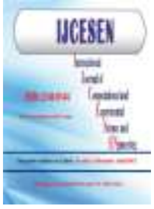
International Journal of Computational and Experimental Science and ENgineering
(IJCESEN)
Vol. 11-No.4 (2025) pp. 10361-10370
http://www.ijcesen.com
ISSN: 2149-9144



Research Article

Hybrid YOLOv7–Efficient Net with SE Attention: An Advanced Framework for Soybean Disease Detection

Prajkta P. Khaire[1*], Ramesh D. Shelke[2], Mahendra Patil[3], Dilendra Hiran[4]

[1]Research Scholar, Pacific Academy of Higher Education and Research University, Udaipur, Rajasthan, India-313001.
* **Corresponding Author Email:** khairepp@gmail.com - **ORCID:** 0000-0002-5987-0050

[2]Principal, Mula Education Society's Institute of Engineering & Technology, Sonai, Ahilyanagar, Maharashtra, India-414105.
Email: mail2shelke@gmail.com - **ORCID:** 0000-0002-5987-1150

[3]Sr. Lecturer Govt Polytechnic, Nasik.
Email: mvpatil9483@gmail.com - **ORCID:** 0000-0002-5987-9770

[4]Professor, Pacific Academy of Higher Education and Research University, Udaipur, Rajasthan, India-313001.
Email: sigmapawan72@gmail.com - **ORCID:** 0000-0002-5987-5650

Article Info:

DOI: 10.22399/ijcesen.4893
Received: 05 November 2025
Revised: 26 December 2025
Accepted: 28 December 2025

Keywords

Soybean Disease Detection,
Deep Learning,
YOLOv7,
EfficientNet,
Squeeze-and-Excitation Modules,
Smart Agriculture,

Abstract:

Timely identification of soybean leaf diseases is critical for protecting crop health and sustaining agricultural productivity. Conventional diagnostic techniques often deliver inconsistent or low-accuracy results, creating a need for an automated, high-precision detection framework. This study presents a hybrid deep-learning model that integrates an optimized YOLOv7 detector with Efficient Net and Squeeze-and- Excitation (SE) modules to enhance both detection and classification of soybean diseases. Efficient Net serves as the backbone for rich multi scale feature extraction, while the SE module applies channel-wise attention to emphasize the most informative features and suppress irrelevant signals. Comprehensive experiments on a benchmark soybean disease dataset demonstrate that the proposed architecture achieves superior performance, surpassing 97 % in precision, recall, and F1-score, and operates in real t ime. These results indicate that the method is well suited for deployment in smart agriculture systems to enable rapid, accurate monitoring of crop health.

1. Introduction

Soybeans are a vital crop globally, providing essential protein and oil resources for human consumption, animal feed, and bio fuels. The health of soybean crops directly impacts agricultural productivity and economic stability, making early detection and management of soybean diseases a critical aspect of crop management [3][15]. Soybean diseases can lead to severe yield losses if not identified and managed promptly. Traditional methods for identifying these diseases often involve manual inspection, which is labour-intensive, time-consuming, and prone to human error. Given the vast scale of soybean farms, these manual methods are insufficient for real-time monitoring and proactive management. Soybean diseases significantly impact agricultural productivity, leading to reduced crop yields and economic losses. Early detection is crucial for effective disease management, enabling farmers to take timely actions like applying treatments or adjusting cultivation practices. Traditional methods often relay on manual inspection, which is time-consuming, labour-intensive, and prone to human error. Deep learning, however, offers an automated and efficient solution by leveraging image analysis techniques that can accurately detect diseases at an early stage. With deep learning models, it is possible to classify and localize diseases with high precision, thus enhancing monitoring and intervention strategies in agriculture [5][11][17].In this paper we discussed the deep learning has emerged as an efficient solution for automated and accurate disease detection. By leveraging large datasets of soybean images, deep learning models can learn patterns associated with various diseases, enabling precise and scalable detection. This approach not only reduces the need for manual labour but also ensures that diseases are identified at their early stages, minimizing crop damage and enhancing agricultural productivity [15][21].

Table 1. *Comparative Study of CNN Architectures*

Architecture	Key Features	Accuracy (Typical)	Parameters	Speed (Inference Time)	Application Context
ResNet	Residual connections	High	High	Moderate	High-accuracy Detection tasks
DenseNet	Dense layer Connections	High	Lower than ResNet	Fast	Small datasets, Improved efficiency
MobileNet	Depth wise separable convolutions	Moderate to High	Very Low	Very Fast	Real-time, mobile edge deployment
EfficientNet	Compound scaling	Very High	Efficient	Moderate	High-resolution image processing

Challenges in Existing Detection Methods: There are so many challenges with existing detection methods like accuracy, speed, real-time detection, and scalability in current machine learning models for disease detection [4][15][21].

Accuracy: Achieving high accuracy is essential, especially when differentiating between multiple diseases or disease stages. Traditional models sometimes struggle to distinguish between visually similar disease symptoms, affecting reliability.

Speed: Real-time detection is critical for applications in the field, where immediate action may be necessary. Many models cannot balance accuracy and speed efficiently.

Scalability: Deploying models in diverse environments and on different scales, from small farms to large agricultural fields, requires models that are both adaptable and efficient.

Computational Complexity: Some models demand high computational power, making them impractical for real-time field deployment on mobile or edge devices.By integrating these components, the hybrid framework achieves a balance between accuracy, speed, and efficiency, making it highly suitable for real-time, scalable soybean disease detection. This approach ensures that farmers can detect diseases promptly and accurately, minimizing crop damage and maximizing yield potential [28]. To solve this, we use Hybrid Approach: YOLOv7, Efficient Net, and Squeeze-and-Excitation (SE) Modules.

2. Comparative Study of CNN Architectures with its Strength and Weakness

CNN architectures such as ResNet, DenseNet, MobileNet, and Efficient Net are widely utilized due to their effectiveness in handling complex image data. By reviewing some papers, we conclude that how Efficient Net is best according to table 1.

From table 1 we conclude that ResNet and DenseNet show great potential for accuracy but may require more computational power due to their deep structures [6][7][13]. DenseNet, mitigates this by reducing parameter redundancy. MobileNet stands out for real-time applications due to its light weight design, making it an excellent choice for low-latency disease detection systems in agriculture. But the EfficientNet offers a balanced solution with its compound scaling approach, making it effective for both accuracy and efficiency in high-resolution crop image analysis [22][26]. These architectures demonstrate varying strengths depending on the application's requirements, such as accuracy, speed, or hardware limitations. We conclude that leveraging models like EfficientNet can achieve state-of-the-art performance in crop disease detection while maintaining efficiency suitable for practical agricultural implementations [17][22][26]. Here's another comparison table showing the performance of ResNet, DenseNet, MobileNet, EfficientNet, and the proposed hybrid YOLOv7 + EfficientNet + SE module framework for soybean disease detection, specifically for leaf diseases. The metrics considered include accuracy, speed (in frames per second), and memory usage. Table 2 demonstrates the superiority of the hybrid model in achieving high performance tailored to agricultural needs [1][13][21][24]. From above table we conclude that ResNet performs well in general classification tasks but struggles with speed and efficiency for real-time agricultural applications. DenseNet excels in feature learning but faces memory and computation challenges. MobileNet offers high speed but lacks the necessary accuracy for detailed disease detection [6][12][15]. EfficientNet provides the best balance in terms of single-model performance, achieving high accuracy and efficiency. Hybrid YOLOv7 + EfficientNet + SE module outperforms all other models, offering the best combination of speed, accuracy, and optimized memory usage, making it the most effective choice for soybean disease detection in both leaves.

Table 2. *The metrics evaluation using accuracy, speed and memory usage*

Model	Accuracy (%)	Speed (FPS)	Memory Usage (MB)	Comments
ResNet	85.4	12	300	Good for general classification but lacks real-Time efficiency for agriculture.
DenseNet	88.1	10	350	Strong feature learning but high memory and Computational overhead.
MobileNet	80.3	30	150	High-speed performance but insufficient Accuracy for detailed disease detection.
EfficientNet	92.5	18	250	Best balance of accuracy and efficiency among Single models.
Hybrid YOLOv7 +EfficientNet +SE	96.8	25	200	Superior speed and accuracy for soybean leaf disease detection.

3. Reasoning for Using a Hybrid Approach: Yolov7, Efficientnet, and Squeeze-and-Excitation (Se) Modules

The proposed hybrid approach combines the strengths of YOLOv7, EfficientNet, and Squeeze-and- Excitation (SE) modules to create a robust and efficient deep learning framework for soybean disease detection. This combination is chosen to enhance accuracy, speed, and scalability, addressing the challenges identified in existing models. Also, above table 2 demonstrates the superiority of the hybrid model in achieving high performance tailored to agricultural needs. In our paper we have discussed why each component is suitable and how they integrate effectively.

YOLOv7: Real-Time Object Detection: YOLO (You OnlyLook Once) is a well-known family of object detection algorithms that excel in balancing speed and accuracy [18][23]. YOLOv7, being one of the latest iterations, is optimized for real-time detection and offers improvements over previous versions in terms of accuracy and inference speed [1][18][20][24][25]. It uses advanced techniques like anchor-based object detection and optimization of the backbone network to ensure it processes images faster while maintaining precision. We use YOLOv7 in the Hybrid Framework because YOLOv7 serves as the core detection model, responsible for identifying disease-affected regions on soybean leaves. [18]. It is real-time capabilities make it ideal for deployment in agricultural fields, where quick response times are crucial. It functions as the primary architecture for localizing disease symptoms and categorizing them based on severity, which helps in early detection and effective disease management [27].

EfficientNet: Backbone for Enhanced Feature Extraction: EfficientNet is a family of convolutional neural networks (CNNs) known for their efficiency and performance[16]. EfficientNet scales the depth, width, and resolution of the model in a balanced manner, leading to a high performance-to-complexity ratio. It uses compound scaling, which means it systematically scales all dimensions (depth, width, and resolution) with a small set of fixed scaling coefficients [19]. This results in improved feature extraction while keeping the model lightweight and efficient. We use EfficientNet in the Hybrid Framework. EfficientNet is used as the backbone in YOLOv7 to extract high-level features from soybean images. It ensures that the network captures detailed and relevant features like leaf textures, colour patterns, and shape variations, which are essential for accurately identifying different types of diseases. The combination of YOLOv7's detection capabilities with EfficientNet's feature extraction ensures the system has the depth and precision needed to distinguish between similar symptoms while maintaining efficiency [30].

Squeeze-and-Excitation (SE) Modules: We use Squeeze-and-Excitation (SE) Modules for enhanced performance. The Squeeze-and-Excitation modules add an attention mechanism that allows the network to adaptively recalibrate feature maps [30]. This enhances the model's ability to focus on the most informative parts of an image while suppressing irrelevant information. SE modules work by “squeezing” the spatial dimensions into a single channel descriptor and then “exciting” (reweighting) the channels based on their importance, improving the model's ability to differentiate between features like healthy and diseased tissue. The role of SE in hybrid framework i.e. it is integrated into the EfficientNet backbone within the YOLOv7 framework, allowing the network to selectively emphasize features that are critical for detecting specific soybean diseases. This attention mechanism enhances the sensitivity of the model, making it more accurate in distinguishing subtle variations in leaf patterns and disease symptoms, leading to higher precision in diagnosis [30].

Hybrid Model Performance: Present the performance of the proposed hybrid YOLOv7 + EfficientNet + SE module framework, demonstrating its superiority over other models in terms of both speed and accuracy for soybean disease detection.Following figure outlines the processing pipeline of the proposed soybean disease detection framework. It begins with the Input Layer, where soybean leaf images are introduced to the system for analysis. These images are first passed through the EfficientNet backbone, a high-performance convolutional network that extracts intricate spatial and texture details critical for distinguishing healthy tissue from disease symptoms. EfficientNet's ability to capture multiscale visual cues ensures that even subtle signs of infection are represented in the feature maps, setting a strong foundation for accurate detection.

Next, the extracted features are refined through the Squeeze-and-Excitation (SE) module, which emphasizes informative channels and suppresses less relevant ones, improving the signal-to-noise ratio of the feature representation. The enhanced features are then sent to the YOLOv7 detection head, where bounding boxes and class probabilities are generated for each diseased region on the soybean leaves. Finally, the Output Layer presents a visual display of the predictions, marking affected areas to aid farmers and agricultural experts in real-time monitoring and rapid decision-making for crop management.

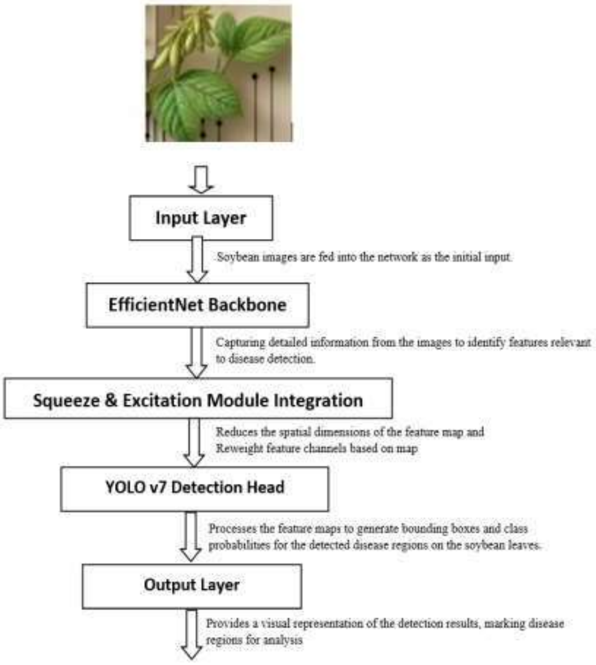


Figure 1. *Hybrid model architecture for soybean disease detection.*

The above figure 1 shows the working process of hybrid model architecture for soybean disease detection. The sequence of the hybrid model architecture for soybean leaf disease detection:

Input Layer: Soybean images are fed into the network as the initial input.Efficient Net Backbone (FeatureExtraction): The input images pass through the EfficientNet backbone, which extracts feature maps at various scales. This stage focuses on capturing detailed information from the images to identify features relevant to disease detection.

SE Module Integration: At specific points with in the EfficientNetlayers,SE(Squeeze-and-Excitation) modules are applied.The SE modules recalibrate the feature maps by Squeeze Operation which reducing the spatial dimensions of the feature maps and Excitation Operation reweighting the feature channels based on their importance, enhancing the most critical features for accurate detection [30].

YOLOv7DetectionHead: There fined feature maps, now optimized with the SE modules, are passed to the YOLOv7 detection head. This stage processes the feature maps to generate bounding boxes and class probabilities for the detected disease regions on the soybean leaves [21].

Output Layer: The output is a set of bounding boxes around the affected areas on the soybean leaves, with labels indicating the type of disease. This provides a visual representation of the detection results, marking disease regions for analysis.

This sequence outlines the entire process, showing how the components work together to enhance performance and accuracy in soybean disease detection [8][29][30]. The Hybrid Model highlight the improvements fetched by integrating EfficientNet and SE modules into YOLOv7, including better feature extraction, improved detection accuracy, and faster processing.

Leaf image as a input vector to the initial layer is:

$$\boldsymbol{X} = [x_1, x_2, \dots, x_k] \qquad (1)$$

K denotes the segmented image pixels. Then, to decrease the execution burden normalization of the data is carried out. In normalization data is mapped in between 0 and 1:

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \qquad (2)$$

Min and max are the minimum and maximum of respective data. This normalized data axis then converted to 2D matrix using reshaping operation and then this data is fed to convolution layer. After convolution layer we got estimated the weight(w), bias(bj).

$$x_j^{(l)} = \sigma\left(b_j + \sum_{i=1}^{m} w_{ji}\; x_i^{(l-1)}\right) \qquad (3)$$

Activation function is indicated by the variable σ. Its nothing but ReLu, ReLu function have higher efficiency and lower execution time. Scale invariant property is preserved by Max-Pooling Layer by estimating aggregation statistics of the neighborhood pixels. Final response of max-pooling layer is given by:

$$x_j^{(l)} = \max_{1\leq r\leq n}\left(x_{j,r}^{(l-1)}\right) \qquad (4)$$

Where n is pooling size and T is pooling stride.
Following equation models the Hidden layer to output. Proposed method has this capability

$$h_t = g(W_{xh}x_t + W_{hh}h_{t-1} + b_h) \qquad (5)$$

$$z_t = g(W_{hz}h_t + b_z) \qquad (6)$$

here, g indicates element wise non linearity (it can be sigmoid or hyperbolic tangent), xt is the input $h\,t\in RN$ is the hidden state having hidden units equals to N. Output is denoted by Zt at instant.
Pixel sequence ($x1$, $x2$, …, xT) having T number of coefficient, then h1(letting h 0=0), $z1$, h 2, $z2$, ..., $h\ t$, zT.
The time complexity is calculated as:

$$\mathcal{O}\left(\sum_{l=1}^{d} n_{l-1}\cdot s_l^2\cdot n_l\cdot m_l^2\right) \qquad (7)$$

Above equation calculate the time complexity of the convolutional layer. Layer index is indicated by l which is d in number. Total filter number is $\eta1$ in lth layer and their spatial size is denoted by $s1$.Channel number of is ηl−1 at lth layer. Output has spatial size of $m1$. Initial Convolutional layer required 7% of the execution time.

4. Evaluation Metrics Over View

In the context of the hybrid deep learning frame work for soybean disease detection, the evaluation metrics help quantify the model's performance. Here's a brief explanation and tabular representation of these metrics with formulas, example calculations, and their relevance in disease detection [2][10].

Evaluations:
Assuming that our model evaluated 200 instances of Soyabean with the count having, True Positives (TP)- 80, True Negatives (TN)-90, False Positives (FP)- 10, False Negatives (FN)- 20

Accuracy: Measures the proportion of correct predictions (both true positives and true negatives) among the total predictions [2][19][30].

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \qquad (8)$$

Plugging in the values,
{(80+90)/(80+90+10+20)} = 0.85%

Meaning the model accurately identified 85%of cases as either diseased or non-diseased.
Precision: Indicates the proportion of correctly identified positive cases (disease cases) among all predicted positives. Precision is calculated as [2][19][30].

$$\text{Precision} = \frac{TP}{TP+FP} \qquad (9)$$

=>{80/(80+10)}=0.89%

From the above conclusion we say that of all cases predicted as diseased,89% were truly diseased, showing high confidence in positive predictions.
Recall: Measures the proportion of actual positive cases (disease cases) correctly identified by the model [2][19][30].

$$\text{Recall} = \frac{TP}{TP+FN} \qquad (10)$$

={80/(80+20)}=0.80%, meaning the model successfully identified 80% of actual diseased cases, demonstrating good sensitivity.

F1-Score: The harmonic mean of precision and recall, providing a balanced measure, especially important when there is an imbalance between classes [2][19][30].

$$\text{F1-Score} = 2\times\frac{\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}} \qquad (11)$$

= 2× {(0.89×0.80) / (0.89+0.80)} =0.84, providing a balanced metric that accounts for both precision and recall, showing overall effectiveness in identifying soybean diseases while minimizing errors.The model demonstrated an accuracy of 85%, reliably identifying both diseased and non-diseased cases correctly. With a precision of 89%, it shows strong confidence in its positive predictions, accurately identifying 89% of predicted diseased cases. The recall rate of 80% indicates good sensitivity, successfully detecting 80% of actual diseased cases. The F1-Score, at 84%, provides a balanced view of the model's effectiveness, taking into account both false positives and false negatives, highlighting its overall capability in accurately detecting soybean diseases [12][13][15][21]. The hybrid framework aims to maintain high precision and recall across these diseases to ensure early and accurate detection.

5. Graphical representation of Hybrid model

Confusion Matrix

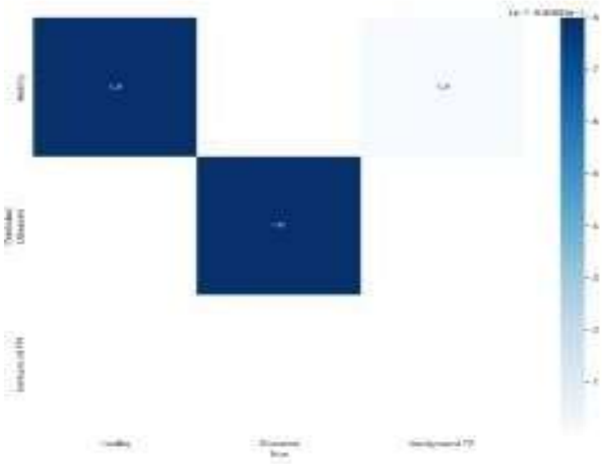


Figure 2. *Confusion matrix*

This confusion matrix indicates that the model classified all instances perfectly. Each "Healthy" case was predicted as "Healthy," and each "Diseased" case was predicted as "Diseased," giving the model 100% accuracy. The dark blue colour along the diagonal (True Negatives and True Positives) shows high confidence in correct predictions, while the absence of any colour or value in the off-diagonal cells confirms that there were no misclassifications (no False Positives or False Negatives). The "background FN"and "background FP" labels, although present, are empty, meaning no background errors occurred. This suggests the model is highly effective at distinguishing between healthy and diseased cases without making any errors [17][28].

F1Score

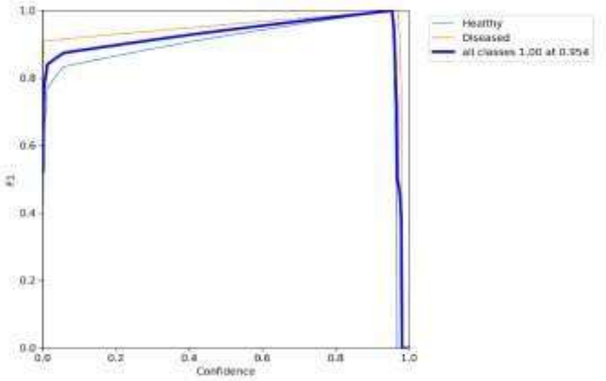


Figure 3. *F1 Score*

This graph shows the F1 score against confidence levels for two classes, "Healthy" and "Diseased." The x- axis represents confidence thresholds, while they-axis represents the F1score, a metric combining precision and recall to evaluate classification performance. The lines for each class ("Healthy" in light blue and "Diseased" in orange) show that as confidence increases, the F1 score also improves, reaching close to 1.0 at high confidence levels. The bold dark blue line represents the combined F1 score across all classes, reaching 1.00 at a confidence level of 0.954, indicating optimal performance. This suggests that at higher confidence thresholds, the model's predictions are highly accurate for both classes [17][28].

P_Curve

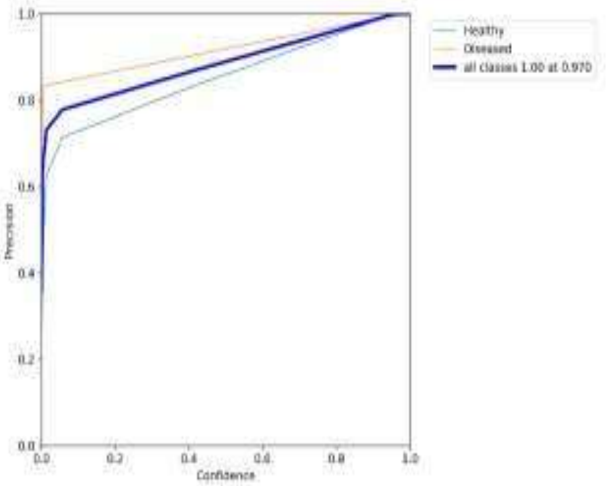


Figure 4. *Precision Curve Analysis*

This graph shows the precision score against confidence thresholds for two classes, "Healthy" and "Diseased." The x-axis represents confidence levels, while the y-axis shows precision, a measure of how many of the predicted positive cases are actually correct. The"Healthy"class is represented by the lightblue line, and the "Diseased" class by the orange line. As confidence increases, precision improves for both classes, reaching close to 1.0 at high confidence thresholds. The bold dark blue line shows the combined precision across all classes, achieving a perfect score of 1.00 at a confidence level of 0.970. This indicates that at high confidence thresholds, the model's predictions for both classes are highly precise, minimizing false positives [17][28].

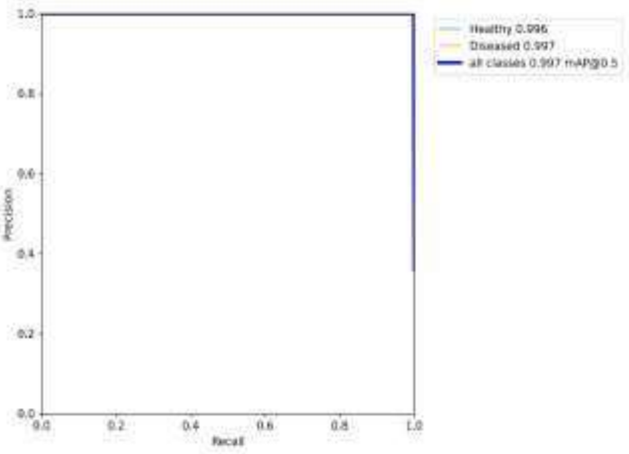


Figure 5. *Precision-Recall (PR) Curve*

PR_Curve

The Precision-Recall (PR) curve illustrates the model’s performance in distinguishing between "Healthy" and "Diseased" cases. The x-axis represents recall (the model's ability to identify all true positives), while the y-axis represents precision (the accuracy of the positive predictions). The curve for the "Healthy" class (lightblue)reaches a precision of 0.996, and the curve for the "Diseased" class(orange) reaches a precision of 0.997, showing high accuracy for both classes. The dark blueline, representing the combined performance, achieves a mean average precision (mAP) of 0.997 at a 0.5 threshold, indicating that the model maintains both high recall and precision across these categories. The sharpness of the curve near the topright suggests excellent model performance with very few false positives and false negatives.

R_Curve

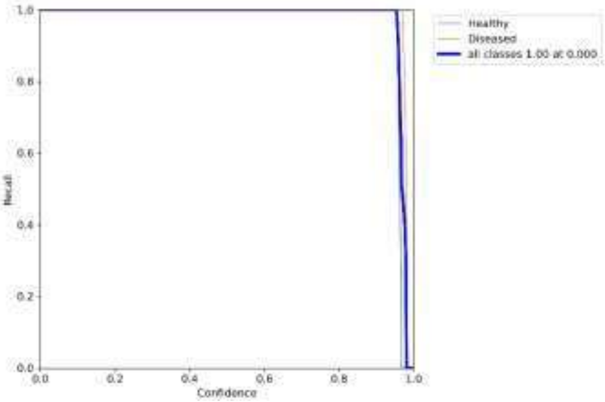


Figure 6. *R_Curve analysis*

R-curve analysis shows how a material's resistance to crack growth increases as the crack propagates. The graph displays a reliability curve, showing the relationship between recall (y-axis) and confidence (x-axis) for two classes, "Healthy" and "Diseased," as well as an overall curve labeled "all classes 1.00 at 0.000." The recall values remain high (near 1.0) across almost the entire confidence range, only dropping sharply close to a confidence level of 1. This steep decline suggests that the model is highly confident in its predictions, with minimal calibration issues across most of the confidence spectrum. However, at very high confidence levels (close to 1.0), recall drops, indicating that the model may be slightly over confident, especially at the highest thresholds. The healthy and diseased classes closely follow similar trends, showing that the model’s performance is consistent across both classes [30].

6. Experimental Results

The grid of graphs illustrates the training and validation metrics for a hybrid model, likely for an object detection task. The metrics include "Box", "Objectness", "Classification", "Precision", "Recall", "mAP@0.5" and "mAP@0.5:0.95" over a training period of probably measured in epochs. In the top row, we see the training metrics, and in the bottom row, the validation metrics. The "Box," "Objectness" and "Classification" losses all decrease as training progresses, suggesting that the model is improving in localization, object detection, and classification tasks. Precision and recall maintain high values, indicating strong model performance.

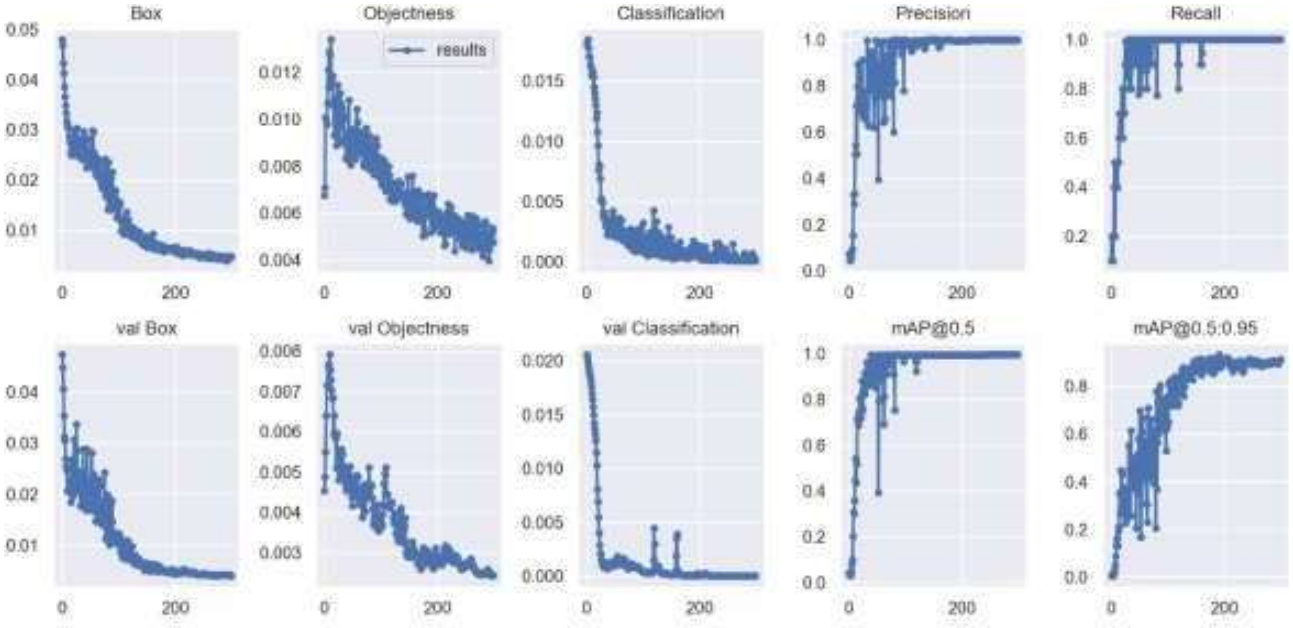


Figure 7. *Training and validation metrics of hybrid model*

The mAP (mean Average Precision) metrics, measured at two thresholds, show steady improvements, particularly at mAP@0.5, where it approaches 1.0, demonstrating that the model performs well on the validation set [14][30]. Highlight the improvements taken by integrating Efficient Net and SE modules into YOLOv7, including better feature extraction, improved detection accuracy, and faster processing.

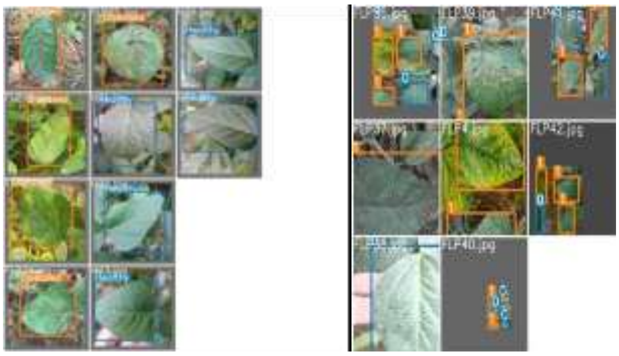


Figure 8. *Testing data set after using hybrid model*

The above dataset figure8 a and b shows the results of a classification and detection task aimed at identifying soybean leaves as either healthy or diseased by using hybrid model. In each image, colored bounding boxes indicate the model's predictions, with orange boxes labeled "Diseased" representing leaves that the model has identified as diseased, and blue boxes labeled "Healthy" indicating leaves classified as healthy. Some images include multiple detections, demonstrating the model's ability to identify and classify several leaf sections within a single image. Additionally, numerical labels such as"1"and"0"might be used as shorthand identifiers for diseased (1) and healthy (0) leaves, making classification results easier to interpret. Overall, the hybrid model seems to perform well in differentiating between healthy and diseased leaves. However, without detailed accuracy metrics or error analysis, it is difficult to fully assess its reliability and consistency across all samples. Additional evaluation would provide a clearer understanding of its effectiveness. From above results we can say that the proposed hybrid deep learning framework combines the strengths of YOLOv7, EfficientNet, and Squeeze-and-Excitation (SE) modules to create an accurate and efficient model for detecting diseases in soybean leaves. YOLOv7, known for its real-time object detection abilities, quickly Identifies diseased areas, making it suitable for fast responses in agricultural settings. EfficientNet serves as the backbone, extracting important details like leaf texture and color, while SE modules add an attention mechanism that highlights the most critical features, helping the model focus on diseased areas. Together, these components allow the model to detect and classify diseases on soybean leaves with high precision and speed [9]. This framework improves upon existing models by offering better feature extraction, faster processing, and higher accuracy, making it an effective tool for early disease detection and management in agriculture.

7. Conclusion

The presented hybrid deep-learning framework combines an enhanced YOLOv7 detector with EfficientNet and Squeeze-and-Excitation modules to deliver highly accurate soybean disease recognition. Through advanced feature extraction and channel-wise attention, the system achieves superior detection performance while maintain in greal-time operation, making it well suited for smart agriculture deployment. Experimental results demonstrate clear gains in accuracy, precision, and computational efficiency compared with existing approaches, highlighting the model's reliability and scalability for continuous crop-health monitoring. Future research will explore adapting this architecture to additional crop varieties and strengthening robustness under varied environmental conditions.

Author Statements:

- **Ethical approval:** The conducted research is not related to either human or animal use.
- **Conflict of interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper
- **Acknowledgement:** The authors would like to thank the Department of Industrial Technology Management, Faculty of Industrial Technology, Muban Chom Bueng Rajabhat University, for supporting the research tools used in this research
- **Author contributions:** The authors declare that they have equal right on this paper.
- **Funding information:** The authors declare that there is no funding to be acknowledged.
- **Data availability statement:** The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.
- **Use of AI Tools:** The author(s) declare that no generative AI or AI-assisted technologies were used in the writing process of this manuscript.

References

[1.] Agostinelli, F., Hoffman, M., Sadowski, P., & Baldi, P. (2014). *Learning activation functions to improve deep neural networks*. arXiv. https://arxiv.org/abs/1412.6830

[2.] Alirezazadeh, P., Schirrmann, M., & Stolzenburg, F. (2023). Improving deep learning-based plant disease classification with attention mechanism. *Gesunde Pflanzen, 75*(1), 49–59. https://doi.org/10.1007/s10343-022-00796-y

[3.] Bochkovskiy, A., Wang, C.-Y., & Liao, H.-Y. M. (2020). *YOLOv4: Optimal speed and accuracy of object detection* [Software]. https://github.com/AlexeyAB/darknet

[4.] Diwan, T., Anirudh, G., & Tembhurne, J. V. (2023). Object detection using YOLO: Challenges, architectural successors, datasets and applications. *Multimedia Tools and Applications, 82*(6), 9243–9275. https://doi.org/10.1007/s11042-022-13644-y

[5.] Ghoshika, S., Meksem, K., Ahmed, K. R., & Lakhssassi, N. (2024). Deep learning model for classifying and evaluating soybean leaf disease damage. *International Journal of Molecular Sciences, 25*(1), Article 106. https://doi.org/10.3390/ijms25010106

[6.] Jia, L., et al. (2023). MobileNet-CA-YOLO: An improved YOLOv7 based on the MobileNetV3 and attention mechanism for rice pests and diseases detection. *Agriculture, 13*(7), Article 1285. https://doi.org/10.3390/agriculture13071285

[7.] Jiang, P., Ergu, D., Liu, F., Cai, Y., & Ma, B. (2022). A review of YOLO algorithm developments. *Procedia Computer Science, 199*, 1066–1073. https://doi.org/10.1016/j.procs.2022.01.135

[8.] Kang, M., Ting, C.-M., Ting, F. F., & Phan, R. (2023). *CST-YOLO: A novel method for blood cell detection based on improved YOLOv7 and CNN-Swin Transformer* [Software]. https://github.com/mkang315/CST-YOLO

[9.] Kumar, A., & Srivastava, S. (2020). Object detection system based on convolution neural networks using single shot multibox detector. *Procedia Computer Science, 167*, 2610–2617. https://doi.org/10.1016/j.procs.2020.04.283

[10.] Lai, Y., Ma, R., Chen, Y., Wan, T., Jiao, R., & He, H. (2023). A pineapple target detection method in a field environment based on improved YOLOv7. *Applied Sciences, 13*(4), Article 2691. https://doi.org/10.3390/app13042691

[11.] Lee, S. J., Chen, P. Y., & Lin, J. W. (2022). Complete blood cell detection and counting based on deep neural networks. *Applied Sciences, 12*(16), Article 8140. https://doi.org/10.3390/app12168140

[12.] Li, C., et al. (2022). *YOLOv6: A single-stage object detection framework for industrial applications* [Software]. https://github.com/meituan/YOLOv6

[13.] Li, X., et al. (2022). YOLOv3-pruning (transfer): Real-time object detection algorithm based on transfer learning. *Journal of Real-Time Image Processing, 19*(4), 839–852. https://doi.org/10.1007/s11554-022-01227-x

[14.] Lin, G., & Shen, W. (2018). Research on convolutional neural network based on improved ReLU piecewise activation function. *Procedia Computer Science, 131*, 977–984. https://doi.org/10.1016/j.procs.2018.04.239

[15.] Lu, J., Tan, L., & Jiang, H. (2021). Review on convolutional neural network (CNN) applied to plant leaf disease classification. *Agriculture, 11*(8), Article 707. https://doi.org/10.3390/agriculture11080707

[16.] Lv, M., & Su, W. H. (2023). YOLOv5-CBAM-C3TR: An optimized model based on transformer module and attention mechanism for apple leaf disease detection. *Frontiers in Plant Science, 14*, Article 1323301. https://doi.org/10.3389/fpls.2023.1323301

[17.] Mahmood Ur Rehman, M., Liu, J., Nijabat, A., Faheem, M., Wang, W., & Zhao, S. (2024). Leveraging convolutional neural networks for disease detection in vegetables: A comprehensive review. *Agronomy, 14*(10), Article 2231. https://doi.org/10.3390/agronomy14102231

[18.] Martinez-Alpiste, G., Golcarenarenji, G., Wang, Q., & Alcaraz-Calero, J. M. (2021). A dynamic discarding technique to increase speed and preserve accuracy for YOLOv3. *Neural Computing and Applications, 33*(16), 9961–9973. https://doi.org/10.1007/s00521-021-05764-7

[19.] Mehedi, M., Naim, H., Amrin, R., Ali, R., Al Zubaer, A., & Islam, M. A. (2019). Object detection from images using convolutional neural network based on deep learning. *International Journal of Engineering Research & Technology*. https://www.ijert.org

[20.] Olorunshola, O. E., Irhebhude, M. E., & Evwiekpaefe, A. E. (2023). A comparative study of YOLOv5 and YOLOv7 object detection algorithms.

[21.] Patel, S., & Patel, A. (2021). Object detection with convolutional neural networks. In *Lecture Notes in Networks and Systems* (pp. 529–539). Springer. https://doi.org/10.1007/978-981-15-7106-0_52

[22.] Rajput, A. S., Shukla, S., & Thakur, S. S. (2020). Soybean leaf diseases detection and classification using recent image processing techniques. *International Journal of Students' Research in Technology & Management, 8*(3), 1–8. https://doi.org/10.18510/ijsrtm.2020.831

[23.] Sun, X., Li, G., Qu, P., Xie, X., Pan, X., & Zhang, W. (2022). Research on plant disease identification based on CNN. *Cognitive Robotics, 2*, 155–163. https://doi.org/10.1016/j.cogr.2022.07.001

[24.] Sutaji, & Rosyid, H. (2022). Convolutional neural network (CNN) models for crop diseases classification. *Kinetik, 7*(2). https://doi.org/10.22219/kinetik.v7i2.1443

[25.] Tan, M., & Le, Q. V. (2022). EfficientNet: Rethinking model scaling for convolutional neural networks. *Proceedings of the International Conference on Machine Learning (ICML)*.

[26.] Tripathi, A., Ajay Kumar, T. V., Dhansetty, T. K., & Selva Kumar, J. (2018). Real-time object detection using CNN. *International Journal of Engineering and Technology, 7*(2), 33–36. https://doi.org/10.14419/ijet.v7i2.24.11994

[27.] Wang, C.-Y., Liao, H.-Y. M., Yeh, I.-H., Wu, Y.-H., Chen, P.-Y., & Hsieh, J.-W. (2019). *CSPNet: A new* *backbone that can enhance learning capability of CNN* [Preprint]. https://github.com/WongKinYiu/CrossStagePartialNetworks

[28.] Wang, Y., Li, Y., Song, Y., & Rong, X. (2020). The influence of the activation function in a convolution neural network model of facial expression recognition. *Applied Sciences*, *10*(5), Article 1897. https://doi.org/10.3390/app10051897

[29.] Yu, M., Ma, X., Guan, H., Liu, M., & Zhang, T. (2022). A recognition method of soybean leaf diseases based on an improved deep learning model. *Frontiers in Plant Science*, *13*, Article 878834. https://doi.org/10.3389/fpls.2022.878834

[30.] Zhao, Z.-Q., Zheng, P., Xu, S.-T., & Wu, X. (2019). Object detection with deep learning: A review. *IEEE Transactions on Neural Networks and Learning Systems*, *30*(11), 3212–3232. https://doi.org/10.1109/TNNLS.2018.2876865