Applied Fruit Science (2025)67:18
https://doi.org/10.1007/s10341-024-01239-w

ORIGINAL ARTICLE

E-AppleNet: An Enhanced Deep Learning Approach for Apple Fruit Leaf Disease Classification

Ameen Banjar[1] · Ali Javed[2] · Marriam Nawaz[2] · Hussain Dawood[3]

Received: 8 April 2024 / Accepted: 21 October 2024

Abstract
Apples are among the most widely consumed fruits globally due to their numerous health benefits. However, their production is significantly impacted by various leaf diseases. Accurate and timely detection of these diseases is challenging due to the similarities between healthy and diseased leaves, as well as issues like image blurring, clutter, and varying conditions. This paper presents E-AppleNet, an advanced approach based on EfficientNetV2, which incorporates attention mechanisms and additional dense layers at the end of the model structure to enhance disease classification. To tackle class imbalance and prevent model overfitting, we utilize transfer learning and multi-class focal loss. Our model, tested on the PlantVillage dataset, a complex repository with real-world images, achieves a remarkable 99% accuracy. Additionally, we generated the heatmap visualizations, which confirm the model's robustness in handling diverse image distortions. This method provides a reliable solution for automated apple disease detection, with potential benefits for precision agriculture.

Keywords Agriculture · Attension mechanism · Computer vision · EfficientNet · Plant abnormalities

Introduction

The apple is regarded to be among the most significant plant fruits because of its ranking as the second-most produced fruit worldwide (Dubey and Jalal 2016). Apple's output peaked in 2017 at 83.10 thousand tons, and the fruit was highly utilized globally (Ayaz et al. 2021; Omrani et al. 2014). Apples are widely consumed because they are inexpensive and have a variety of beneficial medicinal attributes, such as a high fiber, nutrient, vitamin, and oxidant composition. Furthermore, due to their taste, they can be consumed naturally or used to create a variety of derivative items. According to estimates, 33% of the apple fruit grown globally is converted into items such as drinks, ciders, sauces, alcohol, and dry apples. However, severe damage has been witnessed in the yield of apple fruit in the last few years because of the occurrence of various diseases. The National Institute of Food and Agriculture (USA) estimates that insects and numerous field plant illnesses cause the loss of 35–50% of the globe's agricultural production. Particularly at a crucial moment when requirements and consumption are sharply increasing globally, it is projected that this concern will destroy 15–25% of Asia's annual crop yields. Therefore, to practice sustainable farming, it is critical to track crops and identify diseases (Vishwakarma et al. 2024). Diagnosis of plant pathogens can play a crucial part in agricultural protection in the early phases. The infected apple fruit plants can be easily distinguished from the others with a simple examination of the human eye. However, a personal assessment is very sensitive to subjectivity and susceptible to errors too.

Traditional techniques, including polymerase chain reaction (PCR), necessitate extensive molecule sampling, making them inefficient from a monetary perspective (Bracino et al. 2020). The vast of such production areas reside in underdeveloped countries where most people are unable to employ such costly approaches for field crop disease inspection. Artificial intelligence (AI) has recently been employed to assist professionals in the autonomous recognition of illnesses that afflict crops and trees. AI-based techniques are quicker, affordable, and more effective. The de-

✉ Ameen Banjar
abanjar@uj.edu.sa

[1] College of Computer Science and Engineering, Department of Information Systems and Technology, University of Jeddah, Jeddah, Saudi Arabia

[2] Department of Software Engineering, University of Engineering and Technology-Taxila, Taxila, Pakistan

[3] School of Computing, Skyline University College, University City Sharjah, 1797 Sharjah, United Arab Emirates

Published online: 03 January 2025

velopment of computerized approaches for diagnosing crop viruses is being aided by the quick development of machine vision technologies. Several conventional machine learning (ML) predictors like K-nearest neighbors (KNN) and support vector machines (SVM) are extensively used by existing methods. Even though conventional ML approaches need a minimum amount of data for tuning the classifiers, these techniques typically include non-automatic pattern selection process steps when deployed to complicated samples. Moreover, there is always a compromise in the classification effectiveness and time complexity of such techniques (Li et al. 2024). Contrarily, deep learning (DL) is a specialized type of ML that enables computers to learn automatically and increase generalization ability (Nawaz et al. 2023). Because of the emergent properties of DL approaches, these methods are highly employed in the area of image processing and computer vision. Various DL frameworks like convolutional neural networks (CNNs) and recurrent neural networks (RNNs) have recently been extensively admired in the area of food security. Several studies have discussed various DL approaches for tasks like estimating grain size, locating plant heads, quantifying fruits, etc. These techniques can display outstanding detection performance while putting forth the least amount of computational burden since they can make use of the architectural and morphological details from the examined samples.

Although scientists have put immense effort into recognizing and classifying the infected regions of apple plant diseases (Azgomi et al. 2023; Gong and Zhang 2023; Liu et al. 2023; Hassan et al. 2023; Chen et al. 2023), there is still a gap for performance improvement. The classification of diseases at the early stage is a complicated task due to the existence of massive similarities in the normal and diseased portions of apple plant leaves. Moreover, the volume, shape, and orientation of infected areas also vary a lot. Apart from this, various types of image distortions also introduce a huge hindrance to the accurate recognition of infected areas of apple plant leaves. In this work, an effort has been put to overcome the mentioned problems by presenting a robust DL framework called E-AppleNet. More clearly, an improved EfficientNetV2 approach is presented by using the attention mechanism (AM) and adding the extra dense layers at the end of the model construction. We have trained the improved EfficientNetV2 model on the apple leaf images from the PlantVillage repository and given it the name E-AppleNet, which can classify the samples into four types: healthy, apple scab, rot, and rust, respectively. The nominative contributions of the suggested approach are listed as:

- A novel DL framework called E-AppleNet is proposed that is capable of extracting a more resilient set of image characteristics and exhibits better apple leaf disease classification results.
- The framework incorporated the pixel and channel attention technique during the keypoints engineering stage, which progresses its ability to join the cross-channels and space-wise features to boost the recognition of apple fruit abnormalities in complex scenarios.
- We employed transfer learning and introduced multi-class focal loss to address the issues of class imbalance and model over-fitting.
- We introduced three extra dense layers in our DL framework to better nominate the most relevant features.
- The recommended approach accurately classifies apple leaf abnormalities regardless of the difficult external conditions such as distortion, imbalanced exposure, and fluctuations in the contour, hue, and positioning of infected signs.

Method

In this part, the details of the proposed strategy namely the E-AppleNet are given by discussing the architectural description of its modules used to classify the numerous types of apple fruit diseases.

E-AppleNet

For the accurate identification of different apple fruit leaf abnormalities, we have proposed a DL framework called the E-AppleNet, a light framework designed by employing the Efficientnet-V2 network with an AM strategy and added dense layers. The E-AppleNet approach is proficient in capturing the relevant information of diseased areas by utilizing the AM strategy that permits it to emphasize the infected areas of the plant while the added dense layers optimize its power to accurately classify the input samples by passing the most relevant features to the classification layer. The elaboration of the E-AppleNet is described in Fig. 1. The model consists of a CNN unit named the EfficientNet-V2, a spatial-channel AM module next linked to a layer called adaptive average pooling (AAP), and lastly, the classification module comprises dense layers and a prediction unit. The elaboration of the E-AppleNet is described in Fig. 1.

CNN Module

In the presented work, we have employed the EfficientNet-V2 approach as the base CNN module which belongs to the family of EfficientNet DL frameworks (Tan and Le 2019)

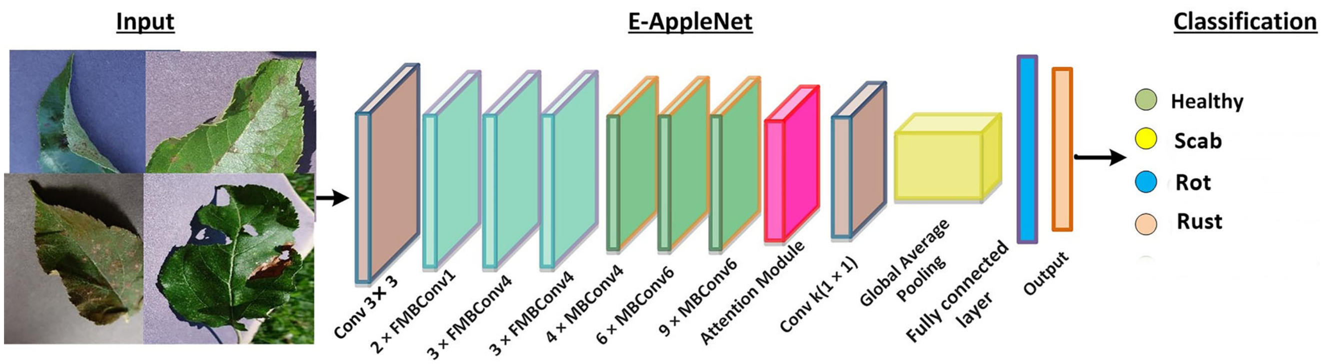


Fig. 1 Structure of presented E-AppleNet framework

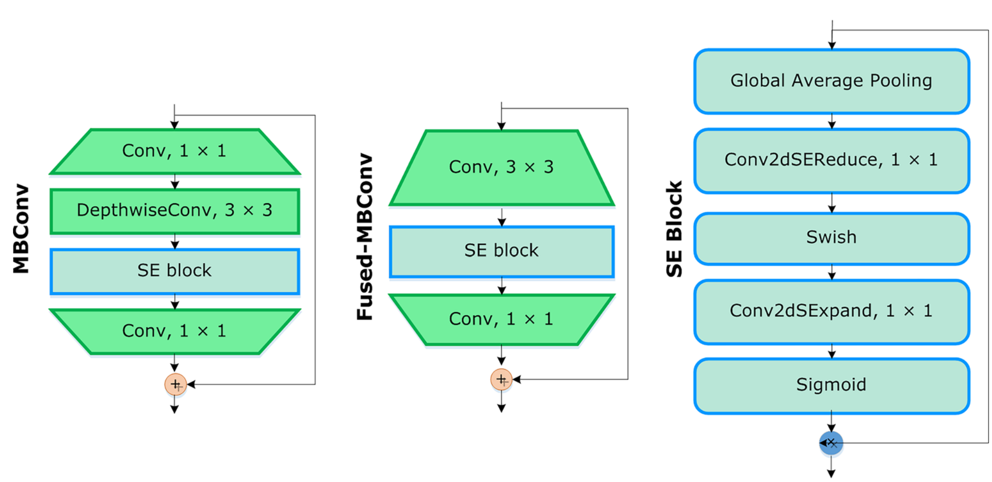
and is famous for its lightweight architecture (Tan and Le 2021).

The major reason to nominate the EfficientNet-V2 approach as the base CNN approach for recognizing the various disorders of the apple fruit plant leaves is due to its shallow structure, less training, and inference time. The EfficientNet-V2 framework contains the MBConv and Fused-MBConv units employing squeeze and excitation (SE) optimization to build channel-oriented attention and boost the frameworks' keypoints nomination ability. A visual description of the SE, MBConv, and Fused-MBConv (FMBConv) is provided in Fig. 2. The inclusion of the Fused-MBConv at the starting phases of the EfficientNet-V2 results in the high proficiency of the model parameters which conveniently fastens the training procedure in comparison to the EfficientNet-V1 model (Tan and Le 2021). The internal structure of the MBConv unit follows a 1×1 conv layer along with a depth-oriented convolution having a window size of 3×3, whereas for the Fused-MBConv unit, the traditional 3×3 conv layers are utilized. A 1×1 pixel-oriented conv layer is utilized next to the SE unit for both MBConv and Fused-MBConv units to regulate the channel sizes. Afterward, a drop linkage is used by employing a skip connection from the input (Fig. 2). Further, the EfficientNetV2 primarily uses the Swish activation function in its convolutional layers and MBConv blocks, enhancing gradient flow for better performance. Additionally, the Sigmoid activation is applied in the SE blocks to generate attention weights, improving channel-wise feature importance.

AM Unit

The feature engineering phase of the CNN module generates an extensive set of unwanted details of the examined samples, like background. Such details greatly impacted the correct recognition of the various sample abnormalities. To overcome this challenge, we have introduced the AM unit in the E-AppleNet that focuses the model's attention on the relevant sample characteristics by overwhelming clutter and background details and ultimately increases the recall power of the approach. The AM unit works by assigning varied weight values to the computed features from the CNN module, like assigning a high-weight value to the lower-level features accompanying the infection details and using a small weight value for the background details. The spatial attention mechanism is famous for identifying the location of the target (that is, infected regions of plant leaves for our case) in the keypoints map, whereas the chan-

Fig. 2 Inner details of MBConv, Fused-MBConv, and the SE units



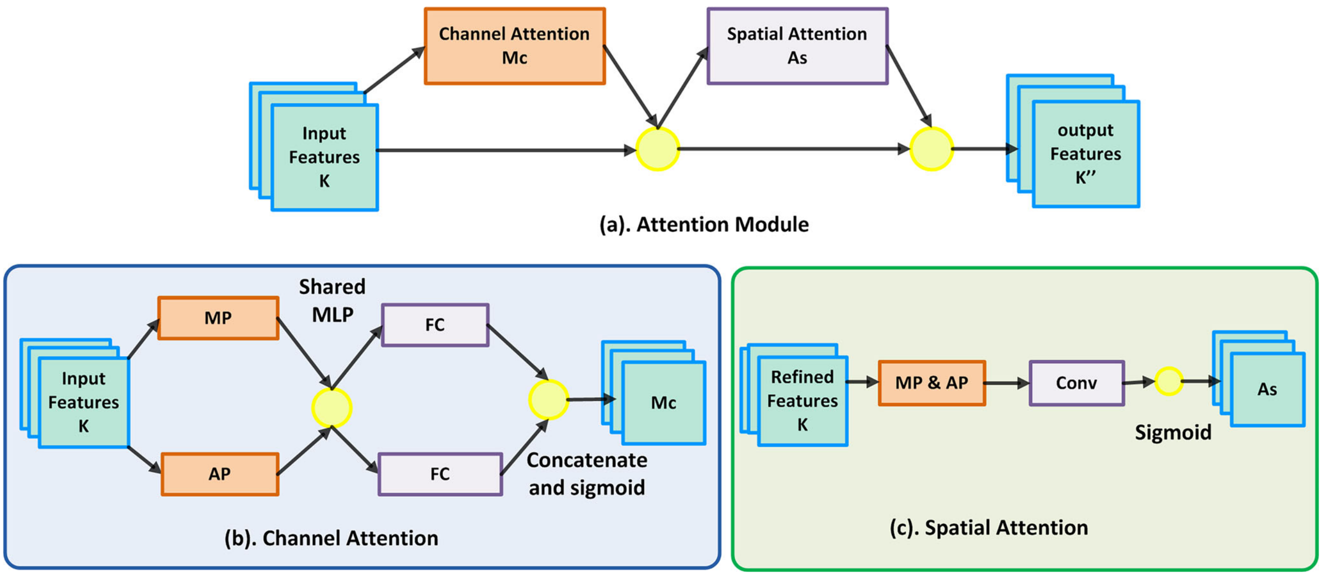


Fig. 3 Architectural details of the attention module strategy comparing the channel attention and units

nel attention approach robustly identifies a particular target through numerous keypoints maps. The incorporation of both results in a more optimized performance of the framework by showing effective results for real-world scenarios. Therefore, we have employed the spatial-channel AM module in the E-AppleNet framework to improve the feature nomination capacity of the approach. A visual depiction of the spatial-channel AM unit is provided in Fig. 3.

Let us assume that $K \in L^{c \times w \times h}$ is presenting an intermediary feature map along having the sizes $c \times w \times h$ computed by the CNN unit (EfficientNet-V2) and propagated to the spatial-channel AM module. The channel attention (CA) unit of the AM strategy produces a 1-D channel keypoints map designated by $M_c \in L^{1 \times 1 \times c}$, while the spatial attention (SA) unit generates a two-dimensional spatial keypoints map $M_s \in L^{1 \times w \times h}$. The mathematical description of the spatial-channel AM strategy is elaborated in Eq. 1.

$$\mathrm{K} = \mathrm{CA(K)} + \mathrm{SA(K)} = \mathrm{M_c(K)} * \mathrm{K} + \mathrm{K} * \mathrm{M_s}\,(\mathrm{K}) \tag{1}$$

In Eq. 1, $M_{s(K)}$ represents the spatial attention map applied to the feature map K. This function generates a 2-dimensional spatial attention map which emphasizes the important spatial regions within K. The output of $M_{s(K)}$ is essentially a matrix that highlights key spatial areas in the feature map K, aiding the model in focusing on crucial regions that contribute to the classification task. The * sign indicates the dot product of elements. For input features computation K, the max, and average pool layers are incorporated in the channel and spatial attention modules, respectively. Within each CA unit, the values of both pool layers are combined to generate the resultant feature map as described in Eqs. 2, and 3 respectively.

$$\mathrm{M_c(K)} = \sigma(\mathrm{MLP(MaxPool(K))} + \mathrm{MLP(AvgPool(K))}) \tag{2}$$

$$= \sigma\left(\mathrm{l_1}\left(\mathrm{l_0}\left(\mathrm{K^c_{max}}\right)\right) + 1\left(\mathrm{l_0}\left(\mathrm{K^c_{avg}}\right)\right)\right) \tag{3}$$

Here, σ signifies the sigmoid activation function, while the l_1 and l_0 denote learning weights, the K^c_{max} and K^c_{avg} are features obtained from the average and max pool layers. Moreover, MLP is denoting neurons. The SA unit produces the keypoints map by joining the resultant features obtained from CA. The feature maps against average (AP) and max pool (MP) layers designated by K^s_{avg} and K^s_{max} along each channel focus on disease-specific information, which is joined in this phase, and a conv layer is employed to accomplish the convolution function. The internal working of the SA module is depicted in Eqs. 4 and 5.

$$\mathrm{A_s}\,(\mathrm{K}) = \sigma(\mathrm{o}^{7\times7}\,([\mathrm{AvgPool}\,(\mathrm{K})\,;\mathrm{MaxPool}\,(\mathrm{K})])) \tag{4}$$

$$\mathrm{A_s}\,(\mathrm{K}) = \sigma(o^{7\times7}\left(\left[\mathrm{K^s_{avg}}; \mathrm{K^s_{max}}\right]\right)) \tag{5}$$

Here o presents the convolutional method having a window size of 7×7. Further, $\mathrm{A_s(K)}$ shows the SA map attained by merging the resultant keypoints attained from the CA phase.

Next, the AAP layer acquires the linkage among different channels recursively and modifies the keypoints map to $1 \times 1 \times 1280$. Then, the computed features are passed to the added three dense layers which assists in the better selection of the highly related features which are later classified

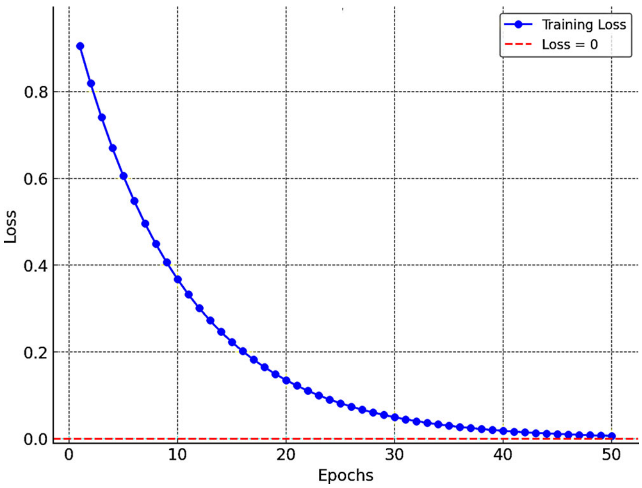


Fig. 4 Loss curve of the work

by the prediction unit employing the softmax classification method.

Loss Method

We have used the focal loss in our model training phase which assists us to overcome the category imbalance problem and improves the prediction performance. Focal loss addresses class imbalance by reducing the weight of easy examples, thus allowing the model to prioritize learning from underrepresented or challenging samples. The mathematical details of the focal loss are mentioned in Eq. 6.

$$L = -\sum_{j=1}^{t} \eta_k \left(1 - o_j\right)^{\gamma} \log(o_j) \tag{6}$$

Here, t denotes the total classes, and o shows the probability distribution of the prediction. The η and γ denote the weighting and modulating parameters with a value of 1 and 1.5 for our case. Further, the loss curve is provided in Fig. 4 which shows effective model learning during the training phase.

Results and Discussion

This part describes the details of the dataset, different experiments, and results, along with the discussion. Further, a thorough investigation accompanying various latest approaches is performed to prove the efficacy of the model.

Dataset

To validate the classification results of the suggested model, a standard and online accessible data sample designated as the PlantVillage (Hughes and Salathé 2015) is applied in this technique. This data sample contains 54,306 pictures of 14 categories. The current work is focused on categorizing the leaf abnormalities of only apple fruit. For Apple, there are a total of 3168 samples with the distribution of 631, 621, 271, and 1645 images from the apple scab, apple rot, apple rust, and healthy classes, respectively. For our work, 70% of samples, which are 2217 images, are used for model training, while the remaining 951 are used for model testing. Further, all images are set to a resolution of 224 as per model requirements. The main purpose of using this dataset for results evaluation is that it includes images that differ in mass, architecture, dimensions, and alignment of both leaves and diseased areas. A few samples are reported in Fig. 5.

Fig. 5 Some images of apple fruit leaves from the PlantVillage

Proposed Work Assessment

This section comprises the evaluation of the presented approach where we performed two types of experiments in which the first test elaborates the category-wise model's performance discussion, while the other discusses the classification values attained on the entire data sample.

Group-wise Performance

The major characteristic of a robust automated plant disease recognition approach is measured in terms of its capability to discriminate the various kinds of infections from a given type of crop. To test the skill of the E-AppleNet in separating the normal samples from the apple leaf images suffering from scab, rot, and rust diseases, we have performed an evaluation.

In the first phase, we calculated the precision and recall scores of the E-AppleNet in terms of all four categories, and the obtained results are given in Fig. 6. The scores shown in Fig. 6 prove the supremacy of the proposed methodology in recalling all the classes of the apple leaves with high recognition ability. Our E-AppleNet approach obtains precision scores of 99.50%, 98.98%, 98.97%, and 98.95% for the apple healthy, scab, rot, and rust categories. Further, in the aspect of the recall, the E-AppleNet approach shows recall scores of 99.26%, 98.81%, 98.52%, and 98.61% for the healthy, scab, rot, and rust forms of apple leaves.

In the next phase, the categorization performance of the E-AppleNet is computed in terms of F1 and error rates against all four types of apple leaves. The obtained F1-scores against all four classes of apple fruit leaves are given in Fig. 7 from where it is quite evident that the E-AppleNet approach has reached an average value of 98.95% with the uppermost and smallest error values of 0.62, and 1.26% against the apple healthy and rot classes, respectively.

Next, the accuracy values computed for each class are reported in Fig. 8 in the form of box plots as these graphs are empowered to exhibit the detailed elaboration of acquired performance values by depicting the extreme, lower, and average scores. The group-oriented accuracy numbers mentioned in Fig. 8 show the efficacy of the E-AppleNet approach for cataloging the diseased regions of apple fruit leaves. More clearly, the E-AppleNet model acquires the accuracy numbers of 99.51%, 98.88%, 98.89%, and 98.72% against the apple healthy, scab, rot, and rust groups.

Lastly, the confusion matrix is reported for the E-AppleNet approach as this graph is capable of presenting the class-wise recall power of an approach by showing the results in the aspect of true positive rates (TPR). The attained confusion matrix is presented in Fig. 9 indicating that the E-AppleNet approach is effective in recognizing all four types of apple fruit leaves. The E-AppleNet model obtains an average TPR of 98.80% along with the largest and smallest error rates of 0.1 and 0.8% for the healthy and rot classes of apple plant leaves.

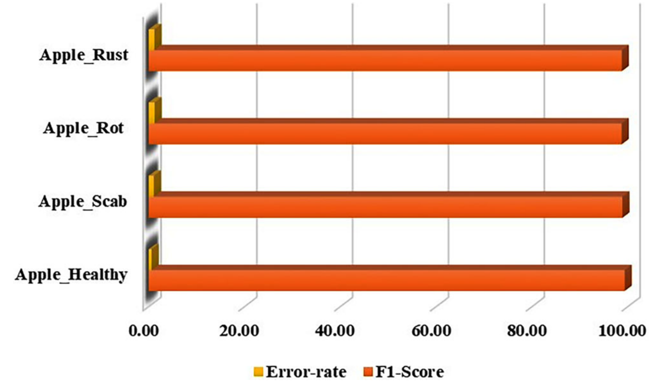


Fig. 7 Category-oriented F1 and error-rates results

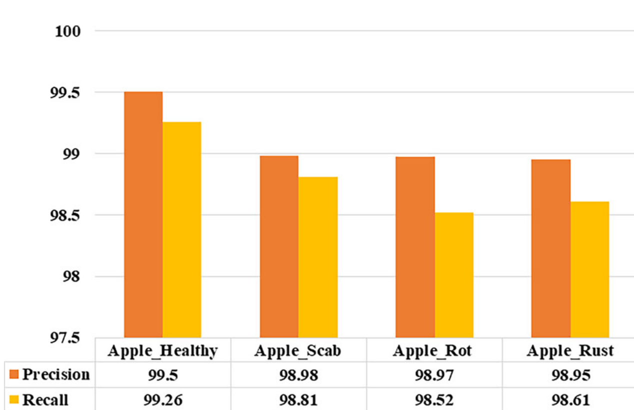


Fig. 6 Category-oriented precision and recall results

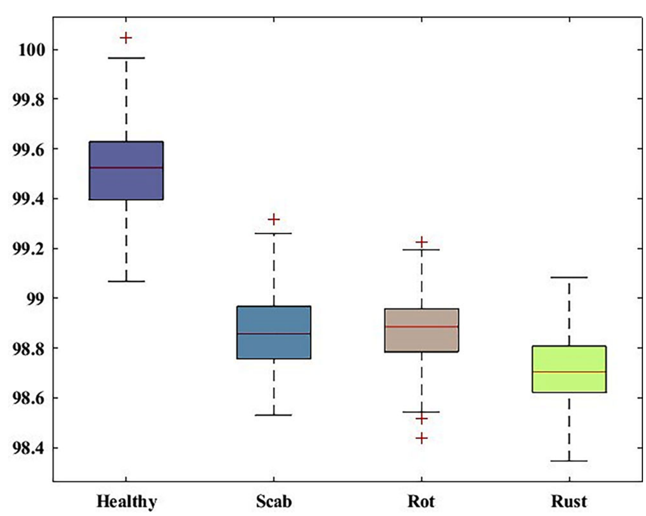


Fig. 8 Category-oriented accuracy results of the E-AppleNet

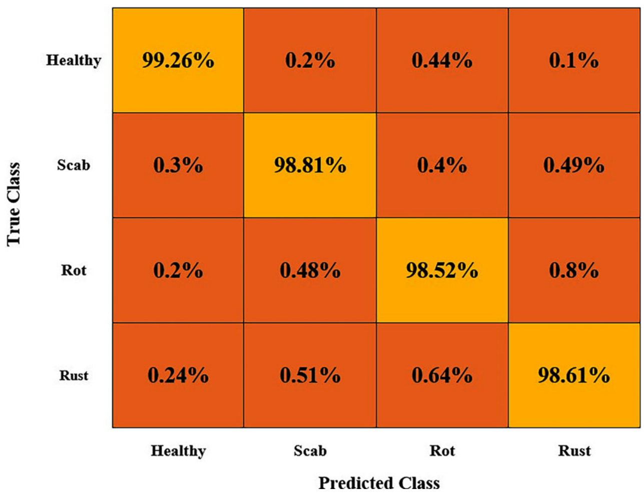


Fig. 9 Obtained confusion matrix

Overall Performance Discussion

In this phase, we have provided an elaboration of the overall performance of the E-AppleNet approach. For this, we have computed the performance of the model on the entire data sample as given in Fig. 10 which proves the effectiveness of the E-AppleNet approach. The E-AppleNet framework acquires a precision score of 99.10%, along with a recall of 98.80%. Moreover, the suggested work has attained an accuracy number of 99%, with an F1 measure of 98.95% which proves the proficiency of our model in robustly executing the task of distributing the leaves images of apple fruit in the respective classes.

Moreover, we have visualized the inner functionality of the E-AppleNet approach by generating the heatmaps which assist the readers to check whether the model is taking the right infected areas to perform the classification task. For heatmap analysis, Grad-CAM was employed to visualize the model's focus on specific regions of apple plant leaf images during disease classification. Grad-CAM generates heatmaps by utilizing the gradients of the target class flowing into the last convolutional layer, highlighting important areas for prediction. These heatmaps reveal the regions most associated with disease symptoms, helping identify whether the model focuses on the correct parts of the leaf (e.g., affected spots or textures). A few samples are shown in Fig. 11. The red color in Fig. 11 shows the infected regions of the examined samples which signifies that the E-AppleNet approach is considering the exact diseased region to perform the cataloging of apple fruit leaf disorders. The basic reason for the better explainability power of the E-AppleNet model is its high recall ability which enables it to recognize all types of leaf disorders and successfully distribute the input samples into relevant classes.

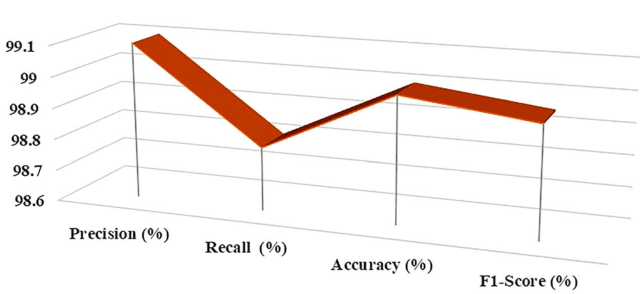


Fig. 10 Overall performance results of the E-AppleNet

Comparison with the Base Approaches

Here, a performance comparison of the introduced work is conducted in contrast to different base models like EfficientNet frameworks with base B0 to B5 (Table 1). The values stated in Table 1 indicate that the E-AppleNet model is more reliable in terms of all performance computing metrics as compared to its base approaches and takes less time, only 32.01 ms, to test a given sample. More clearly, for the accuracy, F1, precision, and recall, the E-AppleNet approach attains scores of 99%, 98.95%, 99.10%, and 98.80%, which are the largest in contrast to all other peer frameworks. The lowest classification results are exhibited by the EfficientNet-B0 model with an accuracy of 95.88%, while the second minimum score is reported by the EfficientNet-B1 approach with an accuracy value of 96.59%. The EfficientNet-B5 approach secures comparable classification results with an accuracy of 98.69%. Clearly, for the precision metric, the base models have shown an average value of 97.49%, which is 99.10 for the E-AppleNet approach, with a performance enhancement of 1.61%. Then, in the aspect of recall and accuracy, the E-AppleNet has reported performance gains of 2.06 and 1.58%. Lastly, for F1, the base works have secured an average score of 97.12%, which is 98.95% for the suggested E-AppleNet model, and secured a performance enhancement of 1.83%.

The distinguishing factor that causes the E-AppleNet approach to perform effectively from the base approaches is the inclusion of the AM unit in the EfficientNet-V2 approach, which allows the approach to compute the low-level attributes of the examined images to robustly identify the various types of infected regions from the leaves of the apple fruit. Further, the introduced dense layers right before the classification module further incorporate the propagation of the most reliable sample features to perform the categorization job.

Comparative Evaluation with DL Frameworks

We executed a comparison of the proposed work with numerous DL frameworks which have been heavily explored in history for the categorization of plant leaf infec-

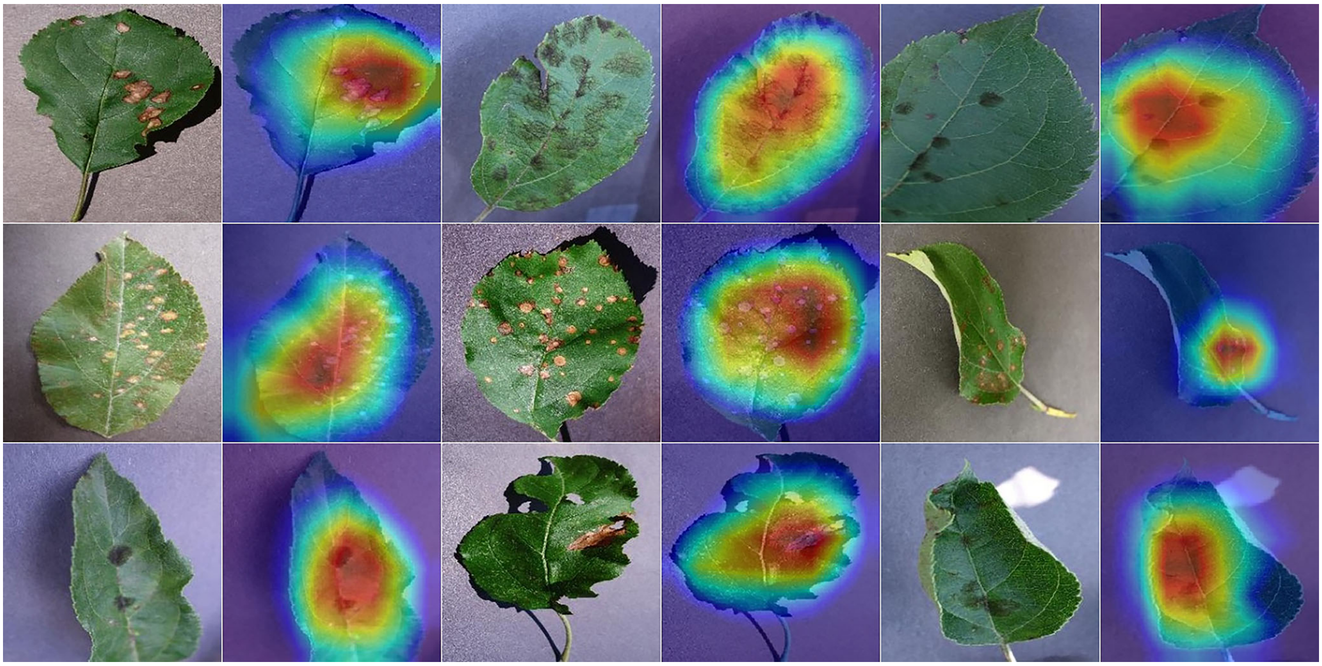


Fig. 11 Heatmaps obtained for the E-AppleNet approach

Table 1 Performance comparison with the base frameworks

Framework	Test time/image (ms)	Precision (%)	Recall (%)	Accuracy (%)	F1 (%)
EfficientNet-B0	46.22	95.98	94.41	95.88	95.19
EfficientNet-B1	49.41	96.65	95.92	96.59	96.28
EfficientNet-B2	49.82	97.03	96.68	97.11	96.85
EfficientNet-B3	55.01	97.68	97.44	97.56	97.56
EfficientNet-B4	57.09	98.75	98.09	98.72	98.42
EfficientNet-B5	63.06	98.87	97.91	98.69	98.39
Proposed	*32.01*	*99.10*	*98.80*	*99*	*98.95*

tions. For this, we have taken the following approaches: VGG19, ResNet152, DenseNet201, ResNet50, VGG16, InceptionV3, and Xception for classifying the plant abnormalities of the apple fruit as mentioned in Vishnoi et al. (2022) and obtained comparison is given in Table 2.

The scores given in Table 2 prove that the E-AppleNet approach is more effective and efficient as compared to all other DL approaches. Clearly, in the aspect of the model complexity, the E-AppleNet approach comprises the smallest range of framework parameters and attained the lowest test time of 13 s. Comparatively, the VGG19 approach comprises the highest parameters and a training time of 25 h. Moreover, for the classification results, again the proposed work is more effective in aspects of all results measuring parameters and shows an average accuracy rate of 99%. Our approach provides a proficient compromise between the framework computing burden and categorization results because of its less deep model architecture and improved ability to compute a powerful set of sample characteristics which provides both the computational and classification overhead to our technique as compared to other DL frameworks.

Comparative Evaluation with the Latest Works

In this work, a comparison of the presented strategy is performed with several new techniques by using two types of evaluations. First, several studies are taken to evaluate the class-wise results, whereas, in the next phase, numerous works are chosen to compare the results on the entire data sample.

Initially, some recent works (Bracino et al. 2020; Hasan et al. 2022; Wiesner-Hanks et al. 2018; Khan et al. 2019) are considered to perform the category-oriented performance comparison of our approach against them. The obtained comparison is shown in Table 3. Comparatively, the proposed approach has shown the accuracy scores for apple scab, black rot, apple rust, and healthy classes of 98.54%,

Table 2 Performance comparison with the deep learning works

Framework	Parameters	Train time (Hrs)	Test time (sec)	Precision (%)	Recall (%)	Accuracy (%)	F1 (%)
VGG19	143,667,240	25	257	94	92	96	92
ResNet152	58,772,356	19	73	95	91	95	93
DenseNet201	20,242,984	17	48	95	92	94	93
ResNet50	25,636,712	12	20	90	88	90	88
VGG16	138,357,544	15	16	94	95	96	94
InceptionV3	23,851,784	12	14	95	94	95	94
Xception	22,910,480	13	113	93	94	95	93
Proposed	*19,718,910*	*11*	*13*	*99.10*	*98.80*	*99*	*98.95*

99.71%, 98.26%, and 99.49%, which is the largest of all other selected works. Clearly, for the apple scab class, the comparison approaches have attained an average accuracy number of 89.06%, which is 98.54% for our work, and we have registered a performance gain (PG) of 9.48%. While for the Black Rot class, the competitor works have given an average accuracy value of 95.41%, which is 99.71% for our model. So, for the Black Rot class, we have shown a PG of 4.3%. Moreover, for the Cedar Apple Rust and healthy classes of the PlantVillage dataset, the comparison works have given the average accuracy numbers of 92.91 and 95.52%, while our approach has obtained the accuracy numbers of 98.26 and 99.49% for the mentioned classes and presents the PGs of 5.35 and 3.97%. The main reason for this effective performance result is due to the improved feature learning competency of our technique, which enables it to recognize the various classes of apple fruit plant diseases accurately. The approach (Hasan et al. 2022) is not proficient enough in dealing with the images suffering from cluttering and distortion problems, while the work mentioned in (Wiesner-Hanks et al. 2018) is unable to perform well for samples with small spots of abnormalities. Moreover, other works (Bracino et al. 2020; Khan et al. 2019) are unable to tackle the brightness variations of a sample. Our model has the capability of effectively dealing with the transformation changes of the examined images by computing the disease-specific features, which improve the recall power of our model and provide it the performance overhead from the comparative techniques.

Table 3 Class-wise performance comparison

Reference	Category	Accuracy (%)
Hasan et al. (2022)	Apple scab	97.94
	Black rot	99.68
	Cedar apple rust	98.31
	Healthy	99.73
	Average	98.63
Wiesner-Hanks et al. (2018)	Apple scab	96.90
	Black rot	98.10
	Cedar apple rust	97.02
	Healthy	98.81
	Average	97.50
Bracino et al. (2020)	Apple scab	64.10
	Black rot	85
	Cedar apple rust	81.70
	Healthy	85.55
	Average	83.30
Khan et al. (2019)	Apple scab	97.30
	Black rot	98.86
	Cedar apple rust	94.62
	Healthy	98.00
	Average	97.20
Proposed	Apple scab	98.54
	Black rot	99.71
	Cedar apple rust	98.26
	Healthy	99.49
	Average	99

In the next step, we have taken some recent approaches (Vishnoi et al. 2022, 2022, pp. 6594–6609; Khan et al. 2019; Gao et al. 2023; Assad et al. 2023; Tahir et al. 2021; Kodors et al. 2021; Rehman et al. 2021; Bi et al. 2022; Chakraborty et al. 2021; Akuthota et al. 2024; Wani et al. 2024; Dong et al. 2024) to compare the results of our approach in the aspect of the entire data sample. The attained comparison in terms of the accuracy measure is given in Table 4. The values in Table 4 prove the effectiveness of our approach as compared to other latest techniques that have attempted to resolve the same problem of recognizing the various infections of leaves of the apple fruit plant. Descriptively, the works in Vishnoi et al. (2022, pp. 6594–6609), Khan et al. (2019), Gao et al. (2023), Assad et al. (2023), Tahir et al. (2021), Kodors et al. (2021), Rehman et al. (2021), Bi et al. (2022), Chakraborty et al. (2021), Akuthota et al. (2024), Wani et al. (2024), and Dong et al. (2024) exhibit an average accuracy number of 94.78%, which is 99% for the suggested approach. So, our suggested approach has given a PG of 4.22% due to its high recall ability and empowerment to tackle the various transformational alterations of samples. Further, the usage of the AM unit along with the EfficientNet-V2 approach permits the E-AppleNet model

Table 4 Performance comparison with new works over the entire data sample

Reference	Accuracy (%)
Gao et al. (2023)	95.64
Assad et al. (2023)	96
Tahir et al. (2021)	97
Kodors et al. (2021)	87
Rehman et al. (2021)	96
Bi et al. (2022)	78
Khan et al. (2019)	97
Chakraborty et al. (2021)	96
Vishnoi et al. (2022)	98
Vishnoi et al. (2022, pp. 6594–6609)	98
Akuthota et al. (2024)	98
Wani et al. (2024)	97.25
Dong et al. (2024)	98.28
Proposed	99

to extract the low-level characteristics of the samples to reliably recognize the various diseases of the apple plant leaves. Moreover, the addition of dense layers right before the classification layer assists in passing the highly related sample feature set and effectively accomplishing the classification task.

Conclusion

This article has presented an effective DL framework named the E-AppleNet to classify the various types of diseases from apple fruit leaves. Descriptively, an enhanced EfficientNetV2 model is suggested by using the attention mechanism (AM) approach and adding the extra dense layers at the end of the model structure. Further, to address the issue of class imbalance and model over-fitting, transfer learning and multi-class focal loss are introduced. The E-AppleNet framework is competent in computing the structural description of the suspected samples by effectively investigating the diseased regions of apple plant leaves. The approach performed well over the apple fruit leaf samples from the PlantVillage dataset due to its high recognition and better feature nomination power. The E-AppleNet approach secures an accuracy score of 99% over the PlantVillage dataset, which proves the efficacy of the suggested framework. However, the proposed work struggles to perform well with fine-grained disease distinctions under complex backgrounds. Future directions could focus on integrating other latest attention mechanisms, like transformer-based models, to improve the model's ability to distinguish fine-grained disease patterns. Additionally, expanding the dataset with more diverse disease examples and background variations can enhance model robustness and accuracy in real-world scenarios.

Acknowledgements This work was funded by the University of Jeddah, Jeddah, Saudi Arabia, under grant No. (UJ-23-DR-24). Therefore, the authors thank the University of Jeddah for its technical and financial support.

Author Contribution A.B., A.J., M.N., and H.D. played a pivotal role in this study. A.B., A.J, and M.N. performed the conceptualization and finalized the methodology of the work, while, software and validation were accomplished by A.B., A.J, and H.S. Formal analysis, and the investigation was performed by A.J. and M.N. Further, A.J. and H.D. executed the resources and data curation phase. A.B., A.L., M.N, and H.D. wrote, reviewed, edited, and visualized the draft. A.B. arranged the funding, while the project supervision and administration were handled by A.J.

Data Availability Statement In the paper, we have used the large-scale, publicly available standard dataset that can be found at Hughes and Salathé (2015).

Conflict of interest A. Banjar, A. Javed, M. Nawaz, and H. Dawood declare that they have no competing interests.

References

Akuthota UC, Abhishek, Bhargava L (2024) A lightweight low-power model for the detection of plant leaf diseases. SN Comput Sci 5(4):327

Assad A et al (2023) Apple diseases: detection and classification using transfer learning. Qual Assur Saf Crop Foods 15(SP1):27–37

Ayaz H, Rodríguez-Esparza E, Ahmad M, Oliva D, Pérez-Cisneros M, Sarkar R (2021) Classification of apple disease based on non-linear deep features. Appl Sci 11(14):6422

Azgomi H, Haredasht FR, Motlagh MRS (2023) Diagnosis of some apple fruit diseases by using image processing and artificial neural network. Food Control 145:109484

Bi C, Wang J, Duan Y, Fu B, Kang J-R, Shi Y (2022) MobileNet based apple leaf diseases identification. Mob Netw Appl: 1–9

Bracino AA, Concepcion RS, Bedruz RAR, Dadios EP, Vicerra RRP (2020) Development of a hybrid machine learning model for apple (Malus domestica) health detection and disease classification. 2020 IEEE 12th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM). IEEE, pp 1–6

Chakraborty S, Paul S, Rahat-uz-Zaman M (2021) Prediction of apple leaf diseases using multiclass support vector machine. 2021 2nd International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST. IEEE, pp 147–151

Chen Y, Pan J, Wu Q (2023) Apple leaf disease identification via improved CycleGAN and Convolutional neural network. Soft Comput: 1–14

Dong Q, Gu R, Chen S, Zhu J (2024) Apple leaf disease diagnosis based on knowledge distillation and attention mechanism. IEEE Access

Dubey SR, Jalal AS (2016) Apple disease classification using color, texture and shape features from images. Signal Image Video Process 10:819–826

Gao Y, Cao Z, Cai W, Gong G, Zhou G, Li L (2023) Apple leaf disease identification in complex background based on BAM-net. Agronomy 13(5):1240

Gong X, Zhang S (2023) A high-precision detection method of apple leaf diseases using improved faster R-CNN. Agriculture 13(2):240

Hasan S, Jahan S, Islam MI (2022) Disease detection of apple leaf with combination of color segmentation and modified DWT. J King Saud Univ Inf Sci 34(9):7212–7224

Hassan NMH, Elshoky B, Hassan N, Elshoky B, Mabrouk A (2023) Quality of performance evaluation of ten machine learning algorithms in classifying thirteen types of apple fruits. Indones J Electr Eng Comput Sci 30:102–109

Hughes D, Salathé M (2015) An open access repository of images on plant health to enable the development of mobile disease diagnostics. arXiv preprint arXiv, vol 08060

Khan MA et al (2019) An optimized method for segmentation and classification of apple diseases based on strong correlation and genetic algorithm based feature selection. IEEE Access 7:46261–46277

Kodors S, Lacis G, Sokolova O, Zhukovs V, Apeinans I, Bartulsons T (2021) Apple scab detection using CNN and transfer learning

Li F, Zheng Y, Liu S, Sun F, Bai H (2024) A multi-objective apple leaf disease detection algorithm based on improved TPH-YOLOV5. Appl Fruit Sci: 1–17

Liu S et al (2023) An apple leaf disease identification model for safeguarding apple food safety. Food Sci Technol 43:

Nawaz M, Nazir T, Baili J, Khan MA, Kim YJ, Cha J-H (2023) CXray-EffDet: chest disease detection and classification from X-ray images using the EfficientDet model. Diagnostics 13(2):248

Omrani E, Khoshnevisan B, Shamshirband S, Saboohi H, Anuar NB, Nasir MHNM (2014) Potential of radial basis function-based support vector regression for apple disease detection. Measurement 55:512–519

Rehman ZU et al (2021) Recognizing apple leaf diseases using a novel parallel real-time processing framework based on MASK RCNN and transfer learning: an application for smart agriculture. IET image process 15(10):2157–2168

Tahir MB et al (2021) WITHDRAWN: recognition of apple leaf diseases using deep learning and variances-controlled features reduction. Elsevier

Tan M, Le Q (2019) Efficientnet: rethinking model scaling for convolutional neural networks. International Conference on Machine Learning. PMLR, pp 6105–6114

Tan M, Le QV (2021) Efficientnetv2: smaller models and faster training. arXiv preprint arXiv, vol 00298

Vishnoi VK, Kumar K, Kumar B, Mohan S, Khan AA (2022) Detection of apple plant diseases using leaf images through convolutional neural network. IEEE Access 11:

Vishwakarma PK, Chander S, Nimbolkar P (2024) Efficient water management tactics for mitigating fruit crop diseases. Appl Fruit Sci: 1–9

Wani OA, Zahoor U, Shah SZA, Khan R (2024) Apple leaf disease detection using transfer learning. Ann Data Sci: 1–10

Wiesner-Hanks T et al (2018) Image set for deep learning: field images of maize annotated with disease symptoms. BMC Res Notes 11(1):1–3