Automatic Field Monitoring in Smart Agriculture Using Segmentation and Classification Based on YOLO-V4 and Xception Networks

T. Akilan[1]* , S. Premkumar[2] , K M Baalamurugan[3]

[1] School of Computing Science and Engineering, Galgotias University, Greater Noida 203201, India
[2] Department of Computer Science and Engineering, Galgotias University, Greater Noida 203201, India
[3] School of Engineering, IILM University, Greater Noida 201306, India

Corresponding Author Email: agilanmecse@gmail.com

ABSTRACT

The economic development of every nation depends heavily on agriculture. It leads to industrial and services sector expansion as it generates income, increases consumer demand, and provides raw material for various industries. Fulfilling the current population's food needs is becoming increasingly difficult due to population growth, frequent climatic changes, reliance on manual inspection, poor generalization, time-consuming, and limited resources. To address this issue, a smart agriculture system based on IoT and deep learning has been developed. Acquired images from sensors, such as plant diseases and intruder images, are used as input. The image preprocessing stages include Median Modified Wiener Filter (MMWF) for noise reduction and Joint Histogram Equalization (JHE) to enhance contrast and visual quality. After processing, YOLO-V4 is used to segment the images, followed by classification via a modified Xception network. When a plant disease is detected, a pesticide sprayer can be manually connected to the Android application. In the event of intruder detection, a buzzer alert system is activated to notify farmers of potential threats. The disease classification model tested on diseases predicted an accuracy of 96% while the intruder prediction model reached 98% accuracy. Compared to existing models (InceptionV2, ResNeXt50, ResNet50, and VGG16) with respective accuracies of 94%, 88%, 83%, 78% and 95%, 89%, 86%, 82% (for intruder prediction), the proposed system provides reliability for improving crop health monitoring.

1. INTRODUCTION

Agriculture is the source of 70% of the population in India. India's main line of work is agriculture. The Indian economy is strong; that much is true. Supporting the farm industry is crucial if we want to prevent food shortages. Agriculture must be improved to fulfill the fast-rising global food demand brought on by population expansion [1]. Farm productivity can be raised by anticipating and considering natural conditions. Applying manure effectively requires accurate disease discovery and recognition. If the infections are not identified at the outset, crop growth will diminish. Agriculture is undergoing a significant transition when it comes to obtaining and using data to inform practical farming decisions. Since the majority of cultivators spend an excessive amount of time in the crop fields, manual farming takes time as well [2]. Additionally, using the manual monitoring method prevents farmers from being aware of intruders, whether people or animals, who might harm or steal their plants.

The use of modern information and communication strategy in agriculture, including the machine learning (ML) approach, as well as an explanation of how natural resources are used in capital-based systems and sophisticated technology in clean, sustainable approaches, is known as smart farming [3]. At this

time, the Internet of Things (IoT) and data analysis techniques like big data analytics and data science have started to play a significant role in people's daily lives, enhancing their capacity to modify their surroundings [4]. IoTs and data processing are typically used in the agro-industrial and ecological sectors to control and diagnose smart farming systems while also giving consumers and final farmers important information about the origins and characteristics of agricultural goods and systems [5].

Recent years have seen the development of a number of strategies to deal with the problems and barriers that effective farming presents, including identifying diseases, yield prediction, recognizing species, crop growth difficulties, drought, and managing irrigation [6]. Some studies have been on applying data analysis techniques to improve farming data decision support systems. Most of them work using conventional methods without considering how well the model performs, and some skip data preprocessing in the early stages. IoT provides solutions in several industries, including farming, medical care, safety, smart homes, sales, and smart cities [7]. The use of IoT in agriculture is the best option because this industry requires constant monitoring and control. IoT is utilized in agriculture for cattle, greenhouses, and micro-agricultural applications, which are organized into

several control regions [8]. Through the use of wireless sensor networks (WSNs), which assist farmers in gathering pertinent data from IoT-based sensors, with the help of numerous Internet-based information sensors and devices, all of these applications can be monitored. Some IoT-based devices analyze data remotely using cloud services, allowing researchers and agriculturalists to make better-informed decisions [9]. Researchers use various ML techniques to attain disease identification and pattern-based disease classification.

Additionally, the approaches Support Vector Machine (SVM), Principal Component Analysis (PCA), Convolution Neural Network (CNN), decision trees, back-propagation neural networks, and K-nearest neighbors' classification for disease classification were created by other researchers [10]. But although intrusion is getting increasingly complex and diversified, the straightforward ML and deep learning (DL) approach has numerous drawbacks. Better learning techniques are required, particularly for automatic intrusion feature extraction and analysis. In this paper, an approach based on IoT and DL called the modified Xception network is used for predicting disease and intruders for precision agriculture.

The key contribution of this proposed approach is listed below:
• Precision agriculture is achieved by developing a combined IoT and DL approach for predicting diseases and intruders to support farmers.
• Preprocessing techniques such as resizing, Median Modified Wiener Filter (MMWF), and Joint Histogram Equalization (JHE) are utilized for resizing an original image, noise removal, and contrast enhancement.
• Improved YOLO-V4-based segmentation is employed for converting an image to a group of regions of pixels, each constituted by a mask.
• A modified Xception network classification approach is used for the final classification, such as different types of intruders and various plant diseases.

The rest of the paper contains: Section 2 presents reviews related to smart agriculture. The proposed techniques and their framework are presented in Section 3. Results and discussion of the proposed method are given in Section 4. Finally, Section 5 contains the conclusion.

2. LITERATURE REVIEW

Agriculture is a key source of income and means of subsistence for any nation with a big population, such as China or India. Agriculture is hampered as a result of population shifts from rural to urban areas. To overcome this problem, many IoT and smart agriculture systems have been developed. Some of the existing techniques related to smart agriculture systems are given below:

An expert system that can simulate the judgment of a human expert regarding the illness and deliver warning signals to users ahead of the onset of the illness was suggested by Khattab et al. [11] using artificial intelligence and prediction algorithms. Field tests demonstrated that this strategy enhances agricultural goods with no chemical residues and high crop yields by reducing the number of chemical treatments. However, this approach has a poor precision value.

Ale et al. [12] have developed a densely connected convolutional network (DenseNet) based transfer learning system to identify plant diseases, and it anticipates running on edge servers with increased computational power. Utilizing a lightweight deep neural network (DNNs) approach that may be used on IoT devices with low resource availability. But gathering and producing well-labeled data takes a lot of work and time.

A decision support system (DSS), designed by Tripathy et al. [13], controls and oversees all of the operations of a smart agriculture system. This method also takes into account the various difficulties in raising greenhouse roses, and it was well adapted to the changing environment, redefining what sustainability means in the process. However, village sides weren't a good place for it.

Region proposal network (RPN) and Chan-Vese (CV) approaches were suggested by Guo et al. [14]. The RPN was used to identify and localize the leaves in a complicated environment. Then, using the CV method, images that have been segmented based on the output of the RPN algorithm contain the feature symptoms. This method yields an accuracy rate of 83.5%. However, this strategy had a slow training pace.

A low-cost farmland digital twin framework called AgriLoRa for smart agriculture was developed by Angin et al. [15]. AgriLoRa consists of a WSN deployed in the farmland and cloud servers that run artificial intelligence algorithms to locate clusters of weeds, sick plants, and nutritional deficiencies in plants. But in this case, security was not taken into account.

Kumar et al. [16] have suggested a Privacy-Enhanced Federated Learning (PEFL), a deep privacy encoding-based federated learning (FL) architecture that uses a long short-term memory autoencoder approach and perturbation-based encoding to achieve privacy. Then, using the encoded data, an FL-based gated recurrent unit neural network for intrusion detection was built. However, it generates iteration in advance.

A secure privacy-preserving framework (SP2F) was designed by Kumar et al. [17] for smart agriculture UAVs. An anomaly detection engine based on DL and a two-level privacy engine are the two core engines of this SP2F architecture. A sparse auto-encoder (SAE) is used to convert data into a new encoded format in order to thwart inference attacks. The anomaly detection engine employs stacked long short-term memory (SLSTM). This suggested strategy, however, had limited effectiveness.

Aburasain and Balobaid [18] introduced a hybrid DL-enabled intrusion detection method with particle swarm hyperparameter optimization (HDLID-PSHO) for IoT-based smart farming. The model includes preprocessing, feature selection, and a hybrid DL with transfer learning for intrusion detection. Particle swarm optimization (PSO) is employed for hyperparameter tuning. Experiments on Network-Internet of Things (ToN-IoT) and Network Intrusion Detection System-Knowledge Discovery in Databases (NSL-KDD) datasets demonstrate strong performance.

Zhukabayeva et al. [19] demonstrated a cybersecurity architecture for WSNs in smart grids, combining traffic analysis, node categorization, and ML. Using the Random Forest model, it effectively predicts traffic load and outperforms other models in intrusion detection across various attack types, which are tested on the Wireless Sensor Network Blackhole Flooding Selective Forwarding (WSNBFSF) dataset.

Maurya et al. [20] designed a meta-ensemble of lightweight Multi-layer Perceptron (MLP)-mixer and fast LSTM models for plant disease detection on low-powered IoT microcontrollers. It uses a two-level structure where predictions from the first-level models train a second-level

classifier to enhance accuracy. Tested on diverse plant datasets, the approach achieves high classification performance with low prediction time and significantly fewer parameters than CNNs and transformer-based models, making it efficient for resource-constrained environments.

Table 1 presents various research works on plant disease detection and intrusion detection in smart agriculture using IoT and AI techniques.

According to the literature, numerous systems are designed for IoT-based smart agriculture systems. Several significant challenges arise in the above-mentioned articles:

Low precision [11], collecting and creating well-label data requires considerable effort and time [12], not suitable for village agriculture [13], low training speed [14], security was considered [15], produce iteration ahead of time [16] and low ability [17], high computational cost [18], scalability issues [19], limited model stability [20].

Therefore, improved learning techniques are required, particularly for the automatic extraction and analysis of incursion features. Therefore, the proposed work focuses on IoT and a DL-based approach for disease prediction and intruder detection for precision agriculture.

Table 1. Overview of the literature survey

Author	Techniques	Dataset	Performances
Khattab et al. [11]	IoT-based monitoring system for plant diseases	-	-
Ale et al. [12]	DenseNet in plant diseases	-	Accuracy: 89.70%
Tripathy et al. [13]	DSS enables a smart greenhouse for sustainable agriculture	-	Accuracy: 91%
Guo et al. [14]	RBN for plant disease identification	Plant Photo Bank of China (PPBC)	Accuracy: 83.57%
Angin et al. [15]	AgriLoRa to detect plant diseases	-	Accuracy: 95%
Kumar et al. [16]	PEFL to detect intrusion detection	-	-
Kumar et al. [17]	SLSTM to detect anomalies in the agricultural	ToN-IoT dataset	Accuracy: 91.93%
Aburasain et al. [18]	HDLID-PSHO to find intrusion detection	ToN-IoT and NSL-KDD datasets	-
Zhukabayeva et al. [19]	Random Forest	WSNBFSF dataset	Precision: 99%
Maurya et al. [20]	Lightweight MLP-mixer and faster LSTM models	Corn, Maize, and TPP datasets	Accuracy: 94.27%

3. PROPOSED METHODOLOGY

For many established countries and emerging ones, farming and agriculture make up a significant component of the GDP. As a result, it is urgently necessary to adapt and improve current farming technologies. It will support the flourishing, sustainable development of people, flora, and fauna, in addition to assisting in addressing global crises like climate change and epidemics like drought. Better technology increases yield, which helps avoid conditions like famine and malnutrition. Crop yields could grow significantly through the automation of traditional methods. This study suggests a DL and Internet of Things strategy for automated and ongoing field monitoring.

The workflow of the proposed approach is illustrated in Figure 1. Sensor nodes are initially deployed in the agricultural area for gathering information regarding the land. The data gathered from sensors is utilized as input in this proposed approach. The acquired images, such as plant diseases and intruders, are preprocessed to convert raw data into meaningful data, which is required for better classification. Image resizing, MMWF, and JHE algorithms are used as a preprocessing technique in this proposed approach for resizing the original image, removing noise present in the input image, and contrast enhancement, respectively. After being preprocessed, photos are segmented using the YOLO-V4 approach, which breaks an image up into a number of regions of pixels that are each represented by a mask. These segmented images are then given as input for the Xception network classifier for the final prediction of disease and intruders. Two trained models are employed to train two different datasets, such as predicting different types of intruders and various types of plant diseases. The predicted data are further used for the purpose of monitoring agricultural areas. The built mobile app will also show the information pulled from the database. Farmers may manually oversee and control the pesticide sprayer process using a linked Android application. A buzzer that alerts to potential hazardous behaviours can also be manually activated.
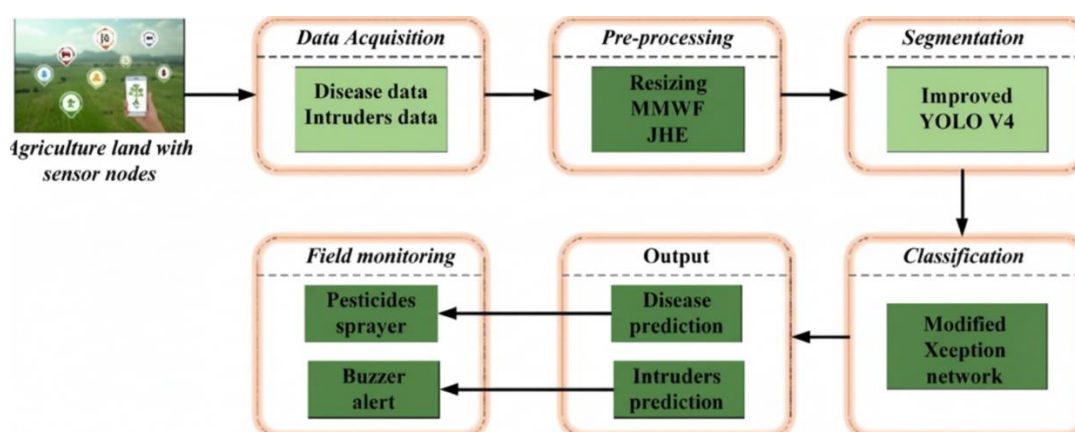

Figure 1. Proposed method workflow

3.1 Data collection

IoT sensors for smart agriculture can collect raw data. The core hardware for the suggested system is an ESP32 NodeMCU, one passive infrared (PIR) sensor positioned at the front of the field for intrusion detection within a 13-meter range, and one soil moisture sensor embedded near the crop roots per monitored segment. Sensor data, formatted as JSON, is collected every 30 seconds, with motion-triggered events transmitted in real time. The microcontroller, the system's brain, receives digital data from the infrared and soil moisture sensors and processes it. The output from these sensors will be used to activate or deactivate the buzzer or motor pump, depending on the given instruction. The soil moisture sensor is buried in the soil close to the crop to look for plant diseases. In front of the crop field, a PIR sensor is mounted to look for any intrusions.

3.2 Data analysis

Following the collection of all data, an agricultural dataset is produced and sent to AI algorithms for classification research.

3.2.1 Preprocessing

Image processing is the process of translating an image into a digital format and running particular operations on it to extract useful information from it. Image resizing, MMWF, and JHE are used as preprocessing techniques in this proposed method to resize the original image, removing noise and improving the contrast of the image.

3.2.2 MMWF

In damaged images, the noise distribution is reduced using the MMWF approach. By applying the median filter to the background area of a damaged image, this approach aims to enhance the image. Additionally, the Wiener filter is generally used in this method to maintain the edge signal. The MMWF method, which depends on the Wiener filter, decreases noise in the degraded image by changing the pixel values in the mask matrix to the median values [21]. In the Wiener filter formula, the median value $\tilde{\mu}$ is utilized rather than the average value μ. As a result, the MMWF is illustrated as follows:

$$b_{mmwf}(n, m) = \tilde{\mu} + \frac{\sigma^2 - v^2}{\sigma^2} \cdot (a(n, m) - \tilde{\mu}) \qquad (1)$$

where, $a(n, m)$ stands for each pixel in the region η, v^2 is the noise variance setting of the mask matrix for the application of the Wiener filter, and σ^2 is the variance of the Gaussian noise in the image.

With the MMWF approach, it is possible to improve the image quality of damaged images in the following ways: The edge signal is better retained as compared to the drop-off effect when the median and Wiener filter approaches are used. In conclusion, the MMWF technique can achieve a denoising effect significantly superior to traditional filters. Additionally, it can maintain the edge signal while removing the signal caused by ambient noise.

3.2.3 JHE

The standard image can be improved using the JHE method. While enhancing the contrast, the qualities of brightness preservation are applied. The proposed method applies equalization while taking spatial information into account. An average image is produced by computing the mean measurement for all pixels in a pixel's neighbourhood. The pixel pair population in the joint histogram is determined using both the pixel strength and its spatial data. The minimal subband component of the input picture (X) of size $M \times N$ in the wavelet domain is $A = \{k(i, j) | 1 \leq i \leq M, 1 \leq j \leq J\}$, as was already mentioned. The grey levels in the range of "0 to L-1" make up image A.

After that, A's image mean value (A_m) is calculated. Two sub-images (A_L and A_U) are created from image A based on this value.

$$A_L \in k(i, j) \big| k(i, j) \leq A_m, \quad \forall k(i, j) \in A \qquad (2)$$

$$A_U \in k(i, j) \big| k(i, j) > A_m, \quad \forall k(i, j) \in A \qquad (3)$$

The sub-images are made up of $A_U = \{A_m + 1, \dots, L - 1\}$ intensity values and $A_L = \{0, \dots, A_m\}$ intensity values. The following illustrates the connection among image A and the picture after it has been broken up.

That is, the sub-images consist of $A_L = \{0, \dots, A_m\}$ and $A_U = \{A_m + 1, \dots, L - 1\}$ intensity values. The relationship between image A and the decomposed image is represented as:

$$A = A_L + A_U \qquad (4)$$

Now, these sub-images are processed using the suggested method for contrast enhancement. In the image A_L, let $f_L(i, j)$ represent the intensity value of a pixel at coordinate (i, j), where, i = 1 to I, and n = 1 to N. Let the spatial information image created from A_L with intensity values $g_L(i, j)$ be represented by $(\widehat{A_L})$. A $w \times w$ averaging kernel is employed to calculate these values of intensity. With intensities ranging from [0, …, A_m], the images A_L and $\widehat{A_L}$ are both $I \times J$ in size. In the spatial information image, the intensity value $g_L(i, j)$ is calculated as follows:

$$g_L(i, j) = \left\lfloor \frac{1}{w \times w} \sum_{u=-k}^{k} \sum_{v=-k}^{k} f_L(i + u, j + v) \right\rfloor \qquad (5)$$

where, $k = \lfloor w/2 \rfloor$, the floor controller is indicated by $\lfloor . \rfloor$. w is normally an odd number. The intensity values $f_L(i, j) = x$ and $g_L(i, j) = y$ are extracted from the photos (A_L) and $(\widehat{A_L})$'s spatial information picture, respectively, for constructing the joint histogram [22].

3.3 Segmentation

The technique of segmenting an image involves separating it into several segments or sections based on the traits and attributes of the individual pixels. An improved YOLO-V4 approach is used as a segmentation technique in this proposed method to partition the image. Preprocessed leaf diseases and intruder images are given as input for this segmentation process.

3.3.1 Improved YOLO-V4

The YOLO-V4 model has been enhanced based on the YOLO-V3 design. Compared to the network structure of

YOLO-V3, the DarkNet53 architecture in YOLO-V3 has been modified to CSP DarkNet53, which is employed as the backbone network. The Head is the prediction portion of the model, and the Neck structure comprises Spatial Pyramid Pooling (SPP) and Path Aggregation Network (PAN) [23].

DarknetConv2D's winding machine changes from DrknetConv2D_BN_Leaky to DrknetConv2D_BN_Mish when the activation function is altered from LeakyReLU to Mish.

$$Mish = x \times \tanh\left(In(1 + e^x)\right) \tag{6}$$

where, x denotes the input of the activation function, tanh shows the hyperbolic tangent function. A modified resblock_body structure separates the residual network into two sections: one section stacking the residual system, and the other serves as a residual border. It goes straight to the end after a brief processing phase. This section avoids a large number of leftover structures, sometimes known as the CSPNet structure. SPP and PANet structures should be adopted.

Boundary dimensions and chances are produced for each class by the YOLO network, which converts the finding issue into a regression problem. The network identifies targets if the centre of the identified objects falls in an artificially specified area. The bounding box loss (L_{CIoU}), confidence loss $(L_{confidence})$, and classification loss (L_{class}) are among the YOLO-V4 loss functions used to train this network.

$$Loss = L_{CIoU} + L_{confidence} + L_{class} \tag{7}$$

$$CIOU = IOU - \frac{\rho^2\left(b, b^{gt}\right)}{c^2} - \alpha\upsilon \tag{8}$$

The Euclidean distance between the prediction frame's centre points is represented by $\rho^2(B, b^{gt})$, and c stands for the diagonal length of the smallest area necessary to encompass both the forecast frame and the actual frame:
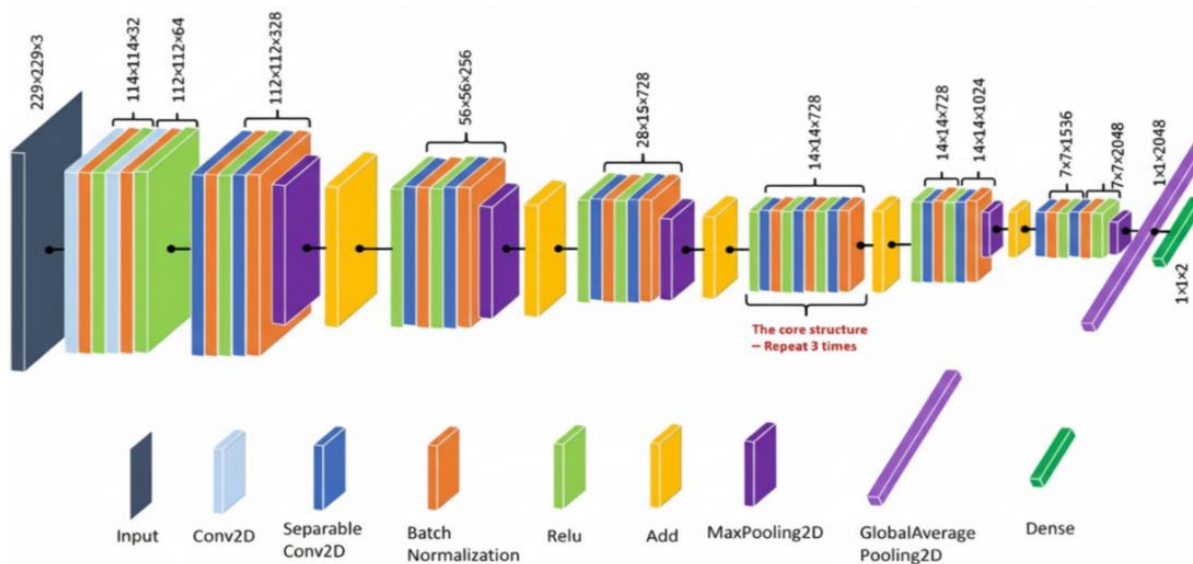
$$\alpha = \frac{\upsilon}{1 - IOU + \upsilon} \tag{9}$$

$$L_{CIOU} = 1 - IOU + \frac{\rho^2\left(b, b^{gt}\right)}{c^2} + \alpha\upsilon \tag{10}$$

The Intersection Over Union (IOU) standard, which also determines the accuracy of detecting targets, displays the intersection ratio between the projected area and the ground truth boundaries.

The YOLO-V4 is quite adept at detecting. In addition to reducing the number of variables and processing costs, DenseNet also lowers gradient disappearance and enhances feature transfer. As a result, the loss function is changed to Leaking ReLU in YOLO-V4, and DenseNet is used in place of the CSP DarkNet53 structure. These promote feature propagation, layer density, feature fusion, and reuse. The network topology becomes more complex by increasing the density between the layers, which is useful in disease prediction and intruder detection.

3.4 Classification

Segmented leaf disease and intruder images are given as input for this classification approach. An algorithm known as a classifier groups data into labelled informative categories. In this designed model, a modified Xception network is used for predicting types of plant disease and the type of intruders.

3.4.1 Modified Xception network

This strategy's attention component and Xception model were integrated. For feature extraction, the segmented data is fetched by the Xception model. The ImageNet dataset served as pre-training for the weights of the Xception model. Attention mechanisms, which are divided into two portions and are utilized to refine the features, are used. Both spatial and channel attention are required [24]. Separable convolution of the Xception network is shown in Figure 2. Figure 2 shows a CNN architecture for binary image classification. It starts with Conv2D and SeparableConv2D layers, followed by batch normalization and ReLU. The core structure, repeated three times, includes separable convolutions, residual connections, and max pooling.

Figure 2. Xception model with separable convolution

(1). Xception model

A model called Xception receives the segmented images. It is a radical interpretation of the 71-layer deep CNN model called Inception from Google. It uses a convolutional layer that is depth-wise separable. It begins with a 1–1 point–wise convolution, progresses to a 3–3 depthwise convolution, and then does a logistic regression.

(2). Channel attention

The proposed model has enhanced the deep CNN's capabilities by channeling attention to the Xception's output. Considered an attribute detector, every channel in a feature map concentrates channel attention on "what" is crucial, provided an input visual. Reducing the input feature map's spatial dimension will improve the computation of channel attention. Channel attention is, therefore, predominantly focused on the most important channels over the others. Increasing the value of the more pertinent channels is the simplest way to do this.

Combining global max pooling and average pooling for channel attention results in a multiplication of the input characteristics (F_x) obtained from the Xception framework. Following that, the spatial attention module receives the channelled refined features (F_c).

(3). Spatial attention

After passing via the spatial attention module, the channelled refined features (F_c) are subsequently output. It complements the channel's primary focus on determining "where" is an instructive section. The spatial attention map is constructed using Con2Dvlayer by concatenating max-pooling and average pooling along the channel axis in Eqs. (11) and (12). The calculation for spatial attention is displayed.

$$M_s(F) = \sigma\big(f^{3\times3}\big[Avg_{Pool(F)}; Max_{Pool(F)}\big]\big) \quad (11)$$

$$M_s(F) = \sigma\Big(f^{3\times3}\big(\big[F_{avg}^s; F_{max}^s\big]\big)\Big) \quad (12)$$

Here, the spatial attention map features are represented by $M_s(F)$, the sigmoid function is denoted by σ, and $f^{3\times3}$ is a convolution procedure with a filter size of 3×3. The spatially refined features (F_s) are produced by multiplying the spatial attention map features by the (F_c). After that, a Global Average Pooling (GAP) layer receives F_s. The classification unit and a sigmoid activation function are then utilized for this purpose.

3.5 Field monitoring

The data is saved on the Firebase database system and in IoT apps like cloud storage. Information is delivered directly to the farmer's mobile and other devices for monitoring after being stored at a specific location. The farmer will be notified if an intruder enters the crop field or if a disease spreads there.
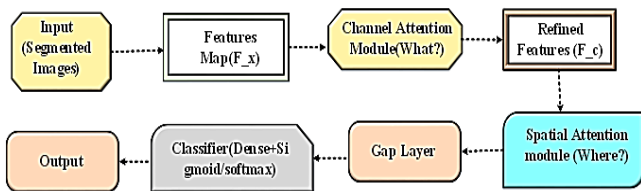

Figure 3. The dual attention mechanism of the model

Figure 3 shows an image classification model using dual attention mechanisms. It extracts features from segmented images, applies channel attention to focus on important features ("what"), and spatial attention to locate key regions ("where"). The refined features are pooled and passed to a classifier for the final output, enhancing accuracy by improving feature focus.

3.5.1 Disease detection and pesticide sprayer

This technique can identify three different leaf diseases. Apple scab disease develops when little soil moisture and the daytime temperature is chilly. Applying silicon fertilizers for 60 to 100 seconds will help to control this problem. Black rot is brought on by strong winds and humidity levels that are above 70%. This disease can be controlled by giving plants a balanced supply of nutrients, particularly nitrogen, for 1 minute. Toxic elements in the soil cause cedar apple rust, which can be controlled by adding the appropriate fertilizers to the soil and applying Neem cake at 150 kg/ha.

The microprocessor and the motors are connected, and the motors will supply pesticides to the plant based on the disease. A healthy leaf image is initially recorded as input, and the image that is obtained is compared to the original leaf. After this, the type of disease is identified, and the specific motor will turn on and off. The soil moisture sensor controls the motor ON and OFF.

Algorithm 1. Pseudocode

A: Disease dataset, B: Intruder dataset
 # Preprocessing
{
M = MMWF
$$b_{mmwf}(n,m) = \tilde{\mu} + \frac{\sigma^2 - v^2}{\sigma^2} \cdot (a(n,m) - \tilde{\mu})$$
$M_1 = MMWF(A)$
$M_2 = MMWF(B)$
J = JHE
$$g_L(i,j) = \left|\frac{1}{w \times w}\sum_{u=-k}^{k}\sum_{v=-k}^{k} f_L(i+u, j+v)\right|$$
$J_1 = JH(M_1)$
$J_2 = JH(M_2)$
Preprocessed disease and intruder images
}
#Segmentation
{
Y=YOLO-V4
$$L_{CIOU} = 1 - IOU + \frac{\rho^2(b,b^{gt})}{c^2} + \alpha v$$
$Y_1 = YOLO - V4(J_1)$
$Y_2 = YOLO - V4(J_2)$
Segmented disease and intruder images
}
#Classification
{
X=Xception-network
$X_1 = Xception(Y_1)$
$X_2 = Xception(Y_2)$
Disease and intruder prediction
}
O: Different types of diseases and intruders are classified

3.5.2 Intrusion detection and buzzer system

In order to identify any intrusions in the crop field, this research uses PIR sensors. The PIR is capable of identifying living objects like people and animals. The sensor works by determining an object's wavelength and temperature. The

sensor will provide information to the buzzer and camera module if an object or obstacle it detects satisfies the necessary conditions, such as the presence of people or animals. The camera will take a picture of the present scene and send real-time data on the picture to the cloud.

Therefore, farmers will receive a smartphone notification about the intruder. Farmers may always keep an eye on their crops using these techniques to see if there are any trespassers who could steal, hurt, or destroy their plants. When an obstacle contains a wavelength between 1 and 13 m, respectively, the PIR sensor detects the intrusion. Any barrier discovered within this range is considered an intrusion since the PIR can identify the infrared rays generated by the thing and fill up any gaps between the detector and the crop field. About 6 seconds pass before the buzzer turns on. Algorithm 1 provides the overall pseudocode for the proposed system.

4. RESULTS AND DISCUSSIONS

Improved production is necessary to meet the growing demand for food across many industries, especially agriculture. There will be instances where supply and demand are not balanced, though. The management and upkeep of resources like capital and labour continue to be challenging for increasing agricultural output. Intelligent farming is a better option for boosting food production, resource management, and labour. This approach develops a smart agriculture system based on IoT and DL. The proposed modified Xception-based algorithm is tested on Python 3.8 with CPU: Intel Core i5, GPU: Nvidia GeForce GTX 1650, RAM: 16 GB. Simulation parameters used for the modified Xception network are listed in Table 2.

The installation of various sensors in the agricultural field is the first step of this proposed approach. Sensor node deployment for the proposed approach is illustrated in Figure 4. The assembled nodes capture information regarding soil moisture, disease, and intruders. The collected information is either stored locally in the nearby fog node or transmitted to the cloud server based on IoT.

Table 2. Simulation parameter for modified Xception

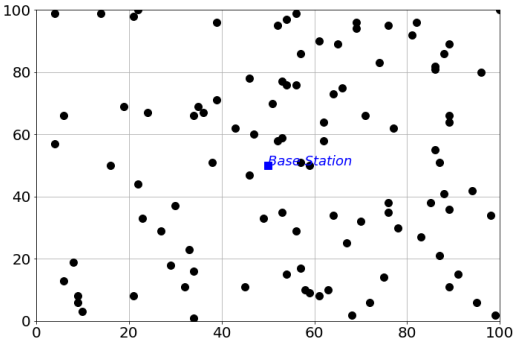| Parameters	Value
Optimizer	adam
Loss	categorical_crossentropy
Activation	softmax

Figure 4. Sensor node deployment

4.1 Disease prediction

The augmented new plant diseases dataset, comprising approximately 87,000 RGB images across 38 classes and totaling 1.43 GB in size, was used for plant disease classification. The dataset was divided into training (80%), validation (20%), and a separate test set of 33 images, with annotations based on the original dataset. For evaluation, 167 test images were selected, including 20 samples of Apple Scab, 21 of Black Rot, 26 of Cedar Apple Rust, and 100 healthy leaf images. These images were processed through a pipeline consisting of MMWF noise reduction, JHE contrast enhancement, YOLO-V4 segmentation, and classification using a modified Xception network. Out of 180 predictions made, all samples, 20 Apple Scab, 21 Black Rot, 26 Cedar Apple Rust, and 100 healthy leaves were correctly classified. Crop images [25] acquired from sensor nodes are used to predict the presence of disease in agricultural land. The acquired leaf images are initially preprocessed and then segmented to partition the image. Finally, it is fed into a modified Xception network for classification purposes.
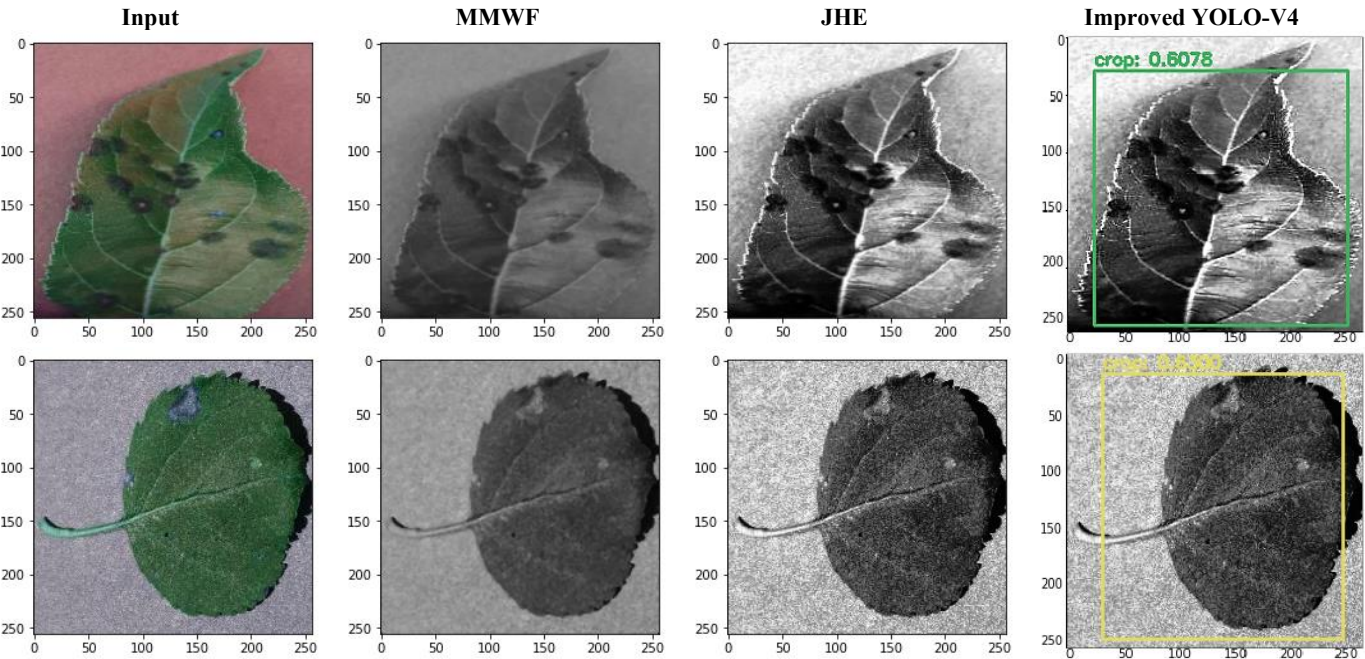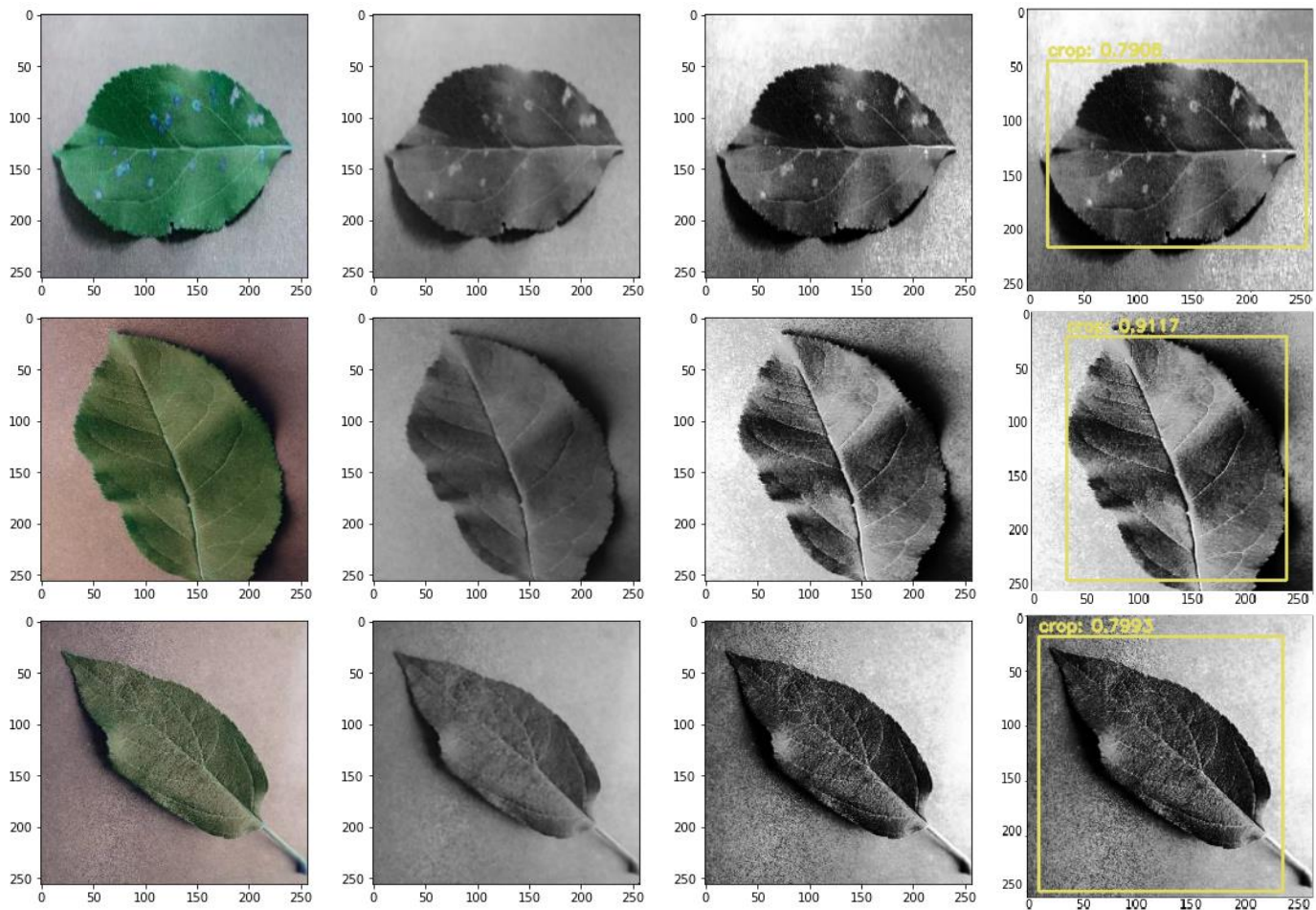
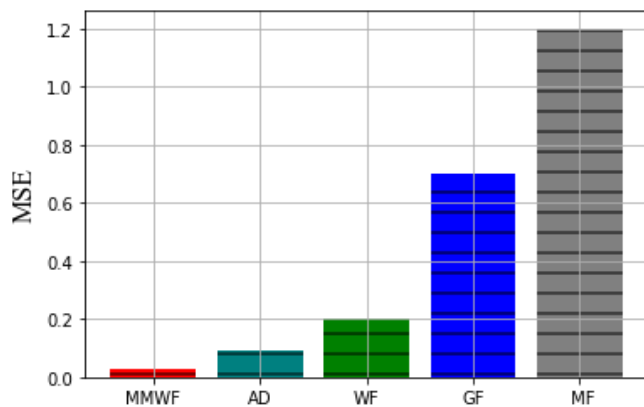**Figure 5.** Sample of preprocessing and segmentation process outputs

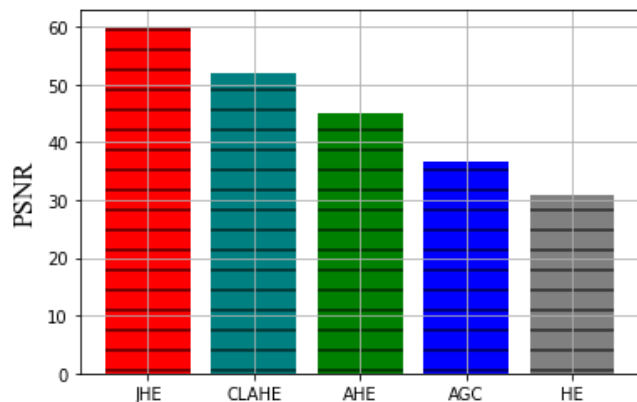**Figure 6.** Comparison of noise removal algorithms

Figure 7. Contrast enhancement algorithm analysis

Figure 5 shows the sample outputs of the preprocessing and segmentation process of this proposed approach. The acquired crop image is given as input here. The noise in the input image is removed using MMWF, and JHE is applied to boost the brightness. After that, this preprocessed output is given to a segmentation technique called improved YOLO-V4 for image portioning. The segmented output is fed into a modified Xception network to predict various diseases.

Examination of various noise removal algorithms is depicted in Figure 6. This graph explains various noise removal algorithms' mean square error (MSE) values. MSE produced by MMWF, anisotropic diffusion (AD), Wiener filter (WF), Gaussian filter (GF), and median filter (MF) is 0.02%, 01%, 0.2%, 0.7%, and 1.2%, respectively. The proposed MMWF has a lower MSE than current approaches. Figure 7 shows the contrast enhancement techniques analysis in terms of peak signal-to-noise ratio (PSNR). JHE, contrast-limited adaptive histogram equalization (CLAHE), AHE, and HE have PSNR values of 59%, 52%, 45%, 37%, and 31%, respectively. JHE has a higher PSNR performance than other algorithms.

The confusion matrix plot for rice leaf disease detection is depicted in Figure 8. Four diseases, such as apple scab, back rot, cedar apple rust, and healthy, are considered four classes. This plot shows that fewer samples from each class are wrongly predicted, so the error is minimal, and better accuracy is achieved using the proposed modified Xception network.

This proposed algorithm's performance is analyzed and compared with current algorithms such as Inception-V2, Renext50, ResNet50, and VGG16. The performance metrics plot used for comparison is given in Figure 9.
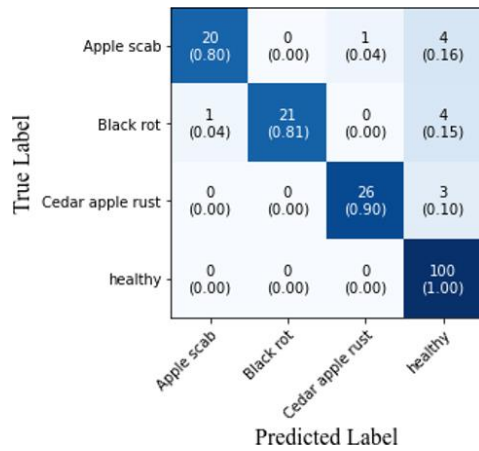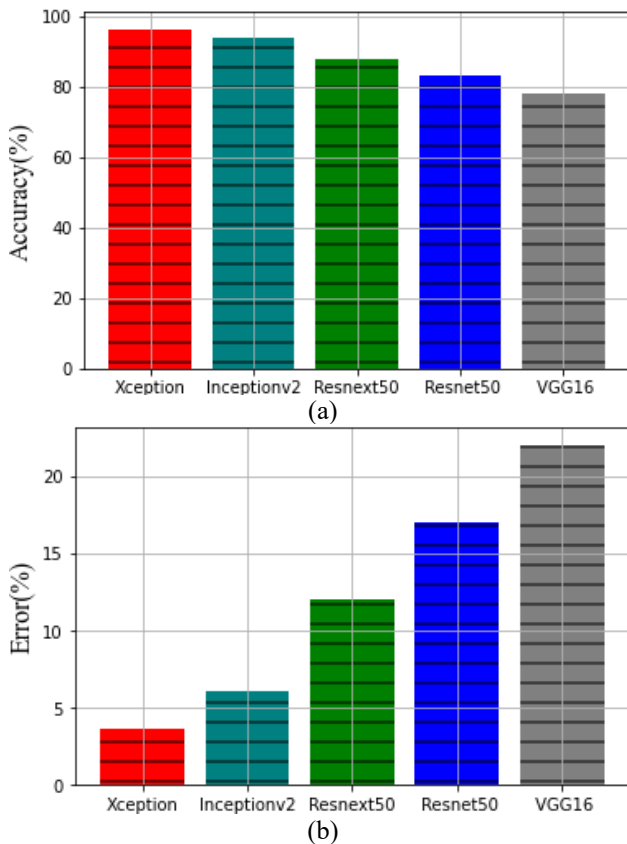
Figure 8. Confusion matrix plot

Figure 9. Evaluation of (a) Accuracy (b) Error

Accuracy analysis for various disease predictions is illustrated in Figure 9(a). Xception, InceptionV2, ResNeXt50, ResNet50, and VGG16 are the classifiers that have an accuracy value of 96%, 94%, 88%, 83%, and 78%, respectively. Similarly, Figure 9(b) explains the error metric examination of proposed and existing classifiers. 4%, 6%, 12%, 17%, and 22% are the errors produced by the Xception, InceptionV2, ResNeXt50, ResNet50, and VGG16, accordingly. The proposed modified Xception network performs better than other algorithms.

Analysis of the F1-score statistic is depicted in Figure 10(a). Xception, InceptionV2, ResNeXt50, ResNet50 and VGG16 classifiers which has a F1-score value of 91.3%, 86%, 82%, 76% and 73% respectively. Likewise, Figure 10(b) represents a precision metric examination of various algorithms. This plot explains the precision rate achieved by the proposed and existing classifiers. Xception, InceptionV2, ResNeXt50,

ResNet50, and VGG16 algorithm has a 95.4%, 92%, 89%, 82% and 80% precision value. This indicates that the proposed modified Xception network has a higher F1-score and precision value than current approaches.
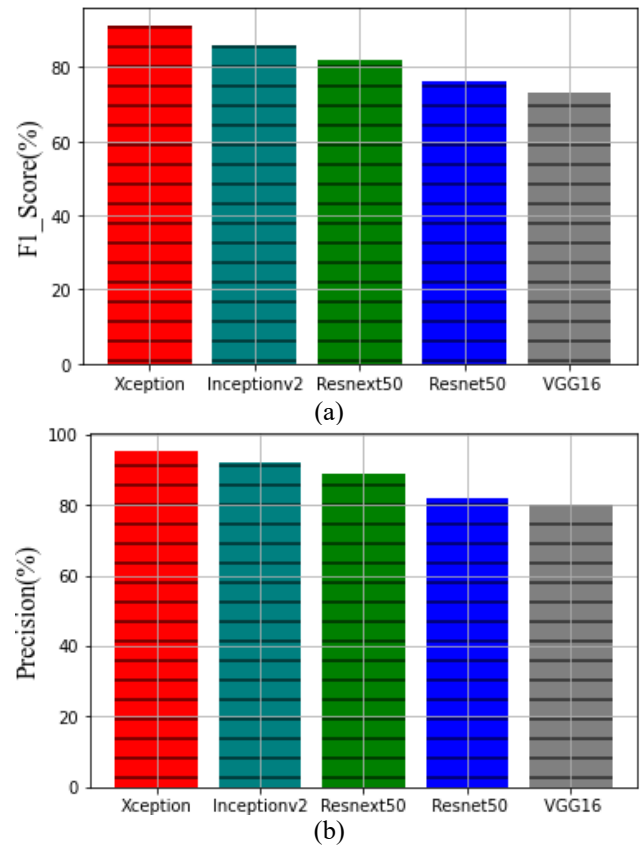

Figure 10. (a) F1-score (b) Precision metric comparison
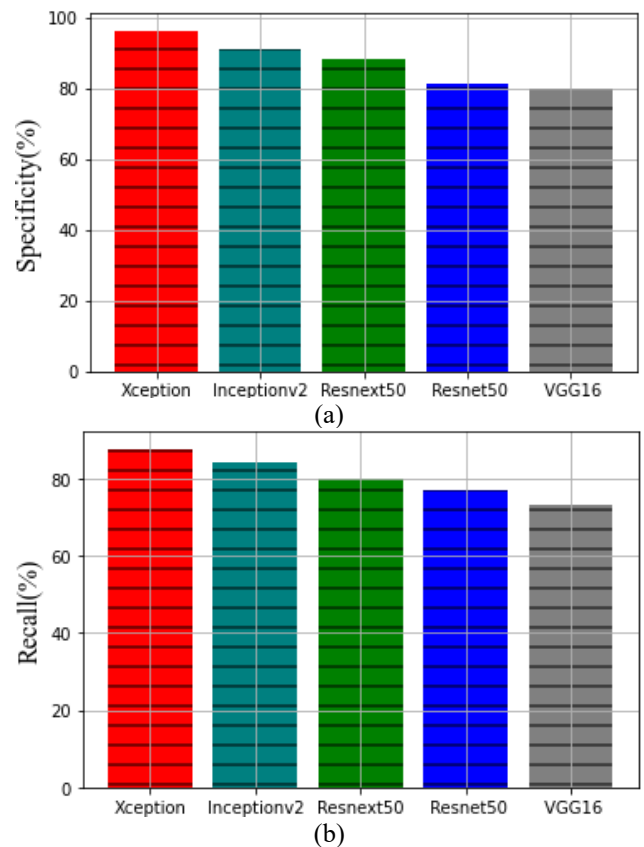

Figure 11. Analysis of (a) Specificity (b) Recall

Figure 11(a) illustrates the specificity evaluation. The Xception network has a 96.2% specificity value. It is higher than existing techniques such as Xception, InceptionV2, ResNeXt50, ResNet50, and VGG16. Because it has a specificity value of 91%, 88%, 81%, and 80%, respectively, the recall metric comparison of various classifiers is depicted in Figure 11(b). 87.6%, 84%, 80%, 77%, and 73% are the recall values produced by Xception, InceptionV2, ResNeXt50, ResNet50, and VGG16 classifiers. It explains that the proposed modified Xception network performs better than current algorithms for predicting various diseases.

4.2 Intruder detection

Dataset 2 contains the animal detection by You Only Look Once Common Objects in Context (YOLO COCO) model, which refers to using a YOLO object detection algorithm, which is known for its speed and accuracy, to detect animals in images. The model is trained on the COCO dataset, which includes over 80 object categories, including animals such as cats, dogs, and horses. This setup allows for real-time identification and localization of animals within visual data. It uses tools such as OpenCV (cv2) for image processing, Matplotlib for visualization, and NumPy for handling numerical data, all implemented in Python [26]. Datasets like animals10 may be used for testing or fine-tuning the detection model, while yolo-coco-data refers to the pretrained weights and configuration files from the COCO-trained YOLO model. For the intruder detection task, test frames were drawn from the COCO dataset, comprising a total of 1,469 labeled frames. The class-wise breakdown includes Elephant (686 frames), Bear (257), Cow (352), Dingo (171), and Human (403). Intruder images were detected from the PIR sensor used to detect intruders in the agricultural area. The acquired images are preprocessed for enhancement. Then it is given to the segmentation process for portioning the image. Finally, these segmented images are fed into a modified Xception network to predict different intruders.
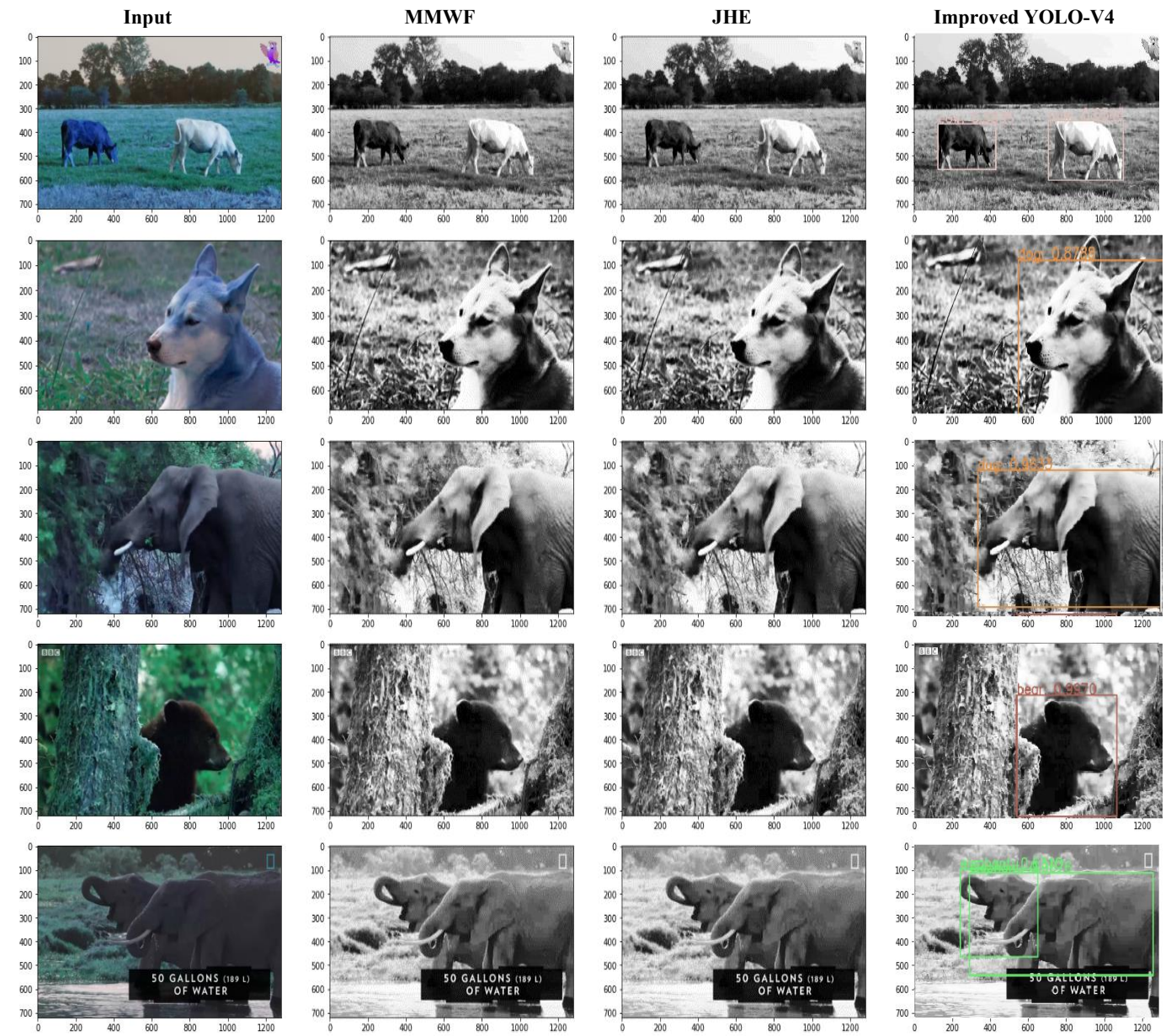

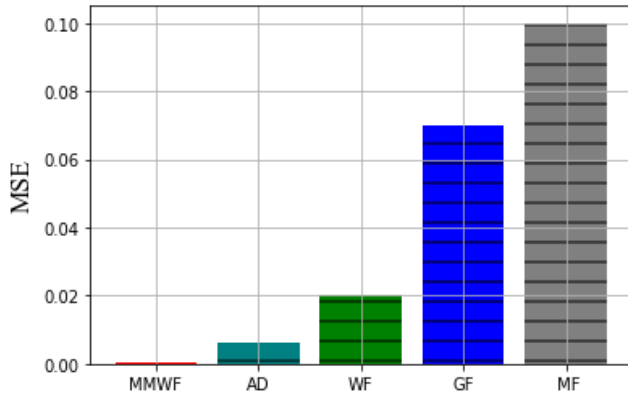**Figure 12.** Sample of preprocessing and segmentation process outputs

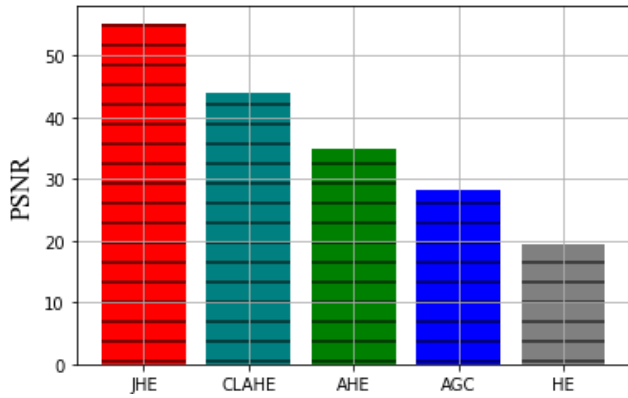**Figure 13.** Analysis of various noise removal approaches

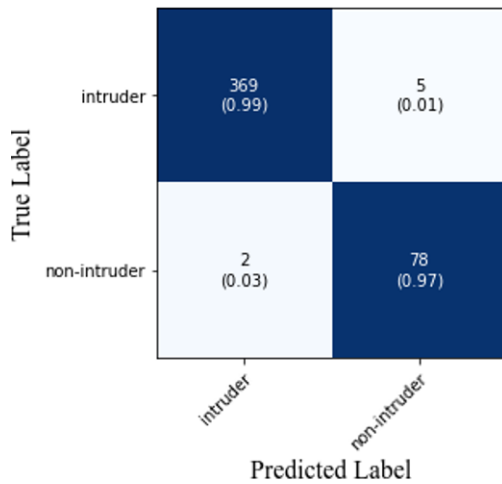**Figure 14.** Examination of different contrast enhancement algorithms

Figure 15. Confusion matrix plot

in terms of PSNR is illustrated in Figure 14. 55%, 43%, 35%, 28%, and 19% are the PSNR values obtained by JHE, CLAHE, AHE, AGC, and HE.

The confusion matrix plot for intruder detection is shown in Figure 15. Intruders and non-intruders are taken as two different classes. It plots a table of all a classifier's predicted and actual values and shows that only a small number of samples are misclassified.

The performance of the modified Xception network is compared to that of other already-in-use methods, including InceptionV2, ResNeXt50, ResNet50, and VGG16. The performance metrics used to compare this proposed and existing approaches are shown in Figure 16.

Accuracy evaluation is illustrated in Figure 16(a). Xception, InceptionV2, ResNeXt50, ResNet50, and VGG16 have an accuracy value of 98%, 95%, 89%, 86%, and 82% respectively, for intruder detection. Similarly, Figure 16(b) depicts the error metrics evaluation. Error rates produced by Xception, InceptionV2, ResNeXt50, ResNet50, and VGG16 classifiers are 2%, 5%, 11%, 14%, and 18%, accordingly.
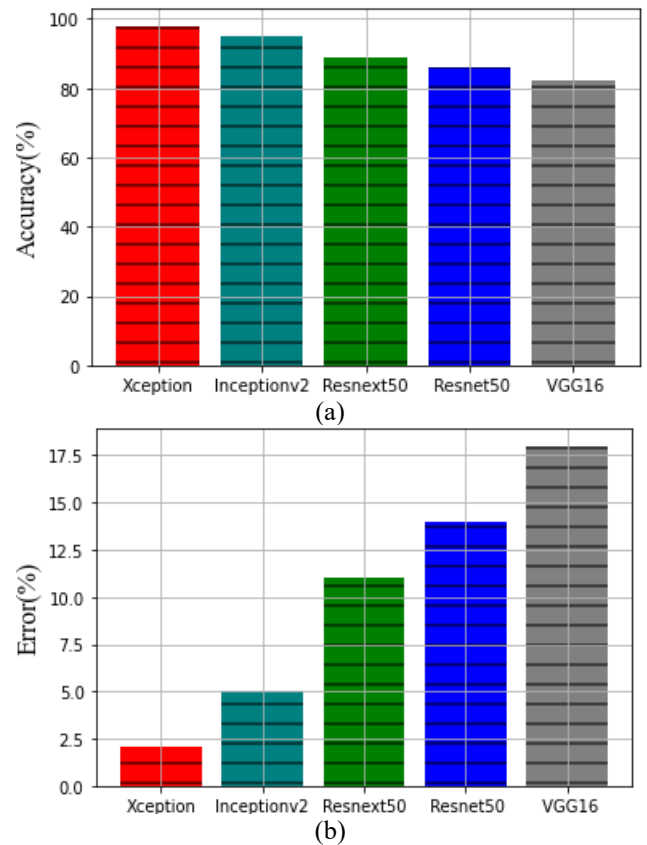

(a)

(b)

Figure 16. Comparison of (a) Accuracy (b) Error

(a)

Figure 12 depicts the output samples for the preprocessing and segmentation process. Initially, MMWF is done in the input image to remove noise present in the image. Then JHE is utilized to improve the contrast of an image. This preprocessed image is then fed into improved YOLO-V4 to segment the intruder's image. Finally, a modified Xception network is used to predict different types of intruders in the segmented images.

Figure 13 shows the proposed and existing noise removal techniques. MMWF, AD, WF, GF, and MF are the noise removal algorithms that have an MSE of 0.001%, 0.008%, 0.02%, 0.07%, and 0.1%, respectively. This explains that the proposed MMWF has a lower MSE than current techniques. The performance of various contrast enhancement algorithms
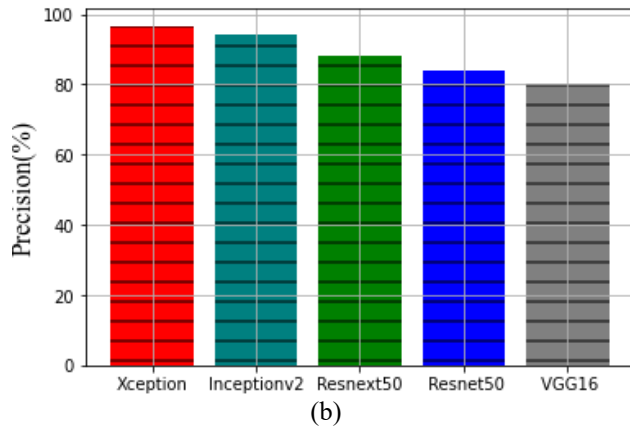
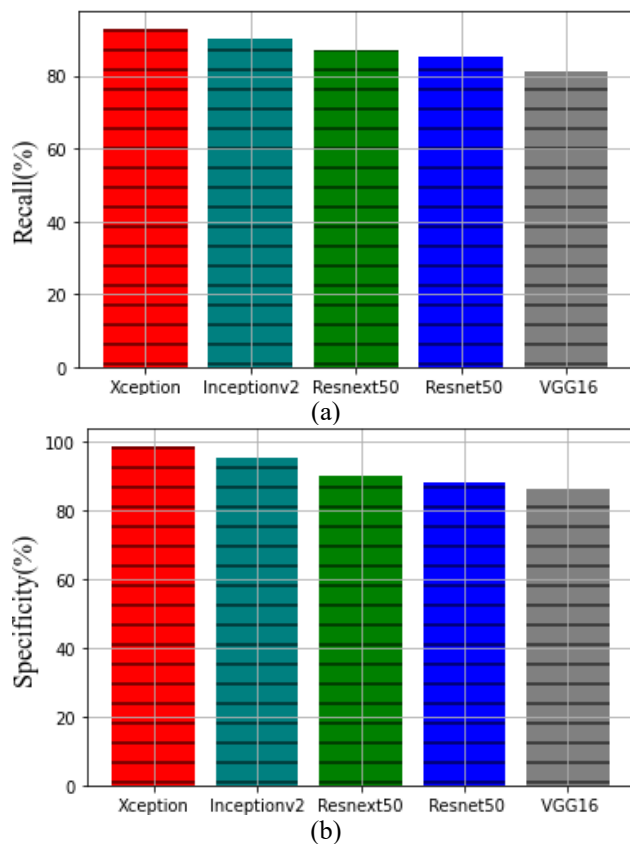**Figure 17.** (a) F1-score (b) Precision metrics analysis

Figure 18. Examination of (a) Recall (b) Specificity

Evaluation of F1-score metrics for different classifiers is shown in Figure 17(a). Xception, InceptionV2, ResNeXt50, ResNet50, and VGG16 classifiers are used for this comparison which has an F1-score value of 94.7%, 91%, 88%, 86%, and 82% respectively. Similarly, Figure 17(b) depicts the comparison of proposed and current algorithms in terms of precision statistics. 96%, 94%, 88%, 84%, and 80% are the precision value of Xception, InceptionV2, ResNeXt50, ResNet50, and VGG16. This indicates that the proposed Xception network is better than other existing techniques.

Figure 18(a) explains the recall values for proposed and existing approaches. Different classifiers, such as Xception, InceptionV2, ResNeXt50, ResNet50, and VGG16, which have a recall of 93%, 90%, 87%, 85%, and 81%, are used for this comparison. Likewise, the specificity statistic evaluation is depicted in Figure 18(b). The specificity rate of proposed and existing classifiers named Xception, InceptionV2, ResNeXt50, ResNet50, and VGG16 is 98%, 95%, 90%, 88%

and 86%. This indicates that the proposed Xception network performs better than the current approaches.

The performance comparison Table 3 highlights the effectiveness of three models, such as Proposed, JAMVO-RNN + DBN, and ResNet-50, based on sensitivity, precision, and F1-score.

Table 3. State of the art for the proposed and existing algorithms

Performances	Proposed	JAMVO-RNN + DBN [27]	ResNet-50 [28]
Sensitivity	87.6%	0.94	99.05%
Precision	91.3%	0.98	98.96%
F1-score	95.4%	0.96	98.98%

5. CONCLUSION

The principal provider of nourishment and raw materials for agriculture is recognized as the cornerstone of existence for the human species. The evolution of the nation's economic situation depends on the agricultural sector. Sadly, a lot of farmers still rely on old-fashioned farming techniques, which reduces food yields. The production grew in those situations where automation was applied, and workers were substituted by mechanical equipment. Therefore, in order to increase yield, the agricultural industry needs to use current science and technology. A smart agriculture system based on IoT and DL was proposed in this paper. Sensors were initially deployed in the agricultural land to collect information. Plant diseases and intruder images acquired from sensors were utilized as input. MMWF and JHE algorithms were employed as pre-processing techniques for removing noise and enhancing the contrast of an original image. Pre-processed images were then segmented using improved YOLO-V4 for dividing a picture into a number of pixel-rich sections. These segmented images were fed into a modified Xception network to predict diseases and intruders. Two trained models were employed to train two datasets, such as various types of intruders and plant diseases. These data were further used for monitoring purposes. In addition, the built mobile app will show the database-retrieved data. Farmers may monitor and manually control the pesticide sprayer operation using an associated Android device. A buzzer that alerts to potential hazardous behaviours can also be manually activated. The graphical representation shows that the proposed approach attains 96% and 98% accuracy for the disease and intruder detection process compared to existing models such as Inceptionv2, ResNet150, ResNet 50, and VGG16. Thus, this proposed approach is the best choice for precision agriculture. However, the system faces certain limitations, such as a relatively small dataset, which may limit the model's generalizability across different environmental conditions and crop types. Additionally, the potential latency in real-time applications in rural connectivity constraints was identified challenge. To address these issues and extend system capabilities, the future will focus on drone integration for automated surveillance and expanding the dataset in crop varieties.

REFERENCES

[1] Chehri, A., Chaibi, H., Saadane, R., Hakem, N., Wahbi, M. (2020). A framework of optimizing the deployment

of IoT for precision agriculture industry. Procedia Computer Science, 176: 2414-2422. https://doi.org/10.1016/j.procs.2020.09.312

[2] Mekonnen, Y., Namuduri, S., Burton, L., Sarwat, A., Bhansali, S. (2019). Machine learning techniques in wireless sensor network-based precision agriculture. Journal of the Electrochemical Society, 167(3): 037522. https://doi.org/10.1149/2.0222003JES

[3] Marcu, I.M., Suciu, G., Balaceanu, C.M., Banaru, A. (2019). IoT based system for smart agriculture. In 2019 11th International Conference on Electronics, Computers and Artificial Intelligence (ECAI), Pitesti, Romania, pp. 1-4. https://doi.org/10.1109/ECAI46879.2019.9041952

[4] Akhter, R., Sofi, S.A. (2022). Precision agriculture using IoT data analytics and machine learning. Journal of King Saud University Computer and Information Sciences, 34(8): 5602-5618. https://doi.org/10.1016/j.jksuci.2021.05.013

[5] Wang, P., Hafshejani, B.A., Wang, D. (2021). An improved multilayer perceptron approach for detecting sugarcane yield production in IoT based smart agriculture. Microprocessors and Microsystems, 82: 103822. https://doi.org/10.1016/j.micpro.2021.103822

[6] Yang, J., Sharma, A., Kumar, R. (2021). IoT-based framework for smart agriculture. International Journal of Agricultural and Environmental Information Systems, 12(2): 1-14. https://doi.org/10.4018/IJAEIS.20210401.oa1

[7] Boursianis, A.D., Papadopoulou, M.S., Gotsis, A., Wan, S., Sarigiannidis, P., Nikolaidis, S., Goudos, S.K. (2020). Smart irrigation system for precision agriculture-The AREThOU5A IoT platform. IEEE Sensors Journal, 21(16): 17539-17547. https://doi.org/10.1109/JSEN.2020.3033526

[8] Symeonaki, E., Arvanitis, K., Piromalis, D. (2020). A context-aware middleware cloud approach for integrating precision farming facilities into the IoT toward agriculture 4.0. Applied Sciences, 10(3): 813. https://doi.org/10.3390/app10030813

[9] Sanjeevi, P., Prasanna, S., Siva Kumar, B., Gunasekaran, G., Alagiri, I., Vijay Anand, R. (2020). Precision agriculture and farming using Internet of Things based on wireless sensor network. Transactions on Emerging Telecommunications Technologies, 31(12): e3978. https://doi.org/10.1002/ett.3978

[10] Naresh, M., Munaswamy, P. (2019). Smart agriculture system using IoT technology. International Journal of Recent Technology and Engineering, 7(5): 98-102.

[11] Khattab, A., Habib, S.E., Ismail, H., Zayan, S., Fahmy, Y., Khairy, M.M. (2019). An IoT-based cognitive monitoring system for early plant disease forecast. Computers and Electronics in Agriculture, 166: 105028. https://doi.org/10.1016/j.compag.2019.105028

[12] Ale, L., Sheta, A., Li, L., Wang, Y., Zhang, N. (2019). Deep learning-based plant disease detection for smart agriculture. In 2019 IEEE GlobeCom Workshops (GC Wkshps), Waikoloa, HI, USA, pp. 1-6. https://doi.org/10.1109/GCWkshps45667.2019.9024439

[13] Tripathy, P.K., Tripathy, A.K., Agarwal, A., Mohanty, S.P. (2021). MyGreen: An IoT-enabled smart greenhouse for sustainable agriculture. IEEE Consumer Electronics Magazine, 10(4): 57-62. https://doi.org/10.1109/MCE.2021.3055930

[14] Guo, Y., Zhang, J., Yin, C., Hu, X., Zou, Y., Xue, Z., Wang, W. (2020). Plant disease identification based on deep learning algorithm in smart farming. Discrete Dynamics in Nature and Society, 2020(1): 2479172. https://doi.org/10.1155/2020/2479172

[15] Angin, P., Anisi, M.H., Göksel, F., Gürsoy, C., Büyükgülcü, A. (2020). AgriLoRa: A digital twin framework for smart agriculture. Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, 11(4): 77-96. https://doi.org/10.22667/JOWUA.2020.12.31.077

[16] Kumar, P., Gupta, G.P., Tripathi, R.A. (2021). PEFL: Deep privacy-encoding-based federated learning framework for smart agriculture. IEEE Micro, 42(1): 33-40. https://doi.org/10.1109/MM.2021.3112476

[17] Kumar, R., Kumar, P., Tripathi, R., Gupta, G.P., Gadekallu, T.R., Srivastava, G.B. (2021). SP2F: A secured privacy-preserving framework for smart agricultural unmanned aerial vehicles. Computer Networks, 187: 107819. https://doi.org/10.1016/j.comnet.2021.107819

[18] Aburasain, R.Y., Balobaid, A. (2024). Hybrid deep learning with optimized hyperparameters based intrusion detection in Internet of Things for smart farming. In 2024 International Symposium on Networks, Computers and Communications (ISNCC), USA, pp. 1-8. https://doi.org/10.1109/ISNCC62547.2024.10758946

[19] Zhukabayeva, T., Pervez, A., Mardenov, Y., Othman, M., Karabayev, N., Ahmad, Z. (2024). A traffic analysis and node categorization-aware machine learning-integrated framework for cybersecurity intrusion detection and prevention of WSNs in smart grids. IEEE Access, 12: 91715-91733. https://doi.org/10.1109/ACCESS.2024.3422077

[20] Maurya, R., Mahapatra, S., Rajput, L. (2024). A lightweight meta-ensemble approach for plant disease detection suitable for IoT-based environments. IEEE Access, 12: 28096-28108. https://doi.org/10.1109/ACCESS.2024.3367443

[21] Park, C.R., Kang, S.H., Lee, Y. (2020). Median modified wiener filter for improving the image quality of gamma camera images. Nuclear Engineering and Technology, 52(10): 2328-2333. https://doi.org/10.1016/j.net.2020.03.022

[22] Agrawal, S., Panda, R., Mishro, P.K., Abraham, A. (2022). A novel joint histogram equalization-based image contrast enhancement. Journal of King Saud University-Computer and Information Sciences, 34(4): 1172-1182. https://doi.org/10.1016/j.jksuci.2019.05.010

[23] Gai, R., Chen, N., Yuan, H. (2023). A detection algorithm for cherry fruits based on the improved YOLO-v4 model. Neural Computing and Applications, 35(19): 13895-13906. https://doi.org/10.1007/s00521-021-06029-z

[24] Upasana, C., Tewari, A.S., Singh, J.P. (2023). An attention-based pneumothorax classification using modified Xception model. Procedia Computer Science, 218: 74-82. https://doi.org/10.1016/j.procs.2022.12.403

[25] New Plant Diseases Dataset. https://www.kaggle.com/datasets/vipoooool/new-plant-diseases-dataset.

[26] Animal Detection by YOLO COCO Model. https://www.kaggle.com/code/stpeteishii/animal-detection-by-yolo-coco-model.

[27] Ampavathi, A., Saradhi, T.V. (2021). Multi disease-

prediction framework using hybrid deep learning: An optimal prediction model. Computer Methods in Biomechanics and Biomedical Engineering, 24(10): 1146-1168.
https://doi.org/10.1080/10255842.2020.1869726

[28] Islam, M.M., Adil, M.A.A., Talukder, M.A., Ahamed, M.K.U., Uddin, M.A., Hasan, M.K., Debnath, S.K. (2023). DeepCrop: Deep learning-based crop disease prediction with web application. Journal of Agriculture and Food Research, 14: 100764. https://doi.org/10.1016/j.jafr.2023.100764