

AI-Augmented Soil Fertility Index Generation Using Temporal and Visual Deep Learning Pipelines

P. Logaiyan*, R. Suriya**

*(Department of Master of Computer Applications, Sri Manakula Vinayagar Engineering College, Pondicherry, India
Email: logaiyan.mca@smvec.ac.in)

** (Department of Master of Computer Applications, Sri Manakula Vinayagar Engineering College, Pondicherry, India
Email: suriyaradha1314@gmail.com)

Abstract:

Soil fertility is a key determinant of sustainable agriculture, influencing nutrient absorption, crop performance, and long-term soil health. Yet, many farmers struggle to apply fertilizers correctly due to limited access to scientific soil assessment and reliance on guesswork. Conventional soil testing requires laboratory infrastructure, multiple sampling cycles, and delayed reporting, making it impractical for timely decision-making. As a result, excessive chemical usage remains common, contributing to nutrient imbalance, declining organic matter, and ecological damage. To address these issues, this paper presents an AI-Assisted Soil Condition Intelligence Framework that integrates multimodal data sources to evaluate soil health and nutrient adequacy. The system processes historical Soil Health Card (SHC) records along with current environmental parameters such as temperature, humidity, and rainfall using a sequence-learning model based on Long Short-Term Memory (LSTM) networks. In parallel, visual cues from crop leaves are analyzed using an image-efficient deep vision architecture derived from EfficientNet-B7 to detect early-stage nitrogen variation through Leaf Color Index (LCI) estimation. The predictions from the temporal soil model and the image-based nitrogen classifier are fused through a dynamic weighting mechanism that generates a unified Soil Vitality Index (SVI). This index provides farmers with real-time, crop-specific nutrient recommendations. By merging data-driven soil analytics with lightweight plant-visual assessment, the proposed system offers rapid, accessible, and field-ready guidance, enabling farmers to reduce fertilizer misuse, preserve soil ecosystems, and improve agricultural productivity.

Keywords — Soil Fertility, LSTM, EfficientNet-B7, Leaf Color Index, Soil Vitality Index, Precision Agriculture, Deep Learning, Fertilizer Recommendation.

I. INTRODUCTION

Modern agriculture faces a growing challenge: how to assess and manage soil fertility efficiently, accurately, and in a timely manner. Traditional laboratory soil testing remains the gold standard for nutrient analysis, but it demands costly infrastructure, skilled personnel, and multiple days or weeks before results are delivered. For small-scale and rural farmers — who form the backbone of food production in developing economies — this process is largely inaccessible. The result is a pervasive reliance on guesswork and generalized fertilizer guidelines that fail to account for the

spatial and temporal variability inherent in real agricultural land.

This paper proposes an AI-Augmented Soil Fertility Index system that bridges this gap by integrating two complementary deep learning pipelines. The first pipeline employs Long Short-Term Memory (LSTM) networks to analyze historical Soil Health Card (SHC) data combined with real-time environmental parameters — temperature, humidity, and rainfall — to predict current soil nutrient conditions. The second pipeline uses EfficientNet-B7 to classify nitrogen levels in crop leaves by analyzing Leaf Color Index (LCI) images. The outputs of both pipelines are fused through a dynamic weighting mechanism to

compute a unified Soil Vitality Index (SVI), which drives crop-specific fertilizer recommendations delivered through a web-based dashboard.

The primary contributions of this work are: (i) a multimodal AI framework combining temporal soil analytics with visual leaf-based nitrogen diagnosis; (ii) a dynamic fusion mechanism adapting prediction weights based on crop type, season, and data reliability; and (iii) a deployable web application enabling real-time, farmer-centric decision support without laboratory access.

II. LITERATURE SURVEY

Khan et al. [1] demonstrated that proximal soil sensing combined with machine learning can effectively delineate soil management zones, confirming that spatial variability in soil nutrient content is far greater than commonly assumed, making generalized fertilizer recommendations systematically ineffective.

Molin and Tavares [2] surveyed sensor-based soil mapping approaches in Brazilian tropical soils and identified key limitations of conventional methods: high cost, limited spatial resolution, and inability to integrate with real-time environmental conditions — limitations that directly motivate the real-time, AI-driven approach proposed in this paper.

Peng et al. [3] demonstrated that visible and near-infrared reflectance spectroscopy can predict soil nutrient content with high accuracy, establishing the scientific basis for non-invasive, image-based soil assessment and informing the Leaf Color Index component of the proposed system.

Yin et al. [4] reviewed soil sensor and plant wearable technologies for precision agriculture, identifying deep learning-based computer vision as the most promising direction for scalable, field-deployable crop health monitoring — supporting the choice of EfficientNet-B7 as the visual backbone of the proposed system.

Hartemink [5] reviewed long-term soil fertility trends and concluded that neither static soil testing nor single-season models are adequate to capture nutrient depletion dynamics — a finding that motivates the LSTM-based temporal modeling approach, which explicitly captures long-range dependencies and seasonal patterns in historical SHC data.

Maity et al. [6] proposed a multi-sensor fusion system for precision agriculture and demonstrated that fusion substantially outperforms single-sensor approaches in accuracy and reliability. Their architecture directly informs the dynamic weighted fusion mechanism used to combine LSTM and EfficientNet outputs into the unified Soil Vitality Index.

Smith [7] and Vasques [8] further highlighted the ecological urgency of reducing fertilizer misuse, reinforcing the need for intelligent, data-driven nutrient management systems that minimize chemical overuse while maximizing crop productivity.

III. THEORETICAL FRAMEWORK

The proposed system is grounded in three theoretical foundations. First, soil science establishes that fertility is a dynamic, time-varying property influenced by cumulative nutrient inputs and withdrawals, seasonal weather patterns, microbial activity, and crop residue decomposition [5]. No single static measurement can characterize the trajectory of soil health; only longitudinal data analyzed with methods capable of capturing temporal dependencies can provide reliable fertility forecasts. This principle justifies the LSTM-based temporal modeling pipeline.

Second, plant physiology confirms that leaf color is a reliable early indicator of nitrogen status. Chlorophyll content, which determines leaf color, is tightly coupled to nitrogen availability; nitrogen-deficient plants exhibit characteristic yellowing that precedes yield loss by several weeks [3]. The Leaf Color Index exploits this relationship to provide a non-invasive, real-time proxy for soil nitrogen adequacy.

Third, the ensemble learning literature demonstrates that combining predictions from models trained on different data modalities consistently outperforms any single-modality model [6]. The dynamic fusion mechanism applies crop- and season-specific weights to the LSTM and EfficientNet outputs before computing the final Soil Vitality Index, enabling the system to adapt its reliance on each source based on expected reliability in context.

IV. EXISTING SYSTEM

Agricultural soil management today relies primarily on three approaches, each with significant limitations that the proposed system directly addresses.

Laboratory soil testing provides the most accurate characterization of soil nutrient content, but its practical value is severely constrained by cost and latency. Sending soil samples to a laboratory and waiting days or weeks for results is incompatible with the rapid pace of crop development. By the time results are available, the optimal window for corrective action may have already passed.

Visual crop inspection — assessing soil health through observations of plant growth, leaf color, and plant vigor — is widely practiced but lacks scientific rigor. Without a calibrated reference, consistent and accurate nitrogen deficiency detection is difficult, particularly in early growth stages before symptoms become visually obvious.

Existing machine learning approaches including Random Forest, SVM, CNN, and KNN classifiers address some limitations but introduce new ones. These models typically rely on tabular soil data alone, ignoring real-time environmental conditions and visual crop indicators. They struggle to capture seasonal patterns and long-term nutrient depletion trends.

TABLE I

Comparison of Existing System vs. Proposed System

Feature	Existing System	Proposed System
Data Source	Lab tests only	SHC + env. data + leaf images
Analysis	Static single-point	LSTM + EfficientNet fusion
N Detection	Lab chemical analysis	Leaf Color Index (LCI)
Prediction	None; retrospective	Real-time SVI forecasting
Recommendation	Generic guidelines	Crop-specific dosage
Accessibility	Requires lab	Web/mobile, field-ready
Env. Factors	Not integrated	Temp, humidity, rainfall

V. PROPOSED SYSTEM

The proposed AI-Augmented Soil Fertility Index system replaces fragmented, static soil assessment with an integrated, real-time intelligence framework. Every time a farmer inputs soil health data, current environmental readings, or a new crop leaf image, the system immediately updates its soil fertility assessment and delivers actionable recommendations.

The system is organized around three functional layers: the Temporal Analysis Layer using an LSTM network trained on historical SHC records; the Visual Diagnosis Layer using EfficientNet-B7 to classify leaf images into nitrogen status categories; and the Fusion and Recommendation Layer combining outputs through a dynamic weighting mechanism to produce the SVI and crop-specific fertilizer guidance.

A. LSTM-Based Soil Condition Prediction

The LSTM network ingests sequential soil health and weather data — nitrogen, phosphorus, and potassium levels, pH, soil texture, temperature, humidity, and rainfall — and produces a forward-looking soil nutrient adequacy score. The LSTM architecture is particularly well-suited because soil fertility is a fundamentally temporal process; current nutrient levels are the product of accumulated history, seasonal cycles, and progressive depletion patterns that no static model can capture.

B. EfficientNet-Based Leaf Nitrogen Diagnosis

The visual pipeline uses EfficientNet-B7 — selected for its exceptional balance of accuracy and computational efficiency — to analyze uploaded crop leaf images. The model extracts color, texture, and spectral features that serve as proxies for chlorophyll content and nitrogen status. Leaf images are classified into one of five Leaf Color Index categories corresponding to nitrogen levels from severely deficient to more than sufficient.

The five LCC classification levels are: (1) Light Yellow-Green — Severely Deficient, apply full nitrogen dose immediately; (2) Yellowish-Green — Deficient, apply high nitrogen dose; (3) Light Green — Moderately Deficient, apply moderate dose; (4) Green — Adequate/Sufficient, no immediate action; and (5) Dark Green — More than Sufficient, monitor only.

C. Dynamic Fusion and Soil Vitality Index

The Soil Vitality Index is computed by combining the LSTM soil score and the EfficientNet nitrogen classification through a dynamic weighted fusion algorithm. Weights are adjusted based on crop type, growing season, and data source reliability. The resulting SVI (0–100%)

drives three output categories: fertilizer type and dosage recommendations, soil improvement strategies, and crop selection guidance.

VI. VI. SYSTEM ARCHITECTURE

The architecture follows a layered, service-oriented design separating data collection, model inference, and user presentation into independently maintainable components. The system is implemented in Python 3.8 using Flask as the backend framework, MySQL for relational data storage, Bootstrap 4 for the responsive frontend, TensorFlow/Keras for deep learning models, and OpenCV/Pillow for image preprocessing.

The System Admin uploads datasets and trains models on the cloud-based server. Environmental sensors feed temperature, humidity, moisture, and NPK values wirelessly. The prediction pipeline performs eight sequential operations: data exploration, preprocessing, feature selection, feature extraction, classification, model training, prediction, and recommendation generation.

VII. VII. DATA FLOW DIAGRAM

The Data Flow Diagrams (DFD) illustrate how data moves through the system across three levels of abstraction — from the high-level overview to the detailed operational flows between the Admin, System, and Farmer entities.

D. A. DFD Level 0 (Context Diagram)

At Level 0, the system is represented as a single process — the Soil Fertility Predictor — receiving login requests from the Admin and register/login requests from the Farmer. The system responds with login acknowledgments and accesses the `tb_admin` and `tb_farmer` database tables.

E. B. DFD Level 1

Level 1 expands the central process into its major functional components: dataset upload and model training by the Admin, and data input and prediction retrieval by the Farmer. The system receives soil health data, weather data, and leaf color data from the Farmer, processes them through the trained SoilFerNet and LLCNet models, and returns soil fertility predictions and fertilizer recommendations.

VIII. VIII. UML DIAGRAMS

F. A. Use Case Diagram

The use case diagram presents the functional scope of the Soil Fertility Predictor Web App from the perspectives of the two primary actors — Admin and Farmer. The Admin use cases include: Register/Login, Upload Dataset, Train SoilFerNet Model, Train LLCNet Model, and System Maintenance. The Farmer use cases include: Login, Input Soil Health Data, Input Leaf Color Data, Preprocessing, Feature Extraction, Soil Fertility Prediction, Crop Recommendation, Pesticides Recommendation, and Fertilizer Recommendation.

G. B. Activity Diagram

The activity diagram illustrates the parallel workflows of the Admin, Soil Fertility Predictor, and Farmer. The System workflow performs Preprocessing, Feature Extraction, Soil Fertility Prediction, and Recommendation generation, storing results in the database. The Farmer workflow concludes with receiving Fertilizer Recommendations.

IX. IX. SYSTEM MODULES

The system is organized around six functional modules, each with a clearly defined responsibility in the overall soil fertility assessment pipeline.

H. A. Soil Fertility Predictor Dashboard

The main interface for farmer and administrator interaction, built with Flask, MySQL, Bootstrap, and WampServer. Farmers input soil health card data, weather patterns, and leaf images through intuitive forms. Administrators manage users, upload datasets, train updated models, and monitor system performance.

I. B. SoilFerNet Model (LSTM)

The LSTM-based soil prediction module processes historical SHC data and environmental time series through five stages: dataset import, preprocessing, feature extraction (NPK levels, pH, texture, moisture, climate variables), LSTM classification across soil fertility categories, and model deployment. The trained model achieves 98.93% classification accuracy.

J. C. LLCNet Model (EfficientNet-B7)

The leaf color classification module uses EfficientNet-B7 to analyze Leaf Color Chart images. The pipeline includes noise removal and color correction, resizing to 224×224 pixels, extraction of hue, saturation, brightness, and texture features, five-class nitrogen status classification, and deployment with associated intervention recommendations.

K. D. Soil Fertility Detector and Recommendation Engine

The fusion module accepts soil health card data, weather patterns, and leaf images simultaneously, applies the SoilFerNet and LLCNet models in parallel, and combines their outputs through the dynamic fusion algorithm to produce the Soil Vitality Index. The recommendation engine generates fertilizer type and dosage suggestions, soil improvement strategies, and crop recommendations.

X. X. PERFORMANCE EVALUATION

System performance was evaluated on a held-out test set of 500 soil profiles spanning sandy, clayey, loamy, red, and black soil types across five crop categories. The LSTM model was trained on 2,000 historical SHC records using 5-fold cross-validation. The EfficientNet-B7 model was fine-tuned on 3,500 labeled Leaf Color Chart images using transfer learning from ImageNet weights.

Table II presents a comparative performance analysis of the proposed models against existing baseline algorithms. The proposed LSTM-EfficientNet fusion model consistently outperforms all baselines, achieving 99.1% accuracy, 98.9% precision, 98.6% recall, and 98.7% F1 score.

TABLE II
Algorithm Performance Comparison

Algorithm	Accuracy	Precision	Recall	F1 Score
Random Forest	91.2%	90.8%	89.5%	90.1%
SVM	87.4%	86.9%	85.3%	86.1%
CNN (baseline)	89.7%	88.4%	87.9%	88.1%
KNN	83.5%	82.1%	81.6%	81.8%
LSTM (proposed)	98.9%	98.5%	98.1%	98.3%
EfficientNet-B7	98.7%	98.2%	97.8%	98.0%
Fused SVI	99.1%	98.9%	98.6%	98.7%

XI. XI. RESULTS AND DISCUSSION

The system was deployed as a web application and tested with 50 farmer users across varied soil conditions. Farmers input NPK values, environmental parameters, and select a Leaf Color Chart category. The system returns the Fertility

Score, its interpretation, and specific fertilizer and crop recommendations in real time.

For a low-fertility profile (N=102, P=79, K=45, Temperature=27.66°C, Humidity=79.68, pH=6.4, Rainfall=100.66 mm, LCC=Light Yellow-Green), the system returned a Fertility Score of 23.69% (Low Fertility) with a recommendation to use balanced fertilizers like NPK 10-26-26 at 40 kg/acre. For a moderate-fertility profile, the Fertility Score was 45.01% with NPK 15-15-15 or compost at 39 kg/acre, recommending maize, sunflower, or cotton.

The most significant performance gap appears on the minority low-fertility and severely nitrogen-deficient classes. The LSTM model's recall on the low-fertility class is 97.8%, compared to 81.2% for Random Forest and 74.6% for SVM. This improvement is directly attributable to the temporal modeling capability of LSTM, which captures early warning signals in nutrient depletion trends that static classifiers cannot detect.

In user evaluation, 88% of farmer participants rated the system as more useful than their existing soil assessment methods. 82% reported that they would follow the system's fertilizer dosage recommendations without additional verification. Average task completion time was 3.2 minutes, compared to 7–14 days for a full laboratory soil test cycle.

XII. XII. BENEFITS

The primary benefit of the system is the conversion of soil fertility assessment from a laboratory-bound, weeks-delayed process into a real-time, field-accessible service. By identifying low-fertility conditions and nitrogen deficiencies before they manifest as yield loss, the system creates an intervention window that does not exist in conventional practice.

Secondary benefits include: (i) reduced fertilizer misuse through precise dosage recommendations; (ii) reduced ecological impact from over-application of fertilizers; (iii) improved crop selection by matching crops to measured soil conditions; (iv) enhanced data-driven decision-making; and (v) reduced dependence on costly laboratory infrastructure.

XIII. CONCLUSION

This paper has presented an AI-Augmented Soil Fertility Index system that integrates LSTM-based temporal soil analysis with EfficientNet-B7-based leaf nitrogen diagnosis into a unified, deployable platform for real-time soil fertility assessment and fertilizer recommendation. The system addresses the core limitations of conventional soil testing — cost, latency, and inaccessibility — and the core limitations of existing ML approaches — static modeling, single-modality input, and lack of real-time environmental integration.

The proposed fused model achieved 99.1% accuracy and 98.7% F1 score, outperforming all baseline algorithms across every evaluation metric. The dynamic fusion mechanism producing the Soil Vitality Index demonstrated improved accuracy over either constituent model alone. In user evaluation, 88% of farmer participants rated the system as more useful than existing methods, with an average assessment time of 3.2 minutes versus the 7–14 day laboratory alternative.

Future enhancements will integrate IoT-based soil sensors for continuous real-time data collection; develop a dedicated mobile application for field use; implement blockchain-based data security; and explore cloud-based deployment for scalability across regions.

XIV. ACKNOWLEDGMENT

The authors would like to thank the Department of Master of Computer Applications, Sri Manakula Vinayagar Engineering College, Pondicherry, India, for providing the resources and support necessary to conduct this research.

XV. REFERENCES

- [1] [1] H. Khan, A. A. Farooque, B. Acharya, F. Abbas, T. J. Esau, and Q. U. Zaman, "Delineation of management zones for site-specific information about soil fertility characteristics through proximal sensing of potato fields," *Agronomy*, vol. 10, no. 12, p. 1854, Nov. 2020.
- [2] [2] J. P. Molin and T. R. Tavares, "Sensor systems for mapping soil fertility attributes: Challenges, advances, and perspectives in Brazilian tropical soils," *Engenharia Agricola*, vol. 39, no. spe, pp. 126-147, Sep. 2019.
- [3] [3] Y. Peng, L. Zhao, Y. Hu, G. Wang, L. Wang, and Z. Liu, "Prediction of soil nutrient contents using visible and near-infrared reflectance spectroscopy," *ISPRS Int. J. Geo-Inf.*, vol. 8, no. 10, p. 437, Oct. 2019.
- [4] [4] H. Yin, Y. Cao, B. Marelli, X. Zeng, A. J. Mason, and C. Cao, "Soil sensors and plant wearables for smart and precision agriculture," *Adv. Mater.*, vol. 33, no. 20, May 2021, Art. no. 2007764.
- [5] [5] E. Hartemink, "Assessing soil fertility decline in the tropics using soil chemical data," *Adv. Agronomy*, vol. 89, pp. 179-225, Jan. 2006.
- [6] [6] T. Maity, A. N. Bhawani, J. Samanta, P. Saha, S. Majumdar, and G. Srivastava, "MLSFDD: Machine-learning-based smart fire detection device for precision agriculture," *IEEE Sensors J.*, vol. 25, no. 5, pp. 8921-8928, Mar. 2025.
- [7] [7] P. Smith, "Global change pressures on soils from land use and management," *Global Change Biol.*, vol. 22, no. 3, pp. 1008-1028, Mar. 2016.
- [8] [8] G. M. Vasques, "Field proximal soil sensor fusion for improving high-resolution soil property maps," *Soil Syst.*, vol. 4, no. 3, p. 52, Aug. 2020.
- [9] [9] H. D. Foth and B. G. Ellis, *Soil Fertility*. Boca Raton, FL, USA: CRC Press, 2018.
- [10] [10] F. Wang, "Technologies and perspectives for achieving carbon neutrality," *Innovation*, vol. 2, no. 4, Nov. 2021, Art. no. 100180.
- [11] [11] Q. Xie, Z. Wang, F. Wen, J. He, and T.-K. Truong, "Coarray tensor train decomposition for bistatic MIMO radar," *IEEE Trans. Antennas Propag.*, vol. 73, no. 8, pp. 5310-5323, Aug. 2025.
- [12] [12] J. J. Stoorvogel, M. Bakkenes, A. J. A. M. Temme, N. H. Batjes, and B. J. E. T. Brink, "S-World: A global soil map for environmental modelling," *Land Degrad. Develop.*, vol. 28, no. 1, pp. 22-33, Jan. 2017.

