J. Mountain Res. P-ISSN: 0974-3030, E-ISSN: 2582-5011
Vol. 21(S2), (2026) 53-61

DOI: https://doi.org/10.51220/jmr.v21-S2.6

Optimized Deep Learning and Mobile Deployment for Real-Time Sugarcane Leaf Disease Diagnosis in Precision Agriculture

Madhulika Mittal • Md. Iqbal

Department of Computer Science and Engineering, Quantum University, Roorkee, Uttarakhand, India

*Corresponding Author's Email Id: madhulikamittal24@gmail.com

Received: 20.02.2025; Revised: 23.11.2025; Accepted: 8.1.2026

Abstract: This study proposes a novel attention-driven multilevel deep learning framework for robust sugarcane leaf disease classification, focusing on precise plant pathology identification. The architecture incorporates spatial and channel attention mechanisms to enhance feature saliency detection while hierarchically fusing low-to-high-level visual patterns. Evaluated on a proprietary dataset, the model achieves an accuracy of 86.53%, outperforming state-of-the-art benchmarks including VGG19, ResNet50, XceptionNet, and EfficientNet_B7. The results underscore the significance of multiscale feature integration in image classification tasks and demonstrate improved performance even with limited training data. By enabling early-stage disease recognition, this approach proposes a practical solution for timely agricultural disease management, potentially reducing crop losses. Furthermore, the deployment of the SLDDCNN model in an Android application illustrates the feasibility of smartphone-based disease diagnostics in agriculture, highlighting mobile technology's role in democratizing access to precision farming tools.

Keywords: Deep learning • Disease classification • Agriculture • Sugarcane database

Introduction

The rapid evolution of digital technologies has streamlined acquisition of visual data, catalyzing the development of smartphone-driven real-time systems across diverse sectors such as healthcare, finance, and agriculture. In farming, digitization has revolutionized practices through smart cultivation, growth monitoring, and automated harvesting. Despite these advancements, plant diseases persist as a critical challenge, undermining crop yields and global food security. Early detection and intervention are vital to mitigating these losses (Amos, et. al.,2014) . While farmers increasingly rely on digital platforms to consult experts, manual disease identification remains prone to inaccuracies and inefficiencies. Consequently, researchers have prioritized automated solutions leveraging image processing and machine learning to enhance diagnostic speed and precision (Vijai et.al.,2022).

Recent strides in deep learning, particularly convolutional neural networks (CNNs), have demonstrated superior performance in image classification tasks due to their robust feature extraction and learning capabilities. Architectures such as SegNet, FCN, and U-Net have addressed specific use cases, outperforming traditional methods in efficiency and scalability. However, deeper CNNs often grapple with vanishing gradients, a limitation mitigated by residual networks (ResNets) through skip connections (Bock, et.al, 2010). The efficacy of these models hinges on comprehensive datasets, though existing repositories like Plant Village—comprising lab-condition images—face criticism for limited real-world applicability. To bridge this gap, studies emphasize field-captured datasets with environmental variability, while transfer learning aids in scenarios with sparse data.

Current research focuses on refining model performance and reducing complexity. Innovations include integrating attention mechanisms with ResNets to prioritize salient features, as seen in channel attention hybrids and pyramidal feature extraction (Pagoal, Yansheng, et.al., 2009). These adaptations

DOI: https://doi.org/10.51220/jmr.v21-S2.6

enhance noise resilience and classification accuracy. Nevertheless, the scarcity of real-time field imagery remains a hurdle, as controlled-environment datasets often fail in practical settings.

This study contributes three key elements: (1) A Sugarcane based Leaf Disease Detection CNN (SLDDCNN) that synthesizes hierarchical features for sugarcane disease recognition while maintaining operational simplicity; (2) A residual architecture fused with attention blocks to optimize disease classification; and (3) A real-time Android application deploying the proposed model.

Methodology

In this, we provide an overview of the architecture of the proposed method before discussing its key components in detail. SLDDCNN is proposed method to extract deep features and classify sugarcane disease. All the parts of the proposed approach will be detailed in the following subsections.

Dataset: The study collected pictures of sugarcane leaves from different cultivated fields at different weather conditions. The original dataset for sugarcane leaf disease composed of five classes, which include four classes of sick sugarcane leaves and one class of healthy sugarcane leaf samples. For the implementation, we used total 2600 photos (approximately 500 in each category). Summary of the sugarcane database, with the sample picture distributions at all of the classes, is presented in Table 1. The source images used were downscaled to 256x256 for computational purposes. To improve overall performance, the dataset was divided into training, validation, and test sets in a ratio of 70:15:15. This splitting gave a sufficient number of images per category. We generated a test set, training set, and validation set of images, with uniform parameters throughout all permutations in the experiments so the results could be easier to interpret. Training data were used for training, and the validation data were derived from it and used to evaluate the model, while the test data were used to measure predictive ability. An example from each class is presented in Fig. 1, and summary statistics for the database are reported in Table 1. The database is publicly available to the research community (Daphel, 2022). To set a baseline for the model and database, all experimentation in the following section is done sans augmentation.

Fig. 1a depicts a rust-infected leaf, its brown lesions clearly visible upon the olive green surface. Alongside sits a mosaic-afflicted foliage in Fig. 1b, its mottled yellow and light green patches betraying the viral invasion within. Health remains in Fig. 1c, its deep green hue unmarred by imperfection. Red rot brings ruin in Fig. 1d, the leaf decaying from the outer edges inward, crimson in color. A sickly yellow pervades the leaf displayed in Fig. 1e, foreshadowing its demise. These sampled images from the database represent the variations encompassed within each designated class: rust, mosaic, healthy, red rot, and yellow leaf, as seen in Fig. 1.

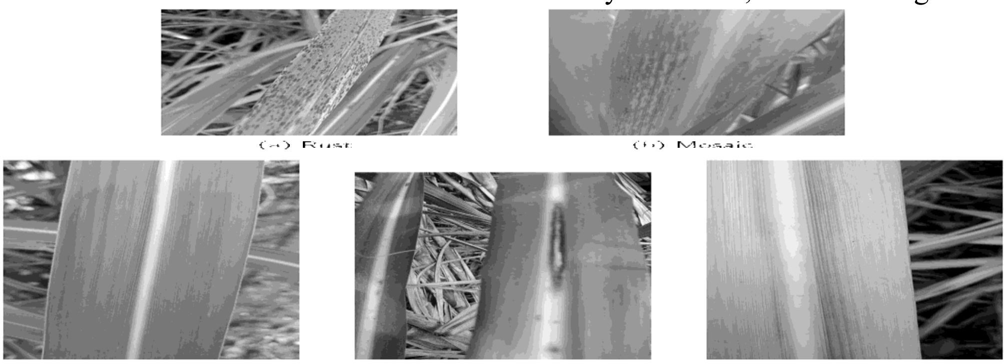

Fig. 1. Images of each class: (a) Rust, (b) Mosaic, (c) Healthy, (d) Red rot, (e) Yellow leaf.

 DOI: https://doi.org/10.51220/jmr.v21-S2.6

Table 1: Dataset statistics: Originally collected.

Class_Name	Total_Images	Outcome (in %)
Healthy	530	20.2%
Rust	521	20.0%
Red rot	519	20.2%
Yellow	510	19.7%
Mosaic	520	19.9%
Total	2600	100%

Sugarcane Leaf Disease Detection Convolutional Neural Network (SLDCNN)

CNN improve picture classification performance by stacking convolutional layers, but this comes at the expense of more parameters and computational complexity. To improve picture classification performance, a feature extraction module composed of SLDCNN blocks from the perspective of network breadth has been developed.

When classifying data, the suggested design, as shown in Fig. 2, takes into account both mid-level and high-level attributes. These feature maps are extracted at each level of the network's residual blocks. The information from the first residual block is not used right away since it comprises basic filters for classifying sugarcane diseases. It is well known that the early layers of a deep CNN are capable of extracting low-level features or the fundamental structural elements of a picture. The basic and important features of sugarcane illness cannot be well conveyed by the networks at this level, which can only learn and understand simple textures. The network's first layers capture low-level features, which roughly include edges, corners, and color patches. The mid-level features, which have textures and object forms, are patterns that the network's middle learns. High-level elements aid in clarifying the situation depicted in the picture.

However, midlevel and high-level characteristics (from the second ResBlock to the third ResBlock) become more complicated by combining learned information from the previous levels.

Additionally, the high-level layers of the third ResBlock are essential for detecting sugarcane illness using global characteristics, while the mid-level layers generate local features. Keypoints, descriptors, and some innate pattern resembling local binary patterns are examples of local features. Moments and activations are examples of global features. Other efforts usually employ the CNN's global properties, or the complete picture characteristic, for categorization. The proposed SLDDCNN in this study carefully combines local and global parameters to identify sugarcane illness. This technique aids in drawing attention to key elements of the supplied image. Our networks, in contrast to traditional CNNs, offer numerous crucial characteristics that strengthen the system against various sugarcane illnesses and image quality issues, respectively.

As illustrated in Fig. 2, the inclusion of every block in the suggested approach is essential for lowering the loss of partial feature information pertaining to plant diseases and raising recognition rates. A residual structure has been used to solve problems with gradient disappearance and model optimization that arise when there are several convolution layers. Because it can solve the issue of growing network depth deterioration, the residual structure is beneficial. With a minor adjustment in weight values, the network may be returned to its ideal state using the mapping relationship between the convolutional layers of the residual structure, which is $H_1(x) = F_1(x) + x$.

Sugarcane disease spots have low intra-class variances and high inter-class variations, which makes it difficult to tell one type of

DOI: https://doi.org/10.51220/jmr.v21-S2.6

sample from another in a deep CNN. In order to precisely detect the disease area in the image and raise the model's identification accuracy, we suggest inserting the improved attention mechanism after each residual block of the network. By doing this, the network classification performance can be less affected by redundant information in the input image. A channel and spatial attention module are also included in the suggested approach, and they are simple to integrate into the suggested network. The attention mechanism is explained in detail in the following subsections.

Finally, using an element-wise addition operation, the attentive features of three blocks were combined and fed into fully-connected layers as input for additional feature extraction and categorization. It is necessary to keep feature maps the same size in order to merge each level of deep attentive features. A down sampling technique is used in conjunction with convolutional and max-pooling operations to extract more intuitive characteristics that are pertinent to sugarcane disease while maintaining the constant size of each block's attentive features. Fig. 2 depicts the three-level down sampling method's architecture. The suggested architecture has two fully-connected layers that are 512 and 256 in size. Softmax, the final layer of SLDDCNN, has five classes.

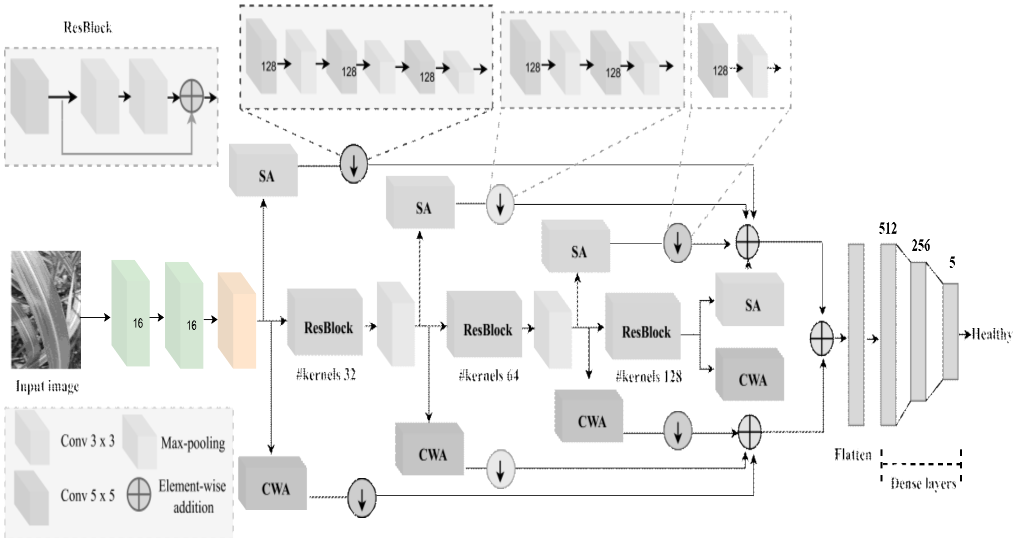

Fig. 2. Detailed description of proposed approach

Attention Mechanism

One of the key functions of the Human Visual System (HVS) is the capacity for attention. The whole scene is not processed at once by HVS. In spite of this, proposed approach concentrates on specific infected spots in order to extract detailed information from the picture and produce useful findings. Proposed approach makes effective use of both spatial and channel attention.

Focus on the channel

A channel attention feature map is produced by considering the inter-channel relationship of feature space. As a feature detector, each channel in the feature space assesses "what" is important in the scene based on its attention. For the channel attention to be calculated effectively, the spatial dimension of the feature space needs to be shrunk. Together, average pooling and maximum pooling are employed to compile spatial data. The representational capability of networks is greatly enhanced by this technique (Woo, Park, et.al., 2018).

The features acquired by max pooling and average pooling are represented by two distinct spatial descriptors, $F_{\max}{}^{channel}$ and $F_{avg}{}^{channel}$, respectively. To generate the channel attention map $Map_{\text{channel}} \in R^{C\times1\times1}$, those descriptors are fed into a standard multi-layer perceptron network. Setting the hidden

J. Mountain Res. P-ISSN: 0974-3030, E-ISSN: 2582-5011
Vol. 21(S2), (2026) 53-61
DOI: https://doi.org/10.51220/jmr.v21-S2.6

activation size $C \times 1 \times 1$ to R r, where r is the reduction factor, reduces parameter overhead. To create output features, the results of a multi-layer perceptron network are combined utilizing summarization by element. Equation (1) provides the mathematical computation for the channel attention map.

$$Map_c(F) = sigmoid(MLP(AveragePool(F)) + MLP(MaxPool(F))) = sigmoid(Wt_1(Wt_0(F^{channel}_{average})) + Wt_1(Wt_0(F^{channel}_{max}))) \quad (1)$$

in which $Wt0 \in R^{C/r \times C}$ and $Wt1 \in R^{C \times C/r}$, Inputs and the ReLU activation share weights, followed by W $t0$. Figure 3 displays the detailed channel-wise attention method.

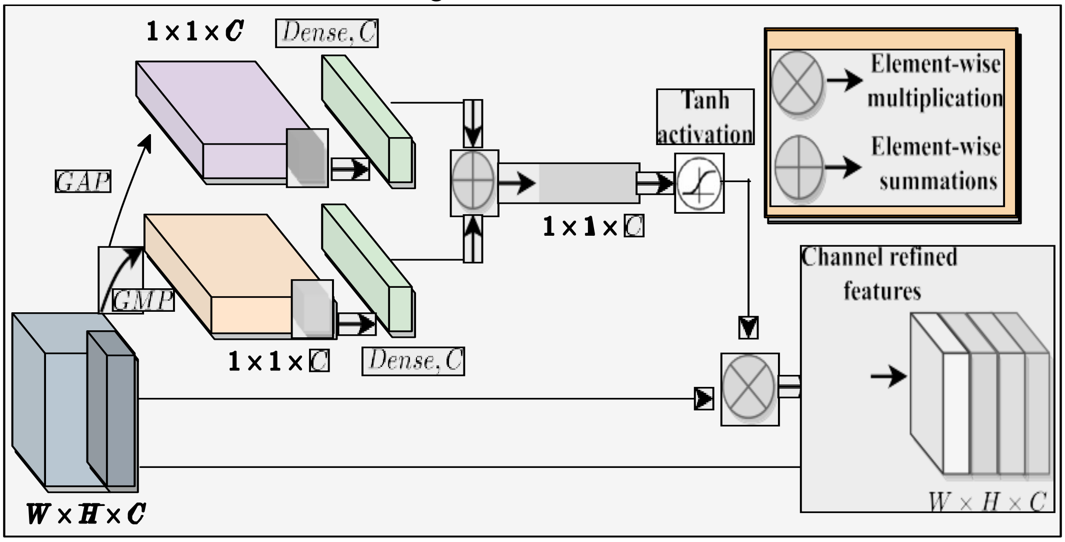

Fig. 3. Channel wise attention.

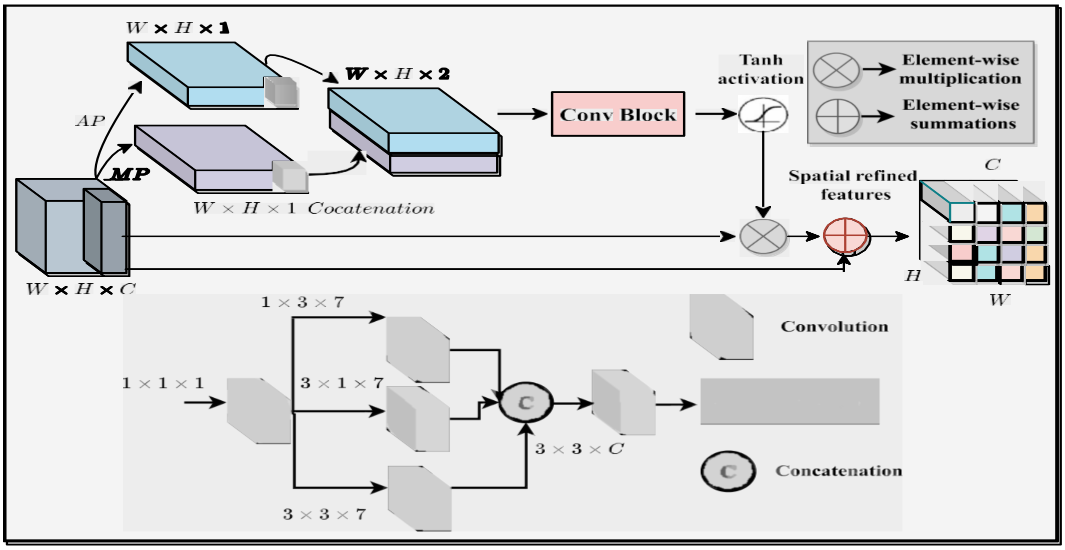

Fig 4: Spatial Attention

SLDDCNN Parameters

Regardless of how individual characteristics would change the network's optimal parameters, all significant network architecture parameters were specified to show that the network is ubiquitous.

- The parameter is set as the starting value, and Adam optimization is selected as the network's optimizer algorithm.
- For the sugarcane classification task, categorical cross-entropy is selected as the loss function.
- A learning rate of 0.0001 is established.

- The activation function LeakyReLU is used in every convolutional and fully-connected layer. The value of alpha was set at 0.02.

Real-time system proposal

A rapidly developing field of study creates mobile applications for agricultural disease classification using machine learning and deep learning techniques. One of the most popular tools for this is TensorFlow Lite, a condensed version of the well-known machine learning framework TensorFlow. TensorFlow Lite can be used as a backend in Android Studio by developers to create accurate and efficient models for categorizing different sugarcane illnesses. The Android project folder now contains the TFLite model file (.tflite). Using a TFLite interpreter, the provided model is then loaded into an Android application, and predictions are created based on the input data that was taken using the smartphone's camera. In this way, the model is trained using CNN and other supervised learning algorithms, which also involves deploying the model on mobile devices with limited resources and preparing the input data (Madhulika Mittal & et.al., 2022). TensorFlow Lite provides a number of features to help you maximize the model's performance, including quantization algorithms that help decrease the model's size and speed it up. The end product is a smartphone application that enables agronomists and growers to quickly and accurately identify sugarcane diseases, improving crop management and leading to higher yields and reduced expenses. Fig. 5 provides a quick overview of the suggested real-time system utilizing an Android application.

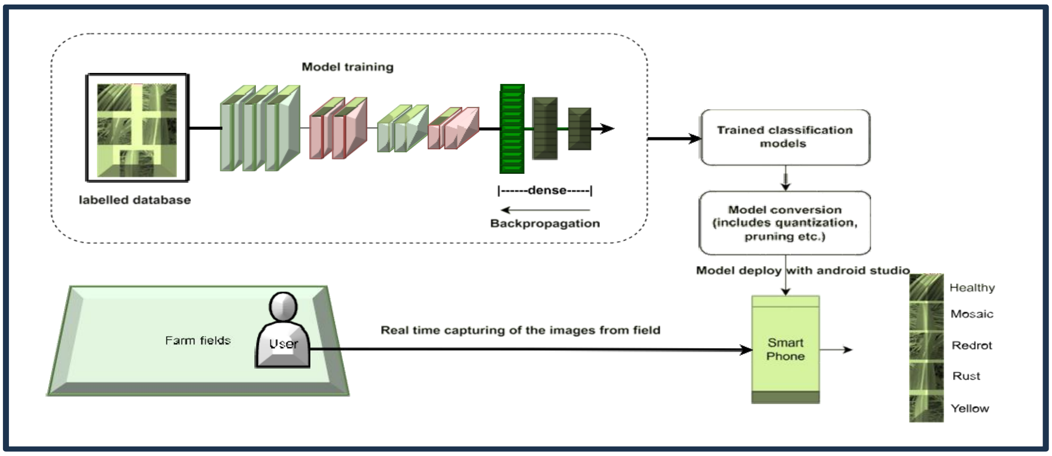

Fig. 5: A complete sugarcane plant disease diagnosis system block diagram.

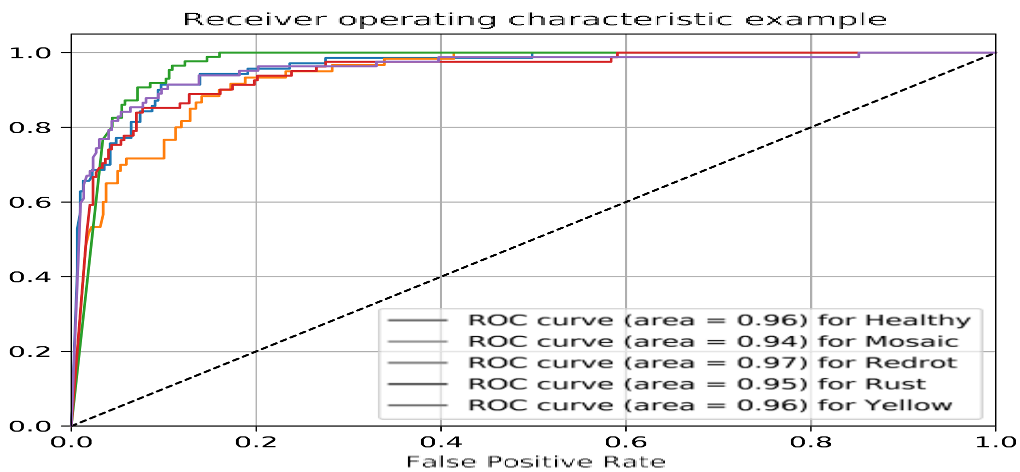

Fig. 6. ROC curves with SLDDCNN.

J. Mountain Res. P-ISSN: 0974-3030, E-ISSN: 2582-5011
Vol. 21(S2), (2026) 53-61

DOI: https://doi.org/10.51220/jmr.v21-S2.6

The suggested SLDDCNN model is contrasted with a number of benchmark deep learning architectures that are currently in use. The comparison was conducted using four popular deep learning architectures: VGGNet, ResNet50-v2, XceptionNet, and EfficientNet-B7 (Simonyan, 2014). The first layers of these structures were trained using a transfer learning-based methodology, and they were then refined using a unique sugarcane leaf disease database. Table 2 provides information on these models' performance. Because of deep interpretations and efficient handling of vanishing gradients, ResNet outperforms the other deep learning architectures. The suggested SLDDCNN design outperforms the baseline architectures in spite of this. A receiver operating curve is commonly used in classification as shown in Fig. 8. It plots the classifiers false positive rate and the true positive rate. A perfect classifier will have ROC-area under a curve equal to 1 and a random classifier will have AUC equal to 0.5. It is given in Fig. 6 that AUC values are near 1, indicating SLDDCNN is a good classifier. This finding suggests that a pyramid- based network that integrates channel and spatial attention pays more attention to the salient part of the images and neglects the noise component. This network fares comparatively well even though the network is merely trained on small databases. It is worth noting that, to a certain extent required performances can be achieved with architectural modifications even if the database is small. Fig. 7, gives a pictorial comparison of all methods used in the experimentation and also highlights the slight improvement observed with the proposed architecture.

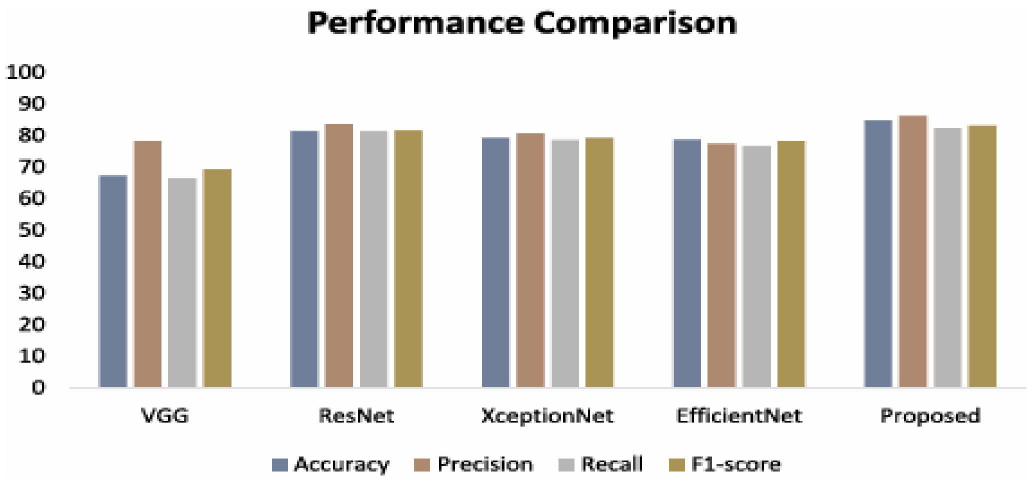

Fig. 7. Performance comparison of SLDDCNN framework with various architectures.

Table 2: Comparison of SLDDCNN framework results with different benchmark methodologies.

Model	Method	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
VGG [12]	**VGG19**	68.34	78.33	67.27	69.34
ResNet [13]	**ResNet50_v2**	81.42	83.64	81.33	81.61
XceptionNet [14]	**XceptionNet**	79.21	80.6	78.46	79.15
EfficientNet [15]	**EfficientNet_B7**	78.66	77.43	76.56	78.31
Proposed	**SLDDCNN**	86.53	86.34	82.32	83.15

The performance of model quantization

The following is the general method for quantifying the models utilized in the work:

1. TensorFlow 2.0 API was used to train the models using the default float32 datatype.

J. Mountain Res. P-ISSN: 0974-3030, E-ISSN: 2582-5011
Vol. 21(S2), (2026) 53-61
DOI: https://doi.org/10.51220/jmr.v21-S2.6

2. Afterwards, models were quantized to float16 by directly casting tensors to float16 using TensorFlow's API.
3. To further minimize the size, tensors were cast to int8 using TensorFlow's API.

Instead of using selective layers, the aforementioned options were used to quantize the entire model in the manner described above, and the quantization strategy was consistent throughout all layers. This procedure shrinks the model's size and could affect how well it performs. The necessary model size and the intended performance must be traded off.

Table 3: Real time performance and Model quantization results of proposed model against simple CNN.

model name	Converted to	size	Accuracy	inference time in ms
simple CNN	Original	84.2 MB	78.23%	865
	float 32	47.59 MB	75.23%	655
	float 16	26.64 MB	73.44%	435
	int 8	7.34 MB	71.23%	375
Proposed model: SLDDCNN	Original	153 MB	87.54%	920
	float 32	78.68 MB	84.65%	700
	float 16	39.67 MB	83.87%	524
	int 8	10.63 MB	81.22%	450

Restrictions

The research conducted thus far in the aforementioned areas has certain limitations.

1. Diversity and size of datasets: Small datasets, such as the one used in the study, carry a high chance of the model picking up odd patterns and occasionally failing to forecast new data correctly. In this investigation, samples were obtained with a desirable degree of variety in terms of rotation, varied light conditions, and even with different collecting equipment in order to somewhat offset this risk. Nonetheless, a larger database would undoubtedly improve generalization and give the model greater robustness.
2. Model complexity versus dataset size: While basic models may suffer from underfitting and ineffective generalization, complex models with limited datasets frequently result in overfitting. In contrast to the provided dataset, the SLDDCNNmodel in this study is rather complex. To get the required degree of performance, SLDDCNN's hierarchy level is limited to just four levels. However, when adding more intricacy, care must be taken.
3. Evaluation of a single dataset: To define the baseline, a single self-created database is used to test the suggested model and other well-known architectures. The results' repeatability may occasionally be impacted.

Conclusion

In this paper, we proposed SLDDCNN approach for real-time field photo classification. The problem of over-fitting can also be addressed with a data augmentation approach. By increasing the quantity of samples in the database, this strategy has improved the network's ability to generalize. The attention modules' ability to detect saliency based on features taken from many levels significantly improved the classification performance during the training phase. Experimental results showed that the proposed SLDDCNN achieved 86.53% classification accuracy, outperforming state-of-the-art VGGNet, ResNet, XceptionNet, and EfficientNet models. The network's performance can still be further enhanced,

 DOI: https://doi.org/10.51220/jmr.v21-S2.6

though, and it is expected that future research will examine architectural developments, adjust the hyperparameters, and increase the amount of the dataset. All things considered, our findings suggest that the SLDDCNN is a strong candidate for accurate and efficient real-time sugarcane leaf disease classification.

References

Amos PK Tai, Maria Val Martin, Colette L. Heald (2014). Threat to future global food security from climate change and ozone air pollution, *Nat. Clim. Change* 4 (9) 817–821.

Clive Bock, Gavin Poole, P.E. Parker, Timothy Gottwald (2010). Plant disease severity estimated visually, by digital photography and image analysis, and by hyperspec- tral imaging, *Crit. Rev. Plant Sci.* 29

François Chollet (2016). Xception: deep learning with depthwise separable convolutions, CoRR, arXiv:1610.02357 [abs].

Javaid Ahmad Wani, Sparsh Sharma, Malik Muzamil, Suhaib Ahmed, Surbhi Sharma, Saurabh Singh (2022).Machine learning and deep learning based computational techniques in automatic agricultural diseases detection: methodologies, applications, and challenges, *Arch. Comput. Methods Eng.* 29 (1): 641–677.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun (2015). Deep residual learning for image recognition, CoRR, arXiv:1512.03385 [abs].

Karen Simonyan, Andrew Zisserman (2014). Very deep convolutional networks for large-scale image recognition, CoRR, arXiv:1409.1556 [abs],

Madhulika Mittal, & Mittal, M. (2022). An Electronic Health Record Management System Based on Blockchain Technology. In 2022 *International Conference on Fourth Industrial Revolution Based Technology and Practices (ICFIRTP)* (pp. 285-290). IEEE.

Miguel Pagola, Rubén Ortiz, Ignacio Irigoyen, Humberto Bustince, Edurne Barrenechea, Pedro Aparicio-Tejo, Carmen Lamsfus, Berta Lasa (2009). New method to assess barley nitrogen nutrition status based on image colour analysis: comparison with spad-502, *Comput. Electron. Agric.* 65 (2): 213–218.

Mingxing Tan, Quoc V. Le (2019). Efficientnet: rethinking model scaling for convolutional neural networks, CoRR, arXiv:1905.11946 [abs].

Swapnil Daphal, Sanjay Koli (2022).Sugarcane leaf disease dataset, https://data.mendeley.com/datasets/9424skmnrk/1,

TensorFlow Lite, https://www.tensorflow.org/lite/models/convert/, 2024.

Vijai Singh, Namita Sharma, Shikha Singh (2020). A review of imaging techniques for plant disease detection, *Artif. Intell. Agric.* 4 : 229–242.

Xiao Xiang Zhu, Devis Tuia, Lichao Mou, Gui-Song Xia, Liangpei Zhang, Feng Xu, Friedrich Fraundorfer (2017). Deep learning in remote sensing: a comprehensive review and list of resources, *IEEE Geosci. Remote Sens. Mag.* 5 (4) ; 8–36.

Xin Yang and Tingwei Guo (2017). Machine learning in plant disease research, *Eur. J. BioMed. Res.* 3: 6.

Yansheng Li, Yongjun Zhang, Zhihui Zhu (2021). Error-tolerant deep learning for remote sensing image scene classification, *IEEE Trans. Cybern.* 51 (4) 1756–1768.

