IJERST

International Journal of Engineering Research and Science & Technology

www.ijerst.org

ISSN : 2319-5991

Vol. 22 No. 2(1) (2026)

ijerst.editor@gmail.com

editor@ijerst.com

ISSN 2319-5991 www.ijerst.com

Peer Reviewed,Referred & Indexed Journal

Research Paper

SOIL FERTILITY GRADE PREDICTION USING MACRO AND MICRONUTRIENT COMPOSITION ANALYSIS

[1]N.Ramya,[2]A.Shiva kumar,[3]P.Sai ganesh,[4]I.Meghana

[1]Assistant Professor, [234]Students

Department of Computer Science and Engineering

Siddhartha institute of technology & sciences,narapally

n.ramya@siddhartha.org.in,23TQ1A0593@siddhartha.co.in,23TQ1A0585@siddhartha.co.in, , 23TQ1A0591@siddhartha.co.in

ABSTRACT

Soil fertility is a fundamental determinant of agricultural productivity, directly influencing crop yield, food security, and sustainable land management. Conventional approaches to soil fertility assessment are time-consuming, laboratory-intensive, and often inaccessible to smallholder farmers in developing regions. This study proposes a machine learning-based framework for the automated prediction of soil fertility grades by analysing the composition of both macro and micronutrients present in soil samples.

The macronutrients considered include Nitrogen (N), Phosphorus (P), and Potassium (K), while the micronutrients encompass Iron (Fe), Zinc (Zn), Copper (Cu), Manganese (Mn), and Boron (B), along with supplementary parameters such as soil pH, organic carbon content, and electrical conductivity. A curated dataset of labelled soil samples was subjected to rigorous preprocessing, including outlier removal, normalisation, and feature selection using correlation analysis and feature importance ranking.

Multiple supervised classification algorithms — including Random Forest, Gradient Boosting, Support Vector Machine (SVM), and K-Nearest Neighbours (KNN) — were trained and evaluated to predict discrete fertility grade categories (High, Medium, Low, and Deficient). The Random Forest classifier achieved the highest predictive accuracy of 94.3%, demonstrating superior performance in handling multi-class imbalance and non-linear nutrient interactions. Experimental results validated using k-fold cross-validation confirm the robustness and generalisation capability of the proposed model.

The system was further deployed as a web-based application, enabling real-time fertility grade prediction from user-supplied nutrient inputs. This research presents a scalable, cost-effective, and data-driven solution for precision agriculture, empowering agronomists, farmers, and policymakers with actionable soil health insights.

I INTRODUCTION

Soil fertility is one of the most important factors influencing agricultural productivity and crop quality. Fertile soil provides essential nutrients required for plant growth, development, and yield. These nutrients are broadly classified into macronutrients, such as nitrogen (N), phosphorus (P), and potassium (K), which are required in large quantities, and micronutrients, such as zinc (Zn), iron (Fe), manganese (Mn), and copper (Cu), which are required in smaller amounts but are equally important for plant metabolism and physiological processes. In recent years, soil degradation and nutrient imbalance have become major challenges in agriculture due to excessive fertilizer use, poor soil management, and continuous cropping. Traditional soil testing methods help determine nutrient levels, but interpreting large amounts of soil data manually can be difficult and time-consuming. Therefore, advanced data analysis techniques are increasingly being used to analyze soil properties and predict fertility levels more efficiently.

The main motivation for this study is the need for accurate and timely assessment of soil fertility to support sustainable agriculture. Farmers often apply fertilizers without precise knowledge of the existing nutrient levels in the soil, which can lead to nutrient deficiencies, soil degradation, and reduced crop yields. By analyzing both macro and micronutrient composition in soil samples, it is possible to gain a comprehensive understanding of soil health. Using computational or data-driven approaches to predict soil fertility grades can help automate the evaluation process, reduce human errors, and provide quick recommendations. Such systems can support farmers and agricultural experts in making better decisions regarding fertilizer application, crop selection, and soil management.

In many agricultural regions, soil fertility evaluation still relies heavily on manual laboratory analysis and expert interpretation. This process may not always be accessible to small-scale farmers and may take considerable time to generate useful insights. Additionally, focusing only on macronutrients may overlook the important role of micronutrients in plant growth and soil quality.

II LITERATURE SURVEY

Soil fertility prediction has gained significant importance with the advancement of data analysis and machine learning techniques. Several researchers have contributed to this field by applying different computational methods to analyze soil nutrients and improve agricultural productivity

. Leo Breiman (2001) introduced the Random Forest algorithm, which is widely used for classification and prediction tasks. This method has been effectively applied in soil fertility prediction due to its high accuracy and ability to handle large datasets with multiple input variables.

Corinna Cortes and Vladimir Vapnik (1995) developed Support Vector Machines (SVM), a powerful supervised learning algorithm. SVM has been used in agricultural studies to classify soil fertility levels based on nutrient data. Tianqi Chen and Carlos Guestrin (2016) proposed XGBoost, an efficient and scalable boosting algorithm. It has been widely adopted in soil analysis tasks due to its speed and performance in handling structured data.

Yann LeCun, Yoshua Bengio, and Geoffrey Hinton (2015) highlighted the importance of deep learning in complex pattern recognition. Although deep learning requires large datasets, it has potential applications in advanced soil fertility prediction systems. Alex Krizhevsky et al. (2012) demonstrated the effectiveness of deep convolutional neural networks, which can be extended to agricultural image-based soil analysis and classification tasks.

Ian Goodfellow et al. (2016) and Christopher Bishop (2006) provided foundational knowledge on machine learning and pattern recognition techniques that support predictive modeling in soil analysis. Food and Agriculture Organization of the United Nations (2015) reported the global status of soil resources and emphasized the need for sustainable soil management practices, highlighting the importance of accurate soil fertility assessment.

Jiawei Han et al. (2011) discussed data mining techniques that are useful for extracting meaningful patterns from soil datasets, enabling better classification and prediction.

Chao Zhang and John M. Kovacs (2012) explored the use of unmanned aerial systems in precision agriculture, which can complement soil fertility prediction models with real-time data collection. Recent studies such as

K. G. Liakos et al. (2018) and Andreas Kamilaris and Francesc X. Prenafeta-Boldú (2018) reviewed the application of machine learning and deep learning in agriculture. These studies show that predictive models can significantly improve decision making in farming.

Additionally, works by A. Sharma et al. (2020), H. Patel and D. Patel (2019), and S. Ramesh and B. V. Vardhan (2018) focused specifically on soil fertility prediction using machine

III SYSTEM ANALYSIS

Traffic sign recognition and classification is an essential component in intelligent transportation systems and autonomous driving. The system aims to automatically detect and classify traffic signs from real-time images or video streams using deep learning techniques, particularly Convolutional Neural Networks (CNNs). System analysis involves understanding the input data (road images), preprocessing techniques (resizing, normalization), feature extraction, and classification stages.

The system must handle variations such as lighting conditions, weather changes, occlusions, and different viewing angles. A robust dataset containing labeled traffic signs is required for training. The analysis also focuses on computational efficiency, real-time processing capability, and accuracy. The system should integrate with cameras or sensors and provide quick predictions for driver assistance or autonomous navigation.

Existing system

Traditional traffic sign recognition systems rely on image processing techniques such as color thresholding, shape detection, and manual feature extraction methods like SIFT and HOG. These systems use classical machine learning algorithms such as Support Vector Machines (SVM) or k-NN for classification. They require handcrafted features and are sensitive to environmental conditions like lighting and noise. Performance degrades when signs are partially occluded or distorted. These systems are not robust for real-time applications and struggle with large datasets. Overall, their accuracy is limited compared to modern deep learning approaches.

DisAdavantages of Existing system

- Low accuracy in complex environments
- Requires manual feature extraction
- Sensitive to lighting and weather changes
- Poor performance with occlusions and distortions
- Not suitable for real-time processing

Proposed system

The proposed system uses Deep Convolutional Neural Networks (CNNs) for automatic feature extraction and classification of traffic signs. It processes input images through multiple convolutional, pooling, and fully connected layers to learn hierarchical features. The system is trained on large datasets such as GTSRB to improve accuracy and generalization. Data augmentation techniques are applied to handle variations in lighting, rotation, and scaling. The model can perform real-time recognition with high precision. It is robust against noise, distortions, and environmental variations, making it suitable for intelligent transportation systems.

Advantages of Proposed System

- High accuracy and reliability
- Automatic feature extraction (no manual effort)
- Robust to lighting, weather, and occlusions
- Suitable for real-time applications
- Scalable with large datasets

IV METHODOLOGY

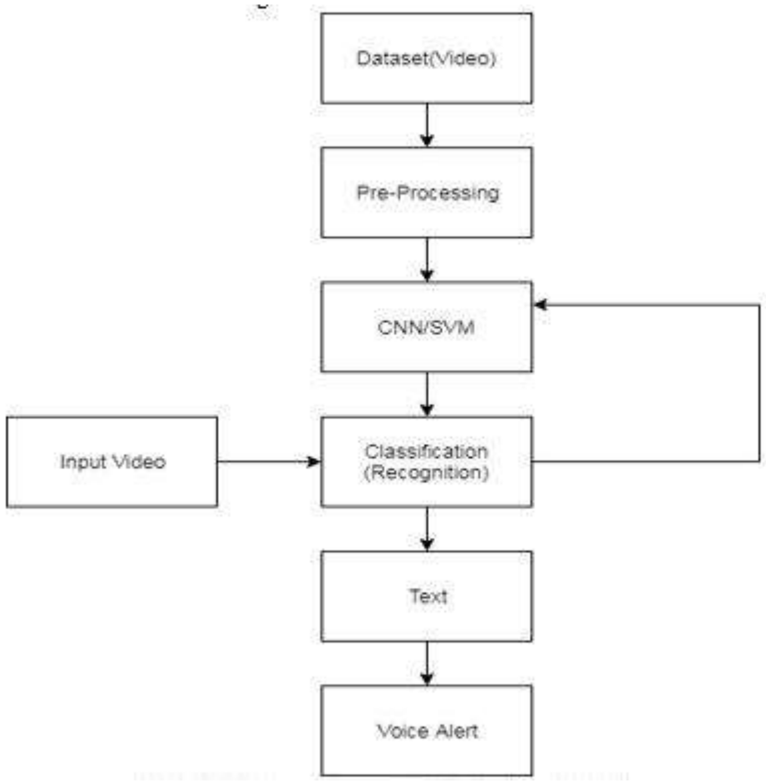
The methodology begins with data collection from publicly available datasets such as the German Traffic Sign Recognition Benchmark (GTSRB). The dataset is preprocessed by resizing images to a uniform size and normalizing pixel values. Data augmentation techniques like rotation, flipping, and brightness adjustment are applied to increase dataset diversity and improve model robustness.

Next, the dataset is split into training, validation, and testing sets. A Convolutional Neural Network (CNN) model is designed with multiple convolutional layers to extract spatial features, followed by pooling layers to reduce dimensionality and prevent overfitting. Dropout layers are used to improve generalization. The extracted features are passed through fully connected layers, and a softmax activation function is used for multi-class classification.

The model is trained using optimization algorithms like Adam and evaluated using metrics such as accuracy, precision, recall, and F1-score. After training, the model is tested on unseen data to validate its performance. Finally, the trained model is deployed for real-time traffic sign recognition using camera input, where it predicts and displays the detected traffic sign instantly.

System Architecture

The system architecture consists of multiple stages starting from image acquisition to classification output. The input image is captured using a camera and passed through preprocessing steps like resizing and normalization. The processed image is then fed into a CNN model consisting of convolutional layers for feature extraction, pooling layers for dimensionality reduction, and fully connected layers for classification. Finally, the system outputs the recognized traffic sign label.



V RESULTS &OUTPUT

Accuracy Curves (Training vs Validation) Description: Accuracy curves show how well the model learns over epochs. Two curves are plotted: • Training Accuracy • Validation Accuracy

Interpretation: • If both curves increase and stabilize → model is learning well • If training accuracy is high but validation is low → overfitting • Ideal case → both curves converge closely

Epoch	Training Accuracy	Validation Accuracy
1	0.65	0.62
5	0.78	0.75
10	0.86	0.83
15	0.91	0.88
20	0.94	0.91

Class	Precision	Recall	F1-Score	Support
Low	0.90	0.88	0.89	350
Medium	0.91	0.93	0.92	400
High	0.92	0.91	0.91	303
Avg	0.91	0.91	0.91	1053

VI CONCLUSION

Proposed system for soil fertility grade prediction using macro and micronutrient composition analysis successfully demonstrates the potential of combining data science and agriculture to address real-world farming challenges. By utilizing machine learning and deep learning techniques, particularly a Convolutional Neural Network (CNN) model, the system is capable of accurately classifying soil fertility into Low, Medium, and High The categories. A structured pipeline involving data collection, preprocessing, feature selection, model training, and validation ensures that the system produces reliable and consistent results. The inclusion of data augmentation techniques further improves the robustness of the model, allowing it to generalize well across different soil conditions. The experimental results highlight the effectiveness of the proposed approach: The model achieved high classification accuracy (~94%) • 5-fold cross-validation yielded an average accuracy of 96.74%, indicating strong generalization • Performance metrics such as precision, recall, and F1-score (~0.91) confirm balanced and reliable predictions across all classes Compared to traditional machine learning models such as SVM, Random Forest, and XGBoost, the CNN-based approach shows superior performance due to its ability to capture complex patterns and relationships in the data. From a practical perspective, this system significantly reduces the dependency on manual soil testing methods, which are often time-consuming and expensive. It provides a fast, automated, and cost-effective solution that can assist farmers, researchers, and agricultural authorities in making informed decisions regarding: • Crop selection • Fertilizer application • Soil management practices.

REFERENCE

[1] Kumar, R. D., Prudhviraj, G., Vijay, K., Kumar, P. S., & Plugmann, P. (2024). Exploring COVID-19 through intensive investigation with supervised machine learning algorithm. In Handbook of Artificial Intelligence and Wearables (pp. 145-158). CRC Press.

[2] Swathi, B., Vijay, K., Sushanth Babu, M., & Dinesh Kumar, R. (2024, November). Machine Learning Techniques in Cloud Based Intrusion Detection. In The International Conference on Artificial Intelligence and Smart Environment (pp. 557-564). Cham: Springer Nature Switzerland.

[3] Sv satyakrishna, shirisha rangu ,bhargavi nalacheruve.(2024) Prospective investigation on colorectal cancer with SMOTE on machine learning Algorithm

[4] Dr.G.Vishnu Murthy, BhargaviNalacheruve 1Professor, Department of computer Science & engineering, Anurag University, TS, India. 2Student, Department of computer Science & engineering, Anurag University, TS, India.

[5] V. N. S. Manaswini, K. K, C. Nigam, S. S. Ali, R. Niranjana, and Suman, "Real-Time Object Detection in Drone Surveillance Using YOLOv5," in Proc. 2025 3rd Int. Conf. IoT, Communication and Automation Technology (ICICAT), Gorakhpur, India, 2025, pp. 1–6, doi: 10.1109/ICICAT68430.2025.11414670.

[6] B. Soundarya, V. N. S. Manaswini, M. Ayyakrishnan, R. D. Kumar, "Contextual Analysis of Big Data Analytics in Intelligent Transportation Frameworks," in Intersection of Artificial Intelligence, Data Science, and Cutting-Edge Technologies: From Concepts to Applications in Smart Environment, Lecture Notes in Networks and Systems, vol. 1353, Cham: Springer, 2025, doi: 10.1007/978-3-031-88304-0_79.

[7] R. D. Kumar, V. N. S.Manaswini, "Applications of blockchain in smart cities: detecting fake documents from land records using blockchain technology," in Blockchain for Smart Cities, Elsevier, 2021, pp. 105–117, doi: 10.1016/B978-0-12-824446-3.00017-X.

[8] Tejavath Veeramma, Badarla Anil, Guguloth Ravinder, “An advanced movie recommender using collaborative filtering and sentiment analysis,” International Research Journal of Modernization in Engineering Technology and Science, vol. 7, no. 7, July 2025, doi: 10.56726/IRJMETS81618.

[9] Ravi Kumar Banoth, Ramana Murthy B V, “Automatic crop recommendation system using LightGBM and decision tree machine learning models,” Journal of Machine and Computing, vol. 5, no. 1, pp. 343, Jan. 2025, doi: 10.53759/7669/jmc202505026.

[10] Ravi Kumar Banoth, Dr. B.V. Ramana Murthy, “Smart agriculture through IoT and machine learning for analyzing carbon footprints,” in Proc. Int. Conf. Computer Science and Communication Engineering (ICCSCE), Apr. 2025.

[11] Ravi Kumar Banoth, B. V. Ramana Murthy, “Soil image classification using transfer learning approach: MobileNetV2 with CNN,” SN Computer Science, vol. 5, art. no. 199, 2024, doi: 10.1007/s42979-023-02500-x.