Plant Disease Detection: Electronic System Design Empowered with Artificial Intelligence

Jiayi Wu
Northwest University &
University of Essex
Xi'an, China, Colchester, UK
jw19483@essex.ac.uk

Usman Dar
Wilkin & Sons
Tiptree, UK
usman@tiptree.com

Mohammad Hossein Anisi
School of Computer Science &
Electronic Engineering
University of Essex
Colchester, UK
m.anisi@essex.ac.uk

Vahid Abolghasemi
School of Computer Science &
Electronic Engineering
University of Essex
Colchester, UK
v.abolghasemi@essex.ac.uk

Chris Newenham
Wilkin & Sons
Tiptree, UK
cwn@tiptree.com

Andrey Ivanov
Wilkin & Sons
Tiptree, UK
ai@tiptree.com

***Abstract*— Today, plant diseases have become a major threat to the development of agriculture and forestry, not only affecting the normal growth of plants but also causing food safety problems. Hence, it is necessary to identify and detect disease regions and types of plants as quickly as possible. We have developed a plant monitoring system consisting of sensors and cameras for early detection of plant diseases. First, we create a dataset based on the data collected from the strawberry plants and then use our dataset as well as some well-established public datasets to evaluate and compare the recent deep learning-based plant disease detection studies. Finally, we propose a solution to identify plant diseases using a ResNet model with a novel variable learning rate which changes during the testing phase. We have explored different learning rates and found out that the highest accuracy for classification of healthy and unhealthy strawberry plants is obtained with the learning rate of 0.01 at 99.77%. Experimental results confirm the effectiveness of the proposed system in achieving high disease detection accuracy.**

Keywords—Plant disease detection, Deep learning, Crop monitoring, Computer vision

I. Introduction

In recent years, plant diseases have become a major challenge to today's forestry development. Plant diseases result in damage to a part of the tissue or organ of the plant, until it is destroyed, killed, or aesthetically ruined. Plant diseases affect plants from the outside to the inside, from top to bottom, from flowers and fruits to the root system in a comprehensive, all-round manner. This not only affect the normal growth of plants but can also cause a reduction in the yield and quality of agricultural products and, in serious cases, food safety problems [1]. Therefore, rapid identification and diagnosis of plant diseases can reduce the economic losses caused by plant diseases to the agricultural industry in the shortest possible time. Plant disease identification is a technique for processing, analysing, and understanding plant image datasets to identify potential kinds of disease objects. It is a key process for the timely and effective control of plant diseases [2]. Today, there are many different types of plant diseases in different stages of growth and in several growing areas, which makes it difficult for laymen to accurately identify the types of disease in a short period of time and in a large scale. Besides, manual identification has the disadvantages of slow identification speed and low accuracy which poses a major challenge in containing the outbreak of diseases in agriculture.

With the continuous development of deep learning from one hand, and increasing the computation power from another hand, many researchers have started to study plant disease identification based on deep learning with either sensors or image data (or both). Using computer vision technology to identify plant disease areas and species can effectively reduce time costs and improve the efficiency of agricultural production [2]. Furthermore, with the advances in internet of things (IoT) technology, effective and continuous monitoring of various systems has become easier and more accessible. This has led to greater autonomy of systems in several industries as well as agriculture. The solutions which IoT offer are complemented by machine learning and computer vision-based techniques to improve the classification and detection performance.

This paper provides a comprehensive overview of the state-of-the-art in plant disease detection with focus on recent developments in hardware design and deep learning approaches. Furthermore, it presents the current frequently used plant disease datasets, and summarises the current cutting-edge development techniques in the field of plant disease identification. It also showcases the details of an electronic IoT-based device for acquisition of strawberry plant data as well as the results of two deep learning models on the captured plant images to identify the disease.

II. State-Of-The-Art

2.1 Plant Diseases Datasets

According to our investigation there exist around 14 available image datasets for plant disease diagnosis. The size and quality of the dataset will affect the accuracy of the deep learning model. Mostly, a large and high-quality dataset will improve the quality of the training process and the accuracy of monitoring of the deep learning model, allowing for more accurate identification of different types of plant diseases [3]. Due to space restriction in the paper, we have provided a summary of these datasets with the key details in Table 1.

2.2 Existing plant disease detection systems

There exist limited works reported on developments of plant disease diagnosis with on-board hardware. Pankaj et al. proposed an IoT hardware sensor-based Cotton Disease prediction using convolutional neural network (CNN). Their

IoT gadget has different sensors such as temperature, humidity, and PH to collect the data to be used for classification [4]. In another work by Mora et al., a plant disease detection using the Raspberry Pi 4 was implemented. Not many results are reported in this work, however, accuracy around 90% was obtained for prediction of plant disease on a private dataset [5]. A diagnostic system implemented on Raspberry Pi was proposed for scab and leaf disease detection. The authors used a CNN model and four classes of Scab, Black Rot, Cedar Rust, and Healthy were detected [6]. A smart crop growth monitoring using edge artificial intelligence (AI) was developed in [7] where a complex system was implemented to monitor health plants and classify the pest and disease severity. They used cryptographic hardware modules, including RTEA32, RTEA64, XTEA32 and XTEA64, and used the binarized neural network and achieved 76.57% accuracy for disease detection on dragon fruits.

We can categorise the type of plan disease detection into three key types: direct object detection, multiclass classification, and binary segmentation. Direct objection detection is typically disease identification on a single plant type. A self-constructed SPIKE dataset from images of relevant complex wheat fields was used in an object detection method based on identifying diseased plants (or parts affected by diseases or pathogens) proposed by Hasan, M.M. et al. [8] The model used was an R-CNN architecture that generated four different models, four different datasets of training and test images based on four different datasets to capture plant diseases at different growth stages with an accuracy of 93.4%.

Toda Y et al. used the YOLOV3 - DenseNet algorithm for direct object detection, focusing on disease object detection concerning growing apple leaves, with an accuracy of 95.75%. And using human intervention to validate the authenticity of the model and the training dataset, a CNN trained using publicly available plant disease image datasets, various neuron and layer visualisation methods were applied [9]. Zhang, S. et al. used the GPDCNN algorithm for multiclass classification detection of cucumber images, i.e., using different stages of the plant for possible disease detection. An accuracy of 94.65% was achieved [10]. Hari et al. in 2019 used the PDDNN algorithm for the detection of various plant disease images, using TensorFlow as the framework, with an accuracy of 86% [11]. As a comparison, Picon et al. also published a paper in 2019 using the RESNET-MC1 algorithm for the detection of various plant disease images using TensorFlow and Keras as the framework of choice, with an accuracy of 98% [12]. Howlader et al. use the AlexNet algorithm to detect plant diseases on guava leaves with an accuracy of 98.74% [13]. Nagasubramanian et al. used the 3D-CNN algorithm to detect plant diseases in soybean using a binary classification method, i.e., only diseased or healthy, without distinguishing between specific growth regions and growth stages, with an accuracy of 95.73% [14]. Arunangshu Pal proposed an Agricultural Inspection (AgriDet) framework, The framework combines the traditional Inception-Visual Geometry Group network (INC-VGGN) and the Kohonen-based deep learning network to detect plant diseases and classify the severity of diseased plants where the performance of the statistical analysis is validated to demonstrate the effectiveness of the results in terms of accuracy, specificity, and sensitivity [15].

In the article by Amal Mathew et al., the support vector machine (SVM) classifier was replaced with a voting classifier to classify the data into multiple classes. The accuracy of voting and SVM classifiers are compared. The results show that the accuracy of the proposed method is improved by 10% [16]. Punam Bedi et al. proposed a hybrid system based on convolutional auto-encoder (CAE) and CNN that can achieve automatic detection of plant diseases. In the experiment, CAE is used to compress the parameters required for training, and the parameters required for the hybrid model are reduced. The proposed hybrid model used only 9914 training parameters. The experiment uses a public dataset called PlantVillage to obtain leaf images of peach plants with the training and testing accuracies reported at 99.35% and 98.38%, respectively [17]. Abdalla et al. used the VGG16 Encoder algorithm to detect binary segmentation of 400 oilseed images in two different environments with an accuracy of 96% [18].

TABLE 1 SUMMARY OF AVAILABLE PUBLIC PLANT DISEASE DATASETS

Dataset Name	No. of images	No. of categories	Vegetable type	Comments
New Plant Diseases[1]	200,000	38		One of the most widely used datasets in the field, consisting of approximately 87K RGB images of healthy and diseased crop leaves
PlantVillage[1]	50,000			Downloaded over 7,000 times.
Rice leaf[1]	120	3	Rice	Leaf smut, Brown spot, and Bacterial leaf blight, with 40 images in each category
Plant disease[1]	1530	3		healthy, powdery and rusty
Cassava plant[1]	--	4	Cassava	Cassava Bacterial Blight (CBB); Cassava Brown Streak Disease (CBSD), Cassava Green Mottle (CGM); Cassava Mosaic Disease (CMD), Healthy
Cucumber plant[1]	695	2	Cucumber	--
PlantifyDr[1]	125,000	10	Fruits	10 different plant types (apple, pepper, cherry, citrus, maize, grape, peach, potato, strawberry, and tomato), with a total of 37 plant diseases.
IDADP[2]	17,726	15	Cereal	rice, wheat, and maize
Corn rust[3] [27]			Corn	Hefei Institute of Intelligent Machinery, Chinese Academy of Sciences
Rice false smut[3]	1,056		Rice	Chinese Academy of Sciences, Hefei Intelligent Machinery Research
Cotton plant disease[1]	--	5	Cotton	Aphids, Army worm, Bacterial Blight, Powdery Mildew and Target spot
Sugarcane Leaf[4]	2,569	5	Sugarcane	Healthy, Mosaic, maculopathy, rust and jaundice
Corn/Maize Leaf[1]	4,188	4	Corn	SMARANJIT GHOSE: Common Rust, gray leaf spots, wilted leaves and healthy
Hops[1]	1,102	2	Hop	Google images and individual growers from social media forums

1 HTTPS://WWW.KAGGLE.COM/

2 http://www.icgroupcas.cn/website_bchtk/index.html

3 https://www.scidb.cn/en

4 https://data.mendeley.com/datasets/9424skmnrk

Lin et al. used the U-Net segmentation algorithm to segment cucumber leaves with an accuracy of 96.08% [19]. Wiesner-Hanks et al. implemented a binary segmentation task to identify maize diseases using the ResNet - Crowdsourced algorithm for binary segmentation, which divides the image into homogeneous regions according to defined criteria and generates a binary image of the plant disease with the highest accuracy rate, i.e. 99.79% [20]. A collective summary of the existing methods with relevant details are provided in Table 2.

TABLE 2 ALGORITHM AND DETECTION METHOD OF THE CITED ARTICLE

Author	Algorithm	Methods	Accuracy	Crop Type
Hasan, M.M.	R-CNN	Object Detection	93.40%	Wheat
Toda Y	YOLOV3 - DenseNet	Object Detection	95.75%	Apple
Zhang, S.	GPDCNN	Multiclass Classifier	94.65%	Cucumbers
Hari	PDDNN	Multiclass Classifier	86.00%	Diverse
Picon	RESNET-MC1	Multiclass Classifier	98.00%	Diverse
Howlader	AlexNet	Multiclass Classifier	98.74%	Guava
Arunangshu Pal	AgriDet	Multiclass Classifier	96.00%	Diverse
Amal Mathew	Voting classifier	Multiclass Classifier	92.00%	Diverse
Punam Bedi	CNN& CAE	Multiclass Classifier	98.38%	Diverse
Nagasubramanian	3D-CNN	Binary Classifier	95.73%	Soybean
Abdalla	VGG16 Encoder	Binary Segmentation	96.00%	Oilseeds
Lin	U-Net Segmentation	Binary Segmentation	96.08%	Cucumbers
Wiesner-Hanks	ResNet - Crowdsourced	Binary Segmentation	99.79%	Corn

2.3 Challenges, Perspectives, and our proposed solution

Looking at the plant disease target detection algorithms in recent years, the overall accuracy is high, basically above 90%, and accurate detection of cucumber, wheat and various plants can be achieved regardless of which detection method is used. In some studies, due to the different data sets used, the proposed algorithms ignore disease areas without obvious boundaries when identifying them, i.e., they are unable to accurately detect the extent of the disease. However, most of the current papers only address plant disease detection in a single environment or photographs taken of individual leaves in the laboratory, and there is less disease target detection for photographs taken in complex environments or in natural scenarios. Difficulties such as lighting, shading, superimposition and background bias exist in practical applications, therefore disease identification in complex natural conditions is an area of ongoing research. Another key limitation of existing system is using pre-developed embedded systems are which negatively affects adaptability and flexibility of the system.

In this paper, we develop a fully scalable system from scratch (including sensors, cameras, connectivity, etc.) which can be generalised for several agricultural applications. This system mitigates the aforementioned challenge of collecting data/images from an operational farm. Using this system, we have collected and processed the "Strawberry Dataset" which their details and the results will be discussed next.

III. PROPOSED SYSTEM

In this section, we showcase the details and results of our plant disease detection system. As shown in Figure 1 (a) and (b), we develop a network of sensors and cameras that are wirelessly connected to a base station, continuously monitor the conditions of plants, and seamlessly transmit the images and sensors' data. In the following, first, the hardware specifications and design for data/image capturing and communication is described. Then, the results of applying deep learning models on both collected dataset and also existing datasets are provided.

3.1 Hardware

The imaging system is composed of an SVC3 camera that is able to capture images at 2560 x 1920 resolution. The camera features 20x optical zoom as well as 255 degrees pan and 120 degrees tilt that enables the capture of high-quality close-up images of the plant matter over a large area. A Raspberry Pi based camera controller which is deployed on the same Wi-Fi Network as the cameras, requests images from each camera at fixed intervals during the day before uploading them via a Wi-Fi access point. In contrast to the imaging system, the sensor network has been custom designed to meet the needs of this application. A 3D rendering of the edge node's PCB can be seen in Figure 1 (c). The ATMega644p microcontroller is responsible for interfacing with 7 sensor modules; temperature, pressure, humidity, ambient light, U.V light, soil moisture and leaf wetness. The microcontroller samples the sensors roughly once every 30 minutes and uses a Semtech SX1262 LoRa

(a) sensors data collection module (b) Integrated sensors and image data collection module (c) PCB design of the sensing and communication hardware

Figure 1 The proposed IoT-based plant disease detection system implemented at Wilkin & Sons in Tiptree.

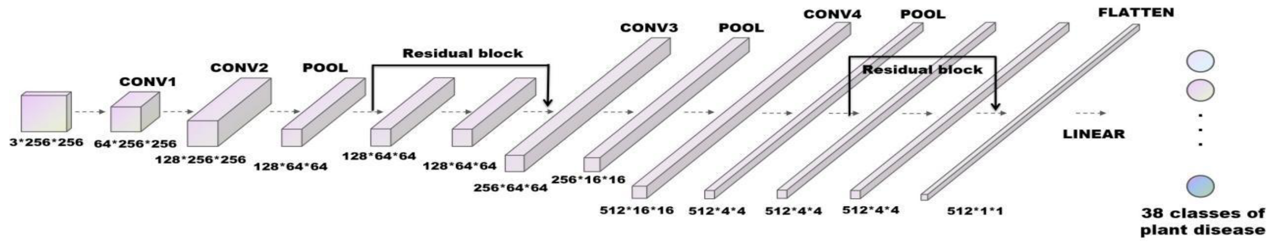
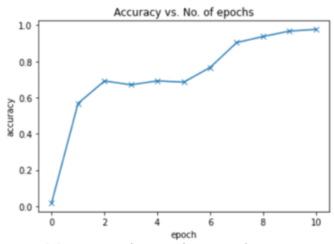
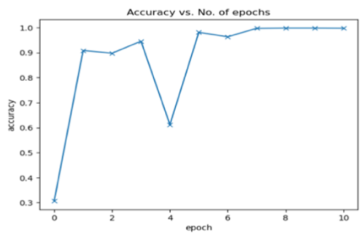


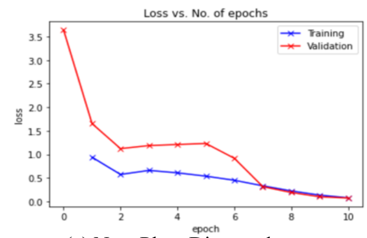
Figure 2 Structure of the proposed model based on ResNet.



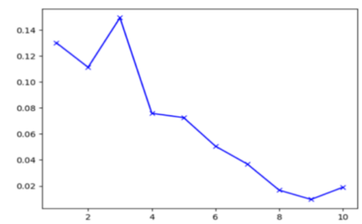
(a) New Plant Disease dataset



(b) Strawberry dataset



(c) New Plant Disease dataset



(d) Strawberry dataset

Figure 3 Accuracy ((a) and (b)) and loss function ((c) and (d)) performance using the proposed model with different datasets.

Transceiver to transmit the data to a Dragino LG01-N LoRa gateway which pushes this data to a privately hosted server that is responsible for parsing the data and storing it in a database hosted on Amazon Web Services.

3.2 Software

We used the 9-layer structure of the ResNet as a baseline. In this model, each layer feeds into the next layer and directly into the layers about 2-3 hops away. Conventional pre-processing such as image re-sizing was applied to input images where required. The network in this project uses a combination of two convolutional layers and two residual blocks, regularising each layer first, then using ReLU as the activation function and Max Pooling to reduce the size of the data and increase the speed of computation. Furthermore, the Adam Optimiser and the Cross-Entropy loss function were employed in this model. The changes made to the dimensions of each layer are indicated in the model structure diagram in Figure 2. In this model, the learning rate can change with the training rounds, and only the maximum learning rate needs to be set when setting the parameters. The model will constantly change the learning rate during training and obtain different accuracy rates in each training process. After many tries, when training the New Plant Diseases Dataset, we set the max learning rate to 0.04 to obtain highest accuracy. When training the strawberry dataset, set the learning rate to 0.01 to achieve the highest accuracy. In the last layer, the data is flattened, and linear regression is used to classify the different types of plant diseases. A total of 6,589,734 parameters were calculated to be trained.

IV. PERFORMANCE EVALUATION

4.1 Selection of dataset

To provide a comprehensive evaluation and the generalisability power of the proposed system, we use a widely used public dataset, i.e., “New Plant Diseases Dataset” as well as data collected by ourselves at Wilkin & Sons in Tiptree [21] (see Figure 1). “New Plant Diseases Dataset” is created using offline augmentation from the original PlantVillage Dataset. This dataset consists of about 87K RGB images of healthy and diseased crop leaves which is categorized into 38 different classes of 14 unique plants. The total dataset is divided into 80/20 ratio of training and validation set preserving the directory structure. And a new directory containing 33 test images is created for prediction purposes. The other dataset, called “strawberry dataset”, contains healthy strawberries and 2 types of diseases of strawberry, including Strawberry Leaf scorch and Strawberry Mildew. Each type contains about 2000 pictures of strawberries.

4.2 Experimental Results

To measure the accuracy, we have used cross-validation and considered 80% of data for training and 20% for testing. We used batch size 32 and run the model for 10 number of epochs. Figure 3 (a) and (b) show the changing trend of accuracy with rounds for the two datasets, respectively. It can be observed that the accuracy fluctuates but the overall trend is increasing and finally reaches 97.36% and 99.77%. Figure 3 (c) and (d) depict the change trend of loss with rounds, and the overall loss is low and stable. Instead of using a fixed learning rate in this project, we use a learning rate scheduler which will change the learning rate after each batch of training. There are several strategies for changing the learning rate during training, i.e., starting with a low learning rate, gradually increasing it to a high learning rate in batches over about 30% of the cycles, and then gradually reducing it to a very low value over the remaining cycles, so only the maximum learning rate needs to be set when setting the parameters. The results in these figures show that in the 10 running epochs, the accuracy of the validation set for New Plant Diseases Dataset finally reaches 97.69% and 99.77% for Strawberry dataset.

V. CONCLUSION

In this paper, a comprehensive collection and comparison among existing plant disease datasets were provided. We designed and implemented a full embedded electronic systems including sensors’ data capturing as well as RGB camera image acquisition from strawberry plants. A new modified ResNet model was proposed and applied to both collected strawberry plants’ data as well as a public dataset. The obtained results show high detection accuracy and hence the effectiveness of the proposed system. We further aim to expand the number of data collection nodes throughout the farm, enriching the dataset and developing a fusion model to analyse both sensors’ data and image data simultaneously to provide early and accurate prediction of potential diseases.

ACKNOWLEDGEMENT

This research was funded by a Knowledge Transfer Partnership (KTP) from Innovate UK (Partnership No: 12298), between Wilkin & Sons Ltd and the University of Essex.

REFERENCES

[1] Navrozidis I, Alexandridis T K, Dimitrakos A, et al. Identification of purple spot disease on asparagus crops across spatial and spectral scales.Computers and Electronics in Agriculture, 2018, 148:322-329.

[2] Shoaib Muhammad, Shah Babar, EI-Sappagh Shaker, Ali Akhtar, Ullah Asad, Alenezi Fayadh, Gechev Tsanko, Hussain Tariq, Ali Farman, An advanced deep learning models-based plant disease detection: A review of recent research, Frontiers in Plant Science, Vol. 14, 2023

[3] SHAO Mingyue, ZHANG Jianhua, FENG Quan, CHAI Xiujuan, ZHANG Ning, ZHANG Wenrong. Research progress of deep learning in plant foliar disease detection and identification. Intelligent Agriculture (in Chinese and English),2022,4(01):29-46.

[4] G. Sahitya, C. Kaushik, P. S. Simha, D. Reddy, S. Kumar and V. Vinay, "Leaf Disease Detection Using Raspberry Pi," *2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, Kharagpur, India, 2022, pp. 1-7

[5] G. Sahitya, C. Kaushik, P. S. Simha, D. Reddy, S. Kumar and V. Vinay, "Leaf Disease Detection Using Raspberry Pi," *2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, Kharagpur, India, 2022, pp. 1-7

[6] G. Sahitya, C. Kaushik, P. S. Simha, D. Reddy, S. Kumar and V. Vinay, "Leaf Disease Detection Using Raspberry Pi," *2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, Kharagpur, India, 2022, pp. 1-7

[7] C. -H. Huang, B. -W. Chen, Y. -J. Lin and J. -X. Zheng, "Smart Crop Growth Monitoring Based on System Adaptivity and Edge AI," in *IEEE Access*, vol. 10, pp. 64114-64125, 2022

[8] Hasan, M.M., Chopin, J.P., Laga, H., Miklavcic, S.J., 2019. Detection and analysis of wheat spikes using convolutional neural networks. Plant Meth. 15

[9] Toda Y, Okura F. How Convolutional Neural Networks Diagnose Plant Disease. Plant Phenomics. 2019 Mar 26;2019:9237136.

[10] Zhang, S., Zhang, S., Zhang, C., Wang, X., Shi, Y., 2019. Cucumber leaf disease identification with global pooling dilated convolutional neural network. Comput. Electroni. Agric. 162, 422–430.

[11] Hari, S.S., Sivakumar, M., Renuga, P., karthikeyan, S., Suriya, S., 2019. Detection of plant disease by leaf image using convolutional neural network. In: 2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN), pp. 1–5.

[12] Picon, A., Seitz, M., Alvarez-Gila, A., Mohnke, P., Ortiz-Barredo, A., Echazarra, J., 2019. Crop conditional convolutional neural networks for massive multi-crop plant disease classification over cell phone acquired images taken on real field conditions. Comput. Electroni. Agric. 167

[13] Howlader, M.R., Habiba, U., Faisal, R.H., Rahman, M.M., 2019. Automatic recognition of guava leaf diseases using deep convolution neural network. In: 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), pp. 1–5.

[14] Nagasubramanian, K., Jones, S., Singh, A.K., Sarkar, S., Singh, A., Ganapathysubramanian, B., 2019. Plant disease identification using explainable 3d deep learning on hyperspectral images. Plant Meth. 15, 98.

[15] Arunangshu Pal, Vinay Kumar, AgriDet: Plant Leaf Disease severity classification using agriculture detection framework, Engineering Applications of Artificial Intelligence, Volume 119, 2023, 105754

[16] Amal Mathew, Anson Antony, Yash Mahadeshwar, Tanisha Khan, Apeksha Kulkarni, Plant disease detection using GLCM feature extractor and voting classification approach, Materials Today: Proceedings, Volume 58, Part 1, 2022, Pages 407-415.

[17] Punam Bedi, Pushkar Gole, Plant disease detection using hybrid model based on convolutional autoencoder and convolutional neural network, Artificial Intelligence in Agriculture, Volume 5, 2021, Pages 90-101

[18] Abdalla, A., Cen, H., Wan, L., Rashid, R., Weng, H., Zhou, W., He, Y., 2019. Fine-tuning convolutional neural network with transfer learning for semantic segmentation of ground-level oilseed rape images in a field with high weed pressure. Comput. Electroni. Agric. 105091, 167.

[19] Lin, K., Gong, L., Huang, Y., Liu, C., Pan, J., 2019. Deep learning-based segmentation and quantification of cucumber powdery mildew using convolutional neural network. Front. Plant Sci. 10, 155.

[20] Wiesner-Hanks, T., Wu, H., Stewart, E., DeChant, C., Kaczmar, N., Lipson, H., Gore, M.A., Nelson, R.J., 2019. Millimeter-level plant disease detection from aerial photographs via deep learning and crowdsourced data. Front. Plant Sci. 10, 1550.

[21] D. Usman, A. Hossein, A. Vahid, N. Chris and I. Andrey, "Visual sensor network based early onset disease detection for strawberry plants," *2023 IEEE Applied Sensing Conference (APSCON)*, Bengaluru, India, 2023, pp. 1-3, doi: 10.1109/APSCON56343.2023.10101121.