A Novel Classification Approach for Grape Leaf Disease Detection Based on Different Attention Deep Learning Techniques

S Phani Praveen[1], Rajeswari Nakka[2], Anuradha Chokka[3],
Venkata Nagaraju Thatha[4], Sai Srinivas Vellela[5], Uddagiri Sirisha[6]
Department of Computer Science & Engineering, Prasad V Potluri Siddhartha Institute of Technology, Andhra Pradesh, India[1]
Department of Computer Science & Engineering, Seshadri Rao Gudlavalleru Engineering College,
Gudlavalleru, Andhra Pradesh, India[2]
Department of Computer Science & Engineering, Koneru Lakshmaiah Education Foundation,
Vaddeswaram, Andhra Pradesh, India[3]
Department of Information Technology, MLR Institute of Technology, Hyderabad, Telangana, India[4]
Department of Computer Science & Engineering, Chalapathi Institute of Technology, Guntur, Andhra Pradesh, India[5]
School of Computer Science & Engineering, VIT-AP, Andhra Pradesh, India[6]

***Abstract*—Preventing and controlling grape diseases is essential for a good grape harvest. With the help of "single shot multibox detectors", "faster region based convolutional neural networks", & "You only look once-X," the study improved grape leaf disease detection accuracy with effective attention mechanisms, which includes convolutional block attention module, squeeze & excitation networks, & efficient channel attention. The various attention techniques helped to emphasize important features while reducing the impact of irrelevant ones, which ultimately improved the precision of the models and allowed for real-time performance. As a result of examining the optimal models from the three types, it was found that the Faster (R-CNN) model had a lower precision value, while You only look once-X and SSD with various attention techniques required the fewest parameters with the highest precision, with the best real-time performance. In addition to providing insights into grape diseases & symptoms in automated agricultural production, this study provided valuable insights into grape leaf disease detection.**

Keywords*—Grape leaves; faster region-based convolutional neural networks; you only look once (x); single shot detection attention techniques

I. Introduction

Preventing and controlling crop diseases is crucial for producing safe and healthy vegetables, minimizing losses, and reducing the use of pesticides in the production of crops [1]. Thus, early detection & prevention of diseases are crucial. Grape plants can be affected by various diseases, such as powdery mildew, brown blotch, and anthracnose, which can significantly impact the yield and quality of the fruit. Traditional methods of detecting grape diseases rely on the experience of the growers or the guidance of experts, which can be slow, inefficient, and lack real-time performance. Images of grape leaves are used to detect, identify, and provide guidance about diseases infected with grape leaves [2] because disease-infected grape leaves often have visible spots.

Grape leaf disease detection is crucial for several reasons. Firstly, it allows growers to monitor the health of their grapevines and take appropriate actions to prevent or manage diseases effectively. Early detection enables timely interventions, minimizing potential damage and crop losses. Different grape leaf diseases require specific treatments, and accurate identification helps growers implement targeted control measures. This optimizes the use of pesticides, reduces environmental impact, and ensures effective disease management.

Grape leaf diseases can significantly impact the yield and quality of grapevine production. Some diseases cause defoliation, reducing the vine's ability to photosynthesize and produce energy, leading to decreased fruit quality, delayed ripening, and reduced yield. Early disease detection enables growers to protect the crop and implement measures to minimize yield losses. Early identification of grape leaf diseases is essential for preventing their spread within vineyards. Prompt isolation and treatment of infected vines help prevent diseases from affecting healthy plants. Additionally, preventive measures such as pruning, canopy management, and cultural practices can be implemented to reduce the likelihood of disease occurrence and spread. Economically, grapevines are valuable crops, and detecting diseases in grape leaves allows growers to make informed decisions on disease management, optimizing resource utilization, and reducing unnecessary costs. This helps preserve the economic viability of vineyards and sustain profitability in grape production.

Efficient disease detection and management practices also contribute to sustainable agriculture. Early identification minimizes the use of broad-spectrum pesticides, reducing their negative impact on the environment and non-target organisms. Targeted treatments based on accurate disease detection help reduce chemical inputs, promote ecological balance, and support sustainable cultivation practices for grapevines. In summary, grape leaf disease detection is vital for crop health monitoring, disease management, yield protection, disease prevention, economic considerations, and sustainable agriculture. Early detection allows for timely interventions, optimization of disease control measures, minimization of crop losses, and the long-term sustainability of grapevine production.

Due to the rapid development of artificial intelligence technologies, a wide variety of vision approaches are utilised in the processing of photos for various crop diseases [3][4][5]. Research into classifying agricultural diseases uses a wide range of approaches, including “genetic algorithms” [6], “support vector machines” [7],“K-means clustering” [8], “ensemble learning” [9], “Bayesian classification” [10], “radial basis functions” [11], & “filter segmentation” techniques [12]. Unfortunately, conventional approaches to crop disease classification and identification rely on labour-intensive, environment-dependent manual feature selection. In particular, the development of deep learning’s Convolutional Neural Network (CNN) has led to vast improvements in the field of autonomous detection and identification of agricultural diseases.

An object detection system that uses a convolutional neural network (CNN) has made great strides recently. Several applications make use of this technique, including recognition of faces [13], navigation [14], detection of road obstacles [15], detection of pedestrians, abnormal activity recognition[16], monitoring of physical activity[17][**?**],[18] detection of fruits, and detection of weeds [19]. Despite complex backdrops, crop leaf diseases can be detected using object detection algorithms due to CNN’s ability to extract high-dimensional properties from object images.

As a result, scientists in China and others have studied object detection algorithms to develop models for detecting crop diseases. For instance, Some authors have applied various models for object detection to the tomato disease dataset, including the Faster(R-CNN), and the Single Shot Multibox Detector. Faster (R-CNN) as well as VGG16, produced the best disease detection results. Dynamic identification of grape leaf illnesses was accomplished by using Faster (R-CNN) on time-series images of grape leaves. Using an enhanced Faster (R-CNN) model, the authors of [20] detected diseases in bitter gourd leaves with excellent results. Using an in-house dataset, The authors of [21] trained the SSD model to identify agricultural diseases with an overall accuracy of 83.90%. An enhanced model based on MobileNetv2 & YOLOv3 was proposed by the authors [22], which allowed for the early detection of grey speck disease in tomatoes. This refined model benefits from a number of desirable characteristics, including a low memory size, outstanding detection accuracy, and lightning-fast identification.

Previous studies have shown that using object detection technology to detect grape leaf diseases is feasible. Existing grape detector models, however, operate slowly and have low detection precision, which severely limits their application. This research included the attention methods of “convolutional block attention module,” “efficient channel attention,” & “squeeze &excitation attention” into the models of “Faster(R-CNN),” “SSD,” & “YOLO-X” to boost their accuracy and speed. The goal was to boost the feature extraction network’s efficiency and put more emphasis on health issues. Experiments were run on a plant village dataset of grape diseases, and the findings revealed that models based on diverse attention mechanisms, such as “Faster(R-CNN),” “SSD,” & “YOLO-X,” significantly improved detection accuracy and operation performance with only little parameter tweaks. The findings of this study can be used as a foundation for future work on grape disease control measures. The main objectives of the paper are to enhance the accuracy and speed of grape leaf disease detection, improve the efficiency of feature extraction networks, validate the performance improvements on a grape disease dataset, and provide a foundation for future grape disease control measures.

However, the existing literature lacks research on incorporating attention mechanisms, such as the ”convolutional block attention module,” ”efficient channel attention,” and ”squeeze & excitation attention,” into grape detector models like ”Faster(R-CNN),” ”SSD,” and ”YOLO-X.” There is a gap in knowledge regarding the potential impact of attention mechanisms on improving detection accuracy and processing speed for grape leaf diseases. The main objectives of the paper are:

- Enhance the accuracy and speed of grape leaf disease detection: The purpose of this work is to enhance the efficiency of previously developed grape detection models by incorporating attention mechanisms such as ”convolutional block attention module,” ”efficient channel attention,” and ”squeeze & excitation attention” into the models of ”Faster(R-CNN),” ”SSD,” and ”YOLO-X.” The objective is to achieve higher detection accuracy and faster processing speeds, addressing the limitations of slow operation and low detection precision in existing models.
- Improve the efficiency of feature extraction networks: By integrating attention methods into the models, the paper aims to enhance the efficiency of the feature extraction networks. The attention mechanisms help to prioritize relevant features and emphasize health issues related to grape leaf diseases, leading to more effective and accurate detection.
- Validate the performance improvements on a grape disease dataset: The research conducts experiments using a dataset specifically focused on grape diseases. By evaluating the models based on diverse attention mechanisms, such as ”Faster(R-CNN),” ”SSD,” and ”YOLO-X,” the paper aims to demonstrate significant improvements in detection accuracy and operation performance. The experiments involve minimal parameter tweaks, ensuring that the observed enhancements are primarily attributed to integrating attention mechanisms.
- Provide a foundation for future grape disease control measures: The findings of this study serve as a basis for future work on grape disease control measures. By demonstrating the effectiveness of attention mechanisms in improving detection accuracy and speed, the paper offers valuable insights and guidance for the development of advanced and efficient techniques for managing and controlling grape leaf diseases.

II. Related Work

Detecting plant diseases in a timely manner is crucial for effectively managing plant losses. However, relying on manual diagnosis by humans is a time-consuming process that is prone to errors and can be costly. To address these challenges, researchers have been actively exploring automated techniques for disease detection and classification in plants. The utilization of automated equipment and methods has emerged as a promising approach for monitoring crop fields. In this section, we will delve into the specifics of computer vision methods employed to identify and diagnose diseases in plant leaves.

The authors of [23] conducted a study focusing on detecting black rot in grape leaves using a YOLOv3-SPP-based deep learning method. The researchers employed a combination of super-resolution image enhancement and convolutional neural network techniques to identify the disease in grape leaves. The initial step involved upsampling the input image through bilinear interpolation. After enhancement, the processed inputs were fed into the YOLOv3-spatial pyramid pooling model, resulting in a remarkable detection accuracy of 95.79%. However, when tested in real field conditions, the precision of this method dropped to 86.69%. In a separate study, The authors of [24] proposed a deep learning approach specifically for accurately detecting black rot spots on grape leaves. They employed the DeepLab V3+ model, which incorporates feature maps from different levels and utilizes ResNet 101 as the backbone network. The test results demonstrated that the improved DeepLab V3+ model outperformed conventional methods.

The authors of [25], [26]developed a novel support vector machine & image processing-enabled technique for identifying and categorizing grape leaf disease. The authors of [27] employed a CNN-SVM-based approach to classifying five different species of grapevine leaves. They utilized the MobileNetv2 CNN model for leaf-type classification. Initially, features were extracted from the pre-trained MobileNet2 logits layer, and classification was performed using SVM with various kernels. The Chi-squares method was applied for feature selection, resulting in an impressive classification accuracy of 97.60%. The use of feature selection techniques significantly contributed to the improved accuracy of classification.

The authors of [28] focused on detecting grape black measles disease. They utilized the ResNet-50-based DeepLabV3 segmentation model in combination with fuzzy logic to determine the severity of the disease. The input image provided region of interest features and the percentage of infections. A fuzzy rule-based reference system was developed based on each feature, which was then used to grade the grape disease. The grading system allowed for the classification of healthy, mild, medium, and severe cases, specifically for measles disease.

III. Materials and Methods

A. Image Acquisition

The plant-Village dataset provides 4,062 images of grape leaves displaying common symptoms. In this dataset, 1,180 images were found to be affected by Black Rot, 1,383 by Esca measles, 1,076 by Leaf spot, and 423 by healthy leaves, all with a resolution of 256 × 256 pixels. A leaf with black rot, a leaf with black measles, a leaf with blight, and a healthy leaf is displayed in Fig. 1.

The data set contains varying quantities of images for each category, indicating significant imbalances. Esca is the most common classification, accounting for roughly 34% of the images, while Black Rot and Leaf Blight make up 29% and 26% respectively. There are also 1076 Healthy images and 10.4% Normal images in the collection.

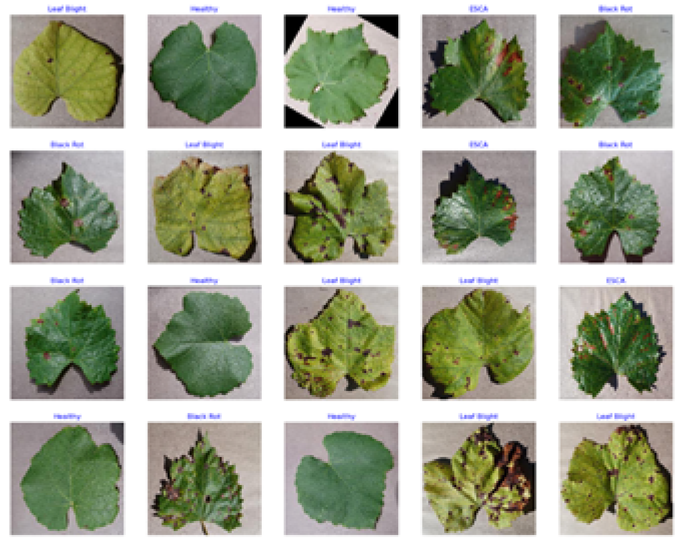

Fig. 1. Sample images from dataset.

TABLE I. List of Parameters used for Image Enhancement

Enhancement method	Parameters
gaussian filter	sigma-range(0.4,1.3)
mean filter	kernel-size-range(3,5)
median filter	kernel-size-range(2,5)
image acutance	alpha-range(0,0,2)
brightness	gamma(2.0)
contrast	alpha(1.0)

B. Image Pre-processing and Augmentation

The size of the dataset must be increased by using data augmentation techniques in order to prepare grape leaves disease images for disease identification. Training the recognition model in this way ensures that it will be more resilient and can generalize more effectively. Using standard data augmentation methods, the experiment compared the effectiveness of the data augmentation method proposed in this study. The traditional methods included flipping the image horizontally and vertically, rotating the image, applying different types of filtering (Gaussian, mean, and median) with a probability of 0.2, enhancing image contrast, sharpening images with a probability of 0.3, and adjusting image brightness. According to Table I, the parameters used for each image enhancement method are listed. Table II provides more details about the dataset and can be accessed at https://www.kaggle.com/datasets/rm1000/augmented-grape-disease-detection-dataset.

C. Proposed Methodology for Grape Leaf Disease Detection

This study focuses on the detection of grape leaf diseases using three specific models: faster-rcnn, YOLOx, and SSD.The training process for these models to detect diseases in grape leaves is depicted in Fig. 2. The process begins with inputting

TABLE II. INFORMATION ABOUT THE DATASET

Class	No of images without augmentation	No of images with augmentation
Healthy	423	3000
Esca measles	1383	3000
Leaf spot	1076	3000
Black rot	1180	3000
Total	4062	12000

the selected grape leaf disease images. Next, classification features are extracted from the input images. Output is then derived from the findings of disease identification using the faster-rcnn, YOLOx, and SSD models.

A loss function is used throughout to quantify the degree to which the projected disease species deviates from the true disease species. This enables the models to learn and improve their detection accuracy over time. The optimization of the final output result is achieved through the utilization of the Adam optimizer, a widely used optimization algorithm in deep learning.

By following this approach, the study aims to leverage the capabilities of faster-rcnn, YOLO:x & SSD models to detect grape leaf diseases effectively. The training flow chart provides a systematic framework for the feature extraction and disease detection process, facilitating the accurate identification of different disease species in grape leaves.

D. Attention Mechanism Models

The study utilizes three attention mechanisms: “Squeeze & Excitation”, “efficient channel attention”, and “Convolutional Block Attention” spatial attention mechanism. We chose the SE attention mechanism because it is simple and adds only a few new parameters. With ECA attention, models become more accurate without significantly increasing model complexity. It is an enhanced version of the SE attention mechanism. Finally, the CBAM attention mechanism is useful because it connects the spatial domain and the channel domain, leading to more effective improvement in network performance.

1) Squeeze & Excitation Attention: In order to extract features, the SE channel attention mechanism employs the CNN channel. It requires re-calibrating features so that the model can pick up and remember relevant details from all of the available feature channels. Fig. 3 depicts the two steps involved in this mechanism: squeezing and excitement. After the feature image has been spatially compressed using the squeeze technique, the feature channel’s relative relevance can be determined using the excitation technique; a model is created based on the correlation between the channels. In doing so, the original feature images are excited into matching channels. The SE mechanism has few additional parameters and is computationally simple.

The “efficient channel attention” attention mechanism is utilized to enhance cross-channel interaction and reduce model complexity, while the Squeeze & Excitation attention mechanism is used to prioritize the most informative channel features for disease identification. For end-to-end training of the grape leaf disease detection model, the “Convolutional Block Attention Module” attention mechanism is introduced to take into account the importance of pixels in different places. All three attention methods contribute significantly to improving the model’s efficiency and precision.

2) ECA Attention Module: It uses local cross channel interaction methods without reducing the magnitude of the dimensionality can be accomplished without using reduced-dimension SE. The functionality of the attention module is enhanced while its complexity is decreased thanks to this mechanism. In Fig. 4 we can observe the construction of the efficient channel attention mechanism.

3) CBAM Attention Module: The “CBAM Spatial Attention Module” is made up of 2 modules, first one is the “spatial attention module”, second one is the “channel attention module” and is designed to optimize input feature maps by inferring attention maps on both channel and spatial dimensions. These attention maps are then multiplied with the input feature map, resulting in self-adaptive feature optimization. The CBAM mechanism is effective in enhancing useful features while suppressing those that are not useful, making it a popular tool in practical applications. Fig. 5 illustrates the network structure of CBAM.

E. Dection Models for Disease Detection in Grape Leaves with Attention Mechanism

CNN-based object detection can be categorized into two main types. The first type uses a regional proposal to detect objects. This involves identifying candidate regions in the image, which are then divided to detect objects. This two-stage approach is exemplified by methods such as “R—CNN”, “Fast(R-CNN)”, & “Faster(R-CNN)”. The second type of object detection does not use a regional proposal and is referred to as one-stage object detection. An image is analyzed based on a CNN prediction of an object’s position & properties. There are a variety of algorithms available for this type of detection, such as SSDs and YOLOs.

The study used three models, namely the “Faster R–CNN model”, “YOLO-X model”, & “SSD model” for detecting grape leaves disease. The input of the selected grape leaf disease images, extraction of classification features, and use of the three disease detection models were involved in the process. The output was an analysis of the disease detection results. For optimizing the final output, an Adam optimizer was used to predict the difference between reality and the prediction of disease species.

Researchers found that the ” Faster(R-CNN)” model boosts high detection accuracy and can detect targets end-to-end. However, its running speed is relatively slow. On the other hand, the “YOLO-X” model runs quickly, but it doesn’t detect small objects. The “SSD” technique has faster running speed and higher detection accuracy than the “YOLO-X” model, but its training process heavily relies on prior experience, and its performance in detecting small targets is not as good as the “Faster(R-CNN)” model. The characteristics of these models are elaborated as follows:

1) Grape Leaves Disease Detection using Faster (R-CNN) Model: This model is comprised of three main components: the “Extraction of features”, the “Region Proposal Network”

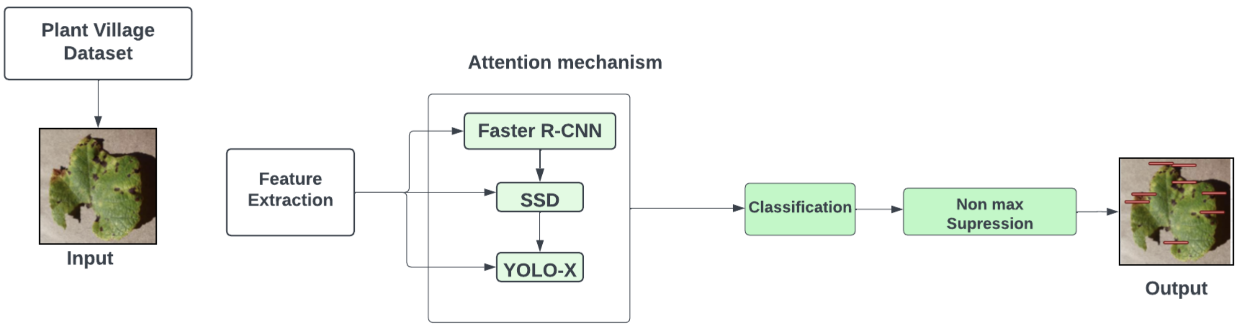

Fig. 2. Proposed attention model for grape disease detection.

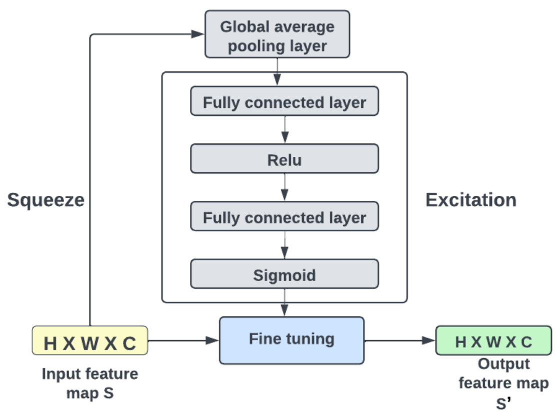

Fig. 3. SE attention mechanism.

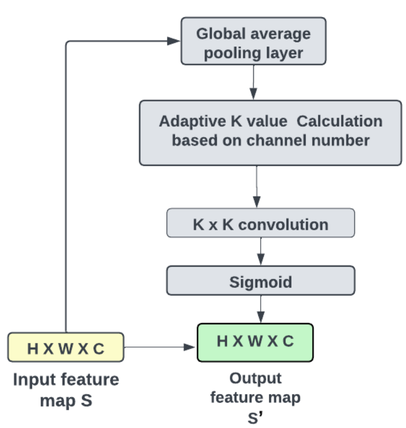

Fig. 4. ECA attention mechanism.

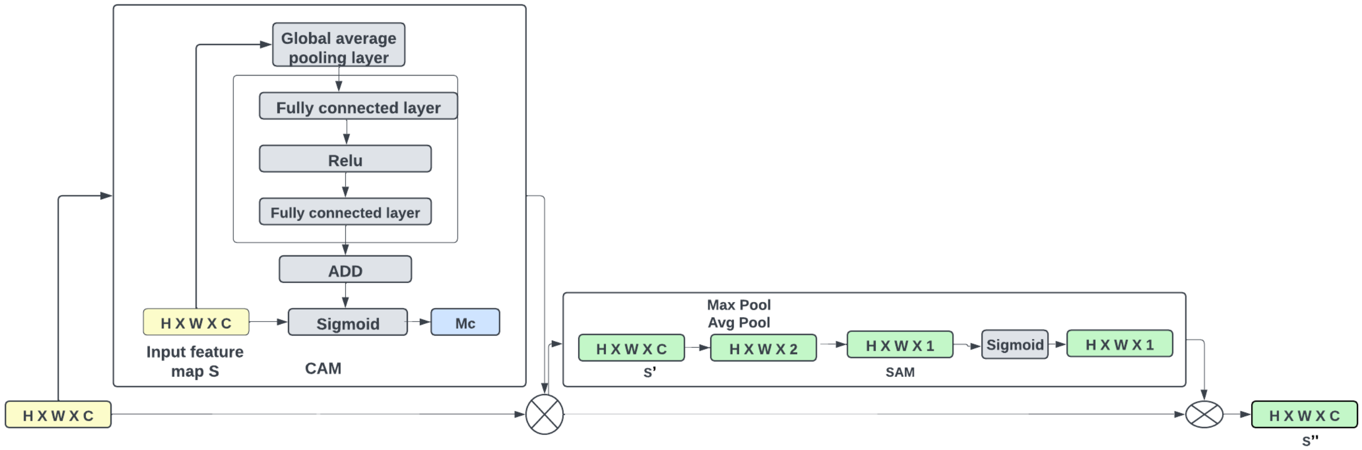

Fig. 5. CBAM attention mechanism.

, and the "Region with Convolutional Neural Network Features". Fig. 6 depicts the Faster (R-CNN) model with attention techniques. A Faster (R-CNN) method is used to detect grape leaf diseases in four primary steps: generating candidate disease regions, extracting disease characteristics, categorizing the disease, and performing bounding box regression. The Faster (R-CNN) model utilizes convolutional neural networks for the extraction of features and then generates feature maps for corresponding images. However, the convolution kernel's inherent locality means that only local information of disease images is retained, leading to information loss and reduced detection accuracy. To address this issue, the study introduced attention mechanisms, namely SE, ECA, and CBAM, without changing the feature extraction network's structure or backbone features. As a result of forward propagation after the last identity block, these mechanisms were introduced to enhance the model.

2) Grape Leaves Disease Detection using YOLO-X Model:: The YOLO-X with Darknet53 network is a model with high operational speed and flexibility. It includes four primary components: the input end, Backbone network, Neck, and Prediction. Fig. 7 illustrates the YOLO-X model based on various attention mechanisms. In the YOLO-X model, the YOLO Head has been changed to a decoupled head in the prediction section, the anchor-based approach has been replaced with an anchor-free method, and the SimOTA method has been introduced for dynamic matching with positive samples. The model's detection accuracy and speed have both been enhanced by these revisions, and the models' parameter sizes have been significantly decreased. The YOLO-X model is known for its high detection speed and precision, but it has some limitations when applied directly for disease detection in different environments. For instance, its backbone lacks the ability to extract features and integrate high-quality contextual feature information, leading to a reduction in the model's detection precision. Therefore, in this study, the Darknet53 network structure of the YOLO-X model remained unchanged, allowing pre-training weights to be directly loaded into model training. The YOLO-X model can selectively strengthen key features while suppressing irrelevant ones based on the branches of the backbone network, namely "Darknet53", "convolutional block attention module", "efficient channel attention" and "squeeze & excitation attention mechanisms.

3) Grape Leaves Disease Detection using SSD Model:: Using a tiny convolution kernel and multi-dimensional feature prediction, the model combines the anchor mechanism of Faster (R-CNN) with the regression mechanism of "YOLO" for fast and accurate detection. Fig. 8 depicts the SSD model that includes attention mechanisms. The first component is an enhanced capability for disease detection based on the deep learning network model used to collect baseline disease features. The multi-scale feature detection network is the second part, and it uses cascaded-neural-networks to categorize features at various scales in order to learn about the disease's category and location, as well as low-layer convolutional layer features to enhance detection precision and Non–Maximum suppression to generate the final detection results. Using a multi-dimensional prediction strategy, the SSD model is able to distinguish between small and large objects; the front-end deep-learning models are responsible for the former, while at the back-end multi-dimensional feature detection models handle them. Although the front-end network delivers precise coordinates and geometry, it has a limited range of perception and isn't great at representing abstract concepts. Whereas the frontal network has a narrow receptive field and poor representational capacity for geometric data, the posterior network has a wide receptive field and excellent representational ability for semantic data. Because of this, the SSD model may overlook some diseases or incorrectly identify others. Six feature images of varying sizes were collected from the "SSD" model and supplied into the various attention modules in order to better represent critical feature information and identify disease object features. With this method, the SSD model is better able to recognize diseased items.

IV. Results and Experiments

A. Evaluation Metrics

Results were evaluated based on standard measures for evaluating target detection. One class of targets will be evaluated using "Precision," "Recall," "Average Precision," and

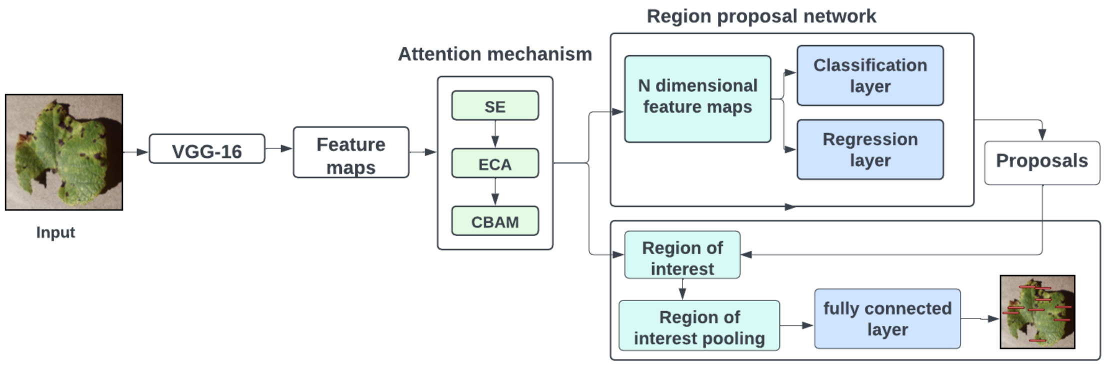

Fig. 6. Faster (R-CNN) model with attention techniques.

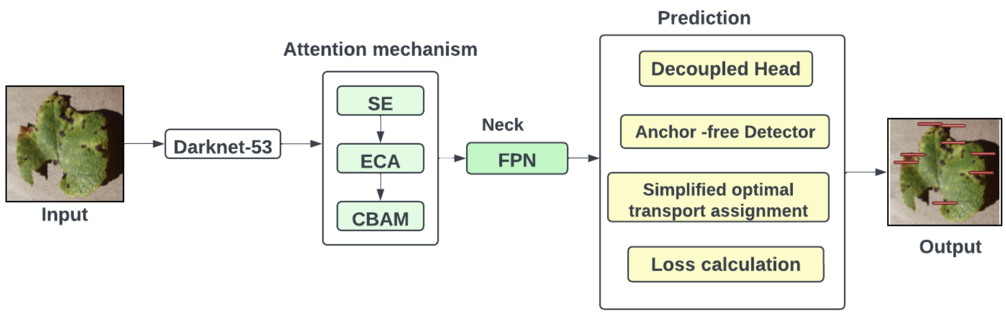

Fig. 7. YOLO-X model with attention mechanism.

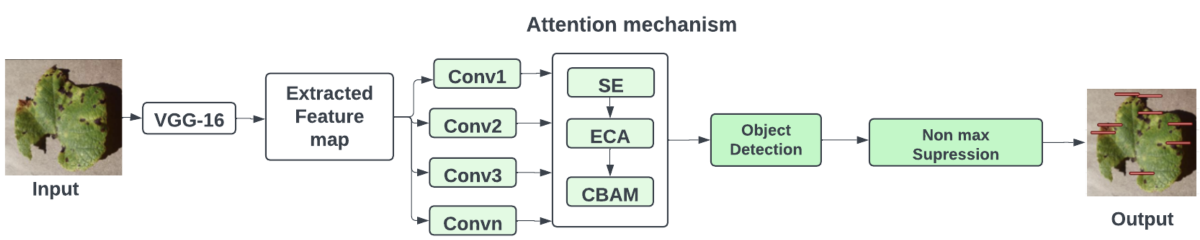

Fig. 8. SSD model with attention mechanism.

"Mean Average Precision," while all targets will be evaluated using "Mean Average Precision." However, in this study, we evaluated the grape leaf disease detection model's performance on a wider set of metrics, including the mean absolute percentage (mAP), the frame rate (FPS), the parameters, and the precision (P) and recall (R) values. The Eq. 1,2 and 3 were used to calculate P, R, and F1.

$$Precision = \frac{TruePositives}{TruePositives + FalsePositives} * 100 \quad (1)$$

$$Recall = \frac{(TruePositives)}{(TruePositives + FalseNegatives)} * 100 \quad (2)$$

$$F1score = \frac{(2 * Precision.Recall)}{(Precision + Recall)} \quad (3)$$

In Eq. 4, the variables P, TP, FP, R, FN, and F1 represent various metrics used to evaluate the performance of a model. P is the precision, which measures the percentage of correct positive predictions. The probability that grape disease leaves are accurately detected is denoted by true positives ('TP'), whereas the probability that they are mistakenly categorised as positive is denoted by false positives ('FP'). Recall, or the proportion of true positives that were correctly detected, is denoted by the letter R. The likelihood of mislabeling a positive sample as negative is known as the "Fasle negatives" ('FN') rate. F1 is a measure of accuracy that is the harmonic mean of two other metrics, recall and precision.

$$\int_0^1 PRdR \quad (4)$$

TABLE III. COMPARISON ANALYSIS OF FASTER (R-CNN) MODELS WITH DIFFERENT ATTENTION TECHNIQUES FOR DETECTING GRAPE DISEASES

Model	Precision	Recall	F1-Score	mAP
Faster (R-CNN) model	75.06	74.42	74.74	79.12
Faster (R-CNN) with SE Attention	79.80	84.23	81.96	85.39
Faster (R-CNN) with ECA Attention	76.54	78.71	77.61	81.93
Faster (R-CNN) with CBAM Attention	75.75	75.89	75.82	79.65
Faster (R-CNN) with SE, ECA,CBAM Attention	84.52	86.32	80.79	84.31

A higher value for TP indicates a more accurate prediction & better performance of the model. A model's performance can be measured using mAP, which is a metric that averages the average precision of all diseases. Eq. 5 defines mAP as the average of all AP values. FPS stands for the number of pictures handled each second. The algorithm's ability to recognize items improves as the FPS increases.

$$mAP = \frac{1}{N}\sum_{m=1}^{N} AP \tag{5}$$

A computer with 16 GB of RAM is used for this research, which runs Windows 10. Model parameters and hardware configuration are considered in Pytorch 1.10.1.

B. Experiment Results and Analysis

The grape disease dataset was utilized to compare the Faster(R-CNN), YOLO-X, and SSD models with the classical versions based on different attention mechanisms. The models were all trained and detected with the same configuration information and training platform.

1) Faster (R-CNN) Result Analysis: The "Faster(R-CNN)" model can be combined with different attention mechanisms to create different versions. Also we have combined the three attention mechanisms i.e. Faster (R-CNN) with SE, ECA,CBAM Attention. To test their performance in detecting grape diseases, all these versions were used in the same experimental setup, and the results are presented in Table III and in Fig. 9. Table III presents a comparison between the Faster (R-CNN) model and four modified versions: "Faster (R-CNN) with SE Attention", "Faster (R-CNN) with ECA Attention", and "Faster (R-CNN) with CBAM Attention". The results indicate that the Faster (R-CNN) with SE Attention model outperformed the original model with an increase in P, R, and F1 values by 4.74%, 9.81%, and 7.22% respectively, and an increase in mAP by 6.27%. Similarly, the Faster (R-CNN) with ECA Attention model showed improvements over the original model with an increase in P, R, and F1 values by 1.48%, 4.29%, and 2.87% respectively, and an increase in mAP by 2.81%. Finally, the "Faster (R-CNN) with CBAM Attention" model showed slight improvements over the original model with an increase in P, R, and F1 values by 0.69%, 1.47%, and 1.08% respectively, and an increase in mAP by 0.53%.

Based on the analysis above, it is evident that the performance of Faster (R-CNN) improved after the inclusion of attention mechanisms, despite a slight increase in parameters for "Faster (R-CNN) with SE Attention and Faster (R-CNN) with CBAM Attention". Enhanced precision and accelerated speed of detection are achieved through the attention mechanism for grape leaves images. Among the various models, "Faster (R-CNN) with SE, ECA, CBAM Attention" displayed the best detection effect when compared with "Faster (R-CNN) with SE Attention". The "Faster (R-CNN) with SE Attention" model demonstrated a 3.26%, 5.52%, and 4.35% increase in P, R, and F1 values, respectively, with an increase of 3.46% in mAP. In comparison with " Faster (R-CNN) with CBAM Attention", " Faster (R-CNN) with SE Attention" increased P, R, and F1 by respectively 4.05%, 8.34%, and 6.14%. When precision is considered, the " Faster (R-CNN) with SE, ECA, and CBAM Attention" model shows optimal results. It focuses on channel features with the most significant information while suppressing un-important features, making it ideal for detecting grape diseases in the dataset.

TABLE IV. COMPARISON ANALYSIS OF YOLO-X MODELS WITH DIFFERENT ATTENTION TECHNIQUES FOR DETECTING GRAPE DISEASES

Model	Precision	Recall	F1-Score	mAP
YOLO-X model	82.35	74.85	78.42	83.22
YOLO-X with SE Attention	82.46	82.21	82.33	84.02
YOLO-X with ECA Attention	87.77	86.07	86.91	88.66
YOLO-X with CBAM Attention	85.81	77.91	81.67	84.21
YOLO-X with SE, ECA,CBAM Attention	89.77	86.97	85.91	88.96

2) YOLO-X Result Analysis: The YOLO-X model has been enhanced with different attention mechanisms: SE, ECA, and CBAM. To compare their performance, all the models (including the original YOLO-X model) were tested on the dataset under the same configuration. The results are shown in Table IV and in Fig. 10. Table IV shows that the "YOLO-X with SE Attention" model has improved performance compared to the YOLO-X model. Specifically, the precision, recall, and F1 values of the "YOLO-X with SE Attention" model increased by 0.11 %, 7.36 %, and 3.91 %, respectively, while the mAP increased by 0.8%. Similarly, the "YOLO-X with ECA Attention" model also outperformed the YOLO-X model, with increases of 5.42%, 11.22%, and 8.49% in precision, recall, and F1 values, respectively. The mAP also increased by 5.44% respectively. The "YOLO-X with CBAM Attention" model also showed improvements, with increases of 3.46%, 3.06%, and 3.25% in precision, recall, and F1 values, respectively, and a 0.99% increase in mAP. Based on the analysis above, it was found that the detection performance of the YOLO-X model was improved with the introduction of attention mechanisms, despite a slight increase in the parameters of the "YOLO-X with SE Attention" and "YOLO-X with ECA Attention" models. Models were able to identify disease objects more accurately due to the attention mechanisms that allowed them to extract more comprehensive and rich features.

Out of all the models, the "YOLO-X with SE, ECA,CBAM Attention" model had the best detection performance. Compared to the "YOLO-X with SE Attention" model, the "YOLO-X with ECA Attention" model had a 5.31%, 3.86%, and 4.58% increase in P, R, and F1 values, respectively, a 4.64% increase in mAP, a 4.8 increase in FPS value, and a 0.49 MB expansion in parameters. Compared to the "YOLO-X with CBAM Attention" model, the "YOLO-X with ECA Attention" model had a 1.96%, 8.16%, and 5.24% increase in P, R, and F1 values, respectively, a 4.45% increase in mAP, a 1.8 increase in FPS value, and a 0.66 MB expansion in parameters. Compared to other models YOLO-X with SE, ECA,CBAM Attention

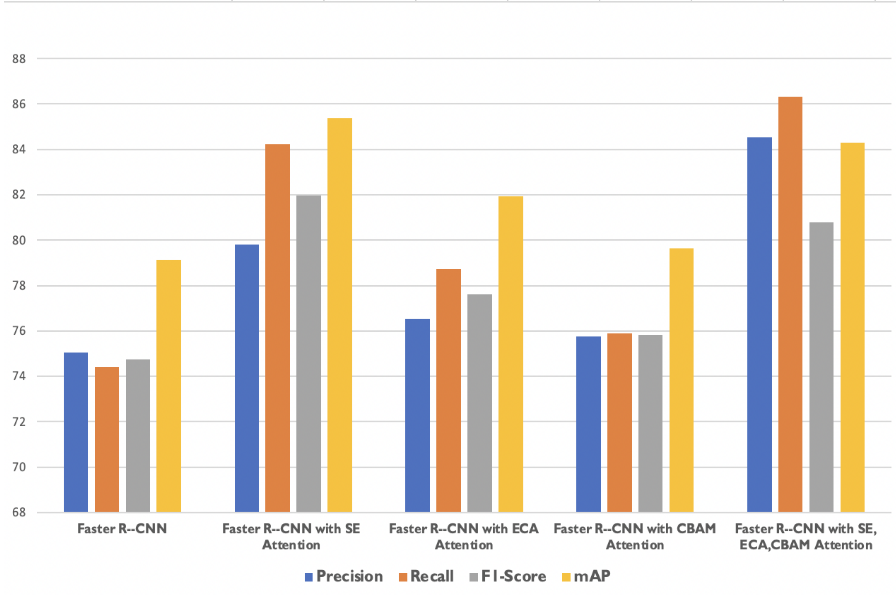

Fig. 9. Comparison analysis of Faster (R-CNN) models with different attention techniques.

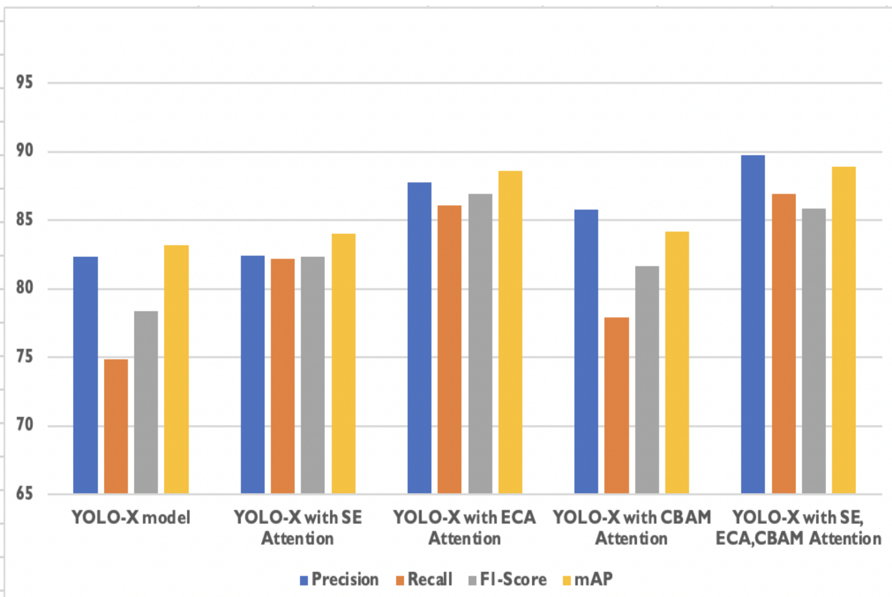

Fig. 10. Comparison analysis of YOLO-X models with different attention techniques.

models had outperformed than previous models. In conclusion, even though the "YOLO-X with SE, ECA,CBAM Attention" model had more parameters than the other three models, it achieved the best detection results with fast operation speed on the grape disease dataset, partially due to its ability to achieve cross-channel interaction.

3) SSD Result Analysis: Under the same experimental conditions, all the models were utilized to detect diseases on the plant village dataset, & the results of the experiment can be found in Table V and in Fig. 11.
Table V displays the results of various different models. The comparison is based on various metrics, including precision (P), recall (R), and F1 values, mean average precision (mAP). Compared to the SSD model, the SSD with SE Attention model showed significant improvements in P, R, and F1 values by 2.72%, 15.23%, and 9.45%, respectively.

The SSD with ECA Attention model also showed improvements over the SSD model, but to a lesser degree. A relative increase of 1.35%, 8.77%, and 5.47% in P, R, and F1 values was experienced, while a relative increase of 6.67% was seen in mAP.

The SSD with CBAM Attention model showed the smallest improvements over the SSD model. There is an increase of 0.94 %, 3.61 %, and 2.48 % in P, R, and F1 values, respectively, as well as a 4.91% increase in mAP and a

TABLE V. COMPARISON ANALYSIS OF SSD MODELS WITH DIFFERENT ATTENTION TECHNIQUES FOR DETECTING GRAPE DISEASES

Model	Precision	Recall	F1-Score	mAP
SSD model	80.74	68.87	74.33	76.23
SSD with SE Attention	83.46	84.10	83.78	86.96
SSD with ECA Attention	82.09	77.64	79.80	82.90
SSD with CBAM Attention	81.68	72.48	76.81	81.14
SSD with SE, ECA,CBAM Attention	85.46	84.90	84.78	83.96

3.38 MB increase in the model parameters. ”SSD with SE Attention”, ”SSD with ECA Attention”, and ”SSD with CBAM Attention” models were all enhanced by the incorporation of attention modules in the network architecture. However, the three models were able to effectively identify important information in feature images while filtering out irrelevant information based on feature importance. As a result, the detection performance of the three attention mechanisms with SSD was superior to that of the SSD model.

We have applied the different attention mechanism but, the “SSD with SE, ECA,CBAM Attention model” demonstrated the best detection performance with significantly faster real-time processing than the other three models. Compared to the “SSD with ECA Attention” model, the “SSD with SE Attention” model showed a 1.37%, 6.46%, and 3.98% improvement in P, R, & F1 values, respectively. Compared to the “SSD with CBAM Attention” model, the “SSD with SE Attention” model showed a 1.78%, 11.62%, and 6.97% improvement in P, R, and F1 values, respectively.
These experimental results demonstrate that the SE attention mechanism optimized feature images, resulting in significantly better detection performance and real-time processing compared to the other three models. Therefore, the “SSD with SE, ECA,CBAM Attention” model can be effectively applied in the detection of various grape diseases with superior comprehensive performance.

4) Comparison Analysis : After screening, the three optimal disease detection models were compared to present their disease detection performance. The analysis above showed that “Faster(R-CNN)”, “YOLO-X”, and “SSD” models when combined with multiple attention mechanisms were the optimal models of their respective detection methods. Fastest R-CNN models exhibited the lowest overall detection accuracy, the slowest operating speed, and the most parameters. The ”SSD” models’ rapid operation speed and great accuracy made them ideal for near-instantaneous disease diagnosis in vineyards. Strong robustness was demonstrated by the ”YOLO-X” models, which achieved the maximum detection precision with the fewest parameters and performed well while identifying both small objects and items hidden by background clutter.

V. CONCLUSION

After initial screening, three top disease detection models were selected and their performance was compared. The results of the foregoing investigation demonstrated that the ” Faster(R-CNN),” ”YOLO-X,” and ”SSD” models, when enhanced with numerous attention mechanisms, provided the most accurate detection results. Overall, ” Faster (R-CNN) ” models exhibited the lowest detection precision, the slowest operating speed, and the most parameters of the three types of models. Due to its excellent accuracy and quick processing speed, the ”SSD” model was found to be ideal for monitoring field grapes in real time. The ”YOLO-X” models demonstrated the highest detection accuracy with the fewest parameters, and they performed well while recognising both small objects and items that were partially obscured.

REFERENCES

[1] W. Baudoin, A. Nersisyan, A. Shamilov, A. Hodder, D. Gutierrez, D. PASCALE S, S. Nicola, N. Gruda, L. Urban, J. Tanny *et al.*, *Good Agricultural Practices for greenhouse vegetable production in the South East European countries-Principles for sustainable intensification of smallholder farms*. FAO, 2017, vol. 230.

[2] G. A. Carlson, “A decision theoretic approach to crop disease prediction and control,” *American Journal of Agricultural Economics*, vol. 52, no. 2, pp. 216–223, 1970.

[3] U. Sirisha and B. S. Chandana, “Privacy preserving image encryption with optimal deep transfer learning based accident severity classification model,” *Sensors*, vol. 23, no. 1, p. 519, 2023.

[4] U. Sirisha and S. C. Bolem, “Aspect based sentiment & emotion analysis with roberta, lstm,” *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 11, 2022.

[5] U. Sirisha and B. Sai Chandana, “Semantic interdisciplinary evaluation of image captioning models,” *Cogent Engineering*, vol. 9, no. 1, p. 2104333, 2022.

[6] R. Ghaffari, J. Laothawornkitkul, D. Iliescu, E. Hines, M. Leeson, R. Napier, J. P. Moore, N. D. Paul, C. N. Hewitt, and J. E. Taylor, “Plant pest and disease diagnosis using electronic nose and support vector machine approach,” *Journal of plant diseases and protection*, vol. 119, pp. 200–207, 2012.

[7] D. Zhang, G. Chen, H. Zhang, N. Jin, C. Gu, S. Weng, Q. Wang, and Y. Chen, “Integration of spectroscopy and image for identifying fusarium damage in wheat kernels,” *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 236, p. 118344, 2020.

[8] Z. Wang, K. Wang, S. Pan, and Y. Han, “Segmentation of crop disease images with an improved k-means clustering algorithm,” *Applied engineering in agriculture*, vol. 34, no. 2, pp. 277–289, 2018.

[9] R. Kamath, M. Balachandra, and S. Prabhu, “Crop and weed discrimination using laws’ texture masks,” *International Journal of Agricultural and Biological Engineering*, vol. 13, no. 1, pp. 191–197, 2020.

[10] C. Bi and G. Chen, “Bayesian networks modeling for crop diseases,” in *Computer and Computing Technologies in Agriculture IV: 4th IFIP TC 12 Conference, CCTA 2010, Nanchang, China, October 22-25, 2010, Selected Papers, Part I 4*. Springer, 2011, pp. 312–320.

[11] K. P. Ferentinos, “Deep learning models for plant disease detection and diagnosis,” *Computers and electronics in agriculture*, vol. 145, pp. 311–318, 2018.

[12] F. Schroff, D. Kalenichenko, and J. Philbin, “Facenet: A unified embedding for face recognition and clustering,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 815–823.

[13] W. Yang, S. Fan, S. Xu, P. King, B. Kang, and E. Kim, “Autonomous underwater vehicle navigation using sonar image matching based on convolutional neural network,” *IFAC-PapersOnLine*, vol. 52, no. 21, pp. 156–162, 2019.

[14] S. Sivaraman and M. M. Trivedi, “Active learning for on-road vehicle detection: A comparative study,” *Machine vision and applications*, vol. 25, pp. 599–611, 2014.

[15] L. Zhang, L. Lin, X. Liang, and K. He, “Is faster r-cnn doing well for pedestrian detection?” in *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part II 14*. Springer, 2016, pp. 443–457.

[16] U. Sirisha and B. S. Chandana, “Gitaar-git based abnormal activity recognition on ucf crime dataset,” in *2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT)*. IEEE, 2023, pp. 1585–1590.

[17] P. N. Srinivasu, G. JayaLakshmi, R. H. Jhaveri, and S. P. Praveen, “Ambient assistive living for monitoring the physical activity of diabetic adults through body area networks,” *Mobile Information Systems*, vol. 2022, pp. 1–18, 2022.

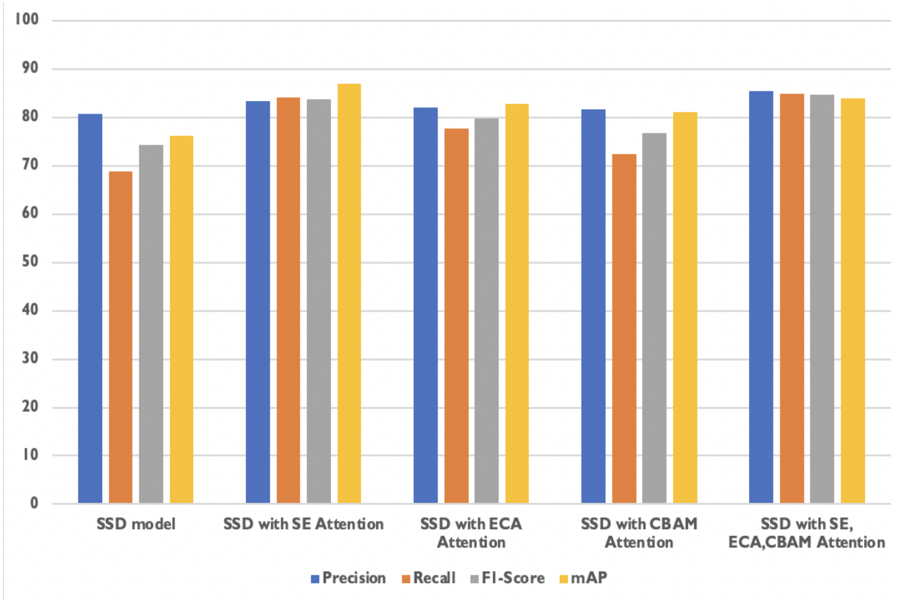

Fig. 11. Comparison analysis of SSD models with different attention techniques.

[18] N. R. Sai, B. S. Chandana, S. P. Praveen, S. S. Kumar *et al.*, "Improving performance of ids by using feature selection with ig-r," in *2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC)*. IEEE, 2021, pp. 1–8.

[19] A. Fuentes, S. Yoon, S. C. Kim, and D. S. Park, "A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition," *Sensors*, vol. 17, no. 9, p. 2022, 2017.

[20] Z. Liu, X. Yuan, J. Weng, Y. Liao, and L. Xie, "Application of bitter gourd leaf disease detection based on faster r-cnn," in *Advancements in Mechatronics and Intelligent Robotics: Proceedings of ICMIR 2020*. Springer, 2021, pp. 191–198.

[21] J. Qi, X. Liu, K. Liu, F. Xu, H. Guo, X. Tian, M. Li, Z. Bao, and Y. Li, "An improved yolov5 model based on visual attention mechanism: Application to recognition of tomato virus disease," *Computers and electronics in agriculture*, vol. 194, p. 106780, 2022.

[22] R. Polly and E. A. Devi, "A deep learning-based study of crop diseases recognition and classification," in *2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS)*. IEEE, 2022, pp. 296–301.

[23] J. Zhu, M. Cheng, Q. Wang, H. Yuan, and Z. Cai, "Grape leaf black rot detection based on super-resolution image enhancement and deep learning," *Frontiers in Plant Science*, vol. 12, p. 695749, 2021.

[24] H. Yuan, J. Zhu, Q. Wang, M. Cheng, and Z. Cai, "An improved deeplab v3+ deep learning network applied to the segmentation of grape leaf black rot spots," *Frontiers in Plant Science*, vol. 13, 2022.

[25] A. S. Ansari, M. Jawarneh, M. Ritonga, P. Jamwal, M. S. Mohammadi, R. K. Veluri, V. Kumar, and M. A. Shah, "Improved support vector machine and image processing enabled methodology for detection and classification of grape leaf disease," *Journal of Food Quality*, vol. 2022, 2022.

[26] S. P. Praveen, T. B. Murali Krishna, C. Anuradha, S. R. Mandalapu, P. Sarala, and S. Sindhura, "A robust framework for handling health care information based on machine learning and big data engineering techniques," *International Journal of Healthcare Management*, pp. 1–18, 2022.

[27] M. Koklu, M. F. Unlersen, I. A. Ozkan, M. F. Aslan, and K. Sabanci, "A cnn-svm study based on selected deep features for grapevine leaves classification," *Measurement*, vol. 188, p. 110425, 2022.

[28] M. Ji and Z. Wu, "Automatic detection and severity analysis of grape black measles disease based on deep learning and fuzzy logic," *Computers and Electronics in Agriculture*, vol. 193, p. 106718, 2022.