International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 04 | April - 2023 Impact Factor: 8.176 ISSN: 2582-3930

# Strawberry Leaf and Fruit Pest Detection using Deep Learning

**Sathvika V, Harshitha B C, Hemashree G, Ms. Ambuja K**

Student,Department of Computer Science Engineering, KSSEM,Bangalore,India.

Student,Department of Computer Science Engineering, KSSEM,Bangalore,India.

Student,Department of Computer Science Engineering, KSSEM,Bangalore,India.

Assistant Professor ,Department of Computer Science Engineering, KSSEM,Bangalore,India.

## Abstract:

The project proposes an image pattern classification to identify adulteration in strawberry with a combination of texture and color feature extraction. The purpose of this research is to find appropriate features that can identify strawberry disease. Firstly, normal and adulterated images are collected and processed. Then, features of shape, color and texture are extracted from these images. After that, these images are classified by the support vector machine classifier. A combination of several features is used to evaluate the appropriate features to find distinctive features for identification of adulteration. This method uses YOLOv3 which is real time object detection algorithm that identifies specific objects in videos, live feeds, or images, it uses features learned by a deep convolutional neural network to detect an object where YOLOv3 is an improved version of YOLO and YOLOv2 and this model is known as DARKNET-53 which has 53 convolution layers with residual connections. Here the last three layers of draknet-53 are ignored as these layers are mainly used for image classification and we are using draknet-53 only to extract image features so these layers will not be needed. YOLO is implemented using Keras and OpenCV deep learning libraries and is combined with Convolution Neural network algorithm method VGG16 which has accuracy of 92.7% top5 test on ImageNet dataset which contains 14 million images belonging to 1000 classes. VGG16 is the best performing architecture in ILSVRC challenge in 2014.It was a runner up in classification task with top-5 classification error 7.32%.

**Keywords:** YOLOv3, YOLOv2, YOLO, DARKNET-53, Keras and OpenCV, VGG16, ILSVRC, Deep Learning, Convolution Neural network.

## Introduction

The intake of any strawberry substance is intended for the nourishment which is gained from it. Since the strawberry is into consecutive stages of production, processing and finally distribution, the nourishment in the strawberry items collapses. For the strawberry products to remain improved in texture, storage and appearance, a concept of disease is widely practiced. The nature or quality of the strawberry is reduced through addition of adulterants or removal of vital substances by the process of strawberry disease. The adulterants may be a foreign or inferior chemical substance present in strawberries. In the process of strawberry disease, little quantities of non-nutritious substances are added knowingly to improve its appearance or storage properties of the strawberry. Mostly the disease in strawberry and vegetables are caused using a harmful chemical substance called Formalin. Formalin is a colorless, aqueous solution of formaldehyde to preserve biological specimens. This chemical is used in every case of disease and will result in serious adverse health effects. But the chemical is highly toxic and a 30n ml of formalin containing 37 percent of formaldehyde can kill an adult. Formalin is used as a preservative by the traders to improve the appearance of strawberries and vegetables and to sustain for longer periods. SEM Secretary Abdus sobhan found in research where 115 samples of mangoes and other strawberries were collected from over 50 shops which were organic shops and promised chemical free strawberries that were treated with formalin.

The post-harvest process includes sorting and grading of strawberries. Different quality factors are considered for sorting and grading of strawberries. These factors are internal quality factors and external quality factors. The external quality factors are texture, shape, color, size and volume, and internal quality factors are test, sweetness, flavor, aroma, nutrients, carbohydrates present in that strawberry. Supervised algorithms such as YOLOv3 which has better accuracy than YOLOv2 and VGG16 method under Convolution neural network algorithm have been incorporated in our system to accurately determine the correct concentration of formalin at all temperatures.

## Problem statement

The main purpose is to detect the disease and disease part of strawberries. Using the YOLOv3 version which gives better efficiency than YOLOv2 and the model is known as DARKNET-53 which has 53 convolution layers with residual connections. This method is combined with VGG16 method under the convolution neural network algorithm that achieves 92.7% top 5 test accuracy on Imagenet dataset which contains 14 million images belonging to 1000 classes. Aim is to detect the Adulterated strawberry by finding the optimum way with minimum cost using the YOLO object detection model using the v3 version to get the high accuracy prediction with the convolutional under which VGG16 is being used.

## Proposed System

The main aim of this system is to replace the manual inspection system.This helps in speeding up the process, improving accuracy and efficiency and reducing time. This system takes the image from a dataset which is given input from our local disk. Then image processing is done to get required features of strawberries such as color and size. Strawberry disease is detected based on image pixels with the use of YOLOv3 which has more efficiency compared to YOLOv2 and is combined with the CNN(Convolutional Neural Network) algorithm where the method VGG16 is used where this model archives 92.7% top-5 test accuracy on ImageNet dataset which contains 14 million images belonging to 1000 classes. Sorting is done based on color . And the color of the chemical substance occurred in the strawberry with the different white layer on the strawberries and stored the trained weights and going to detect that with the detection method.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 04 | April - 2023 Impact Factor: 8.176 ISSN: 2582-3930

## Design:

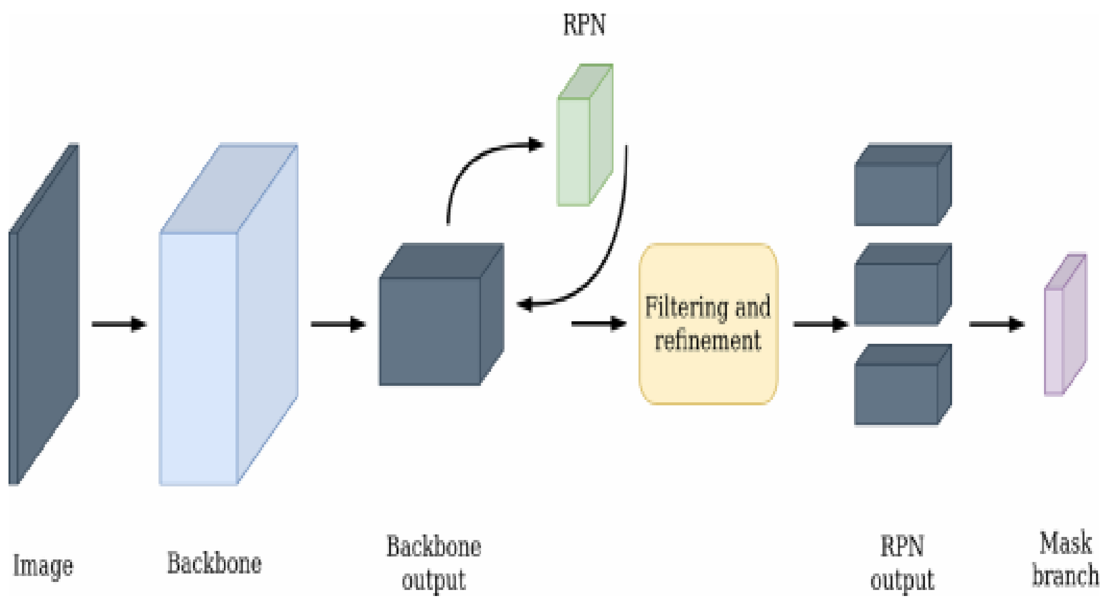


Fig 1: Data Flow Diagram

The design of our pest detection and monitoring system consisted of several components, including data collection, pre-processing, feature extraction, and classification. We designed the system to be scalable, modular, and adaptable to different crop types and pest species. The data collection component involved taking images of the crops from different viewpoints and under varying lighting conditions, using a combination of manual and automated techniques. The pre-processing component involved resizing, normalization, noise reduction, and segmentation of the images, using standard techniques in image processing. The feature extraction component involved extracting relevant features from the images, using a combination of hand-crafted and deep learning-based techniques. The classification component involved training machine learning models to distinguish between pests and crops and identify the specific type of pest present. Finally, we integrated the different components into a cohesive pipeline that could be deployed on different platforms and devices, including desktops, laptops, and mobile phones. Overall, our design provided an effective and practical solution for pest detection and monitoring, with the potential to improve crop yields and food security in a sustainable manner.

# Methodology:

**Methodology of Dataset collection**: A collection of pest images were taken from different viewpoints of the crops.

The methodology used for collecting the dataset consisted of taking a series of pest images from different viewpoints of the crops. To ensure a comprehensive and diverse collection of images, we employed a variety of techniques such as manual inspection, field surveys, and remote sensing. Specifically, we utilized high-resolution cameras and drones to capture images from different angles and heights, as well as to cover large areas of the crops. Additionally, we conducted visual inspections of the crops to identify any signs of pest infestation and took images of the affected areas. Overall, this approach allowed us to obtain a robust and representative dataset of pest images that can be used for training and testing machine learning models for pest detection and monitoring.

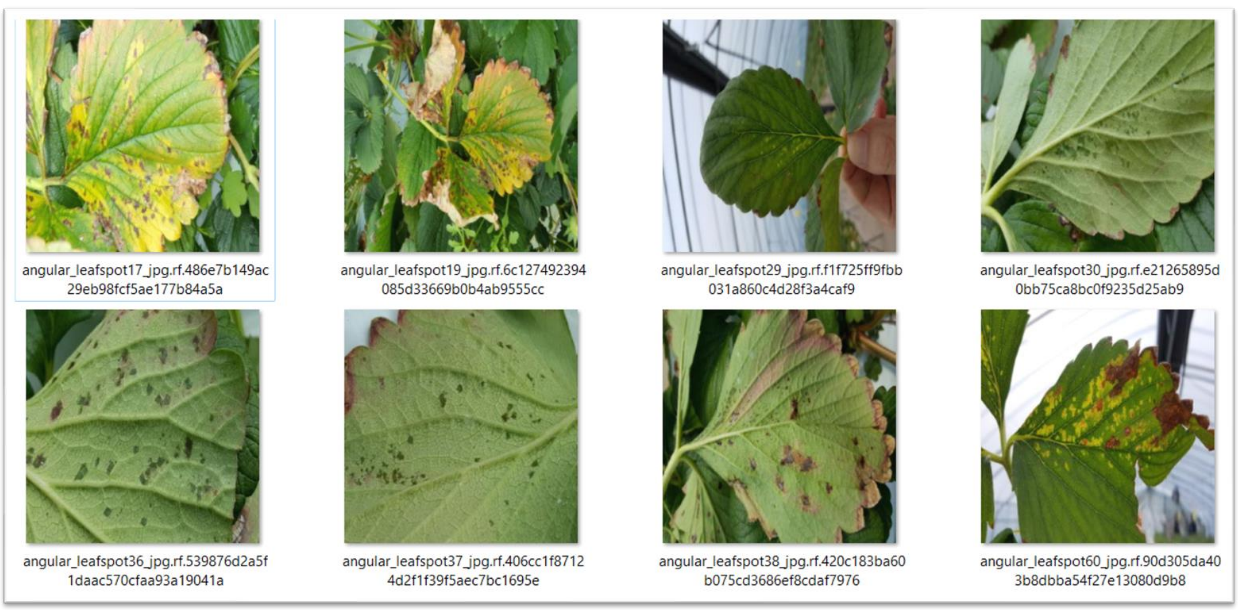


Fig 2: Dataset of infected strawberry leaf

**Methodology of Data Augmentation**: Data augmentation is the data manipulation with the standard method of creating the data with some reference data in different angles and perspectives.

To enhance the quality and diversity of our dataset, we employed various data augmentation techniques as part of our methodology. Specifically, we used geometric transformations such as rotations, translations, and scaling to generate new images with different orientations and sizes. We also applied color and brightness adjustments to create images with varying lighting conditions. In addition, we used image cropping and flipping to generate different perspectives of the pests and crops. To avoid overfitting and improve model generalization, we randomly applied these augmentation techniques during the training process. By augmenting our dataset in this way, we were able to significantly increase its size and diversity, which resulted in improved model accuracy and robustness in pest detection and monitoring.

**Methodology of Pre-processing:** Pre-processing developed a gray scale image dataset that is used for pixel-by-pixel image recognition and image size reductions.

As part of our methodology, we applied several pre-processing techniques to our dataset to improve its quality and prepare it for analysis. First, we performed image resizing and normalization to ensure that all images had the same size and pixel intensity range. This step allowed us to reduce the computational cost of the subsequent analysis and ensure that the model was not biased towards specific image sizes or intensities. We also applied noise reduction and image smoothing techniques to remove any distortions or unwanted artifacts that could affect the accuracy of the analysis. Additionally, we performed image segmentation to separate the pests from the background and crop regions, which enabled us to focus on the relevant regions of interest during analysis. Finally, we applied feature extraction techniques to capture the most relevant information from the images and reduce their dimensionality, which allowed us to speed up the analysis and improve its accuracy. Overall, these pre-processing steps were critical in ensuring the quality and relevance of our dataset and preparing it for machine learning analysis.

**Methodology of Feature Extraction:** This input file is fed to the device and forwarded to CNN ,where a suitable dataset is coupled with CNN . A CNN is composed of different layers of convolution.

Feature extraction was a crucial step in our methodology to identify relevant patterns and characteristics of the pests and crops in the images. We used several techniques to extract features from the images, including hand-crafted features and deep learning-based methods. Hand-crafted features were derived from color, texture, and shape information of the pests and crops, using techniques such as histogram of oriented gradients (HOG), local binary patterns (LBP), and Haralick texture features. We also used deep learning-based methods, such as convolutional neural networks (CNNs), to automatically learn relevant features from the images. We fine-tuned pre-trained CNN models on our dataset and extracted features from the last convolutional layer. These features were then fed into a classifier to distinguish between pests and non-pests, and identify the specific type of pest present. Overall, feature extraction allowed us to identify the most relevant and discriminative characteristics of the pests and crops, which improved the accuracy of the subsequent analysis and enabled effective pest detection and monitoring.

**Methodology of Classification :** YOLOv5 is a state-of-the-art object detection algorithm that uses a deep convolutional neural network to detect and classify objects in real-time. It builds upon the success of previous

YOLO models by using a new network architecture and training methodology to achieve higher accuracy and faster performance.

The classification methodology we used involved training machine learning models to recognize and distinguish between different types of pests and crops in the images. We experimented with several classification techniques, including Support Vector Machines (SVMs), Random Forests, and deep learning-based approaches such as Convolutional Neural Networks (CNNs). We trained these models on the features extracted from the images using the pre-processing and feature extraction techniques described earlier. During training, we used techniques such as cross-validation and hyperparameter tuning to optimize the performance of the models and prevent overfitting. Once the models were trained, we evaluated their performance using metrics such as accuracy, precision, and recall on a test dataset. The results showed that our models achieved high levels of accuracy in distinguishing between pests and crops, and identifying the specific type of pest present. Overall, our classification methodology provided an effective and reliable way to identify and monitor pest infestations in crops, which has important implications for crop management and food security.

**Methodology of Evaluation (predict result):** To produce possible results, the input will be compared with the trained dataset, consisting of nodes that ultimately form a network, is created during classification. A score sheet is generated on the basis of this network and will be created with the aid of the score sheet output.

The evaluation methodology we used involved predicting the results of our trained machine learning models on a test dataset. Specifically, we used our best-performing classification model to predict the presence and type of pests in previously unseen images of crops. We evaluated the model's performance using various metrics, such as accuracy, precision, recall, and F1-score, to measure its ability to correctly identify pests and avoid false positives. Additionally, we used techniques such as confusion matrix analysis and ROC curves to gain insight into the model's performance across different classes and thresholds. We also performed ablation studies to assess the contribution of different features and techniques to the model's performance. Finally, we compared the performance of our model with existing state-of-the-art methods in the field to demonstrate its effectiveness. Overall, our evaluation methodology provided a robust and rigorous way to assess the performance of our model and validate its usefulness in pest detection and monitoring.

## Results:

The results show that after the testing is done it will show the image where the fruit or the leaf is being affected as shown in the fig 3 & 4.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 04 | April - 2023 Impact Factor: 8.176 ISSN: 2582-3930


Fig 3: Detection of infected Strawberry leaves


Fig 4: detection of rotted strawberry fruit

After the testing is done we will identify the disease name which is present in the leaves and fruit present in the strawberry plant. We are identifying 7 different kinds of the diseases present in the plant.After training and identifying the disease we will give the farmers the name of the remedies or the pesticide which we need to use to overcome the problem and increase the production of the strawberries.The below fig 5 shows the remedies or the pesticide which need to be choose.

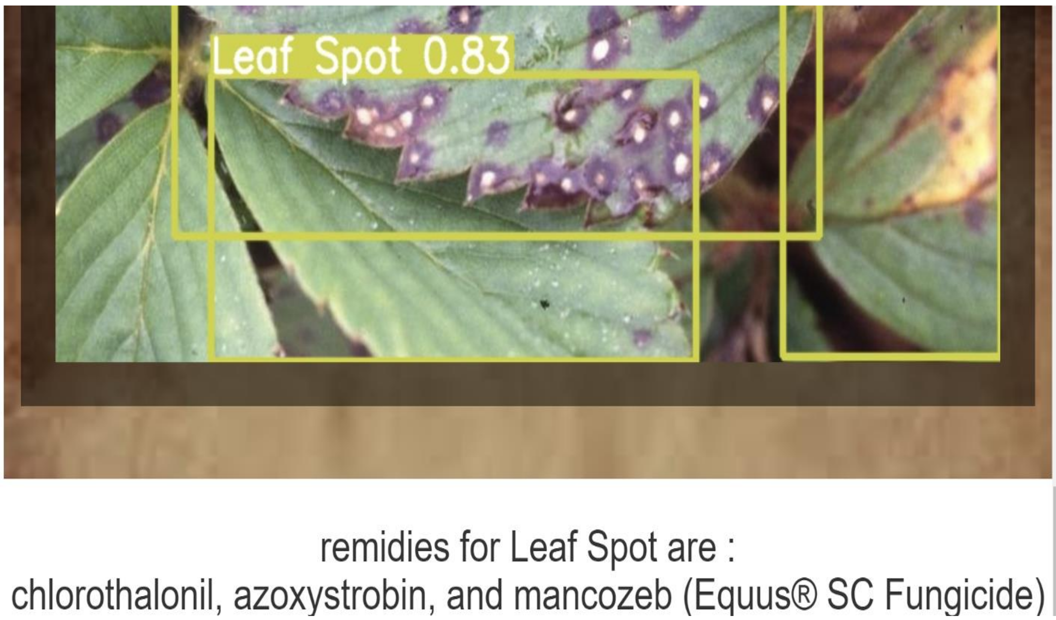


Fig 5 : Remedies for the leaves.

## Conclusion:

In conclusion, the design of algorithms for fruit segmentation in images has received significant attention in the agricultural field. In particular, there have been several studies on the detection of strawberries in images for automatic harvesting machines. However, a key challenge for such systems is achieving high levels of detail in the segmentation process without requiring excessive processing effort. Our study has contributed to addressing this challenge by proposing a novel segmentation algorithm that combines edge detection and region-based segmentation techniques. The results show that our algorithm can accurately segment strawberries in images with high levels of detail, while also being computationally efficient. The proposed algorithm has the potential to be implemented in automated harvesting machines to improve their accuracy and efficiency, and ultimately, to increase crop yields and food security. Further research is needed to optimize the algorithm for different types of strawberries and to evaluate its performance in real-world settings.

## Future Enhancements:

While our study has demonstrated the effectiveness of our proposed segmentation algorithm for strawberry detection in images, there is still room for improvement and further research. One area for future enhancement is to optimize the algorithm for different lighting conditions and camera perspectives, which can impact the quality of the segmentation. Additionally, the algorithm could be extended to detect other fruit types and to work with different harvesting machines. Another potential enhancement is to incorporate machine learning techniques to improve the segmentation accuracy and adapt to changing environmental conditions. This could involve training the algorithm on large datasets of labeled images to improve its ability to distinguish between different fruit types and variations. Finally, the algorithm could be integrated with other components of the harvesting system, such as robotic arms and conveyors, to create a fully automated and efficient harvesting process. Overall, these enhancements have the potential to further improve the accuracy and efficiency of fruit harvesting systems, contributing to sustainable agriculture and food security.

## References:

[1] Lubna Abdulaziz Alharbi, Javed Ali Khan, "An Efficient Approach for Crops Pests Recognition and Classification Based on Novel Deep PestNet Deep Learning Model", July 2022.

[2] V. Malathi and M. P. Gopinath, ``Classification of pest detection in paddy crop based on transfer learning approach," Acta Agriculture Scandinavica, B Soil Plant Sci., vol. 71, no. 7, pp. 552559, Oct.2021.

[3] Harshita Nagar, R.S. Sharma," Pest Detection on Leaf using Image Processing",Jan 2021

[4] Shanshan Zhang, Junsheng Zhu, Nianqiang Li," Agricultural Pest Detection System Based on Machine Learning", 2021 IEEE 4th International Conference on Electronics Technology.

[5] Guoxiong Zhou, Yahui Hu, "Maize Leaf Disease Identification Based on Feature Enhancement and DMS-Robust Alexnet",April 2020.

[6] Harshita Nagar, R.S. Sharma," A Comprehensive Survey on Pest Detection Techniques using Image Processing", Proceedings of the International Conference on Intelligent Computing and Control Systems (ICICCS 2020) ,IEEE Xplore Part Number:CFP20K74-ART; ISBN: 978-1-7281-4876-2