



High-Performance Deep learning to Detection and Tracking Tomato Plant Leaf Predict Disease and Expert Systems

Azhee Wria Muhamad^a, Yaser A. Jasim^b, Mostafa Ismael Shukri Windi^c, Mustafa Ghanem Saeed^d, Sadeeer Dhyaa AbdulAmeer^d

^a Computer Science Department, College of Basic Education, University of Sulaimani, Iraq

^b Accounting Department, College of Administrative and Financial Science, Cihan University-Erbil, Kurdistan Region, Iraq

^c Software Engineering Department, College of Computer Sciences and Mathematics, Mosul University, Mosul, Iraq

^d Computer science Department, College of Science, Cihan University, Slimani, Iraq

Computer science Department, College of Science, Cihan University, Slimani, Iraq
azhee.muhamad@univsul.edu.iq, Yaser.jasim@cihanuniversity.edu.iq, mostafa.i.windi@uomosul.edu.iq, mustafa.saeed@sulicihan.edu.kr, sadeer.dheyaa@sulicihan.edu.kr

KEYWORDS ABSTRACT

Deep Learning, Plant Disease, Expert Systems, Artificial Intelligence, ANN

Nowadays, technology and computer science are rapidly developing many tools and algorithms, especially in the field of artificial intelligence. Machine learning is involved in the development of new methodologies and models that have become a novel machine learning area of applications for artificial intelligence. In addition to the architectures of conventional neural network methodologies, deep learning refers to the use of artificial neural network architectures which include multiple processing layers.

In this paper, Convolutional neural network models were designed to detect and diagnose plant disorders by applying samples of healthy and unhealthy plant images analyzed by means of methods of deep learning. The models were trained using an open data set containing 18,000 images of ten different plants, including healthy plants. Several model architectures have been trained to achieve performance of 97 percent when the respectively [plant, disease] paired are detected. This is a very useful information or early warning technique and a method that can be further improved with the substantially high-performance rate to support an automated plant disease detection system to work in actual farm conditions.



1. Introduction

One of the most important problems facing the improvement and development in the agricultural field is the lack of modern and advanced technological methods in the identification and diagnosis of diseases that affect the plant [1], where it will spend more time diagnosing the disease which leads to the death of the plant, however, when applying programs and technologies in the analysis process of a plant images this became easier to know the type of disease and what are the methods in combating the disease and reducing its spread through the use of a trained deep convoluted neural network to adequately accurately plant leaves database collected independently for various plant diseases; and by process some symptoms and features such as pale colors that indicate particular diseases. Therefore, the quality and speed to show results that will increase the possibility to get benefit from agriculture [13].

Deep Learning strategies are used in this study, and while mass production and accessibility of food has improved several times, food safety continues to be undermined by numerous factors including more than 50 percent plant diseases [2, 21]. Fortunately, mostly with rise of technology and the mobile industry with strong technologies around the world, diagnostic data on certain types of diseases can be obtained quickly by detecting diseases until they occur on the plant to treat them [2, 3]. This analysis is based on a broad variety of new nodes (artificial neurons) used to roughly imitate the way a human brain resolves issues of large clusters of biological neurons attached to axons [12][15].

2. Related Works

Identification of plant diseases is a vital topic studied for years and is informed by the fact that nutritious are required. However, the cost efficiency, usability, sensitivity, and accuracy should be some of the elements to be taken into account [4]. Through the last decade many studies have recommended several unstructured methods to resolve these facts [5]. In the assessment of plant stress to environmental factors, a super specimen proximal sensing technique was employed, where optical technologies are realistic instruments for tracking plant health, for example [6], in the estimation of plant stress, created by increased gas, radiation, water status and insect attack; methods for thermal and fluorescence imaging have been developed to solve the problems mentioned above. In these related works sample used were huge which reduces the performance of the system and accuracy. The study of plant protection to respond to the presence of pathogens also involves another significant field, therefore, in this paper the authors will show how the software detects, diagnose and give treatment recommendations; by developing the software using expert systems to cope the issues that faced the other systems from the previous works. The proposed system has developed some features to improve the software, such as, firstly the proposed model reduces mathematical complexity and calculations via down sampling the input images from (256*256) to (64*64); secondly the convolution neural network architecture used to diagnose tomato diseases has achieved an accuracy of 98% with a limited number of layers and complexity, and thirdly the tomato disease diagnostic system has been developed with expert system to define the right treatment and the procedures to prevent infection in the future, as well as providing a friendly graphical user interface [15].

3. Plant and Diseases

The agriculture field is one of the most fundamental parts for adding to national pay in numerous nations, because of the expanded interest in the farming business, agriculture produce must be



developed utilizing current and efficient methods, where real income is low, focus is put on agriculture and other primary industries [23, 24].

Diseases and defects found in plants and yields greatly affect creation in the agriculture industry and lead to significant economic impacts as mentioned above [25], diseases must be recognized at the first stage, otherwise, it will be more damaging to productivity; while using deep learning to find out specific diseases helps to detect disease and provide appropriate treatment for specific diseases that need to prevent disease kinds of pesticides.

3.1. Phenotypes of Tomato Leaves

Each plant disease has various symptoms known by specialists and botany experts, but in this study will focus about the symptoms that appear on the leaves only because most of the symptoms appear on the leaves, which represent the very important part of plants, as shown in the figure 1 that shows examples of different phenotypes of tomato plants as following:

- A. Healthy leaf
- B. Precocious Blight (*Alternaria Solani*)
- C. Overdue Blight (*Phytophthora Infestans*)
- D. Septoria Leaf Spot (*Septoria Lycopersici*)
- E. Yellow Leaf Curl Virus (Family Geminiviridae genus Begomovirus)
- F. Bacterial Spot (*Xanthomonas Campestris Pv. Vesicatoria*)
- G. Target Spot (*Corynespora Cassiicola*)
- H. Spider Mite Damage (*Tetranychus Urticae*)

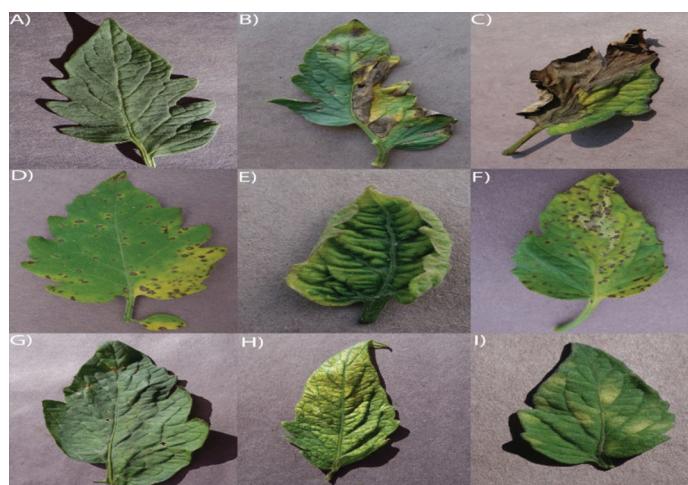


Fig. 1: Examples of different phenotypes of tomato plants

When this hit any of the diseases, and more diseases known in fruits and vegetables such as Bacteria, Fungi, Virus, Mold, and Mite diseases, the research reveals these diseases with their symptoms and pictures so that they can be identified and categorized accordingly.

3.2. Types of Plant Leaf Diseases

Plant diseases, especially those that appear on the leaves; therefore, the species will be explained separately for some of the disease types with the symptoms that characterize each disease, in addition to suggesting how to treat it and limit its spread.

A. Bacterial Diseases

Bacteria are single-celled organisms that are microscopic. By simple cell division, bacteria rapidly divide and absorb nutrients from their immediate surroundings. New sites on contaminated seeds or transplants are introduced to bacterial pathogens., bacteria persist in and around tomato plantings. Bacteria have several types of diseases such as [7]:

- Bacterial spot (*Xanthomonas Euvesicatoria*)

Bacterial spot is a popular tomato & pepper disease, where levels of rainfall and moisture promote the development of the disease. Through defoliating and spotting fruit bacterial spot decreases the yield and value of tomatoes. Through biological openings or injuries, bacteria enter plants. Leaf spots are dark brown, greasy when the leaves are wet and rarely exceed 1/8 inch in thickness figure 2.



Fig. 2 Bacterial spot (leaf spots)

Every diseased plant material must be discarded as the appropriate cure for this disease. Prune air circulation trees, spray with a fungicide of copper.

B. Fungi Disease

The plant infections are the most common group of fungus. Thousands of fungi can cause diseases of plants. Usually, microscopic are these multicellular plants. A fungus shell is made up of filament-like fibers called hyphae. Paroles of hyphae are called mycelia. Such masses can be seen without a microscope when they are big enough. Popular fungal symptoms include spots of the leaf wilts, blights, tumors, rots of the fruit, and dieback. Fungi cause a wide range of diseases such as the Septoria leaf spot, Alternaria-Solani, Corynespora and Fulvia-Fulva, these types will be briefly explained:

❖ Septoria leaf spot

Septoria is a pycnidia-producing ascomycete fungus that causes various leaf spot diseases on crops, forages and plentiful vegetables, including tomatoes, and is responsible for crop losses. Septoria capicola causes celery's late blight. It is distinguished within pycnidia via development of conidia. The symptoms involve chlorotic spots as the following figures that turn brown and necrotic as shown in figure 3.



Fig. 3 Septoria leaf spot

Do not use overhead irrigation to reduce this disease, though repetitive fungicides using chlorothalonil or copper fungicides or mancozeb will manage the disease.

❖ Early blight (Alternaria-Solani)

Alternaria Solani is a fungal pathogen that initiate a disease known as early blight in tomato. The pathogen causes unique patterned "bullseye" leaf spots and on tomato and tuber blight on potato can also cause stem spots and fruit rot. Notwithstanding the name "early" foliar symptoms usually appear on older leaves as in the following figure 4; if unchecked, then early blight may result in a significant reduction in production.



Fig. 4 Early Blight fungi disease

The following should be applied to reduce the spread of the disease, such as, use of pathogen-free seeds, using crop swap, weeding and voluntary tomato plants, and do not wet the tomato plants with irrigation water, if the disease is severe to justify chemical treat, then use fungicides.

❖ Target Spot (*Corynespora cassiicola*)

Corynespora cassiicola is a well-popular pathogenic fungus species known as a shell fungus in the *Corynesporascaceae* group. The disease is detected by leaf exposure with light centers and dark borders in the style of target-shaped spots. *Corynespora cassiicola* ring-spots symptoms on the leaves of tomato plants as shown in figure 5.



Fig. 5 target spot on the leaves of Tomato

It can be treated by spraying with protective fungicides after the first symptoms appear. Removing damaged debris to prevent it from being carried over to the next crop.

❖ Tomato Leaf Mold (*Fulvia-fulva*)

Recently, *P. fulva* is an ascomycete named *Passalora Fulva*, a non-compulsory pathogen that leads to tomato disease called as the tomato leaf mold. *P. fulva* mainly affects tomato plants, particularly leaves, and is a common illness in greenhouses, but may occur in the field. The symptoms of this infection usually occur on the leaves and on the adaxial and abaxial surfaces it transpires on both sides of the leaf. The more adult leaves are first affected and then the disease spreads to new leaves as demonstrated in figure 6.



Fig. 6 Tomato Leaf mold fungus (*Fulvia-fulva*)

In order to protect tomato plants from this disease, crop residues must be eliminated from the field. While pruning and trimming process to increase air circulation helps in controlling the disease. Avoid moistening the leaves when watering, with a preventive fungicide program.

C. Virus

Viruses are objects smaller than a single cell that cannot be seen by a light microscope. Some viruses transmit by humans, but some are mechanically transmitted by exposure to infected sap from plant injuries. Sampling and feeding behaviors of insects such as aphids, mealybugs and leafhoppers bearing viruses to infect plants in insect transmission. There are many diseases of viral plants, and there are two important types of them in the tomato plant, namely tomato pale leaf curl virus, and tomato mosaic virus.

❖ Tomato Yellow Leaf Curl Virus

The term (TYLCV) stands for tomato yellow leaf curl virus which is a Begomovirus family and Geminiviridae type of DNA virus. Here, TYLCV causes the most damaging tomato disease and severe economic losses can be seen in tropical and subtropical areas. This virus is distributed and spread by an insect vector of the Aleyrodidae family and controlled by the whitefly *Bemisia tabaci*, Hemiptera, frequently known as the silver leaf whitefly. Symptoms of TYLCV disease involves an extreme stunting, leaf downsizing, outward curled and chlorosis of leaves, as shown in figure 7.

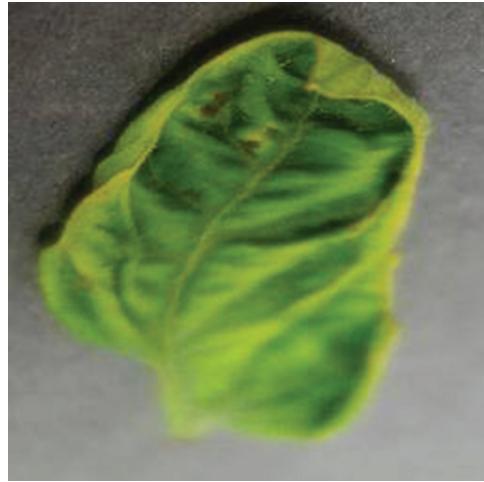


Fig. 7 Tomato Yellow Leaf Curl Virus

To reduce the spread of the disease, reflective thalamus aluminum or silver can be used in rows to reduce whitefly feeding, after removing the plants with primary symptoms.

❖ Tomato Mosaic Virus

Tomato mosaic virus (TMV) is a pathogenic virus in plants. It is present throughout the world and impacts tomatoes and several other plants. The leaves of infected tomato plants display variegation, with contrasting paled and dark green regions, the latest sometimes being deeper and higher, delivering a blister-like appearance. The leaves imply to occur fern-like with pointed hints and maybe bent with younger leaves. The fruit may be bent, pale blotches and gangrenous spots on both mature and green fruit may appear, and the fruit surface may be browned internally. The virus decreases the collection of fruit in young plants and can lead to defects and imperfections, where the whole plant can be eclipsed, as shown in figure 8.



Fig. 8 Tomato Mosaic Virus

In addition, viruses which immediately eliminate and eradicate infected plants, eliminate weed in the garden, do not have chemically controls.

D. Mites

Mites are microscopic arthropods that relate to the Arachnida class and the Acari subclass. Mites are not a clade since they cover two separate arachnid organizations: the Acariformes are sister to the spiders of the horse, while the Parasitiformes are sister to the false scorpions [7], there are several types of mites, the most famous of which is *Tetranychus urticae*.

❖ Spider mites (*Tetranychus Urticaceae*)

Tetranychus Urticaceae (popular terms involve red spider mite and two-spotted spider mite) is a plant-fed mite genus that is commonly regarded as a pestilence. Tetranychidae is the most common family member or spider mites. The genome was comprehensively sequenced in the year 2011 and was the first sequencing of any chelicerate genome. *Tetranychus urticae* is notably small, barely visible as reddish or greenish dots on stems and leaves clearly; grown-up females are approximately 0.4 mm long. The red spider mite seen in greenhouses and tropic and temperate areas spins a little web on and under leaves, as shown in the figure 9:



Fig. 9 Tomato Spider mites

It is better to use hexagon, which mainly eliminates the egg and nymphatics of two spotted spider mites without action in adults to eradicate a spider mite disease. While organophosphate insecticides (Orthene, Malathion, Dursban, and Diazinon) are not used, they appear to promote the reproduction of female spider mites [8, 9].

4. Machine Learning

Machine learning is another tool of Artificial Intelligence, that enables a computer framework to gain from the environment, through re-iterative procedures and develop itself. Machine learning methods sort out the information, gain from it, gather insights and make predictions dependent on the data it analysed without the need for extra definite programming [22]. Machine learning is applied to teach machines and how to deal with the data in an efficient manner. Some of the time after viewing the information, It is not possible to explain the pattern or focus of the data from the information. Taking into account all things, machine learning can be applied with abundant data sets accessible [10, 21]. The interest in Machine learning is in the ascent. Plentiful ventures from medication for agriculture applies machine learning to split significant data and classification for that data.

❖ Deep Learning

In term of, deep learning known as hierarchical learning or deep machine learning is a subset of machine learning constructed on specific algorithms that aim to module high-level data abstractions that uses multiple layers of complex processing with systems or otherwise consisting of many non-linear transformations. An analysis (e.g., an image) can be depicted in various approaches as a vector of intensity values for every pixel, or as a set of edges, or areas of a particular shape, etc. In a more abstract manner. Many characterizations make learning tasks easier (e.g., recognition of plants leaves diseases, recognition of the face, or recognition of facial expression) [21]. So, it is a term that includes a specific approach to neural network construction and learning. Deep learning techniques are dependent on representations that are distributed. The fundamental theory underlying distributed representations is that experimental information is produced by factor correlations structured in layers [11].

Deep learning introduces the presumption that the abstraction or composition levels refer to these layers of factors. Various layer numbers and sizes of layers can be used to provide various abstraction amounts [11]. Deep learning uses this concept of explanatory hierarchical variables in which higher-level, more abstract principles are learned from lower-level ones. Such architectures are often designed using a greedy model of layer-by-layer. Deep learning tries to disassociate these concepts and to determine what features are appropriate for learning. In supervised learning tasks, deep learning technologies prevent feature engineering, by converting the data into compressed intermediate formats akin to principal components, and obtain layered structures that eliminate redundancy in representation. Deep learning algorithms are many and the most famous of these algorithms are Artificial Neural Networks (ANNs) and Convolution Neural Network (CNN).

4.1. Convolution Neural Network (CNN)

Convolutional Neural Networks is one of the most common deep learning techniques with several layers being trained rigorously. It was found to be incredibly powerful and is still the most often used in different computer vision applications. The general CNN machine structure is shown in the figure 10.



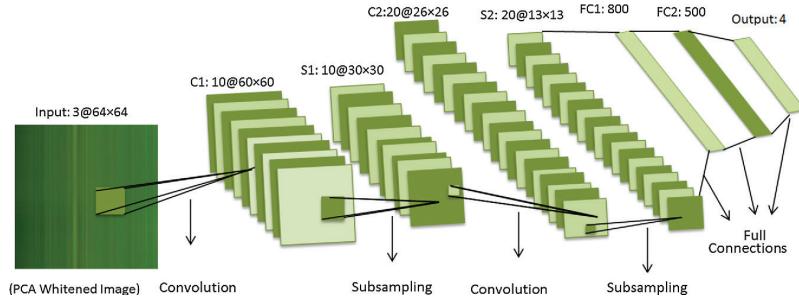


Fig. 10 The general structure of the CNN system [13]

Commonly, CNN consisting of three basic neural layers, convolutional layers, pooling layers, and layers that are fully linked, while various types of layers has different roles. As shown in figure 10, a general CNN structure for classifying image is presented layer by layer. Training the network involves two stages: forward and backward. The forward stage's main objective is to specify the input image in every layer with the proposed arguments weights and bias. Later the output of the prediction is used with the ground truth labels to calculate the cost of loss. Second, relative to the cost of loss, the backward stage determines with chain rules the gradients of per parameter. All parameters are modified and prepared for the next forward calculation depending on the gradients; after the system has a saturated various step forward and backward, then the learning network can be ended.

CNN and traditional Neural Networks (NNs) have major differences; the best is CNN, due to these differences, which are among the advantages of CNN:

1. CNN input can accommodate either 3 dimension or 2D volume and 1D based on specification and requirements although NN input is just a 1D array.
2. CNN lists more layers than regular Neural Network.
3. CNN connections are local when NN is fully connected.

❖ Basic structure of CNN

Convolution neural network is identical to artificial neural network, as both are composed of self-optimized neurons, introduced by inputs and executing non-linear transformation. Convolution neural network is commonly utilized in pattern recognition on objects relative to artificial neural networks, because it encodes image-specific features in system architecture [13]. There are five fundamental elements in the convolution neural network: an input layer, a convolution layer, a nonlinear layer, a pooling layer and a fully - connected layers.

- Input layer

The input layer contains the pixel values of the image.

- Convolution layers

A CNN network applies various kernels in the convolution layers to transform the entire object as well as the optimal feature maps, creating many feature maps as seen in figure 11.

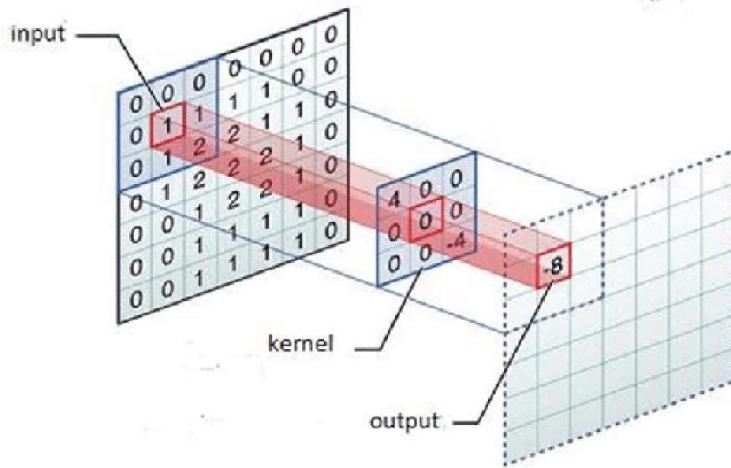


Fig. 11 Convolution layers [14]

Convolution of a function x in one dimension with a kernel w .

$$s(t) = \int Zx(a)w(t-a)da \quad (2.1)$$

Or as a discrete function

$$s(t) = \sum_{a=-\infty}^{\infty} x(a)w(t-a) \quad (2.2)$$

In another meaning a flip and gathering over the kernel, extended to two dimensions [14]:

$$s(i, j) = \sum_m \sum_n I(m, n)F(i-m, j-n) \quad (2.3)$$

In another meaning flip the filter (F) up-down and left-right and sum over all products see an example about convolution layer in figure 11. The convolution method has three key advantages:

- 1) The weight distribution function in a similar feature map reduces the number of parameters.
- 2) Spatial communication learns associations between nearby pixels.
- 3) Invariance to the object's position.

The main idea is to replace the conventional convolution layer with a small multilayer Soft max made up of multiple fully linked layers with nonlinear functions, substituting linear with nonlinear filters of neural networks. This approach produces good results when classifying images.

4.2. Non-Linear Layer (Activation Function)

A non-linear transformation or called activation function; also, is applied to the input by the convolution neural network, the object of which is to classify the features within per hidden layer. The non-linear translation function in the artificial neural network is sigmoid or hyperbolic tangent. Though, if the sparsity of the data is greater, the outcome will be better for image processing. This layer uses many types of activation functions, a most popular is Rectified Linear Units (ReLU).

4.3. Pooling Layers

Typically, a pooling layer approaches a convolution layer which can be used to minimize the dimensions of maps and network parameters. In interpretation, pooling layers are also symmetric, as are convolution layers, since their predictions take adjacent pixels in account. Average pooling and maximum pooling are the most common techniques. The following figure 12 summarizes the mechanisms of the two kinds of pooling.

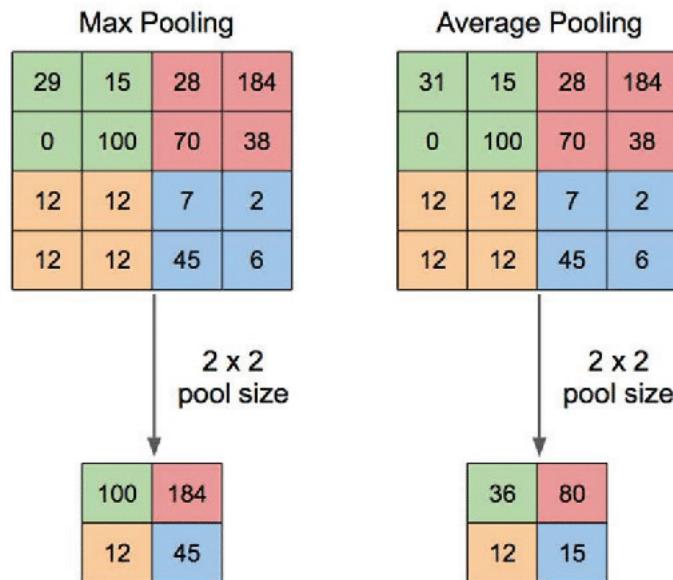


Fig. 12 Two Classic Pooling Methods [16]

The pooling layer applies the max-pooling function with a (2 x 2) kernel and a step of 2 along the spatial dimensions of the input, according to the reference cases. It is based on the care about the pooling layer's destructive functionality. Through this process, the feature map is reduced to 25% of the previous size while keeping the depth capacity at its standard size. It deducts the layer to use the intervening region to move through the whole spatial dimensionality of the data. If the stage is adjusted to 3 with a kernel size set to 3, the model's output will be effectively reduced.

4.4. Fully Connected Layer

Upon the several iterations of previous layers, the data enters the final layer of the neural network of convolution, which is also the completely connected node. Nodes are linked directly to neurons in the two adjacent layers within the fully linked network [17, 20]. The goal of this layer is to summarize the weight of the functionalities in previous layers and display the probability by class. If there is a neural convolution network of neuronal classification for gender and thus the outcome matrix is [0.8, 0.2], so there is a likelihood of 80 percent male gender and 20 percent female gender. There are also a couple of other CNN-related concepts, such as MLP and Soft max.

4.5. Training a Network

A network training is a method of obtaining distributions in a convolution layers and weights in fully connected layers to decrease discrepancies between output and specified regression coefficients labels in a training dataset [18, 19]. The backpropagation algorithm is a widely applied approach to the development of neural networks with the principle of loss and optimization of descending gradient algorithms. An efficient model in different kernels and weights is calculated by a loss function by forwarding propagation over a training dataset and by adjusting the loss value from a back-propagation.

5. Methodology

The main objective of the proposed system is to detect and treat tomato plant diseases, and the work of this system includes several sequential stages.

According to this graph, the work has been divided into several major steps in this system as follows:

The design model of the proposed system depends on the CNN which is contain different stage start with preparing data set, designing, training, testing, online system

1. In the first step of the proposed system, it is within deep learning, which is the algorithm of the convolutional neural network (CNN). It also has different stages before the classification process and depends on building this network.
2. This system classifies diseases and takes into account the most common diseases such as bacteria, fungi, viruses, mold, mite, etc.
3. The plant leaf diseases system, have been built in five stages as shown in the general diagram of the system and its stages, in figure 13.



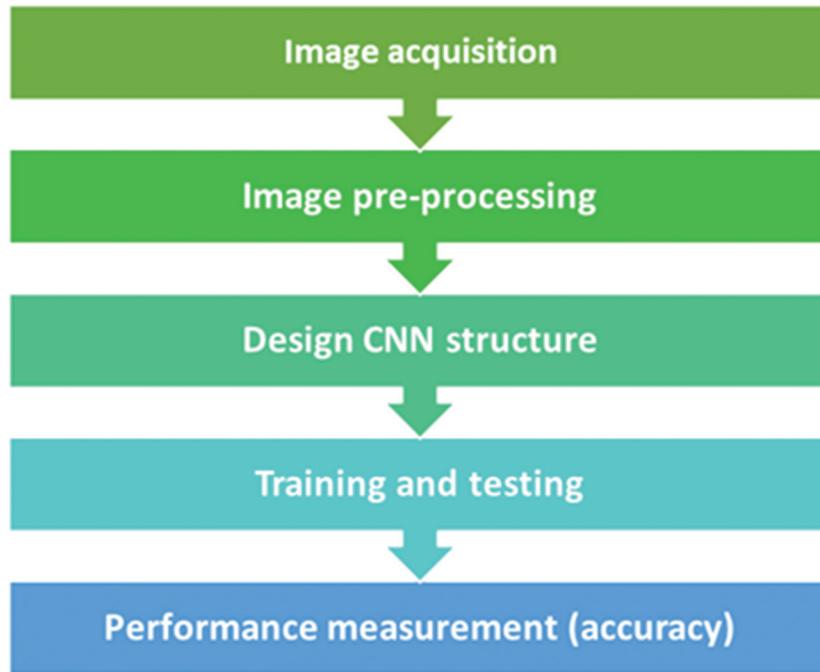


Fig. 13 Plant leaves diseases detection diagram

5.1. Image Acquisition Stage

The dataset (Plant Village) has provided the researcher with a lot of effort in searching for thousands of pictures of all diseases of tomato plant¹. The process of searching for these diseases requires a great effort because he needs to visit many farms and examine a lot of tomato crops to collect the largest possible number of pictures of these diseases.

The case study in the proposed system have been applied on tomato plant disease because it is considered one of the primary plants in the world. Moreover, the reason for choosing these species is their great importance and their spread in the world and in our country Iraq in particular. The data collection collected consists of approximately 18,000 images related to six different categories as shown in table 1. This six categorize in total consist of ten classes of plant disease which has been studied in this proposed system illustrated in figure 14. The data set includes pictures of the healthy plant leaf in addition to all the main types of diseases that can affect the tomato plant. These diseases are bacteria, viruses, fungi, mold and mites. Each sample image in the dataset it is RGB color space with sizes (256 * 256) pixels and the file format of this images is JPG.¹

¹<https://www.kaggle.com/emmarex/plantdisease>

Table 1 Data-Set of tomato plant leaf disease

Categories	Types	No. of Samples
Fungi	1. Alternaria solani 2. Septoria lycopersici 3. Corynespora cassicola 4. Fulvia fulva	1. (1000) 2. (1771) 3. (1404) 4. (952)
Bacteria	Xanthomonas campestris pv. Vesicatoria	(2127)
Mold	Phytophthora Infectants	(1910)
Virus	1. Tomato Yellow Leaf Curl Virus. 2. Tomato Mosaic Virus	1. (5357) 2. (373)
Mite	Tetranychus urticae	(1676)
Healthy	Healthy plant	(1592)

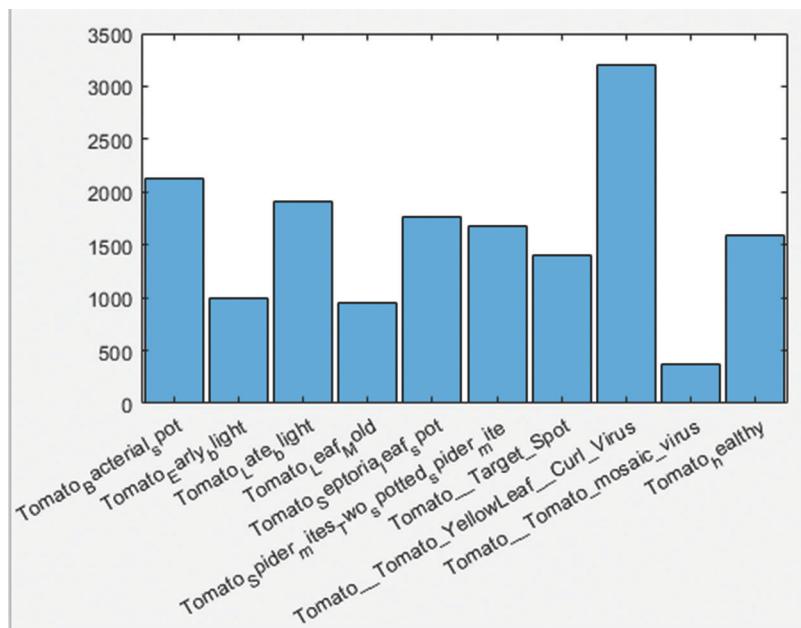
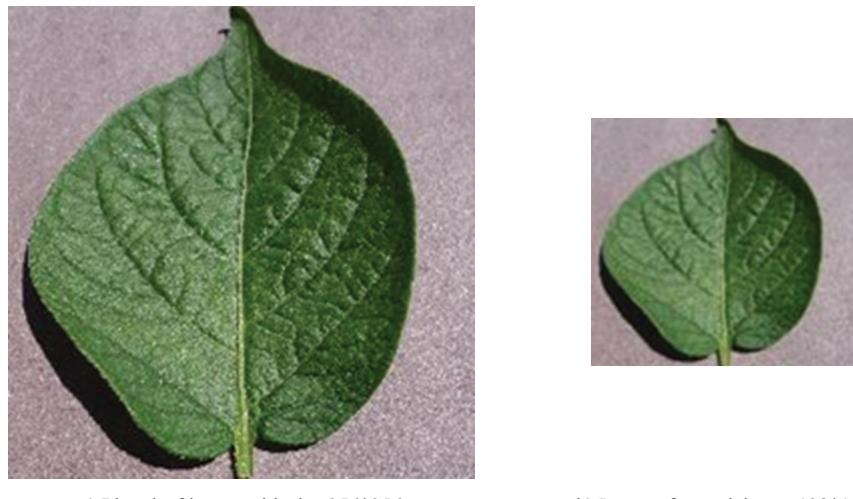


Fig. 14 The histogram of tomato plant disease size of each sample

5.2. Image Pre-Processing

This stage includes the process of reducing the size of the image, as the image is resized (256*256) to (128*128) and this process takes place before the image enters the convolutional neural network. The purpose of reducing the size of the images is to increase the speed of the training method and

obtaining a test account of the model in real terms, this, yields to improve both inputs and variables which speed up the training process; while maintaining the compliments of the data without damaging the original image data. As an example, depicted in figure 15 of resizing images from the proposed system, these images are one of the leaves of the plants used in this system, where images are converted using the image scaling function for each width and height of (256*256) to (128*128).



a) Plant leaf image with size 256*256

b) Image after resizing to 128*128

*Fig. 15 (a) Original image with a size of (256*256), (b) The image after resize (128*128)*

5.3. Design Convolution Neural Network Structure

In the proposed System the neural net has been design as demonstrated in figure 16, which is consist of seven layers. Firstly, the Input layer that has an RGB image with 128 * 128 pixels size. Secondly, the convolution layer that has ten filters were used with size 3*3 with "same" padding. Thirdly, the activation layer which using ReLU activation function. Fourthly, pooling layer that have been using max-pooling layer with each iteration of convolution layer as shown in figure 16 started with ten filters then it became twenty, sixty-four, thirty and so on. Fifthly, normalize layer which use batch normalize layer. Sixthly, fully connected layer which in turn connect all layers to become full connected. Seventhly, Soft max layer that applying Soft max function and the result will be as nine classes according to the types of plant diseases.

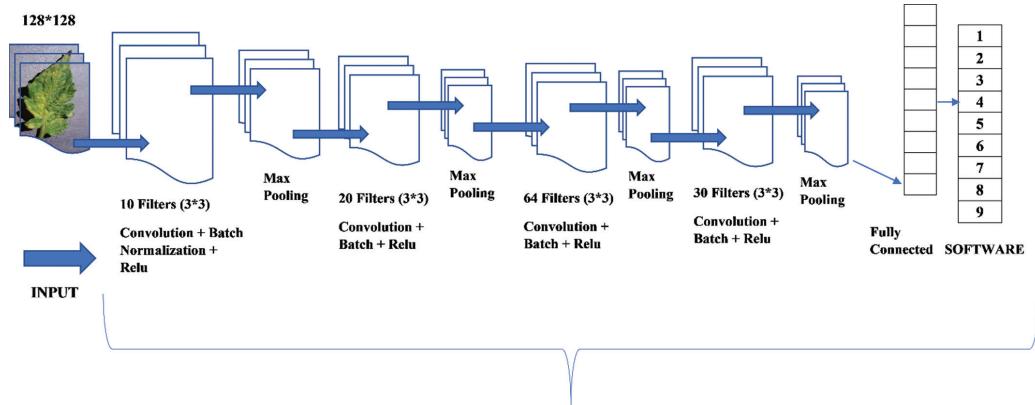


Fig. 16 Structure of the CNN Algorithm

5.4. Training

The network training process is one of the most important stages in the disease classification process because it trains the network on how to deal with the image of tomato plant diseases. The network is trained by extracting the weights in the fully connected layers and these weights are updated every time to bring the result closer to the specific disease.

The training process can take a long time because it deals with a large data set consisting of 18,162 images. After completing the training process the neural network with the optimum weight is saved for used online system. This training can be evaluated by using the "loss function", as shown in figure 17 and 18.

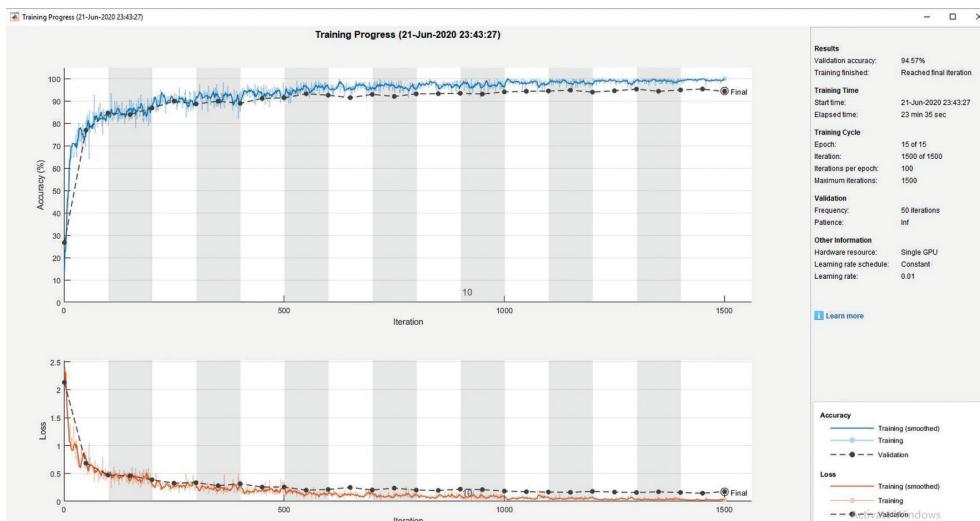


Fig. 17 Training progress of Accuracy and loss function of each iteration



The performance of the training process through the comparing between the accuracy and the loss per each iteration, the system reaches the performance goal at 930 iteration in 15 minutes with accuracy of 99.7% by using a single GPU.

OUTPUT CLASS											97.4%	2.6%
	1472 13.1%	9 0.1%	9 0.1%	0 0%	4 0%	1 0%	2 0%	14 0.1%	0 0%	1 0%		
Tomato Bacterial Spot	1472 13.1%	9 0.1%	9 0.1%	0 0%	4 0%	1 0%	2 0%	14 0.1%	0 0%	1 0%	97.4%	2.6%
Tomato Early Blight	3 0%	627 5.6%	17 0.2%	2 0%	4 0%	1 0%	0 0%	1 0%	0 0%	0 0%	95.7%	4.3%
Tomato Late Blight	1 0%	35 0.3%	1272 11.3%	4 0%	10 0.1%	0 0%	0 0%	2 0%	0 0%	0 0%	96.1%	3.9%
Tomato Leaf Mold	0 0%	2 0%	17 0.2%	647 5.8%	11 0.1%	4 0%	1 0%	1 0%	0 0%	0 0%	94.5%	5.5%
Tomato Sclerotinia Leaf Spot	5 0%	2 0%	6 0.1%	7 0.1%	1197 10.7%	2 0%	4 0%	1 0%	5 0%	0 0%	97.4%	2.6%
Tomato Spider Mites	0 0%	1 0%	2 0%	0 0%	2 0%	1123 10%	10 0.1%	2 0%	1 0%	0 0%	98.4%	1.6%
Tomato Target Spot	0 0%	15 0.1%	3 0%	2 0%	7 0.1%	31 0.3%	956 8.5%	0 0%	1 0%	3 0%	93.9%	6.1%
Tomato Yellow Leaf Curl Virus	8 0.1%	7 0.1%	5 0%	4 0%	1 0.1%	7 0%	2 0%	2227 19.9%	0 0%	0 0%	98.5%	1.5%
Tomato Mosaic Virus	0 0%	2 0%	0 0%	0 0%	4 0%	3 0%	0 0%	0 0%	254 2.3%	0 0%	96.6%	3.4%
Tomato Healthy	0 0%	0 0%	5 0%	0 0%	0 0%	1 0.1%	6 0.1%	0 0%	0 0%	1108 9.9%	98.9%	1.1%
	98.9% 1.1%	89.6% 10.4%	95.2% 4.8%	97.1% 2.9%	96.5% 3.5%	95.7% 4.3%	97.3% 2.7%	99.2% 0.8%	97.3% 2.7%	99.5% 0.5%	97.1% 2.9%	

Fig. 18 Confusion matrix for the training set result

The measurements of confusion matrix in figure 18 illustrates the rate of overall performance accuracy for each disease type of output class training to get the target classes for the training set.

5.5. Testing

The testing process is the final stage in the classification process of tomato leaf diseases and is used to determine the final evaluation and indicate the extent to which the results can be trusted in the classification of the specific disease. The test process differs from the training process as it uses unrated images to know what kind of diseases this image belongs to. This process is done easily because it is pre-trained in the training phase and this function is used in Matlab, CLASS = classification (model, training, group), and this classifies each row from the data set to one of the training groups. The figure 19 shows the ratio of confusion matrix for testing set result.

Rate the overall performance accuracy of CNN[92.9426 in Testing]												
Output Class	Tomato_Bacterial_spot	1220 12.7%	4 0.0%	3 0.0%	1 0.0%	2 0.0%	0 0.0%	2 0.0%	23 0.2%	0 0.0%	2 0.0%	97.1% 2.9%
	Tomato_Early_blight	18 0.2%	525 5.5%	80 0.8%	14 0.1%	26 0.3%	14 0.1%	8 0.1%	11 0.1%	0 0.0%	0 0.0%	75.4% 24.6%
	Tomato_Late_blight	7 0.1%	33 0.3%	998 10.4%	19 0.2%	7 0.1%	2 0.0%	0 0.0%	7 0.1%	0 0.0%	2 0.0%	92.8% 7.2%
	Tomato_Leaf_Mold	0 0.0%	0 0.0%	13 0.1%	480 5.0%	11 0.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	95.2% 4.8%
	Tomato_Sepalaria_leafspot	8 0.1%	12 0.1%	33 0.3%	40 0.4%	981 10.2%	8 0.1%	8 0.1%	2 0.0%	1 0.0%	0 0.0%	89.8% 10.2%
	Tomato_spider_mites_Web_potted_spider_mite	0 0.0%	1 0.0%	4 0.0%	8 0.1%	4 0.0%	899 9.4%	23 0.2%	8 0.1%	0 0.0%	2 0.0%	94.7% 5.3%
	Tomato__Target_Spot	6 0.1%	13 0.1%	4 0.0%	3 0.0%	26 0.3%	74 0.8%	797 8.3%	7 0.1%	1 0.0%	9 0.1%	84.8% 15.2%
	Tomato__Tomato_YellowLeaf_Curl_Virus	12 0.1%	9 0.1%	4 0.0%	3 0.0%	0 0.0%	6 0.1%	0 0.0%	1867 19.4%	0 0.0%	0 0.0%	98.2% 1.8%
	Tomato__Tomato_mosaic_virus	0 0.0%	2 0.0%	1 0.0%	0 0.0%	5 0.1%	3 0.0%	1 0.0%	0 0.0%	222 2.3%	0 0.0%	94.9% 5.1%
	Tomato_healthy	5 0.1%	1 0.0%	5 0.1%	3 0.0%	1 0.0%	0 0.0%	3 0.0%	0 0.0%	0 0.0%	940 9.8%	98.1% 1.9%
	95.6% 4.4%	87.5% 12.5%	87.2% 12.8%	84.1% 15.9%	92.3% 7.7%	89.4% 10.6%	94.7% 5.3%	97.0% 3.0%	99.1% 0.9%	98.4% 1.6%	92.9% 7.1%	
	Target Class	Tomato_Bacterial_spot	Tomato_Early_blight	Tomato_Late_blight	Tomato_Leaf_Mold	Tomato_Sepalaria_leafspot	Tomato_spider_mites_Web_potted_spider_mite	Tomato__Target_Spot	Tomato__YellowLeaf_Curl_Virus	Tomato__Tomato_mosaic_virus	Tomato_healthy	

Fig. 19 Confusion matrix for the testing set result

5.6. GUI Implementation

The system used in this research has three interfaces which in turn facilitates the communication process between the user and the system. The main GUI contain the general information about the system, and from this GUI the user can access the training GUI. The main interface has the basic information about the system title and its developers as usual and the option to run the system, as shown in figure 20:

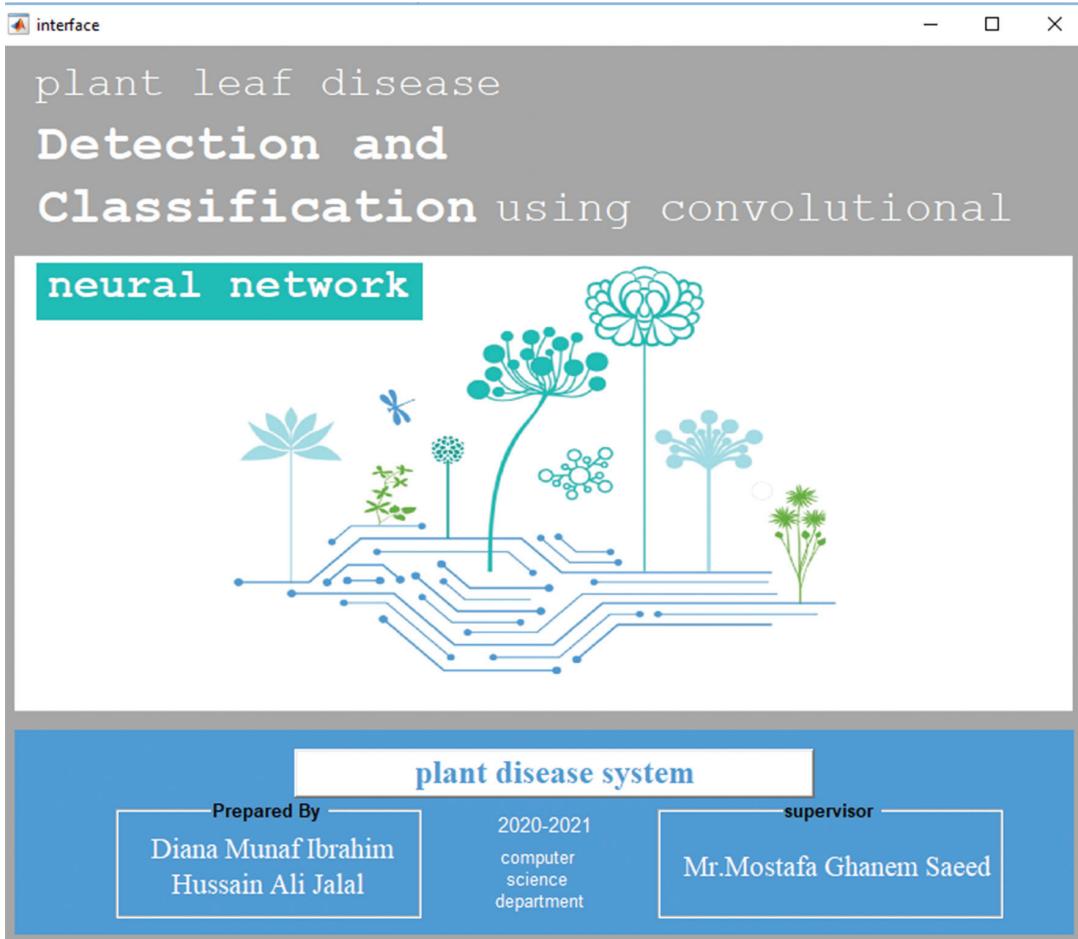


Fig. 20 The main interface of the system

The second interface that appear to the user directly is the training then testing interface in figure 21 and 22 which include all the options about training and testing rate, input image dimension, number of data set records and CNN network name, all these options is relate to the training process that accounts the accuracy and the process time for each process. The interface includes also the options that related to the testing process that needs to select which CNN network that the system has to test. It shows also the accuracy and time for the testing process. In addition to that the interface has a window to show all the data-set used and the amount images for each disease's types.

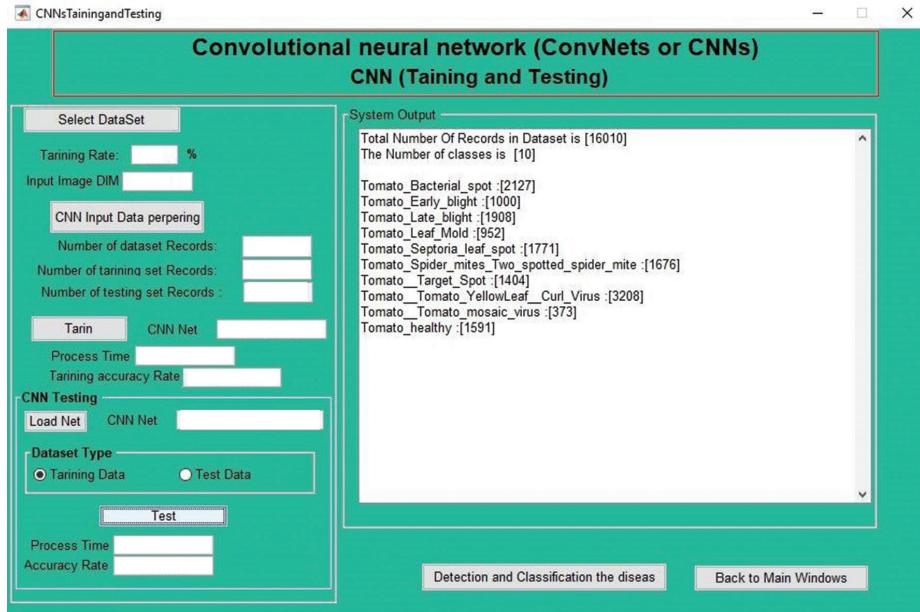


Fig. 21 Training and Testing CNN Interface

After the training progress and confusion interfaces that has been achieved, the system shows in the important interface in which detecting diseases process and suggesting the best solution for the disease are performed.

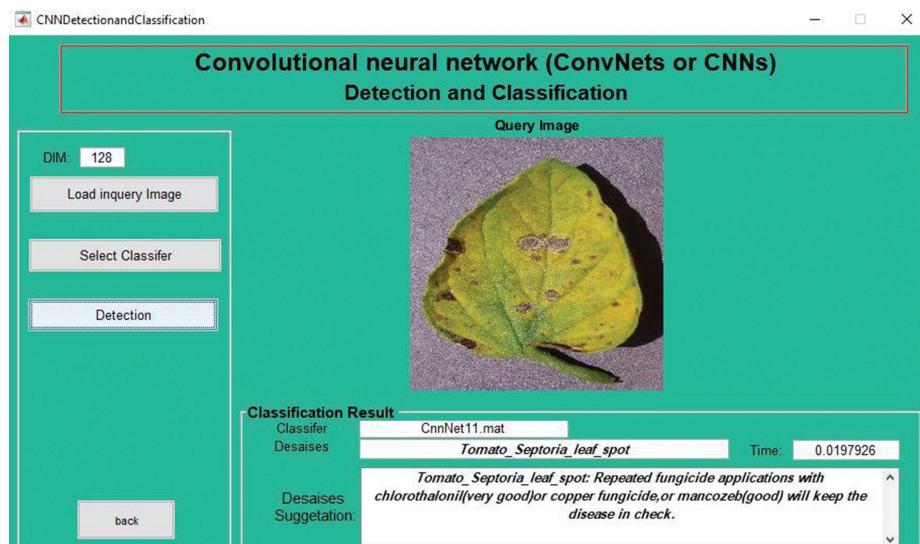


Fig. 22 Detecting diseases and suggesting solution interface

In this interface of figure 21, the user has to upload his image of plant disease through the option ‘Load Inquiry Image’ then he has to select which classifier is appropriate with this testing process after that the system will show all the information about which disease the image is related to through the option ‘Detection’. In addition to that the system will suggest the appropriate solution or therapy that leads to treat the crops and stop spread it to other plants. In addition, the system supports the recommendation and treatment to each disease that has been classified as shown in the table 2.

Table 2 Tomato leaf disease and treatment recommendation

Disease	Treatment Recommendation
Bacterial Spot	1. Eliminate all diseased plant material. 2. Plant pruning to advance air circulation process. 3. Spray unhealthy plants with a copper fungicide.
Septoria Leaf Spot	1. Avoid using over-head irrigation. 2. Repeat fungicide applications on unhealthy plants with chlorothalonil or copper fungicide, or maybe with mancozeb that keeps the disease in check.
Early Blight Fungi	1. Use pathogen-free seeds, use crop rotation. 2. Weeding and voluntary tomato plants. 3. Avoid irrigation water process to the tomato plants. 4. If the disease is severe to manage by chemical control, then use fungicides.
Target Spot	1. Spraying with protective fungicides after the first symptoms appear. 2. Removing damaged debris to prevent it from being carried over to the next crop.
Leaf Mold (Fulvia-Fulva)	1. Crop residues must be eliminated from the field. 2. Pruning and trimming is applied to increase air circulation which controls the disease. 3. Avoid moistening the leaves when watering, with a preventive fungicide program.
Yellow Leaf Curl Virus	Reflective thalamus used in rows to minimize whitefly feeding, after removing the plants with primary symptoms.
Mosaic Virus	Instantly remove and destroy infected plants, uninstall weeds in the garden.
Late Blight	1. Keep the foliage dry. 2. Leave additional space between plants. 3. Avoid top irrigation.
Spider Mites	1. Using hexagon kills the egg and nymphal stages for two spotted spider mites. 2. Avoid applying organophosphate-based insecticides.

6. Result Discussion

The system uses mixing matrix to calculate the accuracy of the algorithm in the training and testing stages, thus showing the total accuracy of each training or test. After that, the specific type of disease that the network is referring to is detected after testing the image. The system will give the best result according to the proportions that have been used for both, the testing and training stages as shown in table 3 and 4:

Table 3 Measuring training rate by accuracy per time

Training rate	50*50	60*40	70*30	80*20	90*10
Accuracy rate	99.512	99.656	99.634	99.515	99.798
Process time	1001.26	1236.05	1399.36	1639.05	1393.62

The results that shown in table 3 means that the system has been well trained with a training accuracy ratio of 99.7%, which mean the system has error ratio of 0.3% that considered as neglected ratio statistically.

Table 4 Measuring testing rate by accuracy per time

Testing rate	50*50	60*40	70*30	80*20	90*10
Accuracy rate	92.768	97.397	97.100	92.785	97.383
Process time	19.53	23.33	28.19	44.81	36.10

After performing many testing operations using different proportions as shown in table 4, it became clear that the system was have testing accuracy rate of 97.3% in about 23 milliseconds by using the proportion 60*40. It is considered a very high ratio in measuring the effectiveness of a system that applied a deep learning approach in studying and applying the structure of the convolutional neural network based on approximately 18,000 images and classifying each one separately. After taking into consideration that the accuracy of the testing process is responsible for determining the extent to which this system can be trusted to classify plant diseases.

7. Conclusions

While developing the tomato plant leaf disease diagnosing system and the process of suggesting the recommended therapy to treat it via applying deep learning and convolutional neural networks, the following conclusion has been delivered:

1. Using deep learning will eliminate the most important agriculture constraints and the traditional ways of detecting plant diseases.
2. The training models were implemented using a dataset of 18,000 images that were taken in the fields of cultivation in laboratory and actual conditions.
3. The optimal model design in the neural convolution networks obtained a success rate of 97.39 percent, which means the error percentage is 2.61 percent previously unheard in the classification of 18,162 testing set.
4. The value of actual images (captured in crop fields) is shown by the training data in the presented findings.
5. The conceptual deep learning method showed a high ability of the data to develop the method in terms of quantity and consistency.



References

- [1] Raut, S., & Ingole, K., 2017, “*Review on Leaf Disease Detection Using Image Processing Techniques*”, International Research Journal of Engineering and Technology (IRJET), 4(04), pp.2044-2047.
- [2] Hanson, A. M. G. J., Joel, M. G., Joy, A., & Francis, J., 2017, “*Plant Leaf Disease Detection Using Deep Learning and Convolutional Neural Network*”, International Journal of Engineering Science.
- [3] Cortes, E., 2017, “*Plant Disease Classification Using Convolutional Networks and Generative Adversarial Networks*”.
- [4] Irudayaraj, J., 2009, “*Pathogen Sensors*”, Vol. 9, pp. 8610–8612.
- [5] Meroni, M.; Rosini, M.; Picchi, V.; Panigada, C.; Cogliati, S.; Nali, C.; Colombo, R. Asse, 2008, “*Assessing Steady-State Fluorescence and PRI From Hyperspectral Proximal Sensing as Early Indicators of Plant Stress: The Case of Ozone Exposure*”, Vol. 8, pp. 1740–1754.
- [6] Wah Liew, O.; Chong, P.; Li, B.; Asundi, K., 2008, “*Signature Optical Cues: Emerging Technologies for Monitoring Plant Health*”, Vol. 8, pp. 3205–3239.
- [7] Fiallo-Olivé, Elvira; Navas-Castillo, Jesús, 2019, “*Tomato Chlorosis Virus, An Emergent Plant Virus Still Expanding Its Geographical and Host Ranges*”, Molecular plant pathology, Vol.20 (9), p.1307-1320, Wiley, England.
- [8] Romero, Ana M; Vega, Damián; Pizzorno, Romina; Cordon, Gabriela; Correa, Olga S, 2018, “*Hydraulic and Leaf Reflectance Alterations Induced by Clavibacter Michiganensis Subsp. Michiganensis On Tomato Plants*”, European journal of plant pathology, Vol.152 (2), p.567-572, Springer Netherlands.
- [9] Fenni, Soumia; Hammou, Habib; Astier, Julien; Bonnet, Lauriane; Karkeni, Esma; Couturier, Charlène; Tourniaire, Franck; Landrier, Jean-François, 2017, “*Lycopene and Tomato Powder Supplementation Similarly Inhibit High-Fat Diet Induced Obesity, Inflammatory Response, And Associated Metabolic Disorders*”, Molecular nutrition & food research, Vol.61 (9), Wiley, Germenay.
- [10] Ko, Jina; Baldassano, Steven N; Loh, Po-Ling; Kording, Konrad; Litt, Brian; Issadore, David, 2018, “*Machine Learning to Detect Signatures of Disease in Liquid Biopsies – A User's Guide*”, Lab on a chip, Vol.18 (3), p.395-405, Royal Society of Chemistry (RSC), England.
- [11] Lauzon, Francis Quintal, 2012, “*An Introduction to Deep Learning*”, 11th International Conference on Information Science, Signal Processing and their Applications (ISSPA). IEEE, 2012. p. 1438-1439.
- [12] Thabit Thabit, Yaser Jasim, 2015, “*A Design of ‘Windows 7 Troubleshooting ‘Software Using Hybrid Intelligence Systems*”, International Journal of Engineering Research & Management Technology, Vol. 2, Issue. 2, India.
- [13] Priyadarshini, R. Ahila et al., 2019, “*Maize Leaf Disease Classification Using Deep Convolutional Neural Networks*”, Neural Computing and Applications, pp. 1-9.
- [14] Jeny, Afsana Ahsan; Junayed, Masum Shah; Atik, Syeda Tanjila, 2018, “*Passnet-Country Identification by Classifying Passport Cover Using Deep Convolutional Neural Networks*”, 21st International Conference of Computer and Information Technology (ICCIT). IEEE, pp. 1-6.
- [15] Safwan O Hasoon, Yaser A Jasim, 2013, “*Diagnosis Windows Problems Based on Hybrid Intelligence Systems*”, Journal of Engineering Science & Technology (JESTEC), Vol. 8, Issue. 5, pp. 566-578, Malaysia.



- [16] Yani, Muhamad, 2019, “*Application of Transfer Learning Using Convolutional Neural Network Method for Early Detection of Terry’s Nail*”, Journal of Physics: conference series. IOP publishing, 2019.
- [17] Turaga, S. C., Murray, J. F., Jain, V., Roth, F., Helmstaedter, M., Briggman, K., Denk, W., and Seung, H. S., 2010, “*Convolutional Networks Can Learn to Generate Affinity Graphs for Image Segmentation*”, Neural Computation.
- [18] Mathieu, Michael; Henaff, Mikael; Lecun, Yann, 2013, “*Fast Training of Convolutional Networks Through Ffts*”, *arXiv preprint arXiv:1312-5851*.
- [19] Mustafa G Saeed, Yaser Abdulaali Jasim, 2018, “*Developing a Software for Diagnosing Heart Disease via Data Mining Techniques*”, Ediciones Universidad de Salamanca (España), Spain.
- [20] Thabit H Thabit, Yaser A. Jasim, 2015, “*A Manuscript of Knowledge Representation*”, International Journal of Human Resource & Industrial Research, Vol. 4, Issue. 4, pp. 10-21, India.
- [21] Alsaaign, M. O., & Saeed, M. G. Jasim, Y. A, 2020, “*Designing and Implementation of a Security System Via UML: Smart Doors*”, CSRID (Computer Science Research and Its Development Journal), 12(1), pp. 01-22.
- [22] Alsaaign, M. O., Flaih, T. M., & Saeed, M. G., Jasim, Y. A., 2020, “*On Announcement for University Whiteboard Using Mobile Application*”, CSRID (Computer Science Research and Its Development Journal), 12(1), pp. 64-79.
- [23] Peter Timmer, 2002, “*Agriculture and economic development*”, Handbook of Agricultural Economics, Volume 2, pp. 1487-1546, Elsevier.
- [24] Praburaj L., 2018, “*Role of Agriculture in the Economic Development of a Country*”, Shanlax International Journal of Commerce, Vol. 6, No. 3, pp. 1–5.
- [25] Khorshed Alam and John Rolfe, 2006, “*Economics of Plant Disease Outbreaks*”, Agenda, Vol. 13, No. 2, pp. 133-146.

