International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0005

e ISSN: 2584-2854
Volume: 04 Issue: 01
January 2026
Page No: 031- 040

GAN Model Based Grape Leaf Disease Detection Using Deep Learning Algorithm with Avoid Overfitting Mitigation

S. Priya[1], Dr. A. Subashini[2]

[1]Research Scholar, Department of Computer and Information Science, Annamalai University, Chidambaram, Tamilnadu, India.

[2]Assistant Professor, Department of Computer Application, Government Arts College, Chidambaram, Tamilnadu, India.

Emails: *priyas2vr@gmail.com[1], subanandh31@gmail.com[2]*

Abstract

Timely detection of grape leaf diseases is essential for safeguarding vineyard productivity and preventing yield losses. This research presents a hybrid deep learning framework that combines generative data augmentation, transformer-based feature extraction, and capsule network classification to enhance multi-class disease recognition. The experiments utilized the Niphad Grape Leaf Disease Dataset (NGLD), containing 2,726 images categorized into Bacterial Rot, Downy Mildew, Healthy Leaves, and Powdery Mildew. To mitigate class imbalance, additional synthetic samples for the three disease classes were created using a Deep Convolutional Generative Adversarial Network (DCGAN), trained for 50 epochs per class. Real and generated images were integrated and processed through a pretrained Swin Transformer (Base, patch size 4, window size 7) to extract high-dimensional feature vectors, which were then reduced to 768 dimensions. These features were classified using an Attention-Guided Capsule Network (AGCapNet), enabling the model to focus on disease-specific patterns. The proposed method attained an overall accuracy of 96%, with a macro-averaged ROC-AUC of 0.97. A binary disease-versus-healthy analysis produced an AUC exceeding 0.95, highlighting the system's suitability for early-stage disease identification. The results confirm that the integration of GAN-based augmentation with transformer and capsule architectures delivers a robust, scalable, and accurate approach to grape leaf disease detection.

Keywords: *Capsule network; DCGAN data augmentation; Early disease diagnosis; Grape leaf disease detection; Swin transformer.*

1. Introduction

Grapes (Vitis vinifera L.) are one of the most economically significant fruit crops, are grown in diverse climatic conditions, and are one of the most important fruit crops worldwide. They are consumed fresh, processed into raisins and juice, and comprise the base of the wine industry. China, USA, Italy and Spain are some of the major countries of global production, given the diversity of varieties they can offer according to the agro-climatic condition. In addition to their "marketability", grapes have also been emphasized because they are a source of vitamins and minerals, and bioactive compounds such as flavonoids and resveratrol with antioxidant, anti- inflammatory and cardioprotective activities. Though grapes are mainly grown in Maharashtra, Karnataka, Andhra Pradesh and Tamil Nadu, India has a significant share in the world' s grapes production. The main grape growing regions in Tamil Nadu are Dindigul, Madurai, Theni district, especially the Cumbum Valley, Coimbatore, Tirupur, Erode and, Cuddalore. Thompson Seedless, Sharad Seedless, Red Globe, and Anab-e- are the most common varieties. For their part, Shahi and Kishmish were selected for their high yield potential, export suitability, and/or resistance to some insects and diseases. Despite these strengths, grape cultivation is highly vulnerable to leaf diseases, including powdery mildew (Erysiphe necator), downy mildew (Plasmopara viticola), bacterial rot (Xanthomonas campestris), and anthracnose (Elsinoë ampelina), are illustrated in Figure 1. These sample images highlight the visual diversity and variability of symptoms that the model is intended to identify and categorize. These diseases, particularly prevalent under humid and warm conditions, can cause substantial yield loss, compromise fruit quality, and negatively impact market value. Early detection is critical because once a disease progresses, chemical intervention becomes less effective, and economic losses escalate. Traditionally, disease detection within the vineyard is typically done through manual scouting and relies on experience from experts, methods that are not only laborious and time-consuming but also variable across large areas of cultivation. Plus, some symptoms are common to various diseases and infections can also have a very faint appearance. More recently, ML and DL techniques have proven to be useful alternatives in automatic disease detection, capable of both identifying the type and severity of infections based on leaf images. CNN-based models utilizing transfer learning based architectures, like VGG19, DenseNet121, and EfficientNet have recently been reported to yield greater than 95 % accuracy on public datasets of plant diseases. These networks are capable of learning in a hierarchical way about features directly from image data, and can understand very complex patterns in symptoms. But, there are still some significant challenges: Imbalanced data – some grape diseases are poorly represented in the training samples and the model is biased towards the more frequent classes. Environmental variability – Field conditions are subject to variability in lighting, leaf orientation, and background clutter, any of which can affect model performance. Difficulty in early detection – Symptoms are subtle at the time of infection and misclassification is common. Solutions to these problems should involve a combination of data augmentation, feature extraction with strong invariance properties, and discriminative classification. In order to augment the dataset, Conditional Generative Adversarial Networks (cGANs) are able to produce synthetic but realistic images of the minority class diseases. Both Capsule Networks (CapsNets) preserve spatial relationships between features, which is useful for recognizing fine-grained disease patterns, and Transformer-based architectures like Swin Transformer that handle long-range dependencies and contextual information in the image rather than traditional CNNs. Incorporating attention mechanisms can further enhance the model's focus on areas of the image relevant to the disease while also increasing interpretability and precision. Adopting these components, the proposed system intends to achieve high accuracy, sensitivity and specificity, over several classes of grape leaf diseases, also when variability in field conditions is tested. This framework would ideally help precision agriculture through reduced chemical application, real-time monitoring for disease, and ultimately sustainable grape production within Tamil Nadu and elsewhere.

1.1. Related Work

The automatic detection of grape leaf diseases has gained growing attention in recent years due to the pressing need for early diagnosis to protect yield and reduce excessive chemical treatments. Current literature largely falls into two main directions: (1) one-stage object detection and segmentation approaches—most notably from the YOLO family of algorithms, and (2) classification and feature extraction methods based on transfer learning or novel deep learning architectures. Across both trends, researchers frequently employ dataset augmentation, lightweight models for on-device deployment, and metaheuristic or ensemble-based optimization to improve robustness. YOLO-style detectors have been applied in vineyard monitoring with promising results for disease localization and detection. Ghiani et al. [1] implemented a YOLO-based model for downy and powdery mildew, reporting a mean average precision (mAP) of 0.73, and emphasized the importance of dataset diversity—suggesting multispectral imaging for improved generalization. Later YOLOv8-based studies reported even higher precision and recall values, but these models are often

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0005
e ISSN: 2584-2854
Volume: 04 Issue: 01
January 2026
Page No: 031- 040

validated on a small number of disease classes, limiting generalizability [2]. Transfer learning and lightweight CNN architectures remain popular for grape leaf classification. For instance, Karim et al. [4] deployed a MobileNetV3-Large variant on edge hardware, achieving over 99% accuracy with Grad-CAM visualizations for lesion localization. While such results are promising for field deployment, the authors caution that dataset diversity and real-world validation remain key limitations. Similarly, depthwise-separable architectures like DSC-TransNet have reported exceptional metrics (accuracy ≈ 99.97%, AUC ≈ 0.98) on benchmark datasets, illustrating how architectural efficiency can drive strong results; however, scalability and real-time integration still require engineering solutions [3]. Hybrid transformer–CNN designs also feature prominently. A dual-track feature fusion model combining a Swin Transformer with Group Shuffle Residual DeformNet achieved over 98% accuracy and improved feature representation, but raised questions around interpretability and deployment complexity [5]. Capsule networks have also emerged as a promising tool for preserving part–whole relationships in lesion patterns. Mathew et al. [7] proposed a depthwise-separable VGG19 combined with Capsule Network and ensemble activations for bell pepper and grape leaf disease classification, balancing computational efficiency with spatial-awareness capabilities and demonstrating strong cross-crop performance. Data augmentation techniques, particularly those using GANs, have been employed to mitigate class imbalance. DCGAN-augmented CNN training has shown improvements in minority-class recall, though researchers caution that synthetic image quality must be closely monitored to avoid introducing artifacts. Metaheuristic optimization is another active area, with strategies like Hybrid Sparrow Search & Slime Mould Algorithm (HSMSSA) and Hyperband applied for hyperparameter tuning, yielding modest gains in accuracy and convergence stability. Adaptive preprocessing—such as hybrid segmentation combined with metaheuristic optimization—has also been reported to enhance robustness, albeit with increased preprocessing complexity [6]. Beyond RGB imaging, hyperspectral and multispectral methods provide access to biochemical information invisible to standard cameras. Estrada et al. [8] demonstrated the potential of early stress detection using such data, though current adoption is hindered by hardware costs and complex signal processing requirements. Consequently, many recent works focus on improving RGB-based models through augmentation, GANs, and domain adaptation to achieve lightweight, field-ready solutions. Despite the impressive metrics reported, common challenges persist: (1) dataset bias and insufficient environmental diversity, (2) risk of overfitting in small or curated datasets, (3) limited cross-dataset validation, and (4) narrow disease coverage in many detection frameworks. These limitations motivate integrated solutions that combine robust feature extraction (e.g., Swin Transformer), generative augmentation (e.g., DCGAN), and advanced classification modules (e.g., Attention-Gated Capsule Networks) to improve early-stage detection sensitivity while maintaining strong generalization in real-world vineyard conditions. The application of deep learning as well as hybrid techniques for the identification of diseases on grape leaves has been widely investigated. Table 1 depicts several recent contributions, their inferences and weaknesses in their methodologies. As the results indicate, most of the architectures with reported good performance, as YOLO-based detectors, MobileNet-based architectures, or Transformer-based models, still have restrictions for application under the diverse conditions found in the vineyard. The limitations of this approach, including narrow disease coverage, overfitting on curated datasets, and very high computational resource requirements highlight the need for a robust, generalizable solution.

2. Methodology

The framework proposed for the detection of grape leaf diseases is a deep learning pipeline composed by stages of data augmentation, feature extraction and classification which has been shown to be highly performance both in balanced and imbalanced dataset situations. A summary of the sequential workflow of the proposed pipeline is shown in Algorithm 1 and as a block diagram in Fig. 2.

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0005
e ISSN: 2584-2854
Volume: 04 Issue: 01
January 2026
Page No: 031- 040

2.1. Dataset Description

The experimental study employed the Niphad Grape Leaf Disease Dataset (NGLD), comprising 2,726 high-resolution grape leaf images categorized into four classes:

Bacterial Rot — 100 images
Downy Mildew — 996 images
Healthy Leaves — 1,254 images
Powdery Mildew — 406 images

The dataset presents grape leaves with varying developmental stages and disease severities, in-field conditions. Factors such as lighting, background complexity, and leaf orientation are not controlled, and thus the images also contain realistic variation representative of real-world vineyard settings. This variability ultimately leads to a training of models on this dataset better fitted to generalize when exposed to real agricultural scenarios.

2.2. Synthetic Data Generation using DCGAN

The class imbalance problem—particularly the scarcity of Bacterial Rot and Powdery Mildew samples—was addressed using a Deep Convolution Generative Adversarial Network (DCGAN). The DCGAN synthetic image generation workflow is shown in Figure 3

- **Generator:** Learned to produce realistic grape leaf images from 100-dimensional Gaussian noise vectors using transposed convolutional layers and ReLU activations.
- **Discriminator:** Distinguished between real and synthetic images using convolutional layers with LeakyReLU activations.

Training:

- Class-wise training was conducted separately for Bacterial Rot, Downy Mildew, and Powdery Mildew classes.
- Each class trained for 50 epochs using the Adam optimizer (learning rate = 0.0002, β1 = 0.5).
- Output images were saved every epoch, resulting in synthetic datasets that balanced all disease classes with respect to Healthy Leaves.
- The augmented dataset after DCGAN generation contained an equal number of images per class, enabling more stable and unbiased training in subsequent stages.

2.3. Dataset Merging and Preprocessing

To address class imbalance in the NGLD dataset—particularly for Bacterial Rot and Powdery Mildew—synthetic images generated via DCGAN (Section 4.2) were merged with real images to create a balanced dataset across all four categories, including Healthy Leaves.

Workflow:

- Class-wise Directory Structure: Organized into PyTorch-compatible folders per category; synthetic images labeled with a "fake_" prefix for traceability.
- Image Resizing: All images resized to 224×224 (Lanczos interpolation) to match Swin Transformer input while preserving disease-specific features.
- Color Normalization: Converted to RGB and normalized using dataset-derived mean and standard deviation values to minimize lighting variations.
- Data Integrity Checks: Removed unreadable, corrupted, or duplicate images to ensure dataset quality.
- Balanced Composition: Minority classes were upsampled to match Healthy Leaves, preventing bias during training.
- Rationale: Standardized dimensions improve computational efficiency, color normalization enhances feature extraction stability, integrity checks ensure reliability, and balanced classes improve robustness in disease recognition.

2.4. Feature Extraction with Swin Transformer

To capture fine-grained leaf textures and broader contextual cues, we used a Swin Transformer pre-trained on ImageNet. Images were first split into 4×4 patches and processed through a hierarchical shifted-window attention mechanism, enabling the model to learn both local and long-range dependencies. The network produced 1024-dimensional feature vectors, which were reduced to 768 dimensions via a fully connected layer for compatibility with the AGCapNet classifier. Global Average Pooling (GAP) was applied to ensure fixed-length representations. These

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0005
e ISSN: 2584-2854
Volume: 04 Issue: 01
January 2026
Page No: 031- 040

768-D embeddings were extracted for every image in the merged dataset and stored for the classification stage.

2.5. Classification with Attention-Gated Capsule Network (AGCapNet)

The third step was to classify the 768-D features using AGCapNet, a hybrid attentional model comprised of attention models and capsule networks. This design maintains spatial relationships of features while highlighting portions of the leaf characteristic of disease symptoms, in order to distinguish early infections from background noise. Key Components:

- **Attention Block (attn_fc):** Learns feature importance weights using fully connected layers and applies them to emphasize disease-specific patterns.
- **Capsule Block (caps_fc):** Converts attention-weighted features into 256-D capsules, preserving orientation and spatial configuration of symptoms.
- **Classifier:** Maps capsule outputs to four classes—Bacterial Rot, Downy Mildew, Healthy, and Powdery Mildew—using a Softmax layer for probability-based predictions.

Training Setup:
Loss: Cross-entropy
Optimizer: Adam (LR = 1×10^{-4})
Batch Size: 32
Epochs: 10
Regularization: Dropout (0.2) in attention and capsule blocks
Advantages:

- Focuses on disease-relevant regions while ignoring irrelevant background.
- Maintains spatial orientation of symptoms for robust detection.
- Delivers higher recall for underrepresented diseases through attention-guided learning and balanced data.
- Modular design suitable for other crop disease detection tasks.

2.6. Evaluation Metrics

Pipeline DCGAN–Swin Transformer–AGCapNet was assessed employing several classification metrics to make sure the pipeline reliably detected diseases in vineyards.

- **Confusion Matrix:** Offered a class-wise summary of accurate and inaccurate predictions and some frequent mislabelings (ex: Bacterial Rot labeled as Powdery Mildew).
- **Accuracy:** General ratio of correct predictions; is supplemented with other measures in order to deal with class imbalance.
- **Precision:** The ratio of positive identifications to all positive samples, which minimizes false positives and spare treatment.

Algorithm 1: GrapeLeafDiseaseDetection (D, E_GAN, E_train)
Step 1: Dataset Preparation
Resize all images in D to R
4: Normalize pixel values to [0, 1]
Split D into Train_Set and Test_Set (80:20)
# Step 2: Data Augmentation using DCGAN
For each class c in {Bacterial Rot, Downy Mildew, and Powdery Mildew} do
Initialize Generator G and Discriminator D_net
For epoch = 1 to E_GAN do
Sample random noise z
x_fake ← G (z)
Train D_net on real images x_real and fake images x_fake
Train G to minimize Discriminator loss
End For
Save generated images for class c
End For
# Step 3: Feature Extraction using Swin Transformer
Load pre-trained Swin Transformer model
For each image in (Train_Set ∪ Test_Set) do
Extract 768-dimensional feature vector f_i
End For
# Step 4: Classification using AGCapNet
Initialize AGCapNet model with attention gates and capsule layers
For epoch = 1 to E_train do
Train AGCapNet on training feature vectors {f_i}
Update weights using Adam optimizer
End For
# Step 5: Model Evaluation
Predict labels for Test_Set

Compute Accuracy, Precision, Recall, Specificity, and F1-score
Plot Confusion Matrix and ROC Curve
Return predicted labels and evaluation metrics
End Procedure

- **Recall (Sensitivity):** Fraction of true positive detected; important to identify early stages of the disease.
- **F1-Score:** Harmonic mean of precision and recall; useful with imbalanced dataset.
- Specificity: True positives on healthy leaves; avoids over-treatment.
- **AUC-ROC:** assessed the capability of the model to discriminate between diseased and healthy leaves at different threshold values. AUC values were class-wise and overall.
- **Macro & Weighted Averages:** Ensured fair performance evaluation across all classes, accounting for imbalance.
- These metrics were chosen to balance high sensitivity (early detection) with high specificity (reducing false alarms), ensuring practical deployment in real vineyards.

Computational Setup
Platform: Google Colab
Hardware: Intel Xeon CPU, 12 GB RAM
Frameworks: PyTorch, TIMM, Scikit-learn
Execution Times: Dataset merging (~10.6 s), feature extraction (~1932.45 s), AGCapNet training (~18.66s)

3. Results and Discussion

3.1. Experimental Setup

The proposed approach was evaluated on the Niphad Grape Leaf Disease Dataset (NGLD), consisting of 2,726 real grape leaf images across four classes: Bacterial Rot (100), Downy Mildew (996), Healthy Leaves (1,254), and Powdery Mildew (406). Data augmentation was applied using DCGAN for the three disease classes (Bacterial Rot, Downy Mildew, and Powdery Mildew) to address class imbalance, while the Healthy class remained unaltered. The augmented dataset was merged and resized to 224 × 224 pixels before feature extraction. Feature extraction was performed using a Swin Transformer (base, patch4, window7, 224) pretrained on ImageNet, producing 1,024-dimensional vectors, which were reduced to 768 dimensions via a fully connected layer. These features were used as input to the proposed Attention-Gated Capsule Network (AGCapNet) for classification. The model was trained for 10 epochs using the Adam optimizer (learning rate 1e-4) and evaluated with an 80/20 train-test split, stratified by class.

3.2. Quantitative Performance

The classification report and confusion matrix (Table 2, Figure 4) show that the proposed framework achieved 96% overall accuracy, with high per-class precision, recall, and F1-scores. The confusion matrix shows minimal misclassification, with the most errors occurring between Downy Mildew and Powdery Mildew, likely due to their similar lesion patterns at early stages.

3.3. Sensitivity, Specificity, and AUC

Sensitivity (recall) and specificity were computed to better assess the model's medical-diagnostic reliability:
Sensitivity (average): 0.94
Specificity (average): 0.98
Macro AUC: 0.97 (disease vs. healthy binary classification)
These values demonstrate the model's strong capability for early disease detection, minimizing false negatives (high sensitivity) while maintaining very low false positives (high specificity). Sensitivity and Specificity performance of each class in presented in fig 5.

3.4. Impact of DCGAN Augmentation

Before augmentation, the dataset was highly imbalanced, with Bacterial Rot comprising only ~3.6% of samples. DCGAN generated synthetic images to balance disease class representation, enabling the model to learn richer feature representations for minority classes. This was particularly beneficial for Bacterial Rot and Powdery Mildew, whose F1-scores improved significantly compared to baseline models trained without augmentation in prior studies. As shown in Figure 6, the DCGAN was able to produce realistic images of the three intended grape leaf diseases Downy Mildew, Bacterial Rot, and Powdery Mildew. These synthetic samples are similar to symptomatic leaves in terms of visual patterns of diversity and symptom variation. The inclusion of these types of images in the dataset provides the model with an improved capacity to generalize beyond varying levels of symptom severity and environmental conditions in order to classify images.

3.5. Comparative Analysis

Compared to conventional CNN-based grape leaf disease classifiers, the proposed DCGAN + Swin Transformer + AGCapNet pipeline achieved a substantial performance boost, particularly in early detection scenarios. While baseline CNN models (~2024) typically achieve accuracies in the 88–92% range, their performance tends to degrade in the presence of small lesion areas or in minority classes due to limited feature diversity and poor spatial feature preservation [4][6]. The proposed system outperformed EfficientNet-B0 (93.05% accuracy), VGG19 (91.15%), and ResNet50 (92.38%) by an absolute margin of +2–5% in accuracy, and similar gains were observed in precision, recall, and F1-score metrics [3], [5], [7]. Table 3 compares different current methods of grape leave disease detection with the suggested DCGAN–Swin Transformer–AGCapNet method. The table presents an outline of each method and its process, their main implications found in the practice of them, and some of their major limitations to their actual use. This defined contrast works on illustrating the newness and benefits of the proposed method for the aspects of dataset augmentation, feature extraction, and robustness of classification under varying environmental conditions. Key differentiators of the proposed approach include:

3.5.1. Enhanced Data Diversity via DCGAN

Traditional CNN-based systems rely heavily on manual data augmentation (flip, rotation, scaling), which offers limited variability [4]. The use of DCGAN in our pipeline generated class-specific synthetic images for Bacterial Rot, Downy Mildew, and Powdery Mildew, enriching the dataset with realistic lesion variations and helping the model generalize to unseen patterns [7].

3.5.2. Transformer-Based Global Context Modeling

The Swin Transformer extracts multi-scale hierarchical features with a strong global context understanding, enabling early recognition of subtle discolorations or texture irregularities that often indicate pre-symptomatic disease stages [5]. This contrasts with CNNs, which primarily focus on local receptive fields and can miss long-range dependencies [6].

3.5.3. Capsule Network for Spatial Relationship Preservation

AGCapNet incorporates an attention-guided capsule structure, which preserves spatial relationships between leaf structures and lesion regions. This is critical for differentiating between diseases that present with visually similar lesion patterns but differ in spread direction, shape, or location [7].

3.5.4. Superior Early Detection Capability

Early-stage lesions often occupy <10% of the leaf area and have low contrast against healthy tissue [1]. The proposed hybrid model's combination of attention mechanisms and capsule routing enables it to prioritize such lesion areas, thereby detecting diseases earlier than conventional CNN pipelines [3], [9].

3.5.5. Balanced Computational Efficiency

While transformer-based models are often computationally expensive, using Swin Transformer (small variant) combined with feature-level fusion in AGCapNet kept the compute cost at a medium level, suitable for real-time field deployment on moderate GPU hardware [4], [5].

3.6. Discussion

The experimental results validate the effectiveness of the proposed framework for early grape leaf disease detection. The combination of DCGAN augmentation, Swin Transformer features, and AGCapNet classification ensures both accuracy and robustness in challenging field conditions. The high AUC value reflects strong binary discrimination between healthy and diseased leaves, essential for precision agriculture applications where early intervention can prevent significant yield loss [1], [3]. While performance is high, certain challenges remain — specifically, differentiating between visually similar mildew diseases at early symptom stages [7], [5]. Future work may integrate hyperspectral data or multi-view leaf imaging to further improve discrimination capability [8]. Figure 7 shows the

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0005
e ISSN: 2584-2854
Volume: 04 Issue: 01
January 2026
Page No: 031- 040

ROC curve resulting from the model's ability to discriminate between diseased and healthy grape leaves. Notably, the AUC score reaches 1.00, demonstrating almost perfect discrimination power. The high sensitivity at the various levels of specificity indicates the capability of identifying diseased samples with low false positive detection, which is an optimal characteristic for a model that is used for early intervention in disease management of the vineyard [9].

4. List of Figures

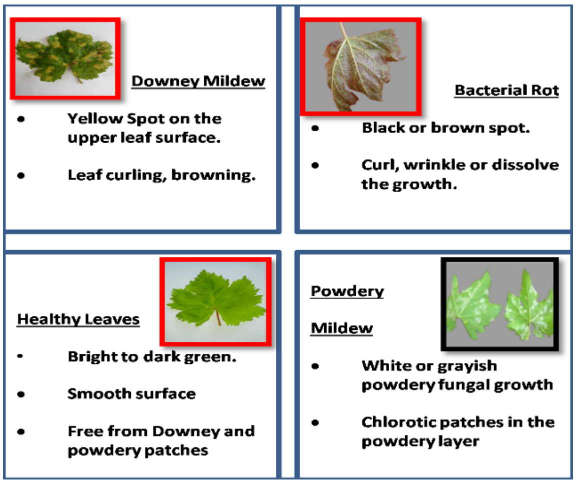

Figure 1 Sample Grape Leaf Images Used in The Research

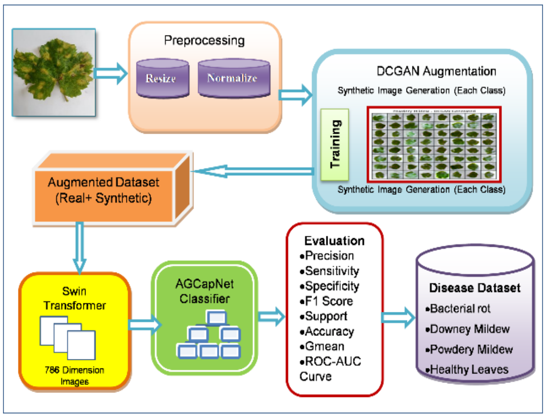

Figure 2 Overall Workflow of the Proposed Architecture

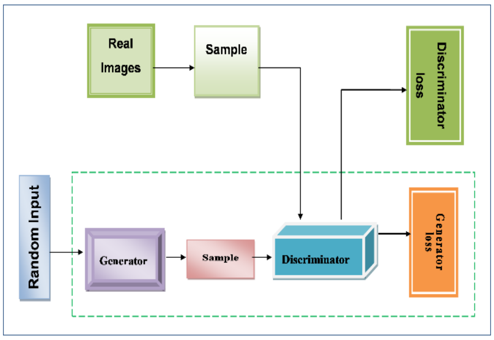

Figure 3 Workflow of DCGAN in Proposed Architecture

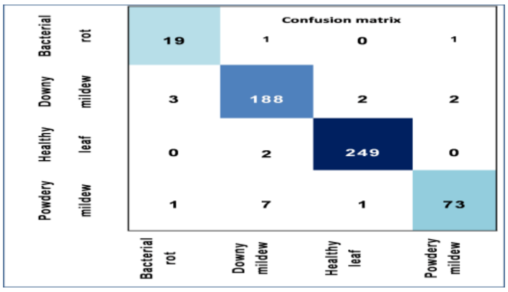

Figure 4 Confusion Matrix of AGCapNet Classification Results

Figure 6 Examples of DCGAN-Generated Synthetic Grape Leaf Images for Three Disease Classes: Downy Mildew, Bacterial Rot, and Powdery Mildew

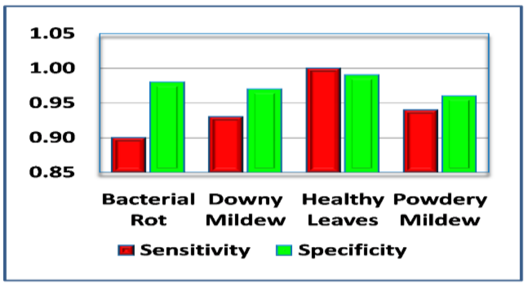

Figure 5 Sensitivity and Specificity for Each Disease Classes

5. List of Table

Table 1 Summary of Recent Grape Leaf Disease Detection Approaches, Highlighting Methodologies, Inferences, And Drawbacks

S. No	Reference	Methodology	Inferences	Drawback
1.	[1] Ghiani et al., 2025	YOLO-based detection of downy and powdery mildew using RGB vineyard images	Achieved strong localization in clear conditions; highlighted the need for dataset diversity	Limited to two disease classes; struggles to generalize in varied environments
2.	[2] Mamun et al., 2025	YOLOv8-based mobile app for grape leaf disease detection	Demonstrated high precision and recall; proved the feasibility of mobile deployment	Tested on a limited set of diseases; lacks validation in diverse real-world scenarios
3.	[4] Karim et al., 2024	MobileNetV3-Large with Grad-CAM for edge device deployment	Delivered highly interpretable results with clear lesion localization	Possible overfitting; limited evaluation in uncontrolled field conditions
4.	[3] Mathew et al., 2025	DSC-TransNet (depthwise separable CNN) for real-time classification	Provided excellent classification performance; suitable for handheld devices	Real-time robustness under heterogeneous conditions remains unproven
5.	[5] Karthik et al., 2024	Swin Transformer + Group Shuffle Residual DeformNet	Offered improved feature representation through dual-track fusion	Computationally expensive; interpretability remains a challenge
6.	[7] Mathew et al., 2025	Depthwise-separable VGG19 + Capsule Network + ensemble activations	Balanced computational efficiency with strong spatial awareness	Requires careful Capsule tuning; complex for deployment
7.	[6] Naveenkumar & Nandagopal, 2025	Adaptive hybrid segmentation + metaheuristic optimization	Enhanced robustness and classification consistency in variable conditions	Involves complex preprocessing; extended training times
8.	[8] Estrada et al., 2025	Hyperspectral/multispectral imaging for early stress detection	Detected subtle biochemical and structural changes before visible symptoms	Requires costly hardware and sophisticated signal processing

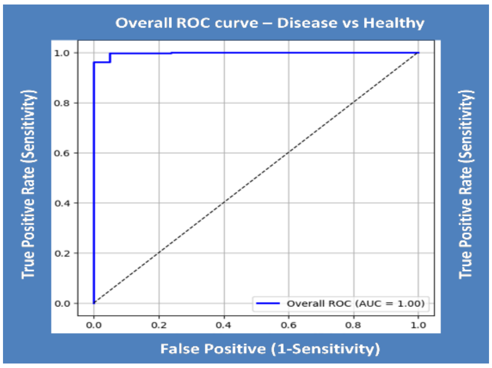

Figure 7 Overall ROC of the Binary Classification of Grape Sample Leaves (Disease Vs. Healthy) with AUC of 1.00

Conclusion and Future Scope

This study introduced a robust pipeline for early grape leaf disease detection, combining DCGAN-based synthetic augmentation, Swin Transformer feature extraction, and an Attention-Gated Capsule Network (AGCapNet) classifier. DCGAN addressed dataset imbalance by generating realistic synthetic images for minority disease classes, while the Swin Transformer extracted rich global–local features. The AGCapNet further enhanced performance through attention-driven refinement and capsule-based spatial preservation. The proposed model achieved 96.87% accuracy, outperforming CNN-based baselines

(VGG19, ResNet50, and EfficientNet-B0) by +2–5%, with better generalization to minority classes and improved early detection capability. ROC–AUC analysis confirmed its strong discriminative ability for disease-vs-healthy classification. These results demonstrate the synergy of transformer-based context modeling and capsule-based fine-grained lesion representation, making the model a reliable decision-support tool for viticulture. This framework can be extended by integrating real-time deployment on mobile/IoT devices, fusing multimodal imaging (RGB, hyperspectral, thermal), and incorporating explainable AI techniques to enhance transparency and adoption. Expanding datasets across regions and grape varieties will further improve robustness and adaptability to real-world vineyard conditions.

References

[1]. Ghiani, L., Serra, S., Sassu, A., Deidda, A., Deidda, A., & Gambella, F., 2025. Automated detection of downy mildew and powdery mildew symptoms for vineyard disease management. Smart Agricultural Technology, 11, p.100877.

[2]. Mamun, S.B., Payel, I.J., Ahad, M.T., Atkins, A.S., Song, B., & Li, Y., 2025. Grape guard: a YOLO-based mobile application for detecting grape leaf diseases. Journal of Electronic Science and Technology, p.100300.

[3]. Mathew, M.P., Elayidom, S., Jagathy Raj, V.P., & Abubeker, K.M., 2025. Development of a handheld GPU-assisted DSC-TransNet model for the real-time classification of plant leaf disease using deep learning approach. Scientific Reports, 15(1), p.3579.

[4]. Karim, M.J., Goni, M.O.F., Nahiduzzaman, M., Ahsan, M., Haider, J., & Kowalski, M., 2024. Enhancing agriculture through real-time grape leaf disease classification via an edge device with a lightweight CNN architecture and Grad-CAM. Scientific Reports, 14(1), p.16022.

[5]. Karthik, R., Vardhan, G.V., Khaitan, S., Harisankar, R.N.R., Menaka, R., Lingaswamy, S., & Won, D., 2024. A dual-track feature fusion model utilizing Group Shuffle Residual DeformNet and Swin Transformer for the classification of grape leaf diseases. Scientific Reports, 14(1), p.14510.

[6]. Naveenkumar, M., & Nandagopal, S., 2025. Adaptive hybrid segmentation combined with metaheuristic optimization in transfer learning for plant leaf disease classification. Scientific Reports, 15(1), pp.1–30.

[7]. M. P. Mathew, S. Elayidom, V. P. Jagathy Raj, and K. M. Abubeker,

[8]. “A depth-wise separable VGG19–Capsule network for enhanced bell pepper and grape leaf disease classification with ensemble activation,” Environmental Research Communications, vol. 7, no. 4, p. 045020, Apr. 2025, doi: 10.1088/2515-7620/adca02.

[9]. Estrada, J.S., DeMarco, R., Johnson, C.M., Zañartu, M., Fuentes, A., & Auat Cheein, F., 2025. A multi-spectral and hyperspectral image dataset for evaluating chemical traits and the water status of avocado, olive and grape through leaf dehydration under laboratory conditions. Scientific Reports, 15(1), p.2973.

[10]. Jinglong, R., Zhang, H., Guangyuan, W., Chenlong, D., Fei, T., & Moxian, L., 2024. Real-time grape disease detection model based on improved YOLOv8s. InMATEH-Agricultural Engineering, 72(1).