ISSN 2063-5346

Disease and Pest Detection in crops using Computer Vision: A Comprehensive Study

Harsh Thakkar
Department of Information Technology,
Pune Institute of Computer Technology, Pune, Maharashtra, India -411043
Email: harshnj27@gmail.com

Aditi Pingle
Department of Information Technology,
Pune Institute of Computer Technology, Pune, Maharashtra, India -411043
Email: aditipingle20@gmail.com

Sarth Kulkarni
Department of Information Technology,
Pune Institute of Computer Technology, Pune, Maharashtra, India -411043
Email: sarthkulkarni23@gmail.com

Rutwik Saraf
Department of Information Technology,
Pune Institute of Computer Technology, Pune, Maharashtra, India -411043
Email: rutwiksaraf3@gmail.com

Radhika V. Kulkarni
Associate Professor,
Department of Information Technology,
Pune Institute of Computer Technology, Pune, Maharashtra, India -411043
Email: radhikavikaskulkarni@gmail.com
ORCID ID:- https://orcid.org/0000-0001-6614-6297

ABSTRACT

Detecting pests and diseases in crops has always been difficult for farmers. They can significantly reduce crop yield if not detected at an early stage. Pests and diseases occur in different phases of crop development. As a result, a continuous monitoring system is essential. Various advancements in the field of deep learning and machine learning have aided in the monitoring and diagnosis of crop- related diseases, resulting in improved production. This research aims to work in the area of disease and pest recognition using machine learning and deep learning. This paper examines and summarizes different techniques that researchers used in their studies on disease and insect detection for the health monitoring of crops. The research paper puts forward two methodologies: First to identify diseases in cotton leaves with the help of three image classification models namely VGG16, Resnet50, and Inception V3, and attained an accuracy of 99.58%, 85.03%, and 95.38% respectively on a custom dataset of 875 images. Second to detect pests in crops using the VGG16 and Inception V3 and achieved an accuracy of 99.78% and 97.96% respectively on the pest dataset available on Kaggle.

Keywords: Crop Disease Detection, Pest Diagnosis, Deep Learning, Machine Learning, Computer Vision

1.0 Introduction

Agriculture is important in sustaining any economy. The presence of pests and diseases significantly hamper the quality of the crops. This leads to losses in crop production that not only affect the farmers but also impact the economy. Identification of pests and diseases in crops has always posed a challenge to not just farmers but to researchers as well. Traditional methods of detection include manually inspecting each crop. This process is very time-consuming, troublesome, and less accurate [1, 2]. Therefore, the need for automated systems for disease and pest detection arises. machine learning and deep learning play a crucial role in this field. It reduces the time consumed and significantly improves accuracy.

There are various challenges that automated systems for disease and pest detection in crops experience. Among the most serious issues are the mild symptoms of several plant diseases during their early life cycles. Another

ISSN 2063-5346

issue in this field is detecting multiple infections in single or multiple leaves [3]. Other difficulties include the small size of the dataset and recognition performance when light and occlusion are present [4]. Some of the challenges are handled by the researchers. The goal of this study is to provide a thorough overview of current research in the field of disease and pest identification using deep learning and machine learning.

This paper is divided into the following five sections: The second section summarizes several crop health monitoring studies. The research methodology is discussed in Section 3. Section 4 discusses the implementation and outcomes. Finally, Section 5 concludes with significant research impacts.

2.0 Literature Review

The below section discusses different machine learning and computer vision techniques for disease and pest detection in crops.

2.1 Crop Disease Detection

Rastogi A. et al. [5] propose a method to identify diseases in the Maple and Hydrangea leaves. This study uses pre-processing techniques like resizing and reshaping, and feature extraction techniques and also uses k- means to group similar pixels on their dataset that was prepared using a predefined library to extract images. An Artificial Neural Network (ANN) is employed on this dataset using MATLAB and depending upon the percentage of infection, the leaves are classified from very low risk to very high risk. In [6], the study classifies the rice leaf images as diseased and healthy on a custom dataset. Several pre- processing techniques along with k-means clustering for image segmentation are used. Support Vector Machine (SVM) and ANN were used, with 92.5% and 87.5% accuracy, respectively.

The study in [7] discusses how k-means clustering is used to identify maturity and disease in tomato leaves. Using specific threshold values, grayscale images are converted to RGB images, and the disease's percentage penetration is calculated. [8] discusses an approach towards the detection and classification of two kinds of disease namely leaf spot and leaf blotch for turmeric using K-means clustering and SVM and the results are visualized using the MATLAB GUI. Md. Arifur Rahman et al. [9], propose a study to improve the image segmentation techniques used for the categorization of plant diseases using a sequential model achieving an accuracy of 99.25% compared to 93.25% achieved using k-means clustering. Several image enhancements, image segmentation, and feature extraction techniques are being used.

Kulkarni Omkar et al. [10] present a study regarding disease detection in the crops on the PlantVillage dataset. Two transfer learning models namely MobileNet and InceptionV3 have been used for the same and an accuracy of 99.62 % and 99.74 % has been achieved respectively. Several pre- processing techniques like resizing and segmentation have also been employed.

The implementation of a Convolutional Neural Network (CNN) on a custom dataset of 3663 images consisting of healthy and diseased leaves of tomato and apple plants is presented in [11]. Several pre-processing techniques like normalization, resizing and feature extraction are used and an accuracy of 88.7 % is achieved. In [12], a custom dataset of the sugarcane crop consisting of around 13842 images classified as healthy and infected is prepared and fed to a CNN. The study achieved a validation accuracy of around 95%. Militante S. V. et al. [13] propose a method to identify diseases in 6 different types of crops. The PlantVillage dataset, which contains approximately 35000 images, is used for this purpose. The dataset is further passed on to the CNN model and an accuracy of 96.5 % is achieved.

The research community has described various machine learning and deep learning techniques used to identify various crop diseases, and it was discovered that deep learning models performed better than machine learning techniques [14]. [15] develops a mathematical model based on deep learning for plant disease diagnosis and recognition that improves accuracy. The Region Proposal Network (RPN) is used to detect and locate leaves in complicated environments. Images segmented using the RPN algorithm's findings incorporate the characteristics of symptoms through the Chan-Vese (CV) method. The data reveal that the approach is 83.57% accurate. Table 1 gives a summary of the various methods employed in the detection of diseases in crops.

Table 1: Summary of Crop Disease Detection Automated System

Reference	Technology used	Dataset used	Accuracy
[10]	Deep learning ,Transfer learning, InceptionV3,MobileNet	PlantVillage	MobileNet: 99.04%, InceptionV3: 99.45%
[6]	ANN and SVM	40 pictures from the Rice Research Facility for every class	SVM: 92.5%, ANN: 87.5%
[7]	CNN	Custom dataset of 3663 images	88.7%

ISSN 2063-5346

[9]	Deep neural network, K-means Clustering	PlantVillage	99.25%
[12]	CNN	Custom dataset of 13,842 images	95%
[13]	CNN	PlantVillage	96.5%

2.2 Pest Detection

[16] examines the unsupervised segmentation approach for surveilling crop development and health. Individual segments are employed to evaluate crop changes such as the appearance of a flower, fruit, deficit, disease, or pest are examined for size, colour, and texture. The results show 94% accuracy in segmenting cabbage with Black Moth pest, 81% accuracy in acquiring segments afflicted by Helopeltis pest on tea leaves, and 92% accuracy in order to detect fruit spotting bugs on a citrus tree. The work in [17] employs shape feature classification and nine-fold cross-validation on both datasets. The feature importance techniques used in SVM provide improved accuracy for insects with comparable structures in their wings and bodies. Due to feature independence, the Gaussian Nave Bayes classifier performs poorly on both datasets and equally weights all features. For 2 datasets consisting of 9 and 24 class insects, the CNN model gets the highest classification rates of 91.5% and 90%, respectively. A study in [18] presents an identification model for the detection of cotton diseases and pests like bacterial blight, spider mites, and leaf miners in the region of Ethiopia. Using the CNN transfer learning technique, the model's accuracy is around 96.4%. The automated approach towards the identification of leaf affectedness and the type of pest presence on various plant leaf photos are investigated in [19]. The work in [20] recommends the use of models such as faster-RCNN for the detection of crop pests such as aphids, flea beetles, Cicadellidae, flax budworms, and red spider mites, as well as the development of mobile-based applications for pest detection in the cloud. The results are close to 99.0% accurate.

Li Y. et al. [21] propose a method in which a manually executable and substantiated crop pest dataset is introduced, and a fine-tuned GoogLeNet model is proposed to deal with the complex backgrounds provided by agricultural scenes, with pest classification results that outperform the original model. When compared to the state-of-the-art technique, the fine-tuned GoogLeNet model performed much better. They looked at ten common agricultural pests: Gryllotalpa, Leafhopper, locust, Oriental fruit fly, Pieris rapae Linnaeus, Snail, Spodoptera litura, Stinkbug, Cydiapomonella, and Weevil.

Selvaraj M. G. et al. [22] discovers several banana pests and diseases with symptoms in varied components of the plant. It creates a smartphone-assisted disease identification tool for banana farmers. Performance and validation measures are also used to assess the accuracy of various models in automated disease detection approaches. The study in [23] used the widely used Apriori algorithm to determine the relationship rules between weather conditions and the occurrence of cotton bugs. It creates an LSTM-based technique for predicting the occurrence of cotton pests and illnesses, with an Area Under Curve (AUC) of roughly 0.97.

Reddy K. A. et al. [24] discuss several strategies and algorithms for pest identification and categorization using a computer vision approach. This research provides a useful and precise methodology for detecting the impacted image. The proposed approach makes use of K-Means clustering since it has higher precision than other approaches and takes less time to analyse. The research work in [25] proposes a method for autonomous monitoring and identification of insect pests. The study discusses various pest detection methods, with a focus on infrared sensors, audio sensors, and image-based classification. Vinushree N. et al. [26] propose a kernel-based Fuzzy C-Means clustering algorithm (KFCM) that was used to assess pest density in plants. The ANN is postulated for classification purposes, with a focus on obtaining plant properties. Table 2 summarizes the various methods used to detect crop diseases. Many studies have majorly concentrated on the accuracy, but because datasets can be imbalanced or balanced, we will concentrate on other parameters such as loss in our study.

Table 2: Summary of Pest Detection Automated System

Reference	**Technology used**	**Dataset used**	**Accuracy**
[16]	CNN	Custom dataset	Black Moth pest on 94%, Helopeltis pest on Tea leaves: 81%, Spotting fruits on a Citrus tree: 92%
[18]	CNN	Custom dataset of 2400 images	96.4%
[21]	FineTuned GoogLeNet	Custom dataset of 5629 images	98.91%
[4]	Improved Yolo V3 CNN	Custom dataset	92.39%
[13]	CNN	PlantVillage 75,000 images	96.5%

ISSN 2063-5346

3.0 Research Methodology

We discovered that deep learning models outperform traditional machine learning models based on our review of the literature. In this research paper, a custom dataset for disease detection is of approximately 875 cotton leaf images classified as fresh and diseased leaves, and for pest detection, the pest dataset available on Kaggle comprising a total of 3150 images divided among nine different classes of the pest was used. The datasets were categorized into train, testing, and validation for training and evaluation purposes. The images are accessible in RGB format. The dataset is pre-processed using techniques such as normalization and scaling. To address the issue of the dataset's small size, data augmentation techniques such as flipping are used. Three transfer learning models are tested for disease identification, and two transfer learning models were used for pest detection. The accuracy and loss of the results are compared. This study compares three cutting- edge transfer learning models for crop disease and pest recognition. Fig. 1 shows implementation steps. Fig. 2 presents pests present on crop leaves and Fig. 3 shows fresh and diseased cotton leaves.

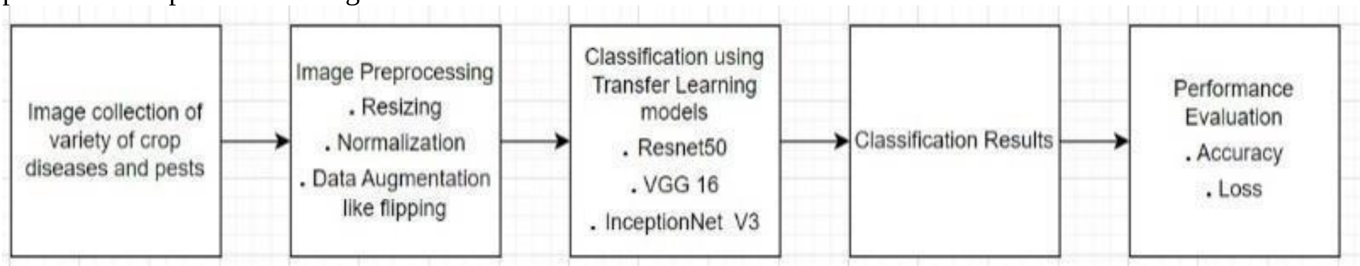
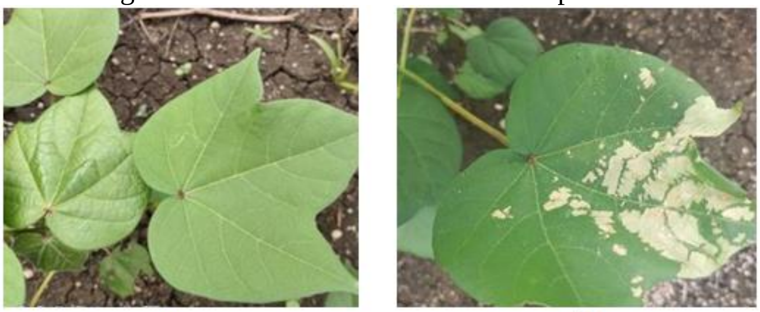


Fig.1 Implementation Steps

A B

Fig. 2A and 2B Pest Present on the Crop Leaves



A B

Fig. 3 A. Fresh Cotton Leaf; B. Diseased Cotton Leaf

4.0 Results and Discussions

The images were available in RGB format. The dataset was pre-processed by employing techniques like normalization and resizing. Data augmentation techniques like flipping were performed to tackle the issue of scarcity of images. For disease detection, three transfer models were employed, and for pest detection, two transfer learning models were used. The results were compared based on accuracy and loss. The summary of the implementation results for disease and pest detection are shown in Table 3 and Table 4, respectively. Accuracy and loss corresponding to different transfer learning models for disease detection in Cotton leaves are graphically depicted in Fig. 4 and 5 respectively. The similar results of accuracy and loss corresponding to different transfer learning models for pest detection are graphically presented in Fig. 6 and 7 respectively. These empirical results show that VGG 16 outperformed all three transfer learning models for disease detection in cotton leaves and two image classification models for pest identification in crops in this study.

ISSN 2063-5346

Table 3: Comparison of Image Classification Models for Disease Detection in Cotton Leaves

Sr. No.	Model	Loss	Accuracy
1.	Resnet50	0.3665	85.03%
2.	VGG-16	0.2549	99.58%
3.	Inception V3	0.0210	95.38%

Table 4: Comparison of Image Classification Models for Pest Detection in Crops

Sr. No.	Model	Loss	Accuracy
1.	VGG-16	0.0211	99.78%
2.	Inception V3	0.1718	97.96%

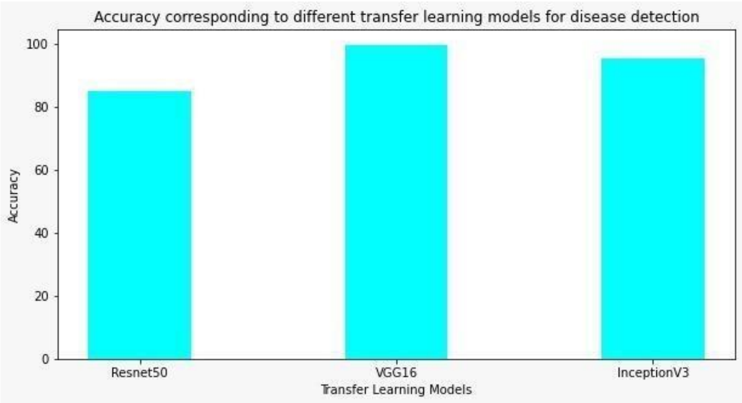


Fig. 4 Accuracy Corresponding to Different Transfer Learning Models for Disease Detection

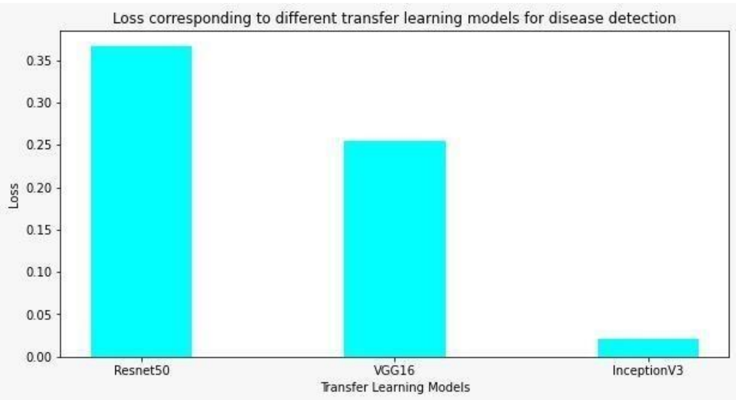


Fig. 5 Loss Corresponding to Different Transfer Learning Models for Disease Detection

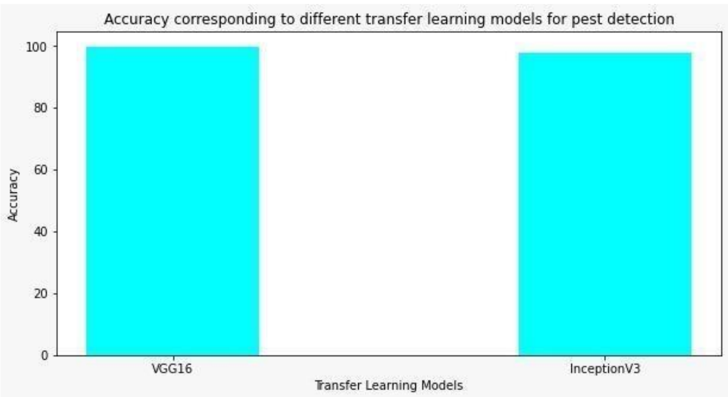


Fig. 6 Accuracy Corresponding to Different Transfer Learning Models for Pest Detection

ISSN 2063-5346

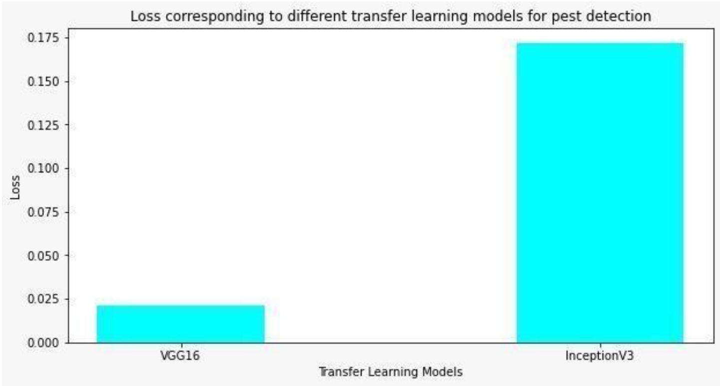


Fig. 7 Loss Corresponding to Different Transfer Learning Models for Pest Detection

5.0 Conclusion

This paper provides a detailed examination of the various machine learning and deep learning approaches used in crop pest and disease identification. The various research studies were analysed on several factors which include the dataset used, evaluation metrics, and the different state-of-the-art techniques which were employed, and a comprehensive guide was prepared for the same. Based on the study deep learning models gave better results. In this work, we investigated three image classification models for disease detection in cotton leaves and two image classification techniques for insect detection in crops, and VGG 16 obtained the maximum accuracy in both cases. The size of the datasets was one of the major challenges that drew our attention. Most studies prepared their datasets from images obtained from the internet or captured in real-time. Due to this, the size of the datasets was relatively small. To achieve more accurate results, we should focus on creating datasets with a large number of images to help neural networks perform better. In the future, we intend to implement various real-time object detection models on a larger dataset for disease and pest detection.

References

1. Deepika P. and Kaliraj S. (2021). A survey on pest and disease monitoring of crops. In: 2021 3rd International Conference on Signal Processing and Communication (ICPSC 2021), May 2021, 156–160. doi: 10.1109/ICSPC51351.2021.9451787.
2. Ngugi L. C., Abelwahab M., and Abo-Zahhad M. (2021). Recent advances in image processing techniques for automated leaf pest and disease recognition – A review. Information Processing in Agriculture, 8(1), 27–51. doi: 10.1016/j.inpa.2020.04.004.
3. Hasan R. I., Yusuf S. M., and Alzubaidi L. (2020). Review of the state of the art of deep learning for plant diseases: A broad analysis and discussion. Plants, 9(10), 1–25. doi: 10.3390/plants9101302.
4. Liu J. and Wang X. (2021). Plant diseases and pests detection based on deep learning: A review. Plant Methods, 17(1), 1–18. doi: 10.1186/s13007-021-00722-9.
5. Rastogi A., Arora R. and Sharma S. (2015). Leaf disease detection and grading using computer vision technology & fuzzy logic. In: 2015 2nd International Conference on Signal Processing and Integrated Networks (SPIN), 500-505. doi: 10.1109/SPIN.2015.7095350.
6. Ghyar B. S. and Birajdar G. K. (2017). Computer vision based approach to detect rice leaf diseases using texture and colour descriptors. In: 2017 International Conference on Inventive Computing and Informatics (ICICI), 1074-1078. doi: 10.1109/ICICI.2017.8365305.
7. Francis M. and C. Deisy (2019). Disease detection and classification in agricultural plants using convolutional neural networks — A visual understanding. In: 2019 6th International Conference on Signal Processing and Integrated Networks (SPIN), 1063-1068. doi:10.1109/SPIN.2019.8711701.
8. Kuricheti G. and Supriya P. (2019). Computer vision based turmeric leaf disease detection and classification: a step to smart agriculture. In: 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), 545-549. doi:10.1109/ICOEI.2019.8862706.
9. Rahman M. A., Islam M. M., Shahir Mahdee G. M. and Ul Kabir M. W. (2019). Improved segmentation approach for plant disease detection. In: 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT 2019), 1-5. doi: 10.1109/ICASERT.2019.8934895.
10. Kulkarni O. (2018). Crop disease detection using deep learning. In: 2018 4th International Conference on Computing Communication Control and Automation (ICCUBEA), 1-4. doi:10.1109/ICCUBEA.2018.8697390.

ISSN 2063-5346

11. Francis M. and Deisy C. (2019). Disease detection and classification in agricultural plants using convolutional neural networks — A visual understanding. In: 2019 6th International Conference on Signal Processing and Integrated Networks (SPIN), 1063-1068. doi:10.1109/SPIN.2019.8711701.
12. Militante S. V., Gerardo B. D. and Medina R. P. (2019). Sugarcane Disease Recognition using Deep Learning. In: 2019 IEEE Eurasia Conference on IOT, Communication and Engineering (ECICE), 575-578. doi: 10.1109/ECICE47484.2019.8942690.
13. Militante S. V., Gerardo B. D. and Dionisio N. V. (2019). Plant leaf detection and disease recognition using deep learning. In: 2019 IEEE Eurasia Conference on IOT, Communication and Engineering (ECICE), 579-582. doi: 10.1109/ECICE47484.2019.8942686.
14. Shruthi U., Nagaveni V. and Raghavendra B. K. (2019). A review on machine learning classification techniques for plant disease detection. In: 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS), 281-284. doi: 10.1109/ICACCS.2019.8728415.
15. Guo Y. et al. (2020). Plant disease identification based on deep learning algorithm in smart farming. Discrete Dynamics in Nature and Society, 2020. doi: 10.1155/2020/2479172.
16. Bhatt P., Sarangi S., and Pappula S. (2019). Unsupervised image segmentation using convolutional neural networks for automated crop monitoring. In: 8th International Conference on Pattern Recognition Applications and Methods (ICPRAM 2019), 887–893. doi: 10.5220/0007687508870893.
17. Kasinathan T., Singaraju D., and Uyyala S. R. (2021). Insect classification and detection in field crops using modern machine learning techniques. In: Information Processing in Agriculture, 8(3), 446–457. doi: 10.1016/j.inpa.2020.09.006.
18. Azath M., Zekiwos M., and Bruck A. (2021). Deep learning-based image processing for cotton leaf disease and pest diagnosis. Journal of Electrical and Computer Engineering. doi: 10.1155/2021/9981437.
19. Suthakaran A. and Premaratne S. (2020). Detection of the affected area and classification of pests using convolutional neural networks from the leaf images. International Journal of Computer Science Engineering (IJCSE), 9(1).
20. Karar M. E., Alsunaydi F., Albusaymi S., and Alotaibi S. (2021). A new mobile application of agricultural pests recognition using deep learning in cloud computing system. Alexandria Engineering Journal, 60(5), 4423–4432. doi: 10.1016/j.aej.2021.03.009.
21. Li Y., Wang H., Dang L. M., Sadeghi-Niaraki A., and Moon H. (2019). Crop pest recognition in natural scenes using convolutional neural networks. Computers and Electronics in Agriculture, 169, December 2019, 105174. doi: 10.1016/j.compag.2019.105174.
22. Selvaraj M. G. et al. (2019). AI-powered banana diseases and pest detection. Plant Methods,15(1), 1–12. doi: 10.1186/s13007-019-0475-z.
23. Xiao Q., Li W., Kai Y., Chen P., Zhang J., and Wang B. (2019). Occurrence prediction of pests and diseases in cotton based on weather factors by long short term memory network. BMC Bioinformatics, 20(Suppl 25), 1–15. doi: 10.1186/s12859-019-3262-y.
24. Reddy K. A., Reddy N. V. M. C., and Sujatha S. (2020). Precision Method for Pest Detection in Plants using the Clustering Algorithm in Image Processing. In: 2020 IEEE International Conference on Communication and Signal Processing (ICCSP 2020), 894–897. doi:10.1109/ICCSP48568.2020.9182190.
25. M. C. F. Lima, M. E. D. de A. Leandro, C. Valero, L. C. P. Coronel, and C. O. G. Bazzo "Automatic detection and monitoring of insect pests—A review," in Agriculture (Switzerland), 2020, vol. 10, no. 5, pp. 1–24. doi: 10.3390/agriculture10050161.
26. Vinushree N., Hemalatha B., and Kaliappan V. K. (2014). Efficient kernel-based fuzzy C-means clustering for pest detection and classification. In: 2014 World Congress on Computing and Communication Technologies (WCCCT 2014), 179–181. doi: 10.1109/WCCCT.2014.61.