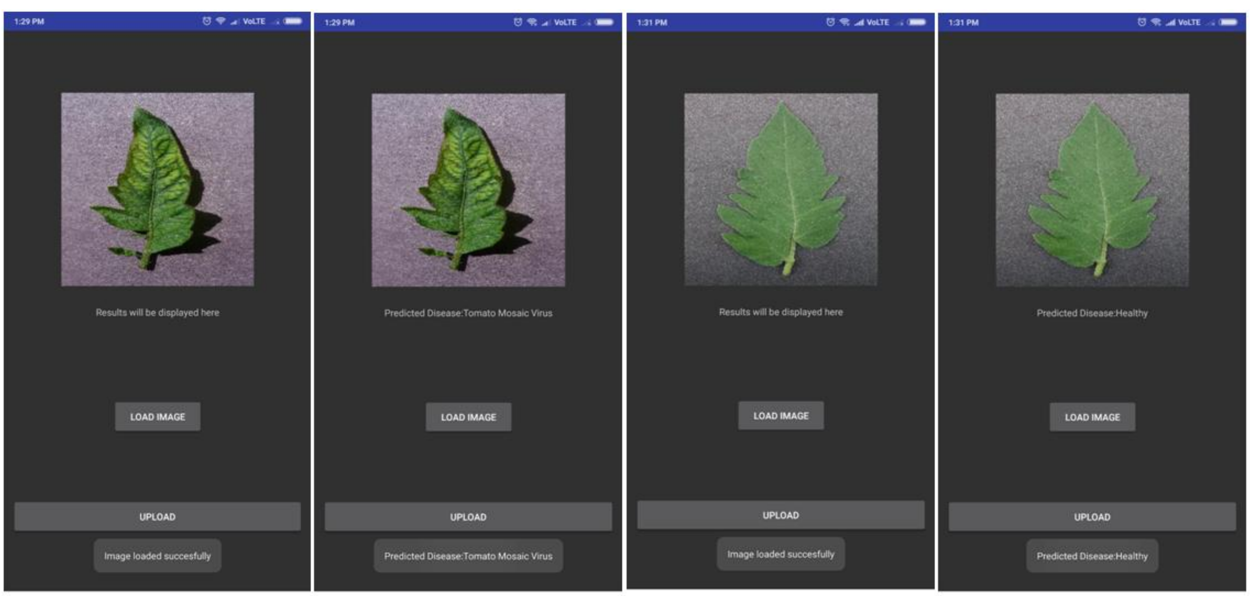

Figure 10. Few snapshots of Classification results

Tech-Stack Used

1. Android Studio - The android application was made from scratch in android studio. This app is compatible with nearly 100% android devices. The primary programming language used was JAVA and XML.

2. Volley - It was used in our application for sending image to our Flask based server in base64 String format as a request, and receiving a response, which consists of predictions made after processing of the input image by the server.

3. Flask Application - The authors have created an API for our machine learning model, using Python along with the lightwork framework Flask. This API will act as an access point for the model across Flask, used to deploy the Resnet-50 Model, trained on images of tomato leaves, as an API service for handling the requests to the server. It receives a POST request with an image in base64 string format sent via Volley from android application and in turn sends a response back to the android application in JSON format which is then displayed on the app as the final result.

Final Note

1. This app can be later launched on PlayStore for people so that anyone can use it and get its benefits.
2. The Flask API is currently deployed on local server for testing purposes, but can be scaled for heavier usage at later stages, by deploying on servers like AWS, Google Cloud Services, Heroku, etc. so that it can handle a large number of requests at same time.

FUTURE RESEARCH DIRECTIONS

The mobile application is in its initial stages of development. The authors aim to replicate it for all crucial crops in the Indian agricultural domain and distribute it for everyday usage. As mentioned earlier, tomato leaf images from the PlantVillage dataset have been utilized for training purposes, as the dataset is openly available and that too in large numbers which is essential for training a deep learning architecture. The authors aim to build a similar database that would be open access as well, and include other parts of a plant in addition to leaves viz. fruit, stem etc. Work needs to be done in order to investigate the visual symptoms appearing on different parts of a plant along with the challenge of capturing images of hidden plants and their diseased elements.

Among other imaging sensors, hyperspectral data and its applications in agriculture, are being researched extensively and combined with deep learning, it can prove to be monumental in differentiating among healthy and diseased plants. Spectral data measurements obtained from the hyperspectral sensors are utilized for computation of vegetation indices (VIs), which hint at the biophysical and biochemical properties of a plant, hence, enabling early disease prediction. These values are used to determine the stress levels and/or disease severity level in a plant. Internet of Things (IoT) based sensors can be supplemented and deployed along with imaging sensors for data collection and subsequent analysis. Implementation of a model, where the inputs are provided by the sensor nodes and intelligent decision-based applications are deployed using machine learning algorithms, a well designed “Smart Plant Disease Surveillance System” could be demonstrated on accurate real time field data. The authors intend to develop a smart system in which various environmental parameters such as temperature, humidity, pressure, CO_2, water level etc. along with images, can be observed and timely action can be taken to prevent any hazards. Images could be captured via the Digital, Infrared and Hyperspectral cameras, and various sensors deployed in the field would monitor the environmental factors. The study would include different parts of the plant where the symptoms might occur. Once the disease has been diagnosed, the concerned would be notified and treatment can be suggested via any communication medium, also basic treatment steps can be incorporated in the mobile application itself.

CONCLUSION

Due to the rapid growth in the overall computational power and researchers constantly striving for faster and accurate problem solving approaches, there is a plethora of non-destructive, non-invasive techniques for disease detection in agriculture. The authors aim to bring these technologies to the field and bridge the gap between experimentation and real life application. This chapter showcased the recent advancements in applications of statistical methods and sensor based technologies to aid in plant disease diagnosis, followed by the implementation details and comparative results of CNN architectures namely Inception ResNet V2, ResNet50, VGG16, VGG19, Xception along with demonstrating a mobile application that can help in timely prediction of diseases in Tomato plant. Precision agriculture and smart farming technologies can assist and provide guidance and support to farmers with the aim to minimize losses, economically as well as ecologically. Hence, this requires for the existing methods to reach the farmers effectively, to be user-friendly, secure and accurate.

REFERENCES

Al-Bashish, D., Braik M. , & Bani-Ahmad, S. (2011). Detection and classification of leaf diseases using K means-based segmentation and neural-networks-based classification, Inform. Technol. J.,10: 267-275.

Brahimi, Mohammed, Boukhalfa, Kamel, & Moussaoui, Abdelouahab (2017). Deep Learning for Tomato Diseases: Classification and Symptoms Visualization, Applied Artificial Intelligence, Vol. 31, No. 4, 299–315, DOI: 10.1080/08839514.2017.1315516.

Camargo, A. & Smith, J.S. (2009). An image-processing based algorithm to automatically identify plant disease visual symptoms, Biosystems Engineering 102, 9-12.

Camargo, A. & Smith, J.S. (2009). Image pattern classification for the identification of disease causing agents in plants. Computers and Electronics in Agriculture, 66, 118–125.

Chung, Chia-Lin, Huang, Kai-Jyun, Chen, Szu-Yu, Lai, Ming-Hsing, Chen, Yu-Chia, & Kuo, Yan-Fu (2016). Detecting Bakanae disease in rice seedlings by machine vision. Computers and Electronics in Agriculture, 121, 404-411.

Durmus, Halil, Günes, Ece Olcay, & Kirci, Mürvet (2017). Disease Detection on the Leaves of the Tomato Plants by Using Deep Learning, Proceedings of 6th International Conference on Agro-Geoinformatics, DOI: 10.1109/Agro-Geoinformatics.2017.8047016.

Ferentinos, Konstantinos P. (2018). Deep learning models for plant disease detection and diagnosis. Computers and Electronics in Agriculture, 145, 311–318.

Food and Agriculture Organization of the United Nations, http://www.fao.org/faostat/en/#country/100.

Ghaffari, Reza, Zhang, Fu, Iliescu, Daciana, Hines, Evor, Leeson, Mark, Napier, Richard, & Clarkson, John (2010). Early Detection of Diseases in Tomato Crops: An Electronic Nose and Intelligent Systems Approach", Proceedings of the International Joint Conference on Neural Networks, PP 1 - 6. 10.1109/IJCNN.2010.5596535.

Heckmann, David, Schluter, Urte, & Weber, Andreas P.M. (2017). Machine Learning Techniques for Predicting Crop Photosynthetic Capacity from Leaf Reflectance Spectra. Molecular Plant, Cell Press, 10, 878–890 (2017).

Huang Kuo-Yi (2007). Application of artificial neural network for detecting Phalaenopsis seed-ling diseases using color and texture features. Computers and Electronics in Agriculture 57, 3–11.

Kaundal, Rakesh, Kapoor, Amar S, & Raghava, Gajendra PS (2006). Machine learning techniques in disease forecasting: a case study on rice blast prediction. BMC Bioinformatics 7(485), 1-16.

Liu, Lianzhong, Zhang, Wu, Shu, Shuangbao, & Jin, Xiu (2013). Image Recognition of Wheat Disease Based on RBF Support Vector Machine. International Conference on Advanced Computer Science and Electronics Information (ICACSEI 2013), pp 307-310, Published by Atlantis Press, China.

Lowe, Amy, Harrison, Nicola, & French Andrew P (2017). Hyperspectral image analysis techniques for the detection and classification of the early onset of plant disease and stress. Plant Methods, 13(80), 1-12.

Lu, Jinzhu, Zhou, Mingchuan, Gao, Yingwang, & Jiang, Huanyu (2017). Using hyperspectral imaging to discriminate yellow leaf curl disease in tomato leaves. Precision Agriculture (online), Springer.

Mahlein, Anne-Katrin (2016). Plant Disease Detection by Imaging Sensors – Parallels and Specific Demands for Precision Agriculture and Plant Phenotyping. Plant Disease, APS Jour-nals,Volume 100, Number 2, 241-251.

Mohanty, S.P., Hughes, D.P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. Frontiers in Plant Science, Volume 7, Article 1419, 1010.

Osroosh, Yasin, Khot, Lav R. , & Peters, R. Troy (2018). Economical thermal-RGB imaging system for monitoring agricultural crops, Computers and Electronics in Agriculture, 147, 34-43.

Pal, Tarun, Jaiswal, Varun, & Chauhan, Rajinder S. (2016). DRPPP: A machine learning based tool for prediction of disease resistance proteins in plants, Computers in Biology and Medicine, http://dx.doi.org/10.1016/j.compbiomed.2016.09.008

Petrellis, Nikos (2015). Plant Disease Diagnosis Based on Image Processing, Appropriate for Mobile Phone Implementation. Proceedings of 7th International Conference on Information and Communication Technologies in Agriculture, Food and Environment (HAICTA-2015), Greece.

PlantVillage homepage: https://plantvillage.org/

Raza, S. A., Prince, G., Clarkson, J.P., & Rajpoot, N.M. (2015). Automatic detection of diseased tomato plants using thermal and stereo visible light images. PLoSONE 10:e0123262. doi:10.1371/journal.pone.0123262.

Rumpf, T., Mahlein, A.-K., Steiner, U., Oerke, E.-C., Dehne, H.-W., & Plümer L. (2010). Early detection and classification of plant diseases with Support Vector Machines based on hyperspectral reflectance. Computers and Electronics in Agriculture 74, 91–99.

Rupanagudi, Sudhir Rao, Ranjani, B. S., Nagaraj, Prathik, Bhat, Varsha G, & Thippeswamy, G (2015). A Novel Cloud Computing based Smart Farming System for Early Detection of Borer Insects in Tomatoes", 2015 IEEE International Conference on Communication, Information & Computing Technology (ICCICT), Jan. 16-17, Mumbai, India.

Sabrol, H. & Kumar, S. (2016). Intensity based feature extraction for tomato plant disease recognition by classification using decision tree. International Journal of Computer Science and Information Security (IJCSIS), Vol. 14, No. 9, 622-626.

Sankaran, Sindhuja, Mishra, Ashish, Ehsani, Reza, & Davis, Cristina (2010). A review of advanced techniques for detecting plant diseases. Computers and Electronics in Agriculture 72, 1–13.

Sankaran, Sindhuja, Mishra, Ashish, Maja, Joe Mari, & Ehsani Reza (2011). Visible-near infrared spectroscopy for detection of Huanglongbing in citrus orchards. Computers and Electronics in Agriculture 77, 127–134.

Shankar, Ravi, Harsha, Seema, & Bhandary, Raj (2014). A Practical Guide To Identification And Control Of Tomato Diseases, Tropica Seeds Pvt.Ltd.

Singh, Chandra B., Jayasa, Digvir S., Paliwala, J., &White, Noel D.G. (2010). Identification of insect-damaged wheat kernels using short-wave near-infrared hyperspectral and digital colour imaging. Computers and Electronics in Agriculture, 73, 118–125.

Singh, V. & Misra, A.K. (2017). Detection of Plant Leaf Diseases Using Image Segmentation and Soft Computing Techniques, Information Processing in Agriculture 4, 41-49.

Wang, Guan, Sun, Yu, & Wang, Jianxin (2017). Automatic Image-Based Plant Disease Severity Estimation Using Deep Learning. Hindawi Computational Intelligence and Neuroscience Volume 2017, Article ID 2917536, 1-8.

Wang, X., Zhang , M., Zhu, J., & Geng, S. (2008). Spectral prediction of Phytophthora infestans infection on tomatoes using artificial neural network (ANN), International Journal of Remote Sensing, Vol. 29, No. 6, PP 1693–1706.
Xie, Chuanqi, Shao, Yongni, Li, Xiaoli, & He, Yong (2015). Detection of early blight and late blight diseases on tomato leaves using hyperspectral imaging. Scientific Reports, 5:16564, DOI: 10.1038/srep16564, 1-11.
Xu , H.R., Ying , Y.B., Fu, X.P., & Zhu, S.P. (2007). Near-infrared Spectroscopy in detecting Leaf Miner Damage on Tomato Leaf, Biosystems Engineering 96 (4), 447–454, doi:10.1016/j.biosystemseng.2007.01.008.
Xu, Huirong, Zhu, Shengpan, Ying, Yibin, & Jiang, Huanyu (2006). Application of multispectral reflectance for early detection of tomato disease, Proceedings Volume 6381, Optics for Natural Resources, Agriculture, and Foods; 63810R (2006); doi: 10.1117/12.685531.
Ye, Xujun, Sakai, Kenshi, Okamoto, Hiroshi, & Garciano, Leroy O. (2008). A ground-based hyperspectral imaging system for characterizing vegetation spectral features, Computers and Electronics in Agriculture 63, 13–21.
Zhang, Minghua, Qin, Zhihao, & Liu, Xue (2005). Remote Sensed Spectral Imagery to Detect Late Blight in Field Tomatoes, Precision Agriculture, 6, 489–508, Springer Science+Business Media, Inc. Manufactured in The Netherlands.
Zhang, Minghua, Qin, Zhihao, Liu, Xue, & Ustin, Susan L. (2003). Detection of stress in tomatoes induced by late blight disease in California, USA, using hyperspectral remote sensing, International Journal of Applied Earth Observation and Geoinformation, 4, PP 295–310.
Zhao, Yan-Ru, Li, Xiaoli, Yu, Ke-Qiang, Cheng, Fan, & He, Yong (2016). Hyperspectral Imaging for Determining Pigment Contents in Cucumber Leaves in Response to Angular Leaf Spot Disease. Scientific Reports, 6:27790, DOI: 10.1038/srep27790, 1-9.

ADDITIONAL READING

Chollet, Francois (2017). Xception: Deep Learning with Depthwise Separable Convolutions, arXiv:1610.02357v3 [cs.CV].
Golhani, Kamlesh, Balasundrama, Siva K., Vadamalai, Ganesan, &Pradhan, Biswajeet (in press). A review of neural networks in plant disease detection using hyperspectral data. Info Proc Agri (2018), https://doi.org/10.1016/j.inpa.2018.05.002.
He, Kaiming, Zhang, Xiangyu, Ren, Shaoqing, & Sun, Jian (2015). Deep Residual Learning for Image Recognition, arXiv:1512.03385[cs.CV].
Lee, Sue Han, Chan, Chee Seng, Mayo, Simon Joseph, & Remagnino, Paolo (2017). How deep learning extracts and learns leaf features for plant classification. Pattern Recognition 71, 1–13.
Rangaswami, G., & Mahadevan, A. (2016). Diseases of Crop Plants in India, PHI Learning Pvt. Ltd, Fourth Edition.
Simonyan, Karen & Zisserman, Andrew (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition, Published as a conference paper at ICLR 2015.
Simonyan, Karen & Zisserman, Andrew (2014). Two-Stream Convolutional Networks for Action Recognition in Videos, arXiv:1406.2199[cs.CV].
Szegedy, Christian, Ioffe, Sergey, & Vanhoucke, Vincent (2016). Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning, arXiv:1602.07261v2 [cs.CV].

KEY TERMS AND DEFINITIONS

Artificial Neural Networks (ANN): It is a computational technique inspired by the human brain. It consists of nodes/neurons and connections also known as synapses, between them to exchange and transfer data. The network learns automatically according to the flow of the data.

Convolutional Neural Networks (CNN): It is a subtype of ANNs with a collection of deep FFNNs, utilized for image analysis and classification generally.

Support Vector Machine (SVM): It is a supervised machine learning tool utilized for data analysis, regression and classification.

Image Processing: Image processing is the area of computer science that deals with the analysis, enhancement and manipulation of digital images for feature extraction, recognition and classification purposes.

Precision: It is a statistical measure of random errors. It is the ratio of valid outputs also known as true positives to retrieved samples only.

Recall: Similar to precision, it is also a statistical measure**.** It is the ratio of valid outputs to total number of relevant samples.

F1 Score: Given Precision and Recall values, F1 score is computed as the harmonic mean of both.

Plant Pathology: It is the study of diseases in plants and their causal factors such as environmental conditions, pathogens etc., that affect the overall plant growth.