# Visual Phytopathometry:
# Ordinal Classification of Diseases in Plants using Computer Vision

**Ashish Gupta** gupta.637@osu.edu

**Alper Yilmaz** yilmaz.15@osu.edu

*Photogrammetric Computer Vision Laboratory*
*The Ohio State University*
*2070 Neil Ave, 470 Hitchcock Hall*
*Columbus, OH 43210*

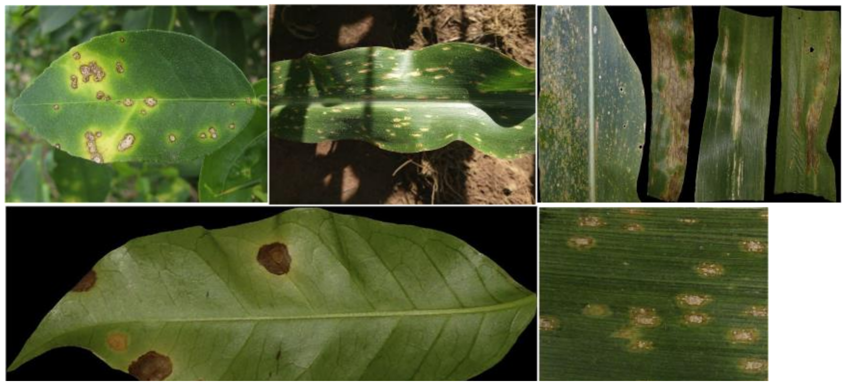
**Abstract**

Constant monitoring of plants is important to detect, classify and measure incidences of diseases to support preventive practice. Manual inspection or laboratory biochemical analysis are time consuming. Images of plants can be analyzed sufficiently quickly for automated disease identification, and several approaches have been researched in the past two decades. These approaches currently work in restrictive conditions like absence of background visual clutter, visually discernible diseased plant tissue and manual configuration of visual features adapted to different species of plants and types of diseases. We propose an automated disease identification and measurement approach that builds on computer vision and machine learning techniques that have proven to be very effective in accurate classification tasks when dealing with very large and complex visual data. In particular, we use Deep Learning which a couple of years ago outperformed all other visual object classification methods and has since then been successfully employed in several different visual classification tasks. Successive layers of a Convolutional Neural Network (CNN) can be used to represent plants images at different spatial scales, which range from simple localized image structures like corners and edges to entire object. Such a representation affords us the ability to be robust to variations in acquired images like rotation, scale and projective transform, ambient light conditions, partial occlusions, etc. We use CNN to automatically learn relevant features of healthy and diseased plant images. CNNs can also be employed for effective semantic segmentation of plant images to ameliorate the issue of background clutter. To measure the severity of disease in plant on an ordinal scale we augment CNN with Support Vector Ordinal Regression. We use a recently published plant disease image dataset that has over 50,000 samples and a variety of plant and disease types to train and validate our approach.

**Keywords:** Plant disease, Visual Pathometry, Computer Vision, Deep Learning, Convolutional Neural Networks, Ordinal classification

## 1 Introduction

Automatic plant disease identification and intensity measurement using visible spectral range images has received attention in the last two decades, however the techniques proposed so far seem typically limited in their scope and mostly dependent on ideal capture conditions to properly function. This apparent lack of significant advancements may be partially explained by some difficult challenges posed by this task. Our focus is on digital images of symptoms in the visible spectral band since although multi and hyperspectral images carry more information, they are usually captured by expensive and bulky sensors, while conventional cameras are ubiquitous. More information on multi and hyperspectral imaging applied to plant pathology can be found in (Sankaran et al. 2010). Some of the significant challenges are:



- The background often contains elements that can make it very difficult to correctly segment the

**Figure 1 Challenges to phytopahtometry using image analysis in visible spectrum**

region of interest where the symptoms are present.
- Capture conditions are difficult to control, which may cause the images to present characteristics that are difficult to predict and make the disease identification more challenging, like specular reflections and shadows.
- Most symptoms do not have well defined boundaries, rather gradually fading into normal tissue, making it difficult to clearly define which are the healthy and diseased regions.
- A given disease may possess very different characteristics depending on its stage of development, and sometimes on where it is located on the plant.
- Symptoms produced by different diseases may be visually similar, which forces the methods to rely on very tenuous differences to discriminate among them.

## 2 Methods for disease detection

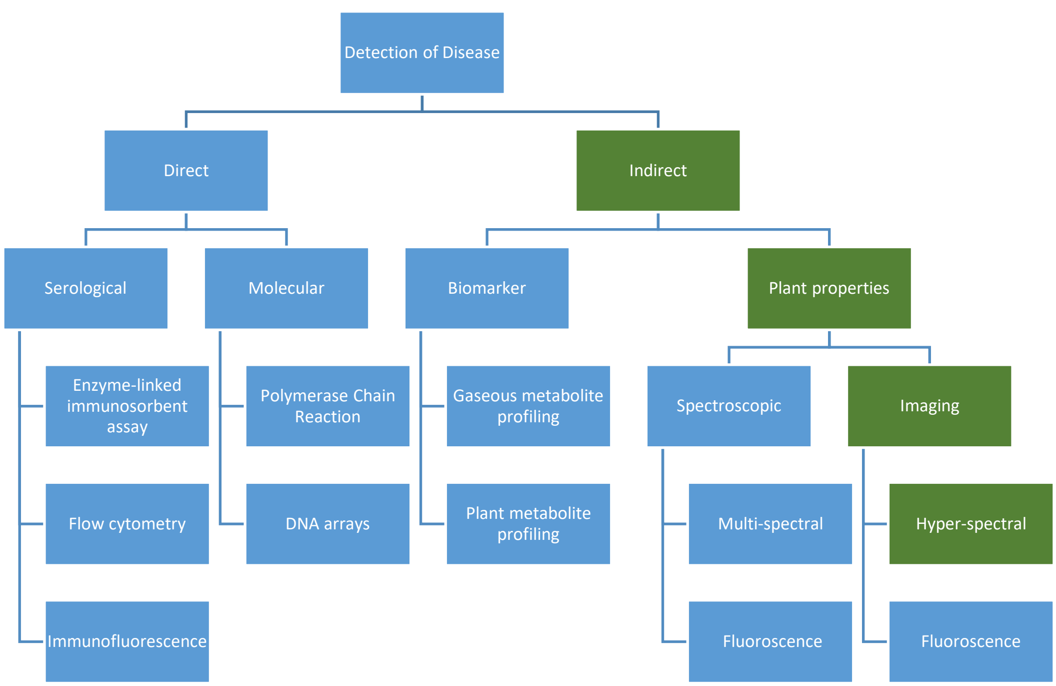



**Figure 2 Analysis of hyper-spectral images of healthy and diseased plants is an indirect method of detecting disease in plant.**

## 3 Multi-Spectral Analysis

Visual analysis of Phytopathometry should typically utilize multi-spectral analysis since the symptoms of pathogens on plants can be sensed beyond the visible spectrum. Many fungal pathogens exhibiting necrotrophic lifestyle rely on an active degradation of host plant tissues, causing in consequence, well visible disease symptoms, referred to as necrosis. Other pathogens, like biotrophic, may trigger intense activation of defense mechanism known as hypersensitive response (HR). Both necrosis and HR result in a change of leaf or other tissue appearances. Those morphological differences can be easily visualized using the visible spectrum. Necrosis and HR are however, not the only possible outcome of a pathogen attack. Upon recognition of a pathogen plants may close their stomata and therefore restrict the access to mesophyll tissue, which results in an increase of leaf temperature. Those differences can be assessed using infrared (IR) imaging. In the same manner, pathogens affecting plant metabolism can influence the content of plants chlorophyll and other pigments, which in turn changes the plants' auto-fluorescence and can be visualized using the near-UV spectrum imaging. The plant physiology and their reactions to pathogen attack predicates the type of sensors and spectral band we choose for our Phytopathometry.

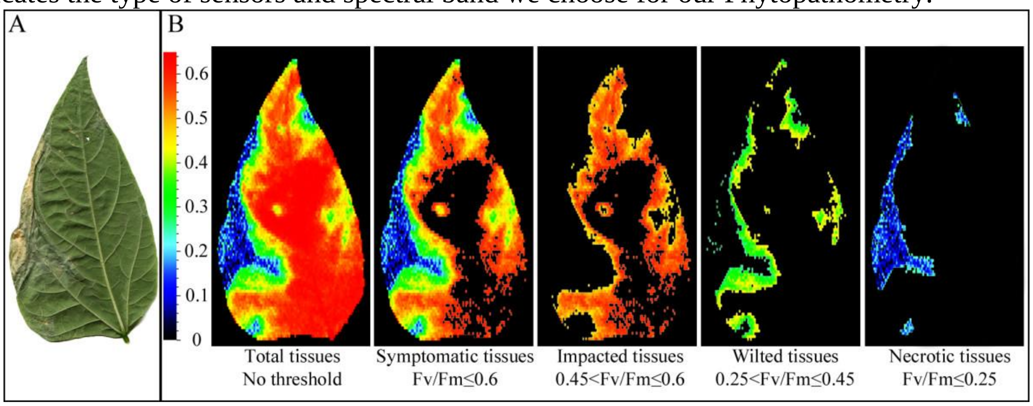



**Figure 3 Type and extent of pathogenic induced decay in leaf image(A). The $\frac{F_V}{F_M}$ image (B) is obtained by chlorophyll fluorescence images. We use image segmentation to separate different parts of leaf exhibiting different symptoms to pathogenic activity.**

## 4 Variety of plants and diseases

The variety of plants and diseases presents a challenge to learning visual features and classifier for diseased plant tissues. This problem is further compounded by the pathogen attacked tissue exhibiting different symptomatic features at different stages of its development. Towards modeling this we utilize an open crowd-sourced dataset *PlantVillage* dataset (Hughes and Marcel Salathe 2015). A study using 50,304 images of 14 crop species and 26 diseased/healthy labeled disease classes from this dataset has been reported in (Mohanty, Hughes, and Salathé 2016). An example of this dataset is shown in Figure

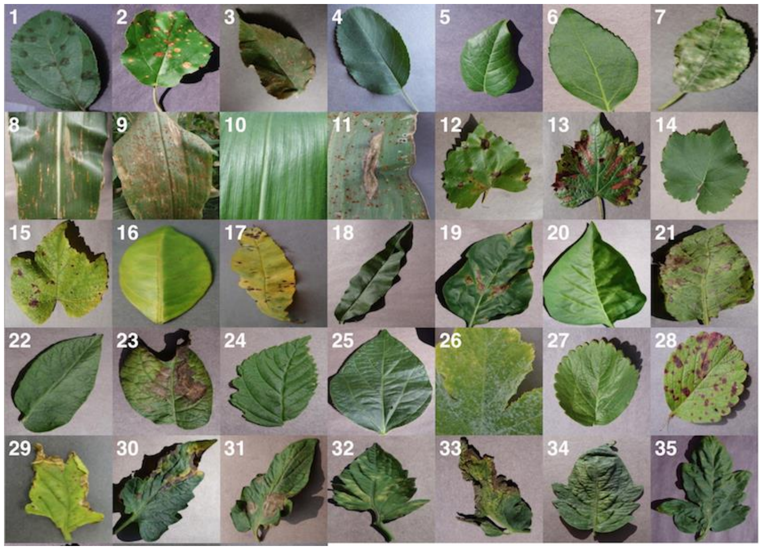



**Figure 4 Example of leaf images from the PlantVillage dataset, representing crop-disease pairs. 1) Apple Scab, Venturia inaequalis 2) Apple Black Rot, Botryosphaeria obtusa 3) Apple Cedar Rust, Gymnosporangium juniperi-virginianae 4) Apple healthy 5) Blueberry healthy 6) Cherry healthy 7) Cherry Powdery Mildew, Podosphaera spp. 8) Corn Gray Leaf Spot, Cercospora zeae-maydis 9) Corn Common Rust, Puccinia sorghi 10) Corn healthy 11) Corn Northern Leaf Blight, Exserohilum turcicum 12) Grape Black Rot, Guignardia bidwellii, 13) Grape Black Measles (Esca), Phaeomoniella aleophilum, Phaeomoniella chlamydospora 14) Grape Healthy 15) Grape Leaf Blight, Pseudocercospora vitis 16) Orange Huanglongbing (Citrus Greening), Candidatus Liberibacter spp. 17) Peach Bacterial Spot, Xanthomonas campestris 18) Peach healthy 19) Bell Pepper Bacterial Spot, Xanthomonas campestris 20) Bell Pepper healthy 21) Potato Early Blight, Alternaria solani 22) Potato healthy 23) Potato Late Blight, Phytophthora infestans 24) Raspberry healthy 25) Soybean healthy 26) Squash Powdery Mildew, Erysiphe cichoracearum, Sphaerotheca fuliginea 27) Strawberry Healthy 28) Strawberry Leaf Scorch, Diplocarpon earlianum 29) Tomato Bacterial Spot, Xanthomonas campestris pv. vesicatoria 30) Tomato Early Blight, Alternaria solani 31) Tomato Late Blight, Phytophthora infestans 32) Tomato Leaf Mold, Fulvia fulva 33) Tomato Septoria Leaf Spot, Septoria lycopersici 34) Tomato Two Spotted Spider Mite, Tetranychus urticae 35) Tomato Target Spot.**

## 5 Deep Learning

Computer vision, and object recognition in particular, has made tremendous advances in the past few years. The PASCAL VOC Challenge, and more recently the Large Scale Visual Recognition Challenge (ILSVRC) based on the ImageNet dataset have been widely used as benchmarks for numerous visualization-related problems in computer vision, including object classification. In 2012, a large, deep convolutional neural network achieved a top-5 error of 16.4\% for the classification of images into 1,000 possible categories. In the following three years, various advances in deep convolutional neural networks lowered the error rate to 3.57\%. While training large neural networks can be very time-consuming, the trained models can classify images very quickly. In order to develop accurate image classifiers for the purposes of plant disease diagnosis, we needed a large, verified dataset of images of diseased and healthy plants. Until very recently, such a dataset did not exist, and even smaller datasets were not freely available. To address this problem, the PlantVillage project has begun collecting tens of thousands of images of healthy and diseased crop plants, and has made them openly and freely available. Here, we report on the classification of 26 diseases in 14 crop species using 54,306 images with a convolutional neural network approach.

Previously, the traditional approach for image classification tasks has been based on hand-engineered features such as SIFT, HoG, SURF, etc., and then to use some form of learning algorithm in these feature spaces. This led to the performance of all these approaches depending heavily on the underlying predefined features. Feature engineering itself is a complex and tedious process which needed to be revisited every time the problem at hand or the associated dataset changed considerably. This problem has occurred in all traditional attempts to detect plant diseases using computer vision as they leaned heavily on hand-engineered features, image enhancement techniques, and a host of other complex and labor-intensive methodologies. A few years ago, AlexNet showed for the first time that end-to-end supervised training using a deep convolutional neural network architecture is a practical possibility even for image classification problems with a very large number of classes, beating the traditional approaches using hand-engineered features by a substantial margin in standard benchmarks. The absence of the labor-intensive phase of feature engineering and the ability to generalize the solution makes them a very promising candidate for a practical and scalable approach for computational inference of plant diseases.

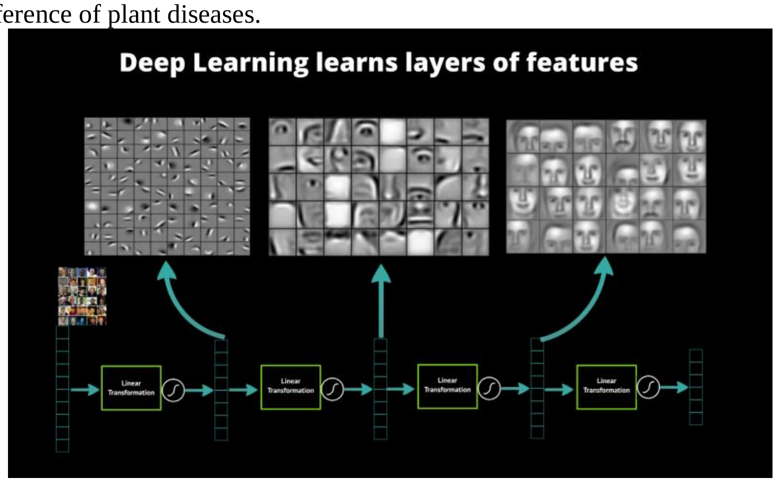



**Figure 5 A Deep Neural Network (DNN) learns features from a large corpus of training images in multiple layers. Each successive layer represents features with greater spatial extent and semantic meaning, thereby providing low-level, mid-level, and high-level features. In this figure we use face images as an illustrative example. However, any visual entity, including plant and leaf images can be used with DNN.**

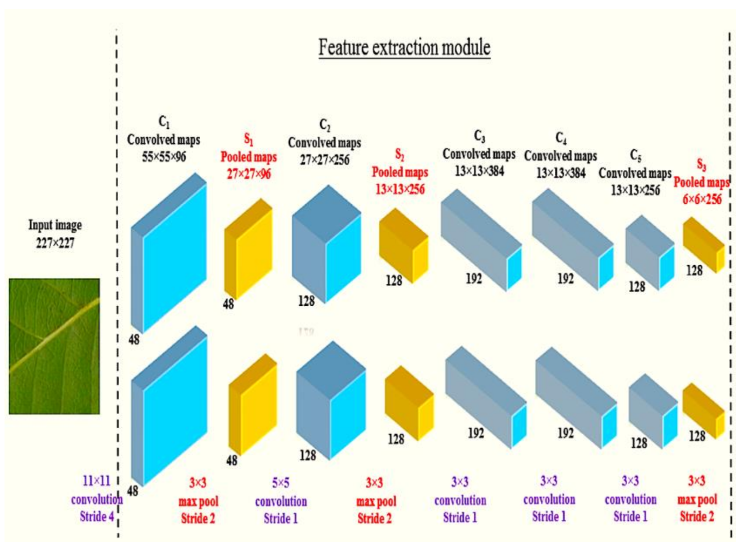



**Figure 6 Learning plant image features automatically using Convolutional Neural Network (CNN). Different layers of CNN learn plant features at different spatial scales so that the classifier can identify diseases tissue in plant images of different spatial extents.**

## 6 Image Segmentation

Image segmentation seeks to split an image into foreground and background, where foreground is comprised of regions of interest (RoI) in the image. Image segmentation is particularly important in images acquired in the field (as opposed to laboratory conditions with high contrast background), where background clutter introduces significant challenges to the image analysis algorithms to be applied to the visual content in the foreground. We leverage the expressive power of CNNs to segment plant images.

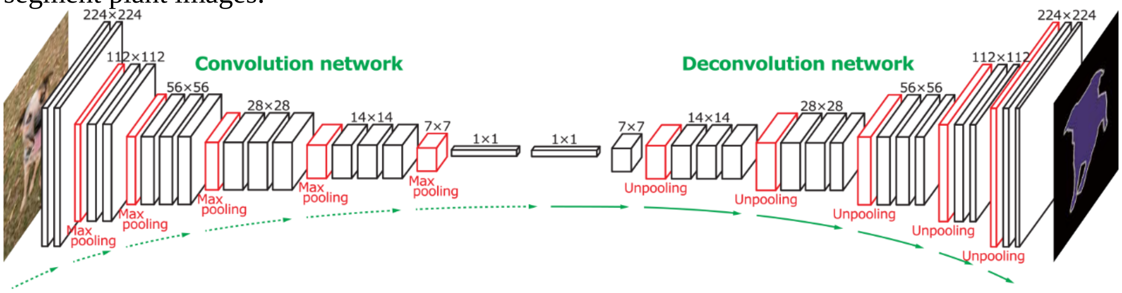



**Figure 7 Convolutional Neural Network for semantic segmentation. The CNN learns associations between semantic classes ({leaf}, {diseases}) at multiple scales from coarse to fine. It can distinguish between two adjacent image pixels that belong to different classes (healthy plant and diseased plant) and consequently can achieve remarkably high quality of segmentation.**

## 7 Summary

Our objective is to design and develop a Deep Neural Network (DNN) computational framework towards automatically learning image features in the visible and multi-spectral bands. These learned features is used in our Convolutional Neural Network (CNN) towards effective segmentation of (i) relevant parts of plants (leaf, stem) acquired image and (ii) diseased tissues from healthy tissues and other types of diseased tissues. These features are also used with a Convolutional Neural Network with an output layer using Support Vector Ordinal Machine that pools results of the CNN to compute intensity of pathogenic influence. Importantly a DNN learns relevant features on its own, which obviates the need to hand engineer features for different types of plants and diseases. Once the framework is built and tested for a few types of plants and diseases, it can be reliably extended to self-train and classify other types of plant and associated disease.

## 8 Bottomline

Computer Vision provides the expertise to quickly analyze images acquired in the field and deal with the associated issues like noise, background clutter, occlusions, shadows, poor lighting, etc.
Deep Learning is a revolutionary new approach that has delivered improvement in learned visual models significantly beyond next best methods, which has renewed interest in exploring viability of using an indirect approach to plant disease detection using image analysis.
Hyper-spectral image analysis can potentially detect disease affliction in the early stages that can benefit timely pest and disease management.

## References

Hughes, David, and Marcel Salathe. 2015. "An Open Access Repository of Images on Plant Health to Enable the Development of Mobile Disease Diagnostics." pennsylvania.

Mohanty, Sharada Prasanna, David Hughes, and Marcel Salathé. 2016. "Using Deep Learning for Image-Based Plant Disease Detection." Switzerland.

Sankaran, Sindhuja, Ashish Mishra, Reza Ehsani, and Cristina Davis. 2010. "A Review of Advanced Techniques for Detecting Plant Diseases." *Computers and Electronics in Agriculture* 72 (1): 1–13. doi:10.1016/j.compag.2010.02.007.