



Soybean Leaf Disease Identification Through Smart Detection using Machine Learning-convolutional Neural Network Model

Bong-Hyun Kim¹, Ssang-Hee Seo²

10.18805/LRF-801

ABSTRACT

Background: Soybean, a vital global crop, faces threats from diverse leaf diseases impacting yield and quality. By utilizing cutting-edge Machine Learning-Convolutional Neural Network (CNN) models, this study develops a Smart Detection System for the precise identification of soybean leaf diseases.

Methods: Convolutional Neural Networks (CNNs) are used in this study to identify soybean leaf diseases. Different images that depict different diseases are used to train the CNN model. The dataset obtained from Mendeley includes three essential categories: Diabrotica Speciosa, Caterpillar and Healthy soybean leaves. Labeling, grayscale conversion and scaling are all part of image processing. 80% of the dataset is used for training, 20% is used for validation and the accuracy of the model is assessed.

Result: The CNN model showcases exceptional capabilities, achieving an impressive 95 per cent accuracy in precise soybean leaf disease classification. The Smart Detection System emerges as a powerful and timely tool for disease identification, holding significant implications for advancing precision agriculture. This study underscores the transformative potential of advanced machine learning in reshaping sustainable soybean crop management practices.

Key words: Caterpillar, Convolutional neural network (CNN) models, Diabrotica speciose, Machine learning, Soybean.

INTRODUCTION

Soybeans [*Glycine max* (L.) Merrill] are one of the most important seed legumes in the world today. It is an important source of edible oil, accounting for almost 25% of global edible oil production and acting as a major reservoir (Agarwal *et al.*, 2013). Its nutritional composition helps lower the risk of diabetes and heart disease. According to predictions, there will be more than 9.1 billion people on the planet by 2050 and at the same time, food demand is expected to increase by 60% (Food Security Statistics, accessed 27 December 2009). It is therefore imperative that efforts to raise and improve the quality of crop yield be intensified. According to recent scientific evaluations, infectious biotic and abiotic diseases have a negative impact on yield potential, typically reducing it by 40%. The unequal impact on farmers in developing nations is particularly concerning, with instances of yield losses reaching an unsettling 100% (Karlekar and Seal, 2020). Continuous crop monitoring in conjunction with prompt and accurate disease diagnosis is essential. Reducing output losses and preserving agricultural sustainability are considered to require effective disease management and control strategies.

Food production must significantly expand to keep up with the growing global population (FAO) (Food Security Statistics, accessed 27 December 2009). High-yield food production must be combined with sustainable agricultural methods to protect natural ecosystems. Global food security must coexist with a high nutritional content (Carvalho, 2006). For the production of healthy crops with a high yield, cutting-edge scientific techniques are essential to employ for crop management and leaf disease detection. Comprehensive ecosystem monitoring is one area in which AI-based state-

¹Department of Computer Engineering, Seowon University, 377-3 Musimseo-ro, Seowon-gu, Cheongju-si, Chungbuk-do, 28674, Republic of Korea.

²Department of Computer Engineering, Kyungnam University, 7Kyungnamdaehak-ro, Masanhappo-gu, Changwon-si, Gyeongsangnam-do, 51767, Republic of Korea.

Corresponding Author: Ssang-Hee Seo, Department of Computer Engineering, Kyungnam University, 7 Kyungnamdaehak-ro, Masanhappo-gu, Changwon-si, Gyeongsangnam-do, 51767, Republic of Korea. Email: shseotwin@kyungnam.ac.kr

How to cite this article: Kim, B.H. and Seo, S.H. (2025). Soybean Leaf Disease Identification Through Smart Detection using Machine Learning-convolutional Neural Network Model. Legume Research. 48(6): 1043-1050. doi: 10.18805/LRF-801.

Submitted: 27-01-2024 **Accepted:** 10-12-2024 **Online:** 28-01-2025

of-the-art technologies are being applied. It is known that soybean legumes are used as a worldwide feed crop and may be processed into a variety of cuisines (Jianing *et al.*, 2022). Soybean disease is an important factor restricting the high quality and high yield of the soybean plant (Cen *et al.*, 2020; Meng *et al.*, 2022). Ensuring the health of soybean plants is critical for optimizing yields and maintaining sustainable agricultural practices. One pivotal aspect of soybean plant health is the early detection of leaf-related issues, including diseases, pests and nutrient deficiencies. Traditional methods of visually inspecting soybean fields for leaf-related concerns are often time-consuming and may lack the precision needed for timely intervention. However, recent advancements in technology, particularly in the field of computer vision, offer innovative solutions to address these challenges. Automated leaf detection

involves leveraging image processing algorithms and machine learning to analyse digital images of soybean plants. This technology enables the rapid and accurate identification of leaves, providing valuable insights into plant health and facilitating early intervention when issues arise. Agricultural research has used a range of machine learning techniques, such as support vector machine (SVMs), ANNs, decision tree architectures, K-means and k- nearest neighbors, CNN (Panigrahi *et al.*, 2020; Min *et al.*, 2024). Traditionally, image classification problems have been solved with hand-engineered features, including SURF (Untari and Satria, 2022; Pranata *et al.*, 2019), HoG (Dalal and Triggs, 2005) and SIFT (Lowe, 2004), together with some sort of learning technique in these feature spaces. However, taught representation are more effective and successful, according to a new advancement in machine learning. The main advantage of representation learning is its capacity to automatically search through large image datasets and identify features that permit the lowest possible level of error in the classification of images (Tang *et al.*, 2023; Wihardjo *et al.*, 2024; Sankaran *et al.*, 2010).

CNNs are a class of deep learning algorithms specifically designed for image analysis tasks, making them well-suited for the complex visual patterns associated with plant diseases. These neural networks can learn intricate features and patterns from large datasets, enabling them to differentiate between healthy and diseased soybean leaves with remarkable accuracy. Concerning the detection of soybean leaf disease, CNNs offer various advantages. By employing massive databases of annotated images, these networks can identify subtle signs associated with specific diseases. The automated capabilities of CNNs allow for the fast and reliable analysis of vast agricultural fields, providing farmers with a useful tool for early disease identification.

In this paper, soybean leaf diseases are identified using Convolutional Neural Networks (CNNs). To train the CNN model, a variety of images depicting different diseases are used. *Diabrotica Speciosa*, Caterpillar and Healthy soybean leaves are the three main categories included in the Mendeley dataset. Labelling, scaling and grayscale conversion are all part of the image processing process. The model's accuracy is then assessed after 80% of the dataset is used for training and 20% for validation. The performance of the model is evaluated by confusion matrix, classification report and ROC (AOC). The accuracy and loss of the model is obtained for unseen data. This work can be utilized as a reference to identify diseases in the leaf and other parts of the plants.

Related work

The most recent advancements in CNN architectures and deep learning applications for agricultural applications are covered in this section. Before deep learning, several plant diseases were categorized by machine learning and image processing techniques (Cho, 2024; Barbedo, 2013; Pydipati, *et al.*, 2005; Camargo and Smith, 2009b; 2009a).

With a digital camera, the original digital photos are taken. The images are then prepared for the following stages by using image processing techniques like segmentation, color space conversion, filtering, as well as picture enhancement. Subsequently, salient characteristics of the picture are taken out and fed into a machine learning model (Al-Hiary *et al.*, 2011). The overall accuracy of the classification will therefore be contingent upon the kind of image processing and the feature extraction techniques chosen. Conversely, more recent studies suggest that generic data may be used to provide state-of-the-art performance that is limited by the network. CNNs are supervised multi-layer networks that can automatically learn features from datasets. CNNs have shown to be at the cutting edge of performance in almost every notable classification task during the past few years. Atabay (2016) asserts that it is capable of both feature extraction and categorization by using identical architecture. Plant diseases were classified using convolutional neural networks, or CNNs. Cortes (2017) utilized an openly available dataset comprising 86,147 images of healthy and diseased plants. They developed a network of neural networks employing deep learning and semi-supervised methods to identify crop varieties and diseases across 57 distinct classifications. The successful experiment with unlabeled data resulted in Russ-net, which became operational in less than five epochs. It achieved a detection rate in the vicinity of $1e^{-5}$ and an initial training score of 80%.

Recently, object recognition and picture classification have been accomplished with convolutional neural networks (CNNs) (Atabay, 2016b; Hanson, *et al.*, 2017; Mohanty, *et al.*, 2016). Inspired by the visual system of humans, convolutional neural networks (CNNs) are a type of deep neural networks, or DNNs, that analyze images. It was suggested that several CNN architectures be used for object recognition. Among them, Alex Net (Krizhevsky, *et al.*, 2012) and LeNet (Le Cun *et al.*, 1998) have been regarded as a baseline for a variety of tasks (Atabay, 2016).

A unique class of neural networks known as ACNN has been extensively used to address a range of pattern identification issues in computer vision, speech recognition and other areas. Three architectural techniques shared weights, spatially-temporal subsampling and local receptive fields are used by CNNs to provide a degree of shift, scale and distortion invariance (Le Cun *et al.*, 1998). Many CNN architectures, like as LeNet, Alex Net, Google Net and others, have been utilized for object recognition.

A review for CNN-based plant disease categorization was presented by Lu *et al.* (2021). They assessed the major issues and fixes with CNN, which is used to classify plant diseases, as well as the DL criteria. They found that to get a more satisfying outcome, more study with more complicated datasets was needed. Golhani *et al.* (2018) outlined the advantages and disadvantages of using hyperspectral data for the diagnosis of plant leaf diseases. Within a brief period of time, they also introduced CNN

techniques for SDI development. They found that tests of SDIs on a variety of hyperspectral imaging devices at plant leaf size are necessary as long as they are relevant for appropriate crop protection. With an emphasis on potato leaf disease, Bangari *et al.* (2022) provided a review of disease detection with CNN. After looking over several studies, they concluded that convolutional neural networks are more effective in identifying the illness. It was also found that CNN significantly improved the highest level of diseases identification accuracy. Korea is known for producing a wealth of agricultural assets and is especially skilled at growing soybeans. Soybeans are key to the country's agriculture as an essential source of vegetable protein for the local diet, a raw material for numerous businesses and a valuable commodity on the market. Even with its bright future, soybean farming is not without its difficulties. One major problem is plant diseases, which can significantly reduce productivity and result in losses for farmers.

MATERIALS AND METHODS

The branch of machine learning known as "deep learning" uses deep neural networks to interpret complicated data. Many different forms of data, including text, audio, video, images, series of times, sensors, as well as Internet of Things data, may be processed using deep learning. Processing picture data is a common application for deep learning techniques such as convolutional neural networks (CNN). CNN model may be trained to distinguish between disease-infected and healthy soybean leaves with high accuracy by utilizing a dataset that includes images of both types of leaves. The methodologies involve four important stages. Data gathering is the first step and a suitable model is developed in the second stage. The training of the data is the third step and testing the model is the last step.

Dataset

All phases of the research object recognition process, from the training to the performance evaluation stages, require the correct dataset to recognize the algorithms. The training dataset used in this work was obtained from the Mendeley database (Mignoni, 2021). The dataset includes three

categories of soybean leaf images: Caterpillar, Diabrotica Speciosa and Healthy.

The images represent soybean leaves affected by caterpillars, Diabrotica Speciosa and undamaged (healthy) leaves (Fig 1). The dataset comprises a total of 6,410 images distributed among three categories: Caterpillar (3,309 images), Diabrotica Speciosa (2,205 images) and Healthy (896 images).

Soybeans have unique taxonomical characteristics. They are categorised as *Glycine max* in the plant kingdom (Plantae) and family Fabaceae. These upright, bushy plants usually grow to a height of 0.2 to 2 metres. These leaves have ovate leaflets of different sizes and are characterised by trifoliate leaves that are placed alternately along the stem. Soybean flowers are small and self-pollinating. They have the papilionaceous structure that is characteristic of legumes. Their fruit takes the form of a legume pod that contains two to four seeds of different sizes and shapes. Depending on the cultivar, soybean seeds themselves come in a variety of colors, such as yellow, green and black. Geographically, soybeans grow well all over the world, but especially in areas with mild weather. China, Brazil, Argentina and the United States are among the major agricultural hubs. It is important to recognize that subtle differences could occur between various cultivars of soybeans.

Image processing and labeling

The raw images required for the CNN classification algorithm to analyze were improved or altered using image pre-processing before training the model. Pictures from different sources can have different sizes, so it is necessary to resize and rescale them to ensure they are all the same. This method is essential for maintaining consistency and accelerating the training process because larger-sized images incur more processing costs. Before entering the data into the network, it is essential to standardize it to an identical size and format. In this work, the images were resized to 256'256'3, Then the images were subsequently turned into grayscale. The resized images of soybean leaves were sorted by kind in the following step and each image was labeled with the relevant health acronym. Three

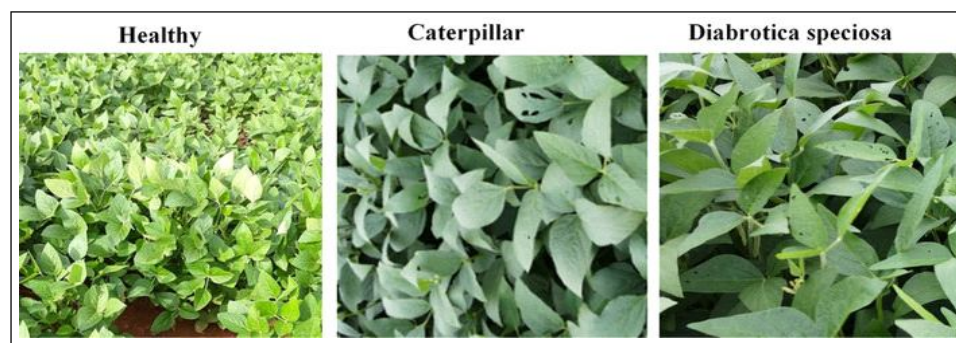


Fig 1: Images of soybean plant categories into (a) Healthy (b) caterpillar (c) diabrotica speciosa disease images.

classes (healthy, caterpillar, diabrotica speciosa) are identified using the training and test dataset.

Training dataset

The steps involved in developing a CNN model for a soybean plant disease detection system are as follows:

The first steps in creating a model are importing the required libraries and compiling a dataset of soybean leaf diseases (Fig 2). After that, the dataset will be split at an 8:1:1 ratio into three categories: training, validation and testing. This division is followed by pre-processing and data augmentation of the dataset. The data will be trained using tensor pictures from the training set as input during the training phase. The accuracy of the model can then be assessed by fine-tuning the learning outputs to the tensor labels.

Architecture of convolutional neural network (CNN) model

The following architecture is utilized to train the model (Fig 3). Many components of the architectural paradigm of CNN improve the efficiency of the learning process. Specifically, convolutional layers are used for feature extraction, pooling layers are used for feature reduction and softmax layers are used for classification. The hyperparameters used in experiments are listed in Table 1.

Convolution operations are carried out by filters (kernels) in the shape of tiny matrices found in convolutional layers. The input picture or feature map from the preceding layer is gradually covered by the filter. Information will be extracted from sections that overlap the filter at each stage. To get a value at a particular place on the feature map, this procedure entails multiplying the filter elements by the matching picture elements and adding the results. The convolution operation is defined over a two-dimensional kernel (K) and input image (I) as follows:

$$S_{(i,j)} = I * K_{(i,j)} = \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} I_{(i+x,j+y)} K_{(x,y)}$$

Where the image (I) coordinates are represented by i and j and the kernel (K) coordinates are denoted by m and n.

To lower computational complexity and efficiently manage the danger of overfitting, pooling layers work to reduce spatial dimensions. This model uses the max pooling layer as the pooling layer, which is the layer that can extract the highest value from the available feature maps. Understanding the pattern of the preceding layer can be facilitated by the fully connected layer. The fully connected layer classifies images using the features that were extracted. To forecast classes using features taken from the preceding layer, this study used a fully connected layer with a Softmax function. Given n the numerical values $x_1, x_2, x_3, \dots, x_n$ of the neurons in the previous layer, the softmax function converts these values into probabilities $P_1, P_2, P_3, \dots, P_n$.

$$P_i = \frac{e^{x_i}}{\sum_{j=1}^n e^{x_j}}$$

Where,

p_i = Probability of class i after applying softmax.

x_i = Numerical value of the i^{th} neuron in the previous layer.

Next, using test data as input, the training model will be evaluated to see the way it performs with fresh data. Once the model has completed training, its overall accuracy, recall, precision and F1 score will be determined by analyzing the classification report and confusion matrix.

Evaluation parameters

One popular algorithm for handling classification problems, such as binary and multiclass classification, is the confusion matrix. The four variables comprising the confusion matrix are True positive (TP), False positive (FP), False negative (FN) and True negative (TN) (Kulkarni *et al.*, 2020).

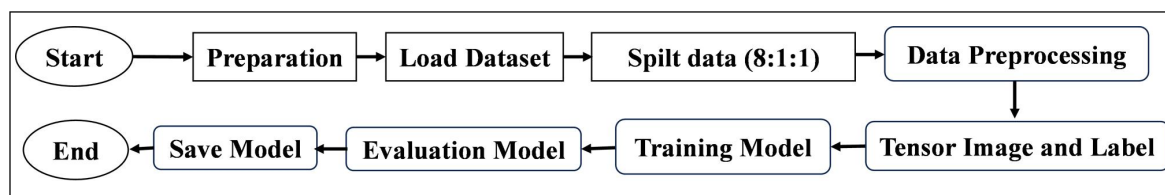


Fig 2: Model making process.

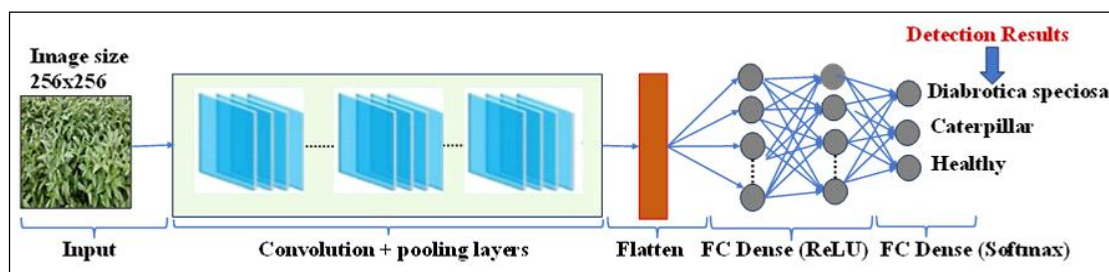


Fig 3: CNN model Architecture along with convolution and pooling layer specifications.

Actual Values	Predicted Values	
	Positive	Negative
Positive	TP	FN
Negative	FP	Tn

The evaluation stage additionally includes computations for accuracy, precision, recall and F1-score. A model's overall performance and the number of labels that an equation accurately predicts are two common ways to conceptualize accuracy.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

By comparing the total number of images predicted in the soybean leaf disease category with the number of accurately predicted images in that category, as determined by the following equation, precision is used to assess prediction accuracy:

$$\text{Precision} = \frac{TP}{TP + FP}$$

Recall compares all actual observations from a soybean disease category with a ratio representing the proportion of photos that are correctly predicted in that category.

$$\text{Recall} = \frac{TP}{TP + FN}$$

The F1 score is a measurement that integrates recall and precision.

$$\text{F1 - Score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

RESULTS AND DISCUSSION

This section presents results for training and validation with the full dataset. This analysis focuses on the results gained during training with the augmented dataset since it acknowledges that convolutional networks may learn features effectively when exposed to larger datasets. Fig 4 (a, b) displays the accuracy and loss comparison for training and validation. It is clear that, following 100 epochs, the training phase gained an accuracy rate of 95.83%. It can be concluded that additional iterations would lead to improved data accuracy under the research technique. It is important to note that as the training period gets longer, the number of epochs increases correspondingly.

In addition, each class underwent an individual assessment to determine the trained model's effectiveness and confidence score. In Fig 5, few images with actual and predicted disease are shown. This result indicates a highly confident and accurate prediction.

Fig 6 shows the anticipated classes within the confusion matrix as columns, while the rows correspond to the real classes. True positives build up the diagonal, whereas false negatives fill up the remaining part of the matrix. The confusion matrix for soybean leaf disease

Table 1: Hyper-parameter of CNN model.

Parameter	Description
No. of convolution layer	6
No. of max pulling layer	6
Dropout rate	0.5
Network weight assigned	Uniform
Activation function	Relu
Flatten	(32, 2048)
Epoch	100
Batch Size	32
Learning rate	0.000001

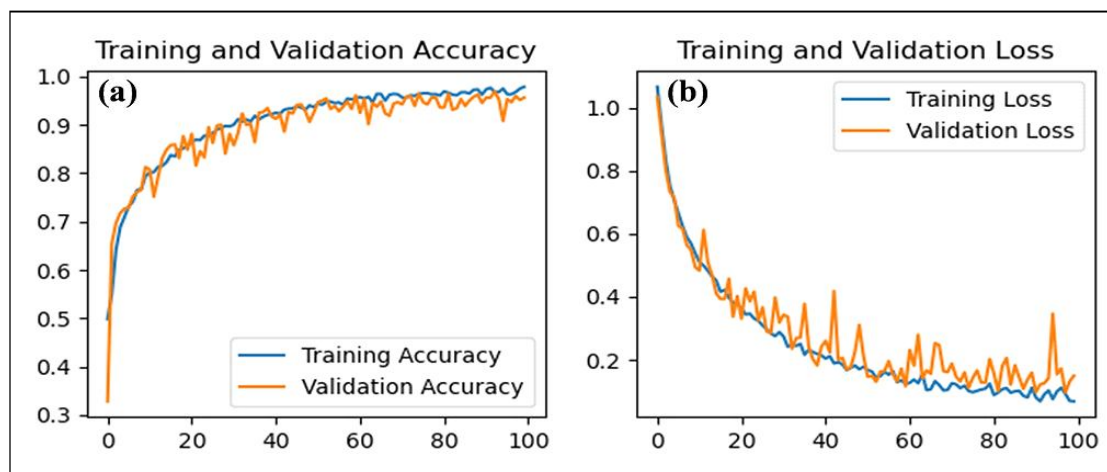


Fig 4: Training and validation (a) accuracy and (b) loss after 100 epochs.

categorizing in Caterpillar, Diabrotica Speciosa and Healthy is shown in Fig 6.

In the presented classification report, three classes are used to assess a machine learning model's performance: "Caterpillar," "Diabrotica Speciosa," and "Healthy." High precision scores, which range from 0.90 to 0.98 and show the accuracy of positive predictions for each class, illustrate the model's good predictive ability. Furthermore, recall values, which range from 0.93 to 0.99 and indicate the capacity to record true positive experiences, are consistently high. Notably good are the F1-Scores, which are a balance between precision and recall and range from 0.94 to 0.96 across classes (Table 2). The 95% overall accuracy rate suggests that the model does an excellent task of accurately identifying cases. The model's efficacy is further supported by the macro and weighted average measures, which show

balanced performance across classes and consider potential class imbalances. The classification report as a whole indicates that the model performs exceptionally well in detecting instances within each class, demonstrating its resilience in managing a variety of patterns and retaining a high level of overall predicted accuracy.

Class	True positive rate	False positive rate	AUC
Caterpillar	0.95	0.014	0.97
Diabrotica speciosa	0.99	0.071	0.97
Healthy	0.93	0.014	0.99

In addition, a measure of the overall discriminatory capacity of the model is provided by the Area Under the Curve (AUC) of the Receiver Operating Characteristic (ROC)



Fig 5: The disease kind and score of recognition confidence for the image of dataset confusion matrix.

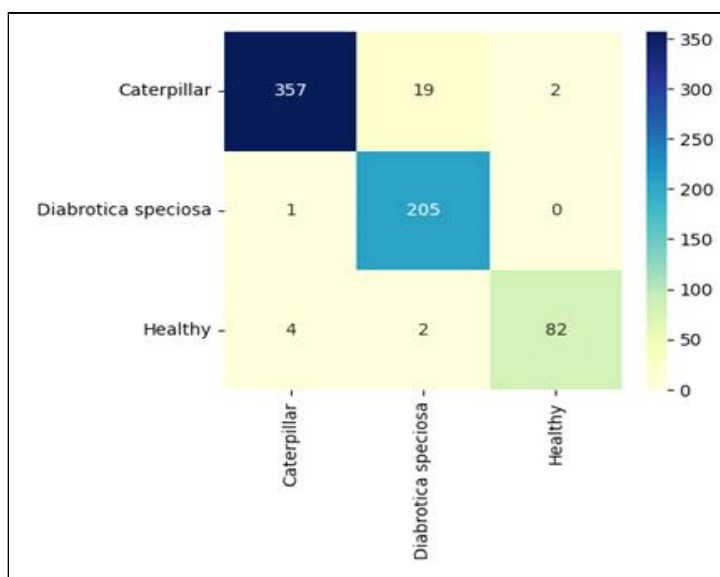


Fig 6: Confusion matrix of soybean leaf disease.

Table 2: Classification accuracies, precision, F1 score and support.

Classification	Precision	Recall	F1-score	Support
Caterpillar	0.98	0.94	0.96	378
Diabrotica speciosa	0.90	0.99	0.94	206
Healthy	0.97	0.93	0.95	88
Accuracy			0.95	672
Macro avg	0.95	0.95	0.95	672
Weighted avg	0.96	0.95	0.95	672

curve. Caterpillar, Diabrotica speciosa and Healthy had AUC values of 0.97, 0.97 and 0.99, respectively. The combined effect of these measures offers a thorough assessment of the model's performance in categorising soybean diseased and healthy classes.

LIMITATIONS FOR FUTURE WORK

- Although this study concentrated on two specific diseases that affect soybean leaves, it is important to recognize that a wide range of diseases can damage soybeans, affecting not just the leaves but also the stems and roots. Therefore, it is recommended that future research cover a wider range of diseases that may impact different sections of the soybean plant.
- Furthermore, the concatenated model's fully connected layer classifier was employed in this investigation. The total performance of the model may be compared and improved in future studies through the use of hybrid machine learning techniques such as support vector machines, decision trees and random forests. Gaining a deeper understanding of the soybeans disease detection model and improving diagnostic abilities may come from utilizing these diverse approaches.

This study's main focus is on the identification and categorization of leaf diseases that impact soybean plants; diseases that affect other plant species are not taken into account. Future research projects must investigate the wider range of plant diseases in order to ascertain the model's applicability to a variety of plant species.

CONCLUSION

In this work, photos of soybean leaves were used to automatically detect and classify plant illnesses using deep learning techniques. The created model could identify when leaves were present and differentiate between two diseases that could be identified visually. The whole method was explained, in turn, starting with obtaining the training and validation photos, followed by image preprocessing and augmentation and concluding with deep CNN training and fine-tuning. The offered retrieval technique performed well, providing confidence and correct identification results for query images containing various leaf diseases. The testing findings demonstrated an accuracy of 90% to 98% in different class assessments. The final overall accuracy of the trained model was 95.83%. Overall accuracy has not

changed significantly with fine-tuning; instead, the augmentation process had a stronger influence to provide outcomes that were acceptable.

ACKNOWLEDGEMENT

Funding statement

None.

Authors' contributions

All author contributed toward data analysis, drafting and revising the paper and agreed to be responsible for all the aspects of this work.

Data availability statement

The database generated and /or analysed during the current study are not publicly available due to privacy, but are available from the corresponding author on reasonable request.

Declarations

Author(s) declare that all works are original and this manuscript has not been published in any other journal.

Conflict of interest

The author declare that they have no conflicts of interest to report regarding the present study.

REFERENCES

- Agarwal, D.K., Billore, S., Sharma, A., Dupare, B., Srivastava, S., (2013). Soybean: Introduction, improvement and utilization in India-problems and prospects. *Agric. Res.* 2(4): 293-300. <https://doi.org/10.1007/s40003-013-0088-0>.
- Al-Hiary, H., Bani-Ahmad, S., Reyat, M., Braik, M. and AL Rahamneh, Z. (2011). Fast and accurate detection and classification of plant diseases. *International Journal of Computer Applications.* 17(1): 31-38.
- Atabay, H.A. (2016a). Binary shape classification using convolutional neural networks. *IIOABJ7.* (5): 332-336.
- Atabay, H.A. (2016b). A convolutional neural network with a new architecture applied on leaf classification. *IIOABJ.* 7(5): 226-331.
- Bangari, S., Rachana, P., Gupta, N., Sudi, P.S. and Baniya, K.K. (2022). A survey on disease detection of a potato leaf using CNN. In *2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS)*. pp. 144-149.

- Barbedo, J.G.A. (2013). Digital image processing techniques for detecting, quantifying and classifying plant diseases. *Springer Plus*. 2(1): 660.
- Camargo, A. and Smith, J. (2009a). Image pattern classification for the identification of disease-causing agents in plants. *Computers and Electronics in Agriculture*. 66(2): 121-125.
- Camargo, A. and Smith, J. (2009b). An image-processing based algorithm to automatically identify plant disease visual symptoms. *Biosystems Engineering*. 102(1): 9-21.
- Carvalho, F.P. (2006). Agriculture, pesticides, food security and food safety. *Environmental Science and Policy*. 9(7-8): 685-692.
- Cen, H., Zhu, Y., Sun, D., Zhai, L., Wan, L., Ma, Z. and He, Y. (2020). Current status and future perspective of the application of deep learning in plant phenotype research. *Transac Chinese Soci. Agricult. Eng.* 36(9): 1-16.
- Cortes, E. (2017). Plant disease classification using convolutional networks and generative adversarial networks. *Stanford University Reports*, Stanford.
- Cho, O.H. (2024). An Evaluation of Various Machine Learning Approaches for Detecting Leaf Diseases in Agriculture. *Legume Research*. <https://doi.org/10.18805/LRF-787>.
- Dalal, N. and Triggs, B. (2005, June). Histograms of oriented gradients for human detection. In 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05) (1): 886-893.
- FAO, Food Security Statistics, accessed 27 December 2009 at www.fao.org/economic/ess/food-security-statistics/en/.
- Golhani, K., Balasundram, S.K., Vadamalai, G., Pradhan, B. (2018). A review of neural networks in plant disease detection using hyperspectral data. *Inf. Process. Agric.* 5: 354-371.
- Hanson, A.M.G.J., Joel, M. G., Joy, A. and Francis, J. (2017). Plant leaf disease detection using deep learning and convolutional neural network. *International Journal of Engineering Science*. 5324: 2-4.
- Jianing, G., Zhiming, X., Rasheed, A., Tiancong, W., Zhao, Q.I.A. N., Zhang, Z.H.U.O. and Yuhong, G. (2022). CRISPR/Cas9 applications for improvement of soybeans, current scenarios and future perspectives. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca*. 50(2): 12678-12678.
- Karlekar, A., Seal, A. (2020). SoyNet: Soybean leaf diseases classification. *Comput. Electron. Agric.* 172: 105-342. <https://doi.org/10.1016/j.compag.2020.105342>.
- Krizhevsky, A., Sutskever, I. and Hinton, G. E. (2012). Image net classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems*. 1097-1105.
- Kulkarni, A., Chong, D. and Batarseh, F. A. (2020). Foundations of data imbalance and solutions for a data democracy. In *Data Democracy*. pp. 83-106. Academic Press.
- Le Cun, Y., Bottou, L., Bengio, Y. and Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE* 86. (11): 2278-2324.
- Lowe, D.G. (2004). Distinctive image features from scale in variant key points. *International Journal of Computer Vision*. 60(2): 91-110.
- Lu, J., Tan, L., Jiang, H. (2021). Review on convolutional neural network (CNN) applied to plant leaf disease classification. *Agriculture*. 11: 707.
- Meng, F. L. *et al.* (2022). Research progress on occurrence and management of soybean cyst nematode in Northeast China. *J. Northeast. Agric. Univ.*
- Mignoni, M.E. (2021). Images of soybean leaves. *Mendeley Data*, 1.
- Mohanty, S.P., Hughes, D.P. and Salathé, M. (2016). Using deep learning for image-based plant disease detection. *Frontiers in Plant Science*. 7: 1419.
- Min, P.K., Mito, K. and Kim, T.H. (2024). The evolving landscape of artificial intelligence applications in animal health. *Indian Journal of Animal Research*. <https://doi.org/10.18805/IJAR.BF-1742>.
- Pydipati, R., Burks, T. and Lee, W. (2005). Statistical and neural network classifiers for citrus disease detection using machine vision. *Transactions of the ASAE*. 48(5). doi: 10.13031/2013.19994.
- Panigrahi, K. P., Das, H., Sahoo, A. K. and Moharana, S.C. (2019). Maize leaf disease detection and classification using machine learning algorithms. In *Progress in Computing, Analytics and Networking: Proceedings of ICCAN*. pp 659-669.
- Pranata, Y.D., Wang, K.C., Wang, J.C., Idram, I., Lai, J.Y., Liu, J., Hsieh, I. (2019). Deep learning and SURF for automated classification and detection of calcaneus fractures in CT images. *Computer Methods and Programs in Biomedicine*. 171: 27-37. <https://doi.org/10.1016/j.cmpb.2019.02.006>.
- Sankaran, S. Mishra, A. Ehsani, R. and Davis, C. (2010). A review of advanced techniques for detecting plant diseases, *Computers and Electronics in Agriculture*. 72(1): 1-13.
- Tang, J., Yu, M., Wu, M. (2023). Detection of railway catenary insulator defects based on improved YOLOv5s. *Peer J. Computer Science*. 9: e14-74. <https://doi.org/10.7717/peerj-cs.1474>.
- Untari, D.T. and Satria, B. (2022). Repositioning Culinary "Betawi Ora" as Bekasi Eco-Culinary Tourism Icon. *International Journal of Environmental Sciences*. 8(2): 15-24. <https://www.theaspd.com/resources/IJES%20Paper02-%20Dhian%20Tyas.pdf>.
- Wihardjo, R.S.D., Muktiono, A. and Suwanda (2024). Exploring Critical Thinking in Environmental Education: APBL Approach to Reaction Rate Material at Public Senior High Schools in Indonesia. *International Journal of Environmental Sciences*. 10(1): 14-25. <https://www.theaspd.com/resources/2.%20Exploring%20Critical%20Thinking%20in%20Environmental%20Education.pdf>.