ORIGINAL PAPER

Nutrient deficiency identification and yield-loss prediction in leaf images of groundnut crop using transfer learning

Kummari Venkatesh[1] · K. Jairam Naik[1]

Abstract

Nutrient deficiencies in the plant can significantly reduce agricultural productivity. A sufficient ratio of nutrient elements like nitrogen (N), phosphorus (P), and potassium (K) is essential for plant growth. Deficiencies in these elements lead to a significant decline in crop yield. The primary objective of this work is to identify the nutrient deficiency in the leaf images and determine how severe the nutrient deficiency is in the plant. We have enhanced the VGG16 transfer learning model and integrated it with another proposed nutrient severity identification module. The proposed VGG16 model is trained with a groundnut dataset and classified the images as N, P, and K deficient. Predicting yield loss in the crop based on nutrition deficiency is the secondary objective of this work. To identify deficiencies, we have used two different datasets: rice plant and groundnut images. The proposed methodology classifies the plant leaf images into N, P, and K classes of deficiencies. Based on this information, crop yield loss due to each nutrient deficiency is estimated. The proposed enhanced (VGG16) classification accuracy is 98% on the groundnut dataset (i.e., our dataset collected from the field). The performances were compared with the existing state-of-the-art models, which found that the proposed model performs better than the current models trained on the same datasets.

Keywords Nutrient deficiency · Severity identification · Crop yield predicting · VGG16 · Agricultural productivity · Plant leaf images · Raspberry Pi

1 Introduction

Providing food sources for the exponentially growing population is one of the world's biggest challenges. According to [1], the population is expected to increase by 9700 million by 2050. It can lead to water scarcity, arable areas decline, global climate change, and environmental pollution. There are no conventional agricultural methods to provide food facilities for this increasing population. One of the biggest challenges of traditional agriculture is the farmers' need for knowledge about farmland's soil and nutrition content. As a result, the farmers select inappropriate crops for cropping in their fields, which causes nutrition deficiency resulting in a decreased crop yield. An alternative way of avoiding such problems

is by automating the process through precision farming [2]. It utilizes several advanced technologies like the Internet of Things (IoT), machine learning (ML), deep learning (DL), and many more to monitor the crop field regularly. It predicts the plant's nutrition contents (nitrogen, potassium, phosphorous, etc.). Nutrition management is essential in improving the quality of the crop outcomes and the gain in the yield. Plants require organic and inorganic elements for growth [3]. Deficiencies in maintaining the appropriate nutrition ratio in the plant can create productivity loss in the final crop. Lack of nutrition is the difference between the required and available nutrition in the plant.

Precision farming must become an active part of agriculture to feed the world's rapidly growing population. It uses IoT sensors for collecting the soil characteristics of the field. Imaging equipment like cameras is used to manage the specific characteristics of the plant like color, size, and shape monitor. Plant phenomics [4] is the study of these visible qualities of a plant using a variety of sensors and imaging equipment, which is essential to this transformation.

✉ Kummari Venkatesh
  Kvenkatesh.phd2020.cse@nitrr.ac.in

  K. Jairam Naik
  Jnaik.cse@nitrr.ac.in

1 Dept. of Computer Science and Engineering, National Institute of Technology Raipur, Raipur, India

Using these visible plant characteristics to identify physiological changes brought on by abiotic stresses like lack of water, nutrients, and poor lighting is extremely beneficial for plant and crop management. Plants' most important element of nutrient-induced stress is that they can significantly lower agricultural production after moisture-induced stress [5]. A lack of minerals in the plant system can cause nutrient-induced stress. But, assessing nutrient deficiencies requires an expert to traverse the field and diagnose problematic areas. Also, this process is overburdened because of repeating the same process in several iterations [6].

Early identification and measurement of plant nutrient deficiency are important for obtaining eventual output. *Nitrogen* is one of the most significant, growth-limiting nutrients in plants, and it is a necessary nutrient for forming amino acids, proteins, nucleic acids, and chlorophyll. Plants also require nutrition, called *phosphorus*, to absorb and convert solar energy into biomolecules like adenosine triphosphate (ATP), which facilitate biological reactions like photosynthesis. The challenge for plant scientists is to overcome these constraints while considering the different effects of climate change on food security, especially for breeders and agronomists. In this sense, cost-effective technologies that can track crop performance, enhance yield prediction, or measure phenotypic variability for breeding reasons are meant to get over the obstacles standing in the way of this technology's full potential.

The smart and automated plant phenotyping methods that can determine the plant's stress level are required for precision agriculture because most traditional phenotyping techniques damage the plants, which is labor-intensive and time-consuming. Recent advances in computer vision and object recognition have made the process of taking high-resolution pictures in the visible spectrum easier [7, 8]. These developments have made it possible to perform image-based plant phenotyping in a small-scale environment, like in a greenhouse or the field, making it a viable choice for plant study and management [9–11]. Such imaging-based techniques are highly automated, significantly quicker, and non-destructive. However, combining high-resolution imaging and computer vision methods for plant phenotyping presents challenges with data collection and interpretation [12]. The interpretation process requires a basic knowledge of engineering and mathematics for appropriate utilization. ML techniques offer several benefits when analyzing big data in various research areas, including healthcare, business, agriculture, etc. But still, it is challenging to extract nutrient deficiency information from images of agricultural plants due to the absence of texture in multivariate plant images, occlusion, and complicated and invisible plant sections [13]. The raw images generated in imaging-based phenotyping applications typically include vast information for standard ML algorithms. Much traditional ML research focuses on pre-computing domain-specific picture key points and features, then training a classifier in the associated representation space [14]. The DL model, a subset of ML innovation, employs numerous convolutions in acquiring the deeper details of the plant/leaf images and presents data hierarchically through gaining ground truth knowledge about nutrient deficiency [15, 16].

Recent efforts by many precision agriculture groups to identify and categorize flaws in large-scale agricultural plants have drawn considerable attention to this field of research. According to the [17] review paper, there are three main problems with current methods: (i) lack of generality (methods are tailored for particular datasets), (ii) difficulty applying in variable field conditions (methods require consistent lighting, angle of capture, and altitude), and (iii) lack of tool sophistication that restricts their applicability. The same author identifies the following two difficulties directly related to the three significant limitations that have not yet been adequately resolved [17]. By resolving these concerns and conducting thorough testing on identifying nitrogen (N) deficiency in groundnut crop plants, we amend, considerably alter, and extend our previously presented technique in this work.

The contributions of this paper are as follows:

1. The paper discusses classifying the plant images based on nutrient deficiencies.
2. Unlike other published efforts that mostly employ traditional ML algorithms for classifying nitrogen shortage in plants, we propose to use a VGG16-layered CNN architecture.
3. This approach improves accuracy, faster speed, and fewer training samples per iteration.
4. We provide a comprehensive performance study of the traditional ML methods employing combinations of critical point detectors and descriptors.

Using the methodology's final product, the farmer user can automatically identify stressed leaves and determine the relative severity of N deficiency in a field. This knowledge can help farmers create a more precise plan for applying fertilizer, which significantly positively affects the economy and the environment. Images of maize that we gathered from experimental fields throughout two cultivation seasons serve as examples of the benefits of the suggested methods. Our findings support using the proposed plan as a stand-alone entity to characterize one of the most severe corn deficits using an unmanned aerial vehicle-assisted visual assessment.

This work uses a self-collected groundnut dataset for nutrient deficiency identification and crop yield prediction analysis. The enhanced version of Visual Geometric Group-16 (eVGG16) architecture was used to automate the process of nutrient (nitrogen (N), phosphorus (P), and potassium (K))

deficiency classification and performance analyze groundnut datasets [18]. We also automated the system through *nutrient severity identification method (NSIM)* to identify the crop yield loss due to the nutrient mentioned above imbalances. Hence, classifying the plant images based on nutrient deficiencies, estimating the severity of the nutrient deficiency, and predicting the yield loss in the crop, are the main objectives of this work.

The rest of the paper is arranged as follows. The literature review is described in Sect. 2; the proposed eVGG16 methodology is provided in Sect. 3. Section 4 details the performance study and results analysis. Section 5 includes the conclusion and future scope of the research.

2 Literature review

Atmosphere temperature impacts crop yield because plants require a specific range of temperatures during their growth stages. Nowadays, crop yield is decreasing yearly due to several factors, including nutrient deficiency, poor water management, failure to detect diseases, etc. Many authors offered solutions to these issues, some discussed here. The authors suggest a crop yield prediction method [1] in soybean crops using the CNN-RNN based on environmental data and management practices. The CNN component of the model was developed to capture the spatial relationships of soil data and the internal temporal relationships of weather data. The RNN component of the model was created to accurately represent the rising crop production trend over time due to ongoing advancements in plant breeding techniques. CNN-RNN model training root-mean-square error (RMSE) is 14.49%, training correlation coefficient percentage is 92%, and validation RMSE is 24.10%. This work can be extended to similar methods that might be used to categorize a hybrid as high-yielding or low-yielding, depending on how it performed relative to other combinations in the same area.

Some farmers need to gain awareness of their agricultural fields, due to which they cannot choose a suitable crop to grow in their agricultural fields. This leads to a reduction in crop yield. This issue can be solved by introducing IoT based [2] crop suggestion system, which suggests the right crop for farmers to grow in their agricultural fields. This crop suggestion system uses a multilayer perceptron (MLP) algorithm. This system provides a graphical user interface (GUI) for friendly user interaction. This MLP system achieves 97% accuracy.

The authors [3] proposed a disease identification and classification system in soybean crops using a support vector machine (SVM). This process categorizes the disease identification procedure into two parts. They are (1) healthy or diseased crop identification and (2) diseases classification.

The SVM classifier categorizes crop plants into different diseases based on the leaf texture, color, and pattern. It only permits four viral diseases: downy mildew, Septoria, Frogeye, and leaf blight. The accuracy reported is 90%. The limitation of this system is that it is restricted to identifying the fungal disease only. [19] Using RGB UAV phenotyping techniques, researchers can assess the performance of the maize genotype under low-nitrogen conditions. [20] presented a study of identifying high-yielding and late-leaf spot disease-tolerant families in groundnut. [21] This study aimed to assess the efficacy of various portable sensors as inexpensive phenotyping instruments for testing late leaf spot (LLS) and groundnut rosette disease resistance in groundnut breeding. [22] This study uses image processing techniques to assess the percentage of groundnut leaf area impacted across five locations in Andhra Pradesh.

The crop needs a specific type of environment during its growth stages. The climatic changes impact the crop growth rate. The Pearson correlation coefficient analyzes the relationship between the crop growth rate and climatic changes. These climatic conditions help to analyze the crop yield rate (Li et al., 2019). The nonlinear multivariate regression equation calculated for plant height at the cotton flowering stage, cotton stalk weight, and lint percentage models ranges from 0.08–0.62, 0.07–0.56, and 0.1–0.8, respectively. The advantage of this model is the accurate determination of the crop growth impact factors. However, this method only identifies some of the crop's growth rate factors.

The plant phenotyping conditions can measure the plant's health. Many researchers have developed advanced tools to measure plant phenotypes, like imaging and sensors. The reviewed plant phenotyping work is presented here [4], and its technologies, like data acquisition, handling, and processing, are essential. Sensor information translated into a knowledgeable format during the early years of plant phenomics research is considerably less. To meet this challenge, it is necessary to consider the close relationship between environmental factors and plant structure, functions, and metabolism. As a result, all the phenotyping steps must include environmental characterization, from data collection to meta-analyses.

Water and nitrogen are essential entities for plant growth. The radial basis function (RBF) is used to analyze water and nitrogen stress levels. The RBF works better than the SVM and backpropagation methods [5]. The water stress level classification accuracy is 85%, and the nitrogen recognition rate was slightly lower. The research proposed the water and nutrition level classification using the chlorophyll fluorescence kinetic curve. This model's future enhancement includes more things to monitor, such as temperature, fertilizers, and humidity.

Nutrient deficiency identification can be monitored using computer vision technology, enabling real-time monitoring of corn. Nitrogen nutrition is observed using crop field images using this technology. The crop field images are collected using the drone camera in RGB color space, and these images are used to analyze plant diseases [6] and nitrogen nutrient deficiency [8]. The SVM classifier is used to classify nitrogen-deficient and healthy pixels in the image, and this classifier achieves an average precision score is 82%. This methodology advantage is the real-time nitrogen monitoring in the corn field. The limitation of this methodology is monitoring only the nitrogen nutrient deficient.

Crop analysis using advanced technologies requires plant temperature, leaf images, soil data, etc. The atmospheric temperature and relative humidity are measured using digital humidity temperature sensors placed in the crop. The two types of plant images used for the early detection of plant diseases are hyperspectral images [7] and RGB images. These images are processed using computer vision technology [10]. A normalized difference vegetation index (NDVI) sensor is also used to obtain the crop field's green color index. Collecting these types of crop-related data is a tedious and time-consuming process. To avoid this issue, an integrated sensor system is introduced [9]. This integrated sensor system collects enormous plant phenotypic data for plant phenotypic analysis. The sensor-based features were correlated with both early- (when Pearson's correlation coefficient R varied from 0.41 to 0.55) and late-season grain yield (when R from 0.55 to 0.70). The drawback of this system is that it needs to collect soil moisture data. These sensor data are used for the classification of plant species. IB1 classifier [11] is used for plant species classification with an accuracy of 79%.

Plant phenotyping is the study of plant characteristics. The recently developed methodologies have become crucial for resolving issues related to plant growth, development, and environmental reactions and selecting optimal genotypes in molecular breeding strategies in basic and applied science. The phenotyping review is discussed here [12]. This brief review concentrates on recent developments in creating integrative automated platforms for high-throughput plant phenotyping that simultaneously use several sensors to analyze plant shoots. The usage of ML in plant phenotyping is a new area where the plant community focuses on learning algorithms like SVM, KNN, ANN, etc. This review article presented plant phenotyping [13]. It provides a comprehensive overview of the best practices using ML tools for identifying plant stress. As part of this review, the author identified several future ways for using ML techniques that showcase huge promise but need utilization by the plant phenotyping research community.

Deep learning is used for identifying plant stress. In DL, CNN uses images, whereas RNN [15] uses time series data as input. Plant leaves images provide the most information

for plant recognition and are used for plant stress classification. Feature extraction is also crucial for plant stress classification, which employs the global image descriptor (GIST) method [14]. The classification accuracy using the GIST features with cosine k-nearest neighbor (KNN) is 79%, and using shape and texture features is 56%. Another benefit of the GIST feature extraction algorithm is the low computational complexity, which boosts calculation speed. The 23-layered CNN is used for identifying nitrogen deficiency. It divides the nitrogen nutrient deficient levels into three categories that are low, medium, and high [23] in sorghum plants. The obtained nitrogen classification accuracy is 75%, with sensitivity, specificity, and precision at 52%, 81%, and 64%, respectively. The advantage of this architecture is nitrogen nutrition monitoring, while a disadvantage is limited to one nutrient monitoring.

[24] Multi-locus genome-wide association analyses revealed a genomic area and potential candidate genes related to leaf spot disorders in African groundnut [25], [26]. To improve the bambara groundnut crop in response to drought stress, it is crucial to understand how genetic features regulate photosynthetic qualities in Bambara groundnut under drought-stressed and well-watered situations. [27] developed a novel method for the classification of different plant leaf diseases. [28] This study aimed to determine how well the RGB-image approach performed as a high-throughput phenotyping tool for evaluating groundnut leaf spot diseases and yield in a breeding program in Ghana.

Fruit defects occur for several reasons: nutrient deficiency, disease attack, and poor soil selection for crop cultivation. Fruits can be classified into two classes, i.e., healthy and defective, and it simplifies the defect stage identification of fruits in apples, pears, pomegranates, and litchi. The results are observed along with the fault phase by KNN, Google Net, and SVM [29]. The SVM method performs better as a fruit classifier for defect and stage prediction. The classification accuracy is 98%, which is better than the existing algorithms. The proposed system needs a network of a trained algorithm, which takes a long prediction time.

We have reviewed relevant research articles published on crop yield prediction through nutrient deficiency identification using different ML and DL algorithms. Though they are performing well on some problems, we have identified a few research gaps, as listed below.

1. No efficient method is available for estimating nutrient deficiency severity percentage accurately.
2. No appropriate methodology is available for crop yield prediction based on nutrient severity.

Based on these identified challenges, we enhanced the VGG16 deep learning model for deficiency classification and integrated it with the NSIM for yield loss prediction based
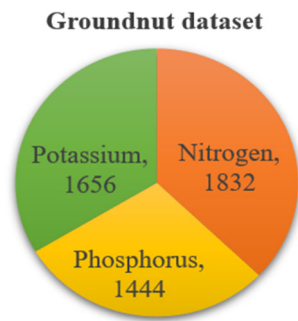
Groundnut dataset

Fig. 1 Groundnut dataset details [18]

on the low severity. The solution for the issue is illustrated in the next section.

3 Proposed methodology

3.1 Data acquisition

The Raspberry Pi (RPI) camera version 2 used to capture images has an attached RPI with the following specifications: aperture size—f/2.9; focal length—3.29 mm; resolution—1080 pixels; FOV—72.4 degrees; model B, 4 GB RAM, 64-bit quad-core CPU with 1.5 GHz frequency. The experiment was conducted during the cultivation period (January 2022–March 2022) by taking high-resolution imagery in the visible spectrum with an angle of 30 Degrees; the distance between the camera and to plant is 40 cm. The proposed methodology was tested on good and nutrient-deficient images of groundnut datasets [18, 30]. These images were captured at Kalwakurthy in Telangana, India, during the growth stage of the same crop for 30 days, 60 days, and 90 days of plant life.

3.2 Data preprocessing

This section focused on preprocessing the dataset and image for this proposed methodology. We analyzed our own image dataset of groundnut plant leaves [18]. The dataset details are explained in Fig. 1. This dataset was generated by collecting various nutrient-deficient plant leaf images deficient in N, P, and K.

The first stage of image classification methodology involved image preprocessing. The image preprocessing involved four steps: (1) background removal, (2) image resizes, (3) noise removal from the image, and (4) image augmentation. Image resizing is the second step that applies the FastStone image resizer software. The third stage is the noise removal stage. It extracted the purified image from the noise, using the total variation filter (TVF), as it smoothens the noise even with a minor signal-to-noise ratio while preserving the

edges [31]. Histogram equalization was used to modify the uniform contrast in the entire image. In this experiment, image augmentation is the final step of image preprocessing. Figure 1 shows the total images used in a dataset, and Fig. 2 shows the nutrient-deficient sample images.

3.3 Classification with the machine learning model

3.3.1 Feature extraction

The first phase of the standard ML method consisted of extracting the features, with the feature extraction steps detecting the key points and then detailing the neighboring local patches. The computer detected key points with areas of high information content. The number of data points forms the critical factor that considerably impacts the computing cost. A well-known image descriptor was utilized with *scale-invariant feature transform (SIFT)* features for this study. These features were separately extracted from images of plant leaves in the dataset. The significant difference between traditional ML and DL methods includes automatic feature extraction by the DL model.

3.3.2 Classification using SVM3

The SVM3-supervised learning algorithm generated a separating hyperplane to function the class's decision surface. The shape of the hyperplane is denoted in Eq. (1).

$$X^T W + b = 0 \tag{1}$$

where W = the input vector, X = the weight vector, and b = the bias. Using the procedure, a labeled training data were obtained during the training phase to make an ideal hyperplane. This hyperplane categorized the new test data. As a result, SVM3 was defined as a non-probabilistic, linear classifier for the binary data. Binary SVM3 was used to extend as a multi-class classifier function. The two most popular strategies known were one-versus-one and one-versus-all. Hence, this method used a one-versus-all model.

3.4 Classification with deep learning model (*CNN*)

The DL methods have a structure specifically designed for image inputs, such as the introduction of CNN. This CNN was introduced by [32]. The CNN retrieves the data from the images directly utilizing the convolution and pooling layers.

CNN can do feature extraction and classification of images. A feature vector is recognized as a collection of arrays that includes input and output for each stage. Each feature vector maps the output layers at specific locations in the information. At each stage, three layers are included: the filter bank, the nonlinearity, and the feature map. A typical
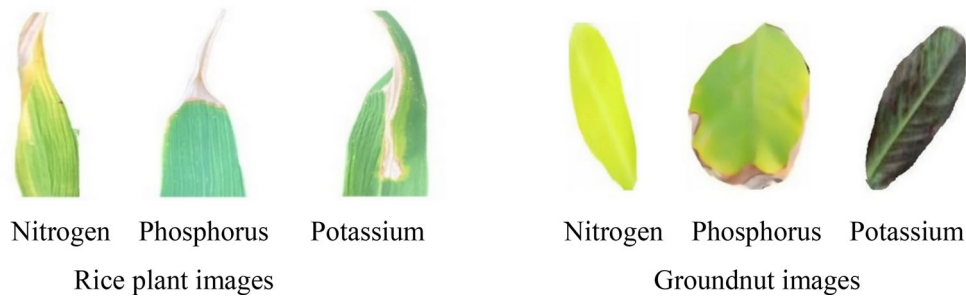
Nitrogen   Phosphorus   Potassium

Rice plant images

Nitrogen   Phosphorus   Potassium

Groundnut images

Fig. 2 Sample dataset images with a nutrient deficiency in leaf

Table 1 Dataset partition

	Training (80%)		Testing (10%)		Validation (10%)		Total	
	Rice	Groundnut	Rice	Groundnut	Rice	Groundnut	Rice	Groundnut
Nitrogen(N)	1602	1473	198	180	196	179	1996	1832
Phosphorus (P)	1234	1162	153	142	151	140	1538	1444
Potassium (K)	1422	1328	170	164	168	164	1760	1656

convolution network has a classification layer usually present after two or three stages. CNN has a 23-layer structure, which is contributed by batch normalization (BN), convolution, max pool, rectified linear unit (ReLu), fully connected layer (FCL), and classification layer. A summary of each layer's operations is described in this section. The convolution operation is expressed in Eq. (2).

$$C(r, s) = I * w(r, s) = \sum_p \sum_q I(r - p, s - q)w(p, q) \tag{2}$$

where I = input image with size (r, s), w = convolutional kernel with size (p, q). Mathematically, the expression for BN is given in Eq. (3).

$$\hat{x} = \frac{x_i - \mu_{mB}}{\sqrt{\sigma_{mB}^2 + \varepsilon}} \tag{3}$$

where x_i = input, $\mu_{mB,}$ and σ_{mB} depict the mean and variance in mini-batch for each input channel, respectively. The ϵ is applied to improve the numerical stability when the mini-batch variance is very small.

$$yi = \gamma \bar{x}i + \beta \tag{4}$$

During network training, the learning parameters β and γ are both updated. The activations are further scaled by factor γ and shifted by offset β. Equation (4) shows the mathematical expression for parameter updating and Eq. (5) shows the

expression of the ReLu activation function.

$$\text{ReLu}(x) = \max\{x, 0\} \tag{5}$$

The ReLu activation output function ranges from zero to one. The result flips to zero if the output value is less than zero. When the ReLu activation function is used, convergence occurs more rapidly. Equation (6) depicts the mathematical expression for maximum pooling (6).

$$m_{i, j} = \max_{a, b} x_{i+a-1, j+b-1}, 1 \leq a, b \leq m \tag{6}$$

where m = kernel width and the maximization are simultaneously over a and b, the SoftMax layer restraints with a range (0,1) and is denoted in Eq. (7).

$$\text{soft}(x_i) = \frac{\exp x_i}{\sum_{j=1}^{n} \exp x_j} \tag{7}$$

Table 1 shows the details of the dataset partition for training, testing, and validation.

3.5 Proposed eVGG16 for nutrient deficiency classification

The proposed transfer learning methodology included two blocks, they are (1) nutrient classification block (eVGG16) and (2) nutrient severity identification method (NSIM). A state-of-art image classification task [33–35] was used to classify the images of nutrient deficiencies with the help of pre-trained eVGG16 architecture.
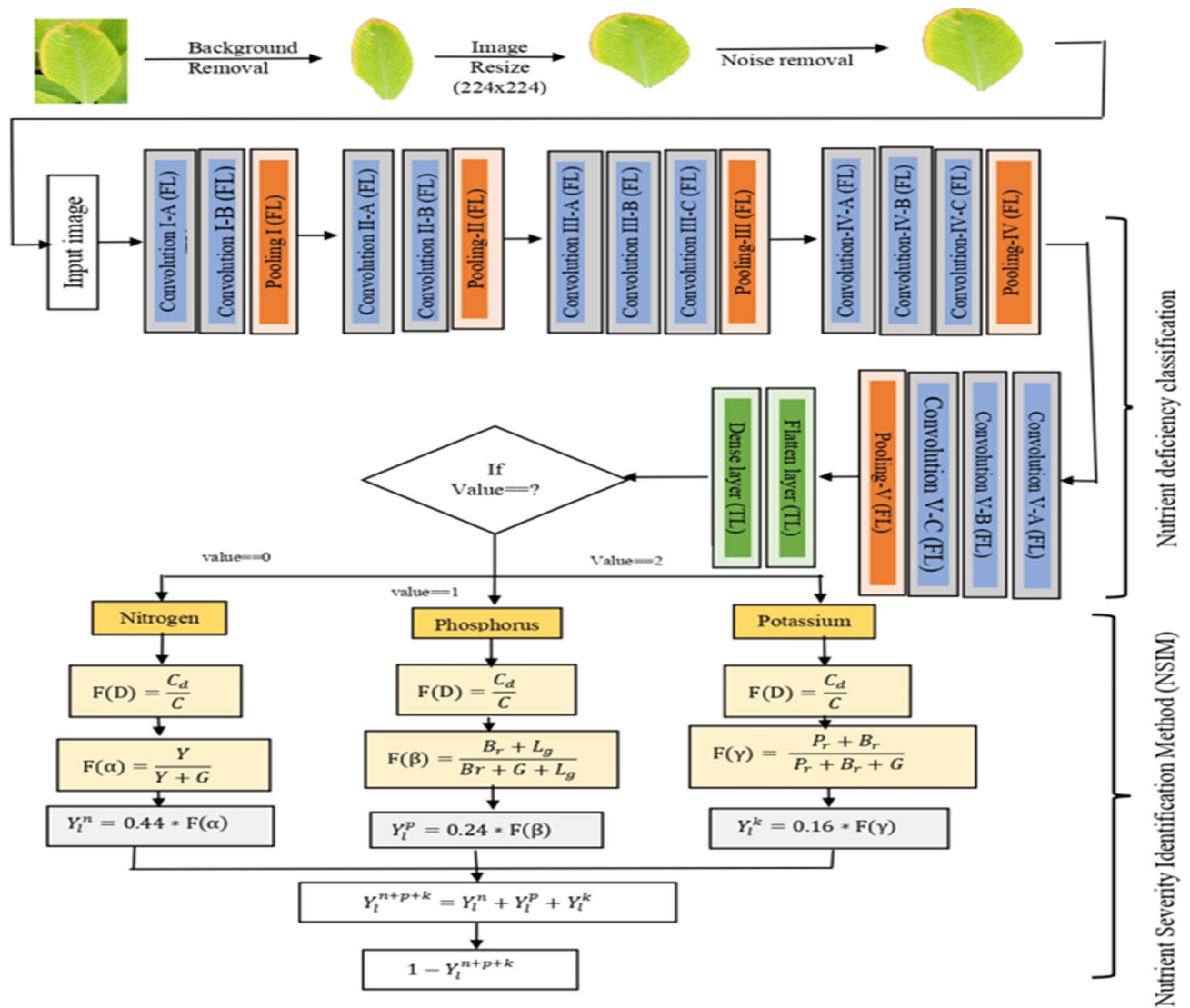
Fig. 3 Integrated architecture of the proposed system

This model was performed with the 3×3 convolution and 2×2 pooling from start to end and had a homogeneous structure. The model was built using Kera's higher-level Python module running on top of the open-source DL framework TensorFlow to classify the agricultural field images with nutrient deficiencies. The convolution steps in the x and y directions were (1, 1) pixels, and these images were processed through a stack of convolution layers of convolution filter size of (3×3). The spatial size is preserved with hyperparameter padding one after convolution. A pooling size of (2, 2) pixels with a maximum of five max-pooling layers is utilized to accomplish the spatial pooling. ReLu (rectified linear unit) activation function was utilized to activate all hidden layers.

In contrast, the SoftMax function activated the final layer to ensure the anticipated likeability output values were between 0 and 1. The network utilizes 32 batch sizes and 70 epochs. The network was improved using a logarithmic loss function of categorical cross-entropy and the gradient descent optimization algorithm. The model was fed with resized images of an input size of 224×224 pixels. Then, a series of two convolutional and pooling layers operated as feature extractors, followed by an output layer and a fully connected layer that interpreted the features. The input image was reflected by the three neurons in the output layer that are categorized as three nutritional deficient categories. Figure 3 shows the proposed system's architecture, which has TL, which stands for the trainable layer, and FL, which stands for the frozen layer. The proposed model classified the plants into N, P, and K deficiencies. These classified plant images were sent to the NSIM block, which has been explained in a later section.

3.5.1 Classification metrics

A comparison of the proposed model with the existing models [6, 23] by utilizing the different performance analysis metrics has been given below. For evaluating the model, the following metrics were considered. Accuracy could be defined as the percentage of the complete set of properly categorized images. The accuracy formula is given in Eq. (8). Where TN = true negative, TP = true positive, FN = false negative, FP = false positive.

$$\text{Accuracy} = \frac{\text{TN} + \text{TP}}{\text{TN} + \text{FP} + \text{FN} + \text{TP}} \tag{8}$$

The number of correctly classified samples that are classified as positive are represented in the precision score, as denoted in the Eq. (9).

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{9}$$

The recall score was the ratio between the positive samples categorized correctly as favorable to the number of positive examples. The recall formula was given as Eq. (10)

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{10}$$

The harmonic means of precision and recall values are written in F1-score. The F1-score formula is given in Eq. (11).

$$F1\text{score} = 2 * \frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}} \tag{11}$$

3.5.2 Nutrient severity identification and yield prediction in crop

Nutrient severity quantifies the rate at which a plant lacks essential nutrients (nitrogen, potassium, and phosphorus) for its growth and development. This evaluation helps in determining the significant effect of nutrient deficiency on the overall health and yield of the crop. Traditionally, the formers/experts can estimate the severity of nutrient deficiencies in plants by observing the plant's physical characteristics like color symptoms. This normal process requires the farmer to analyze the plant's leaves for identifying the characteristic indicators of nutrient deficiencies, such as yellowing leaves, stunted growth, discoloration, it's stems, and overall growth. This visual assessment might be supplemented with frequent laboratory soil testing to ascertain soil nutrient levels, that provides further insight into identifying nutrient deficiency severities.

The classification block bifurcates the plant images to N, P, and K nutrient deficiencies according to the plant nutrient

Table 2 Nutrient symptoms information [36], [37]

Sno	Nutrient	Symptoms
1.	Nitrogen	The yellow color pigmentation is present in the plant leaves
2.	Phosphorus	The brown and light green color pigmentation is present in the plant leaves
3.	Potassium	The purple and brown color pigmentation is present in the plant leaves

deficiency symptoms. The plant nutrient deficiency symptoms are shown in Table 2 [36], [37]). In the nutrient severity block, nutrient deficiency severity is calculated. This nutrient deficiency severity identification, nutrient severity identification method is proposed (NSIM). This NSIM method results are compared with the SLIKM [6] and RDBSCAN [38] algorithms. The nutrient severity percentage is evaluated using Eq. (12). For example, if a plant is classified as 'N' nutrient deficient. According to Table 2, yellow color pigmentation is estimated using Eq. (13). This yellow color pigmentation percentage is the nitrogen deficiency severity percentage. According to Table 2 and Eqs. (14) and (15), the 'P' and 'K' severity percentage is calculated.

$$F(D) = \frac{C_d}{C} \tag{12}$$

$$F() = \frac{Y}{Y + G} \tag{13}$$

$$F(\beta) = \frac{B_r + L_g}{Br + G + L_g} \tag{14}$$

$$F(\gamma) = \frac{P_r + B_r}{P_r + B_r + G} \tag{15}$$

where F = function for identifying the nutrient deficiency severity, Cd = nutrient deficiency percentage, C = total crop, $\alpha, \beta, \gamma,$ = nitrogen, phosphorus, and potassium deficiency, respectively, Y, Br, Lg, Pr, and G = yellow, brown, light green, purple, and green colour, respectively. The loss of yield due to N, P, and K deficiency was distinct because crop yield has differently impacted by N, P, and K [36]. The yield loss percentages for the N, P, and K are mathematically denoted below.

$$Y_l^{n+p+k} = Y_l^n + Y_l^p + Y_l^k \tag{16}$$

where, Y_l^n, Y_l^p, Y_l^k are the yield loss due to the N, P, and K deficiency.

Table 3 Final crop yield loss and obtained yield results

S.L no	Method	Yield loss percentage	Obtained yield
1.	SLIKM [6]	27.17	72.82
2.	RDBSCAN [38]	28.20	71.8
3.	**NSIM [proposed]**	**29.70**	**70.3**

The proposed model results are highlighted in bold

Table 4 Groundnut crop nutrient deficiency—Classification accuracy

S.l no	Model name	Accuracy (groundnut dataset) (%)
1.	**Proposed (eVGG16)**	**98**
2.	SVM3 [6]	86
3.	TLCNN [23]	96

The proposed model results are highlighted in bold

3.5.3 Crop yield analysis on groundnut dataset [18].

This section aims to comprehensively explain the relations among crop yield loss and the severity percentage of nutrition deficiencies in groundnut plants. The impact of nutrient deficiencies on crop yield loss is proportional to the severity of the deficiencies and can vary among plant species. The Eqs. (17) to (19) states the relations among the deficiency severity of nutrient in groundnut plants and the corresponding percentage of yield loss. According to these equations, the severity of P substantially influences the decrease in crop yield loss because it involves seed formation and root development, whereas the severity of N has a relatively smaller impact on crop yield because it involves leaf growth and photosynthesis.

The groundnut crop images as a private dataset were used for further analysis of crop yield. The plants were classified into N, P, and K classes based on nutrient deficiency. Then, the nutrient-deficient severity was identified based on the data of the symptoms in Table 2 and Eqs. (12) to (15). This severity percentage was used as an input to estimate crop

yield loss. A varied crop yield loss was due to N, P, and K. The mathematical expressions for yield loss were represented in Eqs. (17) to (19).

$$Y_l^n = 0.18 * F() \tag{17}$$

$$Y_l^p = 0.31 * F(\beta) \tag{18}$$

$$Y_l^k = 0.29 * F(\gamma) \tag{19}$$

Integrated nutrient management in groundnut—a farmer's manual [39] was utilized for the coefficient constant values in Eqs. (17–19). Where, the Y_l^n, Y_l^p, and Y_l^k represents the yield loss due to N, P, and K deficiency, respectively, whereas the $F(\alpha)$, $F(\beta)$, and $F(\gamma)$ represents the Computation function for N, P, and K severity percentages in the plant, respectively. The variability of yield loss percentage can be explained by the severity of nitrogen (N), phosphorus (P), and potassium (K), as evidenced by Eqs. (17) to (19). The crop yield loss caused by nitrogen is 18% in total, while phosphorus is responsible for a loss of 31%, and potassium adds to a reduction of 29% in overall yield. The relationship between the severity percentage and the ensuing yield loss is crucial in determining the crop yield decline rate. These equations represented the crop yield loss due to the individual elements of N, P, and K deficiency. The total crop yield loss was derived using Eq. (16). The obtained crop yield was defined as the total yield subtracted from crop yield loss due to N, P, and K elements.

4 Result analysis

The software tools for performing the experiment are Python programming language in the anaconda environment. The hardware configurations of the system are Intel(R) Core (TM) i5-1035G1 CPU @ 1.00 GHz, 1.19 GHz (Min, Max), 8 GB RAM, and 512 GB HDD.

Table 5 Classification performance—precision, recall, and F1-score of rice dataset

	SVM3 *[[6]]*			TLCNN *[[23]]*			Proposed (eVGG16)		
	Precision	Recall	F1-score	Precision	Recall	F1-score	Precision	Recall	F1-score
Nitrogen	0.88	0.97	0.88	0.91	1.00	0.95	**0.98**	**1.00**	**0.99**
Phosphorus	0.98	0.98	0.98	1.00	0.99	1.00	**1.00**	**1.00**	**1.00**
Potassium	0.97	0.79	0.87	1.00	0.94	0.97	**1.00**	**0.99**	**0.99**

The proposed model results are highlighted in bold

Table 6 Nutrient Severity Percentage of Rice plant

Method	Nutrient	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
SLIKM [6]	N	35.84	26.77	28.40	53.19	48.07	55.20	39.28	33.70	36.93	42.68
	P	54.14	46.65	53.17	41.53	36.83	31.11	21.19	26.56	35.93	15.40
	K	15.39	14.39	15.78	12.85	28.30	18.14	13.95	26.56	35.93	12.88
RDBSCAN [38]	N	36.45	27.23	28.50	53.39	52.8	57.75	41.98	34.10	38.75	43
	P	57.46	46.92	53.49	41.92	37.21	31.40	22.01	26.80	36.19	15.90
	K	16.12	15.41	16.12	12.91	28.90	18.98	15.20	16.90	11.02	12.89
Proposed (NSIM)	N	**38.26**	**29.12**	**30.01**	**54.15**	**53.15**	**59.16**	**43.01**	**35.82**	**39.92**	**44.17**
	P	**58.15**	**47.42**	**54.15**	**43.19**	**38.30**	**33.30**	**24.17**	**27.30**	**37.43**	**16.40**
	K	**17.51**	**15.90**	**16.90**	**13.52**	**31.22**	**20.20**	**16.19**	**18.93**	**12.41**	**12.99**

The proposed model results are highlighted in bold

Table 7 Yield loss and obtained yields of Rice crop

	Method	Nutrient	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
Yield loss	SLIKM	N	15.76	11.77	12.49	23.40	21.15	24.2	17.28	14.82	16.24	18.75
	RDBSCAN		16.03	11.98	12.49	23.49	21.15	25.41	18.47	15.00	17.05	18.92
	NSIM		**16.83**	**12.81**	**13.20**	**23.82**	**23.38**	**26.03**	**18.92**	**15.76**	**17.56**	**19.43**
Obtained yield	SLIKM		84.24	88.23	87.51	76.6	78.85	75.72	82.72	85.18	83.76	81.25
	RDBSCAN		83.97	88.02	87.51	76.51	78.85	74.59	81.53	85.0	82.95	81.08
	NSIM		**83.17**	**87.19**	**86.8**	**76.18**	**76.62**	**73.97**	**81.08**	**84.24**	**82.44**	**80.57**
Yield loss	SLIKM RDBSCAN NSIM	**P**	12.99	11.96	12.76	9.96	8.83	7.46	5.08	6.37	8.62	3.69
			13.79	11.26	12.83	10	8.93	7.5	5.2	6.4	8.68	3.81
			14	**11.38**	**12.99**	**10.36**	**9.1**	**8.0**	**5.78**	**6.65**	**9.0**	**4.0**
Obtained yield	SLIKM RDBSCAN NSIM		87.01	88.04	87.24	90.04	91.17	92.5	94.9	93.6	91.3	96.3
			86.21	88.74	87.17	90	91.07	92.5	94.8	93.6	91.3	96.2
			86	**88.62**	**87.01**	**89.64**	**90.9**	**92**	**94.2**	**93.3**	**91.0**	**96.0**
Yield loss	SLIKM	**K**	2.46	2.30	2.52	2.05	4.52	2.90	2.23	2.43	1.74	2.06
	RDBSCAN		2.57	2.46	2.57	2.06	4.62	3.03	2.43	2.70	1.71	2.08
	NSIM		**2.80**	**2.54**	**2.70**	**2.16**	**5.0**	**3.2**	**2.59**	**3.02**	**1.98**	**2.20**
Obtained yield	SLIKM		97.54	97.7	97.48	97.95	95.48	97.1	97.7	97.5	98.2	97.9
	RDBSCAN		97.43	97.54	97.43	97.94	95.38	96.9	97.5	97.3	98.2	97.9
	NSIM		**97.2**	**97.46**	**97.3**	**97.84**	**95**	**96.8**	**97.4**	**96.9**	**98.0**	**97.8**

The proposed model results are highlighted in bold

4.1 Classification of nutrient deficiency with SVM3, TLCNN, and eVGG16

This section describes the classification of groundnut nutrient deficiency. The proposed classification model classifies the groundnut plant into various nutrient-deficient plants, such as N, P, and K. Classification metrics fall into four categories: accuracy, precision, recall, and F1-score. Using the proposed model, these classification parameters are compared with existing SVM3 and TLCNN models. Table 4 shows the accuracy of the classification models.

The SVM3 model classification accuracy is 86% due to the working of SVM3 on limiting the constraint optimization and handcrafted features. It is a supervised ML algorithm.

Table 8 Classification performance—Precision, recall, and F1-score of Groundnut dataset

	SVM3			TLCNN			Proposed (eVGG16)		
	Precision	Recall	F1-score	Precision	Recall	F1-score	Precision	Recall	F1-score
Nitrogen	0.88	0.97	0.88	0.91	1.00	0.95	**0.98**	**1.00**	**0.99**
Phosphorus	0.98	0.98	0.98	1.00	0.99	1.00	**1.00**	**1.00**	**1.00**
Potassium	0.97	0.79	0.87	1.00	0.94	0.97	**1.00**	**0.99**	**0.99**

The proposed model results are highlighted in bold

Table 9 Severity identification method

Input image (N):	Severity detection	Percentage: 9.95%
Input image (P):	Severity detection	Percentage: 2.61%
Input image(K):	Severity detection	Percentage: 3.57%

The TLCNN model's classification accuracy is 96% because it can reduce dimensionality without losing information. The proposed model was pre-trained on the ImageNet challenge with 1000 categories with fewer data, providing accuracy and learning the backbone features using the convolution and pooling layers. This results in 98% classification accuracy. The principle behind the operation of this model is the number of layers that increases the model's classification accuracy. Table 3 represent a final crop yield loss of rice plant. Tables 4 and 5 represent a Groundnut and rice crop dataset in terms of classification performance. Tables 6 and 7 show the final crop yield loss and obtained crop yield results. For better understanding, the process of severity calculation on images and the identified severity with their % is shown in Table 9.

4.1.1 Confusion matrix and the classification performances on groundnut dataset

The confusion matrix provides a summary of the performance of the classification model. Figure 4 details the confusion matrix of image classification of the SVM3, TLCNN, and proposed models. Out of 100% N images, The SVM3 model classified correctly 97% of N images due to the poor classification accuracy of SVM3, and it misclassified 3% of the N images. The TLCNN model classified 100% of N images and has misclassified 0% of N images due to its classification accuracy of 96%. The proposed model classified 100% of N images correctly, as it had 98% of classification accuracy. Table 8 represents a Classification performance—-Precision, recall, and F1-score of the Groundnut dataset.
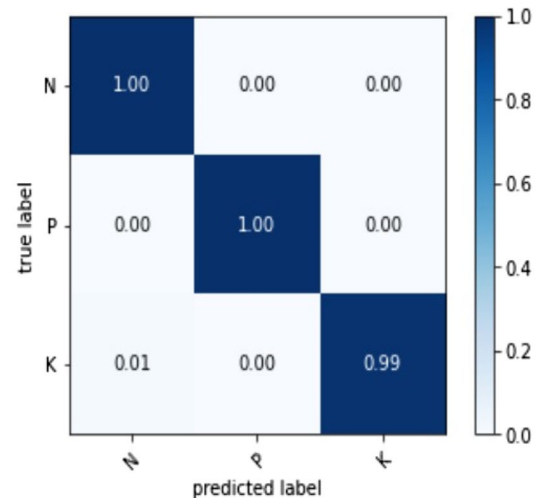
Table 9 comprehensively explains the methodology employed for estimating nutrition severity percentages. Let us examine the initial row of this Table as an illustrative example. The provided image in this context shows a visual representation of a nitrogen deficit. The image undergoes the
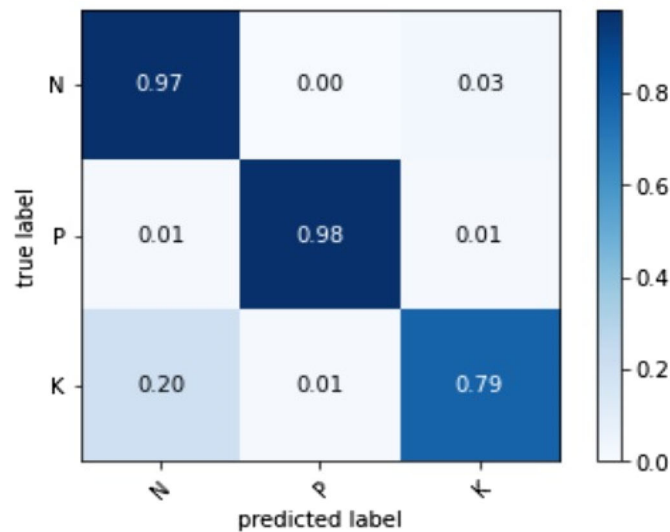
(a) SVM3 confusion matrix.
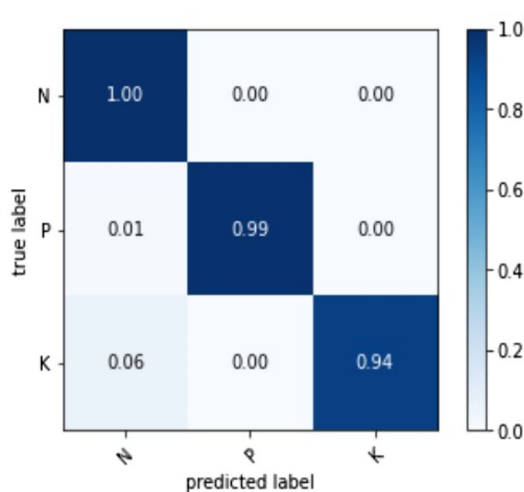
(b) TLCNN confusion matrix.

(c) Proposed model confusion matrix.

Fig. 4 Confusion matrix of groundnut dataset

severity identification process to discover the severity percentage of N. Thermal image of the leaf that shows the portion of leaf area affected with N severity findings are shown in the second column of the first row, whereas the calculated severity percentage is determined as 9.95% as shown in the third column of this row. Similarly, this table also presents the percentages indicating the severity of P and K are provided in the second and third row of the table.

Table 4 lists the SVM3, TLCNN, and proposed classification models' accuracy, recall, and F1-score for test data sets. N precision in the SVM3 model is 88%, while P precision is

98%. N has a lower recall value (97%) than P (98%), which has a higher recall value. When precision metrics were considered, P and K outperformed N. Recall metrics showed that P performed better than K and N. The TLCNN model outperformed the SVM3 model in testing. The SVM3 model's N precision score is 88%, which is lower than the N precision score of the TLCNN model (91%). The F1-score results for the TLCNN and SVM3 models range from 95% (for N) to 100% (for P) and 88% (for N) to 98% (for P), respectively. The proposed model's F1-score values vary from 99% (K) to
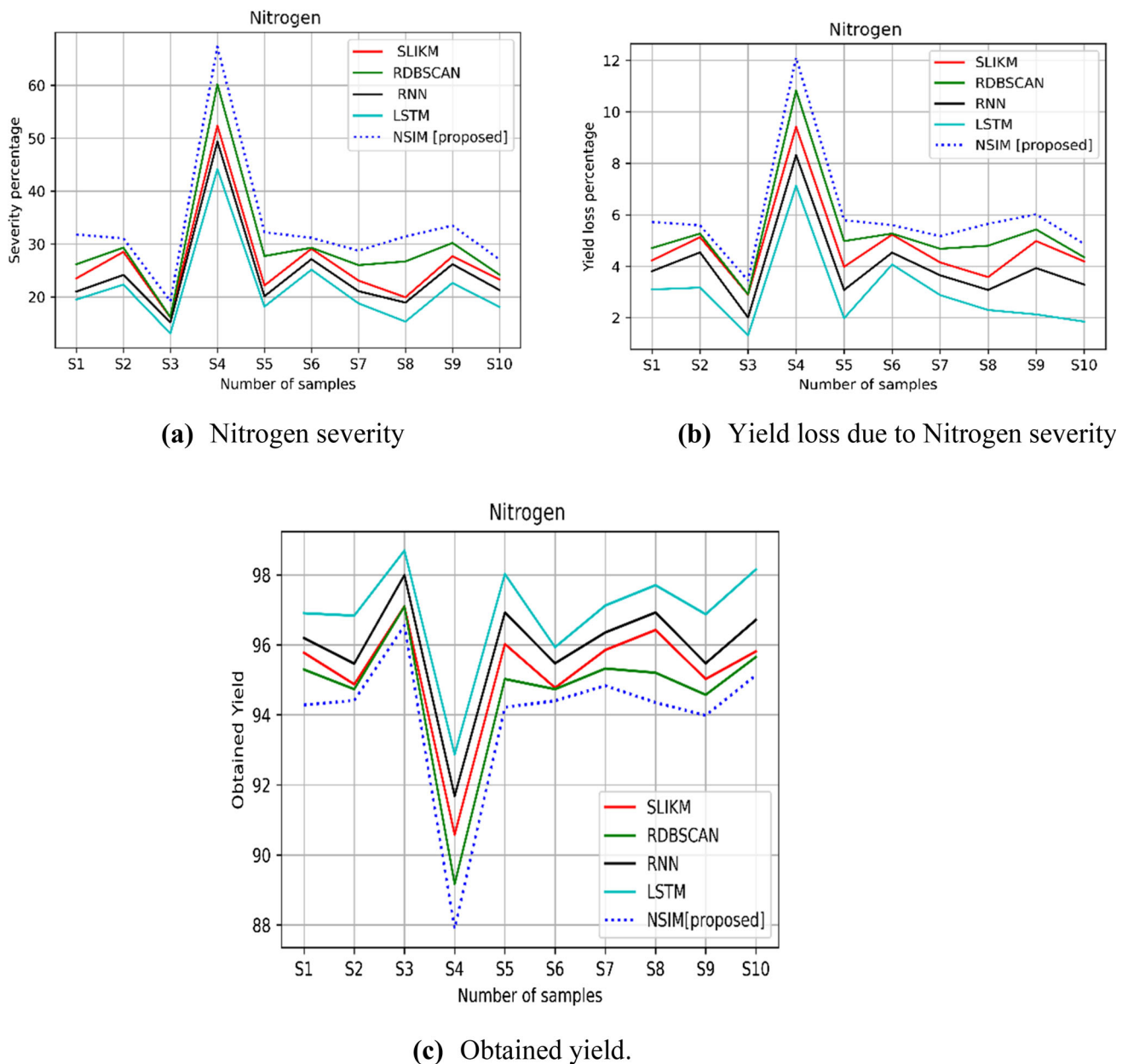
(a) Nitrogen severity

(b) Yield loss due to Nitrogen severity

(c) Obtained yield.

Fig. 5 Analysis of the nitrogen effect

100%. (P). The suggested model outperforms existing models when precision, recall, and F1 scores are considered for all models.

4.1.2 Nutrient severity and crop yield prediction analysis on the groundnut dataset

The study of the groundnut crop's nutritional severity is described in this section. The symptoms of groundnut crop nutrient deficiency [37] are listed in Table 2. Figure 5a displays the results for the severity of the N percentage. The number of samples is represented on the x-axis, and the

severity percentages are shown on the y-axis in this graph. Equation (17) provides the agricultural production loss due to N deficiency. Similarly, yield loss is obtained using Eqs. (18), and (19) due to nutrient deficiency of P and K. Figure 5(c) shows the results of the yield loss percentage due to N deficiency and the results of obtained crop yield percentage. The constant values in Eqs. (17), (18), and (19) are taken from [39]. We have used a different set of methods for comparing the prediction of yield because the set of DL models used for comparing the accuracy of the proposed eVGG16 is just the classification method, and they have yet to be used by any researchers previously for crop yield predicting. These new
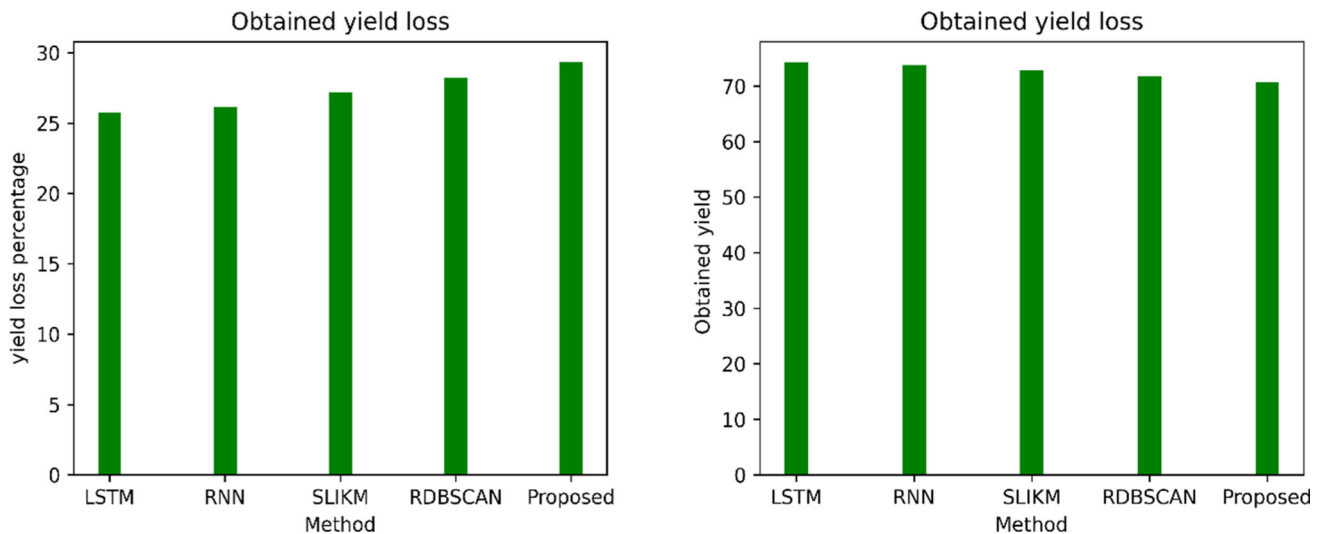
Fig. 6 Final yield loss and obtained yield

sets of approaches have been used previously by researchers for predicting crop yield. Hence, these new sets of methods (SLIKM, RDBSCAN) are applied to the classification results of the proposed eVGG16 for crop yield prediction along with the proposed NSIM method.

In Fig. 5a, the percentage N nutrient severity of several samples (S1, S2, S3…S10) is considered—the nutrient severity identification compared with other existing methods like SLIKM, RDBSCAN, RNN, and LSTM. The percentage nutrient severity of sample one (S1) using SLIKM, RDBSCAN, RNN, and LSTM is 23.51%, 26.17%, 31.92%, 22.17%, and 20%, respectively. Similarly, sample two (S2) using SLIKM, RDBSCAN, RNN, and LSTM are 28%, 29%, 24%, 22%, and 31%, respectively. When the number of samples increased based on the proposed model achieved 67% of the highest percentage. The NSIM algorithm demonstrated superior performance compared to the SLIKM and RDBSCAN algorithms in accurately determining severity percentages. This performance can be mainly attributed to the proposed method's improved classification accuracy. As a result, the model can effectively categorize images depicting nutrient deficiencies, providing a higher and more reliable measure of severity. For all sample results, the NSIM method predicts accurate results. Figure 5b shows the percentage of crop yield loss due to nitrogen deficiency. The yield losses using this yield, SLIKM, RDBSCAN, and NSIM obtained using Eq. (17) are 4.23%, 4.71%, and 5.74%, respectively. The crop yield loss by the NSIM method is higher than SLIKM and RBSCAN due to the higher value of N severity in the NSIM method. Figure 5c details the obtained crop yield from a different sample of a crop field. This crop yield obtained is calculated by subtracting the yield loss rate (1-yield) from the total yield.

The next stage in the proposed model is to identify the final yield loss due to all nutrients. All sample values are averaged and used in Eq. (16) to identify this crop yield loss. The final yield loss and obtained yield results are shown in Fig. 6.

Yield losses due to N, P, and K deficiency using the SLIKM, RDBSCAN, and NSIM methods are 27.17%, 28.20%, and 29.32%, respectively. The yields obtained using the SLIKM, RDBSCAN, and NSIM methods are 72.83%, 71.8%, and 70.68%, respectively. The yield loss of the NSIM method is more significant than other methods because it accurately predicts nutrient levels. Yields obtained using the NSIM method are lower than others due to the higher yield loss rate. Therefore, the NSIM method works efficiently.

5 Conclusion

The study demonstrates the applications of transfer learning in categorizing the nutrient deficiency of leaf images in groundnuts crop (dataset). The proposed methodology performs better than the existing TLCNN and SVM3 algorithms in identifying nutrient deficiency in plant leaves. To identify deficiencies, we have used two different datasets: rice plant and groundnut images. The proposed methodology classifies the plant leaf images into N, P, and K classes of deficiencies. Based on this information, crop yield loss due to each nutrient deficiency is estimated. After comparing both datasets, groundnut crops nutrient deficiency classification accuracy achieved 98%. To determine the crop yield loss, the NSIM method is proposed. This method performed better than the existing SLIKM and RDBSCAN methodologies. The experimental results of the proposed transfer

learning method performance were better for the classification of nutrients than the other existing methods. The proposed NSIM achieved a comparable performance over the prevailing methods. In the future, the researchers may use this enhanced transfer learning-based model for the nutrition deficiency identification of other dataset images and live plant images. The proposed NSIM can also be extended for the severity nutrition identification and yield loss prediction of other crop fields' live images.

Author contributions KV performed concepts, development of methodologies, dataset collection and creation, experimentation, results analysis, and writing of the original draft; KJN analyzed concepts, experimentation, results analysis, writing, document review, editing, and overall supervision. All authors read before submission and approved the final manuscript for submission.

Data availability statement The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Declarations

Conflict of interests The authors declare no competing interests.

Ethical approval: All authors have seen and agreed with the contents of the manuscript and are looking forward to publishing this paper on this journal.

Consent to participate All authors gave explicit consent to participate in this work.

Consent for publication: All authors gave explicit consent to publish this manuscript.

References

1. Khaki, S., Wang, L., Archontoulis, S.V.: A cnn-rnn framework for crop yield prediction. Front. Plant Sci. **10**, 1750 (2020)
2. Priya, P.K., Yuvaraj, N.: An IoT based gradient descent approach for precision crop suggestion using MLP. J. Phys. Conf. Ser. **1362**(1), 012038 (2019)
3. Kaur, S., Pandey, S., Goel, S.: Semi-automatic leaf disease detection and classification system for soybean culture. IET Image Proc. **12**(6), 1038–1048 (2018)
4. Tardieu, F., Cabrera-Bosquet, L., Pridmore, T., Bennett, M.: Plant phenomics, from sensors to knowledge. Curr. Biol. **27**(15), R770–R783 (2017)
5. Zhou, C., Le, J., Hua, D., He, T., Mao, J.: Imaging analysis of chlorophyll fluorescence induction for monitoring plant water and nitrogen treatments. Measurement **136**, 478–486 (2019)
6. Zermas, D., Nelson, H.J., Stanitsas, P., Morellas, V., Mulla, D.J., Papa Nikolopoulos, N.: A methodology for the detection of nitrogen deficiency in corn fields using high-resolution RGB imagery. IEEE Trans. Autom. Sci. Eng. **18**(4), 1879–1891 (2021)
7. Asaari, M.S.M., Mishra, P., Mertens, S., Dhondt, S., Inzé, D., Wuyts, N., Scheunders, P.: Close-range hyperspectral image analysis for the early detection of stress responses in individual plants in a high-throughput phenotyping platform. ISPRS J. Photogramm. Remote Sens. **138**, 121–138 (2018)
8. Neumann, K., Klukas, C., Friedel, S., Rischbeck, P., Chen, D., Entzian, A., et al.: Dissecting spatiotemporal biomass accumulation in barley under different water regimes using high-throughput image analysis. Plant Cell Environ. **38**(10), 1980–1996 (2015)
9. Bai, G., Ge, Y., Hussain, W., Baenziger, P.S., Graef, G.: A multisensor system for high throughput field phenotyping in soybean and wheat breeding. Comput. Electron. Agric. **128**, 181–192 (2016)
10. Minervini, M., Scharr, H., Tsaftaris, S.A.: Image analysis: the new bottleneck in plant phenotyping [applications corner]. IEEE Signal Process. Mag. **32**(4), 126–131 (2015)
11. Panwar, R., Goyal, K., Pandey, N., Khanna, N.: Imaging system for classification of local flora of Uttarakhand region. In: 2014 International Conference on Power, Control and Embedded Systems (ICPCES) (pp. 1–6). IEEE (2014)
12. Humplík, J.F., Lazár, D., Husičková, A., Spíchal, L.: Automated phenotyping of plant shoots using imaging methods for analysis of plant stress responses–a review. Plant Methods **11**(1), 1–10 (2015)
13. Singh, A., Ganapathysubramanian, B., Singh, A.K., Sarkar, S.: Machine learning for high-throughput stress phenotyping in plants. Trends Plant Sci. **21**(2), 110–124 (2016)
14. Kheirkhah, F. M., & Asghari, H. (2018). Plant leaf classification using GIST texture features, IET Comput. Vis. 13 (4) (2018) 369–375.
15. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature **521**(7553), 436–444 (2015)
16. Venkatesh, K., Naik, K. J.: Deep learning for macro-nutrient deficiency identification in the groundnut plants. In: 8th International Conference on Computing in Engineering and Technology (ICCET 2023), pp.193–198. IET (2023)
17. Arnal Barbedo, J.G.: Digital image processing techniques for detecting, quantifying and classifying plant diseases. Springerplus **2**(1), 1–12 (2013)
18. Venkatesh, K., Naik, J.: Groundnut nutrient deficiency dataset (G_N_ dataset) (2022). Accessed on May 2, 2022. https://drive.google.com/file/d/1xNRX9gAlqM-ToWaJ9VbhIXh725dIcyCT/view?usp=sharing
19. Buchaillot, M.L., Gracia-Romero, A., Vergara-Diaz, O., Zaman-Allah, M.A., Tarekegne, A., Cairns, J.E., Kefauver, S.C.: Evaluating maize genotype performance under low nitrogen conditions using RGB UAV phenotyping techniques. Sensors **19**(8), 1815 (2019)
20. Savithramma, D.L., Gangappa, E., Shankar, A.G.: Identification of high-yielding and late leaf spot disease-tolerant families in groundnut (*Arachis hypogaea* L.). Int. J. Environ. Clim. Change **13**(6), 461–471 (2023)
21. Chapu, I., Okello, D.K., Okello, R.C.O., Odong, T.L., Sarkar, S., Balota, M.: Exploration of alternative approaches to phenotyping of late leaf spot and groundnut rosette virus disease for groundnut breeding. Front. Plant Sci. **13**, 912332 (2022)
22. Mahmud, A., Esakki, B., Seshathiri, S.: Quantification of groundnut leaf defects using image processing algorithms. In: Proceedings of International Conference on Trends in Computational and Cognitive Engineering: Proceedings of TCCE 2020, pp. 649–658. Springer Singapore, Singapore (2020)
23. Azimi, S., Kaur, T., Gandhi, T.K.: A deep learning approach to measure stress level in plants due to nitrogen deficiency. Measurement **173**, 108650 (2021)
24. Oteng-Frimpong, R., Karikari, B., Sie, E.K., Kassim, Y.B., Puozaa, D.K., Rasheed, M.A., Ozias-Akins, P.: Multi-locus genome-wide association studies reveal genomic regions and putative candidate genes associated with leaf spot diseases in African groundnut

(*Arachis hypogaea* L.) germplasm. Front. Plant Sci. **13**, 1076744 (2023)

25. Gao, X., Chai, H.H., Ho, W.K., Mayes, S., Massawe, F.: Deciphering the molecular basis for photosynthetic parameters in Bambara groundnut (*Vigna subterranea* L. Verdc) under drought stress. BMC Plant Biol. **23**(1), 287 (2023)

26. Venkatesh, K., Naik, K. J.: An IoT framework for groundnut crop yield prediction using K-means algorithm. In: 2021 International Conference on Data Analytics for Business and Industry (ICDABI), pp. 266–271. IEEE (2021)

27. Kaur, P., Harnal, S., Gautam, V., Singh, M.P., Singh, S.P.: A novel transfer deep learning method for detection and classification of plant leaf disease. J. Amb. Intell. Humaniz. Comput. **14**, 1–18 (2022)

28. Sie, E.K., Oteng-Frimpong, R., Kassim, Y.B., Puozaa, D.K., Adjebeng-Danquah, J., Masawudu, A.R., et al.: RGB-image method enables indirect selection for leaf spot resistance and yield estimation in a groundnut breeding program in Western Africa. Front. Plant Sci. **13**, 957061 (2022)

29. Yogesh, Dubey, A.K., Ratan, R., Rocha, A.: Computer vision based analysis and detection of defects in fruits causes due to nutrients deficiency. Clust. Comput. **23**, 1817–1826 (2020)

30. Weeraphat Raksarikon (2020) Nutrient deficiency symptoms in rice cover image. Accessed 17 May 2022. https://www.kaggle.com/datasets/guy007/nutrientdeficiencysymptomsinrice/metadata

31. Chambolle, A., Caselles, V., Cremers, D., Novaga, M., Pock, T.: An introduction to total variation for image analysis. Theor. Found. Numer. Methods Sparse Recov. **9**(263–340), 227 (2010)

32. LeCun, Y., Kavukcuoglu, K., Farabet, C.: Convolutional networks and applications in vision. In: Proceedings of 2010 IEEE International Symposium on Circuits and Systems, pp. 253–256. IEEE (2010)

33. Sunil, G.C., Zhang, Y., Koparan, C., Ahmed, M.R., Howatt, K., Sun, X.: Weed and crop species classification using computer vision and deep learning technologies in greenhouse conditions. J. Agric. Food Res. **9**, 100325 (2022)

34. Tan, W., Liu, P., Li, X., Liu, Y., Zhou, Q., Chen, C., et al.: Classification of COVID-19 pneumonia from chest CT images based on reconstructed super-resolution images and VGG neural network. Health Inf. Sci. Syst. **9**(1), 1–12 (2021)

35. Akther, J., Harun-Or-Roshid, M., Nayan, A. A., & Kibria, M. G. (2021, December). Transfer learning on VGG16 for the classification of potato leaves infected by blight diseases. In *2021 Emerging Technology in Computing, Communication and Electronics (ETCCE)* (pp. 1–5). IEEE.

36. Government of Tamilnadu: Expert system for paddy, "Nutrient management" (2020). http://www.agritech.tnau.ac.in/expert_system/paddy/nutrientmanagement.html#disorders-NoValue-

37. Singh, A.L., Basu, M.S., Singh, N.B.: Mineral disorders of groundnut. National Research Centre for groundnut (ICAR), Junagadh (2004)

38. Chen, H., Liang, M., Liu, W., Wang, W., Liu, P.X.: An approach to boundary detection for 3D point clouds based on DBSCAN clustering. Pattern Recogn. **124**, 108431 (2022)

39. Singh, A.L., Basu, M.S.: Integrated Nutrient Management in Groundnut-A Farmer's Manual. Government of India, New Delhi (2005)