International Journal of Multidisciplinary Engineering in Current Research - IJMEC
Volume 10, Issue 6, June-2025, http://ijmec.com/, ISSN: 2456-4265

A CNN Based System For Tomato Disease Detection

Ms. K Srinidhi Reddy, B.Ramya, M.Renuka, B.Rushmitha

[1]Assistant Professor, Department Of ECE, Bhoj Reddy Engineering College For Women, India.

[2,3]B. Tech Students, Department Of ECE, Bhoj Reddy Engineering College For Women, India.

ABSTRACT

Rapid human population growth requires corresponding increase in food production. Easily spreadable diseases can have a strong negative impact on plant yields and even destroy whole crops. That is why early disease diagnosis and prevention are of very high importance. Traditional methods rely on lab analysis and human expertise which are usually expensive and unavailable in a large part of the undeveloped world. Since smartphones are becoming increasingly present even in the most rural areas, in recent years scientists have turned to automated image analysis as a way of identifying crop diseases. This paper presents the most recent results in this field, and a comparison of deep learning approach with the classical machine learning algorithms. One of the important and tedious task in agricultural practices is detection of disease on crops. It requires huge time as well as skilled labor. This paper proposes a smart and efficient technique for detection of crop disease which uses computer vision and machine learning techniques. The proposed system is able to detect 20 different diseases of 5 common plants with 93% accuracy.

1- INTRODUCTION

Agriculture plays a vital role in constructing and developing of the economy of any nation (Huang *et al.* 2020). As the world population is expected to reach approximately 10 billion of people by 2050, it is a necessity to increase the agricultural productivity (Fess *et al.* 2011). Among the various fruits, tomato fruit is incredibly useful for health benefits and the livelihood of farmers (Dimatira *et al.* 2016). Tomato is rich in nutrition and it has efficiency in health care (Yinli *et al.* 2011). The biological name of tomato is 'Solanum Lycopersicum'. According to the report of food and agriculture organization of the united nation (FAOSTAT 2017), the worldwide production of tomato is approximately 182,301,395 tons and India had produced 20,708,000 tons of tomatoes. India is the second-largest producer of tomato globally after China. Ripe tomato has high-level antioxidant compound, and it can greatly decreasethe risk of some severe human diseases like cancer and cardiovascular diseases (Ciaccheri *et al.* 2018). Tomato is a major cash crop. So, reducing diseases in tomato fruit is important for increasing the quality and output of the tomato agriculture (Zhao & Qu 2019).

To meet high-scale production with quality of production and to meet the consumer's expectations with the market's standard, it is necessary to perform an accurate and reliable grading method (Ireri *et al.* 2019). According to the regulations (EU 2011), tomatoes should be in fresh condition, and the fruits must reach the market without any damage for further processing. Numerous biological and natural parameters are utilized to examine the quality of tomatoes after harvesting by using grading and sorting. Both grading and sorting can be done by measuring the size, shape, defects, colour, and maturity (Arjenaki *et al.* 2013).

Among the various fruits, tomato fruit is incredibly useful for health benefits and the livelihood of farmers (Dimatira *et al.* 2016). Tomato is rich in nutrition and it has efficiency in health care (Yinli *et* *al.* 2011). Tomato is a major cash crop. So, reducing diseases in tomato fruit is important for increasing the quality and output of the tomato agriculture.

Usually, the consumption of fresh tomatoes is mostly high due to highnutritional properties in the market (Soto Zamora *et al.* 2005). Fresh tomatoes can be identified by their appearance and firmness. The consumer acceptance rate of tomatoes can be confirmed by the major external characteristics like colour. In tomato fruit, colour is a major feature to identify the quality which canindicate the grading factor. Green colour tomato has high chlorophyll content and both orange and red colour tomato contains β-carotene and lycopene content, respectively (Bathgate *et al.* 1985). The major characteristics of lycopene content in the tomato is that it acts against different serious human disease like cancer (Basu & Imrhan 2007). The consumers can also check the firmness of tomato by using their fingers during purchase and based on the firmness, the consumer can purchase the tomatoes (Ali 1998). The firmness of tomato is changed by tissue softening due to weight loss and turgor loss. The weight loss is mainly due to the variation of temperature in the storage place (Alia Tejacal *et al.* 2007).

The growth and usage of tomatoes are highly affected by different factors such as environmental changes and virus or bacteria-based diseases. Due to the disease in tomatoes, the quality, and the economic yield of the architecture are rapidly decreased. Therefore, the prediction of the diseases and reducing the disease factors are highly considered by the farmers while cultivating tomatoes (Wang & Qi 2019). During the last decade, different researchers have proposed different kinds of machine learning-based algorithms for grading and sorting tomato fruits. Estimation of maturity using infrared spectroscopy was done by Sirisomboon *et al.* (2012). This algorithm was based on principal component analysis (PCA). To estimate carotenoid in tomato fruits Vazquez *et al.* (2013) provided a prediction system. A cost-effective grading system was suggested by Rupanagudi *et al.* (2014) to grade the fruits under three categories such as red, turning, and green. Pavithra *et al.* (2015) developed K nearest neighbour (KNN) and support vector machine (SVM) image classification scheme to identify the maturity of tomato. This is operated by using Euclidean distance-based separation algorithm and different features like texture, colour, and shape. This algorithm was based on principal component analysis (PCA).

2-OPTIMIZING CNN FOR TOMATO DISEASE PREDICTION

Tomato is the most popular fruit in many nations over the world due to its unique nutritional value, taste, and visual appeal (Wang *et al.* 2018). Identification of diseases in tomatoes will be useful for getting high-profit tomato harvest. The traditional method for revealing and classification of fruit infections is centered on the human eye inspection by the specialists. In a few emerging countries, conferring specialists are lavish and time spending because of faraway sites of their accessibility. Automatic recognition of fruit infections is vital to spontaneously identify the indications of infections as soon as they occur on the rising fruits. Fruit infections can affect the main benefits in the crop, quality emerged in harvesting (Gaikwad *et al.* 2017) and in marketing.

Virus disease, black spot, late-blight, early-blight, blotchy ripening, canker are common infections in tomato. In the olden days, fruit samples were accumulated and analyzed in laboratory system with aids of some mechanisms like spectroscopy for identifying the infection in tomato (Gu *et al.* 2019). But these techniques needed experienced and skillful expertise and a certain type of spectroscopy methods that are destructive systems (Brahimi *et al.* 2017). Therefore, a precision disease prediction technology is mandatory (Zhang *et al.* 2018). Considerable investigations have been conducted over the years to predict the diseases in tomato plants and fruits by utilizing non-destructive methods as discussed in chapter 2. For this research, the aim is to detect tomato disease using an optimized convolutional neural network (CNN). Neural networks (NN) are widely used systems of disease identification because they perform classification without using explicit recognition concept (Bishop 1995). CNN is a technology that utilizes a lot of simple processing modules, termed neurons. A neuron commonly uses complex input vectors and produces output signal accordingly. The operation of the input vector is defined by using the non-linear function as well as the neuron’s weight. The potential of CNN is that the ability to train itself as well as perform with a concept of pattern recognition and provide a conclusion in an unbiased manner (Iglesias *et al.* 2004). CNN constitutes a specialized tool for data modeling (Goyal 2013). CNN technology is the most widely used technology in various fields especially in agriculture, where it can be used to evaluate a large volume of data and relate these data to a particularly desirable characteristic (Alves *et al.* 2017).

In modern years, abrupt improvements can be observed in evaluating the quality of agricultural foods, along with the sophisticated optical imaging methods like NIR hyperspectral and multispectral imaging, and NIR- spectroscopy. Both spectroscopy and image processing methodologies have developed a considerable platform to perform the faster, economic, accurate, andreliable evaluation that has performed efficiently to excellence and protection evaluations and sorting as well as the grading of fruits based on shape, colour, size, as well as texture recognition (Chandrasekaran *et al.* 2019).

He *et al.* (2005) utilized spectroscopy to measure the quality of tomatoes by using visible infrared reflectance. It was mainly used to measure fruit firmness. This device was constructed by using monochromator, NMOS photodiode and Teflon disk. The firmness of tomato fruits was measured by using maximum compression force. The tomato fruits were determined by reluctance (R) range between 350 nm and 2500 nm. The characteristics ofabsorbance were determined by (1/R), was evaluated for identifying the finest calibration pattern of each attributes utilizing principal component regression (PCR) and partial least square regression (PLSR). Both PCR and PLSR had the possibility to determine the element concentration, biological and physical characteristics of tomatoes by using infrared spectra. This method produced the regression coefficient (RC) of 0.88 with the 16.017 standard error prediction (SEP). Both spectroscopy and image processing methodologies have developed a considerable platform to perform the faster, economic, accurate.

Jackman *et al.* (1990) proposed a measurement system of testing the firmness of tomatoes by using different concepts of spectroscopy such as flat- plate compression, constant area compression, and puncture. By using flat-plate compression, the firmness was achieved as 4.91 for the RMSE of 9.70. The constant area compression provided the firmness as 2.72 with 6.89 RMSE. Also, the puncture identified the firmness as 0.37 with the RMSE of 0.56. Jackman *et al.* (1990) revealed that minor changes in tomato firmness are caused by chilling which could not be found while flat-plate compression checks were conducted incomparison to puncture and constant area compression analysis.

The ability of CNN is that it can solve any complex problems with less time and less computing resources. CNN can provide superior performance by comparing with conventional statistical modules, and they are non-parametric and not requiring more information for the physical process (Laurindo *et al.* 2017). Numerous investigations have been reported in developing disease prediction system by applying the concept of CNN. However, many of those techniques utilized conventional CNN algorithms and the performance was not up to the level of yielding high accuracy. This work aims to optimize the CNN algorithm by using the principle of whale optimization and improves the quality of detection capability.This chapter briefs the related CNN techniques utilized in recent years to identify diseases in tomato plants and fruits, describes the concept of firefly algorithm (FA) of image segmentation, explains the principles of whale optimization, and describes the implementation of whale optimized CNN for classifying fruits. Detailed comparison of the proposed algorithm with the existing state-of-the-art algorithms are given in this chapter.

3-DUAL SUPPORT VECTOR MACHINE CLASSIFIERS FOR TOMATO GRADING SYSTEM

Advanced machine vision technologies motivated the researchers to use fast and accurate non-destructive assessment schemes for implementing reliable fruit grading and sorting systems over the few decades (Srivastava & Sadistap 2018). Non-destructive quality assessment scheme utilizes several properties like structure and physical state of the fruits to classify them (Dolatowski *et al.* 2007). Non-destructive grading systems are widely used for measuring stiffness, firmness, and ripeness stage of the fruits (Molina Delgado *et al.* 2009). Further, based on the concept of non-destructive and various artificial intelligence (AI), various fruit classification algorithmssuch as ANN classifiers, maximum likelihood classifiers (MLPs), and support vector machine classifiers (SVMs) have been proposed over the years. Even though the outcome of these classification schemes is usually positive, yet these methods have some limitations.

Based on the detailed review, it is noticed that any individual classification algorithm needs to be optimized by using some specialized optimization procedures to provide the highest accuracy in classification, for example, whale optimized ANN classifier discussed in the previous chapter. Another way of optimizing the classification scheme is using multi-level classifiers (MLCs). Various classifiers have various accuracies for various classes, because of the complementary benefits of various classifiers. This method utilized polynomial functions to convert the constrained QPP of DSVM into unconstrained minimization problems and it solved the Newton-Armijo algorithm to improve the effectiveness of the classification process. The classification accuracy of non-linear PSDSVM is higher than the linear PSDSVM. This system produced 96.52% accuracy by performing classification on UCI dataset.

This system worked with less than 0.84 of root mean square error of prediction (RMSEP). Acharya *et al.* (2017) suggested that a handheld visible- NIR spectrophotometer was appropriate for field evaluation and it could be utilized to decide the maturity of tomato fruits based on their pigmentation. Further, by updating the calibration model by adding the existing models to the training inhabitants to expand the scope of the characteristic of concern and the scope of biological and natural matrices to be realized, linkage with tomato fruitsis suggested to accomplish robust functional use.

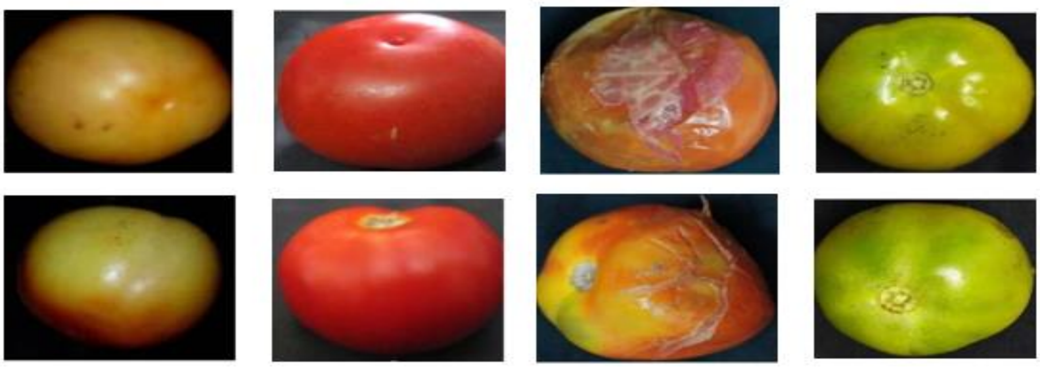
Feng *et al.* (2019) proposed a spectrometer-based tomato quality measurement system to grade the cherry tomatoes. Different CV-basedalgorithms like partial least square (PLS), extreme learning machine (ELM) and support vector machine (SVM) were utilized to predict the quality of tomatoes. Among the different CV-based algorithm, ELM outperformed in terms of RMSE and residual predictive deviation. The ELM algorithm operated with RMSEP of 0.3141 and RPD of 5.6118 for the firmness of cherry tomato. The ELM algorithm can reduce the error rate in training. The outcomes suggested that NIR spectroscopy technology merged with chemometric evaluation can offer a precise estimate of quality characteristics of cherry tomato fruits for the duration of postharvest storing.

Skolik *et al.* (2019) developed a spectroscopy-based tomato fruit grading system for identifying different ripening stages. Different features used for the experimentation were colour, weight, size, and internal composition. SVM and chemometrics are used to identify six different ripening stages. The ripening identification processes were split into two phases such as pre- processing spectra and Linear discriminant analysis (LDA) by using PCA factors. The average accuracy of the system was 99% to 100% for six different stages of ripening.

The different levels of maturity were very early harvest, early harvest, middle harvest, and late harvest. By using fuzzy classifiers, the fruits were sorted by five different categories like immature, early mature, mature, fully ripe, and ripe. By using this method, 60% of the fruits were classified correctly.

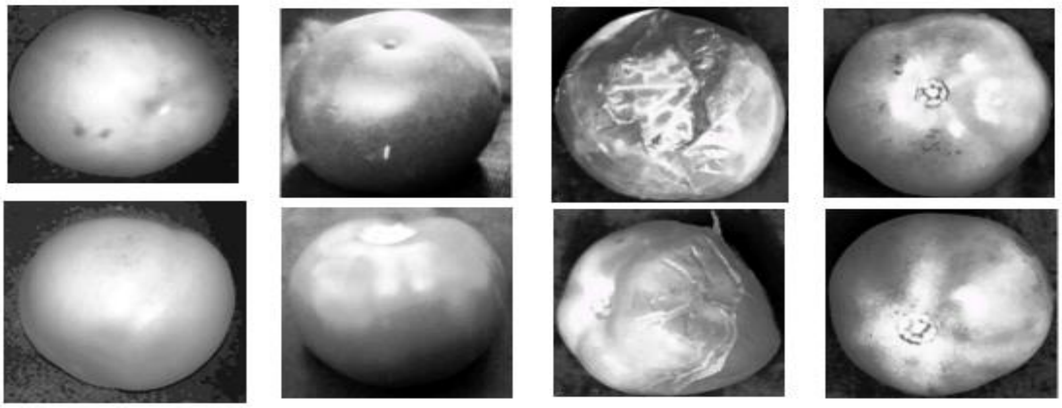
4- OUTPUT AND DISCUSSION

To perform experimental analysis on the proposed WOCNN system, MATLAB is sed. As shown in Figure 3.3 and Figure 3.4, input images and preprocessing output images are considered as two samples per category of tomato.



Blotchy Ripe Sour rot Unripe

Fig 1 Input images of different categories of tomato



Blotchy Ripe Sour rot Unripe

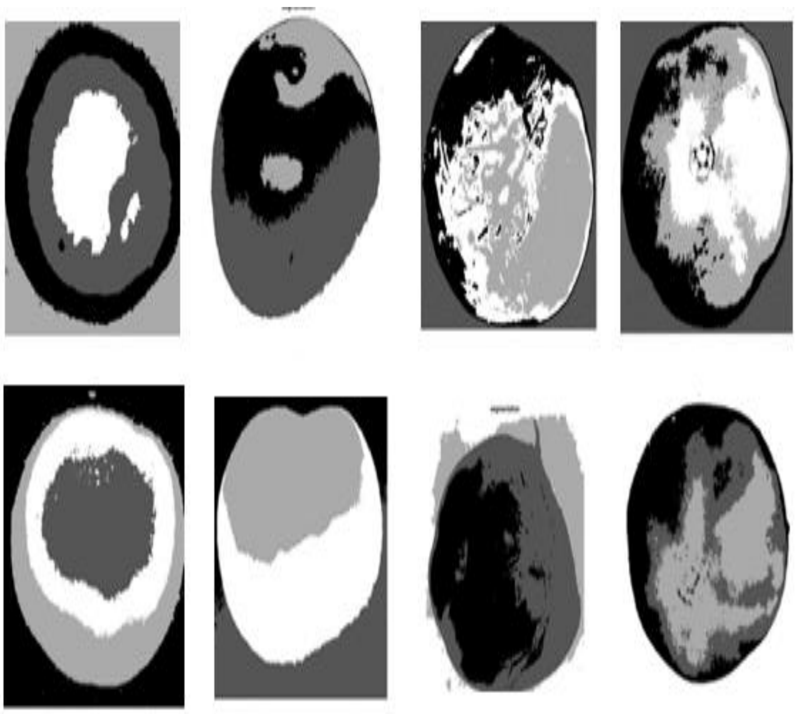
Fig 2 Pre-processing output images of different categories of tomato

As illustrated in Figure 3.5 and Figure 3.6, the image segmentation output is the best and is considered for further experimentation. Various performance factors such as accuracy, precision, false-positive rate, specificity, F1 score, error rate, sensitivity, Matthews's Correlation Coefficient (MCC), and Kappa are measured, as illustrated in Equation (3.24) to (3.32).

Precision = (TP)/ (TP + FP)

False Positive Rate(FPR) = (FP) / (FP + TN)

Specificity =(TN)/ (FP + TN)



Blotchy Ripe Sour rot Unripe

Fig 3 Output of image segmentation using FA algorithm where, TP is the number of cases predicted correctly when the case is positive, TN is the number of cases predicted correctly when the case is negative, FN is the number of cases predicted incorrectly when the case is negative, and FP is the number of cases predicted incorrectly when the case is positive. The project also provides economic benefits by minimizing losses and enabling data-driven decision-making. Additionally, automated disease detection increases efficiency, saving time and labor, while promoting sustainable agriculture practices by reducing chemical usage and enabling targeted treatments. Various performance factors such as accuracy, precision, false-positive rate, specificity, F1 score, error rate, sensitivity, Matthews's Correlation Coefficient (MCC), and Kappa are measured, as illustrated in Equation. Overall, the project has a positive impact on the agricultural industry, supporting farmers and promoting sustainable practices.

Comparison of performance metrics for different classification algorithms

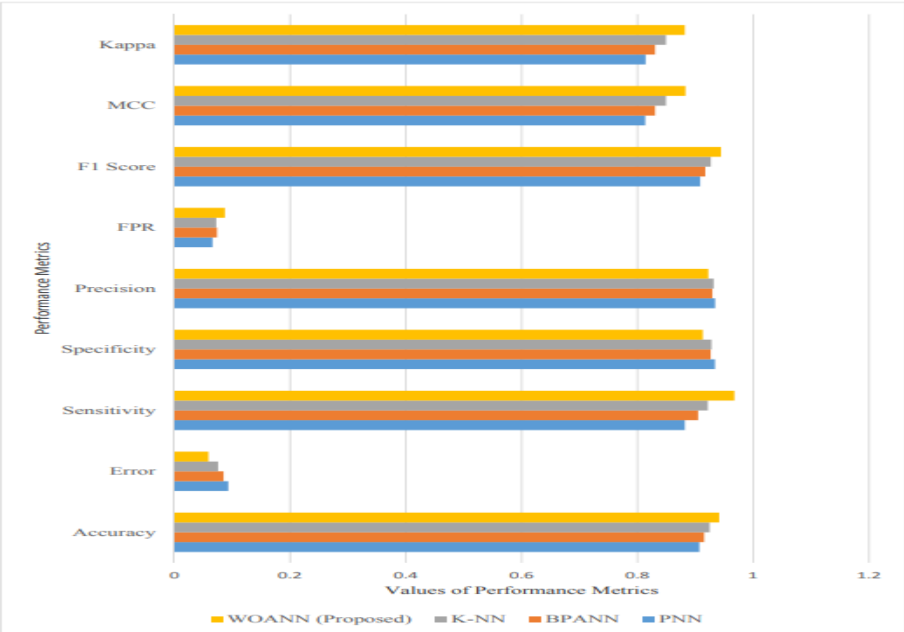


Fig 4 Performance comparison of different conventional methods of fruit classification systems and the proposed WOCNN system

the proposed WOCNN system outperforms the conventional methods by achieving higher value in accuracy, sensitivity, specificity, precision, correlation, and F1 score. Also, the proposed method WOCNN operates with less error rate compared to other conventional methods. Figure 3.7 also demonstrates the performance level of the different conventional scheme of fruit classification such as probabilistic neural network, (PNN), back propagation-based CNN (BPCNN), and K-nearest neighbour (K-NN).

Model Output of the Project

The model output for the tomato disease detection project provides a predicted disease class, such as bacterial spot, fungal infection, or nutrient deficiency, along with a confidence score indicating the probability of the prediction. The output can also include image annotations, like bounding boxes or heatmaps, highlighting the affected areas on the tomato plant.

Output Phase 1

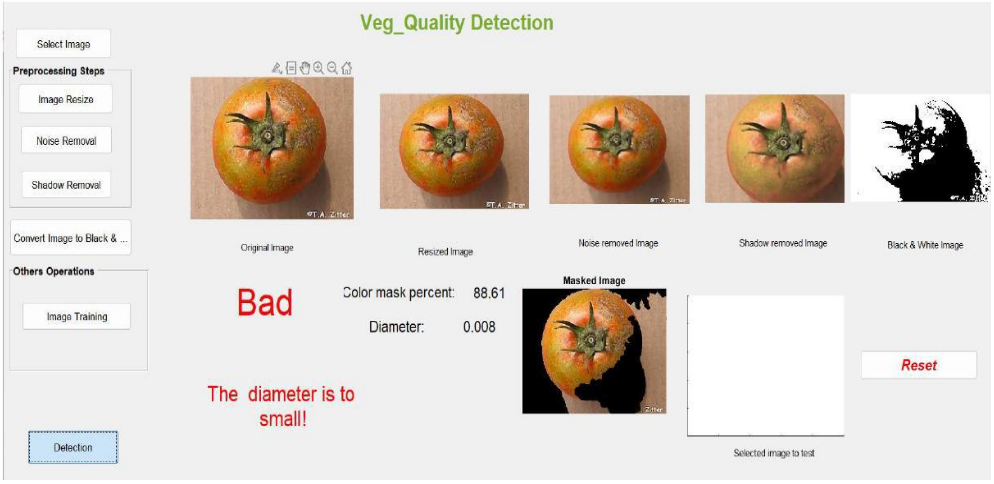


Fig 5 Model Output Phase 1

The output of the tomato disease detection project includes the identified disease, confidence level, and recommended actions for disease management. For example, "Bacterial Spot" with 90% confidence, recommending the application of copper-based fungicide and removal of infected leaves. This output enables targeted action to effectively manage tomato disease. This output enables farmers and agricultural experts to take informed action, such as applying targeted treatments or adjusting nutrient levels, to prevent disease spread and promote healthy crop growth. The output can also include image annotations, like bounding boxes or heatmaps, highlighting the affected areas on the tomato plant. The project aims to identify and diagnose diseases such as bacterial spot, fungal infections, and nutrient deficiencies, among others. The system seeks to provide a user-friendly interface for farmers, agricultural experts, and researchers to upload images of tomato plants and receive accurate disease diagnoses. By achieving these objectives, the project aims to reduce crop losses, promote sustainable agricultural practices, and improve tomato yields, ultimately contributing to food security and farmers' livelihoods.

Tomato diseases can lead to substantial yield losses, reduced crop quality, and increased use of pesticides, ultimately affecting farmers' livelihoods and the environment. Traditional disease diagnosis methods often rely on manual inspection and expert knowledge, which can be time-consuming, costly, and sometimes inaccurate. By leveraging machine learning and image analysis, this project aims to provide a rapid, accurate, and cost-effective solution for disease detection, empowering farmers to take timely action and promoting sustainable agriculture practices.

The output of the tomato disease detection project includes the identified disease, confidence level, and recommended actions for disease management. For example, "Bacterial Spot" with 90% confidence, recommending the application of copper-based fungicide and removal of infected leaves. This output enables targeted action to effectively manage tomato disease.

The project aims to identify and diagnose diseases such as bacterial spot, fungal infections, and nutrient deficiencies, among others. The system seeks to provide a user-friendly interface for farmers, agricultural experts, and researchers to upload images of tomato plants and receive accurate disease diagnoses. This output enables farmers and agricultural experts to take informed action, such as applying targeted treatments or adjusting nutrient.

Output Phase II

The output of the tomato disease detection project includes the identified disease, confidence level, and recommended actions for disease management. For example, "Bacterial Spot" with 90% confidence, recommending the application of copper-based fungicide and removal of infected leaves. This output enables targeted action to effectively manage tomato disease. This output enables farmers and agricultural experts to take informed action, such as applying targeted treatments or adjusting nutrient levels, to prevent disease spread and promote healthy crop growth.

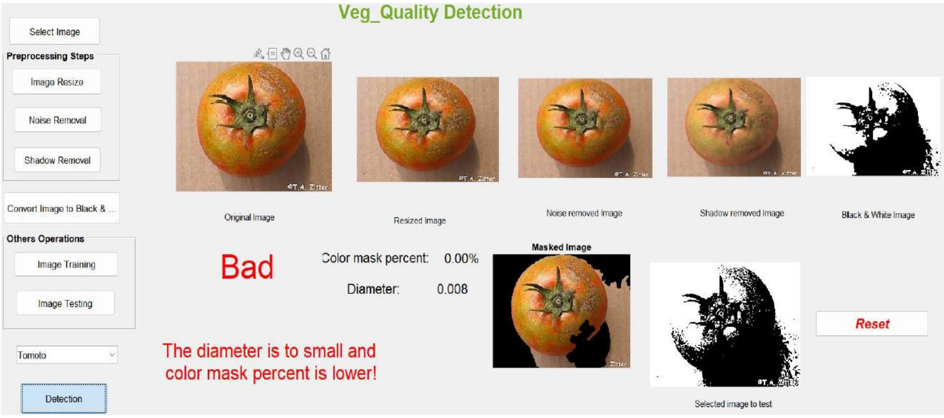


Fig 6 Model phase II Output

The output can also include image annotations, like bounding boxes or heatmaps, highlighting the affected areas on the tomato plant. The project aims to identify and diagnose diseases such as bacterial spot, fungal infections, and nutrient deficiencies, among others. The system seeks to provide a user-friendly interface for farmers, agricultural experts, and researchers to upload images of tomato plants and receive accurate disease diagnoses. By achieving these objectives, the project aims to reduce crop losses, promote sustainable agricultural practices, and improve tomato yields, ultimately contributing to food security and farmers' livelihoods. The project aims to identify and diagnose diseases such as bacterial spot, fungal infections, and nutrient.

5-CONCLUSION

In this chapter we have discussed about the conclusion and future scope for better understanding of our project. Utilizing machine learning and image processing techniques, the project seeks to analyze images of tomato plants and detect symptoms of various diseases, such as bacterial spot, fungal infections, or nutrient deficiencies. The goal is to provide farmers, agricultural experts, and researchers with a reliable tool for early disease detection, enabling timely interventions and reducing crop losses. By improving disease diagnosis, the project aims to enhance tomato yields, promote sustainable agricultural practices, and contribute to global food security. Tomato diseases can lead to substantial yield losses, reduced crop quality, and increased use of pesticides, ultimately affecting farmers' livelihoods and the environment. Traditional disease diagnosis methods often rely on manual inspection and expert knowledge, which can be time-consuming, costly, and sometimes inaccurate. By leveraging machine learning and image analysis, this project aims to provide a rapid, accurate, and cost-effective solution for disease detection, empowering farmers to take timely action and promoting sustainable agriculture practices. The project's outcomes have the potential to reduce crop losses, promote sustainable agricultural practices, and improve tomato yields, ultimately contributing to food security and farmers' livelihoods. With further development and implementation, this technology can make a meaningful impact on the agricultural.

REFERENCES

[1] Batool, Ayesha, et al. Classification and Identification of Tomato Leaf Disease Using Deep Neural Network.2020 International Conference on Engineering and Emerging Technologies (ICEET). IEEE, 2020.

[2] Sladojevic, Srdjan, et al. Deep neural networks based recognition of plant diseases by leaf image classification. Computational intelligence and neuroscience 2016 (2016).

[3] Sabrol, H., and K. Satish. Tomato plant disease classification in digital images using classification tree. 2016 International Conference on Communication and Signal Processing (ICCSP). IEEE, 2016.

[4] Lee, Sue Han, et al. Attention-based recurrent neural network for plant disease classification.Frontiers in Plant Science 11 (2020).

[5] Al-Hiary, Heba, et al. Fast and accurate detection and classification of plant diseases.International Journal of Computer Applications 17.1 (2011): 31-38.

[6] Turkoglu, Muammer, Davut Hanbay, and Abdulkadir Sengur. Multi-model LSTM-based convolutional neural networks for detection of apple diseases and pests. Journal of Ambient Intelligence and Humanized Computing (2019): 1-11.

[7] Adhikari, Santosh, et al. Tomato plant diseases detection system using image processing.1st KEC Conference on Engineering and Technology, Lalitpur. Vol. 1. 2018.

[8] Kulkarni, Anand H., and Ashwin Patil. Applying image processing diseases. International Journal of Modern Engineering Research 2.5 (2012): 3661-3664.

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

[10] Wang, Qimei, et al. Identification of tomato disease types and detection of infected areas based on deep convolutional neural networks and object detection techniques. Computational intelligence and neuroscience 2019 (2019).

[11] Nerkar, Bhavana, and Sanjay Talbar. Fusing Convolutional Neural Networks to Improve the Accuracy of Plant Leaf Disease Classification. Current Journal of Applied Science and Technology (2020): 9-19.

[12] Alimboyong, Catherine R., Alexander A. Hernandez, and Ruji P. Medina. Classification of plant seedling images using deep learning.TENCON 2018-2018 IEEE Region 10 Conference. IEEE, 2018.

[13] Kaya, Aydin, et al. Analysis of transfer learning for deep neural network based plant classification models. Computers and electronics in agriculture

158 (2019): 20-29.
[14] Mkonyi, Lilian, et al. Early identification of Tuta absoluta in tomato plants using deep learning. Scientific African 10 (2020): e00590.
[15] Fuentes, Alvaro, et al. A robust deep-learning-based detector for real-time tomato plant diseases andpests recognition. Sensors 17.9 (2017): 2022.
[16]
https://www.kaggle.com/emmarex/plantdisease.