GLDICCNN Model: Groundnut Leaf Diseases Identification and Classification for Multiclass Classification Using Deep Learning

Anna Anbumozhi * and A. Shanthini

Department of Data Science and Business Systems, School of Computing,
SRM Institute of Science and Technology, Kattankulathur, Tamilnadu, India
Email: annaa2@srmist.edu.in (A.A.); shanthia@srmist.edu.in (A.S.)
*Corresponding author

***Abstract*—Plant diseases must be identified early to protect crop harvests, as agriculture plays a crucial role in ensuring global food security. This paper introduces an advanced deep-learning approach utilizing a conventional Convolutional Neural Network (CNN) for the multiclass classification of groundnut leaf diseases. The research focuses on constructing a robust deep learning model, named Groundnut Leaf Disease Identification Classification Convolution Neural Network (GLDICCNN), to rapidly identify, classify, and predict groundnut leaf diseases. The methodology encompasses comprehensive data collection from agricultural fields, preprocessing, model development, and rigorous evaluation. The proposed model, Groundnut Leaf Diseases Identification, Classification with Convolution Neural Network (GLDICCNN) demonstrates impressive performance metrics after extensive experimentation. The training accuracy reaches 99.73%, while validation accuracy stands at 97.06%. Correspondingly, the training and validation loss values are 0.0035 and 0.1649, respectively. Evaluation metrics, including precision (96%), recall (96%), and F1–Score (96%), highlight the effectiveness of the proposed model. Moreover, the test accuracy attains a commendable 96%. Comparative analysis with pre-trained models such as ResNet50, ResNet101, and ResNet152 underscores the superior accuracy achieved by the GLDICCNN model. In summary, this research establishes a deep learning framework that excels in groundnut leaf disease identification, classification, and prediction. The findings underscore the potential of the proposed model for practical applications in agriculture, contributing to enhanced crop yield protection and global food security.**

***Keywords*—agriculture, augmentation, deep learning, groundnut leaf, hyperparameter tuning, Residual Neural Network (ResNet) models, transfer learning**

I. INTRODUCTION

Agriculture, as the bedrock of human civilization, has evolved from a fundamental survival activity to a sophisticated science driving global sustenance. Within the dynamic tapestry of agriculture, the persistent threat of leaf diseases poses a significant challenge to crop health and global food security. The comprehensive detection of leaf diseases has become more critical due to the changing climatic circumstances and the interconnection of global trade. A Machine Learning or Deep Learning model can change crop management which can be represented by Convolution Neural Networks [1] in Artificial Intelligence [2], which provide data-driven insights, automation, and precision that have the potential to completely transform conventional agricultural methods and increase the productivity, resilience, and sustainability of agriculture. The integration of AI (Artificial Intelligence) applications in crop management exemplifies a fundamental change in agriculture, emphasizing data-driven decision-making, resource optimization, and sustainable practices. These technologies contribute to increased efficiency, reduced environmental impact, and the overall resilience of agricultural systems. In the dynamic landscape of agriculture, groundnuts, or peanuts, emerge as a versatile and indispensable crop.

Identification of groundnut leaf diseases, classification of groundnut leaf diseases [3], and Prediction of leaf diseases can bring into line with contemporary agricultural trends that emphasize technology-driven solutions for optimizing crop management, ensuring food security, and promoting sustainable farming practices. The susceptibility of groundnut plants to various diseases, including foliar diseases, significantly impacts their health and yield. The timely identification and categorization of these diseases are of utmost importance in ensuring efficient agricultural administration and mitigating potential reductions in crop productivity. The conventional approaches to identifying groundnut leaf diseases include visual inspection, field surveys, Symptom Catalogs and Guides, Local Knowledge and Experience, Leaf Sampling and Laboratory Analysis, and Expert Consultation. However, this method is associated with certain drawbacks including time consumption, labor intensiveness, and susceptibility to human fallibility. The limitations of the existing approaches can be overcome by integrating AI into agriculture. There are several AI methods are employed to identify the groundnut leaf

Manuscript received December 22, 2023; revised February 2, 2024; accepted February 18, 2024; published July 8, 2024.

diseases such as Machine Learning Algorithms, Image Recognition, Computer Vision, Deep Learning, Transfer Learning, and decision support systems. Deep learning approaches have demonstrated significant success in accurately identifying groundnut leaf diseases. To identify, classify, and predict the image data in deep learning [4] there are two approaches are available. Two widely employed techniques in this context are Convolutional Neural Networks (CNN) and Transfer Learning utilizing pre-trained network models. The disadvantages of incorporating two approaches for groundnut leaf disease identification, classification, and prediction are computational complexity, Overfitting issues, lower accuracy, limited memory consideration, and difficulty in handlining sequential data. The proposed GLDICCNN models identify, classify, and predict groundnut leaf diseases and solve the overfitting issues.

The contributions of this research work are:

- Acquire and preprocess a comprehensive dataset of groundnut leaf images, including images of both diseased and healthy foliage.
- To design and develop a new model Groundnut Leaf Diseases Identification Classification and Prediction Convolutional Neural Network (GLDICP-CNN) model tailored for the groundnut leaf diseases problem.
- For the training and evaluation of the GLDICP-CNN model in disease identification, classification, and prediction, precise and efficient performance metrics will be employed.
- For a thorough evaluation, the proposed model was compared with existing methods and established benchmarks to prove better performance.
- To provide insight into the expected uses of the produced technology and the limits of such a system.

This paper aligns the content as Segment II provides a comprehensive related work on the detection, classification, and prediction of leaf diseases. Segment III delves into the methodologies employed for data collection and preprocessing. Segment IV expounds on the methodology, detailing the overall architecture of the proposed GLDICCNN model. Segment V presents the experimental results and performance metrics with a discussion of the findings. Finally, Segment VI details the conclusion and future work of the research.

II. Related Work

The researcher used a deep learning-based CNN technique called modified InceptionResNet-V2 [5] for the detection and classification of plant leaf diseases. The MIR-V2 model achieved accuracy of 98.92% and 97.94%, this model was compared with the tomato leaf disease detection algorithm and outperformed than previous model. This model cannot be refined still more accuracy and needs correct diagnosis technique to implement. The author incorporates the Ant Colony optimizer [6] and CNN for feature extraction and classifies the infected leaf and healthy leaf from the class images. Early Weed Identification with Deep Learning: Utilizes Keras and PyTorch to train models on an annotated weed dataset, demonstrating accurate image classification and object detection for early-season weeds [7]. Acknowledges potential misclassification, suggesting the need for additional field testing. Sarkar *et al.* [8] Researchers assessed recent papers on Machine Learning (ML) and Deep Learning (DL) for Image Classification, noting that many ML approaches fall short in accuracy despite large datasets. In contrast, the study finds that deep learning models consistently achieve satisfactory performance metrics in image classifications [9]. Combining Channel Attention and Pruning Network (CACPNET), Efficient Plant Disease Detection: Introduces CACPNET, a model combining channel attention and pruning, achieving high accuracy (99.7% on PlantVillage, 97.7% locally) with reduced parameters. Outperforms ResNet-18, emphasizing efficiency for plant disease detection in precision agriculture.

Chen *et al.* [10] proposed a Neural Network (NN) based method for plant leaf disease identification, employing segmentation algorithms and extracting Haralick features and textures. The neural network achieves acceptable accuracy but is limited by its dependence on clear input images. Thaiyalnayaki and Joseph [11] focused on plant disease classification, and the paper combines Support Vector Machine (SVM) and deep learning. SVM achieves 94.14% accuracy in classifying 19 diseases based on plant attributes, while the deep learning Multi-Layer Perceptron (MLP) classifier achieves 88.73% accuracy, highlighting its performance on soybean data [12]. DL Architectures and Optimizers for Plant Disease: Compares DL architectures and optimizers on the PlantVillage dataset, highlighting F1–Score. Recommends Xception for high performance but notes the need for further research to validate results and explore additional performance measures. Inception-based DL Model for Plant Diseases: Proposes an inception-based DL model achieving high accuracy (up to 99.66%) on various plant disease datasets. Limited by the scope of three datasets and lacks comparison with traditional ML methods [13]. X-caption-based Model for Crop Diseases: Introduces a deep learning model with X-caption architecture achieving 99.69% accuracy in identifying peanut leaf diseases. Successful application to various crops demonstrates versatility but lacks examination of model limits and real-world implementation [14].

Automated Peanut Leaf Spot Disease Detection: Uses a pre-trained DenseNet-169 model for peanut leaf spot disease identification, showing potential for agricultural automation. Emphasizes high accuracy and precision, particularly useful for peanut disease diagnosis [15]. DenseNet for Groundnut Leaf Disease Classification: Presents a groundnut leaf dataset, employing DenseNet for classification with a 99.83% success rate. Recognizes dataset importance and ethical considerations but lacks insight into model limitations and real-world applications [16]. The PGCNN (Progressive Groundnut Convolutional Neural Network) model is a novel approach to groundnut leaf disease classification. It achieves a training accuracy of 99.39%. shows potential for generalization, but more study is required to confirm performance in a variety of contexts [17]. PGCNN (Progressive Groundnut Convolutional Neural Network) Model for Groundnut Leaf Diseases: Introduces PGCNN model achieving 99.39% training accuracy for groundnut leaf disease classification. Shows promise in generalization but needs additional research to verify performance across diverse settings [17]. Deep CNN for Groundnut Leaf Spot Illness: This method uses deep CNN to classify groundnut leaf spot illness, with 0.844 accuracy using DFT pictures [18]. Constraints include a narrow emphasis on particular illnesses and a lack of details on the approach and dataset size [19]. CNN Models for Predicting Groundnut Diseases: Introduces CNN models for predicting groundnut diseases, with ResNet achieving high accuracy and precision. However, the study lacks a detailed dataset and validation information, necessitating additional investigation.

In summary, the review highlights the effectiveness of deep learning models, particularly in image classification, plant disease diagnosis, and groundnut leaf identification. The application of Keras and PyTorch in identifying early-season weeds demonstrates the potential for site-specific agricultural weed management. While neural networks, SVM, and deep learning architectures prove valuable in plant disease identification, the need for clear input photos and further research to validate results is emphasized. The proposed inception-based model showcases superior accuracy with fewer parameters, and the impressive performance of models like X-caption, Densely Convolutional Networks (DenseNet), PGCNN, and ResNet in groundnut leaf disease diagnosis underscores the potential of advanced topologies. However, the overall lack of standardized information on dataset sizes, validation methodologies, and performance measures calls for additional research to validate these models across diverse scenarios and address real-world implementation challenges and ethical considerations.

III. DATA COLLECTION AND PRE-PROCESSING

A. Data Source Unit

In this subsection, the author provides an account of the origins of the groundnut leaf image data collected for experimental purposes. Leaves were gathered from the agricultural field, encompassing diverse climatic conditions in the vicinity of Venkatakulam village in Pudukkottai district, Tamil Nadu. The collected samples include Early_Leaf_Spot (ELS), Late_Leaf_Spot (LLS), Rust, and Rosette, alongside a category representing healthy leaves. To get pictures of groundnut leaves, high-resolution digital cameras and smartphones—which are Internet of Things (IoT) devices—were used [20]. The parameters of the smartphone used to take pictures of the healthy and diseased leaves from the farmed area are listed in Table I.

B. Data Acquisition

The data acquisition process encompassed the systematic collection of groundnut leaf images from the sources above [21]. To ensure the quality and consistency of the data, the subsequent measures were implemented. Below is Table II. details the image specification of the groundnut dataset.

TABLE I. CAMERA SPECIFICATIONS

Camera	VIVO V2029
Processor	1.8GHz Snapdragon SDM460 Octa-core
OS	Android 10
RAM	6.00 GB
Phone Storage	64.00 GB
Front Camera Res.	8 MP
Rear Camera Res.	13 MP + 2 MP + 2 MP
Battery Capacity	5000 mAh

TABLE II. GROUNDNUT LEAF IMAGE DETAILED SPECIFICATIONS

Image Property	Values	Camera Property	Values
Dimensions	1380×1380	Camera Maker	Vivo
Width	1380	Camera Model	V2029
Height	1380	F-stop	f/2.2
Horizontal Resolution	72 dpi	Exposure time	1/100 s
Vertical Resolution	72 dpi	Focal length	4 mm
Bit Depth	24	Max aperture	2.27
GPS			
Latitude	10;19;28.0599000000002263		
Longitude	78;55;36.6300000000048		

C. Data Cleaning and Augmentation

To optimize the dataset for deep learning purposes, a sequence of data cleaning and augmentation methodologies was employed. Below Fig. 1 shows the sample images of a dataset with classes of 5. The cleaning and data augmentation techniques used to enhance the results are rotation, flipping, shifting, noise removal, shearing, and class balancing.

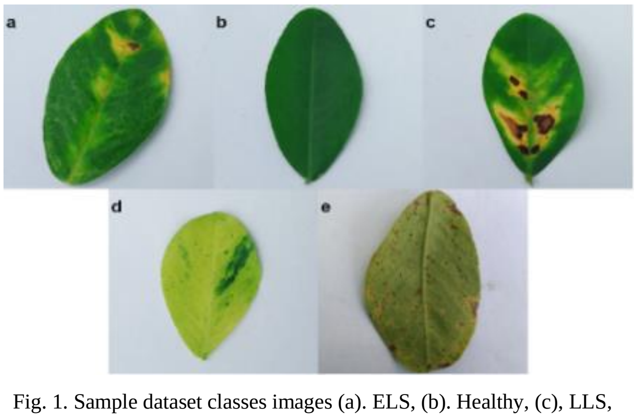

Fig. 1. Sample dataset classes images (a). ELS, (b). Healthy, (c), LLS, (d). Rosette, (e). Rust.

- **Rotation:** Introduces diversity by changing the orientation of images, facilitating the model's ability to recognize objects from various angles.
- **Flipping:** Enhances dataset diversity through horizontal or vertical mirroring of images, providing additional perspectives for model training.

- **Shifting**: Introduces spatial variability by translating image positions horizontally or vertically, aiding the model in handling changes in object location.
- **Noise Removal:** To enhance the quality of the data, images exhibiting substantial noise, such as blurriness or inadequate lighting, backgrounds, and weeds in the agricultural land were excluded.
- **Shearing:** The employed techniques encompassed a random rotations range of 20, a width shift was 0.2, a height shift range was 0.2, a shear range was 0.2, a zoom range was 0.2 with the horizontal flip with true and the nearest of fill mode.
- **Class Balancing:** The implemented strategies such as oversampling and undersampling to address the imbalanced class distribution caused by varying sample sizes across disease categories.

D. Data Splitting

The dataset is divided into three sets: the training set, the validation set, and the test set. This separation is to assess the model's performance, preventing overfitting, and fine-tuning the hyperparameters. Table III provides a detailed breakdown of the crop dataset before and after augmentation for each class. It suggests that data augmentation has been implemented to enlarge the dataset size to achieve better performance, and it shows the distribution of images across different leaf states.

TABLE III. NUMBER OF LEAVES COLLECTED FOR GLDIC-CNN MODEL BUILDING

Sl. No.	Class Names	No. of Original Images	No. of Leaves after Augmentation
1	Early Leaf Spot	155	1000
2	Healthy	180	650
3	Late Leaf Spot	108	1010
4	Rosette	24	1000
5	Rust	48	1010
Total		**515**	**4670**

IV. METHODOLOGY

A. Architecture Overview

The CNN-based model employed in this study is specifically built to acquire hierarchical features from images of groundnut leaves. By doing so, it can effectively capture and analyze patterns and traits diagnostic of different diseases affecting the crop. Fig. 2 shows the architecture and overall workflow of the GLDICCNN model.

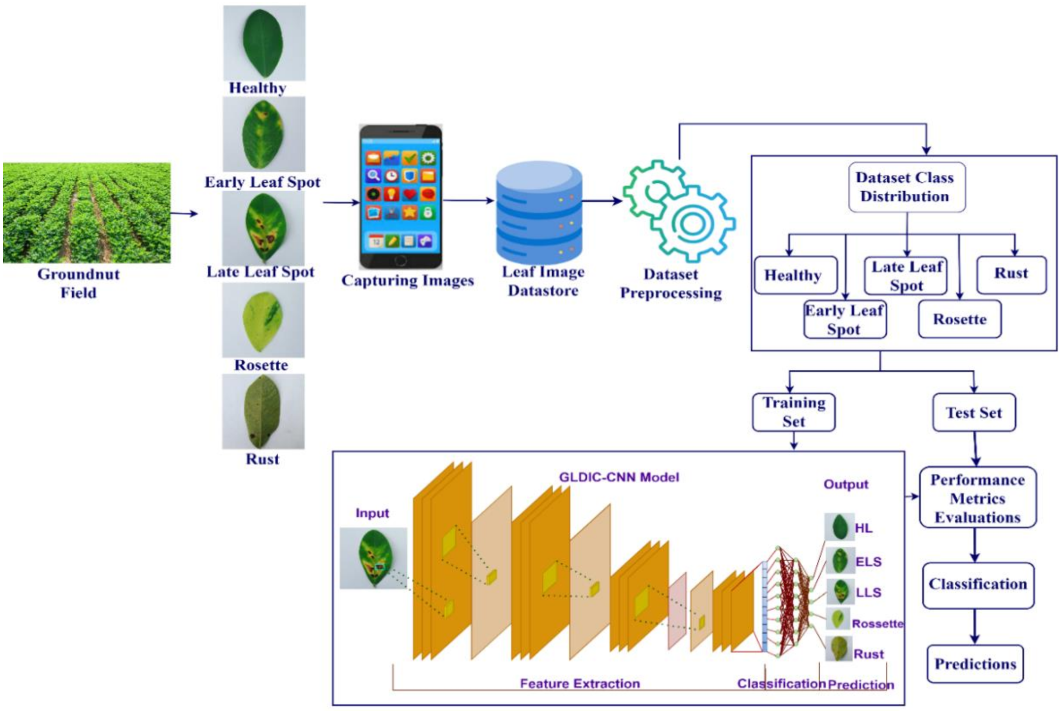

Fig. 2. Overall architecture of multiclass classification using deep learning techniques.

The architecture consists of six convolution layers tasked to extract primary to higher features, Max-pooling layers to reduce the spatial dimension, and one fully connected layer dedicated to classification. Groundnut leaves have been collected from agricultural fields with various climatic conditions. Collected leaves were segregated according to the diseases encountered in the leaf which the professional agricultural specialists ensured. The collected dataset leaves were cross-checked with the leaf available in the Tamilnadu Agricultural University (TNAU AGRITECH) portal and split into five classes. To capture the images, the smartphone is used to collect the healthy and diseased types of groundnut leaf images. Images are stored in the local Graphics Processing Unit (GPU) system for the train-test split of the dataset for further processing. The five classes are Early Spot, Healthy, Late Spot, Rosette, and Rust. To obtain high-performance accuracy, the dataset should be pre-processed before the procedure begins. It involves several steps to enhance the quality of the data. Typical preprocessing steps for image classification involve loading the data, resizing the images, normalizing pixel values to a specific range, and augmenting the dataset through transformations like rotation, flipping, zooming, and shifting. Data augmentation serves to enhance the model's generalization by introducing variations in the input data, alongside operations like splitting the data. For preprocessing, some techniques are used to they are organizing the dataset in a structured manner, resizing the images to ensure that all the images are in the same dimensions to make the neural network easier, Scale pixel values to a specific range, typically [0, 1] or [−1, 1].

Normalization is utilized to converge faster while the model is in training. Then, the leaves are distributed into five classes according to the diseases infected in the leaves. The dataset was divided into train-test sets with an 80%–20% ratio. These are done to improve model accuracy, reduce variance, assess robustness, and effectively utilize the data. The proposed model was derived with six convolution layers with the iterative number of filter sizes and the same filter size, followed by the Rectified Linear Unit (ReLU) as the activation function. The batch normalization and a dropout layer are included for utilizing the hyperparameters tuning to enhance the results of the proposed model. The learning rate has been tested for 0.1, 0.001, and 0.001 for 50 epochs. The classified images were predicted with the new images to predict the classified images.

B. GLDICCNN Model Architecture

Fig. 3 provides details about the specific architecture of the proposed GLDICCNN model for groundnut leaf diseases.

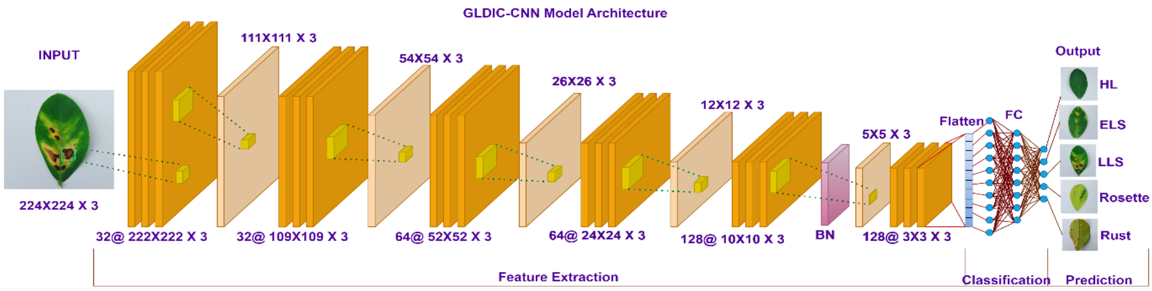

Fig. 3. Groundnut leaf disease identification and classification prediction of CNN.

- Convolution Layer 1: Initiates spatial feature extraction with 32 filters, (3, 3) kernel, and ReLU activation. MaxPooling reduces spatial dimensions for hierarchical learning.
- Convolution Layer 2: Continues extraction with 32 filters and ReLU. L2 regularization (0.01) mitigates overfitting, refining features for robustness.
- Convolution Layers 3 and 4: Introduces 64 filters, (3, 3) kernel in each layer. ReLU maintains non-linearity, and MaxPooling enhances complexity for enriched visual understanding.
- Convolution Layer 5: Utilizes 128 filters, (3, 3) kernel with L2 regularization (0.01). Batch Normalization enhances depth, and controls overfitting, ensuring stable training.
- Convolution Layer 6: Employs 128 filters, (3, 3) kernel, applying Batch Normalization for feature refinement and stability.
- Flatten Layer: Transitions from convolution layers to fully connected layers by flattening output, preparing for densely connected neural network layers.
- Fully Connected Layer 1: Integrates spatial features with 256 neurons, ReLU activation, Batch Normalization, and Dropout (50%) for stable training and prevention of overfitting.
- Output Layer: Final layer with Dense configuration and softmax activation for classification, transforming features into probability distributions for the model's prediction.

Table IV details the parameters used while training the proposed model to get higher accuracy and lower loss values. Here three optimization algorithms were used to compare which suits the proposed model for classification and prediction.

TABLE IV. PARAMETERS DURING TRAINING

Hyperparameters	Values
Loss Function	Categorical cross-entropy
Optimization Algorithm	Adam, SGD, RMSprop
Batch Size	32×64
Epochs	50

B. Hyperparameter Tuning

Hyperparameter tuning is a critical aspect of optimizing the model performance and generalization capabilities [22]. The hyperparameters and the tuning selected for the GLDICCNN are described in Table V.

TABLE V. HYPERPARAMETERS FOR GLICCNN MODEL

Hyperparameters	Hyperparameters Values
Learning Rate tested	0.001
Dropout Ratio selected	0.5 (50%)
Batch Normalization	L1 & L2
Architectural Variants	Fully connected, Convolution, Filters

In summary, the methodology describes the overview of the proposed model used for groundnut leaf disease identification, classification, and prediction using CNNs, including model architecture, training procedure, and hyperparameter tuning.

V. Experimental Results

A. *Interpretation of Results*

The results were obtained from the GLDICCNN model for deep learning-based approach and other comparative methods. The interpretation includes both quantitative and qualitative aspects, providing insights into the performance and significance of the applied approach. This research infers that the proposed model performance gives better values than the existing network of pre-trained models of the Resnet family. Here the trail of all optimizers was utilised to get rid of the overfitting problem. Adam optimizer gives the best result compared to another optimizer. An iterative process called epochs 10, 20, 30, 40, and 50 was also tested and concluded with 50. After executing the proposed model with 50 epochs, the result achieved higher accuracy and was also compared with the Resnet 50, 101, and 152 models to evaluate the proposed model. The results and graph for the achieved results are displayed and tabulated below for a detailed perspective.

Table VI outlines a classification scheme with corresponding labels for different classes, likely intended for a deep learning or data classification task.

TABLE VI. Groundnut Leaf Classes and Labels

Name of the Classes	Label of Classes
Early Leaf Spot	0
Healthy	1
Late Leaf Spot	2
Rosette	3
Rust	4

B. *GLDICCNN Model Training*

Table VII illustrates GLDICCNN's performance metrics across epochs (10 to 50), depicting improvement in both training and validation accuracy. At the 50th epoch, the training accuracy is 99.73%, with minimal loss (0.0035), and validation accuracy is commendable at 97.06%, with a relatively low loss (0.1649). The consistent improvement suggests effective learning and generalization, justifying the pragmatic decision to conclude training at 50 epochs to balance accuracy gains with the risk of overfitting.

Table VIII displays GLDICCNN's performance metrics across epochs (10 to 50), revealing fluctuations in both training and validation accuracy. Starting at 91.16% accuracy and 0.3767 loss at the 10th epoch, training accuracy rises to 98.80% at the 50th epoch.

However, validation accuracy shows less consistent improvement, peaking at 94.39%. Fluctuations in validation loss, reaching 0.2247 at the 50th epoch, suggest challenges in generalizing to new data and potential issues in model performance. Table IX is considered inferior due to the model's inconsistent accuracy improvement and higher validation loss. In contrast, the previous table displayed a more stable and superior performance with higher training accuracy (99.73% at the 50th epoch), minimal training loss (0.0035), and relatively stable, high validation accuracy (97.06%) with a lower validation loss (0.1649). This indicates better generalization and performance on unseen data, making the model in the previous table preferable.

Table IX compares the performance metrics of Adam, RMSprop, and Stochastic Gradient Descent (SGD) optimizers. Adam outperforms, achieving a remarkable training accuracy of 99.83% with low loss (0.0048) and high validation accuracy (97.33%, loss 0.1941). Root Mean Square Propagation (RMSprop) trails slightly with lower training accuracy (98.73%) and higher loss (0.03671), resulting in 95.72% validation accuracy and 0.1948 loss. SGD lags significantly with training accuracy at 48.55%, high loss (1.1530), and validation metrics of 49.49% accuracy and 1.1542 loss. Test accuracy and learning rate further support Adam's superior performance.

TABLE VII. Accuracy, Loss Value of GLDICCNN Model Training and Validation Without Hyperparameter

Epochs	Training Accuracy (%)	Training Loss (%)	Validation Accuracy (%)	Validation Loss (%)
10	95.48	12.78	95.99	12.10
20	97.66	07.18	97.06	10.71
30	99.30	01.54	97.46	13.20
40	98.73	03.67	95.99	19.86
50	**99.73**	**00.35**	**97.06**	**16.49**

TABLE VIII. Accuracy, Loss Value of GLDICCNN Model Training and Validation with Hyperparameter

Epochs	Training Accuracy (%)	Training Loss (%)	Validation Accuracy (%)	Validation Loss (%)
10	91.16	37.67	78.88	79.02
20	95.75	19.73	79.55	86.77
30	97.19	12.56	92.11	29.58
40	96.02	16.91	91.58	33.03
50	**98.80**	**7.92**	**94.39**	**22.47**

TABLE IX. Training and Validation Accuracy and Loss For Adam, RMSPROP, And SGD Optimizers

Optimizers	Training Accuracy (%)	Training Loss (%)	Validation Accuracy (%)	Validation Loss (%)	Test Accuracy (%)	Learning Rate
Adam	**99.83**	**00.48**	**97.33**	**19.41**	**96.38**	
RMSprop	98.73	03.67	95.72	19.48	95.07	**0.001**
SGD	48.55	01.15	49.49	11.54	53.26	

Figs. 4–7 show the training and validation accuracy that assesses the performance of the proposed model. The curves in Figs. 4–7 illustrate the evolution of accuracy metrics throughout the training process. The x-axis denotes the number of training epochs, while the y-axis represents accuracy values. The training accuracy curve reveals an initial rise, followed by stabilization, suggesting a convergence of the model.

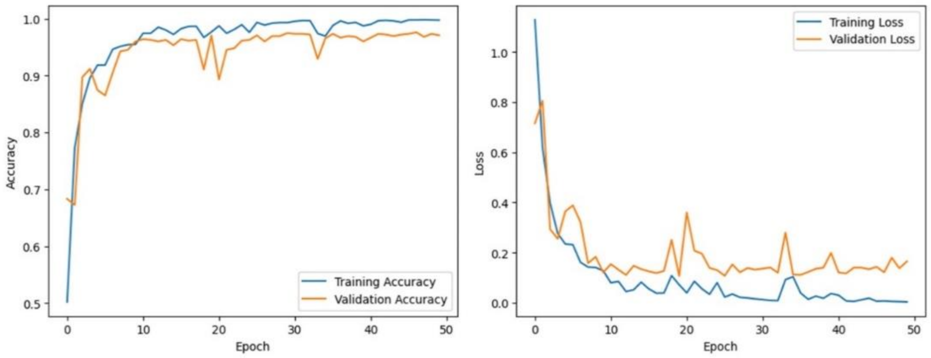

Fig. 4. Adam optimizers accuracy and loss.

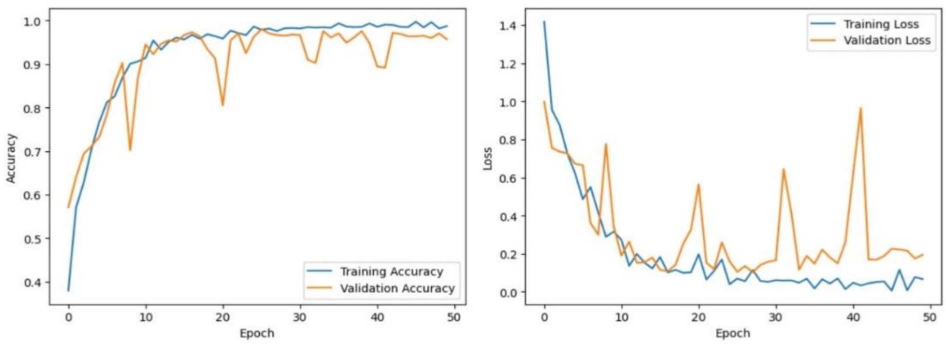

Fig. 5. RMSprop optimizers accuracy and loss.

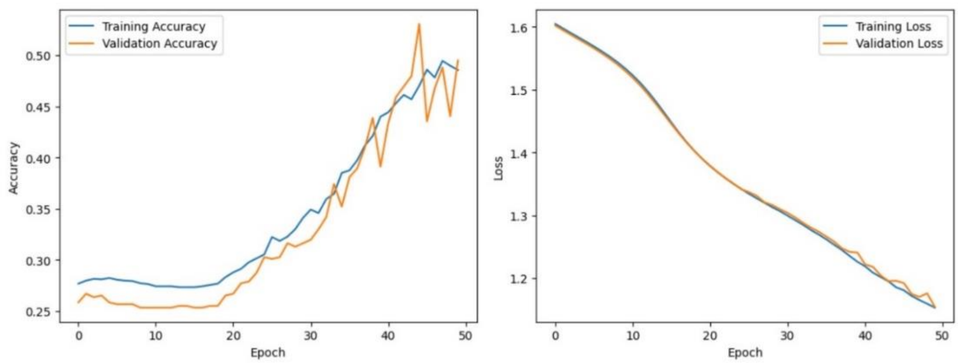

Fig. 6. SGD Optimizers accuracy and loss.

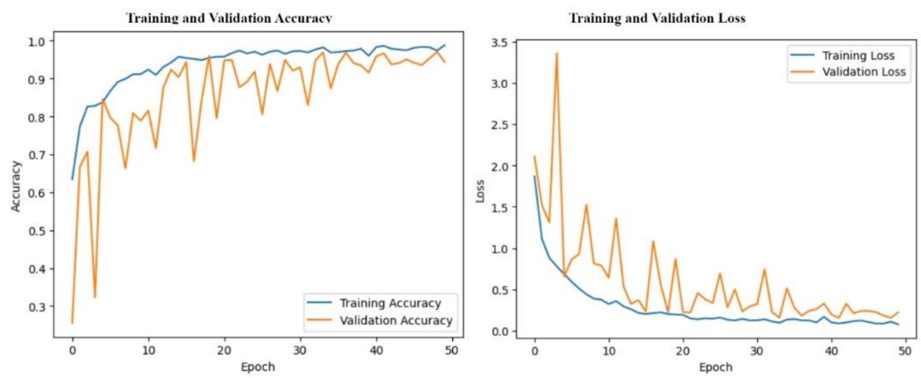

Fig. 7. Accuracy and loss after hyperparameter tuning.

Notably, fluctuations or plateaus in accuracy may be observed, potentially influenced by factors such as learning rate and batch size. Concurrently, the validation accuracy curve provides insights into the model's generalization to unseen data. The examination focuses on the correlation between training and validation accuracy, aiming to detect potential indicators of overfitting or underfitting. Key observations include notable increases or decreases in accuracy and the identification of the epoch with the highest validation accuracy. The findings gleaned from this analysis offer valuable insights into the model's behavior, informing discussions on its overall performance [23] and providing a foundation for future research and improvements. Supplementary visuals and tables are included to enhance the comprehensive understanding of the model's intricacies.

C. GLDICCNN Classification Report

Table X shows the classification report and provides a detailed summary of the performance of a GLDIC-CNN classification model on each class in the groundnut crop dataset. The following details the key metrics of the classification report. Precision values for classes 0, 1, 2, 3, and 4 are 0.99, 0.95, 0.86, 0.99, and 0.99, respectively, showcasing high accuracy in positive predictions for each class. Recall values for classes 0, 1, 2, 3, and 4 are 0.96, 0.90, 0.98, 0.99, and 0.98, respectively, indicating the model's effectiveness in capturing actual positive instances, with values ranging from 90% to 99%. The F1–Scores for classes 0, 1, 2, 3, and 4 are 0.98, 0.92, 0.91, 0.99, and 0.99, respectively, highlighting a commendable balance between precision and recall for each class.

TABLE X. CLASSIFICATION REPORT OF GLDICCNN

	Precision (%)	Recall (%)	F1–Score (%)	Support
ELS	99	96	98	210
Healthy	95	90	92	205
LLS	86	98	91	124
Rosette	99	99	99	186
Rust	99	98	99	209
Accuracy			96	934
Macro avg.	96	97	96	934
Weighted avg.	97	96	96	934

Support values for classes 0, 1, 2, 3, and 4 are 210, 205, 124, 186, and 209 instances, respectively, providing insights into the distribution of occurrences across different classes in the dataset. The model's accuracy of 96% reflects the proportion of correctly predicted observations out of the total, offering a comprehensive measure of correctness. The precision, recall, F1–Score, support, and accuracy values collectively demonstrate the model's ability to make accurate and balanced predictions across various classes, providing a comprehensive assessment of its performance.

D. Classification Report of GLDICCNN with Hyperparameters

Table XI shows the classification report with hyperparameters.

TABLE XI. CLASSIFICATION REPORT OF GLDICCNN WITH HYPERPARAMETERS

	Precision (%)	Recall (%)	F1–Score (%)	Support
Early Leaf Spot	96	94	95	210
Healthy	96	87	91	205
Late Leaf Spot	78	99	88	124
Rosette	99	91	95	186
Rust	98	99	99	209
Accuracy			94	934
Macro avg.	93	94	94	934
Weighted avg.	95	94	94	934

These metrics collectively provide a comprehensive evaluation of the model's performance across the specified classes.

E. Confusion Matrix

Figs. 8 and 9 show that the confusion matrix provides a detailed analysis of GLDICPCNN model classification accuracy for five categories with and without hyperparameter tuning. Analyzing the confusion matrix reveals the model's class-specific performance. For Class 1, it had 202 True Positives, 8 False Positives, and 3 False Negatives. Class 2 showed 185 True Positives, 1 False Positive, and 17 False Negatives. In Class 3, 121 True Positives were identified with 3 False Positives and no False Negatives, highlighting a precision-recall trade-off.

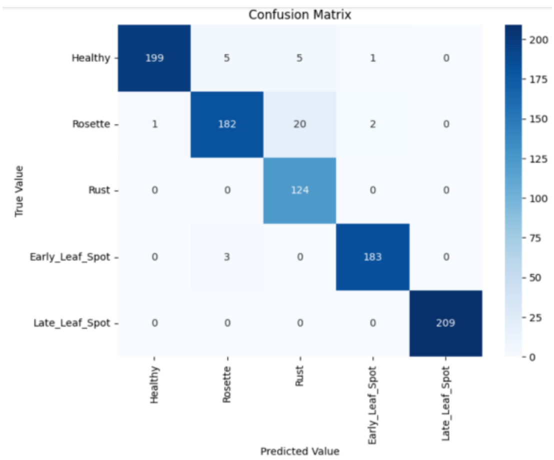

Fig. 8. Confusion matrix of GLDICCNN model.

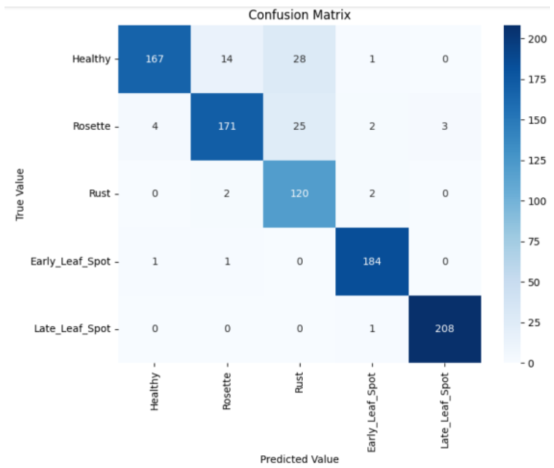

Fig. 9. Confusion matrix of GLDICCNN model with hyperparameter.

TABLE XII. COMPARISON OF PROPOSED GLDICCNN MODEL WITH RESNET MODEL

Sl. No.	Pretrained Network	Training Accuracy (%)	Training Loss (%)	Validation Accuracy (%)	Validation Loss (%)
1	GLDICCNN	99.73	00.35	97.06	16.49
2	ResNet50	98.99	02.46	97.88	11.04
3	ResNet101	99.26	02.44	97.77	08.11
4	ResNet152	98.52	03.68	97.01	17.25

For Class 4, the model achieved high precision (184 TP, 2 False Positives) and zero False Negatives. Class 5 exhibited exceptional performance with 209 TP, perfect precision, and recall. The detailed confusion matrix analysis not only calculates standard metrics but also offers nuanced insights for refining and optimizing class-specific predictions, crucial for overall model enhancement. This detailed analysis of the confusion matrix facilitates a granular evaluation of the model's precision, recall, and accuracy for each class, providing valuable insights for further refinement and optimization.

F. Comparison with Pretrained Models

The comparison of the proposed model GLDICCNN with pre-trained models [24], existing methods, and prior research in the domain of groundnut leaf identification, classification, and prediction. Table XII detailed the comparison of the proposed models, GLDICCNN emerges as a strong contender based on its exemplary performance metrics.

The model's endorsement is rooted in its robust training accuracy, successful generalization, and superior performance across key metrics, making it a compelling choice for diverse deep-learning applications.

VI. CONCLUSION

In conclusion, this research demonstrates the transformative potential of deep learning, specifically CNNs, in revolutionizing groundnut leaf disease identification and classification. The GLDICCNN model exhibits high accuracy, robust generalization, and computational efficiency, offering innovative solutions for disease management in agriculture. This work lays the foundation for future research and development, providing practical tools for disease diagnosis and monitoring to enhance the sustainability and productivity of agricultural practices.

The GLDICCNN architecture, with a training accuracy of 99.73%, loss of 0.0035, and validation accuracy of 97.06% with a loss of 0.1649, showcases meticulous design and adaptability. The comparison with the ResNet family architecture demonstrates the superior performance of the proposed model, emphasizing its value over pre-trained networks.

Looking forward, future work involves expanding the dataset by collecting more diseased images across the Pudukkottai district. This will make it easier to enhance the accuracy of disease identification, providing farmers with accurate diagnostic tools. Implementing recommendations derived from these findings aims to effectively mitigate diseases, ultimately enhancing groundnut yield and contributing to agricultural sustainability.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Anna Anbumozhi conducted the research work, collected the data, and wrote the paper. A. Shanthini supervised the work and approved the final version; all authors had approved the final version.

ACKNOWLEDGMENT

The author would like to acknowledge the supervision of the guide A. Shanthini, for her full guidance to do the research work. This manuscript is declared as original, not having been published previously, and is not presently under consideration for publication elsewhere.

REFERENCES

[1] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015. doi: 10.1038/nature14539

[2] M. Wakchaure, B. K. Patel, and A. K. Mahindrakar, "Application of AI techniques and robotics in agriculture: A review," *Artif. Intell. Life Sci*., vol. 3, no. November 2022, 100057, 2023. doi: 10.1016/j.ailsci.2023.100057

[3] A. Pal and V. Kumar, "AgriDet: Plant leaf disease severity classification using agriculture detection framework," *Eng. Appl. Artif. Intell.*, vol. 119, no. May 2022, 105754, 2023. doi: 10.1016/j.engappai.2022.105754

[4] A. Arshaghi, M. Ashourian, and L. Ghabeli, "Potato diseases detection and classification using deep learning methods," *Multimed. Tools Appl*., vol. 82, no. 4, pp. 5725–5742, 2023. doi: 10.1007/s11042-022-13390-1

[5] P. Kaur, S. Harnal, V. Gautam, M. P. Singh, and S. P. Singh, "A novel transfer deep learning method for detection and classification of plant leaf disease," *J. Ambient Intell. Humaniz. Comput*., 2022. doi: 10.1007/s12652-022-04331-9

[6] S. M. Hassan, A. K. Maji, M. Jasiński, Z. Leonowicz, and E. Jasińska, "Identification of plant-leaf diseases using CNN and transfer-learning approach," *Electron*., vol. 10, no. 12, 2021,

[7] A. Ahmad, D. Saraswat, V. Aggarwal, A. Etienne, and B. Hancock, "Performance of deep learning models for classifying and detecting common weeds in corn and soybean production systems," *Comput. Electron. Agric.*, vol. 184, 106081, 2021. doi: 10.1016/j.compag.2021.106081

[8] C. Sarkar, D. Gupta, U. Gupta, and B. B. Hazarika, "Leaf disease detection using machine learning and deep learning: Review and challenges," *Appl. Soft Comput.*, vol. 145, 110534, 2023. doi: 10.1016/j.asoc.2023.110534

[9] R. Chen, H. Qi, Y. Liang, and M. Yang, "Identification of plant leaf diseases by deep learning based on channel attention and channel pruning," *Front. Plant Sci*., vol. 13, pp. 1–15, 2022. doi: 10.3389/fpls.2022.1023515

[10] S. K. Kumar, "Neural network based plant leaf," *International Journal of Advanced Computer Science*, vol. 9, no. 4, pp. 1349–1358, 2021.

[11] K. Thaiyalnayaki and C. Joseph, "Classification of plant disease using SVM and deep learning," *Mater. Today Proc*., vol. 47, pp. 468–470, 2021. doi: 10.1016/j.matpr.2021.05.029

[12] M. H. Saleem, J. Potgieter, and K. M. Arif, "Plant disease classification: A comparative evaluation of convolutional neural networks and deep learning optimizers," *Plants*, vol. 9, no. 10, pp. 1–17, 2020. doi: 10.3390/plants9101319

[13] S. M. Hassan and A. K. Maji, "Plant disease identification using a novel convolutional neural network," *IEEE Access*, vol. 10, pp. 5390–5401, 2022. doi: 10.1109/ACCESS.2022.3141371

[14] L. Xu, B. Cao, S. Ning, W. Zhang, and F. Zhao, "Peanut leaf disease identification with deep learning algorithms," *Mol. Breed.*, vol. 43, no. 4, pp. 1–12, 2023. doi: 10.1007/s11032-023-01370-8

[15] U. B. Patayon and R. V. Crisostomo, "Peanut leaf spot disease identification using pre-trained deep convolutional neural network," *Int. J. Electr. Comput. Eng*., vol. 12, no. 3, pp. 3005–3012, 2022. doi: 10.11591/ijece.v12i3.pp3005-3012

[16] A. M. P. and A. P. Reddy, "Dataset of groundnut plant leaf images for classification and detection," *Data Br*., vol. 48, 2023. doi: 10.1016/j.dib.2023.109185

[17] A. Anbumozhi and A. Shanthini, "Leaf diseases identification and classification of self-collected dataset on groundnut crop using progressive convolutional neural network (PGCNN)," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 2, pp. 364–373, 2023. doi: 10.14569/IJACSA.2023.0140244

[18] S. Muthukumaran, P. Geetha, E. Ramaraj, and C. Head, "Leaf spot disease image classification for groundnut crop using deep convolutional neural network," *Turkish J. Comput. Math. Educ.*, vol. 12, no. 14, pp. 3166–3179, 2021.

[19] N. Suresh and D. A. Giridharan, "Predicting groundnut disease using CNN models," *J. Univ. Shanghai Sci. Technol*., vol. 23, no. 06, pp. 756–766, 2021. doi: 10.51201/jusst/21/05335

[20] G. Al, M. Anbar, S. Manickam, T. A. Al-amiedy, and A. Al, "Computer and a systematic literature review of recent lightweight detection approaches leveraging machine and deep learning mechanisms in Internet of Things networks," *J. King Saud Univ. - Comput. Inf. Sci*., vol. 36, no. 1, 101866, 2024. doi: 10.1016/j.jksuci.2023.101866

[21] K. S. Devi, P. Srinivasan, and S. Bandhopadhyay, "H2K—A robust and optimum approach for detection and classification of groundnut leaf diseases," *Comput. Electron. Agric*., vol. 178, no. October, 105749, 2020. doi: 10.1016/j.compag.2020.105749

[22] N. Vasudevan and T. Karthick, "A hybrid approach for plant disease detection using E-GAN and CapsNet," *Comput. Syst. Sci. Eng*., vol. 46, no. 1, pp. 337–356, 2023. doi: 10.32604/csse.2023.034242

[23] B. M. Joshi and H. Bhavsar, "Deep learning technology based night-CNN for nightshade crop leaf disease detection," *Orig. Res. Pap. Int. J. Intell. Syst. Appl. Eng*., vol. 2023, no. 1, pp. 215–227, 2023.

[24] Y. Gulzar, "Fruit image classification model based on mobilenetv2 with deep transfer learning technique," *Sustain.*, vol. 15, no. 3, 2023. doi: 10.3390/su15031906