CINEFORUM
ISSN: 0009-7039

Vol. 66. No. 2, 2026

A Deep Learning Based RefineNet Model for Chilli Plant Disease Classification

SRINIVASULU MARAMREDDY[1], SANDIPAN MAITI[2]
[1]Research Scholar, VIT-AP University, School of Computer Science and Engineering, India.
[2]Associate Professor, VIT-AP University, School of Computer Science and Engineering, India.
E-mail: [1]srinivasulu.20phd7154@vitap.ac.in, [2]sandipan.maiti@vitap.ac.in

Abstract

Diseases of the chili leaf are a challenge in chili crop production and quality, hence the reduced quality and quantity. Early detection and control of diseases in the chili leaf are important to ensure the quality and quantity of the produce. However, early detection of diseases in the natural image of the chili leaf is challenging, especially when the leaves have irregularly shaped yellow spots and curls. Current models problem, this paper presents the use of the RefineNet model, which is specifically tailored to ensure effective early detection of diseases in the chili leaves. The proposed method involves the use of the RefineNet model, whose composition includes four blocks: the encoder/decoder block, the residual convolutional block (RCU), the multi-scale fusion block (MRF), and the chained resolution pooling (CRP) block. While the encoder block produces low-resolution maps, the other block produces high-resolution maps. However, the use of residual connections in the RCU block helps reduce vanishing problems, hence facilitating information transfer from different levels. In the MRF block, the feature map is processed at different scale levels. Dilated convolution is used in this block to enhance the receptive field, hence facilitating effective extraction of contextual information. Additionally, the leaky ReLU activation function replaces the standard ReLU activation function, enhancing the feature extraction ability for diseased leaves with irregularly shaped yellow spots and curls. Moreover, the auxiliary classifier is simplified, reducing the model's complexity without compromising the recognition rate. Experimental results demonstrate that compared to conventional models and advanced multi-scale models, RefineNet exhibits superior classification performance for chili leaf diseases, achieving an average accuracy of 99.78%. This evaluation confirms that RefineNet is robust, stable, and highly accurate in practical applications for detecting chili leaf diseases. Furthermore, the evaluation reveals that RefineNet demonstrates a good recall rate, F1 score, and confusion matrix.

Keywords: Deep-Learning, CNN, Encoder-Decoder, Residual Convolution Block, Multi-Scale Fusion Block, And Chained Resolution Pooling.

1. Introduction

India has the second largest arable land in the world, including all 15 major climates in the world, partitioned into 20 agri-climatic regions Sharma et al. [1] Plant infections exist, as their impact is very important, as they tend to cause detrimental effects on the quality as well as the quantity of chilies raised in the agricultural sector. A large number of organisms like Fungi, bacteria, viruses, molds, and many others tend to cause plant infections, ultimately hampering plant production. Then there comes the declining percentage of crop production in India, which has become a major point of concern, as the effect of quality as well as the reduced quantity of chilies produced in the Indian agricultural sector has caused a tremendous impact on the living standards of the Indian Farmers, as 60% of the total Indian population suffers because of loans. It is well recognized that crop loss is primarily marked by the presence of a few exceptional factors, with diseased crops being a strong contributor to these losses, which keep varying from season to season. Today, the scenario has created various hurdles to crop development, which, in turn, provides an increased need to boost the crop development processes through the adoption of new-age technology. In this scenario, the name of the "state of Andhra Pradesh" emerges as a leading crop producer of "Rice," "Black Gram," "Green Gram," and "Chilis," with about "70% of the population" of the state being associated with agriculture-related work. It can be noted that the production of the "CHILI" crop in the respective state has been recorded to be very high; however, various "Leaf Diseases" have been causing hurdles to the development of the CHILIs. The imperatively in the current context revolves around the early diagnosis of various leaf diseases of the Chili plant, including Damping off, Die-back, Anthracnose, wet rot, Gemini virus, and Mosaic diseases. Various researchers have come up with the solution to the problem of identifying the diseases present in the green leaf of the plant in the initial stage using the deep learning approach in various datasets of crops, as seen in the study carried out by the works of Sujatha et al. [2], Vallabhajosyula et al. [3], Tiwari et al. [4] and Kanaparthi KR et al. [5]. There has been limited focus on the solution to the problem of the training accuracy of the models associated with the choice of the optimizer and the number of epochs. The choice of the pre-trained models itself carries a never-ending challenge while working with a limited number of classes. The most prominent pre-trained models, viz. Squeeznet and Google net, have shown higher training accuracy with limited as well as less complex classes. A study carried out by Naik et al. [6], validated the training accuracy in a Chili plant with the context of 5 classes of leaf diseases using a total of 12 pre-trained models like Alex Net, DarkNet53, DenseNet201, EfficientNetb0, InceptionV3, MobileNetV2, ResNet101, Shuffle Net, Squeeze Net, VGG19, and Xception Net Naik et al. [6], with an observed training accuracy of 83.54%-99.28%, with and without the effect of augmentation.

2. Literature Review

Chili plants are susceptible to a variety of diseases that can significantly impact yield and quality. The advent of image processing and machine learning techniques has enabled the development of automated systems for detecting and classifying these diseases. This

CINEFORUM
ISSN: 0009-7039
Vol. 66. No. 2, 2026

survey reviews recent research efforts in the field of chili plant disease classification. At the initial stage of image processing, techniques include feature extraction, image segmentation, thresholding, and morphological operations. Traditionally, color, texture, and shape feature extraction methods have been employed in image processing. For instance, Ahmad et al. employed the color and texture features to identify the various leaf diseases of chilies by isolating the affected areas in the image using image segmentation Ahmad et al. [7]. Additionally, Wahab et al. (2019) employed the k-means clustering method to separate the images into segments and extract features to detect the various diseases in chilies Wahab et al. [8]. Preprocessing techniques like thresholding and morphological operations are quite often employed while handling images. A technique involving adaptive thresholding and morphological operations has been proposed by Kurmi et al. [9] for segregating diseased regions of leaves Kurmi and Gangwar [9]. K-Nearest Neighbors (KNN) algorithms that are part of Support Vector Machines (SVM) are well known and preferred machine learning algorithms for disease classification in the Chilli plant. SVM algorithms have been preferred by many researchers for disease classification in plants. SVM classifiers that processed the extracted characteristics of chili plant images provided better accuracy in disease classification Patil and Lad [10].

KNN classifiers have also been used by many researchers. Hossain et al. in 2019 showed how KNN classifiers performed well in disease classification in chili plants through shape and texture characteristics Hossain et al. [11]. Deep learning and its uses in CNNs have contributed greatly towards an advancement in the area. Work on the development of the CNN model that automatically derives the features from the images of the chili leaves, with higher accuracy than other approaches Sajitha et al. [12]. Convolutional Neural Networks (CNNs) have continued to be the backbone for the classification of plant diseases. A new advanced model using the attention mechanism for the detection of plant leaf diseases was created by the researcher Yang et al. [13]. Transformer models, which found their initial application in natural language processing, have also found their way in image classification. Transfer learning, where pre-trained models on bigger data sets are used to learn other tasks, has found its application. This concept of transfer learning has been applied by the author Ferentinos [14] to classify plant diseases, including chilli plant diseases, using pre-trained models such as VGG 16 and Inception V3. The author Tolstikhin et al. [15] presented new Vision Transformers, also known as ViTs, which performed better than conventional CNNs in plant disease detection, applying self-attention mechanisms to model the long-distance dependencies in the image. Another type of hybrid approach that combines CNNs and Recurrent Neural Network (RNN) has been investigated for the implementation of temporal data incorporation. A combination of conventional image processing techniques and deep learning approaches has also been investigated. A hybrid approach was introduced by the author Malik et al. [16], which includes the implementation of image processing techniques for initial feature extraction, followed by a CNN classifier, resulting in the improvement of the approach's efficiency. A CNN-RNN hybrid approach was introduced by the author Liu et al. [17] that relies on the use of a sequential image for the progression of a disease. Edge computing devices can now be used for real-time plant disease detection. The author Thai et al. [18] proposed a lightweight CNN model suitable for edge devices for real-time plant disease detection. The integration of plant disease detection models in mobile applications has made plant disease diagnosis feasible for farmers. The author Sharma et al. [19] proposed a mobile application based on a pre-trained MobileNetV2 model for efficient plant disease identification in real time. The application of multispectral imaging has improved the efficiency of disease identification, which could not be detected by the naked eye in the RGB channel. A researcher Zhang et al. [20] has created a multispectral imaging technique coupled with a deep learning algorithm to identify different plant diseases. Hyperspectral imaging, involving data acquisition over a broad spectrum of wavelengths, has also been adopted for the in-depth study of a disease. The author Yang et al. [21] has employed data based on hyperspectra, coupled with 3D CNN, to classify plant diseases with outstanding success and greater accuracy by extracting spectral-spatial information.

Based on the improvements in machine learning, specifically deep learning, there have been improvements in the classification of plant diseases. Improvements in CNN architectures, transformer networks, or hybrid networks have opened new paths in terms of accuracy and efficiency. The application of transfer learning, data augmentation, or generating synthetic data has resolved difficulties caused by data scarcity. Real-time detection by edge computing or mobile apps has practical contributions in this area, while multispectral or hyperspectral imaging technologies have widened the methods of disease detection systems. In another research work carried out by Karthik et al. [22], the use of the InceptionV3 model was considered in the aspect of detecting diseases that affect coffee leaves. The authors compiled a dataset comprising 1,747 images of coffee leaves, which were categorized into five separate groups, consisting of healthy leaves and four different kinds of leaf diseases. Transfer learning was used to fast track the training process, and image augmentation methods were also incorporated to tackle the problem of insufficient image data available in the dataset. Additionally, another author Karim et al. [23] developed a dataset comprising 107,366 images of grape leaves utilizing image enhancement and image augmentation methods. They proposed a new Dense Inception-based Convolutional Neural Network (DICNN) focused on the classification of images into seven unique groups. In addition, author Ramcharan et al. [24] gathered a set of information consisting of pictures of diseases of cassava, with a total of 11,670 pictures for three types of diseases and two types of damage caused by pests. By applying transfer learning, the authors were able to develop an InceptionV3 model and test its accuracy by applying three separate classifiers: the softmax classifier, support vector machines (SVM), and k-nearest neighbor (kNN) algorithms. This thorough assessment contributes significantly to an understanding of the applicability of deep learning methods, specifically CNNs, to a technological transformation of disease classification and identification as concerns precision agriculture. To improve image representation for recognition, a dataset of 15000 images of tomato diseases and pests taken in natural environments, including 12 classes of diseases and pests, was created by the authors of Xu and Wu [25]. They modified the feature layer of the You Only Look Once version 3 (YoloV3) model by using an image pyramid approach to detect features of various scales, thereby improving detection accuracy. In research by Saleem et al. [26], Faster Region-based Convolutional Neural Network (Faster RCNN), Region-based Fully Convolutional Network (RFCN), and Single Shot Multibox Detector (SSD) detectors were used to test and determine the efficiency of Deep Learning for disease and pest detection on a tomato crop. To improve detection accuracy and localization of bounding boxes, the researcher combined each detector with deep feature extractors, including VGG and Residual Network (ResNet). They then tested the efficiency of their approach on a dataset consisting of 5000 images of nine classes of diseases and pests of tomatoes. Another research effort by Alirezazadeh et al. [27] presented a deep learning approach based on AlexNet for monitoring diseases of Apple leaves using a dataset of 13689 images showing diseases of Apple, including four classes of diseases. Besides the above studies related to species-specific research, some researchers have ventured into the classification of various diseases across various species of plants. Ferentinos [14] investigated the classification of 58 various plant diseases from 25 species using various CNN models, such as AlexNet, VGG, GoogleNet. Azour [28] carried out an experiment to evaluate the performance of VGG16, InceptionV4, ResNet50-152, and DenseNet121 on the classification of 38 varied diseased crops, which included healthy images as well. There have been tremendous developments in the field of Deep Learning; however, it has also brought about various challenges.

Deep Learning techniques are characterized by the need for extensive amounts of data to train the deep models. This could pose a challenge for researchers working on deep learning due to the lack of annotated images for training, specifically for tasks associated with the classification of diseased plants. In some scenarios, the need for annotated images has been overcome by the modification of existing deep models. Contrary to general beliefs, some researchers have indicated that transfer learning doesn't necessarily produce better results, which could be due to the ineffectiveness of the preservation of the current features across tasks. In addition, there exist certain methods with the objective of improving the performance of deep models for plant disease classification based on complicated scenarios and complex deep models for effective plant disease classification based on deep learning models. However, when working with limited data, a complicated deep learning model with a large number of parameters can potentially cause overfitting. Moreover, although most deep learning models have been able to successfully identify relevant disease attributes, there exist possibilities of capturing irrelevant attributes of an image, for instance, background noise or healthy parts of plants, resulting in a potential conflict between plants belonging to different disease classes. However, to address such a problem, Khanna et al. [29] introduced a region-based deep convolutional neural network based on contaminated regions of leaves, an intensive task requiring human-level annotation of disease regions and a deep understanding of plant disease classification based on deep learning models. A novel approach that was developed by integrating the GoogleNet and Recurrent Neural Network (RNN) for the intention of automatically detecting the areas of infection and extracting relevant features concerning the classification of disease images, including a set of 20 diseases and a single healthy category Salai and Ramapuram [30]. Despite their endeavors, it was noted that their performance was notably more effective when employing GoogleNet as opposed to employing a fusion of GoogleNet with RNN within the PlantVillage dataset.

The utilization of excessively large deep neural network models has been observed to engender the generation of an abundance of superfluous features that essentially replicate one another's essence, either in terms of being shifted versions or bearing significant resemblance, consequently leading to a decline in the overall efficiency of the system. From the advancements in machine learning, particularly deep learning, have driven significant improvements in plant disease classification. Enhanced CNN architectures, transformer models, and hybrid approaches have pushed the boundaries of accuracy and efficiency. Transfer learning, data augmentation, and synthetic data generation have addressed data scarcity challenges. Real-time detection through edge computing and mobile applications has brought practical solutions to the field, while multispectral and hyperspectral imaging have expanded the capabilities of disease detection systems. The motivation of the proposed model to overcome the above stated limitations of the diverse deep learning models i.e., reliance on substantial amounts of labeled data, poor performance in small sample learning, weak model generalization capabilities, model computational complexity, and Aliasing effect. To overcome the limitations, we propose a deep learning based RefineNet model for plant disease classification. The proposed RefineNet-based deep learning model contribution to plant disease classification lies in its ability to accurately segment and identify diseased regions in plant images, which is crucial for precise diagnosis and treatment.

3. Dataset

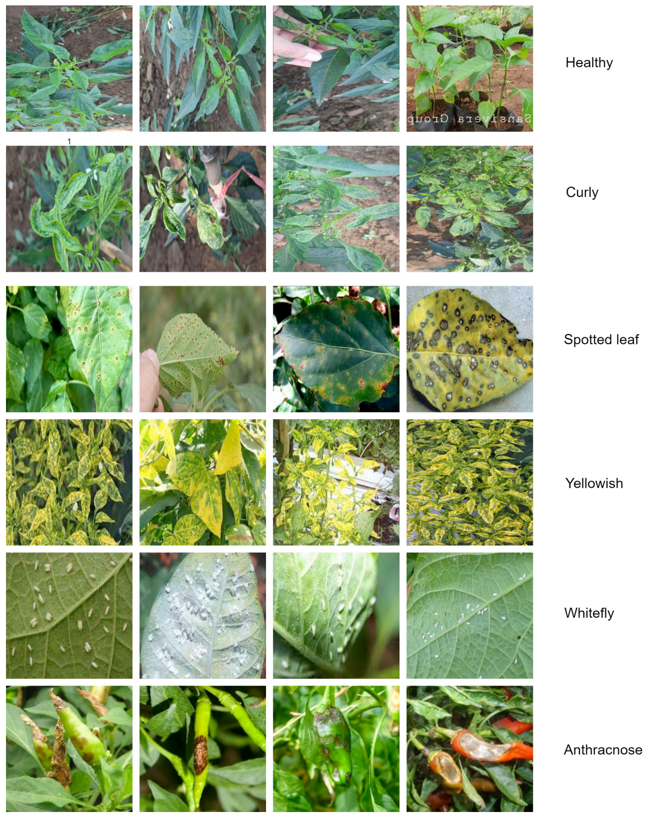
This section explains the importance of the data set used in deep learning approaches, followed by the extraction of the data set considered in this paper. The data set is an effective tool for effective analysis of any information. The data set in the context of deep learning indicates the data collection from which the model can be trained. The data set that is gathered completely can be classified as a training set and testing set. The data from the training set can be used for training and validation purposes.

CINEFORUM
ISSN: 0009-7039

Vol. 66. No. 2, 2026

Figure 1: Different Classes of Chili Plant Disease

The data set termed “Chilli disease detection” is used for the proposed approach and is considered from TugasAkhir [31]; Ullah et al. [32]. The Roboflow Chilli Plant Disease



Dataset contains high-resolution images of chilli plants affected by various diseases plant samples. It includes annotated images with bounding boxes or segmentation masks, supporting formats like COCO JSON and Pascal VOC XML. Collected from diverse environments, the dataset aids in training machine learning models for automated disease detection and classification. Accessible via Roboflow’s platform, it provides tools for visualization, annotation, and dataset splitting. Ideal for applications in precision agriculture and early disease monitoring, its usage is governed by specific licensing terms. It has around 5644 images and divided into three parts (i.e. 70% for training and 20% for validation and the remaining for testing). This data set comprises six versions of data, out of which latest three versions are combined to train the model. The versions that are used for analysis are V4, V5 and V6.

The modified dataset consist of six image classes related to Chilli leaf disease:

- Spotted,
- Yellowish,
- Healthy Leaf,
- Curly Leaf,
- Anthracnose
- Whitefly.

Figure. 1 shows the different classes of sample Chilli plant disease images.
Table 1 provides the total number of images used for training, testing, and validation.

Table 1: Chilli Plant Diseases

Chilli Plant Diseases	Training	Testing	Validation
Spotted	680	265	65
Yellowish	700	235	60
Healthy Leaf	580	242	64
Curly Leaf	620	258	67
Anthracnose	570	248	65
Whitefly	630	232	63

4. Description of the Proposed Model

The block diagram of proposed RNDDNet model is shown in Image 3. The proposed model is a combination of the seven different blocks i.e., Image processing, Augmentation, RNDDNet, Dense, Global average pooling, Softmax and Output. In image processing we gathered the Chilli plant disease images in RGB colour format and then divided into six different classes (i.e., Healthy, Curly, Spotted Leaf, Yellowish, Anthracnose and Whitefly).

4.1 Preprocessing

A preprocessing operation improves the image quality that is needed for classification. During preprocessing, images are rotated, scaled, and translated to make geometric adjustments. Here, we reduced all images with various pixel resolutions in dataset are converted into 100×100 pixels. The size and resolution of all photos must be the same.

4.2 Image Data Augmentation

Training data must be large enough to produce better results. The augmentation is required to improve the performance of deep CNN models that require minimal training data. The image augmentation process expands the dataset by including distorted images and decreases over fitting by including a few distorted images. A network overfits if it learns from the dataset itself rather than the patterns contained within. Image augmentation creates training images by using a flipping (rotating a image), blurring a picture, relighting a picture, and randomly cropping it. We manipulated the images by scaling, shearing, zooming, and flipping them horizontally. Augmentation techniques are commonly utilized on a substantial portion of the training images with the aim of guaranteeing a dataset that is diverse in nature. Nevertheless, the specific ratio employed may differ based on the particular implementation and the dataset that is being utilized for the process. It is common practice for a random subset of images to be chosen for augmentation, with the range usually falling between 50% of the total training images. This method is crucial as it ensures that the neural network encounters augmented versions of images on a regular basis throughout the training phase, thereby aiding in the acquisition of robust features. The application of such techniques contributes to the enhancement of the network's ability to generalize well to unseen data, ultimately improving its performance in various tasks. The utilization of augmented images aids in the development of a more adaptive and resilient network that is capable of handling variations and noise present in real-world data.

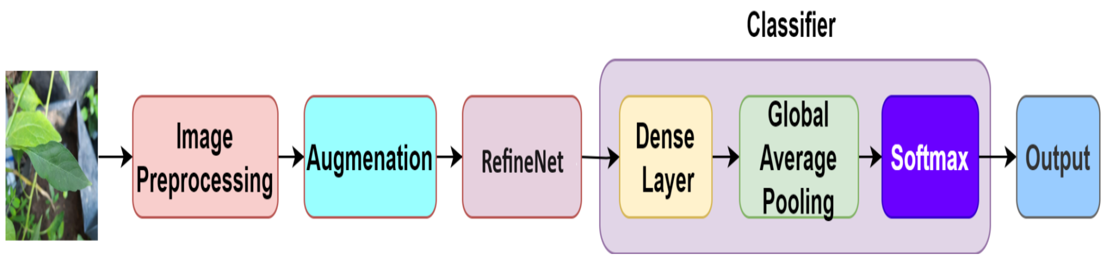


Figure 2: Block diagram of the RefineNet model

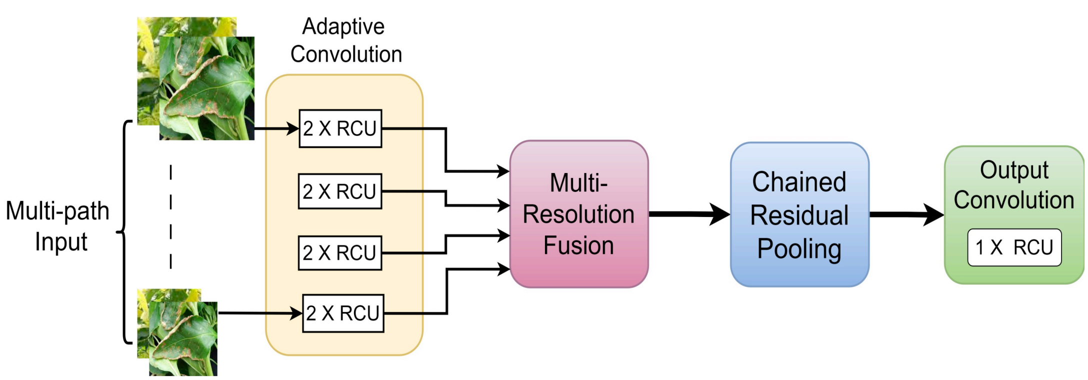
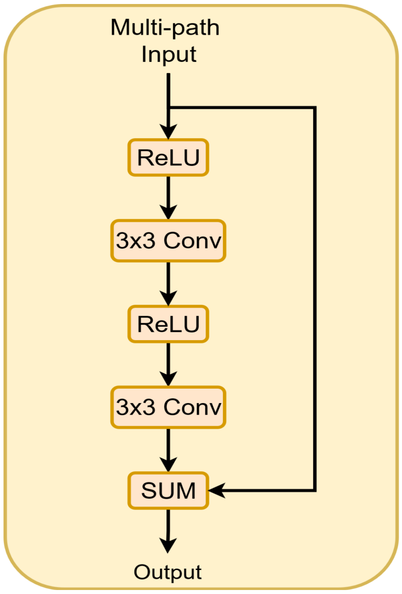


Figure 3: Architecture of the RefineNet model



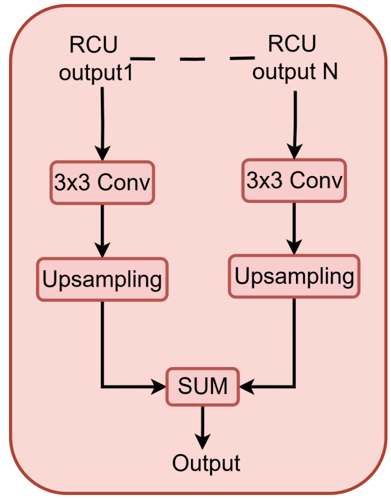


Figure 4: Residual convolutional unit block **Figure 5: Architecture of Multi-resolution fusion block**

4.3 RefineNet

The RefineNet architecture is a highly efficient and effective approach to the problem of semantic segmentation. Currently, there are no other methods that are as efficient as the RefineNet architecture in this application of semantic segmentation. It can be noted that there are four major sub-components to the RefineNet architecture: the encoder-decoder, the residual convolutional block (RCU), the multi-scale resolution fusion block (MRF), and the Chained resolution pooling (CRP). The encoder part of the architecture can be used to calculate the feature maps for a low resolution image, while the decoder part of the architecture can be used to calculate the feature maps for a high resolution image.

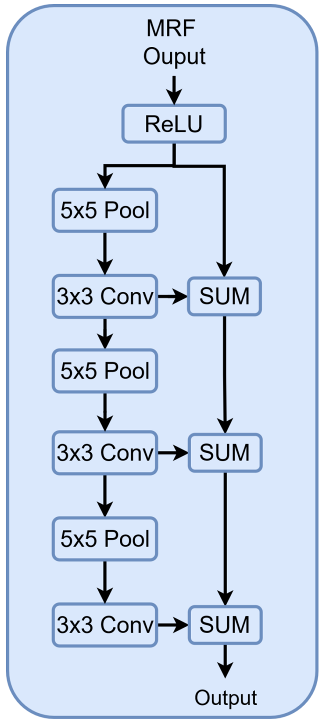


Figure 6:Architecture of Chained residual pooling block.

.

The RCU part of the architecture makes use of a residual connection to ensure that the vanishing gradient problem does not occur, thus ensuring that information can be passed across levels with ease. The MRF part of the architecture works on multiple scales of the feature map. It makes use of dilated convolution to increase the size of the receptive field for a larger contextual understanding of the scene.

4.4 Adaptive Convolution

Adaptive Convolution is an innovative approach of the neural networks domain that performs the processing and learning tasks of the convolutional layer according to the down-sampled information, through the use of Residual Convolutional Units (RCUs). The method is most beneficial in object recognition tasks where the input information may vary in resolution and dimensions, thus requiring the object recognition system itself to adjust and optimize the designed convolutional filters and maps according to the input dimensions. An important aspect of the proposed method is its ability to offer multiresolution fusion and chained resolution pooling within the RefineNet block that encourages the neural networks system itself to understand diverse levels of resolution in the input information and enhance the output classification through a stepwise process of fusion and pooling operations. Through the use of Adaptive Convolution, there are tremendous opportunities of improved accuracy, increased processing speeds, and optimal usage of resources that serve as an imperative approach in solving complex machine learning problems. The Residual Convolutional Unit block is shown in Image 5. The use of RCUs in the adaptive convolution makes it easier to implement the Multi-Resolution Fusion and the Chained Resolution Pooling techniques of the RefineNet block. The use of the Multi-Resolution Fusion technique enables the model to combine information of different scales of the image and hence enables it to extract information from the finer and coarser details of the image proficiently. On the other hand, the use of the Chained Resolution Pooling technique helps in the retention of the high-resolution feature space by chain pooling the feature space. Both techniques are able to improve the overall performance of the adaptive convolutional neural networks by using the RCUs. The use of the RCUs further improves the overall gradient flow and the speed of the training of the networks due to their nature of using the connections of the residuals. Thus, it prevents the problem of the vanishing gradient. Overall, the use of RCUs in the adaptive convolution makes it easier to implement the highly effective techniques and enhances the ability of the model to process information proficiently.

4.5 Multi-Resolution Fusion Within Refinenet Block

In a RefineNet block, Multi-Resolution Fusion is applied to fuse features at different resolution levels. This approach is implemented by stacking the outputs of layers with different receptive resolution levels through a process of channelwise fusion. As features from both layers are stacked, spatial resolution of a resulting feature map remains unchanged but its depth gets increased. This enables a new feature map that captures high semantic information as well as rich spatial information to be generated. This further aids in boosting the acquisition of spatial and temporal information within a network by applying Chained Resolution Pooling within its architecture. This further helps by suppressing spatial resolution at different levels through a series of down-sampling processes within a network. Therefore, by applying Multi-Resolution Fusion and Chained Resolution Pooling within a RefineNet block, a network is enabled to acquire information at a different resolution and scale levels, resulting in better performances in image segmentation and classification. The Multi-Resolution Fusion block is shown in Image 6.

4.6 Chained Resolution Pooling In RefineNet

This step-wise process of resolution integration, where lower resolution attributes are enhanced through their higher resolution counterparts, ensures adaptability as well as efficacy for the RefineNet block. Illustrated in Image 7, below, is the Chained Residual Pooling block. Through completing multiresolution fusion, along with chained resolution pooling, it becomes possible for the RefineNet block to arrive at a deeper level of complexity while simultaneously reducing network structural complexity. As opposed to independent resolution processing, chained resolution pooling helps arrive at features that have a better level of robustness for spatial variations, as a consequence of which it significantly alleviates the possibility of overfitting. Additionally, these RefineNet blocks promote spatial consistency, where features pertaining to input resolution are combined with up-sample layer output, ensuring a focused attention level on spatial attributes of importance. Adopting a chained resolution pooling approach designates a faster convergence process for the RefineNet block, pertaining to leading-edge deep learning frameworks, while requiring smaller structural parameters. Thereby, the RefineNet block represents an efficient, adaptable design for a series of deep learning applications. Chained Resolution Pooling is a complex function used in the RefineNet block to effectively combine features of multiple resolution levels. This process involves creating a sequence of pooling layers, where each pooling layer is provided input from a different resolution level. The result produced in each pooling layer is then passed to the next pooling layer, which eventually results in a continuous process of pooling. sub-classifier = Model composed of Dense, GlobalAveragePooling, and SoftMax activation. The dense layer takes RefineNet model output and organizes various classes of channels. The result of a global average pooling layer is aggregate values of channel, and finally, SoftMax activation function determines classified output for classifying plant leaf diseases at last stage where calculation metrics are carried out.

5. Flowchart of Proposed Method

The detailed flow chart of the proposed classification model based on RefineNet is given in Image 8. The flow chart of the modified RefineNet layout for the requirement of classification begins with the image input, followed by the processing of the input image through the basic CNN structure to acquire the features at various levels, thereby incorporating all the differing levels of detail.

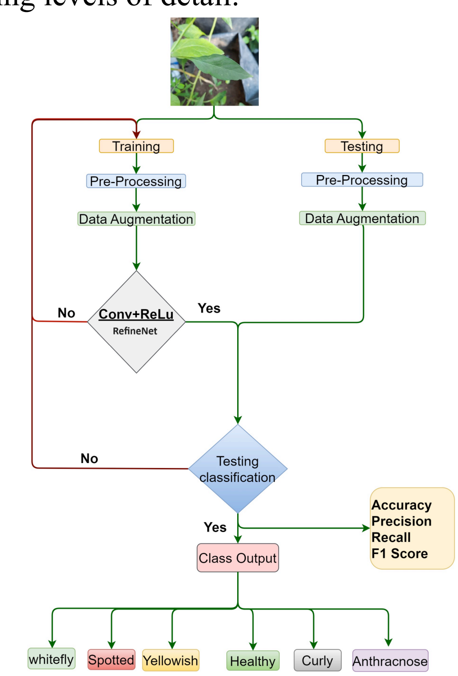


Figure 7:Flow diagram of Proposed model.

.

During the refinement stage, the features go through the process of multi-resolution fusion, where all the features are merged to incorporate all the details as well as the contextual information. This process is followed by the application of chained residual pooling to capture the long dependencies, thereby increasing the overall understanding of the global structure of the image. Residual convolutions are then employed to refine the features, ensuring that they are of a superior class. With the completion of the refinement process, the features are compressed into a single vector based on the global feature aggregation, thereby retaining all the comprehensive information present in the image. This single vector, including all the features, is then passed through fully connected layers, where the softmax filter is applied to generate the resultant probability distribution over all possible classes. This concludes the process of prediction, where the classification class corresponding to the maximum probability is chosen as the predicted class.

6. Experimental Analysis

6.1 Experimental Setup

The RefineNet architecture was developed utilizing a computational setup comprising 64 GB of RAM, an INTEL(R) Xeon(R) W-125 CPU with a processing frequency of 4 GHz, 2 TB of read-only memory, and a 16 GB NVIDIA GeForce GTX 1080 Ti graphics processing unit. The RefineNet received input images of dimensions 256x256x3. Throughout the training process, we observed several performance metrics including mean absolute error, accuracy, loss, precision, recall, and F1-score. An Adam optimizer was employed with a learning rate fixed at 0.02. The model underwent training for a total of 70 epochs, during which the optimal weights were preserved based on the observed training accuracy and loss. In Graph 1, the curves depicting accuracy and loss for the RefineNet model are illustrated. Both training and validation phases utilized a batch size of 32. A total of 75% of the dataset was randomly designated for training purposes, while 15% was allocated for validation, with the remainder reserved for testing. The model was preserved upon achieving a minimum loss threshold. Additionally, the validation loss functioned as a determinant for early stopping. If the validation loss did not exhibit a decrease over 5 successive iterations, the model with the most favorable weights was selected. This section elaborates on the results obtained from the data set that is considered for disease identification of chili leaf. The data set explained in the previous section is used, which consists of 10.9 million trainable parameters. These parameters are used in reducing the prediction time and memory size. The simulation results show that the loss is reduced and accuracy is high. Furthermore, a comparison is made with the CNN algorithm used in [19] to show the efficacy of the proposed method by considering the trainable parameters. The proposed classification model is compared with the following CNNbased baselines i.e., AlexNet, LeNet-5, SqueezeNet, VGG16, Inception V3, ResNet152V2, and DenseNet201.

To evaluate the proposed method with other baselines with the following metrics: Accuracy, Precision and Recall.

6.2 Loss and Accuracy

Here, the method used is RefineNet, as explained in the previous section. By using the proposed method in the identification of chili leaf disease, the simulation results indicate the loss and accuracy. The proposed RefineNet model loss and accuracy simulation results for the are depicted in Graph 1. Both the training loss and validation loss are settling to zero at the end of the last epoch. Further, it is observed that the loss values are gradually decreasing with the increase in the count of the data set. The sudden spike in the validation loss at one value of the data set indicates the combination of affected leaves. Similarly, Graph 1 shows the training accuracy and validation accuracy for the proposed model, the training and validation accuracy values are gradually increasing with the increase in the count of the data set. The mixed combination of affected leaves produces sudden spike in the validation accuracy at one value of the data set. Table 2 shows the accuracy and loss experimental results of proposed model with baselines. The proposed RefineNet outperforms the all baselines in terms of accuracy and loss. It achieved the highest accuracy in training and validation is 99.29% and 99.27% respectively. Similarly, the proposed model gets lowest loss in training and validation is 0.0946, 0.0911 respectively. To enhance the analysis, we scrutinize the results obtained from the all five models. However, our evaluation of the architecture relies on two key factors: validation accuracy and the confusion matrix. Validation accuracy gauges how faithfully the trained model aligns with the training data. Conversely, the confusion matrix breaks down the performance metrics.

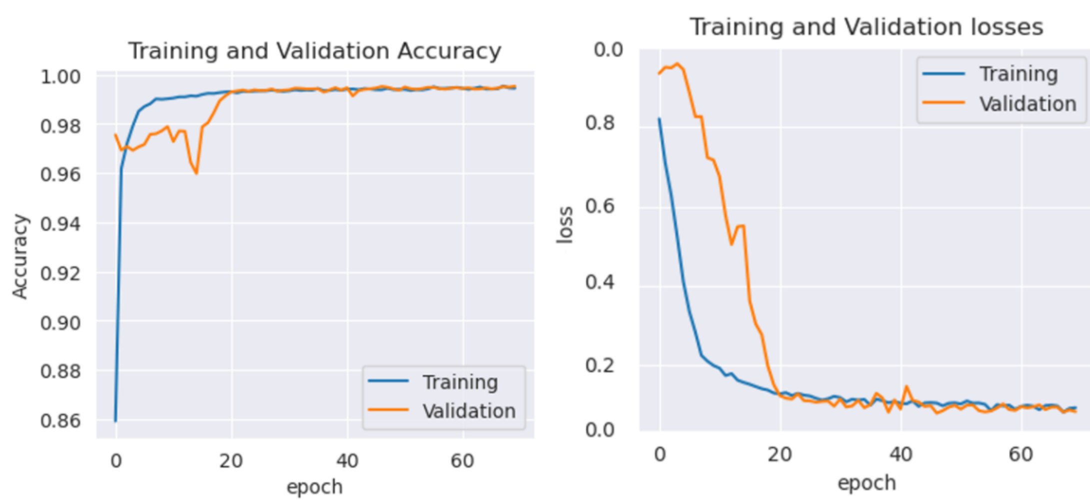


Figure 8:RefineNet model Accuracy and Loss curves

performance metrics: the sum of each column corresponds to the false positive rate (*FP*), while the sum of each row signifies the false negative rate (*FN*) for each class. The sum of the remaining diagonal elements represents the true negative rate (TN), while the diagonal elements themselves denote the true positive rate (TP). Additionally, we utilize equations (1), (2), (3), and (4) to compute the architecture's accuracy, precision, recall, and F1 score.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (1)$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

$$\text{F1} = \frac{2 \times \ Precission \times \ Recall}{Precission + \ Recall} \quad (3)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (4)$$

One commonly utilized performance metric for evaluating the classifier's effectiveness is classification accuracy. This metric takes into account four key components: (FP) (false positive), representing incorrectly identified negative samples; (FN) (false negative), denoting incorrectly identified positive samples; (TN) (true negative), indicating accurately identified negative samples; and (TP) (true positive), signifying correctly identified positive samples. To evaluate the efficiency of different models, we have generated performance metrics such as Classification Accuracy, Precision, Recall, and F1-Score, which are presented in Table 3. Notably, the results obtained from the test dataset closely mirror those obtained from the validation dataset.

Table 2: Accuracy and Loss comparison table with baselines

Model	Parameters	Training Accuracy	Validation Accuracy	Training Loss	Validation Loss
VGG16	15	96.69	94.47	0.046	0.404
Inception V3	22.3	95.60	92.58	0.943	2.147
LeafNet	0.32	85.90	79.61	0.45	0.66
AlexNet	61.3	96.89	95.78	0.294	0.182
MobileNetV2	3.4	92.21	93.25	0.52	0.226
ResNet 152V2	59.34	97.09	94.03	0.442	4.575
DenseNet 201	19.30	97.90	95.08	0.308	1.980
Proposed Model	9.19	99.29	99.27	0.094	0.091

Table 3: Accuracy, Precision, Recall and F1-score of proposed and baselines

Model	Accuracy	Precision	Recall	F1 Score
LeafNet	0.85	0.79	0.79	0.79
AlexNet	0.871	0.82	0.84	0.83
MobileNetV2	0.935	0.89	0.91	0.91
VGG16	0.9447	0.94	0.93	0.93
Inception V3	0.9258	0.93	0.92	0.92
ResNet 152V2	0.9603	0.95	0.95	0.95
DenseNet 201	0.9698	0.96	0.96	0.96
Proposed Model	0.9921	0.98	0.9817	0.9817

7. Conclusion

The proposed model, called the RefineNet-based deep learning model, has shown a remarkable improvement compared to the traditional baseline methods in chilli plant disease classification. Employing the mechanism of multipath refinement networks, the model can achieve a large receptive field to extract a large number of features. The cascaded structure used in RefineNet combines high-level and low-level features, thereby generating large-size feature maps. RefineNet architecture delivers better performance for the accurate detection of diseases in plants. This is a more effective and computationally efficient method for disease identification in plant leaves. Using 5644 images of chilli leaves of six different classes-five representing various disorders and one of healthy chilli plants-the proposed model detects Chilli plant diseases. The RefineNet model produces an accuracy of 99.29%, with precision of 98%, recall rate 98.17%, and F1 score of 98.17%. However, the main drawback with the RefineNet model is its poor scalability, causing it to have limited generalization to different tasks. In the future, mobile applications integrated with drones could be the way to come up with the classification of diseases in plants, changing precision agriculture to bring environmental sustainability in farming.

8. Declarations

Conflict of interest:

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Availability of data:

The data that support the findings of this study are available in Roboflow Universe: https://universe.roboflow.com/tugasakhir-zdajz/Chillidisease-detection1

References

Sharma K, Kaur M, Singh G (2022) Adoption status of various rice residue management technologies in sri muktsar sahib district of punjab. Journal of Community Mobilization and Sustainable Development p 387.

Sujatha R, Chatterjee JM, Jhanjhi N, et al. (2021) Performance of deep learning vs machine learning in plant leaf disease detection. Microprocessors and Microsystems 80:103,615.

Vallabhajosyula S, Sistla V, Kolli VKK (2022) Transfer learning-based deep ensemble neural network for plant leaf disease detection. Journal of Plant Diseases and Protection 129(3):545–558.

Tiwari V, Joshi RC, Dutta MK (2021) Dense convolutional neural networks based multiclass plant disease detection and classification using leaf images. Ecological Informatics 63:101,289.

Kanaparthi KR, Ilango SS (2023) A survey on training issues in chili leaf diseases identification using deep learning techniques. Procedia Computer Science 218:2123–2132.

Naik BN, Malmathanraj R, Palanisamy P (2022) Detection and classification of chilli leaf disease using a squeeze-and-excitation-based cnn model. Ecological Informatics 69:101,663.

Ahmad N, Asif HMS, Saleem G, et al. (2021) Leaf image-based plant disease identification using color and texture features. Wireless Personal Communications 121(2):1139–1168.

Wahab AHBA, Zahari R, Lim TH (2019) Detecting diseases in chilli plants using k-means segmented support vector machine. In: 2019 3rd international conference on imaging, signal processing and communication (ICISPC), IEEE, pp 57–61.

Kurmi Y, Gangwar S (2022) A leaf image localization based algorithm for different crops disease classification. Information processing in agriculture 9(3):456–474.

Patil A, Lad K (2021) Chili plant leaf disease detection using svm and knn classification. In: Rising Threats in Expert Applications and Solutions: Proceedings of FICR-TEAS 2020, Springer, pp 223–231.

Hossain E, Hossain MF, Rahaman MA (2019) A color and texture based approach for the detection and classification of plant leaf disease using knn classifier. In: 2019 international conference on electrical, computer and communication engineering (ECCE), IEEE, pp 1–6.

Sajitha P, Andrushia AD, Anand N, et al. (2024) A review on machine learning and deep learning image-based plant disease classification for industrial farming systems. Journal of Industrial Information Integration p 100572.

Yang G, He Y, Yang Y, et al. (2020) Fine-grained image classification for crop disease based on attention mechanism. Frontiers in Plant Science 11:600,854.

Ferentinos KP (2018) Deep learning models for plant disease detection and diagnosis. Computers and electronics in agriculture 145:311–318.

Tolstikhin IO, Houlsby N, Kolesnikov A, et al. (2021) Mlp-mixer: An all-mlp architecture for vision. Advances in neural information processing systems 34:24,261–24,272.

Malik A, Vaidya G, Jagota V, et al. (2022) Design and evaluation of a hybrid technique for detecting sunflower leaf disease using deep learning approach. Journal of Food Quality 2022(1):9211,700.

Liu H, Wang S, Jing G, et al. (2023) Combined cnn and rnn neural networks for gpr detection of railway subgrade diseases. Sensors 23(12):5383.

Thai HT, Le KH, Nguyen NLT (2023) Towards sustainable agriculture: A lightweight hybrid model and cloud-based collection of datasets for efficient leaf disease detection. Future Generation Computer Systems 148:488–500.

Sharma V, Kannan S, Tanya S, et al. (2024) Detecting plant diseases at scale: A distributed cnn approach with pyspark and hadoop. Procedia Computer Science 235:1044–1057.

Zhang G, Xu T, Tian Y (2022) Hyperspectral imaging-based classification of rice leaf blast severity over multiple growth stages. Plant Methods 18(1):123.

Yang W, Liu C, Zeng S, et al. (2024) Rapid identification of moldy peanuts based on three-dimensional hyperspectral object detection. Journal of Food Composition and Analysis p 106400.

Karthik R, Alfred JJ, Kennedy JJ (2023) Inception-based global context attention network for the classification of coffee leaf diseases. Ecological Informatics 77:102,213.

Karim MJ, Goni MOF, Nahiduzzaman M, et al. (2024) Enhancing agriculture through real-time grape leaf disease classification via an edge device with a lightweight cnn architecture and grad-cam. Scientific Reports 14(1):16,022.

Ramcharan A, Baranowski K, McCloskey P, et al. (2017) Deep learning for image-based cassava disease detection. Frontiers in plant science 8:1852.

Xu D, Wu Y (2020) Improved yolo-v3 with densenet for multi-scale remote sensing target detection. Sensors 20(15):4276.

Saleem MH, Khanchi S, Potgieter J, et al. (2020) Image-based plant disease identification by deep learning meta-architectures. Plants 9(11):1451.

Alirezazadeh P, Schirrmann M, Stolzenburg F (2023) Improving deep learningbased plant disease classification with attention mechanism. Gesunde Pflanzen 75(1):49–59.

Azour F (2022) Breast abnormality diagnosis using transfer and ensemble learning. PhD thesis, Université d'Ottawa/University of Ottawa.

Khanna M, Singh LK, Thawkar S, et al. (2024) Planet: a robust deep convolutional neural network model for plant leaves disease recognition. Multimedia Tools and Applications 83(2):4465–4517.

Salai B, Ramapuram C (2024) Deep learning approaches for disease detection based on plant leaf image: A review. Intelligent Systems Design and Applications: Deep Learning, Volume 2 p 323.

TugasAkhir (2023) chili disease detection1 dataset. URL https://universe.roboflow.com/tugasakhir-zdajz/chili-disease-detection1, visited on 2023-11-26.

Ullah I, Carrion D, Escalera S, et al. (2022) Meta-album: Multi-domain metadataset for few-shot image classification. In: Thirty-sixth Conference on Neural Information Processing Systems Datasets and Benchmarks Track.