

# A Recognition Method for Cassava Phytoplasma Disease (CPD) Real-Time Detection based on Transfer Learning Neural Networks

Irma T. Plata<sup>1</sup>, Edward B. Panganiban<sup>2</sup>  
Darios B. Alado<sup>3</sup>

Faculty, CCSICT, Isabela State University  
San Fabian, Echague, Isabela, Philippines

Allan C. Taracatac<sup>4</sup>, Bryan B. Bartolome<sup>5</sup>  
Freddie Rick E. Labuanan<sup>6</sup>

Technical Staff, CCSICT, Isabela State University  
San Fabian, Echague, Isabela, Philippines

**Abstract**—Object detection technology aims to detect the target objects with the theories and methods of image processing and pattern recognition, determine the semantic categories of these objects, and mark the specific position of the target object in the image. This study generally aims to establish a recognition method for Cassava Phytoplasma Disease (CPD) real-time detection based on transfer learning neural networks. Several methods and procedures were conducted, such as the testing of two methods in transmitting long-distance high definition (HD) video capture; establishment of a compact setup for a long-range wireless video transmission system; the development, testing of the real-time CPD detection and quantification monitoring system, providing the comparative performance analysis of the three models used. We have successfully custom-trained three artificial neural networks using transfer learning: Faster Regions with Convolutional Neural Networks (R-CNN) Inception v2, Single Shot Detector (SSD) Mobilenet v2, and You Only Look Once (YOLO) v4. These deep learning models can detect and recognize CPD in actual environment settings. Overall, the developed real-time CPD detection and quantification monitoring system was successfully integrated into the wireless video receiver and seamlessly visualized all the incoming data using the three different CNN models. If the consideration is the image processing speed, YOLOv4 is better compared to other models. But, if accuracy is the priority, Faster R-CNN inception v2 performs better. However, since CPD detection is the main purpose of this study, the Faster R-CNN model is recommended for adoption to detect CPD in a real-time environment.

**Keywords**—Cassava phytoplasma disease; faster regions with convolutional neural networks (R-CNN) inception v2; you only look once (YOLO) v4; object detection; precision agriculture

## I. INTRODUCTION

The agricultural industry plays an important role in the economy. Plant illness is also caused by climatic circumstances, exacerbated by the exponential trend of population growth. The major challenges of sustainable development include reducing pesticide use, the expense of preserving the environment, and the cost of building quality. Exact, exact, and timely decisions may reduce pesticide use [1]. Cassava is a key crop that has been produced in the Philippines. Nowadays, innovation is commonly used for plant disease prediction. The concept of image processing combined with information mining improvements aids in identifying plant diseases. With the development of intelligent

devices, the data bulk on Internet has grown with high speed. As an important aspect of image processing [2][3][4], object detection has become one of the popular international research fields. In recent years, the powerful ability with feature learning and transfer learning of Convolutional Neural Network (CNN) has received growing interest within the computer vision community, thus making a series of important breakthroughs in object detection [5][6]. So, it is a significant application to apply CNN to object detection for better performance.

Real-time object detection is the task of doing object detection in real-time with fast inference while maintaining a base level of accuracy [7]. Object detection technology aims to detect the target objects with the theories and methods of image processing and pattern recognition [8][9], determine the semantic categories of these objects, and mark the specific position of the target object in the image. It is a very challenging task in the actual application to use computer technology to detect objects automatically. Complex background, noise disturbance, occlusion, low-resolution, scale, and attitude changes, and other factors will seriously affect the object detection performance. The conventional object detection method was based on the hand-crafted feature. It is not robust to illumination change, lacking good generalization ability. Using Google's TensorFlow and YOLOv4 as well as transfer learning, we were able to custom-trained three separate artificial neural networks [10]. These deep learning models can detect and recognize CPD in actual environment settings.

This study established a recognition method for Cassava Phytoplasma Disease (CPD) real-time detection based on transfer learning neural networks. Specifically, this paper is presented to test the two methods in transmitting long-distance high definition (HD) video capture. Likewise, it established a compact setup for a long-range wireless video transmission system. Furthermore, the project developed a real-time CPD detection and quantification monitoring system, performed a wireless video transmission test, and compared the test results in each model used.

As compared with the recent papers and technologies related to plant disease detection, this paper presented a disease detection technology specifically for Cassava

Phytoplasma Disease. It demonstrated which real-time detection technology is the most appropriate to use, which is more practical and can have a great impact on society. This research provides a visual object identification framework capable of processing pictures at high detection rates while processing images at a fast pace.

Hence, the authors conducted a thorough experiment on what innovative technology will be used. To befittingly implement the selected technology to our partner agency, the EDCOR, a cassava development cooperative located in San Guillermo, Isabela. Their cassava farms are situated in remote locations where internet connectivity is scarce. We developed a remote image processing server system utilizing digital First-Person View (FPV) [11]. This remote server can perform real-time image processing for CPD detection and recognition with a dedicated ground station for long-range high-definition video transmission without using the Internet.

The next sections discussed the related works, methods, and results that transpired during the conduct of the study. It explained the details on how the entire project was done and the results of the undertakings.

## II. LITERATURE REVIEW

Cassava plants exhibiting characteristic phytoplasma signs, such as witches' broom, general stunt, chlorosis, distortion, and decreased size [12], were seen in farms near San Guillermo, Isabela. CPD is a severe danger to cassava producers' food security [13].

Object detection is a critical computer vision problem that involves detecting visual objects in digital photos of a specific type (such as humans, animals, or autos). Object detection aims to create computational models and approaches that give one of the most fundamental bits of data required by computer vision applications [14]. Object detection is essential in identifying the disease type in the specific video sequence performed in agricultural farming. Object detection reduces computing time while increasing detection accuracy [15].

Several popular deep learning-based object identification algorithms are available that can accurately identify which section of a farm field is infected with Cassava Phytoplasma Disease. Some frameworks will need a lot of computing power, while some will provide less accuracy. By surveying several neural network frameworks, the authors identified the following frameworks with the best performance in object detecting technology.

Transfer learning is the enhancement of learning in a new task by transferring knowledge from a previously learned related activity [16]. While most machine learning algorithms are designed to solve particular tasks, the development of algorithms that promote transfer learning is a continuing area of study in the machine-learning field. The following are the transfer learning neural networks considered in this paper:

### A. *Faster Regions with Convolutional Neural Networks (R-CNN) Inception v2*

The Region Proposal Network (RPN) is a fully convolutional network that predicts object limits and

objectness scores at each place simultaneously. RPNs are trained from start to finish to create high-quality region suggestions, which Fast R-CNN uses for detection [17]. RPN and Fast R-CNN may be taught to share convolutional features using a simple alternating optimization. RPN was added to the design of faster R-CNN. It means that it uses a quick neural network to handle a sluggish search selection process.

RPN comes after the last convolution layer of a Convolutional Neural Network. Inception V2 was a module created to minimize the complexity of a convolution network. This module causes the convolution network to be broader rather than deeper. There are three types of modules in Inception V2 [18].

### B. *You Only Look Once (YOLO) v4*

YOLO is a cutting-edge, real-time object detecting technology. It is a real-time object identification system that recognizes several items in a single frame. YOLO takes a completely new approach to earlier detection methods. It uses a single neural network to process the entire picture. This network separates the picture into regions and predicts each bounding boxes and probabilities. The projected probabilities are used to weigh these bounding boxes [19]. YOLO offers excellent real-time performance in multi-scale object recognition [20].

### C. *Single Shot Detection (SSD)*

Single Shot Detection is a technique for detecting objects in a single shot. By capturing a single photograph, you can analyze many items. A single frame is used to recognize and analyze several items in a picture. Compared to Convolutional Neural Networks, this is a considerably faster analysis. For  $p$  channel analysis, a feature layer of  $m*n$  is obtained. For each of the  $k$  areas, a bounding box is generated. SSD is sometimes called a Multibox detector since it computes each bounding box and offsets relative to the initial bounding boxes [21].

## III. METHODS

### A. *Conceptual Framework*

The different processes and methodologies in the wireless video transmission system, its establishment, the development of a real-time CPD detection monitoring system, the performance testing were realized based on the conceptual framework of the study as presented in Fig. 1. The Real-time CPD detection and quantification Framework Model.

### B. *Methods in Transmitting Long-distance High Definition Video Capture using the Long-Range Wireless Video Transmission System*

We have used two methods in transmitting long-distance high definition (HD) video capture. These methods involve using 2.4Ghz and 5Ghz frequencies. Both had their advantages and disadvantages. 2.4GHz offers a wide network coverage and is superior at bypassing substantial impediments like trees. However, it has a smaller data rate and is more susceptible to interference; more devices normally utilize this frequency. 5Ghz, on the other hand, offers a greater data rate and is less susceptible to interference; it is typically used by

fewer devices. However, it has a smaller coverage area and is not performing well when penetrating solid/thick layered obstacles.

To implement the 2.4Ghz concept, we adapted the method called Wi-Fi Broadcast illustrated in Fig. 2. Wi-Fi-broadcast activates the monitor mode on the Wi-Fi cards. This mode allows you to transmit and receive arbitrary packets without having to associate them with anything. Furthermore, it is possible to get incorrect frames (where the checksum does not match). In this method, a real unidirectional connection is produced, simulating the benefits of an analog link. They comprise the following: the transmitter broadcasts data regardless of whether or not there are any receivers nearby. As a result, there's no possibility of video stalling due to a loss of association; the receiver gets video as long as it's within range of the transmitter. The video quality declines as it moves out

of range, but it does not stall. Even if frames are incorrect, they will be shown rather than discarded; the classic "one broadcaster – numerous receivers" method will function right out of the box. Bystanders who wish to view the video stream on their devices only need to "shift to the correct channels"; Wi-Fi-broadcast permits the simultaneous use of numerous low-cost receivers and the combining of their data to enhance the likelihood of correct data reception. This so-called software diversity enables the use of identity and complementary receivers to increase dependability (imagine one reception with a 360° omnidirectional antenna and multiple directional antennas for long distances, all functioning in tandem). To archive high dependability at minimal bandwidth needs, Wi-Fi-broadcast employs Forward Error Correction. It can recover packets that have been lost or corrupted at the receiver.

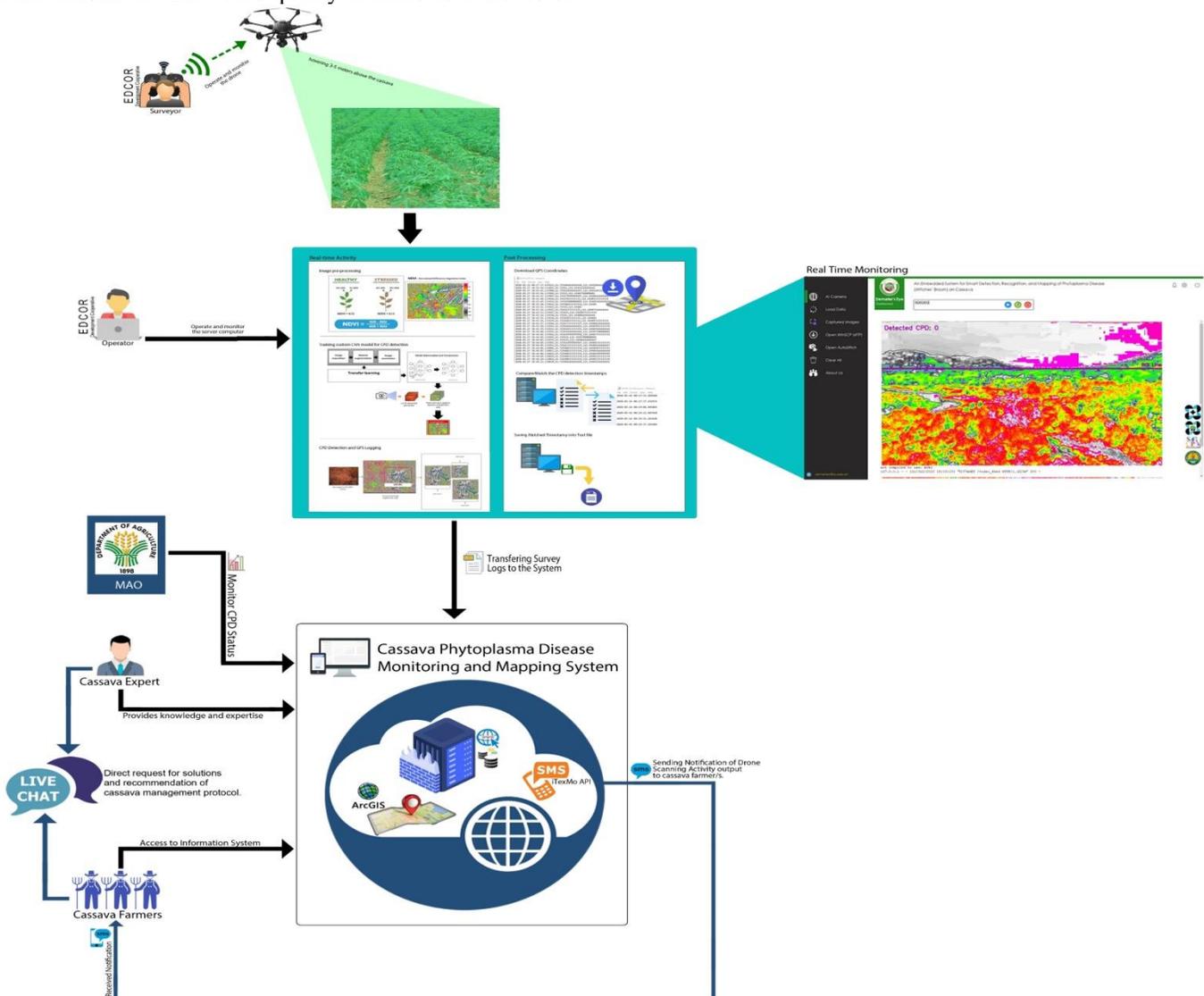


Fig. 1. The Structural Framework of Faster R-CNN(a), YOLOv4(b), SSD Mobilenet v2(c).

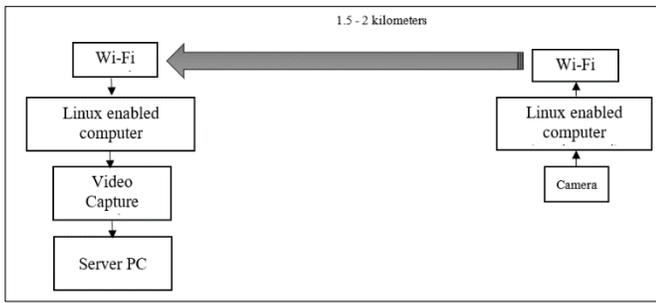


Fig. 2. Wi-Fi Broadcast Block Diagram.

Using standard Wi-Fi dongles with Atheros AR9271 802.11 chip for the transmitter and receiver, and a 2.4Ghz Yagi antenna for our ground station presented in Fig. 3 and 4. This setup can process camera feed from either USB Video or through RTSP/TCP URLs.

The ground station uses a sturdy tripod with a height of 2.5 meters. The Wi-Fi dongle, Yagi antenna, and a single board computer are securely installed on it, as shown in Fig. 5 and 6.



Fig. 3. Atheros AR9271 802.11 Wi-Fi Dongle.



Fig. 4. 2.4Ghz Yagi Antenna.

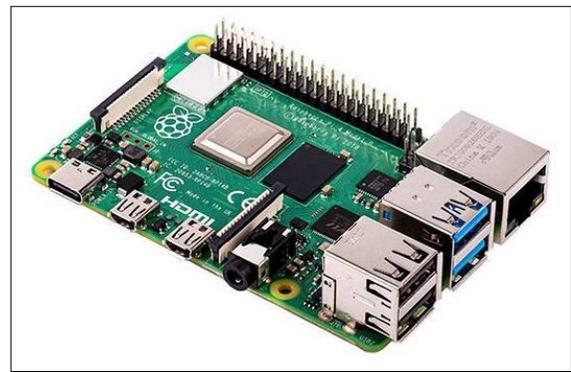


Fig. 5. Single Board Computer; Raspberry pi 4.



Fig. 6. A 2.5 Meters Tripod Stand where the Wi-Fi Dongle, Yagi.

On the other hand, we used a commercially available FPV high-definition video transmitter using 5Ghz presented in Fig. 7, which also supports HDMI all format video input up to 1920\*1080@60fps, and output is up to 1920\*1080@30fps. A 5dB omnidirectional antenna is installed on the ground unit (standard). According to the manufacturer, the effective transmission distance is over 2km with an output delay of 80ms/0.08s.



Fig. 7. Commercially Available full HD Digital Video System.

The receiver/ground unit is also installed on the tripod, just beside the 2.4Ghz Yagi antenna. This method takes advantage of the commercially available HD Video TX/RX and its dedicated windows application. The commercially available video TX/RX works perfectly fine using its dedicated application (Insight.exe) on a windows computer. However, the downside is that we cannot tap on the video feed from the transmitter into our developed system. Given that its dedicated application is not open-source and there is no available software development kit (SDK) for this device, it is hard to analyze and reverse engineer. On the other hand, the RTSP URL that supposedly works based on its documentation turns out to be incorrect (or, if not, maybe they no longer support this feature).

The alternative concept uses the “HD Video Capture” method, as illustrated in Fig. 8, which works by connecting the NDVI camera through HDMI into the Video TX and connected to the Video RX wirelessly (5Ghz band). The Video RX will also act as a Wi-Fi access point using the 5Ghz band. A lightweight, small single-board computer (SBC) running Windows 10 Pro x64 with Insight.exe configured will be linked to the Video RX over Wi-Fi (5Ghz) and display the live camera feed from the TX side. Then a video capture device will be used to mirror the preview display on the SBC

and send it to our server computer via wired connection in digital HD resolution (720P - 1080P) in real-time (or near real-time).

After which, our developed Demeter’s Eyes Monitoring System will then be able to adapt to this alternative method and process each frame from the incoming digital HD resolution camera pass through our custom-trained CNN model for CPD detection and quantification. These two different videos transmission configurations work independently of each other, and only one setup can be used at a time since there is only one camera installed to the transmitter.

### C. Establishment of a Compact Setup for a Long-Range Wireless Video Transmission System

The long-range wireless video transmission system is equipped with 16 megapixels NDVI camera, GPS module, a single-board computer, and a swappable video transmitter (2.4Ghz and 5Ghz), as shown in Fig. 9. This compact setup also includes a portable power supply, mini-fan, and LED light indicators. The LED lights indicate the status of each component on this embedded system. It is very useful when conducting pre-flight checkups and during troubleshooting.

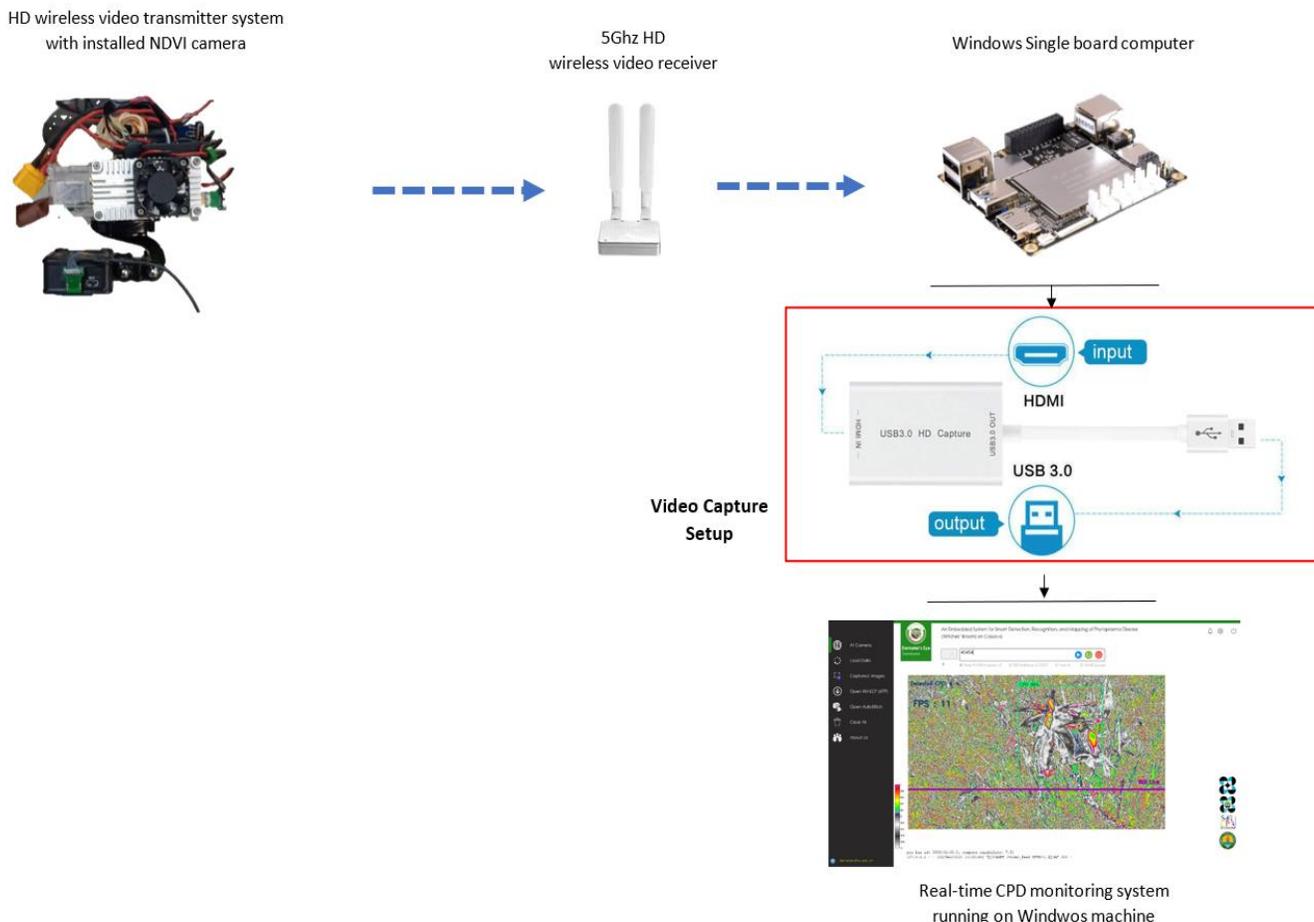


Fig. 8. Video Capture System Layout.

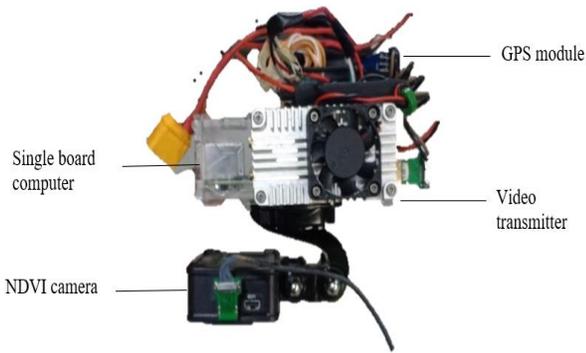


Fig. 9. HD Wireless Video Transmitter System with Installed NDVI Camera and GPS Module.

#### D. Development and Testing of the Real-Time CPD Detection and Quantification Monitoring System

The real-time CPD detection [22][23] and quantification system work by switching on the NDVI camera, video transmitter and receiver (including the single-board computer from the ground station), GPS logging module (will automatically start logging time stamps and GPS coordinates once switched on and successfully linked to a satellite) and, the drone itself. After which, the video transmitter and receiver will automatically link each other (the green light indicator for data on the video receiver must be steady).

On the server-side, the CPD real-time monitoring system should now be able to detect and preview incoming live camera feed from the video transmitter, as shown in Fig. 10.

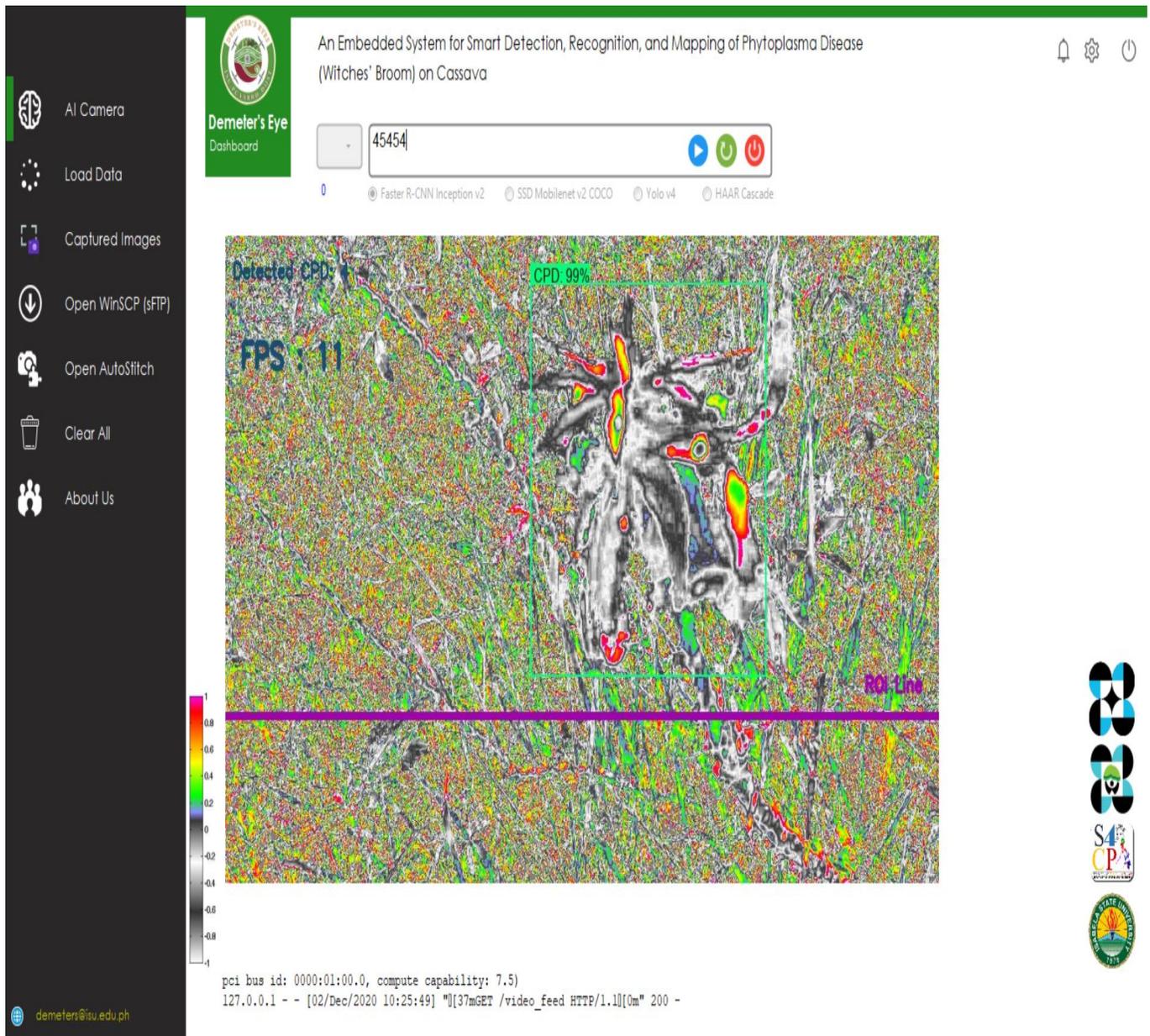


Fig. 10. Real-Time CPD Detection and Quantification System Dashboard.

Once the video link is established, the surveyor can now start hovering above the cassava field (3-5 meters above the cassava). While hovering over the cassava field, the CPD real-time monitoring system performs image pre-processing procedures on each incoming frame. This includes applying a Fastie color map to visualize the NDVI values in the image. The pre-processed frame(s) is passed through our CNN model (custom trained CPD detection model using transfer learning with Faster RCNN for prototype 2, SSD MobileNet V2 [1] COCO for prototype 3, and YOLO v4 for prototype 4) for CPD detection and quantification based on the TensorFlow object detection API. The real-time CPD detection is displayed on the system preview window with an overlaid running count of CPD detected. The system also automatically records each CPD detected timestamp (the latency between the transmitter and receiver was considered; thus, an adjustment with the time delay was done).

After completing the survey and the drone has returned, the GPS logging module will automatically connect to the access point located on the operator side. This will allow the operator to download GPS coordinates and corresponding timestamp Logs from the drone onboard GPS logging module through sFTP and upload it into the server computer. The system will now compare/match the CPD detection timestamps from the server computer and the timestamps from the GPS logging module. For each timestamp that matches, save its corresponding GPS coordinate into a .txt file. This .txt file will be uploaded into the web-based information system for visualization. These processes are illustrated in our conceptual framework shown in Fig. 1.

#### E. Testing the Long-range Wireless Video Transmission System

To verify the actual performance of the video transmission system, we conducted a stress test to determine the following:

working time per battery charge of the NDVI camera; working time per battery charge of the video receiver; the actual working distance of the video TX/RX; actual latency of the video transmission; actual latency of the video transmission when adapted into the monitoring system.

We conducted the test at an open field inside the Isabela State University located at San Fabian, Echague, Isabela, as shown in Fig. 11 and 12. The test location is ideal since it has an almost identical topography with the actual cassava farms. It has a combination of wide-open fields with surrounding trees in various heights and densities.

#### F. Portable Handheld Prototype Field Testing

With favorable weather, we conducted the onsite pilot testing of our prototype at Villa Teresita, San Guillermo Isabela, on a cassava farm owned by one of the EDCOR members between 10:00 in the morning up until 3:00 in the afternoon. The planted cassava was at around three months old at the time of testing. We used PVC pipes to hold the prototype steady, facing the camera down at a height of 2-3 meters from the top-most part of the cassava. Thus, mimics the drone altitude and flight movement.

Fig. 13 shows the actual field testing of the prototype. This was performed in the cassava farm in Villa Sanchez, San Guillermo, Isabela. This testing makes sure that the prototype works well as per its functionalities.

#### G. Drone and Prototype Payload Field Testing

After the initial test, we embedded the prototype as a payload into our custom-built drone and went back to the same site to conduct another series of tests. This milestone is presented in Fig. 14 and 15.

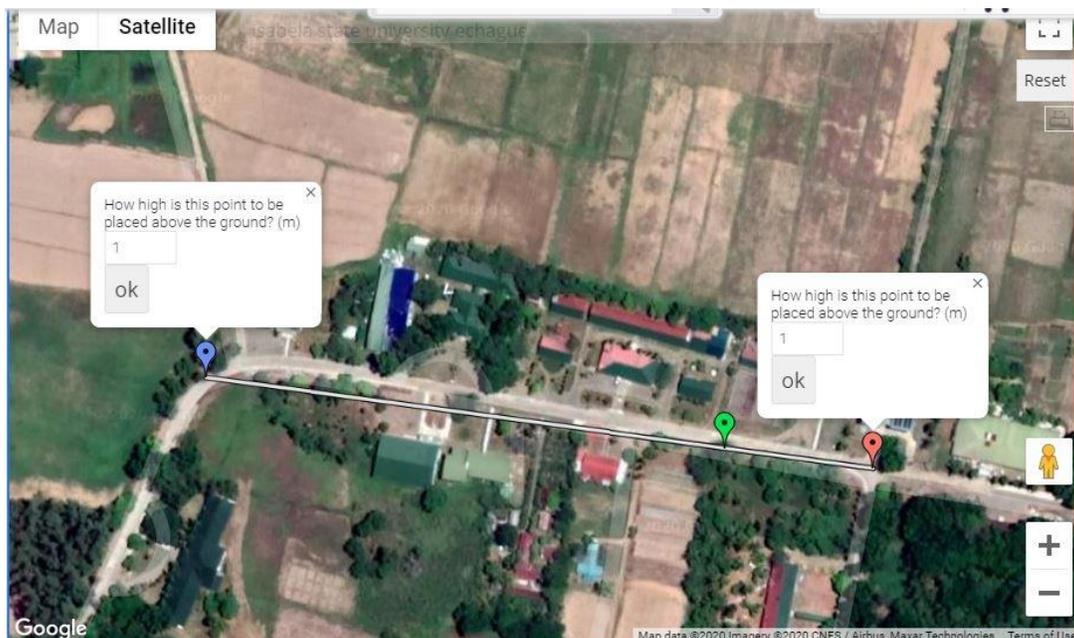


Fig. 11. Testing Site.



Fig. 12. Actual Wireless Video Transmission Test Conducted.



Fig. 13. Actual Field Testing of the Prototype.



Fig. 14. Installed Prototype as Payload underneath the Drone.



Fig. 15. Actual Field Testing using the Drone with the Prototype Payload on it

#### H. Comparison of the Test Results in each Model used

The ground-truthing [24] procedure was done on a selected area from a cassava farm with identified CPD infection. Our plant pathologist selected two lanes of planted cassava by which our prototypes camera can cover at one pass at a height of 2-2.5 meters. The cassava planted on the selected test area was counted manually. We got a total of 40 cassava stalks, wherein our plant pathologist identified 18 as CPD infected, and 22 were healthy/normal/or with a different disease. As the prototype passes through each cassava plant, our plant pathologist verifies if it can detect CPD infected or has no detection.

### IV. RESULTS AND DISCUSSION

For objective one, on testing two methods in transmitting long-distance high definition (HD) video capture. The wireless video transmission results show that the working time per

battery charge of the NDVI camera (while installed in the video TX/RX) is capable of reaching three hours on preview mode with a remaining one bar on the battery indicator. Actual working distance of the Video TX/RX. Both 2.4Ghz and 5Ghz setups reached 900 meters in the actual ground test with few trees in between. Also, both setups lose connection after a total loss of line of sight between the TX and RX. The actual latency of the video transmission is at 90-100 milliseconds.

The second aim is to construct a small configuration for a long-range wireless video transmission system. The tests conducted between the two wireless configurations showed promising results with the given environmental condition during the field testing. However, we opt to use the 5Ghz setup moving forward simply because it is very straightforward to deploy both its transmitter and receiver. Also, given the fact that the end-users are not technically knowledgeable on setting up the 2.4Ghz configuration, which may be too complex for non-technical users, using it may pose problems during the turnover and maintenance of the system.

The creation and testing of a real-time CPD detection and quantification monitoring system and wireless video transmission test performance are goals 3 and 4. The developed real-time CPD detection and quantification monitoring system was successfully integrated into the wireless video receiver and seamlessly visualized all the incoming data using the three different CNN models, as shown in Fig. 16.

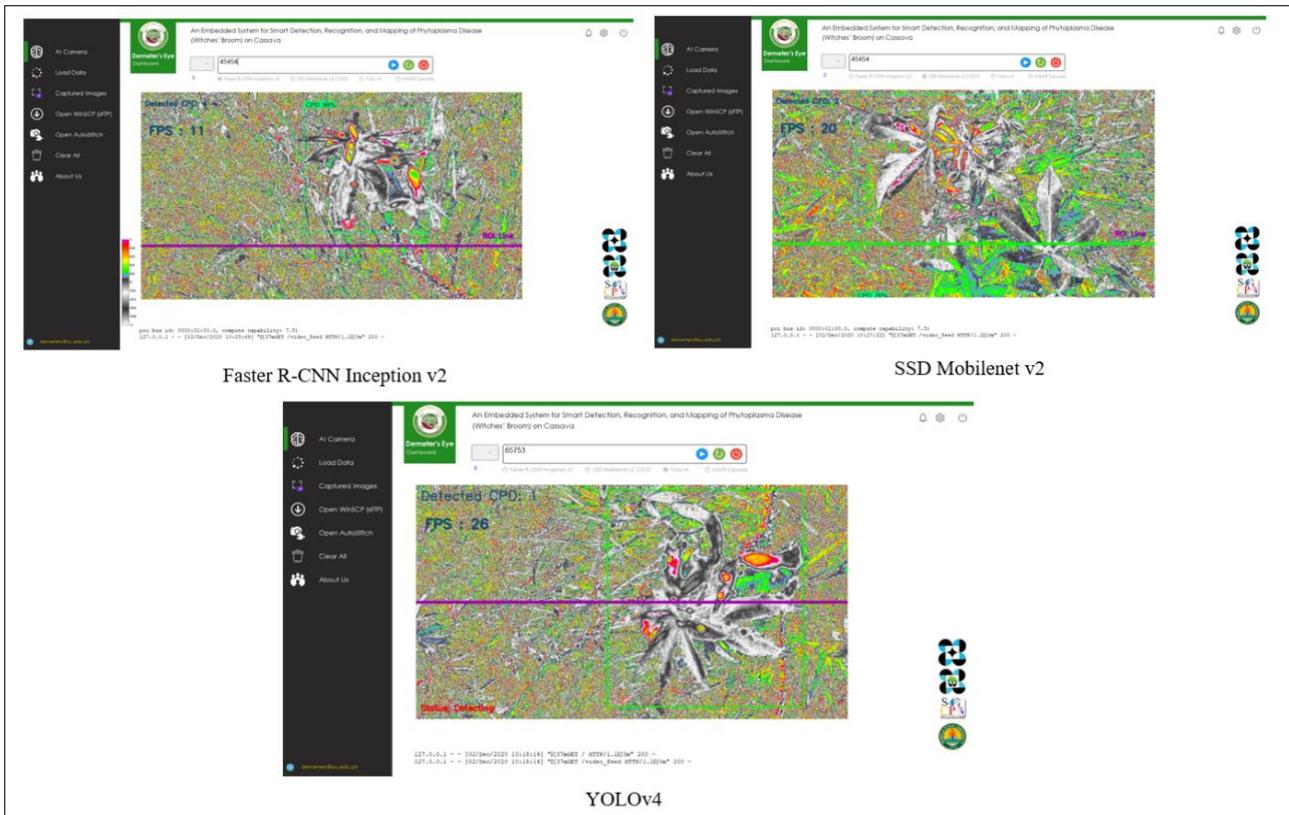


Fig. 16. Deploying the Three Custom Trained CNN Models (Faster R-CNN, SSD Mobilenet v2, and YOLOv4).

For objective 5, the comparison of the test results in each model used, the tabulated results using each custom-trained CNN model at different frame resolutions presented in Tables I to VI show that the Faster R-CNN model has the highest detection accuracy across different frame resolutions with a top detection rate of 13 out of 18 CPD samples. However, this has the lowest processing frame count among the three CNN models used. On the other hand, the YOLOv4 model has a top processing speed of 25 frames per second and a 9 out of 18 detection rate on CPD samples. In contrast, the SSD Mobilenet v2 model has a top processing frame count of 20 frames per second, which is second to the YOLOv4 model in terms of processing speed. However, it has the lowest detection rate, with just 6 out of 18 CPD samples being detected.

Table I shows the detection results using Faster R-CNN Inception v2. The video resolution was 1080 x 720 at eight frames per second (FPS). Negative samples mean that during the detection process of the system, these are samples that are not infected with CPD. In contrast, those positive samples are samples that were infected with CPD. This particular table showed that from the 25 negative samples, it recognized 20 True Negative (TN) from the samples while only 5 were False Negative (FN). Furthermore, the Faster R-CNN framework recognized 13 True Positive (TP) and 2 False Positive (FP) samples from 20 samples infected with CPD.

Table II is the detection tests using Faster R-CNN Inception v2 960 x 640 at 11FPS. The 24 negative samples recognized 17 as TN and seven as FN. From the 16 positive samples, the result is 11 as TP and five as FP.

Table III shows the result of tests using YOLOv4 1080 x 720 @22FPS. From this table, 18 were detected as TN while nine were detected as FN from the 27 negative samples. For the positive samples, the system saw nine as a TP, while 4 were FP.

Table IV also presented the results using YOLOv4 960 x 640 @25FPS. The result is not much desirable since from the 26 negative samples. It only detects 14 as TN, which is close to detection of FN, which is 12. The same is through with its detection on positive samples. It detects 8 FP and 7 TP from 15 positive samples.

Table V presents the results using SSD Mobilenet v2 1080 x 720 @18FPS. From this table, the detection accuracy can be seen as low. It has 13 TN and 12 FN from 25 negative samples, while 9 FP and 6 TP from 25 positive samples.

Table VI shows the result using SSD Mobilenet v2 960 x 640 @20FPS. Its accuracy is less good than the previous methods. From the 25 negative samples it detects 11 TN and 14 FN. While from the 15 positive samples, it detects 11 FP and 4 TP.

Based on these results, if the image processing speed is considered, YOLOv4 is better than other models. Faster R-CNN inception v2 performs better if accuracy is a priority. Hence, these two models can be used depending on the purpose of the detection of the CPD. However, the most important factor to be considered must be its accuracy since CPD detection is the main objective of this study.

TABLE I. FASTER R-CNN INCEPTION v2 1080 X 720 @8FPS

	Positive	Negative
Negative	FN=5	TN=20
Positive	TP=13	FP=2

TABLE II. FASTER R-CNN INCEPTION v2 960 X 640 @11FPS

	Positive	Negative
Negative	FN=7	TN=17
Positive	TP=11	FP=5

TABLE III. YOLOV4 1080 X 720 @22FPS

	Positive	Negative
Negative	FN=9	TN=18
Positive	TP=9	FP=4

TABLE IV. YOLOV4 960 X 640 @25FPS

	Positive	Negative
Negative	FN=12	TN=14
Positive	TP=7	FP=8

TABLE V. SSD MOBILENET v2 1080 X 720 @18FPS

	Positive	Negative
Negative	FN=12	TN=13
Positive	TP=6	FP=9

TABLE VI. SSD MOBILENET v2 960 X 640 @20FPS

	Positive	Negative
Negative	FN=14	TN=11
Positive	TP=4	FP=11

## V. CONCLUSION

During the field testing/actual testing, the 5Ghz set-up was used because it is straightforward to deploy both its transmitter and receiver. The developed real-time CPD detection and quantification monitoring system was successfully integrated into the wireless video receiver and seamlessly visualized all the incoming data using the three different CNN models. If the consideration is the image processing speed, YOLOv4 is better compared to other models. Faster R-CNN inception v2 performs better if accuracy is a top requirement. Hence, these two models can be used depending on the purpose of the detection of the CPD. However, the most important factor to be considered must be its accuracy since CPD detection is the main objective of this study.

## VI. RECOMMENDATION

The test's discussed results using the three methods - SSD Mobilenet v2, Faster R-CNN Inception v2, and YOLOv4, shows that Faster R-CNN Inception v2 has the highest accuracy. However, the accuracy rate needs to be improved to achieve optimal accuracy. The suggestion is to increase the training datasets and modify the hyperparameters to achieve maximum accuracy.

#### ACKNOWLEDGMENT

We want to extend our sincerest gratitude to the Department of Science and Technology (DOST) for funding this project under the Science for Change Program, particularly under the CRADLE PROGRAM. The Philippine Council for Agriculture, Aquatic and Natural Resources Research and Development (PCAARRD) as the monitoring body, particularly the Agricultural Resources Management Research Division (ARMRD). To the Isabela State University, Quirino State University, EDCOR Cassava Development Cooperative, and the LGU San Guillermo, Isabela.

#### REFERENCES

- [1] P. K. Sethy, N. K. Barpanda, A. K. Rath, and S. K. Behera, "Image Processing Techniques for Diagnosing Rice Plant Disease: A Survey," *Procedia Comput. Sci.*, vol. 167, no. 2019, pp. 516–530, 2020, doi: 10.1016/j.procs.2020.03.308.
- [2] S. Jiddi, P. Robert, and E. Marchand, "Reflectance and Illumination Estimation for Realistic Augmentations of Real Scenes," *Adjunct. Proc. 2016 IEEE Int. Symp. Mix. Augment. Reality, ISMAR-Adjunct 2016*, pp. 244–249, 2017, doi: 10.1109/ISMAR-Adjunct.2016.0085.
- [3] A. Miyatra, D. Bosamiya, and N. Kamariya, "A Survey on Disease and Nutrient Deficiency Detection in Cotton Plant," *Int. J. Recent Innov. Trends Comput. Commun.*, vol. 1, no. 11, pp. 812–815, 2013.
- [4] W. Zhu, H. Chen, I. Ciechanowska, and D. Spaner, "Application of infrared thermal imaging for the rapid diagnosis of crop disease," *IFAC-PapersOnLine*, vol. 51, no. 17, pp. 424–430, 2018, doi: 10.1016/j.ifacol.2018.08.184.
- [5] S. Lange, T. Gabel, and M. Riedmiller, "Transfer Learning for Reinforcement Learning Domains: A Survey," *J. of Machine Learn. Res.*, vol. 10, no. 2009, pp. 1633–1685, 2009, doi: 10.1007/978-3-642-27645-3\_2.
- [6] S. Panigrahi, A. Nanda, and T. Swarnkar, "A Survey on Transfer Learning," *Smart Innov. Syst. Technol.*, vol. 194, pp. 781–789, 2021, doi: 10.1007/978-981-15-5971-6\_83.
- [7] D. Qiao and F. Zulkernine, "Vision-based Vehicle Detection and Distance Estimation," *2020 IEEE Symp. Ser. Comput. Intell. SSCI 2020*, no. December, pp. 2836–2842, 2020, doi: 10.1109/SSCI47803.2020.9308364.
- [8] Y. Lu, S. Yi, N. Zeng, Y. Liu, and Y. Zhang, "Identification of rice diseases using deep convolutional neural networks," *Neurocomputing*, vol. 267, pp. 378–384, 2017, doi: 10.1016/j.neucom.2017.06.023.
- [9] H. Gassoumi, N. R. Prasad, and J. J. Ellington, "Neural Network-Based Approach For Insect Classification In Cotton Ecosystems," *Int. Conf. Intell. Technol.*, no. January 2000, pp. 1–7, 1994.
- [10] J. Lyu et al., "Extracting the tailings ponds from high spatial resolution remote sensing images by integrating a deep learning-based model," *Remote Sens.*, vol. 13, no. 4, pp. 1–17, 2021, doi: 10.3390/rs13040743.
- [11] A. C. Taracatac and R. Q. Camungao, "Rice insect classification and quantification (RICQ) using portable neural network model with EFPV-image processing algorithm," *J. Adv. Res. Dyn. Control Syst.*, vol. 11, no. 11 Special Issue, pp. 979–985, 2019, doi: 10.5373/JARDCS/V11SP11/20193124.
- [12] D. Flóres, I. C. Haas, M. C. Canale, and I. P. Bedendo, "Molecular identification of a 16SrIII-B phytoplasma associated with cassava witches' broom disease," *Eur. J. Plant Pathol.*, vol. 137, no. 2, pp. 237–242, 2013, doi: 10.1007/s10658-013-0250-3.
- [13] K. D. Kra, Y. M. N. Toualy, A. C. Kouamé, H. A. Diallo, and Y. A. Rosete, "First report of a phytoplasma affecting cassava orchards in Cote d'Ivoire," *New Dis. Reports*, vol. 35, no. 1, pp. 21–21, 2017, doi: 10.5197/j.2044-0588.2017.035.021.
- [14] Z. Zou, Z. Shi, Y. Guo, and J. Ye, "Object Detection in 20 Years: A Survey," *arXiv Prepr. arXiv1905.05055*, pp. 1–39, 2019, [Online]. Available: <http://arxiv.org/abs/1905.05055>.
- [15] M. Xu, "Robust object detection with real-time fusion of multiview foreground silhouettes," *Opt. Eng.*, vol. 51, no. 4, p. 047202, 2012, doi: 10.1117/1.oe.51.4.047202.
- [16] C. C. Aggarwal, "Transfer Learning," *Data Classif. Algorithms Appl.*, pp. 657–665, 2014, doi: 10.1201/b17320.
- [17] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, 2017, doi: 10.1109/TPAMI.2016.2577031.
- [18] D. Alamsyah and M. Fachrurrozi, "Faster R-CNN with inception v2 for fingertip detection in homogenous background image," *J. Phys. Conf. Ser.*, vol. 1196, no. 1, 2019, doi: 10.1088/1742-6596/1196/1/012017.
- [19] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," 2020, [Online]. Available: <http://arxiv.org/abs/2004.10934>.
- [20] J. Yu and W. Zhang, "Face Mask Wearing Detection Algorithm Based on Improved YOLO-v4," *Sensors*, vol. 21, no. 3263, pp. 1–21, 2021.
- [21] R. Deepa, E. Tamilselvan, E. S. Abrar, and S. Sampath, "Comparison of Yolo, SSD, Faster RCNN for Real Time Tennis Ball Tracking for Action Decision Networks," *Proc. 2019 Int. Conf. Adv. Comput. Commun. Eng. ICACCE 2019*, pp. 2019–2022, 2019, doi: 10.1109/ICACCE46606.2019.9079965.
- [22] I. T. Plata, A. C. Taracatac, and E. B. Panganiban, "Development and testing of embedded system for smart detection and recognition of witches' broom disease on cassava plants using enhanced viola-jones and template matching algorithm," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 8, no. 5, pp. 2613–2621, 2019, doi: 10.30534/ijatcse/2019/113852019.
- [23] I. T. Plata, E. B. Panganiban, B. B. Bartolome, F. E. R. Labuanan, and A. C. Taracatac, "A concept of cassava phytoplasma disease monitoring and mapping system using GIS and SMS technology," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 8, no. 6, pp. 3357–3361, 2019, doi: 10.30534/ijatcse/2019/108862019.
- [24] M. G. Selvaraj et al., "Machine learning for high-throughput field phenotyping and image processing provides insight into the association of above and below-ground traits in cassava (*Manihot esculenta* Crantz)," *Plant Methods*, vol. 16, no. 1, pp. 1–26, 2020, doi: 10.1186/s13007-020-00625-1.