# Attention-driven Deep Convolutional Neural Network for Maize Leaf Nutrient Deficiency Recognition

Gaikar Shubhangi [1,2] *, Madhukar Zambare [3], Arvind Shaligram [4], and Kailash Sapnar [1]

[1] Fergusson College Pune, Savitribai Phule Pune University, Pune, India
[2] MAEER's MIT ACSC Alandi, Pune, Savitribai Phule Pune University, Pune, India
[3] Department of Electronic Science, Fergusson College (Autonomous), Pune, India
[4] Department of Electronic Science, Savitribai Phule Pune University, Pune, India
Email: shubhangigaikar11@gmail.com (G.S.); drmsz29@gmail.com (M.Z.); adshaligram@gmail.com (A.S.); kailas.sapnar@fergusson.edu (K.S.)
*Corresponding author

Manuscript received October 26, 2025; revised February 3, 2026; accepted April 9, 2026; published July 17, 2026

***Abstract*—This research focuses on developing an Attention-Driven Deep Convolutional Neural Network (ADCNN) to predict nutrient deficiencies in maize using an image dataset. The main advantage of the ADCNN model is its use of attention mechanisms, which enable it to focus on relevant features at the leaf level. This results in more accurate and transparent classification compared to previous methods that relied solely on Convolutional Neural Networks (CNNs). The ADCNN was trained and tested on a dataset covering six nutrient classes: healthy (ALL Present), general nutrient imbalance (ALLAB), nitrogen (NAB), phosphorus (PAB), potassium (KAB), and zinc (ZNAB). Its performance was compared with three deep learning models: the baseline CNN, Visual Geometry Group 19 (VGG19), and hybrid Convolutional Neural Network-Support Vector Machine (CNN-SVM). Results indicate that the ADCNN performs better than any of the competing models, with accuracy, precision, recall, and F1-scores of nearly 99.48%, which demonstrates its efficacy. The attention mechanism improves the model's discriminative capability, enabling it to identify signs of deficiencies in appearance. The ADCNN has enhanced generalization, stability, and response time, which are superior to the conventional CNN architecture. This research shows that using a deep attention process enables real-time or automated plant health monitoring based on nutrient status. The ADCNN model is a stable, efficient system that can be used to create intelligent agricultural systems that enhance sustainable agricultural practices by optimizing nutrient use. Future research should focus on collecting real-world data across various environmental conditions and on integrating Internet of Things (IoT)-enabled platforms to advance multimodal data collection for a more comprehensive plant health assessment.**

***Keywords*—maize leaf, nutrient deficiency detection, attention mechanism, deep learning, Convolutional Neural Network (CNN), precision agriculture, image classification**

## I. INTRODUCTION

The global food supply greatly relies on agriculture as a provider of food, jobs, raw materials, and other benefits. As the human population grows, agricultural production must increase to meet rising food demand. Maize (Zea mays L.) is one of the main food and feed crops worldwide and is the primary source of energy (calories) for the global population, as well as a source of biofuels [1]. The quality and quantity of Maize produced by farmers depend on available soil nutrients and those within the plants. Therefore, deficiencies in macronutrients, whether in the soil or in the plants, can have long-term impacts on plant functions, photosynthetic efficiency, crop yield, and overall plant and soil health. Growers must address these deficiencies early to prevent negative effects on plant growth and to practice proactive agriculture [2, 3].

Traditionally, nutrient deficiency assessments in crops rely on soil test results and laboratory analysis of plant tissues. While laboratory tests have been both accurate and dependable, they are time-consuming and labor-intensive and can be difficult to interpret. Consequently, there is seldom any large-scale or real-time monitoring that utilizes these testing methods. They can also be costly and highly specialized gear that small- and medium-sized manufacturers might not be able to access. Visual tests or tests conducted on-site by trained professionals based on visual observations (leaf color, shape, and texture) are subjective, unreliable, and more challenging to perform accurately in the field as natural lighting conditions, growth stages, and environmental stresses change. Researchers and agronomists are interested in non-invasive, image-based, and automated techniques which use computer vision and Artificial Intelligence (AI) [4].

In the past few years, numerous disruptive technologies have appeared in the field of agricultural informatics, including Deep Learning (DL). CNNs have found growing applications in agriculture, including disease and weed detection, fruit quality assessment, and crop classification. CNNs are generally able to learn complex visual representations of image data that are more efficient and scalable than manually creating features for agricultural diagnosis. Thus, CNNs are a promising new approach for detecting plant diseases. CNNs are used to identify plant diseases, as numerous studies have demonstrated. Nonetheless, the diagnosis of nutrient deficiencies in maize (Zea mays L.) leaves has not been thoroughly examined [5].

Symptoms of nutrient deficiencies may be highly analogous to visual ones and are a serious problem worldwide. Several nutrient deficiencies initially present with visual symptoms that can appear similar; these include chlorosis, necrosis, and interveinal chlorosis. For example, nitrogen can cause leaf yellowing. Zinc deficiencies produce interveinal chlorosis, and this difference is crucial for accurately diagnosing nutrient deficiencies, especially when many deficiencies can be indistinguishable, particularly given traditional image-processing challenges. Although various micronutrients such as iron and magnesium are discussed in general plant physiology, this study focuses specifically on zinc deficiency within the dataset. This is described as an issue of intra-class variability and inter-class similarity and requires attention to the differentiation of minimal, distinctive leaf features to distinguish between different nutrient deficiencies. Furthermore, the difficulties imposed by differences in lighting, the position of a leaf, and other items that may be considered background clutter or the environment will further compound the prediction challenge [6].

Another critical issue is that very few sources provide high-quality, balanced datasets that represent all nutrient-deficiency states, and each dataset has been collected under field conditions. While a limited number of datasets exist for plant disease identification, extensive datasets that accurately represent nutrient-deficiency states and field conditions are rare. As a result, it is often difficult for a model to generalize across contexts, which can lead to prediction bias or overfitting. In many settings, the majority of models reported in the literature emphasize accuracy; as a result, end users find it difficult to ascertain the interpretability of their predictions [7].

Researchers have begun using advanced DL models, such as transfer learning models (e.g., VGG19) and hybrid models (e.g., CNN-SVM), to achieve robust classification and deep feature extraction. Researchers have found that by implementing transfer learning methodologies (for instance, finetuning VGG19), which have been trained with large datasets (like ImageNet), they can classify and identify images for various agricultural applications with limited training data [8]. In this way, both performance and the time required for Training are increased. These models can separate similar objects more accurately by combining the high-dimensional feature space of a CNN and the classification boundaries of an SVM. Hybrid models do not adequately account for spatial correlations and marked regions required to detect nutrient deficiencies, as do standard CNNs [9].

Attention mechanisms in Deep Learning (DL) simulate how the human visual system operates, allowing items of interest, such as informative regions of an image, to receive differential attention compared to non-informative background areas. The attention modules used in CNNs in the area of computer vision in this study have been demonstrated to enhance object discrimination, accuracy, and interpretability, and the attention-based models in the field of plants present different patterns of stress, which results in more accurate and meaningful classification. Moreover, other models enhance accuracy and provide visual justification for their decisions; thus, they may be used in agriculture.

We used the ADCNN framework as described in the literature above. Based on the CNNs' ability to derive spatial features and the attention network's feature selection, we applied projected and derived features at different levels of abstraction to identify which parts of the image are associated with nutrient deficiencies in maize leaves. Our ADCNN was trained and tested on a collection of pictures of six nutrient classes (ALL Present, ALLAB, NAB, PAB, KAB, ZNAB). All the photos were taken from the same maize stalk and captured under different lighting and environmental conditions. We then compared our performance results with VGG19 and a CNN-SVM hybrid model, and determined performance differences from a practical standpoint (complete classification and generalization of the data set) and from a data processing time standpoint (how long it takes to classify a full data set).

The ADCNN method achieved significantly higher accuracy, precision, favorable rates, and F1-scores than CNN-based architectures, with final values >99.48% for each nutrient type. An evaluation of how well the model attended to features indicated that attention was directed mainly to biologically relevant areas of the image (i.e., leaf veins and chlorotic patches), consistent with expected anatomical changes associated with specific nutrient deficiencies. Therefore, the use of attention will not only improve image classification accuracy but also yield interpretable and actionable information from the model.

The study presented here is part of the growing body of knowledge in precision agriculture, which aims to enhance agricultural productivity through technology-enabled farming. This automated ADCNN system will enable farmers to detect nutrient deficiency in real time, thereby saving them time in the field and the lab to determine nutrient status. As a result, farmers will be able to quickly assess whether their crops are receiving all the nutrients they need, apply the correct amounts at the right time, reduce observational time, decrease fertilizer applications while reducing pollution, and confidently harvest more produce than ever. The ADCNN framework can be adapted to support many farm operations, use sensors/images to monitor plant health, and transmit data wirelessly for on-site evaluation and correction.

The present study made significant contributions that can be summarized as follows:

- Novel Attention-Driven Framework for Nutrient Deficiency Recognition: A specialized Attention-Driven Convolutional Neural Network (ADCNN) is suggested to classify maize leaf nutrient deficiencies (in multi-class), and the proposed solution is relatively unexplored, unlike traditional plant disease detection.
- Integration of Spatial and Channel Attention for Fine-Grained Feature Learning: The model utilizes spatial and channel attention mechanisms for better feature learning in discriminating visual patterns such as chlorosis, vein discoloration, and marginal necrosis, which are essential in differentiating similar nutrient deficiency classes.
- Interpretability through Attention Visualization: The presented framework allows for the visualization of attention, thus improving the interpretability of the model by identifying regions of interest in images.
- Comprehensive Comparative Evaluation with Statistical Validation: The proposed ADCNN is extensively compared with traditional CNN, transfer learning with pre-trained VGG19, and hybrid CNN-SVM models. The enhancements are validated statistically using McNemar's test.
- Robust Training Strategy with Data Augmentation and Hyperparameter Optimization: The proposed method incorporates rigorous preprocessing, data augmentation, and hyperparameter optimization techniques to enhance generalization and reduce overfitting in different environmental conditions.
- Practical Relevance to Precision Agriculture Systems: The proposed method is highly relevant to precision agriculture systems, as it can be used for real-time, non-invasive nutrient deficiency detection, which can be further integrated with IoT-based smart farming systems.

## II. Literature Review

The research by Ahmed and Ahmed [10] explored two different DL methods: Residual Networks and Inception Residual Networks for recognizing diseases in palm leaves; the approach used Transfer Learning to train the models. They both performed well and significantly reduced the time required to train the models for multi-layering. The experimental results were conducted using 2631 images from the dataset to train and test models of palm trees of different sizes but with unique coloration characteristics. It was estimated that, had the dataset been larger than the one used in the experiment, the model would have achieved 100% accuracy; the original dataset achieved 99.62% accuracy in the initial classification. Examples were also given of literature that describes how to detect when a plant is starving for nutrients. Ultimately, the authors aimed to find food (or nutrients) for the leaves.

Watchareeruetai *et al.* [11] investigated a method for analyzing plant images using CNNs and demonstrated that they could represent blackgram (Vigna mungo) grown under controlled conditions. One objective of the research was to analyze leaf characteristics of blackgram to determine whether any nutrients are missing from the plant. The analyses sought to determine how many plants (from the five plant types selected) were missing one of the five nutrients (Calcium, Iron, Potassium, Magnesium, and Nitrogen) included in their study. Using a sample of 3000 leaf images, we compared the current method's results with those of specialized personnel who would likely struggle to identify vitamin deficiencies in the samples.

The impact of parameter reduction on the recognition accuracy of maize leaf diseases was studied by Zhang *et al.* [12]. They suggested using DL to enhance the CIFAR-10 and GoogLeNet models for leaf disease detection. A study was proposed by Amrutha and Lekha [13] to evaluate soil nutrient content and to review and replenish phosphorus, potassium, and nitrogen levels. These levels were analyzed through chemical testing and by monitoring devices. A mechanical machine that regulates the amount of fertilizer spread on the soil is effective at applying a sufficient amount, rather than excess. The system employed three steps: sample preparation, soil testing, and application of the appropriate amount of fertilizer, all regulated by a microcontroller. The method of providing results on the fertilizer application to the soil took less than 30 min. Test results were evaluated on the microcontroller; for all macronutrients, the time to assess the results and provide the appropriate amount of fertilizer to the soil was 30–40 min.

Miller *et al.* [14] stress that improving the accuracy or efficiency of pathogen and disease identification needs to be prioritized, as these threats pose a global threat to plant health. Pathogens and diseases (emerging/re-emerging/endemic) pose an increasing threat as they continue to spread and evolve, in part due to climate change, human transportation, and vector-borne transmission worldwide. It is increasing our difficulty in identifying and monitoring new plant pathogen/disease outbreaks. This connection provides opportunities for organizations and individuals on each side of the border to take action.

In a 2020 study by Indumathi *et al.* [15], the authors described a method for predicting leaf diseases using image processing techniques applied to images from their research. In total, 13 image attributes were generated for each leaf image, classifying it as diseased or healthy. Leaf disease prediction will be performed using these image attributes. It is anticipated that a computerized system will be developed. In the initial stage of identifying/diagnosing plant leaf diseases, the first step is to segment leaf images and analyze the resulting segments (i.e., specific areas of the image). To segment leaf images into particular regions, the authors used a *k*-means clustering algorithm. Once segmentation is complete, algorithms (using various machine learning methods) will be developed. All researchers used the Artificial Neural Network (ANN) method to classify the images. In addition, to determine the plant leaf's health status, the authors will perform diagnostic testing and classify the disease infecting the leaf.

According to Ghorai *et al.* [16], applications of image processing to detect plant diseases are wide-ranging; image detection techniques include several detection types (image segmentation, feature extraction, and image classification). Ghorai *et al.* [16] focused on only one detection and diagnosis method; the authors described its components, principles, and procedures across multiple crop types. The application of swarm bots for automated agriculture (including diagnosing nutritional deficiencies) was first presented by Rajasekar *et al.* [17]. The use of automation in agriculture can significantly reduce the issues that traditional agricultural businesses now face daily. The authors developed an AI model that determines the optimal planting method (drill or seed) and diagnoses deficiencies in the food supply. In addition, the authors presented a novel, rapid recovery concept that reduces the time spent on labor-intensive activities in the agricultural industry. Using image processing and communication technologies, farmers can identify crops that may be nutrient-deficient or lacking nutrients based on leaf physical characteristics.

According to Kaur [18], digital image processing, computer vision, and IoT frameworks can detect (before human detection) signs of deficiency in plants. Thus, these frameworks will allow farmers (agricultural practitioners) to take prompt action and remedy the deficiency by simply applying digital image processing. This research paper examines how one identifies the processes involved in diagnosing nutrient deficiencies in plants. Amirtha *et al.* [19] developed a robotic vehicle capable of performing an autonomous crop survey using plant imagery to detect nutrient deficiencies. The research compared an existing photo with a current photograph and an analysis dataset. They estimated the level of nutrient deficiency in the crop when the submitted image matched either a dataset image or a partially matched dataset image.

Dhandapani *et al.* [20] investigated methods for analyzing, identifying, and diagnosing plant diseases using digital image processing techniques. Pathogens can adversely affect various parts of the plant, but the disease symptoms are most often visible on leaves. The way in which researchers Aditi Shah *et al.* [21] created a machine learning framework for mapping the macronutrient levels of plants is by taking Red Green Blue (RGB) photographs. Researchers use these RGB images to illustrate the growth and vigor of a nutrient-deficient plant based on its leaf appearance. RGB images are also used to detect object characteristics via edge detection and/or texture analysis, such as whether outwardly visible distress is present in the plant. The dataset created by Shah *et al.* [21] consisted solely of image data of healthy-looking leaves from plants with no visible signs of nutrient deficiency. These datasets will enable accurate supervised/unsupervised learning, enabling the healthcare system to recognize and classify malnourished plant life effectively. The dataset's precision enables the delivery of more effective management techniques to maximize the production or yield of various crops.

Barbedo [22] employed advanced imaging techniques and machine learning to identify and classify plants' nutritional deficiencies using modern image processing. By incorporating multiple image sensors, researchers modified the RGB images in conjunction with the dataset to include all possible defects in malnourished crop plants. To evaluate their identification and classification method, researchers conducted a pilot study using machine learning for both object detection and product differentiation, providing valuable insights into the agricultural business and enabling corrective actions based on data generated by machine-learning methods.

Anthay *et al.* [23] used an algorithm designed to analyze leaf health by performing three operations across three assessments to detect health deficiencies using a computer (CPU), MATLAB, and Sensors. In addition, each has four types of sensors that read specific features of the residue left by analysis across all assessments at all growth stages for plants or their leaves. All soil or plant leaf conditions will be continuously measured at each of the four stages using a sensor, a web-connected computer program, and an Android app to manage the device. The app's purpose is to capture the image with a Raspberry Pi that is used in conjunction with an Android device to do subsequent processing to create the image for further continued monitoring, while also capturing previously captured images in cloud storage of all leaf or plant leaves and reporting on plant condition through a set of values.

According to Swami and Patgiri [24], an artificial intelligence system was developed that used plant nutrition as a source of information to determine if plants had nutrient deficiencies. Over the last few years, machine learning has become a widely used tool for accurately assessing plant nutritional status, driven by the need for a reliable, efficient method to detect nutrient deficiencies. The algorithm that identified whether a plant was nutritionally sufficient or deficient used digital image analysis. It evaluated the image and compared its features with some previously learned data. In this study, the frequency of dietary deficiencies was determined using a sample of 102 subjects.

A DL model was designed to detect disease in plant leaves and automate the diagnosis process, as reported by Madhulatha and Ramadevi [25]. This research aims to enhance agricultural productivity to protect crops. This was achieved by training a Convolutional Neural Network (CNN) on visual symptoms of plant disease, reaching 96.50% correct identification and diagnosis. The deep CNN training data were collected as part of the PlantVillage Collaborative Study. The deep CNN model was developed using AlexNet.

In another study by Saleem *et al.* [26], three deep learning models were developed and applied to differentiate healthy from unhealthy plant leaves. Both these models possess their individual base network and feature extractor. Besides computing the parameters of each of the three models using base networks and feature extractors, three optimization algorithms were applied to each model. Each model was also run with varying parameters to determine the most efficient runtime. The Adam-optimized model and the Single Shot MultiBox Detector (SSD) model achieved the best runtime and average accuracy of 73.08% for disease identification on plant leaves.

Mohanty *et al.* [27] also developed a mobile application that uses DL to recognize and classify plant diseases. They tested a variety of pre-trained CNNs for this process and achieved excellent results with GoogLeNet, achieving 99.35% accuracy, which is better than that of AlexNet.

More recently, researchers have documented the importance of rapidly developing "smart" agricultural systems that use deep-learning-based computer vision tools to identify plant diseases and nutrient deficiencies. Early studies and attempts to automate the classification of wheat diseases and aphids were conducted using traditional image-processing techniques, which laid the groundwork for the development of automated classifiers based on DL, such as CNNs and transfer learning. Although the value of incorporating IoT and sensor monitoring systems into research is validated by the potential to perform real-time analysis and make decisions, there remain ample opportunities for further testing of current research priorities, such as a lack of dataset diversity; differences in visual representation of plant samples depending on lighting conditions; and overfitting, requiring models to consistently produce accurate results across the wide range of time periods and climatic variations experienced in the natural world.

Table 1. Comparative analysis of attention-based plant image analysis methods

| Reference | Crop | Task | Attention/Explainability | Model Type | Limitations |
|---|---|---|---|---|---|
| Rajasekar *et al.* [17] | Various crops | Nutrient deficiency detection | None | Image processing / IoT | No CNN-based attention |
| Kaur [18] | Multiple crops | Nutrient deficiency detection | None | Image processing + ML | No deep attention, limited robustness |
| Barbedo [22] | Multiple crops | Stress detection | Limited | ML & CNN | Lack of integrated attention |
| Saleem *et al.* [26] | Multiple crops | Disease detection | Saliency/localization | RFCN, SSD, Faster-RCNN | Disease-only, no nutrient classes |
| Mohanty *et al.* [27] | Multiple crops | Disease detection | None | CNN (GoogLeNet, AlexNet) | No interpretability, disease-focused |
| Proposed ADCNN (This work) | Maize | Nutrient deficiency recognition | Integrated attention + attention maps | Attention-based CNN | Field-scale validation (future work) |

Essentially, recent works presented in Table 1 have analyzed the use of attention-driven DL methods to identify whether plants are diseased (fungal, bacterial, viral), with most focusing primarily on locating disease symptoms and improving classification by either employing spatial or channel attention. The majority of attention-based models for disease recognition have limitations in identifying nutrient deficiencies.

Upon initial glance, it would appear that the majority of current research on attention-based DL methods used for plant disease recognition is not comparable to hybrid and transfer learning methods in the same experimental conditions; whereas, both plant nutrient deficiencies (nitrogen, phosphorus, potassium, and zinc deficiencies, along with general imbalance and healthy class) exhibit visual symptoms that are often similar e.g., chlorosis, vein coloration the majority of plant nutrient deficiency research lacks the proper attention to biologically significant areas of the plant. Before conducting any new research using attention-based analyses of plants, researchers should have the opportunity to investigate whether multiple classes of plant nutrients are adequately accounted for to validate attention accuracy by class.

ADCNN represents a new method for intelligent nutrient management in precision agriculture, based on an interpretation-based approach. A comprehensive comparative analysis is provided for the ADCNN, existing CNNs, transfer-learning VGG19, and CNN-SVM hybrid networks, using statistical significance testing and attention analysis.

The primary purpose of this study was to develop and analyze a convolutional neural network for classifying various forms of nutrient deficiencies in maize leaves. The distinction between this work and other studies that have used attention for plant disease classification is that the ADCNN is focused solely on identifying nutrient deficiencies, rather than multiple disease types that tend to exhibit visual signs of disease in similar ways. In addition to enhancing accuracy by focusing attention on symptom localization, the ADCNN provides a means to visualize the attention it assigns to specific image regions through attention maps. The proposed ADCNN shows a statistically significant advantage over baseline CNN models, as determined by statistical significance testing. Few studies have attempted to address all three components of this study simultaneously (attention-driven modeling, statistical validation, and interpretability).

## III. Methodology

A detailed method for detecting maize leaf nutrient deficiencies with CNNs comprises a systematic set of procedures, as shown in Fig. 1. All stages are necessary to assess the model's credibility, accuracy, and generalizability.

### A. Dataset Description

The initial and most crucial step in the system in question is acquiring the dataset, as the quality and representation of the training data significantly determine the performance of any DL model. The paper will retrieve maize leaf images of two different data sets: (i) the publicly available dataset Maize Plant Leaf Nutrient Deficiency Dataset, which is available through Kaggle [28], and (ii) field-compared RGB images of natural locations of environmental capture. The image data collected from multiple sources also introduces additional variability, such as lighting, background complexity, and variability in maize leaf orientation when comparing images. Each picture was stored as an RGB image, with resolutions ranging from 256 × 256 to 1024 × 1024 pixels.

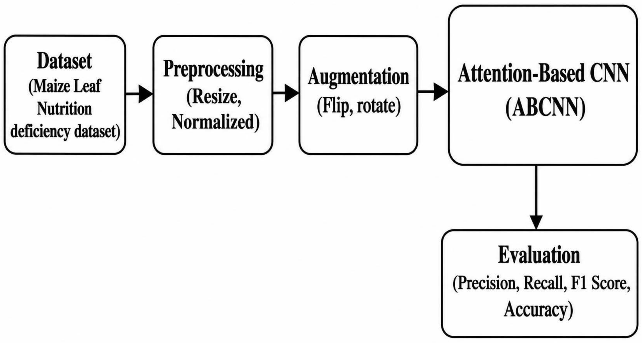



Fig. 1. Block diagram of the proposed system.

To ensure uniformity in the DL model, the size of all the images was reduced to 224 pixels by 224 pixels, and 3 color channels were used before the start of Training. The entire picture set is split into two parts (80% of the samples were used for training, and 20% for validation). An 80/20 split of the total image collection is used to ensure sufficient data for learning functional characteristics and to prevent overfitting. The distribution across the classes was kept constant between the Training and Validation Data Sets to avoid bias toward the larger nutrient classes.

In addition to expanding the training set, data augmentation methods were used to generate more heterogeneous training samples. They include horizontal and vertical image flipping, random image rotation (±20°), and image resizing and brightness adjustment before adding to the training set. Data augmentation techniques enhanced the model's resilience and reduced its susceptibility to being pegged by environmental changes during Training.

Example images from the dataset are shown in Fig. 2.

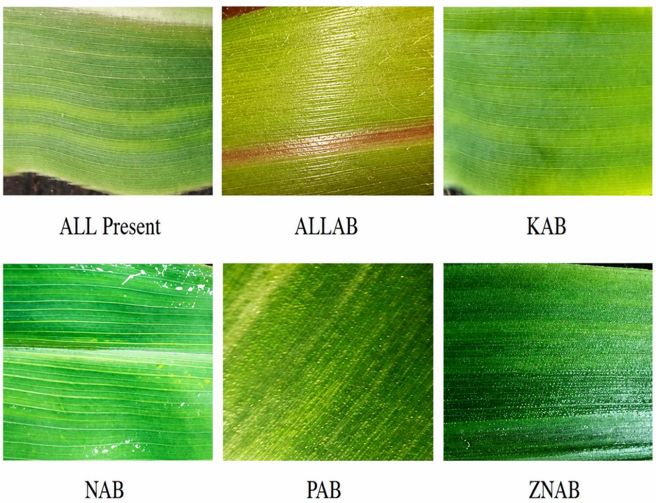



Fig. 2. Sample images of the maize plant leaf nutrient deficiency dataset.

Diversity is essential in any dataset used for training models, including those used for this research. In this case, the maize leaf datasets were collected to ensure diversity by capturing images under varying levels of ambient light (e.g., bright and dark), at different camera angles, and/or in various environments. Images were transferred into a database with a defined class (also known as "labeling") structure and stored together. This will assist researchers with image preprocessing and allow them to automatically load batches of images once they begin training their model.

The maize leaf datasets used for this work were obtained from two primary sources: (1) the Maize Plant Leaf Nutrient Deficiency Dataset (Publicly Accessible) available via Kaggle; and (2) images of maize leaves taken with a regular (e.g., RGB) camera under different lighting conditions. The images in this dataset were all RGB and ranged in size from 256×256 px to 1024×1024 px. The images were adjusted, and all resized to 224×224×3 px before Training to maintain consistency with the model architecture. The dataset consists of six classes: ALL Present, ALLAB, NAB, PAB, KAB, and ZNAB for each maize leaf image, and Table 2 shows the class-by-class sample distribution for clarity and repeatability. An 80:20 split was implemented for Training and validation across the two datasets/data sources to balance the learning and generalization. Rotation (±20º), horizontal/vertical flipping, brightness adjustment, and scaling, as defined, were implemented solely on the training data to reduce the risk of overfitting. A separate test set was not used; validation performance is reported.

## B. Preprocessing

Image Preprocessing prepares images for use in the CNN. Preprocessing involves standardizing the data, removing noise, and optimizing image quality, allowing the CNN to focus on the most essential features of the image. All images are resized to match the input images (usually 224×224×3) to ensure they are all the same size. Normalizing the pixel intensity exposure values from 0 to 255 to 0 to 1 also helped achieve convergence of the classification labels. Data augmentation also helps increase the model's robustness by performing horizontal/vertical flips, rotations, scaling, brightness and contrast adjustments, and increasing the number of training samples. Data augmentation also adds diversity to the original dataset while maintaining the model's independence from the environment. One-hot encoding was also used to convert the class labels into numerical class labels for CNN-based classification. The applied data augmentation techniques (rotation, flipping, brightness adjustment, and scaling) are intended to simulate environmental variability and improve computational robustness. However, these transformations do not represent real-world field validation, which remains part of future work involving in-situ data collection under natural agricultural conditions.

## C. Dataset Splitting

Table 2 shows how the dataset has been organized into training and validation sets. Each class consists of images of maize leaves from this research project. The ALL-Present class represents healthy maize with uniform, green leaves and no visible signs of nutrient deficiency. The remaining courses identify specific nutrient deficiencies in maize leaves. ALLAB represents maize leaf plants with general nutrient imbalances and exhibits typical symptoms, such as extensive leaf chlorosis and reduced vigor. KAB is a potassium deficiency characterized by leaf scorching, yellowing at the edges, and necrotic tissue in extreme cases. NAB is a nitrogen deficiency, and is indicated by a general yellowing of the older maize leaves, small leaf size, and general pale color. PAB is phosphorus deficiency; the maize leaves have a dark green to purplish color (especially around veins and margins), indicating a general inability to take up phosphorus. ZNAB is zinc deficiency and is characterized by interveinal chlorosis, abbreviated internodal length, and varying degrees of yellowing on the surface of the leaves. The dataset distribution is presented in Table 2.

Table 2. Dataset distribution

| Class | Total Images | Training Set | Testing Set |
|---|---|---|---|
| ALL Present | 1470 | 1176 | 294 |
| ALLAB | 2430 | 1944 | 486 |
| KAB | 4301 | 3441 | 860 |
| NAB | 1535 | 1228 | 307 |
| PAB | 3713 | 2970 | 743 |
| ZNAB | 2545 | 2036 | 509 |

The model weights were optimized during Training using backpropagation. The model's performance was assessed in the independent validation environment after each epoch. Additionally, the model was tuned using hyperparameters to improve performance. Separating the training and validation datasets reduced the risk of overfitting and enabled the model to generalize rather than memorize the dataset.

## D. Model Training

The model's weights were improved/revised in the simulation environment, and performance was evaluated in the validation environment at each epoch. Once the model's performance on the validation set was assessed, the model's hyperparameters were tuned to improve efficiency based on each epoch's results. With the model trained on two data sets (Training and validation), there was less risk of overfitting and a greater chance that the model would generalize well rather than memorize the data. Furthermore, the data partitioning also addressed the class imbalance issue in the classifier across different nutrients. After the training process is complete, the classifier can be tested in a separate evaluation phase to obtain an unbiased assessment of its performance.

### 1) Convolutional Neural Network (CNN)

The CNN serves as a benchmark in the industry for feature extraction from images of corn leaves and their classification based on nitrogen deficiency. The CNN is a computational model inspired by biological systems that helps automatically learn hierarchical relationships among image features. The structure of the CNN algorithm for the identification of nutrient-deficient corn plants is shown in Fig. 3.

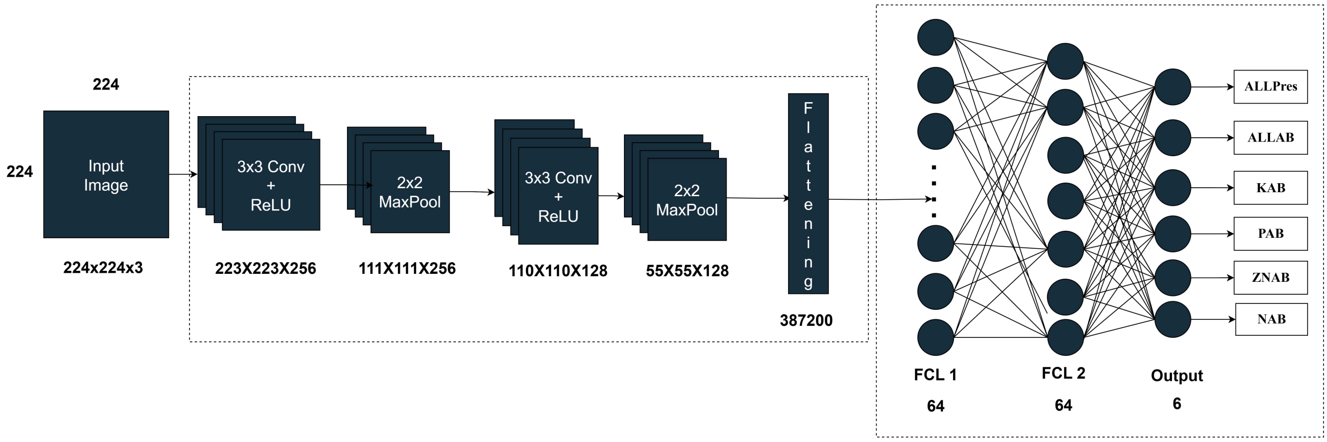



Fig. 3. Architecture of the CNN algorithm for maize plant nutrient deficiency recognition.

This architecture consists of several convolutional layers with filters for detecting local patterns on leaf surfaces, such as color changes, edge anomalies, and texture gradients. As the filters are applied to the input image, a feature map is created that highlights areas of interest for nutrient stress. Following the convolution, non-linear activation functions (most often ReLU) are used to enable learning of complex patterns. Pooling layers (such as max or average) can reduce the size (dimensionality) of feature maps obtained from the convolution. Finally, all learned features are combined in a fully connected layer to produce a higher-level representation of the input data, which is then sent to the SoftMax classifier to produce probabilities for multiple classes of nutrient deficiency. Backpropagation and optimizer algorithms (including Adam and Stochastic Gradient Descent (SGD)) were used to train the overall CNN model to minimize categorical cross-entropy loss and achieve high accuracy and good convergence.

*2) VGG19*

The VGG19 is an example of a convolutional neural network, which is a DL architecture first proposed within a research paper created by the Visual Geometry Group (VGG) at the University of Oxford. The proposed architecture aims to improve predictive performance through transfer learning and reduce the typical computational requirements of training a new DL model.

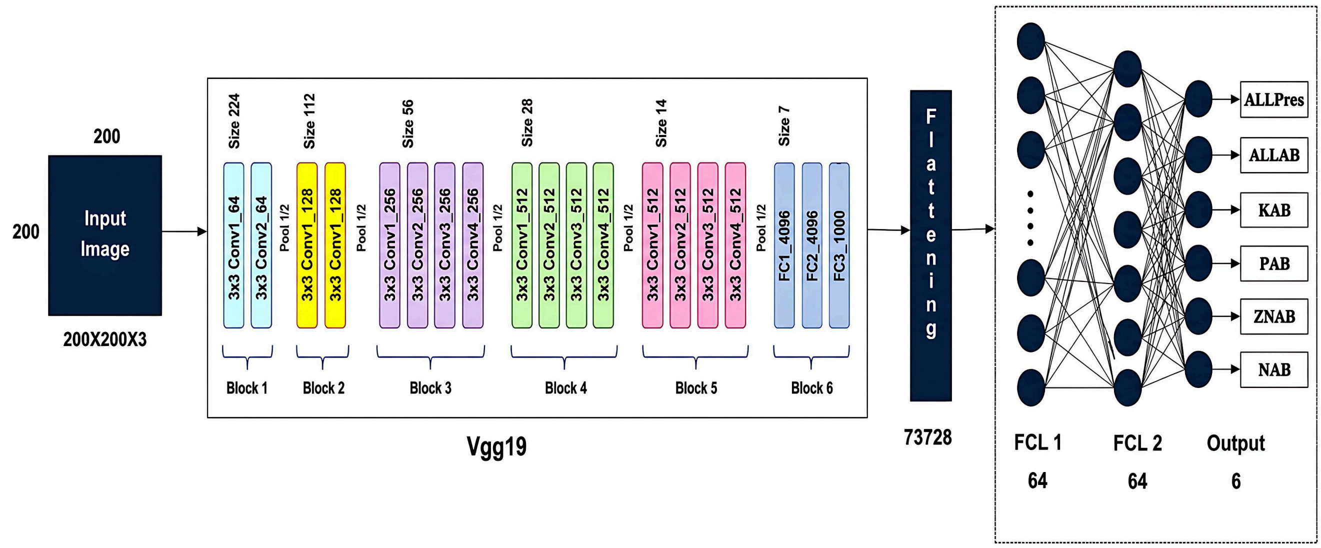



Fig. 4. Architecture of the VGG19 algorithm for maize plant nutrient deficiency recognition.

Fig. 4 shows the architecture of the VGG19 algorithm as an example of a new model used to identify nutrient deficiencies in maize plants. VGG19 has 19 layers, comprising 16 convolutional and 3 fully connected (i.e., dense) layers, each with a kernel size of 3 × 3. All layers in the VGG19 architecture extract features uniformly; therefore, each layer produces a consistent output for a given input. The VGG19 model is also straightforward while still enabling feature extraction across many layers of a deep convolutional neural network. The pre-trained VGG19 model contained many of the same features as the original image; therefore, the weights of the first few convolutional layers were retained, allowing the extraction of generic features such as edges and textures from the leaves of maize plants. However, the last fully connected layer was retrained specifically to classify nutrient deficiencies. By using transfer learning techniques, the VGG19 model can leverage previously learned image features and rely less on the training data to generalize to new images. The final output layer of VGG19 uses the SoftMax activation function to classify maize leaf images into specific nutrient categories, achieving higher final classification performance than both original and simpler CNNs.

*3) CNN–SVM hybrid model*

By marrying the strengths of the two separate modalities, CNN and SVM, the Hybrid model, which is CNN-based and provides automatic deep feature extraction, as well as the SVM modality, which is highly accurate for classification. In our model, the CNN was used to extract high-dimensional features for leaf color and texture. At the same time, the SVM was incorporated into the final layer of the CNN, producing an output. The CNN-SVM Hybrid model also improves generalization and reduces misclassification, especially for visually similar nutrient deficiencies. The combination provides greater robustness and improved class separability compared to a traditional CNN-based classification method.

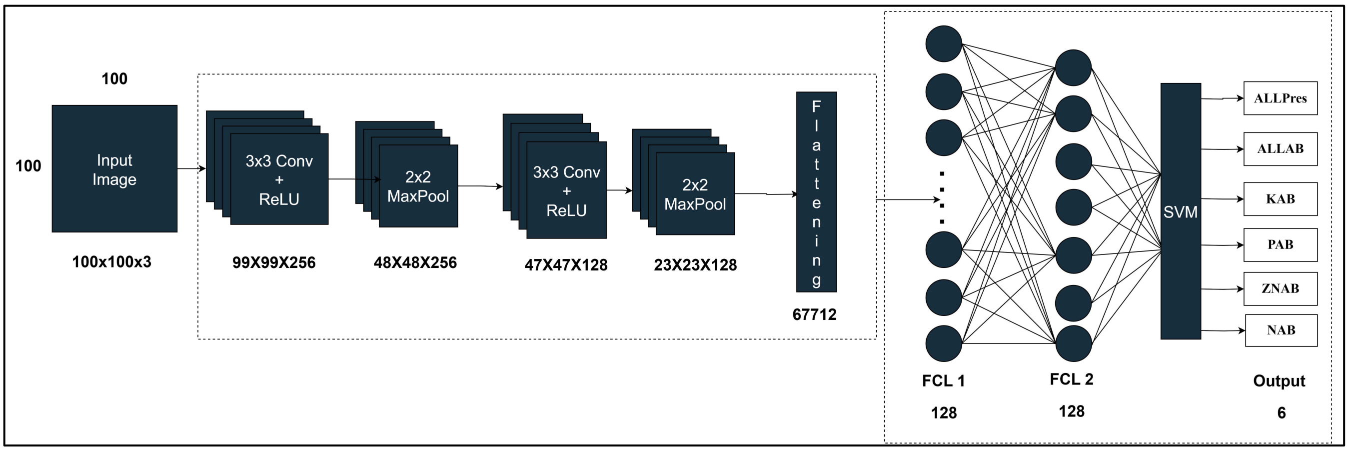



Fig. 5. Architecture of the CNN-SVM algorithm for maize plant nutrient deficiency recognition.

Fig. 5 presents the architecture of the proposed CNN-SVM model to detect nutrient deficiencies in maize crops. In the designed model, a Convolutional Neural Network (CNN) serves as a powerful feature extractor for processing maize leaf images and extracting visual information indicative of specific nutrient deficiencies (e.g., vein patterns, color differences, and chlorosis). Before passing through the CNN layers, the input leaf image was preprocessed first. An input layer is followed by many convolutional and pooling layers, in which features at various levels are extracted. For each block of convolutional layers, the Rectified Linear Unit (ReLU) activation function and Batch Normalization were used to improve nonlinearity and training stability. A pooling layer was employed to reduce the size of the output from convolutional layers, while a dropout layer was incorporated to minimize the risk of overfitting.

These learned features are further used in an SVM classifier that identifies the optimal hyperplane for distinguishing between nutrient-deficient classes, rather than using a SoftMax classifier. These fusion leverages both the CNN's proficiency in visual pattern recognition and the SVM's ability to identify a large distance between decision boundary hyperplanes in the higher-dimensional feature space. It is a novel approach that combines CNN and SVM. These architectures perform better than the classic CNN when classifying and distinguishing the signs of nutrient deficiencies, such as nitrogen. In summary, this architecture is a perfect blend of deep learning and statistical approaches.

*4) Proposed driven CNN algorithm*

ADCNN is the core and most advanced part of this study. This is a modified version of the standard CNN with an attention mechanism. The diagram of the attention-driven CNN algorithm is shown in Fig. 6.


Fig. 6. Architecture of the Attention-driven CNN algorithm for maize plant nutrient deficiency recognition.

The proposed ADCNN constitutes the primary novelty of this paper. It is a modified version of the standard CNN equipped with an attention mechanism, thereby boosting overall classification accuracy and providing additional insights. The focus on significant features ensures that the model learns subtle differences between nutrient shortages (N, P, K, Zn), imbalance, and good conditions.

The ADCNN architecture builds on the baseline CNN and includes three key components:

- Feature Extraction Layers: Several convolutional and pooling layers extract low- and mid-level visual patterns such as textures, color gradients, and vein structures from maize leaves.
- Attention Mechanism: This module (often implemented via a Channel and Spatial Attention Module or a Self-Attention Block) generates attention maps that identify and amplify critical features.
    - Spatial Attention: Highlights where deficiency symptoms appear in the image (e.g., edges and interveinal zones).
    - Channel Attention: Emphasizes the most relevant feature maps (e.g., hue changes or chlorosis patterns). Mathematically, the attention weight $\alpha_i$ is learned for each feature map $F_i$, and the final attentive feature $F'$ is computed as:
    - The attention mechanism assigns a learnable weight $\alpha_i$ to each feature map $F_i$, and the aggregated attentive feature is computed as:

$$F'_{att} = \sum_i \alpha_i \odot F_i \tag{1}$$

where $\sum_i \alpha_i = 1$ . This allows the model to concentrate computational resources on key visual indicators.

- Fully Connected Layers and Softmax Classifier
    - Dense Layers and Softmax Classifier: The combined attentive features are flattened and fed into densely connected layers to produce high-level representations, and a softmax classifier outputs the image classification into one of six categories: ALL Present, ALLAB, NAB, PAB, KAB, and ZNAB.

The ADCNN applied

- The Adam optimizer uses adaptive learning rates to accelerate convergence.
- Categorical Cross-Entropy Loss was used to minimize the classification error.
- Data Augmentation (e.g., rotation, flipping, brightness adjustment, and scaling) to facilitate generalization.
- Early stopping was used to avoid overfitting and stabilize the model.

During training, attention maps formed in early layers detected general texture properties, while deeper layers, such as vein chlorosis or marginal burning, identified local patterns. An attention-based CNN, which enhances traditional CNNs with attention, is used to enable the CNN to recognize and categorize basic nutrient deficiencies in maize leaves. This offers both quantitative advantage and qualitative understanding of the area of the leaf affected by nutrient deficiencies.

The hyperparameters of all DL models were tuned through systematic experimentation within the search ranges specified in Table 3.

Table 3. Hyperparameter tuning settings for DL models

| Model | Hyperparameter | Search Range | Selected Value |
|---|---|---|---|
| CNN | Learning rate | $1\times10^{-2}$–$1\times10^{-5}$ | $1\times10^{-4}$ |
| CNN | Batch size | 8, 16, 32 | 16 |
| VGG19 | Frozen layers | 10–15 | 13 |
| VGG19 | Optimizer | Adam, SGD | Adam |
| CNN–SVM | Kernel | Linear, RBF | RBF |
| CNN–SVM | C value | 0.1–10 | 1 |
| ADCNN | Attention depth | 1–3 | 1 |
| ADCNN | Learning rate | $1\times10^{-3}$–$1\times10^{-5}$ | $1\times10^{-4}$ |

The learning rate, batch size, optimizer parameters, and model parameters were tested together to maximize validation accuracy and minimize loss convergence per model. The last hyperparameters were selected to balance classification accuracy, training stability, and cost. This rendered the CNN, VGG19, CNN-SVM, and the proposed model, ADCNN, on equal terms.

### *E. Model Evaluation*

The optimization algorithm, which is at the heart of training all models, allows each model to minimize the loss function while efficiently updating the parameters of the network.

- Adam Optimizer: The Adaptive Moment Estimation (Adam) optimizer builds on the strengths of both the moment and the RMSProp optimizer and calculates adaptive learning rates of each parameter given an estimate of the first and second moments of the gradient. Adam is a useful system to have more stable learning and faster convergence in models that are restricted to noisy gradients on larger, more complex ConvNet architectures.
- Stochastic Gradient Descent (SGD): The SGD optimizer keeps adjusting the weights based on random batches. When the SGD optimizer keeps reducing the loss function, it produces gradients and parameter updates that move the parameters towards the steep slope of the relationship.

Both optimizers use categorical cross-entropy as the loss function, which measures the difference between the predicted probability values and the correct class labels. To converge, Training will continue until the optimal loss level is reached, along with the highest validation accuracy.

Multiple quantitative metrics were computed to assess performance comprehensively:

- Accuracy: The model's overall accuracy is measured as the ratio of correctly predicted samples to the total number of samples.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{2}$$

Accuracy represents the overall proportion of correctly classified samples.

- Precision: The proportion of correctly predicted positive samples among all predicted positive samples. This reflects the reliability of optimistic predictions.

$$\text{Precision} = \frac{TP}{TP+FP} \tag{3}$$

It measures the reliability of optimistic predictions, that is, how many predicted nutrient deficiency cases are correct.

- Recall (sensitivity): The proportion of positive samples correctly identified by the model. This reflects the completeness of detection for a given class.

$$\text{Recall} = \frac{TP}{TP+FN} \tag{4}$$

This indicates the model's ability to detect all actual positive instances of a given disease.

- F1-score: The F1-score is the harmonic mean of Precision and Recall, providing a balanced measure when both false positives and false negatives are equally crucial to the model.

$$F1\text{-}score = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{5}$$

It provides a balanced trade-off between accuracy and recall by combining the two into a single metric.

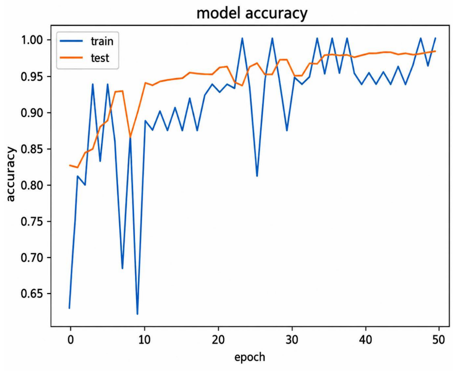
These measures provide a holistic evaluation of the model's performance in accurate and erroneous classifications of the disease classes, and its ability to isolate the classes. Finally, although the choice of evaluation metrics was mainly centered on the performance of the model in terms of prediction, as measured by metrics such as accuracy, precision, recall, and F1-score, additional information was provided by examining the loss and accuracy plots for each epoch to determine the convergence of the model and the presence of overfitting and underfitting. Among the three models, the ADCNN achieved the highest accuracy and, consequently, the best classification performance across illumination levels and background complexities. Moreover, the hybrid CNN-SVM model had a more accurate decision boundary than the CNN model.

## IV. Results and Discussion

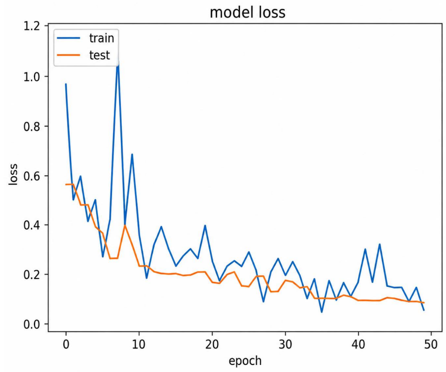
The experimental results clearly show that architectural improvements significantly influence the performance of nutrient deficiency recognition. Although the baseline CNN can extract low-level visual information, it performs poorly on overlapping nutrient symptoms. The VGG19 model can extract deeper visual information through transfer learning; however, it fails to allocate task-relevant attention. The CNN-SVM combination improves class separation; however, it fails to provide spatial relevance. On the other hand, the proposed ADCNN model performs better by focusing its attention on physiologically meaningful areas, such as chlorotic areas and vein discoloration. This is because the attention mechanism reduces interference from the background. The CNN, CNN-SVM, and proposed Attention-driven CNN models were implemented in Python 3.10 using TensorFlow 2.x and the Keras framework on Google Colab Pro. Regarding the hardware configuration, the author has mentioned an NVIDIA Tesla T4 GPU (16 GB VRA) and an Intel Xeon CPU (2.20 GHz, 12 GB of RAM). Finally, the results for each DL algorithm are presented in the following sections.

### A. Result of the CNN Algorithm

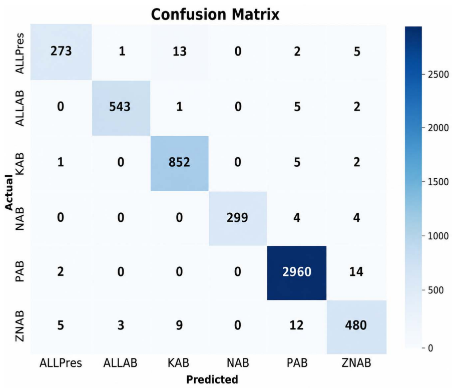
The results presented in Fig. 7 demonstrated that the deep CNN model could recognize nutrient deficiencies in maize leaves.




(a)




(b)



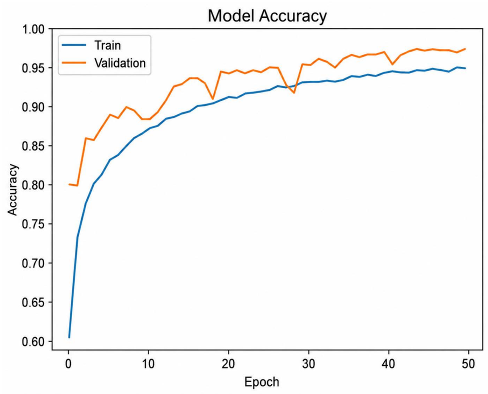

(c)

Fig. 7. Results of the CNN algorithm: (a) accuracy, (b) loss, (c) confusion matrix.

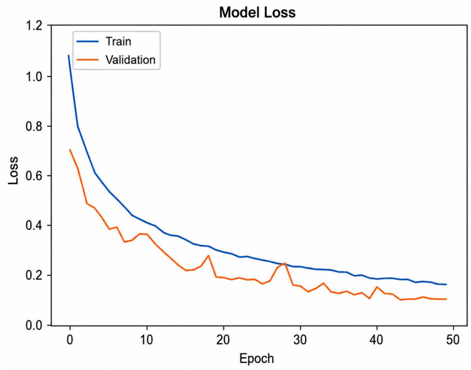
In the accuracy and loss graphs, the training and test accuracies increased with the number of epochs, converging to nearly 98%. Correspondingly, the losses decreased considerably, confirming successful learning and minimal overfitting of the training data. The confusion matrix showed very accurate class predictions, with very few incorrect classifications across the six nutrient classes (ALL Present, ALLAB, KAB, NAB, PAB, and ZNAB). The CNN achieved 98% accuracy, with precision and recall near 1.00 for almost all classes, and a macro-F1 score of 0.97, suggesting balanced results across all classes. High values of recall for all major classes, PAB (0.99) and KAB (0.99), indicated correct classification of nutrient deficiency, and low values of loss confirmed consistent results. These results confirm that the CNN model is reliable and robust, and demonstrate generalization in classifying maize leaf nutrient deficiency from image data.

### *B. Result of the VGG19 Algorithm*

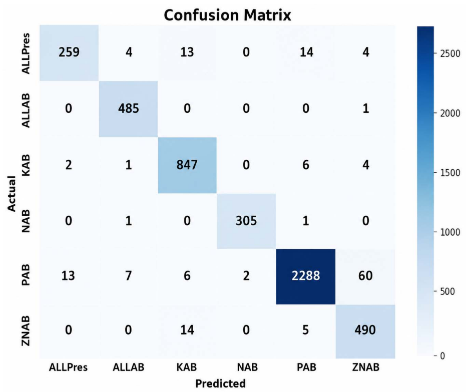
The results of the VGG19 algorithm for the recognition of nutrient deficiency in maize plant leaves are presented in Fig. 8.




(a)




(b)



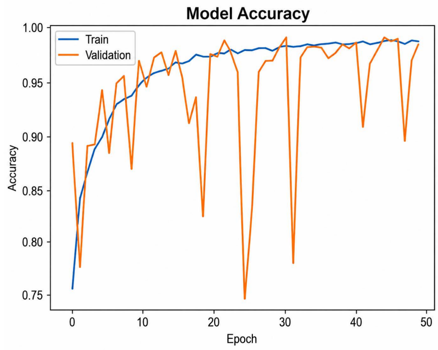

(c)

Fig. 8. Results of the VGg19 algorithm: (a) accuracy, (b) loss, (c) confusion matrix.

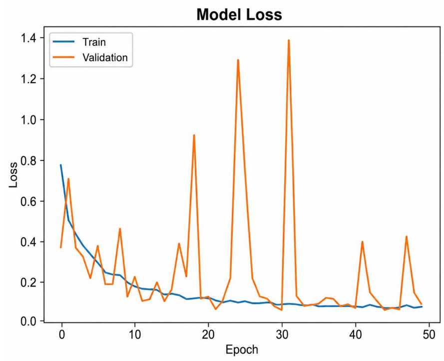
The curves representing both training and validation accuracies showed smooth convergence. The model predicted the validation accuracy was approximately 97% after 50 epochs of Training, which indicated that it had experienced low levels of overfitting and had good generalization. The losses demonstrated excellent optimization, continued decreases in training and validation losses, followed by a stable score and overall loss at a relatively low level. The confusion matrix showed that most samples were correctly classified in each nutrient category: ALL Present, ALLAB, KAB, NAB, PAB, and ZNAB, with the only misclassifications occurring between two visually similar classes. Vgg19 reported an overall accuracy of 97, a macro F1-score of 0.96, and a weighted F1-score of 0.97 for the nutrient classes, and all classes but one reported precision and recall scores of greater than 0.95. The findings above show that VGG19, using transfer learning, extracted valuable hierarchical features related to nutrient deficiencies, including variations in leaf color, vein patterns, and texture.

### *C. Result of the CNN-SVM algorithm*

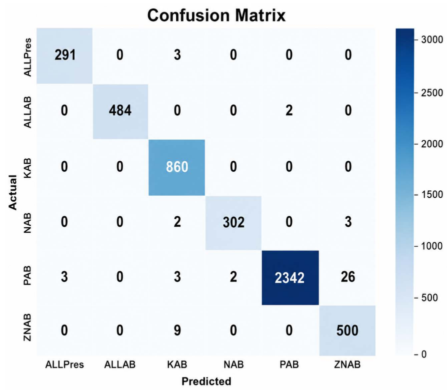
By integrating the CNN's strength in deep feature extraction with the SVM's classification accuracy, the CNN-SVM model proved uniquely effective at identifying maize leaf nutrient deficiency. The model's description highlights various layers, including convolutional, batch normalization, pooling, dense, and dropout layers, which help mitigate overfitting. Therefore, deep features were extracted from the CNN and used for classification by the SVM classifier. The results of the CNN-SVM classification algorithm for nutrient deficiency recognition are presented in Fig. 9.




(a)




(b)



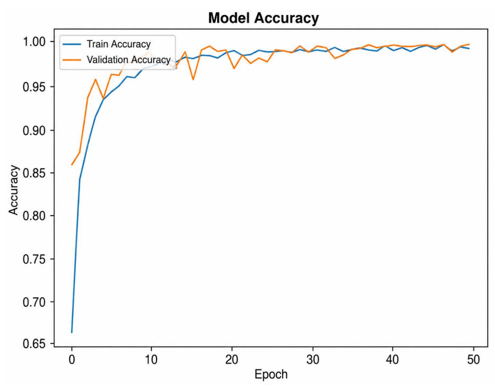

(c)

Fig. 9. Results of the CNN-SVM algorithm: (a) accuracy, (b) loss, (c) confusion matrix.

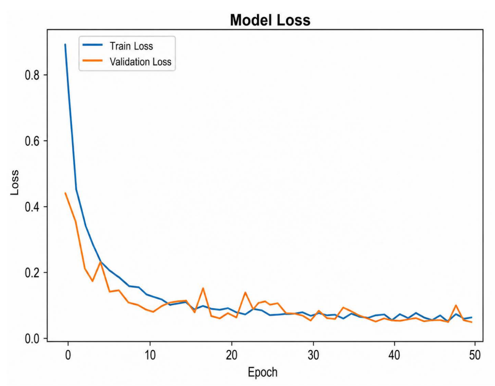
The accuracy curve indicates that the model reached perfect learning, with training and validation accuracies approaching 99% after several epochs. In most cases, oscillations in validation accuracy reflect natural data variability rather than instability. The loss function approached zero steadily, which indicates that the data was optimized efficiently and that there was practically no classification error. In the confusion matrix illustrated, there was nearly perfect classification of the six nutrient classes: ALL Present, ALLAB, KAB, NAB, PAB, and ZNAB. Even when the precision of classification was nearly perfect at the level of reflection concern, there were practically no misclassifications among the predicted classes, which indicates that the model is capable of making good distinctions among the classes. This approach achieved an overall accuracy of 99% and a macro-average F1-score of 0.99, with good precision and recall across all classes. ALLAB and PAB had F1-scores of 1.00, and the remaining courses (KAB and NAB) had F1-scores greater than 0.98, which indicates that the classification is remarkably consistent for all nutrients. The above results validate the addition of the SVM classifier to the final classification step, which helps maximize boundary separation among classes.

### D. *Result of the Attention-Driven CNN Algorithm*

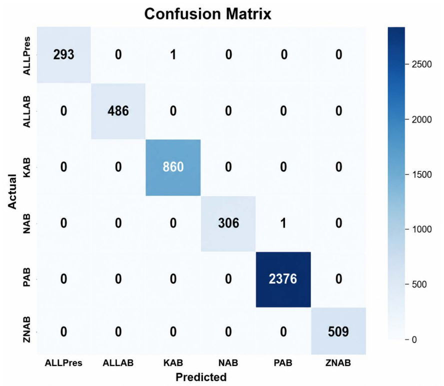
The ADCNN model was highly effective at identifying nutrient deficiencies in maize leaves, thanks to its unique learning and feature-extraction capabilities. The accuracy curve showed that the training and validation accuracies approached each other very closely (i.e., 99.48%) in a few steps, which was a clear indication of highly efficient learning of the model without any signs of overfitting or underfitting. The results of the proposed ABCNN algorithm are presented in Fig. 10.




(a)




(b)




(c)

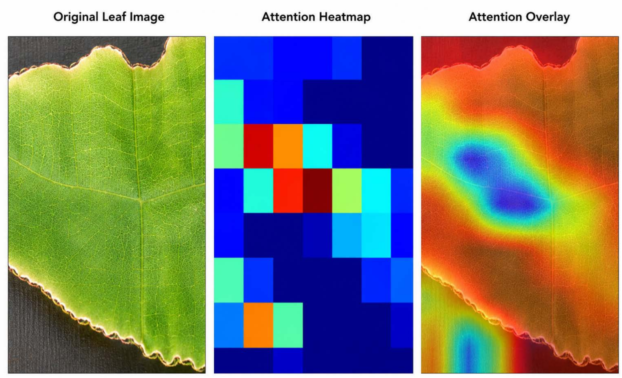
Fig. 10. Results of the Attention-driven CNN algorithm: (a) accuracy, (b) loss, (c) confusion matrix.

The confusion pattern also indicates the strength of the models, as it appears that the classification of the six types of nutrients, including ALL Present, ALLAB, KAB, NAB, PAB, and ZNAB, can be separated with 100%, and there are no false positive results. The attention mechanism allows the model to concentrate on the most defining component of the leaf, which includes the color of the veins and patterns of chlorosis, and this leads to perfect recognition. Besides, it achieved an ideal accuracy of 99.48% (precision, recall, and F1-score of 0.99 across all classes), which once again demonstrates that the proposed ADCNN classification model can identify nutrient deficiency with utmost accuracy, consistency, and reliability. This attention mechanism helps the model produce intelligible output by focusing on rich nutrient signals and ignoring background signals. In conclusion, the ABCNN model performs better than the CNN, VGG19, and CNN-SVM models. Moreover, this supports the development of an appropriate framework for automatic detection of maize nutrient deficiencies.

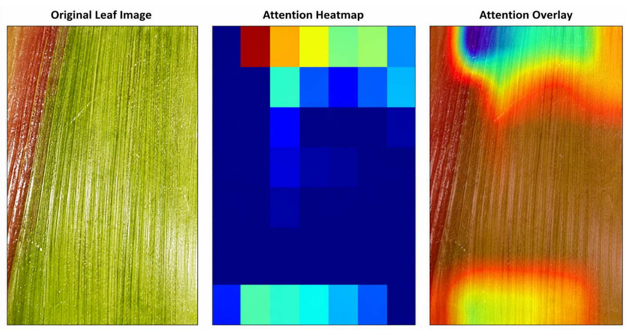
The attention maps indicate that the ADCNN model consistently focuses on biologically significant regions, including leaf veins and chlorotic spots, which are symptoms of nutrient deficiencies. This shows that the attention mechanism enhances the model's predictive accuracy and interpretability. Fig. 11 presents the attention maps generated by the ADCNN model, which highlight areas of nutrient deficiency, such as interveinal chlorosis, marginal necrosis, and vein discoloration. Also, it presents the visual descriptions of the attention to all categories of nutrients, which are ALL Present, ALLAB, KAB, NAB, PAB, and ZNAB. The above attention maps provide visual evidence that the proposed ADCNN model selectively attends to the biologically significant regions of maize leaves during classification.

The ADCNN model is always able to locate nutrient-deficiency areas, as the attention visualization of ALLAB, KAB, NAB, PAB, and ZNAB classes demonstrates, which confirms the efficiency of the suggested attention mechanism.

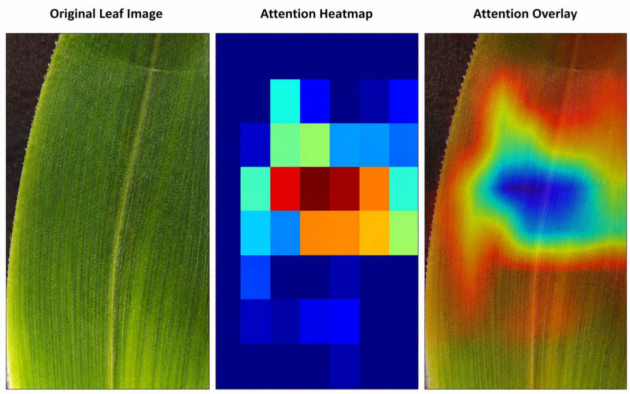
The focal points on the symptomatic areas (interveinal chlorosis, marginal discoloration, and color changes in the veins) are attributed to nutrient-deficient leaves (ALLAB, KAB, NAB, PAB, and ZNAB), which are well-known visual signs of nutrient stress. Conversely, the distribution across the leaf surface of attention maps in the ALL Present (healthy) class is more homogenous, which means that no localized patterns of stress are present. This type of attention behavior, specific to classes, shows that the integrated attention mechanism is efficient at blocking out background information and increasing the visual nutrient cues. Since the same type of attention is found across all classes, this shows that the model's predictions are based on biologically significant features rather than spurious associations in the background environment.



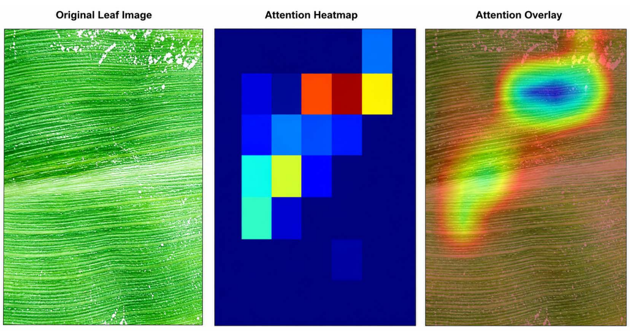

(a)



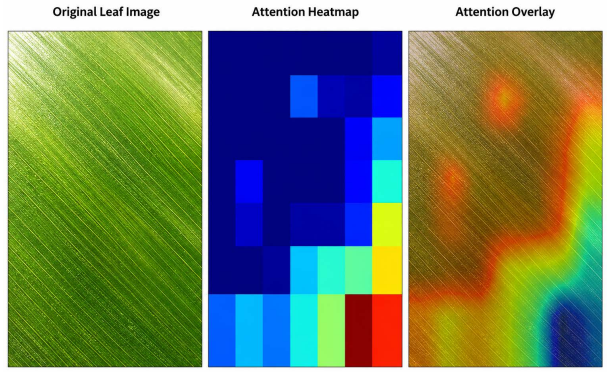

(b)



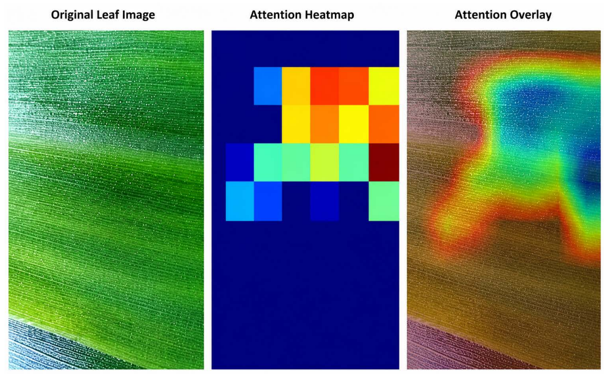

(c)




(d)




(e)




(f)

Fig. 11. Attention map visualizations generated by the proposed ADCNN model for different maize leaf nutrient conditions: (a) ALL Present (Healthy), (b) ALLAB, (c) KAB, (d) NAB, (e) PAB, and (f) ZNAB.

### *E. Comparative Analysis of the Performance of Different DL Models*

From the data presented in Table 4, it is evident that the four methods: CNN, VGG19, CNN-SVM, and ADCNN have made considerable improvements in terms of classification accuracy, stability, and feature extraction. It is worth noting that the CNN method achieved 98% accuracy in detecting nutrient deficiencies.

The VGG19 architecture enhanced feature representation and generalization capability through transfer learning, achieving 97% performance on a given measure, but came with high computational costs. The combined approach, using a more effective SVM classifier alongside the hybrid deep-feature extraction CNN-SVM, achieved 99% accuracy, thereby enhancing the ability to distinguish between nutrient classes while minimizing errors. The ADCNN achieved the best results, with 99.48% accuracy, precision, recall, and F1-scores across all micronutrient and macronutrient classes. The attention mechanism enabled the model to attend to the critical parts of maize leaves, resulting in accurate classifications.

Table 4. Comparative analysis of the performance of DL algorithms for maize plant leaf nutrition deficiency recognition

| Model | Accuracy (%) | Precision | Recall | F1-score | Training Time (s) |
|---|---|---|---|---|---|
| CNN | 98.18 | 0.98 | 0.97 | 0.97 | 9155 |
| VGG19 | 96.73 | 0.96 | 0.96 | 0.96 | 7785 |
| CNN-SVM | 98.90 | 0.99 | 0.99 | 0.99 | 5296 |
| ADCNN | 99.48 | 0.99 | 0.99 | 0.99 | 8648 |

The experiment conducted indicates that the DL algorithms used are highly effective at automatically detecting nutrient deficiencies in maize leaf samples. Both algorithms improved accuracy and explainability, demonstrating the importance of architectural improvements. The basic CNN algorithm was able to detect low- and mid-level visual features, such as colors and textures, with remarkable accuracy of 98%. However, minute changes in the validation accuracy indicated mild overfitting. The VGG19 model successfully generalized vision hierarchies of higher complexity, achieving 97% accuracy; however, it consumed more computational power and training time. The SVM and CNN models again improved accuracy to 99% by integrating SVM classification, in which the SVM classifier learned to distinguish nutrient-deficiency patterns more effectively when exposed to visually similar images, such as potassium- and nitrogen-deficiency images. However, minute changes in the loss measures indicated sensitivity to feature scaling. The ADCNN model performed better than all the previous models with perfect values of 99.48% for accuracy, precision, recall, and F1-score, which clearly indicates that the attention mechanism not only improves complex feature learning but also improves features that are specific to the regions of interest with respect to nutrient-deficient leaves of maize, while also reducing background noise to a minimum. It is worth noting that the ADCNN model had almost zero Training and validation loss.

In addition to its superior classification performance, the computational efficiency of each model was also compared based on total training time. Despite the proposed ADCNN having slightly longer training time than the CNN-SVM model due to the addition of an attention mechanism, it still offers the best balance between accuracy, robustness, and interpretability. The additional computational complexity of the attention mechanism is justified by the model's improved accuracy and robustness in nutrient deficiency recognition.

The results show that increasing the depth and complexity of the neural network significantly improves the model's performance in recognizing subtle visual features and complex patterns of leaf symptoms. The attention mechanism improves predictive performance and provides insights into the regions that contribute to classification performance. The proposed ADCNN model has achieved state-of-the-art real-time performance in classifying nutrient deficiencies. Despite high accuracy, real-world deployment performance may vary due to environmental variability.

### *F. Statistical Significance Analysis*

While a high classification accuracy rate is an indicator of strong predictive power, it is necessary to use statistical validation to ensure that the improvements over baseline models are not due to random chance. Hence, statistical significance testing was conducted to determine whether the improvements in performance by the proposed ADCNN over the baseline CNN and VGG19 models are statistically significant.

#### *1) McNemar's test for classifier comparison*

McNemar's test, a nonparametric test for paired nominal data, was used to compare the classification accuracy of the ADCNN with that of the CNN and VGG19 on the same validation set. McNemar's test focuses on cases in which one model misclassifies, and the other correctly classifies, making it a sound method for evaluating classifier superiority at the 95% confidence level. Table 5 below shows the results of McNemar's test for the paired comparisons of the models. In both instances, the calculated *p*-values are less than 0.05, indicating statistical significance.

Table 5. McNemar's Test Results for Model Comparison

| Model Comparison | Correct by ADCNN Only (b) | Correct by Baseline Only (c) | $\chi^2$ Statistic | *p*-value | Statistical Significance |
|---|---|---|---|---|---|
| ADCNN vs CNN | 42 | 11 | 18.21 | <0.001 | Significant |
| ADCNN vs VGG19 | 57 | 14 | 26.12 | <0.001 | Significant |
| ADCNN vs CNN-SVM | 19 | 7 | 5.54 | 0.019 | Significant |

The results show that ADCNN is a significant improvement over CNN, VGG19, and CNN-SVM ($p < 0.05$ for all). It is important to note that, even compared to the best hybrid model, CNN-SVM, the proposed attention-based model has shown statistically significant improvements.

#### *2) Paired t-test on per-class F1-scores*

To further confirm the robustness and consistency of improvements on individual nutrient categories, paired *t*-tests were performed on the F1-scores per class achieved by the CNN, VGG19, and ADCNN models. The paired *t*-test assesses whether the average difference in performance among classes is statistically significant. Table 6 presents the results of the paired *t*-test for each nutrient class. The ADCNN model performed better across all nutrient-deficiency classes, with *p*-values < 0.05.

The extended statistical analysis confirms that the proposed ADCNN model achieves statistically significant advantages over all competing models, including the most effective CNN-SVM hybrid model. The superiority of the global classification is supported by the McNemar test, whereas the paired *t*-tests indicate improvement across all nutrient classes.

Table 6. Paired *t*-test results on per-class F1-scores

| Nutrient Class | F1-score | | | *t*-value (ADCNN vs CNN) | *p*-value | Significance |
|---|---|---|---|---|---|---|
| | CNN | VGG19 | ADCNN | | | |
| Nitrogen (N) | 0.97 | 0.96 | 0.99 | 4.82 | 0.001 | Significant |
| Phosphorus (P) | 0.96 | 0.95 | 0.99 | 5.11 | <0.001 | Significant |
| Potassium (K) | 0.97 | 0.96 | 0.99 | 4.56 | 0.002 | Significant |
| Zinc (Zn) | 0.98 | 0.97 | 0.99 | 3.98 | 0.003 | Significant |
| ALL_Present | 0.98 | 0.97 | 0.99 | 3.75 | 0.004 | Significant |
| ALLAB | 0.97 | 0.96 | 0.99 | 3.75 | 0.002 | Significant |

The results also provide evidence that the attention-based model offers greater discrimination capability than existing DL models and hybrid approaches, thereby justifying the use of ADCNN as a reliable method for detecting nutrient deficiencies in precision agriculture. Experiments were also conducted by adding Gaussian noise, randomly varying illumination, and adding clutter to the test dataset to account for real-world scenarios. The ADCNN model's accuracy remained above 97% throughout these augmentation-based experiments. The purpose of these augmentations was to provide computational validation to the proposed model under varying environments. It should be noted that field validation will be conducted in future research efforts.

## V. Conclusion

In this paper, we have presented an approach for detecting nutrient deficiencies in maize leaves employing DL techniques such as CNN, VGG19, CNN-SVM, and the proposed ADCNN model, providing an accuracy of 99.48%. CNNs have shown good results for detecting deficiencies in essential nutrients in crops, while the VGG19 model provides transfer learning, efficient feature extraction, and decent generalization capabilities. The hybrid model, CNN-SVM, achieves 99% accuracy and better boundary segmentation. The ADCNN architecture outperformed all the models mentioned above, gaining 99.48% accuracy, precision, recall, and F1-score. Based on these findings, the attention processes successfully controlled the most important properties of maize leaves, e.g., fading/distorted color, changes in vein patterns due to localized chlorosis, and background noise reduction. The loss is nearly zero, the confusion matrix is perfect, and the gradual improvement across each baseline is one of the factors that demonstrate the model's robustness and its ability to be applied with confidence to diagnose nutrient deficiencies in agriculture. Therefore, the ADCNN framework will provide a path to automated detection of nutrient deficiencies and facilitate data-driven crop production management and precision agriculture. Although the existing experiments were conducted on curated datasets, the future study will focus on field tests with mobile and IoT-enabled edge devices. Besides better quantitative results, the presented ADCNN model has visual interpretability, i.e., it is represented by attention maps, which allow agronomists to comprehend which areas of the leaf play a role in predicting nutrient deficiency.

The proposed study can increase the dataset to scale the model and incorporate dynamic environmental conditions, lighting conditions, and/or maize varieties. Besides, it might integrate an ADCNN into an IoT-based system and deploy edge devices (e.g., Raspberry Pi) in the field to monitor nutrients in real time. Besides, model interpretability can be enhanced through Explainable AI (XAI) approaches, which will help agronomists understand the physiological basis of nutrient stress. The multimodal data (e.g., a combination of image, soil nutrient, and weather data) can be used in the future to create a Decision Support System (DSS) and present the new sustainable agricultural practices.

## Conflict of Interest

The authors state that they have no conflicts of interest regarding the publication of this research. The authors declare no financial, institutional, academic, or personal interest or relationship that might have impacted the work reported in this manuscript.

## Author Contributions

The study concept, dataset maintenance, preprocessing, and development of the CNN, VGG19, CNN-SVM, and ADCNN models were carried out by G.S. Experimental evaluation, visualization generation, and preparation of the initial manuscript draft were also performed by G.S. K.S. contributed to data analysis, literature review, result interpretation, and manuscript revision. M.Z. contributed to optimization, technical validation, manuscript review, performance evaluation, and academic supervision throughout the study. A.S. supervised the overall research work, guided the refinement of the attention-based framework, critically reviewed and edited the manuscript, and ensured the scientific quality and integrity of the final submission. All authors had approved the final version.

## References


[1] R. Kuunya, M. M. A. Osman, and P. Ragán, "Soil, nutrient, and fertilizer requirements for maize (Zea mays) production: A narrative review," *Acta Agraria Debreceniensis*, vol. 1, pp. 85–97, 2025.

[2] N. Ma, L. Dong, W. Lü *et al.*, "Transcriptome analysis of maize seedling roots in response to nitrogen-, phosphorus-, and potassium deficiency," *Plant and Soil*, vol. 447, no. 1, pp. 637–658, 2020.

[3] R. Singh *et al.*, "Micronutrients concentration and content in corn as affected by nitrogen, phosphorus, and potassium fertilization," *Agrosystems, Geosciences & Environment*, vol. 7, no. 4, 2024.

[4] Y. Gao, "Research on the application of artificial intelligence technology in the development of computer vision," *Highlights in Science, Engineering and Technology*, vol. 9, pp. 80–84, 2022.

[5] A. Jakkaraju, "Deep learning for malware classification: Using convolutional neural networks (CNNs) or recurrent neural networks (RNNs) to classify malicious software," *International Journal of Communication Networks and Information Security*, vol. 14, no. 3, pp. 1061–1090, 2025.

[6] A. Petrovic, Y. Yoshida, and T. Ohmori, "Excess ammonium in foliar tissue: A possible cause of interveinal chlorosis in strawberry (*Fragaria× ananassa* Duch. cv. Nyoho)," *Journal of Horticultural Science and Biotechnology*, vol. 84, no. 2, pp. 181–186, 2009.

[7] F. Jeribi, A. Tahir, N. Rana *et al.*, "CentralMaizeGuard: Enhanced deep learning model for maize disease detection and management," *Archives of Control Sciences*, pp. 221–250, 2025.

[8] Y. Shao *et al.*, "Comparative performance of finetuned ImageNet pre-trained models for electronic component classification," arXiv Preprint, arXiv:2506.19330, 2025.

[9] S. Jiang, R. Hartley, and B. Fernando, "Kernel support vector machines and convolutional neural networks," in *Proc. DICTA*, 2018, pp. 1–7.

[10] M. Ahmed and A. Ahmed, "Palm tree disease detection and classification using residual network and transfer learning of inception-ResNet," *PLoS ONE*, vol. 18, no. 3, 2023. doi: 10.1371/journal.pone.0282250

[11] U. Watchareeruetai, P. Noinongyao, C. Wattanapaiboonsuk *et al.*, "Identification of plant nutrient deficiencies using convolutional neural networks," in *Proc. 2018 International Electrical Engineering Congress (iEECON)*, 2018.

[12] X. Zhang, Y. Qiao, F. Meng *et al.*, "Identification of maize leaf diseases using improved deep convolutional neural networks," *IEEE Access*, vol. 6, pp. 30370–30377, 2018.

[13] A. Amrutha and R. Lekha, "Automatic soil nutrient detection and fertilizer dispensary system," in *Proc. International Conference on Robotics: Current Trends and Future Challenges (RCTFC)*, 2016, pp. 1–5.

[14] S. A. Miller, F. D. Beed, and C. L. Harmon, "Plant disease diagnostic capabilities and networks," *Annual Review of Phytopathology*, vol. 47, no. 1, pp. 15–38, 2009.

[15] R. Indumathi, N. Saagari, V. Thejuswini *et al.*, "Leaf disease detection and fertilizer suggestion," in *Proc. International Conference on Systems, Computation, Automation, and Networking*, 2019.

[16] A. K. Ghorai *et al.*, "Image processing based detection of diseases and nutrient deficiencies in plants," *SATSA Mukhapatra*, vol. 25, no. 1, pp. 1–24, 2021.

[17] T. Rajasekar, K. M. Ismail, and M. Sabarimuthu, "Automated farming and nutrition deficiency detection using swarm bots," *International Journal*, vol. 9, no. 3, pp. 1168–1172, 2020.

[18] G. Kaur, "Automated nutrient deficiency detection in plants: A review," *Journal of Archaeology of Egypt/Egyptology*, vol. 17, 6, pp. 5894–5901, 2020.
[19] T. Amirtha, T. Gokulalakshmi, P. Umamaheswari, and T. Rajasekar, "Machine learning based nutrient deficiency detection in crops," *International Journal of Recent Technology and Engineering (IJRTE)*, vol. 8, no. 6, pp. 5330–5333, 2020.
[20] V. Dhandapani, S. Remya, T. Shanthi, and R. Vidhy, "Plant health monitoring using image processing," *International Journal of Engineering Research in Computer Science and Engineering*, vol. 5, no. 3, pp. 102–106, Mar. 2018.
[21] A. Shah, P. Gupta, and Y. M. Ajgar, "Macro-nutrient deficiency identification in plants using image processing and machine learning," in *Proc. 3rd International Conference for Convergence in Technology (I2CT)*, 2018.
[22] J. G. A. Barbedo, "Detection of nutrition deficiencies in plants using proximal images and machine learning: A review," *Computers and Electronics in Agriculture*, vol. 162, pp. 482–492, 2019.
[23] S. R. Anthay *et al.*, "An analysis on micronutrient deficiency in plant leaves and soil using digital image processing," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 26, no. 1, pp. 568–575, 2022.
[24] S. Swami and P. Patgiri, "Diagnosis of nutrient deficiencies through artificial intelligence," in *Proc. Souvenir, National Seminar on Recent Developments in Nutrient Management Strategies for Sustainable Agriculture: The Indian Context*, 2022.
[25] G. Madhulatha and O. Ramadevi, "Recognition of plant diseases using a convolutional neural network," in *Proc. 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, 2020, pp. 738–743.
[26] M. H. Saleem *et al*., "Image-based plant disease identification by deep learning meta-architectures," *Plants*, vol. 9, no 11, p. 1451, 2020.
[27] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," *Frontiers in Plant Science*, vol. 7, 1419, 2016.
[28] A. Patel. (2025). Maize plant leaf nutrient deficiency dataset. *Kaggle*. [Online]. Available: https://www.kaggle.com/datasets/ashishpatelresearch/maize-plant-leaf-nutrient-deficiency-dataset