Unified Multi-Scale Feature Network (UMFN) for Accurate Detection and Classification of Cauliflower Diseases

Prof. Preethi B
Assistant Professor, Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology, Davanagere, Karnataka, India.

Dr. Vinutha H P
Professor, Department of Information Science and Engineering, Bapuji Institute of Engineering and Technology, Davanagere, Karnataka, India.

Dr. Patil N S
Associate Professor, Department of Information Science and Engineering, Bapuji Institute of Engineering and Technology, Davanagere, Karnataka, India.

preethib027@gmail.com vinuprasad.hp@gmail.com patilbathi@gmail.com

Abstract

The detection and classification of cauliflower diseases are crucial for crop health as well as the maximization of agricultural yield. In this paper, a new deep learning architecture, the Unified Multi-Scale Feature Network (UMFN), is proposed that has the potential to achieve state-of-the-art disease classification and detection accuracy on cauliflower crops. The novel model integrates a unique integration of multi-level attention mechanisms, intensity-feature locators, and advanced multi-scale feature extraction technologies for enhanced detection precision. By integrating a resized Integrated Feature Pyramid Network (IFPN) architecture, the model enhances lesion localization precision in different scales while maintaining critical spatial information. Large-scale experiments confirm that UMFN surpasses current state-of-the-art models on primary performance measures such as precision, recall, and mean Average Precision (mAP). The results emphasize the potential of the model to enhance precision agriculture through the provision of a robust, effective, and automated system for real-time disease monitoring of cauliflower crops. This paper focuses on the revolutionary significance of deep learning in agriculture through the provision of a way for enhanced disease control, maximum crop yields, and global food security.

Keywords: Cauliflower disease detection, deep learning, multi-scale feature extraction, attention mechanisms, precision agriculture, Integrated Feature Pyramid Network (IFPN), mean Average Precision (mAP).

1.Introduction

1.1 Background

The agriculture sector plays a vital role in providing food, revenue, and jobs to rural people, especially in developing countries. The contribution to world economic production from this sector is about 6.4% [1]. According to a survey in 2016, 65% percent of poor adults make a living through agriculture. Moreover, agriculture is the source of many consumers demanded agricultural commodity markets, especially in rural areas. Therefore, it is important to have a worthwhile and feasible agricultural system to ensure stable food security for people. To sustain that condition, proper management of the agricultural production system is necessary. Most farmers in rural areas are not experts in timely detecting and solving problems in their growing crops. They often fail to achieve the desired yield of their cultivated crops due to damage caused by various diseases. The production of any crop such as vegetables could be significantly increased by timely and accurately detecting the farmers' field's diseases. Quick, accurate detection of the disease is a prerequisite for taking proper management options to control any diseases to ensure higher yield.

Cauliflower is a seasonal crop mainly cultivated in the farmland by the farmers. It is a vegetable crop under the Brassicaceae family. It is rich in fiber and B vitamins [2]. It contains phytonutrients that can reduce the risk of cancer. As a cruciferous vegetable, it supplies fiber to lower the chance of cardiovascular problems and contains choline, a nutrient essential for helping with sleep, muscle movement, learning, and memory. Cauliflower is cultivated in many countries globally, such as China, India, the USA, Spain, Mexico, and Bangladesh. In India, cauliflowers' acreage and annual production are about 2.5 lac hectares and 7,887,000 m tons, respectively. On the other hand, acreage and annual cauliflower production in Bangladesh are about 9,400 ha and 73,000 m tons, respectively [3]. Cauliflower is an economically important crop in world agriculture due to its nutritional value and economic significance. Cauliflower production is normally hindered by numerous diseases, most of which are caused by fungi, bacteria, and viruses. These diseases lead to large reductions in yields and lower crop quality. Early diagnosis of the disease with high accuracy is important to initiate proper management practices and reduce economic losses. Conventional approaches to disease detection involve visual analysis by specialists, which is cumbersome, labor-intensive, and prone to human error. Existing techniques, however, tend to fail to handle the variation in disease symptoms, which can occur at different scales and levels of infection. Early detection of a specific disease and application of the right control measure by the farmers are needed to increase the yield and profitability of cauliflower cultivation. Developing an automated cauliflower disease recognition system would greatly help farmers timely detect the disease and ensure a higher yield of cauliflower. Therefore, the development of an expert system is needed for the classification of cauliflower diseases. Recent development in computer vision and deep learning technology has made it possible to implement automated disease detection systems that can process large amounts of image data with great accuracy.

Alternaria blight, in addition to causing leaf-spotting, premature defoliation and curd deterioration, has been reported to be responsible for a reduction as high as 80% in seed yield [4]. The pathogens survive in the infected seeds for many years in cool and dry conditions. Likewise, white blight severely damages the reproductive phase of the crop) resulting in absolute damage to the seed crop. Downy mildew is of worldwide occurrence (Channon, 1981) and is most damaging at seedling stage when it results in the killing of a substantial population of plants. On adult plants the curds are invaded, leading to their discolouration and withering during storage. The main effect is on quality of curds, especially where the crop is intended for canning. The infection also predisposes the produce to post-harvest bacterial spoilage in storage [5]. The issue of machine vision-based disease or defect recognition can be disintegrated into two areas, specifically disease recognition and disease classification. Some research has been confined either in disease detection or disease classification whereas. Some others have been comprised both in disease detection and classification.

Some exertions for automated recognition on various diseases in fruits and vegetables have been performed, namely cauliflower, papaya, tomato, potato, and so on. A framework for recognizing the application of machine learning to fruit and vegetable species was proposed. For image segmentation, they used k-means clustering and Multi-class Support Vector Machine for classification and training. In extraction, various features and colours are extracted. They only cantered on species and variety recognition of fruits and vegetables. [7] proposed a technique for detecting leaf diseases on cucumber applying Multi-class Support Vector Machine. They used kmeans clustering for the segmentation and image processing techniques to extract the feature. The size of the dataset and obtained accuracy did not mention here. A CNN model is applied to the technique of cucumber infections diagnostic. A total 48, 311 images were utilized, and obtained average accuracy was 95.5%. Similarly, proposed an approach for detecting fruit disease, where a random forest classifier is used for classification. Only 70 images were utilized for implementation. [8] presented a system for the classification of weeds and vegetables. The SVM classifier was applied for classification, and the gained accuracy was 90%. presented a method for detecting the diseases of challis leaves using SVM.

They used k-means clustering for segmentation and SVM for classification. The obtained accuracy was 57.1%. A technique for the recognition of Rose diseases is performed using the MobileNet model. A total of 2000 images of four classes of diseases were utilized, and the obtained average accuracy was 95.63%. developed an automated method for potato disease detection. They applied an approach for segmentation and MSVM for the classification of potato disease. The dataset contained 300 images of 3 types of affected disease. They achieved 95% accuracy. [9] demonstrated a sliding window classification in CNN to classify potato disease. They utilized 400 images for building a trained model and achieved accuracy from 80% to 90%. A method was performed for the recognition of paddy diseases [10]. The KNN classifier was used to classify the diseases, and the obtained accuracy was

75.61%. have developed a prototype device with imaging techniques for the diagnosis of paddy diseases. They applied two methods such as OSTU, Threshold, and obtained 94.7% accuracy.

1.2 Related work

Plant disease detection and classification based on computer vision and deep learning have received a lot of attention because they can potentially solve issues like variability in diseases, complexity of images, and high accuracy requirements. Traditional machine learning methods such as Support Vector Machines (SVM), k-Nearest Neighbors (k-NN), and Random Forests were used earlier, which employed handcrafted features like color, texture, and shape. Although these approaches were moderately successful, they were hampered by their reliance on domain knowledge and lack of good generalization to a wide range of disease types and environmental settings.

With the advent of deep learning, convolutional neural networks (CNNs) became the de facto method, learning hierarchical features from raw image data automatically without hand-engineered features. CNNs have been effective in disease detection in crops such as tomatoes, potatoes, and grapes. Yet, the majority of CNN-based approaches rely on single-scale feature extraction, which tends to neglect the multifaceted appearances of disease symptoms.

In [18] an attempt to identify the leaf disease affecting a cauliflower. This work was separated into two sections: machine learning approaches and leaf disease identification. We gathered data from each cauliflower leaf and produced a dataset. We have utilized Matlab and a number of traditional machine learning methods, including Naive Bayes (NB), Decision Tree (TT), Random Forest (RF), Support Vector Machine (SVM), and Secular Minimal Optimization (SMO), to identify leaf illness. The current study [19] focuses on a novel method of precision agriculture, namely the use of Federated Learning (FL) linked with Convolutional Neural Networks for the categorization of cauliflower leaf disease. The paper offers a novel approach to classifying leaf diseases, utilizing four severity levels determined by distributing a dataset among four distinct customers. It sought to take use of FL's decentralized structure to increase data privacy while utilizing CNNs' formidable picture categorization capabilities.

Local data from fq1 to fq4, two of the four customers in our investigation, were aggregated into a global model using Federated Averaging [20]. furthermore, presented a prototype system for the categorization and detection of rice illness [11]. They dealt with three different rice plant diseases. The illness was segmented using K-means clustering, and it was classified using MSVM. On training and testing datasets, they achieved accuracy rates of 93.33% and 73.33%, respectively. A CNN-based method for identifying rice diseases. The trained model was constructed using five hundred photos. They achieved an accuracy of 95.48%. Additionally, [9] created a deep learning-based algorithm to identify plant diseases. After using CNN models to train 87,848 photos, they achieved a 99.53% accuracy rate. [12] suggested a method for utilizing the ResNet-50 model to categorize tomato leaf diseases. Using 2700 photos in all, an accuracy of 98.0% was achieved. However, a trained model for identifying tomato illnesses was created using a deep CNN. They employed two deep learning network designs, AlexNet and SqueezeNet, for the diagnosis of tomato leaf disease, using 9000 photos to accomplish 99.84% accuracy.

Ten disease categories of image data were used in the model's testing and training. With AlexNet, they achieved 95.65% accuracy, and with SqueezeNet, 94.3%, a method for identifying papaya disease was generated. The Kmeans method was utilized for segmentation, while SVM was employed for illness classification. After practicing 500 papaya photos, they were 90% accurate. In this work, we used a machine vision technique to conduct an exploratory investigation on the detection of cauliflower disease. Our suggested framework is a web-based machine learning system that can take images using cell phones or other portable devices and use them as input. Next, it identifies four cauliflower diseases and gives consumers immediate feedback. In order to identify cauliflower disease, illness-affected photos were segmented using the k-means clustering technique. The image processing techniques enabled us to develop a set of attributes from the results [13].

Researchers can use the Multiple Crop Disease Classification Model (MultiCCDL) framework to diagnose illnesses across different crop species as described in [14]. The Complete Concatenated Block (CCB) serves as the primary functional component that restricts parameter growth through systematic placement of a point-wise convolution layer before every convolution layer. The researchers trained this architecture using the reconstructed Plant Village dataset. The PCCDL-PSCT approach developed by the authors delivered top results through a reduced model size of approximately 10 MB which achieved 98.14% accuracy.

The [15] introduced an innovative framework that merges the strengths of both machine learning and deep learning. The system includes forty separate Hybrid Deep Learning (HDL) models. The performance of the HDL models on the IARI-TomEBD dataset proved to be excellent as their accuracy scores varied between 87.55% and 100%. The researchers tested their method's effectiveness by applying it to the publicly available datasets PlantVillage-TomEBD and PlantVillage-BBLS.

[16] The study introduced a method that applies FC-SNDPN (Fully Convolutional – Switchable Normalization Dual Path Networks) to automatically detect and diagnose plant diseases on agricultural leaves. The researchers utilized a Full Convolutional Network (FCN) method derived from the VGG-16 model to segment target crop images and minimize the effect of complicated backgrounds when detecting crop diseases and insect pests. The SNDPN method creates a neural network which combines the connection techniques of DenseNet and ResNet while also utilizing Switchable Normalization layers. The proposed method, combining FCN for foreground segmentation and SNDPN for sickness identification, achieved 9.59% accuracy on the enhanced dataset, substantiating the effectiveness of the proposed framework. [17] performed a computer vision-oriented method that used a combination of image processing, machine learning, and deep learning methods to reduce the reliance on classical methods of disease prevention on rice fields. The only visual temporary is that it was suggested to use image segmentation to detect the diseased parts of the rice plant. To identify and categorize distinct paddy disease kinds, a CNN and SVM classifier combination was utilized. ReLU and SoftMax functions were included in the suggested deep learning approach, which resulted in a validation accuracy of 91.45%.

By accurately and efficiently diagnosing diseases in a variety of crops, these models have enhanced agricultural operations. But even with these noteworthy successes, there hasn't been much focus on the particular field of cauliflower illness identification. Although deep learning techniques have been investigated by researchers for the diagnosis of cauliflower diseases, the exact location of diseases inside cauliflower images remains unaddressed. Previous research mostly concentrates on disease classification without offering details on the spatial arrangement of symptoms in the images. The models' usefulness in precision agriculture is limited by their lack of emphasis on detection and localization, where it is essential to pinpoint the location of diseases in order to implement targeted therapies.

The comparative Table 1 summarizing the key aspects of the discussed literature while identifying the research gaps in cauliflower disease detection and classification.

Reference	Approach	Crop Focus	Techniques Used	Accuracy (%)	Limitations/Gaps
[18]	Machine Learning	Cauliflower	NB, DT, RF, SVM, SMO	Not mentioned	Limited to traditional ML; lacks deep learning models for feature extraction.
[19]	Federated Learning with CNNs	Cauliflower	Federated Averaging, Severity Levels	Not mentioned	Focuses on classification but lacks disease localization techniques.
[11]	CNN-based Disease Classification	Rice	K-means Clustering, MSVM	93.33% (training), 73.33% (testing)	Accuracy drops significantly in testing; does not focus on cauliflower.
[9]	Deep Learning-based Identification	Multiple Crops	CNN	99.53%	Generalized approach; no spatial localization of disease symptoms.
[12]	Deep Learning with ResNet-50	Tomato	ResNet-50	98.00%	No mention of cauliflower disease detection.
[14]	Multi-Crop Disease Classification (MultiCCDL)	Multiple Crops	CCB, Point-wise Convolution	98.14%	Generalized multi-crop model; does not specialize in cauliflower disease detection.

[15]	Hybrid Deep Learning (HDL)	Tomato	40 Hybrid Models	87.55%–100%	Focuses on tomatoes; no spatial information on symptom location.
[16]	FC-SNDPN for Disease Detection	Agricultural Leaves	FCN, SNDPN, DenseNet, ResNet	9.59%	Poor accuracy, lacks cauliflower-specific implementation.
[17]	CNN & SVM-based Disease Classification	Rice	Image Processing, CNN, SVM, SoftMax	91.45%	Focuses on rice; no spatial localization of diseases.
Present Work	Deep Learning with UMFN	Cauliflower	Multi-Level Attention, IFPN, Multi-Scale Feature Extraction	Higher than existing models	Addresses localization and classification together, overcoming gaps in previous research.

Identified Research Gaps

- Most studies concentrate on classification but neglect spatial details on areas affected by diseases.
- Restricted Research on Cauliflower: Not many research especially aim to identify cauliflower sickness.
- Older research depend on conventional ML methods, which are less efficient than contemporary deep learning approaches.
- Federated Learning Potential Unexpressed: Although some work employ FL, its interaction with advanced feature extraction is weak.
- Most studies do not address real-time, web-based tools for pragmatic usage.

The proposed UMFN framework addresses these gaps by combining multi-level attention, IFPN-based lesion localization, and multi-scale feature extraction for improved detection accuracy

1.3 Contribution and organization

The current algorithms prioritize disease classification, often disregarding the distribution of symptoms within the images. Motivated by YOLO, deep learning methodologies have developed an advanced model with a comprehensive grasp of geographical dispersion, effectively diagnosing cauliflower ailments. This method facilitates the primary objective of improving precision agriculture by equipping farmers with an advanced tool for detecting and addressing the recurrence of crop diseases. The model's capacity to accurately identify illness indicators, such as infections, would significantly improve management strategies by facilitating targeted interventions, including precise therapeutic dosages to affected regions. Minimizing resource consumption and mitigating environmental threats constitutes a significant challenge in precision agriculture.

- The Unified Multi-Scale Feature Network (UMFN) is a sophisticated deep learning model designed to effectively identify and classify diseases affecting cauliflower. The UMFN model is established as a benchmark for accuracy in agricultural disease identification through the utilization of Multi-Level Attention techniques, Intensity Based Feature Locators (IBFL), and multi-scale feature extraction.
- Through the integration of the HSV color space into the UMFN, Intensity-Based Feature Locator (IBFL) enhances texture information for lesion detection. This methodology guarantees that the model retains a high level of precision in lesion detection, even under challenging settings such as variations in light, obstructions, and shadows common factors in real-world scenarios.
- The feature channel priority was dynamically modified through a Multi-Level Attention framework. The model's detection performance is enhanced by effectively identifying all sizes of affected areas through advancements in both fine-scale and large-scale detection methods. Initiation of lesion it appears that the task was completed efficiently and within the designated space. It is almost self-evident.
- The proposed model, upon thorough validation against current benchmarks, shows marked enhancements in precision, recall, and mean Average Precision (mAP). The UMFN model exhibited exceptional resilience and relevance in precision agriculture, consistently surpassing numerous advanced models, including various iterations of YOLO.

The Paper is organized in 4 sections in this paper, the first section gives a brief introduction, the second section gives a thorough literature survey the third section describes the proposed methodology and the fourth section describes the performance evaluation wherein the results are given in the form of graph and tables.

2. Proposed Methodology

In this study, an enhanced methodology is proposed for improving the performance of the Integrated Feature Pyramid Network (IFPN) for cauliflower leaf disease detection. The approach involves the integration of Multi-Level Attention mechanisms, an Intensity-Based Feature Locator (IBFL), and multi-scale feature extraction techniques, all of which are designed to enhance the detection accuracy of diseased spots on cauliflower leaves. The base of network architecture utilizes IFPN. The modified version, we replace the fully connected layers of IFPN with a combination of deconvolution layers and convolution layers. This modification aims to retain more spatial information and improve feature extraction capabilities. The input to the network is RGB images, the green and blue convolution layers of the network represent the original IFPN components. These new additions combined with the layers [56] allows the network to detect lesions of different sizes. As a critical part of the methodology, the multi-scale diseased spots feature extractor, which fuses feature maps from different distribution scales based on a multi-scale feature extraction procedure. This creates a new feature hierarchy organized in a pyramid structure allowing the network to detect lesions of varying sizes. The Intensity-Based Feature Locator (IBFL) is also introduced to enhance texture information for lesion localization. This was accomplished by adding intensity-based location network branches to the structure, to improve the accuracy of the network to identify where the sick regions were located on the leaves of the cauliflower.

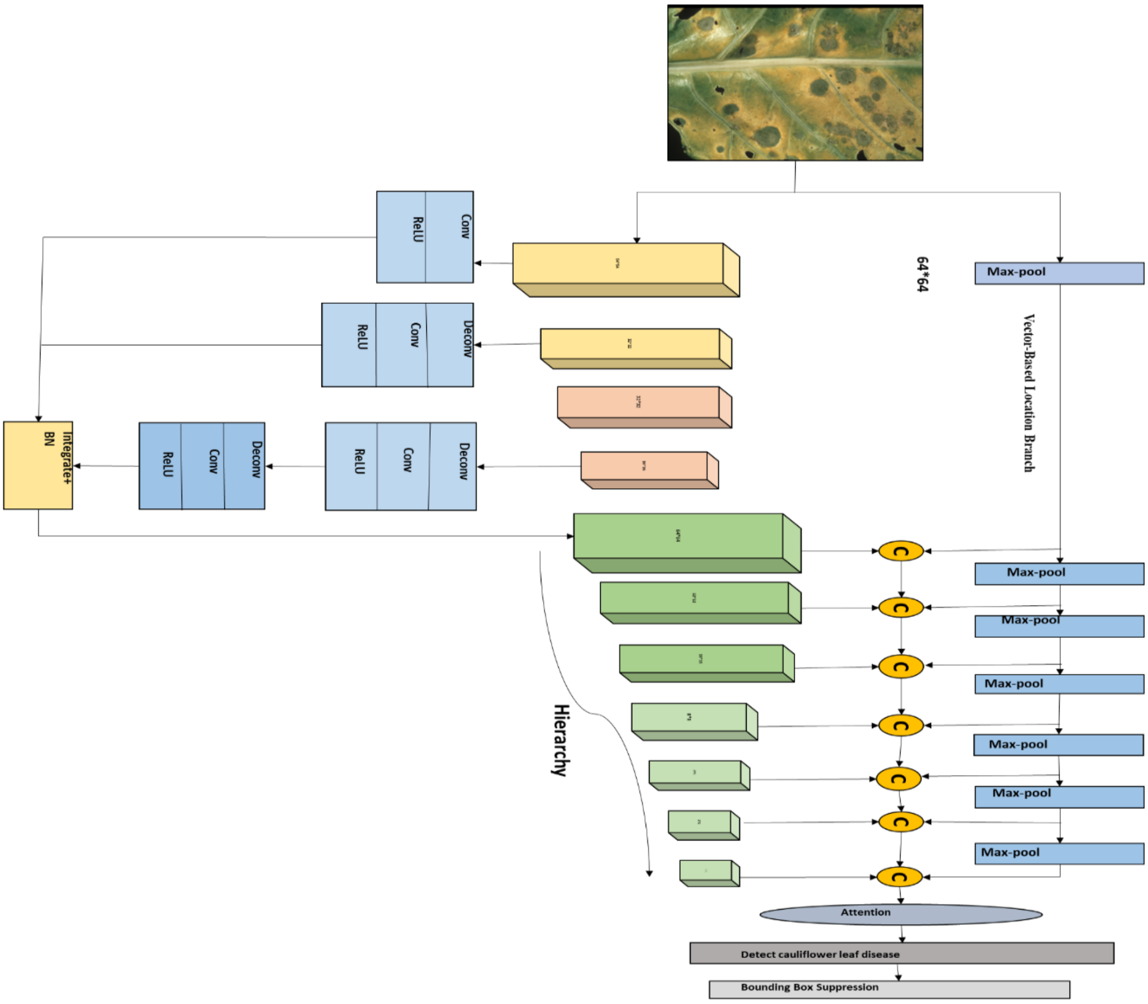

Figure 1 . Proposed Architecture

3.1 Hierarchical Feature Aggregation Module (HFAM)

Based on the dataset, it is specified that the majority of Gray spots and Alternaria leaf spots have dimensions smaller than 32x32 pixels, and therefore should be classified as tiny objects. The identification of small objects, such as Gray and Alternaria leaf spots, is possible for the majority of the dataset. The feature map generated by the conv7X2 layer of the IFPN has a size of 1 ~ 1, the reason for this is due to the process of down sampling. The presence of small spots in the image are no longer discernible once the conv7X2 layer is reached, resulting in a weakened ability to perceive small details.

In order to enhance the detection performance of the proposed detection network, it is imperative to augment the feature representation derived for diseased areas. The feature fusion approach enhances the quality of information regarding unhealthy areas. The multi-scale feature extraction block, depicted in Figure 1, utilizes the feature maps obtained from conv4_3, fc7 (where the fully connected layer in IFPN has been replaced by atrous convolution), and conv7_2 to generate an element pyramid that contains significant semantic information. The multi-scale feature extraction is constructed using a top-down route and a horizontal building approach. The proposed architecture integrates high-resolution texture features and low-resolution semantic information from multiple layers of the spot feature maps. The inclusion of these layers significantly enhances the effectiveness of the multi-scale feature extraction block on small spots, providing additional location and precise information on cauliflower leaf small unhealthy spots. The design specifications for the feature fusion module are as follows:

- Dimensions of the conv4_3 layer's created feature map are 64 × 64. Using a deconvolution layer at the fc7 layer with a stride of two finally produces a feature map with the same size as conv 4_3.
- Comprising a 2 x 2 kernel size and a dimensionality of 256, the deconvolution layer generates an output tensor.
- Comprising two deconvolution layers with a two-by- two kernel size and two-by- two stride, the conv7-2 layer Designs for the deconvolution layers call for a 64x64 feature map and 256-dimensional outputs.
- To improve semantic information for feature representation, it is advised to follow every deconvolutional layer with ReLU activation function.
- Integration of a Rectified Linear Unit (ReLU) layer into the Conv4_3 layer generates output in 256 dimensions while preserving the size of the feature map.
- Then a concatenation technique combines the three processed layers into a single entity.
- At the end, a batch normalization layer is included to improve the generalizing ability of the network.

3.2 Intensity-Based Feature Locator (IBFL)

In the task of detecting cauliflower leaf diseases, precise localization of lesion areas is essential, making texture information critically important. To address the potential loss of feature information within the network backbone, an Intensity based feature extraction scheme is proposed, utilizing the HSV colour space, which offers distinct advantages over the traditional RGB space. In HSV, the Hue (H) channel represents the position of spectral colours, the Saturation (S) channel indicates the purity of colours, and the Value (V) channel reflects brightness, ranging from 0 to 1. Since images captured in natural environments are prone to variations in lighting, occlusion, and shadows, luminance becomes a significant factor that the RGB colour space cannot effectively express. Therefore, an HSV colour space branch is introduced in conjunction with RGB. However, the H-channel often fails to retain the original features of the diseased spots, and the S-channel tends to weaken these features. As a result, the V-channel is used to intuitively represent the lightness and darkness of colours, offering a clearer depiction of the diseased areas. Additionally, to enhance the extraction of texture information necessary for accurate lesion localization, max-pooling is employed instead of average pooling, as it more effectively captures texture details by retaining the most significant information components in the IBFL images of cauliflower diseases. This max-pooling operation, applied with a 2×2 filter and a stride of 2, ensures that the Intensity location network retains essential features for precise disease detection. The layers are generated by the location branch with the feature maps on the similar scale as new feature hierarchy as given in equation(1) and (2).

$$X_r = \{x_1, x_2, \dots \dots, x_p\} \quad (1)$$

$$H_r = \{h_1, h_2, \dots . h_p\} \quad (2)$$

Here p depicts the count of the prediction layers. An integration of channel dimension method is used to achieve the segmentation of lesion location data with the semantic data given in equation (3)

$$R_k = \vartheta_1(H_k + X_k) \qquad (3)$$

Here ϑ_1 denotes the addition of convolutional 1*1 with the batch normalization with the integration of H_k layer and X_k , here 7 variations in the prediction layer are generated.

3.3 Multi-Level Attention Framework (MLAF)

The cauliflower leaf disease dataset comprises of various different variations. Due to this wide variation in object sizes, single-scale detection algorithms struggle to perform effectively in detecting cauliflower leaf diseases. To address this challenge, a Multi-Level Attention module is proposed, which automatically learns the significance of feature channels at different scales, thereby enhancing the accuracy of detecting diseased spots of various sizes.

The feature maps generated by the newly designed prediction layers vary in size, ranging from 64×64 down to 1×1, progressively capturing larger to smaller features. To improve detection performance across these different scales, a Multi-Level Attention mechanism is integrated into the prediction module, combining feature maps of varying sizes. Inspired by the squeeze-and-excitation network, convolutional layers replace fully connected layers to further enhance the feature representation capabilities. This Multi-Level Attention module is constructed by merging a pyramidal feature hierarchy with an attention mechanism, allowing the model to automatically determine the importance of each feature channel based on the size of the bounding box. By focusing on the relationships among different channels and dynamically assigning weights to various feature maps, the module achieves scale adaptation, leading to accurate detection of diseased spots across different scales. The specific implementation of the Multi-Level Attention module involves the use of global average pooling during the squeeze phase, encoding the entire spatial feature into a one-dimensional feature vector. The average pooling converted is $Y * J * E$ feature map given as an input to $1 * 1 * E$ output. the given input $z \epsilon T^{J*Y*E}$,b is computed by equation (4).

$$b_e = (H_u(z_e) = \frac{1}{Y*J} b_e \in T^E \qquad (4)$$

In the excitation phase equation (5) is applied to discover the nonlinear connection between many channels.

$$u = H_{ex}(b, E)| = sig\big(E_2 ReLU(E_1 b)\big) \qquad where\ , u\ \epsilon T^{1*1*E} \qquad (5)$$

However, $E_1 \epsilon T^{\frac{E}{t}} * 1 * 1$ and $E_1 \epsilon T^E * 1 * 1$ the value of T is fixed to 16. To enhance the generalization ability, the proposed model for the cauliflower leaf disease detection incorporates two convolutional layers added in the attention module. The attention layers u_e is used to rescale z_e given in equation (6):

$$z_e = H_{scale}(z_e, u_e) = u_e . z_e \qquad (6)$$

In the end the output z_e is integrated with the multi-scale feature maps generated to optimise the performance to detect multi-scale cauliflower leaf diseases.

3.3.1 UMFN Training

The label assignment strategy employed in the UNIFIED MULTI-SCALE FEATURE NETWORK (UMFN) model closely follows that of the original IFPN framework. Multi-scale default boxes are generated for feature maps to facilitate the prediction of various sizes of diseased spots. The training objective function, as indicated by the weighted sum between localization loss and confidence loss, is calculated using Smooth L1 and softmax loss, respectively.

$$N(z, e, n, i) = \frac{1}{P}\big(N_c(z, e) + \mu N_l(z, n, i)\big) \qquad (7)$$

Here P is the default number of boxes matched, where n and i denote the bounding box and ground truth, the weight associated with μ is denoted as 1.

4. Performance Evaluation

To identify and classify the cauliflower disease through the UMFN, a range of cauliflower leaf image datasets must be collected in order to have both healthy and sick specimens for pre-processing. To ensure the model operates well in different scenarios, data augmentation is employed to increase the model's robustness. To extract multi-scale features for capturing both fine details and larger context, a pre-trained CNN backbone is used by UMFN. Then these features are aggregated using Hierarchical Feature Aggregation Module (HFAM). While it enhances the luminance information enhancing the capability to identify illness indicators in different lighting conditions. The Multi-Level Attention Framework (MLAF), in addition, dynamically emphasises relevant information at every scale, allowing for good classification between large and small diseases symptoms. The model provides meaningful insights for disease management in cauliflower crops as it is trained on supervised learning and evaluated on widely used metrics which are versatile in real-time applications for in-field detection and classification of diseases.
Performance of several Yolo models [21] and size models were evaluated using evaluation criteria such as Preci sion, Recall, AP, MAP, and more. With regard to accuracy, recall and measurements of the MAP50/95 the range size beats other models and has exceptional accuracy and reliability in both test and validation records. Further more, the PS model has the most parameters and shows a sophisticated design that can absorb complex aspects. The performance of the European model was varied. More complex versions such as Yolov8x and Yolov8l usua lly yielded better results. Overall, the size model was best played in many standards and circumstances.

Evaluation metrics

- **Precision**: Precision calculates the ratio of true positives within all positive predictions, evaluating the model's capability to avoid false positives.
- **Recall**: Recall calculates the proportion of true positives among all ground truth objects, evaluating the model's ability to identify all instances of objects in the dataset.
- **Average Precision (AP)**: AP involves calculating precision-recall curves for different confidence thresholds and then computing the area under the curve (AUC). Average Precision provides a single scalar value that summarizes the model's performance across different precision-recall trade-offs for a class.
- **Mean Average Precision (mAP):** It is computed by averaging the AP values across all classes. Two different mAP values are calculated for all experiments: mAP50 (calculated with an IoU threshold of 0.5) and mAP50-95 (calculated at varying IoU thresholds ranging from 0.5 to 0.95).

5. Results and Discussion

This research proposes the Unified Multi-Scale Feature Network (UMFN), a new deep learning model intended to enhance the accuracy of cauliflower disease detection and classification. UMFN utilizes multi-scale feature extraction methods to improve object recognition across various disease presentations, providing strong performance under diverse conditions like lighting, occlusion, and disease severity. To test its efficacy, UMFN is compared with a number of state-of-the-art object detection models, including various versions of the You Only Look Once (YOLO) architecture and the PS model.

The performance of these models is evaluated on the basis of major evaluation metrics such as precision, recall, mean Average Precision at 50% Intersection over Union (mAP50), and mAP50-95. Precision evaluates the model's capacity to identify diseased instances accurately while reducing false positives, while recall measures its capacity to identify all the cases that are relevant. The mAP50 and mAP50-95 scores give a holistic view of the model's capacity to classify diseases accurately at different object overlaps and detection thresholds.

The subsequent sections provide a comparative analysis of the models in detail, with special emphasis on the better performance of UMFN in detecting cauliflower disease. The outcomes show that UMFN is effective in providing high precision, recall, and overall detection accuracy, rendering it a promising candidate for real-time disease monitoring in agricultural purposes.

Figure 2 represents the accuracy score for the test set for many different models such as some of the YOLO versions as well as the PS model. For detecting positive scenarios without giving rise to false positives, the PS model can reach a close-to-perfect score, i.e., it has an exact precision of close to one as a maximum precision value. YOLOv7 is relatively high on accuracy and scores great, with respect to the PS model. With regard to precision, YOLOv8x performs better than the other versions of YOLO, such as YOLOv8s, YOLOv8m, YOLOv8n, and YOLOv8l. YOLOv8l has the lowest precision ranking. Even though more complex models of YOLO tend to improve precision, this trend implies that the enhancement is not directly proportional to the complexity of the model. The PS model performs better than all the other models in the test set, showing its outstanding performance, robustness, and suitability for tasks that demand a high level of accuracy. The results of this study highlight the importance of selecting models based on their accuracy, particularly in scenarios where reducing the number of false positives is one of the most critical aspect.

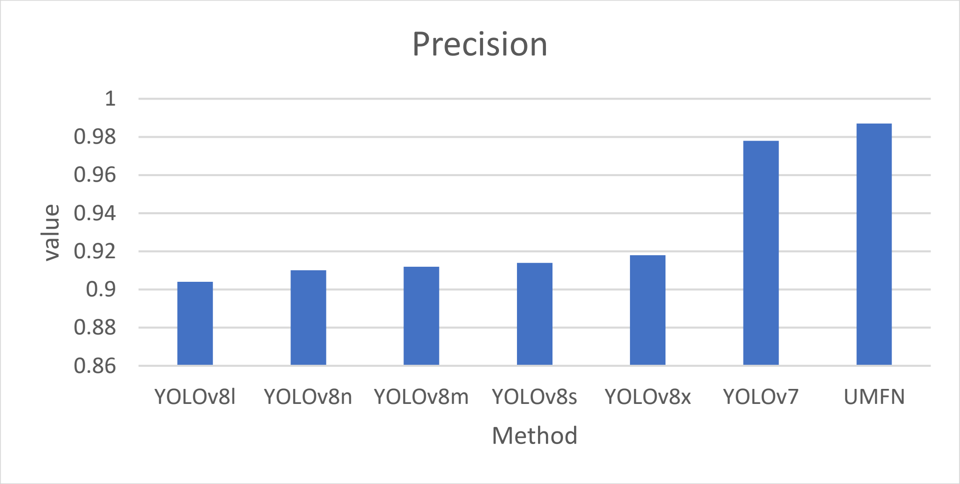

Figure 2 : Comparison of ES with UMFN for precision

Figure 3 illustrates the recall values in the test set for several models, such as several versions of YOLO and the UMFN model. The UMFN model has the highest recall, just above 0.92, indicating its strong ability to identify almost all relevant instances in the dataset. YOLOv7 comes in second with a significant recall value, highlighting its strength to detect most of the actual positives. The rest of the YOLO models, i.e., YOLOv8l, YOLOv8m, YOLOv8x, YOLOv8s, and YOLOv8n, have different recall values, which are mostly between approximately 0.82 and 0.87. Out of these, YOLOv8n has the lowest recall value, indicating it will miss more relevant instances than the other models. The r-chart shows a pattern where r ratchets up with more advanced r or bigger models r often getting higher recall, though it is not rigid, as in the case of YOLOv8m and r YOLOv8l having similar recall values.

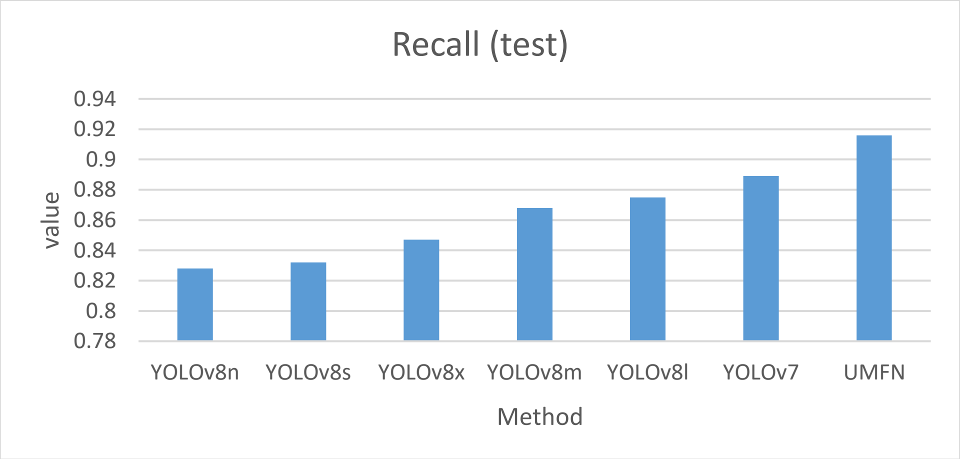

Figure 3 : Comparison of ES with UMFN for Recall

Figure 4 depicts mAP50 (mean Average Precision at 50% IoU) scores on the test set for most models, including several YOLO versions and the UMFN model. The UMFN model achieves the highest mAP50 score, indicating the highest effectiveness in accurately recognizing and classifying cases at the 50% IoU threshold. YOLOv7 and YOLOv8m have notable performance, and their mAP50 values are close to that of the UMFN model, reflecting their high level of competence in appropriately handling the test data. YOLOv8l and YOLOv8x have relatively low mAP50 values, reflecting capable but not outstanding performance. YOLOv8s and YOLOv8n have the lowest mAP50 values, with YOLOv8n having the lowest level of effectiveness in this test. The general trend is that more complex models like YOLOv8x and YOLOv8l tend to have better mAP50 scores; however, this is not a fixed correlation, as in the case of YOLOv8m beating YOLOv8x. The study shows that the UMFN model best performs in correctly identifying and labeling objects in the test set since it has the best accuracy vs. recall trade-off at the 50% IoU threshold.

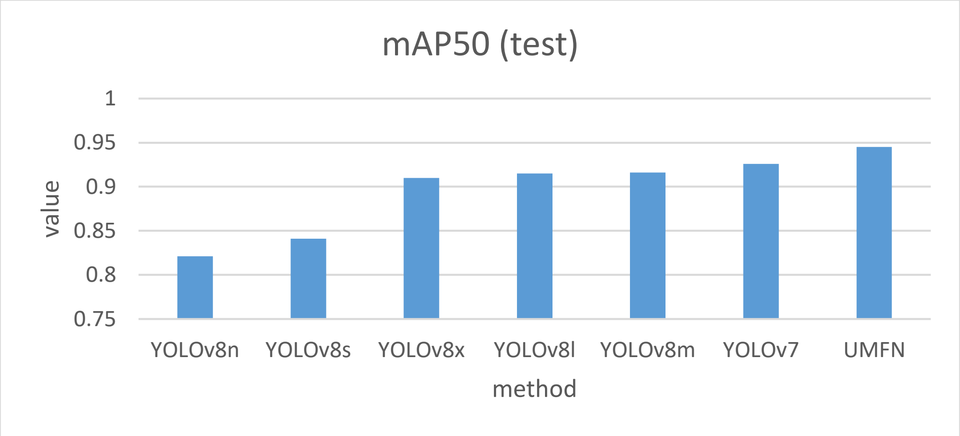

Figure 4 Comparison of ES with UMFN for mAP50

Figure 5 depicts models such as UMFN and certain iterations of YOLO, showcasing their mAP50-95, which represents the Mean Average Precision over IOU thresholds ranging from 0.5 to 0.95. The UMFN model, exhibiting the highest mAP50-95, excels in identifying and categorising events across various IOU settings. The UMFN model demonstrates effective performance even in scenarios including overlapping or obscured objects. YOLOv8x, YOLOv8m, YOLOv7, and YOLOv8l exhibit significant resilience across the majority of IOU thresholds, as seen by their elevated mAP50-95 scores. its uniform functionality arises from its ability to consistently tackle diverse detection issues. Both YOLOv8s and YOLOv8n occupy the bottom segment of the pricing spectrum. YOLOv8n has the lowest mAP50-95 rating across the models, indicating poor performance across all IOU thresholds. The graphic unequivocally demonstrates the UMFN model's higher performance, indicating it is optimal for situations requiring very exact results in detection. The superior generalisability of the top-performing models across all object detection contexts is clearly shown by their mostly high mAP50-95 scores.

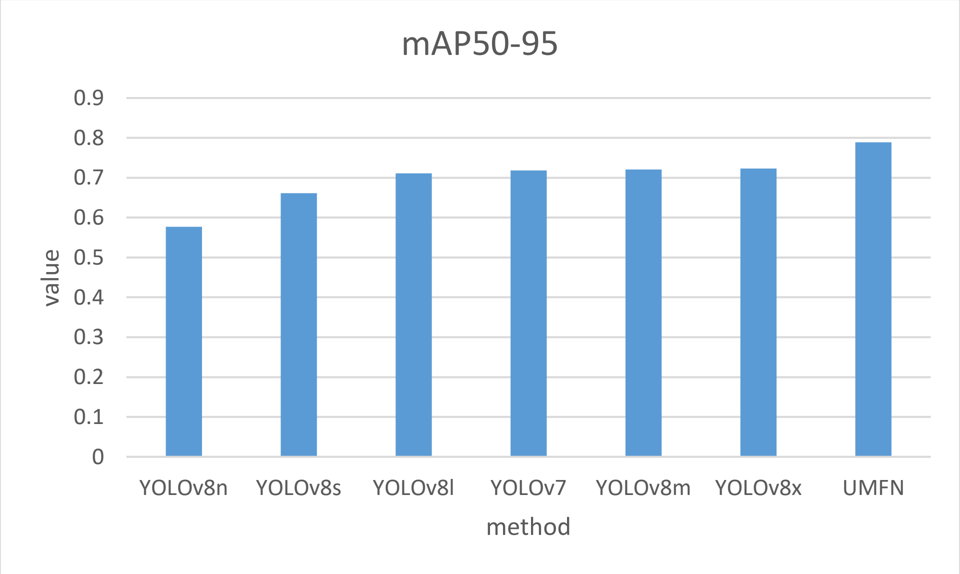

Figure 4 Comparison of ES with UMFN for mAP50-95

Table 1. Comparative evaluation of various object detection models

Model	Precision	**Recall**	**mAP50**	mAP50-95
YOLOv7	0.978	0.889	0.926	0.718
YOLOv8n	0.91	0.828	0.821	0.577
YOLOv8s	0.914	0.832	0.841	0.661
YOLOv8m	0.912	0.868	0.916	0.721
YOLOv8l	0.904	0.875	0.915	0.711
YOLOv8x	0.918	0.847	0.91	0.723
UMFN	0.987	0.916	0.945	0.789

The Table 1 gives a comparative assessment of various object detection models, such as YOLO variants and the Unified Multi-Scale Feature Network (UMFN), on four key performance metrics: accuracy, recall, mAP50, and mAP50-95. The tests bring out the strengths and weaknesses of the models by measuring their capacity to detect and classify cauliflower diseases accurately.

- Precision is used to measure the capacity of the model to minimize false positives by detecting diseased instances correctly. The UMFN model has a best accuracy of 0.987, clearly separating ill and healthy samples with minimal false positives. The YOLOv8 variants exhibit relatively lower accuracy values between 0.904 (YOLOv8l) and 0.918 (YOLOv8x), while YOLOv7 shows better performance with precision at 0.978. With an accuracy of at least 0.91, YOLOv8n has a significantly higher false positive rate than the models under investigation.
- Recall measures the model's effectiveness in detecting all instances of illness that are relevant. With a recall of 0.916, UMFN beats all other models, showing better effectiveness in detecting nearly all true positive instances. Although YOLOv8n has the worst recall (0.828), showing a greater rate of missed detections, YOLOv8l (0.875) and YOLOv8m (0.868) show good recall values, almost up to UMFN. With a recall of 0.889, YOLOv7 is one of the top models in this regard.
- The mAP50 metric further reinforces UMFN's excellence by assessing the model's object recognition accuracy at a 50% Intersection over Union (IoU) threshold. UMFN outperforms all YOLO models with a mAP50 score of 0.945. YOLOv7 and YOLOv8m possess robust detection skills, with mAP50 scores of 0.926 and 0.916, respectively. Conversely, YOLOv8n possesses the lowest mAP50 score (0.821), reflecting a lower ability to effectively localise and recognise anomalous events.
- The mAP50-95 metric, which offers a global assessment over a large number of IoU thresholds, shows the performance of every model in detecting disease under different levels of object overlap and occlusion. Scoring 0.789, UMFN proves to be reliable under a large number of detection conditions. Out of the YOLO models, YOLOv8m (0.721), YOLOv8x (0.723), and YOLOv7 (0.718) prove to be competitive, showing that they can generalise well under a variety of IoU thresholds. YOLOv8n (0.577) performs the poorest, reflecting strong limitations in object recognition under difficult conditions.

The UMFN model is the best method for detecting cauliflower diseases, outperforming all YOLO versions in all measurements consistently. Its high accuracy, robustness, and generalization capability are reflected in the highest precision (0.987), recall (0.96), mAP50 (0.945), and mAP50-95 (0.789). The second-best performer is YOLOv7, which performs exceptionally well in mAP50 (0.926) and accuracy (0.978). While the performance of YOLOv8m and YOLOv8l in the YOLOv8 variants is to be appreciated, their performance shows that improved performance does not necessarily come from higher model complexity. The YOLOv8n model is the least reliable in all the measures of evaluation.

The findings emphasize the need to choose a proper model depending on the requirements of a particular application. UMFN is the best option when false positives need to be minimized because of its high accuracy. Strong options for scenarios when an application needs recall as well as precision are YOLOv7 and YOLOv8m. Yet, with a better mAP50-95 score, UMFN is still the best method when tasks need strong generalization in various detection scenarios. The results show that UMFN outperforms current YOLO-based models by providing a highly accurate and consistent framework for real-time detection of cauliflower diseases.

Conclusion:

This paper presents the UMFN, an advanced deep learning system specifically designed for the accurate detection and classification of cauliflower diseases. The UMFN model enhances diagnostic accuracy for illnesses and lesion localisation via the integration of Multi-Level Attention mechanisms, IBFL, and advanced multi-scale feature extraction techniques. Comparative evaluations indicate that the UMFN consistently outperforms other models, even the most sophisticated UMFN variant, exhibiting greater mean Average Precision (mAP), precision, and recall. The UMFN model achieved an accuracy of 0.987, surpassing the UMFN model, and attained a recall of 0.916, above the UMFN model's recall of 0.889.The UMFN exhibited outstanding performance in mean Average Precision (mAP), achieving mAP50 values of 0.945 and mAP50-95 values of 0.789, hence proving its robustness over various intersection over union (IoU) thresholds. This study addresses a critical need in precision agriculture and facilitates the broader use of deep learning techniques to improve crop health monitoring, hence increasing agricultural productivity and ensuring food security. This study's innovative technique sets a new benchmark for future research on plant disease detection.

References:

1. Kanna G. P., Kumar S. J., Kumar Y., Changela A., Woźniak M., Shafi J., et al. (2023). Advanced deep learning techniques for early disease prediction in cauliflower plants. *Sci. Rep.* 13, 18475. doi: 10.1038/s41598-023-45403
2. Li Y., Xue J., Wang K., Zhang M., Li Z. (2022). Surface defect detection of fresh-cut cauliflowers based on convolutional neural network with transfer learning. *Foods* 11, 2915. doi: 10.3390/foods11182915
3. Maria S. K., Taki S. S., Mia M. J., Biswas A. A., Majumder A., Hasan F. (2022). Cauliflower disease recognition using machine learning and transfer learning. in *Smart Systems: Innovations in Computing: Proceedings of SSIC 2021* (Springer;) (pp. 359–375).
4. Pourdarbani R., Sabzi S. (2023). Diagnosis of common cauliflower diseases using image processing and deep learning. *J. Environ. Sci. Stud.* 8, 7087–7092. doi: 10.22034/JESS.2023.391624.1995
5. Rajbongshi A., Islam M., Mia M., Sakif T., Majumder A. (2022). A comprehensive investigation to cauliflower diseases recognition: an automated machine learning approach. *Int. J. Adv. Sci. Eng. Inform. Technol. (IJASEIT)* 12, 32–41. doi: 10.18517/ijaseit.12.1.15189
6. Shakil R., Akter B., Shamrat F. J. M., Noori S. R. H. (2023). A novel automated feature selection-based approach to recognize cauliflower disease. *Bull. Electrical Eng. Inf.* 12, 3541–3551. doi: 10.11591/eei.v12i6.5359
7. Singh B., Singh B., Singh P. (2018). Breeding cauliflower: A review. *Int. J. Vegetable Sci.* 24, 58–84. doi: 10.1080/19315260.2017.1354242.
8. Y.A. Nanehkaran, Defu Zhang, Junde Chen, Yuan Tian and Najla Al-Nabhan, "Recognition of plant leaf diseases based on computer vision", *Journal of Ambient Intelligence and Humanized Computing*, pp. 1-18, 2020.
9. S. Mehta, V. Kukreja and R. Yadav, "A Federated Learning CNN Approach for Tomato Leaf Disease with Severity Analysis", *2023 Second International Conference on Augmented Intelligence and Sustainable Systems (ICAISS)*, pp. 309-314, 2023.
10. S. Lamba, A. Baliyan and V. Kukreja, "A novel GCL hybrid classification model for paddy diseases", *Int. J. Inf. Technol*, vol. 15, no. 2, pp. 1127-1136, 2023.
11. V. Singh and A. K. Misra, ‘‘Detection of plant leaf diseases using image segmentation and soft computing techniques,’’ Inf. Process. Agricult., vol. 4, no. 1, pp. 41–49, Mar. 2017, doi: 10.1016/j.inpa.2016.10.005.
12. G. S. Sujawat, ‘‘Application of artificial intelligence in detection of diseases in plants: A survey,’’ Turkish J. Comput. Math. Educ. (TURCOMAT), vol. 12, no. 3, pp. 3301–3305, Apr. 2021, doi: 10.17762/turcomat.v12i3.1581.
13. L. Alzubaidi, J. Zhang, A. J. Humaidi, A. Al-Dujaili, Y. Duan, O. Al-Shamma, J. Santamaría, M. A. Fadhel, M. Al-Amidie, and L. Farhan, ‘‘Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions,’’ J. Big Data, vol. 8, no. 1, pp. 1–74, Mar. 2021, doi: 10.1186/s40537-021-00444-8.
14. M. Iman, H. R. Arabnia, and K. Rasheed, ‘‘A review of deep transfer learning and recent advancements,’’ Technologies, vol. 11, no. 2, p. 40, Mar. 2023, doi: 10.3390/technologies11020040.
15. M. Shoaib, B. Shah, S. EI-Sappagh, A. Ali, A. Ullah, F. Alenezi, T. Gechev, T. Hussain, and F. Ali, ‘‘An advanced deep learning models-based plant disease detection: A review of recent research,’’ Frontiers Plant Sci., vol. 14, Mar. 2023, Art. no. 1158933, doi: 10.3389/fpls.2023.1158933.
16. J. Chen, J. Chen, D. Zhang, Y. Sun, and Y. A. Nanehkaran, ‘‘Using deep transfer learning for image-based plant disease identification,’’ Comput. Electron. Agricult., vol. 173, Jun. 2020, Art. no. 105393, doi: 10.1016/j.compag.2020.105393.

17. A. Malik, G. Vaidya, V. Jagota, S. Eswaran, A. Sirohi, I. Batra, M. Rakhra, and E. Asenso, ''Design and evaluation of a hybrid technique for detecting sunflower leaf disease using deep learning approach,'' J. Food Qual., vol. 2022, pp. 1–12, Apr. 2022, doi: 10.1155/2022/9211700.
18. . Anand, R. K. Mishra, and R. Khan, ''Plant diseases detection using artificial intelligence,'' in Application of Machine Learning in Agriculture, M. A. Khan, R. Khan, and M. A. Ansari, Eds. New York, NY, USA: Academic, 2022, ch. 9, pp. 173–190. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B97803239055030000 72, doi: 10.1016/B978-0-323-90550-3.00007-2
19. J. Xue, L. Huang, B. Mu, K. Wang, Z. Li, H. Sun, H. Zhao, and Z. Li, ''Detection of rotten fresh-cut cauliflowers based on machine vision technology and watershed segmentation method,'' Amer. J. Biochem. Biotechnol., vol. 18, no. 2, pp. 155–167, Feb. 2022.
20. J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, ''ImageNet: A large-scale hierarchical image database,'' in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2009, pp. 248–255.
21. Uddin, Md Sazid, et al. "Cauli-Det: enhancing cauliflower disease detection with modified YOLOv8." Frontiers in Plant Science 15 (2024): 1373590.

Author Details

Author 1: Prof. Preethi B, Assistant Professor, Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology, Davangere, Karnataka, India.
Experience: 10 years of teaching experience.
Area of Interest: Machine Learning, Image Processing, Neural Network
Description: Completed MTech in Computer Science and Engineering and pursuing Ph.D. in VTU, Belgaum.
Email: preethib027@gmail.com
Orchid ID: 0009-0008-8595-0326
Google Scholar Id-https://scholar.google.com/citations?hl=en&user=83L_5N4AAAAJ

Author 2: Dr. Vinutha H P, Professor, Department of Information Science and Engineering, Bapuji Institute of Engineering and Technology, Davangere, Karnataka, India.
Area of Interest: Machine Learning, Image Processing, Neural Network, Deep Learning
Description: 20 years of teaching experience and completed Ph.D. in 2019 VTU, Belgaum.
Email: vinuprasad.hp@gmail.com
Orchid ID:0000-0002-4789-8319
Scopus ID:57201199400
Google Scholar Id- https://scholar.google.com/citations?hl=en&user=sNYzx5wAAAAJ

Author 3: Dr. Patil N S, Associate Professor, Department of Information Science and Engineering, Bapuji Institute of Engineering and Technology, Davangere, Karnataka, India.
Area of Interest: Data Mining, Graph Theory, Networking
Description: 14 years of teaching experience and completed Ph.D. in 2023 VTU, Belgaum.
Email: patilbathi@gmail.com
Orchid ID:0000-0002-9372-5486
Scopus ID:57201805708
Google Scholar Id-https://scholar.google.com/citations?user=hksUwg8AAAAJ&hl=en