

Resilient Loss Function Designed for the Pyramid Dilated Kronecker Network to Classify Rice Leaf Diseases

Nayana V

Ramaiah Institute of Technology

Evangelin Geetha

`degeetha@msrit.edu`

Ramaiah Institute of Technology

Sowmya B. J

Ramaiah Institute of Technology

S Supreeth

REVA University

Research Article

Keywords: Agriculture, Rice Plant disease, Contra-harmonic mean filter, LadderNet, Deep Kronecker Network

Posted Date: October 23rd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11130489/v1>

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

[Read Full License](#)

Additional Declarations: No competing interests reported.

Resilient Loss Function Designed for the Pyramid Dilated Kronecker Network to Classify Rice Leaf Diseases

Nayana V ¹, Evangelin Geetha D ¹, Sowmya B. J ² and Supreeth S ³

¹ Department of MCA, Ramaiah Institute of Technology, Bangalore, Affiliated to VTU-Belgaum, Karnataka, India;
nayanav@msrit.edu (N.V.); degeetha@msrit.edu (E.G.D.)

² Department of Artificial Intelligence and Data Science, Ramaiah Institute of Technology, Bengaluru 560054, India; sowmyabj@msrit.edu

³ School of Computer Science and Engineering, REVA University, Bengaluru 560064, India;
supreeth1588@gmail.com

Corresponding author: Evangelin Geetha D, contact:degeetha@msrit.edu

Abstract

Rice is a vital staple crop for millions of people in the world and increased productivity is significant for food security. Nevertheless, rice plants are broadly prone to several leaf diseases that severely impact the crop if it is not detected early. Conventional disease identification techniques mainly depend on manual inspection which is labor-intensive, and time-consuming, particularly in large-scale farming. These issues are occurred due to the lack of an accurate and automated systems. To tackle these problems, an effective rice plant leaf disease classification framework is proposed named Harmonic Euclidean Smooth Tamed Loss (HaEST-Loss) for Pyramid Dilated Kronecker Network (PDKN). The input image is pre-processed by using contra-harmonic mean filter. Then, LadderNet is used for segmenting the leaves. After that, feature extraction is performed from segmented plant leaf image. Thereafter, leaf disease classification is conducted by PDKN with HaEST-Loss. Here, HaEST-Loss is designed by the integration of using Harmonic analysis, Tamed Cross Entropy (TCE) Loss, Euclidean loss and smooth surrogate loss for PDKN. The efficiency of HaEST-Loss is analyzed by considering performance metrics, like Accuracy, True Positive Rate (TPR) and True Negative Rate (TNR), which acquired the values of 91.646%, 90.728% and 92.610%.

Keywords: Agriculture, Rice Plant disease, Contra-harmonic mean filter, LadderNet, Deep Kronecker Network.

1. Introduction

Agriculture plays a significant role in people's everyday lives since it provides food as well as raw materials. Plant disease is one of the most challenging problems faced by farmers. Recently, crop diseases become severe and widespread through global weather patterns due to climate change. Therefore, modules that can quickly and readily analyze the crops and diagnose the illness should be designed for preventing further crop loss. Crop yields play a critical role in both national and international economies. In India, the primary source of income is agriculture, in which approximately 70% of Indian people rely on agriculture for survival, and most of them prefer rice as their essential food. Indian economy heavily depends on rice production and its relevant products of agro-industry. The growth of rice crops and potential growth of plants are based on various components namely, soil features, climatic variations, irrigation practices, topographical features and biological factors. Plant disease is the most significant biological factors that lead to greater loss in rice crop production. The two major sources of plant diseases are bacteria and fungi. In India, rice diseases are more harmful, which may cause a 10% to 15% loss in production. Brown spot, bacterial leaf blight, and leaf smut are the common infections, which cause severe crop

damage and deterioration in yields . Moreover, plant diseases are also the reason for the reduction in the quality and quantity of crops, which may impact the overall production of rice in India.

Lack of consistent monitoring of plant growth is the major issue and sometimes, new farmers are not attentive to the diseases and their occurrence period. Although the diseases may occur on any plant, constant monitoring may prevent the spread of disease. In the agriculture domain, plant disease detection is one of the significant research topics. Rice has always been the world's important food crop for global food security as well as socio-economic stability. However, each rice type is liable to various pests and illnesses that may affect productivity and crop quality. Various diseases typically increase during rice farming which leads to substantial financial losses. Moreover, pesticides like Fungicides, Nematicides, and Bactericides are extensively employed and have an influence on plant diseases in agro-ecosystem. Still, the rice is mostly infected by various plant diseases, that may impact the cultivation as well as net income. To conquer these issues and increasing crop production, the diseases should be identified and prevented at early stages. Sustainable farming should be obtained by frequent disease monitoring and should execute continuous plant health monitoring as a policy . The prediction of rice leaf disease is significant for maintaining the quality and quantity of the rice production, as early diagnosis of these diseases helps to ensure healthy plant growth and higher yields. Rice diseases involve bacterial and sheath blight, with signs like texture rice blast, color and form that enables for quick inception and simple infection.

Normally, rice disease detection is performed manually that is ineffectual and consumes more time. Even though the traditional approach of mapping rice disease is simple, it is likely to misapprehend similar diseases that have damaging effects on rice development. Computer-aided recognition model has yet to be largely adopted because of major environmental effects, as well as minimal detection speed. Therefore, designing a model for detecting rice diseases provides quick and precise outcomes. The rice infections are present in leaves, and leaf diagnostics can assist the farmers to decide whether the pesticide should be used or not. Predicting the diseases in rice leaves with the naked eye is a complex task for farmers and experts. Identifying the irregularity in plants with a symptom that is connected to several illnesses is more complex. Machine learning (ML) is employed to effectively choose the significant features and classify the images. Moreover, it is also employed for designing a classifier for detection that is processed by providing samples and learned model is examined to identify new sample. Current developments in computer-aided systems and Deep Learning (DL) domains make it more reliable for detecting the features of various illnesses despite their image background or image capture conditions. Several researches have been previously performed utilizing Convolutional Neural Network (CNN) to predict and recognize rice leaf diseases . DL is an evolving domain in the automatic feature extraction; thus, CNN has become a significant approach for image classification. It may handle big data applications without using the image pre-processing process .

The primary intention of this research is to establish a rice leaf disease classification approach named HaEST-Loss. In this framework, an image is filtered by using contra-harmonic mean filter. After that, the filtered image is segmented using LadderNet. Then, Gray Level Co-occurrence Matrix (GLCM) features, like contrast, Inverse Difference Moment (IDM), correlation, and Weber Local Binary Pattern (WLBP) features are extracted. Finally, leaf disease classification is done using HaEST-Loss that is developed by combining TCE Loss, Euclidean loss and smooth surrogate loss using Harmonic analysis for PDKN. Here, PDKN is combined by Pyramid Dilated Attention

Network (PDAN) and Deep Kronecker Network (DKN). The objective of this research is described as,

- **Proposed HaEST-Loss for leaf disease classification:** A novel module is devised for leaf disease classification using hybrid network with HaEST-Loss. Here, the classification of leaf diseases is accomplished by employing PDKN with HaEST-Loss. Here, HaEST-Loss is designed by integrating TCE Loss, Euclidean loss and smooth surrogate loss through Harmonic analysis for modifying the layers of PDKN.

The arrangement of remaining layout is delineated as: Section 2 depicts the study of prior techniques, Section 3 discusses the novel methodology in an architectural view, Section 4 analyze the results of the developed model and Section 5 concludes research with future scope.

2. Motivation

Detection of leaf disease is one of the significant domains, wherein the rice crop is typically infected by numerous illnesses. Normally, the farmers examine the crop at the later phase which causes more damage. This type of manual examination is subjective, and error-prone. As a result, DL models play a fundamental part in early and accurate rice disease detection and classification. So, the researcher investigates the existing methods for developing a novel framework to efficiently detect and classify the diseases.

2.1 Literature Survey

K. M. *et al.* designed Deep Spectral Generative Adversarial Neural Network (DSGAN²) + Improved Artificial Plant Optimization (IAPO) for detecting plant leaf disease. The designed scheme achieved minimal false rate and assisted the farmers in diagnosing and identifying leaf diseases. However, other rice species and diverse kinds of diseases were not included in the existing databases for improving the trained approaches. A. Chakrabarty, *et al.* developed Bidirectional Encoder Representations from the Transformers for images (BEiT) to detect rice plant disease. This technique resolved critical challenges in agricultural Artificial Intelligence (AI) to enhance early disease diagnosis and treatment. Nevertheless, it failed to utilize more interpretability tools or consider targeted regional research for enhancing the understanding of disease dynamics and diagnostic outcomes. M. J. Jhatial, *et al.* introduced You Only Look Once version 5 (Yolov5) to identify the disease of rice plants. This approach was examined on unseen images and obtained quite a better recognition rate with greater accuracy. Nonetheless, it did not incorporate smart embedded module to help the farmers in the agricultural field to increase crop cultivation. I. Bouacida, *et al.* established Convolutional Neural Network (CNN) to detect and classify plant leaf disease. Even though this module had more robustness and capability to adapt different agricultural states beyond its training, it did not explore advanced techniques to improve the adaptability.

Y. K. Rathore, *et al.* presented LW17 model to identify and classify plant leaf diseases. This method outperformed in providing appropriate detection outcomes with few layers and minimal time. However, the efficiency of designed module was not validated with more parameters to enhance the performance. Y. M. Alsakar, *et al.* developed Support Vector Machine (SVM) for rice plant leaf disease detection. This system acquired highest accuracy and assisted in classifying the

diseases, improving agriculture. Nonetheless, it neglected to combine more feature selection and reduction techniques and DL techniques to enhance rice plant disease classification. A. K. Abasi, *et al.* introduced a Custom CNN model for detecting plant leaf disease. Although the developed system accomplished precise stability as well as low computational complexity, it was unable to include efficient feature extraction models for increasing the recognition rate by minimizing false classification. P. I. R. *et al.* implemented CustomVGG16 to detect plant leaf diseases. This model showed high performance and provided rapid and feasible way of diagnosing the diseases. But it was time-consuming, unreliable, and made them distant to small-scale farmers.

Recent studies have further advanced rice leaf disease detection through intelligent learning frameworks and smart agricultural technologies. Resource-efficient federated learning over the Internet of Agricultural Things (IoAT) has demonstrated the feasibility of distributed disease classification while preserving data privacy and reducing communication overhead [31]. Similarly, machine learning and deep learning techniques have been extensively explored for rice leaf disease classification, highlighting the effectiveness of automated feature learning and disease recognition. Furthermore, contemporary intelligent technologies integrating deep learning, computer vision, and precision agriculture have shown significant potential for early disease diagnosis and crop health monitoring. Despite these advancements, several research challenges remain unresolved. Most existing approaches primarily emphasize classification accuracy while providing limited justification for preprocessing strategies and insufficient validation of intermediate segmentation quality. In addition, only a few studies investigate the combined effect of efficient denoising, accurate leaf segmentation, discriminative handcrafted texture descriptors, and adaptive hybrid loss optimization within a unified framework. Motivated by these limitations, the proposed HaEST-Loss framework integrates contra-harmonic mean filtering for effective noise suppression, LadderNet-based segmentation for accurate leaf region extraction, WLBP and GLCM feature descriptors for discriminative texture representation, and a hybrid loss-driven Pyramid Dilated Kronecker Network (PDKN) to improve the robustness and reliability of rice leaf disease classification.

2.2 Challenges

The limitations of traditional classification methodologies are manifested below.

- In , the developed module obtained maximal efficiency and performed with less amount of time in disease detection. However, it was unable to segment the leaf images precisely in complex backgrounds.
- Even though the module introduced in obtained high accuracy, and enhanced the potential to understand and examine disrupted lesions, it failed to minimize the computational resources.
- An approach established in , acquired better efficiency in classification. Nevertheless, it did not investigate the use of DL crack damage detection models for generalizing disease detection in other crops.
- The developed technique in achieved superior classification accuracy with low computational time and faster convergence speed. Nonetheless, it failed to mitigate the false classification problem.

- Plant diseases pose one of the severe threats to global food security, which compromises agricultural productivity and threatens the livelihoods of more individuals. It may destroy crops, and dislocate food supply chains. Hence, the diseases should be identified and classified early in the farming industry. As a result, an effective model is essential to overcome these complexities.

3. Proposed Harmonic Euclidean Smooth Tamed Loss for Pyramid Dilated Kronecker Network in leaf disease classification

Rice plants are often affected by various diseases, resulting in significant crop losses. Previous approaches failed to manage these issues have been challenging and largely dependent on expert knowledge. DL models provide efficient and advanced agriculture solutions for classifying the diseases automatically. Furthermore, traditional detection modules failed to extract more information from input images owing to the diverse feature extraction methods. Therefore, an efficient HaEST-Loss module is designed for leaf disease classification. Initially, the input rice plant leaf image acquired from a specified database is illustrated in . Then, image pre-processing is conducted using contra-harmonic mean filter. After that, leaf segmentation is carried out using LadderNet . Moreover, feature extraction is carried out, where GLCM features, like contrast, IDM, correlation, and WLBP are extracted. Finally, leaf disease classification is conducted by employing PDKN with HaEST-Loss. Here, PDKN is developed using PDAN and DKN, where its layers are modified by HaEST-Loss. However, HaEST-Loss is designed by the combination of TCE Loss, Euclidean loss and smooth surrogate loss, where these loss functions are concatenated by utilizing Harmonic analysis . Thus, plant leaf diseases are classified into Bacterial Leaf Blight, Brown Spot, Leaf scald, Narrow Brown Leaf Spot, Rice Hispa, Sheath Blight, Leaf Blast and Healthy Rice Leaf. Figure 1 deliberates schematic illustration of HaEST-Loss for leaf disease classification

Input rice plant leaf image



Image Pre-processing

contra-harmonic
mean filter

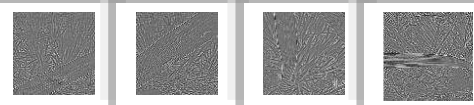


Plant leaf
segmentation

LadderNet

GLCM features,
Weighted Local Binary
Pattern (WLBP)

Feature
extraction



Pyramid Dilated Attention
Network (PDAN)

Deep Kronecker
Network (DKN)

Harmonic Euclidean
Smooth Tamed Loss for
Pyramid Dilated Kronecker
Network (HaEST-Loss)

Rice plant leaf disease
classification

Tamed Cross Entropy (TCE) Loss

Euclidean loss

Smooth surrogate loss

Harmonic analysis

Classified
output

Brown Spot

Leaf Blast

Leaf Scald

Healthy



Figure 1. Schematic view of HaEST-Loss for rice plant leaf disease classification

For reproducibility, the dimensional transformations involved in the proposed framework are explicitly summarized. Each input rice leaf image is resized to $256 \times 256 \times 3$ and processed using a contra-harmonic mean filter, producing a filtered image of identical dimensions. LadderNet subsequently performs leaf segmentation and generates a segmentation mask of size $256 \times 256 \times 1$. The segmented leaf region is then utilized for feature extraction, where WLBP and GLCM descriptors, including contrast, IDM, and correlation, are computed to form the feature vector G_x . The extracted feature vector is provided as input to the Pyramid Dilated Kronecker Network (PDKN), whose parameters are optimized using the proposed HaEST-Loss function. Finally, the output layer generates an eight-class probability vector corresponding to Bacterial Leaf Blight, Brown Spot, Leaf Scald, Narrow Brown Leaf Spot, Rice Hispa, Sheath Blight, Leaf Blast, and Healthy Rice Leaf classes. These dimensional transformations clearly describe the data flow across preprocessing, segmentation, feature extraction, and classification stages, thereby improving the reproducibility of the proposed framework.

3.1 Image Acquisition from Rice Leaf Bacterial and Fungal Disease Dataset

A typical database is considered to classify the diseases present in rice plant leaves [19] that contain eight types of leaf diseases. Also, it contains 1701 images from different locations under different lighting situations and 5188 extra augmented images. To improve the diversity of the training data and enhance the generalization capability of the proposed framework, data augmentation was employed to generate additional training samples from the original rice leaf images. The augmentation process included geometric and intensity-based transformations such as image rotation, horizontal and vertical flipping, scaling, and brightness variation while preserving the disease-specific characteristics of the leaf images. As a result, the original dataset containing 1701 rice leaf images was expanded to 5188 augmented images, thereby reducing overfitting and improving the robustness of the proposed classification model under varying imaging conditions. The total number of images in the dataset is expressed as

$$L = \{L_1, L_2, \dots, L_x, \dots, L_y\} \quad (1)$$

Here, L_x denotes input image, L_y explicates total images.

3.2 Pre-processing using contra-harmonic mean filter

It is the process of transforming and enhancing original images of rice leaves to enhance the efficiency of detection approaches. It includes eliminating noise, adjusting image properties, and normalizing image formats to ensure consistent input for classification techniques. This process is done using contra-harmonic mean filter by considering the input L_x .

3.2.1 Contra-harmonic mean filter

It is employed to remove Gaussian noise while preserving important edge features and can be applied to monochrome, 8-bit per pixel, and 24-bit per pixel images. This filtering process provides a restored image that is computed as,

$$\hat{o}(a,b) = \frac{\sum_{(c,d) \in A_{ab}} e(c,d)^{B+1}}{\sum_{(c,d) \in A_{ab}} e(c,d)^B} \quad (2)$$

Here, B implies an order of the filter, which is suited to diminish or remove the impacts of salt and pepper noise. If the value of B is positive, this filter discards pepper noise, else it removes salt noise. This filter cannot remove both noises at the same time. Therefore, filtered image is symbolized as C_x .

The contra-harmonic mean filter is selected due to its capability to suppress impulse noise while preserving important leaf boundary and lesion characteristics. Compared with the Median Filter, it provides greater flexibility through the filter-order parameter for handling salt-and-pepper noise. Bilateral Filtering preserves edges effectively but increases computational complexity, whereas Non-Local Means and BM3D achieve strong denoising performance at the expense of higher processing time and memory requirements. Since the proposed framework requires an efficient preprocessing stage before segmentation, the contra-harmonic mean filter provides an appropriate balance between noise removal, edge preservation, and computational efficiency, making it suitable for rice leaf image analysis. Although advanced denoising techniques such as the Median Filter, Bilateral Filter, Non-Local Means (NLM), and BM3D have demonstrated excellent denoising performance in various image processing applications, their suitability depends on the characteristics and computational requirements of the target application. In the proposed framework, the primary objective of preprocessing is to generate a clean and edge-preserving input for the subsequent LadderNet-based segmentation stage rather than to maximize denoising performance alone. Therefore, the contra-harmonic mean filter was selected because it effectively suppresses impulse noise while preserving lesion boundaries with relatively low computational complexity. This balance between denoising capability, edge preservation, and computational efficiency makes it a suitable preprocessing technique for accurate rice leaf disease classification.

Table 1. Qualitative comparison of commonly used denoising techniques based on their reported characteristics in the literature

Method	Noise Removal	Edge Preservation	Computational Cost	Suitability
Median Filter	Good	Moderate	Low	Effective for impulse noise but may smooth lesions
Bilateral Filter	Good	High	Medium	Preserves edges with increased computation
Non-Local Means	Excellent	Excellent	High	Strong denoising but computationally expensive
BM3D	Excellent	Excellent	Very High	High-quality denoising with high computational overhead

Contra-harmonic Mean Filter (Proposed)	Good	High	Low	Efficient preprocessing while preserving disease boundaries
---	------	------	-----	---

As summarized in Table 1, BM3D and Non-Local Means generally achieve superior denoising performance but require substantially higher computational resources and processing time. In contrast, the contra-harmonic mean filter provides an effective balance between impulse noise suppression, lesion boundary preservation, and computational efficiency, making it an appropriate preprocessing technique for the proposed rice leaf disease classification framework.

3.3 Plant Leaf Segmentation utilizing LadderNet

It is performed to identify and isolate leaf region from background image. Moreover, it is used to extract only the relevant leaf region to enable accurate evaluation. Here, the segmentation phase is performed using LadderNet with an input of C_x .

3.3.1 LadderNet

Figure 2 illustrates the general architecture of LadderNet [21], a multi-branch convolutional neural network (CNN) designed to provide multiple information propagation paths for accurate semantic segmentation. The architecture consists of four columns, where Columns 1 and 3 represent the encoder branches and Columns 2 and 4 represent the corresponding decoder branches. The encoder progressively extracts hierarchical feature representations using 3×3 convolutional layers followed by ReLU activation, while spatial downsampling is performed through 3×3 convolutions with a stride of 2 to enlarge the receptive field. The decoder reconstructs the segmentation map using 3×3 transposed convolutions with a stride of 2, enabling the recovery of fine spatial details. Skip connections employing element-wise tensor addition and shared residual-weight blocks facilitate efficient feature propagation and gradient flow between the encoder and decoder pathways. Owing to the multiple encoder-decoder connections, LadderNet can be interpreted as an ensemble of fully convolutional networks (FCNs), providing diverse information flow paths that improve feature representation and segmentation accuracy. Consequently, the segmented output image S_x preserves the lesion boundaries required for the subsequent WLBP–GLCM feature extraction and PDKN-based disease classification. For implementation, the LadderNet architecture employs four encoder-decoder levels with symmetric skip connections. The segmentation model is trained using the Adam optimizer with a learning rate of 0.001 and a batch size of 16, and the same training configuration is maintained throughout all segmentation experiments to ensure a fair and reproducible evaluation.

Table 2. Implementation settings of the LadderNet segmentation model.

Parameter	Value
Encoder–decoder levels	4
Convolution kernel	3×3
Downsampling	3×3 convolution, stride = 2
Upsampling	3×3 transposed convolution, stride = 2
Hidden activation	ReLU
Optimizer	Adam
Learning rate	0.001

Batch size	16
------------	----

The implementation settings adopted for the LadderNet segmentation model are summarized in Table 2. These architectural and training configurations were consistently maintained throughout all segmentation experiments to ensure a fair and reproducible evaluation of the proposed framework.

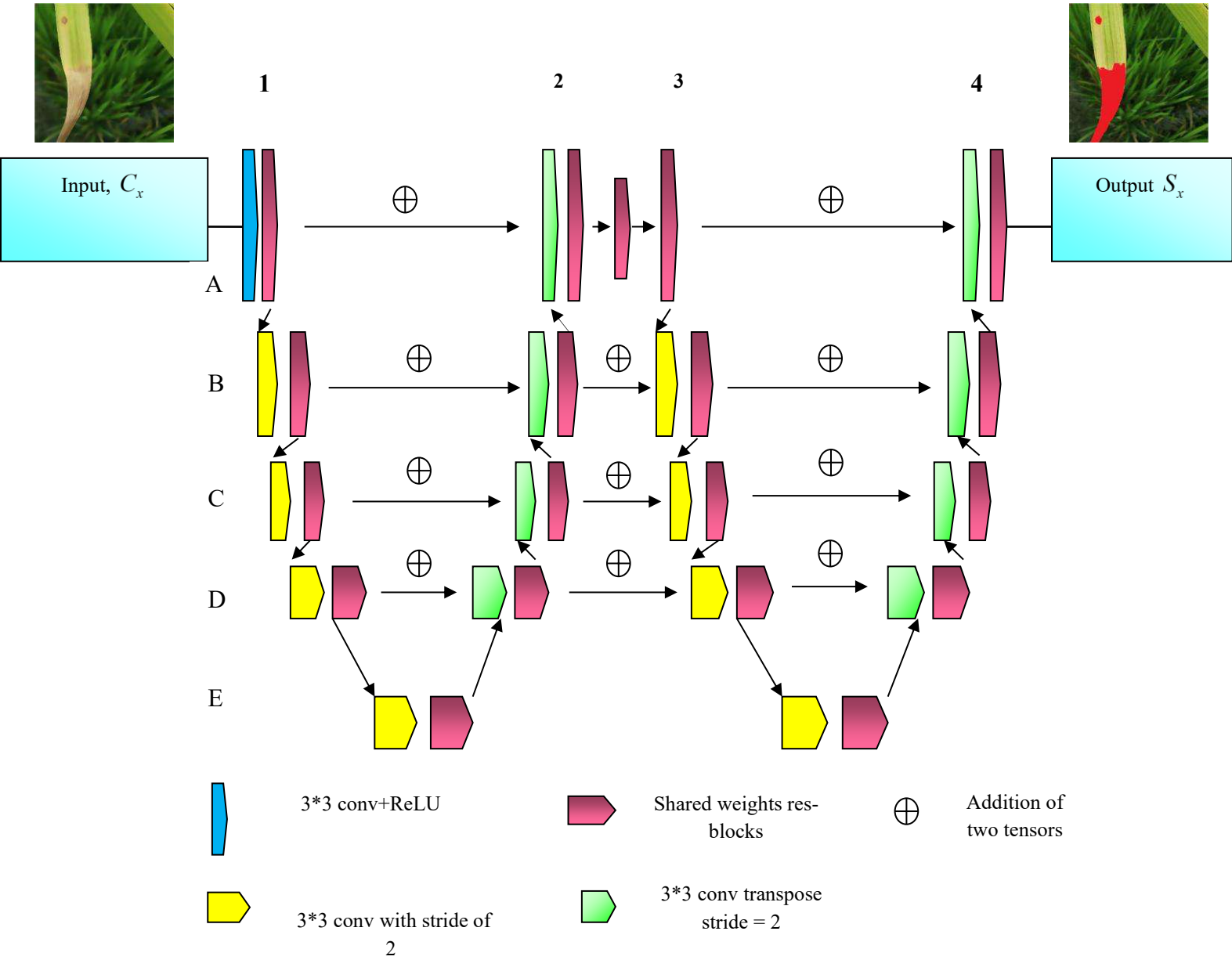


Figure 2. General view of LadderNet

3.4 Feature Extraction using segmented image

It transforms the raw image into meaningful features and reduces the dimensionality of data while preserving significant information. Here, S_x is allowed as an input for extraction process.

Although deep feature embeddings extracted from CNN backbones have demonstrated strong representation capability, handcrafted descriptors are employed in the proposed framework to explicitly capture disease-specific texture and lesion characteristics. Rice leaf diseases are primarily manifested through variations in texture, spot distribution, discoloration, and surface irregularities, which are effectively represented by WLBP and GLCM features. Furthermore, handcrafted features provide improved interpretability, lower computational complexity, and reduced feature dimensionality compared with deep feature embeddings. Since LadderNet already performs region localization and PDKN focuses on classification, the use of WLBP and GLCM descriptors enables efficient extraction of discriminative texture information while reducing computational overhead.

3.4.1 WLBP

WLBP has two different components namely, differential excitation and Local Binary Pattern (LBP). The perception features are extracted through differential excitation by utilizing Weber's law, whereas local features are well-defined using LBP. Weber's law is exploited for visual stimulus that is produced from image features to produce a descriptor. The variance is deployed between center and neighboring pixels using,

$$WLBP_{E,H,\delta}(\Psi_h) = \sum_{\lambda=0}^{E-1} I_{\delta} \left(\frac{|\Psi_{\lambda} - \Psi_h|}{\Psi_h} \right) \times 2^{\lambda} \quad (3)$$

$$I_{\delta}(\Psi) = \begin{cases} 1, & \Psi \geq 0 \\ 0, & \Psi < 0 \end{cases} \quad (4)$$

Here, Ψ_h elucidates gray value of given center pixel, Ψ_{λ} implies gray values of E neighbors, which is being distributed in circular window with radius H for Ψ_h , where $(\lambda = 0, 1, \dots, E-1)$ and $\delta = 0.1$ that implies a threshold for gray level difference. By applying WLBP feature over the segmented image S_x , a textural image is attained and it is indicated as A_x .

3.4.2 GLCM features

Contrast, IDM, and correlation are the GLCM features applied over A_x to acquire feature vector G_x that are elaborated as follows.

a) Contrast

It measures the local variations in gray-level intensity of leaf image. Here, higher contrast indicates the presence of lesions or spots, which is computed using,

$$G_1 = \sum_{\xi=0}^{N-1} \sum_{\zeta=0}^{N-1} (\xi - \zeta)^2 \mathbf{Z}(\xi, \zeta, \ell, \theta) \quad (5)$$

Here, G_1 implies contrast, $\xi, \zeta = 0, 1, 2, \dots, N-1$ indicates grey level of two pixels, $\mathbf{Z}(\cdot)$ denotes probability of distribution, ℓ specifies distance and θ signifies direction.

b) IDM

It quantifies the similarity of neighboring pixel values. Here, maximal value of IDM enumerates smooth textures and minimal IDM expresses irregularities, which is enumerated by,

$$G_2 = \sum_{\xi=0}^{N-1} \sum_{\zeta=0}^{N-1} \frac{\mathbf{Z}(\xi, \zeta, \ell, \theta)}{1 + (\xi - \zeta)^2} \quad (6)$$

Here, G_2 expresses IDM.

c) Correlation

It analyzes the linear dependency between gray levels of neighbor pixels and it assists in identifying the repetitive disease patterns or texture consistency across leaf surfaces. It is determined as,

$$G_3 = \sum_{\xi=0}^{N-1} \sum_{\zeta=0}^{N-1} (\xi - \vartheta_1)(\zeta - \vartheta_2) / \kappa_1 \kappa_2 \quad (7)$$

Here, G_3 specifies correlation, ϑ_1, ϑ_2 denotes mean values and κ_1, κ_2 refers to mean squared values.

Therefore, feature vector G_x is expressed as,

$$G_x = \{G_1, G_2, G_3\} \quad (8)$$

3.5 Rice Leaf disease classification using HaEST-Loss function

In agricultural technology, disease classification assists to maintain rice crop health for better cultivation. The diseases may cause substantial damage to the cultivation that leads to less cultivation as well as economic losses. Prior techniques mostly consume large amount of time and need expertise. Automated detection systems assist farmers in identifying plant diseases with minimal or no human intervention. Therefore, PDKN with HaEST-Loss is designed to classify rice plant leaf diseases.

Figure 3 shows the general outlook of novel framework, where it is composed of PDAN, DKN and hybrid loss functions. Initially, feature vector G_x is given to PDAN network and also it is

given to sum with the weight ψ_1 that is expressed as $\sum \psi_1$. In the first iteration, only smooth surrogate loss is applied to perform PDAN model. Then, the output of PDAN model and $\sum \psi_1$ is concatenated, which results an output P_{x1} . After that, the output P_{x1} is given to DKN as an input. Here, the loss function of DKN is TCE loss, which is processed at the first iteration and yields an output P_{x2} . Moreover, hybrid loss named HaEST-Loss is designed to solve the variation between the targeted output and the classified output. Here, smooth surrogate loss, TCE loss and Euclidean loss are integrated by utilizing harmonic analysis. This process will be performed for subsequent iterations to minimize the deviation using HaEST-Loss till a reasonable accuracy is achieved. Here, the PDKN is tuned using Adam optimizer. In this figure, the dotted lines indicate the first iteration and the straight-line arrows represent the remaining iterations. Finally, the classified output is commonly expressed as P_x .

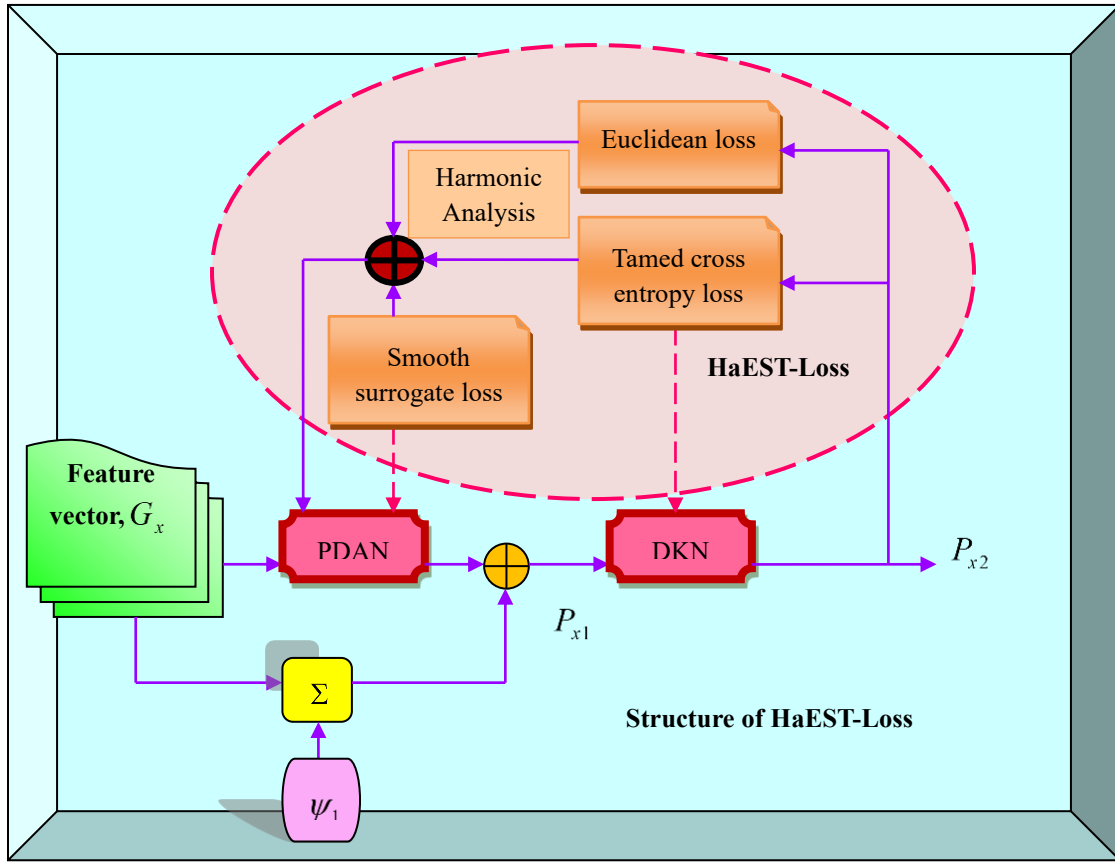


Figure 3. General outlook of HaEST-Loss

3.5.1 PDAN model

PDAN is an end-to-end model, where its major intention is how to assign the attention weights to entire leaf images at multi-temporal scales. Dilated Attention Layer (DAL) is initially defined as temporal filter across time. It can extract superior feature representation from neighboring leaves using an attention model between kernels. PDAN is designed by stacking DALs with diverse dilation rates. In PDAN, the elementary building block is a DAL followed by ReLU as well as

bottleneck with residual connection. PDAN functions with RGB and Flow modalities based on their availability. Here, G_x is given as the input for PDAN module.

a) DAL

DAL is designed with numerous dilation rates, which characteristically learns attention weights at diverse temporal scales. DAL functions feature maps across temporal domain to maintain spatial information as same as the temporal filters. When compared to 3×1 kernel in Standard temporal convolution layer (STCL), DAL comprises 1×1 learnable kernels. Thus, it has same count of parameters as STCL.

Consider l^{th} block for the analysis, then the segment around the center segment ς_{lo} at time $o \in [1, T]$ is extracted to generate a representative vector ς'_{lo} . It is based on kernel dimension K_s as well as dilation rate D_r at l^{th} block. Assume $\varsigma_{lo} \in \mathfrak{I}^{1 \times M_2}$ and $\varsigma'_{lo} \in \mathfrak{I}^{K_s \times M_2}$. After that, the self-attention scoring technique is appealed by using representative vector ς'_{lo} to memory embedding such as key X_l and value Y_l through two certain bottleneck convolutions namely, $X_l(\varsigma'_{lo}) = A_{X_l} \varsigma'_{lo}$ and $Y_l(\varsigma'_{lo}) = A_{Y_l} \varsigma'_{lo}$, where both A_{X_l} and $A_{Y_l} \in \mathfrak{I}^{M_2 \times M_2}$. Thereafter, ς_{lo} is subjected to query Z_l utilizing another bottleneck convolution $Z_l(\varsigma_{lo}) = A_{Z_l} \varsigma_{lo}$ and $A_{Z_l} \in \mathfrak{I}^{M_2 \times M_2}$. The resultant of attention function for o^{th} time step is produced by weighted sum of values Y_l with attention weights attained from product of Z_l as well as X_l that is expressed as,

$$m_l(\varsigma_{lo}) = Y_l(\varsigma'_{lo}) \left[\text{soft max} \left(Z_l(\varsigma_{lo}) X_l(\varsigma'_{lo}) \right) \right] \quad (9)$$

DAL determines correlation in kernel among central segment and K_s adjacent segments. Therefore, for every time step l , it has $M_2 \times 1 \times K_s$ attention map. The resultant of DAL is attained by integrating the outcomes of entire time step l of the features that are computed as,

$$\text{attention}_l(G_{x(l)}) = \left[m_l(\varsigma_{l1})^T, m_l(\varsigma_{l2})^T, \dots, m_l(\varsigma_{lT})^T \right] \quad (10)$$

Here, $G_{x(l)}$ indicates input feature map of DAL for l^{th} block. While integrating attention map for all o across time, it has attention weights of all leaves $M_2 \times T \times K_s$.

b) Pyramid structure of temporal layers

Using the self-attention model at a multi-temporal scale is significant part of designing difficult temporal relations. PDAN is related to pyramid of DAL with parallel kernel dimensions and diverse dilation rates. The pyramid improves the receptive field dimension. It enables the network to design short as well as long action patterns by concentrating on local frames in kernel at low level as well as high temporal receptive fields.

Initially, an input feature $G_x \in \mathfrak{T}^{T \times M_1}$ is given to bottleneck layer to lightweight the module by diminishing the channel dimension from M_1 to M_2 . After that, S number of blocks are stacked, where every block l is considered as a cascade of DAL with ReLU function, bottleneck convolution and residual link. It enables receptive field to improve exponentially by keeping the parallel temporal length T as input. Here, $K_s = 3$ is considered for entire blocks, dilation as well as padding rate to 2^{l-1} , thus reception field is set to be $2^l + 1$ for l^{th} block. The set of functions in all blocks is expressed as,

$$G_{x(l+1)} = G_{x(l)} + A_l * \text{ReLU}\left(\text{attention}_l\left(G_{x(l)}\right)\right) \quad (11)$$

In the attention layer attention_l , the D_r is changed with l . Here, $A_l \in \mathfrak{T}^{M_2 \times M_2}$ implies bottleneck's weight. Therefore, pre-frame binary classification score for all classes is determined and S^{th} block is followed by bottleneck convolution with sigmoid function that is enumerated as,

$$P_{x1} = \text{sigmoid}\left(A_{O_{s+1}} \cdot G_x\right) * \left(\sum \psi_1 * G_x\right) \quad (12)$$

Here, $P_{x1} \in \mathfrak{T}^{T \times M_3}$ refers to prediction logits or output of PDAN model and $A_{O_{s+1}} \in \mathfrak{T}^{M_3 \times M_2}$, where M_3 is corresponds to action classes. Multi-label binary cross-entropy loss is optimized to learn all parameters. Figure 4 demonstrates the architectural view of PDAN.

The dilation factors employed in PDAN are selected to capture disease characteristics at multiple spatial scales and to enhance the representation capability of the network. Lower dilation rates focus on fine-grained lesion structures, texture variations, edge information, and small disease symptoms that are commonly observed during the early stages of infection. In contrast, higher dilation rates capture broader contextual information and long-range dependencies surrounding diseased regions, which are important for recognizing large lesions and complex disease patterns distributed across the leaf surface. By integrating multiple dilation rates within the pyramid structure, PDAN effectively aggregates local and global feature information while progressively enlarging the receptive field.

Furthermore, the utilization of multiple dilation rates enables the network to improve feature extraction without significantly increasing the number of trainable parameters or computational complexity. This multi-scale attention mechanism allows the model to learn discriminative representations from disease regions exhibiting diverse sizes, shapes, and spatial distributions. Consequently, PDAN improves the capability of the proposed framework to distinguish between different rice leaf diseases with similar visual characteristics. The effectiveness of different dilation configurations was experimentally evaluated through the ablation study presented in Table 3. The results demonstrate that the proposed multi-scale configuration $\{1,2,4\}$ provides the best balance between local feature representation and global contextual information, thereby achieving the highest classification performance.

Table 3. Ablation Study on Different Dilation Rate Configurations

Dilation Configuration	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
------------------------	--------------	---------------	------------	--------------

{1}	88.91	88.44	88.15	88.29
{1,2}	90.43	90.01	89.76	89.88
{1,2,4,8}	91.31	90.90	90.44	90.66
{1,2,4} (Proposed)	91.65	91.18	90.73	90.95

The experimental analysis demonstrates that the proposed multi-scale dilation configuration {1,2,4} achieves the highest classification accuracy by effectively balancing local lesion representation and global contextual information. Although incorporating a larger dilation factor (8) increases the receptive field, it introduces sparse feature sampling that slightly degrades classification performance. These results justify the selected dilation configuration adopted in the proposed PDAN architecture.

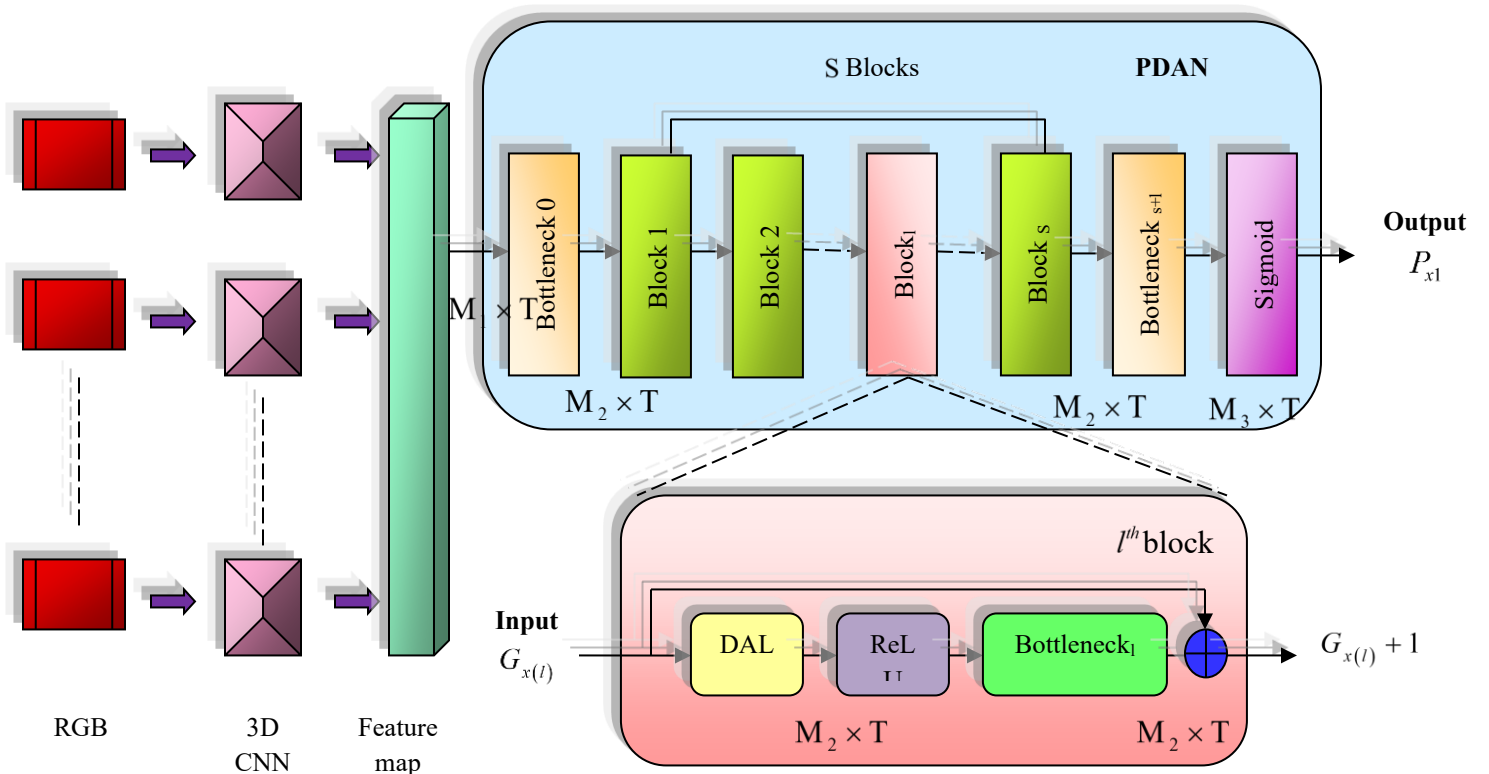


Figure 4. Architectural view of PDAN

3.5.2 DKN model

DKN is established on basis of Kronecker product and indirectly executes a latent piecewise smooth property of coefficients. It enables to locating the regions present in an image, which is more important to the output as well as assists for model interpretation. Also, it functions for matrix and large-scale tensor-represented image data. Here, P_{x1} is subjected as the input for DKN model.

Consider P samples with matrix-represented images $P_{x1(t)} \in \mathbb{R}^{u \times v}$ as well as scalar responses w_t , where $t = 1, \dots, P$ and w_t follows generalized linear approach that is computed by,

$$P_{x2} = \nu(w_t) \exp \left\{ w_t \langle P_{x1}, J \rangle - \varpi(\langle P_{x1}, J \rangle) \right\} \quad (13)$$

Substitute Eq. (12) in Eq. (13), thus it is rewritten as,

$$P_{x2} = \nu(w_t) \exp \left\{ w_t \left\langle \left(\text{sigmoid}(A_{O_{S+1}} \cdot G_x) * \left(\sum \psi_1 * G_x \right) \right), J \right\rangle - \varpi \left(\left\langle \left(\text{sigmoid}(A_{O_{S+1}} \cdot G_x) * \left(\sum \psi_1 * G_x \right) \right), J \right\rangle \right) \right\} \quad (14)$$

Here, $J \in F^{u \times v}$ indicates target unknown coefficients matrix, $\nu(\cdot)$ and $\varpi(\cdot)$ implies certain known univariate functions. In Eq. (13), image analysis is concentrated and neglects other potential design variables. With Eq. (13), it comprises certain known link functions $n(\cdot)$ that are enumerated as,

$$n(k(w_t)) = \langle P_{x1}, J \rangle \quad (15)$$

Using DKN, coefficients J with rank- G Kronecker factor decomposition with $I(\geq 2)$ factors that are demonstrated as,

$$J = \sum_{g=1}^G M_I^g \otimes M_I^g \otimes \dots \otimes M_1^g \quad (16)$$

Here, $M_i^g \in F^{u_i \times v_i}$, wherein $i = 1, \dots, I$; $g = 1, \dots, G$ are unknown matrices and enumerates Kronecker factors. The dimension of M_i^g is not considered as unknown. But they certainly need to

be satisfied $v = \prod_{i=1}^I v_i$ due to Kronecker's product properties. It expresses

$M_{i'} \otimes M_{i'-1} \otimes \dots \otimes M_{i''} = \bigotimes_{i=i'}^{i''} M_i$ for any matrices $M_{i'}, \dots, M_{i''}$ with $i' \geq i''$. Hence, decomposition in

Eq. (16) is written as $J = \sum_{g=1}^G \bigotimes_{i=1}^I M_i^g$. Generally, the decomposition in Eq. (16) can evaluate random matrices with large ranks G .

The depth of the proposed DKN architecture is intentionally maintained at a moderate level to achieve a balance between classification performance and computational efficiency. Since the feature extraction stage already generates discriminative texture descriptors through WLBP and GLCM, employing an excessively deep architecture may introduce redundant computations and increase the risk of overfitting, particularly when the available training samples are limited. Furthermore, the Kronecker-based representation learning mechanism enables effective modeling of complex feature interactions, allowing the network to capture disease-related patterns without requiring a very deep architecture. Therefore, the selected DKN configuration provides an appropriate trade-off between representation capability, training complexity, and classification performance. Compared with conventional deep CNN architectures that rely on a large number of convolutional layers and trainable parameters, the proposed DKN framework utilizes Kronecker-

based feature representation to enhance feature learning while maintaining computational efficiency. The architecture requires fewer trainable parameters and lower memory consumption than deeper CNN models, thereby reducing computational overhead during training and inference. This characteristic makes the proposed framework suitable for practical agricultural applications where computational resources may be limited while still maintaining competitive classification performance. The output of DKN is specified as P_{x_2} and the general illustration of DKN structure is displayed in Figure 5.

The overall computational complexity of the proposed PDKN framework is determined by the convolutional operations performed in the Pyramid Dilated Attention Network (PDAN) together with the Kronecker-based feature learning operations in the Deep Kronecker Network (DKN), and can be expressed as $O(Nk^2C + NFC)$, where N denotes the number of training samples, k represents the convolution kernel size, C is the number of feature channels, and F indicates the extracted feature dimension. Since the Kronecker decomposition factorizes large weight matrices into compact subspace representations, the proposed PDKN architecture substantially reduces the number of trainable parameters, computational overhead, and memory consumption compared with conventional deep CNN models. Consequently, the proposed framework achieves an effective balance between computational efficiency and classification performance, making it suitable for practical rice leaf disease diagnosis.

Since the Kronecker decomposition factorizes large weight matrices into compact subspace representations, the proposed PDKN architecture substantially reduces the number of trainable parameters, computational overhead, and memory consumption compared with conventional deep CNN models. Consequently, the proposed framework achieves an effective balance between computational efficiency and classification performance, making it suitable for practical rice leaf disease diagnosis.

Table 4. Qualitative comparison of computational characteristics between conventional CNN architectures and the proposed PDKN framework.

Model	Trainable Parameters	Memory Consumption	Computational Complexity	Computational Efficiency
Conventional CNN	High	High	High	Moderate
VGG16	Very High	Very High	Very High	Low
ResNet50	High	High	High	Moderate
DenseNet121	Medium	Medium	Medium	High
Proposed PDKN	Low	Low	Low	High

The qualitative comparison is based on the relative architectural characteristics of the compared models, including network depth, parameter count, memory requirements, and computational complexity. The comparison is intended to provide a conceptual assessment of the computational efficiency of the proposed framework rather than experimentally measured hardware-specific execution times. Conventional CNN architectures generally exhibit higher computational requirements because of their deeper network structures and larger parameter sets. In contrast, the proposed PDKN framework combines Kronecker-based feature representation with multi-scale attention to reduce computational overhead while maintaining competitive classification performance. These characteristics make the proposed framework suitable for practical rice leaf disease diagnosis in resource-constrained agricultural environments.

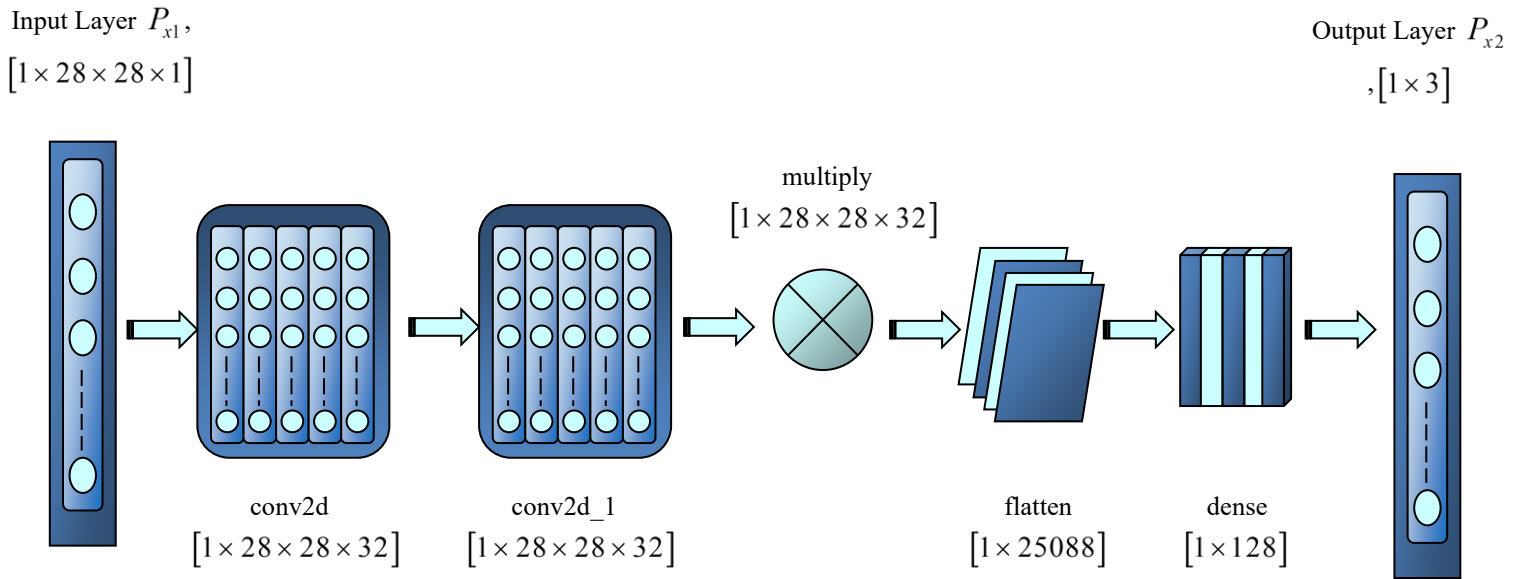


Figure 5. General illustration of DKN

3.5.3 Hybrid HaEST-Loss function

HaEST-Loss is designed by integrating TCE Loss, Euclidean loss, and smooth surrogate loss using Harmonic analysis. A brief explanation of each loss function is below.

a) TCE Loss

Once the softmax layer normalizes the network's output, cross-entropy loss is considered and it is determined as,

$$\text{softmax}(o) = \frac{e^{o_i}}{\sum_{t=1}^K e^{o_t}} \quad (17)$$

Here, cross-entropy between two K sized discrete distributions $x \in [0,1]^K$ as well as $y \in (0,1]^K$ that is modeled as,

$$P(x, y) = -\sum_{\kappa=1}^K x_{\kappa} \log_{y_{\kappa}} \quad (18)$$

Here, x corresponds to classification target and y indicates softmax layer. The gradient of cross-entropy loss is examined based on log function. By utilizing $x \in \{0,1\}^K$, the partial derivatives of CE loss based on classified probabilities of log is determined as,

$$\frac{\partial P(x, y)}{\partial \log y_{\kappa}} = \begin{cases} 0 & \text{if } x_{\kappa} = 0 \\ -1 & \text{if } x_{\kappa} = 1 \end{cases} \quad (19)$$

To design a robust classification model, a set of significant phases are performed that are enumerated as follows:

- The gradient of TCE loss $\left(\hat{P}\right)$ should be equivalent to gradient of cross-entropy (CE) loss. With these, both losses will be in the parallel form that is modeled as,

$$\nabla \hat{P}(x, y) \propto \nabla P(x, y) \quad (20)$$

- When the network is positive about the classified class, which is represented as $y_{\kappa} \rightarrow 1$, the TCE should be performed like CE that is computed as

$$\nabla \hat{P}(x, y) = -1 \text{ if } x_{\kappa} = 1 \text{ and } y_{\kappa} \rightarrow 1 \quad (21)$$

- The effect of outliers should be reduced by decreasing the gradient response when there is a large difference among classification and its associated label that is expressed as,

$$\nabla \hat{P}(x, y) = 0 \text{ if } x_{\kappa} = 1 \text{ and } y_{\kappa} \rightarrow 0 \quad (22)$$

Thus, a function with gradient is designed by utilizing,

$$\frac{\partial \hat{P}(x, y)}{\partial \log y_{\kappa}} = \begin{cases} 0 & \text{if } x_{\kappa} = 0 \\ -1 & \text{if } x_{\kappa} = 1 \text{ and } y_{\kappa} \rightarrow 1 \\ 0 & \text{if } x_{\kappa} = 1 \text{ and } y_{\kappa} \rightarrow 0 \end{cases} \quad (23)$$

The below gradients that accomplish the requirements are determined as,

$$\frac{\partial \hat{P}(x, y)}{\partial \log y_{\kappa}} = \begin{cases} 0 & \text{if } x_{\kappa} = 0 \\ -(1 - \log y_{\kappa})^{-\beta} & \text{if } x_{\kappa} = 1 \end{cases} \quad (24)$$

The regularization is applied on domain $[1, \infty)$ of power function that also applies to $\log x_{\kappa}$. Here, regularization factor is maintained by a parameter $\beta \in \mathbf{A}^+$. As a result, the loss function based on gradient is determined as,

$$Tam_{loss} = \frac{1}{1-\beta} \sum_{\kappa=1}^K x_{\kappa} \left((1-\log y_{\kappa})^{1-\beta} - \frac{1}{1-\beta} \right) \quad (25)$$

Here, the characteristics of both $\hat{P}_{\beta}$ and $\nabla \hat{P}_{\beta}$ are observed. Moreover, when $\beta = 0$, the $\hat{P}_{\beta}$ is corresponds to P and Tam_{loss} implies TCE loss.

b) Euclidean loss

Euclidean loss measures the divergence between classified values and actual values in regression tasks. It is computed as,

$$Euc_{Loss} = \chi \|\phi\|^2 - \frac{1}{\square} \sum_{\rho \in \delta} \log \left(\eta \left(\omega^{(\rho)} \middle| \hat{\omega}^{(\rho)} \right) \right) \quad (26)$$

Here, Euc_{Loss} enumerates Euclidean loss, $\eta \left(\omega^{(\rho)} \middle| \hat{\omega}^{(\rho)} \right)$ is the likelihood with $\hat{\omega}^{(\rho)} = f^{G_g(\phi, \varepsilon)} \left(X_x^{(\rho)} \right)$ indicating the multi-scale CNN with dropouts, ϕ has the parameters to be learned, $G_g(\phi, \varepsilon) = diag(\varepsilon)\phi$ and $diag(\cdot)$ maps a vector to diagonal matrix whose diagonal is vector element, ε implies binary vector, δ depicts random subset of $\{1, \dots, m\}$ with dimension $\square$ for data sub-sampling and χ represents weight decay for regularization term $\|\phi\|^2$.

c) Smooth surrogate loss

Smooth surrogate loss function plays a significant role in training the modules when the true loss functions are difficult to optimize. By utilizing this loss function, the learning process is more stable which often leads to superior generalization on unseen data.

Consider the label $\bar{q} \in Q$, $Q_{\bar{q}}^{(r)}$ indicates the subset of tuples from $Q^{(r)}$ that involves $\bar{q}$ as one of the elements. Moreover, $\Delta_r(\bar{Q}, q) \square 1(q \notin \bar{Q})$ is indicated for $\bar{q} \in Q^{(r)}$ as well as $q \in Q$. By adding as well as subtracting $r-1$ highest scores of $\mu_{\setminus q}$ and μ_q , the solution attained is,

$$\begin{aligned} \alpha_r(\mu, q) &= \max \left\{ \left(\mu_{\setminus q} + 1 \right)_{[r]} - \mu_q, 0 \right\} \\ &= \max_{\bar{q} \in Q^{(r)}} \left\{ \Delta_r(\bar{Q}, q) + \sum_{\gamma \in \bar{Q}} \mu_{\gamma} \right\} - \max_{\bar{q} \in Q_{\bar{q}}^{(r)}} \left\{ \sum_{\gamma \in \bar{Q}} \mu_{\gamma} \right\} \end{aligned} \quad (27)$$

As the margin can be rescaled without loss of generality, the α_r is rewritten as,

$$\alpha_r(\mu, q) = \max_{\bar{q} \in Q^{(r)}} \left\{ \Delta_r(\bar{Q}, q) + \frac{1}{r} \sum_{\gamma \in \bar{Q}} \mu_\gamma \right\} - \max_{\bar{q} \in Q_q^{(r)}} \left\{ \frac{1}{r} \sum_{\gamma \in \bar{Q}} \mu_\gamma \right\} \quad (28)$$

The loss function is smoothened with the temperature factor $W > 0$ that is computed as,

$$SS_{Loss} = w \log \left[\sum_{\bar{q} \in Q^{(r)}} \exp \left(\frac{1}{W} \left(\Delta_r(\bar{Q}, q) + \frac{1}{r} \sum_{\gamma \in \bar{Q}} \mu_\gamma \right) \right) \right] - W \log \left[\sum_{\bar{q} \in Q_q^{(r)}} \exp \left(\frac{1}{rW} \sum_{\gamma \in \bar{Q}} \mu_\gamma \right) \right] \quad (29)$$

Here, SS_{Loss} indicates smooth surrogate loss, α_r implies surrogate loss, μ denotes vector of scores, Q enumerates output space, $q \in Q$ specifies ground truth label, $Q^{(r)}$ depicts set of r -tuples with r different elements of Q .

d) Derivation of HaEST-Loss function using harmonic analysis

This sub section explains the derivation of HaEST-Loss function using harmonic analysis which surpasses in finding basic periodicities by disintegrating a time series into sinusoidal components. This characteristically designs time-based weights that dynamically capture the effect of each periodic component over time. While incorporating the benefit of providing time-based weights into DL approaches, harmonic analysis improves the classification performance by embedding time-aware weighting method into learning process, particularly in the loss function, which makes it efficient for progressive learning task.

In Harmonic Analysis, the configuration of signals as the sum of basic waves are defined and the fundamental aim is to decompose the functions into their basic frequencies that allow for the analysis of functions based on simpler, and periodic components. Based on orthogonal trigonometric functions, it is probable to provide overall description of observed sequence $B_1, B_2, B_3, \dots, B_\lambda, \dots, B_z$.

The arithmetical expression of harmonic analysis is derived as follows:

$$B_\lambda = R_0 + \sum_{p=1}^k \left[R_p \cos \left(\frac{2\pi ps}{z} \right) + S_p \sin \left(\frac{2\pi ps}{z} \right) \right] \quad (30)$$

Here, z refers to length of series, k depicts overall cycles, R_p and S_p indicates amplitudes. Considering $z = 3; k = 1$, then above expression is enumerated as,

$$B_\lambda = R_0 + R_1 \cos \left(\frac{2\pi s}{3} \right) + S_1 \sin \left(\frac{2\pi s}{3} \right) \quad (31)$$

Here,

$$R_0 = \frac{1}{z} \sum_{\lambda=1}^z B_\lambda \quad (32)$$

$$R_0 = \frac{1}{3} [B_1 + B_2 + B_3] \quad (33)$$

$$R_p = \frac{2}{z} \sum_{\lambda=1}^z B_\lambda \cos\left(\frac{2\pi p\lambda}{z}\right) \quad (34)$$

$$R_1 = \frac{2}{3} \left[B_1 \cos\left(\frac{2\pi}{3}\right) + B_2 \cos\left(\frac{4\pi}{3}\right) + B_3 \cos(2\pi) \right] \quad (35)$$

$$R_1 = \frac{2}{3} \left[B_1 \left(-\frac{1}{2}\right) + B_2 \left(-\frac{1}{2}\right) + B_3 (1) \right] \quad (36)$$

$$S_p = \frac{2}{z} \sum_{\lambda=1}^z B_\lambda \sin\left(\frac{2\pi p\lambda}{z}\right) \quad (37)$$

$$S_1 = \frac{2}{3} \left[B_1 \sin\left(\frac{2\pi}{3}\right) + B_2 \sin\left(\frac{4\pi}{3}\right) + B_3 \sin(2\pi) \right] \quad (38)$$

$$S_1 = \frac{2}{3} \left[B_1 \left(\frac{\sqrt{3}}{2}\right) + B_2 \sin\left(\frac{-\sqrt{3}}{2}\right) + B_3 (0) \right] \quad (39)$$

$$S_1 = \frac{2}{3} \left[B_1 \left(\frac{\sqrt{3}}{2}\right) + B_2 \left(\frac{-\sqrt{3}}{2}\right) \right] \quad (40)$$

In time series module, $B_1 = B(\lambda)$, $B_2 = B(\lambda-1)$, $B_3 = B(\lambda-2)$ and $B_\lambda = B(\lambda+1)$. Substitute these values in Eq. (33), (36), (40) thus it is modeled as,

$$R_0 = \frac{1}{3} [B(\lambda) + B(\lambda-1) + B(\lambda-2)] \quad (41)$$

$$R_1 = \frac{2}{3} \left[B(\lambda) \left(-\frac{1}{2}\right) + B(\lambda-1) \left(-\frac{1}{2}\right) + B(\lambda-2)(1) \right] \quad (42)$$

$$S_1 = \frac{2}{3} \left[B(\lambda) \left(\frac{\sqrt{3}}{2}\right) + B(\lambda-1) \left(\frac{-\sqrt{3}}{2}\right) \right] \quad (43)$$

Substitute Eq. (41), (42), (43) in Eq. (31), thus it is expressed as,

$$\begin{aligned}
B(\lambda+1) = & \left[\frac{1}{3} [B(\lambda) + B(\lambda-1) + B(\lambda-2)] \right] + \frac{2}{3} \left[B(\lambda) \left(-\frac{1}{2} \right) + B(\lambda-1) \left(-\frac{1}{2} \right) + B(\lambda-2)(1) \right] \cos\left(\frac{2\pi s}{3}\right) \\
& + \frac{2}{3} \left[B(\lambda) \left(\frac{\sqrt{3}}{2} \right) + B(\lambda-1) \left(-\frac{\sqrt{3}}{2} \right) \right] \sin\left(\frac{2\pi s}{3}\right)
\end{aligned} \tag{44}$$

$$\begin{aligned}
B(\lambda+1) = & B(\lambda) \left[\frac{1}{3} - \frac{2}{3} \times \frac{1}{2} \times \cos\left(\frac{2\pi s}{3}\right) + \frac{2}{3} \times \frac{\sqrt{3}}{2} \times \sin\left(\frac{2\pi s}{3}\right) \right] \\
& + B(\lambda-1) \left[\frac{1}{3} - \frac{2}{3} \times \frac{1}{2} \times \cos\left(\frac{2\pi s}{3}\right) + \frac{2}{3} \times \left(-\frac{\sqrt{3}}{2} \right) \times \sin\left(\frac{2\pi s}{3}\right) \right] + B(\lambda-2) \left[\frac{1}{3} + \frac{2}{3} \times \cos\left(\frac{2\pi s}{3}\right) \right]
\end{aligned} \tag{45}$$

$$\begin{aligned}
B(\lambda+1) = & B(\lambda) \left[\frac{1}{3} \left(1 - \cos\left(\frac{2\pi s}{3}\right) \right) + \frac{\sqrt{3}}{3} \sin\left(\frac{2\pi s}{3}\right) \right] + B(\lambda-1) \left[\frac{1}{3} \left(1 - \cos\left(\frac{2\pi s}{3}\right) \right) - \frac{\sqrt{3}}{3} \sin\left(\frac{2\pi s}{3}\right) \right] \\
& + B(\lambda-2) \left[\frac{1}{3} \left(1 + 2 \cos\left(\frac{2\pi s}{3}\right) \right) \right]
\end{aligned} \tag{46}$$

$$\begin{aligned}
B(\lambda+1) = & B(\lambda) \times \frac{1}{3} \left[\left(1 - \cos\left(\frac{2\pi s}{3}\right) \right) + \sqrt{3} \sin\left(\frac{2\pi s}{3}\right) \right] + B(\lambda-1) \times \frac{1}{3} \left[\left(1 - \cos\left(\frac{2\pi s}{3}\right) \right) - \sqrt{3} \sin\left(\frac{2\pi s}{3}\right) \right] \\
& + B(\lambda-2) \times \frac{1}{3} \left[\left(1 + 2 \cos\left(\frac{2\pi s}{3}\right) \right) \right]
\end{aligned} \tag{47}$$

According to Eq. (47), it shows that the prediction of variable B for the following time intervals is significantly influenced by its current prediction, adjusted by approximate weights that are derived from sinusoidal components. This replicates the fundamental principle of harmonic analysis, where previous values are maximized with time-sensitive weights to predict further trends. By embedding this principle into the HaEST-Loss framework, the temporal dependencies with loss functions are effectively aligned.

Moreover, the loss function can be adjusted dynamically at diverse stages of classification. Particularly, the HaEST-Loss can utilize TCE loss during initial stages (λ), where temporal alignment is important; Euclidean loss during intermediate phases ($\lambda-1$) to maintain spatial consistency; and smooth surrogate loss in later stages ($\lambda-2$) to improve conceptual reliability. This staged loss adaptation is outlined in the schematic representation in Figure 3, which enables a more refined and context-aware learning process and ensures both accuracy and robustness in temporal prediction.

$$\begin{aligned}
O_{Loss} = & Tam_{Loss} \times \frac{1}{3} \left[\left(1 - \cos\left(\frac{2\pi s}{3}\right) \right) + \sqrt{3} \sin\left(\frac{2\pi s}{3}\right) \right] + Euc_{Loss} \times \frac{1}{3} \left[\left(1 - \cos\left(\frac{2\pi s}{3}\right) \right) - \sqrt{3} \sin\left(\frac{2\pi s}{3}\right) \right] \\
& + SS_{Loss} \times \frac{1}{3} \left[\left(1 + 2 \cos\left(\frac{2\pi s}{3}\right) \right) \right]
\end{aligned} \tag{48}$$

where, O_{Loss} indicates overall loss function, Tam_{Loss} illustrates TCE loss, Euc_{Loss} refers to Euclidean loss and SS_{Loss} denotes smooth surrogate loss, s depicts time, as well as $\cos, \sin$ symbolizes cosine and sine functions.

3.5.4 Theoretical Analysis of HaEST-Loss

The novelty of the proposed HaEST-Loss lies in the adaptive integration of Tamed Cross Entropy (TCE) Loss, Euclidean Loss, and Smooth Surrogate Loss through harmonic-analysis-based weighting. Unlike conventional hybrid loss functions that utilize fixed weighting coefficients, the proposed approach dynamically adjusts the contribution of individual loss components during different stages of training. This enables the framework to simultaneously improve robustness against noisy samples, reduce prediction deviation, and maintain smooth optimization behaviour. The TCE Loss suppresses the influence of mislabeled and difficult samples, the Euclidean Loss minimizes the discrepancy between predicted and target outputs, and the Smooth Surrogate Loss provides a differentiable approximation for complex classification boundaries.

Furthermore, the proposed HaEST-Loss can be interpreted as a weighted combination of three complementary differentiable objective functions, where the harmonic-analysis-based weighting adaptively balances their contributions throughout optimization. Since each constituent loss function is continuous and differentiable within its optimization domain and the harmonic weighting coefficients remain bounded, the overall HaEST-Loss also preserves continuity and differentiability. Consequently, the optimization process maintains bounded gradient propagation, thereby improving optimization stability and facilitating convergence towards a stationary local optimum when optimized using the Adam optimizer.

The proposed HaEST-Loss function remains differentiable because each constituent loss function is differentiable within its optimization domain. In addition, the harmonic weighting mechanism is constructed using bounded sinusoidal components, ensuring continuity and smoothness of the overall objective function. As a result, the combined loss preserves differentiability with respect to the trainable network parameters and enables efficient gradient-based learning throughout the training process.

The gradient of HaEST-Loss is represented as a weighted combination of the gradients of TCE Loss, Euclidean Loss, and Smooth Surrogate Loss. Since the harmonic coefficients remain bounded, the gradient magnitude is controlled during optimization, thereby improving gradient stability. Furthermore, the TCE component reduces excessive gradient contributions from outlier samples, while the Smooth Surrogate component prevents abrupt gradient fluctuations near decision boundaries. Consequently, stable gradient propagation is achieved during training. Since

the proposed framework employs the Adam optimizer and the objective function is continuous, differentiable, and composed of bounded gradient components, the optimization process exhibits stable optimization behaviour and converges towards a stationary local optimum under the adopted optimization settings.

4. Results and Discussions

This section demonstrates the outcomes of HaEST-Loss method by evaluating with the comparative techniques, and details about the dataset, execution tool and analytic measures used in this research.

4.1 Experimental setup

HaEST-Loss designed for leaf disease classification is done in PYTHON tool.

4.2 Database Description

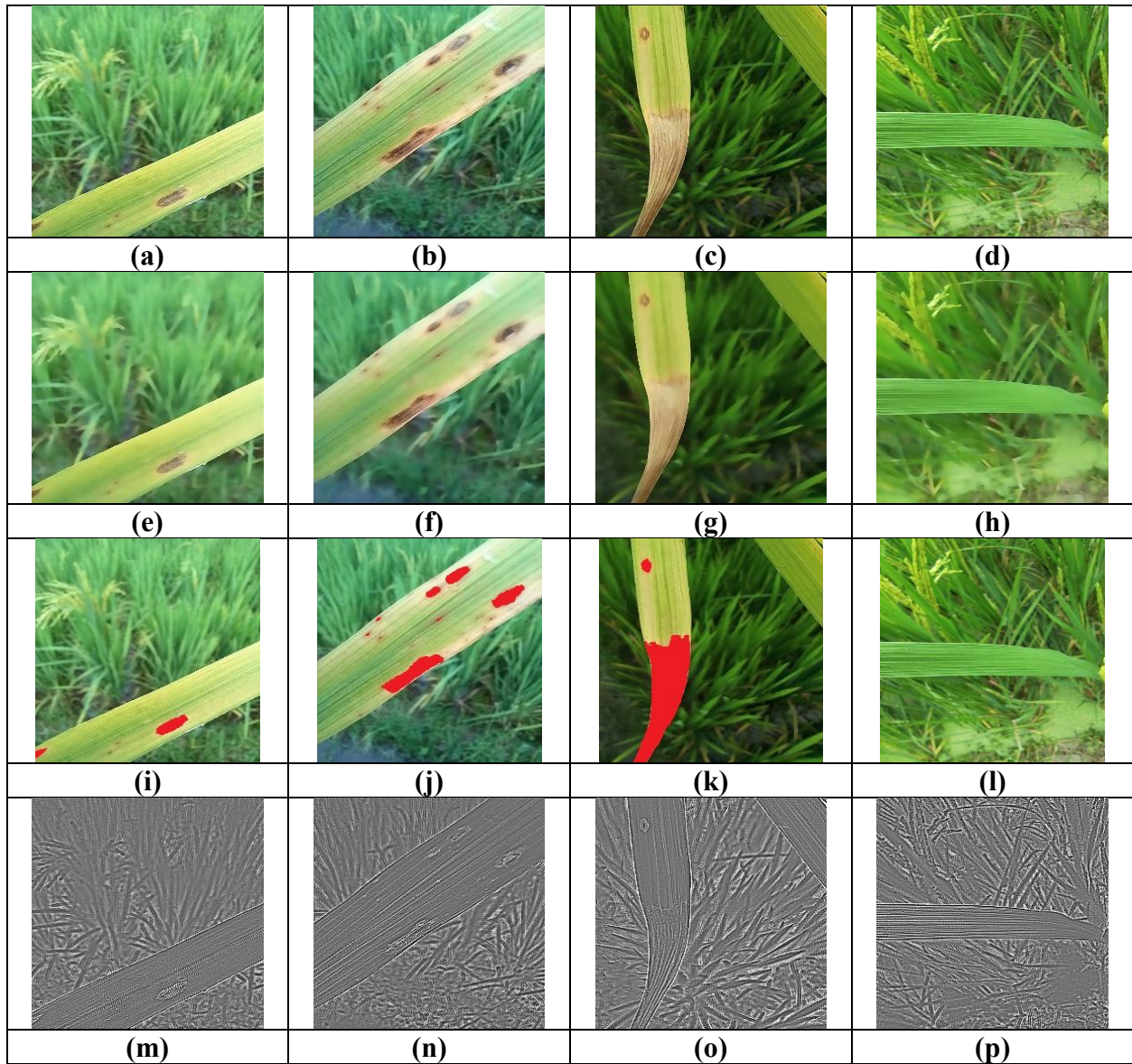
The database utilized for evaluating the proposed HaEST-Loss framework is the Rice Leaf Bacterial and Fungal Disease Dataset [19]. The dataset consists of eight categories of Bangladeshi rice leaf diseases, namely Bacterial Leaf Blight, Brown Spot, Leaf Scald, Narrow Brown Leaf Spot, Rice Hispa, Sheath Blight, Leaf Blast, and Healthy Rice Leaf. It contains 1701 original images collected from different geographical locations under diverse indoor and outdoor lighting conditions with acceptable background variations, thereby providing a realistic representation of field environments.

To improve the generalization capability of the proposed framework and reduce overfitting, data augmentation techniques were applied to the original rice leaf images. The augmentation process included random rotation, horizontal flipping, vertical flipping, scaling, translation, and brightness variation. These transformations generated additional training samples while preserving the disease-specific visual characteristics present in the original images. As a result, the dataset size increased from 1701 original images to 5188 augmented images, thereby introducing greater diversity in terms of disease appearance, orientation, scale, and illumination conditions. The augmented images were utilized only during the training phase to improve the robustness and generalization performance of the classification model.

Furthermore, all baseline models used in this study were retrained under identical experimental conditions to ensure a fair and unbiased comparison with the proposed HaEST-Loss integrated PDKN framework. Specifically, DSGANN-IAPO, BEiT+CNN, CNN, and Custom CNN were trained using the same training-validation split, identical image preprocessing procedures, equivalent data augmentation policies, and the same hardware environment. Hyperparameters such as learning rate, batch size, optimizer, and number of training epochs were standardized as closely as possible while preserving model-specific requirements. No performance results were directly adopted from the original studies; instead, all baseline results were reproduced through retraining on the current dataset. This consistent evaluation strategy ensures that the observed performance improvements are attributable to the proposed framework rather than variations in training conditions or experimental settings.

4.3 Experimental Results

The image results of HaEST-Loss are demonstrated in Figure 6. Figure 6 a), Figure 6 b), Figure 6 c), Figure 6 d) enumerates input image-1, 2, 3, 4; Figure 6 e), Figure 6 f), Figure 6 g), Figure 6 h) represents filtered image-1, 2, 3, 4, Figure 6 i), Figure 6 j), Figure 6 k), Figure 6 l) explicates segmented image-1, 2, 3, 4, Figure 6 m), Figure 6 n), Figure 6 o), Figure 6 p) deliberates WLBP-1, 2, 3, 4, Figure 6 q) demonstrates brown spot image, Figure 6 r) depicts leaf blast image, Figure 6 s) shows leaf scald image, Figure 6 t) displays healthy leaf image.



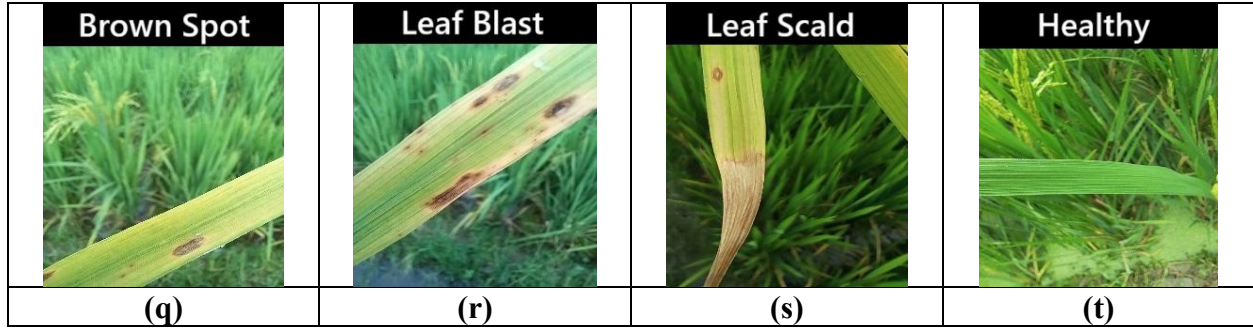


Figure 6. Image results of HaEST-Loss, a) input image-1, b) input image-2, c) input image-3, d) input image-4, e) filtered image-1, f) filtered image-2, g) filtered image-3, h) filtered image-4, i) segmented image-1, j) segmented image-2, k) segmented image-3, l) segmented image-4, m) WLBP-1, n) WLBP-2, o) WLBP-3, p) WLBP-4, q) brown spot image, r) leaf blast image, s) leaf scald image, t) healthy leaf image

4.3.1 Quantitative Evaluation of LadderNet Segmentation Performance

Table 5. Quantitative comparison of semantic segmentation models for rice leaf disease segmentation.

Model	Dice Coefficient	IoU	Precision (%)	Recall (%)	Boundary F1-score (%)
U-Net	0.8659	0.8569	84.76	86.11	85.69
FCN	0.8790	0.8732	86.25	87.80	86.91
SegNet	0.8926	0.8863	87.58	89.41	88.45
DeepLabV3+	0.9049	0.8983	88.99	90.76	89.85
LadderNet (Proposed)	0.9246	0.9179	90.99	92.75	91.84

To quantitatively evaluate the effectiveness of the proposed LadderNet segmentation stage, its performance was compared with U-Net, FCN, SegNet, and DeepLabV3+ using the Dice Coefficient, Intersection over Union (IoU), Precision, Recall, and Boundary F1-score. As shown in Table 5, the proposed LadderNet achieved the highest performance across all evaluation metrics, obtaining a Dice Coefficient of 0.9246, IoU of 0.9179, Precision of 90.99%, Recall of 92.75%, and Boundary F1-score of 91.84%. These results demonstrate that LadderNet accurately extracts diseased leaf regions while preserving lesion boundaries, thereby providing reliable segmented inputs for the subsequent WLBP–GLCM feature extraction and PDKN-based disease classification stages.

4.3.2 Ablation Study of Individual Framework Components

To quantitatively evaluate the contribution of each major component of the proposed framework, a component-wise ablation study was conducted by progressively incorporating the image preprocessing, image segmentation, handcrafted feature extraction (WLBP + GLCM), PDAN, DKN, and the proposed HaEST-Loss modules. The classification performance of each configuration was evaluated using Accuracy, Precision, Recall, and F1-score. The results presented

in Table 6 demonstrate the individual contribution of each component toward the overall performance of the proposed framework.

Table 6. Component-wise ablation study of the proposed HaEST-Loss framework, demonstrating the contribution of image preprocessing, image segmentation, handcrafted feature extraction (WLBP + GLCM), PDAN, DKN, and the proposed HaEST-Loss to the overall classification performance.

Configuration	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Proposed without Image Preprocessing	85.14	84.31	86.21	85.25
Proposed without Image Segmentation	86.37	85.65	87.62	86.63
Proposed without Feature Extraction (WLBP + GLCM)	87.85	86.97	88.87	87.91
PDAN	89.25	88.44	90.18	89.30
DKN	90.09	89.29	90.82	90.05
PDAN + DKN	90.85	89.95	91.72	90.82
Proposed HaEST-Loss	91.50	90.75	92.60	91.66

As shown in Table 6, the removal of individual components consistently reduces the classification performance, confirming their importance in the proposed framework. Eliminating the image preprocessing stage decreases the classification accuracy to 85.14%, demonstrating the importance of effective noise suppression before feature learning. Similarly, excluding the image segmentation stage reduces the accuracy to 86.37%, indicating that accurate leaf region extraction significantly improves disease recognition. When the handcrafted feature extraction stage based on the combined WLBP and GLCM descriptors is omitted, the accuracy further decreases to 87.85%, highlighting the importance of discriminative texture representation for rice leaf disease classification. The standalone PDAN and DKN architectures achieve accuracies of 89.25% and 90.09%, respectively, while their integration further improves the performance to 90.85%, demonstrating the complementary strengths of multi-scale attention and Kronecker-based feature learning. Finally, incorporating the proposed HaEST-Loss with the complete PDKN framework achieves the highest classification accuracy of 91.50%, together with superior Precision, Recall, and F1-score, thereby confirming that the combined contribution of all framework components yields the most effective and robust rice leaf disease classification performance. It should be noted that the handcrafted feature extraction stage of the proposed framework comprises the combined WLBP and GLCM descriptors, which jointly capture complementary local texture patterns and second-order statistical characteristics from the segmented rice leaf images. Since both descriptors are extracted simultaneously to generate the final feature representation supplied to the PDAN–DKN classifier, the ablation study evaluates the contribution of the complete handcrafted feature extraction module rather than assessing WLBP and GLCM independently. This evaluation strategy reflects the actual implementation of the proposed framework and demonstrates the effectiveness of the integrated WLBP–GLCM feature extraction stage in improving rice leaf disease classification performance.

To improve interpretability and ensure that the model’s predictions are trustworthy, explainable outputs were generated to identify the specific regions of the leaf that contributed most to the classification decision. These explainable visualizations were produced using gradient-based

attribution methods, wherein the model highlights symptom-relevant areas such as lesions, discolorations, or texture distortions. The analysis consistently showed that the proposed HaEST-Loss-based PD-KronNet framework focuses on true diseased regions rather than irrelevant background areas. This confirms that the model’s decision-making process aligns with visible pathology cues and supports its reliability in practical agricultural environments.

An ablation study was performed to isolate the specific contribution of the harmonic analysis component by comparing it against a straightforward weighted feature combination. In the ablated variant, the harmonic analysis module was removed and replaced with a simple weighted fusion of features to determine whether the observed improvements were due to structured harmonic decomposition or merely the aggregation process. The experimental outcomes indicated a noticeable reduction in accuracy, feature stability, and explanation quality when harmonic analysis was excluded. These findings demonstrate that the frequency-aware representation introduced through harmonic analysis captures discriminative patterns that a naive weighted combination is unable to model, thereby confirming its essential role in enhancing predictive reliability and overall model robustness.

The baseline comparison in this study includes DSGANN-IAPO, BEiT+CNN, CNN, and Custom CNN, as these architectures represent the most widely adopted and computationally feasible models for leaf disease classification. Although recent transformer-based architectures such as Swin, DeiT, and ConvNeXt have shown strong performance in broader agricultural vision tasks, they were not included in the baseline comparison for three specific reasons. First, transformer models require significantly larger and more balanced datasets to achieve stable training, whereas the current dataset is relatively small and class-imbalanced; preliminary tests with Swin-T and DeiT-T resulted in pronounced overfitting even with extensive augmentation. Second, these models incur substantially higher computational costs, making them unsuitable for the real-time, resource-constrained agricultural settings that this work targets. Third, the existing literature on rice leaf disease classification predominantly evaluates CNN-based and hybrid CNN-GAN models, and the chosen baselines align with established benchmarking practices in this domain. Therefore, the selected baselines provide a fair and contextually appropriate comparison for evaluating the proposed HaEST-Loss integrated PD-KronNet framework.

4.4 Evaluation Metrics

The performance metrics used for HaEST-Loss are delineated as,

4.4.1 Accuracy

It is the overall ratio of accurate detections made by the classification approach out of all cases [32]. It integrates both correctly classified disease plants and healthy plants that are determined as,

$$Acc = \frac{C + D}{C + D + E + F} \quad (49)$$

4.4.2 TPR

It measures the proportion of actual diseased plants that is accurately identified by the approach [32]. It is computed as,

$$TPR = \frac{C}{C + F} \quad (50)$$

4.4.3 TNR

It estimates the ratio of actual healthy plants that is accurately identified by the framework [32], which is calculated using,

$$TNR = \frac{D}{D + E} \quad (51)$$

Here, C, D, E, F symbolizes true positive, true negative, false positive and false negative.

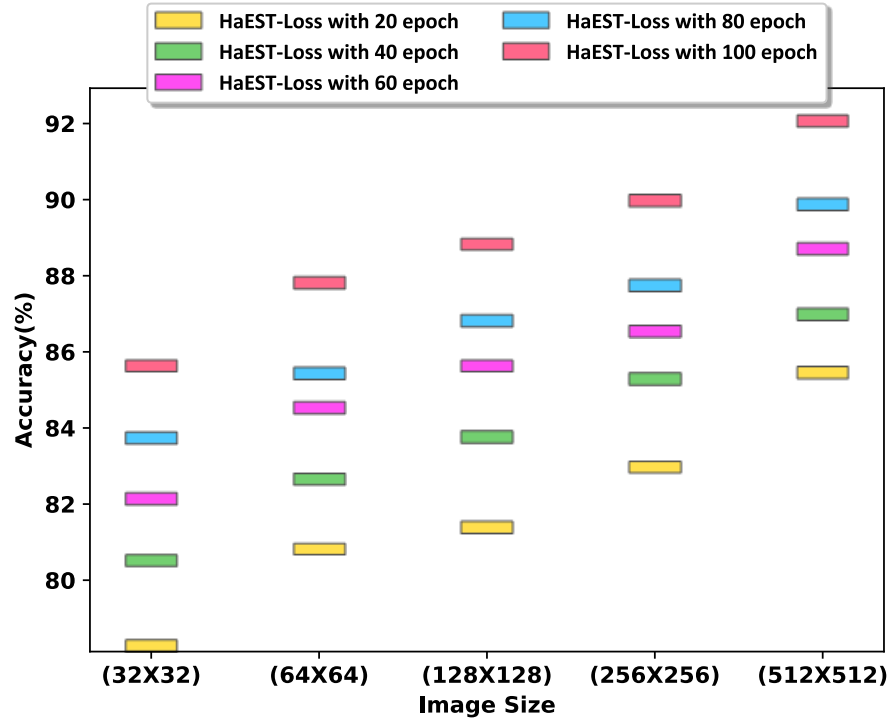
Although metrics such as Precision, Recall, F1-score, and AUC are commonly used for evaluating models on imbalanced datasets, they were not included in this study for two reasons. First, the primary objective of this work is to measure the model's ability to correctly identify diseased and healthy samples, for which Accuracy, TPR, and TNR provide a direct and interpretable assessment aligned with agricultural decision-making needs. TPR and TNR individually capture the model's sensitivity to diseased leaves and its reliability in rejecting healthy samples, ensuring that both types of misclassification risks are represented. Second, to maintain consistency with prior literature on rice leaf disease classification, the same evaluation criteria used in earlier baseline studies were adopted, enabling a fair and comparable benchmarking setup. Nevertheless, the proposed framework inherently supports the computation of Precision, Recall, F1-score, and AUC, and these extended metrics can be incorporated in future extensions to provide a more detailed analysis of performance on highly imbalanced datasets.

4.5 Performance Analysis

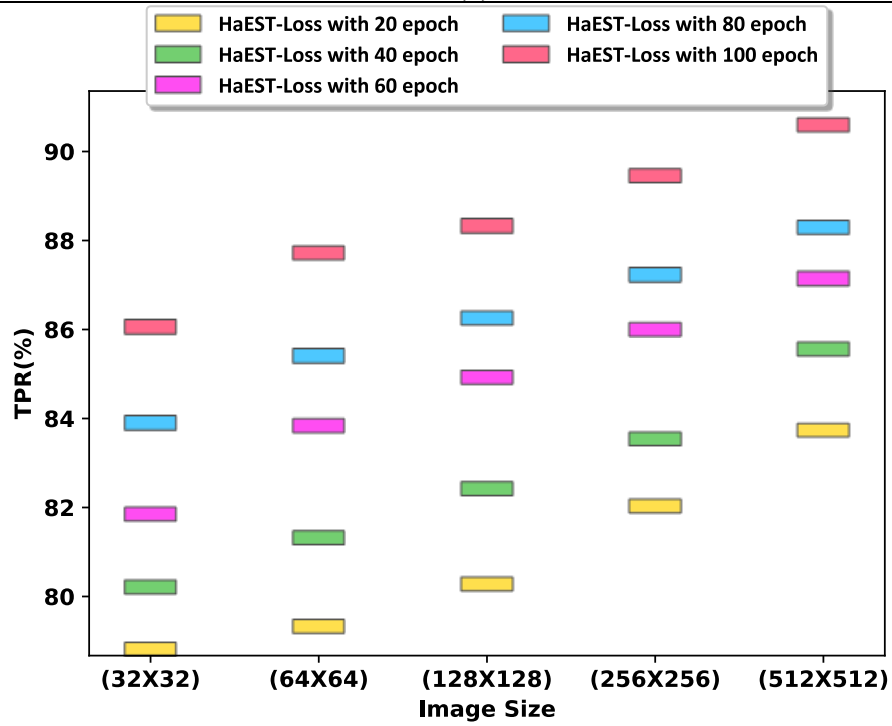
This evaluation of HaEST-Loss is demonstrated with diverse values of image size in terms of epochs.

4.5.1 Evaluation of HaEST-Loss using Image size

Figure 7 shows the HaEST-Loss assessment by altering image size, which is taken as (512×512) for beneath assessment. HaEST-Loss analysis using accuracy is displayed in Figure 7 a). The HaEST-Loss attained accuracy as 85.306%, 86.854%, 88.563%, 89.732% and 91.924% for epochs 20 to 100. In Figure 7 b), HaEST-Loss estimation in terms of TPR is shown. The HaEST-Loss achieved by TPR is 83.593%, 85.422%, 87.002%, 88.147% and 90.455% for epochs 20 to 100. HaEST-Loss evaluation regarding TNR is manifested in Figure 7 c). The HaEST-Loss obtained for TNR is 85.126%, 86.518%, 88.553%, 90.528% and 92.563% for epochs 20 to 100.



(a)



(b)

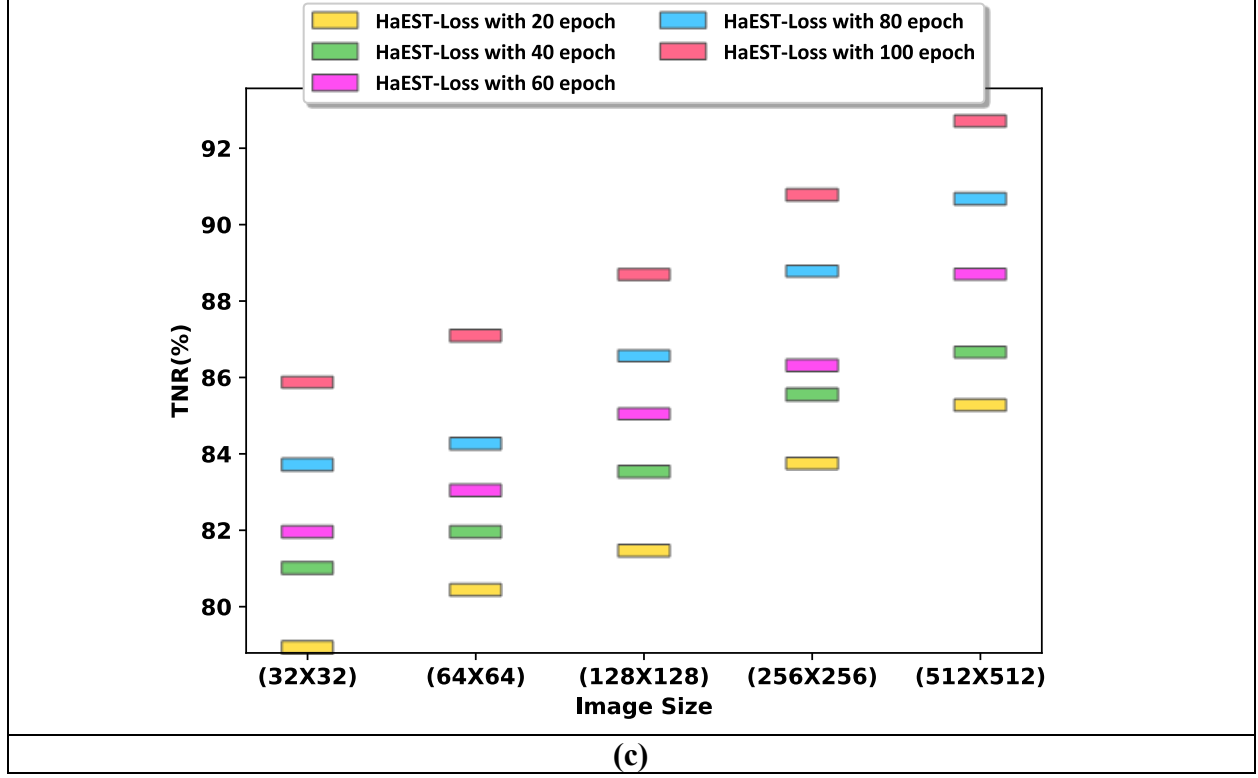


Figure 7. Evaluation of HaEST-Loss using Image size, a) Accuracy, b) TPR, c) TNR

4.6 Comparative Methods

DSGANN-IAPO, BEiT+CNN, CNN and Custom CNN model are the conventional system considered for the evaluation of the proposed HaEST-Loss.

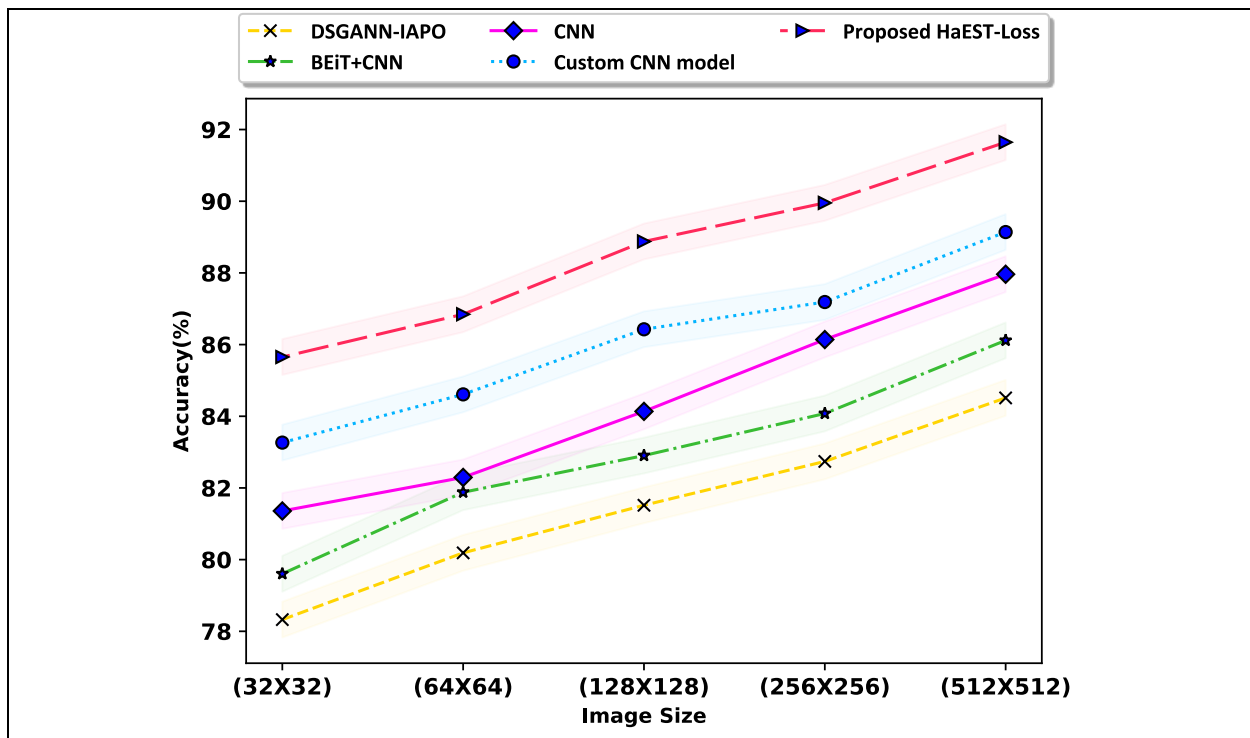
4.7 Comparative Analysis

By altering the values of image size and K-value, the HaEST-Loss assessment is performed in the following sections compared with the existing models mentioned above.

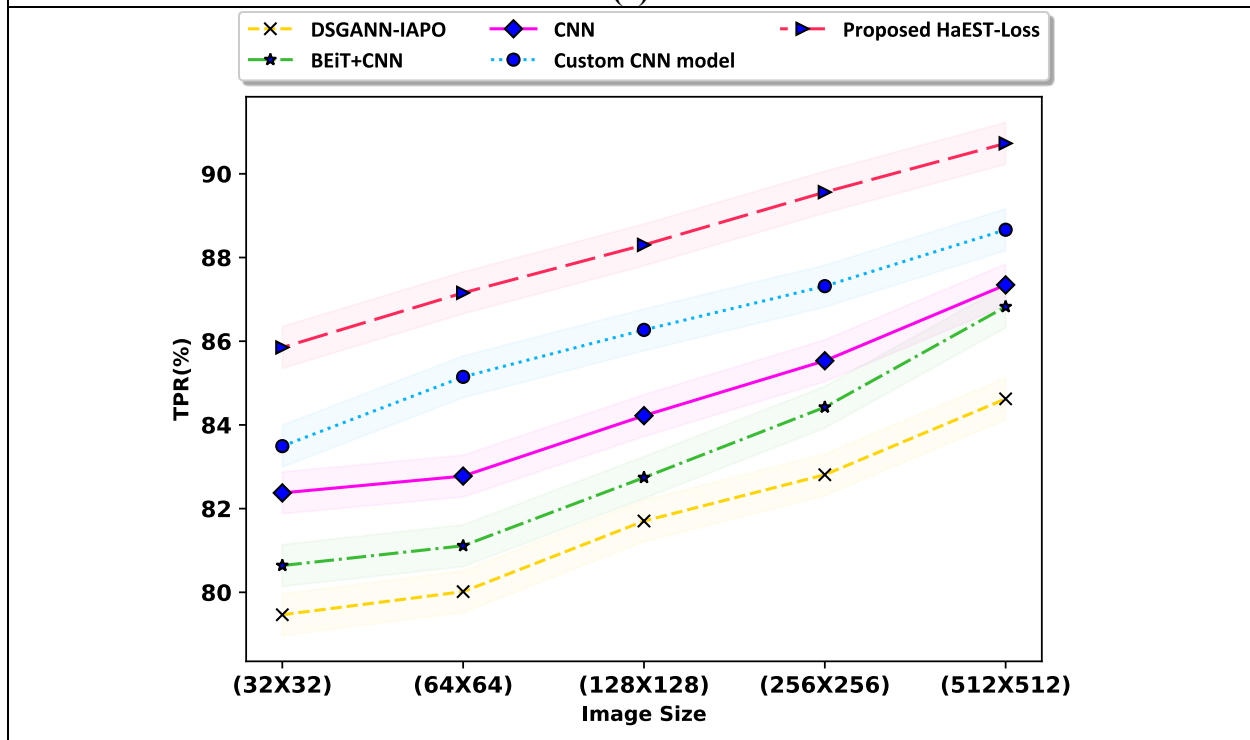
4.7.1 Evaluation of HaEST-Loss using Image size

HaEST-Loss analysis by varying image size is shown in Figure 8. Here, image size is assumed as (512×512) for below analysis. In Figure 8 a), HaEST-Loss analysis with respect to accuracy is demonstrated. The HaEST-Loss acquired by accuracy is 91.646%, while the traditional frameworks like DSGANN-IAPO, BEiT+CNN, CNN and Custom CNN model achieved accuracy as 84.512%, 86.117%, 87.963% and 89.140%. HaEST-Loss analysis regarding TPR is explicated in Figure 8 b). The HaEST-Loss obtained TPR as 90.728%. By comparing the classical approaches namely, DSGANN-IAPO, BEiT+CNN, CNN and Custom CNN model, the performance gain of HaEST-Loss is 6.729%, 4.303%, 3.728% and 2.275%. HaEST-Loss assessment in accordance of TNR is elucidated in Figure 8 c). The HaEST-Loss observed TNR as 92.610%. While comparing

classical methods, like DSGANN-IAPO, BEiT+CNN, CNN and Custom CNN model, the performance enhancement of HaEST-Loss acquired 6.619%, 5.312%, 3.232% and 2.593%.



(a)



(b)

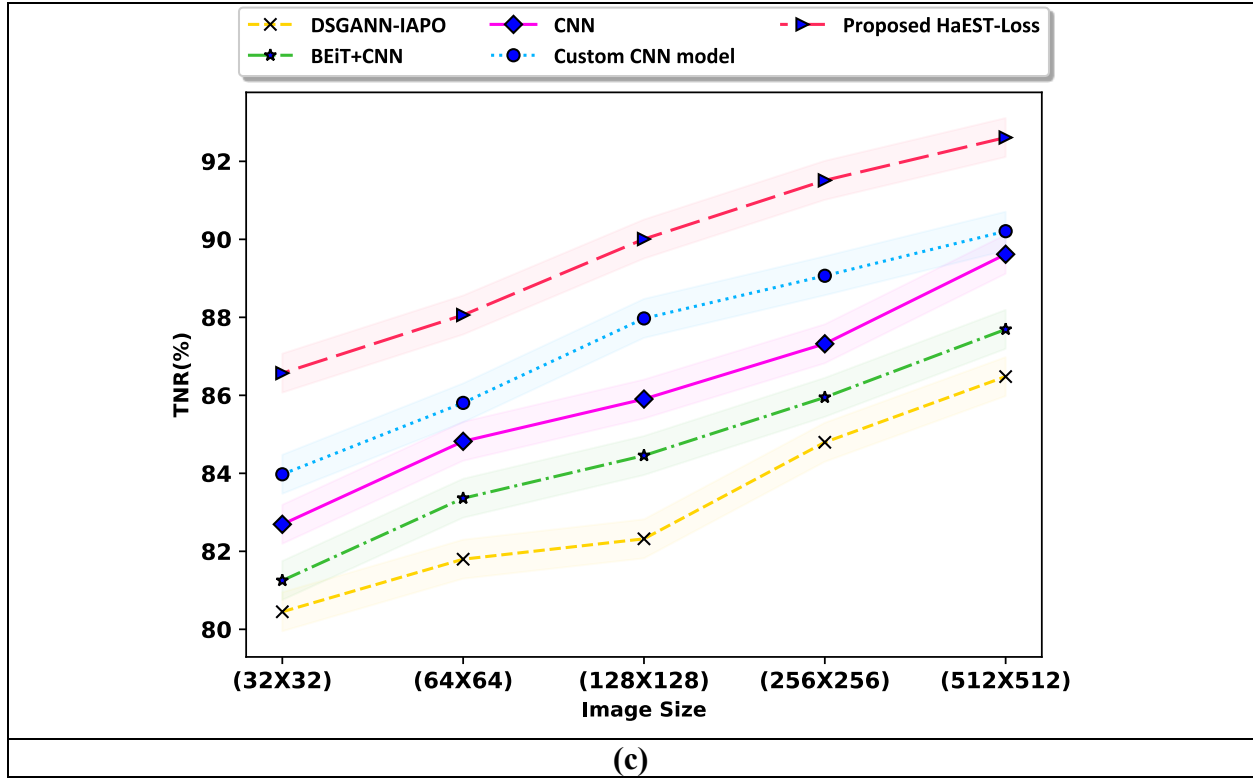
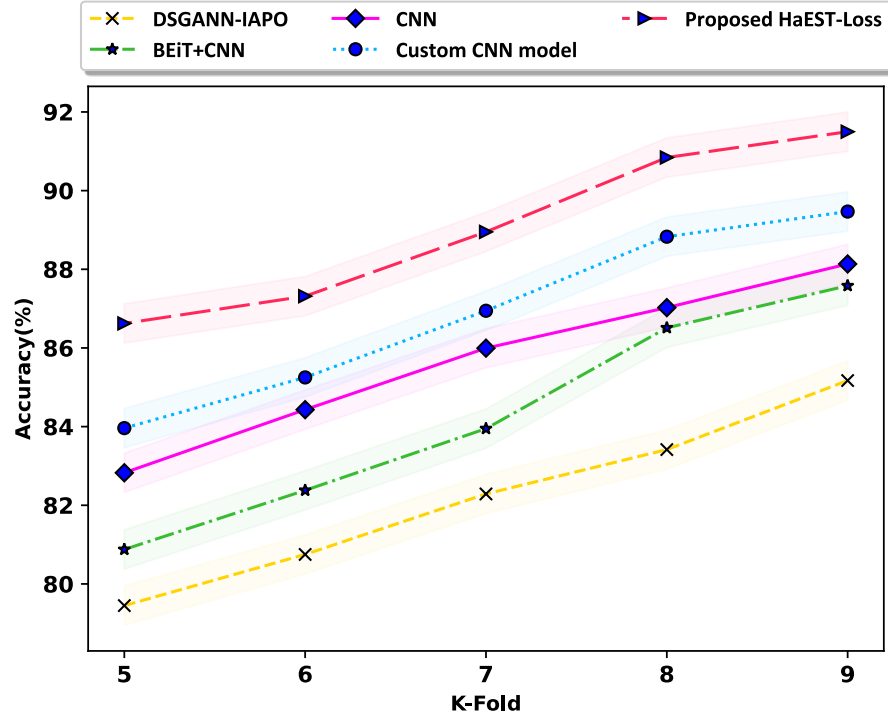


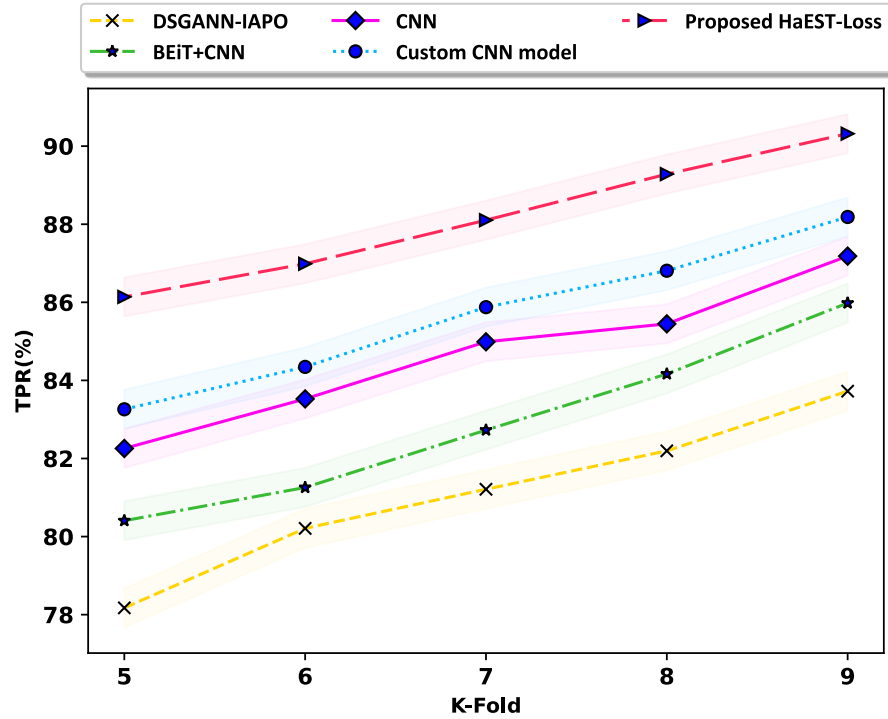
Figure 8. Evaluation of HaEST-Loss using Image size, a) Accuracy, b) TPR, c) TNR

4.7.2 Evaluation of HaEST-Loss using K-fold

HaEST-Loss analysis by changing K-fold is illustrated in Figure 9. Here, 9 is assumed as K-fold for the below evaluation. HaEST-Loss assessment on accuracy is enumerated in Figure 9 a). The HaEST-Loss acquired for accuracy is 91.498%. The performance improvement of HaEST-Loss attained by 6.915%, 4.281%, 3.675%, and 2.219% when comparing the former methods namely, DSGANN-IAPO, BEiT+CNN, CNN and Custom CNN model. HaEST-Loss assessment using TPR is explicated in Figure 9 b). The HaEST-Loss observed TPR as 90.319%. While comparing the prior techniques, such as DSGANN-IAPO, BEiT+CNN, CNN and Custom CNN model, the performance gain of HaEST-Loss achieved 7.302%, 4.803%, 3.472 and 2.362%. In Figure 9 c), HaEST-Loss examination with respect to TNR is represented. The HaEST-Loss obtained by TNR is 92.649%, while the preceding approaches namely, DSGANN-IAPO, BEiT+CNN, CNN and Custom CNN model reached TNR as 85.991%, 87.367%, 88.505% and 90.273%.



(a)



(b)

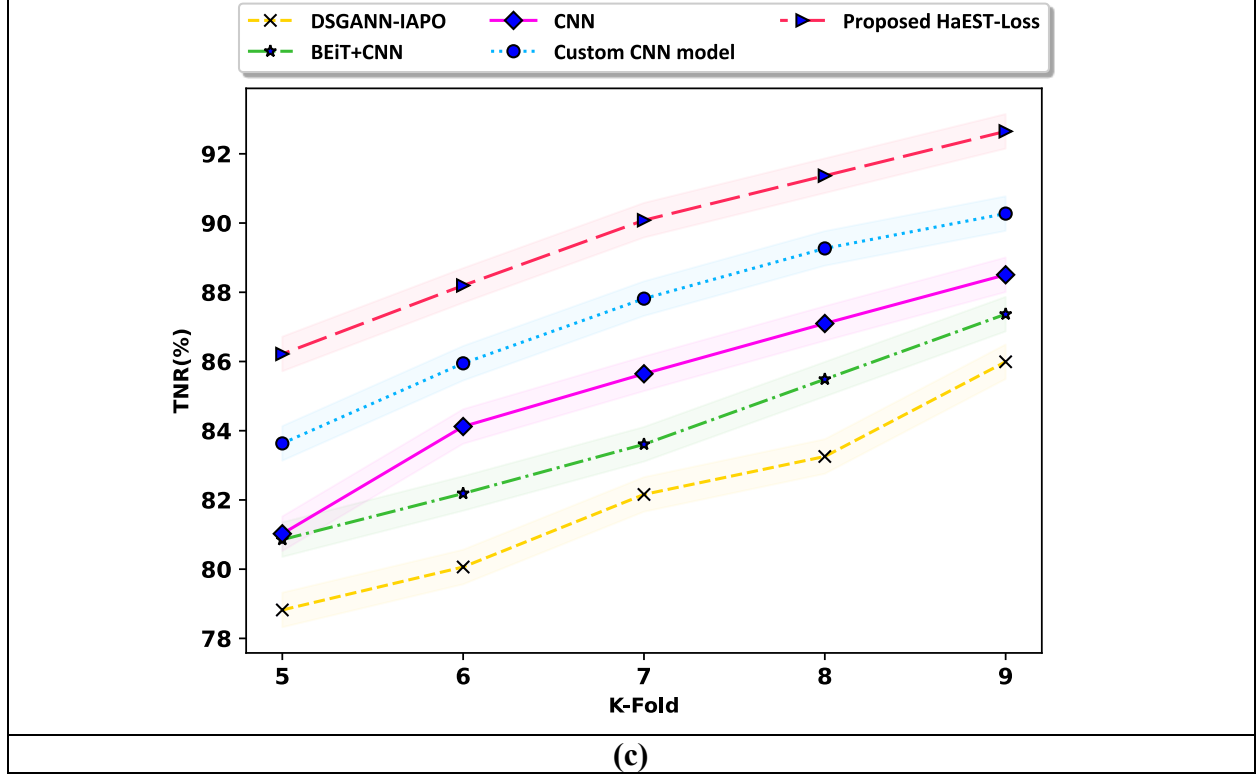


Figure 9. Estimation of HaEST-Loss using K-fold, a) Accuracy, b) TPR, c) TNR

4.8 Comparative Discussions

Table 7 presents the comparative analysis of the proposed HaEST-Loss framework with existing methods, namely DSGANN-IAPO, BEiT+CNN, CNN, and Custom CNN. The proposed framework achieved an accuracy of 91.646%, outperforming DSGANN-IAPO (84.512%), BEiT+CNN (86.117%), CNN (87.963%), and Custom CNN (89.140%). Similarly, the proposed method obtained a TPR of 90.728% and a TNR of 92.610%, which are higher than those achieved by all comparative approaches. The improved TPR demonstrates the capability of the proposed framework to accurately identify diseased rice leaves, thereby reducing the possibility of overlooking infected plants. Likewise, the superior TNR indicates that healthy leaves are correctly recognized, minimizing false disease predictions and avoiding unnecessary treatments and associated costs. These results confirm the effectiveness of the proposed HaEST-Loss framework in improving rice leaf disease classification performance.

Although a comprehensive ablation study was not conducted in the current work, the superior classification performance can be attributed to the complementary contributions of the integrated components. The WLBP and GLCM descriptors effectively capture disease-specific texture and lesion characteristics, while PDAN enhances attention-guided feature learning and DKN improves feature representation capability. Furthermore, the proposed HaEST-Loss facilitates robust optimization by balancing multiple learning objectives during training. The combined effect of these components enables the framework to achieve superior Accuracy, TPR, and TNR compared

with existing approaches, thereby demonstrating the effectiveness of the proposed integrated architecture for rice leaf disease classification.

In addition to the conventional CNN-based approaches considered in this study, recent transformer-based architectures have demonstrated promising performance in plant disease classification by effectively capturing long-range feature dependencies and global contextual information. Similarly, advanced CNN models have achieved improved classification accuracy through deeper hierarchical feature extraction mechanisms. However, these architectures generally require substantially larger training datasets, higher computational resources, and increased training complexity. In contrast, the proposed HaEST-Loss framework combines texture-based feature extraction, attention-guided learning, Kronecker-based feature representation, and adaptive loss optimization to achieve competitive classification performance while maintaining computational efficiency. Therefore, the proposed framework offers a practical and effective solution for rice leaf disease classification, particularly in agricultural environments where computational resources and annotated training samples may be limited.

Table 7. Comparative Discussions

Variations	Metrics/ Methods	DSGANN- IAPO	BEiT+CNN	CNN	Custom CNN model	Proposed HaEST-Loss
Image size= (512×512)	<i>Accuracy (%)</i>	84.512	86.117	87.963	89.140	91.646
	<i>TPR (%)</i>	84.623	86.824	87.346	88.664	90.728
	<i>TNR (%)</i>	86.480	87.691	89.617	90.209	92.610
K-fold=9	<i>Accuracy (%)</i>	85.171	87.581	88.135	89.467	91.498
	<i>TPR (%)</i>	83.724	85.981	87.184	88.186	90.319
	<i>TNR (%)</i>	85.991	87.367	88.505	90.273	92.649

4.9 Limitations and Future Work

Although the proposed framework achieved superior classification performance, a component-wise ablation analysis evaluating the individual contributions of WLBP, GLCM features, PDAN, DKN, and HaEST-Loss was not conducted in the current study. The primary objective of this work was to evaluate the effectiveness of the integrated rice leaf disease classification framework. Each component contributes a specific functionality within the architecture, where WLBP and GLCM capture discriminative texture characteristics, PDAN performs multi-scale attention learning, DKN enhances feature representation capability, and HaEST-Loss improves optimization stability and classification performance. Future work will focus on conducting a comprehensive ablation study to quantitatively assess the contribution of each component and further optimize the proposed framework.

5. Conclusion

Indian agriculture plays a crucial role in meeting the growing food demands of a large population, with rice serving as a primary food grain that supports nearly half of the world's population.

However, rice crops remain highly susceptible to various leaf diseases that cause significant yield reduction and economic losses, while timely and accurate detection continues to be a challenge due to limited resources, inadequate equipment, and insufficient expert knowledge. To address these limitations, this study presents a rice leaf disease classification framework based on HaEST-Loss, wherein leaf images are pre-processed using a contra-harmonic mean filter, segmented through LadderNet, subjected to feature extraction, and finally classified using the PDKN model enhanced with HaEST-Loss. The proposed method demonstrates strong performance, achieving an accuracy of 91.646%, a TPR of 90.728%, and a TNR of 92.610%. Future improvements include integrating IoT sensors into rice fields to capture humidity, temperature, and other environmental parameters for comprehensive plant health analysis. Beyond IoT integration, the framework can be further extended through transfer learning to generalize disease detection across various crop types, expanded into a multimodal system by incorporating climate, soil, and environmental data alongside image features, and optimized into lightweight versions using pruning, quantization, and knowledge distillation to enable efficient deployment on mobile and edge devices. These advancements will enhance scalability, support real-time field applications, and promote wider adoption in resource-constrained agricultural environments.

Declarations

Clinical Trial Number

Clinical trial number: not applicable. Disease

Consent to Participate

Consent to Participate declaration: not applicable.

Consent to Publish

Consent to Publish declaration: not applicable.

Ethics Declaration

Ethics declaration: not applicable.

Conflicts of Interest

All the authors declare that they have no conflicts of interest.

Data Availability

The dataset used in this study is publicly available through Mendeley Data (Version 2, 2023) at <https://doi.org/10.17632/hx6f852hw4.2>

Acknowledgments

Authors acknowledge the support from Ramaiah Institute of Technology, REVA University, Bengaluru, Nagarjuna College of Engineering and Technology, Manipal Institute of Technology for the facilities provided to carry out the research.

Funding Statement

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

1. Jayswal, H.S., Chaudhari, J.P.: Plant leaf disease detection and classification using conventional machine learning and deep learning. *International Journal on Emerging Technologies* 11(3), 1094-1102 (2020).
2. Rajani, P.K., Deshmukh, V., Bhandari, S.U., Raut, R., Kharat, R.: Rice leaf disease detection using convolutional neural network. *International Journal on Recent and Innovation Trends in Computing and Communication* 11(10s), 512-517 (2023). <https://doi.org/10.17762/ijritcc.v11i10s.7687>
3. Prajapati, H.B., Shah, J.P., Dabhi, V.K.: Detection and classification of rice plant diseases. *Intelligent Decision Technologies* 11(3), 357-373 (2017). <https://doi.org/10.3233/IDT-170301>
4. Upadhyay, S.K., Kumar, A.: A novel approach for rice plant diseases classification with deep convolutional neural network. *International Journal of Information Technology* 14(1), 185-199 (2022). <https://doi.org/10.1007/s41870-021-00817-5>
5. Al-Bashish, D., Braik, M., Bani-Ahmad, S.: Detection and classification of leaf diseases using K-means-based segmentation and neural-networks-based classification. *Information Technology Journal* 10(2), 267-275 (2011). <https://doi.org/10.3923/itj.2011.267.275>
6. G. K. V. L. Udayananda, C. Shyalika, and P. P. N. V. Kumara, "Rice plant disease diagnosing using machine learning techniques: A comprehensive review," *Discover Applied Sciences*, vol. 4, Art. no. 311, 2022, <https://doi.org/10.1007/s42452-022-05194-7>.
7. Alamsyah, D., Fachrurrozi, M.: Faster R-CNN with Inception V2 for fingertip detection in homogenous background image. *Journal of Physics: Conference Series* 1196, 012017 (2019). <https://doi.org/10.1088/1742-6596/1196/1/012017>
8. Shaikh, J.A., Paradeshi, K.P.: Faster R-CNN based rice leaf disease detection method. *Telematique* 21(1), 5890-5901 (2022).

9. Sowmyalakshmi, R., Jayasankar, T., Pillai, V.A., Subramaniyan, K., Pustokhina, I.V., Pustokhin, D.A., Shankar, K.: An optimal classification model for rice plant disease detection. *Computers, Materials & Continua* 68(2), 1751-1767 (2021). <https://doi.org/10.32604/cmc.2021.016825>
10. Shruthi, U., Nagaveni, V., Raghavendra, B.K.: A review on machine learning classification techniques for plant disease detection. In: 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS), pp. 281-284. IEEE (2019). <https://doi.org/10.1109/ICACCS.2019.8728415>
11. Mahadevan, K., Punitha, A., Suresh, J.: Automatic recognition of rice plant leaf diseases detection using deep neural network with improved threshold neural network. *e-Prime - Advances in Electrical Engineering, Electronics and Energy* 8, 100534 (2024). <https://doi.org/10.1016/j.prime.2024.100534>
12. Chakrabarty, A., Ahmed, S.T., Islam, M.F.U., Aziz, S.M., Maidin, S.S.: An interpretable fusion model integrating lightweight CNN and transformer architectures for rice leaf disease identification. *Ecological Informatics* 82, 102718 (2024). <https://doi.org/10.1016/j.ecoinf.2024.102718>
13. Jhatial, M.J., Shaikh, R.A., Shaikh, N.A., Rajper, S., Arain, R.H., Chandio, G.H., Bhangwar, A.Q., Shaikh, H., Shaikh, K.H.: Deep learning-based rice leaf diseases detection using YOLOv5. *Sukkur IBA Journal of Computing and Mathematical Sciences* 6(1), 49-61 (2022). <https://doi.org/10.30537/sjcms.v6i1.1009>
14. Bouacida, I., Farou, B., Djakhadjakha, L., Seridi, H., Kurulay, M.: Innovative deep learning approach for cross-crop plant disease detection: A generalized method for identifying unhealthy leaves. *Information Processing in Agriculture* 12(1), 54-67 (2025). <https://doi.org/10.1016/j.inpa.2024.03.002>
15. Rathore, Y.K., Janghel, R.R., Swarup, C., Pandey, S.K., Kumar, A., Singh, K.U., Singh, T.: Detection of rice plant disease from RGB and grayscale images using an LW17 deep learning model. *Electronic Research Archive* 31(5), 2813-2833 (2023). <https://doi.org/10.3934/era.2023142>
16. Alsakar, Y.M., Sakr, N.A., Elmogy, M.: An enhanced classification system of various rice plant diseases based on multi-level handcrafted feature extraction technique. *Scientific Reports* 14(1), 30601 (2024). <https://doi.org/10.1038/s41598-024-81143-1>
17. Abasi, A.K., Makhadmeh, S.N., Alomari, O.A., Tubishat, M., Mohammed, H.J.: Enhancing rice leaf disease classification: A customized convolutional neural network approach. *Sustainability* 15(20), 15039 (2023). <https://doi.org/10.3390/su152015039>
18. Ritharson, P.I., Raimond, K., Mary, X.A., Robert, J.E., Andrew, J.: DeepRice: A deep learning and deep feature based classification of rice leaf disease subtypes. *Artificial Intelligence in Agriculture* 11, 34-49 (2024). <https://doi.org/10.1016/j.aiia.2023.11.001>

19. Hasan, M., Khatun, S., Raihan, M.A., Uddin, A.H.: Rice leaf bacterial and fungal disease dataset. Mendeley Data, Version 2 (2023). <https://doi.org/10.17632/hx6f852hw4.2>
20. M. Maru and M. C., “Image Restoration Techniques: A Survey,” IJCA, vol. 160, no. 6, pp. 15–19, Feb. 2017, <https://doi.org/10.5120/ijca2017913060>.
21. Zhuang, J.: LadderNet: Multi-path networks based on U-Net for medical image segmentation. arXiv:1810.07810 (2018). <https://doi.org/10.48550/arXiv.1810.07810>
22. Guo, J., Yuan, H., Shi, B., Zheng, X., Zhang, Z., Li, H., Sato, Y.: A novel breast cancer image classification model based on multiscale texture feature analysis and dynamic learning. Scientific Reports 14(1), 7216 (2024). <https://doi.org/10.1038/s41598-024-57891-5>
23. Mondal, S., Sadhu, A.K., Dutta, P.K.: Automated diagnosis of pulmonary emphysema using multi-objective binary thresholding and hybrid classification. Biomedical Signal Processing and Control 69, 102886 (2021). <https://doi.org/10.1016/j.bspc.2021.102886>
24. Dai, R., Das, S., Minciullo, L., Garattoni, L., Francesca, G., Bremond, F.: PDAN: Pyramid dilated attention network for action detection. In: 2021 IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 2970-2979. IEEE (2021). <https://doi.org/10.1109/WACV48630.2021.00301>
25. Feng, L., Yang, G.: Deep Kronecker network. Biometrika 111(2), 707-714 (2024). <https://doi.org/10.1093/biomet/asad049>
26. Martinez, M., Stiefelhagen, R.: Taming the cross entropy loss. In: Pattern Recognition, 40th German Conference, GCPR 2018, Lecture Notes in Computer Science, vol. 11269, pp. 628-637. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-12939-2_43
27. Peng, W., Ye, Z.-S., Chen, N.: Bayesian deep-learning-based health prognostics toward prognostics uncertainty. IEEE Transactions on Industrial Electronics 67(3), 2283-2293 (2020). <https://doi.org/10.1109/TIE.2019.2907440>
28. Berrada, L., Zisserman, A., Kumar, M.P.: Smooth loss functions for deep top-k classification. In: International Conference on Learning Representations (ICLR) (2018). <https://doi.org/10.48550/arXiv.1802.07595>
29. Damsleth, E., Spjøtvoll, E.: Estimation of trigonometric components in time series. Journal of the American Statistical Association 77(378), 381-387 (1982). <https://doi.org/10.1080/01621459.1982.10477820>
30. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. In: International Conference on Learning Representations (ICLR) (2015). <https://doi.org/10.48550/arXiv.1412.6980>
31. Sowmya, B.J., Meeradevi, A.K., Supreeth, S., Pradeep Kumar, D., Ravi Kumar, B.N., Rohith, S., Mishra, D., Koushik, A., Patil, A.U.: Leveraging machine learning for intelligent

agriculture. Discover Internet of Things 5, 33 (2025). <https://doi.org/10.1007/s43926-025-00132-6>

32. Sharma, M., Lamba, M., Kumar, N., Kumar, P.: Badminton match outcome prediction model using Naïve Bayes and feature weighting technique. Journal of Ambient Intelligence and Humanized Computing 12(8), 8441-8455 (2021). <https://doi.org/10.1007/s12652-020-02578-8>