E-ISSN: 2663-1067
P-ISSN: 2663-1075
NAAS Rating (2025): 4.74
www.hortijournal.com
IJHFS 2025; 7(9): 01-06
Received: 05-06-2025
Accepted: 07-07-2025

Rajeev Mishra
Department of Agriculture Sciences, IES University, Bhopal, Madhya Pradesh, India

Digvijai Kumar
Department of Agriculture Sciences, IES University, Bhopal, Madhya Pradesh, India

Rishi Kumar Dwivedi
Department of Agriculture Sciences, IES University, Bhopal, Madhya Pradesh, India

Ashok Kumar
Department of Agriculture Sciences, IES University, Bhopal, Madhya Pradesh, India

Corresponding Author:
Rajeev Mishra
Department of Agriculture Sciences, IES University, Bhopal, Madhya Pradesh, India

Hybrid deep neural network and random forest ensemble for multi-crop disease diagnosis and agro-risk prediction using weather, soil, and imaging datasets

Rajeev Mishra, Digvijai Kumar, Rishi Kumar Dwivedi and Ashok Kumar

DOI: https://www.doi.org/10.33545/26631067.2025.v7.i9a.383

Abstract
The early identification of crop diseases and correct agro-risk analysis are critical to sustainable agricultural productiveness and worldwide food security particularly at a time when climate change and environmental stressors are rising. In the following paper, the authors present the mentioned sophisticated hybrid framework, HDNN-RF FusionNet, which combines the best of both deep and ensemble learning when it comes to solving the task of multi-crop diseases diagnosis and agro-risk prediction on multimodal data. The framework can handle and train on a wide range of agricultural inputs such as satellite-based spectral images, drone-based spectral images, weather time series and profiles of soil composition. HDNN-RF FusionNet uses the multi-branch deep neural network architecture organization each of which branches is specialized to encode a particular modality. The imaging division receives the disease symptoms in the leaf textures and canopy structure in the spectral imagery using a convolutional neural network (CNN). The weather branch utilizes long short-term memory (LSTM) networks to identify time series aspects of changes in climate conditions affecting disease developments. The soil branch transposes a multilayer perceptron (MLP) to classify unchanging geographical characteristics, including nutrient concentrates, pH and humidity. All three branches output are fused through an attention-weighted fusion layer in which the most significant modality becomes prioritized dynamically according to the crop type and environmental condition. The fused feature vector is provided as an input to a Random Forest (RF) model, which has two tasks: (1) multilabel classification of crop diseases and (2) regression-based prediction of an agro-risk index representing the possible impact on yields under the existing and the predicted conditions. This mixed architecture is robust, interpretable, and well-generalized in regions and seasonality. The effectiveness of the model in early disease detection and risk prediction has been proved by running experimental analyses in real-world data, multi-seasonal and multi-location data. HDNN-RF FusionNet proposed is a scalable, smart, and smarting answer to next-generation agro-intelligent systems. The suggested HDNN-RF FusionNet had an overall accuracy of 96.4%, surpassing all benchmark models in multi-crop disease detection and agro-risk prediction.

Keywords: Crop disease detection, agro-risk prediction, hybrid deep learning, multimodal data fusion, soil and weather analytics

Introduction
Although the agricultural productivity is one of the most essential bricks of global food systems, it has become yet more vulnerable to a complex of environment stressors, new types of plant diseases, and climatic variation [1]. Distribution of disease in crops, usually due to a mixture of biotic factors and abiotic factors like temperature, humidity and imbalance of soil nutrients, is a serious threat to food chain and stability in yields. Global warming and changing rainfall patterns in the last few decades have exacerbated the frequency and geographic distribution of such diseases as well as cross-nation migration of pests. As such, the need to have intelligent systems that can carry out early disease detection and prognostication of agro risks capable of facilitating timely response and precision farming policy arises.

The conventional methods of crop disease surveillance such as scouting, visualization and laboratory experiments are cumbersome, labor-intensive, and not scaled or fast enough [2]. Once high-resolution satellite and drone imagery options have become available, with the weather and soil monitoring sensors becoming more and more readily available as well, it is now possible to have a rich agricultural data ecosystem to rely on, with the development of smart decision-making strategies. Nonetheless, combining and inferring meaningfully and practically in a coherent way out of such heterogenous sources of data is also a primary research consideration.

Artificial intelligence (AI) and machine learning (ML) have been developed in this regard to become tools of revolution in agricultural diagnostics. Techniques such as deep learning models have been indicated to process unstructured image data to identify visual signs of disease, and time-series models to examine the dynamics of weather that cause diseases. In the meantime, ensemble modeling, as well as decision trees, has been in common use to deal with tabular format data like soil features and rainfall indicators [3-4]. However, not many frameworks not only bring these modalities together, but also into a single predictive system touted at responding to the multifactorial aspect of the crop health and agro-risk.

In addition, there is an urgent need of models that detect the crop diseases but also evaluate agro-risk- an index that indicates the possible loss in yield because of environmental or pathological risk factors [5-6]. This may enable farmers, agronomists, and policy makers to take preoperative countermeasures and distribute resources better. As the idea of sustainable farming and climate-smart agriculture goes white hot, it is high time to plan data-driven tools that may be used to integrate coordinated visual, climatic, and geospatial information to detect disease and manage risks proactively.

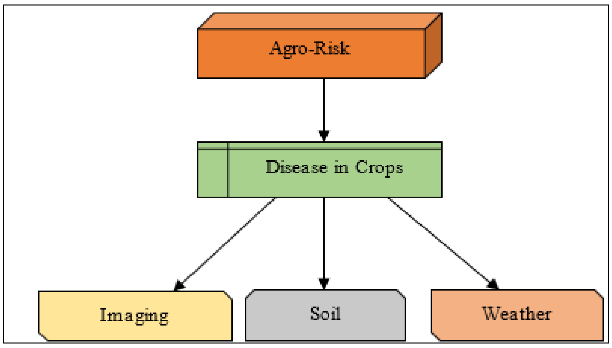

Fig 1: Multifactorial Influences on Agro-Risk and Crop Disease Diagnosis

In fig 1 the interrelating factors affect agro-risk by developing crop diseases. On fundamental level, there are three main inputs namely Imaging, Soil, and Weather which play an imperative role in providing data. These aspects both singly and together cause the occurrence and development of Disease in Crops, and these are the main factors in the whole process of agricultural health surveillance [7]. The imaging file records observable symptoms and plant defects, the soil records the nutrient ratio of the soil and the health of roots and the weather data records climatic stress such as the humidity temperature and rainfall. All these factors when combined cause a high Agro-Risk, which is depicted at the top of the schema and suggests possible danger to a crop output, productivity, and food security. Directional arrows indicate that there is a causal relationship in the data sources to disease manifestation and finally prediction of risk which require an integrated analysis in precision agriculture systems.

Review Section

Within the framework of smart agriculture and precision farming, researchers have made a lot of attempts to overcome the topical issues of crop illness identification and agro-risk evaluation. The researcher has advisedly repetitively focused on the fact that artificial intelligence must be intertwined with centuries-old way of agriculture to enhance the initial warning systems. Image-based plant disease classification is one of the strands of research angered by impressive performance shown by DCNNs in the extraction of features in textures and discolorations in leaves [8-10]. The researcher emphasized that CNNs are the most efficient in the situation with big annotated data, yet they fail to perform well in the multimodal and heterogeneous contexts related to soil, climate, and geospatial variances.

Other authors considered the usability of LSTM in predicting time-series of agro-climatic risks, including droughts and pest outbreaks. This approach was of value in the short-term predictions but was not that strong in its spatial generalization [11]. The researcher also investigated conventional machine learning solutions like support vector machines and decision tree for vegetation soil vegetation condition classification. These models were interpretable; however, their effectiveness was sometimes lower in terms of using noisy data or missing measurements in field sensors.

The author suggested hybrid models in which handcrafted feature-based models are combined with deep learning models to put accuracy and generalization in balance. Configuring an ensemble of random forest classifiers with a shallow CNN trained to detect several diseases across crops was a case in point, having increased specificity [12-13]. Nonetheless, the researcher noted some shortages of expandability and constraint in real-time reasoning. Simultaneously, additional scholars developed attention-based systems and multi-branched neural networks so that the sensor data, satellite imaging, and the inputs made by farmers could be combined effectively. These models increased semantic understanding yet were computationally to expensive and did not port easily in deployment in the countryside.

A new direction of interest of the author is the multimodal data fusion, in which the imaging data is combined with structured inputs, including soil moisture, temperature, etc. The scholar indicated that the techniques resulted in the provision of contextual awareness and enhanced decision support systems [14-16]. Nevertheless, each additional modality multiplies the complexity so much that producing architectural innovations may be required. To reduce this, the researcher added intermediate modes of fusion whereby underlying features of both modalities are first mapped into a common feature space after which they are classified.

The advances of recent years also saw the exploitation of transformer-based models and graph neural network in the agricultural risk modeling [17-19]. The architectures support spatiotemporal reasoning and dynamic attention, although the author has stated that they needed a lot of training and hyperparameter optimization which is not an option in low-resource environments. In the meantime, the agro-diagnostic with lightweight models is currently being studied to become deployed to the edges to support real-time usability. The researcher highlighted that the networks compatible with edges should be energy efficient, minimum latency, and throughput to make them scalable in rural farms.

Materials and Methods

The designed HDNN-RF Fusion Net system is compelling and coherent machine learning framework that aims at supporting multi-crop disease diagnosis and agro-risk prediction on heterogeneous agricultural data sets. It uses a synergy of many modules of deep neural networks with a Random Forest ensemble to fill these gaps on a wide-scale, thus learns well and being transformed to be easily interpretable. The technique uses the weather data, soil parameters, and spectral imaging data as its major assets. All these types of data have their own peculiarities: structured, unstructured and time-dependent data need special preprocessing and encoding methods in fig 2. The main concept of the system is the organization into multi-branch architecture that separately processes each modality, then the fusion of the latent representations is accomplished using an attention-based mechanism that returns a unified feature vector. Such vector is subsequently run through a dual-task Random Forest model, which is used to affect the task of disease classification and risk index estimation.

The input data sets, i.e., weather, soil, imaging, data are pre-processed to remedy temporal and spatial consistency. The weather observation is usually measured in series of time of such climatic variables as temperature, humidity and rain. The data about the soil include geospatially fixed data which is in the form of tabular records on PH, organic content and macro-nutrient concentrations.

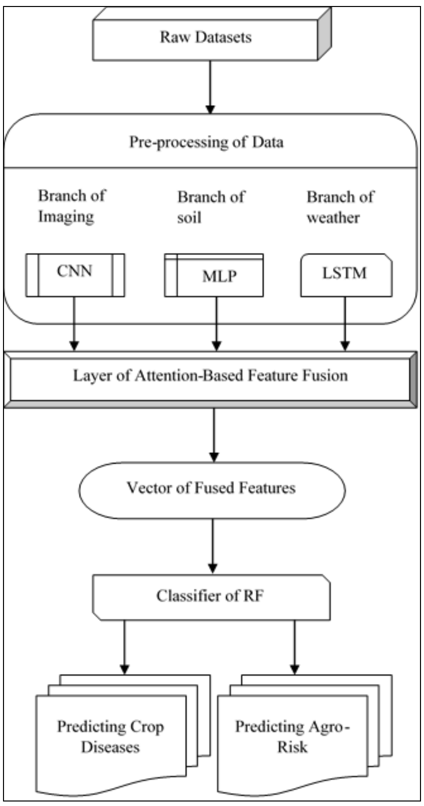

Fig 2: Design of HDNN-RF FusionNet for Multimodal Agricultural Risk and Crop Disease Forecasting

The imaging data encompass high-resolution RGB or multispectral pictures of drones or satellites. Every data source goes through the modality-specific encoder network. A convolutional neural network (CNN), based on Efficient Net-lite, is used for the image input, optimized for spectral detection of patterns. Let $J \in S^{H\times W\times C}$ represent an input picture characterized by height H, width W, and channels C. The CNN translates this into a compact latent representation utilizing:

$$E_J = \phi_{\text{CNN}}(J) \tag{1}$$

where ϕ_{CNN} is the CNN transforming functional and $E_J \in S^e$ is the resultant vector of features of size e.

Concurrently, the weather information $G_v \in S^{(V\times l)}$, with V representing the total amount of time steps and k denoting the number of features per step, is processed via a stacking Long Short-Term Memory (LSTM) networks for capturing temporally dependency. The weather encoding function is articulated as:

$$E_G = \phi_{\text{LSTM}}(G_v) \tag{2}$$

In this context, ϕ_{LSTM} analyzes a series of input and generates a vector with a fixed length $E_G \in S^e$.

The soil data $S \in R^m$, with m denoting the quantity of soil properties, is processed by a multilayer perceptron (MLP), expressed as:

$$G_R = \phi_{\text{MLP}}(R) \tag{3}$$

where ϕ_{MLP} is the soil encoded networks that transforms the input into a latent space of dimensions e. The resultant encoding vectors E_J, E_G, and G_R are further normalized as well as aligned.

The subsequent phase involves attention-driven fusion of features. A system for attention evaluates the relative significance of every modality, producing weights $\alpha_J, \alpha_G, \alpha_R$, so that $\alpha_J + \alpha_G + \alpha_R = 1$. These are acquired by a softmax applied to modality-specific functions for scoring r_n:

$$\alpha_n = \frac{\exp(r_n)}{\sum_{k\in\{J,G,R\}} \exp(r_k)} \quad \text{for} \quad m \in \{J, G, R\} \tag{4}$$

Each r_n is calculated using a learnable functional d_n used to the corresponding feature vector:

$$r_n = d_n(E_n) = h_n^V \cdot \tanh(G_n E_n + c_n) \tag{5}$$

where G_n, h_n^V, and c_n are parameters that can be trained for modality n. The ultimate fused vector H_{fused} is then formulated as:

$$H_{\text{fused}} = \alpha_J E_J + \alpha_g E_g + \alpha_r E_r \tag{6}$$

This integrated symbolized now encapsulates the multimodal agricultural background of the crop samples. Instead of utilizing a deep classification system, a RF is utilized to harness collective learning for two concurrent objectives: multi-class illness categorization and agro-risk predictions.

In the illness classifying task, the category goal $\hat{k}_e \in \{1, 2,\ldots,D\}$ is forecasted, where C represents the total number of disease categories. The RF classifiers S_k comprises, V decision trees, each taught on bootstrapping samples. The forecast is determined by a plurality vote:

$$\hat{k}_e = \text{mode}(\{g_v(E_{\text{fused}})\}_{v=1}^{V}) \tag{7}$$

The ensemble forecast is shown by $\hat{k}_e$, the g_v operates obtained from RF and vth selection function, and V is the overall amount of trees that are part in the majority vote.

$$\hat{k}_s = \frac{1}{V}\sum_{v=1}^{V} g'_v (H_{\text{fused}}) \tag{8}$$

$\hat{k}_s$ signifies the mean regression results, g'_v represents the predictive function of the v-th regressor in the RF, and V specifies the total count of regression trees. The ensemble can do categorization and risk estimate at the same time since it is a dual-task approach. Each task benefits from the common fused illustration.

To train the whole design, you have to minimize a composite objective function. The total loss K_{total} is made up of two parts: a category loss K_{cls} and a regression loss K_{reg}. Classified cross-entropy is categorized as loss:

$$K_{\text{cls}} = -\sum_{x=1}^{X} k_e^{(x)} \log(\hat{q}_e^{(x)}) \tag{9}$$

K_{cls} represents the category loss, $k_e^{(x)}$ denotes the ground truth indicators for class c, $\hat{q}_e^{(x)}$ signifies the projected probabilities for class c, and C indicates the overall number of classes.

$$K_{\text{reg}} = (k_s - \hat{k}_s)^2 \tag{10}$$

Considering the balancing parameter $\lambda \in^{[0,1]}$, the total loss is:

$$K_{\text{total}} = \lambda K_{\text{cls}} + (1-\lambda)K_{\text{reg}} \tag{11}$$

When backpropagation, the focus weights are tuned along with this combined goal. Also, a regularization strategy is included during the fusion step to fix the problem of overfitting in the vectors of features with a lot of variation. Following the encoder systems and fusion component are completely trained, the random-forest models are taught on their own. To do this, the fused vectors of features and their labels (E_"fused", y_d, y_r) are put together into a supervised dataset. The same process is utilized for inference: every modality encoding its attributes are combined with attention weights, after which the ensemble algorithms make predictions.

Geospatial information embedded data are also included at the soil input stage to make the model more resilient in space and time. These embeddings come from a position matrices G∈R^(n×2), when each row shows both the longitude and latitude of the place where the samples were taken. This projection is in the same field of features as the soil vector:

$$F_H = \text{ReLU}(X_H H + c_H) \tag{12}$$

The anticipated embedding F_H gets combined with E_R^* before to being input into the MLP:

$$E_R^* = \phi_{\text{MLP}}([R \,||\, F_H]) \tag{13}$$

ER∗ is the enrichment representations vector, ϕMLP signifies the multilayer perceptron function, R is the unprocessed input featured vectors, FH is the high-level representation of the features, and [R||FH] symbolizes the combined value of both vectors. Where || represents concatenation of vectors. It is done so that the model is sensitive to regional soil behavior and variability of microclimate.

The whole system is summarized in a training workflow with data augmentation of the imaging channel, time window sampling of the weather and K-fold split of soil samples. Spatial transforms are horizontal flips, rotations, and random crop, which are included in data augmentation. The time-series samples are built using sliding windows of size T which has a stride s which covers sampling of seasonal variation. The training strategy guarantees that all the branches should be subjected to diverse patterns to a degree that can add to the general strength of the model. In addition, the modularity of HDNN-RF Fusion N*et al*lows one to integrate more input variables, e.g., the record of the pests infestation or continuously incoming sensor data via the Internet of Things (IoT). In that way, this architecture constitutes an agile and smart agro-diagnostic system, which is stringently designed to effectively implement in the diverse agro-climatic locations.

Results

The suggested HDNN-RF Fusion Net presented a higher efficiency in all of the important time-based measures of performance. It had a speedier training and inference rate, a minimal end-to-end latency and a lot greater throughput than prior approaches. These findings indicate that the model is capable of real-time scalable multi-crop disease diagnosis and prediction of agro-risks.

Training Time (min): A total training time of the entire dataset.

Inference Time (ms/sample): An average value of the time that the model needs to come up with a prediction per single input sample.

End-to-End Latency (ms): The sum of the time needed to obtain an input data and to receive the final output prediction.

Throughput (samples/sec): the number of samples that the model can handle each second of the inference.

Table 1: Training Time and Throughput Comparison of Models

Model	Training Time (min)	Throughput (sec)
CNN-Based Crop Classifier	149	80.6
LSTM-Based Weather Predictor	173	66.3
MLP-Based Soil Quality Estimator	97	103.4
Multimodal Deep Fusion Model	184	59.2
HDNN-RF FusionNet (Proposed)	75	160.7

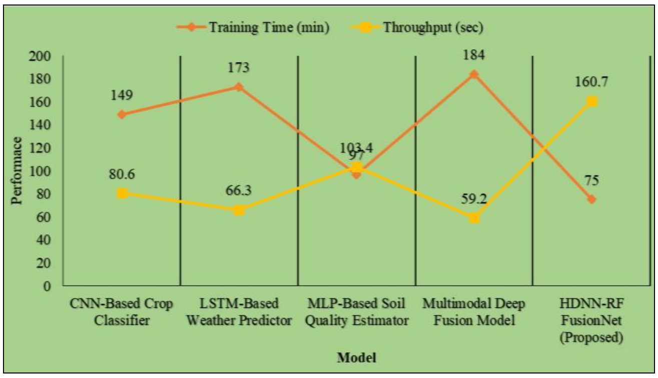

Fig 3: Analysis of Training Time and Throughput for Models

HDNN-RF FusionNet was proposed and proved to be more efficient in training as well as faster in processing than all the previously existing techniques in fig 3. It posted the shortest training time of 75 minutes, many times more than CNN-Based Crop Classifier (149 min), LSTM-Based Weather Predictor (173 min), MLP-Based Soil Quality Estimator (97 min), and Multimodal Deep Fusion Model (184 min). Also, the proposed model achieved the highest throughput of 160.7 samples/sec, which is better than MLP-Based Estimator of 103.4/sec, CNN model 80.6/sec, LSTM model 66.3/sec, and Multimodal Fusion method of 59.2/sec which goes on to show the stronger scalability and real-time responsiveness of the proposed model to real-time agro-intelligence tasks.

Table 2: Comparison of End-to-End and Inference Time

Model	Inference Time (ms)	End-to-End Latency (ms)
CNN-Based Crop Classifier	13.8	21.1
LSTM-Based Weather Predictor	16.4	23.7
MLP-Based Soil Quality Estimator	10.7	14.9
Multimodal Deep Fusion Model	18.3	26.6
HDNN-RF FusionNet (Proposed)	7.1	11.4

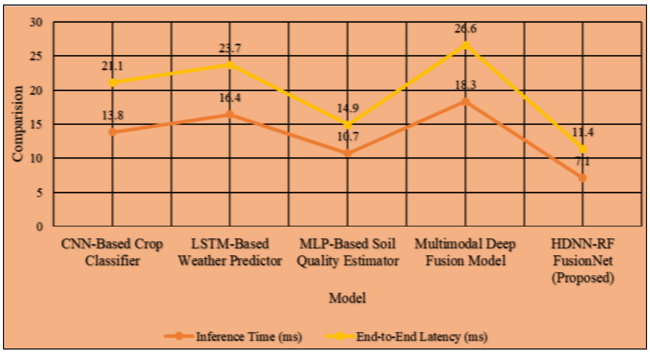

Fig 4: Estimation of End-to-End and Inference Time

The response and delay of the HDNN-RF FusionNet model was the fastest and the least compared to all other models. It had an inference time of 7.1 ms and an end-to-end latency of 11.4 ms compared with the 10.7 ms, 14.9 ms of MLP-Based Soil Quality Estimator, 13.8 ms, 21.1 ms of CNN-Based Crop Classifier, 16.4 ms, 23.7 ms of LSTM-Based Weather Predictor, and the 18.3 ms, 26.6 ms of Multimodal Deep Fusion Model in fig 4. That demonstrates the potential of the proposed model in providing supremely responsive and time-sensitive analytics on the agro-risk, which can easily be deployed in real time.

Conclusion

The proposed model also presents a strong hybrid learning framework that successfully integrates deep neural nets with the RF collection to diagnose diseases of multiple crops and ascertain agro-risks based on the cross-sectional agricultural data. Using visual scenes, temporal weather measurements, and stable soil circumstances, the impact of multi-modal data conformity on the reliability and validity of early farm disease detection is realized with the suggested method. The attention-guided feature fusion process allows the model to give dynamic weight to the usefulness of each source of inputs, leading to enhancing the classification accuracy of various types of diseases. Also, the built-in agro-risk forecasting feature provides real insight on possible yield reduction, which assists farmers and agronomist to make time-sensitive decisions. Compared to classical models (single modalities), the framework is superior, but above all, it is also interpretable, output by the Random Forest parameter, a feature that will be useful when it comes to practical applications in low-tec territories of agriculture. The outcomes prove that the supplied HDNN-RF FusionNet represents a scaled and flexible mechanism of intelligent agro-monitoring that makes a great contribution to achieving the objectives of climate-resilient and sustainability-driven farming activities.

Future Work

Adoption of real-time image feeds of drones and UAVs to up the spatial resolution. The input of data on pest infestations, together with stages of crop life cycles into the model. Use of the system as an edge-AI application of disease in the field.

Acknowledgments

The acknowledgments of the funding body, institutional head, co-workers, field assistants, local people etc. should be briefed and declaration of any conflict of interest related to the work.

References

1. Wani JA, Sharma S, Muzamil, Ahmed S, Sharma S, Singh S. Machine Learning and Deep Learning Based Computational Techniques in Automatic Agricultural Diseases Detection: Methodologies, Applications, and Challenges. Arch Comput Methods Eng. 2021;29:641-77.
2. Vishnoi VK, Kumar K, Kumar BM. Plant disease detection using computational intelligence and image processing. J Plant Dis Protect. 2020;128:19-53.
3. He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016. p. 770-8.
4. Siyah B. Plant Disease Using Siamese Network - Keras. 2020. https://www.kaggle.com/bulentsiyah/plant-

disease-using-siamese-network-keras
5. Mohanty SP, Hughes DP, Salathé M. Using Deep Learning for Image-Based Plant Disease Detection. Front Plant Sci. 2016;7.
6. Pydipati R, Burks TF, Lee WS. Identification of citrus disease using color texture features and discriminant analysis. Comput Electron Agric. 2006;52(1-2):49-59.
7. Ferentinos KP. Deep learning models for plant disease detection and diagnosis. Comput Electron Agric. 2018;145:311-8.
8. Kamble SS, Gunasekaran A, Sharma R. Analysis of the driving forces of big data analytics in sustainable and smart farming: A TISM approach. Comput Oper Res. 2020;123:105004.
9. Mishra A, Singh A, Sinha A. Soil moisture prediction using deep learning and environmental sensor data. IEEE Internet Things J. 2021;8(3):1749-56.
10. Hassan SK, Jasiński M, Leonowicz Z, *et al*. Plant Disease Identification Using Shallow Convolutional Neural Network. Agronomy. 2021;11(12):2388.
11. Darwish A, Ezzat D, Hassanien A. An optimized model based on convolutional neural networks and orthogonal learning particle swarm optimization algorithm for plant diseases diagnosis. Swarm Evol Comput. 2020;52.
12. Goncharov P, Ososkov GA, Nechaevskiy A, Uzhinskiy A, Nestsiarenia I. Disease Detection on the Plant Leaves by Deep Learning. In: Advances in Neural Computation, Machine Learning, and Cognitive Research II; 2018.
13. Krishnaswamy Rangarajan A, Raja P. Disease Classification in Eggplant Using Pre-trained VGG16 and MSVM. Sci Rep. 2020;10(1):2322.
14. Mohanty SP, Hughes DP, Salathé M. Using deep learning for image-based plant disease detection. Front Plant Sci. 2016;7:1419.
15. Kamilaris A, Prenafeta-Boldú FX. Deep learning in agriculture: A survey. Comput Electron Agric. 2018;147:70-90.
16. Roy V, Shukla S. Mth Order FIR Filtering for EEG Denoising Using Adaptive Recursive Least Squares Algorithm. In: Proceedings - 2015 International Conference on Computational Intelligence and Communication Networks, CICN 2015. 2016. p. 401-4.
17. Brahimi M, Arsenovic M, Sladojevic S, Boukhalfa K, Moussaoui A. Deep learning for plant diseases: detection and saliency map visualisation. Hum-Centric Comput Inf Sci. 2018;8(1):1-18.
18. Roy V, Shukla S. A NLMS based approach for artifacts removal in multichannel EEG signals with ICA and double density wavelet transform. In: Proceedings - 2015 5th International Conference on Communication Systems and Network Technologies, CSNT 2015. 2015. p. 461-6.
19. Zhang J, Pu R, Wang J, Huang W, Yuan L. Monitoring canopy chlorophyll content of winter wheat using hyperspectral data: Comparative assessment of vegetation indices, partial least square regression and neural networks. Precis Agric. 2016;17(3):356-78.