A Fungus Detection System for Greenhouses Using Wireless Visual Sensor Networks and Machine Learning

Asmaa Ali
School of Computing
Queen's University
Kingston, Canada
ali@cs.queensu.ca

Hossam S. Hassanein
School of Computing
Queen's University
Kingston, Canada
hossam@cs.queensu.ca

***Abstract*—Greenhouses are proliferating across Canada. Greenhouse crop production requires considerable attention. The only way to maintain the production growth is by controlling the greenhouse atmosphere and monitoring the plants so that they remain healthy in the greenhouse. In this paper, we utilize a Wireless Visual Sensor Network (WVSN) with machine learning and image processing to observe any deficiency, pest, or disease presenting on the leaves of the plants. We distribute camera sensors throughout the greenhouse. Each camera sensor node captures an image from inside the greenhouse and use machine learning and image processing techniques to detect the presence of fungus. When a fungus is detected, the camera sensor node sends a message to the sensor node via the wireless sensor network to measure the humidity and then send a message to the actuator to re-set accordingly. This paper demonstrates how Hough forest machine learning and image processing can be successful in detecting fungus present on crop plant leaves from the images taken from camera sensors in the greenhouse. Cross-validation was applied to measure the performance of the system. The results are highly promising. There was a 94% success rate in detecting the fungus.**

***Index Terms*—Wireless Visual Sensor Network, Machine Learning, Hough Forest, Image Processing.**

I. Introduction

Large production crop greenhouses are growing rapidly. With this growth, comes a significant need to maintain production for economic reasons. The only way to do this is by ensuring that plants remain healthy by providing a healthy atmosphere. First, to ensure healthy weather in greenhouses, we must control the atmosphere parameters of the greenhouse, such as temperature, humidity, CO_2, water, and light. We can achieve this by applying a wireless sensor network (WSN) which will measure and control these parameters, then send a message from the sensor node to the actuator to perform the necessary actions if any parameter is not optimum, such as watering and spraying the required herbicides, fungicides and insecticides. Second, to have healthy plants in greenhouses, we must monitor the growth of the plant. We can do this by applying a Wireless Visual Sensor Network (WVSN) with camera nodes that will monitor the growth of the plant and detect any disease in the plant from the captured images. In our case, we are trying to detect fungus which grows on the plant leaves. Powdery mildew is a fungal disease that looks like a powdery white coating on the leaves and stems of infected plants. A powdery mildew infection usually begins as a few spores on the leaves but quickly spreads. The white powdery surface is a thick coating of the fungi spores [1]. It can eventually cause yellowing the leaves and premature leaf drop. This type of fungus increases in 99% humid conditions with moderate 25^oC temperatures. In a greenhouse when the summer is humid, powdery mildew almost always makes an appearance. It can affect any plant. In extreme cases, it results in leaf yellowing and dropping; stunted plant growth; distortion of buds, blooms, and fruit; and eventually, overall weakening of the plant. As in Fig. 1.

Fig. 1. Powdery Mildew

It is well documented that this type of fungus can cause many diseases in the plant. These diseases can reduce crop production, which will lead to economic losses. Several methods can be used to diagnose and determine what harmful agent is affecting the plant leaves. These methods include visual inspection, soil analysis, plant tissue analysis, bioassays, and field tests [2]. These methods are used as the first step in exploration; however, since field tests are expensive, difficult to administer, and they can be done only in a laboratory we need a new in-house, inexpensive technology to detect and diagnose different kinds of diseases with minimal human interaction.

The use of technology in agriculture has been increasing. Depending on its nature agriculture technology can be biochemical (pesticides and fertilizers), or implemented into farm machinery. Mechanical and information technology can

978-1-7281-0960-2/19/$31.00 ©2019 IEEE

be applied to agriculture, for example in monitoring growth and controlling pests, geophysical measurement systems, flood detection and precision agriculture [3] [4]. Moreover, there are sensor systems for monitoring the environment, such as ambient temperature, humidity, wind, to name a few [5]. Monitoring systems are based on Wireless Sensor Network (WSN) technology [6]. WSN technology will not create a new agricultural product but will help improve existing techniques to improve the diagnosis of plant diseases and ensure final product quality [7].

Wireless Sensor Networks (WSNs) have been used in countless applications [22]- [24]. One of these applications is measuring environmental parameters inside a greenhouse. However, in a greenhouse, we need more than WSNs to monitor parameters, we need to maintain healthy plants and find plant diseases. In this paper, we propose a Wireless Visual Sensor Network (WVSN) that uses machine learning with image processing in each wireless node to determine if a plant has unhealthy leaves. When the WVSN detects an unusual status in the plant leaves, it will send a message to the sensor node to measure the humidity inside the greenhouse. If the humidity is high, the sensor node will send a message to the actuator to decrease the humidity in the greenhouse.

This paper is structured as follows: In the next section, we review current work related to the problem. Then, in section III and IV, we present the fundamental methodology used in our proposed solution. Next, in section V, we describe the results of the algorithm applied to a variety of test images, and draw a comparison with other previous works. Finally, in section VI, we conclude our proposed system.

II. Related Works

There is very little research in the area of combining WSNs with image processing and pattern recognition in agriculture.

The authors in [8] mentioned two systems for image recognition. They explained the structure, the recognition algorithms, and the neural classifier. Of many applications created, one of their applications was for image recognition based on an adaptive control system for micro mechanics where a neural classifier was used for texture recognition of metal surfaces. The authors also used pesticides to kill insects by using a web-camera based computer vision system to automate the recognition of larvae. Their system sought to locate the insect and larvae early so that they could reduce the use of pesticides. The system consisted of neural classifiers which would detect the insect from a captured image. Recognizing the larvae and sensing warmth to indicate the larvae was active are not easy tasks because of the existence of different colors, shapes, sizes and positions. They used pre-processing techniques, then trained the system. Their system still was not efficient enough to distinguish between textures related to the larvae and those related to the background of the image.

Research done by the authors in [9] had the same type of the system used in the previous work, but the purpose of this work was to use a back-propagation ANN model to distinguish between weeds and baby corn. The authors used a series of cameras to obtain high-quality images. Each image was preprocessed from the bitmap format with image processing to indexed images based on the (RGB) color system. Then, each index color acted as input for the ANN. The output value was 0 or 1, which represented whether the image was weeds or baby corn. The processing time was 20 hours for training the network. This process can help reduce the use of herbicide sprays if it decreases the training time.

Another work involved in recognizing weeds [10] used a fuzzy logic system to create a weed map that would help determine the location of the weed to use the right amount of herbicide. They also used a digital camera and a personal computer for more testing. Their system was able to locate some of the weeded areas resulting in using less herbicides, reduced soil and water pollution, and cost savings.

The authors of [11] used machine vision to detect a worm in maize plantings. They used a pre-processing technique that converted the image from grayscale to binary images using an iterative algorithm. First, the system segmented the leaves and divided them into pixels. Second, the images were divided into blocks. Blocks that contained a more significant amount of leaf surface were selected. These selected blocks were recognized as damaged or undamaged by counting the objects in each block. Their system performed well in some cases.

In [12], the authors merged three thresholding strategies, fuzzy method, Otsu method, and Isodata algorithm, to determine whether the field was covered with oat or frost. They stated that this merger provided better results than taking each method separately.

In addition, the work in [13] presented the use of image processing to measure the water droplet size and distribution of agricultural sprinklers. They used the properties of Fourier analysis and correlation in the frequency domain. The purpose of this paper was to obtain a direct measurement of sprinkler drops, which would help avoid exceeding the size of the drop that would lead to soil erosion, surface sealing, and infiltration, as well as to minimize the size of the drop to not be affected by wind drift and that alters the pattern of irrigation. This study would help the farmer control the size of the drops and maintain the right amount of water for the crop.

Another use of visible light image processing and machine vision system was presented in [20] and [21] to detect diseases in the field. Their systems achieved a good detection rate with some restrictions on input, such as taking images only from the top view of the plant with uniform background and taking images only of a single centered leaf. These restrictions make the system unsuitable for autonomous detection.

Using a camera provides more information and benefits over sensor networks alone as in [14]. The authors used a camera sensor network for recognition, tracking and detection. Their work introduced low-latency detection, low power, and efficient recognition. However, their work depended on using light image processing which would not be efficient in detecting pests or disease.

Most of the previous works done on detecting diseases and pests used a digital camera with image processing. To the best of our knowledge, no work such as our proposed system has been done in a greenhouse. We utilize a wireless visual sensor network, a wireless sensor network, a machine learning technique and image processing in a greenhouse that is fully cluttered with varying light degrees to detect the diseases, pests and, in our case, powdery mildew fungus on plant leaves.

III. Wireless Visual Sensor Networks (WVSNs)

A wireless Visual Sensor Network has a camera-equipped sensor node. The very small sensor camera can capture visual data, as well as process and transmit image/video information [15]. The greenhouse will have a WSN with actuator that controls the environmental atmosphere and communicate with a WVSN that capture images to monitor the growth of the plants, detect any fungal diseases or pests that can affect the plant. This communication will minimize human interaction, as shown in Fig. 2. In general, the monitored area is an immense place, which means we must deal with a large number of images. To improve the performance (in terms of storage and processing) and reduce the response time of the image processing unit, we must place cameras such that there is no overlap between images taken by those cameras. Each camera has the same focal length, angle of view, and resolution. The placement of the camera sensor in greenhouse is beyond the scope of this paper.

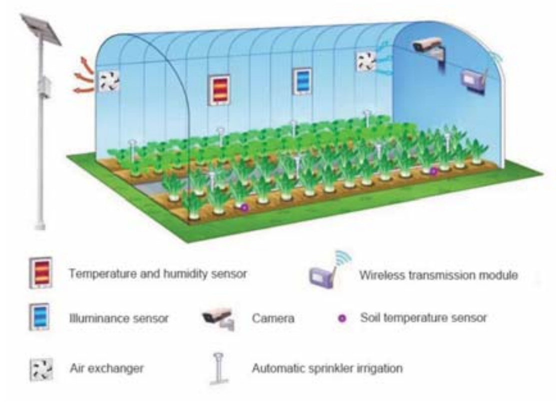

Fig. 2. Typical Greenhouse

IV. Machine Learning

Using WVSN to capture images in a greenhouse internally requires Hough forests machine learning for detection. Hough forest is a combination of the generalized Hough transform with Random forests [16] [25]. The generalized Hough transform is a method of estimating the parameters of a shape from its boundary points [17]. It extends the classic version for simple shapes like lines and circles by parameterizing in an R-table, with no need for an analytical form. Random forest is a supervised learning method that consists of a collection of weaker random classification decision trees. In these decision trees, the mean and the mode of the classes are calculated, and that is the output of the forest [18]. During the day, using many camera sensors, images were taken from greenhouses. Several issues were faced with concerning the detection of fungus in the vision system. Inconsistencies in lighting, occlusion, diffusion, and color similarity were some of the problems in greenhouses. Hough forest was chosen as the detection algorithm due to its robustness to occlusion and noise. The detection of mildew fungus occurs in three steps. First, the images are captured using WVSN. Second, the captured images are processed to remove the background clutter using image processing. Third, the mildew fungus is detected. The following sections describe each of these stages.

A. Captured Images Using WVSN

Camera sensor nodes are distributed in a greenhouse in a way such that there is no overlap between each one. The camera is located in a position that will creat maximum resolution images. Images are taken from different placements and angles of different camera sensor nodes in a greenhouse during the daytime. The sensor camera that has been used has (12MP, 50mm focal length, 1/2.55-inch sensor, dual-pixel PDAF, and f/1.5-2.4 variable-aperture lens), as well as another camera sensor (12MP, 2x focal length, f/2.4 lens, 1/3.6-inch sensor, AF). Also, a single LED flash and OIS is in both cameras. The distance between the camera sensor and the plant was between approximately 30 and 40cm to achieve better resolution at the same time with no overlap. Samples of the plant leaves in the greenhouse images are shown in Fig. 3. The data set consisted of 282 images at 1960 x 4032 pixels/image. Images were taken with different levels of occlusion. The levels varied between images from low to highly occluded and cluttered. The images were visually inspected for fungus (powdery mildew), which was observed in the images. These images were then used to create labeled image samples in which each image sample either had fungus or did not have fungus.

Fig. 3. Plants in the Greenhouse

B. Image Processing

One of the most prominent issues in disease detection techniques is background clutter in a greenhouse setting. The problem with background clutter is the high possibility of false detection, which will decrease efficiency and accuracy by searching to cover an area that does not contain the object of interest. The use of segmentation image processing for the fungus will remove the background clutter. The first step is to take the RGB color spaces from the image. The second step is to take the difference of R color image and G color image. The third step is to convert the images to grayscale. The fourth step is to create the mask. The fifth step is to apply a median filter over the mask to remove the noise. The sixth step is to extract the foreground. The final result will have the region of interest (ROI) image, as shown in Fig. 4. The results are now ready to be used as input into the machine learning detection.

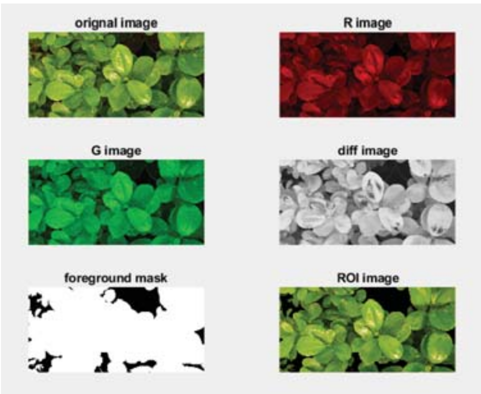

Fig. 4. Preparing Images

C. Machine Learning Detection

In this section, we propose to use Hough forest machine learning for detection. Hough forests combine the learning properties of random forests with the detection properties of Hough transforms. As explained in Section IV. The Hough transform used Hough space, which is an accumulator matrix that counts the vote generated from R-table, for each pixel. Now, random forests use multiple decision trees. Each tree generates an output to create a prediction. A random subset of features was chosen for each split of tree branches. A training process is shown in Fig. 5, and a detection process is shown in Fig. 6.

1) Data Set: Our data set consisted of 282 images at 1960 x 4032 pixels/image. Labeled sample images were prepared using a semi-automatic approach to create patches. Five hundred and two patches were created; 260 positive patches had fungus and 242 negative patches did not. All patches were re-sized to 256 x 256. A Hough forest was trained with positive fungus images with the negative background removed. Samples of the patch images used for training are shown in Fig. 7 and Fig. 8. These images are not part of the testing set.

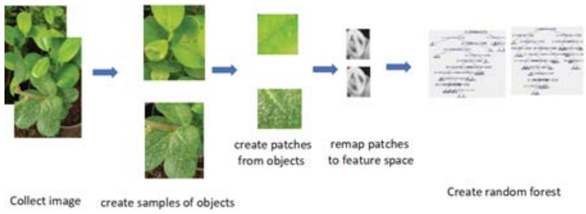

Fig. 5. Hough Forest Training

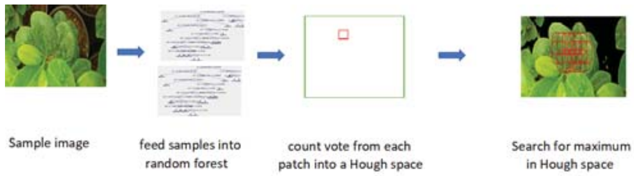

Fig. 6. Hough Forest Detection

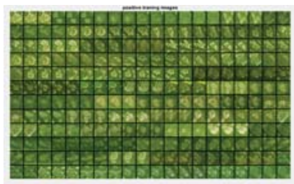

Fig. 7. Positive Training Patches

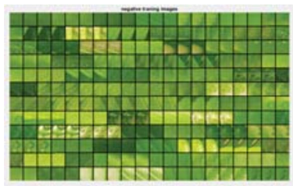

Fig. 8. Negative Training Patches

2) Training Process: A general outline of training and detection is shown in Fig. 5, and Fig. 6. These patches were extracted randomly from each image sample and carried different features. These features had information used for constructing each tree that included each channel of the L*a*b color space, first and second discrete differentials, using the Sobel operator, as well as nine histograms of gradients, as described by Leibe et al. in [19]. J. Gall's re-implementation of the Hough forest described in [16] was used to train the classifier. The patches will be selected randomly with their location and image classification. Then they will be passed along to the root node of the decision tree. It will split into two new nodes which will maximize the information gain. Every node knows the position of the patches relative to the center of the image and image classification. The process will keep repeating until it reaches the stopping point. Many trees were trained using the same steps thus creating a forest. The forest contained ten trees, each with a depth of 18 nodes, as shown in Fig. 9.

The detection of a fungus started by inputting an image into the system, then creating sample patches and their feature spaces. Each patch processed through a tree until it reached a leaf node. The leaf node had the position and the class information about that patch, which would be used to create a vote into a Hough space. All trees have votes in Hough space. The highest number of votes indicates the correct location of the object i.e., fungus.

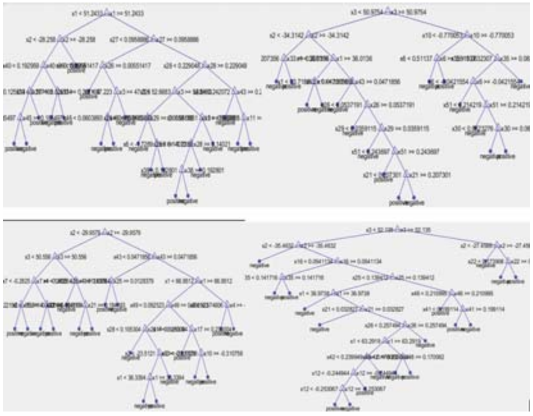

Fig. 9. Tree Forest

V. RESULTS

With a machine learning algorithm, it is necessary to divide the dataset into training, testing, and validation. The training will be in the training set, and the testing will be in the testing set. This process will be repeated n times. Each time, the data will have been randomly selected to create different training and testing subsets. Here, k-fold cross-validation will be applied, and this will divide the data into training and testing. In our case, k=5. This approach ensures that every image sample will be tested and that testing sets will not overlap. Matlab software was used to apply 5-fold cross-validation on 280 images taken from inside the greenhouse. The results of 5-fold cross-validation are presented in Fig. 10. Using the receiver operating characteristic (ROC) parameter, the ROC is calculated by comparing the true positive (TP) rate to the false positive (FP) rate. Also, in Fig. 10, we calculated the area under ROC curve (AUC) which evaluates how good the classifier is, and how accurate the output is. In our case, the AUC was 96.96%.

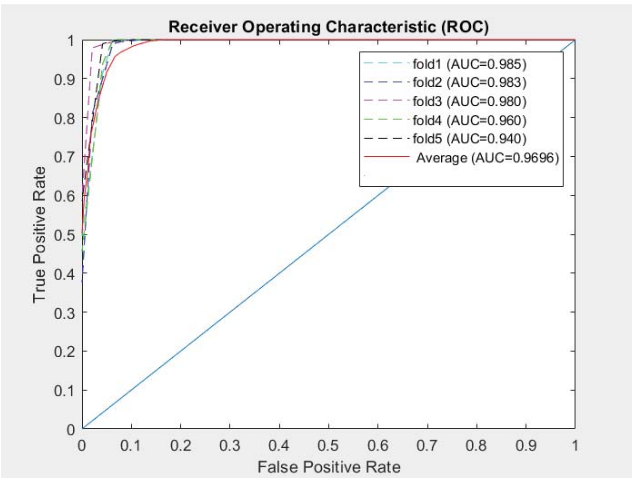

Fig. 10. ROC for Hough Forest Trained with Fungus Image Patches

The results of testing 100 images are shown in Table I. The results show detection rates of 71% true positive, and 23% true negative. The total of all positive detection rates is 94%. The rate of false negatives was low at 5 %, and the rate of false positives was much lower at 1%.

TABLE I
RESULTS OF FUNGUS DETECTION ON THE TESTING DATASET

Test	Fungus in image %	No fungus in image %
Detected fungus	TP (71)	FP (1)
No detected fungus	FN (5)	TN (23)

A. Result Statistics

In Table II, statistics show how well our system performed in predicting fungus based on the images.

TABLE II
STATISTICS RESULTS

Sensitivity	93.4%	**Specificity**	95.8%
Positive Likelihood Ratio	22	**Negative Likelihood Ratio**	0.06
Poistive Predictive Value	98.6%	**Negative Predictive Value**	82.1%

- Sensitivity shows the probability that our test results are positive when fungus is present. In our case, we have high probability 93.4 %.
- Specificity shows the probability that our test results are negative when fungus is not present. In our case, we have high probability 95.8 %.
- A positive likelihood ratio of greater than 1 indicates the test result is associated with fungus. In our case, the result was 22, which conforms with our output results.
- A negative likelihood ratio of less than 1 indicates that the test result is associated with an absence of fungus. In our case, the result was 0.068, which confirms with our output results.
- Positive predictive value shows the probability that the fungus is present in the images when the test is positive. In our case, the probability value was 98.6% (very high).
- Negative predictive value shows the probability that the fungus is not present in the images when the test is negative. In our case, the probability value was 82.1% (very high).

Sample output results of applying the Hough forest machine learning on the images were true negative detection (healthy plant) and true positive detection (fungus found), shown in Fig. 11 and Fig. 12, respectively. False negative detection and false positive detection are shown in Fig. 13 and Fig. 14, respectively.

B. Comparison with Other Works

Table III compares our proposed system of applying Hough forest machine learning on images taken from WVSN from different angles and placement against each image process

Fig. 11. True Negative Detection

Fig. 12. True Positive Detection

Fig. 13. False Negative Detection

Fig. 14. False Positive Detection

used by authors in previous works. The images in the work in [20] were taken from the top view, which minimized the clutter from the background images. The work in [21] cropped the leaf images before applying color-texture detection, which also reduced the clutter of background images.

TABLE III
COMPARISON WITH OTHER WORKS

Author	Method	Images	Detection Rate%
[20]	Color feature detection	Images with top view	70
[21]	Color-texture detection	Images contain leaves only	67-88
This paper	Hough forest, color and background removal	Images with different view and varying light	94

VI. CONCLUSION

Early and fast detection of any diseases or pests in a greenhouse is an essential step and part of an integrated management strategy needed to maintain the health of the plants. Automated plant disease detection in an environment like a greenhouse is complex due to the surroundings are a fully cluttered, large-scale, and uncontrolled environment. This paper shows how Hough forest machine learning can detect powdery mildew fungus in images of the leaves of plants taken from a wireless visual sensor network. The detection rate of 94% is a good indication of our proposed system performance. Maintaining a low false positive rate is very important for a successful detection system, as each positive detection would require sending messages to sensor nodes to measure the humidity of the greenhouse and re-set accordingly.

REFERENCES

[1] D. Jacob,D.R. Sztjenberg, D. Rav David and Y. Elad, "Conditions of development of powdery mildew of tomato caused by Oidium neolycopersici," Phytopathology98, pp. 270-281, 2008.

[2] R.N. Van Den Driessche, "Prediction of mineral nutrient status of trees by foliar analysis," Botan. Rev. , vol. 40, pp. 347-394, 1974.

[3] M. Garcia; D. Bri, S. Sendra, J. Lioret, "Practical deployments of wireless sensor networks: A survey," Int. J. Adv. Netw. Serv., vol. 3, pp. 136-178, 2010.

[4] J. Lloret, M. Garcia, D. Bri and S. Sendra, "A wireless sensor network deployment for rural and forest fire detection and verification," Sensor., vol. 9, pp. 8722-8747, 2009.

[5] C. Anand, S. Sadistap, S. Bindal, B. A. Botre and K. S. N. Rao, "Wireless multi-sensor embedded system for Agro-industrial monitoring and control," Int. J. Adv. Netw. Serv., vol. 3, pp. 1-10, 2010.

[6] Di. D. Palma, L. Bencini, G. Collodi, G. Manes, F. Chiti, R. Fantacci and A. Manes, "Distributed monitoring systems for agriculture based on wireless sensor network technology," Int. J. Adv. Netw. Serv., vol. 3, pp. 11-21, 2010.

[7] D. Sunding and D. Zilberman, "The agricultural innovation process: Research and technology adoption in a changing agricultural industry. In Handbook of Agricultural and Resource Economics," Gardner, B., Rausser, G.C., Eds.; North-Holland: Amsterdam, The Netherlands, 2000.

[8] T. Baidyk, E. Kussul, O. Makeye and, A. Vega, "Limited receptive area neural classifier based image recognition in micromechanics and agriculture," Int. J. Appl. Math. Inform.,vol. 2,pp. 96-103, 2008.

[9] C. C. Yang, S. O. Prasher, J. A. Landry, H. S. Ramaswamy and A. Ditommaso, "Application of artificial neural networks in image recognition and classification of crop and weeds," Can. Agr. Eng., vol. 42, pp. 147-152, 2000.

[10] C. C. Yang, S. O. Prasher, J. A. Landry, J. Perret and H. S. Ramaswamy, "Recognition of weeds with image processing and their use with fuzzy logic for precision farming," Can. Agr. Eng., vol. 42, pp. 195-200, 2000.

[11] D. G. Sena, F. A. C. Pimto, D. M. Queiroz and P. A. Viana, "Algoritmo de processamento de imagens para controle localizado de pragas na cultura do milho," In Proceedings of 2nd International Symposium on Precision Agriculture, Viosa-Minas Gerais, pp. 1214, 2002.

[12] A. Macedo-Cruz, G. Pajares, M. Santos, I. Villegas-Romero, "Digital image sensor-based assessment of the status of Oat (Avena sativa L.) Crops after frost damage," Sensors, vol. 11, pp. 6015-6036, 2011.

[13] P. E. Cruvinel, E. R. Minatel, M. L. Mucheroni, S. R. Vieira, S. Crestana, "An Automatic Method Based on Image Processing for Measurements of Drop Size Distribution from Agricultural Sprinklers," In Proceedings of Simpsio Brasileiro de Computao Grfica e Processamento de Imagens (SIBGRAPI), Caxamb- Minas Gerais, vol. 9, pp. 39-46, 1996.

[14] P. Kulkarni, D. Ganesan, P. Shenoy, Q. Lu, "SensEye, SensEye: A Multitier Camera Sensor Network," In Proceedings of the 13th Annual ACM International Conference on Multimedia (MM05), PP. 611, 2005.

[15] Yap,G. H. Florence, and Y. Hong-Hsu, "A survey on sensor coverage and visual data capturing/processing/transmission in wireless visual sensor networks," Sensors, Switzerland, vo. 14, pp. 3506-3527, 2014.

[16] J. Gall, V. Lempitsky, "Class-specific Hough forests for object detection," pp. 1022-1029, 2009.

[17] D. D. Ballard, D.D., "Generalizing the Hough transform to detect arbitrary shapes, pattern recognition," vol. 13, pp. 111-122, 1981.

[18] L. Breiman,"Random Forests, Machine Learning," vol. 45, pp. 5-32, 2001.

[19] K. Rematas, B. Leibe, "Efficient object detection and segmentation with a cascaded Hough Forest ISM," IEEE international conference on computer vision workshops (ICCV workshops), PP. 966-973, 2011.

[20] L. Velzquez, N. Sasaki, Y. Nakano, K. Meja-Muoz, J.M. Romanchik and E. Kriuchkova, "Detection of powdery mildew disease on rose using image processing with Open CV.," Rev. Chapingo. Ser. Horticult, vol. 17, pp. 151160, 2011.

[21] M. Zhang, Q. Meng, "Automatic citrus canker detection from leaf images captured in field," Pattern Recognition Lett, vol. 32, pp. 2036-2046, 2011.

[22] A. Moubayed, A. Refaey, and A. Shami, "Software defined perimeter: State of the art secure solution for modern networks, To appear in IEEE Network Magazine, 2019.

[23] E. Uchiteleva, A. Shami and A. Refaey, "Virtualization of Wireless Sensor Networks Through MAC Layer Resource Scheduling," in IEEE Sensors Journal, vol. 17, no. 5, pp. 1562-1576, 1 March1, 2017.

[24] T. Yu, X. Wang and A. Shami, "Recursive Principal Component Analysis-Based Data Outlier Detection and Sensor Data Aggregation in IoT Systems," in IEEE Internet of Things Journal, vol. 4, no. 6, pp. 2207-2216, Dec. 2017.

[25] C. Gomez, A. Shami, and X. Wang, "Machine Learning Aided Scheme for Load Balancing in Dense IoT Networks," Sensors 2018, 18, 3779.