RESEARCH



Robust CRW crops leaf disease detection and classification in agriculture using hybrid deep learning models

Check

B. V. Baiju¹, Nancy Kirupanithi², Saravanan Srinivasan³, Anjali Kapoor⁴, Sandeep Kumar Mathivanan⁴ and Mohd Asif Shah^{5,6,7,8*}

Abstract

The problem of plant diseases is huge as it affects the crop quality and leads to reduced crop production. Crop-Convolutional neural network (CNN) depiction is that several scholars have used the approaches of machine learning (ML) and deep learning (DL) techniques and have configured their models to specific crops to diagnose plant diseases. In this logic, it is unjustifiable to apply crop-specific models as farmers are resource-poor and possess a low digital literacy level. This study presents a Slender-CNN model of plant disease detection in corn (C), rice (R) and wheat (W) crops. The designed architecture incorporates parallel convolution layers of different dimensions in order to localize the lesions with multiple scales accurately. The experimentation results show that the designed network achieves the accuracy of 88.54% as well as overcomes several benchmark CNN models: VGG19, EfficientNetb6, ResNeXt, DenseNet201, AlexNet, YOLOv5 and MobileNetV3. In addition, the validated model demonstrates its effectiveness as a multi-purpose device by correctly categorizing the healthy and the infected class of individual types of crops, providing 99.81%, 87.11%, and 98.45% accuracy for CRW crops, respectively. Furthermore, considering the best performance values achieved and compactness of the proposed model, it can be employed for on-farm agricultural diseased crops identification finding applications even in resource-limited settings.

Keywords Plant disease, DL, Convolutional neural network, Slender-CNN, Modern agriculture

*Correspondence:

Mohd Asif Shah

m.asif@kardan.edu.af

- ¹ School of Computer Science and Engineering, Vellore Institute
- of Technology, Vellore, Tamil Nadu, India ² Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu,
- India
- ³ Department of Computer Science and Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai 600062, India
- ⁴ School of Computer Science and Engineering, Galgotias University, Greater Noida 203201, India
- ⁵ Department of Economics, Kardan University, Parwane Du, 1001 Kabul, Afghanistan
- ⁶ Centre for Research Impact & Outcome, Chitkara University
- Institute of Engineering and Technology, Chitkara University, Rajpura, 140401 Panjab, India

⁷ Chitkara Centre for Research and Development, Chitkara University, Baddi, Himachal Pradesh 140401, India



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

⁸ Division of Research and Development, Lovely Professional University, Phagwara, Panjab 140401, India

Introduction

Agriculture has always been a key constituent of the human civilization since its inception and it has not only sustenance food but also supplies vital raw materials that are just a small percentage of the many products that are used in garments, wood's, shelter, medicine and many more. With the advancement of civilization, agriculture began to progress and its consequences are realized well beyond simply food production [1]. Its importance can nowadays be judged from the fact that it has become part of the sources of world economy contributing to the economic development and providing employment opportunities. For instance, in India, the contribution made by the agriculture sector towards the economics of the nation was very large constituting 19.9% of Gross Domestic Product (GDP) in the financial year of 2020-2021 [2]. Besides, a large part of the population was involved in the sector with about 41.49% of the total labour force employed in different facets of agriculture. These statistics are salient to the issue of the economy since they reflect a sector of the economy that assures the survival of millions of people [3]. But the global population growth rate that is estimated to rise from the 7.7 billion people in 2019 to 9.7 billion in 2050 is on the other hand putting up a challenge to agricultural expansion in order to locate food plus raw materials to utilize in production of such expected commodities [4]. It is a common understanding that output optimization and wasting by plant diseases and other factors must be balanced. There will be a need for efficient agricultural enhancement and better management of available resources and technologies with regard to the fact that productivity requirements of the society will increase but bounds to the ecological damage will still be maintained [5]. Plant diseases are one of the primary factors that limit agricultural productivity all over the world and usually result in reduction in quantity and quality of food crops. These diseases together with pests cause 30-40% loss of the entire crop production over the years [6]. These losses are not beneficial to the farmers and the regions, but to food provision around the world especially when the population is growing. To enable the human race to maintain and keep feeding billions, it is very important to detect crop infection and disease timely and accurately [7]. In the past, plant diseases have been detected and diagnosed through the visual assessment of trained plant pathologists. Most farmers or professionals would examine plants for disease symptoms, such as spots on the leaves, death inference, or leaf positioning changes [8]. This direct method while useful in the small scale, has various drawbacks when applied to expansive farming. First, manual intervention is sore be particularly tire some in terms of time and human power when labour farms are gigantic as there's no room for errors as far as trying to stop the spread of a disease is concerned [9].

Unfortunately, by the time an actual inspection reveals the disease, the infection could have already affected a large portion of the crop rendering control and mitigation more expensive and problematic. Second, a physical examination of crops leaves a lot to be desired [10]. Even highly trained plant doctors succumb in the place of correct diagnosis due to the diseases being complex or extreme deficiency symptoms or other stress conditions that may appear similar, with their signs being mild [11]. When mistakes of this nature occur, farmers tend to apply the wrong remedy making things worse to the crop and hence incurring additional expenses to them. Recent advancements and innovations in areas like ML, DL as well as the Internet of Things (IoT) have emerged as efficient alternatives to existing approaches within Agriculture. These technologies can create models that are as good as, if not better than human beings, in the rapid and precise detection of plant infections [12]. Agriculture is necessary for economic stabilization, reducing poverty, solving the problem of food shortages, and reducing unemployment. However, pests and diseases hurt crop productivity and compromise the economy significantly. Therefore, monitoring leads to early identification which is crucial in ensuring optimal crop management and sustainability of resources [13]. The peppers were separated into two sections where one section was infected with the disease and the other was the control. Using multispectral imaging spectral and textural features were extracted and subsequently analyzed with GLCM and LBP techniques. Some feature selection strategies were undertaken and classification models built. More accurate forms of disease detection were accomplished by the application of two-dimensional CNN [14]. The network has convolutional rebalancing, an image augmentation and a feature fusion as its modules. The convolutional rebalancing module encodes the images, with reversed sampling helping enhance feature extraction even for categories with fewer images. In contrast, the image augmentation module is employed to increase the amount of training data, and the feature fusion module is used to merge several features derived from the images. According to the experiments, the network is more effective than existing approaches and thus is helpful for intelligent crop pest and disease management [15]. Imaging methods are analyzed either with respect to the field or

the glasshouse plants, and the techniques are categorized into 'healthy and diseased plants classification' focusing on improving the classification accuracy, stress and disease detection in terms of time and severity. One of the main points of interest in the review is addressing the subject of hyperspectral imaging including its usage to reveal extra information concerning the status of plants and predict the appearance of diseases [16]. Identifying both the disease and the type of disease affecting different crops especially tomato and grape is one of the areas of great emphasis. The primary goal is to predict the type of disease that would occur to grape and tomato leaves in the future and even at an early stage. Multi-Crops Leaf Disease (MCLD) was detected using CNN methods. The features extraction of images collaborated with DL-based model which was able to discriminate between sick and healthy leaves. To obtain good performance measures, an appropriate CNN based model was the Visual Geometry Group (VGG) [17].

ML, as the name suggests, learns complex patterns embedded in the data and discovers disease patterns, DL and neural networks work Magic in identifying diseases through images of the plants. IoT connects devices for sensors and data collection such that realtime surveillance of plants is possible, hence responses can be timely. All in all, these technologies minimize human errors and time wastage while improving crop health management, hence disrupting the agricultural sector. The present study explores the common diseases in CRW crops, which together account for about two-thirds of the calories consumed worldwide. These crops were harvested in 2019 in 575.2 million hectares, which is a considerable area out of the available 1204.4 million hectares of arable land. Considering the magnitude of cultivation, DL and ML techniques have to be employed in disease detection and control measures to reduce the loss of crops. For this objective, many researchers have created disease detection models specialized for the CRW crops.

The number of agriculture sector workers who are digitally literate in India is only 13% and therefore using models for certain crops is not useful. On the other hand, the integration process of specific models into one application would result in creation of a complex model which requires advanced computing powers. Thus, universal models should be adopted in diagnosing several crops as well as in areas with scarce resources. The study focuses on the compact neural network that aims to screen for certain diseases in maize, rice and wheat. These diseases include, but are not limited to, Corn Grey Spot (CGS), Corn Northern Blight (CNB), Corn Rust (CR), Rice Brown Spot (RBS), Rice Hispa (RH), Rice Leaf Blast (RLB), Wheat Leaf Rust (WLR), Wheat Stem Rust (WSR), and Wheat Stripe Rust (WSR).

- In this study, it is hoped to present a Slender-CNN model which is capable of classifying diseases of Corn, Rice, and Wheat crops. This model uses filters with various dimensions at one level, which helps reveal important image features even when the scale of the object differs from image to image. Besides the proposed model, the ability to diagnose diseases for the CRW crops is tested for seven other popular CNN models. This comparative strategy assesses the efficiency of different models in the diagnosis of plant diseases relative to these major crops.
- The proposed model reaches an accuracy of 88.54% with only 387,340 parameters, which makes it a possible case of low resources. On top of that, it is also seen that the rest of the concerning benchmark models which were evaluated do not only fall behind the proposed model regarding their accuracies but also the number of parameters. MobileNetV3 achieves an accuracy level of 77.11% but with 8 times the parameters used in the present model and therefore falling short by 11.43% of the present study.
- Comprehensive experiments proved that the Slender-CNN models proposed are able to detect cropspecific diseases in Corn, Rice, and Wheat. In the case of Wheat planted crop where only healthy and infected crops classified 99.96% accuracy of classification between the two, it was also 87.11% for Rice and 98.45% for Corn. All these results were achieved without changing the model architecture proving its flexibility and efficiency with respect to other crops too.

The structure of this study is kept as follows: "Related work" covers related work, which discusses different existing models related to classifying various leaf diseases. "Material and methods" focuses on materials and methods, detailing the dataset employed for testing the efficiency of the presented model in classification. "Experimental results and discussion" presents experimental results and discussion, highlighting the advantages of the proposed model, providing a summary, and addressing potential challenges. "Discussion" provides the detailed comparison and summary of accuracy comparison of proposed and other existing models. Finally, "Conclusion and future work" concludes the current work and describes its possible continuations to advance promising achievements in the study of different leaf diseases.

Related work

Wasswa Shafik et al. [18] suggested two modes for plant disease identification, AE and LVE, have been proposed and connected with nine distinct DL analyzer for enhanced recognition and randomization. A Series of tests were conducted on 15 categories of the PlantVillage dataset which contains the 54,305 images of a variety of plant diseases. Tuning of hyperparameters focused on available popular pretrained models. Validation of the developed model involved independent testing and testing as part of an ensemble. A logistic regression classifier was used to benchmark various combinations of proposed deep models. Results from the experiments showed that PDDNet-AE and PDDNet-LVE gained higher performance compare with other current CNN models with the performance of 96.74% and 97.79%, respectively indicating high levels of stability and generalization ability.

Arunangshu Pal et al. [19] the AgriDet framework utilized traditional INC-VGGN and Kohonen-based DL networks to identify plant diseases and categorize their severity levels. Image pre-processing was employed in the framework to eliminate limitations, along with a multi-variate grab cut algorithm for segmentation efficiency. The framework utilized an enhanced base network, specifically a pre-trained INC-VGGN model, for precise disease detection and classification. Pre-trained weights and features were moved to the newly developed neural network for the task of plant disease detection. In order to address overfitting, a dropout layer was added, and Kohonen learning was employed for DL of features. After percentage computation, the enhanced base network categorized severity classes in training sets. The framework attained a higher accuracy rate of 98.96% in performance evaluation.

Kumar Satyam Tanti et al. [20] in order to provide training and testing for plant leaf identification, two of the pre-trained models—DenseNet and Efficient-Net—were used in the study. It was found out that the DenseNet model obtained accuracies of 99.17% and 95.26%, while the Efficient-Net model obtained 99.22% and 94.54% accuracy rate. As such, it suggested that the DL models such as CNN utilized in this paper to examine plant leaves have potential enhancement of accuracy which would enhance efficiency by providing faster detection, customized treatment, and better prediction of illnesses.

Andrew et al. [21] this study deployed models such as DenseNet-121, ResNet-50, VGG-16 and Inception V4 for their purpose that required quick and accurate detection of plant diseases based on CNNs. As stated by the authors, the PlantVillage dataset that contains 54,305 image samples of different types of plant diseases was used for the experiments. The results showed that DenseNet-121 classification accuracy was 99.15%, which significantly exceeds the precision provided by the most modern classifiers. Also, to enhance the findings, the study included comparison of the results with some other related researches.

Khalil Khan et al. [22] introduced a model for segmenting plant leaves semantically to identify diseases, employing a deep CNN focused on semantic segmentation. It distinguished between diseased and healthy regions on a particular plant leaf, enabling the classification of ten distinct diseases. Furthermore, it assigned semantic labels to individual pixels, estimating the extent of disease impact. Evaluation of the model was conducted using the Plant-Village database and a set of 20,000 images. Attained an average accuracy of 97.6%, the model showcased a notable performance enhancement compared to prior outcomes.

Mingyue Guo et al. [23] plant diseases can result in great losses as well as extinction of some species. Techniques for automated detection can help raise food output and reduce losses as well. Object recognition systems have been advanced by the introduction of deep neural networks. In this paper, the author proposes an efficient approach to plant disease recognition based on DL without restricting attention to a particular plant. The performance of different Visual Transformer architectures was evaluated on the AI Challenger 2018 Crop Disease Detection and PlantVillage datasets. Experimental results showed that the accuracy achieved is 98.96%, which is commendable.

Aayush Deshmukh et al. [24] proposed model solves the problems of overfitting as well as employs data enhancement. The structure employs a CNN with six layers in addition to an artificial neural network with dense interconnectivity and superimposed on pre-processed data. The model employs kernels and activation functions in such a way as to maximize classification power. Similarly, the model terminates the classification of the plant's leaves mere as diseased or non-diseased but also classifies the diseased leaves in quick succession and with an optimal model configuration. The results obtained from the experiments are accurate and effective with quantitative results of 95.23%, 94.33%, and 94.52% on the PlantVillage data set.

Aadarsh Kumar Singh et al. [25] developed a model for the augmentation of healthy leaves by designing a DL based Generative Adversarial Network called as LeafyGAN which is capable of simulating disease patterns otherwise known as sympathetically viewed. Such a popular plant disease dataset has been augmented by the model by providing additional synthetic samples for such disease categories where ground truth data is scanty. A standard GAN and rotation loss were used to segment the leaf image by separating its foreground from the background using exogenous loss. Following training on this improved dataset, the effective mobile ViT model achieved outstanding accuracy rates 99.92% in PlantVillage dataset and 75.72% in PlantDoc dataset.

Sheng Yu et al. [26] presented a novel transformer block model that uses transformer architecture for longrange features and soft split token embedding for local information. It also incorporates inception architecture and cross-channel feature learning for fine-grained learning. The model outperforms previous convolution and vision transformer-based models, achieving 99.64% accuracy on VillagePlant, 99.22% accuracy on iBean, 86.89% accuracy on AI2018, and 77.54% accuracy on PlantDoc.

Yuzhi Wang et al. [27] proposed that addressed the challenges posed by large amounts of labelled data by combining a convolutional block attention module (CBAM) with a masked autoencoder (MAE). The model's performance was validated using both the collected dataset and the CCMT dataset. Accuracy of 95.35% and 99.61%, recall rates of 96.2% and 98.51%, and F1 scores of 95.52% and 98.62% were achieved by the enhanced model on the CCMT dataset and the collected dataset,

respectively. Notable works are summarized in Table 1 depicting the different models employed in the different tasks of plant disease detection and their performance.

A number of models have been created and reported to be specifically targeted at disease detection in the CRW; however, no single model is in operational status currently. To achieve the cost effectiveness of avoiding the use of several models to address this problem, a Slender-CNN model is brought to this study, which performs disease forecasting in CRW at performance metrics aspired to by the existing models. The proposed Slender-CNN model tackles some of the shortcomings identified in the existing models for plant disease detection shown in Table 1. To deal with overfitting, the model adopts several techniques which include data augmentation, dropout regularization and weight regularization, making it well suited for deployment on unseen data. In contrast to other models, Slender-CNN is tested and compared with leading models, showing its best performance efficiency with fewer parameters. The model improves its robustness using early stopping, advanced regularization and other techniques to reduce the distance between training and testing accuracy. Furthermore, unlike some models that focus on only semantic segmentation, Slender-CNN integrates the entire disease detection procedure, which is an added advantage. The generalization of the model is not locked to only a particular dataset, for example, PlantVillage, but it transfers learning from other diverse datasets to learn under diverse environments. It is also more efficient in terms of computational resources than transformer models while preserving the accuracy of predictions. Hence, Slender-CNN can be employed in devices with low resources making it viable for real time plant disease diagnosis. The model can also generalize on other datasets, including PlantDoc too, which makes it more robust and places it in a position to be an efficient

Author	Year	Model	Limitation
[18]	2024	PDDNet-AE, PDDNet-LVE	No data of handling overfitting
[19]	2023	AgriDet framework with INC-VGGN	Lack of comparison with newer models
[20]	2023	DenseNet and EfficientNet	Minor drop in testing accuracy compared to training accuracy
[21]	2022	DenseNet-121, ResNet-50, VGG-16, Inception V4	No data of solving overfitting or generalization
[22]	2022	Deep CNN	Focused only on semantic segmentation, not end-to-end detection
[23]	2024	Visual Transformer architectures	Limited generalization outside specific datasets
[24]	2024	CNN with six layers and Dense ANN	No information of comparison with modern transformer models
[25]	2024	LeafyGAN	Lower performance on PlantDoc dataset compared to PlantVillage
[26]	2024	Transformer block model with inception architecture	Requires high computational resources due to transformer architecture
[27]	2024	CBAM + MAE	No indication of real-time efficiency or generalizability to other datasets

Table 1 Summary of state-of-the-art (SOTA) models using PlantVillage dataset



Fig. 1 Sample CRW images from PlantVillage dataset

and scalable option for plant disease management in realistic agricultural conditions.

Material and methods

Material

The dataset employed data on plant diseases specific to crops belonging to the CRW class. The images include those associated with maize, rice, and wheat diseases. These images are extracted from official documents that are open to the public and cover a large spectrum of examples for every type of the crop. To maize

Category	Images (%)	Category	Images (%)
W_stripe_rust	9	Healthy_W	7
R_brown_spot	6	Healthy_R	12
R_Hispa	6	Healthy_C	10
R_Leaf_Blast	8	C_Rust	10
W_leaf_rust	8	C_Northern_Blight	8
W_stem_rust	12	C_grey_spot	4



Fig. 2 Proposed Slender-CNN model architecture

disease images, a number of them were obtained from the famous PlantVillage dataset which is a common international resource [28]. The Plant Village Dataset is often regarded as the place where almost all crops can be represented with sufficiently broad collection of labelled image sets to facilitate disease detection and classification. As for images of rice disease, a range was also searched on the Kaggle site [29]. On Kaggle, a large number of datasets have been collected by numerous microbiologists and other medical professionals competent in their subjects, containing well-annotated data supporting different plant disease conditions. Rice diseases covered in this dataset includes the bacterial blight, blast and brown spot diseases among others. Equally, images of wheat disease were retrieved from two different Kaggle datasets [30]. These databases carry informative images on the various diseases of wheat especially filamentous ones like rust, septoria-sheath blight, etc. The inclusion of two separate sources in case of wheat disease ensures that the dataset consists of wheat disease cases across the board improving the quality of the data. The diversity of plant diseases included is helpful in Fig. 1 which shows sample images corresponding to different dataset classification [31]. Each of them refers to a predetermined type of disease for maize, rice or wheat enabling detailed analysis comparison and thoroughness of the analysis. This constructed data set is guite useful for the purposes of creation and testing of ML models enabling diagnosis of plant diseases in actual and time critical agricultural environments. The data was split into three parts with 70% being training, 15% testing and 15% being validation. The largest subdivision was for training the ML models constituting seventy percent of the total data that is towards the process. Out of this fifteen percent was used to carryout testing and the rest fifteen percent was to be needed towards the validation of the work done. The dataset, as seen in Table 2, contained a number of agricultural diseases within it however some diseases were in a larger proportion than others. For example, among the major diseases corn grey spot disease was only about 4% of the cases remaining while healthy rice and most wheat stem rust was each about 12% of the data. Due to this kind of data set distribution, it might affect the performance of models and therefore imposed the use of some special approaches such as normalization or weighting to capture the detection of minority class centers.

Slender-CNN Method

The proposed model illustrated in Fig. 2 is designed to identify diseases in CRW. At the very same level, the salient features of an image are captured using several filters with different dimensions. These filters help the model to capture targets with different sizes present in different images. Owing to this, there are three building blocks that are embedded in the model shown in Fig. 2, each block has almost the same architecture but different in terms of the filter sizes for the convolutions at that particular level. The structure depicted in Fig. 2 shows that each cell represents a single layer of the neural network, each cell contains an input field with the corresponding input size and an output field which shows the size after the operation is performed. Directional arrows are

used to depict how data moves from on cell to another within cells located entirely in the different representative architecture.

Algorithm 1 Slender-CNN model for disease classification

Start
{
#Input layer
Input: X_input ⊗ R^(224×224×3)
Outcome: RGB image input with height, width, and depth.
#2D Conv_layer
Conv1: Apply convolution operation X_conv1 = f(W_conv1 * X_input + b_conv1)
Outcome: X_conv1 ⊗ R^(224×224×64).
Conv2: Next 2D convolution which is used to reduce the size:
$X_conv2 = f(W_conv2 * X_conv1 + b_conv2)$
Outcome: X_conv2 ⊗ R^(222×222×64).
#Separable layer
Apply depthwise separable convolutions to reduce computational cost:
X_sep = f(W_depthwise * X_conv2) + f(W_pointwise * X_depthwise)
Outcome: X_sep
#GlobalMaxPooling layer
Apply Global Max Pooling to reduce dimensionality:
$X \text{ pool} = \max(X \text{ sep})$
Outcome: X pool \otimes R^(1×1×64).
#Dense layer
Pass through fully connected layers:
X dense = $(1/N)^*$ f(W dense * X pool + b dense)
Outcome: X dense $@$ R^(64).
#Average pooling
Pool features from dense layers:
$X_avg = (1/N) * \Sigma X_dense$
Outcome: Reduced features.
#Fully Connected layer
Apply final dense layers for classification:
X_output = f(W_out * X_avg + b_out)
Outcome: X_output @ R^(12) (which means 12 classifications).
#Average layer
It combines the outputs of the three paths by computing the average value for each
corresponding feature across the paths. For outputs O1, O2, O3, the output is:
$O_avg(i,j) = (1/3) * (O1(i,j) + O2(i,j) + O3(i,j)).$
#Softmax layer
It is used to convert the raw scores from the final dense layer into probabilities for
each of the 12 classes. The softmax function for class i is defined as:
$P(y_i) = e^(z_i) / \Sigma_{(j=1)} N e^(z_j).$
}
Stop



Fig. 3 Convolution function

The proposed model has three dimensions, each of which works in performing convolutional operations to obtain different characteristics from the input image. All three blocks have a common start where they have an input layer, which receives an image of size 224×224 with 3 channels.

• Upper block: The first block within the first layer of the network consists of a 2D convolutional layer on the input image. During these layers, the image is filtered and feature maps with deeper hierarchically summarized content of the image are formed. At this point, the spatial sizes of the image are reduced from 224×224 to somewhat smaller sizes (for example, 222×222), but the number of filters rises from 3 to 64. This allows the network to engage complex patterns and edges as the network's depth increases. After this sequence of convolutions, the block performs separable 2D convolutions which is a type of convolution that is energy efficient and is done in two steps rather than one. This essentially lowers the work done without loss of essential spatial information. Now, a global max pooling layer is involved for this task and it further reduces the volume of the generated feature maps by retaining the maximum of every feature map bringing about a great reduction in the size of the features with certain key characteristics sustained. Clinically, the pool layer is followed by a dense layer that has 64 neurons which serves to combine the performed processing on different outputs and resynthesize them with respect to the ultimate decision making.

• Middle block: The second block is similar to the first block in structure. This time, starting from the same input image, the block is subjected to the same operations: 2D convolutions, separable convolutions and a final global average pooling operation. As with the first block, the convolutions in the second block outperform the first ones by increasing the number of filters and sifting spatial dimensions of the intermediate feature maps, enabling important hierarchical features to be learned by the model. As before, following global max pooling, the block adds 64 salient features as an output through a dense layer. This parallel path



Fig. 4 Operation of separable convolution

assists in ensuring that the network obtains different views of aspects of spatial features of the image.

- Lower block: The last block is a repeat of the third and the second blocks. This takes the same input image and processes it through the same series of convolutional layers, separable convolutions, global averaging pooling and a final layer of 64 neurons. In this way, feature extraction is made uniform across the three processes. Each block is, however, built in such a way that, they are all able to unlike features of the input image despite the fact that they hold a similar architecture. The network has three pathways through the images. This means that each pathway can pay attention to various parts of the image to improve the feature extraction and consequently increase the recognition accuracy.
- Output block: In the last block all three outputs after all the three parallel blocks have processed the input image are fused. At this stage, an average layer combines the feature outputs from the three paths. Instead of performing a feature selection from any one particular path, the average layer calculates and outputs the average of course of three output vectors creating a better feature representation. This stage guarantees that the network adopts the different views available from each path that aids better comprehension of the input image.

After the average layer, the network proceeds to apply these defined features to two more dense layers, each containing 64 neurons. This is done with the intention of complementing the knowledge that has been acquired before in the network so that better results can be achieved. The purpose of the dense layers is to serve as classifiers that decode the features and make it appropriate for the last classification. Then the architecture is completed with the output layer that assigns classification for 12 classes, which could be different disease types or other categories like fungi. This is an important step in the network since that is where the network gets to predict based on the features it has learned during the course of the model. The architecture adopts a three-parallel-block-followed-by-asingle-merge strategy in the design, which guarantees diversity and complementation in the features captured from the image, thereby improving the capability of classification of complex image patterns. This, with separable convolutions and global pooling in their design controls computation requirements while conducting effective feature capturing making it applicable in plant disease recognition processes where high accuracy and efficiency are needed.



Fig. 5 GlobalMaxPooling operation

Convolution layer

The filter functioning in one of the convolution layers of the proposed model has been depicted in Fig. 3. The very first component of the model utilizes for its convolution's filters of size 3×3 , which follows from Eq. (1). It follows here that the 3×3 image patch produces a feature map (the output) using matrix's dot product with the filters saved therein. Just the same, 5×5 and 7×7 filter sizes have been employed by Middle Block and Lower Block convolution layers, respectively in order to obtain feature map as an output [32]. This is an operation performed on every convolutional layer where a 2D image is subjected to a moving window called a filter (kernel) which is placed on the input image. It formulates mathematically how the mathematical convolution is carried out on the output feature map with reference to the position (i, j) as follows,

$$S(i,j) = \sum_{m=1}^{M} \sum_{n=1}^{N} I(i+m,j+n) \cdot K(m,n).$$
(1)

Here, I(i + m, j + n) is the input pixel at position (i + m, j + n), and K(m, n) is the kernel value at position (m, n). S(i, j) is the output at position (i, j), and $M \times N$ is the size of the kernel.

The Fig. 3 depicts the procedure of a convolutional operation in a neural network, and a binary matrix measuring 5×5 is applied. A patch measuring 3×3 is picked from this grid and 3×3 mask or filter where its values range from 0 to 3 is placed along with the patch and is superimposed with this area patch. It gives rise to one value which is used in the output grid in the corresponding cell. Here, it is noted that the first cell of the 4×4 output grid has a value of six as a result of the summation of the patch filter region values equally summed to the cell values. Such procedures are carried out moving the filter across the grid and over the patches alternatively until the entire output grid is obtained.

Separable convolution

The best way to conceptualize Fig. 4 is also the most computationally efficient solution owing to the basic structure of a standard convolution layer within any computer vision application [33]. This layer has two parts rather than one as in previous example. In the first part, called depthwise convolution, the single transmitted image is divided into its basic components and retains every color, e.g., red, green and blue. Each of these channels is convolved independently with a 3×3 filter. Depthwise convolutional layers however, treat each channel independently and capture the spatial information within each channel individually. In this way, the number of computations is greatly minimized by not integrating the channels at this stage thus making it a better option of performing feature extraction on the input image. After conducting the depthwise convolution on each channel, all the feature maps produced are combined in a vertical manner. In the second part, this layered combined feature vector is subjected to a pointwise convolution. Pointwise convolution performs a linear mixture of the channels which is done using a 1×1 filter where the output of the depthwise convolution is combined to produce or mix information from the different channels. This step is very important for assimilating all the spatial information from all channels and produces the final set of feature maps of the processed image. In the model proposed in this paper, having the depthwise separable convolution layers fed by the previous convolution layers guarantees that the channels in the depthwise separable convolution core are equal to the number of filters in the previous layer. This is true for this case whereby the number of channels as well as the number of filters is 64. It ensures that the feature maps at a given layer are effectively utilized before a saturation occurs. Additionally, this phenomenon is also performed in the remaining two building blocks of the model but only differs in the filter sizes used in each block. Such multi-stage treatment not only saves the pipes as well as shoots the calculative burden to a great level, on account of the fewer parameters and operations, but also fully holds the spatial and channel information required for the correct representation of an image in detail.

In separable convolution, the convolution operation is split into two parts: a depthwise convolution and a pointwise convolution. A separate filter is applied to each input channel. The depthwise convolution equation for a channel c is,

$$S_{c}(i,j) = \sum_{m=1}^{M} \sum_{n=1}^{N} I_{c}(i+m,j+n) \cdot K_{c}(m,n).$$
(2)

Here, $S_c(i, j)$ is the output for channel *c* at position (i, j). The $I_c(i + m, j + n)$ is the input value for channel *c*. Then, $K_c(m, n)$ is the depthwise kernel for channel *c*.

$$S'(i,j) = \sum_{c=1}^{C} S_c(i,j) \cdot W_c.$$
(3)

Here, $S_c(i,j)$ is the depthwise output for channel *c*. Then, W_c is the weight for the 1×1 pointwise filter. The separable convolution reduces the number of parameters and computations compared to a standard 2D convolution.

Global max pooling

In the proposed model, it is very important that the GlobalMaxPool2D layer is used with the multi-dimensional feature maps that come from the convolution layers in the model [34]. This is necessary because these feature vectors will be useful for the following simple layers that perform either a classification or regression task. Here, the kernel contains the transition between the spatial dimensions that are found inside the convolutional feature maps and the fully connected layers that take in flat, one-dimensional format. GlobalMax-Pool2D is a method that works by taking the maximum value from a single feature map across the various channels and spanning the spatial width and height. In simpler terms, the pooling operation for each channel of the feature map reduces the area matrix that represents the width and height of the feature map into a single number. This is aimed at achieving the reduction of complexity and dimensionality of the data without the underlying important information because it enables reduction of 2d spatial representation to one number per channel only. As designed in the proposed model, every single building block of the network has a GlobalMaxPool2D counterpart. Figure 5 depicts the operation of GlobalMaxPooling layer. There are several such pooling layers attached to the feature map with an equal number of channels to the one in the preceding layers so as to maintain the depth of the feature maps [35]. Nevertheless, the dimensions of the feature maps are different in other building blocks as well due to the differences in filter sizes in the convolutional layers. It is possible that larger filters cut down the spatial dimensions to a larger extent, while more finesse ones capture greater details and thus there are feature maps of various sizes that enter the GlobalMaxPool2D layers. The image demonstrates how GlobalMaxPool2D works. Here's how the calculation is performed,

Table 3 Training specifications			
Aspect	Details		
Training method	Supervised		
Loss function	Categorical cross-entropy		
Purpose of loss	Computes the difference between two probability distributions		
Optimizer	Adam		
Learning rate	0.001		

Table 4 Accurac	y comparison	of normal	l and infected	crops
-----------------	--------------	-----------	----------------	-------

Crop	Acc (%)
Corn (C)	99.81
Rice (R)	87.11
Wheat (W)	98.45

Algorithm 2 GlobalMaxPool2D operation for classification

- **Input**: Consists of multiple feature maps with dimensions defined by height, width, and depth. In the image, there are three 3×3 feature maps, corresponding to the depth dimension of 3.
- **GlobalMaxPool2D**: For each feature map, GlobalMaxPool2D finds the maximum value from the entire matrix. It does this by looking at all values in the 3×3 grid and selecting the largest number.

Start

```
{
    For the first feature map (left matrix)
    {
        Values = 6,3,4,3,5,3,4,9,2
        Maximum = 9
    }
    For the second feature map (middle matrix)
    {
        Values = 8,3,7,5,9,1,1,8,4
        Maximum = 8
    }
    For the second feature map (middle matrix)
    {
        Values = 7,3,1,1,1,8,4,4
        Maximum = 7
    }
    Output
}
Stop
Output: The result is a single value from each feature map, forming a 1D tensor where the number
```

of values corresponds to the number of channels. In this case, the output is: 9, 8, 7.

In spite of such differences in spatial dimensions of the feature maps, there is no much variation on the Global-MaxPool2D operation except that each feature map is reduced to one-dimensional form along the channel axis. This pooling operation is very essential due to its role in decreasing the computations as well as the intensities in conducting dense layers without losing significant features in the process.

$$S = \max_{i,j} F(i,j). \tag{4}$$

Here, F(i,j) is the value at position (i,j) in the feature map and *S* is the single maximum value. In the case of the GlobalMaxPool2D operation, it retrieves one maximum value from each of the channels only in order to allow the model to capture the most powerful features while downplaying less useful ones, that may be useful in reducing overfitting and increasing generalization. So, it is the case when we have convolutional layers with different sized filters and GlobalMaxPool2D layer, the model can extract the proper features and reduce the data volume efficiently for future tasks.

Dense layer

Dense layers are an important element of neural networks, especially in the latter parts of the architectures where these layers take the reasonable features obtained from other layers and combine/process them [30]. Each dense layer receives an input and applies a linear transformation, which is a weighted sum of the inputs plus a bias. Formally speaking, this concept can be written in terms of equations as,

$$y = f(W \cdot x + b), \tag{5}$$

where W is the weight components, x is the vector of input, b is the bias, and f is an activation unit. The weight matrix W isn't just a static matrix. It has adjustable conductors that decide what percent of each input coming into the model is obstructing the final outcome. The bias b causes a shift in the output, increasing the adaptations that the model can make with the data. The activation function f will include but not be limited to ReLU, sigmoid, which is incorporated into the model in order to increase its ability to learn complex structures. In DL, dense outputs are also employed at the last sections of the networks where class probabilities are sought for instance or the most desired output is aimed for every input. In the process that sits between design and construction, these layers undergo weight modification, bias adjustment with the aim of the reduction of the loss and enhancement of the functionality of the model with respect to the responsibilities bestowed on it.

Experimental setup and evaluation parameters

The model is developed in Python—the Pandas library is used for data preparation, the Keras API is used for the construction of the CNN, NumPy for mathematical

Table 5 Performance metric of Corn

Class	Precision (%)	Recall (%)	F1 score (%)	Acc (%)
Corn grey spot	98.12	99.33	97.21	99.81
Corn northern blight	98.33	97.21	99.42	99.80
Corn rust	97.33	99.25	96.51	99.91
Healthy corn	99.22	98.36	98.55	99.92
Average				99.86

functions and TensorFlow for neural networks. The training is carried using the Adam optimizer, cross-entropy as loss function, and early stopping, and learning rate scheduling to avoid overfitting and enhance convergence efficiency. The integration of these procedures and tools is critical in building an effective and precise model for classification of the CRW disease.

$$Acc = \frac{T.positive + T.negative}{T.positive + T.negative + F.positive + F.negatice},$$
 (6)

$$Pr = \frac{T.positive}{T.positive + F.positive},$$
(7)

$$Recall = \frac{T.positive}{T.positive + F.negative},$$
(8)

$$Sp = \frac{T.negative}{T.negative + F.positive},$$
(9)

$$F \, score = \frac{2 * Pr * Recall}{Pr + Recall}.$$
(10)

The diagnosed pests and disease of 12 categories together with a healthy class of CRW crops framework objects aggravates this problem to be a Multi class classification. The above result could be derived from a confusion matrix, which is a better tool for performance evaluation that provides specific performance metrics such as True Positive (T.positive), False Positive (F.positive), True Negative (T.negative) and False Negative (F.negative). These parameters in multi label image

Table 6 Performance metrics of Rice

Class	Precision (%)	Recall (%)	F1 score (%)	Acc (%)
Healthy rice	85.37	86.88	87.12	86.08
Rice brown spot	88.44	86.75	85.98	87.32
Rice hispa	85.87	88.12	87.09	87.61
Rice leaf blast	87.52	87.88	88.02	87.50
Average				87.11

Table 7 Performance metric of Wheat

Class	Precision (%)	Recall (%)	F1 score (%)	Acc (%)
Healthy wheat	98.09	99.15	97.21	98.12
Wheat leaf rust	98.11	97.54	99.04	97.49
Wheat stem rust	97.23	99.33	96.22	97.98
Wheat stripe rust	99.02	98.45	98.76	98.49
Average				98.45







classification T.positive is the number of images that are relevant and have been categorized properly; whereas F.positive is the number of images that are irrelevant but have been placed in the relevant category by the investigators. T.negative refers to the contrary in that it comprises those images which have been correctly classified outside relevant category. F.negative however consists of images that do not fall in the relevant category but have been pushed into it. This structure is developed to augment understanding of how performant the model is in regard to all the classes. All of the models follow a supervised approach for training and use categorical cross-entropy as a loss function which measures the dissimilarity of two distributions along with an Adam optimizer with a learning rate of 0.001. In order to avoid overfitting and enhance generalization on the model, data augmentation techniques are applied during training. A validation set is used to track performance and allow for early stopping if overfitting occurs. After building and validating the model, it is checked for performance over a separate test dataset with new data. Table 3 illustrates the suggested model training specification.

Experimental results and discussion

In the proposed work for the disease classification of CRW, image augmentation is used in order to increase the diversity of the dataset and the robustness of the model. Rotation, flipping, zoom, and shift are some of the augmentation techniques employed to enhance the variability of the images thus the model is able to perform well on unseen data. In the case of CNNs, a constant input size is required, so every single image contained in the dataset is resized into a square of 224×224 pixels. Furthermore, the pixel values were normalized to the range [0, 1] to assist in the input process and improve model training. The dataset contains images of different resolutions, splits into training, validation, and test subsets for proper testing purposes. The evaluation of the model's performance is carried out based on several measures including accuracy (Acc), precision (Pr), recall, F1-score, specificity (Sp) and area under the curve (AUC-ROC) to avoid bias, especially in cases of imbalanced datasets.

Discussion

In order to adequately characterize the proposed Slender-CNN model in the context of the classification of diseases of CRW plants competitive studies have been performed. These tests are meant to determine how far the model can go in comparison with other baseline models and methodologies available in this research area. The evaluation concerns two dimensions: diagnosis of healthy and infected crops and diagnosis of one infected crop within the range of different diseases. In the first part of analysis, the performance of the model in identifying infected and healthy crops is evaluated giving an estimate of the reliability and the possible presence of disease in the crops that are tested. This analysis is important because, as noted, this test will assist in focusing on plants that are scrubbed for infection therefore ensuring prompt action is taken where necessary to manage crops. The second part of the evaluation consists of a detailed assessment of disease classification for the individual crops. In this case, the Slender-CNN model is also evaluated on the ability to differentiate between diseases and tested on the labelled data set consisting of different disease manifestations of the same crop. The results of these experimental findings will be presented in the next chapter with accompanying tables stating accuracy, precision, recall, and the F1 score alongside simple illustrative average confusion matrices. Through these deliberations, we intend to conclude as

Fig. 9 ROC curve of all model accuracy comparison

Fig. 10 Performance metric comparison of proposed and different models

Table 8 Parameter details of all CNN models

Models	Parameters
VGG19	143,667,240
EfficientNetb6	43,040,704
ResNeXt	25,000,000
DenseNet201	20,242,984
AlexNet	61,100,840
YOLOv5	7,000,000
MobileNetV3	5,400,960
Proposed model	387,340

clearly and thoroughly as possible, both the merit and demerit of the model taking the issue of disease classification in CRW plants a step further.

Table 9	Advantages of	proposed stud	y over SOTA models
---------	---------------	---------------	--------------------

sets

Feature/metric **Proposed slender-CNN** Traditional CNNs Hybrid models Parameter count 387,340 (extremely low) High (MobileNetV3:~3 M) Very high 98.45%/87.11%/99.96% Accuracy (corn/rice/wheat) Lower for comparable models High but with increased computational cost Generalization across crops High; no architectural modifications Moderate; often requires re-training High but computationally expensive needed Enabled using multi-scale filters Nil Partially enabled through attention Multi-scale feature extraction in a single layer mechanisms Low; high computational requirements Computational efficiency High; suitable for resource-constrained Moderate devices Flexibility Robust for various diseases and crops Limited to specific conditions Robust but dependent on large-scale resources Deployment feasibility Requires high-end hardware Highly deployable on edge devices Limited to mid-range devices Scalability Excellent; works effectively across data-High but with higher cost

l imited

Classification outcomes of the proposed model

In this section, we demonstrate how the proposed model performs in disease recognition for the CRW under three distinct scenarios. To begin with, the model differentiates healthy and sick crops of each wheat class with a remarkable degree of accuracy allowing for efficient and early detection of crop diseases. Second, it is able to perform discrimination of multi-class classification on many diseases affecting each crop type; the confusion matrices and classification reports support this claim which means that there is effectiveness in distinguishing the disorders and at the same time, there is emphasis on the more specific areas that need to be improved. Finally, our most challenging application, the model demonstrates high confidence in labeling plants that are healthy and those which have been sick in all the crops with a higher

 Table 10
 Proposed and SOTA models metric comparison

Models	Sp (%)	Pr (%)	Valid Acc (%)	Test Acc (%)
VGG19	79.36	80.41	80.91	80.36
EfficientNetb6	78.81	81.39	81.89	79.80
ResNeXt	81.09	83.62	84.12	82.09
DenseNet201	78.87	79.09	79.59	79.87
AlexNet	82.91	83.34	83.84	83.91
YOLOv5	77.72	80.07	80.57	78.72
MobileNetV3	77.13	78.52	79.02	78.13
Proposed model	87.12	86.06	86.56	88.12

level of accuracy than all the other models. In general, the evidence indicates that indeed, the proposed model performs suitably when it comes to disease detection in CRW. These findings can help improve targeting disease and therefore help in site specific agriculture.

CRW-normal versus infected crops The verification of the new and effective model for the purpose of disease detection in various crops has seen many rates of success which suggests the applicability of the model in different agricultural settings shown in Table 4. For Corn (C), the accuracy of 99.81% is recorded indicating that it was able to separate healthy and infected corn plants. Such accuracy suggests that the model could be useful for early detection of diseases and eventually help reduce crop loss and increase yield. For the case of Rice (R), the accuracy was 87.11%. This is a good accuracy however; it implies that there are essential difficulties related to the disease diagnostics of rice crops as compared to corn and wheat. Additional evaluation may help determine the reason for this particular low performance such as the level of involvement of rice diseases or the differences in the data used to train the model. Finally, for Wheat (W), the model was applied with an accuracy of 98.45%. This observation confirms strong distinguishing ability towards healthy versus diseased wheat plants and thus the model shall be of immense help in the control of the wheat diseases. As a last step in this section, each crop will be assessed separately with regards to the previously described model and this information will be provided in the following subsection.

Crop-wise disease classification This subsection presents the results pertaining to disease classification of CRW crops using the proposed model one at a time. In this subsection, the results of the proposed model for disease classification in CRW acceptable models as demonstrated in the above separately evaluated. This helps to examine the effectiveness of the model on each crop in regards to discriminative healthy and diseased plants of these crops. The classification report for Corn is presented in Table 5, and it provides further models' figures in terms of the classification metrics of accuracy, precision, recall, and F1 score on the diseases of corn. This indicates that the model is also efficient in detecting corn diseases which boosts its use in general crop health monitoring in precision agriculture.

The performance achieved in classifying different types of corn diseases and healthy corn is very commendable. With regard to Corn Grey Spot, the model obtained a precision of 98.12%, a recall of 99.33% and F1 Score of 97.21% all of this done with an accuracy of 99.81%. In respect of Corn Northern Blight, precision was 98.33%, recall was 97.21% and F1 Score was 99.42% with an overall classification accuracy of 99.80%. In the case of Corn Rust, precision was 97.33%, recall was 99.25%, F1 Score was in the range of 96.51%, accuracy level was 99.91%. Finally, concerning Healthy Corn, the results were still high with precision of 99.22%, Recall of 98.36%, and F1 score of 98.55%, resulting into an accuracy of 99.86%.

Classification accuracy of various rice conditions has been demonstrated to be quite promising in Table 6. For Healthy Rice, the model reached 85.37% in precision, 86.88% in recall, an F1 score of 87.12% with an overall accuracy of 86.08%. In the case of Rice Brown Spot, precision reached 88.44%, recall was 86.75% and F1 score achieved 85.98% giving rise to an accuracy of 87.32%. For Rice Hispa, the model reached 85.87% precision, 88.12% recall and 87.09% F1 score to achieve 87.61% accuracy. Finally, in the case of Rice Leaf Blast, accuracy of precision 87.52%, recall of 87.88% and F1 scores of 88.02% were achieved, and an F1 score of 87.5 degrees was achieved.

The classification results for various wheat diseases and healthy wheat demonstrate outstanding performance shown in Table 7. For normal wheat, the model achieved a precision of 98.09%, recall of 99.15%, and an F1 score of 97.21%, with an accuracy of 98.12%. In the case of Wheat Leaf Rust, the model attained a precision of 98.11%, recall of 97.54%, and an F1 score of 99.04%, leading to an accuracy of 97.49%. For Wheat Stem Rust, the precision was 97.23%, recall was 99.33%, and the F1 score was 96.22%, resulting in an accuracy of 97.98%. Lastly, for Wheat Stripe Rust, the performance remained high with a precision of 99.02%, recall of 98.45%, and an F1 score of 98.76%, with an accuracy of 98.49%. Figure 6, depicts the overall classification accuracy of all three crops like; Corn, Rice and Wheat obtained by proposed model. The confusion matrix illustrates the performance of a classification model applied to three classes: Corn, Rice, and Wheat depicted in Fig. 7. The matrix shows that the

Table 11Classification accuracy summary of proposed and
other SOTA methods

Author	Accuracy (%)
Wasswa Shafik et al. [13]	97.79
Arunangshu Pal et al. [14]	98.96
Kumar Satyam Tanti et al. [15]	99.22
Andrew et al. [16]	99.15
Khalil Khan et al. [17]	97.60
Mingyue Guo et al. [18]	98.96
Aayush Deshmukh et al. [19]	95.23
Aadarsh Kumar Singh et al. [20]	99.32
Sheng Yu et al. [21]	99.64
Yuzhi Wang et al. [22]	99.61
Proposed model	99.81

model performs exceptionally well in identifying Corn, correctly classifying 997 instances, with only 1 instance misclassified as Rice and 1 as Wheat. For Rice, the model correctly identifies 752 instances out of 864, but misclassifies 56 instances each as Corn and Wheat, indicating some difficulty in distinguishing this class. In the case of Wheat, the model achieves high accuracy by correctly classifying 791 instances, with only minor misclassifications—6 instances as Corn and 7 as Rice.

Performance metric comparison of proposed and SOTA models

The table summarizes four measures, which are Specificity (Sp %), Precision (Pr %), Validation Accuracy (Valid Acc %), and Test Accuracy (Test Acc %), in which the respective models performed. Out of all models, the Proposed model demonstrated overall best performance, that is, a Valid Acc of 86.56% and a Test Acc of 88.12%. At the same time, it is estimated that the model's Sp and Pr are 87.12% and 86.06%, respectively. AlexNet and ResNeXt also give convincing results with a validation accuracy of 83.84% and 84.12%, respectively and test accuracies of more than 82%. YOLOv5 and MobileNetv3, although only slightly lower than this performance limits the Precisions to 80.07% and 78.52%, respectively. In all other aspects, the Proposed model has been shown to outperform all other architectures in regard to the metrics. The Fig. 8 depicts the validation accuracy and loss graph for the proposed model. Figure 9 depicts the ROC curve of all model accuracy comparison. Figure 10 depicts the performance metric comparison of different models. Table 8 provides detailed information on the number of parameters held by all CNN models.

Fig. 11 Graphical representation of accuracy comparison summary

Advantages of proposed study over SOTA models

In order to achieve this goal, we have elaborated on the advantages offered by our Slender-CNN model in particular and our approach in general shown in Table 9. The inclusion of multi-scale filters within a single layer in the Slender-CNN model scaling allows it to excel at extraction of features at varying object scales ensuring robust disease detection at different sizes of some symptoms. In this dimension, it is different from other CNN architectures that depend on the use of fixed scale filters. Additionally, we conducted extensive comparisons with state-of-the-art hybrid models that utilize advanced techniques such as transformer blocks, attention mechanisms, and ensemble learning. Our findings demonstrate that Slender-CNN delivers competitive accuracy while maintaining an exceptionally low parameter count (387,340 parameters), drastically reducing computational and memory demands. This efficiency makes it highly deployable on resource-constrained devices, such as IoT systems or mobile platforms, without sacrificing performance. Furthermore, Slender-CNN achieves high classification accuracy across multiple crops-Corn, Rice, and Wheat-without requiring architectural modifications. The ability of the model to perform so many functions and to learn at such a rate is due to its ability to adapt and learn efficiently across a wide variety of situations and circumstances, making it better than many other DL models. Hence, the proposed model can work under a constrained computing environment and still provide robust performance, making it a useful model in the field of plant disease detection.

Summary

Tables 8 and 10 let us examine two model perspectives and present their results. In the Table 10, comparative graphs of several models were shown which included VGG19, EfficientNetb6, ResNeXt, DenseNet201, AlexNet, YOLOv5, and Mobilenetv3 along with their validation and testing accuracy. Moreover, in highlight of this table, the proposed model achieves a validation accuracy of 86.56% and a test accuracy of 88.12%, thereby outperforming all the other models. In this regard, the validation and test accuracy of ResNeXt is recorded at 84.12% and 82.09% respectively, recursively while it was also found that Alexnet exhibited 83.84% validation and 83.91% test accuracy. Turning to the Table 8, accuracy is further evaluated for one model dataset (Plant-Village) where models' information such as LeafyGAN and visually DenseNet, Visual Transformer, etc., and their performance accuracy for plant disease identification is assessed. In this regard, the proposed model using Slender-CNN exhibited 99.81% accuracy which was found to be more than all other models including transformer-based models and DenseNet based architecture. In both tables, it is worth noting that the model proposed in this work is the best among the rest. In the first table, it performs the best in validation and test accuracy for an altogether different task while in the second table, it performs the best in accuracy for the detection of the plant disease. These findings demonstrate the flexibility and effectiveness of the model as it provides better performance than famous architectures like ResNeXt, AlexNet and DenseNet, which makes it suitable in various ML applications.

Potential challenges

The proposed model aims for balance between efficiency and computational requirements when applied in real agricultural conditions. However, there are some potential challenges in the proposed study, which include:

- The model has reported an impressive performance and accuracy on controlled datasets. However, difficulties may arise in uncontrolled agricultural environments owing to factors such as lighting, noise in the images, or differences in the appearance of the crops. In real-life situations, where a great number of diseases or diseases in a certain environment are present, it is very difficult to collect well-annotated samples.
- Despite the fact that proposed model requires fewer parameters, it is important to emphasize that usage of DL models in practice especially on mobile or edge devices can still be quite computationally expensive. One of the causes of this may be how the model will operate on devices with limited computing power, memory, and available disk space.
- To ensure efficiency in disease detection, the model needs to perform inferences in real-time. For practical applications in particular in vast farms where there may be an emergency situation and there is a threat of spreading disease, it is vitally important to optimize the model in a way that it gives accurate predictions in a very short time without compromising on any performance.

Discussion

The performance of the proposed model is compared with the performance of many models found in previous literature to showcase its effectiveness. The baseline performance of the model as showcased by Wasswa Shafik et al. [13] is 97.79%. Significantly better results were achieved by Arunangshu Pal et al. [14] and Mingyue Guo et al. [18] who managed to report an accuracy result of 98.96%.

Kumar Satyam Tanti et al. [15] achieved an accuracy result of 99.22%, which is lower than Andrew et al. [16] who reported an accuracy of 99.15%. Some walls can be pushed against with what Khalil Khan et al. [17] obtained, which was a 97.60% accuracy. A lower accuracy result was also reported by Aayush Deshmukh et al. [19] with 95.23%, which means the approach taken was less effective. The most competitive models in the literature were developed by Aadarsh Kumar Singh et al. [20], Sheng Yu et al. [21] and Yuzhi Wang et al. [22] with accuracies of 99.32%, 99.64%, and 99.61%, respectively. Likewise, the model framework proposed improves on previous methods, achieving 99.81% accuracy which is higher than any of the previously reported models. The evaluated models in this section positively reflect the enhancements towards the accuracy, effectiveness, and novel aspects of the proposed method of perhaps other methods available and justifies the model's ability to be placed among the better models today. The results can be considered effective when the established model has a level of accuracy which goes above and beyond what previous models achieved. Table 11 demonstrates the overall classification summary of SOTA models and compared with proposed model. Figure 11 depicts the graphical illustration of accuracy comparison result of proposed and other SOTA model and highlighted the proposed model outcome.

Conclusion and future work

Crop diseases have long been one of the reasons for food insecurity across the globe, thereby yielding the need for their early detection. Corn, rice, and wheat (CRW) happen to be among the most widespread food crops in the world, necessitating the need for timely diagnosis of diseases affecting those crops in order to enhance food security. Although there are many DL and ML models for plant disease classification in CRW crops, the deployment of these models is hampered since many farmers are not digital literate in developing countries. In this work, the Slender-CNN model is proposed as a means of accomplishing diseases in CRW crops without creating different models for different plants. The model is also appropriate for use in low resource settings, as it contains a low number of parameters, with only 387,340 parameters. It reliable, since it achieves even an accuracy of 86.56%, which is better than optimization of some benchmark CNN models in both accuracy and parameters. Further experiments reinforce the suitability of the model for a variety of experimental conditions and tasks concerning the diagnosis of diseases in single CRW crops.

The present study has considered some common diseases present in Corn, Rice, and Wheat that are Corn Grey Spot, Corn Northern Blight, Corn Rust, Rice Brown Spot, Rice Hispa, Rice Leaf Blast, Wheat Leaf Rust, Wheat Stem Rust, and Wheat Stripe Rust. On the other hand, the proposed of the novel approach has not been tried out on Downy Mildew disease in Rice, Eyespot disease in Corn, Flag Smut disease in Wheat et cetera. In the same way, the model has not been applied in the diagnosis of diseases in other crops. For practical uses, the slender form factor of the model allows combination with portable gadgets in the future allowing farmers' easy diagnosis of diseases. It may also be used on drones as a way of reducing the manual inspection of plants by hand. The model, by incorporating sensor inputs and images of plants obtained from the crop fields, is expected to improve disease classification and uncover disease patterns influenced by environmental and farming factors.

Likewise, it can also be scaled to crops other than Corn, Rice and Wheat. The architecture's versatility allows it to be set up with additional parameters and expand the model so that it's possible to not only deal with disease image recognition but also include climatic conditions and agricultural management practices that are likely to change the total disease pattern. Such flexibility might provide the basis for future, more advanced and reliable disease detection systems. Besides, Slender-CNN can be improved by adding features which enable it to learn in real time or an enhancement whereby the model will get more accurate with field data collected over time. This could foster the creation of stronger disease prediction models that will have the ability to predict outbreaks based on available early warning signs and environmental conditions.

Acknowledgements

Not applicable.

Author contributions

Author Contributions: "Conceptualization, B.B.V and N.K; methodology, S.K.M.; validation, S.S and S.K.M; resources, M.A.S; data curation, A.P; writing—original draft preparation, B.B.V and N.K; writing—review and editing, S.K.M, and S.S; visualization, M.A.S and S.S; supervision S.K.M and M.A.S; project administration. S.K.M and M.A.S.

Funding

This research received no external funding.

Data availability

No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate Not applicable.

Institutional review board statement Not applicable.

Informed consent Not applicable.

Competing interests

The authors declare no competing interests.

Received: 26 September 2024 Accepted: 25 January 2025 Published online: 13 February 2025

References

- Agri share in GDP hit 20% after 17 years: economic survey. Available: https://www.downtoearth.org.in/news/agriculture/agri-share-in-gdp-hit-20-after-17-years-economicsurvey-75271. Accessed Aug 2021.
- India—Employment in Agriculture (% of total Employment). Available: India – Employment In Agriculture (% Of Total Employment)—2022 Data 2023 Forecast 1991–2020 Historical (tradingeconomics.com). Accessed Aug 2021.
- Kumar M, Kumar A, Palaparthy VS. Soil sensors-based prediction system for plant diseases using exploratory data analysis and ML. IEEE Sens J. 2021;21(16):17455–68.
- Nagasubramanian G, Sakthivel RK, Patan R, Sankayya M, Daneshmand M, Gandomi AH. Ensemble classification and IoT-based pattern recognition for crop disease monitoring system. IEEE Internet Things J. 2021;8(16):12847–54.
- Blesslin Elizabeth CP, Baulkani S. Novel network for medicinal leaves identification. IETE J Res. 2022;69(4):1–18.
- Yang Y, Liu Z, Huang M, Zhu Q, Zhao X. Automatic detection of multi-type defects on potatoes using multispectral imaging combined with a DL model. J Food Eng. 2023;336: 111213.
- 7. Anwar Z, Masood S. Exploring deep ensemble model for insect and pest detection from images. Proc Comput Sci. 2023;218:2328–37.
- Ma R, Wang J, Zhao W, Guo H, Dai D. Identification of maize seed varieties using MobileNetV2 with improved attention mechanism CBAM. Agriculture. 2023;13(1):1–11.
- Srilatha D, Thillaiarasu N. Implementation of intrusion detection and prevention with DL in cloud computing. J Inform Technol Manag. 2023;15:1–8.
- Zhang Y, Yang Y, Zhang J, Wang Y. Sensitivity study of multi-field information maps of typical landslides in mining areas based on transfer learning. Front Earth Sci. 2023;11:1105985.
- Rao DS, Ch RB, Kiran VS, Rajasekhar N, Srinivas K, Akshay PS, Mohan GS, Bharadwaj BL. Plant disease classification using deep bilinear CNN. Intell Autom Soft Comput. 2022;31(1):161–76.
- Li X, Li X, Zhang S, Zhang G, Zhang M, Shang H. SLViT: shuffle-convolution-based lightweight Vision transformer for effective diagnosis of sugarcane leaf diseases. J King Saud Univ Comput Inf Sci. 2023;35(6): 101401.
- Shahid MF, Khanzada TJS, Aslam MA, Hussain S, Baowidan SA, Ashari RB. An ensemble DL models approach using image analysis for cotton crop classification in Al-enabled smart agriculture. Plant Methods. 2024;20:1–22.
- Duan Z, Li H, Li C, Zhang J, Zhang D, Fan X, Chen X. A CNN model for early detection of pepper Phytophthora blight using multispectral imaging, integrating spectral and textural information. Plant Methods. 2024;20(115):1–12.
- Yang G, Chen G, Chen G, Li C, Jiangfan Fu, Guo Y, Liang H. Convolutional rebalancing network for the classification of large imbalanced rice pest and disease datasets in the field. Front Plant Sci. 2021;12:1–14.
- Lowe A, Harrison N, French AP. Hyperspectral image analysis techniques for the detection and classification of the early onset of plant disease and stress. Plant Methods. 2021;13(80):1–13.
- Paymode AS, Malode VB. Transfer learning for multi-crop leaf disease image classification using convolutional neural network VGG. Artif Intell Agric. 2022;6:23–33.
- Shafik W, Tufail A, De Silva C, Liyanage RA, Apong AHM. Using transfer learning-based plant disease classification and detection for sustainable agriculture. BMC Plant Biol. 2024;24(136):1–19.
- Pal A, Kumar V. AgriDet: plant leaf disease severity classification using agriculture detection framework. Eng Appl Artif Intell. 2023;119: 105754.
- 20. Tanti KS, Gupta M, Kumar R, Obaid AJ. Identification of Plant Leaf Disease using Image Augmentation and DL. In: 2024 3rd International

Conference on Computational Modelling, Simulation and Optimization (ICCMSO), 2024. https://doi.org/10.1109/ICCMSO61761.2024.00031.

- Andrew J, Eunice J, Popescu DE, Kalpana Chowdary M, Hemanth J. DL-based leaf disease detection in crops using images for agricultural applications. Agronomy. 2022;12(10):1–19.
- Khan K, Khan RU, Albattah W, Qamar AM. End-to-end semantic leaf segmentation framework for plants disease classification. Complexity. 2022;1168700:1–11.
- Guo M, Li Q, Liu D. An end-to-end dl based approach for plant disease recognition. In: 2024 6th International Conference on Communications, Information System and Computer Engineering (CISCE), 2024. https://doi. org/10.1109/CISCE62493.2024.10653053.
- 24. Deshmukh A, Verma A, Singh VK, Shivhare SN. 'Efficient plant leaf disease detection using a customized convolutional neural network', data science and applications, ICDSA 2023. Lect Notes Netw Syst. 2024;820:383–94.
- Singh AK, Rao A, Chattopadhyay P, Maurya R, Singh L. Effective plant disease diagnosis using vision transformer trained with leafy-generative adversarial network-generated images. Expert Syst Appl. 2024;254: 124387.
- Wang Y, Yin Y, Li Y, Tengteng Qu, Guo Z, Peng M, Jia S, Wang Q, Zhang W, Li F. Classification of plant leaf disease recognition based on self-supervised learning. Agronomy. 2024;14(3):1–13.
- Plant Village Dataset. Available: https://www.tensorflow.org/datasets/ catalog/plant_village. Accessed Nov 2022.
- 28. Rice Leafs. Available: https://www.kaggle.com/datasets/shayanriyaz/Ricel eafs. Accessed Nov 2022.
- 29. Wheat Leaf Dataset. Available: https://www.kaggle.com/datasets/olyad getch/Wheat-leaf-dataset. Accessed Nov 2022.
- 30. CGIAR Computer Vision for Crop Disease. Available:https://www.kaggle. com/datasets/shadabhussain/cgiarcomputer-vision-for-crop-disease. Accessed Nov 2022.
- Shobana M, Vaishnavi S, Gokul Prasad C, Pranava Kailash SP, Madhumitha K P, Nitheesh C, Kumaresan N. Plant disease detection using convolution neural network. In: 2022 International Conference on Computer Communication and Informatics (ICCCI), 2022. https://doi.org/10.1109/ICCCI 54379.2022.9740975.
- Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, Andreetto M, Adam H. Mobilenets: efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861. 2017.
- Chen R, Qi H, Haixia Qi Yu, Liang LY, YangMing M, Yang C. Identification of plant leaf diseases by DL based on channel attention and channel pruning. Front Plant Sci. 2022;13:1–15.
- Pandey A, Jain K. A robust deep attention dense convolutional neural network for plant leaf disease identification and classification from smart phone captured real world images. Eco Inform. 2022;70: 101725.
- Hukkeri GS, Soundarya BC, Gururaj HL, Ravi V. Classification of various plant leaf disease using pretrained convolutional neural network on Imagenet. Open Agric J. 2024;18(1):1–15.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.