Original Article

Modified Convolutional Block Attention Mechanism-Based VGG16 for Apple Plant Leaf Disease Classification

Pooja Chandrakantbhai Gajjar[1], Hetal Nirmit Dalal[2], Hina Kalpesh Patel[1], Nisha Kantilal Prajapati[1]

[1]*Electronics & Communication Engineering Department, Government Engineering College, Gandhinagar, Gujarat, India.*
[2]*Instrumentation & Control Engineering Department, Vishwakarma Government Engineering College, Chandkheda, Ahmedabad, India.*

[1]*Corresponding Author : pooja@gecg28.ac.in*

Abstract - Agriculture plays a significant role in a country's development like India, where most of the population's livelihood is based on agriculture. The production of apples plays a vital role in the agricultural sector by contributing to the economy through domestic consumption and export. However, different diseases exhibit similar visual symptoms on apple leaves, such as spots or discolorations, leading to misclassification. This research proposes the Modified Convolutional Block Attention Mechanism-VGG16 (MCBAM-VGG16) to classify apple plant leaf disease accurately. Global Average Pooling (GAP) layer and CBAM are added after the first convolutional layer which minimizes the number of parameters and helps to avoid underfitting issues during training. Initially, images are obtained from the plant village dataset to evaluate the proposed approach. Then, data augmentation is used to transform images, which helps MCBAM-VGG16 attain better tolerance and generalization ability. Then, W-Net is employed to segment images that capture both global context and fine-grained information within the leaf images. At last, MCBAM-VGG16 classifies apple plant leaf disease accurately. When compared to the existing techniques like Deep Convolutional Neural Networks with three convolutional layers (Conv-3 DCNN), the improved DCNN and Random Sample Consensus (RANSAC), MCBAM-VGG16 achieves a superior accuracy of 0.998.

Keywords - Agriculture, Apple plant leaf disease, Data augmentation, Modified convolutional block attention mechanism-VGG16, W-Net.

1. Introduction

Agriculture is the primary foundation for ensuring global food security and holds a significant share of the world's economy. Moreover, plant disease infections have increased recently due to the continuous deterioration of the environment. The agriculture industry faces a significant impact each year due to plant disease infection, with a recent focus on their effects on various fruit diseases [1]. Many researchers have introduced various approaches for identifying and classifying fruit diseases [2].

Apples are one of the primary fruits worldwide and one of the most commonly planted and consumed fruits worldwide [3]. Faced with a complex and ever-changing natural environment, the growth of apples is affected by numerous diseases that greatly impact their quality and yield [4]. Hence, enhancing the ability to manage apple leaf disease significantly improves yield and quality [5, 6]. Numerous fungal and bacterial diseases primarily manifest themselves on leaf and fruit area surfaces [7]. Lesions caused by pests exhibit complex patterns, which make them challenging to understand [8]. More than 100 kinds of apple diseases exist, with primary diseases like Apple Scab, Black Rot, Cedar Apple Rust, and so on, found in orchards and fruit-growing regions worldwide [9].

Accurately and quickly identifying and classifying these diseases, along with efficient analysis and management, can assist the rational use of agricultural resources like fertilizers, pesticides, and water for apple cultivation [10, 11]. The smart agriculture system depends on automatic crop disease classification and monitoring to increase agriculture's overall protectivity and maintenance [12].

With the establishment of Artificial Intelligence (AI) and computer vision, Deep Learning (DL) has received increasing attention in the image processing field due to its wide application range in agriculture [13]. Neural networks have also contributed extremely to monitoring systems that attain high accuracy and faster efficacy in different agricultural

fields [14, 15]. Moreover, a significant difficulty arises from overlapping visual symptoms between different diseases, like similar discolorations and spots, resulting in misclassification in the apple plant leaf disease classification field. To address this gap, this research proposes MCBAM-VGG16 for accurately classifying apple plant leaf disease by incorporating the GAP layer and CBAM. The GAP layer assists in minimizing parameter numbers, which helps to avoid underfitting issues during training, whereas CBAM enhances feature representation by concentrating on appropriate regions of input images. This approach provides a robust solution for classifying apple plant leaf disease by leveraging advanced methods to accurately manage the complexities of disease symptoms. The primary contributions of this research are described as follows:

- Data augmentation is employed to transform images, which assists the model in attaining better tolerance and generalization ability through exposure to various transformations of leaf images.
- W-Net segments the augmented images effectively by leveraging two interconnected U-Net. The first U-Net generates a coarse segmentation, which the second U-Net refines for higher accuracy. This approach enhances the segmentation quality by iteratively refining the results in apple leaf disease.
- The MCBAM-VGG16 classifies the apple plant leaf disease accurately by adding a GAP layer and CBAM after the first convolutional layer. This minimizes the number of parameters and helps avoid underfitting and overfitting issues during training. By performing this process, the MCBAM-VGG16 achieves superior classification accuracy in apple plant leaf disease.

The research is explained as follows: Section 2 describes a literature survey of the existing methods. Section 3 explains the proposed methodology in detail, Section 4 evaluates the results of existing techniques and the proposed methods, and Section 5 summarizes the paper's overall conclusion.

2. Literature Survey

The related works about apple leaf disease classification and their benefits and limitations are discussed in this section.

Vishnoi et al. [16] suggested a Deep Convolutional Neural Network with three convolutional layers (Conv-3 DCNN) for apple leaf disease detection and classification. The CNN approach involved a smaller number of layers, which led to a smaller computational burden. The hyperparameters were tuned by a random search approach, which assisted in selecting the most appropriate ones. The augmentation techniques like shear, shift, zoom, scaling, and flipping were employed to produce additional samples that enhanced the training set without capturing more images. However, Conv-3 DCNN did not capture complex patterns due to shallow architecture, which minimized the classification accuracy.

The overall analysis of [16] achieved minimized computational burden and enhanced efficiency but faced challenges with capturing complex patterns, which causes less classification accuracy.

Mahato et al. [17] presented an improved DCNN to detect and classify apple plant leaf disease. The chosen dataset had a class imbalance issue managed by employing class weight approaches. After managing class imbalance issues, the improved DCNN was performed to classify apple leaf disease. The presented approach employed image data augmentation and annotation methods, increasing the accuracy. Nonetheless, improved DCNN faced difficulty in generalizing across various environmental conditions and variances in leaf images, which further affected the classification performance due to inconsistent leaf appearance. In an overall examination of [17], increased accuracy via augmentation but faces struggled with a generalization which affects the performance of classification

Kurmi and Gangwar [18] introduced a Random Sample Consensus (RANSAC) to classify different plant leaf diseases. Further, image enhancement and histogram equalization were utilized in the pre-processing stage to enhance the image and provide better discrimination among objects and backgrounds. The RANSAC utilized Region of Interest (RoI) localization followed by feature extraction and classification. A mean and variance for Gaussian distribution were utilized for energy fitting functions, generating enhanced performance. Nevertheless, RANSAC struggled with highly complex or overlapping leaf disease patterns due to its dependence on fitting models to a subset of data that did not accurately represent all complex patterns. In the overall analysis of [18], enhanced performance was achieved but struggles with overlapping patterns, which causes model performances.

Hosny et al. [19] developed a lightweight Deep Convolutional Neural Network (DCNN) to classify plant leaf disease. The traditional convolutional layers were replaced by deeper separable convolutions that minimized the amount of iteration time and model parameters. A combined approach was employed by concatenating deep features and handcrafted Local Binary Pattern (LBP) features. This approach effectively captured the local spatial texture data determined in plant leaf images. However, Lightweight DCNNs have limited capacity to capture complex features because of simplified architectures, leading to decreased classification accuracy. The overall investigation of [19] achieved enhanced efficiency, but the model accuracy was decreased due to limited capability.

Zhang et al. [20] implemented integrating Capsule Networks (CapsNet) and Residual Networks (ResNet) for classifying plant leaf disease images. Initially, by refining and optimizing the traditional ResNet, the ResNet's initial convolutional layer was improved by replacing its kernel with

small convolutional kernels, which effectively extracted plant leaf lesion features. Then, a channel attention approach was developed into ResNet to enhance the model's focus on significant features. At last, improved ResNet was integrated effectively with CapsNet. The implemented approach maintained stable performance while the image changed significant rotation angles and increased the model's robustness. Nevertheless, the implemented approach struggled with constrained training data due to their deep architectures, which resulted in overfitting or generalization issues with insufficient data. In the overall analysis of [20] improved model performance and robustness under rotation but struggled with overfitting issues.

From the overall evaluation of [16] to [20], it is seen that the existing techniques have limitations, such as not capturing complex patterns due to shallow architecture, difficulty in generalizing across various environmental conditions, overfitting issues, generalization, and variances in leaf images due to inconsistent leaf appearance, and visual symptoms on apple leaves like spots or discolorations. The MCBAM-VGG16 is proposed to accurately classify apple plant leaf disease by adding a GAP layer and CBAM in VGG16 to overcome these issues. This approach minimizes the number of parameters and avoids overfitting and underfitting issues during training by concentrating on appropriate regions of input images.

3. Proposed Methodology

The MCBAM-VGG16 is proposed to classify the apple plant leaf disease accurately. Images are gathered from the plant village dataset to evaluate the proposed approach, and data augmentation is utilized to transform the images. W-Net is employed to segment the augmented images, and then MCBAM-VGG16 accurately classifies the apple plant leaf diseases. Figure 1 demonstrates a block diagram for the MCBAM-VGG16 method.
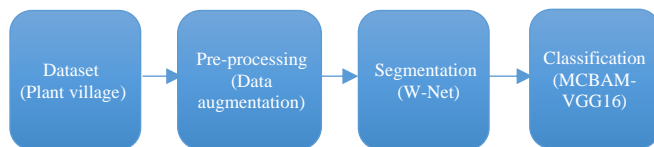

Fig. 1 Block diagram for the MCBAM-VGG16

3.1. Dataset

In this research, apple leaf images are collected from the benchmark dataset of PlantVillage [21], which is available on the PlantVillage project. The dataset has 2536 apple leaf images and is split into 4 classes. Figure 2 shows a sample of apple leaf images. The 3 classes are associated with 3 apple diseases: scab, black rot, and apple cedar rust. The remaining 1 class indicates healthy leaf images and has a 256 x 256 size in apple leaf images for all 4 classes captured with simple backgrounds at different stages. A collected image is transferred into a pre-processing stage for further processing.

(a) Black rot   **(b) Apple Scab**   **(c) Cedar Apple Rust**   **(d) Healthy Leaf**

Fig. 2 Samples for apple leaf images

3.2. Pre-processing

After gathering images, data augmentation [22] is utilized in a pre-processing stage because the original image data is inadequate for effective training. Certain transformation techniques like zoom and scaling are utilized for transforming images. These transformations produce minor variations in images, which assist in creating variety in the training set. It helps minimize overfitting and assists in attaining enhanced tolerance and generalization through exposure to various transformations.

Augmented images assist the model in learning more invariant and discriminative data, leading to improved accuracy. Then, the augmented images are passed to segment the apple leaf disease images using W-net. These transformations enhance the robustness of segmentation models by augmenting the training data with variations in perspective, position, and scale, improving their ability to generalize images.

3.3. Segmentation

The segmentation technique uses W-net [23] after the images are augmented. W-net's multi-resolution architecture facilitates skip connections at various scales, enabling the model to capture both global context and fine-grained information within the leaf images. This increases the segmentation accuracy by effectively delineating disease-affected regions from healthy images. W-net contains a two-stage U-Net with refined and coarse structures derived at each stage, which is reorganized into 8 layers: 5 deep features as structure features and 3 shallow layers as texture features. This depends on insight, which is a shallow skip connection that synthesizes textures and deep sampling layers to produce structures.

The initial part provides a feature map concatenation from the encoder path to associated feature maps from the decoder. It assists in recovering data lost during max pooling and convolutional operations. This 2nd part is identical to the prior; however, the max-pooling layer output is concatenated with the prior U-Net part. Moreover, the final layer is a 1 x 1 convolutional layer that matches feature maps to appropriate classes and employs a softmax activation function. For the training stage, W-Net requires a supervised learning approach and hence, input-target (x, y) is determined. The objective function depends on Intersection over Union (IoU) in error back-propagation, expressed in equation (1).

$$IoU = \frac{I}{U} = \frac{y \cap \hat{y}}{y \cup \hat{y}} \tag{1}$$

Where I and U represent IoU, y and $\hat{y}$ indicate reference map and predicted one, respectively. Also, IoU is represented as False Positive (FP), True Positive (TP), and False Negative (FN), which is formulated in Equation (2).

$$\frac{I}{U} = \frac{TP}{TP+FP+FN} \tag{2}$$

IoU loss is calculated based on objective function by averaging mini batch sample at every learning procedure's updating step, expressed in Equation (3).

$$\mathcal{L}_{IoU} = 1 - IoU = 1 - \frac{1}{N}\sum_{n=1}^{N}\frac{y_n \cap \hat{y}_n}{y_n \cup \hat{y}_n} \tag{3}$$

Where N represents the batch size, y_n indicates $n-th$ target and $\hat{y}_n$ denotes $n-th$ predicted map. Then, the IoU loss into the objective function is considered, represented in Equation (4).

$$argmin_w = 1 - IoU \tag{4}$$

Where, w shows the convolutional layer weight. W-Net performs superior in distinguishing healthy and diseased regions, making it significant for agricultural applications. Figure 3 represents samples for the segmented apple leaf images. Then, the segmented images are fed into apple plant leaf disease classification.
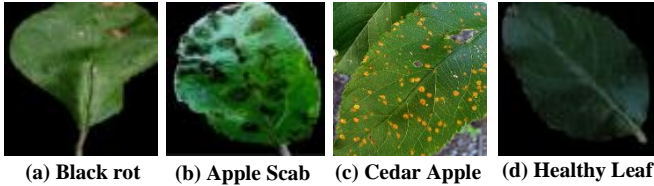

(a) Black rot   (b) Apple Scab   (c) Cedar Apple Rust   (d) Healthy Leaf

Fig. 3 Samples for the segmented apple leaf images

3.4. Classification

After segmenting the images, the MCBAM-VGG16 is employed to classify plant leaf disease. The segmented input is passed to the VGG16 [24], a Convolutional Neural Network (CNN) based model that captures more complex features in the input images. CBAM provides a channel size and spatial operation in the CNN network with 2 modules: SAM and CAM.

The CAM is established by employing the feature relations between channels, focusing on the significance of the given input. Initially, the spatial data available in feature maps is required for the process of average and maximum pooling to attain the values of FCmax and Fcavg. Then, these values are passed into the McRc x 1 x 1 network with 1 hidden layer. Channel attention is acquired and expressed in Equations (5) and (6).

$$Mc(F) = (MLP(AvgPool(F)) + MLP(MaxPool(F))) \tag{5}$$

$$Mc(F) = (W1(W0(Fcavg)) + W1(W0(Fcmax))) \tag{6}$$

Where c represents the channel number, $W0$ and $W1$ indicate a value of weight, and MLP denotes the Multi-Layer Perceptron. SAM employs property relationships between spaces to establish a spatial attention map. It focuses on where significant data is generated in input.

To achieve spatial attention maps, average and maximum pooling operations are utilized. Acquired values are integrated, and convolution like Ms(F)RH x W is performed. The spatial data is mathematically obtained in Equations (7) and (8).

$$Ms(F) = (f7x7([AvgPool(F); MaxPool(F)])) \tag{7}$$

$$Ms(F) = (f7x7([Fsavg; Fsmax])) \tag{8}$$

Where $f7x7$ represents the $7x7$ convolution process, the attention mechanism enhances the model power representation, focusing on significant features and removing unwanted ones. The CBAM generates significant features with spatial and channel axes. Both max-pooled and average-pooled features are employed simultaneously, enhancing model representation rather than utilizing each independently in the channel attention approach.

Initially, the operation of average and max-pooling are applied with the channel axis and integrated into producing an effective feature descriptor. Then, a convolution is employed to provide a spatial attention map which encodes to suppress or emphasize. The sequential arrangement generates an improved outcome than the parallel arrangement. Figure 4 represents a modified VGG16 architecture.

After each initial layer of the convolution block, the GAP layer is added, which minimizes parameter numbers by minimizing network depth. During training, it helps to avoid underfitting and overfitting issues. The GAP layer is a downsampling approach, and it is primarily utilized to increase the network's anti-distortion ability for images when retaining a sample's primary features, which minimizes parameter number. GAP layers of every convolutional block are connected after the CBAM approach, which is utilized to refine the feature maps.

Finally, the CBAM block's output is concatenated and passed into the Feed Forward Neural Network (FFNN). CBAM allows for superior discrimination among healthy and diseased leaves, which improves classification accuracy. Using MCBAM -VGG16, the model demonstrates enhanced feature refinement and attention, capturing complex patterns and disease-related anomalies in apple plant leaf images. Figure 5 indicates a flow chart for the proposed approach.
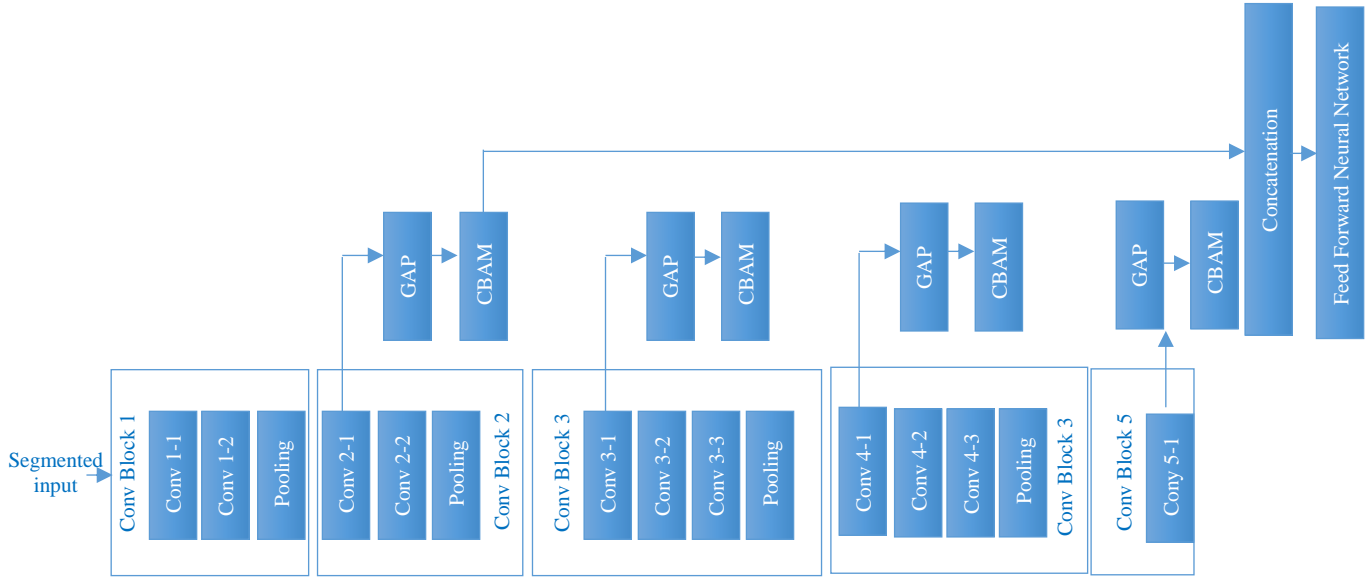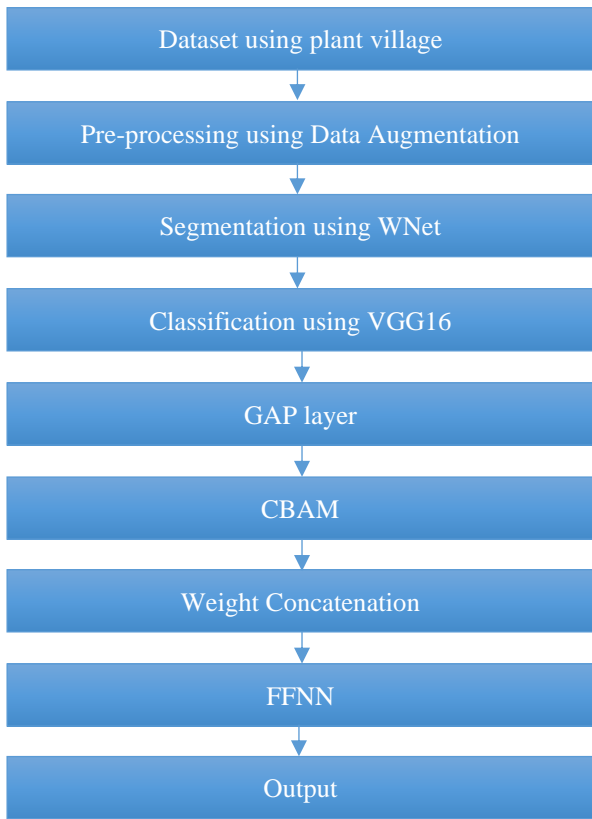
Fig. 4 Modified VGG16 architecture

Fig. 5 Flowchart for the proposed approach

4. Results

The proposed MCBAM-VGG16 is simulated using MATLAB environment with 128 GB RAM, an i5 intel processor, and Windows 10 operating system. The performance measures like accuracy, f1-score, recall, precision, Intersection Over Union (IoU), Mean IoU (MIoU),

and Dice Score Coefficient (DSC) are used to evaluate the proposed approach, which is formulated in Equations (9) to (15).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \qquad (9)$$

$$F1 - Score = \frac{2 \times TP}{2 \times TP+FP+FN} \times 100 \qquad (10)$$

$$Recall = \frac{TP}{TP+FN} \qquad (11)$$

$$Precision = \frac{TP}{TP+FP} \qquad (12)$$

$$IoU = \frac{TP}{TP+FP+FN} \qquad (13)$$

$$MIOU = \frac{1}{k+1} \sum_{i=0}^{k} \frac{TP}{FN+FP+TP} \qquad (14)$$

$$DSC = \frac{2 \times TP}{(TP+FP)+(TP+FN)} \qquad (15)$$

Where FN is the False Negative, FP is the False Positive, TP is the True Positive, TN is the True Negative.

5. Performance Analysis

The proposed MCBAM-VGG16 presents the performance analysis in Figures 6 to 8. Table 1 represents a performance analysis of different segmentation methods. Figure 6 indicates a graphical representation of different segmentation methods using the plant village dataset. The existing methods like Segnet, U-Net, and V-Net are compared with the W-Net. When compared with these existing methods, the W-Net achieves an enhanced DSC of 0.9233, as it leverages both encoder-decoder paths to refine segmentation

by progressively capturing and reconstructing fine details. This method effectively reduces the loss of spatial information, leading to superior segmentation performance.

Table 1. Performance analysis of different segmentation methods

Metrics	Segnet	U-Net	V-Net	W-Net
IoU	0.8057	0.8257	0.8634	0.8957
MIoU	0.8226	0.8468	0.8712	0.9026
DSC	0.8367	0.8536	0.8963	0.9233
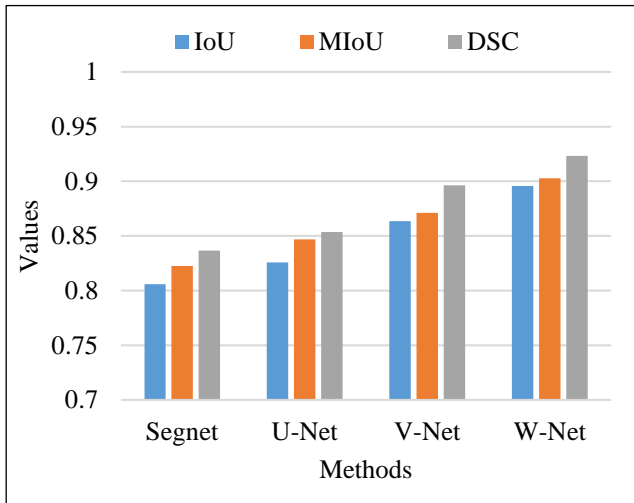
Fig. 6 Graphical representation of different segmentation methods

Table 2 indicates a performance analysis of different classification methods. Figure 7 demonstrates a graphical representation of different classification methods using the plant village dataset. Existing methods like ResNet, VGG16, and CBAM-VGG16 are compared with those of MCBAM-VGG16. The proposed approach achieves a better accuracy of 0.99 compared to the existing techniques due to its enhanced attention mechanism, which selectively focuses on significant features while suppressing unnecessary information. Additionally, GAP minimizes the number of parameters that help to avoid underfitting and overfitting issues during training.

Table 2. Performance analysis of different classification methods

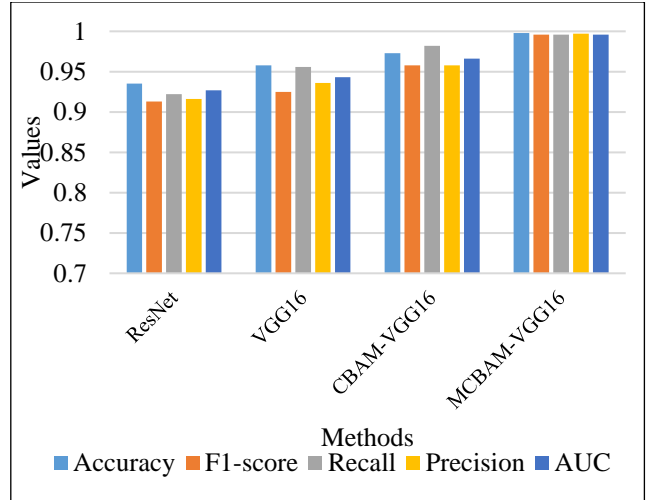| Metrics	ResNet	VGG16	CBAM-VGG16	MCBAM-VGG16
Accuracy	0.935	0.958	0.973	0.998
F1-score	0.913	0.925	0.958	0.996
Recall	0.922	0.956	0.982	0.996
Precision	0.916	0.936	0.958	0.997
AUC	0.927	0.943	0.966	0.996

Fig. 7 Graphical representation of different classification methods

Table 3 displays the performance analysis of different classes. The four classes determined are Apple scab, Black rot, Healthy leaf, and Cedar apple rust. Figure 8 indicates a graphical representation of different classes. The black rot achieves 0.998 average accuracy because of its enhanced attention mechanism, which selectively focuses on significant features when suppressing unnecessary information.

Table 3. Performance analysis of different classes

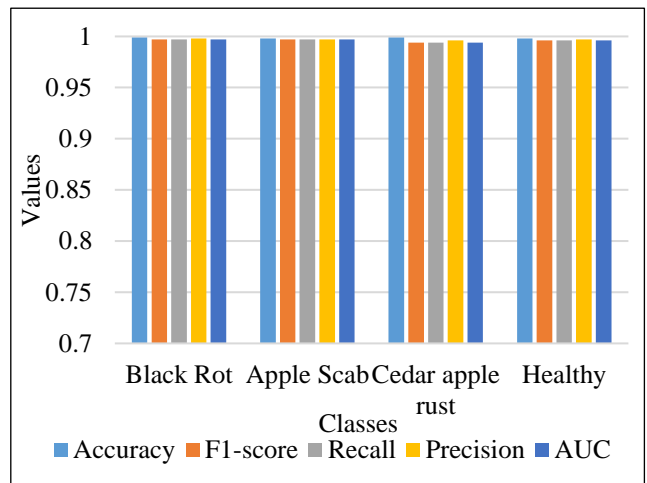| Classes	Accuracy	F1-score	Recall	Precision	AUC
Black Rot	0.999	0.997	0.997	0.998	0.997
Apple Scab	0.998	0.997	0.997	0.997	0.997
Cedar apple rust	0.999	0.994	0.994	0.996	0.994
Healthy	0.998	0.996	0.996	0.997	0.996
Average	0.998	0.996	0.996	0.997	0.996

Fig. 8 Graphical representation of different classes

Table 4 denotes a performance analysis of K-fold validation. Figure 9 indicates a graphical representation of the K-fold analysis. It exhibits a greater performance of 0.998 accuracy when K=5 fold cross-validation because of its balanced trade-off between variance and bias. The data is split into smaller pieces with more representation subsets for training and validation, with 5 folds. This makes effective performance in apple plant leaf disease by reducing variability.

Table 4. Performance analysis of K-fold validation

K Fold	Accuracy	F1-score	Recall	Precision	AUC
3.00	0.925	0.914	0.897	0.906	0.931
5.00	0.998	0.996	0.996	0.997	0.996
7.00	0.948	0.936	0.947	0.968	0.926
9.00	0.907	0.897	0.861	0.847	0.893
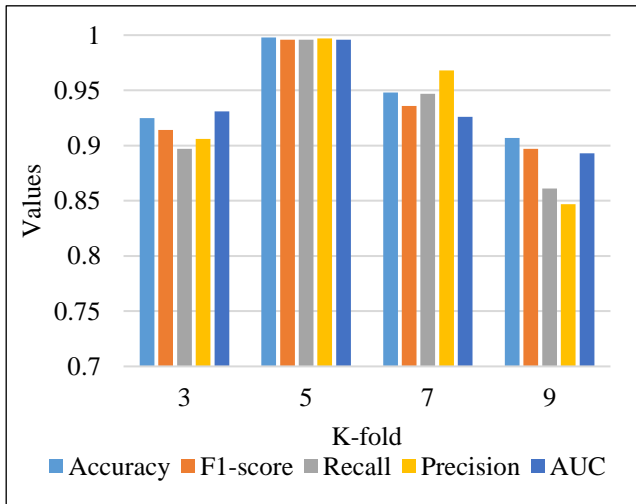
Fig. 9 Graphical representation of K-fold analysis

Figure 10 indicates a training loss vs validation loss over 10 epochs for classifying apple plant leaf disease. The rapid decrease in training loss in the initial few epochs represents that the proposed MCBAM-VGG16 approach quickly learns the patterns in training data. Also, the validation loss is minimized, showing that the model is enhanced to generalize unseen data. The validation and training loss convergence near the end of the epochs represents that the proposed approach achieved better performance on both validation and training data.

Figure 11 represents the proposed approach's training and validation accuracy over 10 epochs. Initially, the training accuracy rapidly increases by the 2nd epoch, whereas validation accuracy rises over 0.998. Both accuracies remain stable and high, representing that the model is efficiently learning to categorize diseases in apple plant leaves and perform effectively in validation data.
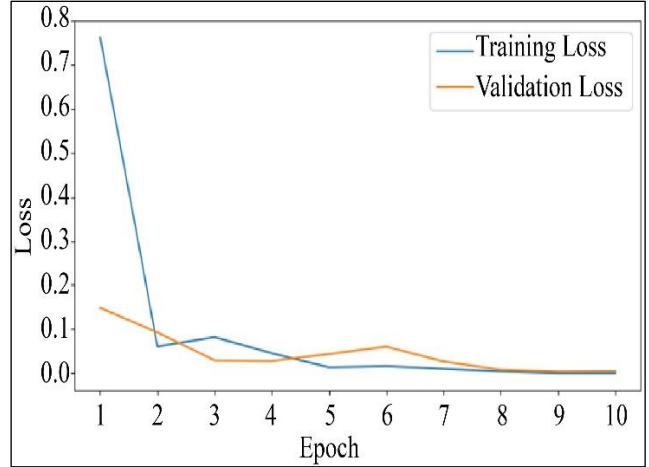

Fig. 10 Training loss vs validation loss for MCBAM-VGG16
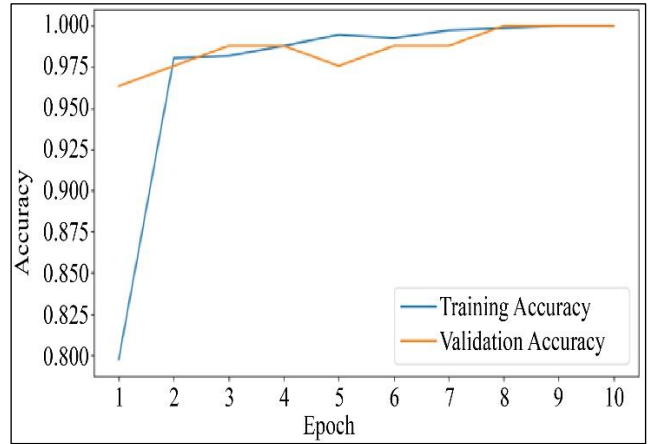

Fig. 11 Training accuracy vs validation accuracy for MCBAM-VGG16

Figure 12 demonstrates the ROC curve for classifying apple leaf disease using the proposed approach. The curve indicates True Positive Rate (TPR) vs False Positive Rate (FPR) at different threshold settings. The area under ROC represents 0.996, representing better model performance and is greatly efficient at classifying healthy and diseased apple plant leaves.
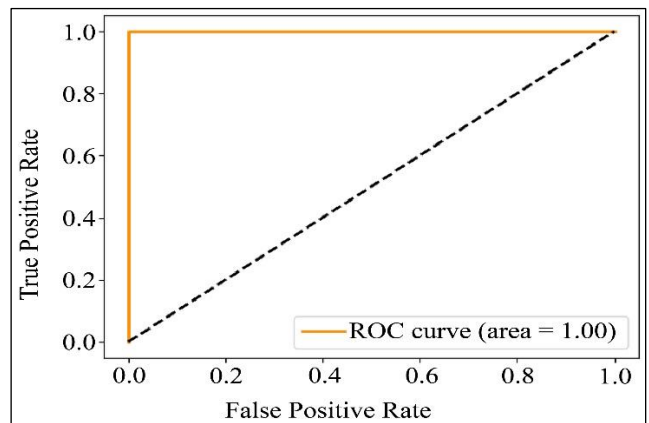

Fig. 12 ROC curve for proposed MCBAM-VGG16

Figures 13 and 15 present the training and validation loss for existing CBAM-VGG and VGG. Figures 14 and 16 indicate training and validation accuracy for existing CBAM-VGG and VGG. When compared to existing approaches, the proposed approach achieves a better performance.
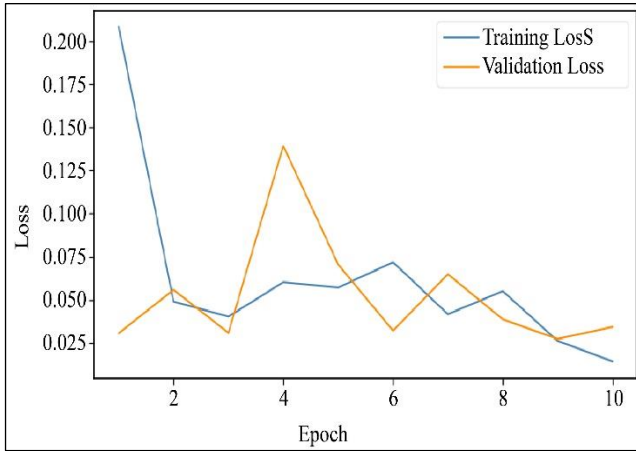

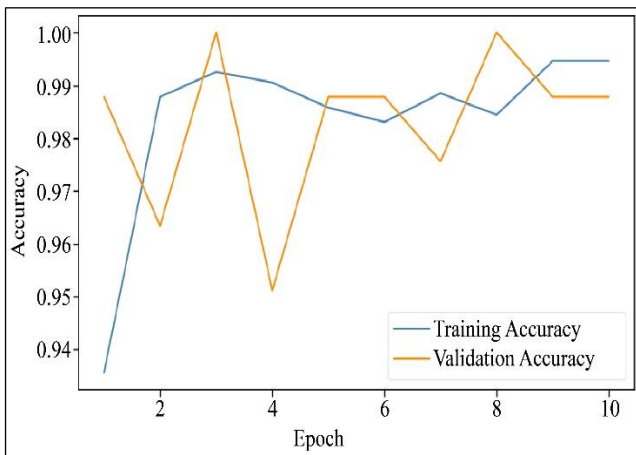**Fig. 13 Training loss vs validation loss for existing CBAM-VGG**

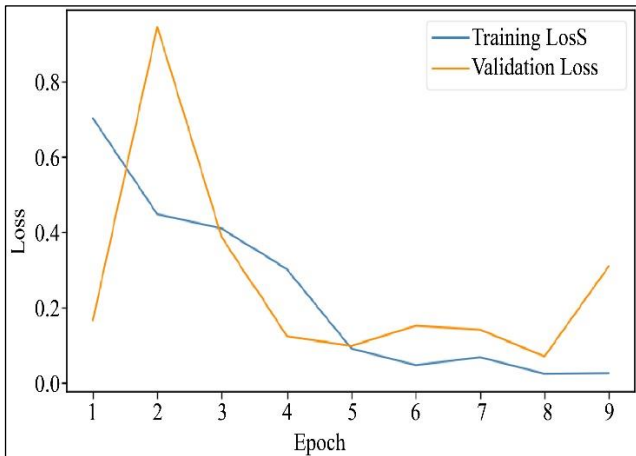**Fig. 14 Training accuracy vs validation accuracy for existing CBAM-VGG**

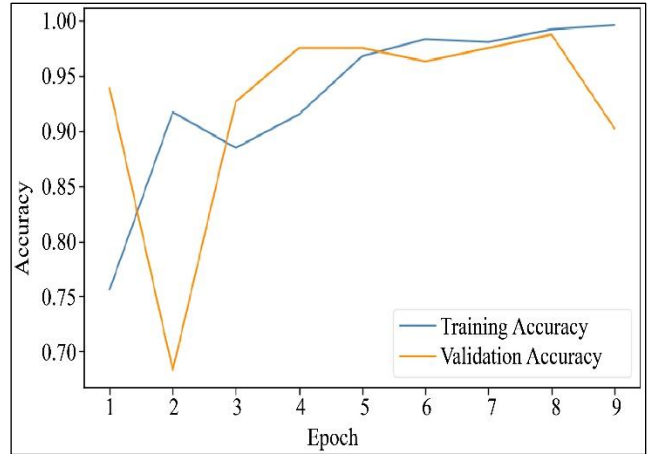**Fig. 15 Training loss vs validation loss for existing VGG**

Fig. 16 Training accuracy vs validation accuracy for existing VGG

Figure 17 illustrates a confusion matrix for four classes for the proposed approach. The predicted labels are represented on the x-axis, and true labels are indicated on the y-axis. The diagonal elements indicate correct predictions for each class with values of 55 for cedar apple rust, 126 for black rot, 32 for healthy leaf, and 124 for apple scab. The proposed MCBAM-VGG16 effectively performs with high accuracy for every class, representing minimal off-diagonal values.
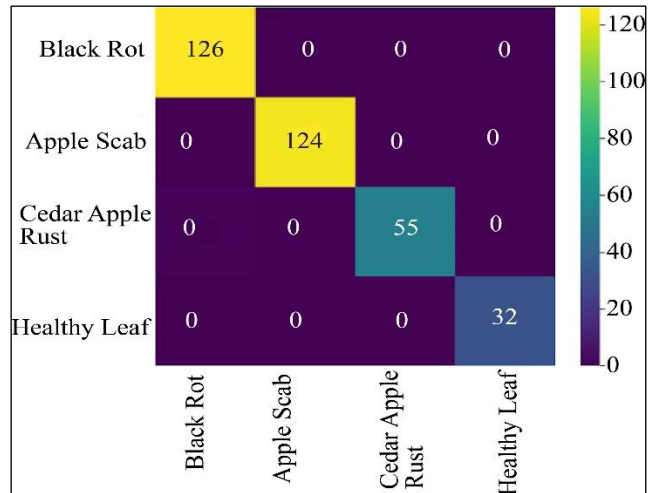

Fig. 17 Confusion matric for four classes

6. Comparative Analysis

Table 5 displays a comparative analysis of existing techniques using the plant village dataset. The existing techniques like Conv-3 DCNN [16], improved DCNN [17], RANSAC [18], and Deep feature +LTP [19] are compared with the proposed approach. The accuracy value of the existing method [19], 98.80%, is converted into decimal point 0.988 according to the proposed approach performance. In relation to these existing techniques, the proposed approach attains 0.998 accuracy as it optimizes the network's ability to capture complex patterns, improving overall performance and generalization.

Table 5. Comparative analysis of existing techniques using plant village dataset

Methods	Accuracy	F1-score	Recall	Precision	AUC
Conv-3 DCNN [16]	0.98	0.97	0.97	0.98	0.99
Improved DCNN [17]	0.993	0.99	0.993	0.99	N/A
RANSAC [18]	0.932	N/A	N/A	N/A	0.903
Deep Feature +LBP [19]	0.988	N/A	N/A	N/A	N/A
MCBAM-VGG16	0.998	0.996	0.996	0.997	0.996

7. Discussion

The advantages of the proposed MCBAM-VGG16 and the limitations of existing techniques are discussed in this section. The limitation of existing techniques like Conv-3 DCNN [16] is that it does not capture complex patterns due to shallow architecture, which minimizes classification accuracy. On the other hand, the Improved DCNN [17] has difficulty generalizing across various environmental conditions and variances in leaf images, which affect classification performance due to inconsistent leaf appearance. RANSAC [18] struggles with highly complex or overlapping leaf disease patterns as it depends on fitting models to a subset of data, which does not accurately indicate the entirety of complex patterns, resulting in misclassification.

Lightweight DCNNs [19] have a limited capacity to capture complex features because of simplified architectures, leading to decreased classification accuracy. CapsNet and ResNet [20] struggled with constrained training data due to their deep architectures, which resulted in overfitting or generalization issues with insufficient data.

The proposed MCBAM-VGG16 overcomes the limitations of the existing methods by adding GAP and CBAM after the first convolutional layer, which minimizes the number of parameters and enables it to be computationally efficient. CBAM enhances feature representation by focusing on significant features through the channel and spatial attention mechanisms. These modifications assist in balancing the model's complexity, reducing overfitting and underfitting issues during training, consequently giving rise to commendable generalization ability in apple plant leaf disease. In contrast, MCBAM-VGG16 enables more accurate attention to appropriate features by adding GAP and CBAM, which improves the model's ability to classify subtle disease symptoms in plant leaves compared to existing techniques like [16-19]. This enhances robustness and classification accuracy in classifying different plant leaf diseases.

8. Conclusion

The MCBAM-VGG16 is proposed to classify the apple plant leaf disease accurately. In VGG16, the GAP layer and CBAM are added after the first convolutional layer, which minimizes the number of parameters and helps avoid underfitting and overfitting issues during training. W-Net segments the augmented images efficiently by applying two interconnected U-Net. The initial U-Net provides a coarse segmentation, which the second U-Net then refines for higher accuracy.

This approach enhances the segmentation quality by iteratively refining the results in apple leaf disease. By performing this process, the MCBAM-VGG16 accomplishes better classification accuracy in apple plant leaf disease. The MCBAM-VGG16 accomplishes a commendable accuracy of 0.998, as opposed to the existing techniques like Conv-3 DCNN, improved DCNN, and RANSAC. In the future, a different fruit leaf disease classification will be considered to improve disease diagnosis's effectiveness across various crops.

References

[1] Sharad Hasan, Sarwar Jahan, and Md. Imdadul Islam, "Disease Detection of Apple Leaf with Combination of Color Segmentation and Modified DWT," *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no. 9, pp.7212-7224, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[2] Prabhjot Kaur et al., "Recognition of Leaf Disease Using Hybrid Convolutional Neural Network by Applying Feature Reduction," *Sensors*, vol. 22, no. 2, pp. 1-16, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[3] Jiye Zheng et al., "RepDI: A Light-Weight CPU Network for Apple Leaf Disease Identification," *Computers and Electronics in Agriculture*, vol. 212, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[4] Lili Fu et al., "Lightweight-Convolutional Neural Network for Apple Leaf Disease Identification," *Frontiers in Plant Science*, vol. 13, pp. 1-10, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[5] Emmanuel Moupojou et al., "FieldPlant: A Dataset of Field Plant Images for Plant Disease Detection and Classification with Deep Learning," *IEEE Access*, vol. 11, pp. 35398-35410, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[6] Waleed Albattah et al., "A Novel Deep Learning Method for Detection and Classification of Plant Diseases," *Complex & Intelligent Systems*, vol. 8, no. 1, pp. 507-524, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[7] Yuxi Gao et al., "Apple Leaf Disease Identification in Complex Background based on BAM-Net," *Agronomy*, vol. 13, no. 5, pp. 1-22, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[8] Saleh Albahli, and Marriam Nawaz, "DCNet: DenseNet-77-Based CornerNet Model for the Tomato Plant Leaf Disease Detection and Classification," *Frontiers in Plant Science*, vol. 13, pp. 1-19, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[9] Kangchen Liu, and Xiujun Zhang, "PiTLiD: Identification of Plant Disease from Leaf Images Based on Convolutional Neural Network," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 20, no. 2, pp. 1278-1288, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[10] K. Indira, and H. Mallika, "Classification of Plant Leaf Disease Using Deep Learning," *Journal of The Institution of Engineers (India): Series B*, vol. 105, no. 3, pp. 609-620, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[11] Adesh V. Panchal et al., "Image-Based Plant Diseases Detection using Deep Learning," *Materials Today: Proceedings*, vol. 80, no. 3, pp. 3500-3506, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[12] Poornima Singh Thakur, Tanuja Sheorey, and Aparajita Ojha, "VGG-ICNN: A Lightweight CNN Model for Crop Disease Identification," *Multimedia Tools and Applications*, vol. 82, no. 1, pp. 497-520, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[13] Gao Ang et al., "Construction and Verification of Machine Vision Algorithm Model Based on Apple Leaf Disease Images," *Frontiers in Plant Science*, vol. 14, pp. 1-13, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[14] Ruchi Gajjar et al., "Real-Time Detection and Identification of Plant Leaf Diseases Using Convolutional Neural Networks on an Embedded Platform," *The Visual Computer*, vol. 38, no. 8, pp. 2923-2938, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[15] Bui Thi Hanh, Hoang Van Manh, and Ngoc-Viet Nguyen, "Enhancing the Performance of Transferred Efficientnet Models in Leaf Image-Based Plant Disease Classification," *Journal of Plant Diseases and Protection*, vol. 129, no. 3, pp. 623-634, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[16] Vibhor Kumar Vishnoi et al., "Detection of Apple Plant Diseases Using Leaf Images through Convolutional Neural Network," *IEEE Access*, vol. 11, pp. 6594-6609, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[17] Dharmendra Kumar Mahato, Amit Pundir, and Geetika Jain Saxena, "An Improved Deep Convolutional Neural Network for Image-Based Apple Plant Leaf Disease Detection and Identification," *Journal of the Institution of Engineers (India): Series A*, vol. 103, no. 4, pp. 975-987, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[18] Yashwant Kurmi, and Suchi Gangwar, "A Leaf Image Localization Based Algorithm for Different Crops Disease Classification," *Information Processing in Agriculture*, vol. 9, no. 3, pp. 456-474, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[19] Khalid M. Hosny et al., "Multi-Class Classification of Plant Leaf Diseases Using Feature Fusion of Deep Convolutional Neural Network and Local Binary Pattern," *IEEE Access*, vol. 11, pp. 62307-62317, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[20] Xin Zhang et al., "A Plant Leaf Disease Image Classification Method Integrating Capsule Network and Residual Network," *IEEE Access*, vol. 12, pp. 44573-44585, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[21] [Online]. Available: https://www.kaggle.com/datasets/abdallahalidev/plantvillage-dataset

[22] Qing Yang, Shukai Duan, and Lidan Wang, "Efficient Identification of Apple Leaf Diseases in the Wild Using Convolutional Neural Networks," *Agronomy*, vol. 12, no. 11, pp. 1-20, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[23] Ruisong Zhang et al., "W-Net: Structure and Texture Interaction for Image Inpainting," *IEEE Transactions on Multimedia*, vol. 25, pp. 7299-7310, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[24] Shagun Sharma et al., "A Deep Learning Based Convolutional Neural Network Model with VGG16 Feature Extractor for the Detection of Alzheimer Disease Using MRI Scans," *Measurement: Sensors*, vol. 24, pp. 1-8, 2022. [CrossRef] [Google Scholar] [Publisher Link]