

Received December 9, 2021, accepted January 3, 2022, date of publication January 6, 2022, date of current version January 14, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3140815

Rice-Fusion: A Multimodality Data Fusion Framework for Rice Disease Diagnosis

RUTUJA R. PATIL AND SUMIT KUMAR^{ID}, (Member, IEEE)

Symbiosis Institute of Technology, Symbiosis International (Deemed University), Pune, Maharashtra 412115, India

Corresponding author: Sumit Kumar (er.sumitkumar21@gmail.com)

ABSTRACT Rice leaf infections are a common hazard to rice production, affecting many farmers all over the world. Early detection and treatment of rice leaf infection are critical for promoting healthy rice plant growth and ensuring adequate supply for the fast-growing population. Computer-assisted rice leaf disease diagnoses are hampered due to strong image backgrounds. Popular Convolutional Neural Network (CNN) architecture extracts the features from images and diagnoses the disease to address the issues above. However, this method is best suitable for segmented images and gives low accuracy with real-time images. In this case, the Internet of Things is a paradigm shift that collects agro-meteorological information that effectively helps diagnose rice diseases. Motivated by the usefulness of CNN models and agricultural IoT, a novel multimodal data fusion framework named Rice-Fusion is proposed to diagnose rice disease. Rice disease diagnosis based on a single modality may not be accurate, and hence the fusion of heterogeneous modalities is essential for robust and reliable disease diagnosis. This gives a new dimension to the domain of rice disease diagnosis. The dataset was collected manually with 3200 rice health category samples using two modalities, namely agro-meteorological sensors and a camera. The Rice-Fusion framework initially extracts the numerical features from agro-meteorological data collected from sensors. Next, it extracts the visual features from the captured rice images. These extracted features are further fused using a concatenation layer followed by a dense layer, which provides single output for diagnosing the rice disease. The testing accuracy of Rice-Fusion is 95.31% as opposed to other unimodal framework accuracies of 82.03% and 91.25% based on CNN and Multi-Layer Perceptron (MLP) architectures, respectively. Experimental results analysis demonstrates that the proposed Rice-Fusion multimodal data fusion framework outperforms the outcome of unimodal frameworks.

INDEX TERMS Rice disease diagnosis, convolutional neural network, multi-layer perceptron, rice-fusion, multimodal data fusion.

I. INTRODUCTION

As per the statistical analysis [1], the two main causes that lead to depletion in food availability are crop diseases and pests that attack the crop and thus resulting in causing significant losses to agricultural production. The leading causes are poor water management, inadequate soil nutrients, unstable climatic conditions that lead to plant diseases and ultimately reduce the yield [2]. The right decisions can be made by developing decision support systems that can assist farmers in taking the right actions and achieving higher crop yields. Therefore, automatic and accurate diagnosis of plant diseases plays an essential role in ensuring high yield and

quality [3]. It also avoids manually identifying plant diseases in the field [4], [5]. The automatic detection and analysis of plant illnesses using image processing techniques is currently a challenging subject that is being actively researched for applications such as early disease diagnosis, disease prediction, pesticide recommendation, etc. Multispectral, hyperspectral [6], [7], and digital images [8] are used extensively in the literature. Using digital photographs is the most prevalent approach among these.

Rice production makes a significant contribution to the agricultural economy. With overall consumption of 486.62 million metric tons in 2018–2019 and 496.30 million metric tons in 2019–2020, it is one of the world's most widely consumed cereal crops. When compared to the metric tons consumed through time, this indicates a rise in rice

The associate editor coordinating the review of this manuscript and approving it for publication was Long Wang^{ID}.

consumption. It is predicted that rising rice consumption will keep pace with increasing production rates. However, disease-related problems frequently destroy a considerable rice volume due to a lack of adequate field monitoring. Several illnesses often occur in rice production, resulting in significant economic losses. Furthermore, the widespread use of pesticides to treat plant diseases has had negative consequences in the agro-ecosystem [9]. The most common rice illnesses are sheath blight, bacterial blight, rice blast, and symptoms characterized by texture, color, and shape, typical of rapid development and uncomplicated infection [10]. Rice disease detection procedures now include artificial identification, querying rice disease maps, and automated detection.

In recent years, a tremendous improvement in computational power efficiency and the extensive amount of data accessible from various sources that may be utilized to learn more about the agriculture industry is witnessed. The domain of Deep Learning and the Internet of Things are explored and have opened up new avenues for diagnosing crop anomalies. Disease surveillance is broadly categorized in three ways: using digital and spectral images, using measurements from soil sensors, and analyzing climatic normal. The different types of Artificial Intelligence (AI) models are built, trained, validated, and tested against the data collected by the ways mentioned above [11]. Thus, to develop an integrated management system for crop diseases, ML and DL algorithms can improve the farmers' profit, and land resources can be conserved. In totality, these techniques provide effective treatments in the appropriate location, at the appropriate time, and at the appropriate rate [12].

The research in the agricultural domain is enriched with a variety of data that can be obtained from various sources like IoT sensors, vegetation indices, images from Unmanned Aerial Vehicle(UAV), satellite images. Data fusion techniques must fuse multiple forms of retrieved data to comprehend crop growth circumstances and disease symptoms development. Furthermore, machine learning-based data fusion has advanced significantly, and when applied to agriculture data, it will have a significant impact on plant protection, particularly early disease diagnosis [13]. The combination of agricultural data from various data collection tools with AI algorithms and fusion algorithms has resulted in widespread research in Precision Agriculture, particularly for crop growth monitoring and protection.

Deep learning is one of the most popular approaches in crop disease detection [14]. However, there are still problems with using deep learning as a plug-and-play formula in identifying crop diseases. While many deep learning-based approaches recognize diseases in different crops, such as potatoes, rice, and tomatoes, some researchers focused on identifying crop diseases in the field where crops are cultivated, and the field scenarios can influence the classifier by discriminating between different types of illness. The paper aims to solve the automatic identification of various rice diseases with images and environmental data collected from the field using AI techniques. The algorithms used to identify

the disorders must be robust enough to face the challenges in diagnosing diseases.

Furthermore, unusual circumstances may pose severe problems in the recorded raw pictures, such as substantial light variations or background clutter. Using low-level visual characteristics may not be an acceptable option in this situation. In this scenario, the Internet of Things (IoT) is gaining traction, with various options for gathering high-level precise soil data that may be retrieved as very relevant characteristics to aid modern identification systems inefficiently identifying agricultural diseases in the field.

Due to geographical and climatic restrictions, the use of unimodal IoT sensing methods in crop disease detection may not achieve the needed accuracy and resilience [15]. One of the most commonly used sensor fusion algorithms in robotics applications such as position and orientation estimates, guided vehicles, and so on is the Kalman filter. It requires data from two sensors in a similar format as input [16]. In the problem under consideration, the sensors produce scalar values while image data is a 2D vector. As a result, in this application of fusion which includes 1D and 2D vectors, the Kalman filter cannot be employed [17]. With the progress and flexibility of AI frameworks, various AI algorithms are utilized to extract essential characteristics more efficiently and accurately, improving classification accuracy [18], [19].

Therefore, the primary motivation for developing a Multimodal Data Fusion model is to extract the features from two dissimilar modalities using two different frameworks. Then apply the proposed Rice-Fusion model to concatenate these extracted features to train the classifier to identify rice diseases. It becomes a robust solution as it utilizes the measurements from agro-meteorological sensors and camera images. It will increase true positives and reduce false positives. Fusion with another modality can assist in identifying the correct conclusion more effectively if one modality produces false negatives. If one modality produces false positives, the second modality can help reduce the aggregate accuracy of fused output, resulting in accurate predictions. The architecture of which is illustrated in Figure 7. Rice-Fusion uses CNN, and MLP approaches to extract robust visual and agro-meteorological characteristics from 3200 illness samples collected from the field. At last, a fully connected network is developed to combine image and numerical data to diagnose rice crop disease. This method might be an effective way to diagnose agricultural diseases in the rice field accurately. It could be utilized as a real-world application that will assist farmers in taking preventive and corrective measures against the disease.

Following are the most significant contributions of the paper:

- 1) A novel multimodal AI-based framework is suggested and demonstrated to merge two distinct modalities for more robust and reliable crop disease detection. This method makes use of continuous data to boost recognition performance.

- 2) The application of early fusion of the outputs of CNN and MLP architectures for rice disease detection has been demonstrated.
- 3) To validate the accuracy and applicability of the proposed model, we present a thorough experimental examination. The results demonstrate that the suggested multimodal framework outperforms the existing methods in diagnosing rice diseases.

The following sections make up the paper: Section 2 presents the available literature on plant disease diagnosis. A brief review of multi modal data fusion techniques based on AI is provided in Section 3. Section 4 discusses the data collecting and preparation frameworks and the proposed architecture of the system. Section 5 delves into the details of the results, while Section 6 concludes the paper along with prospects in the field.

II. RELATED WORK

A. ARTIFICIAL INTELLIGENCE FOR CROP DISEASE DETECTION

A model based on Support Vector Machine(SVM) [20] is proposed to classify three rice classes: blight, brown spot, and smut. The images were captured from a rice farm. It achieved an accuracy of 93.33% on the training dataset and 73.33% on the test dataset. To quantify the rice crop damage that occurred due to hopper infestation [21], Fuzzy C means classifier is used to classify four severity classes of infestation, namely severe, moderate, mild, and no infestation. It reached an accuracy of 87%. [22] investigated a technique that classifies and detects various types of mineral deficiencies in rice crops. The model was built for two types of different inputs, i.e., text and color, with different numbers of neurons in hidden layers. 88.56% accuracy is obtained. One more approach was proposed to identify blast and brown spot disease on rice. The Fractal Fourier method [23] is used to develop an approach that analyses the texture to identify diseases. The raw image is converted to CIELab space. The system achieved an accuracy of 92.5%. The gray-Level Co-occurrence Matrix (GLCM) technique is used to classify whether the rice is healthy or infected. Many researchers have focused on improving the accuracy and speed of identifying rice illnesses by using traditional methods such as pattern recognition techniques, support vector machines, digital image processing techniques, and computer vision. In a study [5], the infected rice pictures were classified using Self Organizing Map (SOM) in which the train images were produced by extracting the characteristics of the infected portions of the leaf while four other types of images were used for testing reasons. The researchers used a Neural network to simulate the results. The classification accuracy is enhanced with the classifier. [24] proposed a new stacked CNN architecture that uses two-stage training to reduce model size while maintaining good classification accuracy significantly. When stacked CNN was used instead of VGG16, the test accuracy was determined to be 95 percent.

B. INTERNET OF THINGS FOR CROP DISEASE DETECTION

The study [25] proposed a model that aims to create a monitoring system that uses a Hidden Markov Model to detect grape disease early and sends notifications to the farmer via SMS. Temperature, relative humidity, moisture, a leaf wetness sensor, and Zig-Bee for wireless data transmission are all part of the system. An accuracy of 90.9 is achieved. In a different study [26], the Goidanich model is used to predict powdery mildew fungal disease in a vine. The parameters used were temperature, moisture, and humidity. K-Nearest Neighbour (KNN) [27] technique is used to monitor the farm and climatic parameters on a daily basis to predict the outbreak of diseases and pests. The results were compared with other machine algorithms like Logistic Regression, Linear Regression, and Random Forest classifiers. Similarly, (Series, 2020) devised a method for predicting tomato plant health. The system included two sensors: a soil moisture sensor and a temperature-humidity sensor because abiotic parameters such as temperature, soil moisture, and humidity can assist detect whether the plant is developing in health conditions or not. The researchers examined two supervised learning algorithms (SVM and Random Forest) and an unsupervised learning strategy (K-means clustering). SVM, Random Forest, and K-means all had test accuracy of 99.3%, 99.6%, and 99.5 [28] proposed a Bi-LSTM model that is used to predict the occurrence of cotton pests and diseases based on climate variables. Bi-LSTM performs well in predicting the occurrence of pests and diseases in cotton fields, with an Area Under the Curve (AUC) of 0.95.

C. DATA FUSION FOR CROP DISEASES DETECTION

The various types of tools used for plant monitoring and specifically for disease diagnosis generate enormous data [29]. There are two alternatives to deal with this data; one way is to execute the execute individual modality and evaluate the validity of the method. The second way is to fuse the features collected from multiple sources related to crop diseases [30]. [31] developed a combined multi-input model based on satellite images and environmental data. Logistic regression was used to extract the features from both modalities. The accuracy was increased from 69% to 78%. A novel multi-context fusion network for crop disease detection is proposed where it exploits the contextual features to increase the performance. CNN is used as a backbone for attribute extraction, and the Bag of word approach is used for contextual information. The data set was collected for 19 different crops and 77 related categories. An accuracy of 97.5% is achieved [32]. However, it suffers from the issue of imbalanced data. In another approach [33], a customized classifier is developed that detects banana fruit diseases in African cultivation fields. Disease detection is a twofold process that classifies the diseases based on pixel values and later based on object detection. Support Vector Machine(SVM) is used to perform the classification. The input to the model was spectral images with vegetation indices and UAV images. RetinaNet is used to train the Object detection modal based on

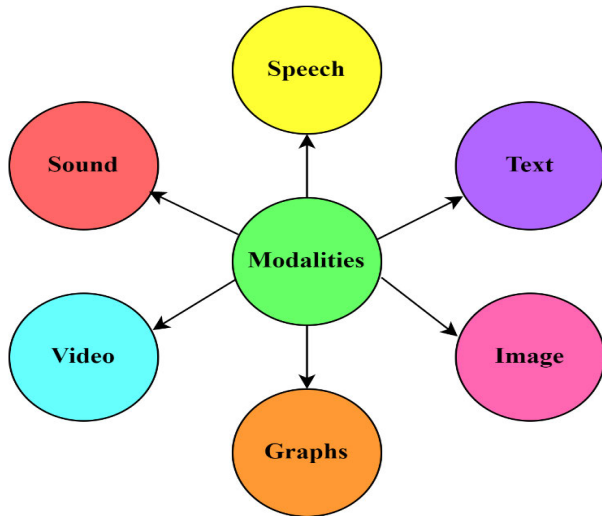


FIGURE 1. Different types of modalities.

UAV images. The proposed approach outperforms the VGG model by achieving an accuracy of 92%. However, the training time required to train the model is compensated.

It can be summarized that even though the above methodologies successfully recognize contemporary agricultural diseases, progress in this area has stagnated in recent years [43], [44]. Table 1 represents the comparative analysis of the state of the art techniques on rice disease diagnosis. Furthermore, due to many hurdles in realistic scenarios during test image inference, such as illumination and intricate background depicted, most of these approaches may not reach acceptable performance in practical crop disease recognition applications. As a result, a novel multimodal data fusion approach named Rice-Fusion is developed to solve difficulties in crop disease diagnosis tasks by combining agro-meteorological data to increase performance, inspired by deep learning breakthroughs in agriculture.

III. THEORETICAL BACKGROUND

When the data from multiple sources are combined, the system becomes more robust, fault-tolerant, and reliable than those that work with a single source. Many data fusion methods are available based on AI paradigms in the literature as a prelude to the proposed system. A brief discussion about these methods is represented in this section.

A. MULTIMODAL DATA FUSION APPROACHES

A modality is a kind of input data used to find a solution to the problem. Different types of inputs, such as image, video, sound, speech, text, graphs, etc., can be given to the model to perform the task. These inputs are called modalities. Multimodality learning means learning from multiple modalities. Figure 1 shows various types of modalities that can be inputted into the model to attain the goal.

Data is a collection of modalities. In AI, multimodality is when the same AI model processes two or more

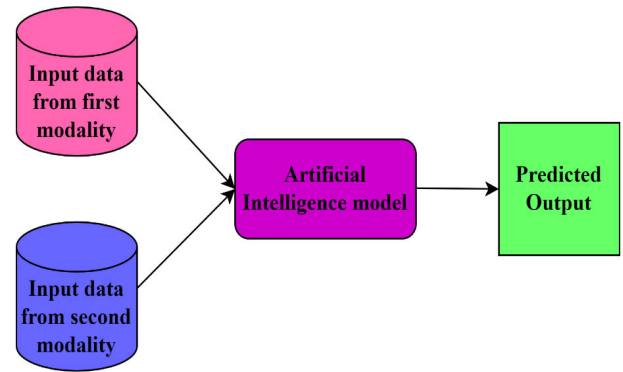


FIGURE 2. Multimodal data fusion concept.

heterogeneous inputs to solve the problem. Multiple data sources like real-time data from sensors, images of diseased leaves, weather data, etc., are often used to predict diseases in crops. These are a few of the modalities for crop disease identification. These sources contain information that complements the overall accuracy and performance of the model. The fusion of these modalities produces a more reliable, consistent, and accurate model to predict rice disease, thus reducing false positive and false negative percentages. However, there are challenges while using different modalities with different representations. It is not easy to fuse two representations in one model because they have different characteristics and dimensions. In addition to this, when the datasets are combined, they may have noisy and missing data. To solve the problem with multimodality, combine two separate models with a single modality at a higher level. Figure 2 conceptualizes the idea of multimodal fusion data.

There are three types of fusion models: Early, late, and hybrid fusion models. Early fusion combines the feature maps extracted from the multiple modalities [45]–[47]. When there is a high correlation between modalities, this is an appropriate technique to use. The features are extracted from independent modalities by using AI algorithms. The features obtained from the modalities are combined by the concatenation method. These fused features are again passed through an AI algorithm. The training of the model is done by using the fused features to obtain the final feature set. The features are fused before classification, where these features are learned together and assist each other in co-learning. Late fusion is a technique that trains unimodal based on specific classifiers. The notion of utilizing multiple inputs of independent types from multiple sources is referred to as multimodal fusion data.

The predictions from the individual modality are fused by using mathematical techniques such as mean, mode, median. It requires multiple training stages. It is known as Decision Fusion Technique as it is a fusion of the decisions predicted by individual modalities. The low-level interaction between the individual modalities is not modeled. The fusion mechanisms can be voting, weighted sum, or any AI model. When the

TABLE 1. Comparison with the state-of-the-art methods on rice crop disease detection methods.

Reference	Type of rice diseases	Features	Real time data / Publicly available dataset	Methodology	Performance Metric
[34]	Blast, brown spot, tungro and bacterial blight	5932 images	Publicly available data	ResNet50 with SVM	F1 score= 0.9438 Training time= 69seconds
[35]	Rice blast	26 pairs of infected and healthy images were collected per sampling day	Real time data collected from field	SVM-based SFFS	Overall Accuracy = 95% for mild infection stage
[36]	Leaf Smut, Bacterial Blight, Brown Spot	120 images (RGB)	UCI public dataset	Ensemble model of extreme gradient boosting and decision tree	Accuracy = 86.58%
[37]	Bacterial blight, brown spot and Smut	900 images	Publicly available dataset	AlexNet	Accuracy=0.9642
[38]	BLB, LB, BS and Sheath Rot	650 images	Collected from field	Author proposed Deep Neural Network_ Jaya Optimization Algorithm (DNN_JOA)	Mean Accuracy = 0.93
[39]	Brown Spot, Rice blast, Rice blight	600 images	Publicly available dataset	SVM, Naive Bayes, Decision Tree, KNN, MLP	SVM=91.3%, 90.7%, 88.7% Naive Bayes= 93%, 92%, 92% Decision Tree= 88.3%, 87.3%, 87% KNN =89%, 88%, 86% MLP =90%, 88.3%, 87.3%
[40]	Bacterial blight, blast, brown spot, Sheath rot	115 images	Publicly available dataset	MDC and KNN	MDC=89.23% KNN=87.02%
[41]	Blast, Brown spot, Blight, Tungro	40 images	Publicly available dataset	PNN	Accuracy=0.83
[20]	Blight, Leaf Smut, and Brown Spot	120 images (40 images of each type)	Publicly available dataset	SVM	Accuracy=0.7333
[21]	Rice plant hopper pest	NA	NA	Fuzzy C Means	Classification Accuracy =63.50% Infestation area detection accuracy =0.87
[22]	Brown spot and Blast	400 rice leaf images	Publicly available dataset	Multi Layer Perceptron	Accuracy=0.8926
[24]	6 disease classes, 2 rice pests classes and 1 healthy class	1426 images	Publicly a available dataset	Ensemble model based on VGG16 and InceptionV3	Accuracy=0.933
[42]	Rice Blast, False Smut, Bacterial Blight, and Brown Spot	Agro-meteorological parameters since 1989-2019	Data collected from Indian Meteorological Department	ANN	Accuracy=0.9215

modalities have a time correlation amongst them, this method is preferable. In hybrid fusion, the benefits of early fusion and late fusion techniques are combined. The feature set obtained after fusing the individual feature set modalities before classification is combined with decisions predicted

by the unimodal. Later they are again united to get a final decision.

The values of agro-meteorological attributes are continuous, so a simple MLP approach is best suitable for this modality. The second modality to the proposed framework is

rice crop images. CNN architecture proves to be excellent in extracting image features. Both the modalities have different characteristics and are not co-related with respect to time. Therefore, in the proposed framework, an early fusion type of multimodal data fusion technique is used to detect rice diseases based on the environmental features extracted from the MLP framework and image features extracted from the CNN framework. The following section discusses Data collection, data pre-processing. Also, it proposes a multimodal data fusion framework to diagnose rice diseases.

IV. MATERIALS AND METHODOLOGY

A. DATA COLLECTION

Rice disease diagnosis is a classification problem as the model will predict the name of the rice disease class based on image data and agro-meteorological data. Initially, categorical names of the classes are defined as Blight, Blast, Brown spot, and Healthy. The model learns to predict the label of the rice disease class based on the input features. The dataset is collected in two folds; first, the numerical values corresponding to the agro-meteorological sensors placed in the farm are collected. The second is simultaneously captured image of rice crop from the field. The data collection process is represented in Figure 3.

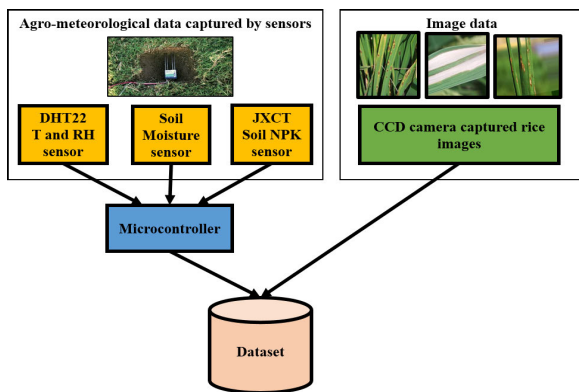


FIGURE 3. Data collection for the proposed framework.

The agro-meteorological dataset includes numerical values of environmental attributes like Temperature(T), Relative Humidity(RH), Soil moisture(M), and N-P-K soil nutrients values collected by sensors. These climatic normals play a vital role in diagnosing crop diseases [48]. DHT22 sensor is used to collect Temperature and Relative Humidity values. DHT22 is selected over DHT11 after doing trial and error. DHT22 more precisely captures the environmental values over DHT11 [49]. The resistive soil moisture sensor is used to collect moisture or water level values from the soil. The JXCT Soil NPK Sensor fetches the nitrogen, phosphorous, and potassium nutrient values from the earth. The NPK is connected to an Arduino microcontroller via Modbus RS485. The Arduino microcontroller is used to interface the sensors. The sketches are written using Arduino Integrated Development Environment. The numerical values corresponding to

the sensors are stored in a .csv file. All the sensors that are used to collect data have characteristics such as low cost, quick responsiveness, high precision, and portable.

Along with agro-meteorological parameters, the images for 3200 rice disease and healthy samples were simultaneously captured. This makes the dataset suitable for multimodal data fusion. The majority of the images were captured using Charged Coupled Device (CCD), a light sensor on an integrated chip [50]. The images are manually filtered to remove noise within the dataset to preserve the consistency of crop disease data and handle incorrect information. The blur and duplicate images are removed. The disease identification capabilities of our dataset-trained model could be significantly improved with these pre-processing techniques. There are no repeated data or missing data. All collected rice crop disease images are thoroughly reviewed by agricultural experts from various Research Extension Centers to ensure the authentication of the image annotations. A dataset in total consists of 3200 samples where each class of rice diseases classification consists of 800 samples. After sample collection, the dataset is divided with a 70-20-10% ratio as training, validation, and testing sets, respectively. The agro-meteorological sensors data is integrated with rice image data and is further used for training and testing phases of the newly developed multimodal data fusion model. Figure 4 represents the training and testing phases of the network.

The building process to develop an independent model is explained in the next subsections. Also, the Rice-Fusion framework that works on fused features and its variants are modeled in the following sub-sections.

B. MULTI-LAYER PERCEPTRON (MLP)

One of the input streams to the multimodal approach is continuous data from agro-climatic sensors. It consists of 3 or more layers to classify the non-linear data. The dataset collected is non-linearly separable. Hence, Multi Layer Perceptron is best suitable to work on continuous data. It is a supervised learning algorithm. It is a fully connected network where each node is connected to all nodes in the next layer. The building blocks of MLP are Neuron, neuron weights, Activation Function, Network of neurons.

1) NEURONS

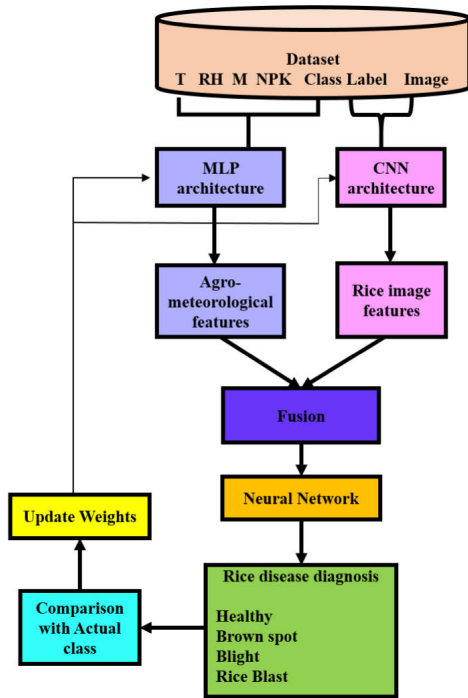
The input features and their corresponding weights are given as input to this computational unit, and it produces the result based on the Activation Function. The input vector is mathematically represented as

$$\text{Input} = [i_1, i_2 \dots, i_n] \quad (1)$$

2) NEURON WEIGHTS

This is the channel through all the inputs that have a weight associated with it. Each input will have a weight associated with it. The weights are smaller random values as it makes the network simpler. Equation 2 is the mathematical

Model Training Process



Model Testing Process

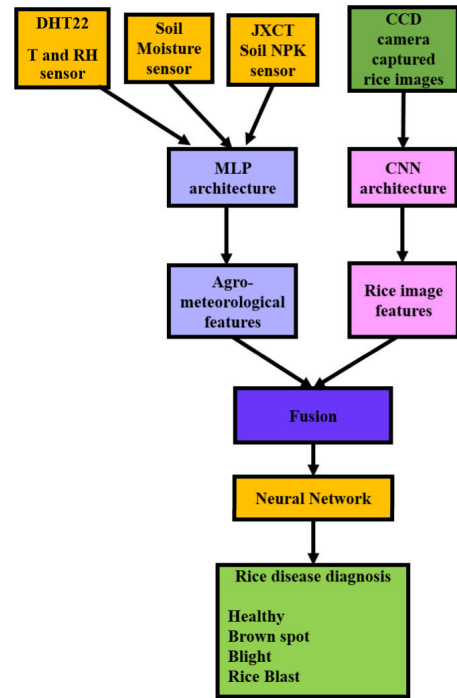


FIGURE 4. Training and testing process of the model.

representation of the weight matrix.

$$Weight = [w_1, w_2 \dots, w_n] \tag{2}$$

The Scalar product of ‘Input’ and ‘Weight’ is passed to the summation unit. The mathematical representation for this is shown in equation 3.

$$Result = [i_1, i_2 \dots, i_n] X [w_1, w_2 \dots, w_n] \tag{3}$$

3) SUMMATION

It multiplies the components of the input features with the weights and ultimately adds all of them. A bias ‘b’ is added to the output of the Summation circuit.

$$Sum(output) = \sum_0^n Result + bias \tag{4}$$

4) ACTIVATION FUNCTION

The result of the Summation circuit is passed to the Activation Function. It maps the summation of weighted input to the output of the neuron. It activates the neuron based on the threshold value. There various non-linear activation functions such as tanH, sigmoid, ReLu available [42]. However, this is a multi class classification problem. A softmax activation function is used. Error correction is done using backpropagation, where new weights are calculated and again passed through the network. This is computed using the following equation :

$$Weight (new) = Weight (old) + lr \times (Actual - Predicted) \tag{5}$$

This process continues till the error value between actual and predicted values become less. This depends on the batch learning rate ‘lr’ used.

5) NETWORKS OF NEURONS

A layer is a row of neurons, and a network can have several layers. The network topology refers to the architecture of the network. It comprises the input layer, hidden layer, and output layer. With this, the perceptron will get trained and perform the required task.

C. RICE DISEASE DIAGNOSIS USING AGRO-METEOROLOGICAL DATA: MLP ARCHITECTURE

Numpy, scikit-learn, pandas, Keras, and Open CV packages are installed. The dataset is loaded. The columns that have continuous data in them are defined. A sequential model is created by adding a series of layers. The 70-20-10% train validation test data split is used to build the model. The MLP architecture [42] is used to construct this model that consists of three layers. The network topology used to construct the model is 4-7-4 as the number of inputs is four parameters and 7 is the number of hidden layers, and 4 is the number of output classes. The dense class defines them with activation function applied is ReLu. In the computations of the hidden layer, the loss function must be specified to evaluate a set of weights, and the optimizer must be specified to search through different weights for the network. In this case, the loss function is categorical cross-entropy as the output is

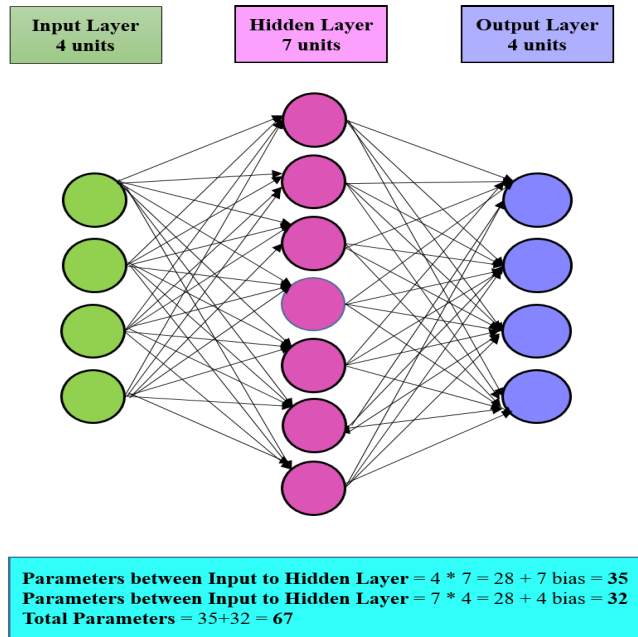


FIGURE 5. Unimodal MLP architecture to extract features from agro-meteorological data.

multiple classes of rice diseases, and the efficiently working optimizer is Adam. The optimized learning rate applied to the model was 0.01. Finally, the Softmax activation function is used to classify multi classes of rice infections. The accuracy, precision, and F1 score for rice disease classification are the metrics used. By invoking the model's fit () function, the model is trained or fit on the loaded data. Training takes place in epochs, with each epoch divided into batches. The batch size is 32, and the number of epochs considered is 500. The model is finally evaluated and achieves an accuracy of 91.25%. Figure 5 shows the MLP architecture network topology used to implement Rice disease diagnosis.

D. CONVOLUTIONAL NEURAL NETWORK (CNN)

The images of rice plants contain thousands of pixels that are stored in Red, Green, Blue (RGB) forms. The features in images are non-linear. CNN is the preferable approach to fetch out the complex features from the images. The basic building blocks to construct a CNN consists of kernel, stride, padding, pooling, and flattening [51]. In the proposed system, the building blocks are utilized to obtain an activation map for 3D images over 2D images.

1) KERNEL

It is a filter that extracts image characteristics. It is a tiny matrix that traverses the image input data, performs the scalar product with the particular cell of the input data, and returns the matrix of scalar products as an output. The stride value causes it to move over the input data. The dimensions of the activation map can be calculated using equation number 6.

$$K_{output} = [sizeof(input) - sizeof(kernel)] + 1 \quad (6)$$

2) STRIDE

The filter rasterizes the rice input image from left to right and top to bottom, changing the one-pixel column horizontally and one-pixel row vertically. Stride is the dimension of the movement matrix that is applied to the input image. The height and width dimension of the stride is symmetrical.

$$S_{output} = \left[\frac{sizeof(input) - sizeof(kernel)}{stride} \right] + 1 \quad (7)$$

3) PADDING

The method in which the number of pixels required by the convolutional kernel to process the edge pixels is added. The data is preserved without sacrificing essential attributes.

$$P_{out} = \left[\frac{sizeof(input) - sizeof(kernel) + (2padding)}{stride} \right] + 1 \quad (8)$$

4) POOLING

Pooling is required to down sample feature detection in feature maps by summarizing features in feature map patches. Average pooling and maximum pooling are two standard pooling methods. Max Pooling summarizes the average presence of a feature and discards unnecessary features.

5) FLATTENING

The feature map obtained from the pooling layer is a multi-dimensional vector. The process to convert this multi-dimensional vector to a 1D vector is called flattening. This 1D vector is then fed to the classifier to predict the rice diseases.

E. RICE DISEASE DIAGNOSIS USING IMAGE DATA: CNN ARCHITECTURE

The input image is a multidimensional matrix of rows, columns, and depth. The size of the input image given to the model is $200 * 200 * 3$. Initially, a CNN model is built using the build () method that takes four parameters: rows, columns, depth, and classes. The input image dimensions are $200 * 200 * 3 * 5$ as 3 is the layers of colors, and the model's output is classifying four classes of rice diseases. The model is defined as Sequential as the layers are added serially to the model. After the initialization of the model, the layers in CNN are added. By applying convolution filters, CNN automatically learns the features, the building blocks for rice diseases. The size of the convolution layer is $200 * 200 * 32$. The first convolution layer learns 32 features. The dimension of the input image is kept intact by adding padding bits. The number of convolution layers used is 2. The second convolution layer learns 64 features. ReLu activation function is activated, followed by the Max-pooling layer. The dimension of the image is reduced to half. The input image is of rice disease, represented as a vast matrix, and the kernel is a thin matrix that slides all over the image, starting from the top left. The kernel size used is $3 * 3$ so that a valid integer number comes at the center of the image. The dropout layer with a value of

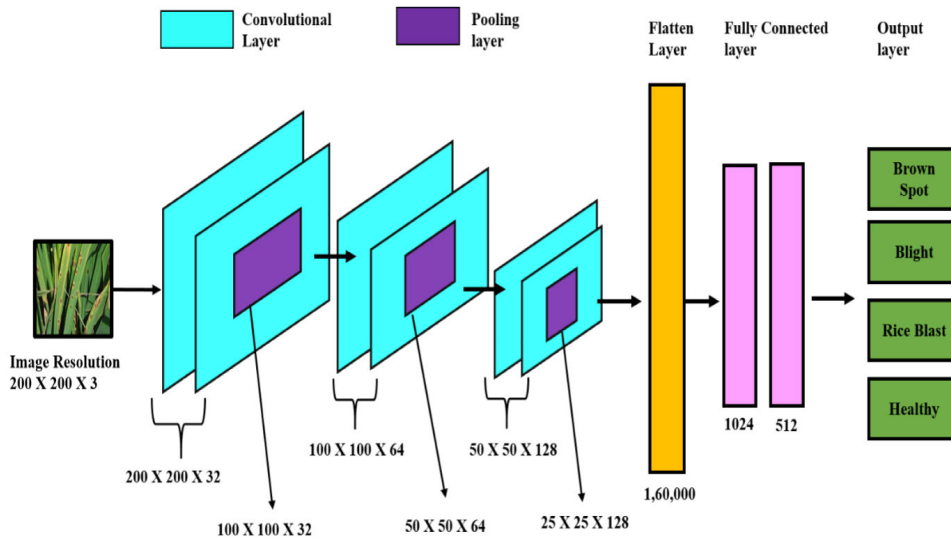


FIGURE 6. Unimodal CNN architecture to extract features from images.

0.2 is added after the max-pooling layer to avoid overfitting, which means some connections between layers are removed. The flattening layer is applied to the previous max-pooling layer output that is $25 \times 25 \times 128$. After flattening, this becomes a 1-dimensional array with 160000 values in a single vector. The number of dense layers is two, the number of neurons in the first fully connected layer is 1024, and the second layer is reduced by half, thus 512. The nonlinear ReLU function further activates this. After this, one more fully connected layer is added, equal to the number of rice disease classes the model is diagnosing, which is 4. This dense layer is then given as input to the Softmax classifier, giving the probabilities for each class type. The highest probability of class means that class is the diagnosis of a rice disease. Figure 6 shows the unimodal CNN architecture used to detect rice diseases on the basis of images.

A deep learning library Keras is used to develop a Rice image classifier. The training, validation, and testing dataset split is 70%, 20%, and 10%, respectively. Fit () is used to train the model. Adam optimizer is used for training the model. As it is a multi-class classification problem, categorical cross-entropy is the loss function used. The best scores of evaluation parameters like Accuracy, Precision, F1 score, and Recall are obtained using 500 epochs along with a learning rate of 0.01 with a batch size of 32 while training the model. The testing accuracy achieved is 82.03%.

F. RICE-FUSION: A MULTIMODAL DATA FUSION FRAMEWORK FOR RICE DISEASE DIAGNOSIS

In the paper, a novel model named R-Fusion is proposed to diagnose rice diseases accurately. R-Fusion is a Keras-based model that accepts inputs from different modalities like numeric or continuous data from sensors and rice image data.

A single network is trained on this multimodal data. The model needs to classify rice diseases accurately by accepting inputs from multiple modes, also called mixed data. These inputs are dissimilar from one another. Both the sources contain complementary information that improves the overall performance of R-Fusion in classifying the rice diseases based on these inputs.

1) BUILDING OF RICE-FUSION MODEL

The Rice-Fusion model building process is comprised of building two sub-models that are capable of handling independent data. The first sub model is constructed using the MLP model that handles agro-meteorological data, which is continuous in nature. The second sub model is building a CNN that operates on rice disease leaf image data. Once these two sub models are constructed, they are concatenated to form the multi modal Rice-Fusion model. Figure 7 shows the Rice Fusion model framework.

Initially, to load the agro-meteorological dataset, Pandas data frame is used. Later Rice image dataset is loaded and scaled to the range of 0 to 1. To convert the text into integer labels in the dataset Label Encoder() from the sklearn.preprocessing module is used. The dataset is split into training, validation, and testing sets with a 70:20:10% ratio, respectively. 70% data is used to train the model. 20% data is used to evaluate the model's performance, and the remaining 10% of test data is used to test the model's performance. Min-Max scaling is performed on continuous features, and on categorical features, one-hot encoding is applied. Later both the features are merged together.

Initially, MLP and CNN models are created. The outputs of MLP and CNN are concatenated using concatenate (). This merged output is now given as input to the final set

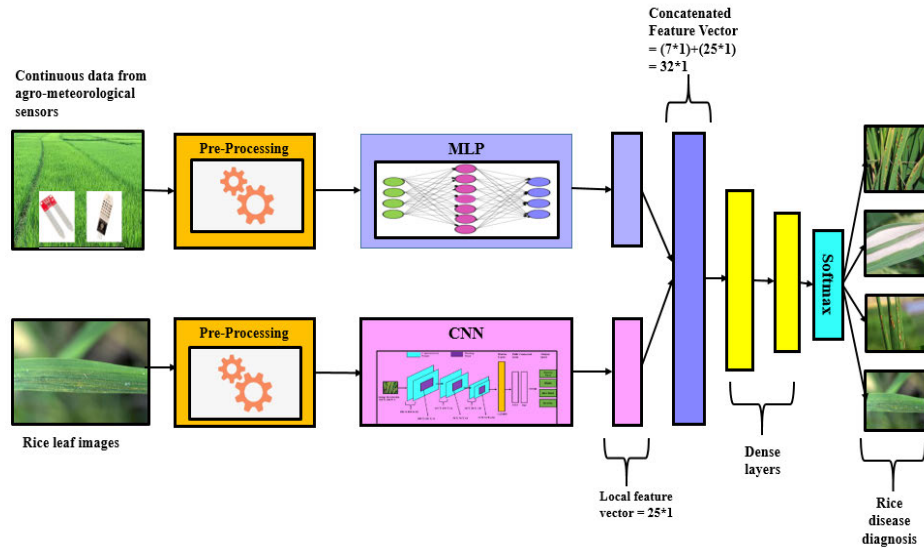


FIGURE 7. Multimodal rice-fusion architecture.

of layers as the output of both MLP and CNN. The model structure of the R-Fusion is based on the outputs of MLP and CNN outputs individually. There are two inputs to the Rice-Fusion model. The output of the MLP network is 4-7, and 200-100-50-25 is the output of CNN. The input that is given to the new Keras model is the feature vector from MLP and CNN architectures. The feature vector extracted from MLP is 7×1 vector, and the feature vector from CNN is 25×1 . These outputs are then concatenated, and a combined 1-dimensional vector of 32 is obtained. After this, two more fully connected layers are applied. The first layer has neurons 20, and the second layer has ten neurons. The nonlinear ReLU function further activates this.

The compilation is the last stage in the Rice-Fusion building process. Keras provides a method called compile () to compile the model. The three most vital arguments for the compile method are loss, optimizer, and metrics. The value of the loss function is set to cross-entropy. Optimizers play a critical role in improving the accuracy of R-Fusion. There are several optimizer options to choose from. The comparison of seven optimizers, namely Stochastic gradient descent (SGD), RMSprop, Adagrad, Adadelta, Adam, Adamax, Nadam, is performed. It can be summarised that Adam outperforms all the optimizers considered for experimental analysis; therefore, Adam is chosen as the best optimizer to perform experimental work as it expedites the model training and minimizes the computational costs. The learning rate applied is 0.001 with a decay of 1×10^{-3} . The performance of Rice-Fusion is evaluated by using accuracy as a metric. After successful compilation of Rice-Fusion, the next step is to train it. Rice-Fusion is trained by using the fit() function and evaluates the performance on the training dataset. The experimental work is carried out for 500 epochs, and the

batch size used is 32. Backpropagation is used to fine-tune all of the weights. After this, one more fully connected layer is added, equal to the number of rice disease classes the model is diagnosing. The last layer will have four neurons as the final output classes “Healthy,” “Bacterial Blight,” “BrownSpot,” “Sheath Blight”. The activation function used to classify the four rice disease classes is Softmax.

G. VARIANTS OF RICE-FUSION FRAMEWORK: MAX RICE-FUSION AND AVERAGE RICE-FUSION

The two variants of the Rice-Fusion model are proposed in this subsection, namely Average Rice-Fusion and Max Rice-Fusion. The framework for Average Rice-Fusion architecture is depicted in Figure 8. The average Rice-Fusion method, in contrast to the Rice-Fusion model, trains individual classifiers for individual modalities. The prediction from individual modalities, i.e., predictions from MLP and CNN classifiers, are obtained. Later these predictions are concatenated, and an average of both predictions is calculated. Finally, the model gives one classification category based on the fused features.

$$\begin{aligned} \text{AverageRiceFusion}(\text{output}) \\ = \left[\frac{\text{Output}(\text{MLP}) + \text{Output}(\text{CNN})}{2} \right] \quad (9) \end{aligned}$$

The average fusion model faces specific challenges, such as high computational time, as it must train two different classifiers independently.

The Max Rice-Fusion model outputs the classification result as the output of the classifier, which has produced a higher accurate prediction. The classifier which classifies more accurately, the result generated from that classifier is the final output. Figure 9 represents the overview of

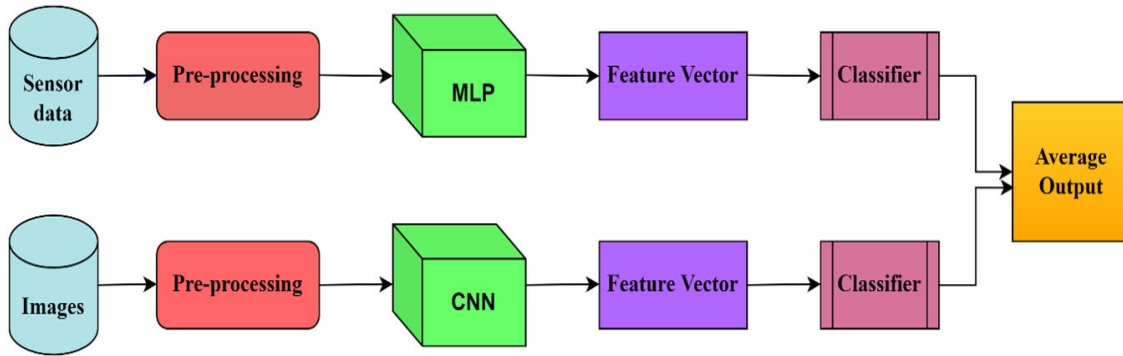


FIGURE 8. Average rice-fusion framework.

Max Rice-Fusion framework.

MaxRiceFusion (output)

$$= \max [Output (MLP) , Output (CNN)] \quad (10)$$

All three models were implemented, and comparative analysis represents that the Rice-Fusion model is the best approach to classify rice diseases. The next section discusses the result analysis.

V. RESULTS AND DISCUSSION

A novel framework based on the concept of multimodal data fusion is proposed in this paper to diagnose rice crop diseases and healthy rice crops. The two diverse modalities are considered for the proposed work, namely agro-meteorological attributes and images of rice crops. The CNN architecture with two dense layers is used to extract the features from the images, whereas MLP is used to extract the features from agro-meteorological data. As rice disease classification involves images as well as numeric data, it consumes huge memory. So Google Colab, an open-source GPU, is used to perform all the experimentation of the proposed model. Python 3.7 programming language is used for the implementation. Table 2 shows the hardware specification of GPU and CPU used for training and testing of the model. The MLP model starts converging at the 40th epoch, and the CNN model stabilizes itself approximately at the 120th epoch. The fused model stabilizes itself approximately at around the 40th epoch. The performance of the individual models is less as compared to fusion models.

The unimodal modalities help each other co-learn the features and learn whether the features complement and support to classify diseases more accurately. Thus the system obtained is more consistent, reliable, and accurate to diagnose the rice disease. The independent models, namely CNN and MLP, have achieved an accuracy of 82.03% and 91.25%, respectively, when tested again on rice disease data. In the proposed system, the features from both the unimodal models are concatenated, and its accuracy is 95.31%, which is greater than that of independent models. The other variants of Rice

TABLE 2. Hardware specification of proposed framework.

GPU Hardware Specification	
GPU	Nvidia K80
GPU Memory	12GB
Performance	4.1 TFLOPS
No. CPU Cores	2
Available RAM	12GB
Disk Space	358GB
CPU Hardware Specification	
CPU	Intel(R) Xeon(R)
CPU Frequency	2.30GHz
CPU Family	Haswell
No. CPU Cores	2
Available RAM	12GB
Disk Space	25GB

Fusion also achieved good accuracy when compared with individual models.

A. PERFORMANCE EVALUATION METRICS

The datasets used are skewed in nature. The classifier should not take advantage of its skewness. The following set of metrics, such as Confusion Matrix, Precision, Recall, F1 score, Specificity, Negative Predictive Value (NPR), False Positive Rate (FPR), False Negative Rate (FNR), Matthews Correlation Coefficient (MCC), and Accuracy, are used to understand whether the designed classifier is taking advantage of data skewness. This helps to understand the performance of the classifier.

B. CONFUSION MATRIX

The confusion matrix for two unimodal architectures and three variants of multimodal data fusion architectures are calculated and is represented in Figure 10. This parameter basically tells that how many times the designed classifier got confused. This is a matrix comprised of rows and columns. The rows are an actual count of the rice diseases, and the column has a predicted count of rice diseases. The predicted count is predicted by the classifier. The classifier is said to be best if it has only true positives and true negatives. The diagonal values in the Confusion matrix should be non-zero.

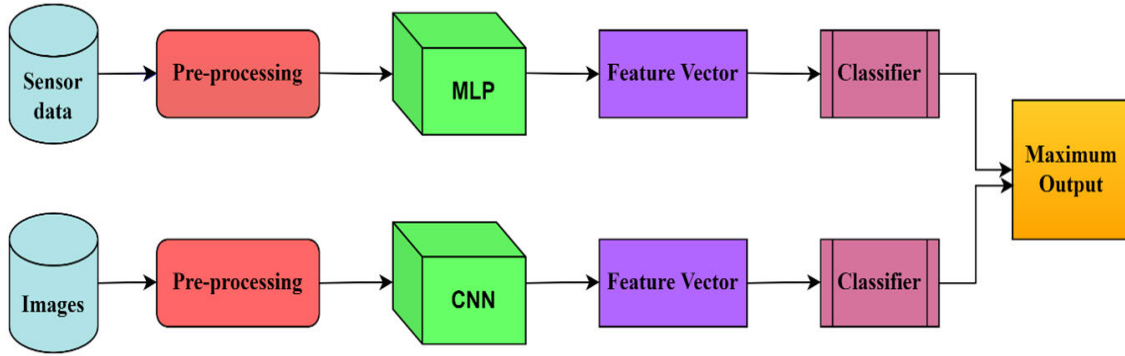


FIGURE 9. Max rice-fusion framework.

Ideally, other values should be zero, which means that the classifier designed is strong. After comparing the confusion matrix of multimodal fusion models and individual models, it is prominent that the percentage of false positives and false negatives are lowered with multimodal models. Thus it can be said that the multimodal data fusion models outperform the independent models. Out of the Brown spot images in the testing set, 14 images have been misclassified as blight disease. The reason for misclassification would be similar geometrical characteristics amongst the rice disease classes. However, it is observed that Rice Fusion architecture is obtaining excellent performance in identifying other classes of rice disease, and the accuracy for most of the rice disease is above 91%. Therefore, it can be summarized that the proposed R-Fusion framework is the best approach to identify rice diseases as there are minimum inaccuracies in identifying rice diseases. The main factor for excellent performance is the usage of environmental attributes and image data, which reduces the possibility of creating confusion for the model.

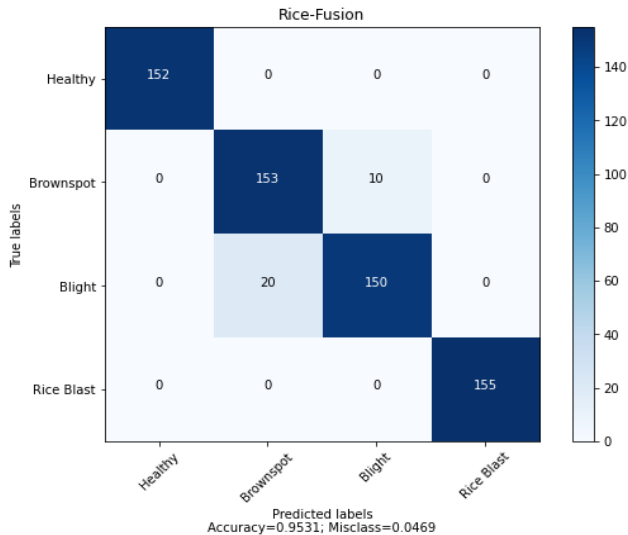


FIGURE 10. Confusion matrix for proposed multimodal rice-fusion framework.

C. PRECISION, RECALL, F1 SCORE, ACCURACY

Once the confusion matrix is calculated, calculating Precision, Recall and F1 score values is easy. Precision is the percentage of correct positive predictions made by the proposed model. The recall is the percentage of classifying the actual positive values classified by the classifier out of the total positive values. It is convenient to combine Precision and Recall, and the metric obtained after combining both metrics is the F1 score.

D. SUPPLEMENTARY PERFORMANCE METRICS FOR RICE DISEASE DIAGNOSIS

Specificity is the proportion of healthy incidences of crop that tested negative when compared to the total number of incidences without the disease. It identifies incidences that do not have a disease. If the value of specificity is higher then the model has the highest capacity to identify the incidences. The test that has specificity value as 1 will identify all of the incidences that do not have the disease. The Negative Predictive Value (NPV) is the probability that the rice disease is not present when the actual result is negative. The False Positive Rate (FPR) is the proportion of negative cases in the data that were incorrectly identified as positive cases. If the false positives rate is 0.0 then the model performs best. $1 - \text{Specificity}$ is another way to calculate it. The False Negative Rate (FNR) is the probability that a diseased incidence is missed by the classifier. Matthew’s Correlation Coefficient (MCC) is a parameter to measure model performance. It calculates the disparity between actual and expected values. True negatives, true positives, false negatives, and false positives are all factored into the coefficient. Only if the forecast delivers good rates for all four of these areas does this trustworthy metric produce high scores. MCC is a number that runs from -1 to 1, with 1 representing the best agreement between actuals and projections and zero representing no agreement at all. Table 3 represents the quantitative comparison of unimodal and multimodal fusion techniques based on various performance metrics Table 4 shows that Rice-Fusion when compared to other unimodal and multimodal architectures, outputs higher F1 scores.

TABLE 3. Quantitative comparison of unimodal and fusion models.

Model	Class	Precision	Recall	F1 Score	Specificity	NPV	FPR	FNR	MCC
MLP	Healthy	0.99	1.0	0.99	0.99	1.0	0.002	0.00	0.99
	Brown spot	0.85	0.85	0.85	0.94	0.94	0.05	0.14	0.79
	Blight	0.85	0.84	0.85	0.94	0.94	0.05	0.15	0.79
	Rice Blast	0.96	0.96	0.96	0.98	0.98	0.01	0.03	0.95
CNN	Healthy	0.97	0.98	0.97	0.99	0.99	0.008	0.01	0.97
	Brown spot	0.84	0.64	0.73	0.93	0.82	0.06	0.35	0.62
	Blight	0.48	0.72	0.57	0.84	0.94	0.15	0.27	0.49
	Rice Blast	0.98	0.98	0.98	0.99	0.99	0.006	0.019	0.97
Rice-Fusion	Healthy	1.0	1.0	1.0	1.0	1.0	0.0	0.0	1.0
	Brown spot	0.94	0.88	0.91	0.97	0.95	0.02	0.11	0.87
	Blight	0.88	0.94	0.91	0.95	0.97	0.04	0.06	0.87
	Rice Blast	1.0	1.0	1.0	1.0	1.0	0.0	0.0	1.0
Max Rice-Fusion	Healthy	1.0	0.86	0.92	1.0	0.95	0.0	0.13	0.90
	Brown spot	0.93	0.92	0.91	0.97	0.98	0.02	0.03	0.93
	Blight	0.88	0.93	0.90	0.95	0.97	0.04	0.06	0.87
	Rice Blast	0.94	0.99	0.97	0.98	0.99	0.016	0.006	0.96
Average Rice-Fusion	Healthy	1.0	0.86	0.92	0.86	0.95	0.0	0.13	0.90
	Brown spot	0.91	0.12	0.91	0.96	0.96	0.03	0.08	0.88
	Blight	0.82	0.93	0.87	0.94	0.97	0.05	0.06	0.83
	Rice Blast	0.94	0.96	0.95	0.98	0.98	0.01	0.03	0.94

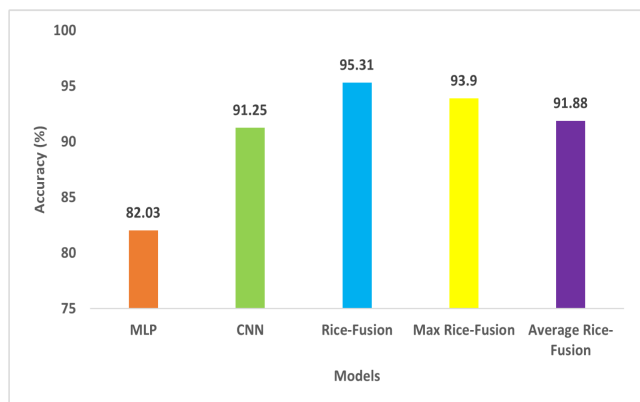


FIGURE 11. Accuracy comparison of different unimodal and multimodal data fusion models.

The overall accuracy scores for unimodal and multimodal models are represented in Figure 11. It states that all the variants of fusion models outperform individual models as the classification is done on the basis of both modalities. Thus, the fused model is stronger, consistent, reliable, and fault-tolerant to perform the rice disease classification task. The features from both modalities are combined before classification, which makes Rice Fusion architecture most accurate for rice disease classification.

E. LOSS ANALYSIS OF RICE-FUSION MODEL

The minimum the value of loss function less is the number of errors in the model. The model aims to have a minimum loss of function. The model computes for 500 epochs, and it can be observed that the model starts converging from the 218th epoch over training data. The total loss is approximately 0.1. The learning rate of 0.01 is set to minimize the loss in the proposed approach. Figure 12 shows the classification loss of the proposed model.

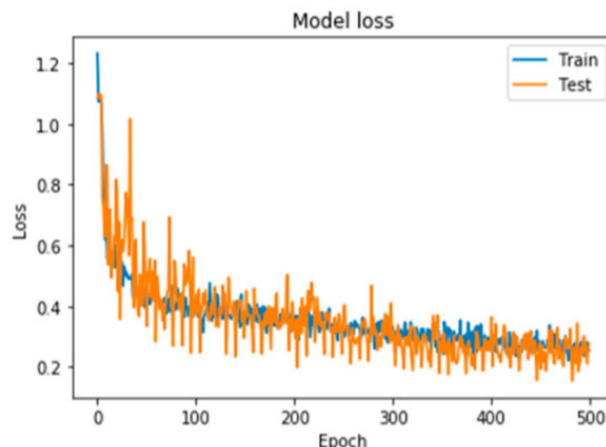


FIGURE 12. Classification loss of rice-fusion model.

F. COMPARISON OF RICE-FUSION FRAMEWORK WITH THE OTHER RELATED MULTIMODAL DATA FUSION FRAMEWORKS

Table 3 compares the Rice-Fusion framework with other existing multimodal data fusion models for crop disease classification in the literature. The existing studies use datasets captured by them or the datasets available in the public domain for validating their models. The dataset is collected in real-time, and the model is validated against this dataset. As rice is considered a staple food all over the world, it needs to be conserved. Very little research has been done on diagnosing multiple types of rice diseases using the multimodal data fusion approach. Keeping this in mind, the proposed model focuses on developing a model based on multimodal fusion that will complement independent modalities to increase classification accuracy.

The overall accuracy achieved by [32] is 97.5% which is slightly higher than that of the proposed model, as the authors have focussed on classifying diseases related to

TABLE 4. Comparison of rice-fusion framework with the other related multimodal data fusion frameworks.

Ref.	Crop	Disease types	Features	Methods	Performance
[33]	Banana	Banana bunchy top, Wilt	Satellite images and UAV images	UAV-RGB mixed model	Overall Accuracy = 90.8%
[32]	19 different crops	77 disease classes related to 19 crops	Real time field images and contextual information	Multi Context Fusion Network	Overall Accuracy = 97.5%
[7]	Wheat	Powdery Mildew	Climatic data and satellite images	Logistic Regression	Overall Accuracy = 78%
[30]	-	-	Vegetation indices and environmental features	FLDA model	Accuracy = 82%
Proposed framework-Rice-fusion	Rice	Blight, Brownspot, Blast and Healthy	Agro-meteorological parameters and rice diseased and healthy images	Multimodal Early Fusion Method	Overall Accuracy = 95.31%

multiple crops. The authors have used the dataset that comprises 50,000 images as the number of images is higher accuracy increases. The proposed model focuses only on different rice diseases. Hence slight disparities are tolerable. Thus, the proposed model is the best fit for rice disease diagnosis as it is based on a fusion approach.

G. LIMITATIONS OF THE PROPOSED WORK

Despite the fact that the paper proposes a unique Rice-Fusion model, a multimodal data fusion technique for rice crop disease diagnosis, and achieves good results for the dataset, it has significant limitations. The following are some of the study's limitations, along with a possible remedy to these issues:

- When image modality is considered, the limitation of rice disease misclassification can occur as there is a similarity between the geometrical features of the rice diseases. Hence to overcome this barrier, more image datasets corresponding to the environmental dataset with similar geometrical properties should be necessary to train the network. It is also recommended to use a deep learning approach that can efficiently classify rice diseases even with tiny feature dissimilarities.
- The issue of an imbalanced dataset has not been adequately addressed. It is very hard to obtain a balanced dataset for diverse types of rice diseases because of climatic and geographical difficulties. Furthermore, the incidence frequency of agricultural diseases in real-time applications may differ. The proposed model is able to diagnose the multiple rice disease categories; however, the estimation of disease severity is not given attention as this is also a major step in an integrated rice disease management system.
- Rice Fusion model could identify crop diseases in various real-time conditions despite certain insurmountable uncertainties such as illumination and noisy background. The overfitting problem is predicted to occur during the training stage due to a lack of appropriate photos with varied agro-meteorological circumstances.

The techniques such as Generative Adversarial Networks (GANs) might be used to address the problem.

- As two different models are executed parallelly for two different modalities, the performance of the Rice Fusion model depends on the performance of the two models running in earlier passes. The expected performance can be achieved by training the model intensively using appropriate datasets.

VI. CONCLUSION AND FUTURE DIRECTIONS

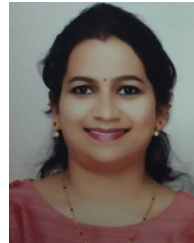
The proposed Rice-Fusion framework is an AI-based multimodal data fusion model in the agricultural domain used to diagnose different rice diseases automatically. The three classes of infections such as Brown spot, Rice blast, Bacterial blight, and one type for the healthy category is considered for the study. The data collected is unique as it has 3600 samples of both modalities, images, and environmental attributes. These two modalities are fused using fusion models such as Early and Late Fusion approaches. This work will contribute towards the agriculture field as an emerging technology that will assist farmers in the decision-making process related to rice crop diseases. The decision provided by the model is a combination of two modalities that enhance the performance and robustness of the system. The Rice-Fusion data fusion architecture outperforms the unimodal data models by co-learning, complementing, or opposing the unimodal data models. The system is more reliable and fault-tolerant. The system is built upon deep learning neural network approach and therefore requires a large number of data samples for appropriate training of the model. The comparative analysis made in the paper shows that the Rice-Fusion data fusion approach achieves an accuracy of 95.31% and thus outperforms other unimodal approaches such as CNN and MLP architectures. The future work would focus on collecting balanced datasets for diverse types of rice diseases because of climatic and geographical difficulties. The proposed study's results are very encouraging in diagnosing healthy and infected leaves with various rice diseases in real

time conditions. However, further, research can focus on segmenting the infected portions of the leaf images. In addition to this, the severity level of the diseases can be quantified and providing fertilizer recommendations based on the type, and the severity of the disease can be interesting future work. Even though the system works on real-time data, it is partially automated. So this can be one of the future contributions to the work. As a result, more research should be conducted to implement an automatic and robust system that the farmers can utilize to detect rice diseases. This system could include applications based on agricultural sensors, which could help to modernize the agricultural industry. Such an integrated management system will help the farmers retain the crop's natural quality, making it more organic.

REFERENCES

- [1] E. C. Oerke, "Crop losses to pests," *J. Agricult. Sci.*, vol. 144, no. 1, pp. 31–43, 2006.
- [2] M. Asif and A. Kamran, "Crop stress and its management: Perspectives and strategies," *Crop Sci.*, vol. 52, no. 4, pp. 1968–1969, Jul. 2012.
- [3] R. Patil and S. Kumar, "A bibliometric survey on the diagnosis of plant leaf diseases using artificial intelligence," *Library Philosophy Pract.*, vol. 3987, pp. 1–25, Aug. 2020.
- [4] X. E. Pantazi, D. Moshou, and A. A. Tamouridou, "Automated leaf disease detection in different crop species through image features analysis and one class classifiers," *Comput. Electron. Agricult.*, vol. 156, pp. 96–104, Jan. 2018, doi: [10.1016/j.compag.2018.11.005](https://doi.org/10.1016/j.compag.2018.11.005).
- [5] S. Phadikar and J. Sil, "Rice disease identification using pattern recognition techniques," in *Proc. 11th Int. Conf. Comput. Inf. Technol. (ICCIT)*, Dec. 2008, pp. 420–423.
- [6] J. G. A. Barbedo, C. S. Tibola, and J. M. C. Fernandes, "Detecting fusarium head blight in wheat kernels using hyperspectral imaging," *Biosyst. Eng.*, vol. 131, pp. 65–76, Mar. 2015.
- [7] J. Zhang, R. Pu, L. Yuan, W. Huang, C. Nie, and G. Yang, "Integrating remotely sensed and meteorological observations to forecast wheat powdery mildew at a regional scale," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 11, pp. 4328–4339, Nov. 2014.
- [8] R. Zhou, S. Kaneko, F. Tanaka, M. Kayamori, and M. Shimizu, "Image-based field monitoring of *Cercospora* leaf spot in sugar beet by robust template matching and pattern recognition," *Comput. Electron. Agricult.*, vol. 116, pp. 65–79, Aug. 2015, doi: [10.1016/j.compag.2015.05.020](https://doi.org/10.1016/j.compag.2015.05.020).
- [9] M. Nagaraju and P. Chawla, "Systematic review of deep learning techniques in plant disease detection," *Int. J. Syst. Assurance Eng. Manage.*, vol. 11, no. 3, pp. 547–560, Jun. 2020, doi: [10.1007/s13198-020-00972-1](https://doi.org/10.1007/s13198-020-00972-1).
- [10] S. S. Zarbafi and J. H. Ham, "An overview of rice QTLs associated with disease resistance to three major rice diseases: Blast, sheath blight, and bacterial panicle blight," *Agronomy*, vol. 9, no. 4, pp. 1–41, 2019.
- [11] D. J. Lary, A. H. Alavi, A. H. Gandomi, and A. L. Walker, "Machine learning in geosciences and remote sensing," *Geosci. Frontiers*, vol. 7, no. 1, pp. 3–10, 2016, doi: [10.1016/j.gsf.2015.07.003](https://doi.org/10.1016/j.gsf.2015.07.003).
- [12] K. Golhani, S. K. Balasundram, G. Vadamalai, and B. Pradhan, "A review of neural networks in plant disease detection using hyperspectral data," *Inf. Process. Agricult.*, vol. 5, no. 3, pp. 354–371, 2018, doi: [10.1016/j.inpa.2018.05.002](https://doi.org/10.1016/j.inpa.2018.05.002).
- [13] W. Li, Z. Wang, G. Wei, L. Ma, J. Hu, and D. Ding, "A survey on multisensor fusion and consensus filtering for sensor networks," *Discrete Dyn. Nature Soc.*, vol. 2015, pp. 1–12, Oct. 2015.
- [14] S. Kumar, R. R. Patil, V. Kumawat, Y. Rai, N. Krishnan, and S. K. Singh, "A bibliometric analysis of plant disease classification with artificial intelligence using convolutional neural network," *Library Philosophy Pract.*, vol. 5777, pp. 1–14, May 2021.
- [15] W. Elmenreich, "A review on system architectures for sensor fusion applications," in *Proc. IFIP Int. Workshop Softw. Technol. Embedded Ubiquitous Syst.*, in Lecture Notes in Computer Science: Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics, vol. 4761, Jun. 2007, pp. 547–559.
- [16] A. M. Wolf, "Means and objectives in multiratio gearing," SAE Tech. Papers 490207, 1949.
- [17] P. Narkhede, R. Walambe, S. Mandaokar, P. Chandel, K. Kotecha, and G. Ghinea, "Gas detection and identification using multimodal artificial intelligence based sensor fusion," *Appl. Syst. Innov.*, vol. 4, no. 1, pp. 1–14, 2021.
- [18] Y. Luo, W. Ye, X. Zhao, X. Pan, and Y. Cao, "Classification of data from electronic nose using gradient tree boosting algorithm," *Sensors*, vol. 17, no. 10, pp. 1–10, 2017.
- [19] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. 3rd Int. Conf. Learn. Represent. (ICLR)*, 2015, pp. 1–14.
- [20] H. B. Prajapati, J. P. Shah, and V. K. Dabhi, "Detection and classification of rice plant diseases," *Intell. Decis. Technol.*, vol. 11, no. 3, pp. 357–373, Aug. 2017.
- [21] Z. Zhou, Y. Zang, Y. Li, Y. Zhang, P. Wang, and X. Luo, "Rice plant-hopper infestation detection and classification algorithms based on fractal dimension values and fuzzy C-means," *Math. Comput. Model.*, vol. 58, nos. 3–4, pp. 701–709, Aug. 2013, doi: [10.1016/j.mcm.2011.10.028](https://doi.org/10.1016/j.mcm.2011.10.028).
- [22] P. Sanyal and S. C. Patel, "Pattern recognition method to detect two diseases in rice plants," *Imag. Sci. J.*, vol. 56, no. 6, pp. 319–325, Dec. 2008.
- [23] A. Asfarian, Y. Herdiyeni, A. Rauf, and K. H. Mutaqin, "A computer vision for rice disease identification to support integrated pest management," *Crop Protection*, vol. 61, pp. 103–104, Jul. 2014, doi: [10.1016/j.cropro.2013.12.044](https://doi.org/10.1016/j.cropro.2013.12.044).
- [24] C. R. Rahman, P. S. Arko, M. E. Ali, M. A. I. Khan, S. H. Apon, F. Nowrin, and A. Wasif, "Identification and recognition of rice diseases and pests using convolutional neural networks," *Biosyst. Eng.*, vol. 194, pp. 112–120, Jun. 2020.
- [25] S. S. Patil and S. A. Thorat, "Early detection of grapes diseases using machine learning and IoT," in *Proc. 2nd Int. Conf. Cogn. Comput. Inf. Process. (CCIP)*, Aug. 2016.
- [26] S. Trilles, J. Torres-Sospedra, Ó. Belmonte, F. J. Zarazaga-Soria, A. González-Pérez, and J. Huerta, "Development of an open sensorized platform in a smart agriculture context: A vineyard support system for monitoring mildew disease," *Sustain. Comput., Informat. Syst.*, vol. 28, Dec. 2020, Art. no. 100309.
- [27] N. Materne and M. Inoue, "IoT monitoring system for early detection of agricultural pests and diseases," in *Proc. 12th SEATUC Symp.*, Mar. 2018, pp. 1–5.
- [28] P. Chen, Q. Xiao, J. Zhang, C. Xie, and B. Wang, "Occurrence prediction of cotton pests and diseases by bidirectional long short-term memory networks with climate and atmosphere circulation," *Comput. Electron. Agricult.*, vol. 176, Sep. 2020, Art. no. 105612, doi: [10.1016/j.compag.2020.105612](https://doi.org/10.1016/j.compag.2020.105612).
- [29] M. Ouhami, A. Hafiane, Y. Es-Saady, M. E. Hajji, and R. Canals, "Computer vision, IoT and data fusion for crop disease detection using machine learning: A survey and ongoing research," *Remote Sens.*, vol. 13, no. 13, p. 2486, Jun. 2021.
- [30] L. Yuan, Z. Bao, H. Zhang, Y. Zhang, and X. Liang, "Habitat monitoring to evaluate crop disease and pest distributions based on multi-source satellite remote sensing imagery," *Optik*, vol. 145, pp. 66–73, Sep. 2017, doi: [10.1016/j.ijleo.2017.06.071](https://doi.org/10.1016/j.ijleo.2017.06.071).
- [31] J. Zhang, L. Yuan, R. Pu, R. W. Loraamm, G. Yang, and J. Wang, "Comparison between wavelet spectral features and conventional spectral features in detecting yellow rust for winter wheat," *Comput. Electron. Agricult.*, vol. 100, pp. 79–87, Jan. 2014, doi: [10.1016/j.compag.2013.11.001](https://doi.org/10.1016/j.compag.2013.11.001).
- [32] Y. Zhao, L. Liu, C. Xie, R. Wang, F. Wang, Y. Bu, and S. Zhang, "An effective automatic system deployed in agricultural Internet of Things using multi-context fusion network towards crop disease recognition in the wild," *Appl. Soft Comput.*, vol. 89, Apr. 2020, Art. no. 106128, doi: [10.1016/j.asoc.2020.106128](https://doi.org/10.1016/j.asoc.2020.106128).
- [33] M. G. Selvaraj, A. Vergara, F. Montenegro, H. A. Ruiz, N. Safari, D. Raymaekers, W. Ocimati, J. Ntamwira, L. Tits, A. B. Omondi, and G. Blomme, "Detection of banana plants and their major diseases through aerial images and machine learning methods: A case study in DR Congo and Republic of Benin," *ISPRS J. Photogramm. Remote Sens.*, vol. 169, pp. 110–124, Nov. 2020, doi: [10.1016/j.isprsjprs.2020.08.025](https://doi.org/10.1016/j.isprsjprs.2020.08.025).
- [34] P. K. Sathy, N. K. Barpanda, A. K. Rath, and S. K. Behera, "Deep feature based rice leaf disease identification using support vector machine," *Comput. Electron. Agricult.*, vol. 175, Aug. 2020, Art. no. 105527, doi: [10.1016/j.compag.2020.105527](https://doi.org/10.1016/j.compag.2020.105527).
- [35] L. Tian, B. Xue, Z. Wang, D. Li, X. Yao, Q. Cao, Y. Zhu, W. Cao, and T. Cheng, "Spectroscopic detection of rice leaf blast infection from asymptomatic to mild stages with integrated machine learning and feature selection," *Remote Sens. Environ.*, vol. 257, May 2021, Art. no. 112350.

- [36] M. A. Azim, M. K. Islam, M. M. Rahman, and F. Jahan, "An effective feature extraction method for rice leaf disease classification," *Telkomnika, Telecommun. Comput. Electron. Control*, vol. 19, no. 2, pp. 463–470, 2021.
- [37] M. M. H. Matin, A. Khatun, M. G. Moazzam, and M. S. Uddin, "An efficient disease detection technique of rice leaf using AlexNet," *J. Comput. Commun.*, vol. 8, no. 12, pp. 49–57, 2020.
- [38] S. Ramesh and D. Vydeki, "Recognition and classification of paddy leaf diseases using optimized deep neural network with Jaya algorithm," *Inf. Process. Agricult.*, vol. 7, no. 2, pp. 249–260, Jun. 2020, doi: 10.1016/j.inpa.2019.09.002.
- [39] G. Yasmin, A. K. Das, and A. Ghosal, "A hierarchical stratagem for rice leaf disease distinction," in *Proc. Int. Conf. Intell. Comput., Instrum. Control Technol. (ICICICT)*, Jul. 2017, pp. 1177–1183.
- [40] A. A. Joshi and B. D. Jadhav, "Monitoring and controlling rice diseases using image processing techniques," in *Proc. Int. Conf. Comput., Anal. Secur. Trends (CAST)*, Dec. 2016, pp. 471–476.
- [41] A. Asfarian, Y. Herdiyeni, A. Rauf, and K. H. Mutaqin, "Paddy diseases identification with texture analysis using fractal descriptors based on Fourier spectrum," in *Proc. Int. Conf. Comput., Control, Informat. Appl. (IC3INA)*, Nov. 2013, pp. 77–81.
- [42] R. R. Patil and S. Kumar, "Predicting rice diseases across diverse agro-meteorological conditions using an artificial intelligence approach," *PeerJ Comput. Sci.*, vol. 7, p. e687, Sep. 2021.
- [43] T. Tawde and A. Reddy, "Rice plant disease detection and classification techniques: A survey," *Int. J. Eng. Res. Technol.*, vol. 10, no. 7, pp. 560–567, 2021.
- [44] S. V. Vasantha, B. Kiranmai, and S. R. Krishna, "Techniques for rice leaf disease detection using machine learning algorithms," *Int. J. Eng. Res. Technol.*, vol. 9, no. 8, pp. 162–166, 2021.
- [45] Y. Dong, S. Gao, K. Tao, J. Liu, and H. Wang, "Performance evaluation of early and late fusion methods for generic semantics indexing," *Pattern Anal. Appl.*, vol. 17, no. 1, pp. 37–50, Feb. 2014.
- [46] D. Lahat, T. Adali, and C. Jutten, "Multimodal data fusion: An overview of methods, challenges, and prospects," *Proc. IEEE*, vol. 103, no. 9, pp. 1449–1477, Sep. 2015.
- [47] K. Liu, Y. Li, N. Xu, and P. Natarajan, "Learn to combine modalities in multimodal deep learning," 2018, *arXiv:1805.11730*.
- [48] S. Yoganand, P. R. P. Narasingaperumal, and S. Rahul, "Prevention of crop disease in plants (groundnut) using IoT and machine learning models," *Int. Res. J. Eng. Technol.*, vol. 7, no. 3, pp. 1164–1169, 2020.
- [49] S. Khan and M. Narvekar, "Disorder detection of tomato plant (*Solanum lycopersicum*) using IoT and machine learning," *J. Phys., Conf. Ser.*, vol. 1432, no. 1, Jan. 2020, Art. no. 012086.
- [50] H. Ren, T. Hu, Y. Song, H. Sun, B. Liu, and M. Gao, "An improved electronic image motion compensation (IMC) method of aerial full-frame-type area array CCD camera based on the CCD multiphase structure and hardware implementation," *Sensors*, vol. 18, no. 8, p. 2632, Aug. 2018.
- [51] S. M. Hassan, A. K. Maji, M. Jasiński, Z. Leonowicz, and E. Jasińska, "Identification of plant-leaf diseases using CNN and transfer-learning approach," *Electronics*, vol. 10, no. 12, p. 1388, Jun. 2021.



RUTUJA R. PATIL received the bachelor's and master's degrees in information technology from Pune University, Pune, India, in 2006 and 2015, respectively. She is currently pursuing the Ph.D. degree with the Symbiosis Institute of Technology, Symbiosis International (Deemed University), Pune. She is currently a Research Scholar at Symbiosis International (Deemed University). Her research interests include machine learning and deep learning.



SUMIT KUMAR (Member, IEEE) received the bachelor's degree in electronics and telecommunication from Kurukshetra University, Kurukshetra, India, in 2005, the master's degree from the Guru Jambheshwar University of Science and Technology, Haryana, India, in 2008, and the Ph.D. degree from Jamia Millia Islamia, New Delhi, India, in 2017. Currently, he is working as an Associate Professor at the Electronics and Telecommunication Department, Symbiosis Institute of Technology, Symbiosis International (Deemed University), Pune, Maharashtra, India. His research interests include antenna design, wireless networks, wireless communication, and computational intelligence.

• • •