Design and Implementation of Cost-Effective IT-based Phenotyping Tools for Research on Eggplant Breeding Techniques

Val Randolf M. Madrid[1], Aaron V. Magnaye[1], Antonette A. Porca[1], Lourdes D. Taylo[2], Desiree M. Hautea[2]
1 Institute of Computer Science, University of the Philippines Los Banos (UPLB)
2 Institute of Plant Breeding, UPLB

フィリピンはナスの生産量で世界10位。農家にとっては安定した収入源だが、病気や害虫に強い野菜をつくるための品種改良には手間と時間がかかる。そこでITを活用して品種改良の手間とコストを減らすことを試みた。

Abstract

The Philippines ranks 10th in terms of eggplant production globally. In the second quarter of 2020, Philippine eggplant production was 104,440 metric tons. Compared to other vegetables, eggplant can be planted year-round. With proper cultural practices, 1 hectare of land can yield 18,000 kilograms giving the farmer about PHP 90,000 (~USD 1,850) net income per hectare. However, eggplant production is not without its challenges. Aside from natural disasters like typhoons, eggplant production is greatly hampered by pest infestation by the leafhopper (*Amrasca biguttula*) and the eggplant fruit and shoot borer (*Leucinodes orbonalis* Guenee), which result in an average of 50%–70% yield loss. Host plant resistance is one of the insect pest management strategies that are sustainable, environment-friendly, and compatible with other control methods. However, it is tedious and time-consuming as researchers would have to conduct phenotyping on many varieties of eggplants. These activities include trichome density counting and observing the feeding preference of pests. This paper presents the development of cost-effective IT-based applications, such as automated trichome density classification and feeding preference of pest larvae through video analysis to speed up the phenotyping process of finding host plant resistance in eggplant varieties.

Keywords eggplant phenotyping, computer vision, pattern recognition, machine learning

Introduction

Eggplant (*Solanum melongena*) ranks third among the vegetables and rootcrops produced in the Philippines, thus, playing a big role in the country's economy. According to the 2018 survey conducted by the Crops Production Survey (CrPS), covering all regions except the National Capital Region, eggplant production increased to 103,000 metric tons compared to the previous year at 99,490 metric tons, showing a growth of 3.6% [1].

In the second quarter of 2020, the Philippines ranked 10th in eggplant production globally, at 104,440 metric tons. Compared to other vegetables, eggplant can be planted year-round, and with proper cultural practices, it can potentially yield 18,000 kg/hectare giving the eggplant farmer about PHP 90,000 (~USD 1,850) net income per hectare [2].

However, eggplant production in the Philippines is not without its challenges. Aside from natural disasters such as typhoons, eggplant production is greatly affected by the two most damaging pests: the leafhopper (*Amrasca biguttula*) and the eggplant fruit and shoot borer (EFSB) (*Leucinodes orbonalis* Guenee). Infestations of these two pests have resulted in an average of 50%–70% yield loss. Aside from yield loss, farmers often spend a big chunk of their money on pesticides that they would spray 60–80 times over a 4-month period to get rid of EFSB, according to local studies [3].

The Bt Eggplant Project, which ended in 2014, was designed to utilize agrobiotechnology approaches, such as genetic modification, in making eggplant resistant to EFSB larvae. The eggplant was modified using a gene from a common soil bacterium (*Bacillus thuringiensis*), which would prove fatal to EFSB larvae but harmless to humans. It was successful during field trials [3] but encountered opposition in the approval of Bt eggplant's use in the Philippines because of the lobbies against genetically modified crops. The ruling was later reversed [4].

Eggplant researchers at the Institute of Plant Breeding in UP Los Banos started to look for a different approach to enhancing eggplant's host resistance based on the lessons learned from the Bt eggplant. Under the leadership of Dr. Desiree Hautea, a multidisciplinary project was approved by the Department of Science and Technology's Philippine Council for Agriculture, Aquatic and Natural Resources Research and Development (DOST-PCAARRD). The project aims to develop new eggplant varieties with new plant defense genes that have multiple insect resistance using innovative technologies that do not require the introduction of genes from other organisms. One of these innovative technologies is to utilize computer vision, machine learning, and other information technology-based tools in assisting researchers in identifying host plant resistance of eggplants based on certain physical characteristics and then identifying the genes responsible for those characteristics.

Cost-Effective IT-Based Phenotyping Tools

Host plant resistance is one of the insect pest management strategies that are sustainable, environment-friendly, and compatible with other control methods. However, it is tedious and time-consuming as researchers would have to conduct phenotyping of varieties of eggplants, such as trichome density counting and observing the feeding preference of the pests.

Trichomes are hair-like epidermal structures that can be found on the aerial parts of the plant. Trichomes can vary in size and have different characteristics that are species-specific. Therefore, it is used as a diagnostic characteristic to identify plant species. The high density of trichomes makes the plant resistant to pests, such as leafhoppers [5].

One research showed that the feeding activity and the number of herbivores present on a leaf is inversely proportional to the trichome density of the leaf [6].

To quantify trichome density, researchers would take a leaf sample from an eggplant variety and place it under a microscope, where they would count the number of trichomes in the image to measure its density.

Fig. 1 shows a sample trichome image taken from a camera mounted on a stereo zoom trinocular microscope.

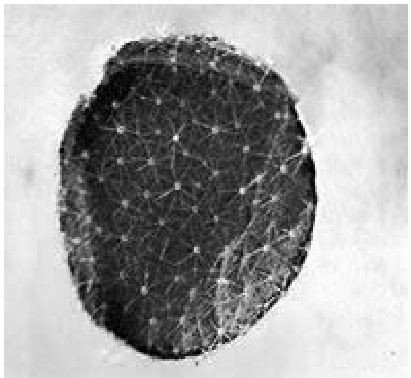

Fig. 1 A sample trichome image taken from a camera mounted on a microscope

After counting the trichomes in the image, researchers would then repeat the process for all the eggplant samples for every eggplant variety they would study, which would entail a lot of work and would be prone to human error in the long run.

Each leaf sample is then categorized based on the number of trichomes counted. There are three categories: few (less than or equal to 49), intermediate (between 50 to 99), and many (greater than or equal to 100) [7]. There is a need to assist eggplant researchers by developing IT-based tools to speed up the counting and analysis of trichome densities in different eggplant varieties.

The feeding preference of pests is already known as a good indicator of host plant resistance. Using video data, insects are analyzed based on their feeding preferences given certain varieties of plants. It has been done on aphids [8] and thrips [9], both destructive pests.

But to do this, access to custom computer hardware and software is needed, such as the ones provided by Noldus [10]. Fig. 2 shows Noldus EntoVision's hardware and software components. These products, however, are not available in the local market. The cost of purchasing such custom technologies can go up to USD 70,000 (~PHP 3.5 million), including shipping and personnel training.

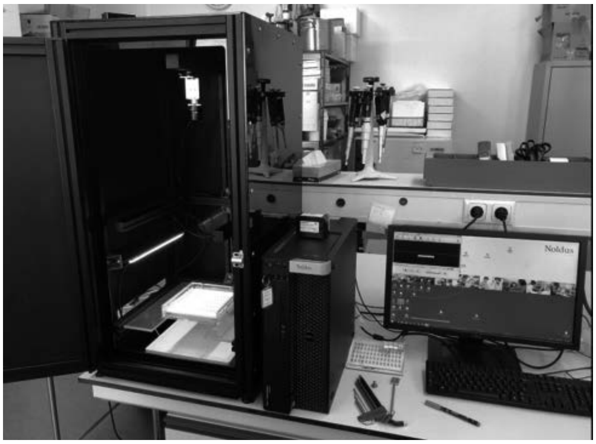

Fig. 2 Noldus EntoVision lab setup

Thus, there is a lot of room for research when it comes to developing low-cost technologies yet still capable of generating high throughput eggplant phenotype data that will be used for genotype analysis.

Objectives

This paper presents the development of cost-effective IT-based phenotyping tools that can allow an eggplant researcher to:

1. Analyze trichome images and automatically provide the trichome densities of different eggplant varieties.
2. Set up an EFSB feeding preference experiment using off-the-shelf commodity equipment and gather feeding preferences data of eggplant varieties via larvae trajectory analysis.

Materials and Methods

Based on the objectives, two types of IT-based phenotyping tools were created. The first one is a computer program that will be able to batch process a set of eggplant trichome images and classify them according to their trichome density categories. The computer program was designed to be modular and uses open-source programming libraries for the image analysis and classification components. The use of image processing and analysis on agriculture-based images is already well established.

Image segmentation techniques such as image feature points and corner detection were used to automatically count rice and wheat grains using an Android mobile device [11]. Another study also used image segmentation techniques to count the leaves of medicinal plants even from uneven and complex backgrounds [12].

Trichome Counting

The trichome counting problem can be reduced to an image segmentation problem where the components in the image you want to segment are the trichomes only, allowing automated counting. Based on the quality and image capture approach provided for this study, the image segmentation approach followed the following process:

1. Image enhancement — The images used in this study employed image sharpening to enhance the edge pixels of the trichomes. Contrast limited adaptive histogram equalization (CLAHE) further enhanced the lighter trichome pixels against the darker green background. CLAHE is already established as a good algorithm to enhance contrast and remove noise pixels in the image [13]. Fig. 3 shows the original image and the sharpened CLAHE image. After this operation, the trichomes would look like circular disks with faded hair-like projections. When counting manually, the researcher looks at the central circular part of a trichome.

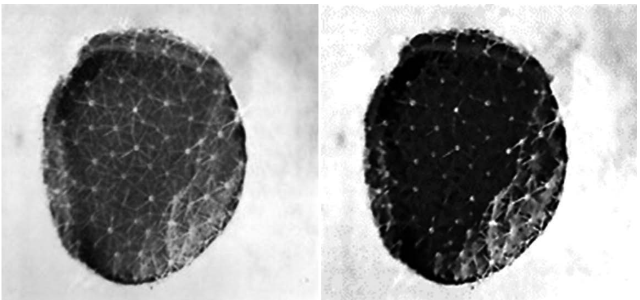

Fig. 3 Original trichome image (left) and the image after CLAHE was done (right)

2. Image morphology operations — After image enhancement, morphological transformation, specifically, opening, was performed. By using an elliptical kernel, the centers were able to maintain their circular shape, whereas the branches lost their linear structure and merged with their surroundings. This also resulted in the image being composed of circular segments. Thresholding was performed to determine the potential central discs representing trichomes among the circular segments. Fig. 4 shows the remaining circles that passed the threshold.

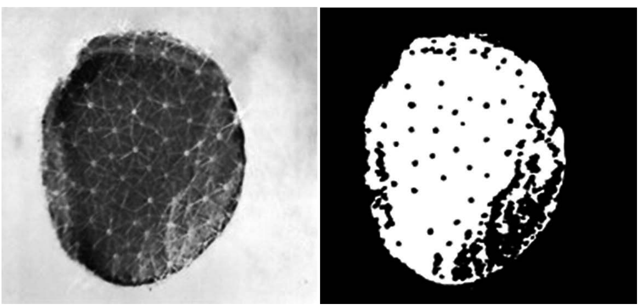

Fig. 4 Original trichome image (left) and image after morphology operations and thresholding were done (right)

3. Trichome segmentation — Using the circle Hough transform algorithm to detect shapes, the central discs were detected and highlighted in the original image. Fig. 5 shows the result of trichome detection based on the thresholded morphological image.

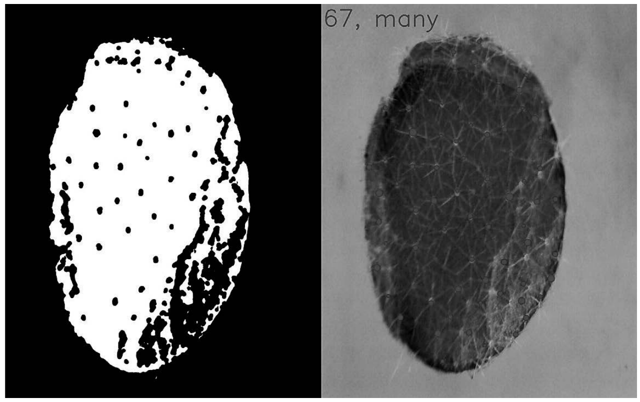

Fig. 5 Trichome images with the detected trichomes labeled

Tracking EFSB Feeding Preference

The second IT-based phenotyping tool for tracking EFSB feeding preference employs a computer program and an experimental setup that uses an ordinary USB-based web camera, a cardboard box covered in white paper, and a laptop. Fig. 6a shows the design of the experimental setup.

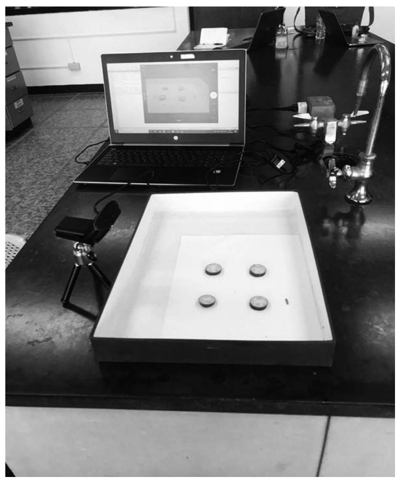

Fig. 6a EFSB feeding preference experiment setup using the USB-based webcam

Alternatively, you can also utilize your mobile device in place of a USB-based camera connected to your laptop. This can be easily set up using a free Android application called the IP Webcam [14], which turns your Android phone into a network camera, as shown in Fig. 6b.

The setup is simple; the camera is positioned above the box containing eggplant fruit slices and EFSB larvae at a top view angle while connected to the laptop. The camera sends video frames to the laptop while the computer program analyzes the video frames and tracks the movement of the larvae, and detects which eggplant variety it prefers over a specified length of time, usually from 30 minutes to up to 2 hours. The computer program also saves the data generated in a simple table format that shows the larvae's position at different time intervals. This would allow eggplant researchers to identify the eggplant varieties that the EFSB larvae would prefer at a given time frame. This approach would allow the researchers to save time as they wouldn't have to manually observe the movements of the EFSB larvae all the time. This is very similar to the studies done using the more expensive automated setups like the Noldus EntoVision [15].

Fig. 6b EFSB feeding preference experiment setup using the camera of the mobile device

Fig. 7 shows the sample video frame, which contains the experimental setup containing eggplant slices from different varieties and EFSB larvae roaming around.

Fig. 7 Sample video frame showing slices of different eggplant varieties and EFSB larvae

Because of the nature of video and its near real-time operation, the main computer vision algorithm for tracking the position of the EFSB larvae is the YOLOv4 framework. The original YOLO framework uses a convolutional neural network (CNN) to detect objects near real-time with good performance in object detection accuracy and fast video frame rate [16]. YOLOv4 is the fourth iteration of the original YOLO framework and has faster video frame processing speed and higher object detection accuracy [17].

To use the YOLOv4 framework as the EFSB larvae tracker, sample videos were collected, and the EFSB larvae in each video frame were labeled. Then an artificial intelligence (AI) model using the YOLOv4 framework was trained from the labeled video frames containing EFSB larvae. Fig. 8 shows the initial training results of the YOLOv4 framework AI; EFSB larvae were automatically detected and marked using colored bounding boxes.

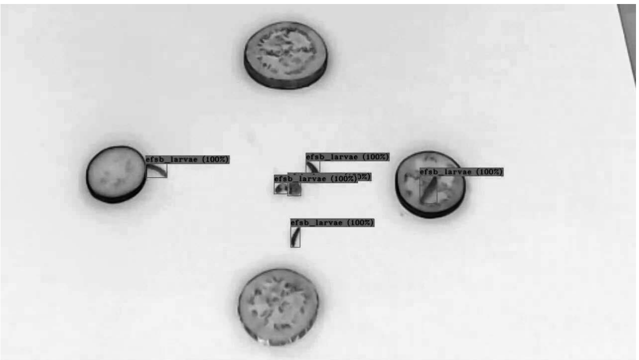

Fig. 8 EFSB larvae tracking using YOLOv4 initial results

Once the AI was done training, the YOLOv4 AI model would then be loaded into the computer program so that it will be able to detect and track EFSB larvae as well as the location of the eggplant slices.

Software Specifications

The two IT-based phenotyping tools developed for this study used readily available computer hardware and peripherals and open-source software libraries.

Table 1 shows the hardware specifications used to create the IT-based phenotyping tools for the automated trichome counting program and the EFSB feeding preference setup.

Table 1. Computer hardware specifications used in the study

Hardware	Specification
Laptop	Core i5 8 GB RAM
Desktop PC for AI Training	Core i7 16GB RAM with NVIDIA GTX 1070
Mobile phone or webcam	720p, 23 fps

Table 2 shows the software specifications used to create the IT-based phenotyping tools for the automated trichome counting program and the EFSB feeding preference setup.

Table 2. Software libraries specifications used in the study

Library	Specification
Computer Vision	OpenCV 4.5.1
Programming Environment	Python 3.8 with Tkinter for GUI
AI Framework	YOLOv4 Framework

Results and Discussions

Automated Trichome Counting

The trichome image dataset provided was pre-classified into its categories: few, intermediate, and many. There were a total of 62 trichome images: 5 images in the 'few' category, 43 in the 'intermediate' category, and 14 in the 'many' category. The limited data available for each category was because of the availability of the eggplant varieties that were planted when the image data were gathered. The computer program for trichome counting was created using the Python programming language and the Tkinter graphical user interface (GUI). The user would only have to load the folder containing the trichome images, and the program would automatically count the number of trichomes in every image and classify each image in its respective categories. Fig. 9a shows the screenshot of the program when loading a folder, and Fig. 9b shows the result of one of the images processed.

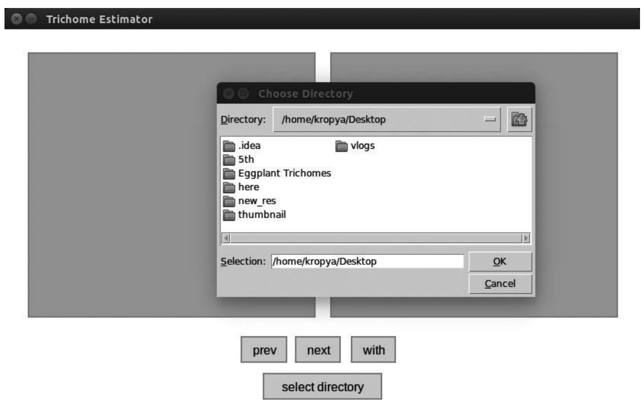

Fig. 9a Program screenshot: loading a folder containing trichome images

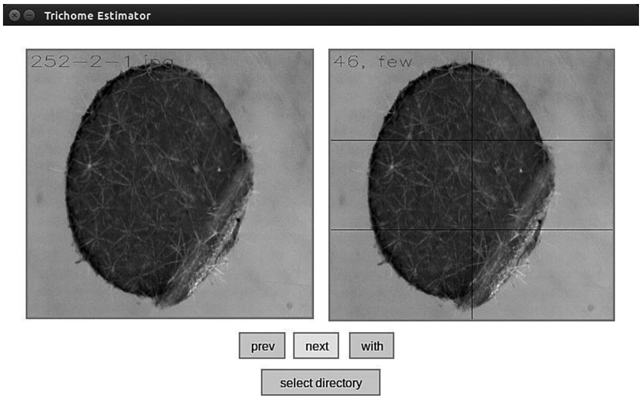

Fig. 9b Program screenshot: viewing one of the processed images and its respective category

The program's overall performance was assessed by comparing the computer program's assessment of each trichome image's category and the assessment by an expert. Table 3 shows the result of the computer vs. the expert's assessment.

Table 3. Trichome category classification accuracy after expert validation

Category	Accuracy (%)
Few	100 %
Intermediate	65.1 %
Many	35.7 %

Based on the results, two factors played a big role in the performance of the program: image quality and image scale. The existing image capture hardware only had a 2-megapixel resolution, and the scale at which the images were taken made some of the trichomes less visible than others. The presence of trichome clusters also made accurate counting difficult, leading to false positives and false negatives. Some trichomes were detected where there shouldn't be one, and trichomes were not detected even if there should be one. Analyzing trichome images with few trichomes was simpler as the trichomes were often separate and distinct. Fig. 10 shows the image samples with the difficult cases present. It seems like pure image processing techniques alone proved insufficient for the task, but the computer program that was developed was designed to be modular, so it was just a simple task to experiment with other techniques such as incorporating a trained CNN AI model to handle the difficult cases. Modifying the computer program to use the trained AI yielded the following results, as shown in Fig. 11. There is a vast improvement in the accuracy of detecting trichomes and classifying them in the right category. It was still able to detect "few" trichomes at 100% accuracy, but it improved detection accuracy for "intermediate" and "many" at 79% and 82%, respectively. Fig. 12 shows how the improved approach even captures the difficult cases.

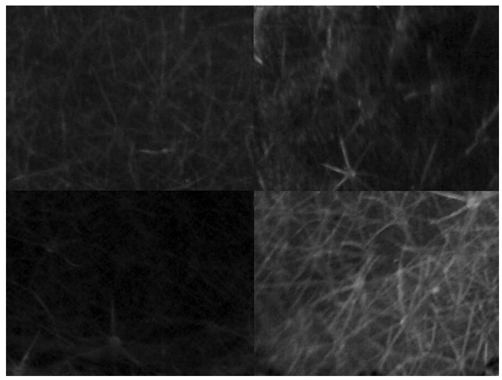

Fig. 10 Examples of difficult images that gave poor results

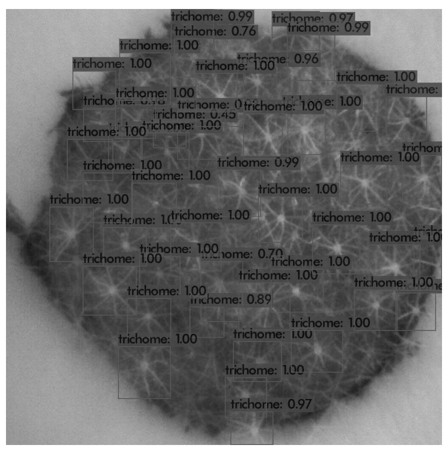

Fig. 11 Using the CNN AI model to detect trichomes

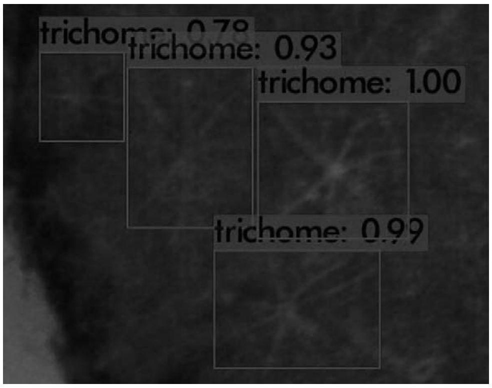

Fig. 12 Detecting trichomes in difficult images using improved approaches

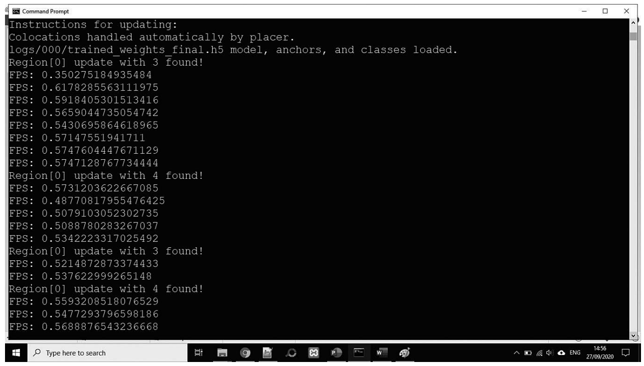

Fig. 13a Screenshot of the EFSB Feeding Preference computer program running

Tracking EFSB Feeding Preference

For the EFSB feeding preference tool, the computer program was created using Python and incorporated a trained YOLOv4 model to track EFSB larvae trajectories and then save the report in a text file. The program successfully generated feeding preference data that will be used for later analyses. Training the YOLOv4 AI model used only one sample video with 1,000 frames. The frames were annotated with the locations of the EFSB larvae and eggplant slices marked. Once the training was done, the AI model was plugged into the computer program for the EFSB feeding preference test to begin.

Fig. 13a shows the screenshot of the EFSB feeding preference program running. The decision to make it a command-line program was to ensure that performance and accuracy in the processing of video frames do not slow down the program since video frame processing takes up a lot of computing resources and memory. Fig. 13b shows the results of the processed video frame.

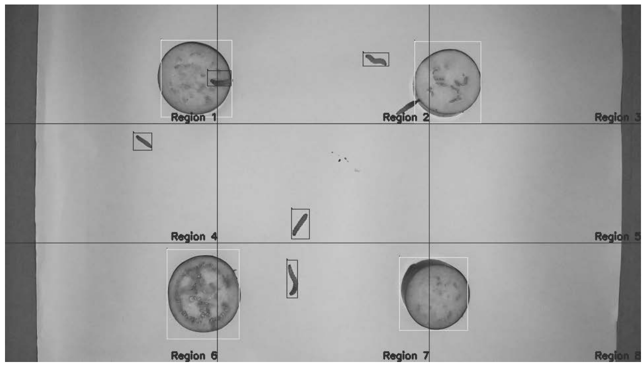

Fig. 13b Processed video frame showing the locations of EFSB larvae and the eggplant slices

The video frame was divided into nine equal regions to add more information regarding larvae behavior, such as to determine whether it is important to know whether there is a reason why certain larvae linger in a particular region between two eggplant slices of certain varieties. As the video frames are processed, the locations of larvae and the region they are in are recorded into a text file. Fig. 14a shows a portion of the text file generated which shows that at the 24-second mark, larvae were seen lurking in regions 4 and 7 in the video frame.

An image containing the trajectories of the EFSB larvae can also be generated once the experiment is over, as seen in Fig. 14b.

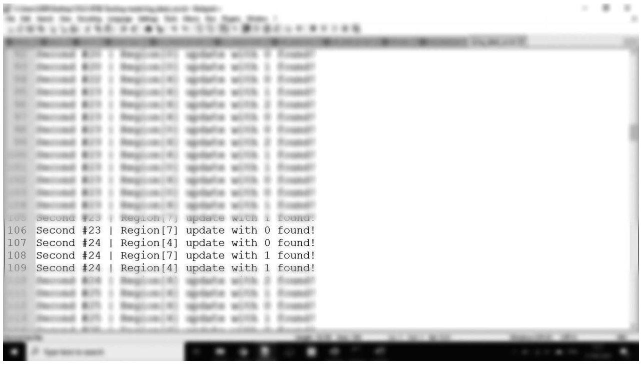

Fig. 14a Text file data containing info on detected EFSB larvae lurking in regions 4 and 7 at the 24-second mark

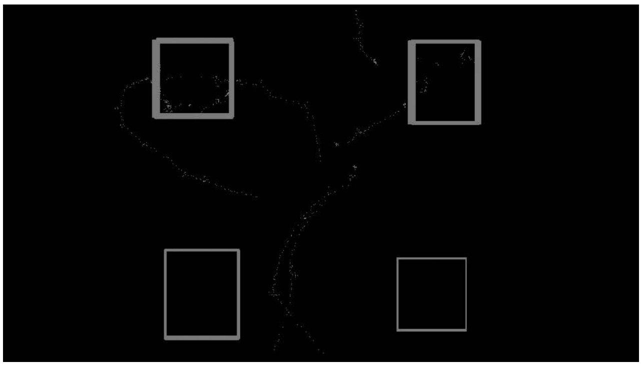

Fig. 14b Trajectories of the EFSB larvae (green dots) relative to the location of eggplant varieties

The EFSB feeding preferences program successfully generated feeding preference data given a time frame.

Most of the cost of developing the two IT-based phenotyping tools centered around the computer hardware and peripherals, which amounted to about PHP 250,000 (~ USD 4,900), a cost that may be more affordable to smaller labs and research projects.

Conclusion

There is an urgent need to develop tools that will assist eggplant researchers in doing high throughput phenotyping to determine host plant resistance. There are existing tools already that allow them to do this, but the availability and cost might be a bit steep for smaller research labs.

This paper is a result of designing and pilot testing cost-effective IT-based phenotyping tools that will be able to generate similar data. Although it was able to perform as intended, it still leaves much to be desired.

There is a need to update the image capture setup for the trichome counting program to produce better results. Based on lessons learned from these studies, there have already been preliminary results in the improved trichome images and the initial AI model that analyzes them. Fig. 15 shows the screenshot of the updated computer program using improved images. The images have a higher image resolution at 5 megapixels and a more close-up shot to address the issues found in this study.

For the EFSB feeding preference program, the challenge is to streamline and integrate the data generated into different data mining techniques and make the computer program easier to use. Preliminary designs have already been started to make the system into a full-stack web application to implement data mining techniques since the generated data will be stored in a database while making it easier to use via a GUI in the web browser. Fig. 16 shows the proposed wireframes planned to achieve this in future improvements.

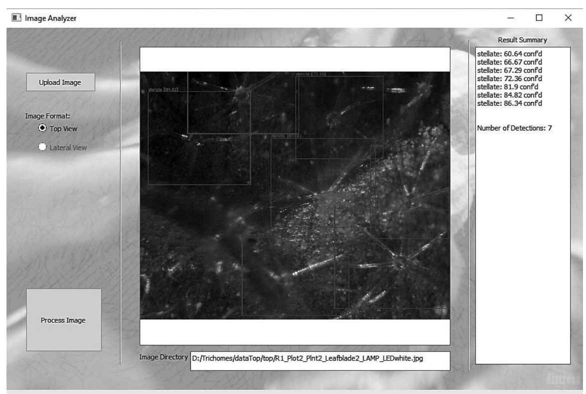

Fig. 15 Updated trichome counting program using the new and improved images, detected trichomes are marked purple

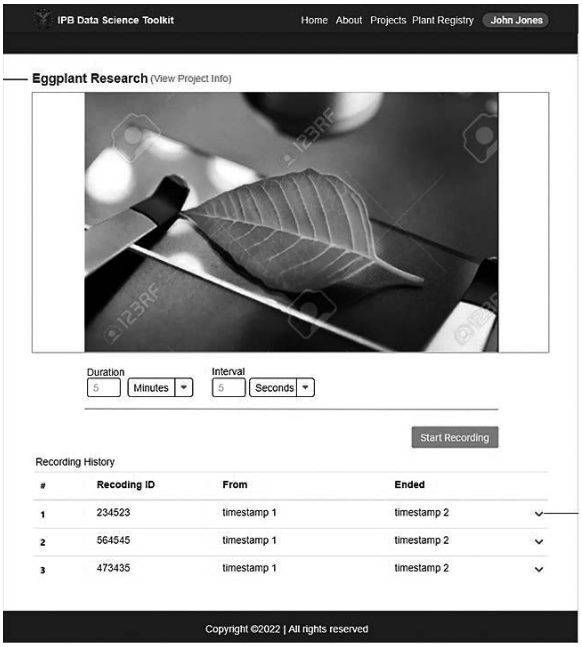

Fig. 16 Proposed web application for the EFSB feeding preference experiment

References

1. Major Vegetables and Rootcrop Quarterly Bulletin. 2018. Vol. 12 No.2Licuanan W.Y., and E.D. Gomez. 2000. *Philippine Coral Reefs, Reef Fishes, and Associated Fisheries: Status and Recommendations to Improve Their Management*. Manila: Global Coral Reef Monitoring Network (GCRMN).

2. Roque, Anselmo. (2005, April 18). 'How to profit from eggplant production.' Philippine Daily Inquirer. https://www.seedquest.com/ News/releases/2005/april/12104.htm

3. Hautea, D.M., Taylo, L.D., Masanga, A.P.L., Sison, M.L.J., Narciso, J.O., Quilloy, R.B., Hautea, R.A., Shotkoski, F.A. and Shelton, A.M. (2016). Field performance of Bt eggplants (Solanum melongena L.) in the Philippines: Cry1Ac expression and control of the eggplant fruit and shoot borer (Leucinodes orbonalis Guenée). *PLoS One*, *11*(6), p.e0157498.

4. Palicte, Che. (2021, July 24). 'PH approves Bt eggplant's use for food, feeds.' Philippine News Agency. https://www.pna.gov.ph/articles/1148086

5. Glas, J.J., Schimmel, B.C., Alba, J.M., Escobar-Bravo, R., Schuurink, R.C., & Kant, M.R. (2012). Plant glandular trichomes as targets for breeding or engineering of resistance to herbivores. *International Journal of Molecular Sciences*, 13(12), 17077–17103.

6. Marcela Carvalho Andrade, Alex Antônio Da Silva, Irã Pinheiro Neiva, Izabela Regina C Oliveira, Evaristo Mauro De Castro, David Merrill Francis, and Wilson Roberto Maluf. (2017). Inheritance of type iv glandular trichome density and its association with whitefly resistance from solanum galapagense accession la1401. *Euphytica*, 213(2): 52

7. Taylo, L. D., & Bernardo, E. N. (1995). Morphological and biochemical bases of resistance of eggplant (Solanum melongena Linn.) and okra (Abelmoschus esculentus (L.) Moench.) to the cotton leafhopper (Amrasca biguttula (Ishida)). In *26. PMCP Anniversary and Annual Scientific Meeting, La Trinidad, Benguet (Philippines), 2-5 May 1995*.

8. Kloth, K. J., Ten Broeke, C. J., Thoen, M. P., Hanhart-van den Brink, M., Wiegers, G. L., Krips, O. E., ... & Jongsma, M. A. (2015). High-throughput phenotyping of plant resistance to aphids by automated video tracking. *Plant Methods*, *11*(1), 1–14.

9. Thoen, M.P., Kloth, K.J., Wiegers, G.L., Krips, O.E., Noldus, L.P., Dicke, M. and Jongsma, M.A., 2016. Automated video tracking of thrips behavior to assess host-plant resistance in multiple parallel two-choice setups. *Plant Methods*, *12*(1), pp.1–12.

10. Noldus Information Technology. (2022). Noldus — Our products. https://www.noldus.com/products

11. Liu, T., Chen, W., Wang, Y., Wu, W., Sun, C., Ding, J., & Guo, W. (2017). Rice and wheat grain counting method and software development based on Android system. *Computers and Electronics in Agriculture*, *141*, 302–309.

12. Gao, L., & Lin, X. (2018). A method for accurately segmenting images of medicinal plant leaves with complex backgrounds. *Computers and Electronics in Agriculture*, *155*, 426–445.

13. Sahu, S., Singh, A. K., Ghrera, S. P., & Elhoseny, M. (2019). An approach for de-noising and contrast enhancement of retinal fundus image using CLAHE. *Optics & Laser Technology*, *110*, 87–98.

14. IP Webcam Android Application https://play.google.com/store/apps/details?id=com.pas.webcam&hl=en&gl=US

15. Goggin, F. L., Lorence, A., & Topp, C. N. (2015). Applying high-throughput phenotyping to plant–insect interactions: picturing more resistant crops. *Current Opinion in Insect Science*, *9*, 69–76.

16. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779–788).

17. Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). Yolov4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*.