

Received March 21, 2022, accepted March 31, 2022, date of publication April 13, 2022, date of current version April 22, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3167061

# Learning Aided System for Agriculture Monitoring Designed Using Image Processing and IoT-CNN

KANDARPA KUMAR SARMA<sup>1</sup>, (Senior Member, IEEE), KUNAL KINGKAR DAS<sup>1</sup>,  
VIKASH MISHRA<sup>1</sup>, SAMADRITA BHUIYA<sup>1</sup>, AND DMITRII KAPLUN<sup>2,3</sup>, (Member, IEEE)

<sup>1</sup>Department of Electronics and Communication Engineering, Gauhati University, Guwahati, Assam 781014, India

<sup>2</sup>Department of Automation and Control Processes, Saint Petersburg Electrotechnical University "LETI," 197022 Saint Petersburg, Russia

<sup>3</sup>Faculty of Information Technologies, Kazakh-British Technical University, Almaty 050000, Kazakhstan

Corresponding author: Dmitrii Kaplun (dikaplun@etu.ru)

This work was supported by the Ministry of Science and Higher Education of the Russian Federation under Grant FSEE-2020-0002.

**ABSTRACT** The Internet of Things (IoT) and artificial intelligence (AI) based methods for monitoring, control, and decision support are combined to design of a smart agriculture assistance system. The proposed system has a sensor pack that provides continuous data capture of temperature records, air and soil moisture and a camera for obtaining near-infrared (NIR) images of the plant leaves for use with an AI decision support system. We identify twelve types of vegetation for the study, out of which five disease classes of the tomato leaves are categorized using a Convolutional Neural Network (CNN). The work also includes experiments conducted with multiple clustering-based segmentation methods and some features namely Gray level co-occurrence matrix (GLCM), Local binary pattern (LBP), Local Binary Gray Level Co-occurrence Matrix (LBGLCM), Gray Level Run Length Matrix (GLRLM), and Segmentation-based Fractal Texture Analysis (SFTA). Out of several AI tools, CNN proves to be effective in providing automated decision support for classifying the plant leaf disease types through a cloud server that can be accessed using an app. Extensive on-field trials show that the system (VGG16 CNN, GLCM and a fuzzy based clustering) is effective in hot and humid conditions and proves to be reliable in identifying disease classes of certain vegetable types, certain usable vegetation cover of farmland and regulation of watering mechanism of crops.

**INDEX TERMS** Artificial intelligence, near-infrared images, CNN, image processing, leaf disease, smart agriculture.

## I. INTRODUCTION

World over agriculture is one such human activity that critically ensures food supply to the ever-increasing human population. This has necessitated the adoption of technology and other modern means to enhance the productivity of the agriculture sector to meet the increasing demand for food required to sustain the expanding human population all over the world [1]. With the proliferation of the internet and digital technology, the agriculture sector has also adopted the best practices of these developments to explore ways and means to increase productivity, ensure efficiency, and contribute to protecting the environment [2], [3]. Lately, the trend has been

towards using data-aided tools, especially artificial intelligence (AI) based methods that are effective in providing yield fore-cast, process monitoring, accurate control, and reliable decision support [4]. The advantage of using AI tools like Artificial Neural Network (ANN), Deep Neural Network (DNN), etc., is the fact that these can learn from the surroundings, retain the learning and use it subsequently [5].

For a majority of the countries, including India, the use of pattern recognition and AI-aided tools in agriculture becomes more pertinent because a sizable section of the productive workforce is engaged in this sector. Subsequently, the mechanism embraces precision approaches increasing efficiency and productivity in agriculture. Over the years, attempts have been made to use many such techniques, including identifying deficiencies in the farm produce and

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

several other critical areas. In [6], color and pattern analysis applied to identify multiple deficiencies in paddy leaf images have been reported. There are also attempts to use internet of things (IoT) platforms and cloud computing as aids to farming. Authors in [7] proposed a framework that combines cloud computing and unified IoT for application in the agriculture sector. IoT has also been a preferred option for water flow control in agriculture [8]. It has been combined with pattern analysis techniques for providing smart agriculture solutions [9], including classification of weeds [10] and crops disease detection [11], [12]. The application of AI-based methods for predictive farmland optimization has been reported in [13], a combination of IoT and machine learning tools for goat monitoring in [14], soil texture classification using support vector machine (SVM) by Barman *et al.* in [15], chlorophyll detection by a class of learning techniques explored in [16], etc. The performance of deep learning methods for agriculture has been highlighted in [17]–[20]. Application of AI based method in combination of spatio-temporal methods for real life situations have been reported [21], [22], [26]–[28]. These are some of the recent works related to the use of AI and IoT tools in agriculture and certain real-life situations.

The discussion above establishes the importance of learning-based tools and IoT in the agriculture sector. As health monitoring of the plants is one of the foremost tasks in smart agriculture, there are ample scopes to explore innovative ways to make the farmer's life easier, especially in a country like India where agriculture and allied sectors employ a huge section of the population. Here we describe the design of a proof-of-concept approach to see the effectiveness of an arrangement based on multiple sensors and AI based decision support system used for process monitoring and control for improving yield of agricultural produce. The proposed system is part of a precision farming approach where the application of IoT devices, image processing techniques and an AI framework are configured to derive decision states and execute process control to help the farmer to maximize the produce. In this work, the decision states obtained from the AI system regarding plant health triggers a pump for water sprinkling. The farmer maybe constrained to keep an eye continuously on a plot of land where certain crops are being grown. So, for continuous monitoring and deriving decision states regarding plant health such a set-up is considered to be effective. It provides an approach to automate certain repetitive tasks and assists the farmer to execute healthy practices for obtaining better agricultural yield. The IoT system is required to capture the state of the geographical elements and facilitate continuous monitoring. A drone might be an extension of an IoT system. The Landsat images used currently to validate the ability of the AI system. The Landsat images can be replaced by real-time feed obtained from a camera mounted on a drone. Ground truth is that the life of the farmer is very miserable. So, any arrangement that makes the misery of the farmer less is always a healthy development. It matters a lot for populous countries like India.

We report the design of a health monitoring system of the plants which is based on the calculation of the Normalized Difference Vegetation Index (NDVI) to distinguish the healthy and non-healthy plants using images that are captured by a near infra-red (NIR) camera as part of an IoT set-up and integrated with a learning-aided platform. The low-cost pack connecting several sensors and the NIR camera forms the IoT and revolves around a programmed microcontroller linked to a computer system and subsequently to a cloud server. Subsequently, the entire framework works together with a learning-based system that identifies diseased and healthy leaves. The framework is formulated after conducting a series of experiments involving clustering, SVM, machine, and deep learning methods. The work involves experiments with two clustering-based segmentation methods: K-means clustering (KMC) and Fuzzy C-means (FCM) clustering used to separate healthy and diseased leaf segments. The specific AI tools used are SVM, a Multi-Layer Perceptron (MLP) that is a feed-forward ANN, a Time Delay Neural Network (TDNN) that is another feed-forward ANN, an Adaptive Neuro-Fuzzy Inference System (ANFIS) that is a fuzzy-based decision-making method, and a Convolutional Neural Network (CNN) that is a popular deep learning tool. The AI aspect of the work is deployed in a cloud server which can be accessed using an app. After extensive on-field trials, it has been found that the system is useful for a class of agricultural produce commonly seen in hot and humid conditions of India. Especially, the deep learning-based decision support system formed by the VGG16 CNN proves to be effective when used in combination with Gray level co-occurrence matrix (GLCM) features and FCM based clustering (which performs segmentation of the region of interest (RoI) of tomato leaves) and adopted as part of the health monitoring system. The system is also used to identify effective usable vegetation cover of farmland. The novelty of the system is the design of a deep learning-based decision support system formed by the VGG16 CNN which proves to be effective when used in combination of GLCM features and FCM based clustering (which performs segmentation of the region of interest (RoI) of tomato leaves) and adopted as part of the health monitoring and process control system which is also used to identify effective usable vegetation cover of a farmland. The rest of the paper is organized as follows. In Section 2, the proposed work is discussed in detail. The results and discussion have been covered in Section 3. The outcomes of the work have been summarized in the conclusion section.

## II. PROPOSED WORK

Here, we discuss the details of the plant leaf health monitoring system deployed over an IoT, a cloud server and the decision states generated using a class of clustering, SVM and learning aided methods. Before discussing the design and working of the complete system, some basic notions related to the different elements of the work are discussed in the sections below.

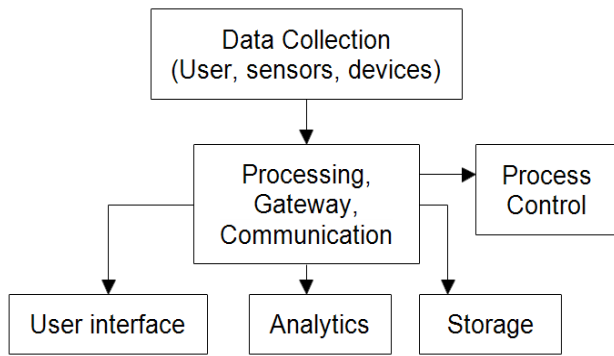


FIGURE 1. Generic IoT setup.

A. THEORETICAL CONSIDERATIONS

In this section, we include brief discussion on some of the theoretical considerations of the techniques related to the work. We specially highlight the key aspects of NDVI, IoT, clustering methods, ANN techniques like MLP and TDNN, ANFIS, SVM and CNN.

1) NDVI

NDVI is used to differentiate between bare soil and grass or forest and differentiate between various crop stages etc. It works using the principle that healthy plants reflect more near infra-red (NIR) and green light compared to other colours. So, more is the chlorophyll (healthy sign), more will be NIR reflection [16]. NDVI is a ratio between the red (R) and near infrared (NIR) values and is expressed as

$$NDVI = \frac{NIR - R}{NIR + R} \tag{1}$$

2) IOT

IoT is meant as a collection of physical devices, sensors, communication systems, user interfaces, etc., in a connected state used to provide analytics, process control and real-time decision support in a range of situations and applications [23]. A generic IoT setup is shown in Fig. 1.

3) CLUSTERING METHODS

These are a type of unsupervised grouping method where, depending upon some similarity, objects are placed in segments or classes. Two types of clustering methods are used. First one is the K-means clustering (KMC) and the second one is the Fuzzy C-Means clustering (FCMC) [5].

4) FEATURE EXTRACTION METHODS

Features provide a concise and relevant description of an input image and assist in appropriate decision-making using classifiers. However, a deep learning network may not require features as feature learning is part of such networks. But other learning-based classifiers require certain features. For agriculture produce and plants, a few popular features are Gray level co-occurrence matrix (GLCM), Local binary pattern (LBP), Local Binary Gray Level Co-occurrence Matrix

(LBGLCM), Gray Level Run Length Matrix (GLRLM), and Segmentation-based Fractal Texture Analysis (SFTA) [24]. Further, GLCM forms a composite set using features like skewness, standard deviation, homogeneity, contrast, smoothness, correlation, kurtosis, energy, entropy, mean, variance, and root mean square (RMS) to make the extracts content-rich. The GLCM feature determines the textural relationship among pixels and is related to the second-order statistics. The GLCM features determines how gray scale intensities are distributed vertically and horizontally and also in different orientations along the diagonals. With the help of the GLCM features certain statistics like contrast, correlation, energy and homogeneity can be calculated which are useful in determining the texture of the image. The LBP features denote statistical and structural model of the textual content of images and demonstrate high invariance to light intensity variations. The LBGLCM feature is generated by combining LBP and GLCCM algorithms. The GLRLM feature is obtained from texture representations that contain each pixel’s spatial plane features connected to certain higher-order statistics. The SFTA algorithm carries out a fractal-based analysis of the image texture. Skewness is associated with a third-order moment and represents a sample’s deviation from an ideal probability distribution. Standard deviation represents the deviation of a sample’s statistical properties from its mean. Homogeneity signifies the variation a sample may have from its average statistical distribution. Contrast denotes the pixel intensity variations with regard to the neighborhood. Smoothness ensures the absence of sudden variation of pixel intensity.

Correlation is linked with the similarity of pixel variations between two different images. It can measure the changes in the pixel intensity within the image compared to the situation when the variations had not taken place. Kurtosis is a fourth-order statistic. Energy is associated with localized changes in an image. Entropy denotes the information richness of an image. RMS outlines the texture variations occurring in an im-age and captures the distortions in statistical distributions.

5) ANFIS

The ANFIS is based on the Sugeno-type fuzzy systems and can be implemented using a multilayered feed forward structure [25] as shown in Fig. 2.

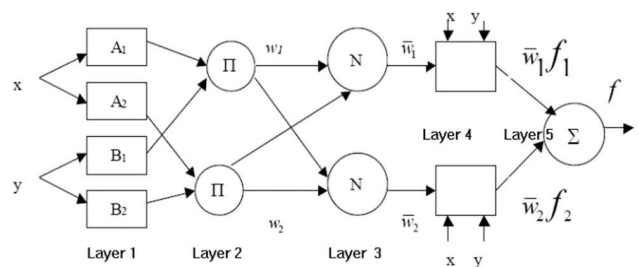


FIGURE 2. ANFIS structure.

In Layer 1 (fuzzification), each node generates the membership grades of a linguistic label. Here a membership function of the type Gaussian Membership Function (MF) is used. The  $A_j$  and  $B_j$  are the linguistic variable with Gaussian MF. In Layer 2 (antecedent), using any other fuzzy AND operation, the nodes calculate the firing strength of each rule. Next, in Layer 3 (normalization), the nodes find out the normalized firing strength. It calculates ratios of the rule's firing strength to the sum of all the rules firing strength. The nodes of Layer 4 (consequent) compute consequent parameters which are dependent on Layer 3. It is an adaptive layer. In the Layer 5 (aggregation), the overall output is calculated where each node aggregates the final output as the sum of all incoming signals. The layers are linked by connectionist links which are adaptive and update during each cycle of learning. The output is compared with the desired goal and the updating of the weights which multiply the inputs to each layer is continued till the objective is attained.

6) MLP

The MLP is a multi-layer feed forward ANN which is made up of one input and one output layer and one or several hidden layers with log-sigmoid, tan-sigmoid or purely linear activation functions. A generic MLP architecture is shown in Fig. 3 [23]. Let  $X$  be an input applied to a MLP. Let  $W_{1ij}$  be the weight matrix between input and hidden layer and  $W_{2jk}$  be the weight matrix required to connect the artificial neurons between the middle layer and the output layer. Here,  $i$  keeps count of number of input samples,  $j$  is related to the indexing of the number of hidden layer neurons and  $k$  is linked to the output layer size. The output of the MLP is expressed as

$$Y_o = \sum_{k=1}^K f_k \left\{ \sum_{j=1}^J \sum_{i=1}^I (X_i W_{1ij} + b_i) W_{2jk} \right\} \quad (2)$$

Let  $Y_T$  be the target output. The mean square error (MSE) is expressed as:

$$E = \frac{1}{2} \sum_{p=1}^P (Y_o^p - Y_T^p)^2 \quad (3)$$

During training, weight update happens following a gradient descent principle. The weight update expression is given as

$$W[n+1]_j = W[n]_j + \mu \frac{\Delta E}{\Delta W[n]} \quad (4)$$

where  $\mu$  is the learning rate of the network usually taken to be a fraction between 0 and 1. The MLP is trained with back propagation algorithm till it learns the patterns completely. The choice of hidden layers' number and the activation functions depends upon the requirements. In our case we have used a one hidden layer MLP with one and half times more numbers of artificial neurons of the input layer. These activation functions are of log-sigmoid type and are used to learn the leaf classification patterns.

7) TIME DELAY NEURAL NETWORK (TDNN)

The TDNN is another multi layered feedforward ANN like MLP but has delayed feed in the input. With this feeding

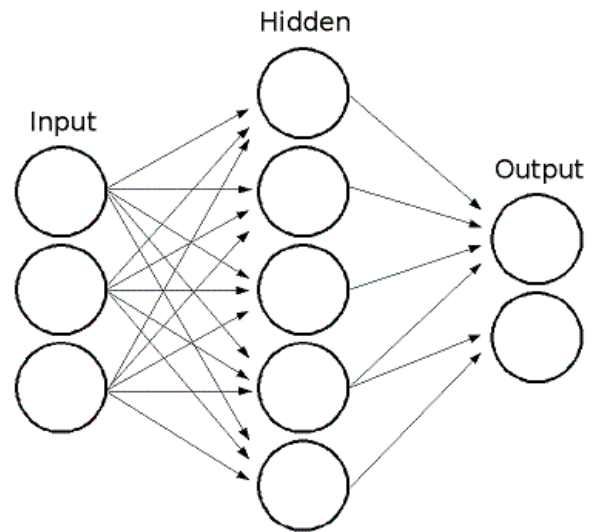


FIGURE 3. Layout of the MLP.

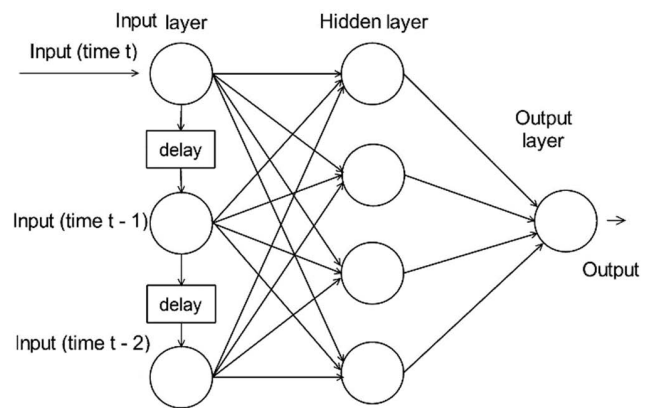


FIGURE 4. Layout of the TDNN.

method, the TDNN is able to track variations in data due to time. Like MLP, it also has multiple layers with input, output and hidden types and is trained with back propagation algorithm. In our case with have used 1 to 2 numbers of positive delay to develop time tracking ability of the ANN. Fig. 4 shows a layout of the TDNN [23]. For the TDNN, the output and other expressions are same as shown in (2)-(4) except that the input shall be a combination of  $X(n)+X(n+T)$  where T denotes the delay used in the input feed layer.

8) SUPPORT VECTOR MACHINE (SVM)

The SVM is a supervised learning-based approach used for classification and regression. The working of the SVM is explained using Fig. 5. The discrimination boundary between the classes is laid by

$$H_0: w \cdot x^i + b = 0 \quad (5)$$

where input  $x^i$  is scaled by weight  $w$  and aided by the bias  $b$ . It classifies the patterns according to the values associated with support vectors across the two sides of the



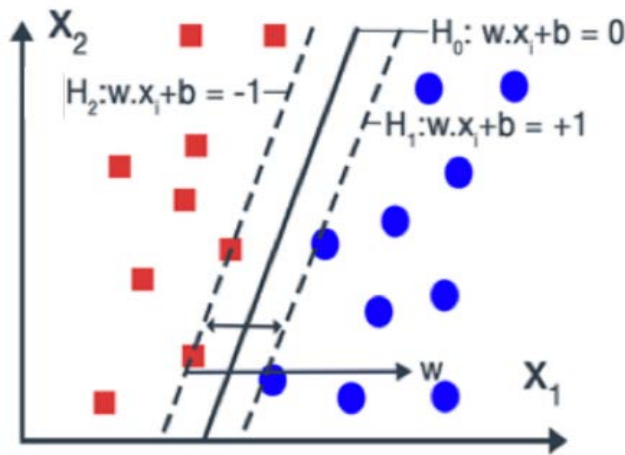


FIGURE 5. Working of the SVM.

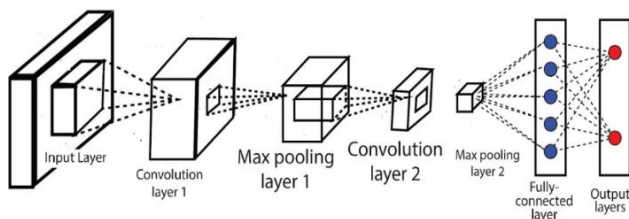


FIGURE 6. Basic layout of CNN.

hyperplane [5]. Here, each sample value is considered as a point in n-dimensional space with the numerical count of each pattern representing a particular coordinate. The closeness of the coordinate to the class-region derived through the use of support vectors determines the classification of the samples. In this work, the Gaussian kernel is used for passing the input  $x^i$  and generating multi-class states which are subsequently scaled by the weight  $w$  for performing the classification. The SVM is used as a benchmark technique in this work.

9) CONVOLUTIONAL NEURAL NETWORK (CNN)

The CNN is a deep learning network used for pattern classification and is constituted by an input, output and a number of middle layers. These include multiple repetitions of max-pooling layer, sub-sampling layer, normalization layer, fully connected layer etc. before connecting to a classification layer which is normally the output layer (see Fig. 6). The CNN doesn't require much of preprocessed and labelled data while feature learning takes place layer by layer without any human intervention. The frontend has a number of convolutional masks of varying sizes which take part in data capture with minute details which greatly enhance its discrimination and internal hierarchical feature extraction contributing towards its superior classification capabilities [5].

B. SYSTEM MODEL AND WORKING

The system model of the proposed design is shown in Fig. 7. The system has both hardware and software components.

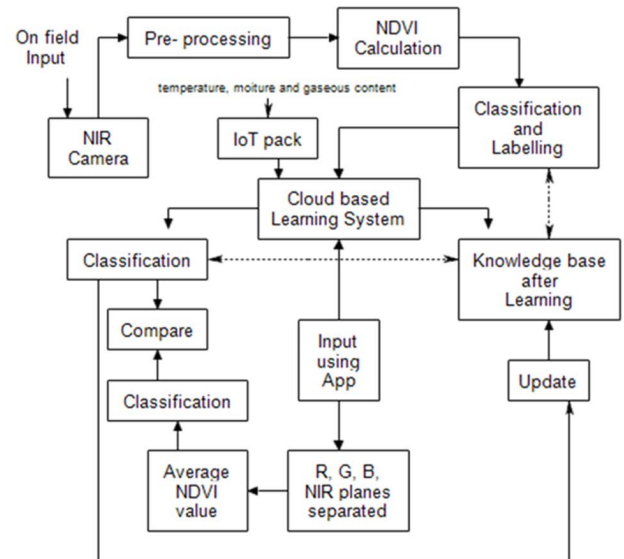


FIGURE 7. Block diagram of the proposed design.

Its hardware elements include a Raspberry Pi 3 microcontroller, temperature and humidity sensor (DHT11), soil moisture sensor, motor pump with a relay, and a NIR camera. Further, an MQ-135 gas sensor pack is also connected to detect spurious gases and initiate certain process control. The sensors and the camera are connected to the microcontroller to form an IoT pack connected to a host computer and subsequently to the internet so that seamless data transfer, web access, and related processing can occur. There are two aspects to the design. First, an on-field feed mechanism as part of which data regarding the leaf of a plant using a NIR camera, moisture content, and temperature of the surroundings are connected using the IoT pack. Next, a user can use a smartphone application (designed for the purpose) to feed the related data online to the system. Leaf health of certain common agricultural produce like cabbage, chilly, cauliflower, carrot, coriander, cucumber, pumpkin, ginger, potato, tomato, ladyfinger, and gourd are taken into consideration.

Initially, through certain on-field visits, training samples covering healthy, diseased, and dead leaves numbering at least 1000 of each of the twelve vegetable types are captured using the NIR camera. Data bootstrapping techniques (Fig. 8) are used to increase each type's training and testing set volumes when required. Simultaneously the reading of the temperature, moisture, and the surrounding gaseous contents are recorded both using the IoT pack and manually. The NDVI value for each type of leaf sample is calculated and the health state labeled. Accordingly, the associated temperature, moisture, and gaseous surrounding attributes are embedded into a file used with the cloud-based learning system. These physical parameters are required for process control like running a water pump etc., as per the vegetation requirements. But the main focus is on designing a learning-based system for monitoring plant health. Here we have explored

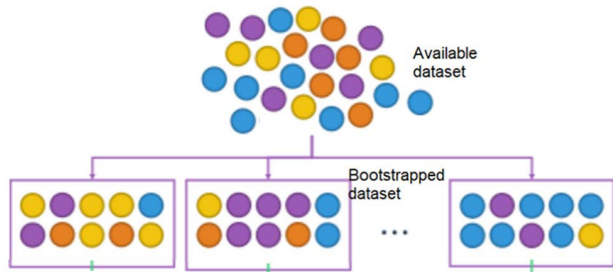


FIGURE 8. Bootstrapping method to increase data samples for training.

the effectiveness of SVM, MLP, TDNN, ANFIS, and CNN in carrying out the training and subsequently providing classification for different leaf states for the twelve identified types of vegetables. The benchmark experiments are carried out with the tomato plant. The NDVI values are calculated using certain steps. First, using the NoIR camera, a visible leaf of a plant is captured. Next, the NoIR image is split into R, G, B, and NIR planes from which NDVI values of individual planes are calculated. An average of all these values is taken from this representation, which becomes the required NDVI value. For a healthy leaf, the NDVI value is 0.96, the diseased has a value of 0.42, and a dead leaf has a value of -0.02. These are taken for apriori classification and labeling of the leaf samples. Now taking the values of the sensor, separate training cycles are carried out for each of the selected learning-based systems.

The updated states of the learning along with the classification decision details are held in a database. An application for the user has been designed using the Blync app to hold the data coming from the field sensors and interfaces in a cloud server. The images of the plant leaves are taken, pre-processed (for noise removal, enhancement, etc.), and subjected to segmentation. Clustering techniques like KMC and FCM are used for this purpose. Next, features using GLCM, LBP, LBGLCM, GLRLM, and SFTA are extracted and applied to SVM, MLP, TDNN, ANFIS, and CNN for carrying out the classification. Further, features like skewness, standard deviation, homogeneity, contrast, smoothness, correlation, kurtosis, energy, entropy, mean, variance, and root mean square (RMS) are considered a composite set derived from the GLCM and used extensively to make the process robust. The processing logic is summarized in Fig. 9.

Further, the learning-based systems use the sensor feeds to automate the process of watering the plants keeping an eye on threshold set. The sensors continuously provide the readings regarding the temperature and the humidity sensor of the surroundings and the soil moisture state. The combination of these three readings is used to decide upon the triggering of the water pump and the duration. All probable conditions involving the temperature, humidity and soil moisture and association with the watering requirements are ascertained beforehand for a particular crop type and applied to the learning system. These time dependent sensor feeds are learnt and the mapping with the watering states are established by the learning systems.

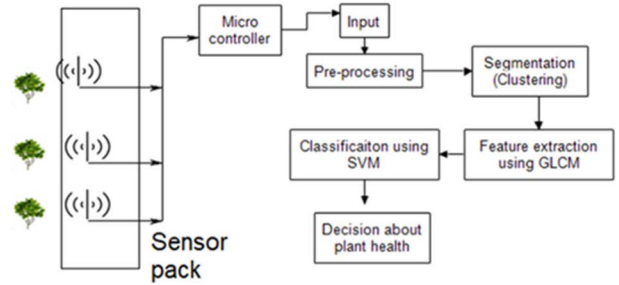


FIGURE 9. Process logic of plant leaf classification.

TABLE 1. Experimental parameters.

Description	Type
Preprocessing	Noise removal (weighted moving average, median filter, low pass filter, Gaussian filter and Weiner filter), enhancement (histogram processing), re-sizing;
Image segmentation	KMC, FCM
Features	GLCM, LBP, LBGLCM, GLRLM, SFTA and composite set
Classifiers	SVM, MLP, TDNN, ANFIS, CNN MLP and TDNN with N size input layer, 1.8 times N middle layer and C size output layer where N is the size of the input pattern and C is number of classes; TDNN with 0.1 delay; Log-sigmoid activation function used; training with scaled gradient conjugate (SCG) back propagation (BP); ANFIS formed with Bell shaped membership functions;
Parameters	NoIR camera input, NVDI, temperature, moisture, gaseous content
Data	Training set of 12,250 on-field samples; Landsat images for training validation of CNN
Training time (max.)	SVM- 67 minutes, MLP- 88 minutes, TDNN- 84 minutes, ANFIS- 48 minutes, CNN- 225 minutes.

### III. EXPERIMENTAL DETAILS AND RESULTS

Extensive experiments have been carried out to check the performance of the proposed design. The relevant experimental specifications are summarized in Table 1.

The experiments are carried out taking into account these experimental specifications covering pre-processing method, segmentation approach, features considered, classifiers used, parameters necessary and details of data applied to train the learning aided decision support system. A GUI has been developed for applying the samples into the system. The outputs of each stage can be obtained using the GUI while the system processes and generates responses. After carrying out the pre-processing operations like noise removal and enhancement (Fig. 10), and re-sizing in some cases, image segmentation is carried out using two clustering techniques which have already been mentioned above.

Weighted moving average filtering provides 8 to 12% improvement in terms of peak signal to noise ratio (PSNR)

**TABLE 2.** Accuracy (Acc), IOU and PA performance of KCM and FCM clustering approaches used for segmentation.

Segmentation using	SVM (in %)			MLP (in %)			TDNN (in %)			ANFIS (in %)			CNN (in %)		
	Acc	IOU	PA	Acc	IOU	PA	Acc	IOU	PA	Acc	IOU	PA	Acc	IOU	PA
KMC	83	80	81	84	80	82	84	80	81	87	82	83	92	85	87
FCM	84	81	82	85	81	82	85	81	82	87	83	85	94	86	88

values compared to spatial filtering during noise removal. Similarly, histogram processing consistently yields 9% better results in terms of PSNR compared to enhancement operations using filter masks. The fuzzy based segmentation approach related to FCM provides better segmentation performance (segmented portion + GLCM feature + classifier) as observed from the experimental results (Table 2).

This is because of the fact that the randomness and minor variations in the pixels of the leaf samples are handled well by the fuzzy attributes in the approach. The classifiers provide outputs in terms of true positive (TP), true negative (TN), false positive (FP) and false negative (FN). The ratio between the sums of TP and TN (TP+TN) and TP, TN, FP and FN (TP+TN+FP+FN) in percentage gives accuracy. Further, as the segmentation block plays a very significant role in determining the reliability of the proposed system, its performance is ascertained in terms of Intersection over union (IOU) and Pixel Accuracy (PA). The IOU and PA values are obtained using

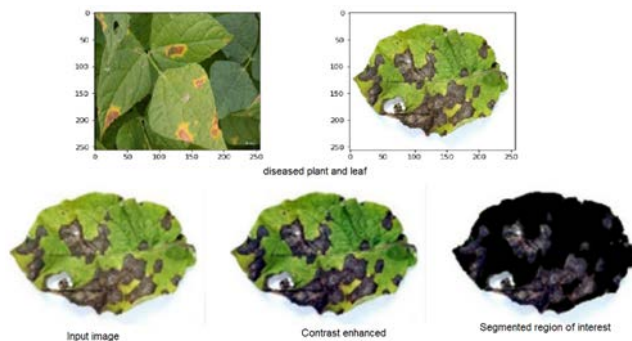
$$IOU = \frac{TP}{TP + FP + FN} \tag{6}$$

$$PA = \frac{TP}{TP + FN} \tag{7}$$

The average values of IOU and PA obtained from both the segmentation methods used with the classifiers are shown in Table 2. Moreover, the clustering-based approach is used for selecting the region of interest (RoI) which is reinforced using manual labelling. The clustering based RoI selection can be used to automate the class labelling process. This is shown in Fig. 11. The GLCM algorithm enables calculation of contrast, correlation, energy, homogeneity, mean, standard deviation, entropy, RMS, variance, smoothness, kurtosis and skewness values (Fig. 12) which are applied to a benchmark SVM classifier. The classifier not only classifies in terms of healthy and unhealthy leaves but also detects disease attacks like bacterial spot, blight, septoria, spider mites and yellow leaf curl. These are common disease associated with tomato plants.

The classification has also been carried out with MLP, TDNN, ANFIS and CNN. In case of the deep learning-based recognition, the CNN based VGG16 model is taken for the purpose of classifying the leaf disease. The configuration of the VGG16 is shown in Fig. 13.

The maximum time consumed by each of the learning based methods while training with the full complement of training data is shown in Table 1. While the CNN based VGG16 model takes the maximum time, it is robust, doesn't require additional feature learning, provides



**FIGURE 10.** Image processing operations over a diseased leaf.

consistent accuracy and demonstrates the best ability to handle variations in the input. The SVM takes least amount of time to train but reliability performance with samples not within the extended data set is the worst amongst the methods considered for the work. The multi layered feature transform taking place in case of the CNN based VGG16 model enables detailed capture of the relevant content and hence provides the best results despite a higher requirement of training time. However, the testing time for all the methods are nearly equal.

The model has a classifier layer at the end which is preceded by a block formed by three fully connected layers two of which have a length of 4096 and the last layer has a length of 1000. The classification layer is formed by a soft-max layer with log-sigmoid activation function. The main body of the VGG16 CNN is formed by five repetitions of two blocks of 3 × 3 convolution mask and a layer of 64 rectified linear unit (ReLU) activations followed by a 2 × 2 max-pooling layers. This thick block carries out the feature learning from the input samples. Experiments are carried out using the feature set and also by feeding the images directly. The soft-max layer is initially provided with labelled class data for carrying out the learning. In another set of experiments the approach is made automated by using the FCM to generate the RoI from the leaves which are used as targets for classification. Fig. 14 shows a training time accuracy performance generated using the feature set.

This accuracy (number of correct pixel classification compared to the total number of pixels of an image sample expressed in percentage) of around 95% (Fig. 14) in the case of tomato leaf samples for the five different disease types is generated after seven epochs of training with a batch size of 50 which is repeated for about 500 training samples. The average performance with all the identified plant types and class groupings suffers a bit but is over 94%. When the images are used as input, the training takes over ten times more



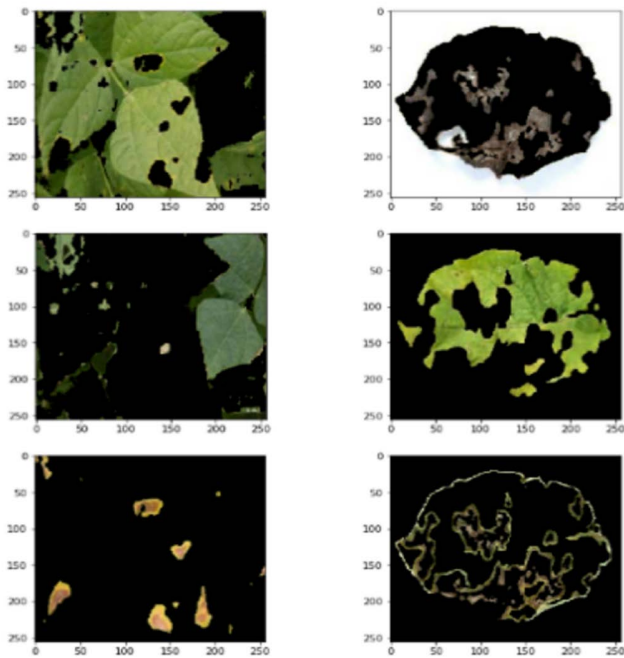


FIGURE 11. GUI showing pre and post pre-processing, segmentation and selection of region of interest operations over a diseased leaf.

Contrast : [[0.18820466]]	Contrast : [[0.50909926]]
Correlation : [[0.90418]]	Correlation : [[0.98103431]]
Energy : [[0.94638166]]	Energy : [[0.52803895]]
Homogeneity : [[0.98483242]]	Homogeneity : [[0.94565149]]
Mean : 7.041798909505208	Mean : 128.97718302408853
Standard Deviation : 33.23482406088039	Standard Deviation : 119.9977207315583
Entropy : 0.6585890945441193	Entropy : 3.1332164654303796
RMS : 2.3426733030267483	RMS : 4.416505822665767
Variance : 1104.5535303576744	Variance : 14399.452980769072
Smoothness : 0.999999277784545	Smoothness : 0.999999685646309
Kurtosis : 5.300034068023861	Kurtosis : 0.6188549002996556
Skewness : 2.245031649067062	Skewness : -0.28050356131727366

FIGURE 12. Feature values of disease affected region from GLCM algorithm.

cycles, but accuracy performance is improved to 96% with no dedicated feature extraction block requirements. The performance obtained with the validation set is better compared to the training set. This is because the validation data set has better diversity than the training dataset. During training, the AI system completely becomes familiar with the training data set. Hence, the training dataset loses diversity, but the validation dataset is formed using bootstrapping technique which has greater diversity. Learning-based systems work well with datasets where there is diversity. The experimental results of the various experiments are explained below.

The accuracy performance of the segmentation block is shown in Table 2. The KCM and FCM approaches are used to extract out the leaf segments of the tomato leaf samples. For these segments, features are extracted using the GLCM approach and classification using SVM, MLP, TDNN, ANFIS, and CNN classifiers. Except for ANFIS in all the classifiers, the FCM based segmentation contributes towards improved performance. In the case of CNN, there is a 2% improvement while in cases of SVM, MLP and the TDNN there is 1% better result due to the use of the FCM. The

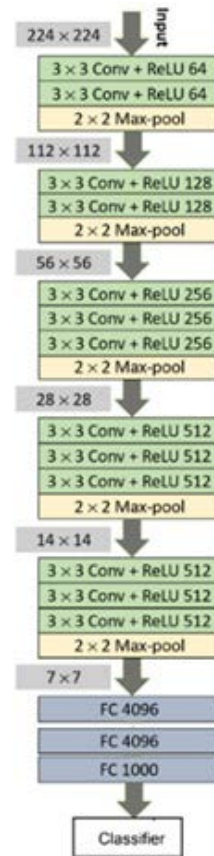


FIGURE 13. VGG16 CNN used for plant leaf disease.

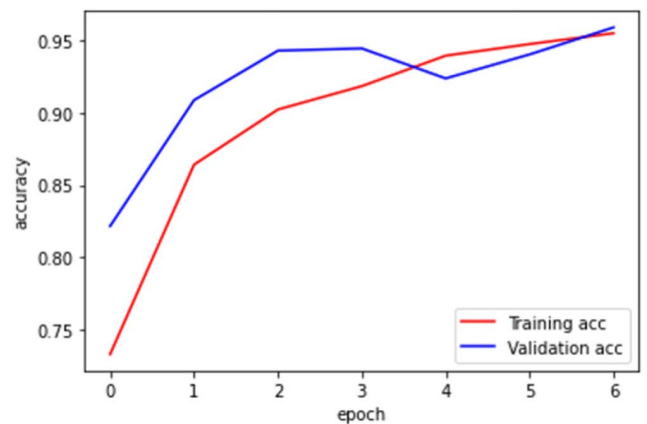


FIGURE 14. Training time performance of the VGG16 CNN.

fuzzy-based approach of capturing the segment details helps to extract the segments better. The use of the fuzzy-based image segmentation also helps in finer discrimination even if there might be a very small amount of variation for extraction out the leaf's healthy and diseased sections, which adds to the improvement in the performance of the classifiers.

Another set of results are generated to show the effectiveness of the features. We have used GLCM, LBP, LBGLCM, GLRLM, and SFTA features to derive decisions regarding the health state of the plants. For about 500 tomato leaf



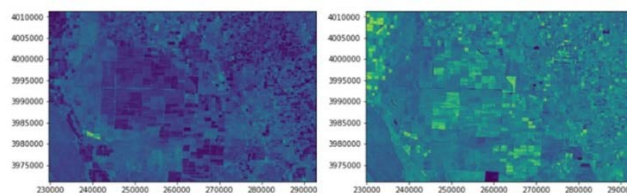
**TABLE 3. Accuracy performance of feature sets with a few identified classifiers as part of health monitoring of tomato leaf samples.**

Segmentation using	SVM	MLP	TDNN	ANFIS	CNN
GLCM	84%	85%	85%	87%	94%
LBP	84.3%	85.5%	86%	88%	94%
LBGLCM	83%	85%	86%	88%	94%
GLRLM	83.5%	85.3%	86%	88%	94%
SFTA	84%	86%	86%	88%	94%
GLCM+GLRLM	85%	86%	87%	88%	94%
GLCM+SFTA	85.2%	86.5%	87.3%	88%	94%
GLCM+GLRLM+SFTA	86%	86.7%	87.5%	88%	94%

samples sets of five different features, their combinations are extracted and applied to SVM, MLP, TDNN, ANFIS and CNN classifiers. This is calculated for a number of trials for the samples used during training and testing. The average accuracy performances are recorded and are shown in Table 3.

The healthy and diseased extracts segmentation is done using the FCM approach, showing improved performance, as noted in Table 2. It is seen that the best performance with GLCM is obtained from CNN, giving 94% accuracy, while with SVM, MLP, TDNN, and ANFIS, the variations are between 84 to 87%. With LBP features, there is minor performance improvement with SVM, MLP, TDNN, and ANFIS but no change with the CNN. There is a minor performance fall with LBGLCM and GLRLM features with SVM and MLP classifiers compared to LBP, while with TDNN, ANFIS and CNN, there are no changes in accuracy. The SFTA features show minor improvement in accuracy with SVM and MLP classifiers, but the performance is constant with TDNN, ANFIS, and CNN. A few combinations of features are explored, out of which GLCM + GLRLM, GLCM+SFTA, and GLCM+GLRLM+SFTA show variations in performances, as shown in Table 3. The CNN and in many cases the ANFIS show very little variation in performance despite changes in the features and the use of combinations. Since the GLCM feature with FCM based segmentation and classification using CNN gives the best performance, this combination (FCM+GLCM+CNN) is used as the standard combination for the subsequent work.

As the proposed system is intended to help the farmers, an additional characteristic incorporated into it is its ability to identify farm vegetation and the geographical distribution of the surroundings. Most agricultural lands are surrounded by a range of vegetation and other geographical elements. During rainy seasons, the distribution of thick vegetation, normal vegetation, vegetation cover less soil, water, and others (not falling into these four) vary. Certain weed types grow very fast and harm agricultural produce. With it comes insects and other life forms. So, the distribution of geographical elements around farmland and its monitoring using an IoT system is essential for the effective management of farmland. The IoT system is required to capture the state of the geographical elements and facilitate continuous monitoring. A drone might be an extension of an IoT system. The Landsat images are currently used to validate the ability of the AI system. The



**FIGURE 15. Red band and NIR band images obtained as part of vegetation estimation.**

**TABLE 4. PDR with varying pack sizes of sensors with Bluetooth, physical and Wi-Fi connectivity.**

Pack size	Bluetooth	Physical	Wifi
4-pack	0.78	0.92	0.86
6-pack	0.78	0.92	0.86
8-pack	0.78	0.92	0.86
10-pack	0.68	0.90	0.85
12-pack	0.67	0.89	0.82
16-pack	0.66	0.86	0.82

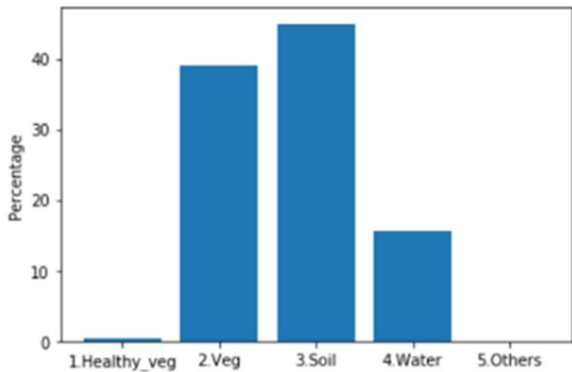
Landsat images can be replaced by real-time feed obtained from a camera mounted on a drone. The VGG16 network is trained with Landsat satellite data to classify thick vegetation, normal vegetation, and vegetation cover less soil, water, and others (not falling into these four). NVDI threshold values are calculated for use with the system to derive the class decisions for these classes. From the satellite images band4 (Red channel) and band5 (Near Infrared band) (Fig. 15), contents of a particular area are generated, which are next used to calculate the NDVI values.

The NDVI image of an area is obtained, which forms the basis of estimating the total vegetation cover of the area. The image dataset consists of 9627 samples separated into a 60:40 ratio where 60% of them have been used for training, and 40% applied for model evaluation. With a batch size of 50, the model took seven epochs completed over 35 seconds on a cloud-based GPU. In this effort, the weights and the features of the datasets are obtained through different processes which amounted to over two hours of computation. A separate validation set generated using bootstrapping method is applied for validating the VGG16 network after training. The output of the FCM+NVDI+CNN combination for a set of experiments related to vegetation estimation is shown in Fig. 16.

Certain experiments have been carried out to see the effectiveness of the IoT arrangement. Experiments involving the packet delivery ratio (PDR) with sensor pack sizes varying between 4 and 16 and connected via Bluetooth, Wi-fi, and physical connection have been carried out. There are one temperature and one moisture sensor in each sensor pack and a NoIR camera connected to a Raspberry Pi microcontroller. This arrangement is for one plant. The system's effectiveness has been tested by connecting the sensor packs to the controller computer in numbers between 4 and 16 to form a robust IoT arrangement. The effectiveness of the arrangement

**TABLE 5. Outcomes of statistical trials involving recall, specificity, type I and type II errors, F1 score, micro F1 and macro F1 values involving leaf disease recognition (C1) with GLCM features and FCM segmentation and geographic distribution identification (C2) using SVM, MLP, TDNN, ANFIS and CNN classifiers.**

Sl no.	Model	Cases	Recall (%)	Specificity (%)	Type I error	Type II error	F1 score (%)	Micro F1 (%)	Macro F1 (%)
1	SVM	C1	83	84	0.07	0.09	81	82	83
		C2	84	85	0.06	0.08	80	82	83
2	MLP	C1	85	87	0.07	0.08	82	84	85
		C2	84	87	0.08	0.08	82	85	85
3	TDNN	C1	85	86	0.09	0.09	84	86	88
		C2	85	87	0.08	0.09	85	87	88
4	ANFIS	C1	86	88	0.06	0.06	85	89	89
		C2	88	88	0.06	0.07	85	89	89
5	CNN	C1	93	95	0.06	0.07	91	93	94
		C2	95	95	0.04	0.06	92	93	94



**FIGURE 16. Percentage distribution of different classes of geographic content extracted as part of vegetation estimation.**

is checked in terms of PDR, and the results are summarized in Table 4. The physical connectivity has the best PDR, while Bluetooth links prove to be less reliable.

We have performed certain statistical experiments to verify the experimental reliability and real-time robustness in on field conditions. The results are summarized in Table 5. It includes the outcomes of statistical trials involving recall, specificity, type I and type II errors, F1 score, micro F1 and macro F1 values involving leaf disease recognition (C1) with GLCM features and FCM segmentation and geographic distribution identification (C2) using SVM, MLP, TDNN, ANFIS and CNN classifiers.

A summary of the performance of the proposed approach and its comparison with certain previous research is included in Table 6. As discussed above, the techniques reported in [22] and [15] have been reproduced and applied to perform plant health monitoring. These are compared with those obtained with pre-trained VGG16 CNN + GLCM and VGG 16 CNN taking the complete leaf sample as input (without features) as discussed above. With around 9000 samples, the experiments are carried out, and the average results are

**TABLE 6. Summary of results and comparison.**

Description	Accuracy Performance (in %)
SVM+GLCM [22]	84%
SVM+SFTA [22]	84%
Boosted tree + SFTA [22]	88%
SCG + ANN [15]	87%
Pre-trained VGG16 CNN + GLCM (proposed)	95%
Pre-trained VGG16 CNN without features	96%

reported. The samples are divided into a 60:40 ratio, with 60% used for training and 40% used for testing. Samples have light intensity variations, and noise-related distortions assumed to be identical to those seen due to faulty sensor pickup. The training time associated with the VGG16 CNN while feeding the complete image and with no feature is considerably higher compared to the case when features are used. The difference is at least two times less in the case when features are used. But when no features are used, the associated computation cycles are saved, and the design complexity is lowered. However, it increases the computational workload of the CNN, but if the weights are trained separately, and one cycle is completed successfully, subsequent iterations become less intensive. The average results clearly indicate the advantage of the proposed approach.

**IV. CONCLUSION**

Here, we have discussed the design of a smart agriculture assistance system that is based on IoT and learning-based decision making. The proposed system has a sensor pack that obtains a readout of temperature, moisture of the air, and the soil and a camera that captures NIR images of the plant leaves. The sensor readings are used to obtain continuous monitoring and to regulate a small-scale water

sprinkling mechanism. Further, the images obtained from the NIR camera are used to train a set of learning-based tools that help monitor the health of the plant leaves. We identify twelve types of vegetation for the study, out of which five disease types of the tomato leaves are categorized into five different types using a pre-trained VGG16 CNN. The work includes experiments conducted with two clustering-based segmentation methods, namely KMC and FCM, which separates the healthy and diseased leaf segments and features, including GLCM, LBP, LBGLCM, and GLRLM SFTA. Specific AI tools like SVM, MLP, TDNN, ANFIS, and the VGG16 CNN are trained to provide automated decision support for the classification of the plant leaf disease types through a cloud server that can be accessed using an app. Extensive on-field trials show that the system is effective in hot and humid conditions, especially the FCM, GLCM feature, and the VGG16 CNN proves to be reliable.

Further, the system has been found to be effective in identifying certain usable vegetation cover of farmland and regulation the watering mechanism of crops. The system proves to be effective in identifying the disease type of certain agricultural products. It proves reliable in discriminating thick vegetation, normal vegetation, vegetation cover less soil, water, and others (not falling into these four) controlling the water spaying mechanism associated with the complete framework. An expanded version of the work with certain added features shall be an effective aid of the farmers.

## REFERENCES

- [1] A. D. Foster and M. R. Rosenzweig, "Microeconomics of technology adoption," in *Economic Growth Centre Discussion*, vol. 2. Annu. Rev. Econ., 2010, pp. 395–424. [Online]. Available: <https://www.annualreviews.org/doi/abs/10.1146/annurev.economics.102308.124433>
- [2] M. A. Shahin and S. J. Symons, "A machine vision system for grading lentils," *Can. Biosyst. Eng.*, vol. 43, pp. 7–14, Jun. 2001.
- [3] R. Shahzadi, J. Ferzund, M. Tausif, and M. Asif, "Internet of Things based expert system for smart agriculture," *Int. J. Adv. Comput. Sci. Appl.*, vol. 7, no. 9, pp. 341–350, 2016.
- [4] K. Jha, A. Doshi, P. Patel, and M. Shah, "A comprehensive review on automation in agriculture using artificial intelligence," *Artif. Intell. Agricult.*, vol. 2, pp. 1–12, Jun. 2019.
- [5] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.
- [6] M. V. Latte and S. Shidnal, "Multiple nutrient deficiency detection in paddy leaf images using color and pattern analysis," in *Proc. Int. Conf. Commun. Signal Process. (ICCSP)*, Apr. 2016, pp. 1247–1250.
- [7] J. Rui and S. Danpeng, "Architecture design of the Internet of Things based on cloud computing," in *Proc. 7th Int. Conf. Measuring Technol. Mechatronics Autom.*, Jun. 2015, pp. 206–209.
- [8] V. V. H. Ram, H. Vishal, S. Dhanalakshmi, and P. M. Vidya, "Regulation of water in agriculture field using Internet of Things," in *Proc. IEEE Technological Innov. ICT Agricult. Rural Develop. (TIAR)*, Jul. 2015, pp. 112–115.
- [9] A. Kapoor, S. I. Bhat, S. Shidnal, and A. Mehra, "Implementation of IoT (Internet of Things) and image processing in smart agriculture," in *Proc. Int. Conf. Comput. Syst. Inf. Technol. Sustain. Solutions (CSITSS)*, Oct. 2016, pp. 21–26.
- [10] M. Dyrmann, A. K. Mortensen, H. S. Midtby, and R. N. Jørgensen, "Pixel-wise classification of weeds and crops in images by using a fully convolutional neural network," in *Proc. Int. Conf. Agricult. Eng., Aarhus*, Denmark, Jun. 2016, pp. 26–29.
- [11] J. Amara, B. Bouaziz, and A. Algergawy, "A deep learning-based approach for banana leaf diseases classification," in *Proc. Stuttgart, BTW Workshop*, 2017, pp. 79–88.
- [12] A. Devaraj, K. Rathan, S. Jaahnavi, and K. Indira, "Identification of plant disease using image processing technique," in *Proc. Int. Conf. Commun. Signal Process. (ICCSP)*, Apr. 2019, pp. 749–753.
- [13] M. O. Adebisi, R. O. Ogundokun, and A. A. Abokhai, "Machine learning-based predictive farmland optimization and crop monitoring system," *Scientifica*, vol. 2020, pp. 1–12, May 2020.
- [14] Y. Rao, M. Jiang, W. Wang, W. Zhang, and R. Wang, "On-farm welfare monitoring system for goats based on Internet of Things and machine learning," *Int. J. Distrib. Sensor Netw.*, vol. 16, no. 7, pp. 1–17, 2020.
- [15] U. Barman and R. D. Choudhury, "Soil texture classification using multi class support vector machine," *Inf. Process. Agricult.*, vol. 7, no. 2, pp. 318–332, Jun. 2020.
- [16] U. Barman and R. D. Choudhury, "Smartphone image based digital chlorophyll meter to estimate the value of citrus leaves chlorophyll using linear regression, LMBP-ANN and SCGBP-ANN," *J. King Saud Univ. Comput. Inf. Sci.*, doi: [10.1016/j.jksuci.2020.01.005](https://doi.org/10.1016/j.jksuci.2020.01.005).
- [17] P. Sharma, Y. P. S. Berwal, and W. Ghai, "Performance analysis of deep learning CNN models for disease detection in plants using image segmentation," *Inf. Process. Agricult.*, vol. 7, no. 4, pp. 566–574, Dec. 2020.
- [18] S. H. Lee, H. Goëaua, P. Bonnetta, and J. Alexis, "New perspectives on plant disease characterization based on deep learning," *Comput. Electron. Agricult.*, vol. 170, Mar. 2020, Art. no. 105220.
- [19] Z. Zhai, J. F. Martínez, V. Beltran, and N. L. Martínez, "Decision support systems for agriculture 4.0: Survey and challenges," *Comput. Electron. Agricult.*, vol. 170, Mar. 2020, Art. no. 105256.
- [20] T. V. Klompenburga, A. Kassahuna, and C. Catalb, "Crop yield prediction using machine learning: A systematic literature review," *Comput. Electron. Agricult.*, vol. 170, Oct. 2021, Art. no. 105709.
- [21] A. Ali, Y. Zhu, and M. Zakarya, "A data aggregation based approach to exploit dynamic spatio-temporal correlations for citywide crowd flows prediction in fog computing," *Multimedia Tools Appl.*, vol. 80, no. 20, pp. 31401–31433, Aug. 2021, doi: [10.1007/s11042-020-10486-4](https://doi.org/10.1007/s11042-020-10486-4).
- [22] N. Awan, A. Ali, F. Khan, M. Zakarya, R. Alturki, M. Kundi, M. D. Alshehri, and M. Haleem, "Modeling dynamic spatio-temporal correlations for urban traffic flows prediction," *IEEE Access*, vol. 9, pp. 26502–26511, 2021, doi: [10.1109/ACCESS.2021.3056926](https://doi.org/10.1109/ACCESS.2021.3056926).
- [23] S. Greengard, *The Internet of Things*. Cambridge, MA, USA: MIT Press, 2015.
- [24] Ş. Öztürka and B. Akdemirb, "Application of feature extraction and classification methods for histopathological image using GLCM, LBP, LBGLCM, GLRLM and SFTA," in *Proc. Int. Conf. Comput. Intell. Data Sci. (ICCIDS)*, 2018, pp. 40–46.
- [25] S. Kaushik and S. Tiwari, *Soft-Computing*, 1st ed. New Delhi, India: MacGraw-Hill, 2018.
- [26] A. Ali, Y. Zhu, Q. Chen, J. Yu, and H. Cai, "Leveraging spatio-temporal patterns for predicting citywide traffic crowd flows using deep hybrid neural networks," in *Proc. IEEE 25th Int. Conf. Parallel Distrib. Syst. (ICPADS)*, Dec. 2019, pp. 125–132.
- [27] A. Ali, Y. Zhu, and M. Zakarya, "Exploiting dynamic spatio-temporal graph convolutional neural networks for citywide traffic flows prediction," *Neural Netw.*, vol. 145, pp. 233–247, Jan. 2022.
- [28] A. Ali, Y. Zhu, and M. Zakarya, "Exploiting dynamic spatio-temporal correlations for citywide traffic flow prediction using attention based neural networks," *Inf. Sci.*, vol. 577, pp. 852–870, Oct. 2021.



**KANDARPA KUMAR SARMA** (Senior Member, IEEE) received the M.Tech. degree in signal processing from the Indian Institute of Technology Guwahati, Assam, India, in 2005, and the Ph.D. degree in the area of mobile communication from the Indian Institute of Technology Guwahati. He currently works as a Professor and the Head of the Department of Electronics and Communication Engineering, Gauhati University, Assam. His research interests include deep learning, mobile communications, speech processing, antenna design, and electronic warfare.



**KUNAL KINGKAR DAS** received the B.Tech. degree in electronics and communication engineering from Gauhati University, in 2020. He specializes in intelligent system design.



**SAMADRITA BHUIYA** received the B.Tech. degree in electronics and communication engineering from Gauhati University, in 2020, where she is currently pursuing the master's degree in microelectronics and VLSI design. She specializes in intelligent system design and intelligent farming.



**VIKASH MISHRA** received the B.Tech. degree in electronics and communication engineering from Gauhati University, in 2020. He specializes in intelligent system design and artificial intelligence. He is currently working as an Assistant System Engineer with Tata Consultancy Service, India.



**DMITRII KAPLUN** (Member, IEEE) received the Ph.D. degree, in 2009. In 2009, he defended his Ph.D. thesis in digital signal processing at Saint Petersburg Electrotechnical University "LETI," Saint Petersburg, Russia. He was an Associate Professor, in 2015. He is currently a Senior Researcher at Saint Petersburg Electrotechnical University "LETI." His current research and academic work are concerned with digital signal and image processing, radio monitoring and hydroacoustic monitoring applications, embedded and reconfigurable systems, computer vision, and machine learning. He regularly takes part in different collaborative projects related to signal processing, neural networks, and software-hardware implementation of digital signal processing algorithms. The most substantial results are in the fields of digital signal and image processing, embedded systems, neural networks, and machine learning. He is the author of more than 50 papers in journals and conference proceedings.

...