Available online at https://www.gyanvihar.org/researchjournals/ctm_journals.php
SGVU International Journal of Convergence of Technology and Management
E-ISSN: 2455-7528
Vol.12 Issue 2 Page No 77-85

Agrovision: Enhanced Grape Leaf Disease recognition with a combined deep learning model from LR images

Rachit Khandelwal [1], Paresh Jain [1]

[1] *Department of Electronics and Communication Engineering, Suresh Gyan Vihar University, Jaipur, India*

Abstract- Effective management of diseases is critical in viticulture to ensure the continuous growth of grape production across the globe. Unfortunately, identification of disorders in grape leaves using either aerial imaging taken by drones or surveillance monitoring on the farm can be challenging due to poor image resolution, complicated background environment, and ambiguous boundaries between healthy and infected leaf areas. This study presents a combined two-stage deep learning architecture intended to boost the accuracy of the diagnosis based on low-quality inputs. In the first stage, a WIN (Wavelet transform and Inception-based Network) Super-Resolution (SR) network capable of recovering fine-grained disease features in low-resolution images is applied. The proposed WIN framework exploits advantages of the discrete wavelet transform domain and employs both global and local cascading strategies. In the second stage, classification of diseases is performed through utilizing the established MobileNetV4 network. The experimental evaluation reveals that the SR pre-processing technique brings remarkable gains in terms of accuracy, with PSNR and SSIM of up to 31.18 and 0.979 respectively, for the proposed SR module. The final performance of the proposed network when tested on the PlantVillage datasets was 99.15% classification accuracy, considerably surpassing well established deep learning baseline models. Consequently, the proposed deep learning model is a practical and effective tool to be applied to intelligent agriculture devices such as unmanned aerial vehicles and hand-held devices, allowing farmers to employ specific measures and prevent losses in crops.

Keywords— Discrete Wavelet Transform, plant leaf diseases (PLDs), super resolution, bicubic interpolation, lightweight architecture

1. INTRODUCTION

Agriculture is a key contributor in a nation's economy due to its ability to create jobs, enhance a country's GDP, and supply important raw commodities. With a growing global population, the agriculture sector is under increasing pressure to develop innovative ways to provide food at an affordable price. However, agricultural practices face several challenges which threaten the success and sustainability of the agricultural sector. One of the key challenges is biotic stresses such as plant leaf diseases (PLDs) which impact the yield and monetary value of the agricultural crop. Crop destruction due to insect infestations and diseases caused by bacteria, fungi, and viruses is the primary source of economic loss in the agricultural sector.

Grapes (Vitis Vinifera) are a very important crop worldwide because of their utilisation in the wine, table grape, and dried fruit industry. Grapes face many PLDs that hinder photosynthesis and result in catastrophic yield deficits[1] particularly when environmental factors are favourable for pathogen spread. As such, it is important that PLDs in grape plants are detected as rapidly and accurately as possible to ensure the health of the crop and the economy.

Over the years, farmers and agronomists have relied on manual visual assessments to diagnose plant illnesses. These methods may be traditional, but they are time-consuming, exhausting, and highly subjective in nature [2]. Manual diagnosis is prone to mistakes caused by the various complexities involved such as poorly defined outlines of symptoms [3], the visual resemblance of different diseases [4], and so on. These factors are often the cause of delayed treatment and spread of infections. Consequently, there is a pressing need to develop automated and specialized diagnostics.

Researchers have adopted computational, and more recently, data-centric methods to overcome the challenges associated with manual diagnosis. Biochemical and remote sensing techniques previously offered alternatives [5], [6],

Correspondence to: Rachit Khandelwal, Department of Electronics and Communication Engineering, Suresh Gyan Vihar University, Jaipur
Corresponding author. E-mail addresses: rachit.23182560@mygyanvihar.com

but they did not have the required scalability for real time use. Deep learning has revolutionized plant pathology in the recent years. With respect to visual data, Convolutional Neural Networks (CNNs) outshine other, more traditional, approaches using machine learning techniques. Unfortunately, the practical use of traditional CNN architectures (i.e. AlexNet, and MobileNet) is hindered by the low resolution (LR) of images often captured by Unmanned Aerial Vehicle (UAV) or surveillance drones due to poor image resolution, complicated background environment, and ambiguous boundaries between healthy and infected leaf areas[7], [8].

Conventional methods of image upscaling often neglect the preservation of details related to the image, resulting in artefacts and loss of detail detracting from the image, which may inhibit correct classification. In contrast, the wavelet transform algorithm presents an improved approach to the representation of images in super-resolution(SR) neural networks. This technique [9] is particularly effective in maintaining fine visual features like edges and textures that are important to human vision.

This study proposes a combined deep learning-based model to empower farming community by detecting and classifying crop diseases from deteriorated images in practical farming environments, enhancing them through real-time super-resolution methods, and improving scalable precision viticulture to minimize crop loss and ensure food security. The proposed model pairs a lightweight WIN based SR model with a Movilenentv4 classification model to provide a resolution enhancement of the images prior to classification, which is most useful in a field environment to enhance the recognition of images.

Following are the key contributions in the proposed study:

1. A novel WIN-based lightweight SR model inspired by the inception model in Discrete Wavelet Transform (DWT) domain, utilizing global and local cascading mechanisms is proposed surpassing performance of various state of the art(SOTA) models, to obtain SR images from LR images.
2. The performance and consistency of the proposed combined deep learning model is extensively evaluated using PlantVillage benchmark dataset.

The rest of the work covers the following aspects. Relevant academic publications regarding the classification of diseases of grape leaves are reviewed in detail in Section 2. Section 3 details the proposed methodology, including the architecture of the WIN model and the design of the combined deep learning framework. In Section 4 the results of the experiments are discussed along with the evaluation of the performance. In Section 5 the key findings are discussed and a comparison of the effectiveness of the model is provided. Section 6 discusses the effectiveness of the proposed model in real-time applicability of smart farming.

II. Literature Review

The identification and classification of plant pathologies have changed a lot with advancements in artificial intelligence. Specifically detecting grape leaf diseases has moved from manual inspection to automated systems using deep learning. However these systems work well only if the input data has high fidelity and the models are designed efficiently. This review looks at the trajectory of progress of research from common deep learning applications to advanced dual-stage super-resolution classification frameworks.

Deep learning approaches have made plant disease detection more efficient, reliable and accurate. Despite this progress Ubbens and Stavness [10] highlight that phenotyping plants is still challenging in uncontrolled environments. Many machine and deep learning techniques have shown positive results but existing models have several limitations[11]. A major limitation is the variation in leaf size and image capture techniques, which often results in low-resolution images[8]. These images lack details necessary for accurate disease identification leading to misclassification[12]. High-quality images are crucial to reveal symptom patterns. However in real-world agricultural settings obtaining such high-fidelity data consistently is often impractical.

Due to the problem of limited resolution, Super Resolution models were introduced to help improve image resolution while retaining the original pixel arrangement thus helping to make the image quality better prior to classification. The study by Subha and Kasturi [13] on disease detection in apple fruit investigated the application of Spectral GAN together with DenseNet. The model exhibited high generalization capabilities and robustness against overfitting by using GAN to synthesize the training data and using DenseNet to extract features. The models similar to the ones proposed by Qiu et al. [14] tried to improve the low-resolution images through adversarial loss and progressive up-sampling. Although they proved to be successful when tested under controlled conditions, the models showed poor results under field testing due to inconsistent inputs. Furthermore, Chudasama & Upla [15] used edge-sensitive features to enhance image quality, whereas Zhou et al. [16] utilized adversarial learning to enable the network to preserve fine structural details.

There is growing research focused on making detection models more deployable on resource-constrained devices. Dheeraj and Chand [17] presented a DL model based on the DenseNet-121 architecture reducing computational overhead by utilizing a pruned and concatenated structure. The shift toward mobile architecture is further established in the works of Z. Wang et al. and Zhou et al. Wang et al. [18] replaced backbone with LCNet and introduced GSConv modules reducing parameters to one-third though this resulted in a slight decrease in mean Average Precision (mAP). Zhou et al. [19] reconstructed the YOLOv11 backbone using a DeShuffleNet module achieving an mAP of 95.2% on a grape leaf disease dataset. However these lightweight strategies

Correspondence to: Rachit Khandelwal, Department of Electronics and Communication Engineering, Suresh Gyan Vihar University, Jaipur
Corresponding author. E-mail addresses: rachit.23182560@mygyanvihar.com

often sacrifice the capacity for modelling -scale semantics and global context. While these models reduce the cost of inference, they may struggle to provide high-quality representations in complex environments where lesions vary from subtle punctuate spots to large-area diffusions.

Transformer-based models and attention mechanisms have introduced a novel paradigm to the field. These approaches attempt to overcome the local-receptive-field limitations of traditional CNNs by leveraging context modelling. Hua et al [20] & Zhang et al. [21] integrated CNNs with transformers achieving high accuracy on rice disease datasets. Despite their accuracy these models are characterized by massive parameter sizes. Researchers like Fu et al. [22] introduced TSCANet, a transformer-based channel attention network to prioritize disease-relevant features and suppress the irrelevant background noise. The model proposed by Catal Reis and Turk [23] utilised transformers with separable convolutions to show promising results for potato PLD. The integration of attention mechanisms serves a dual purpose: it prioritizes disease-relevant features and suppresses irrelevant background noise, such as soil or shadows. This is particularly vital in grape leaf detection, where illumination fluctuations and changing weather conditions can significantly alter the leaf's appearance.

Despite the documented innovations in SR-based enhancement and lightweight classification, a critical gap remains. Most existing models fail to adapt well to the inherent variability of the field, such as inconsistent lighting or occlusions[24]. Furthermore, many enhancement techniques lack built-in mechanisms for attention-based feature prioritization, meaning they might enhance noise just as much as the actual disease symptoms. Models like those proposed by Catal Reis and Turk [23] show promise by combining separable convolutions and transformers but still require improved input clarity to reach peak performance.

The current state of the art suggests that the effective path forward involves a hybrid approach: utilizing SR for initial image enhancement and following this with an effective classification model. The proposed framework combines WIN-based SR enhancement, with MobileNetV4 [25] classification model to diagnose disease with enhanced accuracy in low resolution (LR) images.

III. Methodology

A combined two-stage deep learning model for disease detection in low resolution grape leaf images is proposed in the study as illustrated in Fig 1. This method uses a WIN based SR model in the first phase to enhance features of the low resolution input image (I_{LR}) which are deteriorated due to several factors such as, fog, wind motion blur, camera quality. This phase aims to restore high frequency details, making the image clear and assisting the model to perform better. The second phase involves the MobileNetv4 classification model, which enables precise feature extraction and accurate disease classification, through super resolution (I_{SR}) images constructed in the previous phase. This framework acts as an advanced decision support system as the first stage creates high quality images from degraded ones which enable an excellent disease prediction framework with increased accuracy for weak imaging conditions.

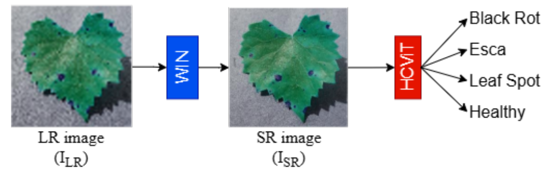


Fig. 1 Proposed combined Architecture for disease detection

A. WIN based Super Resolution

SR with Discrete Wavelet Transform (DWT) is an efficient technique to emphasize the high-frequency component of the image. Applying DWT to the image (I), produce I_{LL} (Average), I_{HL} (Vertical), I_{LH} (Horizontal), and I_{HH} (Diagonal) components [33], as shown in equation 1. In the proposed methodology, the Haar wavelet is used for obtaining DWT.

$$DWT(I) = I_{LL}, I_{HH}, I_{LH}, I_{HH} \qquad (1)$$

The four images from Equation 1 are half the size of the number of rows and columns of the original image I. In the wavelet domain, the SR maps the four components of the LR image to the corresponding four components of the High Resolution (HR) ground truth image. A detailed SR image can effortlessly be created using the wavelet transform since it is an inherently invertible operation.

The input image size in the proposed method is initially up-sampled to match the input image shape and kept constant in subsequent stages. The training speed of the model substantially improves by this. This also enables the model to handle fractional scale factors. However, these approaches also increase noise and deterioration of the up-sampled images compared to methods that adopt post-up-sampling strategies such as pixel convolution or transposed convolution. The novel SR technique utilizes the Cascaded Residual Network CARN [26] model. From the original CARN model dilated convolution blocks are used to extract the patterns and the modification is represented as a modified Cascaded Residual Network (mCARN).

A dilated convolution allows the kernel to sample input features with gaps or dilations between them, thereby increasing the receptive field size without increasing the number of parameters or the computational cost. The dilation rate controls the spacing between kernel elements and determines how far apart the sampled input values are. The key idea behind a dilated convolution block is to stack multiple dilated convolutions with increasing dilation rates to capture information from different scales. By combining dilated convolutions with different dilation rates, the block can capture both local and global contextual information in an image. This network is built on a global and local cascading mechanism and can quickly propagate features

Correspondence to: Rachit Khandelwal, Department of Electronics and Communication Engineering, Suresh Gyan Vihar University, Jaipur
Corresponding author. E-mail addresses: rachit.23182560@mygyanvihar.com

from many paths in different layers. Additionally, the network is free to determine the distribution of the feature information. It utilizes a residual network architecture based on group-wise convolution.

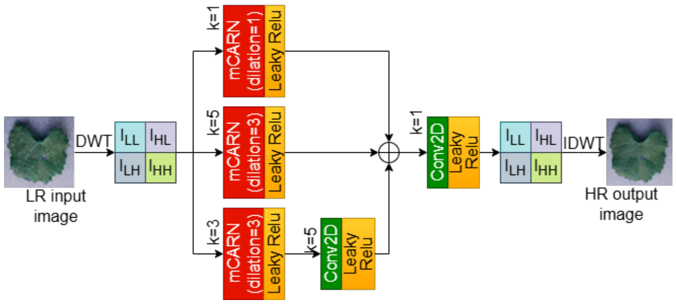


Fig. 1 Architecture of a Wavelet transform and Inception-based Network (WIN)

The multi-path network topology inspired by the Inception network is proposed and demonstrated in Fig.2. The first branch has a dilation rate and kernel size of 1. The second branch has a dilation rate of 3 and a kernel size of 5. The dilation rate and kernel size on the third branch are 3. All the branches of the modified CARN (mCARN) structure are followed by LeakyReLU activation. The mCARN in the last branch has a kernel size of 3, followed by a convolution block of kernel size 5 and LeakyReLU activation. The output of all the branches is then added. Finally, a point-wise convolution with kernel size one and LeakyReLU activation is added to obtain the accurate pattern for super-resolution. IDWT is applied to these features to get the final SR image as the output.

B. Classification of disease based on MobileNetv4

The ability of deep learning algorithms to instinctively learn complex data features makes it suitable for diagnosing leaf diseases under limited resolution requirements. The proposed methodology employs an efficient artificial intelligence model, MobileNetV4, to facilitate field diagnoses effectively. In classifying the disease from the enhanced super-resolution images, the algorithm uses the MobileNetV4 model that incorporates the Universal Inverted Bottleneck (UIB) search block. The model's unified design comprises the inverted bottleneck, convnext, feed forward network, and extra depth-wise (Extra DW) layer designs to achieve efficient feature extraction regardless of the processor used. Moreover, the model contains the Mobile MQA attention block that increases the inference speed by 39% while emphasizing the necessary lesion features and suppressing the noise. With a refined neural architecture search (NAS) recipe and unique distillation strategy, the MobileNetV4 achieves the best trade-off among mobile CPU, DSP, GPU, and edge accelerators.

C. Dataset

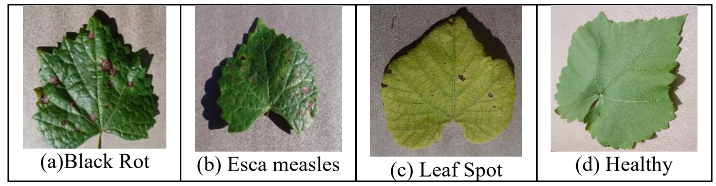


Fig. 2 Multiple classes in Grape PlantVillage Dataset

In this study, the PlantVillage dataset is used as the input data that has gained wide recognition in the classification of diseases in plants [27]. The dataset consists of 54,634 leaf images covering 14 types of plants, including grapes, with 38 distinct conditions of the plants. The grape leaf images in PlantVillage dataset comprises of 4062 image of four classes of healthy and infected leaves. A unique sample of each class is shown in Fig 3. Various classes of grape dataset are described in Fig 4. For the creation of conditions similar to the low-resolution field environment and aerial drone images, bicubic interpolation method is utilized. All images are downscaled to 256×256 pixels in the experiment.

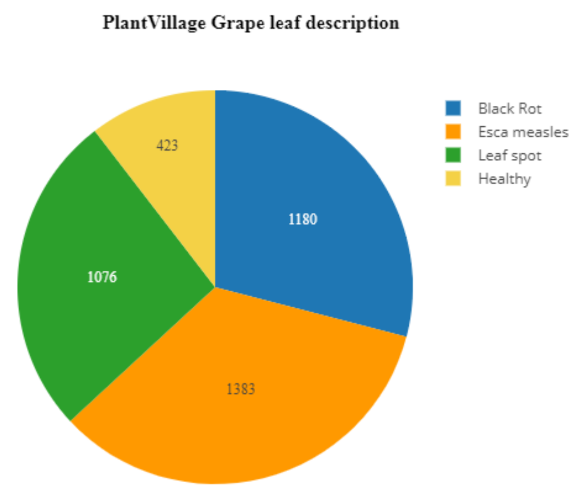


Fig. 4 Grape PlantVillage Dataset

D. Experimental Setup

The experiments assessed the model's accuracy, providing a thorough test of its effectiveness relative to other benchmark models. To ensure consistency during the training and testing processes, the size of the input images was set to 256×256 pixels. The dataset is further divided into 80%, 10%, and 10% portions for training, validating, and testing purposes, respectively. The key hyperparameter settings used in this study for WIN model are discussed Table 1. The loss and accuracy curve for the combined WIN-MobileNetv4 disease detection models is depicted in Fig 5 for scale factor of 2.

Correspondence to: Rachit Khandelwal, Department of Electronics and Communication Engineering, Suresh Gyan Vihar University, Jaipur
Corresponding author. E-mail addresses: rachit.23182560@mygyanvihar.com

TABLE I
HYPERPARAMETER SETTING

Parameter	Value
Number of Epochs	200
Batch Size	14
Optimization Algorithm	Adamax
Learning Rate	0.0001
Loss Function	MAE (Means Absolute Error)

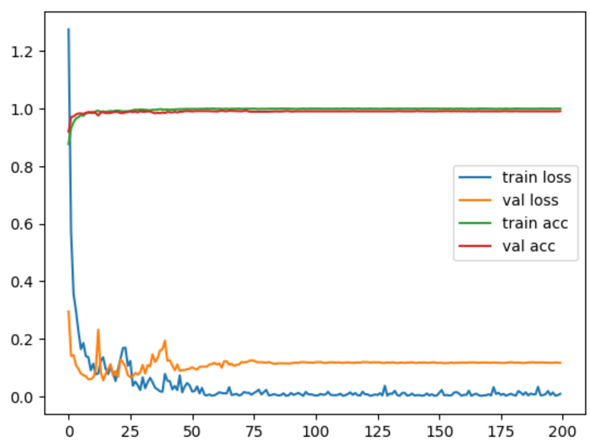


Fig. 5 Loss & accuracy curve for combined WIN-Mobilenetv4 model

IV. RESULTS

In this section, comparison is made with some of the well-established benchmark deep learning algorithms in order to confirm the effectiveness of the proposed model. These experiments illustrate the advantage of proposed dual stage model over existing methods by improving both the input images as well as the disease detection process.

4.1. Performance Evaluation of WIN SR

A key measure of the WIN super-resolution (SR) model's capabilities is the extent to which it can improve image resolution across different scaling factors. In order to effectively appraise the performance of the model, both subjective and objective methods were employed. The subjectivity comes in through the careful observation and scrutiny of the reconstructed grape leaf images across various combinations of the upscaling factors (x2, x4, and x6), to gauge the effectiveness of the model in generating good output. Fig. 6 provides a visual representation of the model's qualitative performance. Here, Grape leaf images from the PlantVillage dataset were utilized as ground truth high-resolution samples while equivalent I_{LR} images were created using bicubic interpolation, which naturally results in edge flattening and the loss of fine texture features.

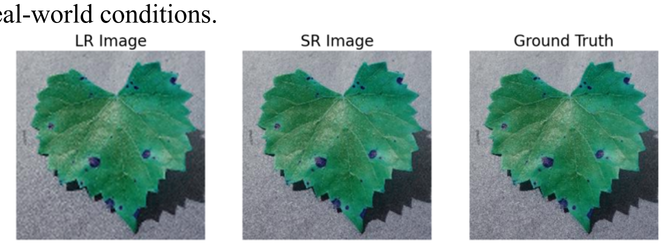
In terms of objectivity, the technique makes use of some commonly accepted measurements such as those used in gauging the quality of image reconstruction and classification tasks. A set of metrics were adopted for the performance evaluation:

- Peak Signal-to-Noise Ratio (PSNR): This metric measures the accuracy of image reconstruction and the size of errors that occurred between the original and generated images.
- Structural Similarity Index Measure (SSIM): It evaluates the structural information, luminance, and contrast of generated images, which is consistent with human perception.

TABLE II: COMPARISON TABLE (PSNR/SSIM) FOR VARIOUS MODELS

Model	No. of parameters (in Millions)	×2 input size (PSNR/ SSIM)	×4 input size (PSNR/ SSIM)	×6 input size (PSNR/ SSIM)
WIN (Proposed SR)	0.5	31.18/ 0.9546	29.39/ 0.8916	27.78/ 0.7784
EWSCNN (2023) [28]	0.15	30.47/ 0.9246	29.14/ 0.8567	27.21/ 0.7716
SRResCGAN (2020) [16]	1.16	29.26/ 0.8923	28.68/ 0.8336	27.06/ 0.7542
CARN (2018) [26]	1.6	28.82/ 0.8927	27.006/ 0.8230	25.21/ 0.7369
DBPN (2018) [29]	4.97	29.14/ 0.8825	28.36/ 0.8308	26.58/ 0.7327
SRCNN (2014) [30]	0.057	26.83/ 0.8231	25.12/ 0.7607	23.74/ 0.7021

Through these techniques, one can effectively determine how well the generated images are compared to the input images across various combinations of the upscaling factors (x2, x4, and x6). Based on the experimental results, the proposed model is superior to the well established SOTA models as depicted in Table 1. The high level of PSNR and SSIM values indicate the model’s ability to successfully reconstruct high-fidelity detailed images in the output from low-resolution (LR) grape leaf images under challenging, real-world conditions.



Correspondence to: Rachit Khandelwal, Department of Electronics and Communication Engineering, Suresh Gyan Vihar University, Jaipur
Corresponding author. E-mail addresses: rachit.23182560@mygyanvihar.com

Fig. 6 Grape leaf image reconstruction

B. Performance Analysis of MobileNetv4

The proposed deep learning framework was tested against several other neural network architectures, including VGG, EANet, DLVTNet across various combinations of the upscaling factors (x2, x4, and x6) as shown in the Table 2. In order to effectively appraise the performance of the model, both subjective and objective methods were employed. The subjective analysis can be done by assessing the classification matrix as shown in Fig 7 for scaling factor of 2. A wide array of performance metrics such as accuracy, precision, recall, F1 score, specificity, and Matthews Correlation Coefficient (MCC) [31] will be of interest to gauge the objective effectiveness of the model across various combinations of the upscaling factors (x2, x4, and x6) for the classification task.

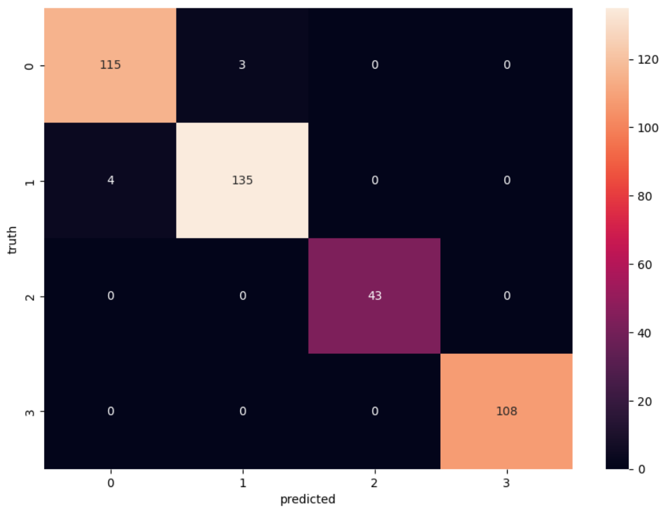


Fig. 7 Confusion matrix of combined model

Based on the experimental results, the proposed model is superior to the well established SOTA models as depicted in Table 2. The high level of classification accuracy indicate the model's ability to accurately differentiate between diseased and healthy grape leaves under challenging, real-world conditions. As a result, the proposed model demonstrates a high level of robustness towards disease detection, making it well-suited for applications in modern agriculture, which relies heavily on accurate disease detection to enhance crop yields and management.

TABLE IIII: COMPARISON OF CLASSIFICATION ACCURACY FOR VARIOUS SCALE FACTORS

Model	No. of parameters (in Millions)	Classification Accuracy for scale factor ×2	Classification Accuracy for scale factor ×4	Classification Accuracy for scale factor ×6
Proposed methodology	2.96	99.15	98.47	98.19
MobileNetv4 (2024) [25]	2.46	96.81	94.83	92.24
DLVTNet (2025) [32]	1.05	96.81	94.02	91.89
Eanet (2022) [33]	6.1	96.32	93.63	92.12
VGG (2015) [34]	138	95.34	94.32	91.42
Inception (2016) [35]	13.6	95.28	92.1	90.82

As shown in Table 3, the proposed approach was benchmarked against various state-of-the-art models using a scaling factor of ×2 and a wide array of performance metrics such as accuracy, precision, recall, F1-score, specificity, and Matthews Correlation Coefficient (MCC)[31].

TABLE IIIV
PERFORMANCE COMPARISON OF THE VARIOUS CLASSIFICATION MODELS

Model	Accuracy (%)	Precision	Recall	F1-score	Specificity	MCC
Proposed methodology	99.15	0.9831	0.9829	0.9865	99.71	0.9989
MobileNetv4 (2024) [25]	96.81	0.9636	0.9738	0.9683	98.92	0.9683
DLVTNet (2025) [32]	96.81	0.9745	0.9733	0.9737	98.85	0.9737
Eanet (2022) [33]	96.32	0.9712	0.9651	0.9679	98.65	0.9679
VGG (2015) [34]	95.34	0.9647	0.9554	0.9584	98.36	0.9584
Inception (2016) [35]	95.28	0.9067	0.9045	0.9214	96.89	0.9578

V. DISCUSSION

The results demonstrate that the proposed combined framework effectively addresses the challenges of precision agriculture by successfully reconstructing high-fidelity details from low-resolution (LR) grape leaf images. By integrating a Wavelet-based Inception Network (WIN) for super-resolution with MobileNetV4-driven classification architecture, the model ensures robust generalization and adaptability across diverse imaging environments. The methodology's reliability was confirmed through rigorous validation on the PlantVillage benchmark dataset.

The experimental data reveals that the super-resolution enhancement significantly elevates the diagnostic performance of the downstream classifier. For instance, the classification accuracy on LR images was initially recorded

Correspondence to: Rachit Khandelwal, Department of Electronics and Communication Engineering, Suresh Gyan Vihar University, Jaipur
Corresponding author. E-mail addresses: rachit.23182560@mygyanvihar.com

at 96.81%, 94.83%, and 92.24% for combinations of the upscaling factors (×2, ×4, and ×6), respectively. Upon applying WIN-generated super-resolution (I_{SR}) images, these figures ascended to 99.15%, 98.47%, and 98.19%. This substantial improvement highlights the capacity of the wavelet-based approach to preserve critical spatial attributes and symptom patterns that are otherwise lost in deteriorated inputs.

Furthermore, the integration of efficient convolution-based embeddings within the classification framework optimized prediction performance without a disproportionate increase in computational overhead. Specifically, the model achieved a performance gain of approximately 2% with the addition of only 0.5 million parameters. Such efficiency confirms that high-accuracy disease detection is attainable even in resource-constrained environments. By mitigating the impact of data scarcity and low hardware specifications, this integrated approach proves to be highly suitable for real-world precision viticulture and scalable agricultural monitoring.

VI. Conclusion

In order to ensure effective management and containment of diseases in addition to maintaining the health and productivity of plants, timely and accurate disease recognition plays a crucial role. An innovative methodology that combines the concept of super-resolution (SR) along with MobileNetV4 classification model is proposed in order to detect grape leaf diseases with enhanced accuracy.

The central component of the suggested framework is the lightweight WIN model, which operates inside the wavelet domain using both local and global cascading mechanisms to accomplish effective and high-quality image reconstruction. The WIN model showed excellent reconstruction ability, with PSNR values of up to 31.18 dB as well as SSIM scores of upto 0.9546. These outcomes validate its ability to provide high-fidelity reconstructed images appropriate for agricultural real-time applications. Through this innovative framework, super-resolution methodology is applied to enhance the quality and resolution of low-resolution images, thus allowing for recovery of lost information. Subsequently, the framework makes use of the superior features in the reconstructed image by the MobileNetV4 network for accurate classification of up to 99.15% accuracy. The entire framework works as an illustrative example of the potential of utilizing image SR as pre-processing along with classification methodology for achieving unprecedented accuracy from degraded images. In conclusion, the combined approach of super-resolution based WIN with MobileNetV4 classification is an effective disease diagnosis solution that can be used as a reliable, scalable, and accurate method in precision agriculture. The application of the proposed approach opens the door towards automating the process of plant disease diagnosis in field conditions.

References

1. G. Chen, Z. Xia, X. Ma, Y. Jiang, and Z. He, "MobileNet-GDR: a lightweight algorithm for grape leaf disease identification based on improved MobileNetV4-small," *Front. Plant Sci.*, vol. 16, p. 1702071, Nov. 2025, doi: 10.3389/fpls.2025.1702071.
2. F. Martinelli *et al.*, "Advanced methods of plant disease detection. A review," *Agron. Sustain. Dev.*, vol. 35, no. 1, pp. 1–25, Jan. 2015, doi: 10.1007/s13593-014-0246-1.
3. J. Zhao *et al.*, "A review of plant leaf disease identification by deep learning algorithms," *Front. Plant Sci.*, vol. 16, p. 1637241, Aug. 2025, doi: 10.3389/fpls.2025.1637241.
4. T. S. Priyadarshana *et al.*, "Crop and landscape heterogeneity increase biodiversity in agricultural landscapes: A global review and meta-analysis," *Ecol. Lett.*, vol. 27, no. 3, p. e14412, Mar. 2024, doi: 10.1111/ele.14412.
5. M. Shoaib, A. Sadeghi-Niaraki, F. Ali, I. Hussain, and S. Khalid, "Leveraging deep learning for plant disease and pest detection: a comprehensive review and future directions," *Front. Plant Sci.*, vol. 16, p. 1538163, Feb. 2025, doi: 10.3389/fpls.2025.1538163.
6. S. Alsubai, A. Almadhor, A. Al Hejaili, and T. R. Gadekallu, "Intelligent smart sensing with ResNet-PCA and hybrid ML–DNN for sustainable and accurate plant disease detection," *Front. Plant Sci.*, vol. 16, p. 1691415, Nov. 2025, doi: 10.3389/fpls.2025.1691415.
7. P. V. Yeswanth, R. Khandelwal, and S. Deivalakshmi, "Super Resolution-Based Leaf Disease Detection in Potato Plant Using Broad Deep Residual Network (BDRN)," *SN Comput. Sci.*, vol. 4, no. 2, p. 112, Dec. 2022, doi: 10.1007/s42979-022-01514-1.
8. P. V. Yeswanth and S. Deivalakshmi, "ASFESRN: bridging the gap in real-time corn leaf disease detection with image super-resolution," *Multimed. Syst.*, vol. 30, no. 4, p. 175, Jun. 2024, doi: 10.1007/s00530-024-01377-x.
9. B. Lowe, H. Salman, and J. Zhan, "GHM Wavelet Transform for Deep Image Super Resolution," 2022, *arXiv*. doi: 10.48550/ARXIV.2204.07862.
10. J. R. Ubbens and I. Stavness, "Deep Plant Phenomics: A Deep Learning Platform for Complex Plant Phenotyping Tasks," *Front. Plant Sci.*, vol. 8, p. 1190, Jul. 2017, doi: 10.3389/fpls.2017.01190.
11. P. V. Yeswanth, R. Khandelwal, and S. Deivalakshmi, "Super Resolution-Based Leaf Disease Detection in Potato Plant Using Broad Deep Residual Network (BDRN)," *SN Comput. Sci.*, vol. 4, no. 2, p. 112, Dec. 2022, doi: 10.1007/s42979-022-01514-1.
12. N. K. Tiwari and S. S. Rajput, "Enhancing Potato Leaf Disease Detection Using Super-Resolution and Multi-path Multi-attention Transformers," *Potato Res.*, Jul. 2025, doi: 10.1007/s11540-025-09890-w.
13. V. Subha and K. Kasturi, "A comparative stratification of Automated Indagation on Apple Disease Using

Correspondence to: Rachit Khandelwal, Department of Electronics and Communication Engineering, Suresh Gyan Vihar University, Jaipur

Corresponding author. E-mail addresses: rachit.23182560@mygyanvihar.com

DeepSpectral Generative Adversarial and Densenet Convolutional Neural Network," in *2024 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI)*, Chennai, India: IEEE, May 2024, pp. 1–6. doi: 10.1109/ACCAI61061.2024.10602298.
14. D. Qiu, Y. Cheng, and X. Wang, "Improved generative adversarial network for retinal image super-resolution," *Comput. Methods Programs Biomed.*, vol. 225, p. 106995, Oct. 2022, doi: 10.1016/j.cmpb.2022.106995.
15. V. Chudasama and K. Upla, "E-ProSRNet: An enhanced progressive single image super-resolution approach," *Comput. Vis. Image Underst.*, vol. 200, p. 103038, Nov. 2020, doi: 10.1016/j.cviu.2020.103038.
16. D. Zhou, R. Duan, L. Zhao, and X. Chai, "Single image super-resolution reconstruction based on multi-scale feature mapping adversarial network," *Signal Process.*, vol. 166, p. 107251, Jan. 2020, doi: 10.1016/j.sigpro.2019.107251.
17. A. Dheeraj and S. Chand, "LWDN: lightweight DenseNet model for plant disease diagnosis," *J. Plant Dis. Prot.*, vol. 131, no. 3, pp. 1043–1059, Jun. 2024, doi: 10.1007/s41348-024-00915-z.
18. J. Wang, C. Qin, B. Hou, Y. Yuan, Y. Zhang, and W. Feng, "LCGSC-YOLO: a lightweight apple leaf diseases detection method based on LCNet and GSConv module under YOLO framework," *Front. Plant Sci.*, vol. 15, p. 1398277, Oct. 2024, doi: 10.3389/fpls.2024.1398277.
19. "Lightweight Grape Leaf Disease Identification for Complex Field Scenes $$-$$ $$-$$ YOLO - Grape Model and Its Boundary - Lesion Coalesce Optimization Policies | Signal, Image and Video Processing | Springer Nature Link." Accessed: May 09, 2026. [Online]. Available: https://link.springer.com/article/10.1007/s11760-025-04976-x
20. H. Yang, X. Deng, H. Shen, Q. Lei, S. Zhang, and N. Liu, "Disease Detection and Identification of Rice Leaf Based on Improved Detection Transformer," *Agriculture*, vol. 13, no. 7, p. 1361, Jul. 2023, doi: 10.3390/agriculture13071361.
21. X. Zhang, H. Dong, J. Yang, Z. Lu, L. Gong, and L. Zhang, "A Multi-Disease Detection Method for Paddy Rice Based on Enhancing Detection Transformer With ConvNeXt - DAM - FFNet Refinement," *J. Phytopathol.*, vol. 173, no. 4, p. e70106, Jul. 2025, doi: 10.1111/jph.70106.
22. Y. Fu, J. Liu, and J. Shi, "TSCA-Net: Transformer based spatial-channel attention segmentation network for medical images," *Comput. Biol. Med.*, vol. 170, p. 107938, Mar. 2024, doi: 10.1016/j.compbiomed.2024.107938.
23. H. Catal Reis and V. Turk, "Potato leaf disease detection with a novel deep learning model based on depthwise separable convolution and transformer networks," *Eng. Appl. Artif. Intell.*, vol. 133, p. 108307, Jul. 2024, doi: 10.1016/j.engappai.2024.108307.
24. Yeswanth et al., "Residual Skip Network-Based Super-Resolution for Leaf Disease Detection of Grape Plant | Circuits, Systems, and Signal Processing." Accessed: Jul. 21, 2024. [Online]. Available: https://link.springer.com/article/10.1007/s00034-023-02430-2
25. D. Qin *et al.*, "MobileNetV4 -- Universal Models for the Mobile Ecosystem," Sep. 29, 2024, *arXiv*: arXiv:2404.10518. doi: 10.48550/arXiv.2404.10518.
26. N. Ahn, B. Kang, and K.-A. Sohn, "Fast, Accurate, and Lightweight Super-Resolution with Cascading Residual Network," in *Computer Vision – ECCV 2018*, vol. 11214, V. Ferrari, M. Hebert, C. Sminchisescu, and Y. Weiss, Eds., in Lecture Notes in Computer Science, vol. 11214. , Cham: Springer International Publishing, 2018, pp. 256–272. doi: 10.1007/978-3-030-01249-6_16.
27. D. P. Hughes and M. Salathe, "An open access repository of images on plant health to enable the development of mobile disease diagnostics," Apr. 12, 2016, *arXiv*: arXiv:1511.08060. doi: 10.48550/arXiv.1511.08060.
28. Yeswanth, P. V., and S. Deivalakshmi., "Extended wavelet sparse convolutional neural network (EWSCNN) for super resolution | Sādhanā." Accessed: Aug. 25, 2025. [Online]. Available: https://link.springer.com/article/10.1007/s12046-023-02108-0
29. M. Haris, G. Shakhnarovich, and N. Ukita, "Deep Back-Projection Networks for Super-Resolution," presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 1664–1673. Accessed: Aug. 25, 2025. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_2018/html/Haris_Deep_Back-Projection_Networks_CVPR_2018_paper.html
30. C. Dong, C. C. Loy, K. He, and X. Tang, "Learning a Deep Convolutional Network for Image Super-Resolution," in *Computer Vision – ECCV 2014*, D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds., Cham: Springer International Publishing, 2014, pp. 184–199. doi: 10.1007/978-3-319-10593-2_13.
31. C. Hong *et al.*, "Analysis of Machine Learning Algorithms for Real-Time Gallbladder Stone Identification from Ultrasound Images in Clinical Decision Support Systems," *Int. J. Comput. Intell. Syst.*, vol. 18, no. 1, p. 73, Apr. 2025, doi: 10.1007/s44196-025-00740-8.
32. N. Zhang, E. Zhang, G. Qi, F. Li, and C. Lv, "Lightweight grape leaf disease recognition method based on transformer framework," *Sci. Rep.*, vol. 15, no. 1, p. 28974, Aug. 2025, doi: 10.1038/s41598-025-13689-7.
33. M.-H. Guo, Z.-N. Liu, T.-J. Mu, and S.-M. Hu, "Beyond Self-Attention: External Attention Using Two Linear

Correspondence to: Rachit Khandelwal, Department of Electronics and Communication Engineering, Suresh Gyan Vihar University, Jaipur
Corresponding author. E-mail addresses: rachit.23182560@mygyanvihar.com

Layers for Visual Tasks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 5, pp. 5436–5447, May 2023, doi: 10.1109/TPAMI.2022.3211006.

34. K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," Apr. 10, 2015, *arXiv*: arXiv:1409.1556. doi: 10.48550/arXiv.1409.1556.
35. C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, "Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning," Aug. 23, 2016, *arXiv*: arXiv:1602.07261. doi: 10.48550/arXiv.1602.07261.

Correspondence to: Rachit Khandelwal, Department of Electronics and Communication Engineering, Suresh Gyan Vihar University, Jaipur
Corresponding author. E-mail addresses: rachit.23182560@mygyanvihar.com