

Artificial Intelligence as a Catalyst for Soil Classification: Deep Learning

¹Nanbal Jibba Ladan., ¹Datti Useni Emmanuel., ²Goteng Kuwunidi Job

¹Computer Department Federal University of Education, Pankshin

²Computer Department, Plateau State Polytechnic, Barkin Ladi

DOI: <https://doi.org/10.51584/IJRIAS.2026.11070125>

Received: 24 July 2026; Accepted: 29 July 2026; Published: 10 August 2026

ABSTRACT

Soil classification is a foundational process in agriculture, environmental management, geotechnical engineering, and land-use planning because it determines soil suitability, fertility, productivity, and ecological sustainability. Conventional soil classification approaches depend heavily on field sampling, laboratory analysis, and expert interpretation, which are often labor-intensive, time-consuming, and difficult to scale across large geographic regions. Recent advances in Artificial Intelligence (AI), particularly Deep Learning (DL), have transformed soil classification by enabling automated extraction of complex patterns from heterogeneous datasets including soil images, hyperspectral data, sensor measurements, and geospatial information. This review examines the role of AI as a catalyst for modern soil classification, focusing on deep learning architectures, emerging trends, applications, challenges, and future directions. Evidence from recent studies demonstrates that convolutional neural networks (CNNs), transformer-based models, hybrid architectures, and multimodal learning frameworks substantially improve classification accuracy and operational efficiency compared with traditional approaches. However, limitations associated with data availability, explainability, computational cost, and model transferability remain major concerns.

Keywords: Artificial Intelligence, Deep Learning, Soil Classification, Digital Soil Mapping, Precision Agriculture, Convolutional Neural Networks.

INTRODUCTION

Soil is a critical natural resource that supports agricultural productivity, ecosystem stability, water regulation, carbon sequestration, and environmental sustainability. Accurate soil classification provides essential information for crop selection, nutrient management, irrigation planning, and environmental monitoring. Traditional soil classification methods typically involve manual observation, laboratory testing, and expert interpretation based on physical, chemical, and morphological characteristics (Han, Xiao-Le, et al.2022).

Soil forms the foundation of agricultural productivity and environmental sustainability. Accurate soil classification enables effective crop management, land-use planning, environmental conservation, and infrastructure development. Historically, soil classification has depended on physical inspection and laboratory characterization of texture, structure, mineral composition, moisture content, and chemical properties (Abimala, T., S. Flora Sashya, and K. Sripriya, 2020).

Although conventional approaches remain valuable, they often suffer from limitations including high operational cost, long processing time, restricted spatial coverage, and inconsistency in classification outcomes. The growing availability of remote sensing platforms, geospatial technologies, unmanned aerial vehicles (UAVs), and sensor networks has generated large-scale soil datasets that exceed the analytical capabilities of traditional statistical methods. Artificial Intelligence (AI), particularly Deep Learning (DL), has emerged as an enabling technology for extracting meaningful knowledge from these datasets and automating soil classification processes (Awais et al., 2023).

Artificial Intelligence has emerged as a computational solution capable of transforming soil classification into an automated and scalable process. Deep learning, as an advanced branch of AI, performs end-to-end feature extraction and predictive modelling directly from raw data, reducing dependence on handcrafted features and expert-driven workflows. Recent literature increasingly identifies AI as a major driver of next-generation digital soil intelligence systems.

Recent developments show that AI-driven soil systems are transitioning from simple predictive models toward integrated decision-support platforms capable of continuous monitoring, high-resolution mapping, and adaptive recommendations for sustainable land management (Kikis & Antoniadis, 2026).

Concept of Artificial Intelligence and Deep Learning in Soil Classification

Artificial Intelligence refers to computational systems capable of simulating intelligent decision-making processes using algorithms that learn patterns from data. Within AI, Deep Learning represents a subset based on multilayer neural networks that automatically extract hierarchical features from raw inputs.

In soil classification, deep learning enables automated recognition of soil characteristics from:

i. Soil Profile Images

Soil profile images are visual representations of the vertical section of soil extending from the land surface to the underlying parent material or bedrock. These images reveal the arrangement of soil horizons and provide important information about soil formation processes, physical properties, and environmental conditions. Soil profile images are commonly captured through field excavation, trenching, soil pits, digital photography, and increasingly through remote sensing and image-based analysis systems (Schillaci et al., 2026). A typical soil profile image displays distinct soil horizons that develop through pedogenic processes. These horizons generally include the O horizon (organic layer), A horizon (topsoil rich in organic matter), E horizon (eluviation or leached layer), B horizon (subsoil with accumulated minerals), C horizon (weathered parent material), and R horizon (bedrock) (Britannica, 2026). The visual differences among these layers are usually expressed through variations in colour, texture, structure, porosity, root distribution, and moisture content. Soil profile images play a critical role in soil science because they provide direct visual evidence of soil morphology and support the classification and interpretation of soil properties. Researchers use soil profile images to evaluate soil fertility, drainage conditions, erosion status, carbon storage potential, and land suitability for agriculture and construction purposes (Hartemink & Krasilnikov, 2020). These images also assist in understanding soil genesis and landscape evolution by revealing historical environmental changes preserved within the soil layers (Schillaci et al., 2026).

Advances in digital imaging and artificial intelligence have transformed soil profile analysis. Modern soil profile imaging now integrates high-resolution photography, machine learning, and deep learning algorithms for automated horizon detection, soil classification, and prediction of soil properties. Image datasets obtained from field profiles are increasingly used to train convolutional neural networks and computer vision systems to support precision agriculture and environmental monitoring (Schillaci et al., 2026).

Furthermore, soil profile images serve as valuable educational and communication tools because they simplify the interpretation of complex subsurface processes. They enable scientists, farmers, engineers, and policy makers to visualize soil conditions and make informed decisions regarding sustainable land management and environmental conservation (Environmental Biology, 2026).

ii. Hyperspectral Imagery

Hyperspectral imagery (HSI) is an advanced remote sensing technology that captures image information across hundreds to thousands of contiguous and narrow spectral bands throughout the electromagnetic spectrum. Unlike conventional RGB images or multispectral imagery, which record information in only a few broad bands, hyperspectral imagery provides detailed spectral signatures for every pixel in an image, enabling accurate identification and characterization of materials based on their unique reflectance properties (Bioucas-

Dias et al., 2023). A hyperspectral image consists of three dimensions commonly referred to as a hyperspectral cube: two spatial dimensions (x and y coordinates) and one spectral dimension (λ). Each pixel within the image contains a complete spectral profile that reflects how an object interacts with electromagnetic radiation across multiple wavelengths. This rich spectral information allows researchers to distinguish between materials that may appear visually similar in ordinary images but exhibit different spectral responses (Chang, 2022). The operation of hyperspectral imaging relies on measuring reflected or emitted radiation from objects over a broad range of wavelengths, typically extending from the visible region (400–700 nm) through near-infrared (700–2500 nm) and sometimes into shortwave infrared regions. The collected spectral signatures can then be analyzed using statistical, machine learning, and deep learning techniques for classification, detection, and prediction tasks (Li et al., 2024). Hyperspectral imagery has become increasingly important in soil science and precision agriculture because it enables non-destructive assessment of soil properties. Researchers use hyperspectral data to estimate soil organic matter, moisture content, mineral composition, nutrient concentration, salinity, pH levels, and soil texture with high spatial and spectral resolution (Dotto et al., 2021). In soil classification studies, hyperspectral images provide detailed spectral information that improves the discrimination of soil classes and supports automated mapping systems. Recent advances in artificial intelligence have significantly improved hyperspectral image analysis. Deep learning approaches, particularly convolutional neural networks (CNNs), recurrent neural networks (RNNs), transformers, and hybrid architectures, have demonstrated strong performance in extracting spectral-spatial features for hyperspectral classification tasks. These methods reduce dependence on manual feature engineering and achieve higher classification accuracy in complex environmental conditions (Zhang et al., 2025). Despite its advantages, hyperspectral imagery presents several challenges. The high dimensionality of hyperspectral datasets often leads to increased computational complexity and storage requirements. Additionally, hyperspectral systems are sensitive to atmospheric interference, sensor noise, and data redundancy, making preprocessing techniques such as normalization, dimensionality reduction, denoising, and feature selection essential before analysis (Li et al., 2024).

Overall, hyperspectral imagery represents a powerful sensing technology for environmental monitoring and intelligent soil assessment. Its ability to capture detailed spectral information combined with modern artificial intelligence techniques continues to expand its applications in soil classification, precision agriculture, and sustainable land management.

iii. Satellite Observations

Satellite observations refer to the acquisition of information about the Earth's surface and atmosphere using sensors mounted on orbiting satellites. These observations provide continuous, large-scale, and non-invasive monitoring of environmental conditions and have become fundamental tools in earth sciences, agriculture, climate studies, and soil assessment. Satellite-based observations generate spatially distributed datasets that enable researchers to analyze land characteristics over extensive geographical areas and long time periods (Wulder et al., 2022). Satellite observation systems collect electromagnetic radiation reflected or emitted from the Earth's surface through passive and active sensors. Passive sensors measure naturally reflected solar radiation or emitted thermal energy, whereas active sensors such as radar emit their own signals and measure the returned energy after interaction with surface features (Liang et al., 2023). The resulting data are transformed into satellite images that contain valuable information regarding vegetation, soil conditions, moisture dynamics, land use, and environmental changes.

In soil science and land resource assessment, satellite observations have emerged as essential sources of spatial information for soil mapping and classification. Satellite platforms capture spectral, temporal, and spatial characteristics of soils, enabling estimation of soil properties such as organic carbon content, soil moisture, texture, mineral composition, salinity, and fertility status (Taghizadeh-Mehrjardi et al., 2021). These observations support digital soil mapping by reducing the dependence on extensive field sampling while improving coverage and efficiency.

Several satellite missions provide data commonly used in soil and environmental studies. Optical satellites such as Landsat and Sentinel-2 provide multispectral imagery suitable for monitoring soil reflectance and land cover changes, while radar satellites such as Sentinel-1 enable observation under cloudy conditions and

support soil moisture estimation (Gorelick et al., 2022). The integration of these multi-source satellite datasets allows comprehensive characterization of soil variability across landscapes.

Recent developments in artificial intelligence and deep learning have significantly enhanced the processing and interpretation of satellite observations. Convolutional neural networks (CNNs), transformers, and hybrid deep learning architectures are increasingly applied to extract spatial-spectral patterns from satellite imagery for automated soil classification, land degradation monitoring, and precision agriculture applications (Zhu et al., 2024). AI-driven satellite analysis improves prediction accuracy and supports near real-time environmental decision-making. Despite their advantages, satellite observations face challenges including atmospheric interference, cloud contamination, sensor calibration issues, temporal inconsistencies, and limitations in spatial resolution. To address these challenges, preprocessing techniques such as atmospheric correction, radiometric normalization, image fusion, and feature extraction are commonly applied before analytical modelling (Liang et al., 2023).

Satellite observations have transformed modern soil assessment by enabling large-scale, cost-effective, and timely acquisition of environmental data. Their integration with artificial intelligence technologies continues to advance intelligent soil classification systems and sustainable land management practices.

iv. Sensor Measurements

Sensor measurements refer to the collection of physical, chemical, and environmental data through sensing devices that detect and quantify changes in specific parameters within a system or environment. In soil science and environmental monitoring, sensors provide direct and continuous observations of soil characteristics and have become essential for modern precision agriculture, digital soil mapping, and intelligent soil classification systems (Adamchuk et al., 2021). A sensor operates by converting physical or chemical properties into measurable electrical signals that can be processed and analyzed. Soil-related sensor measurements commonly include soil moisture, temperature, electrical conductivity, pH, organic matter, nutrient concentration, salinity, and compaction. These measurements are acquired using contact-based sensors placed within the soil or non-contact sensors that rely on electromagnetic interactions with soil surfaces (Viscarra Rossel et al., 2022). Sensor technologies used in soil analysis can be broadly classified into in-situ sensors and remote sensing systems. In-situ sensors collect data directly from the field environment and include capacitance sensors, tensiometers, electrochemical sensors, optical sensors, and dielectric probes. These sensors enable real-time monitoring and provide high temporal resolution for understanding soil variability and dynamic environmental conditions (Nocita et al., 2021).

Recent advances in sensor technologies have expanded the use of proximal soil sensing (PSS), where portable or mounted sensors are used to collect high-density soil information at close distances. Proximal sensing techniques commonly employ visible and near-infrared (Vis-NIR) spectroscopy, gamma-ray sensing, electromagnetic induction, and laser-based systems to estimate multiple soil attributes simultaneously (Kuang et al., 2023). These methods reduce labour requirements and improve the efficiency of soil assessment processes.

The integration of sensor measurements with artificial intelligence and deep learning has significantly enhanced soil classification and predictive modelling. Machine learning algorithms such as support vector machines, random forests, convolutional neural networks (CNNs), and transformer-based architectures are increasingly applied to process large sensor datasets and identify complex patterns associated with soil properties and land conditions (Minasny et al., 2024). AI-driven sensor systems support automated decision-making and improve classification accuracy in agricultural and environmental applications.

Sensor measurements also contribute to precision agriculture by enabling site-specific management practices. Real-time sensor data allow farmers and researchers to optimize irrigation schedules, fertilizer application, crop monitoring, and environmental conservation strategies based on actual field conditions rather than generalized assumptions (Viscarra Rossel et al., 2022). Despite these advantages, sensor-based measurements face challenges including calibration requirements, environmental noise, signal variability, sensor drift, and limitations associated with data integration and interoperability. Therefore, preprocessing techniques such as

filtering, normalization, calibration correction, and feature extraction are often applied to improve data quality before analytical modelling (Kuang et al., 2023). Sensor measurements provide a critical source of high-resolution environmental data that supports intelligent soil assessment and classification. Their integration with artificial intelligence technologies continues to advance precision agriculture and sustainable land management systems.

v. Geospatial Datasets

Geospatial datasets refer to collections of spatially referenced data that describe the location, characteristics, and relationships of features on or beneath the Earth's surface. These datasets integrate geographic coordinates with attribute information, allowing researchers to analyze spatial patterns, environmental conditions, and interactions across landscapes. Geospatial datasets have become fundamental resources in earth observation, environmental management, agriculture, and intelligent soil classification systems (Goodchild et al., 2021). Geospatial datasets are generally composed of two major components: spatial data and attribute data. Spatial data describe the geographic position and geometry of objects using coordinate systems, while attribute data provide descriptive information related to those objects. Spatial information may be represented in raster format, where data are organized into grids of pixels, or vector format, where features are represented as points, lines, and polygons (Longley et al., 2022).

In soil science and land resource assessment, geospatial datasets provide essential information for understanding soil variability and supporting digital soil mapping. These datasets are obtained from multiple sources including satellite imagery, hyperspectral observations, field surveys, sensor networks, geographic information systems (GIS), unmanned aerial vehicles (UAVs), and global environmental databases (Minasny et al., 2023). Common geospatial variables used in soil classification include elevation, slope, land cover, vegetation indices, rainfall, temperature, soil moisture, and geological characteristics. The integration of geospatial datasets with remote sensing technologies has significantly improved the efficiency and accuracy of soil assessment. Satellite-derived geospatial products provide continuous observations over large areas, enabling the generation of spatial models that capture soil distribution and environmental changes over time. Geographic Information Systems further facilitate storage, visualization, manipulation, and analysis of these large and complex datasets (Zhu et al., 2024).

Recent developments in artificial intelligence and deep learning have transformed the analysis of geospatial datasets. Deep learning architectures such as convolutional neural networks (CNNs), graph neural networks (GNNs), recurrent neural networks (RNNs), and transformer-based models are increasingly employed to extract spatial patterns and improve predictive performance in soil classification and environmental monitoring tasks (Ma et al., 2025). These approaches enable automated feature extraction and support high-dimensional geospatial analysis.

Geospatial datasets also support precision agriculture and sustainable land management by enabling site-specific interventions and evidence-based decision-making. Farmers and environmental managers can utilize geospatial information to optimize fertilizer application, irrigation scheduling, crop monitoring, erosion control, and land-use planning (Minasny et al., 2023). Despite their advantages, geospatial datasets present several challenges including data heterogeneity, spatial uncertainty, missing observations, temporal inconsistencies, and computational complexity associated with large-scale analysis. To overcome these challenges, preprocessing techniques such as georeferencing, spatial interpolation, normalization, dimensionality reduction, and data fusion are commonly applied before analytical modelling (Goodchild et al., 2021).

Geospatial datasets serve as critical inputs for modern soil classification and environmental intelligence systems. Their integration with artificial intelligence technologies continues to enhance the accuracy, scalability, and efficiency of soil analysis and sustainable resource management.

vi. Environmental Covariates

Environmental covariates are auxiliary variables that describe the physical, climatic, biological, and geological conditions of a landscape and are used to explain or predict spatial and temporal variations in environmental

processes. In soil science and digital soil mapping, environmental covariates are essential predictors that help model soil properties and soil classes by capturing the factors that influence soil formation and distribution across different landscapes (McBratney et al., 2021). Environmental covariates typically include a wide range of biophysical and climatic variables such as elevation, slope, aspect, curvature, land surface temperature, precipitation, vegetation indices, parent material, land use/land cover, and proximity to water bodies. These variables are often derived from remote sensing products, satellite observations, climate reanalysis datasets, and digital elevation models (DEM) (Hengl et al., 2021).

In the context of soil mapping, environmental covariates are used within the framework of the SCORPAN model, which describes soil as a function of soil-forming factors including climate, organisms, relief, parent material, age, and spatial position. This framework supports the idea that soil properties can be inferred from measurable environmental variables that are spatially continuous and globally available (McBratney et al., 2021). By integrating these covariates, researchers can predict soil attributes even in areas with limited or no direct field observations. Modern digital soil mapping systems rely heavily on high-resolution environmental covariates extracted from satellite platforms such as Sentinel and Landsat, as well as global climate databases and terrain models. These datasets allow for fine-scale prediction of soil variability and improve the accuracy of machine learning and geostatistical models used in environmental analysis (Hengl et al., 2021).

Recent advances in artificial intelligence and machine learning have significantly improved the utilization of environmental covariates. Algorithms such as random forests, gradient boosting machines, support vector machines, and deep neural networks are commonly used to integrate multiple covariates and identify complex nonlinear relationships between environmental conditions and soil properties (Zhang et al., 2024). These models enhance predictive performance and support automated soil classification systems.

Environmental covariates also play a critical role in precision agriculture and environmental monitoring by supporting site-specific management decisions. For example, variations in vegetation indices and soil moisture covariates can be used to optimize irrigation, fertilizer application, and crop selection strategies, thereby improving agricultural productivity and sustainability (Florinsky, 2022).

Despite their usefulness, environmental covariates present challenges such as multicollinearity, spatial autocorrelation, scale mismatch, and uncertainty in derived datasets. To address these issues, preprocessing techniques such as feature selection, principal component analysis (PCA), normalization, and spatial resampling are often applied before modeling (Hengl et al., 2021).

Environmental covariates are fundamental components of modern soil modeling frameworks. Their integration with remote sensing and artificial intelligence has significantly advanced the accuracy and efficiency of soil prediction, classification, and environmental decision-making.

Unlike traditional machine learning methods that require manual feature engineering, deep learning learns representations directly from data and therefore adapts more effectively to complex and nonlinear soil relationships (Pal & Gupta, 2025).

Theoretical Foundation of AI-Driven Soil Classification

Artificial Intelligence and Soil Analytics

Artificial Intelligence refers to computational methods capable of learning patterns and making decisions from data. Soil analytics uses AI to classify soils according to physical, chemical, spectral, and environmental properties.

Modern soil classification systems integrate:

Computer vision

Computer vision refers to a branch of artificial intelligence (AI) that enables computers to acquire, process, analyze, and interpret visual information from images and videos to support automated decision-making. In

soil science, computer vision has emerged as an important technology for soil classification because it allows rapid, accurate, and cost-effective analysis of soil characteristics compared with conventional laboratory and field-based methods (Minasny et al., 2021).

Traditional soil classification relies heavily on manual observation of soil properties such as color, texture, structure, organic matter content, and horizon development. These approaches are often labor-intensive, time-consuming, and subject to observer variability. Computer vision addresses these limitations by extracting meaningful visual features from soil images and integrating them with machine learning and deep learning algorithms to classify soils automatically (Zhang et al., 2022).

The application of computer vision in soil classification generally involves several stages: image acquisition, image preprocessing, feature extraction, model training, and classification. Image acquisition may involve the use of digital cameras, unmanned aerial vehicles (UAVs), smartphones, hyperspectral sensors, or satellite imagery to capture soil surface and profile information. Preprocessing techniques such as image enhancement, normalization, segmentation, and noise reduction improve image quality and prepare the data for analysis (Padarian et al., 2021).

Feature extraction is a critical stage in computer vision-based soil classification. Traditional approaches rely on handcrafted features including color histograms, texture descriptors, edge information, and shape characteristics. Soil color is particularly important because it provides information about mineral composition, moisture content, and organic matter. Texture features extracted using methods such as Gray-Level Co-occurrence Matrix (GLCM) and Local Binary Pattern (LBP) are frequently employed to distinguish soil classes (Abdelkader et al., 2022).

Recent developments have shifted toward deep learning-based computer vision approaches, particularly Convolutional Neural Networks (CNNs), which automatically learn hierarchical features directly from soil images. CNN architectures such as AlexNet, VGGNet, ResNet, EfficientNet, and Vision Transformers (ViTs) have demonstrated superior performance in soil classification tasks due to their ability to capture complex spatial patterns and nonlinear relationships within image data (Khan et al., 2023). These models reduce dependence on manual feature engineering and improve classification accuracy across diverse soil conditions.

Remote Sensing

Remote sensing refers to the acquisition of information about objects or surfaces without direct physical contact through sensors mounted on satellites, aircraft, unmanned aerial vehicles (UAVs), or ground-based platforms. In soil science, remote sensing has become an essential technology for soil classification because it enables rapid, large-scale, and non-destructive assessment of soil properties across different spatial and temporal scales (Minasny et al., 2021).

Conventional soil classification methods typically involve extensive field sampling and laboratory analysis, which are expensive, labor-intensive, and difficult to apply over large geographical areas. Remote sensing overcomes these limitations by providing continuous spatial observations that capture variations in soil characteristics such as texture, moisture, mineral composition, organic matter, salinity, and land surface conditions (Padarian et al., 2021).

Remote sensing for soil classification operates on the principle that soils exhibit unique spectral signatures due to differences in their physical and chemical properties. Sensors detect reflected or emitted electromagnetic radiation from soil surfaces across different wavelength regions, including visible (VIS), near-infrared (NIR), shortwave infrared (SWIR), and thermal infrared (TIR) bands. These spectral responses are then analyzed to distinguish and classify different soil types (Mulder et al., 2022).

Geographic Information Systems (GIS)

Geographic Information Systems (GIS) refer to computer-based systems designed to capture, store, manage, analyze, visualize, and interpret geographically referenced data. In soil science, GIS has become an essential

tool for soil classification because it enables the integration of spatial data from multiple sources to support accurate mapping, monitoring, and analysis of soil properties across different landscapes (Minasny et al., 2021).

Traditional soil classification methods rely largely on field surveys and laboratory analyses, which often require significant time, labor, and financial resources. GIS improves this process by providing a platform that integrates soil observations with spatial variables such as topography, climate, geology, land use, vegetation cover, and remote sensing information to generate detailed and scalable soil classification outputs (Hengl et al., 2022).

GIS supports soil classification through spatial data management and spatial analysis techniques. Spatial datasets are commonly organized into raster and vector formats. Raster data represent continuous environmental variables such as elevation and soil moisture, while vector data represent discrete features including sampling locations and administrative boundaries. These data layers are analyzed collectively to identify spatial relationships and classify soils according to established taxonomic systems (Taghizadeh-Mehrjardi et al., 2022).

One of the major applications of GIS in soil classification is Digital Soil Mapping (DSM). Digital Soil Mapping uses mathematical and statistical models to predict soil classes and properties based on observed soil samples and environmental covariates. GIS serves as the analytical framework that combines field observations with environmental variables to produce spatially continuous soil maps with improved efficiency and accuracy (McBratney et al., 2021).

Hyperspectral Imaging

Hyperspectral imaging (HSI) is an advanced remote sensing technology that acquires image data across hundreds of narrow and contiguous spectral bands throughout the electromagnetic spectrum. Unlike conventional RGB imaging or multispectral systems that capture only a few broad wavelength bands, hyperspectral imaging records detailed spectral information for every pixel, enabling precise identification and characterization of surface materials. In soil science, hyperspectral imaging has become an important technique for soil classification because of its capability to detect subtle variations in soil physical, chemical, and mineralogical properties (Nocita et al., 2023).

Traditional approaches to soil classification depend on field surveys and laboratory analysis, which are often costly, time-consuming, and limited in spatial coverage. Hyperspectral imaging provides a rapid and non-destructive alternative by capturing spectral signatures that reflect the composition and condition of soils. Soil attributes such as texture, organic matter content, moisture level, mineral composition, salinity, and nutrient concentration influence spectral reflectance patterns, making them detectable through hyperspectral analysis (Minasny et al., 2021).

The principle of hyperspectral imaging in soil classification is based on the interaction between electromagnetic radiation and soil materials. Different soil components absorb and reflect radiation differently across spectral wavelengths, creating unique spectral fingerprints. These spectral responses are measured and analyzed to distinguish among soil classes and estimate soil properties with high precision (Mulder et al., 2022).

Hyperspectral imaging systems may be deployed on satellite platforms, aircraft, unmanned aerial vehicles (UAVs), or laboratory-based sensors. Satellite hyperspectral missions and airborne imaging systems provide large-scale observations, while proximal sensing instruments enable detailed local analysis of soil samples. The flexibility of these acquisition platforms has expanded the application of hyperspectral imaging in digital soil mapping and precision agriculture (Taghizadeh-Mehrjardi et al., 2022).

Cloud and Edge Computing

Cloud and edge computing are emerging computational paradigms that support the storage, processing, analysis, and delivery of large-scale data generated from modern agricultural and environmental monitoring

systems. In soil classification, these technologies have become increasingly important because they enable efficient handling of massive datasets obtained from remote sensing platforms, hyperspectral imaging systems, Geographic Information Systems (GIS), sensors, and artificial intelligence models (Li et al., 2024).

Traditional soil classification approaches generally depend on field sampling and laboratory analysis, which can be limited by processing speed, storage capacity, and scalability. The increasing adoption of digital soil mapping, Internet of Things (IoT) devices, and machine learning has generated large volumes of spatial and environmental data that require advanced computational infrastructure. Cloud and edge computing provide solutions for managing these data-intensive workflows while improving the speed and accessibility of soil classification systems (Minasny et al., 2021).

Cloud computing refers to the delivery of computing services—including storage, processing power, databases, networking, and analytics—through remote internet-based infrastructure. In soil classification, cloud platforms allow researchers and agricultural practitioners to process large datasets without relying on local computing resources. Cloud environments support large-scale geospatial analysis, machine learning model training, and integration of multiple environmental datasets for digital soil mapping applications (Hengl et al., 2022).

Cloud-based platforms have become widely adopted for soil analysis because they enable access to satellite imagery, climate data, and computational resources in real time. Technologies such as cloud geospatial processing environments facilitate automated analysis of soil variability, prediction of soil classes, and generation of high-resolution soil maps over extensive geographic regions. The scalability of cloud infrastructure reduces computational constraints associated with deep learning and remote sensing applications (Padarian et al., 2021).

Edge computing complements cloud computing by moving data processing closer to the source of data generation. Instead of transmitting all data to centralized servers, edge devices process information locally or near the sensing environment. In soil classification, edge computing enables real-time analysis of soil images, sensor measurements, and environmental observations directly in agricultural fields. This reduces network latency, minimizes bandwidth requirements, and supports rapid decision-making for precision agriculture applications (Shi et al., 2022).

Transition Toward AI-Based Soil Classification

The integration of AI has introduced automated workflows capable of processing millions of observations simultaneously. Digital Soil Mapping (DSM), combined with machine learning and deep learning, enables predictive classification across broad geographical areas with improved temporal and spatial resolution (Yang & Li, 2026).

AI-driven classification systems increasingly combine:

- i. Remote sensing;
- ii. Geographic Information Systems (GIS);
- iii. Internet of Things (IoT);
- iv. Cloud computing;
- v. Edge AI platforms.

These integrations have accelerated precision agriculture and sustainable land-use management (Sabharwal & Reddy, 2026).

Deep Learning Architectures Driving Soil Classification

Convolutional Neural Networks (CNNs)

CNNs remain among the most widely adopted deep learning models for soil image classification due to their ability to identify texture, color variation, and spatial patterns automatically.

CNN-based systems have demonstrated superior performance in:

i. Soil Texture Recognition

Soil texture recognition refers to the process of identifying and categorizing soil based on the relative proportions of sand, silt, and clay particles that constitute the soil matrix. Soil texture is one of the most fundamental properties used in soil classification because it influences water retention, nutrient availability, aeration, root penetration, permeability, and overall soil productivity. Accurate recognition of soil texture supports agricultural planning, environmental monitoring, and sustainable land management practices (Minasny et al., 2021).

Traditional methods of soil texture recognition rely on field observations and laboratory-based mechanical analysis techniques such as the hydrometer method, pipette method, and particle size distribution analysis. Although these approaches provide reliable measurements, they are often labor-intensive, time-consuming, costly, and difficult to apply across large geographic areas. These limitations have encouraged the adoption of digital and automated approaches for soil texture recognition (Padarian et al., 2021).

ii. Soil Fertility Assessment

Soil fertility assessment refers to the systematic evaluation of a soil's capacity to supply essential nutrients and provide suitable conditions for plant growth and agricultural productivity. Soil fertility is a major indicator in soil classification because it reflects the physical, chemical, and biological characteristics that determine soil quality and land suitability. Accurate assessment of soil fertility supports crop management, precision agriculture, environmental sustainability, and efficient land-use planning (Minasny et al., 2021).

Traditional soil fertility assessment methods depend on field sampling and laboratory analysis to determine nutrient status and soil quality indicators. Common parameters evaluated include soil pH, organic matter content, nitrogen (N), phosphorus (P), potassium (K), cation exchange capacity (CEC), soil texture, moisture content, and micronutrient availability. Although laboratory methods provide reliable results, they are often costly, time-consuming, and difficult to implement across extensive geographic areas (Padarian et al., 2021).

iii. Soil Image Segmentation

Soil image segmentation refers to the process of partitioning soil images into meaningful and homogeneous regions to isolate important soil features for analysis and classification. It is a fundamental stage in computer vision and image-based soil analysis because it enables the identification of distinct soil components, boundaries, textures, and structural characteristics. In soil classification, image segmentation improves the extraction of relevant information from soil images and enhances the accuracy of automated classification systems (Minasny et al., 2021).

Traditional soil classification methods largely depend on manual field observations and laboratory analysis, which are often labor-intensive and susceptible to subjective interpretation. Advances in digital imaging and artificial intelligence have enabled soil scientists to use image segmentation techniques to automate the analysis of soil characteristics such as texture, color, structure, porosity, moisture patterns, and horizon boundaries (Padarian et al., 2021).

iv. Soil type identification

Soil type identification refers to the process of recognizing, distinguishing, and categorizing soils into specific classes based on their physical, chemical, biological, and morphological characteristics. It is a core component of soil classification because it enables the systematic organization of soils according to their properties and supports decision-making in agriculture, environmental management, land-use planning, and ecosystem conservation. Accurate soil type identification is essential for understanding soil behavior and optimizing land productivity (Minasny et al., 2021).

Conventional approaches to soil type identification rely on field surveys, soil profile examination, and laboratory testing. Soil scientists traditionally classify soils by evaluating characteristics such as color, texture, structure, pH, moisture content, organic matter, mineral composition, drainage conditions, and horizon development. While these methods provide reliable results, they are often labor-intensive, time-consuming, expensive, and difficult to apply over large geographic regions (Padarian et al., 2021).

Research indicates that CNNs outperform conventional feature-engineering approaches in image-based soil analysis because of their end-to-end learning capability (Gyasi & Purushotham, 2023).

Applications of Deep Learning in Soil Classification

Digital Soil Mapping

Digital Soil Mapping has become one of the most impactful applications of AI in soil science.

Deep learning supports:

i. Automated Mapping

Automated mapping refers to the use of computational methods, spatial technologies, and artificial intelligence to generate maps with minimal manual intervention. In soil classification, automated mapping involves the automatic collection, processing, analysis, and visualization of soil-related data to produce accurate spatial representations of soil classes and properties. This approach has become increasingly important in modern soil science because it enables rapid, scalable, and cost-effective mapping of soil resources across different geographic regions (Minasny et al., 2021).

Traditional soil mapping methods are commonly based on field surveys, manual delineation, laboratory analysis, and expert interpretation. Although these approaches provide reliable outputs, they are often time-consuming, expensive, and difficult to update frequently over large areas. Automated mapping addresses these limitations by integrating digital technologies that accelerate data acquisition and improve consistency in soil classification processes (McBratney et al., 2021).

ii. Spatial Interpolation

Spatial interpolation refers to a group of statistical and geospatial techniques used to estimate unknown values at unsampled locations based on observed data from surrounding points. In soil classification, spatial interpolation plays an important role in predicting the distribution of soil properties and generating continuous soil maps from limited field observations. Since direct soil sampling across an entire landscape is often impractical and costly, interpolation techniques provide an efficient approach for estimating soil variability and supporting soil classification and digital soil mapping (Minasny et al., 2021).

Traditional soil surveys generally involve collecting soil samples at selected locations and extending conclusions to unsampled areas through expert judgment. However, soil properties often exhibit spatial variability influenced by environmental and topographic conditions. Spatial interpolation addresses this challenge by using mathematical and statistical relationships among neighboring observations to estimate values across geographic space and improve classification accuracy (McBratney et al., 2021).

iii. Prediction of Unsampled Locations

Prediction of unsampled locations refers to the process of estimating soil characteristics at locations where direct measurements or observations are unavailable by using existing sampled data and predictive modelling techniques. In soil classification, this approach is essential because collecting soil samples across an entire landscape is often expensive, labor-intensive, and impractical. Predictive methods therefore enable the generation of continuous spatial information and support efficient soil mapping and land management decisions (McBratney et al., 2021).

Traditional soil classification approaches depend heavily on field surveys and laboratory analyses conducted at selected sampling points. Although these methods provide accurate local measurements, they often fail to represent the spatial variability of soil properties over large areas. Predicting conditions at unsampled locations addresses this limitation by extending information from observed locations to surrounding areas using spatial relationships and environmental predictors (Minasny et al., 2021).

iv. Soil Boundary Identification

Soil boundary identification refers to the process of detecting, delineating, and mapping transitions between different soil units or horizons based on variations in soil properties and spatial characteristics. It is an important component of soil classification because accurate identification of soil boundaries supports the separation of soil classes, improves soil mapping accuracy, and enhances understanding of soil variability across landscapes. Soil boundaries may occur vertically within soil profiles or horizontally across land surfaces and are influenced by soil-forming factors such as climate, topography, parent material, organisms, and time (Minasny et al., 2021).

Traditional approaches to soil boundary identification rely on field surveys, expert interpretation, and manual delineation using observable characteristics including soil color, texture, structure, drainage, horizon depth, and vegetation patterns. Although these methods provide valuable information, they are often labor-intensive, time-consuming, and subject to observer variability, particularly when applied across large geographic areas (McBratney, et al., 2021).

Recent developments in digital soil science have introduced automated and data-driven approaches for identifying soil boundaries. These approaches combine spatial analysis, remote sensing, computer vision, Geographic Information Systems (GIS), and machine learning techniques to improve efficiency and consistency in soil classification (Hengl, et al., 2022)

Large-scale mapping frameworks increasingly rely on AI for real-time soil information generation (Yang & Li, 2026).

Table 1: Comparative Analysis of Deep Learning Models for Soil Classification

Model	Architecture Description	Typical Input Data	Strengths	Limitations	Reported Soil Classification Applications	Recent Literature
Convolutional Neural Networks (CNN)	Deep neural network that extracts hierarchical spatial features through convolution operations	Soil profile images, RGB images, remote sensing imagery	High feature extraction capability; strong image recognition performance;	Requires large labeled datasets; computationally intensive	Soil texture classification, fertility prediction, soil profile identification	Gyasi & Purushotham (2023); Awais et al. (2023)

			reduced manual feature engineering			
Deep CNN (DCNN)	Multi-layer CNN with deeper architecture for improved representation learning	High-resolution soil imagery	Better classification accuracy; captures complex spatial structures	Higher training cost; prone to overfitting	Fine-grained soil category recognition	Pal & Gupta (2025)
Vision Transformer (ViT)	Transformer-based visual architecture using self-attention mechanisms	Satellite imagery, hyperspectral images	Captures long-range dependencies; effective for large datasets	Requires extensive training data and computational resources	Large-scale soil mapping and classification	Recent Transformer studies (2024–2026)
CNN–Transformer Hybrid	Combines CNN local feature extraction with transformer global context modeling	Soil image datasets and multisource data	Improved robustness and generalization	Complex architecture design	Multiscale soil and environmental prediction	Hybrid DL studies (2025–2026)
Recurrent Neural Networks (RNN/LSTM)	Sequential neural architecture designed for temporal pattern learning	Time-series soil sensor data	Handles temporal dependencies effectively	Less efficient for large visual datasets	Soil moisture and temporal soil monitoring	Recent agricultural AI studies (2022–2025)
Autoencoder-Based Models	Unsupervised deep learning model for feature compression and extraction	Hyperspectral soil datasets	Effective dimensionality reduction	Reduced interpretability	Spectral soil classification	Environmental AI reviews (2023–2025)
Generative Adversarial Networks	Dual-network architecture for synthetic	Limited soil image datasets	Improves data augmentation and	Training instability	Synthetic soil dataset generation	Emerging AI applications (2025–2026)

(GANs)	data generation		model robustness			
Multimodal Deep Learning Models	Integrates multiple data modalities into a unified model	Images, GIS layers, sensor data, climate variables	Improved classification accuracy and adaptability	Complex preprocessing requirements	Digital mapping precision agriculture and soil	Recent multimodal studies (2025–2026)

Comparative Interpretation

The reviewed evidence suggests that CNN-based models remain the most widely implemented architectures for soil classification because of their strong performance in image analysis and automated feature extraction. However, Vision Transformers and hybrid CNN–Transformer approaches increasingly demonstrate superior performance for large-scale and heterogeneous soil datasets due to improved contextual learning.

Multimodal architectures appear to represent the current state of the art because they combine complementary data sources and reduce limitations associated with single-input systems. Emerging approaches involving explainable AI and foundation models are expected to further improve transparency and scalability of soil classification systems.

Challenges Limiting Deep Learning Adoption

Despite substantial progress, several barriers persist.

Data Scarcity

Deep learning models require extensive labeled datasets, which remain limited in many regions.

Explainability

Many deep learning systems operate as black-box models, limiting trust and adoption in scientific decision-making.

Transferability

Models trained under specific environmental conditions frequently exhibit reduced performance in different geographical settings.

Computational Complexity

Training advanced architectures requires high-performance computing infrastructure.

These challenges remain active research priorities in AI-enabled soil science (Kikis & Antoniadis, 2026).

CONCLUSION

Artificial Intelligence has emerged as a major catalyst transforming soil classification from manual and laboratory-centered approaches toward intelligent, scalable, and data-driven systems. Deep learning models particularly CNNs, transformers, and hybrid architectures have demonstrated remarkable capability in extracting meaningful representations from complex soil datasets. These advances improve classification accuracy, enable digital soil mapping, and support sustainable agricultural decision-making. Nevertheless, addressing challenges related to interpretability, data scarcity, computational requirements, and model generalization remains necessary to maximize real-world adoption.

The future of soil classification lies in explainable, multimodal, and integrated AI ecosystems capable of delivering accurate and sustainable soil intelligence across diverse environmental conditions.

REFERENCE

1. Abimala, T., S. Flora Sashya, and K. Sripriya. "Soil Classification & Crop Suggestion based on HSV, GLCM, Gabor Wavelet Techniques and Decision Tree Classifier in Image Processing." (2020).
2. Abraham, Shiny, Chau Huynh, and Huy Vu. "Classification of soils into hydrologic groups using machine learning." Data 5. 1 (2019)
3. Bioucas-Dias, J. M., et al. (2023). *Hyperspectral Remote Sensing Data Analysis and Future Perspectives*.
4. Chang, C. I. (2022). *Hyperspectral Imaging: Techniques for Spectral Detection and Classification*.
5. Dotto, A. C., et al. (2021). Hyperspectral remote sensing for soil property prediction.
6. Girme, Shivani, et al. 2023. "SOIL CLASSIFICATION USING CONVOLUTIONAL NEURAL NETWORK AND CROP RECOMMENDATION BASED ON SUPPORT VECTOR MACHINE
7. Li, J., et al. (2024). Deep learning approaches for hyperspectral image processing and classification.
8. Zhang, Y., et al. (2025). Advances in AI-driven hyperspectral image analysis for environmental applications.
9. Gorelick, N., et al. (2022). Remote sensing data platforms and applications in environmental monitoring.
10. Liang, S., et al. (2023). *Remote Sensing of Environment: Principles and Applications*.
11. Taghizadeh-Mehrjardi, R., et al. (2021). Satellite observations and digital soil mapping for precision agriculture.
12. Han, Xiao-Le, et al. "Deep learning based approach for the instance segmentation of clayey soil desiccation cracks." Computers and Geotechnics 146 (2022)
13. Wulder, M. A., et al. (2022). Advances in satellite earth observation systems.
14. Zhu, X. X., et al. (2024). Deep learning and satellite observation for environmental intelligence.