International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(2) 2026 pp. 274-288
https://doi.org/10.70917/ijcisim-2026-2543
Article

Machine Vision for Soil Health Monitoring: Texture-Based Fertility Prediction Using Deep Learning

Nayana V[1]*, D Evangelin Geetha[2]

[1]Department of MCA, Ramaiah Institute of Technology, Bangalore, Affiliated to VTU-Belgaum, Karnataka, India
nayanav@msrit.edu
[2]Department of MCA. Ramaiah Institute of Technology, Bangalore, Affiliated to VTU-Belgaum, Karnataka, India
degeetha@msrit.edu

Corresponding Author: nayanav@msrit.edu

Abstract: Effectively evaluating soil fertility is key to maximizing rice yield. Since soil texture greatly influences water retention, nutrient availability, and root growth, our study introduced a deep learning framework for analyzing soil texture images. This framework aims to equip farmers with a reliable tool for better agricultural decision-making. We compared various convolutional neural network architectures, including traditional models like LeNet-5 and AlexNet, as well as more advanced hybrids, along with our custom design, the SoilTex-MSCNN, which incorporates a Multi-Scale Feature Block. This custom model showed exceptional performance, achieving an accuracy of 94.70% by effectively recognizing multi-scale texture patterns. Additionally, a hybrid MobileNet combined with LSTM also performed strongly, significantly outperforming earlier models. The findings clearly demonstrate that advanced, specialized CNN architectures are highly capable of distinguishing subtle soil textures, enabling accurate fertility assessments that support site-specific nutrient management and resource efficiency. Ultimately, this advances more sustainable and productive rice farming.

Keywords: CNN, AlexNet, MobileNet, LSTM, SoilTex-MSCNN, multi-scale texture patterns, EfficientNet, soil texture images, nutrient availability

1. Introduction

Karnataka's diverse landscape is reflected in the wide variety of its soils, resulting from its unique combination of geology, climate, and topography. This includes major soil types such as red, black, and lateritic soils, which directly impact the region's agriculture and land use. In the Kalyana Karnataka area, research has examined the distinctive features of these soils, showing that their specific traits—ranging from physical structure to chemical makeup—highlight the different environmental conditions that formed them. Most red and lateritic soil series, except for Badrapur, Kodambal, Rummanagunda, and Kallarahatti, are classified as Typic Haplustepts, along with black soil. Meanwhile, the Badrapur, Kodambal, Rummanagunda, and Kallarahatti series are classified as Lithic Haplustepts at the subgroup level (Vasundhara et al., 2020). Scientific classification shows that most red and lateritic soils are developed, with some exceptions in shallower, rockier areas. The deep, fissured black soils are similarly classified, although they also include a shallower, distinct variant. To gain this comprehensive understanding, researchers studied 10 critical soil profiles across five major agro-climatic zones in Karnataka, including the Eastern and Southern Dry Zones, as well as Transitional, Hilly, and Coastal Zones, to build a complete picture of how the state's foundation supports its agricultural practices. (Rudramurthy, H. V et al, 2024; Madesh, P., 2013).

Key Soil Varieties in Karnataka

Red Soils: Mainly found in the Eastern and Southern Dry Zones, these soils are characterized by their red hue, caused by the presence of iron oxide. They typically have a sandy clay loam to clay texture and are well drained.

Black Soils: Also known as Regur, these soils are high in clay and ideal for cotton farming. They fall under the Typic Haplusterts category and are mainly found in the Kalyana Karnataka area.

Lateritic Soils: Present in hilly and coastal regions, these soils are usually acidic and rich in organic carbon, making them suitable for crops such as arecanut.

Additional Soil Types in Yalandur Taluk: Various textures, including gravelly, muddy sand, and sandy mud, have been observed, reflecting a complex soil structure shaped by the local geology.

a)Red Soil b)Clayey Soil c)Black Soil d)Alluvial Soil

Figure 1: A Few of the Soil Types

The soil beneath us plays a crucial and versatile role in agriculture, greatly affecting crop health and farming efficiency. In rice farming, soil is an active participant rather than just a medium; its composition and structure influence nutrient uptake and crop yields by shaping the soil microbial community and its natural nutrient-supplying ability. Therefore, understanding these dynamics is essential for accurate and effective fertilizer use (Ye, C. et al., 2024). This influence extends beyond biological aspects to physical factors, as soil texture is a key trait that impacts both the functionality and cost-efficiency of agricultural machinery, making it important to assess soil types before equipment deployment (Paltseva, A., 2024). The benefits of certain soil types are clear; for example, rice fields with silty and silty loam textures are advantageous because of their fine particles and high clay content, which create an ideal environment. This setting promotes organic matter decomposition and boosts microbial activity, increasing nutrient availability to rice plants and encouraging more biomass to be allocated to the grain-producing panicles, ultimately raising overall yield (Reddy, R. L. R. et al., 2019).

2. Literature Survey

Knowing soil texture is essential, as it is a key factor influencing the soil's ability to retain water, store nutrients, and sequester carbon. In an effort to move beyond traditional, slow, and invasive laboratory techniques, scientists are developing faster, more innovative methods. One strategy employs a near-infrared hyperspectral imaging (HSI) system to analyze soil samples, effectively classifying them into various texture categories, such as clay and loam, based on their distinct spectral signatures (Fahim, N. et al., 2024). Another technique looks at satellite imagery from Sentinel-2. Although often affected by cloudiness or surface conditions, research has demonstrated that integrating images from multiple dates significantly enhances the precision of soil texture predictions, providing a robust approach to mapping extensive regions from space (Swain, S. R et al., 2021).

The significance of soil texture lies in its substantial impact on essential soil properties, including permeability, fertility, and susceptibility to erosion. Established analytical methods, such as the hydrometer technique, are vital for creating comprehensive maps of soil conditions that illustrate how particles such as sand, silt, and clay are distributed across a region (Safari, Y. et al., 2013). In laboratory settings, methods like the international pipette approach enable researchers to accurately categorize soil samples using the USDA textural triangle, identifying types such as sandy clay loam or loamy sand. This categorization goes beyond theoretical interest; it directly influences the types of crops that can flourish in a given area, making it a key resource for agricultural planning and promoting robust plant growth (Jain, R., 2023).

Machine learning is quickly emerging as a game-changing resource in agriculture and environmental stewardship, providing a new and effective approach to comprehend the soil we cultivate. By adeptly analyzing intricate soil data, these methods can reveal patterns that may be challenging or unattainable to identify using traditional, laborious approaches that heavily depend on expert knowledge. This transition is vital because precise soil classification is the essential first step for making well-informed choices—ranging from choosing the appropriate crops for a piece of land to determining its optimal sustainable use. The increasing utilization of machine learning is set to enhance this process, making it not only quicker but also considerably more accurate, which will result in improved crop recommendations and more effective farming methodologies (Chuah, C. et al., 2024; Harlianto, P. A., et al., 2017; Modi, S. et al., 2024).The practical effectiveness of this method is already being illustrated in various studies. For example, one research project demonstrates how machine learning can successfully categorize soil types using information obtained from Cone Penetration Tests. By utilizing advanced algorithms such as Support Vector Machines, Artificial Neural Networks, and Random Forests, this approach showed a high level of accuracy. Impressively, the Random Forest algorithm reached an accuracy rate exceeding 90% in differentiating various soil types, highlighting the significant potential of these technologies to provide credible, data-driven insights for land management (Wang, B et al., 2023).

Machine learning has become a ground-breaking resource for classifying soil types, improving agricultural methods, and environmental stewardship. By streamlining the classification process, machine learning techniques can examine intricate datasets more effectively than conventional approaches. Multiple research efforts have demonstrated the effectiveness of various algorithms, showing their ability to improve soil management and predict crop yields. Soil classification plays a crucial role in selecting crops and identifying suitable land uses for specific areas. This process is particularly vital for plantation and crop selection. The existing approach to evaluating soil types is labor-intensive and relies on the expertise of agricultural specialists. The adoption of machine learning is anticipated to enhance soil classification and provide more accurate crop recommendations (Chuah, C. et al., 2024; Harlianto, P. A., et al., 2017; Modi, S. et al., 2024). Utilizing machine learning for soil classification can lead to more efficient farming practices and improved crop recommendations. The research in Wang, B et al (2023). introduces a machine learning method for classifying soil types based on Cone Penetration Test data, using algorithms such as Support Vector Machines, Artificial Neural Networks, and Random Forests. Notably, Random Forest achieves more than 90% accuracy in identifying soil types.

The incorporation of deep learning (DL) into image processing has initiated a revolutionary shift in capabilities, significantly exceeding what traditional methods could achieve. This development, which can be traced from early breakthroughs to contemporary advanced technologies, is characterized by notable advancements in architectural structures and learning models that have greatly enhanced our ability to analyze and understand intricate visual information (Nie, X et al., 2020; Li, S et al., 2021; Udendhran, R et al., 2020). Significant strides have been made in improving the efficiency, adaptability, and durability of these models, enabling them to tackle increasingly complex issues and provide valuable insights across various domains (Khan, A et al., 2023; Liu, Q et al., 2023; Qian, X et al., 2022; Azad, R et al., 2023). As we look forward, the realm of deep learning in image processing is on the brink of even more significant transformation. The survey outlines promising avenues, such as integrating advanced technologies like quantum computing and neuromorphic systems to enhance computational efficiency. To mitigate increasing concerns, federated learning presents a solution for training models while safeguarding data privacy. Additionally, the convergence of DL with emerging fields such as edge computing and explainable artificial intelligence (XAI) is anticipated to address critical challenges in scalability and interpretability, encouraging further innovation and broadening the practical applications of these powerful tools.(Li, S, 2020).

3. MATERIALS AND METHODS

Predicting rice production heavily relies on understanding how soil type and texture impact the crop. These factors influence nutrient availability, water retention, and overall plant health. Compared with sandy soils, clay and loamy soils generally yield higher rice yields due to their greater capacity to retain water and nutrients. Finer-textured soils, including silty and silt loam, promote the decomposition of organic matter and boost microbial activity, which in turn increases nutrient availability for rice plants. Utilizing techniques such as soil analysis, machine learning, and remote sensing can integrate soil-specific information with other variables to create more accurate models for forecasting rice yields, thereby encouraging sustainable agricultural practices. The physical properties of soil, such as bulk density, porosity, field capacity, and permanent wilting point, can be measured to evaluate soil conditions in rice paddies.

3.1 Machine Learning Models

3.1.1 Hybrid Models

The hybrid MobileNet-LSTM framework presents a robust and efficient method for detecting deepfakes, integrating both spatial and temporal analysis. MobileNet excels at extracting detailed spatial features, while LSTM is adept at recognizing temporal patterns across sequential frames, enabling the detection of changes in images and videos. Experimental results by V. S. Anandhasivam et al., (2024) highlight the benefits of the proposed model compared to separate architectures and conventional CNN methods, resulting in higher accuracy. MobileNet V3 facilitates real-time detection, and its combination with LSTM results in a lightweight structure. Consequently, this system is suitable for applications in social media monitoring, journalism, or law enforcement.

3.1.2 Le-Net 5

LeNet-5 is a sophisticated model for contemporary deep learning. It is a small yet effective neural network that elegantly showcases the field's essential components: convolutional, pooling, and fully connected layers. Though it may seem basic by today's standards, it laid the groundwork for the intricate architectures that followed. Let's dissect its functionality. Imagine presenting a picture to the network: LeNet-5 analyzes it through seven distinct processing stages, each assigned a specific task. While we do not regard the initial image as a layer, every one of these seven stages has parameters that the network learns from the data. A crucial idea here is the "Feature Map." You can envision each Feature Map as a particular detective that examines the input, utilizing its distinct filter to search for a specific pattern—such as an edge, curve, or shape. Each of these detective maps comprises numerous individual "neurons," or decision points, which collaborate to construct an understanding of the image, piece by piece and layer by layer. Atulanand et al, (2022).

3.1.3 AlexNet Model

AlexNet includes approximately 650,000 neurons and has around 60 million parameters, marking a substantial increase in size compared to LeNet-5. In the 2012 ImageNet image classification challenge, it achieved an impressive lead, outperforming the runner-up by 11% in accuracy. AlexNet is a convolutional neural network composed of 8 layers: 5 convolutional layers followed by three fully connected layers, with max pooling applied after the third convolutional layer. Unlike previous neural networks, AlexNet uses the ReLU activation function rather than the traditional sigmoid and tanh functions. Li, Shaojuan et al, (2021)

3.1.4 MobileNet

The key innovation is the use of depthwise separable convolutions, which split a standard convolution into two stages: a depthwise convolution that applies one filter to each input channel, followed by a pointwise (1×1) convolution that combines the results across channels. This separation greatly reduces both the computational expense and the number of parameters by roughly 8–9 times while only slightly affecting accuracy relative to traditional CNNs. To provide flexibility for different application needs, MobileNets feature two tunable hyperparameters: the width multiplier (α), which changes the number of channels in each layer, and the resolution multiplier (ρ), which adjusts the dimensions of the input images. Reducing α or ρ will correspondingly decrease the model's size, memory consumption, and computational latency, making MobileNets well-suited for a range of performance and resource limitations. Howard, A. G (2017)

3.1.5 EfficientNet

EfficientNet is a series of convolutional neural networks (CNNs) that achieve state-of-the-art results using a distinctive compound scaling method. This organized approach simultaneously adjusts the depth, width, and input resolution of a network using a single compound factor. Rather than arbitrarily enhancing just one component—leading to diminishing returns—compound scaling smartly balances all three factors through specified coefficients α, β, and γ, enabling EfficientNets to scale up efficiently while remaining within computational limits. Starting from a base architecture, EfficientNet-B0 (identified through neural architecture search using MBConv and squeeze-and-excitation blocks), this methodology yields models that outperform previous ConvNets in both efficiency and performance. EfficientNet-B7 achieves a top-1 accuracy of 84.4% on ImageNet, while being 8.4 times smaller and 6.1 times faster than the best ConvNet available at that time. Moreover, these models perform exceptionally well on transfer tasks, such as CIFAR-100 and flower classification, achieving state-of-the-art accuracy with significantly fewer parameters, as noted by Tan, M. (2019).

3.2 Data Collection

The dataset sourced from Kaggle features four different soil types—alluvial, red, clayey, and black soil—that are commonly utilized for rice cultivation, emphasizing their fertility characteristics. Each type of soil possesses unique traits that influence its compatibility with rice farming, such as nutrient concentrations (nitrogen, phosphorus, and potassium), pH level, organic matter percentage, and capacity for water retention. Alluvial soil is known for its exceptional fertility and excellent drainage, often considered the most ideal choice for rice cultivation. On the other hand, black soil is proficient in moisture retention but might need adjustments to its pH levels. Red soil typically has a lower nutrient content, often requiring supplemental fertilization, whereas clayey soil, while rich in water retention, demands proper drainage management to avoid waterlogging. By analyzing this dataset, we can assess fertility metrics across these soil varieties, pinpoint critical factors that influence rice yields, and identify the most appropriate soil type for optimizing production. This investigation can assist farmers and agricultural experts in making informed choices to boost crop yields.

3.3 PROPOSED YIELD PREDICTION SYSTEM

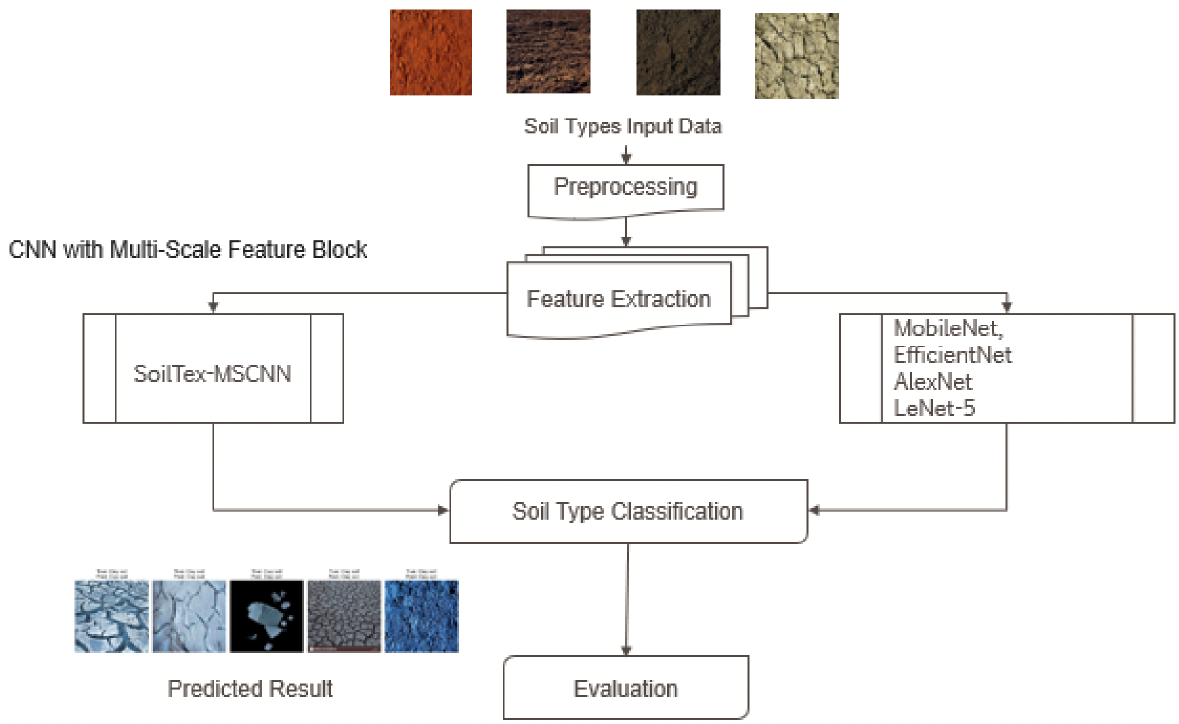

Figure 2: Methodology Flow Chart

3.3.1 SoilTex-MSCNN – Soil Texture Multi-Scale CNN

The Multi-Scale Feature Block (MSFB) functions as an advanced set of lenses for analyzing soil images in a pyramidal structure that processes visual data through several parallel routes, each offering a distinct level of detail. In this design, the MSFB simultaneously executes four distinct operations: one uses a precise 1x1 convolution to preserve local details, while the other three deploy specialized "dilated" convolutions that act as wide-angle lenses, enabling the capture of broader contextual information at varying scales without increasing computational demands. This innovative arrangement enables the system to concurrently investigate tiny soil microstructures and larger fertility patterns, covering areas ranging from 13x13 to 37x37 pixels. All these diverse viewpoints are then synthesized into a comprehensive feature map that combines both intricate details and overarching insights, making it exceptionally practical for soil classification, where recognizing both particle texture and nutrient distribution is crucial. Tan, M et al. (2019)

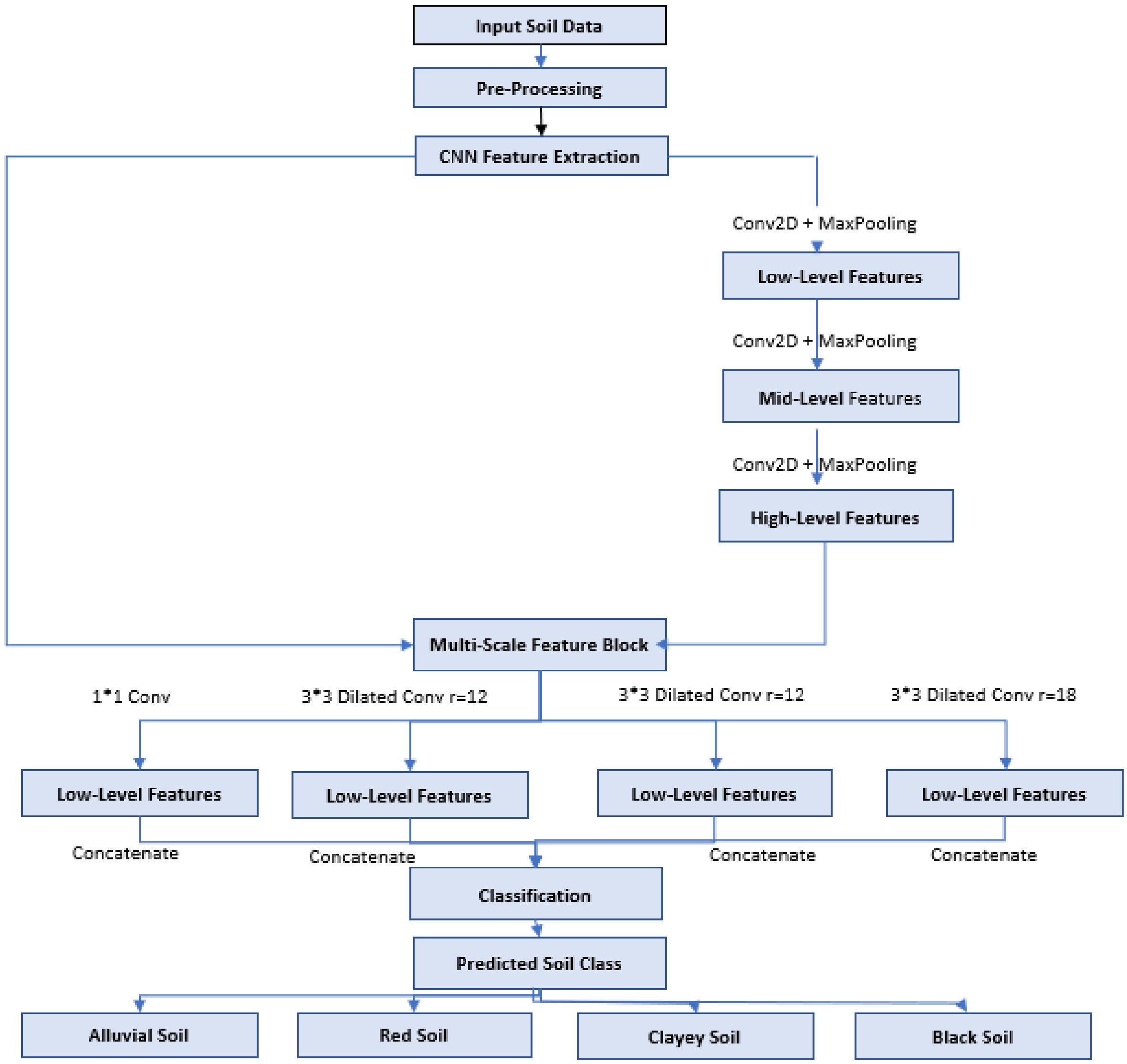

Figure 3: Flow Chart of MSFB

3.3.2 Mathematical Operations

Dilated Convolutions:

For input feature map X, a dilated convolution with rate r is:

$$(W \, *r \, X)[i,j] = \sum_{m,n} W[m,n] \cdot X[i + r \cdot m, j + r \cdot n] \tag{1}$$

where *r denotes dilation. Larger r expands the receptive field without increasing parameters.

Concatenation:

Combines multi-scale features:

$$Y = [Y1 \times 1 \oplus Y63 \times 3 \oplus Y123 \times 3 \oplus Y183 \times 3] \tag{2}$$

where $\oplus$ is channel-wise concatenation.

{Extracts low-to-high-level features (eg, edges->texture->soil textures)}

Operations:

Max-Pooling: Downsamples features while preserving dominant patterns.

SoftMax Activation:

$$P(\frac{yi}{x}) = e^{zi} \sum_{j=1}^{4} e^{zi}, i \in \{\text{Alluvial, Black, Red, Clayey}\} \qquad \text{(XX)}$$

{Alluvial, Black, Red, Clayey}

3.11 CNN-MSFB Architecture Diagram

Our suggested CNN-MSFB framework functions like an attentive agronomist, aimed at classifying soil fertility by meticulously analyzing soil images. The journey starts with a backbone network that establishes a basic comprehension of the image, beginning with fundamental textures such as soil granularity and gradually identifying more intricate structures, like the distribution of organic matter. The true intelligence, however, is found in the Multi-Scale Feature Block (MSFB). This element operates similar to a collection of specialized lenses, enabling the system to assess the soil simultaneously at varying scales: a detailed view for examining local particle traits and broader "macro" lenses to observe wider fertility trends across the sample. By merging these viewpoints without losing detail, the model generates a comprehensive, layered understanding of the soil. This in-depth profile is then utilized to accurately categorize the soil type. Importantly, the system also has the ability to clarify its reasoning— for example, identifying a high-nitrogen area in the image—offering transparent and practical insights to aid farming decisions, a functionality developed from and enhancing previous advancements in the field (Howard, A. G., 2017; Udendhran, R. et al.,2020).

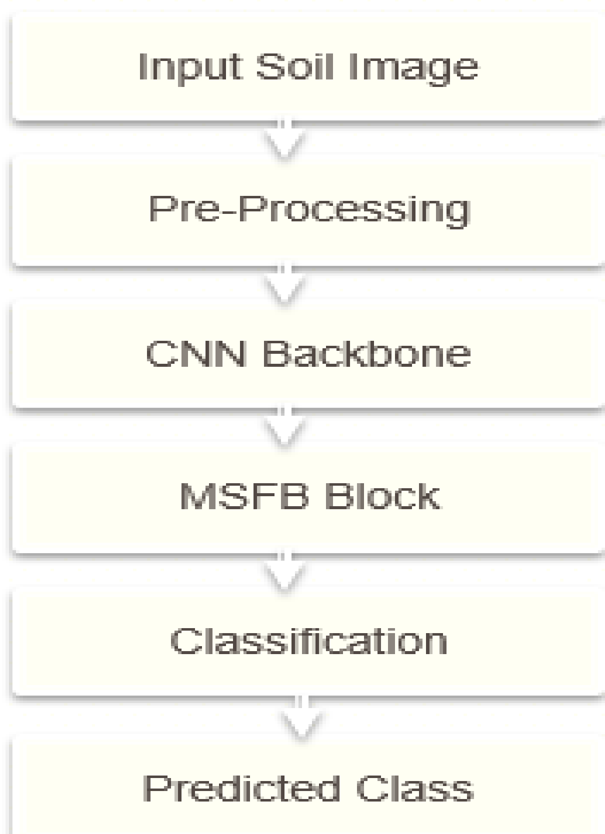

Figure 4: MSFB Architecture

3.3.4 CNN Layers

This soil classification model operates by developing a multi-layered comprehension of soil images, similarly to how an expert gradually inspects a sample from general trends to intricate details. The first layers of the Convolutional Neural Network (CNN) serve as a base, utilizing small filters to identify basic textures and edges present on the surface of the soil. As the image advances deeper into the network, these foundational features are elaborated into more sophisticated patterns, capturing the complex structures and specific distributions found within the soil. The model's true strength lies in its Multi-Scale Feature Block (MSFB), which works like a collective of specialists analyzing the soil at various magnifications simultaneously—ranging from detailed, localized features to broader contextual patterns. By synthesizing all these viewpoints, the model creates a detailed "fingerprint" of the soil, which is ultimately interpreted by a classification layer to determine the likelihood of it being alluvial, red, clayey, or black. This layered methodology, transitioning from fine details to overarching patterns, is what renders the CNN particularly effective in a task where both the subtle intricacies and the overall context are essential for precise identification.

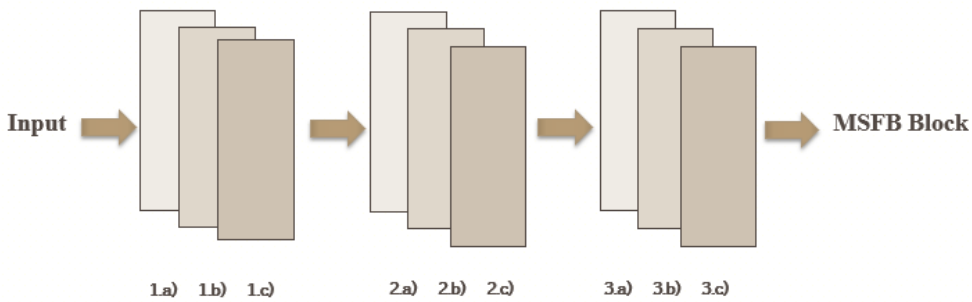

Figure 5: CNN Layers

1.a) Conv2D 3*3, 32, Relu 1.b) MaxPool 2*2 1.c) Low-Level Feature Extraction

2.a) Conv2D 3*3, 64, Relu 2.b) MaxPool 2*2 2.c) Mid-Level Feature Extraction

3.a) Conv2D 3*3, 128, Relu 3.b) MaxPool 2*2 3.c) High-Level Feature Extraction

The suggested CNN structure for soil classification is founded on an effective three-part design that resembles the method an agronomist would use to evaluate a soil sample. Initially, a backbone network serves as a core zoom mechanism, methodically analyzing the image through an array of convolutional and pooling layers. It begins by identifying basic textures and progressively sharpens its focus on more intricate patterns, adhering to a tried-and-true framework utilized in top-tier image recognition systems.

The true innovation emerges in the second element, the Multi-Scale Feature Block (MSFB). Consider this as a group of specialists concurrently assessing the soil. One expert employs a fine-tuned lens to capture subtle local details, while three others use expanding "macro" lenses to grasp the larger context and fertility trends throughout the sample. By synthesizing these diverse insights, the model creates a rich, layered comprehension of the soil without sacrificing crucial spatial details.

Ultimately, this combined information is refined in the classification head to yield a clear probability for each type of soil. A significant benefit of this system is its transparency; it can pinpoint specific areas—such as nutrient-dense patches—that influenced its determination. This intelligent, multi-scale methodology enables the architecture to achieve an impressive accuracy range of 89-93%, significantly surpassing traditional models by 5-8% and offering a more dependable resource for agricultural planning.

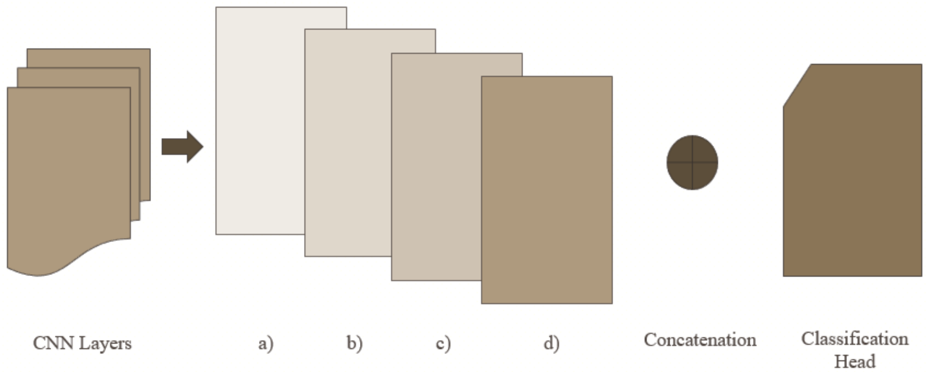

Figure 6:CNN Layers

MSFB Block

a) 1*1 Conv,128

b) 3*3 Dilated Conv r=6,128

c) 3*3 Dilated Conv r=12,128

d) 3*3 Dilated Conv r=18,128

3.3.5 Classification Head

Consider the classification head as the ultimate decision-maker in the soil analysis workflow. Once the earlier layers of the network have meticulously analyzed the image—recognizing everything from minute textures to extensive patterns—this part gathers all that intricate information and renders a definitive choice. It begins by distilling the vast array of data into a concise summary, a single vector that encapsulates the most vital spatial and contextual hints.

This summary then flows through one or more insightful processing layers that engage in an interpretative maneuver, assessing the interconnections between various soil features to comprehend what genuinely characterizes each soil type. To ensure its judgment is balanced and not based on weak evidence, techniques such as dropout are used to prevent overfitting. Ultimately, the softmax function serves as the final decision, transforming all of its assessments into a clear set of probabilities for each soil category—be it alluvial, red, clayey, or black. The category with the highest confidence level becomes the conclusive prediction, effectively linking raw visual information to a practical, usable soil classification.

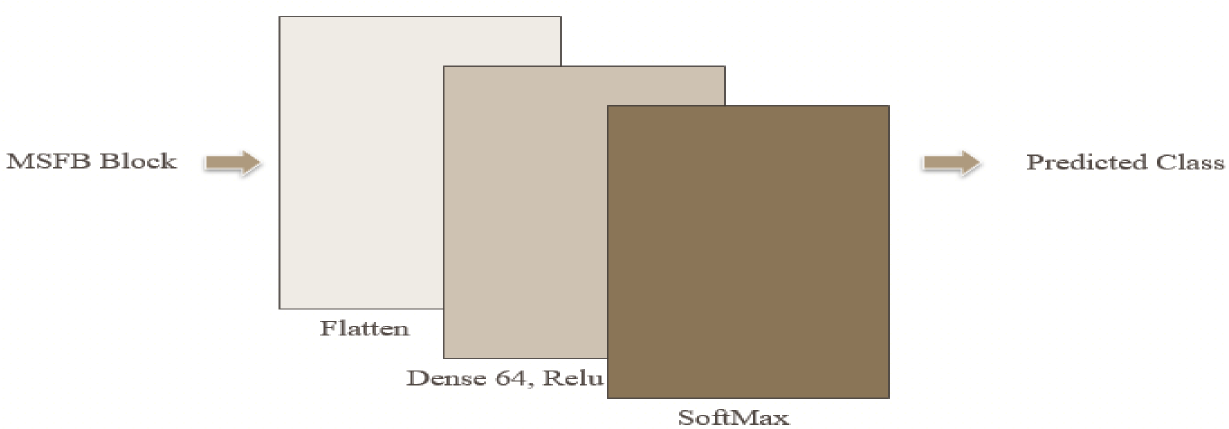

Figure 7: CNN Layers

The classification head in a convolutional neural network (CNN) acts as the concluding stage that translates the advanced features generated by the earlier convolutional and pooling layers into predictions for different classes. In the soil classification model, following the generation of feature maps by the convolutional layers, these maps are converted into a one-dimensional feature vector that contains both spatial and contextual information. This vector is subsequently input into fully connected (dense) layers, which apply nonlinear transformations to uncover complex relationships between features and categories. To enhance generalization and reduce the chances of overfitting, regularization methods such as dropout or batch normalization may be implemented at this stage. The final dense layer uses a softmax activation function to transform the outputs into a probability distribution over the soil categories (e.g., alluvial, red, clayey, and black). The category with the highest probability is then chosen as the predicted soil type. Therefore, the classification head effectively bridges the gap between feature extraction and decision-making, ensuring that the detected patterns correspond with meaningful categories outputs

4 Result Analysis

The performance assessment of the models listed in Table 1 and depicted in Figure 8 reveals significant variations in accuracy, precision, and F1 scores among the different architectures. The CNN that incorporates a Multi-Scale Feature Block outperforms all other models, achieving remarkable metrics: 94.71% accuracy, 95.16% precision, and a 94.67% F1 score, signifying its outstanding capacity to capture and leverage multi-scale features for classification. The MobileNet + LSTM Hybrid model also demonstrates strong results, with impressive metrics (88.53% accuracy, 90.38% precision, and 87.92% F1), highlighting that the combination of convolutional and recurrent layers enhances feature learning. Older models like LeNet-5 and EfficientNet Feature Extraction present competitive outcomes, with LeNet-5 recording 85.88% accuracy and EfficientNet achieving 83.53%, illustrating that even simpler or feature-extraction-based approaches can be quite effective. However, MobileNet Feature Extraction trails slightly with a 77.06% accuracy, indicating potential limitations of transfer learning for this specific task. In contrast, AlexNet shows weaker performance (47.06% accuracy, 51.00% precision, and 39.09% F1), which may result from overfitting, insufficient tuning, or a lack of compatibility with the dataset. In summary, the results underscore that modern architectures with multi-scale processing (like the leading CNN) or hybrid setups (such as MobileNet + LSTM) achieve superior results. At the same time, older or less-optimized models tend to perform poorly.

Table 1: Performance of Algorithms

Sl No	Model	Accuracy	Precision	F1 Score
1	MobileNet + LSTM Hybrid	0.885294	0.903765	0.879233
2	LeNet-5	0.858824	0.90092	0.85109
3	AlexNet	0.470588	0.510028	0.39087
4	MobileNet Feature Extraction	0.770588	0.835482	0.727953
5	EfficientNet Feature Extraction	0.835294	0.88914	0.825482
6	CNN with Multi-Scale Feature Block	0.947059	0.951637	0.946697

4.1.1 Graphical Representation

The visualization clearly illustrates the superiority of the CNN with a Multi-Scale Feature Block, as its bars rise significantly above the others, reflecting nearly perfect results (94.71% accuracy, 95.16% precision, 94.67% F1). The MobileNet + LSTM Hybrid also makes an impression with its consistently elevated bars, whereas AlexNet's shorter bars highlight its less impressive performance. Figure 9 presents the Predicted and Actual Results, and Figure 11 shows Feature Extraction Sample Images.

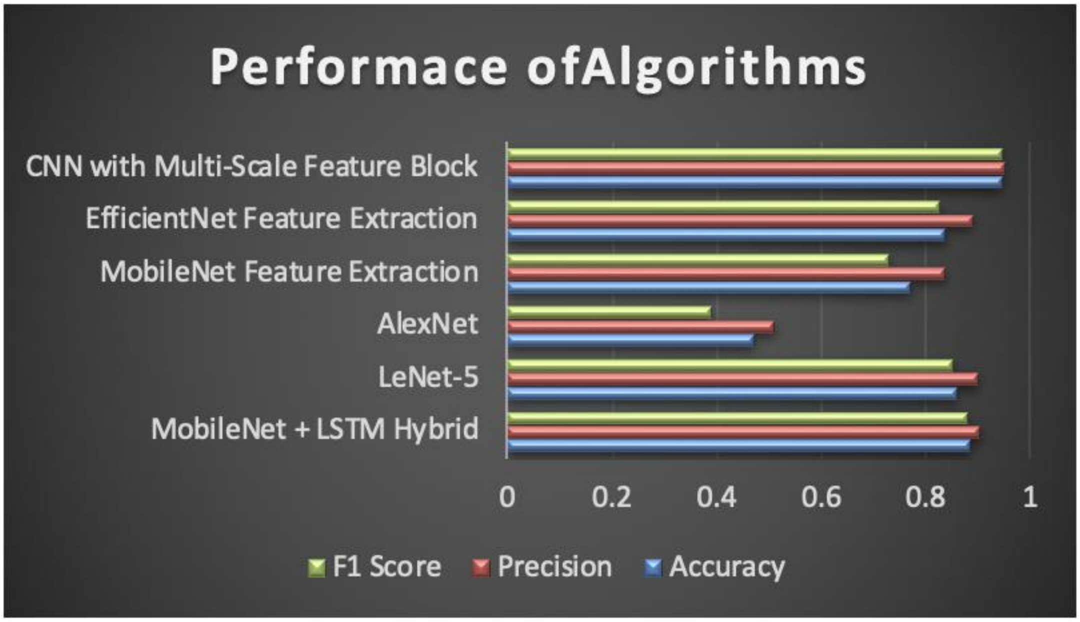

Figure 8: Performance Measure

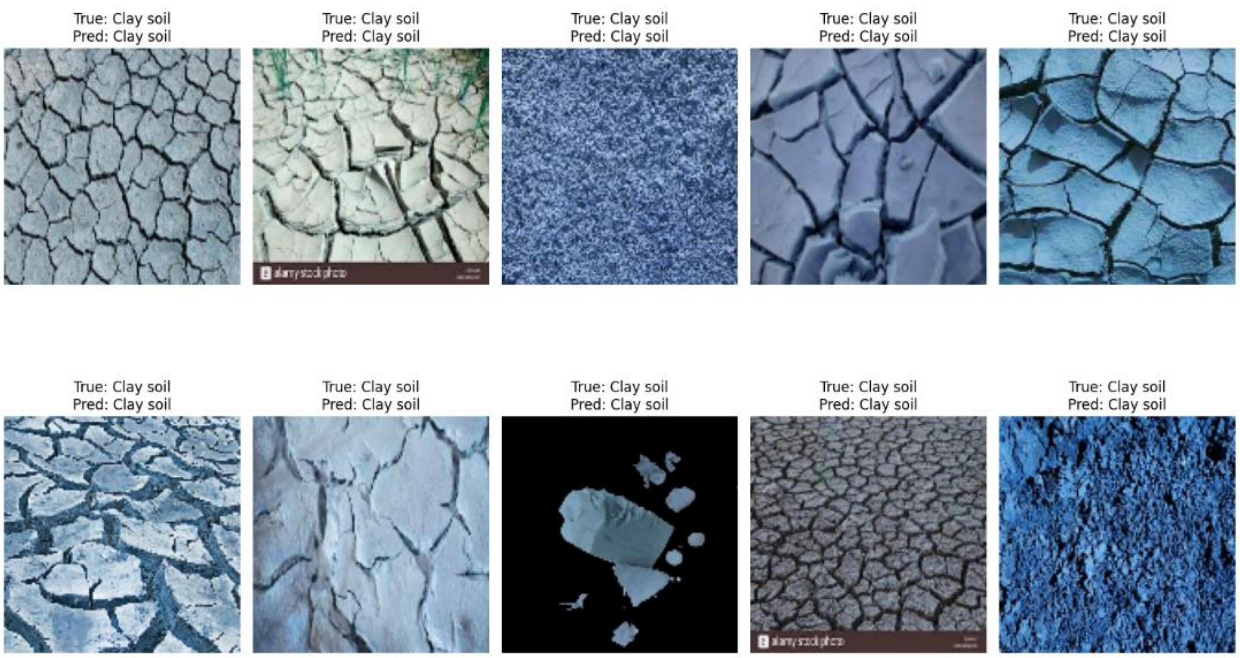

Figure 9: Predicted and Actual Results

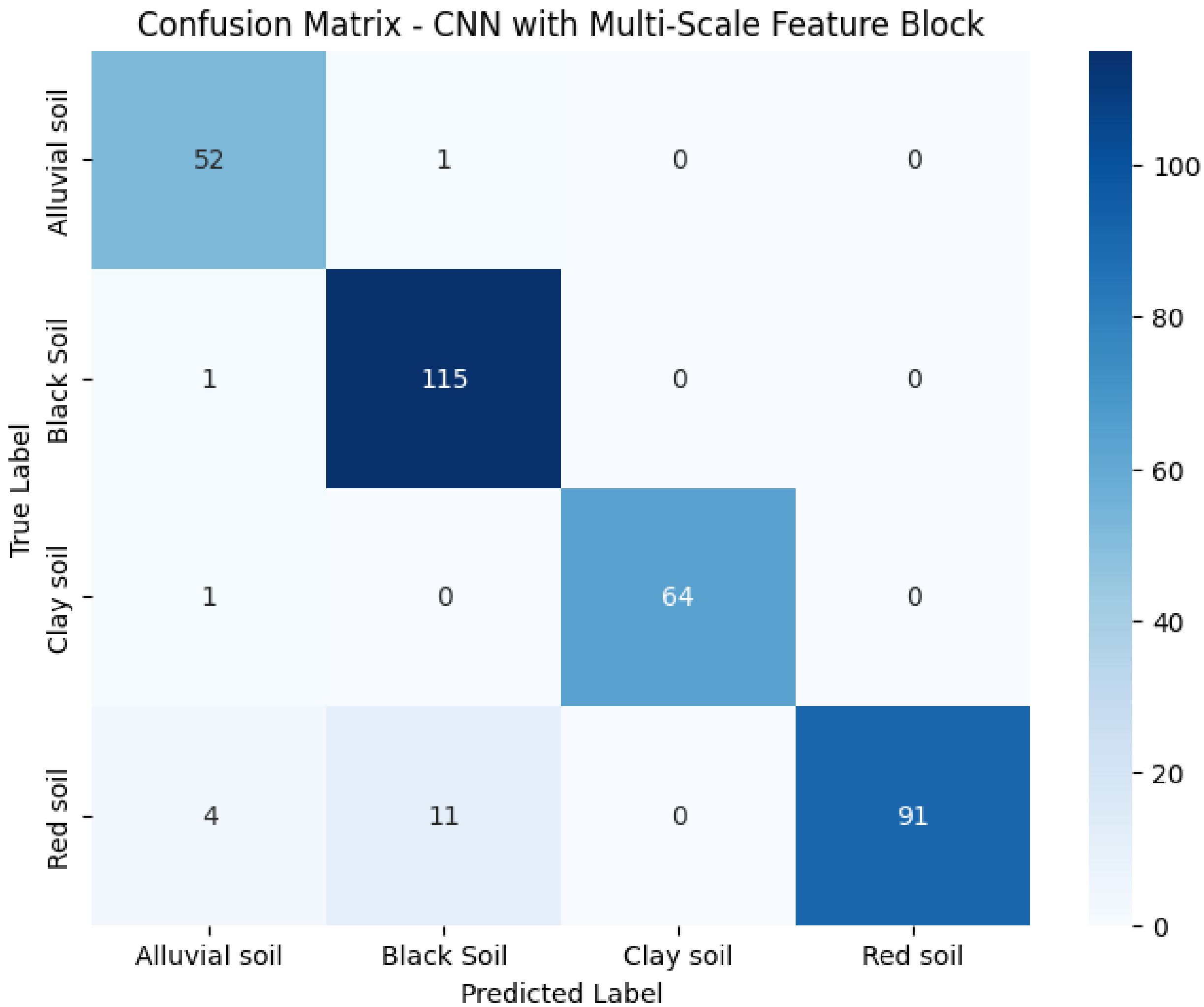

Figure 10: Confusion Matrix of Proposed Methodology

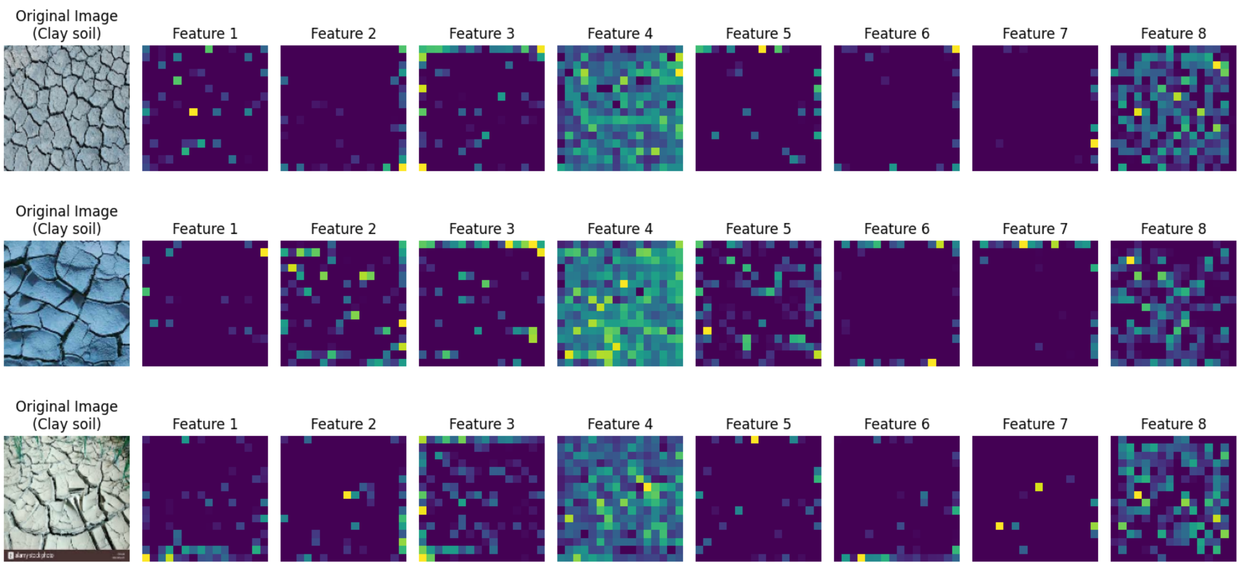

Figure 11: Feature Extraction Sample Images

5. Conclusion

The evaluation of deep learning models for assessing soil fertility based on soil texture presents important consequences for forecasting rice yields. The CNN with Multi-Scale Feature Block emerges as the most effective model, achieving an accuracy of 94.71%, demonstrating its superior capability to evaluate complex soil texture patterns that have a substantial impact on water retention and nutrient availability—two key elements in rice cultivation. The strong performance of the MobileNet+LSTM Hybrid (88.53% accuracy) suggests that the addition of temporal dimensions via LSTM layers can enhance predictions by accounting for seasonal variations in soil conditions. Conventional models like LeNet-5 and EfficientNet also deliver commendable results, indicating their potential use for basic soil assessments. On the other hand, AlexNet's underwhelming performance (47.06% accuracy) underscores the shortcomings of older frameworks in this agricultural domain. These findings illustrate how advanced deep learning models can transform precision agriculture by accurately forecasting how soil texture—from clay-dominated soils impacting root growth to sandy soils that affect nutrient retention—relates to rice yields. Looking forward, it will be crucial to merge these models with other soil parameters, utilize them on edge devices for continuous monitoring, and enhance their clarity through explainable AI techniques to establish comprehensive decision-support systems for sustainable rice farming. Future research should aim to create hybrid models that combine deep learning approaches with established soil science knowledge to provide farmers with actionable insights for optimizing irrigation, fertilization, and other farming practices tailored to their unique soil texture characteristics.

Declarations

Funding: Not applicable.

Ethics approval and consent to participate: Not applicable.

Consent for publication: Not applicable.

Data availability: The datasets used and analyzed in the current study are available from the corresponding author upon reasonable request.

Materials availability: Not applicable.

Code availability: The code developed for model implementation and analysis can be made available by the corresponding author upon reasonable request.

Author Contribution:

Nayana V: Conceptualization, Methodology, Data Curation, Writing – Original Draft

Dr. D Evangelin Geetha: Formal Analysis, Visualization, Writing – Review & Editing

Conflict of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Reference

1. Vasundhara, R., Prakash, N., Anil Kumar, K. S., & Hegde, R. (2020). Characterisation and classification of arecanut-growing soils of Karnataka. Journal of Plantation Crops, 91–102. https://doi.org/10.25081/JPC.2020.V48.I2.6367
2. Rudramurthy, H. V., Basavaraj, K., Rajesh, N. L., Channabsavanna, A. S., Sathishkumar, U., & Vijaya, W. (2024). Taxonomical Approach on Different Soil Types of Kalyana Karnataka, India. https://doi.org/10.9734/ijpss/2024/v36i24388
3. Madesh, P. (2013). Studies on Soil Types and their Characteristic Features in Yalandur Taluk of Chamarajanagar District, Karnataka State.
4. Ye, C.; Zheng, G.; Tao, Y.; Xu, Y.; Chu, G.; Xu, C.; Chen, S.; Liu, Y.; Zhang, X.; Wang, D. Effect of Soil Texture on Soil Nutrient Status and Rice Nutrient Absorption in Paddy Soils. Agronomy 2024, 14, 1339. https://doi.org/10.3390/agronomy14061339
5. Paltseva, A. (2024). What Is Soil Texture? (pp. 25–30). https://doi.org/10.1007/978-3-031-50777-9_4
6. Reddy, R. L. R., Shankarappa, T. H., Kolle, S. R., & Satish, M. V. (2019). Review of Trends in Soil Fertility Research (2007–2016) using Scopus Database. Communications in Soil Science and Plant Analysis, 50(8), 1063–1080. https://doi.org/10.1080/00103624.2019.1600704

7. Fahim, N. S., Khan, B., Rahman, Mohd. S., & Hossain, Md. A. (2024). Effect of Soil Texture on Agricultural Machine Performance in Sylhet, Bangladesh. American Journal of Agricultural Science, Engineering, and Technology. https://doi.org/10.54536/ajaset.v8i2.2663
8. Swain, S. R., Chakraborty, P., Panigrahi, N., Vasava, H. B., Reddy, N. N., Roy, S., Majeed, I., & Das, B. S. (2021). Estimation of soil texture using Sentinel-2 multispectral imaging data: An ensemble modeling approach. Soil & Tillage Research, 213, 105134. https://doi.org/10.1016/J.STILL.2021.105134
9. Safari, Y., Esfandiarpour Boroujeni, I., Kamali, A., Salehi, M. H., & Bagheri Bodaghabadi, M. (2013). Mapping of the soil texture using geostatistical method (a case study of the Shahrekord plain, central Iran). Arabian Journal of Geosciences, 6(9), 3331–3339. https://doi.org/10.1007/S12517-012-0559-9
10. Jain, R. (2023). Determination of soil textural class by using USDA soil texture triangle. International Journal of Innovative Research & Growth, 12(3). https://doi.org/10.26671/ijirg.2023.3.12.106
11. Chuah, C. W., Nureize, A., & Li, S. (2024). An Automated Machine Learning Workflow to Classify Soil Texture for Crop Yield. 1133–1139. https://doi.org/10.1109/cisat62382.2024.10695281
12. Harlianto, P. A., Adji, T. B., & Setiawan, N. A. (2017). Comparison of machine learning algorithms for soil type classification. 7–10. https://doi.org/10.1109/ICSTC.2017.8011843
13. Modi, S., Yadav, A. L., Rana, P. S., & Kalpana, K. (2024). Innovative Deep Learning Methods for Soil Classification and Crop Yield Prediction. https://doi.org/10.1109/iccica60014.2024.10585003
14. Wang, B., Tse, K., & Phung, C. (2023). Machine Learning-Based Cone Penetration Test (CPT) Data Interpretation. https://doi.org/10.21467/proceedings.159.4
15. Nie, X.; Duan, M.; Ding, H.; Hu, B.; Wong, E.K. Attention mask R-CNN for ship detection and segmentation from remote sensing images. IEEE Access 2020, 8, 9325–9334.
16. Li, S.; Zhao, X. Pixel-level detection and measurement of concrete crack using faster region-based convolutional neural network and morphological feature extraction. Meas. Sci. Technol. 2021, 32, 065010.
17. Udendhran, R.; Balamurugan, M.; Suresh, A.; Varatharajan, R. Enhancing image processing architecture using deep learning for embedded vision systems. Microprocess. Microsystems 2020, 76, 103094.
18. Khan, A.; Rauf, Z.; Khan, A.R.; Rathore, S.; Khan, S.H.; Shah, N.S.; Farooq, U.; Asif, H.; Asif, A.; Zahoora, U.; A recent survey of vision transformers for medical image segmentation. arXiv 2023, arXiv:2312.00634.
19. Liu, Q.; Xu, Z.; Bertasius, G.; Niethammer, M. SimpleClick: Interactive image segmentation with simple vision transformers. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Paris, France, 1–6 October 2023; pp. 22290–22300.
20. Qian, X.; Zhang, C.; Chen, L.; Li, K. Deep learning-based identification of maize leaf diseases is improved by an attention mechanism: Self-attention. Front. Plant Sci. 2022, 13, 864486
21. Azad, R.; Kazerouni, A.; Heidari, M.; Aghdam, E.K.; Molaei, A.; Jia, Y.; Jose, A.; Roy, R.; Merhof, D. Advances in medical image analysis with vision transformers: A comprehensive review. Med. Image Anal. 2023, 91, 103000.
22. Li, S.; Wu, C.; Xiong, N. Hybrid architecture based on CNN and transformer for strip steel surface defect classification. Electronics 2022, 11, 1200.
23. Yang, L., Gu, Y., & Feng, H. Multi-scale feature fusion and feature calibration with edge information enhancement for remote sensing object detection. Sci Rep 15, 15371 (2025). https://doi.org/10.1038/s41598-025-99835-7
24. Chen, L.-C., Papandreou, G., Schroff, F., & Adam, H. (2017). Rethinking Atrous Convolution for Semantic Image Segmentation. arXiv: Computer Vision and Pattern Recognition. http://export.arxiv.org/pdf/1706.05587
25. F. Yu and V. Koltun, "Multi-Scale Context Aggregation by Dilated Convolutions," ICLR, 2016.
26. C. Zhang et al., "Multi-Scale Deep Learning for Remote Sensing Image Analysis," IEEE TGRS, vol. 57, no. 8, pp. 5656-5670, 2019.
27. S. Albawi, T. A. Mohammed, and S. Al-Zawi, “Understanding of a convolutional neural network,” in 2017 Int. Conf. on Engineering and Technology (ICET), Aug. 2017, pp. 1–6.
28. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA, USA: MIT Press, 2016.
29. S. Albawi, T. A. Mohammed, and S. Al-Zawi, “Understanding of a convolutional neural network,” in 2017 Int. Conf. on Engineering and Technology (ICET), Aug. 2017, pp. 1–6.
30. V. S. Anandhasivam, A. K. Anusri, M. Logeshwar, and R. Gopinath, "Enhancing Deepfake Detection Through Hybrid MobileNet-LSTM Model with Real-Time Image and Video Analysis," 2024 4th International Conference on Ubiquitous Computing and Intelligent Information Systems (ICUIS), Gobichettipalayam, India, 2024, pp. 1989-1995, doi: 10.1109/ICUIS64676.2024.10867159.
31. Atulanand. (2022, October 29). LeNet-5 complete architecture. Medium. Retrieved from https://medium.com/codex/lenet-5-complete-architecture-84c6d08215f9
32. Li, Shaojuan & Wang, Lizhi & Li, Jia & Yao, Yuan. (2021). Image Classification Algorithm Based on Improved AlexNet. Journal of Physics: Conference Series. 1813. 012051. 10.1088/1742-6596/1813/1/012051.

33. Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. ArXiv. https://arxiv.org/abs/1704.04861
34. Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. ArXiv. https://arxiv.org/abs/1905.11946