International Journal of Research and Technology (IJRT)

International Open-Access, Peer-Reviewed, Refereed, Online Journal

ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529

| An ISO 9001:2015 Certified Journal |

Intelligent Tomato Leaf Nutrient Deficiency Detection and Classification Using Convolutional Neural Networks

Aditi Bhushan

Department of Mathematics for Computer Science, Babasaheb Bhimrao Ambedkar Bihar University, Muzaffarpur

aditibhushan90@gmail.com

Dr. Mala

Assistant Professor (Senior Scale), , Department of Mathematics, M.D.D.M. College, Babasaheb Bhimrao Ambedkar Bihar University, Muzaffarpur

dr.mala.mddmcollege@gmail.com

Abstract

Nutrient deficiency in tomato (Solanum lycopersicum) crops is a major factor limiting yield and fruit quality across the world, and its early diagnosis remains largely dependent on manual inspection by agronomists, a process that is slow, subjective, and difficult to scale to large farms. This paper presents an intelligent, image-based system for the automatic detection and classification of nutrient deficiencies in tomato leaves using Convolutional Neural Networks (CNNs). A curated dataset of tomato leaf images spanning healthy leaves and leaves affected by nitrogen, phosphorus, potassium, magnesium, calcium, and iron deficiencies was preprocessed using resizing, normalization, background segmentation, and augmentation techniques to improve model generalization. A custom CNN architecture is proposed and benchmarked against transfer-learning-based models, namely VGG16, ResNet50, MobileNetV2, InceptionV3, and EfficientNetB0. Experimental results show that the fine-tuned EfficientNetB0 model achieves the highest classification accuracy of 97.4%, outperforming the custom CNN (93.8%) and other baseline models, while MobileNetV2 offers the best accuracy-to-computation trade-off for deployment on resource-constrained edge devices. The proposed system demonstrates strong potential for integration into smartphone-based or drone-assisted precision agriculture tools, enabling farmers to obtain rapid, low-cost, and reliable diagnosis of nutrient stress and take corrective fertilization measures in a timely manner.

Keywords—Convolutional Neural Network, Deep Learning, Tomato Leaf, Nutrient Deficiency, Precision Agriculture, Image Classification, Transfer Learning, Plant Health Monitoring

1. INTRODUCTION

Agriculture continues to be the backbone of food security for a rapidly growing global population, and tomato is one of the most widely cultivated vegetable crops owing to its high nutritional value, economic importance, and versatility in both fresh consumption and processed food industries. However, tomato plants are highly sensitive to fluctuations in soil nutrient availability, and deficiencies in essential macro- and micro-nutrients such as nitrogen (N), phosphorus (P), potassium (K), magnesium (Mg), calcium (Ca), and iron (Fe) manifest as visible symptoms on the foliage long before the effects become irreversible. Symptoms such as chlorosis, necrosis, interveinal yellowing, leaf curling, and stunted growth are often the first indicators of an underlying nutrient imbalance, yet these visual cues are frequently misdiagnosed by inexperienced farmers or confused with disease symptoms caused by fungal, bacterial, or viral pathogens.

Traditional approaches to nutrient deficiency diagnosis rely on laboratory-based soil and tissue analysis or expert visual inspection, both of which are time-consuming, resource-intensive, and impractical for real-time, large-scale farm monitoring. The delay between symptom onset and corrective fertilization can result in substantial yield losses. In recent years, advances in computer vision and deep learning have opened new avenues for automated, non-invasive plant health monitoring using only leaf images captured through smartphones, hand-held cameras, or unmanned aerial vehicles (UAVs).

Tomato is cultivated across millions of hectares worldwide and represents one of the highest-value vegetable crops in terms of both production volume and per-hectare economic returns, making even modest improvements in early deficiency detection economically significant at scale. Nutrient deficiencies in tomato typically arise from a combination of factors, including inherently poor soil fertility, incorrect or imbalanced fertilizer application, unfavorable soil pH that limits nutrient uptake, excessive irrigation leading to nutrient leaching, and competition from weeds or intercropped species. Because different nutrients play distinct physiological roles—nitrogen in chlorophyll and protein synthesis, phosphorus in energy transfer and root development, potassium in water regulation and enzyme activation, magnesium as the central atom of chlorophyll, calcium in cell wall structural integrity, and iron in chlorophyll synthesis and electron transport—each deficiency produces a visually distinct, though sometimes overlapping, foliar symptom pattern that can, in principle, be learned by a sufficiently expressive image classification model.

Smallholder farmers, who collectively produce a substantial share of the world's tomato output, particularly in South Asian and Sub-Saharan African contexts, often lack timely access to agricultural extension officers or laboratory testing facilities. In such settings, a smartphone-based diagnostic tool capable of instantly classifying a photographed leaf into a specific deficiency category—and by extension recommending a corrective fertilizer regimen—could substantially reduce the diagnostic delay, lower dependency on expert visits, and ultimately improve both yield and input-use efficiency by discouraging blanket, non-targeted fertilizer application, which is both economically wasteful and environmentally harmful due to nutrient runoff.

Convolutional Neural Networks (CNNs) have emerged as the state-of-the-art tool for image classification tasks due to their ability to automatically learn hierarchical spatial features directly from raw pixel data, eliminating the need for hand-crafted feature engineering. CNN-based approaches have already demonstrated remarkable success in plant disease classification tasks using datasets such as PlantVillage; however, the specific problem of nutrient deficiency classification—as opposed to pathogen-based disease detection—remains comparatively underexplored, largely due to the scarcity of well-labeled, deficiency-specific datasets and the subtlety and overlap of visual symptoms across different deficiencies.

This paper addresses this gap by proposing an intelligent CNN-based framework specifically designed for tomato leaf nutrient deficiency detection and classification. The key contributions of this work are summarized as follows:

- Compilation and preprocessing of a labeled tomato leaf image dataset covering six nutrient deficiency classes and a healthy control class.
- Design of a lightweight custom CNN architecture optimized for nutrient deficiency classification.
- Comparative evaluation of five state-of-the-art transfer learning architectures (VGG16, ResNet50, MobileNetV2, InceptionV3, EfficientNetB0) fine-tuned on the target dataset.
- Comprehensive performance analysis using accuracy, precision, recall, F1-score, and confusion matrices, along with a discussion of deployment feasibility on edge devices for real-world precision agriculture applications.

The remainder of this paper is organized as follows. Section II reviews related work in plant disease and nutrient deficiency detection using deep learning. Section III presents the problem statement and objectives. Section IV describes the proposed methodology, including dataset preparation, preprocessing, augmentation, and model architectures. Section V outlines the experimental setup. Section VI presents and discusses the results. Section VII compares the proposed approach with existing literature. Section VIII discusses limitations, Section IX outlines future work, and Section X concludes the paper.

2. RELATED WORK

The application of deep learning to agricultural image analysis has grown substantially over the past decade. Mohanty et al. demonstrated that deep CNNs trained on the PlantVillage dataset could classify 26 diseases across 14 crop species with high accuracy, establishing CNNs as a viable tool for automated plant health diagnostics. Following this, numerous studies applied architectures such as AlexNet, VGGNet, GoogLeNet, and ResNet to plant disease classification tasks, generally reporting accuracies above 95% on controlled, laboratory-captured datasets.

Ramcharan et al. explored transfer learning for cassava disease detection using mobile-captured images, highlighting the practical challenges of applying models trained on clean, curated datasets to noisy field conditions with variable lighting, background clutter, and occlusion. Similarly, Ferentinos evaluated multiple CNN architectures across 58 distinct classes of plant diseases and healthy leaves, reporting that deeper architectures such as VGG achieved superior performance at the cost of higher computational requirements.

Compared to disease detection, research specifically targeting nutrient deficiency classification is comparatively limited. Existing efforts have primarily focused on nitrogen deficiency detection in crops such as rice and maize using spectral imaging and hyperspectral sensors, which, while accurate, require specialized and expensive equipment not accessible to smallholder farmers. A smaller body of work has explored RGB-image-based nutrient deficiency classification using traditional machine learning classifiers such as Support Vector Machines (SVM) and Random Forests combined with hand-crafted color and texture features (e.g., color histograms, GLCM texture descriptors), but these approaches tend to generalize poorly to varying illumination and background conditions.

More recent studies have begun applying CNN-based transfer learning to nutrient deficiency problems in crops such as rice, banana, and coffee, reporting promising results in the 90–96% accuracy range. However, tomato-specific nutrient deficiency classification using deep learning remains sparsely studied, with most existing tomato-related CNN research focused on fungal and bacterial disease classification rather than nutrient stress. Furthermore, many existing studies do not provide a systematic comparison across multiple modern CNN backbones with a focus on deployment feasibility for real-world farm use, which is a gap that the present work aims to address.

Fuentes et al. proposed a robust real-time detector for tomato diseases and pests combining region-based CNNs with deep feature extractors, achieving strong detection performance across nine disease and pest classes, but did not address nutrient-related stress conditions. Brahimi et al. applied CNNs to nine tomato disease categories and additionally employed visualization techniques to highlight the leaf regions most responsible for the model's predictions, underscoring the growing importance of interpretability alongside raw accuracy in agricultural deep learning applications. Durmuş et al. compared AlexNet and SqueezeNet architectures for tomato disease classification on embedded platforms, demonstrating that lightweight architectures can achieve competitive accuracy while being suitable for deployment on low-power devices such as Raspberry Pi units used in field settings.

Several works have also investigated the use of attention mechanisms and ensemble strategies to boost classification robustness under field conditions. Too et al. conducted a systematic comparison of fine-tuning strategies across VGG16, InceptionV3, ResNet, and DenseNet architectures for plant disease identification, concluding that deeper architectures with residual or dense connections generally offer better convergence stability, albeit with increased computational cost. Ghosal et al. proposed an explainable deep learning framework for soybean stress phenotyping that combined classification with saliency-based visual explanations, an approach that is highly relevant to nutrient deficiency diagnosis since farmers and agronomists benefit from understanding which leaf regions and visual cues drive

ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529
| An ISO 9001:2015 Certified Journal |

a given prediction. These studies collectively motivate the present work's emphasis on both predictive accuracy and practical deployability.

A parallel body of research has examined the broader challenge of limited labeled agricultural data, which is a recurring bottleneck across nearly all plant image classification tasks, including nutrient deficiency detection. Transfer learning, in which a CNN pretrained on a large, general-purpose dataset such as ImageNet is fine-tuned on a smaller, domain-specific dataset, has emerged as the dominant strategy for addressing this data scarcity, since it allows the network to reuse generic low-level visual features (edges, color gradients, textures) learned from millions of natural images rather than learning them from scratch on a comparatively small agricultural dataset. This strategy underpins the transfer-learning-based models evaluated in the present study and is a key reason why such models are expected to outperform a custom CNN trained entirely from randomly initialized weights on a moderately sized dataset.

Overall, the literature reveals three recurring gaps that this study seeks to address: (i) a scarcity of curated, deficiency-specific tomato leaf datasets suitable for supervised deep learning; (ii) a lack of systematic, head-to-head comparison of modern CNN backbones specifically for nutrient stress classification rather than pathogen-based disease detection; and (iii) limited attention to deployment-oriented metrics such as model size and inference latency, which are critical for real-world adoption by resource-constrained smallholder farmers.

3. PROBLEM STATEMENT AND OBJECTIVES

Manual identification of tomato leaf nutrient deficiencies is inherently subjective and requires domain expertise that is often unavailable to smallholder farmers, particularly in developing regions where agricultural extension services are limited. Visual symptoms of different deficiencies frequently overlap—for instance, both nitrogen and iron deficiency can cause leaf chlorosis—making accurate manual differentiation difficult even for experienced agronomists. There is, therefore, a pressing need for an automated, accurate, and computationally efficient system capable of classifying tomato leaf nutrient deficiencies directly from RGB images captured under real-world field conditions.

The specific objectives of this study are as follows:

- To design and preprocess a labeled tomato leaf dataset representing healthy foliage and six common nutrient deficiency conditions.
- To develop a custom lightweight CNN architecture tailored to the subtle visual patterns associated with nutrient stress.
- To benchmark the custom CNN against established transfer-learning architectures in terms of classification accuracy, training efficiency, and inference speed.

- To identify the most suitable model configuration for deployment in resource-constrained, real-time agricultural monitoring applications such as smartphone apps or drone-mounted embedded systems.

Formally, the problem can be framed as a multi-class image classification task: given an input RGB leaf image x, the goal is to learn a mapping function $f(x) \rightarrow y$, where y belongs to the set of seven mutually exclusive classes {Healthy, N-Deficient, P-Deficient, K-Deficient, Mg-Deficient, Ca-Deficient, Fe-Deficient}. The learned function f is parameterized by a CNN whose weights are optimized to minimize the categorical cross-entropy loss between predicted and true class labels over a labeled training set. The central research question addressed in this paper is therefore: which CNN architecture, among a custom lightweight design and several established transfer-learning backbones, provides the best balance of classification accuracy and computational efficiency for this specific agricultural use case?

4. PROPOSED METHODOLOGY

A. System Overview

The proposed system follows a standard supervised deep learning pipeline comprising four major stages: (1) data acquisition and labeling, (2) image preprocessing and augmentation, (3) CNN-based feature extraction and classification, and (4) performance evaluation. Raw leaf images are first captured or sourced, followed by preprocessing steps to standardize image quality, after which augmented images are fed into the CNN models for training. The trained models are evaluated on a held-out test set to determine the most effective architecture for deployment.

Conceptually, the pipeline can be viewed as a five-stage flow: Image Acquisition → Preprocessing (resizing, normalization, background segmentation) → Augmentation → CNN Feature Extraction and Classification → Deficiency Class Output with confidence score. This modular design allows individual stages, such as the classification backbone, to be swapped or upgraded independently without redesigning the entire pipeline, which is particularly useful as newer, more efficient CNN architectures continue to emerge. The final output of the system is a predicted deficiency class label accompanied by a softmax confidence score, which can optionally be thresholded to flag low-confidence predictions for manual expert review, a safeguard that is important in agricultural decision-support tools where incorrect fertilizer recommendations can have real economic consequences for farmers.

B. Dataset Description and Acquisition

The dataset used in this study comprises RGB images of tomato leaves organized into seven classes: Healthy, Nitrogen-Deficient, Phosphorus-Deficient, Potassium-Deficient, Magnesium-Deficient, Calcium-Deficient, and Iron-Deficient. Images were collected from a combination of controlled greenhouse photography and publicly available agricultural image repositories, supplemented with images captured under natural field lighting to improve robustness. Each image was manually verified and labeled with the guidance of agricultural

International Journal of Research and Technology (IJRT)
International Open-Access, Peer-Reviewed, Refereed, Online Journal
ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529
| An ISO 9001:2015 Certified Journal |

domain experts based on characteristic visual symptoms, such as interveinal chlorosis for magnesium deficiency, purpling of leaf veins for phosphorus deficiency, and marginal leaf scorch for potassium deficiency. Table I summarizes the class-wise distribution of the compiled dataset.

The annotation protocol followed a two-stage verification process: images were first pre-labeled based on the known growing conditions under which they were captured (e.g., leaves from plants grown under intentionally nitrogen-limited soil in the greenhouse trials), after which a second independent visual verification pass was conducted by an agricultural domain expert to confirm the presence of characteristic symptoms and to discard ambiguous or borderline images. Inter-annotator agreement between the initial condition-based labels and the expert visual verification was measured using Cohen's kappa coefficient, yielding a value of 0.91, indicating substantial agreement and lending confidence to the reliability of the ground-truth labels used for model training and evaluation.

TABLE I. CLASS-WISE DATASET DISTRIBUTION

Class	No. of Images	Symptom Description
Healthy	620	Uniform green, no lesions
Nitrogen (N)	540	Uniform yellowing, older leaves first
Phosphorus (P)	480	Purplish veins, dark green leaves
Potassium (K)	510	Marginal scorch, necrotic edges
Magnesium (Mg)	460	Interveinal chlorosis, older leaves
Calcium (Ca)	430	Blossom-end rot, tip burn
Iron (Fe)	400	Interveinal chlorosis, young leaves

C. Image Preprocessing

All images were resized to a fixed resolution of 224×224 pixels to match the input requirements of the CNN architectures used. Pixel intensity values were normalized to the range [0, 1] by dividing by 255. Background segmentation was applied using a combination of HSV color-space thresholding and morphological operations to isolate the leaf region from

International Journal of Research and Technology (IJRT)
International Open-Access, Peer-Reviewed, Refereed, Online Journal
ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529
| An ISO 9001:2015 Certified Journal |

soil, pot, or other background clutter, thereby reducing the influence of irrelevant background features on classification. Images with excessive blur, occlusion, or ambiguous symptoms were manually discarded during a quality-control pass.

D. Data Augmentation

To mitigate overfitting and improve model generalization given the moderate dataset size, extensive data augmentation was applied during training, including random rotations (±25°), horizontal and vertical flips, zoom variation (0.8×–1.2×), brightness adjustment (±20%), and width/height shifts (up to 15%). These augmentations simulate the variability encountered in real-world field image capture, such as differing camera angles, lighting conditions, and leaf orientations.

E. Custom CNN Architecture

A lightweight custom CNN was designed specifically for this task, consisting of four convolutional blocks, each comprising a convolutional layer with ReLU activation, batch normalization, and max-pooling, followed by two fully connected layers with dropout regularization (rate = 0.5) to reduce overfitting, and a final softmax layer producing probability scores across the seven output classes. The architecture is summarized in Table II.

TABLE II. CUSTOM CNN ARCHITECTURE SUMMARY

Layer	Output Shape	Parameters
Conv2D (32, 3×3) + BN + ReLU	222×222×32	896
MaxPooling2D (2×2)	111×111×32	0
Conv2D (64, 3×3) + BN + ReLU	109×109×64	18,496
MaxPooling2D (2×2)	54×54×64	0
Conv2D (128, 3×3) + BN + ReLU	52×52×128	73,856
MaxPooling2D (2×2)	26×26×128	0
Conv2D (256, 3×3) + BN + ReLU	24×24×256	295,168
MaxPooling2D (2×2)	12×12×256	0
Flatten	36864	0

Layer	Output Shape	Parameters
Dense (256) + Dropout(0.5)	256	9,437,440
Dense (7, Softmax)	7	1,799

F. Transfer Learning Models

In addition to the custom CNN, five widely used pretrained architectures—VGG16, ResNet50, MobileNetV2, InceptionV3, and EfficientNetB0—were fine-tuned on the tomato leaf dataset. Each model was initialized with ImageNet pretrained weights, with the top classification layers removed and replaced with a global average pooling layer, a dense layer of 256 units with ReLU activation, a dropout layer, and a final softmax layer for seven-class classification. The initial convolutional base layers were frozen for the first 10 training epochs to allow the newly added classification head to stabilize, after which the last 20% of base layers were unfrozen and fine-tuned at a reduced learning rate to adapt the pretrained features to the tomato leaf domain.

G. Training Configuration

All models were trained using the Adam optimizer with an initial learning rate of 0.001, reduced by a factor of 0.5 upon validation loss plateauing over 5 consecutive epochs. Categorical cross-entropy was used as the loss function. Models were trained for a maximum of 50 epochs with early stopping (patience = 8) based on validation loss, using a batch size of 32. All experiments were conducted using the TensorFlow/Keras framework on a workstation equipped with an NVIDIA RTX 3080 GPU (10 GB VRAM), an Intel Core i7 processor, and 32 GB of system RAM.

H. Hyperparameter Tuning

A grid search was conducted over key hyperparameters, including learning rate, batch size, dropout rate, and optimizer choice, using the validation set to select the best-performing configuration for each model. Table III(a) summarizes the hyperparameter search space explored during tuning. The final configurations reported in Section V correspond to the combination yielding the lowest validation loss for each respective architecture.

TABLE III(A). HYPERPARAMETER SEARCH SPACE

Hyperparameter	Values Explored	Selected
Learning rate	1e-2, 1e-3, 1e-4	1e-3
Batch size	16, 32, 64	32
Dropout rate	0.3, 0.5, 0.7	0.5
Optimizer	SGD, Adam, RMSprop	Adam
Fine-tune layers	10%, 20%, 30%	20%

I. Computational Complexity Analysis

Beyond raw classification accuracy, the computational footprint of each model was assessed in terms of the number of trainable parameters, floating-point operations (FLOPs) per inference, and model file size after export, since these factors directly determine feasibility for deployment on mobile phones or low-power embedded boards commonly used in rural field settings. Models with a smaller memory footprint and lower FLOPs, such as MobileNetV2 and the custom CNN, are better suited for on-device inference without reliance on a stable internet connection, whereas larger models such as VGG16 are more appropriately hosted on a cloud backend accessed via an API call from a lightweight mobile front-end.

J. Software Environment and Reproducibility

All experiments were implemented in Python 3.9 using TensorFlow 2.x and the Keras high-level API for model construction and training, with scikit-learn used for computing evaluation metrics and OpenCV used for image preprocessing and background segmentation operations. Random seeds were fixed for NumPy, TensorFlow, and Python's built-in random module prior to each training run to maximize reproducibility, though minor non-determinism inherent to GPU-accelerated floating-point operations was still observed across repeated runs, motivating the multi-run averaging strategy described in Section V. All trained model weights, along with the preprocessing and augmentation pipeline configuration, were version-controlled to ensure that reported results can be independently reproduced or extended by future researchers working on related agricultural image classification problems.

K. Model Interpretability

To improve the transparency and trustworthiness of model predictions for end users such as farmers and agricultural extension officers, Gradient-weighted Class Activation Mapping (Grad-CAM) was applied to the best-performing model to visualize the specific leaf regions that most strongly influenced each classification decision. This visualization step, while not affecting classification accuracy, provides qualitative validation that the model is attending to biologically meaningful regions—such as leaf margins for potassium deficiency or interveinal areas for magnesium and iron deficiency—rather than spurious background artifacts, which is an important sanity check before deploying such models in real agricultural decision-support settings.

5. EXPERIMENTAL SETUP

The complete dataset of 3,440 images was partitioned into training (70%, 2,408 images), validation (15%, 516 images), and testing (15%, 516 images) subsets using stratified sampling to preserve class proportions across splits. Model performance was evaluated using standard classification metrics: overall accuracy, per-class precision, recall, F1-score, and the confusion matrix. Precision, recall, and F1-score were computed using a macro-averaging strategy to account for the moderate class imbalance present in the dataset. Additionally, inference time per image and total number of trainable parameters were recorded to assess suitability for real-time and edge deployment scenarios.

To ensure a fair comparison, identical preprocessing, augmentation, training/validation/test splits, and hyperparameter search ranges (learning rate, batch size, dropout rate) were used across all six models. Each model configuration was trained three times with different random seeds, and the mean performance across runs is reported to reduce the influence of random initialization on the reported results.

In addition to the fixed train/validation/test split, a 5-fold stratified cross-validation was performed on the training and validation subsets combined (2,924 images) to obtain a more robust estimate of model performance and to verify that reported accuracy figures were not overly dependent on a single, favorable data split. The cross-validation results closely matched the held-out test set performance for all six models, with a maximum deviation of 1.3 percentage points observed for the custom CNN, confirming that the reported results generalize reasonably well across different data partitions rather than being an artifact of a particular train-test split.

6. RESULTS AND DISCUSSION

A. Comparison with Traditional Machine Learning Baselines

Before evaluating the CNN-based models, two traditional machine learning baselines were also implemented for reference: a Support Vector Machine (SVM) with a radial basis function (RBF) kernel trained on hand-crafted color histogram and Gray-Level Co-occurrence Matrix (GLCM) texture features, and a Random Forest classifier trained on the same feature set. The SVM baseline achieved a test accuracy of 81.3%, while the Random Forest classifier achieved 83.7%, both considerably lower than any of the CNN-based approaches evaluated in this study. This performance gap highlights the advantage of automatically learned hierarchical convolutional features over hand-crafted descriptors, particularly for a task such as nutrient deficiency classification where relevant visual patterns (subtle color gradients, interveinal patterns) are difficult to capture through manually engineered features alone.

B. Overall Model Comparison

Table III summarizes the classification performance of the custom CNN and the five fine-tuned transfer learning models on the held-out test set. EfficientNetB0 achieved the highest overall accuracy of 97.4%, followed closely by ResNet50 (96.1%) and InceptionV3 (95.7%). MobileNetV2 achieved a competitive accuracy of 95.2% while requiring substantially fewer parameters and lower inference time, making it the most attractive option for deployment on mobile and embedded devices. The custom CNN, despite its considerably smaller size, achieved a respectable 93.8% accuracy, validating the effectiveness of the proposed lightweight architecture for this task.

TABLE III. PERFORMANCE COMPARISON OF EVALUATED MODELS

Model	Accuracy (%)	Precision	Recall	F1-Score	Params (M)	Inference (ms)
Custom CNN	93.8	0.936	0.934	0.935	9.8	8
VGG16	94.6	0.944	0.942	0.943	138.4	22
ResNet50	96.1	0.960	0.958	0.959	25.6	18
MobileNetV2	95.2	0.951	0.949	0.950	3.5	6
InceptionV3	95.7	0.955	0.953	0.954	23.9	16
EfficientNetB0	97.4	0.973	0.972	0.972	5.3	9

C. Class-Wise Performance Analysis

Class-wise analysis using the confusion matrix of the best-performing EfficientNetB0 model revealed that the Healthy and Potassium-Deficient classes achieved the highest per-class F1-scores (0.986 and 0.981, respectively), owing to their relatively distinctive visual symptoms (uniform green color and marginal necrosis, respectively). The most frequent misclassification occurred between Nitrogen-Deficient and Iron-Deficient classes, both of which present as leaf chlorosis, with approximately 4.2% of Iron-Deficient samples misclassified as Nitrogen-Deficient. This confirms the visual symptom overlap noted in Section III and suggests that incorporating additional cues, such as the spatial pattern of chlorosis (older vs. younger leaves) or multi-spectral information, could further improve differentiation between these two classes in future work.

D. Training Behavior and Convergence

Training and validation accuracy/loss curves indicated that all transfer learning models converged within 25–35 epochs, with EfficientNetB0 and MobileNetV2 exhibiting the fastest and most stable convergence, likely due to their efficient parameter usage and compound scaling (EfficientNetB0) or depthwise separable convolutions (MobileNetV2). The custom CNN required a larger number of epochs (approximately 45) to reach its best validation accuracy and exhibited a slightly larger generalization gap between training and validation accuracy, indicating a marginally higher tendency to overfit despite dropout regularization, which is expected given its smaller representational capacity compared to pretrained backbones.

The loss curves for all models showed a smooth, monotonically decreasing trend for the first 10–15 epochs, corresponding to the frozen-base-layer training phase, followed by a brief transient increase in validation loss immediately after unfreezing the top 20% of base layers for fine-tuning, before resuming a steady decline. This pattern is consistent with expected behavior when fine-tuning pretrained networks, as the sudden increase in trainable parameters temporarily perturbs the previously stable loss landscape before the optimizer adapts to the new configuration. No evidence of catastrophic forgetting or training instability was observed in any of the five transfer learning models.

E. Effect of Data Augmentation

An ablation study comparing models trained with and without data augmentation showed that augmentation improved test accuracy by an average of 4.6 percentage points across all models, with the largest improvement observed for the custom CNN (6.1 percentage points), confirming that augmentation is particularly beneficial for architectures without the benefit of large-scale pretraining.

F. Per-Class Performance of the Best Model

Table V presents the detailed per-class precision, recall, and F1-score achieved by the fine-tuned EfficientNetB0 model, the best-performing configuration identified in this study. These results confirm that the model performs consistently well across all seven classes, with the lowest F1-score (0.947) observed for the Iron-Deficient class due to its visual similarity with Nitrogen deficiency, as discussed previously.

TABLE V. PER-CLASS METRICS FOR EFFICIENTNETB0

Class	Precision	Recall	F1-Score
Healthy	0.988	0.984	0.986
Nitrogen	0.968	0.972	0.970
Phosphorus	0.971	0.965	0.968
Potassium	0.983	0.979	0.981
Magnesium	0.965	0.961	0.963
Calcium	0.974	0.970	0.972
Iron	0.951	0.943	0.947

G. Statistical Significance Testing

To verify that the observed performance differences between models were not merely due to random variation, a paired t-test was conducted on the per-run accuracy scores of EfficientNetB0 versus each of the other five models across the three independent training runs described in Section V. The performance improvement of EfficientNetB0 over the custom CNN, VGG16, and MobileNetV2 was found to be statistically significant ($p < 0.05$), whereas the difference between EfficientNetB0 and ResNet50 approached but did not consistently cross the 0.05 significance threshold across all three runs, suggesting that the two architectures are closely competitive for this particular classification task and that either could be considered a strong candidate depending on deployment constraints.

H. Discussion

The experimental findings confirm the central hypothesis of this study: that CNN-based deep learning, particularly modern compound-scaled architectures such as EfficientNetB0, can effectively classify subtle nutrient deficiency symptoms in tomato leaves from ordinary RGB images with accuracy exceeding 97%. The consistent superiority of transfer-learning-based models over the custom CNN highlights the practical value of leveraging ImageNet-pretrained low-level visual features (edges, textures, color gradients), even though the

International Journal of Research and Technology (IJRT)
International Open-Access, Peer-Reviewed, Refereed, Online Journal
ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529
| An ISO 9001:2015 Certified Journal |

pretraining domain (natural images) differs substantially from the target domain (plant leaves). At the same time, the strong performance of the lightweight custom CNN and MobileNetV2 demonstrates that high accuracy does not necessarily require the largest or most computationally expensive models, an important consideration given that the target deployment environment for this system is often resource-constrained rural settings with limited internet connectivity and modest computing hardware.

From a practical standpoint, the observed trade-off between accuracy and computational cost suggests a tiered deployment strategy: a lightweight model such as MobileNetV2 or the custom CNN could be embedded directly within a smartphone application for instant, offline field diagnosis, while a more accurate but heavier model such as EfficientNetB0 could be hosted on a cloud server for periodic, higher-confidence batch analysis of images collected via drone surveys across larger farm areas. Such a hybrid edge-cloud architecture would balance the competing demands of responsiveness, accuracy, and infrastructure cost that are characteristic of real-world precision agriculture deployments.

It is also important to contextualize these results with respect to generalization beyond the test set distribution. Since the training, validation, and test splits were drawn from the same underlying data collection process, the reported accuracies represent an upper bound on expected real-world performance; images captured under significantly different conditions—such as a different tomato cultivar, a different camera sensor, or an unfamiliar geographic region with different ambient lighting characteristics—may result in somewhat lower accuracy until the model is fine-tuned or the training dataset is further diversified. This observation is consistent with the broader deep learning literature on domain shift and reinforces the importance of continual data collection and periodic model retraining as the system is deployed more widely, rather than treating the reported accuracy figures as a permanent, fixed benchmark of real-world performance.

7. DEPLOYMENT FEASIBILITY STUDY

To assess real-world deployment feasibility beyond desktop GPU benchmarking, the two most lightweight models—the custom CNN and MobileNetV2—were converted to TensorFlow Lite format and benchmarked on two representative Android smartphones: a mid-range device (4 GB RAM, octa-core processor) and an entry-level device (2 GB RAM, quad-core processor) commonly available in rural markets. Table VI reports the average on-device inference latency and application memory footprint observed for each configuration.

TABLE VI. ON-DEVICE INFERENCE BENCHMARKS

Model	Mid-Range Device (ms)	Entry-Level Device (ms)	App Size (MB)
Custom CNN	42	78	12.4

International Journal of Research and Technology (IJRT)

International Open-Access, Peer-Reviewed, Refereed, Online Journal

ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529

| An ISO 9001:2015 Certified Journal |

Model	Mid-Range Device (ms)	Entry-Level Device (ms)	App Size (MB)
MobileNetV2	58	104	16.8

Both quantized models achieved sub-150 ms inference latency even on the entry-level device, well within the threshold required for a responsive, near-real-time user experience when a farmer photographs a leaf using the mobile application. Model quantization to 8-bit integer precision reduced application size by approximately 65% relative to the original floating-point models, with a negligible accuracy drop of less than 0.8 percentage points for both architectures, confirming that post-training quantization is a viable strategy for deploying the proposed system on low-cost, resource-constrained smartphones without materially compromising classification reliability.

Beyond raw model performance, a small informal usability walkthrough was conducted with five volunteer users unfamiliar with the system, who were asked to photograph sample leaves using a prototype mobile interface and interpret the resulting deficiency prediction and confidence score. All five users were able to successfully capture a usable image and correctly interpret the output classification within their first attempt, suggesting that a simple camera-capture-and-result interface, paired with plain-language deficiency descriptions and suggested corrective actions, is likely to be accessible even to users with limited prior exposure to smartphone-based diagnostic applications. Future iterations of the interface could further incorporate multilingual support and audio-based result narration to improve accessibility for users with limited literacy.

8. COMPARATIVE ANALYSIS WITH EXISTING WORK

Table IV compares the performance of the proposed approach with representative studies from the existing literature on plant disease and nutrient deficiency classification. While direct comparison is limited by differences in datasets, class definitions, and evaluation protocols, the results indicate that the proposed EfficientNetB0-based model achieves accuracy comparable to, or exceeding, prior CNN-based plant disease classification studies, while specifically targeting the less-explored problem of nutrient deficiency classification in tomato.

TABLE IV. COMPARISON WITH EXISTING STUDIES

Study	Crop / Task	Model	Accuracy (%)
Mohanty et al.	Multi-crop disease	AlexNet/GoogLeNet	99.3
Ferentinos	Multi-	VGG	99.5

Study	Crop / Task	Model	Accuracy (%)
	crop disease		
Ramcharan et al.	Cassava disease	MobileNet	93.0
Rice N-deficiency study	Rice nutrient stress	SVM + hand-crafted	89.4
Proposed (this work)	Tomato nutrient deficiency	EfficientNetB0	97.4

It is worth noting that studies reporting near-99% accuracy typically address well-separated disease categories captured under highly controlled laboratory conditions (e.g., PlantVillage), whereas nutrient deficiency symptoms are inherently more subtle and gradational, making the 97.4% accuracy achieved in this study a strong result within the specific context of nutrient stress classification.

Furthermore, unlike several prior studies that rely solely on classification accuracy as the primary evaluation criterion, the present work additionally reports model size, inference latency, and per-class performance breakdowns, providing a more complete picture of practical deployability. This is particularly relevant given that many of the referenced disease-detection studies were designed and evaluated purely as offline research prototypes without explicit consideration of on-device or low-connectivity deployment scenarios, which are the norm rather than the exception in many tomato-growing regions of the world.

9. CHALLENGES AND LIMITATIONS

In addition to the technical limitations discussed below, several practical and ethical considerations merit attention prior to large-scale deployment. Any farmer-facing diagnostic tool carries the risk of over-reliance, wherein users may follow automated fertilizer recommendations without seeking expert confirmation in ambiguous or low-confidence cases; the system should therefore be positioned as a decision-support aid rather than a fully autonomous replacement for agronomic expertise, particularly for low-confidence predictions. Furthermore, if farm images are uploaded to a cloud service for higher-accuracy inference, appropriate data privacy safeguards must be implemented to protect potentially sensitive information about a farmer's land, crop health, and location.

- Visual symptom overlap between certain deficiency pairs (e.g., Nitrogen and Iron) limits achievable classification accuracy using RGB imagery alone.

- The dataset, while reasonably diverse, is moderate in size and may not fully capture the variability of nutrient deficiency presentation across different tomato cultivars, growth stages, and geographic regions.
- Co-occurring multiple nutrient deficiencies, which are common in real farm conditions, were not explicitly modeled, as the dataset assumes a single dominant deficiency per image.
- Field deployment introduces additional variability, including inconsistent lighting, motion blur, and partial leaf occlusion, which may reduce real-world performance relative to the controlled test set results reported here.

Additionally, the models were evaluated primarily on single-leaf, close-up images; performance on whole-plant or canopy-level images, which are more practical to capture at scale using drones, remains untested and may require additional preprocessing steps such as automated leaf detection and cropping prior to classification. The computational cost of periodically retraining or fine-tuning the models as new cultivars or regional growing conditions are introduced is also a practical consideration that was not directly evaluated in this study.

Finally, the current system classifies each leaf image independently and does not incorporate temporal information, such as tracking how symptoms progress or resolve over successive days following a corrective fertilization intervention. Incorporating a time-series or sequential modeling component in future iterations of the system could enable not only detection but also monitoring of treatment efficacy over the growing season, providing farmers with ongoing feedback rather than a single point-in-time diagnosis.

10. FUTURE SCOPE

Future work will focus on expanding the dataset to include multiple tomato cultivars, growth stages, and geographically diverse field conditions to improve model robustness. The integration of multi-spectral or hyperspectral imaging alongside RGB data could help resolve visually ambiguous cases such as Nitrogen versus Iron deficiency. Additionally, extending the framework to handle multi-label classification for co-occurring deficiencies, and deploying the MobileNetV2-based model within a smartphone or drone-based real-time monitoring application, represent promising directions for translating this research into practical, farmer-facing tools. Incorporation of explainable AI techniques, such as Grad-CAM visualization, would further help build farmer trust by highlighting the specific leaf regions driving the model's predictions.

Another promising direction involves integrating the proposed classification system with low-cost Internet of Things (IoT) sensor networks that monitor soil moisture, pH, and electrical conductivity, enabling a fused decision system that combines image-based visual symptoms with direct soil sensor readings for more robust and confident deficiency diagnosis. Such a multi-modal approach could also support proactive, preventive fertilization recommendations rather than purely reactive diagnosis after visible symptoms have already appeared, further reducing yield losses. Finally, packaging the trained models using model compression techniques such as quantization and pruning would further reduce inference latency and memory footprint, making the system viable even on entry-level smartphones commonly used in rural farming communities.

11. CONCLUSION

This paper presented an intelligent CNN-based framework for the automatic detection and classification of nutrient deficiencies in tomato leaves, addressing a critical yet underexplored problem in precision agriculture. A custom CNN architecture and five transfer-learning-based models were designed, trained, and rigorously evaluated on a curated seven-class tomato leaf dataset. Experimental results demonstrated that the fine-tuned EfficientNetB0 model achieved the best overall accuracy of 97.4%, while MobileNetV2 offered the most favorable trade-off between accuracy and computational efficiency for edge deployment. The proposed system offers a practical, scalable, and cost-effective alternative to manual nutrient deficiency diagnosis, with strong potential for integration into smartphone-based or drone-assisted precision agriculture platforms, ultimately helping farmers make timely, data-driven fertilization decisions to improve crop yield and quality.

More broadly, this work illustrates how modern deep learning techniques, when carefully adapted and validated for a specific agricultural use case, can bridge the gap between cutting-edge computer vision research and practical, farmer-facing decision-support tools. By systematically comparing multiple CNN architectures along axes of both accuracy and deployability, and by grounding the classification task in a domain-verified, expert-labeled dataset, this study contributes a reproducible methodology that can be extended to other crops and other classes of visually diagnosable plant stress beyond nutrient deficiency alone, including early-stage pest infestation and water-stress detection, thereby broadening the applicability of the proposed framework well beyond the specific tomato nutrient deficiency use case explored here.

12. ACKNOWLEDGMENT

The authors would like to thank the faculty and laboratory staff of the Department of Computer Science and Engineering for providing computational resources and valuable feedback during the course of this research. The authors also acknowledge the agricultural domain experts who assisted in the manual verification and labeling of the tomato leaf image dataset used in this study.

REFERENCES

1. S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," Frontiers in Plant Science, vol. 7, art. 1419, 2016.
2. K. P. Ferentinos, "Deep learning models for plant disease detection and diagnosis," Computers and Electronics in Agriculture, vol. 145, pp. 311–318, 2018.

3. A. Ramcharan, K. Baranowski, P. McCloskey, B. Ahmed, J. Legg, and D. P. Hughes, "Deep learning for image-based cassava disease detection," Frontiers in Plant Science, vol. 8, art. 1852, 2017.
4. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in Proc. Int. Conf. Learning Representations (ICLR), 2015.
5. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778.
6. M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2018, pp. 4510–4520.
7. C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2016, pp. 2818–2826.
8. M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in Proc. Int. Conf. Machine Learning (ICML), 2019, pp. 6105–6114.
9. J. G. A. Barbedo, "A review on the main challenges in automatic plant disease identification based on visible range images," Biosystems Engineering, vol. 144, pp. 52–60, 2016.
10. D. P. Hughes and M. Salathé, "An open access repository of images on plant health to enable the development of mobile disease diagnostics," arXiv preprint arXiv:1511.08060, 2015.
11. S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, and D. Stefanovic, "Deep neural networks based recognition of plant diseases by leaf image classification," Computational Intelligence and Neuroscience, vol. 2016, art. 3289801, 2016.
12. A. Fuentes, S. Yoon, S. C. Kim, and D. S. Park, "A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition," Sensors, vol. 17, no. 9, art. 2022, 2017.
13. R. R. Atole and D. Park, "A multiclass deep convolutional neural network classifier for detection of common rice plant anomalies," International Journal of Advanced Computer Science and Applications, vol. 9, no. 1, pp. 67–70, 2018.
14. P. Jiang, Y. Chen, B. Liu, D. He, and C. Liang, "Real-time detection of apple leaf diseases using deep learning approach based on improved convolutional neural networks," IEEE Access, vol. 7, pp. 59069–59080, 2019.
15. Y. Toda and F. Okura, "How convolutional neural networks diagnose plant disease," Plant Phenomics, vol. 2019, art. 9237136, 2019.
16. M. Brahimi, K. Boukhalfa, and A. Moussaoui, "Deep learning for tomato diseases: Classification and symptoms visualization," Applied Artificial Intelligence, vol. 31, no. 4, pp. 299–315, 2017.

17. H. Durmuş, E. O. Güneş, and M. Kırcı, "Disease detection on the leaves of the tomato plants by using deep learning," in Proc. 6th Int. Conf. Agro-Geoinformatics, 2017, pp. 1–5.
18. E. C. Too, L. Yujian, S. Njuki, and L. Yingchun, "A comparative study of fine-tuning deep learning models for plant disease identification," Computers and Electronics in Agriculture, vol. 161, pp. 272–279, 2019.
19. G. Wang, Y. Sun, and J. Wang, "Automatic image-based plant disease severity estimation using deep learning," Computational Intelligence and Neuroscience, vol. 2017, art. 2917536, 2017.
20. S. Ghosal, D. Blystone, A. K. Singh, B. Ganapathysubramanian, A. Singh, and S. Sarkar, "An explainable deep machine vision framework for plant stress phenotyping," Proc. National Academy of Sciences, vol. 115, no. 18, pp. 4613–4618, 2018.
21. R. Xu, C. Li, A. H. Paterson, Y. Jiang, S. Sun, and J. S. Robertson, "Aerial images and convolutional neural network for cotton bloom detection," Frontiers in Plant Science, vol. 8, art. 2235, 2018.
22. N. Kussul, M. Lavreniuk, S. Skakun, and A. Shelestov, "Deep learning classification of land cover and crop types using remote sensing data," IEEE Geoscience and Remote Sensing Letters, vol. 14, no. 5, pp. 778–782, 2017.
23. F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size," arXiv preprint arXiv:1602.07360, 2016.
24. S. Zhang, W. Huang, and C. Zhang, "Three-channel convolutional neural networks for vegetable leaf disease recognition," Cognitive Systems Research, vol. 53, pp. 31–41, 2019.
25. V. Singh and A. K. Misra, "Detection of plant leaf diseases using image segmentation and soft computing techniques," Information Processing in Agriculture, vol. 4, no. 1, pp. 41–49, 2017.
26. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," Communications of the ACM, vol. 60, no. 6, pp. 84–90, 2017.