Potato Research
https://doi.org/10.1007/s11540-024-09760-x

Early Detection of Potato Disease Using an Enhanced Convolutional Neural Network-Long Short-Term Memory Deep Learning Model

Sarah A. Alzakari[1] · Amel Ali Alhussan[1] · Al-Seyday T. Qenawy[2] · Ahmed M. Elshewey[3]

Received: 13 April 2024 / Accepted: 20 June 2024

Abstract
Potato diseases pose a significant threat to farmers, impacting potato crops' productivity, quality, and financial stability. Among the most notorious diseases is late blight, caused by *Phytophthora infestans*, famously responsible for triggering the Irish Potato Famine in the 1840s. Late blight swiftly devastates potato foliage and tubers, particularly in damp, humid conditions. Another common disease is early blight, attributed to *Alternaria solani*. This disease affects various parts of the potato plant—leaves, stems, and tubers. It mainly shows up in the form of dark stains around the center of a bull's eye on the leaves, bringing down both the yield and the crop quality. A model consisting of a Convolutional Neural Network - Long Short-Term Memory (CNN-LSTM) enhanced for potato disease detection was proposed in our paper. The dataset used was Z-score standardized before the training and testing process using the proposed CNN-LSTM model was started. The performance of the implemented model, CNN-LSTM, was analyzed alongside five traditional machine learning algorithms, namely Random Forest (RF), Extra Trees (ET), K-Nearest Neighbours (KNN), Adaptive Boosting (AdaBoost), and Support Vector Machine (SVM). Accuracy, sensitivity, specificity, F-score, and AUC were the metrics included in the evaluation, confirming the effectiveness of the models. The results of the experiments showed that our CNN-LSTM reached the highest accuracy at 97.1%.

Keywords Convolutional neural network · Detection · Early blight · Late blight · Long short-term memory · Potato disease

✉ Ahmed M. Elshewey
ahmed.elshewey@fci.suezuni.edu.eg

[1] Department of Computer Sciences, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, P.O. Box 84428, 11671 Riyadh, Saudi Arabia

[2] Intelligent Systems and Machine Learning Lab, Shenzhen 518000, China

[3] Department of Computer Science, Faculty of Computers and Information, Suez University, P.O. Box 43221, Suez, Egypt

Introduction

For the farmer, potato diseases present an immense obstacle, as they not only diminish the yield but also decrease crop quality, an outcome that may adversely affect the economic viability of the agricultural endeavors (Athanikar and Badar 2016). These diseases originate from different sources, such as fungi, bacteria and viruses and the diseases can be affected by environmental factors such as soil conditions and weather patterns (Butte et al. 2021). So, understanding and implementing good disease management skills are vital in maintaining agricultural systems and food security. Potatoes are, throughout their life cycle, from planting to the harvest, susceptible to many diseases. Late blight, early blight, potato scab, blackleg, bacterial wilt and viral infections such as potato leaf roll virus (PLRV) and potato virus Y (PVY) are the most common diseases (Hu et al. 2016).

Late blight, a disease caused by *Phytophthera infestans*, takes the lead among the most widespread diseases damaging potato plants worldwide (Lee et al. 2021). Also, early blight, caused by*Alternaria solani*, poses a grave issue because the infection of the foliage causes decreased yield. Potato diseases have a variety of symptoms, such as leaf spots, root rot, stunted growth and entire plant withering. In high humidity and warm temperature conditions, the diseases will propagate quickly, so they need to be managed effectively and immediately if their spread is to be controlled (Sharma et al. 2017). Addressing potato diseases tends to be a complex task that involves employing cultural practices, including the use of rotation, sanitation, and disease-resistant varieties, together with the application of chemical measures that consist of the use of fungicides and pesticides. Integrated Plant Management (IPM) strategies are aimed at reducing chemical reliance and concurrently effectively mitigating diseases (Lee et al. 2021). Ongoing discovery of the resistance mechanisms of potatoes, pathogen biology, and developing environment-friendly agriculture methodologies mainly help develop new strategies against potato diseases. By using very detailed knowledge about the close bonds between pathogens, host plants, and environmental conditions, farmers are able to secure potato crops and have a reliable food supply for the next generations (Sharma et al. 2017). Where national prosperity is dependent on the economic development of the agricultural sector, potato production is usually an important crop, particularly for highly productive countries. The issue of field security for many developing countries is also a very serious problem that is connected to the amelioration of malnutrition (Singh et al. 2020).

Recent data reveals that in Pakistan potatoes are grown across 170,300 ha, yielding an annual production of 4 million tons, with a visible upward trend in per capita consumption rates. Recognizing its crucial economic role, farmers must prioritize the healthy development of potato crops. Potatoes play a crucial role in Bangladesh's agricultural landscape, yet their production falls short compared to that of more developed nations globally. This shortfall is largely due to the prevalence of diseases and diseases that impede potato cultivation. Consequently, Bangladesh encounters challenges in meeting its potato export targets to other countries. Notably, diseases like early blight, leaf roll virus, scab, and tuber defects such as hollow heart are significant obstacles for potato farmers across key growing regions in the country. Consequently, both farmers and traders face substantial difficulties, particularly in exporting to countries such as Russia, Indonesia, Malaysia, Sri Lanka, Thailand, Hong Kong, Turkey, Vietnam, and Maldives. The COVID-19 pandemic has further aggravated the situation, leading to a gradual rise in potato prices (Boas et al. 2023; Gavhale and Gawande 2014). Potatoes are highly regarded for their efficiency as a crop, as they offer greater yields of protein, dry matter, and minerals per unit area compared to cereals. On the other hand, the potato sector faces numerous diseases with the possibility of infection leading to poor harvests and damaged tubers, which can make potatoes expensive in the end. The disadvantages are a high price and a requirement for specialist supervision when local expertise is limited (Balodi et al. 2017; Khalifa et al. 2021).

These challenges are being addressed in many ways, including disease diagnostics and certification methods in potato plants. State-of-the-art image processing and machine-learning methods for automated disease identification have been established (Tsedaley 2014; Tarik et al. 2021; Rizk et al. 2023; Iqbal and Talukder 2020). With the help of leaf inspection and the deployment of VGG16 and VGG19 deep learning models, the researchers were able to provide the highest accuracy in different types of potato disease classification (Sholihati et al. 2020). Early detection of potato diseases means effective management of the disease, which in turn leads to improved crop yield. A transfer learning method has been implemented for potato disease detection with limited leaf photos. Nonetheless, it has yielded good accuracy rates. Here, they usually rely on pre-trained deep-learning models that do not require a sufficiently large dataset (Islam et al. 2019). Automation systems developed for disease detection and classification are the primary focus of ongoing endeavors. These constitute image-processing techniques and machine-learning approaches for the identification and classification of potato leaf diseases (Tiwari et al. 2020).

The use of segmentation of images of healthy and diseased potato leaves and employing the various classifier algorithms have helped these researchers recognize and classify different leaf diseases (Mahum et al. 2023). For this purpose, a deep learning model with high efficiency, known as Convolutional Neural Network-Long Short-Term Memory (CNN-LSTM), will be used for potato disease detection. The CNN-LSTM model contributes effectively to the classification problems. CNN excels at identifying spatialities, which is helpful in classification tasks. On the other hand, LSTMs are better at grasping temporal dependencies that are characteristic of sequential data. The CNN-LSTM model exhibits its effectiveness in terms of extracting both spatial and temporal patterns by fusing CNN and LSTM. CNNs learn the hierarchical feature, where initial layers learn from lower-level features and deeper layers learn from the higher-level features. However, LSTMs focus specifically on temporal dependencies amid the sequential data elements. Once CNNs learn to encode the spatial hierarchies and LSTM the temporal dependencies, the classifier task gets more informative representations. CNNs have an advantage when it comes to extracting spatial local features through convolutional filters. At the same time, LSTMs can be said to excel at long-range temporal dependencies using memory cells. It is this feature that allows the CNN-LSTM model to pull out relevant features from complex input data successfully and effectively, thus increasing its classification result.

Related Work

Identification and classification of potato leaf diseases is another point of disagreement among researchers. Research has put forward different ways of tackling potato leaf disease recognition and detection with the help of image processing and machine learning tools (Ngugi et al. 2021). The subject of the occurrence of potato diseases as the reason for crop decline and their economic impact is another point of contention. Certain literature brings to light the extreme harm caused by potato leaf diseases as well as the influence on earners (Fiers et al. 2012). However, other studies emphasize early detection and action to prevent losses (Devaux et al. 2021). For potato disease detection and classification, a lot of research is focused on building automated systems using complex machine learning and image processing methods. These systems strive to grow more potatoes, avoid economic losses for farmers, and alert the agricultural supply chain (Thomas-Sharma et al. 2016). Potato disease detection aims to use deep learning and image processing approaches in the field of crop disease classification (Powelson and Rowe 1993). The goal is to develop autonomous systems that will be able to correctly identify and distinguish potato leaf diseases by pictures of leaves (Singh and Kaur 2021).

Doing so will lead to the early detection and speedy management of these diseases, which will result in reduced crop losses and, hence, increased potato production (Kreuze et al. 2020). Bienkowski et al. (Bienkowski et al. 2019) conducted a study to illustrate the relationship of the light wavelength to the leaf structure in the identification of potato mold based on image segmentation and machine learning processes. With the help of the results of the greenhouse experiment, disease screening accuracy was found to be 84.6%. The types of diseases in different categories, such as the process of late blight, were correctly detected with a success rate of 92%. The model also showed its capacity to distinguish between post-marked leaves, healthy ones and those with black crown rot, reaching a 74.6% accuracy rate. Some experts, Islam et al. (2017), suggest algorithms based on classification and image processing to be used in detecting illnesses from patterns visualized in stem images. Current phenotyping and disease detection technologies are pivotal in securing food and adapting agricultural production practices to the changing environment. The method described below is fully automated; it categorizes potato diseases from the freely accessible image data from ‘Plant’ Village. With the aid of the segmentation method and support vector machines, the system acquires a high resolution of 95% in the classification of 300 diseased images. Biswas and her colleagues (Biswas et al. 2014) proposed a disease-level assessment system specifically targeting potato diseases, especially late blight of potatoes, which usually attacks the potato leaves. Their approach includes fuzzy C-means clustering as well as convolutional neural networks that are used for disease detection and identification; their machine was able to get a 93% accuracy.

Ranjan and co-workers (Ranjan et al. 2015) demonstrated an approach to recognizing the diseases in cotton plants using leaf images. They captured images of affected regions, which were preprocessed by a neural network, and the analysis was done later on. Through a comparison of healthy and impaired samples, they could achieve an 80% accuracy. The authors (Pinki 2017) Pinki et al. have proposed a three-disease detection model. Their method of analysis relied on the K-means clustering algorithm to classify afflicted regions by considering visual characteristics like texture, shape and color. Resorting to the Support Vector Machine (SVM), they managed to distinguish diseases, reaching an impressive accuracy of 92.06%. (Revathi and Hemalatha 2014): The author made a novel model for checking the diseased leaves of cotton by amalgamating the Particle Swarm Optimization (PSO) with the feed-forward neural network (FFNN). What stands out about this approach is that it managed to achieve an impressive inspection rate of 95%. The result of (Kandel et al. 2023) aimed at using a deep learning framework on a dataset containing more than 4483 images of various fruit leaves, including pear, cherry, peach, apple and grape leaves. The aim of this was to classify these images as any of the 13 categories. Notably, it has recorded accuracy ranges of 91 to 98%. This study makes significant contributions in the field by introducing a new deep neural model termed Convolutional Neural Network-Long Short-Term Memory (CNN-LSTM), which resulted in a vast improvement in testing accuracy, reaching a level of up to 97.1% for the classification of potato diseases. Furthermore, the developed method showed impressive results in terms of accuracy compared to that of existing methods in this area.

The structure of the paper is organized as follows: In the “Materials and Methods” section, the author discusses the materials and techniques. The “Results and Discussion” section contains the results and discussion. Next, the last section is about the conclusion and future works of the study.

Materials and Methods

In this study, a highly efficient deep learning model called Convolutional Neural Network-Long Short-Term Memory (CNN-LSTM) was used for potato disease detection. The CNN-LSTM model plays a crucial role in classification tasks. CNNs are adept at identifying spatial patterns, which is beneficial for classification tasks. There are several key components in the architecture of the CNN-LSTM model. First, there are two convolutional layers, followed by a max-pooling layer. Then, an LSTM layer is used, together with a hidden layer and an output layer. The first convolutional layer consists of 100 filters with a kernel size of 20, while the second convolutional layer consists of 64 filters with a kernel size of 10. The LSTM layer consists of 128 hidden units. There are 64 neurons in the hidden layer, and a batch size of 32 is used during training. The learning rate is set to 0.1, and the Adam optimizer is used. The data is structured in time steps of 32, and training is performed in 100 epochs. The activation function used in the output layer is the sigmoid activation function. Figure 1 depicts the framework of the proposed model in six steps.

The LSTM architecture, shown in Figure 2, summarizes the main concept of a sophisticated solution to the vanishing gradient problem standard for recurrent neural networks (RNNs) that occurs during training on long sequences. It emphasizes the composition of different LSTM cells from diverse layers, which is a crucial characteristic. Each cell comprises input gates, forget gates, cell states, and output gates. These parts collectively control the flow of knowledge within the network by selectively handling and altering critical data over time. To recall memory, LSTMs can adjust the content of cells, which control the mechanism of inputs and forget gates to

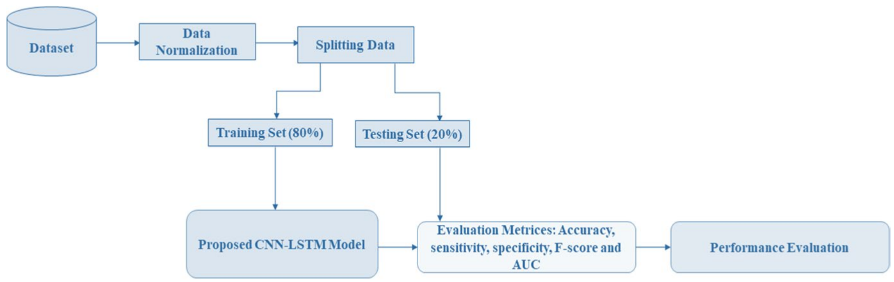

Fig. 1 Framework of the proposed model

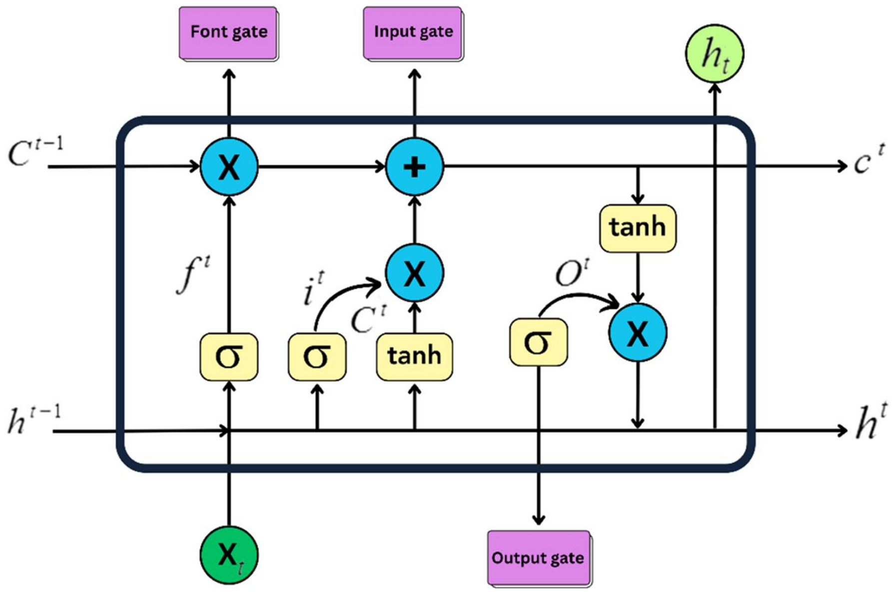

Fig. 2 Long Short-Term Memory architecture

capture long-term connections, contributing to the execution of speech recognition, language modeling, and time series prediction.

The input gate monitors and controls the amount of information entering the cell at a given time, helping select which details from the input and the previous state are saved in memory space. In contrast, the forget gate is responsible for selecting which data should be deleted from the memory cells. The cell state acts as a reading memory to be updated by the input data and memory cells. Consequently, the output gate is used to regulate communication between the memory of the cell and the output so that only relevant information is passed on for the task at hand. This intricate infrastructure thus allows LSTMs to manipulate and use information effectively, ranging from shorter to longer sequences, therefore anchoring them in various fields that rely on the utilization of sequential data analysis.

Dataset

The dataset utilized in this paper is available at Kaggle (n.d.). The dataset contains 4020 records and 7 features, with 6 features representing the inputs and 1 feature representing the output. The input features are temperature, humidity, wind speed, wind bearing, visibility, and pressure, and the output feature is the class. The class feature contains 2066 early blight records and 1954 late blight records.

Table 1 presents a summary of the features of the dataset, providing key statistical insights:

- Feature Name: This column lists the names of the variables in the dataset.
- Count: Indicates the total number of observations available for each feature, which is consistent at 4020 for all features.
- Mean: The mean represents the average value of each characteristic across all observations, obtained by summing all values and dividing them by the total number of observations.
- Standard Deviation (Std): This measure reflects the dispersion, or scatter, of the data around the mean, showing how much values differ from the mean. A higher standard deviation indicates greater variability.
- Minimum (Min): Indicates the smallest observed value for each characteristic in the dataset.
- Median (50%): Represents the median value of each characteristic. The median, which is the middle value when observations are arranged in ascending order, divides the data into two equal parts.
- Maximum (Max): Displays the largest observed value for each characteristic in the dataset.

Figure 3 shows a heat map representation of the connections between the feature variables within the dataset. Heatmap analysis is a very efficient tool for comparing different variables by visualizing those relationships.

Figure 4 illustrates the histogram distribution analysis for the data set's features. This technique provides a graphical way to view the distribution of data

Table 1 Statistical calculation for the dataset features

Feature name	count	mean	std	min	50%	max
Temperature	4020.0	23.580249	2.905305	17.5	23.4	30.4
Humidity	4020.0	67.247264	6.814943	48.0	68.0	83.0
Wind speed	4020.0	5.404602	2.114070	1.2	5.1	11.1
Wind bearing	4020.0	184.280846	89.812636	0.0	178.0	358.0
Visibility	4020.0	10.111443	1.257829	6.9	9.9	14.5
Pressure	4020.0	1015.167040	0.595284	1012.6	1015.2	1017.2
Class	4020.0	0.513930	0.499868	0.0	1.0	1.0

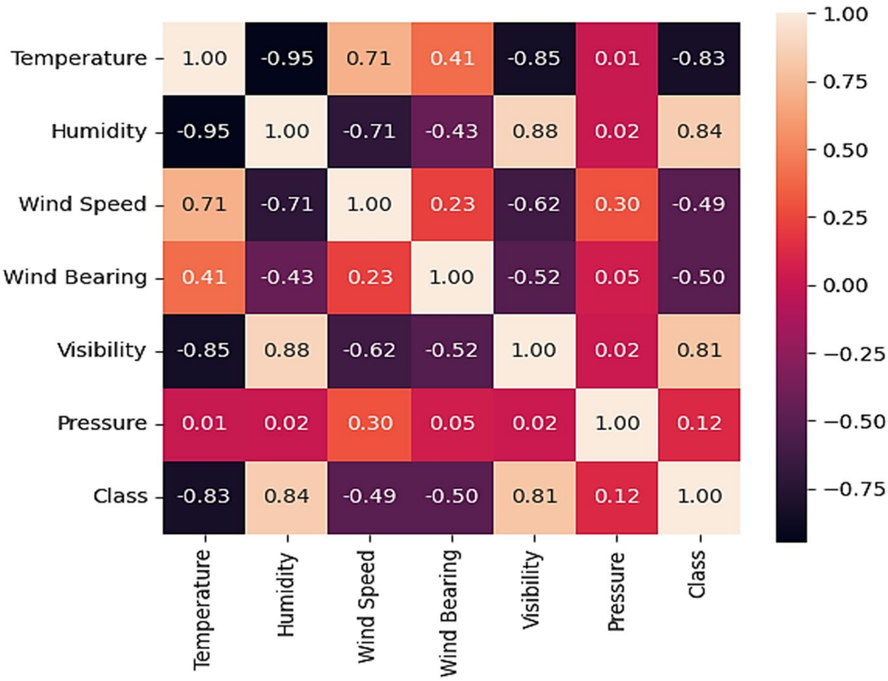

Fig. 3 Heatmap analysis for the dataset features

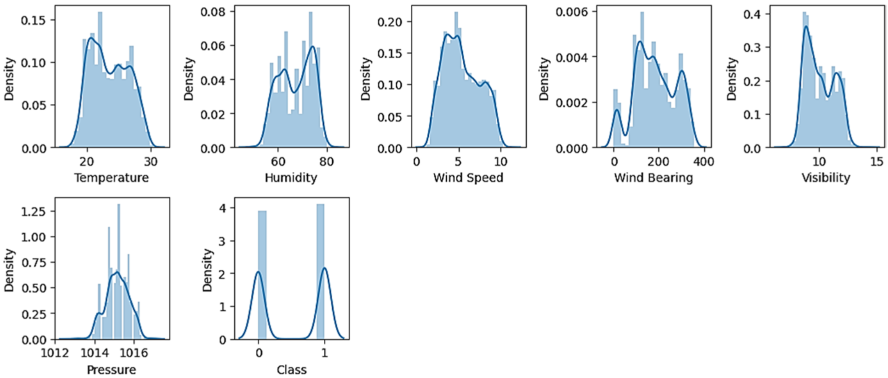

Fig. 4 Histogram distribution for the dataset features

points across different ranges. The shape of the histogram is mainly used to identify statistical features such as skew, bimodality, or symmetry. This analysis gives information on the overall distribution and features of the data.

Z-score Normalization

Normalization is a frequently used preprocessing operation in machine learning. It seeks to tackle the variability phenomenon by consistently performing the analysis on all elements in a dataset. Likewise, it provides for comparisons between data of different scales. Mathematical transformations will not help convert data that do not match in scales into standard format (Fei et al. 2021). The process involves managing minimum and maximum values and then standardizing the range. Essentially, the goal is to assign a value of 0 to the smallest value and a value of 1 to the largest value while distributing other values evenly within this 0 to 1 range. The application of the minimum-maximum formula, also called*Z*-score normalization, is illustrated in Equation (1).

$$z = \frac{x - \min(x)}{(x) - \max(x)} \tag{1}$$

where, in this equation, z denotes the transformed data while x denotes the input value. The terms $\min(x)$ and $\max(x)$ correspond to the minimum and maximum values within the given input data set.

Machine Learning Algorithms

Various Machine Learning (ML) classifiers were trained using the training data to compare their performance with the performance of the proposed CNN-LSTM model. In our study, we examined many machine-learning classification models, incorporating both linear and non-linear machine-learning classification models. The classification models utilized in this study included Random Forest (RF), Extra Trees (ET), K-Nearest Neighbours (KNN), Adaptive Boosting (AdaBoost) and Support Vector Machine (SVM).

Random Forest (RF)

The Random Forest (RF) model is a flexible tool suitable for both batch and non-batch training scenarios, addressing both regression and classification problems (Pal 2005). It works by cultivating numerous decision trees, which are later merged into a unified model. In classification efforts, the final prediction of the model depends on the majority vote of each tree individually predicting a class. The core principle behind RF is to construct a forest of trees that are relatively independent of each other, thereby mitigating overfitting and ensuring robust predictions. Although individual trees are susceptible to noise and inaccuracy, the presence of different trees in the forest facilitates accurate predictions (Parmar et al. 2019). To reduce inter-tree dependencies, the construction process of each tree employs random feature selection and bagging (bootstrap aggregation) techniques, thereby increasing diversity within the forest. RF is characterized as a robust ensemble learning (EL) approach that overcomes the limitations of a single classifier by integrating multiple decision tree classifiers. Consequently, the RF method uses majority voting and multiple trees to determine the final class label rather than relying solely on a single tree.

Extra Trees (ET)

Extra trees (ET), also known as extremely randomized trees, is a tree-based ensemble method similar to random forest, but it is relatively new in the field of machine learning (Sharaff and Gupta 2019). As an extension of the popular random forest technique, its primary goal is to reduce the risk of overfitting (Bhati and Rai 2020). Unlike conventional methods, the extra trees model uses a top-down approach to construct an assembly of unpruned decision or regression tree structures. It differs in that it uses the entire learning sample (rather than a bootstrap replica) and selects cut points entirely at random. Similar to Random Forest, Extra Trees trains each base estimator with a chosen set of features, but it differs by randomly selecting both the features and their corresponding values for node splitting. Through ensemble learning, Extra Trees (ET) combines multiple decision trees to create a more robust model. ET uses a randomized approach by splitting nodes using a randomly selected set of features and a randomized threshold for each feature.

K-Nearest Neighbours (KNN)

K-Nearest Neighbours (KNN) is a simple yet effective algorithm used in both classification and regression tasks (Zaki et al. 2023). In classification, it assigns a label to a data point by examining the majority class among its K nearest neighbours in the feature space. Unlike many algorithms, KNN does not require an explicit training phase; it simply remembers the training data. When presented with a new data point for classification, KNN identifies its K nearest neighbours based on a chosen distance metric, typically the Euclidean distance. It then assigns the class label to the new data point by majority voting among its neighbours. The class that appears most frequently among the K neighbours is considered the predicted class. The parameter K, which represents the number of neighbours to consider, plays a crucial role in KNN. A smaller K value increases the sensitivity to noise and outliers, while a larger K value can lead to smoothing of the decision boundaries, possibly resulting in underfitting (Abu Alfeilat et al. 2019). Determining the optimal K value often involves the use of cross-validation or othesr tuning techniques. In addition, the choice of distance metric has a significant impact on the performance of KNN. Although Euclidean distance is commonly used, alternative metrics such as Manhattan distance, Minkowski distance, or cosine similarity may better fit the data characteristics.

Adaptive Boosting (AdaBoost)

Adaptive boosting, commonly referred to as AdaBoost, is the most widely used boosting algorithm (Rojas 2009). Compared to other versions of boosting algorithms, AdaBoost requires less fine-tuning of algorithmic parameters. In an AdaBoost classification model, the process begins by fitting a replica of the original data set with a version of the previous classifier that has been adjusted to eliminate error-prone and inaccurate data points. This allows subsequent classifiers to focus on instances that contribute to larger inaccuracies. AdaBoost generates a hypothesis based on potential labels, aiming for a weak hypothesis prediction error of less than 0.5 during training (Walse et al. 2017). The goal of the distribution is to select the "hard" trained data and samples for the next iteration. AdaBoost uses weighted majority voting to determine classes from the prediction of each hypothesis and constructs an ensemble and a set of hypotheses. AdaBoost operates as an iterative computational method where multiple classifiers (i.e., weak learners) are trained using a training set and various coordination techniques are employed to create a more robust model.

Support Vector Machine (SVM)

Support Vector Machine (SVM) is a powerful supervised machine learning technique used primarily for classification tasks (Fung and Mangasarian 2001). It excels in scenarios where data can be linearly separated and seeks to identify the optimal hyperplane that effectively divides the data into classes. In SVM classification, the algorithm works by projecting the input data into a higher dimensional feature space. It attempts to find the hyperplane that maximizes the separation margin between classes. This margin is the distance between the hyperplane and the closest data points from each class, called support vectors. The primary goal of SVM is to find the hyperplane that not only separates the data but also maximizes this margin. When data is not linearly separable, SVM uses kernel functions to map it into higher dimensions where separation becomes possible. As SVM learning progresses, a cost function is gradually refined, having a misclassification bias to be penalized while favoring margin maximization. This process of optimization leads to solving the convex optimization problem by applying methods like gradient descent or quadratic programming, among others. The specific strength of SVM in handling high-dimensional data is that it provides for the quick performance of tasks with a large number of features (Amari and Wu 1999). SVM is also robust to overfitting, which is typically prevented by regularization when the model's complexity is managed. As a result, it turns out that SVM is a rather versatile and widely adopted algorithm for classification, with commendable accuracy and generalization abilities in various fields.

Evaluation Metrics

The metrics used to assess our models' performance were various. However, they included accuracy, sensitivity, specificity, *F*-score and Area Under Curve ROC (AUC) (Hossin and Sulaiman 2015). These clearly indicate the performance of the system from different viewpoints, such as TP, TN, FP, and FN. The following equations express the classification performance metrics:

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN}$$
$$Sensitivity = \frac{TP}{TP+FN}$$
$$Specitivity = \frac{TN}{TN+FP}$$
$$F - score = \frac{2*TP}{2*TP+FP+FN}$$
$$AUC = \frac{1}{2}\left(\frac{TP}{TP+FN} + \frac{TN}{TN+FP}\right)$$

The extent evaluation criteria used for assessing our models included rather diverse parameters, seeking to allow one to have a comprehensive scope. These metrics included the accuracy, sensitivity, specificity, *F*-score, and the Area Under Curve ROC. Every one of these metrics will provide an excellent view of the model by allowing us to evaluate the performance in a multifaceted way (Muhammed and Almetwally 2024; Towfek 2023).

- Accuracy is considered the most basic metric that measures the total number of correctly classified cases (true positives and true negatives) versus all cases. It outlines the model's performance as a whole, indicating in which classes it makes its most accurate predictions.
- Sensitivity refers to the model's ability to pinpoint correct instances (positive in this case) within the data. It calculates the rate of correctly predicted cases relative to all the existing positive cases. Sensitivity becomes all the more significant in cases where the correct location of the positive cases is of utmost importance, as in medical diagnoses or anomaly detection.
- Specificity sets off sensitivity in the machine learning model because it accentuates a model's ability to identify negative samples correctly. The TPR reflects the percentage of true negative instances that were recognized as such, compared to all verifiable negative instances. Uniqueness is critical for cases when it is vital to exclude negative cases accurately, like fraud detection or quality control.
- *F*-measure expresses precision and recall for both positive and negative classes of the model as a harmonic mean. It takes into consideration both false positives and false negatives, and through this, it gives out a single metric that summarizes the compromise between the two.
- The Area Under Curve Receiver Operating Characteristic (AUC ROC) measure demonstrates the model's ability to identify between groups. It is plotted to show the true positive rate on the y-axis and the false positive rate on the x-axis as the threshold setting is varied. A larger AUC value means better discrimination ability. Often, values approaching 1 and higher indicate excellent performance.

Together, all the metrics allow the development of a complex evaluation of the model from different aspects, which helps to make educated decisions and reform it to obtain appropriate results.

Table 2 Hyperparameters of the machine learning models used in this study

Models	Hyperparameters
RF	N_estimators = 100, criterion = gini.
ET	N_estimators = 50, criterion = entropy.
KNN	N_neighbors = 20, weights = distance.
AdaBoost	N_estimators = 100, learning_rate = 0.1.
SVM	Kernel = rbf, regularization parameter (C) = 0.2.

Table 3 Performance of the proposed CNN-LSTM model and other individual machine learning classifiers

Models	Accuracy	Sensitivity	Specificity	*F*-score	AUC
CNN-LSTM	**97.1%**	**97.1%**	**97.1%**	**97.2%**	**0.987**
RF	96.2%	96.3%	96.2%	96.2%	0.973
ET	95.8%	95.8%	95.8%	95.8%	0.964
KNN	93.9%	93.9%	93.8%	93.9%	0.946
AdaBoost	89.6%	89.6%	89.7%	89.6%	0.900
SVM	83.3%	83.3%	83.3%	83.4%	0.848

Results and Discussion

This study compares the performance of the CNN-LSTM model with other individual classifiers like Random Forest (RF), Extra Trees (ET), K-Nearest Neighbours (KNN), Adaptive Boosting (AdaBoost) and Support Vector Machine (SVM). Evaluation indicators, i.e., accuracy, sensitivity, specificity, *F*-score and Area Under the ROC Curve (AUC), were used to evaluate the performances of these models in detecting early foliage symptoms of early and late blight.

Table 2 includes hyperparameter values which are used for machine learning models. RF configuration needs 100 estimators and a “gini” criterion to make split quality measurements. For the ET approach, there will be 50 trees, and the “entropy” will be the factor for splitting. For the K-Nearest Neighbours (KNN) algorithm, twenty neighbours are considered for predictions, which means that the nearest neighbours are given more weight through the distance weight function. The max number of AdaBoost estimators is set at 100, while each classifier’s contribution to the final combination update is configured with a learning rate of 0.1. Support Vector Machine, which implements the “rbf” kernel function and C = 0.2, has achieved the best performance.

As seen in Table 3, the best results are obtained by the proposed CNN-LSTM model results with an accuracy of 97.1%, sensitivity of 97.1%, specificity of 97.1%, *F*-score of 97.2%, and AUC of 0.987, which indicates excellent discrimination ability as it is near to 1. The SVM model achieved worthwhile results with an accuracy of 83.3%, a sensitivity of 83.3%, a specificity of 83.3%, an *F*-score of 83.4% and an AUC of 0.848. For the RF model, its accuracy, sensitivity, specificity, *F*-score, and AUC are 96.2%, 96.3%, 96.2%, 96.2%, and 0.973. The ET model achieved an accuracy of 95.8%, sensitivity of 95.8%, specificity of 95.8%, *F*-score of 95.8%, and

AUC of 0.964. The accuracy, sensitivity, specificity, *F*-score, and AUC for the KNN model are 93.9%, 93.9%, 93.8%, 93.9% and 0.946, respectively. For the Adaboost model, its accuracy, sensitivity, specificity, *F*-score, and AUC are 89.6%, 89.6%, 89.7%, 89.6%, and 0.900.

Figure 5 demonstrates the accuracy of the proposed CNN-LSTM model and other classification models, namely, Random Forest (RF), Extra Trees (ET), K-nearest neighbours (KNN), Adaptive Boosting (AdaBoost) and Support Vector Machine (SVM).

Figure 6 presents the training and testing loss values vs a number of epochs using the proposed CNN-LSTM model. Figure 7 presents the training and testing accuracy vs number of epochs using the proposed CNN-LSTM model.

Conclusion

In conclusion, an enhanced CNN-LSTM deep learning model offers a promising solution to early potato disease detection and that impacts on yield and quality. By leveraging a normal dataset and employing *Z*-score normalization techniques, we trained and tested the CNN-LSTM model, comparing its performance against five conventional machine learning models: Random Forest (RF), Extra Tree (ET), K-Nearest Neighbour (KNN), AdaBoost (AB), and Support Vector Machine (SVM). The simulations demonstrated the CNN-LSTM produced the highest accuracy rate of 97.1% and this highlights the effectiveness of deep learning technology in providing precise and reliable early potato disease detection.

It is expected that, apart from classroom theories, farmers will be able to apply our research findings in practical farming and other agricultural districts. Through the application of our CNN-LSTM model to agricultural settings, farmers can enjoy

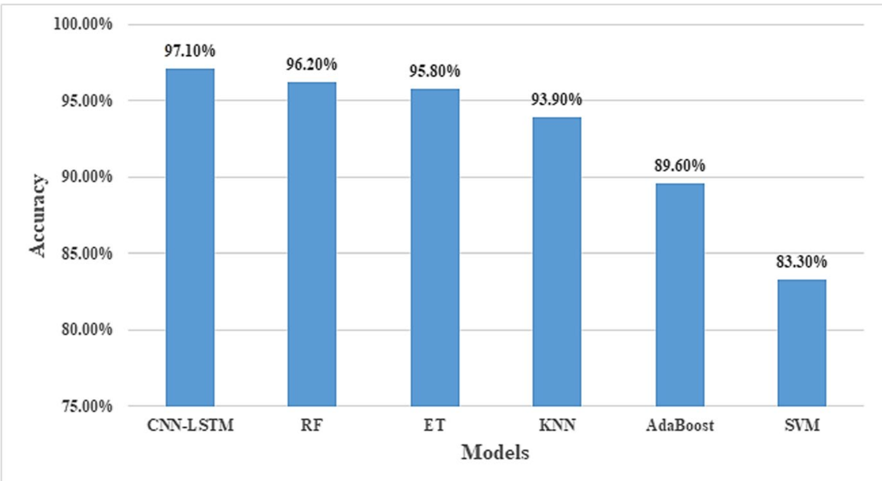

Fig. 5 Accuracy for the proposed CNN-LSTM model and several classification models

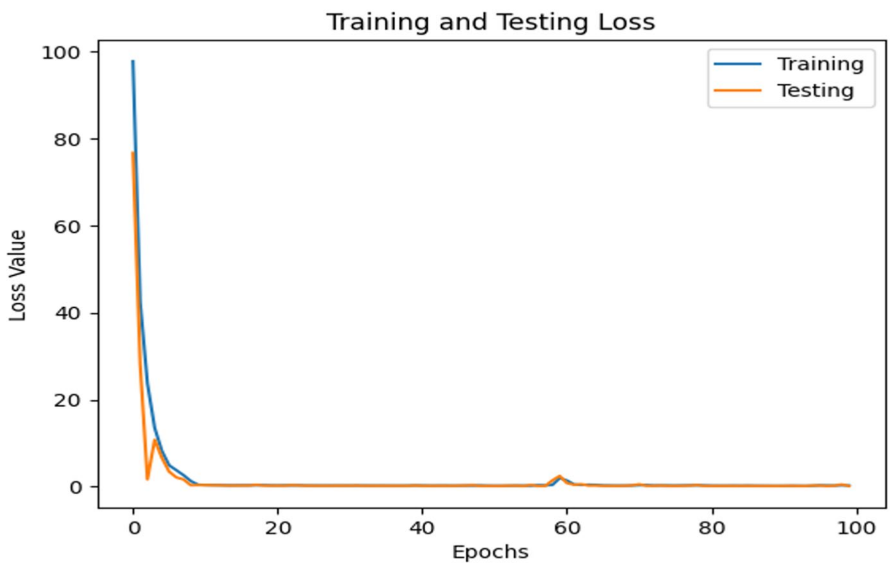

Fig. 6 Training and testing loss values vs number of epochs using the proposed CNN-LSTM model

Fig. 7 Training and testing accuracy vs number of epochs using the proposed CNN-LSTM model

state-of-the-art disease diagnosis service and follow progressive preventive agriculture strategies to reduce crop loss and maintain the harvest yield, which should be of high quality. Furthermore, embedding the model into multidimensional tools that are part of the farmer's decision-making system seems promising for raising the quality of disease management scenarios by providing on-the-fly feedback and expert advice for perfecting crop protection measures. These schemes, therefore, embody the innovation required in the potato monitoring systems that will, in turn, build resiliency in the potato supply chain and guarantee crop security in the event of changing risks to human health.

Regarding future research, the development of virtual reality simulations for early learning environments is just one of the promising routes resulting from this study. Including datasets that involve various pathogenic species that affect potatoes as well as agroclimatic factors ensures the model works properly in different situations, thereby becoming more pertinent. Further investigations modifying the CNN-LSTM architecture and alternative options for the normalization methods are areas where additional research is targeted at improving and enhancing the system performance. Furthermore, an in-depth examination of the model as an interplay with emerging technologies, particularly remote sensing and Internet of Things (IoT) devices, could be promising in developing data-driven solutions for localized potato diseases. Preliminary studies in these domains are almost always followed up with significant advances in agricultural technology and these new technologies will contribute to sustainable farming practices that will feed a growing global population.

Acknowledgements Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2024R716), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Author Contribution All authors have contributed equally.

Data Availability Https://www.kaggle.com/datasets/tamima1530/potato-leaf-disease-based-on-weather-details

Declarations

Conflict of Interest The authors declare no competing interests.

References

Abu Alfeilat HA, Hassanat AB, Lasassmeh O, Tarawneh AS, Alhasanat MB, Eyal Salman HS, Prasath VS (2019) Effects of distance measure choice on k-nearest neighbor classifier performance: a review. Big data 7(4):221–248. https://doi.org/10.1089/big.2018.0175

Amari SI, Wu S (1999) Improving support vector machine classifiers by modifying kernel functions. Neural Networks 12(6):783–789. https://doi.org/10.1016/S0893-6080(99)00032-5

Athanikar G, Badar P (2016) Potato leaf diseases detection and classification system. International Journal of Computer Science and Mobile Computing 5(2):76–88

Balodi R, Bisht SUNAINA, Ghatak A, Rao KH (2017) Plant disease diagnosis: technological advancements and challenges. Indian Phytopathology, 70(3):275-281. https://doi.org/10.24838/ip.2017.v70.i3.72487

Bhati BS, Rai CS (2020) Ensemble based approach for intrusion detection using extra tree classifier. In: Intelligent Computing in Engineering: Select Proceedings of RICE 2019, Springer Singapore, 213-220. https://doi.org/10.1007/978-981-15-2780-7_25

Bienkowski D, Aitkenhead MJ, Lees AK, Gallagher C, Neilson R (2019) Detection and differentiation between potato (Solanum tuberosum) diseases using calibration models trained with non-imaging spectrometry data. Computers and Electronics in Agriculture 167:105056. https://doi.org/10.1016/j.compag.2019.105056

Biswas S, Jagyasi B, Singh BP, Lal M (2014) Severity identification of potato late blight disease from crop images captured under uncontrolled environment. In: 2014 IEEE Canada international humanitarian technology conference-(IHTC), IEEE, 1-5. https://doi.org/10.1109/IHTC.2014.7147519

Boas LVV, Meireles ALD, Rezende RBA, Silva LFL e, Oliveira CL de (2023) Impact of Covid-19 on the supply and prices of lettuce, potato and tomato in Minas Gerais. Colloquium agrariae issn: 1809-8215 19(1):210-225. https://doi.org/10.5747/ca.2023.v19.h525

Butte S, Vakanski A, Duellman K, Wang H, Mirkouei A (2021) Potato crop stress identification in aerial images using deep learning-based object detection. Agronomy Journal 113(5):3991–4002. https://doi.org/10.1002/agj2.20841

Devaux A, Goffart JP, Kromann P, Andrade-Piedra J, Polar V, Hareau G (2021) The potato of the future: opportunities and challenges in sustainable agri-food systems. Potato Research 64(4):681–720. https://doi.org/10.1007/s11540-021-09501-4

Fei N, Gao Y, Lu Z, Xiang T (2021) Z-score normalization, hubness, and few-shot learning. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, 142-151. https://doi.org/10.1109/ICCV48922.2021.00021

Fiers M, Edel-Hermann V, Chatot C, Le Hingrat Y, Alabouvette C, Steinberg C (2012) Potato soil-borne diseases. A review. Agronomy for Sustainable Development 32(1):93–132. https://doi.org/10.1007/s13593-011-0035-z

Fung G, Mangasarian OL (2001) Proximal support vector machine classifiers. In: Proceedings of the seventh ACM SIGKDD international conference on knowledge discovery and data mining, 77-86. https://doi.org/10.1145/502512.502527

Gavhale KR, Gawande U (2014) An overview of the research on plant leaves disease detection using image processing techniques. IOSR journal of computer engineering (iosr-jce), 16(1):10-16. https://doi.org/10.9790/0661-16151016

Hossin M, Sulaiman MN (2015) A review on evaluation metrics for data classification evaluations. International journal of data mining & knowledge management process 5(2):1. https://doi.org/10.5121/ijdkp.2015.5201

Hu Y, Ping X, Xu M, Shan W, He Y (2016) Detection of late blight disease on potato leaves using hyperspectral imaging technique. Guang Pu Xue Yu Guang Pu Fen Xi 36(2):515–519

Iqbal MA, Talukder KH (2020) Detection of potato disease using image segmentation and machine learning. In: 2020 international conference on wireless communications signal processing and networking (WiSPNET), IEEE, 43-47. https://doi.org/10.1109/WiSPNET48689.2020.9198563

Islam M, Dinh A, Wahid K, Bhowmik P (2017) Detection of potato diseases using image segmentation and multiclass support vector machine. In: 2017 IEEE 30th Canadian conference on electrical and computer engineering (CCECE), IEEE, 1-4. https://doi.org/10.1109/CCECE.2017.7946594

Islam F, Hoq MN, Rahman CM (2019) Application of transfer learning to detect potato disease from leaf image. In: 2019 IEEE International Conference on Robotics, Automation, Artificial-intelligence and Internet-of-Things (RAAICON), IEEE, 127-130. https://doi.org/10.1109/RAAICON48939.2019.53

Kaggle (n.d.) Potato leaf disease based on weather details. Available at: https://www.kaggle.com/datasets/tamima1530/potato-leaf-disease-based-on-weather-details.

Kandel MA, Rizk FH, Hongou L, Zaki AM, Khan H, El-Kenawy E-SM (2023) Evaluating the efficacy of deep learning architectures in predicting traffic patterns for smart city development. Journal of Artificial Intelligence and Metaheuristics, 6:26. https://doi.org/10.54216/JAIM.060203

Khalifa NEM, Taha MHN, Abou El-Maged LM, Hassanien AE (2021) Artificial intelligence in potato leaf disease classification: a deep learning approach. In: Machine learning and big data analytics paradigms: analysis, applications and challenges, 63-79. https://doi.org/10.1007/978-3-030-59338-4_4

Kreuze JF, Souza-Dias JAC, Jeevalatha A, Figueira AR, Valkonen JPT, Jones RAC (2020) Viral diseases in potato. In: The potato crop: its agricultural, nutritional and social contribution to humankind, 389-430. https://doi.org/10.1007/978-3-030-28683-5_11

Lee TY, Lin IA, Yu JY, Yang JM, Chang YC (2021) High-efficiency disease detection for potato leaf with convolutional neural network. SN Computer Science 2(4):297. https://doi.org/10.1007/s42979-021-00691-9

Mahum R, Munir H, Mughal ZUN, Awais M, Sher Khan F, Saqlain M, Mahamad S, Tlili I (2023) A novel framework for potato leaf disease detection using an efficient deep learning model. Human and ecological risk assessment: An International Journal 29(2):303–326. https://doi.org/10.1080/10807039.2022.2064814

Muhammed HZ, Almetwally EM (2024) Bayesian and non-Bayesian estimation for the shape parameters of new versions of bivariate inverse Weibull distribution based on Progressive Type II Censoring.

Computational Journal of Mathematical and Statistical Sciences, 3:85-111. https://doi.org/10.21608/cjmss.2023.250678.1028
Ngugi LC, Abelwahab M, Abo-Zahhad M (2021) Recent advances in image processing techniques for automated leaf pest and disease recognition-a review. Information processing in agriculture 8(1):27–51. https://doi.org/10.1016/j.inpa.2020.04.004
Pal M (2005) Random forest classifier for remote sensing classification. International journal of remote sensing 26(1):217–222. https://doi.org/10.1080/01431160412331269698
Parmar A, Katariya R, Patel V (2019) A review on random forest: an ensemble classifier. In: International conference on intelligent data communication technologies and internet of things (ICICI) 2018, Springer International Publishing, 758-763. https://doi.org/10.1007/978-3-030-03146-6_86
Pinki FT, Khatun N, Islam SM (2017) Content based paddy leaf disease recognition and remedy prediction using support vector machine. In: 2017 20th international conference of computer and information technology (ICCIT), IEEE, 1-5. https://doi.org/10.1109/ICCITECHN.2017.8281764
Powelson ML, Rowe RC (1993) Biology and management of early dying of potatoes. Annual review of phytopathology 31(1):111–126. https://doi.org/10.1146/annurev.phyto.31.1.111
Ranjan M, Weginwar MR, Joshi N, Ingole AB (2015) Detection and classification of leaf disease using artificial neural network. International journal of technical research and applications 3(3):331–333
Revathi P, Hemalatha M (2014) Identification of cotton diseases based on cross information gain deep forward neural network classifier with PSO feature selection. International Journal of Engineering and Technology 5(6):4637–4642
Rizk FH, Arkhstan S, Zaki AM, Kandel MA, Towfek SK (2023) Integrated CNN and waterwheel plant algorithm for enhanced global traffic detection. Journal of Artificial Intelligence and Metaheuristics, 6:36-45. https://doi.org/10.54216/JAIM.060204
Rojas, R. (2009) AdaBoost and the super bowl of classifiers a tutorial introduction to adaptive boosting. Tech. Rep, Freie University, Berlin, 1(1): 1-6.
Sharaff A, Gupta H (2019) Extra-tree classifier with metaheuristics approach for email classification. In: Advances in Computer Communication and Computational Sciences: Proceedings of IC4S 2018, Springer Singapore, 189-197. https://doi.org/10.1007/978-981-13-6861-5_17
Sharma R, Singh A, Dutta MK, Riha K, Kriz P (2017) Image processing-based automated identification of late blight disease from leaf images of potato crops. In: 2017 40th international conference on telecommunications and signal processing (TSP), IEEE, 758-762.
Sholihati RA, Sulistijono IA, Risnumawan A, Kusumawati E (2020) Potato leaf disease classification using deep learning approach. In: 2020 international electronics symposium (IES), IEEE, 392-397. https://doi.org/10.1109/IES50839.2020.9231784
Singh A, Kaur H (2021) Potato plant leaves disease detection and classification using machine learning methodologies. IOP Conference Series: Materials Science and Engineering 1022(1):012121. https://doi.org/10.1088/1757-899X/1022/1/012121
Singh, B., Raigond, P., Dutt, S., & Kumar, M. (2020). Potatoes for food and nutritional security. *Potato: Nutrition and food security*, 1-12.
Tarik MI, Akter S, Al Mamun A, Sattar A (2021) February. Potato disease detection using machine learning. In: 2021 Third international conference on intelligent communication technologies and virtual mobile networks (ICICV), IEEE, 800-803. https://doi.org/10.1109/ICICV50876.2021.9388606
Thomas-Sharma S, Abdurahman A, Ali S, Andrade-Piedra JL, Bao S, Charkowski AO, Crook D, Kadian M, Kromann P, Struik PC, Torrance L (2016) Seed degeneration in potato: the need for an integrated seed health strategy to mitigate the problem in developing countries. Plant Pathology 65(1):3–16. https://doi.org/10.1111/ppa.12439
Tiwari D, Ashish M, Gangwar N, Sharma A, Patel S, Bhardwaj S (2020) May. Potato leaf diseases detection using deep learning. In: 2020 4th international conference on intelligent computing and control systems (ICICCS), IEEE, 461-466. https://doi.org/10.1109/ICICCS48265.2020.9121067
Towfek SK (2023) Navigating the storm: cutting-edge risk mitigation and analysis for volatile markets. Journal of Artificial Intelligence and Metaheuristics, 4:36-44. https://doi.org/10.54216/JAIM.040204
Tsedaley B (2014) Late blight of potato (Phytophthora infestans) biology, economic importance and its management approaches. Journal of, Agriculture and Healthcare 4(25):215–225
Walse KH, Dharaskar RV, Thakare VM (2017) A study on the effect of adaptive boosting on performance of classifiers for human activity recognition. In: Proceedings of the International Conference on Data Engineering and Communication Technology: ICDECT 2016, Springer Singapore, 419-429. https://doi.org/10.1007/978-981-10-1678-3_41

Zaki AM, Abdelhamid AA, Ibrahim A, Eid MM, El-Kenawy E-SM (2023) Enhancing K-Nearest Neighbors algorithm in wireless sensor networks through stochastic fractal search and particle swarm optimization. Journal of Cybersecurity and Information Management, 13:76-84. https://doi.org/10.54216/JCIM.130108