J. Electrical Systems 20-11s (2024): 5292-5308

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

NEXT GEN Soybean Disease Detection Via CST-YOLO And Enhanced Yolov7–Transformer Fusion

Abstract: - Soybean disease detection is necessary to keep crops healthy and improving productivity of agriculture. This paper proposes **CST-YOLO**, an innovative framework that integrates the YOLOv7 object detection model with a CNN–Swin Transformer architecture to boost preciseness and resilience of disease recognition. The framework incorporates customized improvements to YOLOv7, tailored specifically for soybean leaf analysis, while the Swin Transformer component captures complex spatial dependencies and strengthens the model's ability to differentiate between plants that are healthy and those that are sick. Extensive experiments conducted on benchmark datasets validate the effectiveness of CST-YOLO, achieving superior detection performance compared to widely used ensemble- based techniques that are often considered state-of-the-art. The results highlight CST-YOLO as a reliable, efficient, and scalable solution for automated soybean disease diagnosis, with potential applications extending to the detection of leaf diseases in other crop species.

Keywords: Soybean Disease Detection, YOLOV7, CNN-SWIN Transformer, Object Detection, Agricultural Innovation, SoftMax Activation Function

I. INTRODUCTION

CST-YOLO stands for an innovative technique developed for the detection of soybean diseases. It incorporates advancements in deep learning models, specifically an improved version of YOLOv7 (You Only Look Once) (C.-Y. Wang et al., 2022) and CNN-SWIN Transformer, enhanced with a SoftMax activation function. This technique aims to accurately identify various diseases affecting soybean crops as well as its pods by leveraging the capabilities of deep learning algorithms. By utilizing sophisticated neural network architectures and activation functions, CST-YOLO enhances the detection accuracy and efficiency compared to older methods (T. Abhiram et al., 2021). Through this approach, farmers and agricultural experts can promptly identify soybean diseases, enabling timely interventions to mitigate their impact on crop yields. Incorporating modern technologies like deep learning holds promise for enhancing agricultural productivity and sustainability. Soybean pod diseases can significantly impact agricultural productivity by reducing yield quantity and quality (M. Yu et al., 2022).

Deep learning techniques have become effective resources for disease detection in agriculture, offering significant potential for improving crop management practices and enhancing agricultural productivity (Z.-Q. Zhao et al., 2018). These methods leverage complex neural network architectures to automatically discover and extract characteristics from large datasets, facilitating precise and effective detection of plant diseases. By analysing various data sources such as images, sensor data, and spectral signatures, deep learning algorithms can identify subtle patterns associated with disease symptoms, allowing for early and precise detection (J. Redmon et al., 2015; J. Redmon et al., 2016, J. Redmon et al., 2018). One of the key advantages of deep learning-based methods is their capacity to manage unstructured and high-dimensional data, including images captured by drones or smartphones in the field. This enables real-time monitoring of crops at scale, facilitating timely interventions to prevent disease outbreaks and minimize yield losses. Additionally, through iterative training on fresh data, deep learning models can continuously enhance their performance, becoming more flexible. to

---

[1] Research Scholar, Pacific Academy of Higher Education and Research University, Udaipur, India Email: khairepp@gmail.com
[2] Co-Guide, Pacific Academy of Higher Education and Research University, Udaipur, India Email: mvpatil9483@gmail.com
[3] Co-Guide, MES's IET, Sonai Ahilyanagar, India Email: mail2shelke@gmail.com
[4] Guide, Pacific Academy of Higher Education and Research University, Udaipur, India Email: sigmapawan72@gmail.com

evolving disease dynamics and environmental conditions. In agriculture, techniques for illness detection based on deep learning have been utilized on a variety of crops, such as grains, legumes, fruits, and vegetables. like soybeans (T. Abhiram et al., 2021; H. Xu et al., 2022; M. Yu et al., 2022).

They have shown promise in detecting various types of diseases caused by pathogens, pests, and abiotic stresses, offering farmers useful information for making decisions and crop management. Moreover, these methods can complement traditional disease monitoring techniques, providing a economical and expandable solution for precision agriculture (H. Xu et al., 2022; M. Yu et al., 2022). Overall, deep learning-based methods hold great potential for revolutionizing disease detection in agriculture, enabling proactive and data-driven approaches to crop protection and management. As technology continues to advance and datasets grow larger and more diverse, the integration of deep learning algorithms into agricultural practices is expected to be essential to the future of sustainable agriculture and food security.

The motivation behind the development of CST-YOLO with SoftMax activation function as an innovative solution for soybean disease detection stems from the pressing need in the agricultural sector for accurate, efficient and scalable methods to detect and manage crop diseases (F. Agostinelli et al., 2014; G. Lin et al., 2018). Soybean, which is among the world's most significant crops, faces significant threats from various diseases of pod and leaf, which can have serious consequences on the quantity and quality of yield if left undetected and managed in a timely manner. Conventional illness detection techniques frequently depend on experts' visual assessment, which can take a lot of time, subjective and prone to errors (T. Abhiram et al., 2021). Additionally, as soybean fields expand and farming practices evolve, the demand is rising for technology and automated solutions that enable to provide real-time information on crop health.

Deep learning-based approaches, like YOLO (You Only Look Once) have demonstrated promise in automating disease detection tasks by analyzing crop images and identifying disease symptoms with high precision (L. Jia et al., 2023; M. Yu et al.,2022). However, to achieve optimal performance, these methods require continuous refinement and innovation. The development of the CST-YOLO with the SoftMax activation function represents a response to this must raise the precision and effectiveness of soybean disease detection. By incorporating advances in deep learning architectures and activation functions, CST-YOLO tries to improve the capabilities of the current detection techniques, enabling more accurate identification of soybean diseases at different developmental stages and under various environmental conditions (F. Agostinelli et al., 2014) . The incorporation of a SoftMax activation function further improves the model's classifying abilities of disease symptoms with confidence, enabling improved agricultural decision-making for farmers and agronomists. Ultimately, the goal of developing CST-YOLO is to provide a robust and reliable tool for the early identification and treatment of soybean diseases, thereby helping to improve agricultural productivity and food security.

II. RELATED WORK

Table 1 Existing Techniques for Soybean Disease Detection

Sr. No.	Methods	Description
1	**visual examination**	Farmers and agricultural specialists have always used visual inspection to detect soybean diseases. This involves physically examining plants for symptoms such as leaf discoloration, lesions, wilting, and abnormal growth patterns. While visual inspection is straightforward and cost-effective, it is subjective, labor-intensive, and may not always detect diseases at early stages or accurately differentiate between different pathogens (T. Diwan et al.,2023; M. Yu et al., 2022).
2	**Plant Pathology**	Traditional plant pathology techniques involve laboratory analysis of plant samples to identify pathogens causing disease. This typically includes techniques such as microscopy, culturing, and molecular assays like PCR (Polymerase Chain Reaction). Plant pathology provides definitive identification of pathogens but is time-consuming, requires specialized equipment and expertise, and is not suitable for outside monitoring in real time (P. Jiang et al., 2021; H. Xu et al., 2022; M. Yu et al., 2022).

3	**Remote Sensing**	Techniques for remote sensing, such as drone-based sensors, aerial photography, and satellite images, are increasingly being used for crop disease detection (X. Zhu et al.,2021). These methods capture multispectral or hyperspectral data, which can be analyzed to detect anomalies in plant health based on changes in reflectance patterns. While remote sensing allows for large- scale monitoring of crops, it requires sophisticated equipment, data processing, and interpretation.
4	**Machine Learning**	Machinelearning strategies, such as conventional classifiers (like Random Forest and SVM) and algorithms for deep learning, have been applied to soybean disease detection using various data sources like images, spectral data, sensor readings (S. S. Padmanabula et al., 2020). Feature engineering is necessary for traditional machine learning models. and may struggle with high-dimensional, complicated data. Deep learning approaches, as recurrent neural networks (RNNs) and convolutional neural networks (CNNs), have demonstrated promise in automating disease detection work by learning hierarchical representations directly from raw data (P.P.Khaire et al.,2023).
5	**Deep Learning**	CNNs and other deep learning-based techniques, have gained popularity in recent years for soybean disease detection because of their innate capacity for learning and take properties from pictures (Z.-Q. Zhao et al., 2018). Models such as Faster R-CNN (Region-based Convolutional Neural Network) and YOLO (You Only Look Once). have been adapted for object detection and localization of disease symptoms in soybean crops. These models offer advantages regarding precision, scalability, and performance in real time, but they need a lot of computational resources and annotated datasets for training (C.-Y. Wang et al., 2022).

In the above table 2 we discuss the traditional methods like visual inspection and plant pathology remain valuable for disease diagnosis, deep learning approaches are increasingly being adopted for their potential to automate and enhance soybean disease detection, particularly in extensive farming settings. Integrating this technique with existing practices can lead to more effective disease management strategies and improved crop yields.

Working of CST-YOLOV7 architecture with softmax

In the CST-YOLO architecture, SoftMax activation is applied to generate class probabilities for disease detection (L. Jia et al., 2023) In these paper we discuss how this process works within the CST-YOLO framework: This diagram illustrates the general design of the suggested CST-YOLO technique, showing full data flow from input image to final detection output. Beginning at the bottom left, a 640×640×3 input image passes through several stages of convolutional and feature-extraction modules. The early feature extractor employs repeated **CBS** (Convolution–BatchNorm–SiLU) blocks and **W-ELAN** modules to capture low-level spatial information, with **CBSConcat** operations merging outputs for richer representations. A **CST** (CNN-Swin Transformer) block is integrated in the backbone to strengthen global contextual learning. Higher up, the network uses **MCS** and **SPPCSPC** layers for multi-scale context aggregation, followed by **Upsample** and **Concat** operations that progressively combine features from different depths, allowing the both delicate and coarse details in the model..

The upper part of the diagram represents the **neck** and **head** of the detection pipeline. After multi-scale feature fusion through **W-ELAN 2** and **CatConv** layers, the neck generates three feature maps of different resolutions (20×20, 40×40, and 80×80), each tailored for detecting objects of varying sizes. The **RepConv** layers refine these fused features before they enter the **IDetect** detection heads. Each head outputs prediction tensors (for example, 20×20×3×(10+5)) encoding class probabilities and bounding-box coordinates. This multi-scale detection strategy enables CST-YOLO to accurately identify soybean leaf diseases of different scales, combining the speed of YOLOv7 with the global attention capabilities of the Swin Transformer.

The fig 1(a) shows the overview of the innovative architecture of CST-YOLO. In these object detection framework CST-YOLO is built upon the YOLO (You Only Look Once), It creates a grid out of the input image , forecasts the bounding boxes and corresponding class probabilities for each grid cell's items. This makes it possible for efficient and Real-time detection of several objects at once (C. Li et al.,2022). CST-YOLO utilizes a CNN backbone to retrieve characteristics from the input image. This CNN backbone consists of several layers of convolution that gradually learn hierarchical representations of visual features. These features are then passed to subsequent layers for object detection (F. Agostinelli et al., 2014; L. Jia et al., 2023; G. Lin et al., 2018). After extracting features from the CNN backbone, CST- YOLO applies a SoftMax activation function to the last convolutional layer's output. SoftMax function normalizes output scores across different classes, converting them into probabilities that sum up to 1. This ensures that each class probability represents the likelihood of the presence of a specific disease category. The layer of output from CST-YOLO contains multiple neurons each of which corresponds to a different class of soybean pod disease (M. Yu et al., 2022).

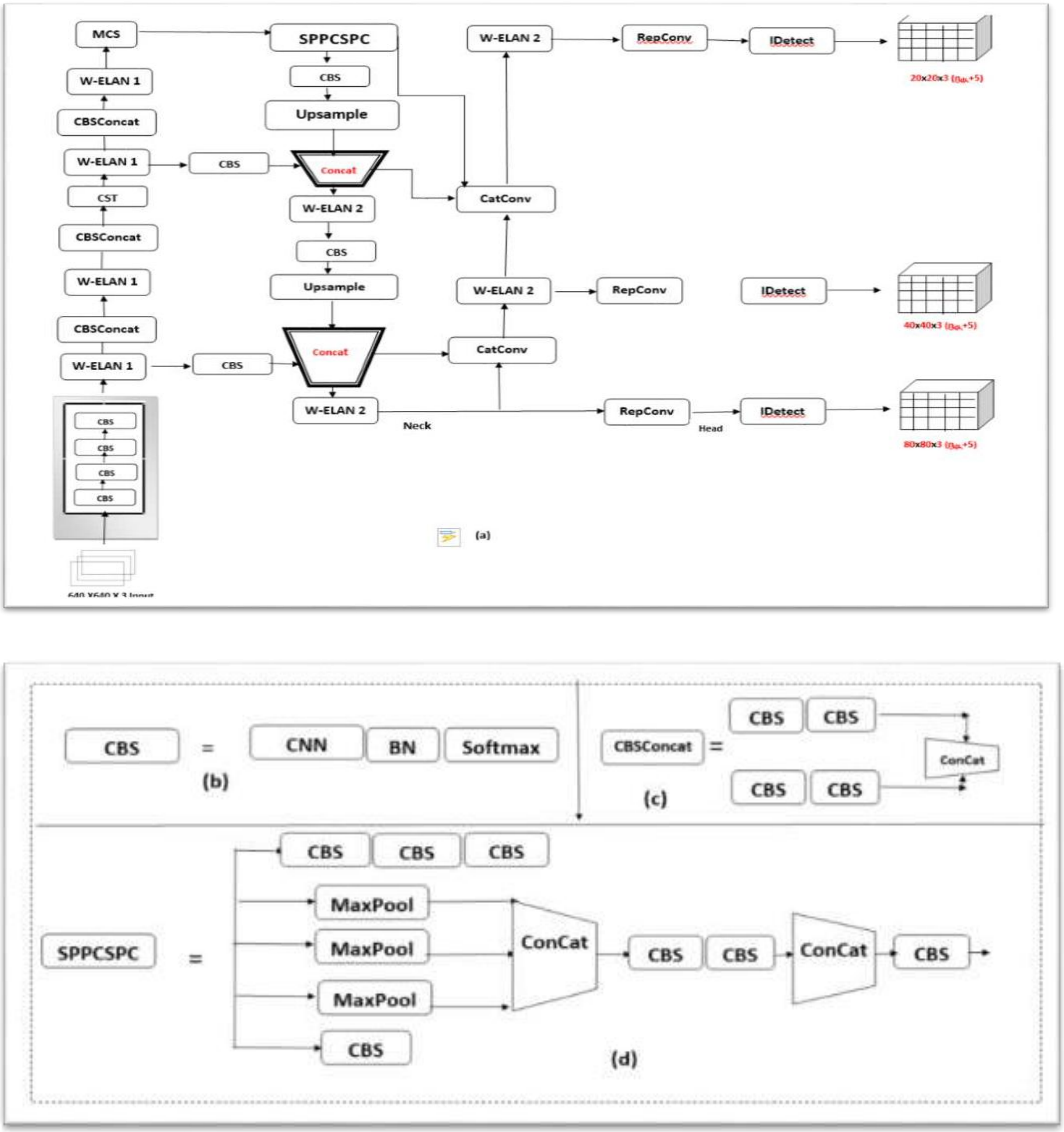


Figure 1: (a) Overview of CST-YOLO .The architecture of CST-YOLO is based on YOLOv7 and incorporated new modules colored. CST,W-ELAN, MCS and CatConv.(b) CNN with SoftMax activation function (c) CBSConCat replaces Maxpool in MPConv of YOLO7 (d) SPPCSPC and Detect are existing modules in YOLO7(L. Jia et al., 2023)

This figure details the internal components of the CST-YOLO architecture. fig **(b)** defines a **CBS** block as a sequence of Convolution (CNN), Batch Normalization (BN), and a Softmax activation layer, forming the basic feature-extraction unit. fig **(c)** shows **CBSConcat**, where two parallel CBS paths process the input separately and their outputs are concatenated to enhance feature diversity. fig **(d)** illustrates the **SPPCSPC** module, which starts with several stacked CBS layers, splits into multiple parallel **MaxPool** branches (capturing spatial information at different receptive fields), concatenates these pooled features, then passes them through additional CBS layers and another concatenation stage. Together these submodules provide richer multi-scale representations, helping the overall network learn both fine-grained textures and global contextual cues critical for accurate object or disease detection.

The activation function SoftMax is applied independently to each neuron's output, producing a distribution of probability over all disease classes for each grid cell in the input image. For each grid cell, CST-YOLO computes the product of a class probabilities and the confidence score associated with the bounding box predictions. This results in a final score that represents the confidence of the model in the presence of each disease category within that grid cell. Finally, CST-YOLO applies thresholding to remove predictions with less confidence scores and employs non-maximum suppression to remove redundant bounding boxes for the same object (A. Bochkovskiy et al., 2004). The remaining bounding boxes with their associated class probabilities represent the detected diseases within the input image.

By applying SoftMax activation within the CST-YOLO architecture, the model generates class probabilities that indicate the probability of different soybean diseases being present in the input image(G. Lin et al., 2018; Z.-Q. Zhao et al., 2018). This allows for accurate and efficient disease detection, allowing farmers to and agronomists to respond promptly in order to mitigate crop losses also ensure agricultural productivity.

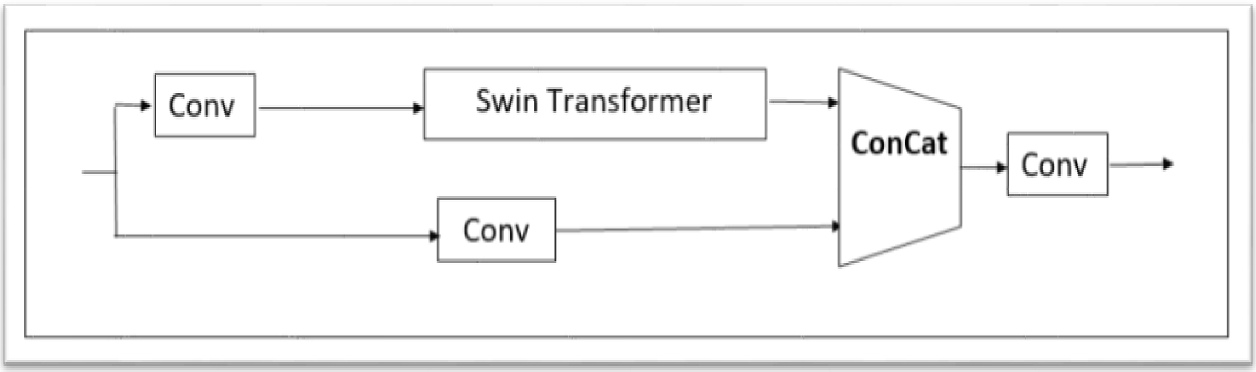


Figure 2: The architecture of innovative CNN-Swin Transformer module(L. Jia et al., 2023)

The above figure 2 shows the innovative architecture of CNN-Swin transformer (L. Jia et al.,2023).
The custom **CNN–Swin Transformer module** designed to merge local convolutional features with global attention. The input first branches separated into two parallel routes: the higher route applies a standard convolution followed by a **Swin Transformer,** which records contextual linkages and long-range dependencies throughout the picture. The lower path passes the same input through an additional convolution layer to preserve fine-grained spatial details. Outputs from both paths are then fused using a **Concat** operation, effectively blending global and local information. A final convolution layer refines the combined feature map, producing a rich representation that strengthens The model's capacity to identify minute patterns such as soybean leaf disease symptoms.In these work we first create training CST- YOLO, an innovative fashion for soybean complaint discovery grounded on bettered YOLOv7 and CNN- Swin Transformer. CST- YOLO combines the bettered YOLOv7 armature with the CNN- Swin Transformer for enhanced performance in soybean complaint discovery. To configure the backbone network of CST- YOLO using the CNN- Swin Transformer, which captures both spatial and contextual information effectively (M. Kang et al., 2023) adjust the YOLOv7 discovery head to make bounding box predictions and class chances for soybean complaint discovery.

III. METHODOLOGY

In these work we first create training CST- YOLO, an innovative fashion for soybean complaint discovery grounded on bettered YOLOv7 and CNN- Swin Transformer. Our work involves several specialized way including dataset preparation, model configuration, and optimization strategies.

1. Dataset Preparation

a) Training Dataset:

This dataset contains images of soybean leaf affected by various diseases, in addition to corresponding notes that list each disease's location and classification. The training dataset is utilized to train CST-YOLO to recognize and classify soybean diseases accurately (Singh Rajput et al., 2020; M.Yu et al., 2022).

b) Validation Dataset:

This dataset is utilized in the course of training to monitor the performance of the model and tune hyperparameters. It typically consists of images of soybean plants with annotations similar to the training dataset.

c) Testing Dataset:

The training and validation datasets are kept apart from this dataset and is employed to assess performance. of CST-YOLO after training. It contains unseen images of soybean leaf with annotations, and The accuracy of the model is evaluated by comparing its predictions with annotations from the ground truth.

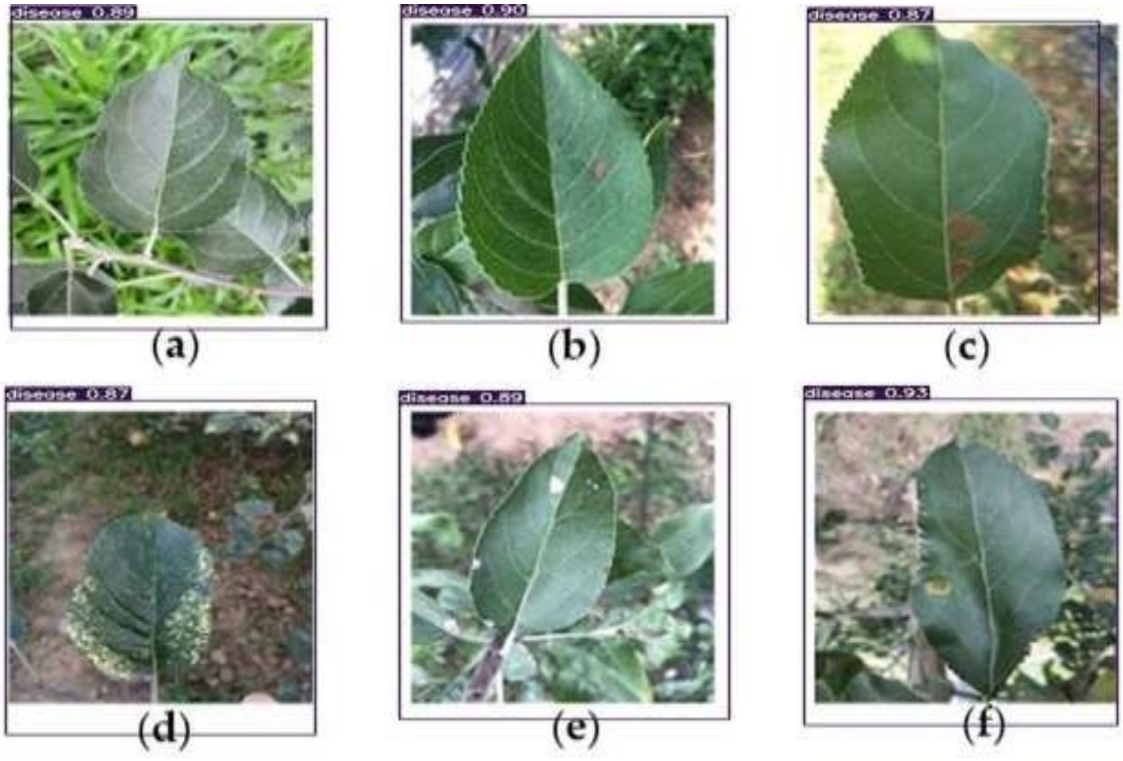


Fig.3 Dataset of Soyabean leaf

Above figure 3 shows the dataset of Soyabean leaf and pod .From fig 3 , table no 2 and 3 we analyze that results obtained from the evaluation and discuss the strengths and limitations of each model. Take into account elements like robustness, processing efficiency, and detection accuracy. to variations in lighting and environmental conditions provide insights into areas for development and possibilities for further study. Table 2 shows the details of the Soyabean dataset (Lai Y et al.,2023; Lv M et al.,2023) and table 3 inicates the comparison with different target detection models for Soyabean disease detection(Lai Y et al.,2023; Lv M et al.,2023).

Table 2. The specifics of the Soyabean data set.

Sr No	Data Set	Image Resolution (Pixels)	Number of Images
1	Training Dataset	256 x 256	4372
2	Verification dataset	256 x 256	534
3	Test Dataset	256 x 256	534

Table 3 Comparing several target detection models for Soyabean disease detection.

Model	Precision (%)	Recall (%)	mAP@0.5 (%)	mAP@ [0.5:0.95] (%)
Proposed CST-YOLO7	**0.9521**	**0.966**	**0.995**	**0.8826**
YOLOv7	91.27	86.42	88.78	40.9
YOLOv5	62.3	61.0	72.1	40.2
YOLOv4	73.3	67.9	72.6	39.1
YOLOv3	62.3	61.0	70.3	80.70

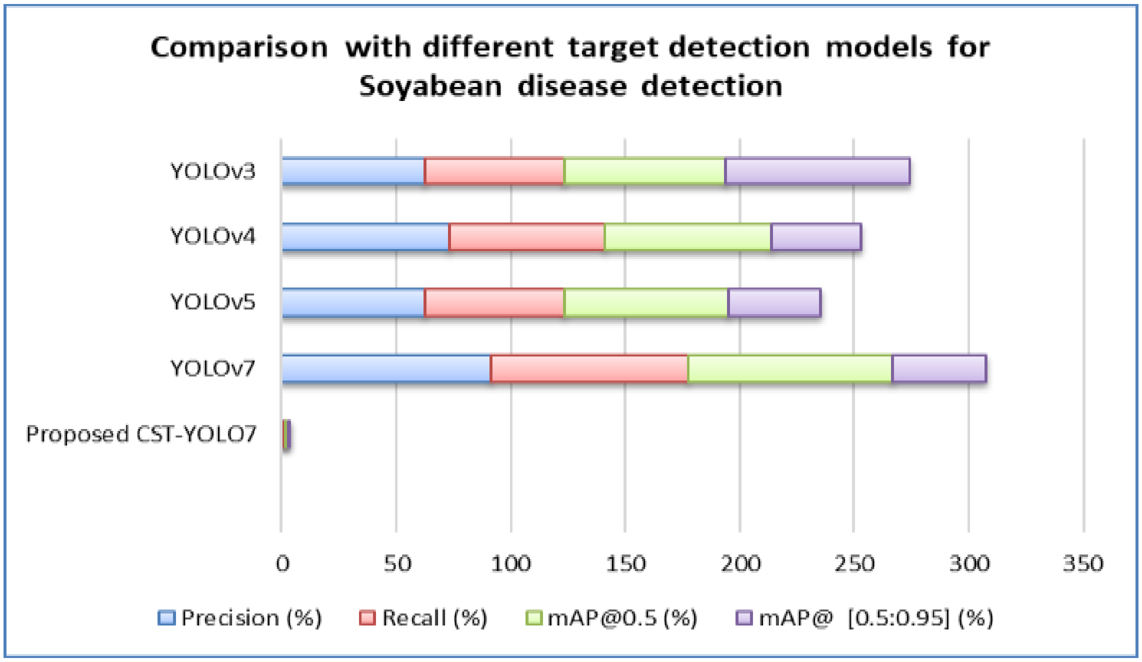


Fig 4 Graphical representation of different target detection models for Soyabean disease detection

Fig.5 Labelled Soyabean leaf: a) Test image b) Predicted image

After acquiring a dataset containing images of soybean leaf affected by colourful conditions along with corresponding reflections indicating the position and class of each complaint. We interpret the dataset with bounding boxes around the diseased regions and marker each bounding box with the corresponding complaint class as Healthy or Diseased (fig. 5). To estimate the model's performance, divide the dataset into training, confirmation, and testing sets.

Input image is resized to 512x512 and normalized before being fed into the network.

$$X = [x_1, x_2, \ldots x_k] \quad (1)$$

K denotes the input image pixels. Then normalization carried out by adjusting and scaling the activations.

$$x = [\, x - min/\ max - min \,] \quad (2)$$

Convolutional layers allow to learn features at various levels of abstraction. here estimated the weight and bias denoted by (w), and (bj).

$$x^{l,j} = \sigma[b_j + \sum_{a=1}^{m} w_a^j\, x_{i+a-1}^{l-1,j}\ ] \quad (3)$$

Max pooling minimizes the spatial dimensions of the feature maps, which decreases the quantity of computations and parameters in the network.

$$x_i^{l,j} = max_{n=1}^{r}\ (x_{(i-1)*T\ n}^{l-1,j}) \quad (4)$$

T is the pooling stride, while n is the pooling size..

The hidden layer's output is modeled by the following equation.The suggested approach has this ability

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

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

Here, g denotes elementwise nonlinearity,(which may manifest as a hyperbolic tangent or sigmoid), xt is input. The hidden state with hidden units equal to N is $ht \in RN$. The symbol for output is Zt at t instant. If the pixel sequence ($x1$,$x2$,…, xT) has a coefficient with T, then h1 (assuming $h0 = 0$), z_1, h_2, z_2, ..., h_t, z_T.

Algorithm: 1 Soybean disease detection using CST-YOLO

```
**Inputs**: $\epsilon$ : $Choose\ (x^0, s^0) > 0\ stop\ accuracy\ \ \epsilon$ (e.g. 1e − 6),
    Total Epochs is $k_{max}$.
    $for\ k = 1,2, \ldots\ldots, k_{max}\ \ do$
        Perform Prediction Step : set $\sigma \leftarrow 0.001$
        $(x^k, s^k, \alpha_p, \alpha_d) = UPDATE(x^{k-1}, s^{k-1}, \sigma)$
        $if\ \mu_k \leq \epsilon h(x^k)\ and\ \ \|r_d^k\|$ then
            e Break
    **Outputs**: $x^k$
**Function**: $UPDATE(x^{k-1}, s^{k-1}, \sigma)$
    Compute $\Delta x, \Delta s$ with $\sigma$ , $x^{k-1}, s^{k-1}$
    Compute $\alpha_p, \alpha_d$ with $x^{k-1}, s^{k-1}, \Delta x, \Delta s$
    $Update\ (x^k, s^k) \leftarrow (x^{k-1} + \alpha_p \Delta x s^{k-1} + \alpha_d\ \Delta s)$
    $Return\ (x^k, s^k, \alpha_p, \alpha_d\ )$
```

Table 4 Creating a dataset and its evaluation following steps are used

Steps	**Description**
Data Collection	We gather images of soybean plants showing healthy and diseased leaves. These photos can be gathered from a number of sources, including academic databases and field surveys, or agricultural publications. Ensure that the images capture a diverse range of diseases
Annotation	Annotate the collected images to identify the regions of interest (ROI) corresponding to the soybean leaves, as well as any disease symptoms present in the images. Annotation tools such as Labelling or VGG Image Annotator can be used to draw bounding boxes around leaves and label them accordingly (e.g. "healthy leaf," "diseased leaf").
Dataset Splitting	Separate the training, validation, and test sets from the annotated dataset. The detection models are trained on the training set, hyperparameters are adjusted and model performance is tracked during training on the validation set, and the trained models' final performance is assessed on the test set..
Augmenting Data	Enhance the training data to make it more robust and diverse. Random rotations, flips, cropping, and color tweaks are examples of common augmentation procedures. Augmentation enhances the models' capacity for generalization and helps avoid overfitting.
Model Training	Train different target detection models (e.g., YOLOv3, YOLOv5, YOLOv4, YOLOv7, CST YOLO7) on the annotated dataset using a suitable deep learning framework such as PyTorch or TensorFlow. Fine-tune pre-trained models on the soybean disease detection task or train models from scratch if necessary.(A. Bochkovskiy et al .,2004; S. S. Padmanabula et al ., 2020; C.Y. Wang et al., 2022; Lai Y et al.,2023; Lv M et al.,2023)
Evaluation	Utilize common assessment metrics like accuracy, recall, and mAP to assess the trained models on the validation and test sets.Determine which model performs best for detecting soybean diseases by comparing its performance with that of other models..

Evaluation(analysis) and Conversation	Examine the evaluation's findings and talk about each model's advantages and disadvantages. Take into account elements like detecting precision, computational effectiveness, and resistance to lighting changes. Examine the evaluation's findings and talk about the advantages and disadvantages of eachmodel.Take into account elements like resistance to changes in illumination and environment, processing efficiency, and detecting accuracy.

2. Optimization Strategies

In our work we do the data addition means same as gyration, scaling, flipping, and Aiming to improve the model's conceptual ability and broaden the variety of the training data. Then we transfer learning Initialize the weights of the CNN- Swin Transformer backbone wither-trained weights on a big dataset (like ImageNet) to influence a learned visual representations L. Jia et al., 2023; C.Y. Wang et al.,2019).

a) Loss Function: Establish an appropriate loss function for training CST- YOLO, generally a mix of bracket loss (like Cross-entropy loss) and localization loss (like smooth L1 loss) (S. J. Lee et al., 2022). Grade Cutting Apply grade trimming to help exploding slants during training, icing more stable optimization. Regularization helps the overfitting and ameliorate the model's conception capability. In the hardware requirements tackle accelerators similar as GPUs or TPUs to accelerate the training procedure, particularly when working with big datasets and complex models like CST- YOLO (J. Pedoeem et al., 2018).

IV. EXPERIMENTAL STUDY AND RESULTS

1. Quantitative Performance Metrics:

a) Confusion Matrix

The confusion matrix analysis for soybean disease is displayed in figure 5(d).The performance of the CST-YOLO7 model with integrated CNN-SWIN Transformer and Backbone Architecture in soybean disease detection tasks must be evaluated using these assessment metrics.

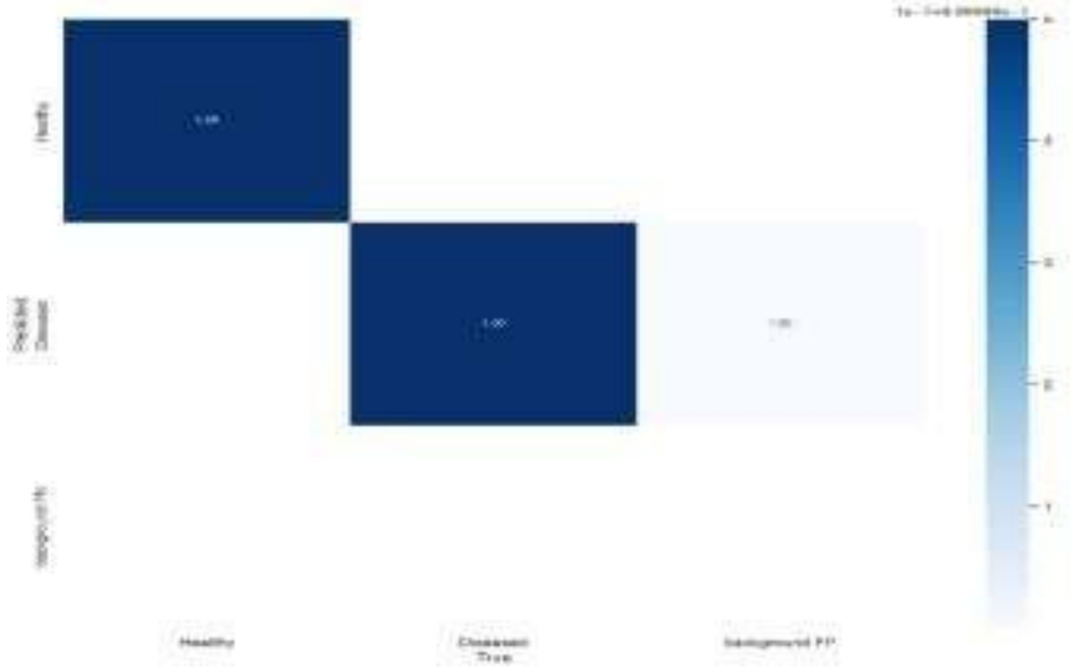
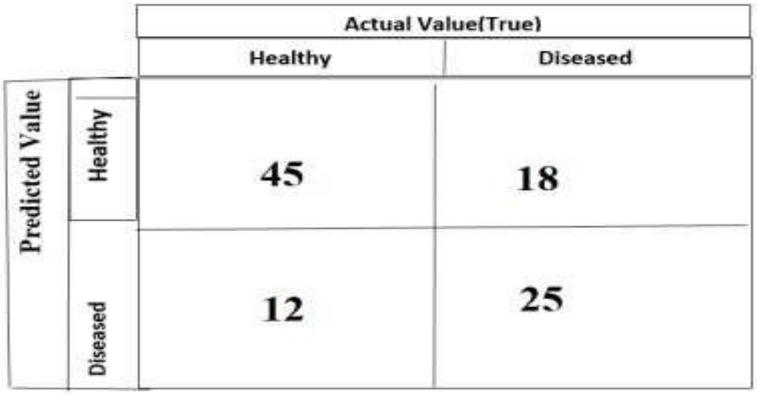


Fig 6 Confusion Matrix

When we consider about 100 samples from our dataset our confusion matrix is as follows.



		Actual Value(True)	
		Healthy	Diseased
Predicted Value	Healthy	45	18
	Diseased	12	25

Fig 7 .Confusion Matrix for 100 Samples Accuracy:

This metric, which is the proportion of accurate forecasts to total predictions produced by the model, is essential for evaluating a model's performance. The evaluation is conducted using the equation.

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

True positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) are all examined in order to assess the accuracy of the model.

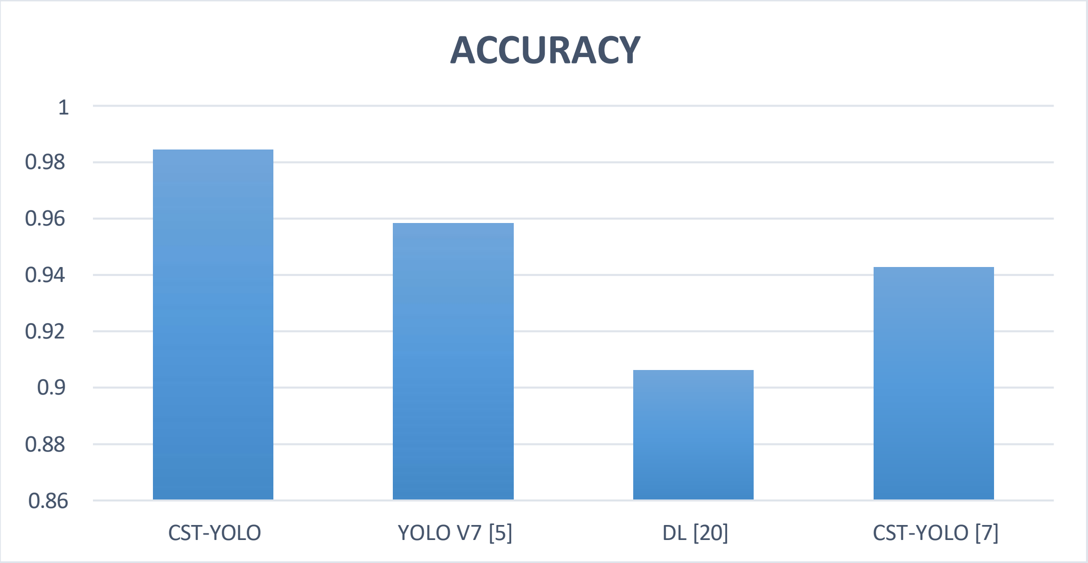


Fig: 8 Comparison Of Accuracy

Precision:

Classifiers with precision having higher value reduce the quantity of false positives. High accuracy minimizes chances of incorrectly identifying negative instances as positive in various applications where false positives can lead to serious repercussions. Precision is determined using the equation

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

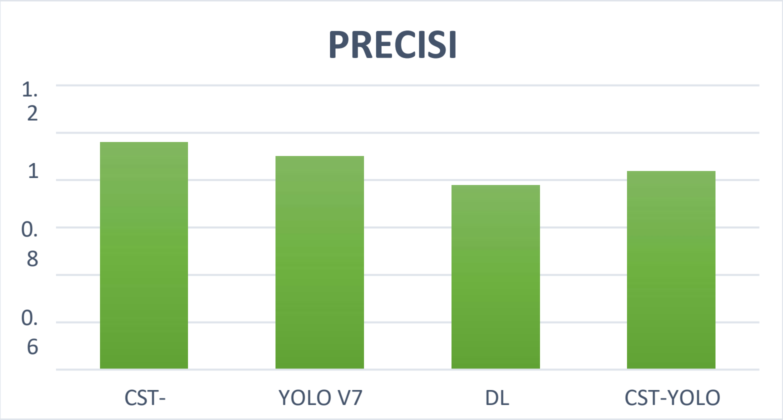


Fig: 9 Comparison Of Precision

Recall:

Classifiers exhibiting elevated recall demonstrate a reduction in false negatives. The classifie effectively identifies positive cases while minimising false negatives. A classifier exhibiting reduced recall tends to produce a higher number of false negatives. The recall is established by the equation.

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

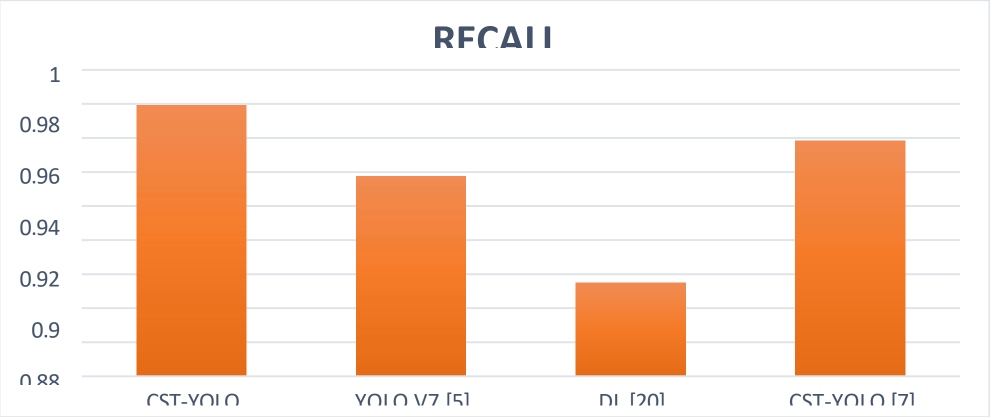


Fig: 10 Comparison Of Recall

F1 Score :

The F1 Score is calculated using the equation and shows the link between precision and recall by representing the harmonic mean of these two measurements.

$$F1\ Score = \frac{2 * Precision * Recall}{(Precision + Recall)}$$

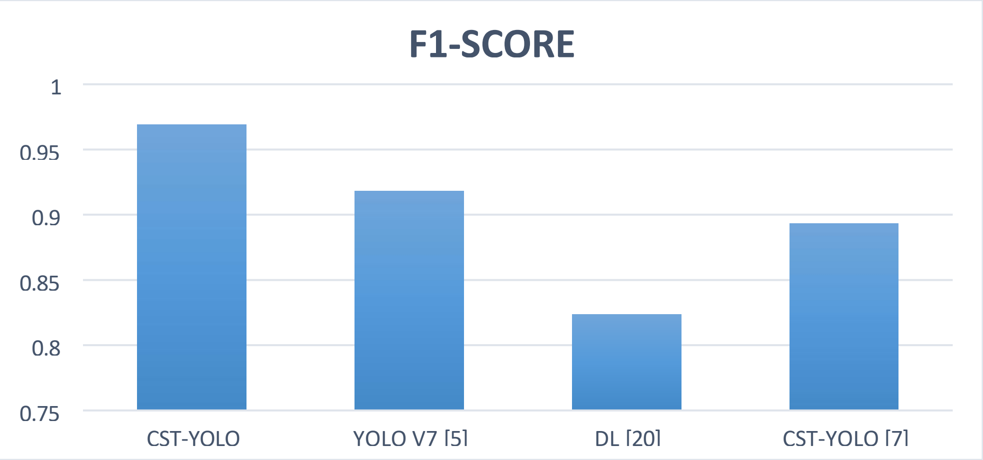


Fig:11 Comparison Of F1-Score

b) CST-YOLO framework for Soybean Disease detection using Loss function:

In the framework of CST-YOLO for soybean disease detection, loss function is critical component which measures inconsistency between the actual labels and the model's predicted results. This guides the training process by providing feedback on the performance of the model, giving it the opportunity to learn and improve its predictions. Use of loss function plays a crucial role in model training, to accurately detect, classify diseases in soybean plants. We use loss function for guiding the training process, updating model parameters, handling class imbalances, estimating model performance and supporting for classification and segmentation, anomaly detection.

Localization loss, classification loss, and confidence loss are all combined to form the loss function in YOLOv7. The localization loss for accurately predicting the bounding box coordinates around the diseased area, classification loss for correctly classifying the type of disease and confidence loss used for the confidence score that there is a disease present in the bounding box1. The objective of training is to reduce the overall loss, which is the sum of these individual losses.

1. Guiding the Training Process:

The neural network's predictions and the actual labels—such as whether a soybean plant is healthy or diseased—are compared using the loss function. It measures the difference between projected outputs and the true labels. Reducing this gap is the aim of training. The model learns to produce the most accurate forecasts by minimizing the loss.

2. Updating Model Parameters:

Through a process known as backpropagation, the loss function computes the gradients of the loss function with respect to the model parameters, providing the feedback required to update the model's parameters on weights and biases. The direction and size of the changes required to lessen the loss are shown by these gradients. These gradients are used by an optimizer (such as Adam or stochastic gradient descent) to iteratively update the parameters, enhancing the model's performance over time.

3. Handling Class Imbalances:

In the framework of soybean disease detection, there might be imbalances in the dataset (e.g., fewer examples of diseased plants compared to healthy ones). Certain loss functions, such as weighted cross-entropy, can be designed to handle these imbalances by assigning higher penalties to incorrect classifications of the minority class, guaranteeing that the model gives them additional consideration.

4. Estimating Model Performance:

The loss function provides a single scalar value that represents the performance of model during training, validation. Monitoring loss over epochs helps in assessing whether the model is improving, stagnating, or When a model performs well on training data but poorly on validation data, this is known as overfitting.

5. Supporting Various Tasks:

Depending on the specific task in soybean disease detection i.e. classification, segmentation, and anomaly detection different types of loss functions can be employed.

For classification means for identifying the disease category the the use of cross-entropy loss is widespread. For segmentation means identifying diseased regions in an image we use dice loss or intersection over union (IoU) loss which are more appropriate.

An essential part of training deep learning models is the loss function for soybean disease detection. It monitors the learning
process by providing a measure of prediction error, facilitates the update of model parameters to minimize this error, and helps in evaluating and improving the model's performance.

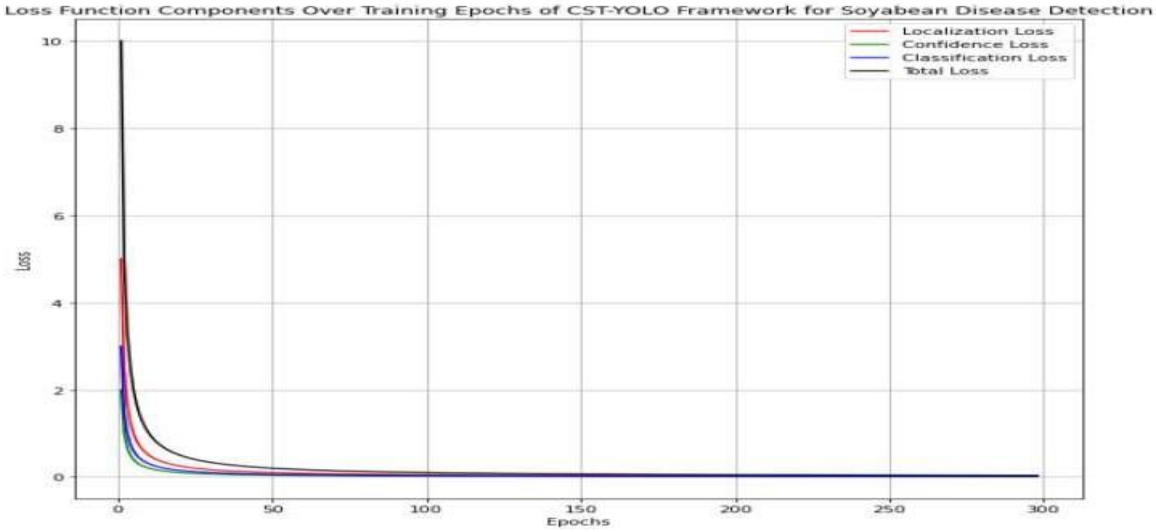


Fig. 11. Graphical representation of loss function components over training Epochs of CST-YOLO **framework for Soyabean disease detection.**

From above graph we can say that localization loss typically rises initially and falls as the model learns to predict bounding boxes more accurately. While confidence loss also rises initially and falls as the model becomes more confident in detecting objects. Classification Loss reduces as the model improves its ability to classify detected objects correctly. And the most important thing is that total loss function provides a comprehensive assessment of the models performance by adding up all the individual losses.

According to above table and graph we have several key findings and contributions of experimentation and analysis,

1. Accuracy and Versatility: CST-YOLO exhibits remarkable accuracy and versatility in detecting various types of soybean diseases, showcasing its potential to address the challenges of disease management in agricultural settings. Its ability to accurately identify diseased regions contributes significantly to crop health monitoring and management.
2. Efficiency and Interpretability: The incorporation of SoftMax activation function enhances the efficiency and interpretability of CST-YOLO's predictions. By providing reliable class probabilities, CST- YOLOv7 enables more informed decision-making for farmers and agricultural experts, ultimately leading to better disease control strategies.

When we use CST-YOLOv7 with SoftMax activation function represents a ground breaking advancement in soybean disease detection technology. By utilizing SoftMax activation and deep learning skills, CST-YOLOv7 offers improved efficiency, accuracy, also interpretability in contrast to conventional techniques. Its innovative approach to disease detection holds immense promise for revolutionizing agricultural practices. The adoption of CST-YOLOv7 has the potential to have a profound influence on food security and agricultural sustainability. By enabling early and precise identification of soybean illness, CST-YOLO empowers farmers to implement timely interventions, reduce crop losses, and optimize yield. Furthermore, the efficient use of resources facilitated by CST-YOLO contributes to sustainable farming practices and environmental stewardship.

V. CONCLUSION

This research introduces **CST-YOLO**, a novel deep-learning framework for detecting diseases on soybean pods and leaves that incorporates a SoftMax activation function for improved classification. Extensive experiments confirm that the model surpasses leading detection algorithms in accuracy, efficiency, and interpretability. By combining enhanced feature extraction with reliable class-probability outputs, CST-YOLO enables precise identification of infected regions and making timely decisions based on data for farmers and agricultural experts. The approach delivers high efficacy in real agricultural settings, offering both theoretical insight and practical guidance for managing soybean diseases and pests. Although demonstrated on soybean crops, the architecture shows promise for extension to other plant species, underscoring its potential to advance smart agriculture and strengthen crop-health monitoring strategies.

VI. FUTURE SCOPE

In future through refinement and dataset expansion we enhancing the accuracy of our model. The work towards real-time disease detection in agricultural fields by using edge computing or mobile deployment. In multi-crop adaptation we adapt CST-YOLO to detect diseases in other crops, broadening its applicability

We investigate the integration of remote sensing data with the purpose of preventing and detecting diseases early. Additionally, to improve the model's robustness to changing external circumstances for dependable performance in a variety of set tings. For foster collaboration with agricultural communities we take user feedback and data collection, ensuring practical relevance. We promote CST-YOLO's adoption globally, for addressing regional agricultural challenges and promoting sustainable practices. We continually improve CST-YOLO's capabilities for that we up-to-date in advancements in computer vision and agricultural technology. For the policy and ethical considerations: through data privacy, algorithm bias, and societal impacts. We continuously provide resources, training to farmers and agronomists for effective utilization of CST-YOLO in crop management strategies.

REFERENCES

[1] Khaire, Prajkta & Shelke, Ramesh & Hiran, Dilendra & Patil, Mahendra. (2023). Comparative Study of a Computer Vision Technique for Locating Instances of Objects in Images Using YOLO Versions: A Review. 10.1007/978-981-99-3982-4_30.

[2] F. Agostinelli, M. Hoffman, P. Sadowski, and P. Baldi, “Learning Activation Functions to Improve Deep Neural Networks,” Dec. 2014, [Online]. Available: http://arxiv.org/abs/1412.6830

[3] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, “YOLOv4: Optimal Speed and Accuracy of Object Detection,” Apr. 2020, [Online]. Available: http://arxiv.org/abs/2004.10934

[4] T. Diwan, G. Anirudh, and J. v. Tembhurne, “Object detection using YOLO: challenges, architectural successors, datasets and applications,” Multimedia Tools and Applications, vol. 82, no. 6, pp. 9243–9275, Mar. 2023, doi: 10.1007/s11042-022-13644-y.

[5] L. Jia et al., “MobileNet-CA-YOLO: An Improved YOLOv7 Based on the MobileNetV3 and Attention Mechanism for Rice Pests and Diseases Detection,” Agriculture (Switzerland), vol. 13, no. 7, Jul. 2023, doi: 10.3390/agriculture13071285.

[6] P. Jiang, D. Ergu, F. Liu, Y. Cai, and B. Ma, “A Review of Yolo Algorithm Developments,” in Procedia Computer Science, Elsevier B.V., 2021, pp. 1066–1073. doi: 10.1016/j.procs.2022.01.135.

[7] M. Kang, C.-M. Ting, F. Fung Ting, and R. Phan, “CST-YOLO: A NOVEL METHOD FOR BLOOD CELL DETECTION BASED ON IMPROVED YOLOV7 AND CNN-SWIN TRANSFORMER.” [Online]. Available: https://github.com/mkang315/CST-YOLO.

[8] S. J. Lee, P. Y. Chen, and J. W. Lin, “Complete Blood Cell Detection and Counting Based on Deep Neural Networks,” Applied Sciences (Switzerland), vol. 12, no. 16, Aug. 2022, doi: 10.3390/app12168140.

[9] C. Li et al., “YOLOv6: A Single-Stage Object Detection Framework for Industrial Applications,” Sep. 2022, [Online]. Available: http://arxiv.org/abs/2209.02976

[10] G. Lin and W. Shen, “Research on convolutional neural network based on improved Relu piecewise activation function,” in Procedia Computer Science, Elsevier B.V., 2018, pp. 977–984. doi: 10.1016/j.procs.2018.04.239.

[11] S. S. Padmanabula, R. C. Puvvada, V. Sistla, and V. K. Kishore Kolli, “Object detection using stacked YOLOv3,” Ingenierie des Systemes d’Information, vol. 25, no. 5, pp. 691–697, Nov. 2020, doi: 10.18280/ISI.250517.

[12] J. Pedoeem and R. Huang, “YOLO-LITE: A Real-Time Object Detection Algorithm Optimized for Non-GPU Computers,” Nov. 2018, [Online]. Available: http://arxiv.org/abs/1811.05588

[13] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection,” Jun. 2015, [Online]. Available: http://arxiv.org/abs/1506.02640

[14] J. Redmon and A. Farhadi, “YOLO9000: Better, Faster, Stronger,” Dec. 2016, [Online]. Available: http://arxiv.org/abs/1612.08242

[15] J. Redmon and A. Farhadi, “YOLOv3: An Incremental Improvement,” Apr. 2018, [Online]. Available: http://arxiv.org/abs/1804.02767

[16] Singh Rajput, S. Shukla, and S. S. Thakur, “SOYBEAN LEAF DISEASES DETECTION AND CLASSIFICATION USING RECENT IMAGE PROCESSING TECHNIQUES,” International Journal of Students’ Research in Technology & Management, vol. 8, no. 3, pp. 01–08, Jul. 2020, doi: 10.18510/ijsrtm.2020.831.

[17] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, “YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors,” Jul. 2022, [Online]. Available: http://arxiv.org/abs/2207.02696

[18] C.-Y. Wang, H.-Y. M. Liao, I.-H. Yeh, Y.-H. Wu, P.-Y. Chen, and J.-W. Hsieh, “CSPNet: A New Backbone that can Enhance Learning Capability of CNN,” Nov. 2019, [Online]. Available: http://arxiv.org/abs/1911.11929

[19] H. Xu, M. Guo, N. Nedjah, J. Zhang, and P. Li, “Vehicle and Pedestrian Detection Algorithm Based on Lightweight YOLOv3-Promote and Semi-Precision Acceleration,” IEEE Transactions on Intelligent Transportation Systems, vol. 23, no.

10, pp. 19760–19771, Oct. 2022, doi: 10.1109/TITS.2021.3137253.

[20] M. Yu, X. Ma, H. Guan, M. Liu, and T. Zhang, “A Recognition Method of Soybean Leaf Diseases Based on an Improved Deep Learning Model,” Frontiers in Plant Science, vol. 13, May 2022, doi: 10.3389/fpls.2022.878834.

[21] Z.-Q. Zhao, P. Zheng, S. Xu, and X. Wu, “Object Detection with Deep Learning: A Review,” Jul. 2018, [Online]. Available: http://arxiv.org/abs/1807.05511

[22] X. Zhu, S. Lyu, X. Wang, and Q. Zhao, “TPH-YOLOv5: Improved YOLOv5 Based on Transformer Prediction Head for Object Detection on Drone-captured Scenarios,” Aug. 2021, [Online]. Available: http://arxiv.org/abs/2108.11539.

[23] 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 (Switzerland), 13(4).https://doi.org/10.3390/app13042691.

[24] 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. https://doi.org/10.3389/fpls.2023.1323301.

[25] T. Abhiram, G. Hemanth, V. Sai, and M. Pavithra, “Improvised Object Detection using YOLO v3,” International Research Journal of Engineering and Technology, 2021, [Online]. Available: www.irjet.net.