Optimization of Plant Disease Detection and Classification Using an Antlion-Optimized VGG16 Model with Fuzzy Rough C-Means Segmentation

T. Kalaiselvi[1]; M. Natarajan[2]; P. Kaviya[3*]

[1]Assistant Professor, Department of Computer and Information Science, Annamalai University, Chidambaram, Tamil Nadu, India - 608 002.

[2]Assistant Professor, Department of Agricultural Extension, Annamalai University, Chidambaram, Tamil Nadu, India - 608 002.

[*3]Research Scholar, Department of Agricultural Extension, Annamalai University, Chidambaram, Tamil Nadu, India - 608 002.

*Corresponding Author

Received:- 08 December 2025/ Revised:- 16 December 2025/ Accepted:- 21 December 2025/ Published: 31-12-2025

Copyright @ 2025 International Journal of Environmental and Agriculture Research
This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (https://creativecommons.org/licenses/by-nc/4.0) which permits unrestricted Non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

***Abstract*—** *The agricultural sector in India, supporting over 65% of the population, faces significant challenges from plant diseases that threaten crop productivity and food security. Traditional disease identification methods are often slow and require expert knowledge. This paper proposes a novel, automated framework for accurate plant disease detection by integrating advanced image processing with deep learning. The methodology employs a Median filter for image pre-processing, the Fuzzy Rough C-Means (FRCM) clustering algorithm for robust segmentation of diseased leaf regions, and a Convolutional Neural Network (CNN) for classification. The core innovation lies in enhancing a standard VGG16 CNN architecture using the Antlion Optimization (ALO) algorithm to optimize its hyperparameters, specifically the number of neurons in a fully connected layer, thereby improving feature learning and classification performance. Trained and tested on a dataset of cotton leaf images encompassing healthy samples and four disease types, the proposed ALO-enhanced VGG16 model achieved a high average classification accuracy of 93.33%. This performance surpasses that of other standard classifiers, including basic CNN, SVM, and ResNet models. The findings demonstrate that the integration of metaheuristic optimization with deep learning offers a powerful, scalable tool for precise plant disease diagnosis, with the potential to aid sustainable agricultural practices.*

***Keywords*— Plant Disease Detection, Image Segmentation, Fuzzy Rough C-Means (FRCM), Convolutional Neural Network (CNN), Antlion Optimization (ALO), VGG16.**

I. INTRODUCTION

Agriculture is a vital global industry and a primary source of livelihood. Plant diseases, caused by pathogens such as fungi, bacteria, and viruses, represent a major threat to crop yield and quality, leading to significant economic losses [1, 2]. Early and accurate detection is crucial for implementing effective management strategies. Conventional visual inspection by experts is subjective, time-consuming, and not scalable. Consequently, there is a growing need for automated, reliable, and rapid diagnostic systems.

Computer vision and artificial intelligence offer promising solutions. Initial approaches utilized basic image processing for colour and texture analysis [6], while machine learning models like Artificial Neural Networks (ANNs) and Support Vector Machines (SVMs) were applied for classification [7, 8]. The advent of deep learning, particularly Convolutional Neural Networks (CNNs), has revolutionized the field due to their superior ability to automatically learn hierarchical and discriminative features from raw image data [9]. Models like VGG16 have become benchmarks in image classification tasks.

Despite their success, standard CNNs may not be optimally configured for specific domains like plant pathology. Their fixed architectures might be suboptimal for learning the distinctive features of various leaf diseases. This creates a research

opportunity to customize and optimize these models. Furthermore, accurate segmentation of the diseased portion from the leaf background remains a challenge, especially under noisy or ambiguous conditions.

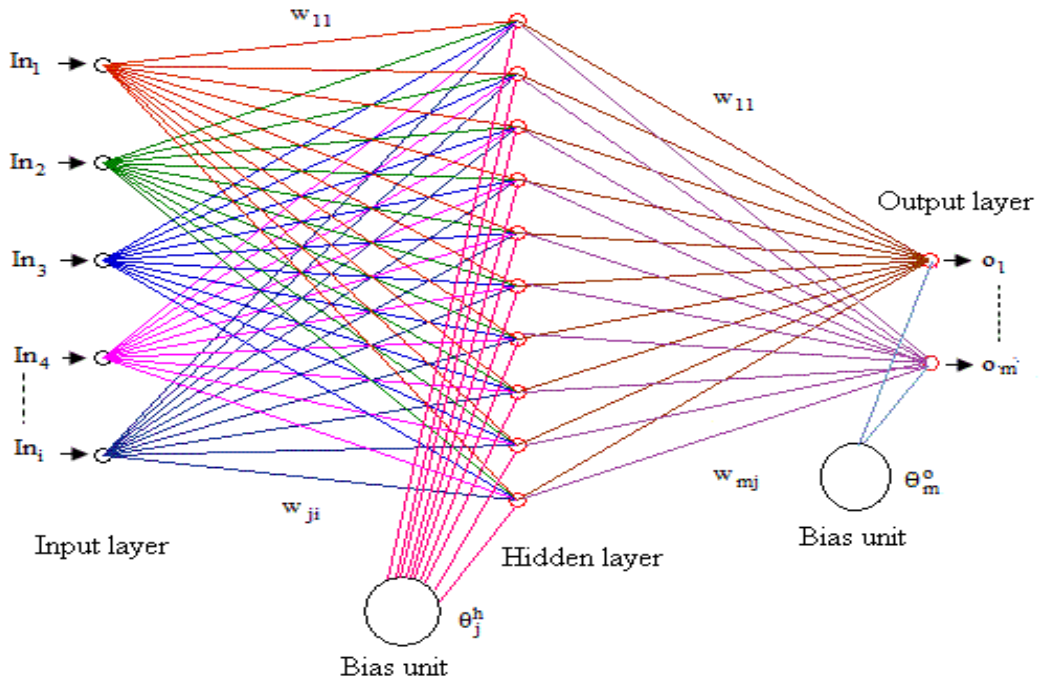

FIGURE 1: Generalized Neural Network Configuration

To address these challenges, this paper proposes a comprehensive framework that combines robust segmentation with an optimized deep learning model. The key contributions are:

1. Application of the Fuzzy Rough C-Means (FRCM) clustering algorithm for effective segmentation of diseased regions, leveraging its strength in handling uncertainty and image noise.
2. A novel hybrid classification model where the architecture of a VGG16 CNN is optimized using the Antlion Optimization (ALO) algorithm to enhance its performance for the specific task of cotton disease identification.
3. A comparative evaluation demonstrating that the proposed ALO-VGG16 model achieves superior accuracy compared to several existing methods.

The remainder of this paper is structured as follows: Section 2 details the proposed methodology, Section 3 presents the experimental results and discussion, and Section 4 concludes the work and suggests future directions.

II. PROPOSED METHODOLOGY

The proposed framework for automated plant disease detection comprises four sequential stages: Image Pre-processing, Segmentation, Feature Extraction, and Classification. The overall workflow is designed to enhance image quality, isolate the region of interest, and accurately classify the disease.

2.1 Image Acquisition and Pre-processing:

A dataset of cotton leaf images, including samples affected by Bacterial Blight, Anthracnose, Cercospora, and Alternaria, along with healthy leaves, was compiled for this study. Raw images captured in field conditions often contain noise from various sources. To improve subsequent analysis, a **Median filter** is applied in the pre-processing stage. This non-linear digital filtering technique is particularly effective at removing 'salt-and-pepper' noise while preserving the edges of the leaves and disease lesions, resulting in a cleaner image for segmentation [10, 11].

2.2 Image Segmentation using Fuzzy Rough C-Means (FRCM) Algorithm:

Accurate segmentation of the diseased lesion from the healthy leaf tissue is critical. This work employs the **Fuzzy Rough C-Means (FRCM)** clustering algorithm for this task [12, 13]. FRCM integrates the principles of fuzzy logic and rough set theory, offering advantages over conventional clustering like K-means.

- **Fuzzy Logic:** Allows pixels to belong to multiple clusters with varying degrees of membership, effectively handling the inherent ambiguity at lesion boundaries.
- **Rough Set Theory:** Models uncertainty by defining upper and lower approximations of clusters, making the algorithm more robust to noise and intensity variations within the lesion and background.

The algorithm initializes cluster centers and iteratively minimizes an objective function that factors in both the distance of pixels to cluster centers (fuzzy membership) and the roughness of the cluster approximations. Upon convergence, each pixel is assigned to the cluster with the highest membership value, producing a segmented image where the diseased region is distinctly isolated.

2.3 Feature Extraction:

Following segmentation, relevant features are extracted from the diseased region to facilitate classification. This process involves quantifying visual characteristics that distinguish one disease from another. While deep learning models automate this in later stages, the initial segmentation focuses on color, texture, and shape attributes. Color-based feature extraction, analyzing changes in color distribution (e.g., chlorosis, necrosis), is a fundamental method for identifying disease symptoms [14].

2.4 Classification using Antlion Optimization (ALO) Enhanced VGG16 CNN:

This stage constitutes the core innovation of the work, involving a hybrid model for disease classification.

2.4.1 Base Classifier: VGG16 Architecture:

VGG16 is a deep CNN model comprising 13 convolutional layers and 3 fully connected (FC) layers [Jaderberg et al., 2015]. Its deep, uniform architecture (using small 3x3 filters throughout) has proven highly effective for image recognition. We adopt VGG16 as our base classifier, modifying its final FC layer to have five neurons corresponding to our four disease classes and one healthy class.

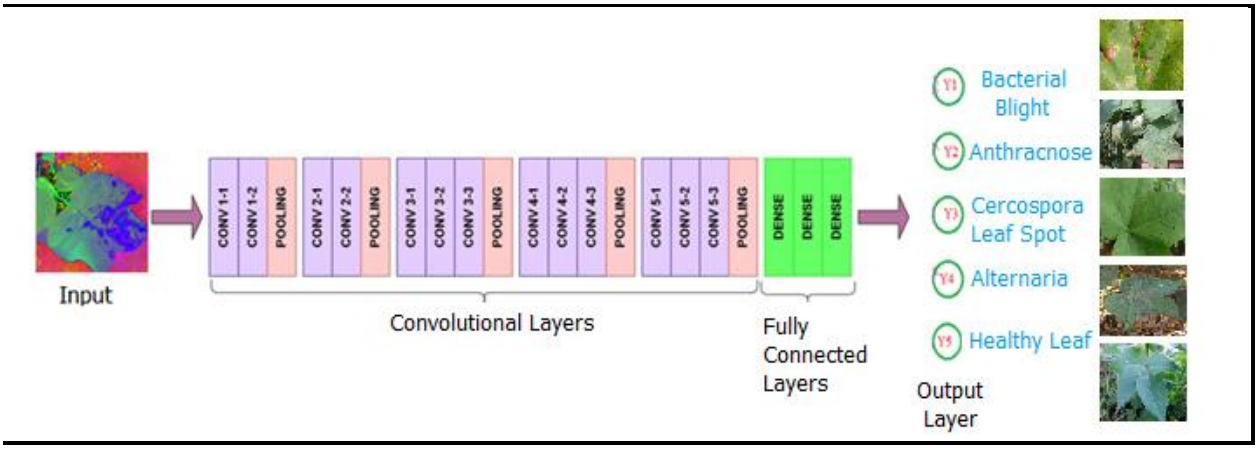

FIGURE 2: VGG16 Architecture for diseased Image Classification

2.4.2 Optimization with Antlion Algorithm (ALO):

The **Antlion Optimization (ALO)** algorithm is a nature-inspired metaheuristic that mimics the hunting mechanism of antlions in nature [15]. The process involves two main phases:

1. **Random Walk of Ants:** Represents the exploration of the search space (potential solutions).
2. **Building Traps by Antlions:** Represents the exploitation of promising areas, where better solutions (antlions) build pits to catch ants (other solutions).

In our work, ALO is employed to **optimize a key hyperparameter of the VGG16 model: the number of neurons in its first fully connected layer**. A population of candidate solutions (different neuron counts) is generated. Each candidate is evaluated by training a VGG16 model with that architecture on a subset of data and measuring its performance (e.g., validation accuracy), which serves as the fitness function. The ALO algorithm iteratively refines the population over generations, guiding the search toward the neuron count that yields the highest model fitness.

2.4.3 The ALO-VGG16 Hybrid Model:

The proposed hybrid model, termed **ALO-VGG16**, is created by integrating the optimization power of ALO into the VGG16 framework. The optimized architecture, determined by the ALO algorithm, is then fully trained on the entire training dataset. This approach tailors the deep learning model specifically to the characteristics of the plant disease image data, aiming to improve feature representation and final classification accuracy compared to the standard, non-optimized VGG16 model.

III. RESULTS AND DISCUSSION

The proposed ALO-VGG16 model was implemented and evaluated using a dataset of cotton leaf images. The model's performance was assessed using standard metrics derived from the confusion matrix: True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN).

Performance Metrics:

- **Precision** (Positive Predictive Value) = TP / (TP + FP)
- **Recall** (Sensitivity) = TP / (TP + FN)
- **F1-Score** = 2 * (Precision * Recall) / (Precision + Recall)
- **Specificity** = TN / (TN + FP)
- **Accuracy** = (TP + TN) / (TP + TN + FP + FN)

3.1 Performance of the ALO-VGG16 Model:

Table 1 presents the detailed performance of the ALO-VGG16 classifier across the five classes. The model achieved high and consistent metrics for all disease types and the healthy class. The average accuracy across all classes was **93.33%**, with individual class accuracies ranging from 93.06% to 93.52%. The high F1-scores (average 93.4%) indicate a strong balance between precision and recall.

TABLE 1
CLASSIFICATION PERFORMANCE METRICS FOR ALO-VGG16

S. No	Disease	Samples	Precision	Recall	F1-Score	Specificity	Accuracy
1	Bacterial Blight	397	92.5	93.6	92.8	93.7	93.45
2	Anthracnose	340	93.4	92.9	93.2	93.1	93.52
3	Cercospora	370	92.4	93.7	93.8	92.9	93.51
4	Alternaria	392	93.5	92.7	93.4	92.7	93.11
5	Healthy	562	93.7	93.5	93.7	93.1	93.06
Average			93	93.3	93.4	93	93.33

3.2 Comparative Analysis:

To validate the efficacy of the proposed optimization, the ALO-VGG16 model was compared against several standard classifiers. The results, summarized in Table 2 and visualized in Figure 3, clearly demonstrate its superior performance.

TABLE 2
ACCURACY COMPARISON OF DIFFERENT CLASSIFIERS

Classifier	Accuracy (%)
ALO-VGG16 (Proposed)	93.33
ResNet-CNN	92.06
ANN	92.5
SVM	91
CNN	91
KNN	82.5

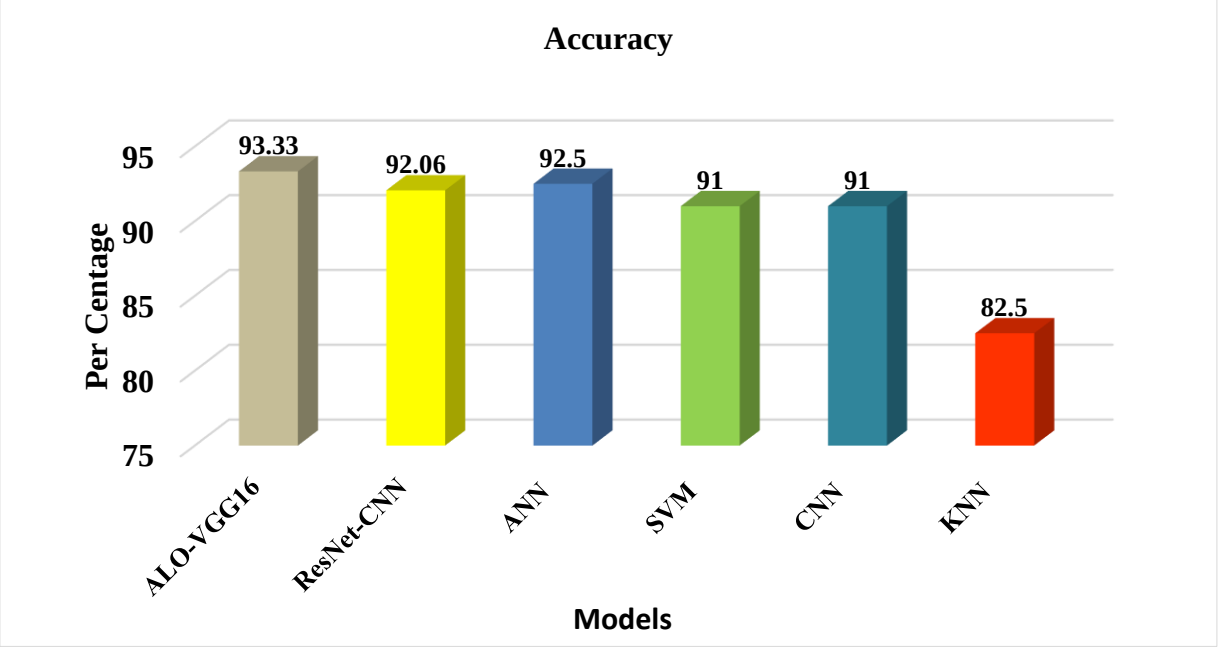

FIGURE 3: Comparative Accuracy of Various Classifiers (A bar chart would visually represent the data from Table 2).

3.3 Discussion:

The experimental results confirm the effectiveness of the proposed framework. The use of the FRCM algorithm for segmentation likely contributed to cleaner feature extraction by accurately isolating lesion boundaries despite noise. Most significantly, the integration of the Antlion Optimization algorithm with VGG16 provided a measurable boost in classification accuracy.

The ALO-VGG16 model outperformed the standard CNN (91.00% vs. 93.33%), highlighting the benefit of hyperparameter optimization. It also surpassed the more complex ResNet-CNN model (92.06%), suggesting that a well-optimized, simpler architecture can be more effective than a deeper, non-optimized one for this specific task. The superior performance over traditional machine learning models like SVM, KNN, and ANN underscores the advantage of deep learning in handling complex visual patterns in plant disease imagery.

The high accuracy and robust metrics across all classes indicate that the model is not only accurate but also reliable and generalizable within the tested dataset. The ALO algorithm successfully identified a more efficient configuration for the VGG16 network, enhancing its capability to discriminate between the subtle visual features of different cotton diseases.

IV. CONCLUSION AND FUTURE WORK

This research presented an optimized AI-based framework for the automated detection and classification of cotton leaf diseases. The methodology combined robust pre-processing using a Median filter, precise segmentation via the Fuzzy Rough C-Means (FRCM) algorithm, and a hybrid deep learning model. The key innovation was the enhancement of a VGG16 convolutional neural network using the Antlion Optimization (ALO) metaheuristic to optimize its architectural hyperparameter.

The proposed ALO-VGG16 model achieved an average classification accuracy of 93.33%, demonstrating superior performance compared to several benchmark models including standard CNN, ResNet, SVM, and ANN. This work validates that integrating nature-inspired optimization algorithms with deep learning architectures can significantly improve the precision of agricultural diagnostic systems.

For future work, the model can be tested on larger, more diverse datasets encompassing multiple crops and diseases captured under real-field conditions with complex backgrounds. Further optimization could explore tuning other hyperparameters (learning rate, filter sizes) using ALO or other metaheuristics. Finally, deploying the trained model as a user-friendly mobile application would be a practical step toward making this technology accessible to farmers for rapid, in-field disease diagnosis.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

[1] Karnaraja, D., & Natarajan, M. (2019). Assessing the knowledge level of ethno-agriculture practices in paddy among the tribal farmers of Kalrayan Hills. *Agriculture Update Journal*, 250–252.

[2] Kaviya, P., & Durairaj, S. (2023). A study on profile beneficiaries of IAMWARM farmers in Tiruchirappalli District of Tamil Nadu. *New Agriculture Era Magazine*, 2(1), 77–82.

[3] Sharma, O. P., & Bambawale, O. M. (2008). Integrated management of key diseases of cotton and rice. In *Integrated management of diseases caused by fungi, phytoplasma and bacteria* (pp. 271–302). New Delhi, India: Publisher.

[4] Wang, Y., Peng, S., Huang, J., Zhang, Y., Feng, L., Zhao, W., & Deng, N. (2022). Prospects for cotton self-sufficiency in China by closing yield gaps. *European Journal of Agronomy*, 133, Article 126437. https://doi.org/10.1016/j.eja.2021.126437

[5] Papageorgiou, E. I., Markinos, A. T., & Gemtos, T. A. (2011). Fuzzy cognitive map-based approach for predicting yield in cotton crop production as a basis for decision support system in precision agriculture application. *Applied Soft Computing*, 11(4), 3643–3657. https://doi.org/10.1016/j.asoc.2011.01.036

[6] Gulhane, V. A., & Gurjar, A. A. (2011). Detection of diseases on cotton leaves and its possible diagnosis. *International Journal of Image Processing*, 5(5), 590–598.

[7] Shah, N., & Jain, S. (2019). Detection of disease in cotton leaf using artificial neural network. In *Proceedings of the Amity International Conference on Artificial Intelligence (AICAI)*. Amity University.

[8] Chen, Y., et al. (2017). Corn yield prediction model using GA-BP neural network. *IOP Conference Series: Earth and Environmental Science*, 72(2), Article 022016. https://doi.org/10.1088/1755-1315/72/2/022016

[9] Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., & Farhan, L. (2021). Review of deep learning: Concepts, CNN architectures, challenges, applications and future directions. *Journal of Big Data*, 8, 53. https://doi.org/10.1186/s40537-021-00444-8

[10] Vidal, M., & Amigo, J. M. (2012). Pre-processing of hyperspectral images: Essential steps before image analysis. *Chemometrics and Intelligent Laboratory Systems*, 117, 138–148. https://doi.org/10.1016/j.chemolab.2012.05.009

[11] Shrestha, A., et al. (2017). Plant disease detection and classification by image processing. *Journal of Software Engineering and Applications*, 10(4), 223–232. https://doi.org/10.4236/jsea.2017.104013

[12] Sivaraman, E., & Arulselvi, S. (2009). Gustafson–Kessel (GK) clustering approach of TS fuzzy model for nonlinear processes. In *Proceedings of the Chinese Control and Decision Conference* (pp. 791–796). IEEE. https://doi.org/10.1109/CCDC.2009.5191864

[13] Kalaiselvi, T., & Narmatha, V. (2023). Cotton crop disease detection using FRCM segmentation and convolution neural network classifier. In *Computational vision and bio-inspired computing: Proceedings of ICCVBIC 2022* (pp. 557–577). Springer Nature. https://doi.org/10.1007/978-981-19-9472-7_47

[14] Abualigah, L., Shehab, M., Alshinwan, M., Mirjalili, S., & Elaziz, M. A. (2021). Ant lion optimizer: A comprehensive survey of its variants and applications. *Archives of Computational Methods in Engineering*, 28, 1397–1416. https://doi.org/10.1007/s11831-020-09432-5.