IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 15 December 2023, accepted 29 January 2024, date of publication 2 February 2024, date of current version 20 February 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3361756

RESEARCH ARTICLE

Improved Commodity Supply Chain Performance Through AI and Computer Vision Techniques

IRFAN AHMED[1], MOHAMMED ALKAHTANI[2], QAZI SALMAN KHALID[3], AND FAHAD M. ALQAHTANI[2]

[1]Ayass BioScience LLC, Frisco, TX 75034, USA
[2]Department of Industrial Engineering, College of Engineering, King Saud University, P.O. Box 800, Riyadh 11421, Saudi Arabia
[3]Department of Industrial Engineering, University of Engineering and Technology, Peshawar 25000, Pakistan

Corresponding author: Irfan Ahmed (irfanahmed@uetpeshawar.edu.pk)

This work was supported by Researchers Supporting Project number (RSP2024R274), King Saud University, Riyadh, Saudi Arabia.

ABSTRACT In the realm of supply chain management, the impact of Artificial Intelligence (AI) tools on optimizing commodity distribution is undeniable. This study presents the transformative potential of AI and computer vision in the field of commodity supply chain management. The capability of AI to reduce yield loss and enhance supply chain efficiency is a growing trend and vision-based commodity defect monitoring can be useful in this regard. We explored the employment of real-time computer vision techniques in supply chain flaw management, which include Detection Transformer (DETR), a type of Vision Transformer (ViT), and compared its performance with the You Only Look Once (YOLO) and other AI models. Computational feasibility is assessed, encompassing various computer vision and AI models, by using a dataset comprising images of commodity items used to substantiate our findings. The obtained results have shown the improved performance of DETR with a detection and classification accuracy of 96%, directly correlating with improved supply chain management. On the other hand, the higher computational burden imposed by DETR makes it less feasible for the higher constrained embedded applications. The practicality of AI algorithms for real-time defect identification reveals promising prospects for integration into supply chain systems. This research underscores AI's potential to revolutionize commodity supply chain management, extending its benefits to various commodity distribution networks.

INDEX TERMS Supply chain management, computer vision, artificial intelligence, smart agriculture.

I. INTRODUCTION

From the very beginning of the human race, food items supply, and management have remained very crucial for survival. The availability of defective food items may cause deterioration on a large scale. To handle such issues regarding the supply chain management of food items, the natural production of crops has to be sustained under the required constraints. Crop diseases are one of the natural causes of defective food items, which has long posed a significant threat to the sustainability of production and supply of these items. Such defects, be they in the form of contaminated crops or low-quality produce, can have far-reaching consequences on a global scale, affecting not only the economic aspects but also the overall health and well-being of communities. A key catalyst for these issues within the commodity supply chain is the integrity of natural crop production. Maintaining crop health and quality following stringent constraints has always been a complex task. Crop diseases harm our ecosystem because they cause damage to crops, soil, and other things [1]. Farmers are having a difficult time combating the many diseases that are affecting their crops, which has led to a significant drop in both production and the product rating of the yield. It is estimated that 85-95 percent of plant diseases manifest themselves specifically on the plant leaves. Nematodes, bacteria, fungi, and viruses are the four main agents responsible for diseases. Plant disease detection is one of the crucial tasks, which a farmer has to carry out for

The associate editor coordinating the review of this manuscript and approving it for publication was Davide Patti.

producing agricultural products to satisfy the needs of his family and society. The disease identification in plants is a tedious task, that takes a lot of time, requires manpower, and is still prone to human error [2]. One of the main causes of a slower rate of plant development is disease attack. According to the findings of a comprehensive study on agriculture, the quantity and quality of agricultural goods may suffer as a result of the many factors that contribute to plant diseases [3]. Wheat crop is one of the most significant crops used for food purposes throughout the entire world.

In recent years, the agriculture sector has been witnessing a growing emphasis on technological advancements to optimize yields and ensure sustainable practices [4], [5], [6]. Smart agriculture, fuelled by breakthroughs in Artificial Intelligence (AI), holds immense potential in effectively addressing the challenges faced by Small and Medium Enterprises (SMEs) in the agricultural sector [7]. The integration of AI techniques, specifically in disease detection, can revolutionize the way farmers operate, enabling them to mitigate the risks of crop diseases and improve overall productivity. This manuscript explores the application of AI-based disease detection systems within the context of small and medium enterprises in the agricultural sector of many countries, intending to enhance agricultural practices, increase yields, and promote efficient resource management [8], [9], [10].

In this work, we aim to deploy modern computer vision techniques to facilitate local farmers to combat the issues concerning plant disease identification and detection process. Therefore, the experimental procedure is adopted for developing efficient computer vision models, which target the timely prevention of crop diseases using modern computer vision techniques. In this work, we propose a Vision Transformer (ViT) for wheat crop disease detection system, which is called the Detection Transformer (DETR) [11], [12], [13], [14]. The baseline framework has utilized three distinct Machine Learning (ML) and Deep Learning (DL) models, which include Convolutional Neural Network (CNN) [15], Random Forest (RF) [16], and You Only Look Once (YOLO) version 5 [17] algorithms for evaluating our proposed DETR model. Similarly, we also developed a wheat plant disease dataset, in our effort to facilitate and encourage further research to deploy smart computational tools in food and agriculture.

The proposed work has a great potential of employing modern hardware accelerators, where on-field cameras can be utilized in farms, greenhouses, or fields to continuously monitor plant health and detect diseases in real-time. Similarly, IoT-based sensor networks can deploy it for smart precision agriculture systems that can monitor, identify, optimize, manage, and continuously treat wheat crop diseases over small and large areas quickly.

The food industry heavily relies upon the products made of wheat, therefore it attained utmost importance in industrial and household food products. The development of an AI-based method to detect and identify diseases before the production stage will enhance quality management and enable a smooth supply of quality end products. This work is proposed by keeping in view the significance of AI and computer vision in commodity supply chain management in developing countries. We aim at the deployment of cutting-edge computer vision methods to reduce disruption in the supply chain which is mostly due to underproduction or low yield, defects, and bad quality. In this era, the strength of AI is trendy in the recognition of factors to yield loss, which can significantly improve the production cycle and boost the efficiency of the supply chain. Therefore, we present a case study based on a food commodity, as a baseline to evaluate the performances of the proposed AI-based computer vision models.

II. LITERATURE REVIEW

Data-driven techniques have been used in many areas because of the inherent benefits they provide in comparison to conventional methods [18], [19]. Some of the prominent advantages offered by the data-driven methods are their structural learning capability, offline training mechanism, and flexibility to adapt according to the available datasets. Therefore, their applicability has also been observed in plant disease detection. In many agricultural-dependent economies, the primary concerns are the identification of plant diseases and the mitigation of economic damage. To accomplish the task of automatic detection and identification of plant disease using computer vision, the work done in [20] has suggested an approach, which is based on DL-based object detection. The deployed techniques included the Faster Region-based Convolutional Neural Network (Faster R-CNN) and primary Architectures of the Neural Network [3], which were capable of accurately detecting a wide variety of diseases and navigating through the challenging tasks related to moving vehicle, mobile user, and number plate detection. The validation results showed that the accuracy was 94.6 percent, which depicted the feasibility of the CNN and presented the path for an AI-based computational intelligence solution to this complex problem [20].

In another study, Support Vector Machine (SVM) [21] was used to achieve early distinction in comparison to other specific diseases (such as powdery mildew, leaf rust, and Cercospora leaf spot) as well as between healthy and vaccine-protected plants [22]. This was accomplished with a high level of accuracy (up to 90 percent). In the age of data-driven applications, the storage and processing of higher amounts of data is a challenging task. Therefore, many applications are devised to tackle the issue of data storage and complexity [23], [24], to offer the deployment of these applications in real-time. The work conducted by S. R. Maniyath et al. included end-to-end implementation phases, such as dataset preparation, preprocessing, training, and validation of a model. The generated datasets of sick and healthy leaves were pooled, trained using Random Forest, and used to classify images of diseased and healthy leaves. Overall, they may use ML to train the big publicly available data sets to accurately detect the disease that is present in plants on a large scale. This work encompassed several phases of implementation, including the creation of datasets, the extraction of features, the training of the classifier, and classification. To classify images of diseased and healthy leaves, the generated datasets of healthy and diseased leaves were combined and trained using Random Forest. In this study, the technique used for image description was a well-known method called Histogram of an Oriented Gradient (HOG) [25]. General models for the training of large datasets, that are freely available to the public, provide us with a clear method for identifying the disease that is present in plants on a massive scale [25], [26].

Similarly, the proposed hybrid model in [27] used photos of the peach plants' leaves to identify the Infectious spot disease, that was present in peach plants. The experimental process for this study has used the Plant Village dataset [28], a publicly accessible resource to get images of peach plants' leaves. With 9,914 training parameters, the suggested system had a 99.35 percent accuracy rate during training and a 98.38 percent accuracy rate during testing. Using a public dataset made up of 54,306 images of healthy and diseased plant leaves taken in controlled settings, the authors of this study train a DL network to distinguish between 14 crop types and 26 illnesses. This method worked because the trained model's accuracy on a test set was 99.35 percent [29].

The work done in [30] describes the applications of computer vision in warehouse supply management. This work illustrated various studies, describing the importance of cutting-edge computer vision and AI models in the concept of warehouse management. The study suggests the use of AI due to the delicacy of this task, which requires special precision to be accomplished without any fault or mistakes. Also, the work in [31] analyzed the applications of AI for supply chain management systems in Industry 4.0. It is shown that the use of AI in warehouse activities improves the logistics, managerial, and coordinating capabilities of warehouse activities. A critical review of state-of-the-art computer vision models in the food industry has been carried out in [32]. This study termed these technologies as computer vision and AI-driven food industry and exhibited the importance of this technology for food supply chain management. Different opportunities including robotics, drone technology, satellite, precision agriculture, sensor technology, remote sensing, plant data analysis, and smart irrigation systems are explored to comply with the food commodity management. Considering rice as a food commodity, the review of technologies has been carried out in [33]. This study discussed and analyzed the potential applications and adoption of computer vision from the perspective of a supplier, focal firm, and customer for the supply chain management systems. This study conducted three (3) research steps that refer to Kitchenham's SLR; planning, implementation, and reporting, and suggested the use of computer vision models. The SLR process obtained the Critical Success Factor (CSF) component and the SCM test component, while four stages of testing must be carried out to evaluate the process, which are: strategic fit, end-to-end focus, simplicity, and integrity.

By going through the previous literature studies, we found the research gap in coming up with computer vision methods to tackle local problems for detecting and identifying the diseases in commodities, for which the selected case studies are the Wheat and Pea crops. We propose a framework that is effective at automatically detecting and identifying diseases that affect food crops, by using cutting-edge computer vision techniques for stored time and real-time applications. The proposed framework is developed to tackle the issues in cultivating Wheat and Pea crops, to enhance effectiveness in their supply chain management.

III. CONTRIBUTIONS

- The primary objective of our research is to address the underlying issues within the commodity supply chain, with a specific focus on ailment detection and identification to enhance flawless production. To achieve this, we propose a comprehensive framework that leverages state-of-the-art AI and computer vision techniques for the identification and recognition of diseases in wheat crops. Our work distinguishes itself by embracing cutting-edge computer vision models that not only deliver exceptional disease identification capabilities but are also highly compatible with existing hardware systems. These models can seamlessly integrate into the overall supply chain management process. The novelty of our work lies in adopting and proposing cutting-edge computer vision models, which are hardware-friendly and may further be embedded as a full-fledged system. For this reason, various aspects of baseline models - real-time (DETR, YOLO) and stored time (CNN, random forest)- are analyzed which include disease recognition and identification accuracy and computational efficiency.
- In this study, we propose a framework that employs off-the-shelf computer vision models for wheat crop disease identification and recognition systems. The novelty of our work lies in adopting and proposing cutting-edge computer vision models, which comply with the resource-constrained environment and may further be embedded as a full-fledged system. For this reason, various aspects of baseline models - real-time (DETR and YOLO) and stored time (CNN and random forest)- are analyzed which include disease recognition and identification accuracy, computational efficiency, and inference speed.
- Given our ultimate goal, we propose AI and computer vision models that are capable of recognizing food crop diseases. The proposed study evaluated the performances of DETR with the baseline computer vision models based on evaluation metrics, which include model precision, model recall, and model accuracy. The primary focus of this study is the identification of diseases that can affect wheat plant leaves.

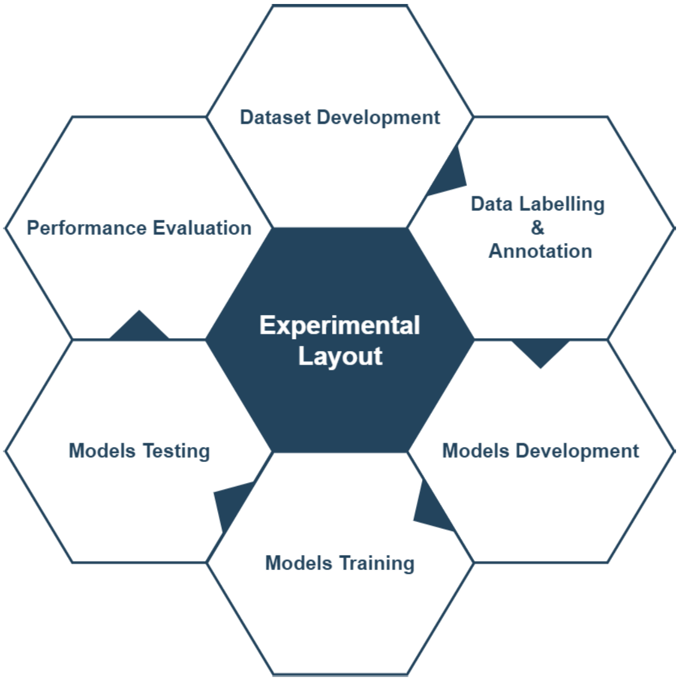

FIGURE 1. Layout of the proposed framework.

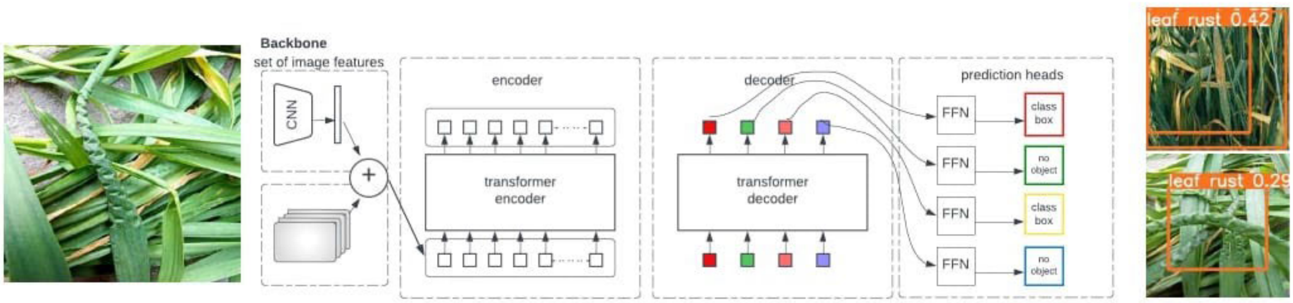

FIGURE 2. Internal architecture of a DETR.

- In our work, four commonly used, vision-based computational models are investigated to demonstrate their deployment in applications, where model accuracy can be compromised over the vision speed and vice-versa, i.e., in real-time applications and stored-time applications respectively. Hence, the achieved objectives are in accordance with the future employments of not only the datasets but also the algorithmic complexity of the model as well.
- The proposed work also carried out a study regarding the computational feasibility of four different ML/DL models for Wheat crop disease detection and identification, comprised of different levels of structural complexities. The CNN and YOLO v5 models are sophisticated models with the capacity to compress images, transforming them into a format that is simpler to process while ensuring that the elements necessary for creating an accurate prediction are not jeopardized. The RF model is a straightforward model with a reputation for quick computation times and simple output interpretation. One of the key elements that contributed to the choice of these models in particular is the variation in application and model complexity across them.

The workflow followed in this research is depicted in the form of a layout shown in Figure 1.

IV. COMPUTER VISION MODELS

In this work, the proposed DETR model is compared with the commonly used, cutting-edge computer vision models to classify and detect diseases in wheat plant images. The baseline models include classical ML models i.e., Random Forest classifier [34], while modern DL algorithms, such as

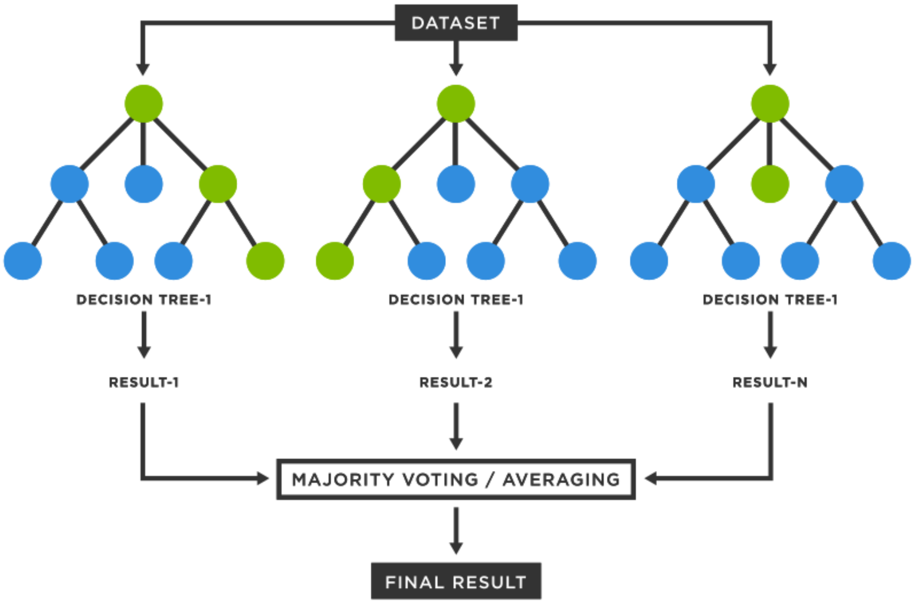

FIGURE 3. Layout of the random forest classifier.

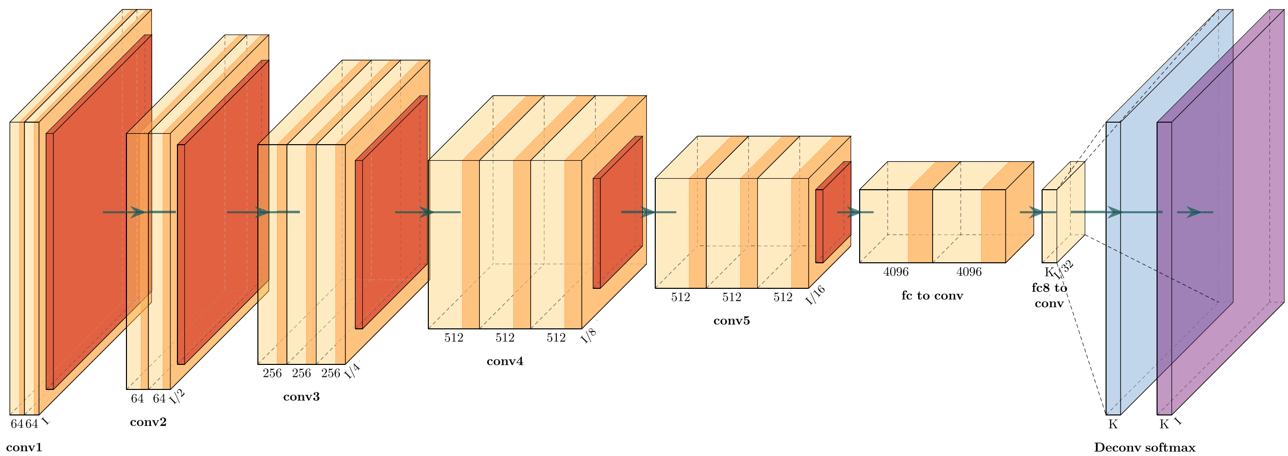

FIGURE 4. Generic architecture of a layer-wise connected convolution NN.

CNN [35], and YOLOv5 [36] are also trained and tested to check their feasibility for the task. A brief description and reason for selecting these algorithms are described in subsequent sections.

A. DETECTION TRANSFORMER (DETR)

The Detection Transformer (DETR) is a transformer architecture, which has drawn a lot of interest since it can deliver cutting-edge performance without relying on custom features or laborious post-processing procedures [37]. It was first introduced by the researchers from Facebook AI Research, by combining the strengths of Vision Transformer (ViT) and CNN [38]. In order to achieve a fully end-to-end framework for object detection, DETR employs a straightforward Transformer encoder-decoder pipeline and does away with the requirement for those manually constructed components. The internal architecture of a DETR is shown in Figure 2.

In Figure 2, we can see that a DETR architecture is composed of an encoder and decoder layers. The input image is processed by the DETR encoder using a CNN, which derives the spatial information in the form of feature maps. The attended features are processed by the feed-forward neural network to provide contextualized representations. The encoder layers aid in the agglomeration of the spatial data and its encoding into a collection of high-level feature representations.

The task of object detection is carried out by the DETR decoder component using the encoded feature representations obtained from the encoder. It is made up of layers of Transformer decoders. Additionally, self-attention and

TABLE 1. CNN model setting.

Block	Layers	Structure		
1	1	Filters= 32	Size=3x3	Activation= relu
	2	Dropout=25%		
	3	Batch Normalization	Max pooling=(2,2)	
2	1	Filters=64	Dimension=3x3	Activation=relu
	2	Dropout=25%		
	3	Batch Normalization	Max Pooling= (2,2)	
3	1	Filters=128	Dimension=3x3	Activation=relu
	2	Dropout=25%		
	3	Batch Normalization	Max Pooling= (2,2)	
4	1	Flattening Layer		

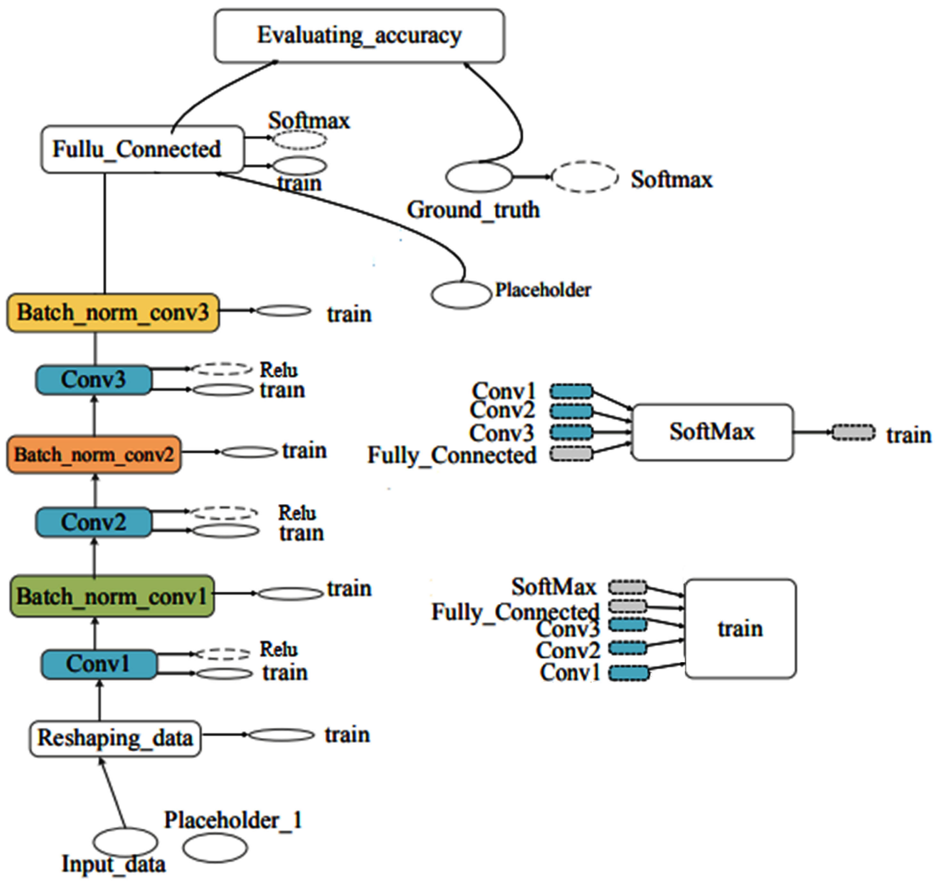

FIGURE 5. Computational graph view of CNN.

feed-forward neural network modules are included in each decoder layer. A novel learnable positional encoding module is also included, which aids in maintaining the spatial information in the object detection process. By paying attention to both the geographical and contextual data, the decoder layers hone the feature representations and provide object queries.

1) RANDOM FOREST

Random Forest is a well-known supervised ML model, which is used for classification and regression applications [39]. Random forest classifiers are ensemble learning, begging algorithms, which build decision trees using a range of samples, then classify data samples using the majority vote of those samples and calculate regression by using

TABLE 2. Dataset breakdown.

S.No.	Disease Name	Number of Pictures
1	Healthy Plants	245
2	Leaf Rust	500
3	Barley Yellow Dwarf	316
4	Powdery Mildew	420

FIGURE 6. (a) Healthy plant (b) Barkley yellow dwarf disease (c) Powdery mildew disease (d) Leaf rust disease.

TABLE 3. Hyperparameters setting for random forest.

Hyperparamter	Settings
n-estimators	100
Random state	42

averaging. The structural layout of a random forest classifier is illustrated in Figure 3. One of the most crucial features of the Random Forest algorithm is its capacity to handle datasets comprising both continuous and discrete variables (e.g., regression tasks), and categorical variables (e.g., classification tasks). The improved accuracy and computational efficiency of classical ML techniques, in general, and Random Forest, specifically, encouraged us to employ them in our research work. The other reason behind the selection of Random Forest for the classification task is the efficiency of the ensemble learning in a variety of tasks.

(a) (b) (c) (d)

FIGURE 7. (a) Healthy plant (b) Barkley yellow dwarf disease (c) Powdery mildew disease (d) Leaf rust disease.

TABLE 4. Parameters used in CNN.

Parameters	Settings
No. of epochs	180
Optimizer function	Stochastic Gradient Descent (SGD)
Learning rate	0.001
Loss function	Sparse Categorical Crossentropy
Batch size	128

2) CONVOLUTIONAL NEURAL NETWORK (CNN)

One of the most widely used DL architectures for images is called the Convolutional Neural Network (CNN). The efficacy of CNN lies in its ability to process images by deploying the convolution operator, which tends to use different operations in which the image is reduced to its essential smaller size structures. These layer-wise reduced structures enable the model parameters to be tuned according to the training images. One such CNN architecture is illustrated in the form of Figure 4. CNN has recently been

TABLE 5. Performance metrics obtained for each model.

Model Name	Accuracy (%)	Precision (%)	Recall (%)
DETR	96	92	90
Random Forest	89	88	88
CNN	92	78	80
YOLO	80	70	70

Model	Computational Burden during Training	Testing Time	Features
Random Forest	Low	Low	Enhanced classification performance with low complexity, but prone to overfitting for a lower amount of training data.
CNN	Moderate	Moderate	Performs complex operations for data processing and an excellent match for higher complexity diseases classification tasks for future work.
YOLO	High	Low	Model training takes long time as compared with its application for object detection, which makes it one of the best models for real-time disease detection.
DETR	High	Low	Similar to YOLO, model training takes long time as compared with its application for object detection, which illustrate its prospects for real-time disease detection.

FIGURE 8. The attained computational complexities for deployed models.

used for plant disease identification and classification tasks as illustrated in [40]. In our work, we implemented CNN due to its efficiency in image and video classification tasks. Our model settings are listed below in the form of Table 1. In this framework, three-channel inputs are applied at the input layer which propagates through all the layers and results in a three-channel labeled output image.

Here it should be noted that a flattening operation is carried out to convert a 3D image into a 1D flattened form before its propagation to the final convolution layer, for a pixel-level classification task.

The computational graph of the employed CNN model is also shown in Figure 5.

3) YOU ONLY LOOK ONCE (YOLO)

You Only Look Once (YOLO) is a cutting-edge, real-time object detection system that is incredibly quick and precise. YOLO excels at real-time target detection due to its speed and detection abilities, making it the ultimate option for real-time object detection.

YOLOv5 is the fifth version of YOLO, which is developed by the Ultralytics LLC team [41]. YOLOv5 features high accuracy both in terms of detection and inference speed [36]. Since YOLOv5 is a single-stage object detector, it functions similarly to all other single-stage object detectors. YOLO v5 is comprised of three essential components.

- Model backbone
- Model head
- Model neck

Extraction of important features from the input image is the main goal of Model Backbone. The CSP (Cross Stage Partial) Networks are used as the foundation of YOLOv5 to extract features from an input image that are rich in informative content.

The weight file in YOLOv5 is of a tiny size and is approximately 90 percent smaller than the weight file in YOLOv4.

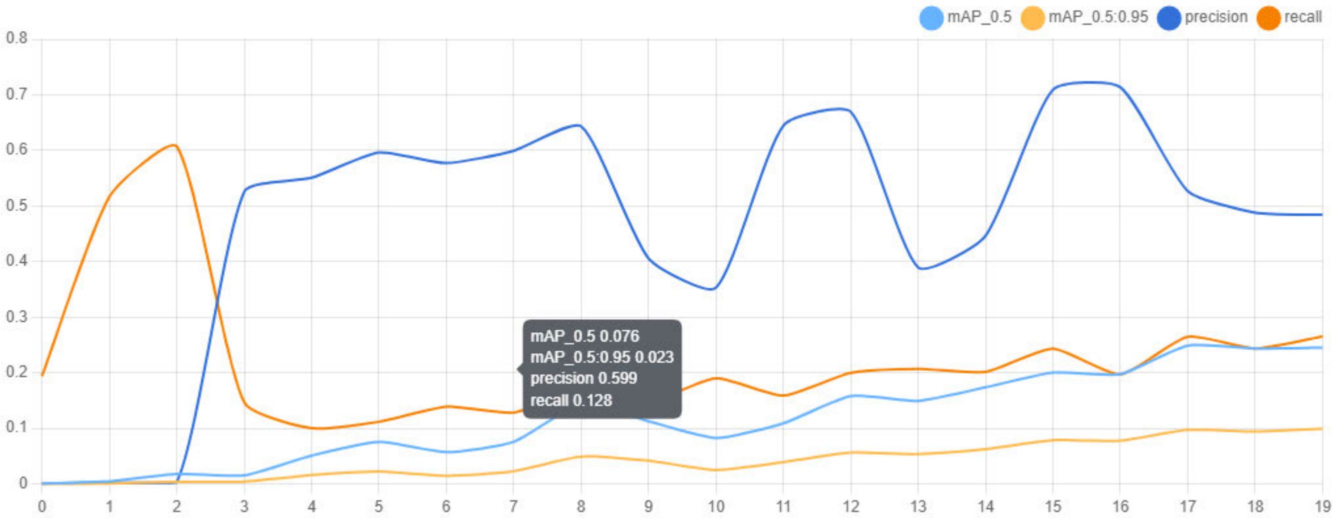

FIGURE 9. Precision VS recall curve.

It is hence appropriate for use in embedded systems for real-time detection. The YOLOv5 delivers increased detection accuracy over earlier versions, a lightweight design, and a short detection time. One of the reasons for choosing it for our study is its prior efficacy, which is crucial in similar tasks.

V. EXPERIMENTAL SETUP

This section is composed of subsections that are related to the experimental environment, settings, and performance evaluation metrics used in our study.

A. EXPERIMENTAL ENVIRONMENT

The experiments were conducted at Google servers to speed up the model training using freely available hardware accelerators i.e., GPU: 1xTesla K80, compute 3.7, having 2496 CUDA cores, 12GB GDDR5 VRAM, and 12GB RAM. The developed datasets were first imported into the environment, which were then used to train and test the models after carrying out the desired changes.

B. PERFORMANCE METRICS

In this work, the prime task is the classification of plant images to detect plant diseases. Therefore, we used the commonly used performance evaluation metrics as the performance metrics. Further, the details of deployed metrics are listed below.

1) MODEL ACCURACY

The accuracy of a model is one of the key performance indicators, which describes how accurate the model is in predicting the true labels. Mathematically it can be written as:

$$Model\_Accuracy = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \tag{1}$$

where:

T_P is True Positive
T_N is True Negative
F_P is False Positive
F_N is False Negative

2) PRECISION

By dividing the entire number of correct predictions (True positive) over the count of total predictions, the precision of the classification model is determined [42]. This measure illustrates how accurate a classifier is by providing an answer to the question of how many samples are correctly identified as "Leaf Rust" were truly "Leaf Rust" out of all the samples. It can be written as:

$$Precision = \frac{T_P}{T_P + F_P} \tag{2}$$

3) RECALL

The number of right forecasts, for positive predictions, are known as true positives. Recall is also another performance metric, which is used to assess the developed model by dividing true positive forecasts over the summation of true positive forecasts and false negative forecasts [42]. This statistic is designed to emphasize the sensitivity of the model, which means that out of all the samples in the dataset that have been labeled as having "Leaf Rust" how many of those were also recognized by the classifier as having "Leaf Rust" in its predictions? It is determined as:

$$Recall = \frac{T_P}{T_P + F_N} \tag{3}$$

Equation 3 is used to calculate model precision for positive classes.

Similarly, for negative classes, there is also another variant of 3, which is defined as:

$$Recall\_Score = \frac{T_N}{T_P + F_N} \tag{4}$$

FIGURE 10. Plant disease classification with DETR.

VI. METHODOLOGY

In this section, a detailed discussion regarding the key methodological steps has been included. We have included the subsections, which are related to the dataset development, labeling, and selection of the parameters and hyperparameters.

A. DATASET DEVELOPMENT

In this work, dataset development was one of the prime objectives, due to the unavailability of publicly available wheat crop diseases dataset. Recently, database development task was carried out in many ML-based research works, due to the unavailability of the requisite datasets [43], [44]. Therefore, we developed our own image dataset for this research work. We collected 1500 images of four different categories of wheat crop images from a variety of fields. Out of these four classes, three classes were labeled with corresponding diseases, while one class was named ‘‘healthy’’. The allotted class names are healthy, barley yellow dwarf, leaf rust, and powdery mildew. For the sake of applying the DL and ML algorithms, the images were reduced to 224 × 224 pixels, to have consistent dimensionality. The dataset images were taken with a high-quality mobile phone camera of 64 MP.

In many areas of the world, the commonly observed diseases in wheat crops are listed in Table 2. Therefore, our problem breakdown has been specifically included for these diseases. Table 2 also contains information about our dataset, including the number of images gathered for each ailment. For this reason, the images dataset for the following classes of wheat plant diseases were collected. The types of images are shown in Figure 6.

FIGURE 11. **Real Time plant disease detection with YOLO.**

In this study, we kept this ratio 70/30 for our train/test data split due to the unavailability of a higher amount of data. Similarly, the available data is balanced and stratified to avoid sampling and class biases.

B. DATA LABELLING AND ANNOTATION

The images used for training and testing, in real-time disease detection, are labeled in Roboflow [45], each image is labeled by drawing the bounding box and class labeling. Two separate

FIGURE 12. Plant disease classification with CNN.

files containing texts and classes are obtained after the image labeling step. The text file consists of a class label and the orientation of the bounding box [46]. Since numbers between 0 and 1 are simple to predict, all of the values in the text file are normalized between these ranges. The number of drawn bounding boxes affects the number of lines in a text file. To label both the train and test datasets, the dataset is split into images for training and test purposes at a rate of 80%

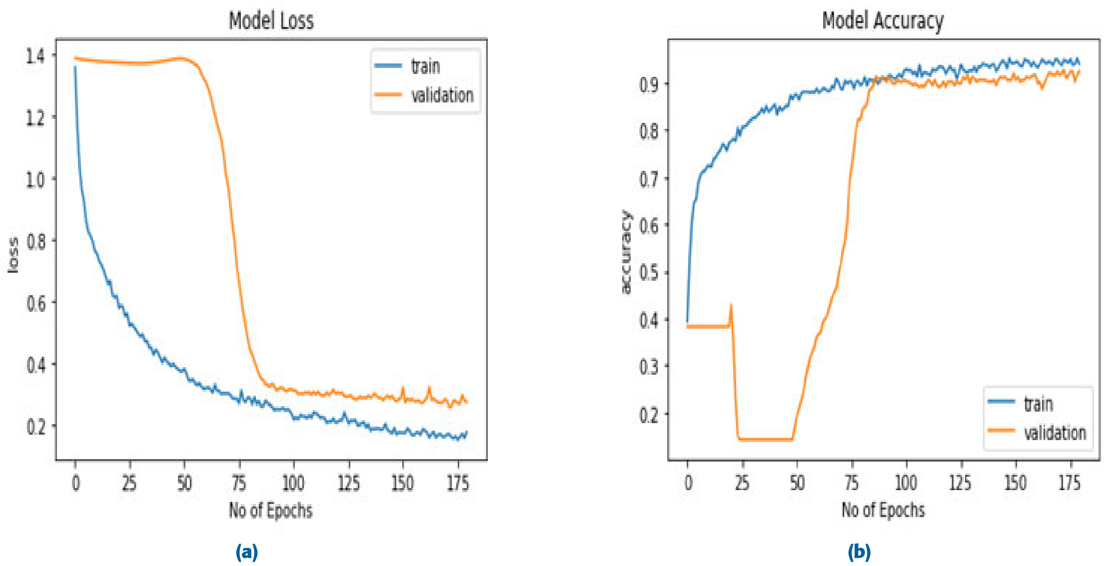

FIGURE 13. (a) Training and validation loss of CNN (b) Training and validation accuracy of CNN.

and 20%, respectively. Some examples of annotated images are shown in Figure 7.

C. PARAMETERS AND HYPERPARAMETER SELECTIONS

In ML, hyperparameters are values that the user directly defines to regulate the learning process. These values are determined before the start of a learning process because a rigorously tuned set of hyperparameters is essential for ML and DL to demonstrate effectively while having the optimized set of parameters. A set of key hyperparameters employed in this study are:

- **n-estimators**: The number of trees the algorithm creates before averaging the result.
- **Random State**: Regulates how random the sample selection process is. The model will dependably produce the same results if it is given the same hyperparameters, training data, and random state with a fixed value.

The important hyperparameters used for this research are as follows. These are the parameters that were used for Random Forest in Table 3and CNN 4, as well.

The selection of hyperparameters is based on a careful tuning of the model in a trial-and-error manner. In this work, our model is trained for various numbers of epochs and the optimum number is selected which provides the approximate accuracy convergence for preserving computational, power, and storage resources. Similarly, learning rate is also one of the crucial hyperparameters, which is selected from a range of hyperparameters based on the speed and performance convergence trade-off.

In our model, the convolutional layer is the initial layer that is utilized to extract the various features from the input images. In this study, we constructed the following convolutional layers with the mentioned parameters:

1) Size = (32, 3, 3), Strides = 2, Padding = same, Activation = Relu.
2) Size = (3,3), Strides = 2, Padding = same, Activation = Relu.
3) Fully connected layer, Dropout = 0.5, Activation = Softmax.

This study does not employ the FC layer without dropout. The activation function Relu is used, and in the final step, we apply the softmax function to classify the data into four different groups.

VII. RESULTS

In our work, we used the commonly used ML and DL algorithms to check the feasibility of the proposed framework, in terms of computation resources utilization, power efficiency, and process time (to check for real-time applications). The attained performances achieved, in terms of accuracy, precision, and recall, are shown in the form of Table 5 and discussed in the subsequent section.

As is evident from Table 5, the DETR outperformed in all avenues of performance comparison. In comparison to our model, RF scored 89 percent, and the YOLO v5 model scored 80 percent, while the CNN model achieved an impressive 92 percent accuracy rate. In comparison to real-time detection in the YOLO model, the DETR has outperformed it in the field of accuracy, precision, and recall while it has some limitations of working which are GPU usage limit.

We found that the model accuracies reported for other plant disease identification are higher as compared to our study. When observed keenly, we came to know that the training data was composed of images of different qualities. Therefore, low-quality training images were one of the sources, which induced accuracy reduction in our methods. Therefore, the achieved model accuracies can be further enhanced by incorporating high-quality image samples.

Similarly, It is clear from looking at Table 5, that the CNN model has a recall of 82 percent, while the YOLO v5 model has a recall of 70 percent, and the RF model has a recall of 88 percent. This indicates that the RF model can return most of the pertinent results - after our proposed framework- in the dataset in comparison to other models. While CNN achieves a precision score of 80 percent, as shown in Table 5, the RF model scores 88 percent, and YOLO v5 scores 70 percent. This demonstrates that when compared to the other two models, the RF model will significantly produce more relevant results than it would return irrelevant ones. The achieved efficacy in terms of classification results of DETR exhibits the prospects of ViTs for the described categories of plant diseases. The categorical labels for the healthy and diseased plants aided the deployment of classical DETR as compared with the classical ML and CNN models, keeping in view the lower complexity of the task at hand. It is, further, investigated that CNN can perform well if more training data is provided for a multi-class complex problem set.

When we compared the computational efficiency of our models, the random forest has shown promising results. Similarly, the begging and voting operations performed by the RF classifier are more computationally feasible as stated in Table 8. Therefore, we found it feasible to be utilized for the applications, where computational complexity is fundamentally important.

The objectness score, the class probability score, and the bounding box regression score all contribute to the YOLO family's calculation of a compound loss. The Binary Cross-Entropy with Logits Loss function, which is available in PyTorch, was used by Ultralytics to determine the loss of class probability and object score [47].

As far as the real-time applications are concerned, the proposed DETR model and YOLO demonstrated higher classification accuracy when tested in real-time tasks for disease detection as listed in the table shown in Figure 8. The faster accuracy of DETR has been achieved at the cost of longer and more tedious offline training, which is also the same case with YOLO v5. Therefore, its accuracy is prone to further improvements if the amount of training data is increased, not only in quantity but in quality as well. The results obtained with real-time plant detection are shown in the form of Figure 10 and 11 The objectness score, the class probability score, and the bounding box regression score all contribute to the YOLO family's calculation of a compound loss. The Binary Cross-Entropy with Logits Loss function, which is available in PyTorch, was used by Ultralytics to determine the loss of class probability and object score [47].

As far as the real-time applications are concerned, the proposed DETR model and YOLO demonstrated higher classification accuracy when tested in real-time tasks for disease detection as listed in Table 8. The faster accuracy of DETR has been achieved at the cost of longer and more tedious offline training, which is also the same case with YOLO v5. Therefore, its accuracy is prone to further improvements if the amount of training data is increased, not only in quantity but in quality as well. For our models, the precision-recall curve has been shown in the form of Figure 9.

The results obtained with real-time plant detection are shown in the form of Figure 10 and 11

Similarly, the classification performance is illustrated in the form of Figure 12.

VIII. TRAINING AND VALIDATION ACCURACY

In ML practices, a model is said to be underfitted when it performs badly on both the datasets it was trained on and the datasets it was validated and tested on. An ML model is said to be overfitted if it performs perfectly on a training sample but badly during the validation/test steps. Accuracy and loss plots concerning the number of epochs are used to track CNN's performance throughout these steps. These plots indicate that the training and validation accuracies/losses are directly proportional, showing that the model is continuously learning and is neither overfitting nor underfitting. A model is said to have a good fit when it can appropriately generalize its findings from the validation dataset and learn from the training dataset. Figure 13 displays the CNN classifier's training and validation loss.

IX. LIMITATIONS, SUGGESTIONS, AND FUTURE DIRECTIONS

In this study, DETR has shown promising results for the detection and identification of wheat crop diseases as defects in a food commodity. However, we observed some limitations and research gaps which can be helpful for future research in this field. These are listed below:

- **Data greedy approach:** Although the proposed work has shown promising results, yet the observed performances can be improved further. As we have observed, DETR constitutes an efficient framework by training it over a mediocre amount of available data. The sole cause for this average performance was the availability of limited training data. It is observed that more training data will enhance the performance of the framework by converging to the optimal performance level. Therefore, the efficiency will further increase by incorporating more data samples in the training set.
- **Higher training time:** Model training is one of the crucial tasks for achieving the required results using data-driven methodologies by using the available dataset. The accuracy vs. epochs chart shows that the model tends to further improve during the experimental process. As a result, its performance level is more susceptible to increased usage of computing, time, and power resources.
- **Deploying more models:** In the future, more computational methods can be deployed for object detection and identification while other data-driven machine learning and deep learning models can also be used for further commodity supply chain management tasks. Furthermore, the scope of this work can also be widened

by using the deployed techniques for other datasets related to supply chain management.

X. CONCLUSION

The proposed work targets the empowerment of local farmers by deploying the cutting-edge object detection model called DETR. Our model is trained on the local crops dataset emphasizing the development of wheat crop dataset from wheat-producing areas and compared with the results obtained from CNN, YOLO v5, and Random Forest classifier based on accuracy, precision, and recall metrics. In all three evaluation criteria, the DETR outperforms for detecting plant illnesses on wheat crop leaves, in terms of model accuracy and demonstrated 4% higher accuracy than the best baseline model. In this work, the utilization of the Random Forest model foresees that our work was more accurate and completed in a shorter amount of time. Additionally, we have observed that CNN stands as a formidable choice for complex, stored-time tasks, hinting at the possibility of refining its architecture to unlock further improvements. As we tread into the future, the integration of these models into broader agricultural contexts holds immense promise, marking a decisive stride toward more effective and responsive crop disease management in the commodity supply chain. This path of progress beckons, and its realization is well within reach.

ACKNOWLEDGMENT

This work was supported by Researchers Supporting Project number (RSP2024R274), King Saud University, Riyadh, Saudi Arabia.

REFERENCES

[1] M. R. Ullah, N. A. Dola, A. Sattar, and A. Hasnat, "Plant diseases recognition using machine learning," in *Proc. 8th Int. Conf. Syst. Model. Advancement Res. Trends (SMART)*, Nov. 2019, pp. 67–73.

[2] S. Poornima, S. Kavitha, S. Mohanavalli, and N. Sripriya, "Detection and classification of diseases in plants using image processing and machine learning techniques," *AIP Conf. Proc.*, vol. 2095, no. 1, Apr. 2019, Art. no. 030018.

[3] J. Shirahatti, R. Patil, and P. Akulwar, "A survey paper on plant disease identification using machine learning approach," in *Proc. 3rd Int. Conf. Commun. Electron. Syst. (ICCES)*, Oct. 2018, pp. 1171–1174.

[4] S. Alqethami, B. Almtanni, W. Alzhrani, and M. Alghamdi, "Disease detection in apple leaves using image processing techniques," *Eng., Technol. Appl. Sci. Res.*, vol. 12, no. 2, pp. 8335–8341, Apr. 2022.

[5] A. Ksibi, M. Ayadi, B. O. Soufiene, M. M. Jamjoom, and Z. Ullah, "MobiRes-Net: A hybrid deep learning model for detecting and classifying olive leaf diseases," *Appl. Sci.*, vol. 12, no. 20, p. 10278, Oct. 2022.

[6] R. Sharma, A. Singh, Kavita, N. Z. Jhanjhi, M. Masud, E. S. Jaha, and S. Verma, "Plant disease diagnosis and image classification using deep learning," *Comput., Mater. Continua*, vol. 71, no. 2, pp. 2125–2140, 2022.

[7] L. M. Cornejo-Bueno, J. Pérez-Aracil, C. Casanova-Mateo, J. Sanz-Justo, and S. Salcedo-Sanz, "Machine learning classification-regression schemes for desert locust presence prediction in Western Africa," *SSRN Electron. J.*, Sep. 2022.

[8] B. Min, T. Kim, D. Shin, and D. Shin, "Data augmentation method for plant leaf disease recognition," *Appl. Sci.*, vol. 13, no. 3, p. 1465, Jan. 2023.

[9] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," *Frontiers Plant Sci.*, vol. 7, p. 1419, Sep. 2016.

[10] Z. Shao, X. Feng, L. Bai, H. Jiao, Y. Zhang, D. Li, H. Fan, X. Huang, Y. Ding, O. Altan, and N. Saleem, "Monitoring and predicting desert locust plague severity in Asia–Africa using multisource remote sensing time-series data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 8638–8652, 2021.

[11] H.-T. Thai, K.-H. Le, and N. L.-T. Nguyen, "FormerLeaf: An efficient vision transformer for cassava leaf disease detection," *Comput. Electron. Agricult.*, vol. 204, Jan. 2023, Art. no. 107518.

[12] Y. Shen, L. Wang, and Y. Jin, "AAFormer: A multi-modal transformer network for aerial agricultural images," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2022, pp. 1704–1710.

[13] V. G. Dhanya, A. Subeesh, N. L. Kushwaha, D. K. Vishwakarma, T. N. Kumar, G. Ritika, and A. N. Singh, "Deep learning based computer vision approaches for smart agricultural applications," *Artif. Intell. Agricult.*, vol. 6, pp. 211–229, 2022.

[14] Z. Yang, J.-H. Lai, J. Zhou, H. Zhou, C. Du, and Z. Lai, "Agriculture-vision challenge 2022—The runner-up solution for agricultural pattern recognition via transformer-based models," 2022, *arXiv:2206.11920*.

[15] L. Jiao, S. Dong, S. Zhang, C. Xie, and H. Wang, "AF-RCNN: An anchor-free convolutional neural network for multi-categories agricultural pest detection," *Comput. Electron. Agricult.*, vol. 174, Jul. 2020, Art. no. 105522.

[16] S. M. Basha, D. S. Rajput, J. Janet, R. S. Somula, and S. Ram, "Principles and practices of making agriculture sustainable: Crop yield prediction using random forest," *Scalable Comput., Pract. Exp.*, vol. 21, no. 4, pp. 591–599, Dec. 2020.

[17] R. Gai, N. Chen, and H. Yuan, "A detection algorithm for cherry fruits based on the improved YOLO-v4 model," *Neural Comput. Appl.*, vol. 35, no. 19, pp. 13895–13906, Jul. 2023.

[18] I. Ahmed, A. Khan, A. Khan, K. Mujahid, and N. Khan, "Efficient measurement matrix for speech compressive sampling," *Multimedia Tools Appl.*, vol. 80, no. 13, pp. 20327–20343, May 2021.

[19] I. Ahmed, A. Khan, N. Ahmad, NasruMinallah, and H. Ali, "Speech signal recovery using block sparse Bayesian learning," *Arabian J. Sci. Eng.*, vol. 45, no. 3, pp. 1567–1579, Mar. 2020.

[20] G. Shrestha, M. Das, and N. Dey, "Plant disease detection using CNN," in *Proc. IEEE Appl. Signal Process. Conf. (ASPCON)*, Oct. 2020, pp. 109–113.

[21] A. Widodo and B.-S. Yang, "Support vector machine in machine condition monitoring and fault diagnosis," *Mech. Syst. Signal Process.*, vol. 21, no. 6, pp. 2560–2574, Aug. 2007.

[22] J. Behmann, A.-K. Mahlein, T. Rumpf, C. Römer, and L. Plümer, "A review of advanced machine learning methods for the detection of biotic stress in precision crop protection," *Precis. Agricult.*, vol. 16, no. 3, pp. 239–260, Jun. 2015.

[23] I. Ahmed and A. Khan, "Genetic algorithm based framework for optimized sensing matrix design in compressed sensing," *Multimedia Tools Appl.*, vol. 81, no. 27, pp. 39077–39102, Nov. 2022.

[24] I. Ahmed, A. Khalil, I. Ahmed, and J. Frnda, "Sparse signal representation, sampling, and recovery in compressive sensing frameworks," *IEEE Access*, vol. 10, pp. 85002–85018, 2022.

[25] S. Ramesh, R. Hebbar, M. Niveditha, R. Pooja, B. N. Prasad, N. Shashank, and P. V. Vinod, "Plant disease detection using machine learning," in *Proc. Int. Conf. Design Innov. 3Cs Compute Communicate Control (ICDI3C)*, Apr. 2018, pp. 41–45.

[26] I. Ahmed and A. Khan, "Learning based speech compressive subsampling," *Multimedia Tools Appl.*, vol. 82, no. 10, pp. 15327–15343, Apr. 2023.

[27] P. Bedi and P. Gole, "Plant disease detection using hybrid model based on convolutional autoencoder and convolutional neural network," *Artif. Intell. Agricult.*, vol. 5, pp. 90–101, 2021.

[28] D. Singh, N. Jain, P. Jain, P. Kayal, S. Kumawat, and N. Batra, "PlantDoc: A dataset for visual plant disease detection," in *Proc. 7th ACM IKDD CoDS 25th COMAD*, Jan. 2020, pp. 249–253.

[29] J. Boulent, S. Foucher, J. Théau, and P.-L. St-Charles, "Convolutional neural networks for the automatic identification of plant diseases," *Frontiers Plant Sci.*, vol. 10, p. 941, Jul. 2019.

[30] H. Patel, "Computer vision for supply chain management optimization," *ScienceOpen Preprints*, 2022.

[31] S. Jagadeesan, D. K. Malik, S. Bharti, S. P. Singh, R. K. Ibrahim, and M. B. Alazzam, "Artificial intelligence in supply chain management in Industry 4.0," in *Proc. 3rd Int. Conf. Advance Comput. Innov. Technol. Eng. (ICACITE)*, May 2023, pp. 2722–2726.

[32] V. Kakani, V. H. Nguyen, B. P. Kumar, H. Kim, and V. R. Pasupuleti, ''A critical review on computer vision and artificial intelligence in food industry,'' *J. Agricult. Food Res.*, vol. 2, Dec. 2020, Art. no. 100033.
[33] P. A. W. Putro, E. K. Purwaningsih, D. I. Sensuse, R. R. Suryono, and Kautsarina, ''Model and implementation of rice supply chain management: A literature review,'' *Proc. Comput. Sci.*, vol. 197, pp. 453–460, Jan. 2022.
[34] W. Lin, Z. Wu, L. Lin, A. Wen, and J. Li, ''An ensemble random forest algorithm for insurance big data analysis,'' *IEEE Access*, vol. 5, pp. 16568–16575, 2017.
[35] J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, B. Shuai, T. Liu, X. Wang, G. Wang, J. Cai, and T. Chen, ''Recent advances in convolutional neural networks,'' *Pattern Recognit.*, vol. 77, pp. 354–377, May 2018.
[36] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, ''You only look once: Unified, real-time object detection,'' in *Proc. IEEE Conf. Comput. Vis. pattern Recognit.*, 2016, pp. 779–788.
[37] D. Meng, X. Chen, Z. Fan, G. Zeng, H. Li, Y. Yuan, L. Sun, and J. Wang, ''Conditional DETR for fast training convergence,'' in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, Oct. 2021, pp. 3651–3660.
[38] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, and S. Zagoruyko, ''End-to-end object detection with transformers,'' in *Proc. Eur. Conf. Comput. Vis.* Cham, Switzerland: Springer, 2020, pp. 213–229.
[39] P. Panchal, V. C. Raman, and S. Mantri, ''Plant diseases detection and classification using machine learning models,'' in *Proc. 4th Int. Conf. Comput. Syst. Inf. Technol. Sustain. Solution (CSITSS)*, Dec. 2019, pp. 1–6.
[40] A. Darwish, D. Ezzat, and A. E. Hassanien, ''An optimized model based on convolutional neural networks and orthogonal learning particle swarm optimization algorithm for plant diseases diagnosis,'' *Swarm Evol. Comput.*, vol. 52, Feb. 2020, Art. no. 100616.
[41] M. P. Mathew and T. Y. Mahesh, ''Leaf-based disease detection in bell pepper plant using YOLO v5,'' *Signal, Image Video Process.*, vol. 16, no. 3, pp. 841–847, Apr. 2022.
[42] M. Hossin and M. N. Sulaiman, ''A review on evaluation metrics for data classification evaluations,'' *Int. J. Data Mining Knowl. Manage. Process.*, vol. 5, no. 2, pp. 1–11, Mar. 2015.
[43] I. Ahmed, H. Ali, N. Ahmad, and G. Ahmad, ''The development of isolated words corpus of Pashto for the automatic speech recognition research,'' in *Proc. Int. Conf. Robot. Artif. Intell.*, Oct. 2012, pp. 139–143.
[44] I. Ahmed, N. Ahmad, H. Ali, and G. Ahmad, ''The development of isolated words Pashto automatic speech recognition system,'' in *Proc. 18th Int. Conf. Autom. Comput. (ICAC)*, Sep. 2012, pp. 1–4.
[45] S. Alexandrova, Z. Tatlock, and M. Cakmak, ''RoboFlow: A flow-based visual programming language for mobile manipulation tasks,'' in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2015, pp. 5537–5544.
[46] G. Yang, J. Jin, Q. Lei, Y. Wang, J. Zhou, Z. Sun, X. Li, and W. Wang, ''Garbage classification system with YOLOV5 based on image recognition,'' in *Proc. IEEE 6th Int. Conf. Signal Image Process. (ICSIP)*, Oct. 2021, pp. 11–18.
[47] E. Stevens, L. Antiga, and T. Viehmann, *Deep Learning with PyTorch*. Shelter Island, NY, USA: Manning, 2020.

IRFAN AHMED received the Ph.D. degree in computer systems engineering, with a focus on the employment of data-driven models for higher dimensional signals from a limited number of information. He is currently a Machine Learning Engineer with Ayass BioScience LLC, Frisco, TX, USA. His research interests include the applications of machine learning/deep learning for developing a predictive analysis system for disease identification and detection systems.

MOHAMMED ALKAHTANI received the B.Sc. degree in industrial engineering from King Saud University (KSU), Riyadh, Saudi Arabia, the M.Sc. degree in industrial engineering from the University of Central Florida, Orlando, FL, USA, and the Ph.D. degree in manufacturing engineering from Loughborough University, Loughborough, U.K. From 2014 to 2018, he was the Chairman of the Industrial Engineering Department, College of Engineering, KSU, for over four years. He is currently a Professor in industrial engineering with the Industrial Engineering Department and the Vice Dean of the Advanced Manufacturing Institute, KSU. His research interests include analysis, modeling, and design of manufacturing systems, supply chain and operations management, responsiveness measurement, leanness, and agility in manufacturing and supply chain.

QAZI SALMAN KHALID received the bachelor's degree in industrial engineering from the University of Engineering and Technology (UET), Peshawar, in 2013, and the master's degree in industrial engineering from UET, Taxila, in 2016. He is currently pursuing the Ph.D. degree with UET, Peshawar. He is a dedicated Lecturer with the Department of Industrial Engineering, UET, Peshawar. His academic journey was underscored by a deep-seated passion for addressing intricate operational research challenges. Notably, he has fostered active collaborations with researchers across various industrial engineering disciplines. His current research interests include the optimization of diverse domains, including vehicle routing, manufacturing system scheduling, combinatorial optimization, and supply chain operations. In addition to his research pursuits, he actively contributes to the academic community as a discerning reviewer of numerous renowned journals in the same field.

FAHAD M. ALQAHTANI received the B.Sc. degree from King Saud University, Saudi Arabia, and the M.S. and Ph.D. degrees from West Virginia University, Morgantown, WV, USA. He is currently an Assistant Professor with the Industrial Engineering Department, College of Engineering, King Saud University. His research interests include system safety, health and safety in construction, industrial hygiene, and human factors engineering. He has been a member of the Saudi Council of Engineers, since November 2016.