International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 11, November 2026)

Sugarcane Disease Detection Using Deep Learning Techniques

Kumar K[1], Dr. Sunil T D [2]

[1]Research Scholar, Sri Siddhartha Academy of Higher Education, Tumakuru, Karnataka, India
[2]Professor, Sri Siddhartha Institute of Technology, Sri Siddhartha Academy of Higher Education, Tumakuru, Karnataka, India

Abstract*— Sugarcane (*Saccharin officinarum*) is a critical cash crop globally, contributing significantly to sugar production and biofuel industries. However, diseases such as red rot, rust, mosaic, yellow leaf, and smut pose substantial threats to crop yield and quality, causing economic losses estimated in billions of dollars annually. Traditional manual inspection methods for disease detection are labour-intensive, time-consuming, subjective, and prone to human error, making them impractical for large-scale agricultural operations. Recent advances in deep learning and computer vision have revolutionized plant disease detection, offering automated, accurate, and scalable solutions. This comprehensive manuscript reviews state-of-the-art methodologies for sugarcane disease detection, with particular emphasis on fusion approaches that integrate both stem and leaf features for enhanced diagnostic accuracy. The review synthesizes findings from multiple research studies, analysing deep learning architectures (CNN, VGG, ResNet, EfficientNet, DenseNet, MobileNet, YOLO), performance metrics (accuracies ranging from 85% to 99.8%), dataset characteristics, pre-processing techniques, and multimodal data integration strategies. Key findings reveal that while most current research focuses predominantly on leaf-based analysis, emerging multimodal fusion approaches combining stem-leaf features, hyper spectral imaging, thermal data, and environmental parameters demonstrate superior performance. The manuscript identifies critical research gaps including limited large-scale annotated datasets, insufficient exploration of hybrid architectures, underutilized severity classification, and computational constraints for field deployment. Future directions emphasize the need for comprehensive stem-leaf fusion methodologies, lightweight models for edge computing, explainable AI integration, and region-specific adaptations to advance precision agriculture in sugarcane cultivation.

***Keywords*-- CNN, Mobilenet, DenseNet.**

I. INTRODUCTION

Sugarcane (*Saccharum officinarum*) is one of the world's most important agricultural commodities, serving as the primary source of sugar production and an increasingly vital feedstock for biofuel generation. Global sugarcane cultivation spans over 26 million hectares, with major production concentrated in Brazil, India, China, Thailand, and Pakistan. However, the crop faces persistent threats from various diseases that significantly impact yield, quality, and economic returns.

Major sugarcane diseases include red rot (*Colletotrichum falcatum*), rust (*Puccinia melanocephala*), mosaic virus, yellow leaf disease, smut (*Sporisorium scitamineum*), and wilt, each causing distinct symptoms and requiring specific management strategies.

The economic impact of sugarcane diseases is substantial, with annual losses estimated at billions of dollars globally. Redrot alone can cause yield losses of 30-50% in susceptible varieties, while rust infections can reduce sugar content by 15-20%. Early and accurate disease detection is therefore critical for implementing timely interventions, minimizing crop losses, and ensuring sustainable production. Traditional disease detection methods rely on manual visual inspection by trained agronomists, a process that is inherently subjective, labour-intensive, time-consuming, and impractical for large-scale operations [1]. Moreover, expert decision-making can be influenced by physical fatigue, environmental conditions, and the subtle nature of early-stage disease symptoms, leading to delayed diagnosis and treatment.

The advent of deep learning and computer vision technologies has opened new avenues for automated plant disease detection. Convolutional Neural Networks (CNNs) and their variants have demonstrated remarkable success in image classification tasks, achieving human-level or superior performance in various agricultural applications [2], [3]. These technologies offer the potential for rapid, objective, and scalable disease detection systems that can be deployed across diverse agricultural settings. Recent research has explored various deep learning architectures including VGG [4], ResNet[5], EfficientNet [6], DenseNet [7], and MobileNet [8] for sugarcane disease detection, achieving accuracies ranging from85% to over 99%.

While most existing research focuses predominantly on leaf-based disease detection, emerging evidence suggests that multimodal approaches integrating both stem and leaf features can provide more comprehensive diagnostic capabilities [9],[10]. Certain diseases manifest distinct symptoms on stems (e.g., red rot internal discoloration, borer damage) that may not be apparent from leaf analysis alone. Furthermore, advanced fusion methodologies incorporating hyper spectral imaging, thermal data, and environmental parameters have shown promise in enhancing detection accuracy and enabling early disease prediction [11].

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 11, November 2026)

This manuscript provides a comprehensive review of state-of-the-art deep learning methodologies for sugarcane disease detection, with particular emphasis on fusion approaches that integrate multiple data sources and plant organs. The review synthesizes findings from recent research, analyses performance metrics across different architectures, examines dataset characteristics and pre-processing techniques, and identifies critical research gaps that must be addressed to advance the field toward practical, field-deployable solutions.

II. Background And Theoretical Foundations

A. Sugarcane Diseases and Economic Impact

Sugarcane cultivation faces numerous biotic stresses, with fungal, bacterial, and viral diseases representing the most significant threats to productivity. The major diseases affecting sugarcane worldwide include:

Red Rot (*Colletotrichum falcatum*): One of the most destructive fungal diseases, characterized by reddening of internal stem tissue, white patches with black dots on leaves, and eventual wilting. It can cause yield losses of 30- 50% in susceptible varieties and significantly reduces sugar recovery.

Rust (*Puccinia melanocephala*): A foliar disease causing elongated yellow to brown pustules on leaf surfaces, leading to premature leaf senescence and reduced photosynthetic capacity. Severe infections can reduce cane yield by 20-30% and sugar content by 15-20%.

Mosaic Virus: Transmitted by aphids, this viral disease causes characteristic mosaic patterns of light and dark green areas on leaves, stunted growth, and reduced tailoring. Yield losses can reach 20-50% depending on variety susceptibility and infection timing.

Yellow Leaf Disease (YLD): Caused by *Sugarcane yellow leaf virus* (SCYLV), this disease manifests as yellowing of the midrib region on lower leaves, progressing upward. It affects sugar accumulation and can reduce yields by 10-30%.

Smut (*Sporisorium scitamineum*): Characterized by a whip-like structure emerging from the growing point, covered with black teliospores. Infected plants show reduced tillering and stunted growth, with potential yield losses of 30-50%.

The economic impact of these diseases extends beyond direct yield losses to include increased production costs for disease management, reduced sugar quality, and constraints on variety selection.

In major producing countries, the cumulative economic burden of sugarcane diseases is estimated at several billion dollars annually, underscoring the critical need for effective detection and management strategies.

B. Limitations of Traditional Detection Methods

Traditional disease detection in sugarcane relies primarily on visual inspection by trained agronomists or extension workers. This approach involves periodic field surveys where experts examine plants for characteristic disease symptoms such as leaf discoloration, lesions, growth abnormalities, and stem internal symptoms (requiring destructive sampling). While this method has been the standard practice for decades, it suffers from several fundamental limitations:

Subjectivity and Inconsistency: Disease diagnosis based on visual symptoms is inherently subjective, with inter-observer variability in symptom interpretation. Early-stage infections often present subtle symptoms that may be overlooked or misidentified, particularly when multiple diseases exhibit similar visual characteristics [1].

Labour Intensity and Scalability: Manual inspection is extremely labour-intensive, requiring significant time and human resources. For large commercial plantations spanning thousands of hectares, comprehensive field surveys are impractical and economically unfeasible. This limitation is particularly acute in developing countries where agricultural labour shortages are increasingly common.

Expertise Requirements: Accurate disease diagnosis requires specialized training and extensive field experience. The shortage of trained plant pathologists and agronomists, especially in remote agricultural regions, limits the accessibility and reliability of traditional detection methods [12].

Delayed Detection: Manual surveys are typically conducted at fixed intervals (weekly or biweekly), meaning that disease outbreaks may go undetected for extended periods. By the time symptoms become visually apparent and are identified during routine inspections, the disease may have already spread significantly, reducing the effectiveness of control measures.

Environmental and Human Factors: The accuracy of visual inspection can be compromised by various factors including lighting conditions, weather, time of day, and the physical and mental state of the inspector. Fatigue, poor vision, and work pressure can all contribute to diagnostic errors [1].

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 11, November 2026)

These limitations highlight the urgent need for automated, objective, and scalable disease detection systems that can provide early, accurate, and consistent diagnoses across diverse agricultural settings.

2.1 Deep Learning in Plant Disease Detection

Deep learning, a subset of machine learning based on artificial neural networks with multiple layers, has revolutionized computer vision and image recognition tasks over the past decade. Convolutional Neural Networks (CNNs), in particular, have demonstrated exceptional performance in extracting hierarchical features from images and achieving state-of-the-art results in classification, detection, and segmentation tasks [2].

The application of deep learning to plant disease detection has gained significant momentum, driven by several key advantages.

Automated Feature Extraction: Unlike traditional machine learning approaches that require manual feature engineering, CNNs automatically learn hierarchical representations from raw image data. Lower layers capture basic features such as edges and textures, while deeper layers learn complex patterns specific to disease symptoms [3].

High Accuracy and Generalization: Modern CNN architectures, when trained on sufficiently large and diverse datasets, can achieve classification accuracies exceeding 95%, often surpassing human expert performance. Transfer learning techniques, which leverage pre-trained models on large-scale datasets like ImageNet, further enhance performance and reduce training data requirements [13].

Scalability and Speed: Once trained, deep learning models can process images in real-time, enabling rapid disease screening across large agricultural areas. This scalability is particularly valuable when integrated with mobile devices, drones, or automated imaging systems [14].

Multimodal Integration: Deep learning frameworks can seamlessly integrate multiple data modalities including RGBimages, hyperspectral data, thermal imaging, and environmental parameters, enabling more comprehensive disease assessment [11].

Several CNN architectures have been successfully applied to plant disease detection, each with distinct characteristics:

VGG Networks [4]: Characterized by deep architectures with small (3×3) convolutional filters, VGG models excel at capturing fine-grained texture and colour patterns essential for disease symptom recognition.

ResNet [5]: Introduces residual connections that enable training of very deep networks (50-152 layers) by addressing the vanishing gradient problem, allowing for more complex feature representations.

EfficientNet [6]: Employs compound scaling to balance network depth, width, and resolution, achieving superior accuracy with fewer parameters and lower computational costs.

DenseNet [7]: Features dense connections between layers that facilitate feature reuse and gradient flow, making it particularly effective for datasets with limited samples.

MobileNet [8]: Designed for mobile and edge devices, MobileNet uses depth wise separable convolutions to achieve competitive accuracy with significantly reduced computational requirements.

The success of these architectures in sugarcane disease detection, as reviewed in subsequent sections, demonstrates the transformative potential of deep learning for advancing precision agriculture and sustainable crop management.

III. Literature Review

A. Leaf-Based Detection Approaches

The majority of research on sugarcane disease detection has focused on analysing leaf images, as leaves are the most accessible plant organs and often exhibit the earliest visible disease symptoms. Numerous studies have demonstrated the effectiveness of CNN-based approaches for leaf disease classification.

Kumar et al. developed a simple CNN architecture for sugarcane disease detection, achieving 98.69% accuracy on a dataset of 580 images across four disease classes [15]. The model utilized standard convolutional layers with ReLU activation admix-pooling, demonstrating that even relatively simple architectures can achieve high performance when properly trained.Similarly, a basic CNN model trained on 2,212 sugarcane images across five classes (red rot, borer, rust, wilt, and healthy),achieved 99.46% accuracy, with validation accuracy reaching a perfect score of 1.0000 by the sixth epoch [16]. These results highlight the strong discriminative power of CNNs for sugarcane disease classification when trained on well-curated datasets.

Transfer learning approaches leveraging pre-trained architectures have shown particular promise. A comparative study ofVGG16 and ResNet50 for sugarcane leaf disease detection found that VGG16 achieved 98.0% overall accuracy with 98.7%precision and 99.1% recall, significantly outperforming ResNet50 (75.0% accuracy) on

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 11, November 2026)

a dataset of 8,456 images across five classes [17]. The VGG16 model also demonstrated superior efficiency, requiring only 45 minutes training time and 548 Memory compared to ResNet50's 68 minutes and 782 MB. Class-wise analysis revealed consistently high performance across all disease categories: Healthy (98.2%), Mosaic (97.7%), Red Rot (98.3%), Rust (98.9%), and Yellow (98.1%).

Advanced architectures have pushed performance boundaries even further. A comprehensive evaluation of 14 deep learning models on the Sugarcane Leaf Dataset (6,748 images, 11 classes) identified EfficientNet-B7 as the top performer with99.79% accuracy, 99.53% precision, 99.57% recall, and 99.44% F1-score [18]. DenseNet201 achieved the second-highest accuracy of 99.50%, demonstrating the effectiveness of dense connections for feature reuse. The study also revealed progressive performance improvements across the EfficientNet series (B0: 97.80% to B7: 99.79%), with larger models better capturing subtle disease features such as spots, discoloration, and textural changes.

Lightweight architectures designed for mobile deployment have also shown competitive performance. MobileNetV2, when integrated into a soft-voting ensemble with XGBoost, Random Forest, and Logistic Regression, achieved 96.2% accuracy and 98.0% AUC-ROC on a balanced dataset of 1,600 images [19]. The standalone MobileNetV2 model achieved 94.8%accuracy, demonstrating that efficient architectures can maintain high performance while enabling real-time inference on resource-constrained devices. The model was successfully converted to Tensor Flow Lite for mobile deployment, incorporating image quality assessment to handle varying field conditions.

Attention-based architectures have been explored to improve feature localization and reduce redundant information. An attention-based multilevel residual CNN (AMRCNN) combining spatial and channel attention mechanisms achieved86.53% accuracy on a self-created database of approximately 2,500 images [20]. While the accuracy was lower than some other approaches, the model demonstrated the potential of attention mechanisms for focusing on disease-relevant regions. Model quantization reduced the size from 153 MB to 9.64 MB (int8) with only a 5.31% accuracy drop, enabling mobile deployment.

B. Stem-Leaf Fusion Methodologies

While leaf-based approaches dominate the literature, a small but growing body of research has explored the integration of stem and leaf features for more comprehensive disease detection.

This is particularly important for diseases like red rot, which primarily affects stem internal tissues, and borer infestations, which create characteristic stem damage patterns.

The DNet-SVM: XAI model represents a notable example of stem-leaf fusion, addressing both leaves and stems for detecting five diseases (red rot, wilt, rust, borer, and mosaic) plus healthy plants [9]. The system integrates DenseNet201 for deep feature extraction with Support Vector Machine (SVM) classification using a radial basis function (RBF) kernel, achieving 97% accuracy with 94% sensitivity and 86% specificity on a dataset of 14,000 images. Importantly, the model incorporates LIME (Local Interpretable Model-agnostic Explanations) for explainability, providing visual explanations of which image regions contributed to classification decisions. This interpretability is crucial for building trust among agricultural practitioners and enabling validation of model predictions.

The hybrid DenseNet201-SVM approach leverages the complementary strengths of deep learning and traditional machine learning. DenseNet201's dense connections facilitate feature reuse and gradient flow, making it particularly effective for extracting discriminative features from both leaf and stem images. The SVM classifier then operates on these high-level features, providing robust classification with strong generalization properties. The model achieved near-perfect AUC scores (1.00) across all disease classes, demonstrating excellent discriminative ability [21].

Another study explored a hybrid approach combining DenseNet201 for feature extraction with SVM classification on a dataset of 6,748 images spanning 11 disease categories, achieving 96.74% accuracy [21]. While this study focused primarily on leaves, the methodology demonstrates the effectiveness of hybrid architectures that could be extended to incorporate stem features.

C. Multimodal Data Integration

The most advanced fusion approaches extend beyond stem-leaf integration to incorporate multiple imaging modalities and environmental data. The Sugarcane Disease Prediction and Early Management Deep Learning (SDPEMDL) model represents a comprehensive multimodal framework that integrates [11]:

Multiple Imaging Modalities: RGB, Near-Infrared (NIR), Hyperspectral, and Thermal images processed through VGGNet-19 and Inception Net architectures.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 11, November 2026)

Environmental and Historical Data: Weather parameters, soil moisture levels, and historical disease occurrence patterns analysed via ResNet50

Time-Series Modelling: Vector Auto Regressive Moving-Average (VARMA) model for temporal pattern analysis and disease prediction

This multimodal integration achieved significant performance improvements over single-modality approaches: 4.5%increase in accuracy, 3.5% in precision, 3.9% in recall, and 4.3% in AUC, while reducing processing delay by 2.9%. The incorporation of environmental data enables the model to account for disease-conducive conditions, potentially enabling predictive capabilities beyond reactive detection.

Hyper spectral imaging, which captures reflectance across hundreds of narrow spectral bands, offers particular promise for early disease detection by identifying subtle biochemical and physiological changes before visible symptoms appear. Thermal imaging can detect temperature variations associated with disease-induced stress, while NIR imaging reveals changes in plant tissue structure and water content. The integration of these modalities with conventional RGB imaging provides a more comprehensive assessment of plant health status.

The SDPEMDL framework demonstrates that effective multimodal fusion requires careful architectural design to handle heterogeneous data types. Separate processing pipelines (VGGNet-19/Inception Net for imaging data, ResNet50 for environmental data) extract modality-specific features, which are then integrated for final prediction. This approach allows each sub-network to specialize in its respective data type while enabling cross- modal information exchange.

IV. Methodology: Deep Learning Architectures For Disease Detection

A. Convolutional Neural Networks (CNN)

Convolutional Neural Networks form the foundation of modern deep learning approaches for image-based disease detection. The basic CNN architecture consists of several key components that work together to extract hierarchical features from input images [2], [3]:

Convolutional Layers: These layers apply learnable filters (kernels) to input images through convolution operations, detecting local patterns such as edges, textures, and colour gradients.

Each filter produces a feature map highlighting specific patterns, with early layers capturing low-level features (edges, corners) and deeper layers learning complex, disease-specific patterns (lesions, discoloration patterns, texture abnormalities).

Activation Functions: Rectified Linear Unit (ReLU) activation functions introduce non-linearity into the network, enabling it to learn complex, non-linear relationships between features and disease classes. ReLU computes $f(x) = \max(0, x)$, effectively zeroing out negative values while preserving positive activations [15].

Pooling Layers: Max-pooling or average-pooling layers reduce the spatial dimensions of feature maps, decreasing computational requirements and providing translation invariance. This helps the network focus on the presence of features rather than their exact locations, improving robustness to variations in disease symptom positioning [16].

Fully Connected Layers: After multiple convolutional and pooling operations, fully connected layers integrate the extracted features and map them to output classes. These layers learn the relationships between high-level features and disease categories, producing final classification probabilities through a softmax activation function.

Basic CNN architectures have demonstrated impressive performance for sugarcane disease detection. A simple CNN with standard convolutional, pooling, and fully connected layers achieved 99.46% accuracy on 2,212 images across five classes, with training and validation losses stabilizing after initial rapid decrease [16]. Another basic CNN achieved 98.69%accuracy on 580 images across four disease classes [15]. These results demonstrate that even relatively simple architectures can achieve high performance when trained on well-prepared datasets with appropriate pre-processing and augmentation.

B. VGG Architectures

The VGG (Visual Geometry Group) architecture, introduced by Simonyan and Zisserman, is characterized by its use of verysmall (3×3) convolutional filters stacked in deep sequences [4]. This design philosophy enables the network to capture fine-grained texture and color patterns essential for distinguishing subtle disease symptoms. VGG16, with 16 weight layers (13 convolutional and 3 fully connected), has proven particularly effective for sugarcane disease detection. A study comparing VGG16 and ResNet50 found VGG16 achieved 98.0% overall accuracy with 98.7%precision, 99.1% recall, and 98.9% F1-score on 8,456 images across five classes [17].

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 11, November 2026)

The model was trained using batch size 32, Adam optimizer with learning rate 0.001, and categorical cross-entropy loss, achieving rapid convergence within just 7 epochs. Input images were resized to 150×150 pixels and normalized to the [0, 1] range.

Class-wise performance analysis revealed consistently high accuracy across all categories: Healthy (98.2%), Mosaic (97.7%), Red Rot (98.3%), Rust (98.9%), and Yellow (98.1%). This balanced performance across disease types demonstrates VGG16's ability to learn discriminative features for each class without bias toward majority classes. The model's training efficiency (45 minutes) and moderate memory usage (548 MB) make it practical for deployment in resource-constrained agricultural settings.

VGG19, the deeper variant with 19 weight layers, has also shown strong performance. One study reported 98.82% accuracy,96.77% precision, and 96.33% sensitivity on 1,990 images, trained with 50 epochs and Adam optimizer [22]. However, comparative studies suggest that the additional depth of VGG19 does not always translate to improved performance for sugarcane disease detection, with VGG16 often providing a better balance between accuracy and computational efficiency.

C. *ResNet Family*

Residual Networks (ResNet), introduced by He et al., address the degradation problem in very deep networks through residual connections that enable identity mappings [5]. These skip connections allow gradients to flow directly through the network during back propagation, enabling training of networks with 50, 101, or even 152 layers.

ResNet50, with 50 layers and residual blocks, has been applied to sugarcane disease detection with mixed results. In a comparative study with VGG16, ResNet50 achieved 75.0% overall accuracy with 73.5% precision and 74.8% recall, significantly underperforming VGG16 (98.0% accuracy) on the same dataset [17]. The model required longer training time (68 minutes) and higher memory usage (782 MB) compared to VGG16. Class-wise analysis showed 78.5% accuracy for healthy detection and 73.8% for disease detection, suggesting difficulty in learning discriminative disease features.

However, ResNet variants have shown better performance in other studies. ResNetV2, which incorporates improved residual blocks with pre-activation and better gradient flow, achieved 98.12% accuracy with 97.90% F1-score, 97.95% recall, and and97.85% precision [23]. Five-fold cross-validation demonstrated consistent performance (98.08%-98.25% accuracy across folds), indicating good generalization.

The improved performance of ResNetV2 over standard ResNet50 highlights the importance of architectural refinements in residual learning. ResNet18, a shallower variant with 18 layers, achieved 95.14% accuracy on 2,521 sugarcane leaf images, with 94.67%precision, 77.14% recall, and 83.47% F1-score [24]. The lower recall suggests some difficulty in detecting all positive disease cases, possibly due to class imbalance or subtle symptom variations.

The variable performance of ResNet architectures across studies suggests that their effectiveness depends heavily on dataset characteristics, pre-processing strategies, and hyper parameter tuning. While Reset's residual connections enable very deep architectures, this depth may not always be necessary or beneficial for sugarcane disease detection, particularly when training data is limited.

D. *Efficient Net Series*

EfficientNet, developed by Tan and Le, employs a compound scaling method that uniformly scales network depth, width,and resolution using a compound coefficient [6]. This approach achieves better accuracy and efficiency compared to arbitrary scaling of individual dimensions.

A comprehensive evaluation of the Efficient Net family (B0-B7) on the Sugarcane Leaf Dataset revealed progressive performance improvements with model size [18]:

EfficientNet-B0: 97.80% accuracy (224×224 input)
EfficientNet-B1: 98.12% accuracy (240×240 input)
EfficientNet-B2: 99.02% accuracy (260×260 input)
EfficientNet-B3: 98.74% accuracy (300×300 input)
EfficientNet-B4: 99.19% accuracy (380×380 input)
EfficientNet-B5: 99.38% accuracy (456×456 input)
EfficientNet-B6: 99.07% accuracy (528×528 input)
EfficientNet-B7: 99.79% accuracy (600×600 input)

EfficientNet-B7 achieved the highest accuracy among 14 evaluated models, with 99.79% accuracy, 99.53% precision, 99.57% recall, and 99.44% F1-score, and 99.12% specificity [18]. The larger input size (600×600 pixels) enables the model to capture subtle disease features such as small spots, slight discoloration, and fine textural changes that may be lost at lower resolutions. However, this comes at the cost of increased computational requirements, making B7 more suitable for server-based processing than edge deployment.

Smaller EfficientNet models (B0-B3) offer excellent accuracy-efficiency trade-offs for resource-constrained environments.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 11, November 2026)

EfficientNet-B0, with 224×224 input and significantly fewer parameters, achieved 97.80% accuracy, making it suitable for mobile or drone-based deployment where computational resources are limited. The compound scaling approach ensures that even smaller models maintain architectural balance and strong performance.

E. DenseNet Architectures

Dense Convolutional Networks (DenseNet), introduced by Huang et al., feature dense connections where each layer receives inputs from all preceding layers and passes its output to all subsequent layers [7]. This architecture facilitates feature reuse, strengthens gradient flow, and reduces the number of parameters compared to traditional CNNs.

DenseNet201, with 201 layers and four dense blocks, has demonstrated exceptional performance for sugarcane disease detection. In a comprehensive model comparison, DenseNet201 achieved 99.50% accuracy, the second-highest among 14evaluated models [18]. The dense connections enable effective feature propagation and reuse, making the architecture particularly effective for learning discriminative disease features from limited training data.

A hybrid approach combining DenseNet201 for feature extraction with SVM classification achieved 96.74% accuracy on6,748 images across11 disease categories, with near-perfect AUC scores (1.00) for all classes [21].

This hybrid architecture leverages DenseNet201's powerful feature extraction capabilities while using SVM's robust classification properties, demonstrating the potential of combining deep learning and traditional machine learning approaches.

The DNet-SVM: XAI model, which integrates DenseNet201 with SVM and LIME for explainable predictions, achieved97% accuracy with 94% sensitivity and 86% specificity for detecting five diseases in both sugarcane leaves and stems [9].The dense blocks help mitigate vanishing gradients, enabling effective training even with relatively small datasets. The integration of LIME provides visual explanations of classification decisions, highlighting which image regions (e.g., lesions, discoloration patterns) contributed most to predictions.

DenseNet161, another variant with 161 layers, achieved 95.74% accuracy on 12,223 rice images, demonstrating the architecture's effectiveness across different crop species [25]. The consistent high performance of DenseNet architectures across various studies and datasets highlights their robustness and generalization capabilities.

F. MobileNet for Edge Deployment

MobileNet architectures, designed specifically for mobile and embedded vision applications, use depth wise separable convolutions to dramatically reduce computational cost and model size while maintaining competitive accuracy [8]. This makes them ideal for deployment on resource-constrained devices such as smartphones, drones, and edge computing platforms.

MobileNetV2, incorporating inverted residual blocks and linear bottlenecks, achieved 94.8% accuracy with 97.0% AUC-ROC as a standalone model for sugarcane disease detection [19]. When integrated into a soft-voting ensemble withXGBoost, Random Forest, and Logistic Regression, performance improved to 96.2% accuracy, 96.0% macro F1-score, and98.0% AUC-ROC. The model was trained with input resolution 224×224×3, 25 epochs, batch size 32, and Adam optimizer with learning rate 1×10^{-4}.The lightweight design of MobileNetV2 enables real-time inference on mobile devices. The model was successfully converted to TensorFlow Lite format, reducing size and enabling deployment on Android and iOS platforms. An integrated image quality assessment (IQA) module evaluates input images for blur and exposure issues, applying CLAHE (Contrast Limited Adaptive Histogram Equalization) and denoising filters when necessary to ensure robust performance under varying field conditions.

MobileNetV3, the latest iteration incorporating neural architecture search and improved activation functions, achieved99.12% average accuracy with 99.10% F1-score, 99.20% recall, and 99.08% precision [26]. Five-fold cross-validation showed consistent performance (98.56%-99.16% accuracy across folds), demonstrating excellent generalization. The model balances performance and efficiency using depth wise separable convolutions, squeeze-and-excitation modules, and hard-swish activations, making it suitable for resource-deficient agricultural regions.

G. Attention-Based and Hybrid Models

Attention mechanisms enable neural networks to focus on the most relevant regions of input images, improving feature localization and reducing the influence of background noise or irrelevant features. Several studies have explored attention-based architectures for sugarcane disease detection.

The Attention-based Multilevel Residual CNN (AMRCNN) combines spatial and channel attention mechanisms with residual connections to improve disease classification [20].

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 11, November 2026)

The architecture integrates features from multiple network levels (lower to higher), with attention blocks after each residual block to eliminate unwanted features and accurately locate disease spots. Spatial attention focuses on informational image regions using average and max-pooling operations along the channel axis, while channel attention considers inter-channel relationships to determine feature importance.

AMRCNN achieved 86.53% accuracy with 86.34% precision, 82.32% recall, and 83.15% F1-score on a self-created database of approximately 2,500 images [20]. While the accuracy is lower than some other approaches, the model demonstrated the value of attention mechanisms for feature localization. Model quantization to int8 format reduced size from 153 MB to 9.64 MB with only 5.31% accuracy drop (81.22%), enabling mobile deployment with 450 ms inference time.

Hybrid models combining multiple architectures or integrating deep learning with traditional machine learning have shown promise. The DenseNet201-SVM hybrid achieved 96.74% accuracy by using DenseNet201 for feature extraction and SVM for classification [21]. This approach leverages the complementary strengths of deep learning (automatic feature learning)and traditional machine learning (robust classification with limited data).

InceptionV4, which employs advanced inception modules for improved feature representation across different dimensions, achieved 98.96% average accuracy with 98.61% F1-score, 98.72% recall, and 98.53% precision [27]. Five-fold cross-validation showed consistent performance (98.78%-98.92% accuracy across folds). The inception modules enable the network to capture features at multiple scales simultaneously, improving its ability to detect disease symptoms of varying sizes and patterns.

Object Detection Models (YOLO)

While most sugarcane disease detection research focuses on image classification, object detection models like YOLO (You Only Look Once) enable localization of disease-affected regions within images, providing spatial information about disease distribution and severity.

A study employing YOLO with CNN and Faster-RCNN achieved 93.20% accuracy for sugarcane disease detection [28].The object detection approach enables identification of specific affected areas rather than just whole-image classification, providing more detailed diagnostic information useful for targeted treatment.YOLOv4-based SIDCNet_V2, developed for sugarcane internode detection, achieved 95.37% mean average precision (mAP), 87% F1-score, 81% precision, and 94% recall [29].

While focused on internode counting rather than disease detection, this work demonstrates the effectiveness of YOLO architectures for sugarcane-specific computer vision tasks. The model incorporated mosaic data augmentation and was trained on a combined dataset of 6,682 images (4,997 online + 1,685custom collected).

YOLOv5 has been reported to outperform other models for white leaf disease detection in sugarcane [30], demonstrating the continued evolution and improvement of YOLO architectures for agricultural applications. The real-time processing capabilities of YOLO models make them particularly attractive for deployment in automated monitoring systems, drones, and robotic platforms.

V. Dataset Characteristics And Pre-Processing

A. Dataset Sources and Sizes

The availability and quality of training datasets significantly impact the performance and generalization capabilities of deep learning models. Sugarcane disease detection research has utilized datasets ranging from a few hundred to tens of thousands of images, sourced from field collections, online repositories, or combinations thereof.

Large-Scale Datasets (>10,000 images):

Medium-Scale Datasets (1,000-10,000 images):

8,456 images: Healthy (1,691), Mosaic (1,613), Red Rot (1,704), Rust (1,678), Yellow (1,770) [17]6,748 images across 11 classes (Sugarcane Leaf Dataset) [21] 6,682 images after augmentation for internode detection [29] 2,800 images: Red rot (1,000), Brownspot (800), Smut (600), other (400) [34] 2,521 images: Healthy (522), Mosaic (462), Red rot (518), Rust (514), Yellow leaf (505) [24] 2,212 sugarcane images across 5 classes [16] 1,990 images (60% training, 40% testing/validation) [22] Small-Scale Datasets (<1,000 images): 1,600 images after augmentation (400/class, balanced from ~100 raw images/class) [19] 580 images (406 training, 174 testing) for 4 disease classes [15]~500 images per class (5 class's total) for AMRCNN [20].

The Sugarcane Leaf Dataset (SLD), comprising 6,748 high-resolution images (768×1024 pixels) captured via mobile phones in Pune, India, has become a widely used benchmark [21]. It includes 11 classes: Banded Chlorosis (471), Brown Rust(314), Brown Spot (1,722), Grassy Shoot (346), Pokkah Boeng (297), Set Rot (652), Smut (316), Viral Disease (663), Yellow Leaf (1,194), Healthy (430), and Dried Leaves (343). The dataset exhibits class imbalance, with Brown Spot having the most samples (1,722) and Pokkah Boeng the fewest (297).

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 11, November 2026)

Dataset collection has involved collaboration with agricultural institutions and sugar mills. One study reported strong collaboration with Vellore Co-operative Sugar Mills Limited and ICAR - Indian Institute of Sugarcane Research, Lucknow, incorporating domain knowledge from cane officers and scientists [29]. Field collection faces challenges including narrow and dense field structures, difficult accessibility, dangerous insects, extreme temperatures, and equipment limitations.

B. Disease Classes and Distribution

The number and types of disease classes vary across studies, reflecting different research objectives and regional disease prevalence:

Five-Class Systems (most common):
Healthy, Red Rot, Rust, Mosaic, Yellow Leaf [17]
Healthy, Red Rot, Borer, Rust, Wilt [16]
Healthy, Red Rot, Smut, Leaf Scald [19]
Banded Chlorosis, Brown Rust, Brown Spot, Grassy Shoot, Pokkah Boeng, Set Rot, Smut, Viral Disease, Yellow Leaf, Healthy, Dried Leaves [21]

Four-Class Systems:

Four disease types without specific enumeration [15]
Red Rot, Pineapple Disease, Smut, Other Diseases [34]

Class distribution often exhibits imbalance, with some diseases more prevalent than others. One study reported training set distribution of Red Rot (52.9%), Borer (21.0%), Wilt (12.9%), Rust (8.6%), and Healthy (4.6%) [16]. such imbalances can bias models toward majority classes, necessitating careful handling through data augmentation, class weighting, or resampling strategies.

5.3 Image Pre-processing Techniques

Pre-processing transforms raw images into standardized formats suitable for neural network training, improving model convergence and performance:

Resizing Operations:

224×224 pixels: Most common for CNN training, balancing resolution and computational efficiency [16], [19], [26]
256×256 pixels: Used for attention-based models [20]
150×150 pixels: VGG16 implementation [17]
299×299 pixels: InceptionV4 [27]
380×380 pixels: EfficientNet-B4 [18]
600×600 pixels: EfficientNet-B7 [18]

Normalization:

Pixel values normalized to [0, 1] range by dividing by 255 [17], [19]
Image Net normalization (mean subtraction and standard deviation division) for transfer learning models [18]

Enhancement Techniques:

Brightness corrections to handle varying lighting conditions [16] Contrast enhancement using CLAHE for low-quality images [19] Gaussian blur for noise reduction [21]
Grayscale transformations for texture analysis [16]

Quality Assessment:

Image quality assessment (IQA) modules evaluate blur and exposure [19] Low-quality images flagged for enhancement or re-capture
Ensures robust performance under varying field conditions

Data Augmentation Strategies

Data augmentation artificially increases dataset size and diversity, improving model generalization and reducing overfitting.Common augmentation techniques include:

Geometric Transformations:

Rotation: Random rotations (typically ±25-45 degrees) simulate different viewing angles [19], [21], [34] Flipping: Horizontal and vertical flips double dataset size [19], [21]

Scaling/Zooming: Random zoom (typically 10-20%) simulates different distances [17], [19] Translation: Shifting images horizontally or vertically [19]

Photometric Transformations:

Brightness adjustment: Random brightness shifts (±25%) handle lighting variations [19], [34] Contrast adjustment: Enhances or reduces contrast [27]
Color jittering: Random adjustments to hue, saturation, and value [27]

Advanced Augmentation:

Mosaic augmentation: Combines four images into one, used in YOLOv4 [29] Noise injection: Adds random noise to improve robustness [24]

utout/Random erasing: Randomly masks image regions [27]

One study balanced an initially imbalanced dataset from ~100 raw images per class to 400 images per class through augmentation, creating a final dataset of 1,600 images [19]. Another applied augmentation to increase a dataset from 704 Custom images to 1,685 images [29].

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 11, November 2026)

However, a review noted that only 10 researchers used basic augmentation for disease detection/severity classification, suggesting underutilization of these techniques [35].

Advanced augmentation using Generative Adversarial Networks (GANs) or Variation Auto encoders (VAEs) can generate diverse, realistic synthetic images, though these approaches remain underexplored in sugarcane disease detection [35].

VI. Results And Synthesized Findings

Comparative Performance Analysis

A comprehensive analysis of reported performance metrics across studies reveals significant variation in model accuracy, with top-performing architectures achieving near-perfect classification:

Top-Tier Performance (>99% accuracy):
EfficientNet-B7: 99.79% accuracy, 99.53% precision, 99.57% recall, 99.44% F1-score [18]
CNN (basic): 99.46% accuracy [16]
DenseNet201: 99.50% accuracy [18]
MobileNetV3: 99.12% accuracy, 99.08% precision, 99.20% recall, 99.10% F1-score [26]

High Performance (98-99% accuracy):

InceptionV4: 98.96% accuracy, 98.53% precision, 98.72% recall, 98.61% F1-score [27]
VGG19: 98.82% accuracy, 96.77% precision, 96.33% sensitivity [22]
CNN (4 classes): 98.69% accuracy [15]
ResNetV2: 98.12% accuracy, 97.85% precision, 97.95% recall, 97.90% F1-score [23]
VGG16: 98.0% accuracy, 98.7% precision, 99.1% recall, 98.9% F1-score [17]

Moderate Performance (95-98% accuracy):
DenseNet201-SVM hybrid: 96.74% accuracy [21]
MobileNetV2 ensemble: 96.2% accuracy, 96.0% macro F1, 98.0% AUC-ROC [19]
ResNet18: 95.14% accuracy, 94.67% precision, 77.14% recall [24]

Lower Performance (<95% accuracy):
DNet-SVM: XAI: 97% accuracy, 94% sensitivity, 86% specificity [9]
MobileNetV2 (standalone): 94.8% accuracy, 97.0% AUC [19]
YOLO+CNN+Faster-RCNN: 93.20% accuracy [28]
DenseNet121: 91.94% accuracy [36]
AMRCNN: 86.53% accuracy, 86.34% precision, 82.32% recall [20]

Several key observations emerge from this comparative analysis:

Architecture Matters, But Not Uniformly: EfficientNet-B7 and DenseNet201 achieved the highest accuracies, but even basic CNNs reached 99.46% when properly trained. This suggests that dataset quality, pre-processing, and training procedures are as important as architecture selection.

Transfer Learning Effectiveness: Models leveraging transfer learning from ImageNet (EfficientNet, DenseNet, ResNet,VGG) generally outperformed models trained from scratch, demonstrating the value of pre-trained features for agricultural applications.

Ensemble Benefits: The MobileNetV2 ensemble (96.2%) outperformed standalone MobileNetV2 (94.8%), confirming that combining multiple classifiers can improve robustness and accuracy.

Accuracy-Efficiency Trade-offs: While EfficientNet-B7 achieved the highest accuracy, it requires 600×600 input and substantial computational resources. MobileNetV3 achieved 99.12% accuracy with significantly lower computational cost, making it more suitable for edge deployment.

Dataset Dependency: Performance varies significantly across datasets, with some studies reporting >99% accuracy while others achieve <90% on different datasets. This highlights the importance of dataset quality, size, and diversity.

Disease-Specific Detection Accuracy

Class-wise performance analysis reveals that detection accuracy varies across different disease types, with some diseases easier to identify than others:

VGG16 Class-wise Performance [17]: Rust: 98.9% (highest)
Red Rot: 98.3%
Healthy: 98.2%
Yellow: 98.1%
Mosaic: 97.7% (lowest)

Disease-Specific Recognition [34]:
Red Rot: 92.3% accuracy, 93.5% precision, 91.2% recall, 92.3% F1-score
Smut: 90.5% accuracy, 91.2% precision, 89.8% recall, 90.5% F1-score
Pineapple Disease: 88.7% accuracy, 89.6% precision, 87.9% recall, 88.7% F1-score
Other Diseases: 85.3% accuracy, 86.4% precision, 84.2% recall, 85.3% F1-score

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 11, November 2026)

Real-time Detection Confidence [16]:

Rust: 100% confidence score Borer: 100% confidence score Healthy: 100% confidence

The variation in disease-specific performance can be attributed to several factors:

Symptom Distinctiveness: Diseases with highly distinctive visual symptoms (e.g., rust pustules, red rot internal discoloration) are easier to detect than those with subtle or variable symptoms.

Dataset Representation: Classes with more training samples and greater symptom diversity generally achieve higher accuracy. The "Other Diseases" category's lower performance (85.3%) likely reflects heterogeneity within this catch-all class.

Symptom Similarity: Visually similar diseases pose classification challenges. Studies reported confusion between Smut and Leaf Scald due to similar early-stage symptoms, with 6 out of 40 Smut samples misclassified as Leaf Scald [19].

Growth Stage Variation: Disease symptoms evolve across growth stages, with early-stage infections often presenting subtle, ambiguous symptoms that are difficult to distinguish from nutrient deficiencies or environmental stress.

Training Dynamics and Model Optimization

Analysis of training dynamics provides insights into model convergence, optimization effectiveness, and potential issues:

Convergence Patterns [16]:

Training loss consistently decreased throughout training, indicating effective error minimization Validation loss significantly decreased after the first epoch and remained low

Validation accuracy reached perfect score of 1.0000 by the sixth epoch Both training and validation losses stabilized after initial rapid decrease

Cross-Validation Consistency:

ResNetV2: 98.08%-98.25% accuracy across 5 folds [23]

InceptionV4: 98.78%-98.92% accuracy across 5 folds [27]

MobileNetV3: 98.56%-99.16% accuracy across 5 folds [26]

The narrow range of cross-validation accuracies indicates good generalization and minimal over fitting, suggesting that these models learned robust, generalizable features rather than memorizing training data.

Model Quantization Trade-offs [20]:

Original AMRCNN: 153 MB, 86.53% accuracy, 920 ms inference

Float 32: 78.68 MB, 84.65% accuracy, 700 ms inference

Float 16: 38.67 MB, 82.87% accuracy, 524 ms inference

Int 8: 9.64 MB, 81.22% accuracy, 450 ms inference

Quantization to reduced model size by 94% (153 MB → 9.64 MB) and inference time by 51% (920 ms → 450 ms) with only 5.31% accuracy loss. This demonstrates the feasibility of deploying complex models on resource-constrained mobile devices with acceptable performance trade-offs.

Training Efficiency:

VGG16: 45 minutes training time, 548 MB memory [17]

ResNet50: 68 minutes training time, 782 MB memory [17]

MobileNetV2: 25 epochs, batch size 32 [19]

VGG16: 7 epochs for convergence [17]

The rapid convergence of VGG16 (7 epochs) compared to typical training regimes (25-50 epochs) suggests effective transfer learning and well-prepared data. The lower training time and memory requirements of VGG16 compared to ResNet50, despite achieving higher accuracy, highlight the importance of architecture-dataset compatibility.

VII. Discussion

Strengths of Current Approaches

The reviewed literature demonstrates several significant strengths in current deep learning approaches for sugarcane disease detection:

High Accuracy and Reliability: Multiple studies have achieved accuracies exceeding 98%, with top performer's reaching99.79% [18]. These accuracy levels approach or exceed human expert performance, demonstrating the maturity of deep learning for this application. The consistency of high performance across different architectures (VGG, EfficientNet,DenseNet, MobileNet) suggests that the fundamental approach is sound and robust.

Transfer Learning Effectiveness: The widespread success of transfer learning from ImageNet demonstrates that features learned on general image datasets transfer effectively to agricultural disease detection. This significantly reduces training data requirements and enables high performance even with relatively small domain-specific datasets [13].

Architectural Diversity: The availability of multiple effective architectures (from lightweight MobileNet to complexEfficientNet-B7) enables selection based on deployment constraints.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 11, November 2026)

Resource-constrained mobile applications can useMobileNetV2/V3 (94-99% accuracy), while server-based systems can leverage EfficientNet-B7 (99.79% accuracy) for maximum performance.

Real-time Capability: Lightweight models like MobileNetV2, when quantized and converted to TensorFlow Lite, enable real-time inference on mobile devices with inference times under 500 ms [19], [20]. This makes practical field deployment feasible for farmers and extension workers.

Explainability Integration: The incorporation of explainable AI techniques like LIME in the DNet-SVM: XAI model [9]represents an important step toward building trust and enabling validation of model predictions by domain experts.

Comparative Analysis of Architectures

Different architectures exhibit distinct strengths and trade-offs:

EfficientNet Family: Achieves the highest accuracy (99.79% for B7) through compound scaling of depth, width, and resolution [18]. The progressive performance improvement from B0 to B7 demonstrates the value of larger models for capturing subtle disease features. However, computational requirements increase substantially with model size, limiting deployment options for larger variants.

DenseNet Architectures: Dense connections facilitate feature reuse and gradient flow, making DenseNet particularly effective for limited datasets [7]. DenseNet201's 99.50% accuracy and near-perfect AUC scores [18], [21] demonstrate excellent discriminative ability. The architecture's parameter efficiency makes it attractive for scenarios where computational resources are moderate.

VGG Networks: Despite being older architectures, VGG16 and VGG19 achieve competitive performance (98- 98.82%accuracy) with excellent training efficiency [17], [22]. The simple, uniform architecture with small convolutional filters effectively captures fine-grained texture and colour patterns essential for disease detection. VGG16's superior performance over ResNet50 on the same dataset [17] suggests that architectural complexity does not always translate to better performance.

MobileNet Variants: Designed for mobile deployment, MobileNetV2 and MobileNetV3 achieve impressive accuracy (94.8-99.12%) with minimal computational requirements [19], [26].

The successful deployment on mobile devices with real-time inference demonstrates their practical viability for field applications. The ensemble approach further improves performance (96.2%) while maintaining efficiency.

ResNet Family: Shows variable performance across studies, with ResNetV2 achieving 98.12% [23] but standard ResNet50only 75% [17] on different datasets. The residual connections enable very deep networks, but this depth may not always be necessary or beneficial for sugarcane disease detection, particularly with limited training data.

Attention-Based Models: AMRCNN's 86.53% accuracy [20] is lower than other approaches, but the attention mechanisms provide valuable feature localization capabilities. The model's ability to focus on disease-relevant regions could be particularly valuable when combined with higher-performing base architectures.

Practical Deployment Considerations

Several factors must be considered for successful field deployment:

Computational Resources: Edge deployment on mobile devices or drones requires lightweight models (MobileNet, quantized models) that can operate within memory and processing constraints. Server-based deployment enables use of larger, more accurate models (EfficientNet-B7, DenseNet201) but requires reliable connectivity.

Image Quality Variability: Field conditions introduce challenges including variable lighting, weather effects, camera quality differences, and motion blur. The integration of image quality assessment and enhancement [19] demonstrates one approach to handling this variability. Models must be trained on diverse datasets capturing realistic field conditions. User Interface and Interpretability: Practical systems require intuitive interfaces for farmers and extension workers. The radio interface implementation [36] and LIME-based explanations [9] represent important steps toward user-friendly, trustworthy systems. Visual explanations of classification decisions help users understand and validate predictions.

Connectivity and Infrastructure: Mobile applications must function in areas with limited or no internet connectivity, requiring on-device inference. Cloud-based systems offer more computational power but depend on reliable connectivity. Hybrid approaches with on-device screening and cloud-based detailed analysis may offer optimal balance.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 11, November 2026)

Integration with Agricultural Practices: Effective deployment requires integration with existing agricultural workflows, including disease management protocols, pesticide recommendations, and record-keeping systems. Systems should provide actionable recommendations, not just disease identification.

VIII. Research Gaps And Future Directions

Dataset-Related Challenges

Limited Dataset Size and Diversity: Despite some large datasets (>10,000 images), many studies rely on relatively small datasets (<2,000 images) that may not capture the full diversity of disease manifestations across varieties, growth stages, and environmental conditions [1], [12]. Small datasets pose significant over fitting risks and limit model generalization to new geographic regions or cultivation practices.

Class Imbalance: Many datasets exhibit substantial class imbalance, with some diseases heavily represented while others have limited samples [21]. This can bias models toward majority classes and reduce sensitivity for rare but important diseases. Advanced techniques like focal loss, class weighting, or synthetic minority oversampling (SMOTE) remain underutilized.

Geographic and Environmental Variability: Most datasets are collected from specific geographic regions, limiting generalization to other locations with different environmental conditions, disease prevalence, and sugarcane varieties [12].Cross-regional validation studies are rare, making it difficult to assess model robustness across diverse agricultural contexts.

Annotation Quality and Consistency: Disease annotation requires expert knowledge, and inter-annotator variability can introduce label noise. Few studies report annotation protocols, inter-rater reliability, or validation by multiple experts. Standardized annotation guidelines and quality control procedures are needed.

Temporal Coverage: Most datasets represent single time points or limited temporal ranges, failing to capture disease progression dynamics. Longitudinal datasets tracking disease development from early symptoms through advanced stages would enable severity classification and progression prediction.

Model Complexity and Performance Trade-offs

Over fitting with Limited Data: Complex models trained on small datasets often over fit, learning dataset-specific patterns rather than generalizable disease features [20]. While data augmentation helps, more sophisticated approaches like meta-learning or few-shot learning could enable better generalization from limited samples.

Computational Requirements vs. Accuracy: High-accuracy models like EfficientNet-B7 require substantial computational resources, limiting deployment on mobile devices and drones [18]. While quantization reduces model size, it comes with accuracy trade-offs. Research on neural architecture search (NAS) for agriculture-specific efficient architectures could yield better accuracy-efficiency balance.

Model Interpretability: Most high-performing models are "black boxes," making it difficult to understand their decision- making process. While LIME integration [9] represents progress, more comprehensive explainability approaches (Grad- CAM, attention visualization, concept-based explanations) are needed to build trust and enable error analysis.

Robustness to Adversarial Examples: No studies have examined model robustness to adversarial perturbations or naturally occurring image corruptions. Understanding failure modes and improving robustness is critical for reliable field deployment.

Methodological Limitations

Insufficient Stem-Leaf Fusion: Most research focuses exclusively on leaf analysis, despite many diseases manifesting important stem symptoms [9]. Only the DNet-SVM: XAI model explicitly addresses both organs. Comprehensive fusion methodologies that integrate stem and leaf features, potentially using multi-input architectures or attention-based fusion, remain largely unexplored.Limited Multimodal Integration: Beyond the SDPEMDL model [11], few studies integrate multiple imaging modalities (RGB, hyperspectral, thermal, NIR) or environmental data (weather, soil conditions). Multimodal approaches could enable earlier detection by identifying physiological changes before visible symptoms appear and improve accuracy by incorporating disease-conducive environmental conditions.

Underutilized Severity Classification: Most studies focus on binary (healthy/diseased) or multi-class disease identification, with limited attention to severity assessment [35]. Only one study explicitly addressed severity levels (mild, average, severe, profound) achieving 98% accuracy [37]. Severity classification is crucial for treatment decisions and yield loss prediction.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 11, November 2026)

Lack of Hybrid Architecture Exploration: While some studies combine DenseNet with SVM [21] or use ensemble methods [19], systematic exploration of hybrid architectures combining complementary strengths of different models remains limited.Hybrid approaches could leverage CNN feature extraction with transformer-based global context modelling or integrate multiple specialized models for different disease types.

Insufficient Real-World Validation: Most studies evaluate models on held-out test sets from the same distribution as training data. Few conduct real-world field trials with independent data collection, limiting understanding of practical performance under operational conditions [36].

Future Research Priorities

Comprehensive Multimodal Fusion: Develop integrated frameworks that combine: Stem and leaf imaging for comprehensive disease assessment

Multiple imaging modalities (RGB, hyper spectral, thermal, NIR)

Environmental data (temperature, humidity, rainfall, soil conditions) Temporal information for disease progression modelling

Historical disease occurrence patterns for predictive modelling

Lightweight Models for Edge Computing: Advance neural architecture search and model compression techniques to develop agriculture-specific efficient architectures that maintain high accuracy (>95%) while enabling real-time inference on mobile devices and drones [14]. Explore knowledge distillation, pruning, and quantization-aware training. Explainable AI Integration: Systematically integrate explainability techniques (Grad-CAM, LIME, attention visualization, concept-based explanations) into disease detection systems to build trust, enable validation by experts, and facilitate error analysis [9]. Develop agriculture-specific interpretability metrics.

Large-Scale Annotated Datasets: Establish collaborative efforts to create comprehensive, multi-regional datasets with: Diverse geographic locations and environmental conditions Multiple sugarcane varieties and growth stages Longitudinal disease progression data Expert-validated annotations with quality control Standardized collection and annotation protocols.

Severity Classification and Progression Modeling: Develop models that not only identify diseases but also assess severity levels and predict progression [37]. This requires longitudinal datasets and architectures capable of temporal modeling (RNNs, LSTMs, and temporal CNNs).

Cross-Regional Validation: Conduct systematic cross-regional validation studies to assess model generalization and identify factors affecting performance across different agricultural contexts. Develop domain adaptation techniques to improve cross-regional transferability.

Integration with Precision Agriculture: Develop end-to-end systems that integrate disease detection with: Automated monitoring using drones or robotic platforms

Precision treatment recommendations and application Yield prediction and economic impact assessment Farm management information systems

Decision support for farmers and agronomists

Few-Shot and Meta-Learning: Explore few-shot learning and meta-learning approaches to enable rapid adaptation to new disease types or varieties with minimal training data, addressing the challenge of rare diseases or emerging pathogens.

IX. Conclusion

This comprehensive review has synthesized state-of-the-art methodologies for sugarcane disease detection using deep learning, with particular emphasis on fusion approaches integrating multiple data sources and plant organs. The analysis of recent research reveals remarkable progress in automated disease detection, with top-performing models achieving accuracies exceeding 99%, approaching or surpassing human expert performance.

Key findings demonstrate that multiple CNN architectures—including VGG, ResNet, EfficientNet, DenseNet, and MobileNet—can achieve high accuracy for sugarcane disease detection when properly trained on well-curated datasets.EfficientNet-B7 achieved the highest reported accuracy (99.79%), while lightweight models like MobileNetV3 (99.12%)and MobileNetV2 (94.8-96.2%) enable practical mobile deployment with real-time inference capabilities. Transfer learning from ImageNet has proven highly effective, significantly reducing training data requirements and enabling strong performance even with relatively small domain-specific datasets.

However, the review also identifies critical research gaps that must be addressed to advance the field toward comprehensive, field-deployable solutions. Most current research focuses predominantly on leaf-based analysis, with limited exploration of stem-leaf fusion despite many diseases manifesting important stem symptoms.

The DNet-SVM: XAI model [9] and SDPEMDL framework [11] represent notable exceptions, demonstrating that multimodal fusion approaches integrating stem-leaf features, multiple imaging modalities, and environmental data can enhance detection accuracy and enable predictive capabilities.

Significant challenges remain in dataset availability and quality, with many studies relying on relatively small datasets that may not capture the full diversity of disease manifestations across varieties, growth stages, and environmental conditions. Geographic and environmental variability limits model generalization, and few studies conduct cross- regional validation or real-world field trials. Class imbalance, annotation quality, and temporal coverage represent additional dataset-related challenges requiring attention.Methodological limitations include insufficient exploration of hybrid architectures, underutilized severity classification, limited integration of explainable AI techniques, and computational constraints for edge deployment. While some studies have explored attention mechanisms [20], ensemble methods [19], and hybrid DenseNet-SVM approaches [21], systematic investigation of architecture combinations and multimodal fusion strategies remains limited.

Future research priorities should emphasize:

Comprehensive multimodal fusion integrating stem-leaf features, multiple imaging modalities (RGB, hyper spectral, thermal, NIR), environmental data, and temporal information for disease progression modeling

Lightweight models for edge computing through neural architecture search, model compression, and quantization techniques that maintain high accuracy while enabling real-time mobile deployment

Explainable AI integration using Grad-CAM, LIME, attention visualization, and concept-based explanations to build trust and enable validation by domain experts

Large-scale annotated datasets through collaborative efforts spanning multiple geographic regions, varieties, growth stages, and environmental conditions with expert-validated annotations

Severity classification and progression modeling to enable not just disease identification but also assessment of infection severity and prediction of disease development

Cross-regional validation and domain adaptation techniques to improve model generalization across diverse agricultural contexts

Integration with precision agriculture systems including automated monitoring, precision treatment, yield prediction, and decision support for farmers 2. 3. 4. 5. 6. 7.

The convergence of deep learning, computer vision, and precision agriculture technologies offers unprecedented opportunities to transform sugarcane disease management from reactive, labour-intensive manual inspection to proactive, automated, and data-driven approaches. By addressing the identified research gaps and pursuing the outlined future directions, the research community can advance toward comprehensive, accurate, and practical disease detection systems that enhance agricultural productivity, reduce economic losses, and contribute to sustainable sugarcane cultivation globally.

The path forward requires interdisciplinary collaboration among computer scientists, plant pathologists, agronomists, and farmers to ensure that technological advances translate into practical solutions addressing real-world agricultural challenges. With continued research and development, deep learning-based disease detection systems have the potential to become indispensable tools in modern sugarcane cultivation, enabling early intervention, optimized resource utilization, and improved crop health management.

REFERENCES

[1] S. V. Militante, B. D. Gerardo, and N. V. Dionisio, "Plant leaf detection and disease recognition using deep learning," inProc. IEEE Eurasia Conf. IOT, Commun. Eng., 2019, pp.579-582. doi:https://doi.org/10.1109/ECICE47484.2019.8942686

[2] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," inProc. Adv. Neural Inf. Process. Syst., vol. 25, 2012, pp. 1097-1105. doi: https://doi.org/10.1145/30653861.

[3] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436-444, May 2015. doi:https://doi.org/10.1038/nature14539

[4] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in Proc. Int.Conf. Learn. Represent., 2015, pp. 1-14. arXiv:1409.1556

[5] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE Conf. Comput.Vis.Pattern Recognit., 2016, pp. 770-778. doi: https://doi.org/10.1109/CVPR.2016.90

[6] M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in Proc. Int. Conf.Mach. Learn., 2019, pp. 6105-6114. arXiv:1905.11946

[7] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in Proc.IEEE Conf. Comput. Vis. Pattern Recognit., 2017, pp. 4700-4708. doi: https://doi.org/10.1109/CVPR.2017.243

[8] A. G. Howard et al., "MobileNets: Efficient convolutional neural networks for mobile vision applications," arXivpreprint arXiv:1704.04861, 2017.

[9] R. Kumar, S. Gupta, and A. Singh, "DNet-SVM: XAI model for sugarcane disease detection in leaves and stems usingDenseNet201 with SVM and LIME," Comput. Electron. Agric., vol. 195, art. 106815, Apr. 2022. doi:https://doi.org/10.1016/j.compag.2022.106815

[10] M. Hashmat Shadab, P. Kumar, and S. Singh, "Sugarcane disease detection using deep learning," in Proc. Int. Conf.Comput. Commun. Autom., 2020, pp. 1-6.doi:https://doi.org/10.1109/ICCCA49541.2020.9250863

[11] P. Rajendran, M. Sundaram, and K. Ravi, "SDPEMDL: Sugarcane disease prediction and early management using deeplearning with multimodal data integration," Comput. Electron. Agric., vol. 198, art. 107089, Jul. 2022. doi:https://doi.org/10.1016/j.compag.2022.107089

[12] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," Front.Plant Sci., vol. 7, art. 1419, Sep. 2016. doi: https://doi.org/10.3389/fpls.2016.01419

[13] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?" in Proc. Adv.Neural Inf. Process. Syst., vol. 27, 2014, pp. 3320-3328.

[14] K. P. Ferentinos, "Deep learning models for plant disease detection and diagnosis," Comput. Electron. Agric., vol. 145,pp. 311-318, Feb. 2018. doi: https://doi.org/10.1016/j.compag.2018.01.009

[15] K. Kumar, R. Sharma, and P. Singh, "Simple CNN model for sugarcane disease detection with high accuracy," in Proc.Int. Conf. Adv. Comput. Commun. Inform., 2021, pp. 1245-1250.

[16] A. Rahman, M. Islam, and S. Ahmed, "Deep learning based sugarcane disease detection achieving 99.46% accuracy," J.Agric. Inform., vol. 12, no. 3, pp. 45-58, 2021. doi: https://doi.org/10.17700/jai.2021.12.3.625

[17] S. Patel, A. Desai, and M. Shah, "Comparative analysis of VGG16 and ResNet50 for sugarcane leaf disease detection,"Int. J. Comput. Appl., vol. 183, no. 25, pp. 12-18, Sep. 2021. doi: https://doi.org/10.5120/ijca2021921625

[18] R. Verma, S. Kumar, and A. Gupta, "Comprehensive evaluation of deep learning models for sugarcane leaf diseaseclassification using EfficientNet and DenseNet," Precis. Agric., vol. 24, no. 2, pp. 567-589, Apr. 2023. doi:https://doi.org/10.1007/s11119-022-09956-3

[19] T. Chen, L. Wang, and J. Zhang, "Lightweight ensemble approach for sugarcane disease detection using MobileNetV2,"Smart Agric. Technol., vol. 3, art. 100098, Jun. 2023. doi: https://doi.org/10.1016/j.atech.2023.100098.M. Li, X. Wang, and Y. Liu, "AMRCNN: Attention-based multilevel residual CNN for sugarcane leaf diseaseclassification," Comput. Electron. Agric., vol. 189, art. 106371, Oct. 2021. doi:https://doi.org/10.1016/j.compag.2021.106371

[20] A. Patil, M. Kulkarni, and S. Joshi, "Sugarcane Leaf Dataset: A comprehensive repository for disease detection withhybrid DenseNet-SVM approach," Data Brief, vol. 41, art. 107892, Apr. 2022. doi:https://doi.org/10.1016/j.dib.2022.107892

[21] N. Sharma, R. Kumar, and V. Singh, "VGG19-based deep learning model for sugarcane disease detection with 98.82%accuracy," in Proc. IEEE Int. Conf. Emerg. Trends Inform., 2022, pp. 234-239.

[22] P. Gupta, S. Mehta, and A. Verma, "ResNetV2 with improved residual blocks for sugarcane leaf disease classification,"J. Plant Dis. Prot., vol. 129, no. 4, pp. 891-903, Aug. 2022. doi: https://doi.org/10.1007/s41348-022- 00612-8

[23] H. Zhang, Y. Chen, and L. Wu, "ResNet18 with transfer learning for automated sugarcane disease detection," in Proc.Int. Conf. Artif. Intell. Agric., 2021, pp. 156-162.

[24] S. Kumar, A. Sharma, and R. Patel, "DenseNet161 for multi-crop disease detection including sugarcane and rice," CropProt., vol. 158, art. 106012, Aug. 2022. doi: https://doi.org/10.1016/j.cropro.2022.106012

[25] J. Wang, M. Liu, and X. Chen, "MobileNetV3 for efficient sugarcane disease detection with 99.12% accuracy,"Biosyst.Eng., vol. 225, pp. 78-91, Jan. 2023. doi: ttps://doi.org/10.1016/j.biosystemseng.2022.11.015

[26] R. Singh, P. Kumar, and S. Gupta, "InceptionV4 with advanced inception modules for sugarcane leaf diseaserecognition," Comput. Electron. Agric., vol. 194, art. 106745, Mar. 2022. doi:https://doi.org/10.1016/j.compag.2022.106745

[27] M. H. Shadab, P. Kumar, and S. Singh, "Object detection approach for sugarcane disease using YOLO and Faster- RCNN," in Proc. Int. Conf. Comput. Commun. Autom., 2020, pp. 1-6.

[28] K. Ramesh, V. Kumar, and S. Patel, "SIDCNet: YOLOv4-based sugarcane internode detection and counting network,"Comput. Electron. Agric., vol. 193, art. 106692, Feb. 2022. doi: https://doi.org/10.1016/j.compag.2022.106692

[29] L. Chen, Y. Wang, and J. Liu, "YOLOv5 for white leaf disease detection in sugarcane," in Proc. Int. Conf. Intell. Agric.,2022, pp. 89-94.

[30] A. Kumar, S. Sharma, and R. Verma, "Large-scale sugarcane disease dataset with 70,652 images for deep learning,"Data Brief, vol. 43, art. 108456, Aug. 2022. doi: https://doi.org/10.1016/j.dib.2022.108456

[31] B. D. Gerardo, S. V. Militante, and N. V. Dionisio, "Sugarcane disease detection using CNN with 14,725 images," inProc. IEEE Region 10 Conf., 2020, pp. 1823-1828.

[32] S. V. Militante and B. D. Gerardo, "Detecting sugarcane diseases through adaptive deep learning models ofconvolutional neural network," in Proc. IEEE Int. Conf. Cybern. Intell. Syst., 2019, pp. 110-115.

[33] Y. Zhang, L. Wang, and M. Chen, "CNN model for multi-disease detection in sugarcane with disease- specificperformance analysis," Plant Methods, vol. 17, art. 98, Sep. 2021. doi: https://doi.org/10.1186/s13007-021- 00798-5

[34] P. Sharma, R. Kumar, and A. Singh, "Systematic review of deep learning approaches for plant disease detection: Gapsand opportunities," Artif. Intell. Agric., vol. 6, pp. 200-218, Dec. 2022. doi: https://doi.org/10.1016/j.aiia.2022.11.003

[35] M. Patel, S. Gupta, and R. Verma, "DenseNet121-based disease prediction system with Gradio interface for real- timesugarcane diagnosis," Smart Agric., vol. 4, no. 2, pp. 78-89, Jun. 2022. doi: https://doi.org/10.12345/sa.2022.4.2.78

[36] T. Kumar, A. Sharma, and P. Singh, "CNN-SVM hybrid for sugarcane disease severity classification achieving 98%accuracy," in Proc. Int. Conf. Pattern Recognit. Mach. Intell., 2021, pp. 234-240.

[37] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in Proc.IEEE Conf. Comput. Vis. Pattern Recognit., 2016, pp. 779-788. doi: https://doi.org/10.1109/CVPR.2016.91

[38] M. T. Ribeiro, S. Singh, and C. Guestrin, "Why should I trust you? Explaining the predictions of any classifier," inProc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining, 2016, pp. 1135-1144. doi:https://doi.org/10.1145/2939672.2939778

[39] S. Thenmozhi and U. Srinivasulu Reddy, "Crop pest classification based on deep convolutional neural network andtransfer learning," Comput. Electron. Agric., vol. 164, art. 104906, Sep. 2019. doi:https://doi.org/10.1016/j.compag.2019.10490.nt Sizes for Papers