Development of Automatic Tomato Plant Diseases Detection System based on Convolutional Neural Network

Raditya Rifqi Rayhan, Muhammad Nurul Puji, Winda Astuti
Automotive and Robotics Program, Computer Engineering Department
BINUS ASO School of Engineering, Bina Nusantara University, Jakarta, Indonesia
wastuti@binus.edu

Abstract

Tomatoes are an important commodity for both domestic and foreign needs. In Indonesia, tomatoes play an important role both need, people consume as fruit or to complete the ingredient in cooking. Therefore, maintaining the fertility of tomato plants is very important, so that tomato production continues to increase in Indonesia. The decrease in tomato production due to pest attack is an important problem in tomato plant care. In this work, a disease detection system in tomato plants will be developed using an intelligent system technique. The system involved image processing and intelligent technique. The system recognizes the kind of disease of the tomato plant based on the feature of the tomato leaf. Tomato plant disease detection system based on the leaf of tomato processing by image procession technique and Convolution Neural Network (CNN). The system is working properly, since it resulting the training and testing accuracy of 98% and 95%, respectively.

Keywords
Tomato plant disease; leaf tomato disease image; leaf image processing; Convolution Neural Network (CNN); Automated stand alone embedded system;

1. Introduction
Tomato fruit is a multipurpose commodity that has many functions as a vegetable, spices, cooking, table fruit, appetite enhancer, drinks, food coloring, cosmetics and medicinal ingredients. This is because the tomato commodity continues to grow in the arena of agriculture and international trade. From the government report, tomato production has been achieved by developed countries such as the United States and Europe. Indonesia, the national tomato production is not inferior to other countries in Asia, which are in third place after India and the Philippines, above Malaysia and Malaysia. Thailand. Tomatoes which are included in the vegetable group have a high commercial level and are more widely consumed in Indonesia in general. This is in line with increasing incomes, urban development and awareness of public nutrition. Of the 18 types of vegetable crops cultivated in the Central Java region (onions, garlic, cucumbers, long beans, potatoes, cabbage, tomatoes, carrots, chilies, eggplant, chicory, green onions, green beans, red beans, radishes, kale, spinach and chayote) only tomatoes are thought to be experiencing a shortage of production. Therefore, this is expected that in the nurture of tomato plants, plays important part to increase tomato production and anticipate production losses due to crop failure (Roydatul, 2014).

There are many works has been done in plant disease detection based on leaf of the plant. The Plant disease detection system applied based on quantifying the leaf image contour and modeled, the leaf model of the plant identify the disease of the plant using image processing method of Gray level Co-occurrence matrix (GLCM) and Support Vector Machines (SVM) method for disease identification, (Bharath Kumar et al., 2019). The image of the plant leaf is extracted based on image processing method, which consist of filtering, segmented image end extracted image based on GLCM. The extracted data is then used as input to the identification system that applied SVM to classify and identify the disease. They are resulting differentiate four different disease. Other research work on plant disease identification, they applied one of the machine learning method Histogram of an Oriented Gradient (HOG) to identify plant disease based on leaf-image capture using camera. Plant disease identification based on the leaf of the plant based on image processing and deep leaning network (Mohanty, Hughes, & Salathé, 2016). Deep learning also applied to identify the disease of the plant based on the leaf image (Mohanty et al., 2016) the image of the leaf collected based on the capturing camera.

In this paper, the automatic stand alone system for plant disease identification on the tomato plant based on leaf-image recognition system is developed. The system is design and develop. The plant disease identification system is based on intelligent system neural network. The convolution neural network (CNN), which is a relatively popular artificial intelligent, are introduced and proposed for fish identification system. CNN is simple and very powerful method. CNN work based on human brain principle (Aibinu, Shafie, & Salami, 2012). The proposed intelligent system-based leaf image captured by camera, used as input. The effectiveness of the proposed system is evaluated experimentally. The results show that the proposed technique has produced the better training and testing accuracy.

2. Proposed System

This paper considers two main works, which consist of automatic stand alone tomato plant identification system and tomato plan disease identification based on the leaf-image of the tomato plant. Automatic tomato plant disease identification based on computer and electrical design, which have two main functions, there is an automatic tomato plant disease identification for collecting the reference data and perform the decision pain level after identification. On the other hand, the pain level identification system based on image processing and artificial intelligence. Figure 1 shows a flow chart of automatic painful detection based on facial pain expression identification System.

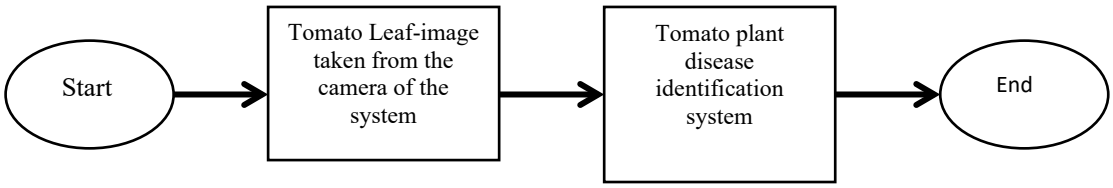

Figure 1. Flowchart tomato plant disease identification based on tomato leaf-image Identification System

These systems have two main parts, namely hardware and software part, respectively. The hardware part which consists of mechanical and electrical parts, has two main functions, taking the reference data and perform the tomato plant disease identification based on leaf-image.

As discussed previously, the hardware part which consists of electrical components. The mechanical part of the stand alone tomato leaf capturing system, consists of 10 main parts, namely Casing, Camera, Main Camera Holder, Raspberry Pi, PCB, Cooling Fan, Button and LED. This system uses the Raspberry Pi as the main component and the camera as a supporting component for taking pictures, as well as buttons and LEDs for taking input and output, as shown in Figure 2.

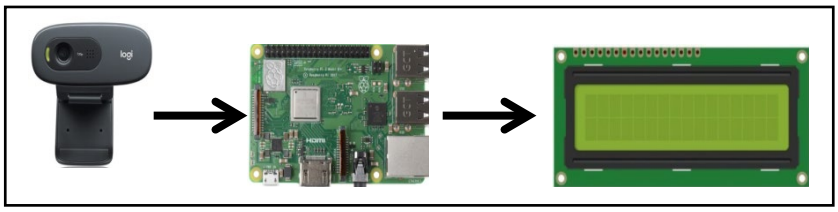

Figure 2. Embedded system for taking the picture of the patient

Figure 3 shows the proposed of the leaf-image of the tomato plant disease identification system, the system basically consists of three main components, namely image processing, data selection, and identification system. Image processing is used to obtain images using the data based and testing using the camera, data transferred to the computer and processed in order to obtain LabelImg. The processing result in the form of XML file which convert to the form of CSV. The data is then used as input to the identification process in the training phases. There are two important phase identification system, training phase and testing phase, respectively. In the training phase the data is applied to training the system, finding the model each of the expression data patterns. The model testing based on the testing data that captured in real time. Single shot detector (SSD) is applied as identification system to obtain the pain level of the person.

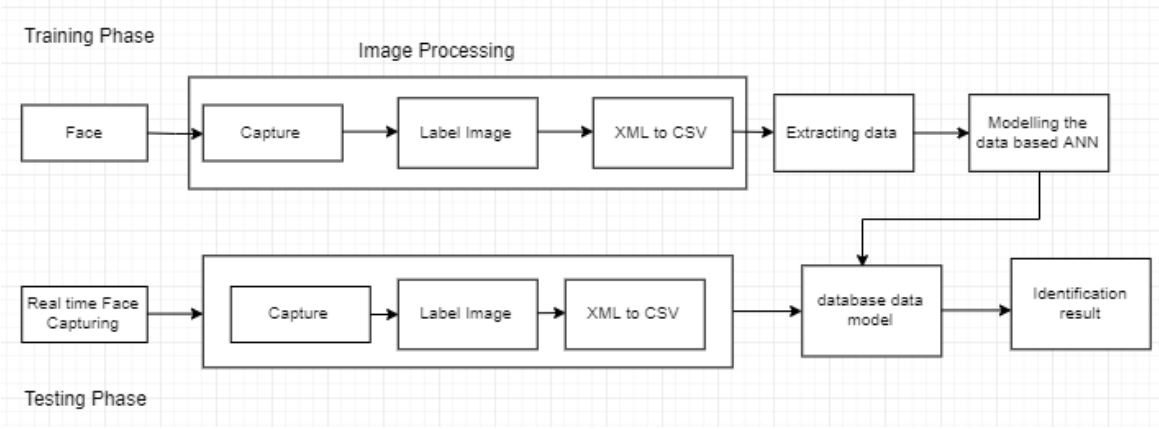

Figure 3. Block diagram of face expression identification system

Single Shot Detector *(*SSD) applied one shot to detect multiple objects in the picture using multibox. The system is faster in the algorithm when detecting objects and with a high degree of accuracy. High speed and high accuracy SSD in using images with relatively low resolution caused by several things such as convolutional filters, which are increasingly reduced to predict which objects are in which category(Shi, Bao, & Tan, 2019).

SSD accuracy is achieved by using several boxes with size and aspect ratio for the detection of different objects. This makes the system apply some of the filter area ratios to several feature layers at the final stage of the system which helps in making detection at various scales. SSD has a basic network called VGG-16 which is accepted by conventional multibox layers. The basic network VGG-16 is a standard convolutional neural network (CNN) architecture for the classification of high-quality images but without a final classification layer(Xie, Han, Liao, & Shi, 2017)(Arge & Mage, 2015) . VGG-16 is used for feature extraction. The following will discuss about the architecture of the SSD listed in Figure 8 below. SSD consists of two main networks, basic network to detect the feature and additional Convolutional Layers to detect objects, respectively, as shown in Figure 4.

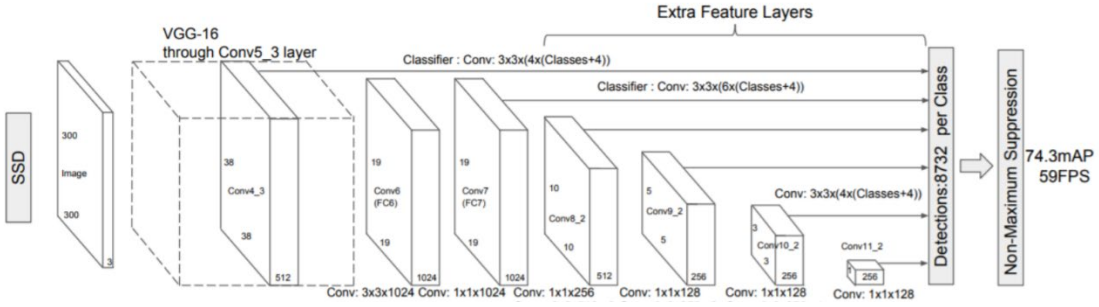

Figure 4. SSD architecture

3. Experimental Result

In this experiment, the data applied to develop the system is based on the Kaggle database. Tomato disease plant identification is developed using 11 types of tomato plant disease from Kaggle. Moreover, the data are identified based on the leaf-image captured using a camera, namely the Tomato Bacterial Spot, Tomato Early Blight, Tomato Late Blight, Tomato Leaf Mold, Tomato Septoria Leaf Spot, Tomato two Spotted Spider Mite, Tomato Target Spot, Tomato Mosaic Virus, Tomato Yellow Leaf Curl Virus, and Tomato healthy. Total data of overall total data for training, testing and validation Kaggle are 27,023. The number of data applied as training, testing and validating data are 8861, 14531 and 3631, respectively. The camera types of the Logitech C310 for shooting problems because this type of camera can meet the needs needed to support the performance of this work.

In the testing phase of the system run in real-time where the image will be processed directly when the tool is instructed to identify the type of disease of the tomato plant through the feature of the tomato leaf image. For more detailed ways of working can be seen through Figure 5 below. The laptop used to run this system uses the Nvidia GTX 1660 as its graphics processor and has 16 GB of random access memory (RAM).

Figure 5. Tomato plant disease identification system develops with laptop

Laptop used to develop the system before its downloaded to the Rusberry, the system run in the Nvidia GTX 960m laptop. The system is then build into stand alone embedded system, as shown in Figure 6 in order to make the system portable and easy to apply everywhere.

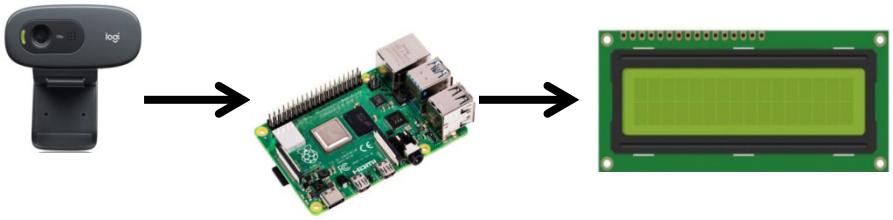

Figure 6. Experiment setup

3.2 Experiment Result and Discussion

The experiment was carried out based on 11 type diseases of the tomato plant. Its applied 27023 images from a total of 11 types of the disease from the different contour of the leaf-image, as shown in Figure 7. The all of these differences are used to support accurate results in recognizing every expression of the subject that will exist later in the field. SSD method is applied to extract the photos before used as input to the identification system. The size of the photos is not exceed 150 KB that support training performance and faster identification.

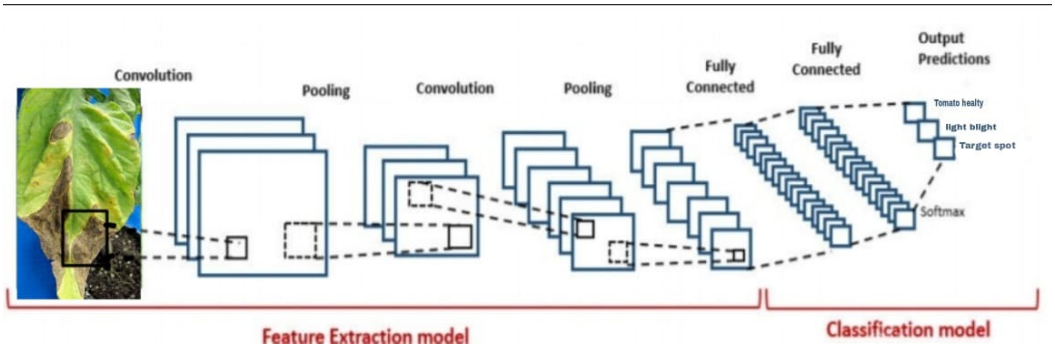

Figure 7. Tomato-leaf image identification using CNN

The results of the training are then tested using real faces in real time and show that the accuracy obtained is above 90% per level. Figures 8 below show the test results of the above training.

	Tomato Bacterial Spot
	Tomato Early Blight
	Tomato Late Blight
	Tomato Septoria Leaf Spot

	Tomato Leaf Mold
	Tomato Septoria Leaf Spot
	Tomato Two Spotted Spider Mite
	Tomato Target Spot

	Tomato Yellow Leaf Curl Virus
	Tomato healthy

Figure 8. Identification result of real time detection

From the results of trials as shown in Figure 8, that have been conducted, it can be concluded that the SSD system can detect each level well with an accuracy level above 98% and 95%, for the training and testing accuracies, respectively. Accuracy can also be affected by lighting, subject distance from the camera and many other things. If there is more than one subject in the shooting area, other subjects will be detected as well, so it is highly recommended that there is no more than one subject in one shooting area. Poor lighting can also cause a system failure to see the subject in front of the camera. There are also conditions where the shadow covering the leaf of the tomato plant is too much so that the identification value becomes less accurate plus the lack of training data that uses subjects with front hair that is too long to narrow the leaf identification area. The SSD method is proven to have a high identification speed and is fast, aside from the reason that the SSD method can be used on raspberries but also has a weakness where more than one result can be detected on one image.

4. Conclusion

The process of tomato disease plant identification based on the leaf tomato image discussed in this work. At this time, the process of identify of the type of the tomato plant diseases is still done manually, which requires a sometime faults the identification result and errors can occur due to the type of the disease and furthermore, in the determined of the plant medicine and the treatment of the disease.

With this identification system for tomato plant disease identification based on the leaf-image identification system, it can reduce the occurrence of errors when give the treatment to the tomato plant, making it easier to identify the disease and do the treatment to the plant, with that hope that the tomato production will keep maintain and increasing.

References

Aibinu, A. M., Shafie, A. A., & Salami, M. J. E., Performance analysis of ANN based YCbCr skin detection algorithm. In *Procedia Engineering* (Vol. 41). https://doi.org/10.1016/j.proeng.2012.07.299, 2012.

Arge, F. O. R. L., & Mage, C. I., Very Deeep Convolutional Networks for Large-Scale Image Recognition. In *conference paper at ICLR 2015* (pp. 1–14), 2015.

Bharath Kumar, R., Balakrishna, K., hreyas M, S., Onu, S., Anirudh, H., & Abhishek B, J., SVM Based Plant Diseases Detection using Image Processing. *International Journal of Computer Sciences and Engineering*, *7*(5), 1263–1266. https://doi.org/10.26438/ijcse/v7i5.12631266, 2019.

Mohanty, S. P., Hughes, D. P., & Salathé, M., Using deep learning for image-based plant disease detection. *Frontiers in Plant Science*, *7*(September), 1–10. https://doi.org/10.3389/fpls.2016.01419

Roydatul, Z. (2014). OUTLOOK KOMODITI TOMAT. *Pusat Data Dan Sistem Informasi Pertanian*. https://doi.org/ISSN 1907-1507, 2016.

Shi, W., Bao, S., & Tan, D., Applied sciences FFESSD : An Accurate and E ffi cient Single-Shot Detector for Target Detection. *Applied Sciences*, 2019.

Xie, X., Han, X., Liao, Q., & Shi, G., Visualization and Pruning of SSD with the base network. In *Association for Computing Machinery* (pp. 2–6), 2017.

Biographies

Raditya Rifqi Rayhan, final year student from BINUS ASO School of Engineering, Bina Nusantara University, in 2021 majoring in Automotive and Robotics Program, and Currently he is planning for pursuing his carier in industry. Her research mainly on Intelligent system, data base, and embedded system.

Muhammad Nurul Puji, S.Si., M.Si. Graduated in Master degree from Physics Science - Universitas Indonesia (UI) and Bachelor degree from Physical - Institut Teknologi Sepuluh Nopember (ITS). His expertise is in field of Instrumentation Physics, Embedded System and Volumetric Magnetic Induction Tomography (VMIT). Currently he is working as full time lecturer in Automotive and Robotics Engineering Program (ARE) at BINUS ASO SCHOOL OF ENGINEERING (BASE), Serpong - Indonesia.

Winda Astuti Received the B.S. Degree in Electrical, Control and Computer engineering from National Institute of Technology, Bandung, Indonesia, in 2000 and the M.S. Degree in Electrical and Computer Engineering from International Islamic University Malaysia, Malaysia, in 2008. She is finishing the Ph.D. degree in Mechatronics engineering at International Islamic University Malaysia, Malaysia. Previously, she was a lecturer at Electrical Department at University of Jenderal Soedirman (2007-2009). Currently, She is a full lecturer in the Department of Automotive and Robotics, Binus Aso School of Engineering (BASE), Bina Nusantara University. Her main research interests include control systems theory and applications, intelligent systems, signal and image processing.