SHM PUBLISHER
SURYA HIJAU MANFAAT

Journal of Soft Computing Exploration

Homepage: htttps://shmpublisher.com/index.php/joscex

JOSCEX

p-ISSN: 2746-7686
e-ISSN: 2746-0991

.

Squeeze-and-Excitation networks and attention mechanism in automatic detection of coffee leaf diseases based on images

Muhammad Izza Iqbal[1], Donny Avianto[2]
[1,2]Department of Informatics, Universitas Teknologi Yogyakarta, Indonesia

Article Info

Article history:

Received Oct 30, 2024
Revised Nov 14, 2024
Accepted Nov 28, 2024

Keywords:

Attention Mechanism
Coffee Leaf Disease Detection
Computer Vision
Deep Learning
Squeeze-and-Excitation Networks

ABSTRACT

This research examines the effectiveness of Squeeze-and-Excitation Networks (SENet) combined with Attention Mechanism for automated detection of coffee leaf diseases. The integration of SENet and Attention Mechanism presents a promising technological opportunity as SENet has proven effective in improving CNN performance by modeling channel interdependencies, while Attention Mechanism enables focused feature extraction on crucial leaf areas - a combination that remains underexplored in coffee leaf disease detection. Using a combination of three datasets: Coffee Leaf Diseases, Disease and Pest in Coffee Leaves, and RoCoLe.Original, comprising 3,177 coffee leaf images divided into four classes (Healthy, Miner, Phoma, and Rust), this study compares the performance of SENet against other deep learning architectures such as InceptionV3, ResNet101V2, and MobileNet. Experiments were conducted with variations in epochs (15 and 30), three data split ratios, and three optimizer types. Results demonstrate that SENet with Attention mechanism performs, achieving a peak accuracy of 96% at 30 epochs with an 80:20 data ratio and RMSprop optimizer. InceptionV3 and MobileNet showed competitive performance with 93% accuracy, while ResNet101V2 achieved 81%. Class-wise analysis reveals SENet's proficiency in detecting various coffee leaf diseases, with F1-scores 91% for all classes.

Corresponding Author:

Muhammad Izza Iqbal,
Department of Informatics,
University of Technology Yogyakarta,
Siliwangi Road, Sleman, 55285, D.I. Yogyakarta, Indonesia.
Email: izza240803@gmail.com
https://doi.org/10.52465/joscex.v5i4.490

1. INTRODUCTION

Indonesia is one of the largest coffee-producing countries in the world and ranks second in ASEAN after Vietnam [1]. According to food researchers from the Center for Indonesia of Policy Studies (CIPS), there are two reasons why Indonesian coffee is not produced enough. First, because old trees are very susceptible to disease, and second, the coffee plant rejuvenation process has not been appropriately done [2]. Coffee is a plant vulnerable to pests and diseases [3]. The lack of knowledge and information from coffee plantations about untreated coffee plant diseases can damage plants and losses [4]. Diseases can be identified through changes in color and patterns on coffee plant leaves. However, because the colors and patterns look similar, coffee farmers find it difficult to detect these diseases [5].

The agricultural sector has begun embracing artificial intelligence as a transformative solution for improving production efficiency and reducing environmental impact [6]. Modern AI technologies, particularly Deep Learning, are transforming agriculture by enabling automated plant disease detection and reducing pesticide costs by up to 90%. The integration of AI, computer vision, and big data is driving sustainable farming innovations and enhancing agricultural efficiency. Convolutional Neural Networks (CNN), a prominent Deep Learning method, have become the go-to solution for image-related challenges in agricultural applications [7], [8]. The evolution of CNN has led to innovative algorithms, including Squeeze and Excitation Networks (SENet) developed by Jie Hu [9]–[11]. SENet's architecture employs specialized squeeze-and-excitation blocks that dynamically rebalance information based on channel importance. This design enables more effective feature selection and enhanced disease diagnosis capabilities through intelligent channel recalibration. The adaptability of SENet's architecture makes it particularly suitable for complex image analysis tasks in agricultural applications.

The complexity of coffee leaf disease detection extends beyond channel feature calibration, as environmental factors and texture variations can significantly impact image analysis [12]. To address these challenges, attention mechanisms such as the Convolutional Block Attention Module (CBAM) have been developed to enhance detection accuracy [13]. CBAM implements a dual-attention approach, incorporating both channel and spatial attention mechanisms [14]. The channel attention component emphasizes significant channel characteristics, while spatial attention highlights critical image regions. This comprehensive attention mechanism allows for more nuanced and accurate disease detection, particularly in varying environmental conditions and complex leaf textures.

Research [15] used a combination of the Haralick method for texture and Color Histogram for color. The classification model utilizing the Random Forest approach performed well, with an accuracy of 98.83% and good recall and precision values for each class. This study successfully identified coffee leaf diseases with a high level of accuracy. Subsequently, research [16] applied Convolutional Neural Network (CNN) techniques to detect diseases in arabica coffee leaves. The dataset included 2,829 images of coffee leaves, categorized as "Leaf Spot Disease," "Leaf Rust," and "Healthy Leaf." Data was divided into 80% for training and 20% for testing. This study explored various hyperparameter settings, including epochs, batch size, and optimizer type. Results indicated that increasing epochs and batch size enhanced CNN model accuracy. The optimal model, trained over 50 epochs with a batch size of 32 and the Adam optimizer, reached an accuracy of 94.33% in final testing, demonstrating the CNN model’s efficacy in accurately detecting coffee leaf diseases.

Further research [17] found diseases in robusta coffee leaves using the ResNet-50 architecture. This study evaluated model performance using a "Confusion Matrix," which calculates accuracy, precision, recall, specificity, and F1-Score metrics. In binary and multi-class classification, the model classified healthy and unhealthy coffee leaves with 92.68% accuracy and 92.88% F1-Score. In multi-class classification, the model distinguished types of diseases on coffee leaves and produced 92.68% accuracy and 92.88% F1-Score. The following research [18] focused on applying the Convolutional Neural Network (CNN) based Squeeze-and-Excitation Network (SENet) method to classify diseases on tomato leaves. This technique analyzes attributes such as the affected area's structure and color to aid detection. The disease classification process was carried out with CNN, and the results showed that the SENet CNN approach provided better performance than other methods, such as ResNet and SVM, with detection accuracy reaching 92-96%.

Further research [19] regarding coffee leaf diseases identified three main types of diseases: Leaf Blight, Leaf Miner, and Leaf Rust. This process uses digital image processing methods with feature extraction using Local Binary Pattern (LBP) and Random Forest algorithm classification. The dataset consisted of 240 images divided into 192 training data and 48 test data. The test results showed the best accuracy of 95.83%, achieved with image size parameters of 128x128 pixels, LBP radius = 1, and Random Forest n-estimators = 100. This research proves that the system built effectively identifies diseases on coffee leaves.

Based on previous research, although methods such as Haralick, Color Histogram, ResNet-50, and CNN show good performance in detecting coffee leaf diseases, there are limitations in capturing essential features in more depth. Squeeze-and-Excitation Networks (SENet) were chosen because they can recalibrate features based on channels, strengthening the focus on relevant features. With the addition of Attention Mechanism such as CBAM, the model can pay more attention to critical areas of the image, increasing the precision of coffee leaf disease detection compared to other methods.

2. METHOD

This study employs a comprehensive methodological approach to develop and evaluate a model for automatic detection of coffee leaf diseases using deep learning techniques. The study procedure is divided into five major phases: data collection, data preprocessing, model training, and evaluation. Each phase aims to systematically meet the study objectives while ensuring the suggested model's robustness. As illustrated in Figure 1, the research flow consists of interconnected phases that form the foundation of our methodological approach. The research flow diagram demonstrates how each phase contributes to the development of our disease detection model, starting with data collection and progressing systematically through preprocessing, training, and evaluation stages. The following sections describe the steps used in each phase in detail.

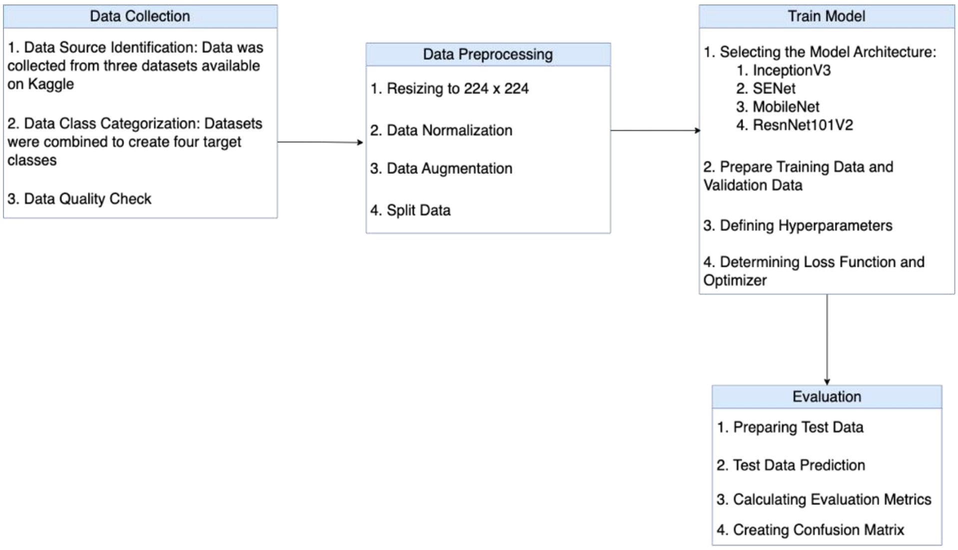


Figure 1. Research Flow

Data Collection

This experiment involved collecting images of healthy and infected coffee leaves from three public Kaggle datasets: Coffee Leaf Diseases by Gaurav Dutta, Disease and Pests in Coffee Leaves by Alvaro Leandro Cavalcante Carneiro and Lucas Brito, and RoCoLe.Original by Diego P. González, covering four classes—healthy, miner, phoma, and rust. The dataset of 3,177 images was split into training and validation sets, tested at three ratios: 90/10, 80/20, and 70/30, respectively. A separate test set of 636 images from an independent source was reserved solely for the final model evaluation, ensuring no data leakage and providing an unbiased performance assessment and Figure 2 illustrates the distribution of data across the four classes.

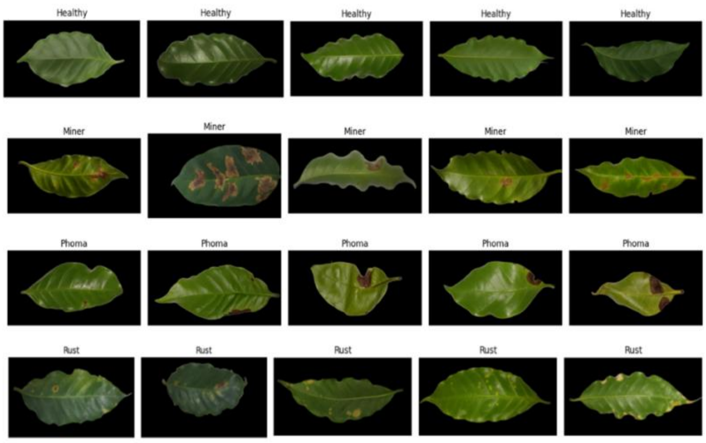


Figure 2. Data from each class

.

Data Preprocessing

Before the image data is used in the model, it undergoes a comprehensive preprocessing process that includes data splitting, resizing, normalization, cleaning, and augmentation. These steps ensure the data is properly prepared for model training, resulting in more accurate and effective outcomes. Table 1 outlines the basic parameters utilized during preprocessing and training.

Table 1. Basic Parameters

Category	Parameter	Value
Input	Image Size	224x224 pixel
	Batch Size	32
Split Data	Ratio Data	90:10, 80:20, 70:30
Data Augmentation	Rescale	1/255
	Rotation Range	20°
	Width Shift Range	0,2
	Height Shift Range	0,2
	Shear Range	0,2
	Zoom Range	0,2
	Fill Mode	'nearest'
	Brightness Range	[0,8, 1.2]
	Horizontal Flip	TRUE
	Vertical Flip	TRUE
Training	Learning Rate	0,0001
	Epochs	15 & 30
	Loss Function	categorical_crossentropy
	Optimizer	Adam, RMSprop, SGD

Table 1 highlights the methodological approach of the current study, which employed a more comprehensive set of data augmentation techniques than those referenced in other articles [20]–[22]. These techniques included not only common methods like rotation and flipping but also brightness adjustments and a wider zoom range. Furthermore, the training hyperparameters differed from those employed in other studies. This work utilized a lower learning rate of 0.0001, in contrast to the higher rates of 0.001 and 0.0002 employed in other studies. Additionally, it explored multiple optimizers, including Adam, RMSprop, and SGD, and evaluated a wider range of training epochs. Overall, the methodological approach in the current study appears to be more extensive and varied than the analogous techniques described in the related literature.

Train Model

Currently, model development based on Squeeze-and-Excitation Networks (SENet) and attention mechanism is being carried out. Using training data, the model will be trained to learn patterns in coffee leaf images, such as distinguishing healthy leaves from infected ones. In addition, a hyperparameter adjustment process is carried out to obtain the best model results.

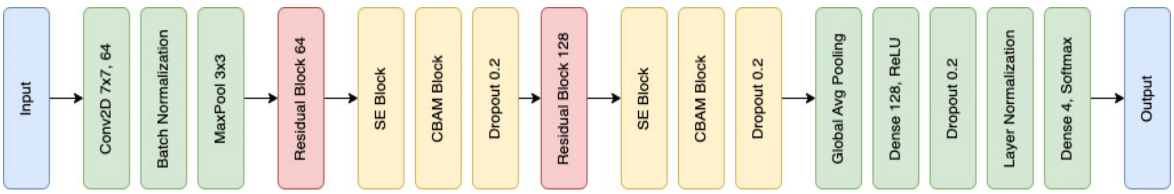


Figure 3. SENet and Attention Mechanism architecture

Figure 3 illustrates the architecture of the model developed in this study, which integrates the Convolutional Block Attention Module (CBAM) and Squeeze-and-Excitation Networks (SENet). This combination creates a synergistic framework for robust coffee leaf disease detection. The model begins with a Conv2D layer (7x7, 64 filters) that performs initial feature extraction, followed by Batch Normalization to standardize these features and a MaxPool layer (3x3) to reduce spatial dimensions while retaining important information. The core strength of this architecture lies in its innovative integration of Residual Blocks with SE and CBAM components. The first Residual Block (64 filters) enhances feature learning through skip connections, addressing the vanishing gradient problem common in deep networks. The SE Block operates through two key operations. First, the squeeze operation performs global average pooling to generate channel-wise statistics using the equation (1).

(1)

Where $\mathcal{Z}_c$ represents the squeeze output for channel c, and $H \times W$ are the feature map dimensions. Second, the excitation operation generates channel-specific weights through equation (2).

$$s = \mathbf{F}_{ex}(\boldsymbol{z}, \mathbf{W}) = \sigma(\mathbf{W}_2\delta(\mathbf{W}_1\mathbf{z})) \tag{2}$$

Where σ is the sigmoid function, δ is ReLU, and W_1, W_2 are learnable parameters. This adaptive feature recalibration significantly enhances the model's representational power [11].

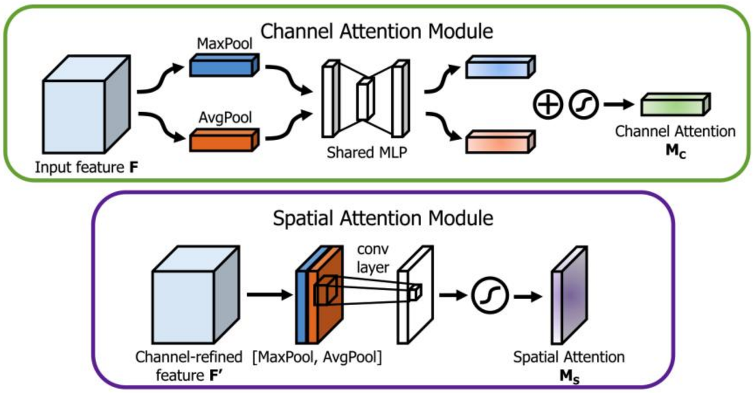


Figure 4. Diagram of Channel and Spatial Attention in CBAM [23]

CBAM generates a 1D channel attention map $\mathbf{M}_c \in \mathbb{R}^{C\times1\times1}$ and a 2D spatial attention map $\mathbf{M}_s \in \mathbb{R}^{1\times H\times W}$ from an intermediate feature map $\mathbf{F} \in \mathbb{R}^{C\times H\times W}$, as illustrated in Figure 4. The complete attention process can be summarized as follows equations (3) and (4):

$$\mathbf{F}' = \mathbf{M}_c(\mathbf{F}) \otimes \mathbf{F} \tag{3}$$

$$\mathbf{F}'' = \mathbf{M}_s(\mathbf{F}) \otimes \mathbf{F}' \tag{4}$$

Where ⊗ represents element-wise multiplication. This dual attention approach is particularly advantageous for detecting subtle disease patterns in coffee leaves, as it can highlight both the relevant features and their spatial locations within the leaf image.The CBAM Block further refines feature processing by implementing a dual attention mechanism. The Channel Attention Module (CAM) combines both max-pooling and average-pooling operations as shown in equation (5).

$$\mathrm{M}_c(\mathbf{F}) = \sigma(\mathbf{W_1}(\mathbf{W_0}(\mathbf{F}^{c}_{\max})) \tag{5}$$

Where σ denotes the sigmoid function, $\mathbf{W}_0 \in \mathbb{R}^{C/r\times C}$ and $\mathbf{W}_1 \in \mathbb{R}^{C\times C/r}$. For the Spatial Attention Module (SAM) focuses on spatial information using equation (6).

$$\mathrm{M}_s(\mathbf{F}) = \sigma(f^{7x7}([AvgPool(\mathbf{F}); MaxPool(\mathbf{F})])$$
$$\mathrm{M}_s(\mathbf{F}) = \sigma(f^{7x7}([\mathbf{F}^{s}_{avg};\ \mathbf{F}^{s}_{max}]) \tag{6}$$

.

Where $f^{7\times7}$ is a convolution operation with a 7 × 7 filter size, and σ stands for the sigmoid function. The architecture then progresses to a deeper Residual Block (128 filters), followed by another SE-CBAM combination, creating a hierarchical feature extraction pipeline. The Global Average Pooling layer reduces spatial dimensions while maintaining feature relationships, and the Dense layers (128 neurons with ReLU activation, followed by 4 neurons with Softmax) perform the final classification. Strategic placement of Dropout (0.2) and normalization layers throughout the network enhances regularization and prevents overfitting.

Evaluation

After the model is created and trained, an evaluation is conducted to assess the model's ability to find and classify coffee leaf diseases. This is done using test data never used in the training phase. The results are typically measured by accuracy, precision, recall, and F1 score metrics. Equations (7) are utilized to compute Accuracy [24]. Accuracy is the percentage of successfully classified data to the total number of data sets [25].

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{7}$$

Precision is calculated by dividing true positive observations by the expected positive observations. Precision is calculated using equation (8) [26].

$$Precision = \frac{TP}{(TP+FP)} \tag{8}$$

Recall is the ratio of true positive predictions (TP) to true positive outcomes. The equation (9) is used to calculate recall [27].

$$Recall = \frac{TP}{(TP+FN)} \tag{9}$$

The F1-score is a weighted mean of recall and precision that attempts to analyze circumstances in which the target distribution exhibits anomalies by employing the positive and negative functions as described. The F1-score is calculated in line with equation (10) [28].

$$F1\text{-}Score = 2\ x \frac{(Presisi\ x\ Recall)}{(Presisi+Recall)} \tag{10}$$

3. RESULTS AND DISCUSSIONS

This study evaluates SENet performance in image classification, especially for tasks requiring detailed feature extraction, by comparing it to InceptionV3, ResNet101V2, and MobileNet. SENet channel-wise attention focuses on relevant features, useful for identifying subtle patterns like plant disease indicators, while ResNet101V2's residual connections enhance gradient flow and feature learning in deep networks. InceptionV3's multi-scale feature extraction captures complex textures, and MobileNet lightweight design balances speed and accuracy, ideal for limited-resource deployments. Recent studies affirm these models' efficacy in agricultural image analysis, showcasing varied network structures for complex data [29]. Tables 2 through 5 summarize the overall experimental results, highlighting performance trends across models, optimizers, data splits, and epochs.

Table 2. Experiments Using InceptionV3 Architecture

No	Model	Data Split	Optimizer	Accuracy		Loss	
				15 epoch	30 epoch	15 epoch	30 epoch
1	InceptionV3	90/10	Adam	87%	92%	0,2998	0,1882
2	InceptionV3	90/10	RMSprop	85%	92%	0,3498	0,2106
3	InceptionV3	90/10	SGD	77%	87%	0,6049	0,4379
4	**InceptionV3**	**80/20**	**Adam**	**89%**	**93%**	**0,2892**	**0,2256**
5	InceptionV3	80/20	RMSprop	86%	91%	0,3335	0,2256
6	InceptionV3	80/20	SGD	77%	84%	0,5972	0,475
7	InceptionV3	70/30	Adam	88%	91%	0,3022	0,2185
8	InceptionV3	70/30	RMSprop	87%	89%	0,3228	0,2819
9	InceptionV3	70/30	SGD	79%	82%	0,5984	0,5051

Table 3. Experiments Using SENet Architecture

No	Model	Data Split	Optimizer	Accuracy		Loss	
				15 epoch	30 epoch	15 epoch	30 epoch
1	SENet	90/10	Adam	90%	95%	0,3149	0,1752
2	SENet	90/10	RMSprop	93%	95%	0,2301	0,2245
3	SENet	90/10	SGD	86%	92%	0,3232	0,196
4	SENet	80/20	Adam	90%	95%	0,319	0,155
5	**SENet**	**80/20**	**RMSprop**	**94%**	**96%**	**0,2335**	**0,1783**
6	SENet	80/20	SGD	80%	94%	0,4953	0,1886
7	SENet	70/30	Adam	94%	92%	0,1841	0,2764
8	SENet	70/30	RMSprop	91%	93%	0,2274	0,308
9	SENet	70/30	SGD	88%	92%	0,3231	0,2184

Table 4. Experiments Using MobileNet Architecture

No	Model	Data Split	Optimizer	Accuracy		Loss	
				15 epoch	30 epoch	15 epoch	30 epoch
1	MobileNet	90/10	Adam	90%	87%	0,2703	0,3555
2	MobileNet	90/10	RMSprop	86%	88%	0,363	0,3024
3	MobileNet	90/10	SGD	74%	72%	0,9322	0,8499
4	**MobileNet**	**80/20**	**Adam**	**92%**	**93%**	**0,2448**	**0,2076**
5	MobileNet	80/20	RMSprop	91%	86%	0,3281	0,323
6	MobileNet	80/20	SGD	73%	74%	0,9525	0,8128
7	MobileNet	70/30	Adam	88%	90%	0,3403	0,3046
8	MobileNet	70/30	RMSprop	84%	89%	0,4376	0,3287
9	MobileNet	70/30	SGD	64%	78%	10.076	0,6588

Table 5. Experiments Using ResNet101V2 Architecture

No	Model	Data Split	Optimizer	Accuracy		Loss	
				15 epoch	30 epoch	15 epoch	30 epoch
1	ResNet101V2	90/10	Adam	78%	72%	0,5934	0,6548
2	ResNet101V2	90/10	RMSprop	67%	76%	0,7731	0,567
3	ResNet101V2	90/10	SGD	52%	53%	12.471	11.403
4	ResNet101V2	80/20	Adam	72%	74%	0,6461	0,6754
5	ResNet101V2	80/20	RMSprop	66%	76%	0,8035	0,5949
6	ResNet101V2	80/20	SGD	53%	55%	12.359	11.716
7	**ResNet101V2**	**70/30**	**Adam**	**74%**	**81%**	**0,6346**	**0,4951**
8	ResNet101V2	70/30	RMSprop	67%	74%	0,7737	0,6336
9	ResNet101V2	70/30	SGD	49%	51%	12.577	11.972

.

Table 6 shows the training results for the best performing model architectures, highlighting SENet's superior accuracy (96%) with an 80:20 data split and the RMSprop optimizer, followed by InceptionV3 and MobileNet, each achieving 93% accuracy with the Adam optimizer. ResNet101V2 had the lowest performance, achieving 81% accuracy with a 70:30 data split. A comparison of the results of this study with previous research reveals some important insights.

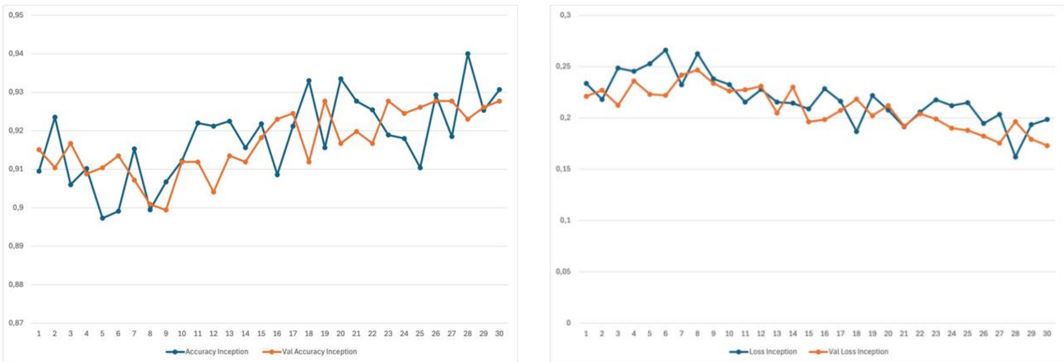
Table 6. Training Results for Each Best Model Architecture

No	Model	Epoch	Data Split	Optimizer	Accuracy	Val Loss
1	InceptionV3	30	80:20	Adam	93%	0,17
2	SENet	30	80:20	RMSprop	96%	0,17
3	MobileNet	30	80:20	Adam	93%	0,21
4	ResNet101V2	30	70:30	Adam	81%	0,50

The comparison of the four articles reveals several noteworthy findings [18], [20]–[22]. First, the current study achieved high validation accuracies with the SENet (96%), InceptionV3 (93%), and MobileNet (93%) models, outperforming or matching previous results, including a 92% test accuracy for SENet, 93.75% for InceptionV3, 79.14% for MobileNet and a lower 61.5% validation accuracy for ResNet101V2 in prior studies. Secondly, this study employed a thorough evaluation approach by testing models across multiple train-validation split ratios, aligning with best practices highlighted in the MobileNet article [22]. Finally, all three studies targeted important crops—tomatoes, rice, coffee, and potatoes—underscoring the real-world relevance and practical value of this research. The strong performances of SENet in Table 6, in particular, suggest promising potential for real-world deployment.

InceptionV3 Model Evaluate

The performance evaluation of the InceptionV3 model can be observed through its accuracy and loss metrics as shown in Figure 5. The accuracy curve, depicted in Figure 5a, shows a stable trend with slight fluctuations throughout training. Initially, the Inception model's accuracy was already high at around 0.9095 and continued to fluctuate around this number until the end of training. In the middle epochs, the model's accuracy peaked at 0.94 during epoch 28. The validation accuracy curve follows a similar trend, with relatively stable values and minor fluctuations, reaching a final result of 0.9277 at epoch 30, demonstrating exemplary performance in training and validation.



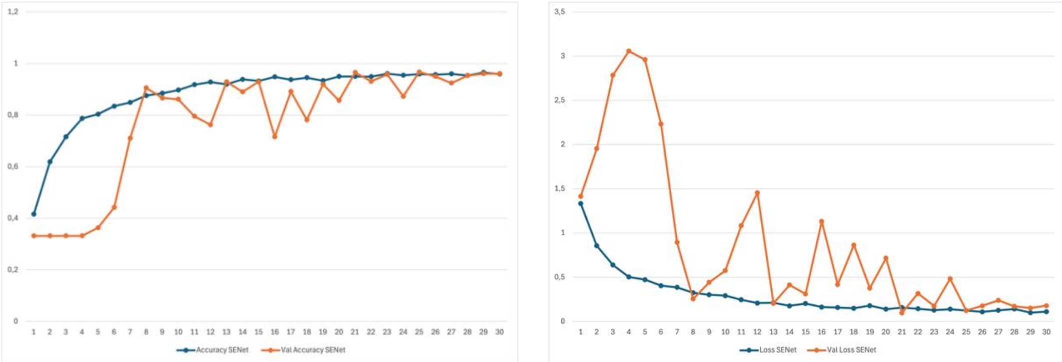
(a) (b)

Figure 5. (a) accuracy and (b) loss InceptionV3 model

As illustrated in Figure 5b, the loss curve for Inception shows a gradual decrease with some fluctuations at the beginning, starting from 0.2336 at epoch one and reaching its lowest point at epoch 28 with a loss value of 0.162. The validation loss follows a similar pattern, showing a consistent decrease with minor fluctuations, with a final result of 0.173 at epoch 30. Overall, this downward trend in loss indicates model improvement in minimizing prediction errors for training and validation data.

SENet Model Evaluate

The accuracy metrics, as illustrated in Figure 6a, shows a significant increase from the beginning to the end of training. At the beginning of training, the accuracy of the model increases consistently, starting from a low value to reach more than 0.9 around the middle of training, with a steady increase. Although the rate of improvement slowed down slightly after the middle, the accuracy continued to approach the highest value of about 0.96 at the end of training. Meanwhile, the validation accuracy curve shows greater variation, where it is initially stagnant, but increases dramatically after a few epochs, peaking at 0.9654. Despite the fluctuations, the general trend shows a steady increase.



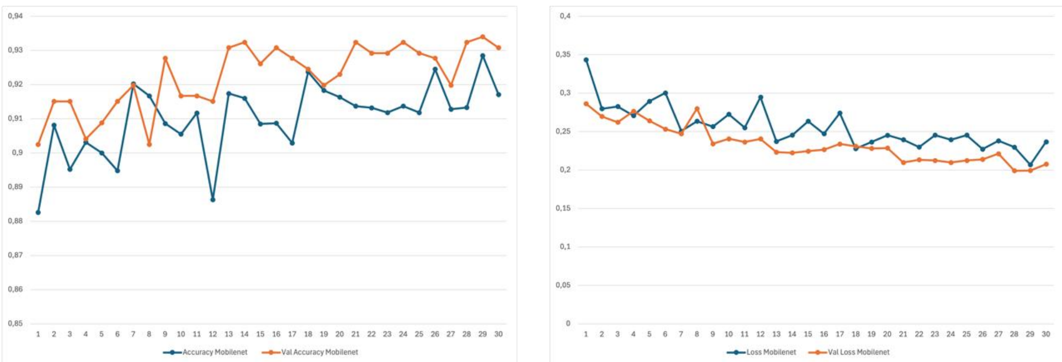
(a) (b)

Figure 6. (a) accuracy and (b) loss SENet model

The loss metric is illustrated in Figure 6b, which shows the progress of the model optimization during the training process. The SENet loss curve shows a consistent decrease during training, starting from a high value and decreasing to around 0.1107 at the end of training. This decrease indicates an improvement in model performance, although the rate of decrease slows down after a few epochs. On the other hand, the validation loss curve experienced larger fluctuations, with high values at the beginning of training and a significant decrease after a few epochs, especially in the middle epochs, with the lowest value reaching 0.0967. Although there is a slight increase thereafter, the overall trend shows a significant improvement in the generalization ability of the model.

MobileNet Model Evaluate

The training results of the Mobilenet model can be observed through the accuracy curve shown in Figure 7a. The Mobilenet accuracy curve graph shown in this figure increases relatively steadily from the beginning to the end of the training, although there are slight fluctuations in some epochs. The accuracy starts from a fairly high number and continues to increase, with a peak at epoch 29 of 0.9285. The trend of validation accuracy is similar, with the value remaining stable around 0.93, indicating good generalization ability of the model, with a final result of 0.9308 at epoch 30.



(a) (b)

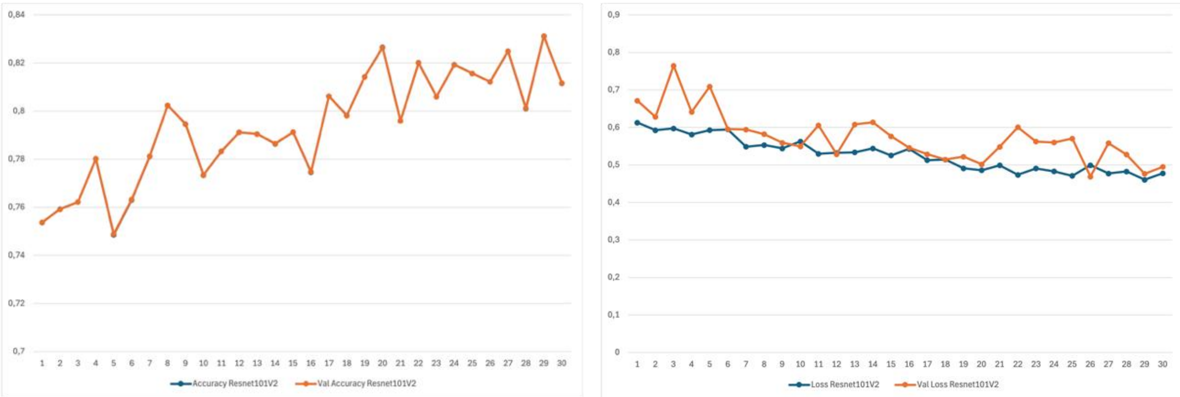
Figure 7. (a) accuracy and (b) loss MobileNet model

.

The loss curve shown in Figure 7b can be visualized to examine the model's performance. This figure's loss curve for Mobilenet shows a steady decline throughout the course of training, suggesting that the model's parameters were effectively optimized. As the model improves its ability to predict the output, the training loss gradually drops from 0.3435 to 0.2068 at the 29th epoch. Even though there are some little variations that are common during deep learning model training, the validation loss curve shows a similar declining trend and settles at 0.2076 by epoch 30. The model has successfully struck a compromise between fitting the training data and generalizing to unknown data, as evidenced by the near alignment of training and validation loss values, with just a tiny gap of 0.0008

ResNet101V2 Model Evaluate

As illustrated in Figure 8a, the accuracy curve of Resnet101V2 demonstrates a gradual and consistent enhancement throughout the training phase. At the outset of the training process, the model exhibits a moderate level of accuracy, which then increases steadily as the number of epochs rises. This growth reaches its peak at epoch 29, where the accuracy level reaches 0.8312. The validation accuracy trend exhibits a comparable pattern, with a final result of 0.8311 at epoch 29, indicating consistency between training and validation performance throughout the training process.



(a) (b)

Figure 8. (a) accuracy and (b) loss ResNet101V2 model

As illustrated in Figure 8b, the Resnet101V2 loss curve exhibits a distinct downward trajectory throughout the training phase, commencing at 0.6129 and culminating at 0.4605 at epoch 29. The validation loss also exhibits a comparable pattern, with a final value of 0.4763 at epoch 29. Despite minor fluctuations, the declining loss suggests that the model is effectively reducing the prediction error in a gradual manner on both the training and validation datasets.

Classification Report For The Best Models

Table 7. Classification Report for InceptionV3 and SeNet

	InceptionV3				**SeNet**			
	Precision	**Recall**	**F1-Score**	**Support**	**Precision**	**Recall**	**F1-Score**	**Support**
Healthy	0,91	0,79	0,85	174	0,93	0,89	0,91	174
Miner	0,87	0,97	0,92	142	1	0,97	0,99	142
Phoma	0,98	0,97	0,97	99	0,91	1	0,95	99
Rust	0,84	0,86	0,85	221	0,93	0,94	0,93	221
Accuracy			0,88	636			0,94	636
Macro Avg	0,9	0,9	0,9	636	0,94	0,95	0,95	636
Weighted Avg	0,89	0,88	0,88	636	0,94	0,94	0,94	636

Table 8. Classification Report for MobileNet and ResNet101V2

	MobileNet				ResNet101V2			
	Precision	**Recall**	**F1-Score**	**Support**	**Precision**	**Recall**	**F1-Score**	**Support**
Healthy	0,91	0,8	0,85	174	0,81	0,38	0,52	265
Miner	0,87	0,96	0,91	142	0,83	0,76	0,8	234
Phoma	0,91	0,99	0,95	99	0,56	0,99	0,72	133
Rust	0,86	0,84	0,85	221	0,67	0,79	0,73	322
Accuracy			0,88	636			0,7	954
Macro Avg	0,88	0,9	0,89	636	0,72	0,73	0,69	954
Weighted Avg	0,88	0,88	0,88	636	0,74	0,7	0,69	954

The evaluation results of four deep learning models for plant disease classification are presented in Tables 7 and 8, which compare the models' performance metrics including accuracy, precision, recall, and F1-score across different plant conditions. The analysis of these performance metrics shows that among the tested models (InceptionV3, SeNet, MobileNet, and ResNet101V2) for classifying plant conditions into Healthy, Miner, Phoma, and Rust categories, SeNet demonstrated superior performance with 94% accuracy. As detailed in Table 7, InceptionV3 and MobileNet achieved comparable performance with 88% accuracy, while ResNet101V2 showed relatively lower performance at 70%. Table 8 further breaks down the performance metrics by class, revealing that SeNet excelled across all categories, notably achieving perfect precision (100%) for the Miner class and maintaining strong performance in other categories. The detailed metrics in Table 8 also indicate that ResNet101V2 consistently underperformed, with particular weakness in Healthy class detection (38% recall). In the comprehensive evaluation shown in both tables, SeNet proved its superiority by achieving 95% and 94% in macro and weighted averages respectively, significantly outperforming the other tested models.

4. CONCLUSION

This study evaluates the effectiveness of SENet (Squeeze-and-Excitation Networks) combined with the Attention mechanism for automatic detection of coffee leaf diseases. SENet achieved the highest accuracy of 96% with a validation loss of 0.18 using an 80:20 data split and RMSprop optimizer, where its performance increased from 94% to 96% when epochs were increased from 15 to 30. Compared with other models, SENet outperformed InceptionV3, ResNet101V2, and MobileNet with testing accuracy reaching 94%, while InceptionV3 and MobileNet achieved 88% each and ResNet101V2 70%. Model optimization showed that Adam and RMSprop optimizers provided the best results with an 80:20 data split consistently producing the best performance, and increasing epochs proved to improve the performance of all models. In per-class disease detection, SENet excelled in detecting all classes, especially Miner and Rust, with F1 scores above 0.91 for all classes. However, there are still challenges in distinguishing healthy leaves from Rust disease. In conclusion, the combination of SENet with the Attention mechanism proves to be highly effective in detecting various coffee leaf diseases, although there is still room for improvement.

REFERENCES

[1] C. P. Ginting and F. Kartiasih, "Analisi Ekspor Kopi Indonesia Ke Negara-Negara ASEAN," *J. Ilm. Ekon. dan Bisnis*, vol. 16, pp. 143–157, Sep. 2019, doi: 10.31849/jieb.v16i2.2922.

[2] D. Irfansyah and others, "Arsitektur Convolutional Neural Network (CNN) Alexnet Untuk Klasifikasi Hama Pada Citra Daun Tanaman Kopi," vol. 6, no. 2, 2021.

[3] L. X. B. Sorte, C. T. Ferraz, F. Fambrini, R. D. R. Goulart, and J. H. Saito, "Coffee leaf disease recognition based on deep learning and texture attributes," in *Procedia Computer Science*, Elsevier B.V., 2019, pp. 135–144. doi: 10.1016/j.procs.2019.09.168.

[4] J. Jepkoech, D. M. Mugo, B. K. Kenduiywo, and E. C. Too, "Arabica coffee leaf images dataset for coffee leaf disease detection and classification," *Data Br.*, vol. 36, Jun. 2021, doi: 10.1016/j.dib.2021.107142.

[5] M. I. Rosadi, L. Hakim, and M. F. A., "Classification of Coffee Leaf Diseases using the Convolutional Neural Network (CNN) EfficientNet Model," *Conf. Ser.*, vol. 4, no. 1, pp. 58–69, Dec. 2023, doi: 10.34306/conferenceseries.v4i1.627.

.

[6] D. Novtahaning, H. A. Shah, and J. M. Kang, “Deep Learning Ensemble-Based Automated and High-Performing Recognition of Coffee Leaf Disease,” *Agric.*, vol. 12, no. 11, Nov. 2022, doi: 10.3390/agriculture12111909.

[7] G. Wan and L. Yao, “LMFRNet: A Lightweight Convolutional Neural Network Model for Image Analysis,” *Electron.*, vol. 13, no. 1, Jan. 2024, doi: 10.3390/electronics13010129.

[8] R. Archana and P. S. E. Jeevaraj, “Deep learning models for digital image processing: a review,” *Artif Intell Rev*, vol. 57, no. 1, Jan. 2024, doi: 10.1007/s10462-023-10631-z.

[9] C. Wen and others, “Recognition of mulberry leaf diseases based on multi-scale residual network fusion SENet,” *PLoS One*, vol. 19, no. 2, Feb. 2024, doi: 10.1371/journal.pone.0298700.

[10] Y. Zhao, C. Sun, X. Xu, and J. Chen, “RIC-Net: A plant disease classification model based on the fusion of Inception and residual structure and embedded attention mechanism,” *Comput Electron Agric*, vol. 193, p. 106644, Feb. 2022, doi: 10.1016/J.COMPAG.2021.106644.

[11] J. Hu, L. Shen, S. Albanie, G. Sun, and E. Wu, “Squeeze-and-Excitation Networks,” *ArXiv*, Sep. 2019.

[12] W. Yang, Y. Yuan, D. Zhang, L. Zheng, and F. Nie, “An Effective Image Classification Method for Plant Diseases with Improved Channel Attention Mechanism aECAnet Based on Deep Learning,” *Symmetry (Basel)*, vol. 16, no. 4, Apr. 2024, doi: 10.3390/sym16040451.

[13] X. Zhang, D. Li, X. Liu, T. Sun, X. Lin, and Z. Ren, “Research of segmentation recognition of small disease spots on apple leaves based on hybrid loss function and CBAM,” *Front Plant Sci*, vol. 14, 2023, doi: 10.3389/fpls.2023.1175027.

[14] P. Alirezazadeh, M. Schirrmann, and F. Stolzenburg, “Improving Deep Learning-based Plant Disease Classification with Attention Mechanism,” *Gesunde Pflanz.*, vol. 75, no. 1, pp. 49–59, Feb. 2023, doi: 10.1007/s10343-022-00796-y.

[15] S. K. Wildah, A. Latif, A. Mustopa, S. Suharyanto, M. S. Maulana, and A. Sasongko, “Klasifikasi Penyakit Daun Kopi Menggunakan Kombinasi Haralick, Color Histogram dan Random Forest,” *J. Sist. dan Teknol. Inf.*, vol. 11, no. 1, p. 35, Jan. 2023, doi: 10.26418/justin.v11i1.60985.

[16] A. I. Saputra, I. Weni, and U. Khaira, “Implementasi Metode Convolutional Neural Network Untuk Deteksi Penyakit Pada Tanaman Kopi Arabika Melalui Citra Daun Berbasis Android,” *Decod. J. Pendidik. Teknol. Inf.*, vol. 4, no. 1, pp. 41–51, Oct. 2023, doi: 10.51454/decode.v4i1.231.

[17] I. Awaludin and others, “Analisis Kinerja ResNet-50 dalam Klasifikasi Penyakit pada Daun Kopi Robusta,” *J. Inform.*, vol. 9, no. 2, 2022.

[18] P. Pragy, V. Sharma, and V. Sharma, “Senet CNN Based Tomato Leaf Disease Detection,” in *International Journal of Innovative Technology and Exploring Engineering*, Sep. 2019, pp. 773–777. doi: 10.35940/ijitee.K1452.0981119.

[19] B. Wahyuningtyas, I. I. Tritoasmoro, and N. Ibrahim, “Identifikasi Penyakit Pada Daun Kopi Menggunakan Metode Local Binary Pattern Dan Random Forest,” *e-Proceeding Eng.*, vol. 8, no. 6, 2022.

[20] S. Sheila, I. P. Sari, A. B. Saputra, M. K. Anwar, and F. R. Pujianto, “Detection of Diseases in Rice Leaves Based on Image Processing Using the Convolutional Neural Network (CNN) Method,” *J. Multimed. Netw. INFORMATICS*, vol. 9, no. 1, pp. 27–34, Apr. 2023.

[21] A. Gheorghiu, I.-M. Tăiatu, D.-C. Cercel, I. Marin, and F. Pop, “Evaluating Data Augmentation Techniques for Coffee Leaf Disease Classification,” *ArXiv*, Jan. 2024.

[22] A. Fuadi and others, “Perbandingan Arsitektur MobileNet dan NasNetMobile untuk Klasifikasi Penyakit pada Citra Daun Kentang,” *JIPI*, vol. 7, no. 3, Sep. 2022, doi: 10.29100/jipi.v7i3.3026.

[23] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, “CBAM: Convolutional Block Attention Module,” *CoRR*, vol. abs/1807.06521, 2018.

[24] M. A. Muslim *et al.*, “New model combination meta-learner to improve accuracy prediction P2P lending with stacking ensemble learning,” *Intell. Syst. with Appl.*, vol. 18, p. 200204, May 2023, doi: 10.1016/j.iswa.2023.200204.

[25] Y. Fu, L. Guo, and F. Huang, “A lightweight CNN model for pepper leaf disease recognition in a human palm background,” *Heliyon*, vol. 10, no. 12, Jun. 2024, doi: 10.1016/j.heliyon.2024.e33447.

[26] A. G. Argho, M. M. S. Maswood, M. I. Mahmood, and N. Mondol, “EfficientCovNet: A CNN-based approach to detect various pulmonary diseases including COVID-19 using modified EfficientNet,” *Intell. Syst. with Appl.*, vol. 21, Mar. 2024, doi: 10.1016/j.iswa.2023.200315.

[27] D. K. Saha, A. M. Joy, and A. Majumder, “YoTransViT: A transformer and CNN method for predicting and classifying skin diseases using segmentation techniques,” *Inf. Med Unlocked*, vol. 47, Jan. 2024, doi: 10.1016/j.imu.2024.101495.

[28] C. Ashwini and V. Sellam, “An optimal model for identification and classification of corn leaf disease using hybrid 3D-CNN and LSTM,” *Biomed Signal Process Control*, vol. 92, p. 106089, Jun. 2024, doi: 10.1016/J.BSPC.2024.106089.

[29] N. Parashar and P. Johri, “Deep Learning for Cotton Leaf Disease Detection,” in *2024 2nd International Conference on Device Intelligence, Computing and Communication Technologies (DICCT)*, Dehradun, India, 2024, pp. 158–162. doi: 10.1109/DICCT61038.2024.10533021.