IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940



International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 11, November 2026

DOI: 10.17148/IJARCCE.2026.1511192

A Review of Deep Learning Approaches for Sugarcane Leaf Disease Detection

Dr Manjusha K K[1], Ann Maria Jose[2]

Assistant Professor, Department of Computer Application, Mar Dionysius College, Pazhanji, Thrissur, India[1]

Assistant Professor, Department of Computer Application, Mar Dionysius College, Pazhanji, Thrissur, India[2]

Abstract: Sugarcane is one of the economically significant crops and diseases in sugarcane leaves might cause substantial damage in terms of crop yields and quality. Visual disease identification is the traditional way of identifying diseases in agricultural practice, but it might be labour-intensive, subjective and challenging to distinguish especially in case of similar symptoms at an early stage of development. Advances in the field of artificial intelligence, computer vision and deep learning allowed automating this process with the help of image-based methods. This paper offers a review of the Convolutional Neural Network (CNN), transfer learning, Vision Transformer (ViT), CNN-Transformer hybrid, Transformer ensemble and hyperspectral imaging techniques used for the purpose of automatic sugarcane disease detection. While CNN algorithms are good at capturing local spatial patterns, the Vision Transformer can identify long-range relationships within the image. The hybrid architectures try to combine these features. The latest papers have offered new developments in this field through the use of the ensemble of MobileNetV2 architectures, adaptive CNN-Transformer fusion, ViT-Swin Transformer ensemble and the uncertainty-aware semi-supervised learning. Some of the reported results are the following: 97.43% accuracy of the CNN-ViT hybrid model, 97.06% test accuracy of the ViT-Swin Transformer ensemble, 96.2% accuracy of the MobileNetV2-based soft-voting ensemble, and 96.30% F1-score of the EnhancedConvFormer architecture. It is worth mentioning that the hyperspectral imaging approach might be useful in case of pre-symptomatic disease detection. As the analysed works used various datasets and disease classes for evaluation, their results cannot be compared directly. The review helps to define current research directions and gaps and the importance of such aspects as the standardized dataset usage, field validation, lightweight architecture design, explain ability and pre-symptomatic disease detection.

Keywords: Sugarcane leaf disease, deep learning, convolutional neural network, Vision Transformer, CNN–ViT, Swin Transformer, hyperspectral imaging, transfer learning, semi-supervised learning.

I. INTRODUCTION

The Saccharum officinarum, which is known as sugarcane, is one of the widely grown commercial crops all over the world. This crop plays an important role as a sugar, ethanol and bioenergy resource. However, sugarcane cultivation is influenced by various diseases including red rot, smut, mosaic, rust, and leaf scald. The disease infection adversely impacts plant health, yield and quality of harvested sugarcane.

Conventionally, the diseases diagnosis in plants was carried out using visual inspection by farmers or specialists in agriculture. Despite the fact that specialist with sufficient experience may identify a great variety of diseases based on their symptoms, the manual diagnosis is rather difficult in case of vast agricultural fields and at the stage, when symptoms are still weak and similar to plant age effects, irrigation, illumination and other factors. Convolutional Neural Network (CNN) architecture enables to learn disease-related characteristics directly from the leaf images. Upadhye et al. studied CNN-based sugarcane disease detection and showed the applicability of deep learning for automated sugarcane disease diagnosis [1].

Later Srinivasan et al. examined various architectures of deep neural networks for the classification of sugarcane leaf diseases using 6,748 images belonging to 11 classes [2]. They tested EfficientNet, DenseNet201, ResNetV2, InceptionV4, MobileNetV3, and RegNetX architectures. These investigations show the relevance of CNN architectures in agricultural image classification.

However, CNNs mainly learn local spatial features by means of convolution operations. Disease symptoms may appear in different parts of the leaf and hence the relationship between distant image regions is essential. On the contrary, the Vision Transformers (ViTs) enable to analyse images by splitting it into patches and learning the relations between them by self-attention mechanism.

 This work is licensed under a Creative Commons Attribution 4.0 International License

DOI: 10.17148/IJARCCE.2026.1511192

Paçal and Kunduracıoğlu researched the data-efficient vision transformers models for the classification of sugarcane disease and reported 93.79% accuracy for the DeiT3-Small architecture on sugarcane dataset with 6748 images and 11 classes [3].

Hybrid CNN and Transformer models offer complementary feature representations. Aswani and Sinha proposed CNN-ViT hybrid model for the classification of five sugarcane leaf categories. They reported 84.3%, 93.07%, and 97.43% accuracy for CNN, ViT and the hybrid CNN-ViT mdoels respectively [4].

Kumar et al. suggested a lightweight MobileNetV2-based solution combined with classical machine learning classifiers and soft-voting ensemble learning. The solution reached 96.2% accuracy, 96.0% Macro-F1, and 98.0% AUC-ROC on the held-out test set and was transformed to TensorFlow Lite for potentially edge deployment [5].

An adaptive hybrid CNN and Vision Transformer architecture called EnhancedConvFormer that combines the CNN and Vision Transformer features via Adaptive Dynamic Token Mixer. The architecture reached 96.30% F1-score with approximately 3.96 GFLOPs, addressing both classification performance and computational efficiency [6].

Saritha and Rasane combining Vision Transformer and Swin Transformer models. The combination reached 97.06% test accuracy, compared to 96.13% for ViT and 96.56% for Swin Transformer separately [7]. This study also included an LLM-based post-prediction explanation module. U-SSDNet, an uncertainty-aware semi-supervised framework, based on Mean-Teacher strategy and supervised contrastive learning reducing dependence on large labelled datasets.[8]. Also evaluated various classical and modern CNN architectures including LeNet-5, AlexNet, VGG16, VGG19, GoogLeNet/InceptionV3, ResNet101, MobileNetV2, DenseNet121, and NASNet Mobile. DenseNet121 reached 86% accuracy in that investigation [9].
An important approach is hyperspectral imaging that enables early disease detection in plants. Contrary to conventional RGB images, hyperspectral imaging systems acquire information about a broad spectrum of wavelengths. Bao et al. investigated the hyperspectral imaging and deep learning for sugarcane disease detection [10]. In another paper hyperspectral imaging with spectral-spatial attention deep neural networks was used for the detection of sugarcane smut and mosaic disease and more than 90% detection prior to appearance of visible symptoms was reported[11]. Thus, the objective of this paper is to provide a comparative review of the development of sugarcane leaf disease detection methods from conventional CNNs to modern Transformer, hybrid, ensemble, semi-supervised and hyperspectral approaches.

II. METHODOLOGY

TABLE II 1 - COMPARISON OF METHODS

Method	Main Feature	Major Strength	Major Limitation
CNN	Local spatial features	Effective and established	Limited long-range context
VGG16	Deep transfer learning	Simple and widely studied	Computationally heavier
MobileNetV2	Lightweight CNN	Suitable for mobile/edge devices	Limited contextual modelling
EfficientNet	Efficient feature extraction	Accuracy-efficiency balance	Model selection required
Vision Transformer	Global attention	Captures long-range relationships	Computational/data requirements
CNN–ViT	Local + global features	Complementary representations	Increased complexity
ViT + Swin	Global + hierarchical attention	Multi-scale representation	Ensemble complexity
EnhancedConvFormer	Adaptive CNN–Transformer fusion	Lightweight hybrid approach	More complex architecture
Semi-supervised U-SSDNet	Labelled + unlabelled data	Reduces annotation dependence	Training complexity

 This work is licensed under a Creative Commons Attribution 4.0 International License

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940



International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 11, November 2026

DOI: 10.17148/IJARCCE.2026.1511192

Hyperspectral + DL	Spectral + spatial information	Potential pre-symptomatic detection	Expensive data acquisition

III. RESULTS

TABLE III 1 THE FOLLOWING TABLE COMBINES REPRESENTATIVE STUDIES FROM THE LITERATURE.

Ref.	Year	Approach	Dataset/Task	Reported Result
[1]	2023	CNN	Sugarcane disease detection	CNN-based classification
[2]	2025	Multiple CNNs	6,748 images, 11 classes	Comparative evaluation
[3]	2024	DeiT3-Small	6,748 images, 11 classes	93.79% accuracy
[4]	2025	CNN	5 classes	84.30% accuracy
[4]	2025	ViT	5 classes	93.07% accuracy
[4]	2025	CNN + ViT	5 classes	97.43% accuracy
[11]	2024	HSI + spectral-spatial attention DNN	Smut and mosaic	>90% before visible symptoms
[5]	2026	MobileNetV2	4 classes	94.80% accuracy
[5]	2026	MobileNetV2 + ensemble	4 classes	96.20% accuracy
[6]	2026	EnhancedConvFormer	Sugarcane diseases	96.30% F1
[7]	2026	ViT	Sugarcane diseases	96.13% accuracy
[7]	2026	Swin Transformer	Sugarcane diseases	96.56% accuracy
[7]	2026	ViT + Swin	Sugarcane diseases	97.06% accuracy
[9]	2026	DenseNet121	Sugarcane disease classification	86% accuracy

IV. DISCUSSION

This literature is a clear progression towards integration of both local and global information in CNNs and beyond. Despite the rise in other architectures, CNNs continue to be used due to the presence of local information in disease symptoms such as lesions, spots and colour variations. Transfer-learning algorithms make them even more applicable if the agricultural dataset is not very big.

Vision Transformers allow one to model the relationships between patches of image which have large ranges. The 93.79% accuracy of DeiT3-Small confirms the ability of efficient transformers to process data [3].

The hybrid of CNN-ViT is another step in the same direction, as it integrates local and global features. Aswani and Sinha reported 97.43% accuracy on their five-class dataset [4].
The 2026 EnhancedConvFormer makes progress in this field through adaptive fusion of CNN and Transformer [6]. The 96.30% F1-score and approximately 3.96 GFLOPs indicate that hybrid architectures become computationally efficient.

The ViT–Swin ensemble introduced by Saritha and Rasane is another step towards hybrid architectures in 2026. It has 97.06% test accuracy, whereas individual ViT and Swin models 96.13% and 96.56% respectively [7].

Deployment of the developed model is also important for practical applications. The mobile model based on MobileNetV2 by Kumar et al. achieves 96.2% accuracy and is implemented in TensorFlow Lite for easier deployment [5].

Semi-supervised learning indicates that in the future researchers will not only use complex models but will pay more attention to data availability and quality. For example, the U-SSDNet addresses the problem of limited amount of labelled agricultural images[8].
Finally, hyperspectral imaging introduces another important distinction between disease classification and disease early detection. If the change in spectra can be seen before visible symptoms, then hyperspectral systems will warn of diseases earlier than RGB image processing [10], [11].

The development of sugarcane disease detection has progressed from traditional visual inspection to CNN-based classification and transfer learning.

 This work is licensed under a Creative Commons Attribution 4.0 International License

DOI: 10.17148/IJARCCE.2026.1511192

This was followed by the adoption of modern CNN architectures and Vision Transformers (ViTs) for improved disease recognition. Recent studies have increasingly explored CNN–ViT hybrid models, ViT–Swin Transformer ensembles, and adaptive CNN-Transformer architectures.

More recent approaches also focus on semi-supervised and uncertainty-aware learning to handle limited labelled data and improve reliability.
In parallel, hyperspectral and multimodal imaging are being investigated for early symptom detection, while computational efficiency, field deployment, and explainability have become important research considerations.

V. CONCLUSION

This paper reviewed the development of artificial-intelligence-based sugarcane leaf disease detection, beginning with CNN-based approaches and progressing toward transfer learning, Vision Transformers, CNN–Transformer hybrids, Transformer ensembles, lightweight deployment models, semi-supervised learning and hyperspectral imaging.

CNN models provide effective local feature extraction and remain an important foundation for automated disease diagnosis. Transfer learning makes these architectures useful when agricultural datasets are relatively small. Vision Transformers extend the modelling capability by capturing long-range relationships between image regions.

Hybrid CNN–ViT models combine these complementary characteristics. The reviewed Aswani and Sinha study reported 84.30% accuracy for CNN, 93.07% for ViT and 97.43% for the CNN–ViT hybrid [4].

MobileNetV2 combined with soft-voting ensemble learning achieved 96.2% accuracy and provided an edge-deployment pathway [5]. EnhancedConvFormer achieved a reported F1-score of 96.30% while emphasizing computational efficiency [6]. The ViT–Swin ensemble achieved 97.06% test accuracy and introduced an LLM-based explanation component [7]. U-SSDNet addressed the challenge of limited labelled data through uncertainty-aware semi-supervised learning [8]. Hyperspectral imaging represents another important research direction because it can potentially identify physiological changes before visible symptoms appear [10], [11].

However, reported accuracy values from separate studies should not be interpreted as a direct ranking because the datasets, disease classes, image conditions and evaluation protocols differ. Future work should therefore conduct standardized experiments in which CNN, VGG16, MobileNetV2, EfficientNet, ViT, CNN–ViT and recent hybrid models are trained and evaluated using the same dataset and identical evaluation procedures.

Future research should also emphasize field validation, lightweight deployment, explainability, semi-supervised learning and multimodal or hyperspectral information. These directions can contribute to the development of reliable, efficient and practical systems for early sugarcane disease detection.

REFERENCES

[1]. S. A. Upadhye, M. R. Dhanvijay, and S. M. Patil, “Sugarcane Disease Detection Using CNN-Deep Learning Method: An Indian Perspective,” Universal Journal of Agricultural Research, vol. 11, no. 1, pp. 80–97, 2023.
[2]. S. Srinivasan, S. M. Prabin, S. K. Mathivanan, H. Rajadurai, S. Kulandaivelu, and M. A. Shah, “Sugarcane leaf disease classification using deep neural network approach,” BMC Plant Biology, vol. 25, 2025, doi: 10.1186/s12870-025-06289-0.
[3]. İ. Paçal and İ. Kunduracıoğlu, “Data-Efficient Vision Transformer Models for Robust Classification of Sugarcane,” Journal of Soft Computing and Decision Analytics, vol. 2, no. 1, pp. 258–271, 2024.
[4]. S. Aswani and A. Sinha, “Early diagnosis of Sugarcane leaf diseases through CNN and Vision Transformer hybrid model,” Research Square, 2025, doi: 10.21203/rs.3.rs-7232200/v1.
[5]. R. Kumar, P. K. Singh, and R. K. Tiwari, “Leaf-level sugarcane disease detection using CNN models: enhancing agricultural intelligence,” Discover Artificial Intelligence, vol. 6, art. no. 257, 2026, doi: 10.1007/s44163-026-00967-0.
[6]. “An adaptive hybrid vision architecture for sugarcane leaf disease identification,” Information Processing in Agriculture, 2026, doi: 10.1016/j.inpa.2026.07.005.
[7]. M. Saritha and K. Rasane, “An ensemble of vision and swin transformers with LLM-based explanations for sugarcane leaf disease diagnosis,” Scientific Reports, vol. 16, art. no. 10707, 2026, doi: 10.1038/s41598-026-45453-w.

 This work is licensed under a Creative Commons Attribution 4.0 International License



International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 11, November 2026
DOI: 10.17148/IJARCCE.2026.1511192

[8]. "Uncertainty-aware semi-supervised framework with supervised contrastive learning for sugarcane leaf disease detection," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 239, pp. 259–275, 2026, doi: 10.1016/j.isprsjprs.2026.06.005.

[9]. "CNNs in Crop Diagnostics: A Performance Comparison of Classic and Modern Architectures for Sugarcane Leaf Disease Detection," Procedia Computer Science, vol. 283, pp. 4900–4909, 2026, doi: 10.1016/j.procs.2026.06.544.

[10]. D. Bao, J. Zhou, S. A. Bhuiyan, R. Ford, and Y. Gao, "Early detection of sugarcane diseases through hyperspectral imaging and deep learning," Proceedings of the Australian Society of Sugar Cane Technologists, vol. 44, pp. 176–183, 2023.

[11]. D. Bao, J. Zhou, S. A. Bhuiyan, P. Adhikari, G. Tuxworth, R. Ford, and Y. Gao, "Early detection of sugarcane smut and mosaic diseases via hyperspectral imaging and spectral-spatial attention deep neural networks," Journal of Agriculture and Food Research, vol. 18, art. no. 101369, 2024, doi: 10.1016/j.jafr.2024.101369.