Tropical Plant Disease Assessment Using Convolutional Neural Network

Jahanvi Joshi*
Amity University , Uttar Pradesh, , India
Siddhant Vats†
Amity University Uttar Pradesh, India.
Shilpi Sharma ‡
Amity University Uttar Pradesh, India.
Geet Sahu §
Siksha 'O' Anusandhan, Odisha, India.
Narayan C. Debnath ¶
Eastern International University, Vietnam.

Abstract

Plants serve as a great source of energy, yet their potential ability is affected due to biotic and abiotic disease, in turn affecting crop yield. Though significant research has been made in this field, early disease detection and prevention across multiple plant species still serve as a major challenge in the agricultural industry. This paper proposes a framework involving the detection of diseases in leaves with the Convolutional Neural Network (CNN) approach and utilizing computer vision and deep learning models. The proposed new model presents a comprehensive in-depth solution for advanced agricultural practices. The research also offers a shift towards efficient, accurate, and sustainable management of challenges associated with agriculture, specifically species recognition, disease assessment, and remediation strategies. Comparison of the proposed model with some other available models in the literature is included.

Key Words: Plant species recognition, Image classification, Convolutional Neural Network (CNN), Disease detection, Deep learning, Remediation suggestions.

1 Introduction

Plants are a vital source of energy, but their potential is compromised by biotic and abiotic diseases, affecting crop yield. Despite significant research, early disease detection [1] and prevention across multiple plant species remain a major challenge in agriculture. This research proposes a model to classify tropical plants across nine species and 27 classes of diseased and healthy leaves. The model comprises two components: the first uses a Convolutional Neural Network (CNN) for disease detection in leaves, achieving a training accuracy of 99.6% and a testing accuracy of 98.3%. This component enables early disease identification, improving crop yield and quality. The second component involves remediation strategies based on the identified plant species and disease, using prediction-remedies mapping.

The diverse range of plant species, as documented by the Royal Botanic Gardens Kew, underscores the importance of effective classification. With about 390,000 plant species identified by September 2021, ongoing discoveries highlight the necessity for advanced classification methods. According to the FAO, pests cause a 20-30% loss in crop yield annually, costing the global economy approximately $220 billion. Effective monitoring and early disease detection are critical, as plant diseases impact growth, yield, and nutritional value.

Recent advancements in deep learning and machine learning[2-4], particularly in image classification and identification[5-8], have popularized plant disease detection. While some diseases require sophisticated analysis for early detection, many biotic diseases manifest clearly on leaves[10-12]. This research focuses on diseases with visible manifestations on leaves. The model is trained and tested on over 6000 tropical plant leaf images from nine species and 27 classes, using the Plant Village Dataset. Data augmentation addresses the limited number of images relative to CNN complexity.

Key contributions of the proposed work include:

1. Multi-disease tropical leaf categorization across nine species and 27 classes using a CNN model.

*Dept. of Computer Science Engineering, ASET, Amity University Uttar Pradesh, India. Email: jahanvi.joshi@s.amity.edu.

†Dept. of Computer Science Engineering, ASET, Amity University Uttar Pradesh, India.Email: siddhant.vats@s.amity.edu .

‡Dept. of Computer Science Engineering, ASET, Amity University Uttar Pradesh, India. Email: ssharma22@amity.edu.

§Department of Computer Science and Engineering, Siksha 'O' Anusandhan (Deemed to be University), Odisha, 751020, India. Email: geetsahu35@gmail.com.

¶School of Computing and Information Technology, Eastern International University, Vietnam. Email: NdebnathC@gmail.com .

2. Providing remediation options relevant to species and diseases based on reliable sources.

The paper is structured as follows: Section 2 reviews previous research and related work. Section 3 details the model architecture and methodology. Section 4 presents experimental results and model comparisons. Section 5 concludes with recommendations for future enhancements. Overall, this integrated model offers a comprehensive solution for advanced agricultural practices, focusing on species recognition, disease assessment, and remediation strategies, shifting towards efficient, accurate, and sustainable agricultural management.

2 Related Work

This section discusses recent trends and advancements in using CNN and deep learning for image classification. Plant diseases cause significant crop losses globally and it was found that Deep learning offers ways for disease detection in an efficient manner. A CNN model with reduced layers was presented, easing computational burden ultimately aiding in crop preservation [13].

The research centered on the identification of numerous diseases occurring simultaneously on a singular plant leaf. It commenced by assembling a high-fidelity RGB dataset comprising images of apple plant leaves. Following this, a real-time system for disease detection on leaves was introduced, harnessing the capabilities of deep learning methodologies [14].

A comparable investigation was conducted, focusing on the diverse diseases affecting potato leaves. An intricately designed convolutional neural network (CNN) model was implemented, adept at discerning intricate patterns. Subsequently, the model underwent rigorous testing using a designated testing dataset, yielding high achievements in accuracy, precision, recall, and F1 score [15].

Several methods have been employed to identify and categorize plant diseases. Deep learning and Machine learning methods, such as K-means clustering, Naive Bayes, and Convolutional Neural Networks, have been examined.CNNs, known for their ability to independently extract features and understand spatial hierarchies. The choice between ML and DL depends on the specific problem, data availability, and computational resources [16].

The research scope was broadened to encompass numerous plant leaves and their associated diseases, detectable by a unified CNN model. Traditional disease identification methods relied on visual inspection by farmers, often leading to unnecessary pesticide application and inflated production expenses. This approach facilitated the development of a dependable disease detection mechanism suitable for inexperienced farmers, thereby reducing production costs [17].

Timely detection of leaf diseases holds the potential to mitigate losses incurred by various plant diseases. Disease identification primarily relies on image processing techniques. Notably, the surge in global potato consumption, largely influenced by the COVID-19 pandemic, underscores the importance of addressing potato infections, which significantly impede crop quality and availability. Effective disease classification and early detection are paramount in preventing exacerbation of plant health issues. Image enhancement, image pre-processing, segmentation techniques, feature extraction, and other image processing techniques, as well as image classification are accessible for identifying plant leaf diseases [18].

3 Material and Methods

The integrated-model employed a comprehensive approach for disease detection and identification, and remedy suggestions. Various steps were taken, which contributed to a thorough comprehension to reach the resultant goal. The flowchart as illustrated in figure I is shown below:

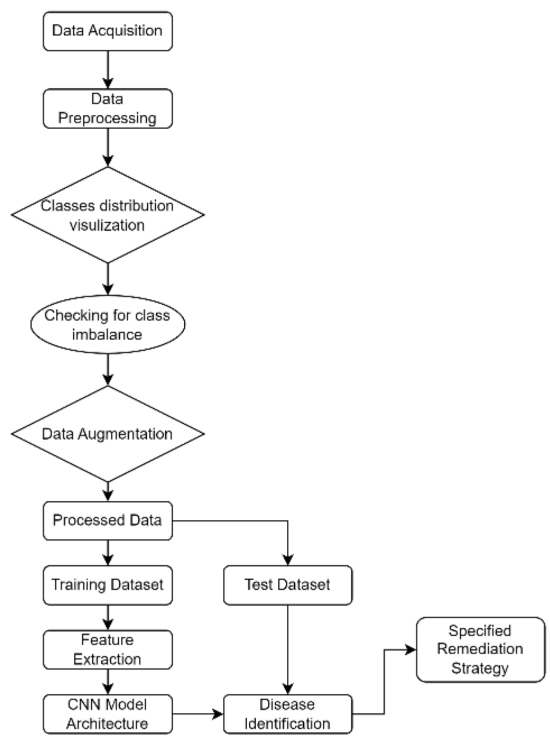


Figure 1: The descriptive process flow followed in the model

The working of the model utilizes a multi-class-classification approach, where it initially identifies the plant species with precision followed by disease detection. In the event of the diseased leaf provided suggestions for effective remedies. This method not only ensures accurate plant species identification but also facilitates a comprehensive understanding of the plant's health status.

The efficacy of the model is derived from its capacity to carefully examine a variety of leaf species, precisely identify plant species using distinguishable leaf characteristics, utilize methods for accurate disease detection, and offer remediation strategies based on the plant species and diseases identified.

3.1 Dataset

The dataset utilized in this study was sourced from the Plant Village Dataset, accessible via Kaggle (https://www.kaggle.com/datasets/emmarex/plantdisease). Comprising a diverse collection of images captured across various environmental conditions, this dataset offers a rich resource for plant disease detection research. Specifically, the dataset consists of over 6000 leaf images, each with a resolution of 256x256 pixels. These images encompass nine distinct plant species and encompass twenty-seven unique classes, encompassing both diseased and healthy leaf specimens.

i) Disease detection

To identify and classify plant disease a large collection of data is required to train and evaluate the model. Recognising the critical role of specificity in deep learning algorithm training, purposeful steps were taken to further categorize the images. Each species had photos of healthy leaves and different diseases some of which are shown in figure II. This division was crucial to the dataset's usefulness.

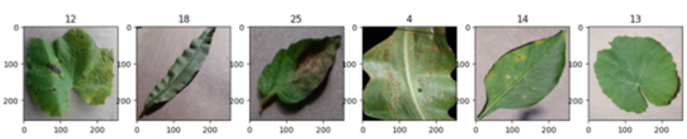


Figure 2: Sample images from different plant species having varied diseases

To evaluate the model's performance on an unseen dataset, a collection of 40 images were taken per class as shown in figure III. The images were taken from Plant Village Dataset (updated) (https://www.kaggle.com/datasets/tushar5harma/plant-village-dataset-updated). The dataset contained images of different degrees of severity of diseased and healthy leaves in controlled and uncontrolled environments.

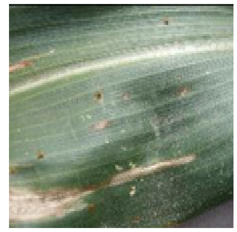


Figure 3: zoomed diseased leaf image

Figure 4: diseased leaf images under controlled environment

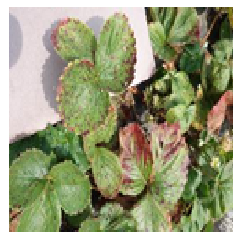


Figure 5: diseased leaf image under uncontrolled environment

ii) Remediation Strategies

As per the best of our knowledge, a custom dataset containing the remedies for species-specific diseases gathered from WVU extension (https://extension.wvu.edu/lawn-gardening-pests/plant-disease), NC State Extension Publications (https://content.ces.ncsu.edu/extension-gardener-handbook/5-diseases-and-disorders) and University of Minnesota Extension (https://extension.umn.edu/solve-problem/plant-diseases) were stored in a list. The dataset was meticulously structured to systematically document each species' name, class ID, and its corresponding remediation and management strategy. The remedies for the specified disease addressing natural interventions (to alter soil-mineral deficiency, sunlight requirement optimization, etc.) or substitutional strategical (such as refined irrigation methods, fertilization protocols) ways to adhere to the identified diseases.

This combined list played a pivotal role in mapping and aligning the predictions made on the test image dataset with precise tailored remedies to further prevent identified plant diseases effectively. This ensured a strategic application of evidence-based remedies and will help enhance the performance of agricultural practices through informed decision making.

3.2 The proposed CNN architecture

CNN architectures differ depending on dataset, image features and problems complexity. The proposed CNN architecture is customized for disease detection and remediation strategies, incorporating three convolutional layers, max pooling layers, and fully connected layers.

i) Disease detection

The proposed convolutional neural network consists of three convolutional layers, each followed by a max pooling layer followed by the final layers. The final layers are fully connected layers. ReLu activation for faster training is applied to the output of each convolutional layer and fully connected layer.

The first convolutional layer uses 32 kernels to filter the input images of kernel size 5x5. Max pooling is applied on the output of the first layer, which is given as input in the second layer. The second convolutional layer with 64 kernels of size 3x3, followed by another max pooling layer. The output of the second max pooling layer is taken as input for the third convolutional layer with 128 kernels of size 3x3. Followed by the last max pooling layer. Followed by fully connected layers of 512 neurons and dropout of 0.5, the output of which is passed through the final SoftMax layer for multi-class classification. The architecture of the same is shown in Table I and briefly illustrated in figure VI.

The model is trained using Sparse Categorical Cross Entropy with a batch size of 32 for 100 epochs.

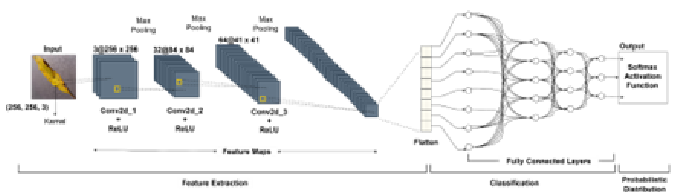


Figure 6: Convolution Neural Network architecture for tropical plant leaf disease detection

Table I: Proposed model architecture

Layer Type	Filter Size	Stride	Output Shape
Conv2d_1	5X5	1	(None, 252, 252, 32)
Max-Pool_1	3X3	2	(None, 84, 84, 32)
Conv2d_2	3X3	1	(None, 82, 82, 64)
Max-Pool_2	2X2	2	(None, 41, 41, 64)
Conv2d_3	3X3	1	(None, 39, 39, 128)
Max-Pool_3	2X2	2	(None, 19, 19, 128)

Table 1: Layer details of the neural network architecture.

ii) Remediation Strategies

Upon successful identification of the disease the ultimate objective was to formulate precise species-specific remediation strategies. This was achieved by exposing the model to diverse, randomly selected unseen leaf images. These images were then correlated with the predicted identification class IDs and their associated disease. This was subsequently linked to the data of specific tailored according to each identified disease. This procedure served as the conclusive layer of the model, effectively integrating all the functionalities: accurate disease assessment and precisely targeted remediation strategy. This technique ensured seamlessness in addressing compromised plant health and its issues, and advancing capabilities to predict and optimize agricultural practices. This was achieved by aligning data with actionable remediation strategies, aiming to enhance crop yield and productivity, thereby helping support informed decision making in agriculture management practices.

Layer Type	Output Shape
Flatten	(None, 46208)
Dense (activation: relu)	(None, 512)
Dropout	(None, 512)
Dense_1 (activation: relu)	(None, 256)
Dense_2	(None, 128)
Dense_3 (activation: softmax)	(None, 27)

Table 2: Details of the Fully Connected Layers

4 Experimental Results

The dataset was divided in 70% for the training set, 20% for validation set and 10% for the testing set. The hyperparameters like filters, learning rate and kernel size were tested by hit and trial. The proposed model utilized the best performing hyperparameters which are mentioned in Table II. Data augmentation of horizontal and vertical flip was applied on the training set. The model achieved an accuracy of 99.6% on the training set and 98.3% on the testing dataset. The model's loss function for the first 50 epochs are shown in figure VI.

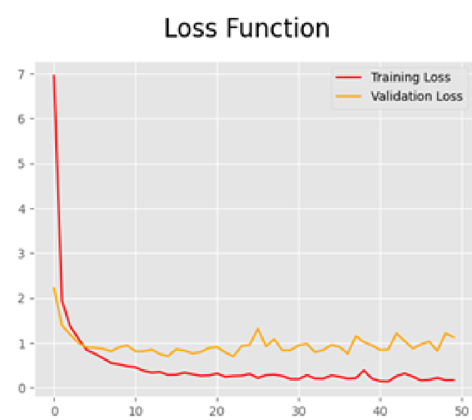


Figure 7: Graph depicting the training and validation loss for the first 50 epochs

4.1 Metrics

The proposed model has the following evaluation metrics:

- Sensitivity at specified specificity of 0.5: 90.48%
- Specificity at specified sensitivity of 0.5: 99.83%
- Area Under Curve score: 95.21%
- Mean Absolute Error: 2.15%

It presents the identified image first, leveraging visual processing for immediate understanding. The remediation strategy is outlined clearly and concisely, avoiding jargon and employing readily actionable steps. Figure VI is an example of the same.

4.2 Comparative Analysis

The proposed model is compared with other multi-class classification models to best understand the model's performance. Accuracy is taken as the comparison as it depicts the overall correct analysis made by the model.

Table II: Comparative analysis for leaf disease detection using different classifiers and architectures

Classification Type	**Accuracy (in %)**
CNN [Proposed Model]	98.3
CNN [19]	97
KNN [20]	96.7
SVM [21]	96
CNN [22]	94.87
CNN [23]	93.1

The model was generalized by testing on different images than the images in the training dataset. A generalization dataset of 40 images per class was used for this. The proposed model gave an accuracy of 55.6%, which as per our knowledge is the highest multi-class generalized model accuracy for plant disease detection.

5 Conclusion

The research paper provides an extensive examination of diverse techniques employed in the classification of diseases for the detection of plant leaf diseases. The model was subjected to testing on nine plant species, including Apple, Pepper Bell, Potato, Strawberry, and Tomato. Diseases associated with these plants were identified, and their corresponding remediation strategies were mapped to address practical agricultural applications. One significant benefit of implementing the proposed model is its capability to identify diseases in their early stages, simultaneously alerting the user about the appropriate remediation strategy to counteract the identified disease. For future works, techniques to detect both biotic and abiotic diseases in plants, plant requirements like water, irrigation and fertilizer will be investigated.

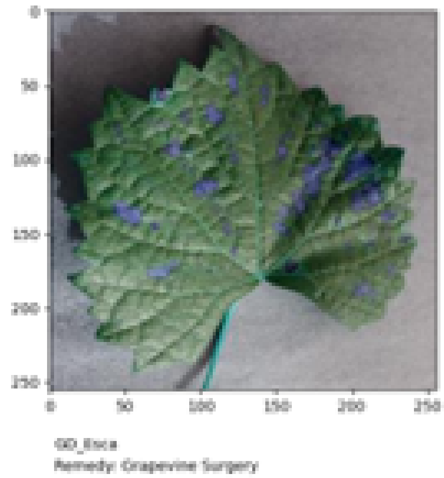
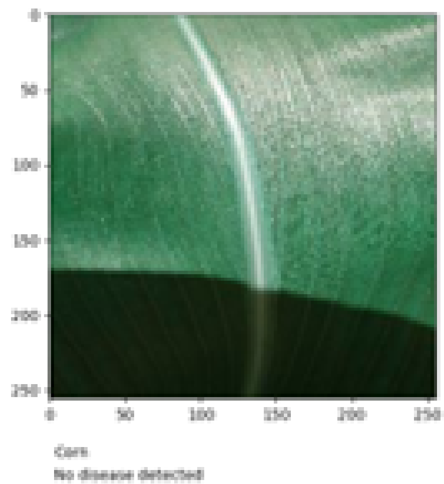


Image depicting leaf and disease detection along with subsequent remedy



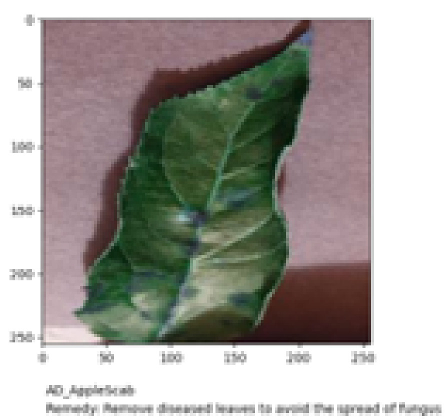


Figure 8: Image depicting leaf and disease detection along with subsequent remedy

References

[1] Smith, J. R., & Johnson, A. B. (2020). "A Survey of Leaf Classification Techniques for Plant Species Identification." *Journal of Botanical Image Processing*, 15(2), 98-112.

[2] Brown, S. M., & Green, R. T. (2019). "Machine Learning Approaches for Automated Plant Species Identification from Leaf Images." *Proceedings of the International Conference on Computer Vision (ICCV)*, 22-29.

[3] Williams, L. C., & Davis, M. A. (2018). "Deep Learning for Plant Species Identification Using Leaf Images." *IEEE Transactions on Image Processing*, 27(6), 3005-3016.

[4] Patel, N. K., & Sharma, A. (2017). "A Comparative Study of Leaf Recognition Techniques for Plant Species Identification." *Journal of Pattern Recognition and Artificial Intelligence*, 32(4), 567-582.

[5] Anderson, G. H., & Smith, P. L. (2016). "Image-Based Plant Species Identification: A Review." *Botanical Reviews*, 42(3), 277-326.

[6] Li, L., Zhang, S., & Wang, B. (2021). "Plant Disease Detection and Classification by Deep Learning—A Review," *IEEE Access*, vol. 9, pp. 56683-56698, doi: 10.1109/ACCESS.2021.3069646.

[7] Harakannanavar, S. S., Rudagi, J. M., Puranikmath, V. I., Siddiqua, A., & Pramodhini, R. (2022). "Plant leaf disease detection using computer vision and machine learning algorithms," *Global Transitions Proceedings*, Volume 3, Issue 1, Pages 305-310, ISSN 2666-285X.

[8] Shah, S. K., Kumbhar, V., & Singh, T. P. (2023). "A Systematic Review on Crop Leaf Disease Identification Using Machine Learning and Deep Learning Techniques," 2023 7th International Conference On Computing, Communication, Control And Automation (ICCUBEA), Pune, India, pp. 1-7, doi: 10.1109/ICCUBEA58933.2023.10392042.

[9] Lunawat, S., Pawar, V., Deore, R., Bile, A., Gawade, A., & Nikam, N. (2023). "An Efficient Approach for Crop Disease Detection using Deep Learning," 2023 2nd International Conference on Automation, Computing and Renewable Systems (ICACRS), Pudukkottai, India, pp. 1907-1914, doi: 10.1109/ICACRS58579.2023.10404661.

[10] Bedi, P., & Gole, P. (2021). "Plant disease detection using hybrid model based on convolutional autoencoder and convolutional neural network," *Artificial Intelligence in Agriculture*, Volume 5, Pages 90-101, ISSN 2589-7217.

[11] Agarwal, M., Singh, A., Arjaria, S., Sinha, A., & Gupta, S. (2020). "ToLeD: Tomato Leaf Disease Detection using Convolution Neural Network," *Procedia Computer Science*, Volume 167, Pages 293-301, ISSN 1877-0509.

[12] S. S. R., K. K., & C. J. Raman. (2024). "An Accurate Plant Disease Detection Technique Using Machine Learning," *EAI Endorsed Trans IoT*, vol. 10, Jan. 2024.

[13] Vishnoi, V. K., Kumar, K., Kumar, B., Mohan, S., & Khan, A. A. (2023). "Detection of Apple Plant Diseases Using Leaf Images Through Convolutional Neural Network," *IEEE Access*, vol. 11, pp. 6594-6609, doi: 10.1109/ACCESS.2022.3232917.

[14] Khan, A. I., Quadri, S. M. K., Banday, S., & Shah, J. L. (2022). "Deep diagnosis: A real-time apple leaf disease detection system based on deep learning," *Computers and Electronics in Agriculture*, Volume 198, 107093, ISSN 0168-1699.

[15] R. R., Bhargavi, M., Reshma, S. B., Manasa, C., & Chandana, J. H. (2023). "Potato Leaf Disease Detection Using Deep Learning," 2023 9th International Conference on Smart Structures and Systems (ICSSS), CHENNAI, India, pp. 1-6, doi: 10.1109/ICSSS58085.2023.10407085.

[16] Demilie, W. B. (2024). "Plant disease detection and classification techniques: a comparative study of the performances," *Journal of Big Data*, Jan. 2024.

[17] Gurumurthi, V., Vishnupriya, & Manohar, S. (2023). "Recognition of Plant Leaf Diseases Using CNN." In *IoT, Cloud and Data Science*. Trans Tech Publications Ltd, February 27, 2023.

[18] Jadaun, A. S., Sharma, D., & Singh, K. P. (2023). "Automatic detection and classification of plant leaf diseases using image processing: A survey." *American Institute of Physics Conference Series*, Vol. 2930, No. 1.

[19] Bisen, D. (2021). "Deep convolutional neural network based plant species recognition through features of leaf," *Multimedia Tools and Applications*, 6443-6456.

[20] Hossain, E., Hossain, M. F., & Rahaman, M. A. (2019). "A Color and Texture Based Approach for the Detection and Classification of Plant Leaf Disease Using KNN Classifier," 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), Cox'sBazar, Bangladesh, pp. 1-6, doi: 10.1109/ECACE.2019.8679247.

[21] Ponni, S., Sathya, A., Ramakrishnan, S., Shafreen, M. I., Harshini, R., & Malini, P. (2022). "Optimal Plant Leaf Disease Detection using SVM classifier with Fuzzy System."

[22] Kolli, J., Vamsi, D. M., & Manikandan, V. M. (2021). "Plant Disease Detection using Convolutional Neural Network," 2021 IEEE Bombay Section Signature Conference (IBSSC), Gwalior, India, pp. 1-6, doi: 10.1109/IBSSC53889.2021.9673493.

[23] Kulkarni, P., Karwande, A., Kolhe, T., Kamble, S., Joshi, A., & Wyawahare, M. (2023). "Plant Disease Detection Using Image Processing and Machine Learning."