DOI: https://doi.org/10.36094/sc.v92.2026.Machine_Learning_Approach_for_Detection.Sinha.8

ARTICLE

Sci. and Cult. 92 (1–2) : 8-10 (2026)

# MACHINE LEARNING APPROACH FOR DETECTION OF TEA LEAF DISEASES USING CONVOLUTIONAL NEURAL NETWORK ARCHITECTURE

SUBRATA SINHA[1]*, SAURAV MALI[1] AND DONALD TANDIA[2]

*India is a major global tea exporter, but tea cultivation faces yield losses of up to 40% due to diseases. Accurate identification of these diseases is vital to reduce production loss. Traditional visual detection methods are ineffective, so this study introduces a Deep Convolutional Neural Network (CNN) designed to classify tea leaf diseases, including Gray Blight, Algal Spot, Brown Blight, Helopeltis, Healthy Leaves, and Red Spot. The model was trained on 5,867 images from Kaggle and achieved a high accuracy of 98.25%, outperformed VGG16*

## *Introduction*

India is the world's second-largest producer of tea, with production reaching 1,374.97 million kg in 2022-23. However, tea crops face significant losses due to diseases and pests—yield losses in some regions can reach up to 40%[1]. Traditional identification methods, such as visual inspection, are time-consuming, require expert knowledge, and are often ineffective[2]. Recent advances in AI, particularly Convolutional Neural Networks (CNNs), have shown promise for automatic, accurate disease detection in crops[3-4]. CNNs mimic the visual system of living beings and can efficiently process image data for tasks like classification and identification. This technology benefits both technical and non-technical users by enabling early detection through visual cues[5].

This study developed a CNN-based machine learning model, consisting of 6 convolutional and 6 pooling layers, to classify tea leaf diseases. The model's performance was compared with VGG16 to validate its accuracy.

Recent research shows rapid progress in tea leaf disease detection using deep learning, especially CNN and attention-based models. Sun et al. (2023) introduced *Tea Disease Net* with multiscale self-attention, outperforming YOLOv3 in detection accuracy and robustness[6]. Datta and Gupta (2023) developed a deep CNN model trained on a large dataset, achieving 96.56% accuracy[7]. Yucel and Yildirim (2023) proposed a hybrid framework using EfficientNetB0, ResNet101, and Shuffle Net, reporting 91.3% accuracy[8], while Rosyidah et al. (2023) demonstrated the effectiveness of CNNs combined with data augmentation under controlled conditions[9]. The latest studies focus on lightweight and attention-enhanced architectures. Yang et al. (2025) proposed *Wave Lite Net*, integrating discrete wavelet transform, MobileNetV3, and CBAM attention, achieving 98.70% accuracy. Suresh et al. (2025) developed a residual CNN with Global Average Pooling and reported 99% accuracy[10].

These studies highlight a clear shift toward highly accurate, lightweight, and attention-driven CNN models, while emphasizing the need for improved robustness and real-world generalization, which the proposed work aims to address.

## *Materials and Methods*

This study presents the methodology for developing a deep learning model for tea leaf disease classification. A dataset of 5,867 images, spanning five diseased classes

1 Department of Computational Sciences, Brainware University, Kolkata, India

2 Center for Biotechnology and Bioinformatics, Dibrugarh University, Assam, India

* Corresponding Author: dss.cs@brainwareuniversity.ac.in

and one healthy class, was compiled from natural environments in Tripura, India and augmented for diversity. Images were resized to 400×400 pixels and split into training (80%), validation (10%), and testing (10%) sets.

The proposed CNN architecture comprises 10 convolutional layers, 6 average pooling layers, 3 fully connected layers, and utilizes ReLU and Softmax activation functions. The model was trained for up to 100 epochs using the ADAM optimizer (learning rate 0.0005, batch size 32), with data augmentation enhancing generalizability.

**Table 2: Performance Metrics of Proposed Model**

| Class | Assessment Metrics | | | |
|---|---|---|---|---|
| | **Acc (%)** | **Pr (%)** | **Sn (%)** | **Sp (%)** |
| Algal Spot | 99.53 | 98.23 | 99.11 | 99.62 |
| Brown Blight | 97.97 | 100 | 87.5 | 100 |
| Gray Blight | 97.81 | 89.76 | 99.13 | 97.52 |
| Healthy | 98.28 | 100 | 89.52 | 100 |
| Helopeltis | 99.06 | 94.83 | 100 | 98.87 |
| Redspot | 98.91 | 93.94 | 98.94 | 98.9 |

Performance was evaluated on several metrics, including accuracy, precision, recall, F1-score, sensitivity, specificity, NPV, MCC, and mean average precision (mAP).

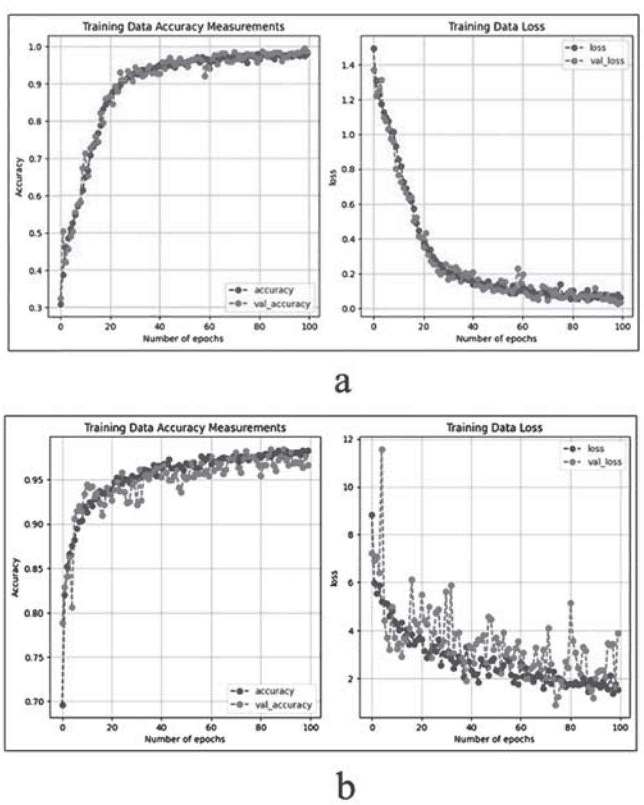


**Fig. 1:** Accuracy and Loss Curves (a) Proposed Model (b) VGG16

**Table3: Performance Metrics of VGG16**

| Class | Assessment Metrics | | | |
|---|---|---|---|---|
| | **Acc (%)** | **Pr (%)** | **Sn (%)** | **Sp (%)** |
| Algal Spot | 99.69 | 100 | 98.18 | 100 |
| Brown Blight | 99.69 | 97.94 | 100 | 99.63 |
| Gray Blight | 99.69 | 99.12 | 99.12 | 99.81 |
| Healthy | 99.38 | 98.11 | 98.11 | 99.63 |
| Helopeltis | 99.53 | 99.08 | 98.18 | 99.81 |
| Redspot | 99.84 | 99.07 | 100 | 99.81 |

Comparative analyses with pretrained models VGG16 were conducted. Experiments were performed on Google Colab with GPU acceleration, using Python, Tensorflow, and Keras to ensure replicability. The methodology emphasizes robust training, rigorous evaluation, and fair comparison with benchmark models.

## Results

***Classification Curves of the Proposed Model Along with VGG16:*** The experimental results demonstrate that the proposed CNN model outperforms the VGG16 architecture in terms of classification accuracy, stability, and overall predictive reliability for tea leaf disease detection. When trained on the same dataset under identical conditions (learning rate = 0.0005, Adam optimizer), the proposed model achieved an overall accuracy of 98.25%, whereas VGG16 achieved 98.05%. Furthermore, the confusion matrix analysis revealed that the proposed model generated fewer false positives and false negatives compared to VGG16, particularly for challenging disease classes such as Brown Blight and Gray Blight.

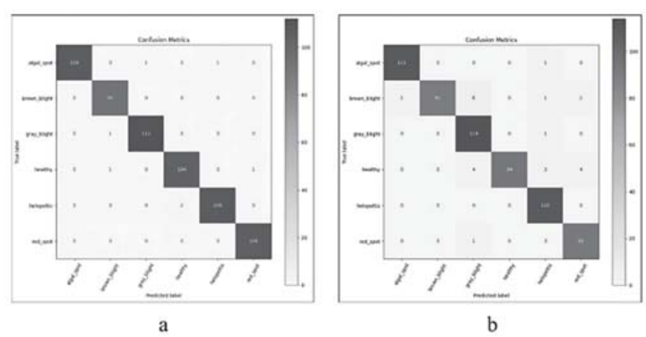


**Fig. 2:** Confusion Matrix (a) Proposed Model (b) VGG16

## Conclusion

This work demonstrates that a Convolutional Neural Network (CNN) can accurately and efficiently identify tea leaf diseases, achieving a classification accuracy of 98.25% and surpassing established models such as VGG16. The proposed model automates disease classification into five categories and offers significant advantages over traditional methods, including speed, scalability, and a non-invasive approach. While the model's performance is promising, further work is needed to address challenges such as dataset quality and validation across diverse environments. Integration of this model into real-time monitoring systems holds potential for early disease detection, reduced chemical usage, and more sustainable agricultural practices. Continued research and refinement of CNN approaches are expected to enhance disease management and support food system sustainability. ❐

## References


1. Tea Board of India, IBEF. Tea industry and exports in India. Indian Tea: Industry, Market & Export Data. 2023. Available from: https://www.ibef.org/exports/indian-tea-industry
2. G.A. Secor, V.V. Rivera, J.A. Abad, I.M. Lee, G.R.G. Clover, L.W. Liefting, et al., *Plant Dis.*, **93(6)**: 574–83, (2009).
3. J. Wu., Introduction to Convolutional Neural Networks, (National Key Lab for Novel Software Technology, Nanjing University, China; 2017).
4. N. Krishnamoorthy, L.V. Narasimha Prasad, C.S. Pavan Kumar, B. Subedi, H. Baraki Abraha, V.E. Sathishkumar, *Environ Res.*, 2021;198:111275. doi:10.1016/j.envres.2021.111275
5. L. Ahmad, F. Nabi, Agriculture 5.0: Artificial Intelligence, IoT and Machine Learning (CRC Press; 2021).
6. Y. Sun, F. Wu, H. Guo, R. Li, J. Yao, J. Shen, *Front Plant Sci.*, **14**:1257212 2023. doi:10.3389/fpls.2023.1257212
7. S. Datta, N. Gupta, *Procedia Comput Sci.*, **218**: 2273–86, (2023).
8. N. Yucel, M. Yildirim, *Int J Appl Math Electron Comput.*, **11(1)**: 30–6, (2023).
9. E.A. Rosyidah, A. F. Hadi, Y.S. Dewi, The Classification of Tea Leaf Disease Using CNN Image Classifier. In: 1st International Conference on Neural Networks and Machine Learning 2022 (ICONNSMAL 2022) (Atlantis Press; 2023). p. 89–112.
10. X. Yang, Y. Zhang, Z. Liu, H. Chen, *PLoS One*, **20(2)**:e0323322 (2025). doi:10.1371/journal.pone.0323322