International Journal of Advanced and Applied Sciences, 11(8) 2024, Pages: 198-210

Contents lists available at Science-Gate

International Journal of Advanced and Applied Sciences

Journal homepage: http://www.science-gate.com/IJAAS.html

A sophisticated approach to soil productivity detection using a convolutional neural network-based model

Saikat Banerjee [1,*], Abhoy Chand Mandol [2]

[1]*Department of Computer Applications, Vivekananda Mahavidyalaya, Haripal, Hooghly, West Bengal, India*
[2]*Department of Computer Science, The University of Burdwan, Golapbag, West Bengal, India*

ARTICLE INFO

Article history:
Received 9 May 2024
Received in revised form
23 August 2024
Accepted 26 August 2024

Keywords:
Agricultural sector
Soil classification
Convolutional neural network
Soil properties
Image analysis

ABSTRACT

India is primarily an agricultural country where the quality of land is crucial for the livelihoods and well-being of its people. The agricultural sector plays a significant role in shaping the current state of the nation's economy. Therefore, it is essential to regularly evaluate our understanding of soil properties, such as its type, texture, color, and moisture content. Many developing countries lack sufficient knowledge and awareness about soil development. Understanding soil behavior helps farmers predict crop performance, monitor nutrient movement, and recognize soil limitations. Traditional methods for classifying soil in laboratories require significant time, staff, and financial resources. In this study, various image features, such as color, particle size, and texture, were randomly extracted and combined to predict soil fertility based on its sand, clay, and silt content using the AlexNet-CNN algorithm. We collected soil images using mobile cameras from regions such as Purulia, Hooghly, Bankura, and Burdwan to build a useful soil image dataset. The research focuses on categorizing productive and unproductive soil using convolutional neural network architectures, such as AlexNet and VGG16. Compared to previous studies, our proposed model showed better performance in terms of precision and recall. This study presents an efficient new convolutional neural network architecture for classifying soil images.

1. Introduction

Food is one of humanity's most essential requirements; farming is one way to meet that demand. India's agricultural sector would be pivotal to the country's economic future. India's economy has relied heavily on the farming industry for many years. About 15.5 percent of India's Gross domestic product comes from the farm sector. It could be a source of employment in many countries, including India. Forty percent or more of India's workforce works in agriculture (Banerjee et al., 2023). Because of the increased global population growth rate, the demand for food is also rising quickly. To keep up with the needs of the growing market, farmers must rely on more than the traditional methods of farming. The soil is the most crucial component of every agricultural area. Every farmer must take great care in choosing the best soil for their crops before planting. The ever-increasing need for food could be affected if this still needs to be done.

Consequently, reliable soil testing for agriculture is necessary. Among the most crucial tasks in farming is the conduct of soil tests. The data it gives about soil nutrients, like calcium (Ca), potassium (K), and nitrogen (N), is invaluable. Due to the ongoing drought, agricultural output and yields have declined in some parts of the Indian state of Maharashtra. The success rate is significantly lower than usual since the farmers need to prepare to provide fertilizers to the land. This means that they cannot repay the loan they used to fund their cultivation (Banerjee and Mondal, 2021). When they cannot repay the loan, they end their lives. More people have taken their own lives in India recently, and this could be a contributing factor. Many countries, India included, continue to employ the old way of farming. Maybe it is the high price tag, their lack of knowledge, or simply not knowing what good can come from employing modern technology, but farmers are hesitant to use it.

Problems arise at every agricultural production stage due to a general lack of knowledge. Pressure

* Corresponding Author.
Email Address: saikat.Banerjee56@gmail.com (S. Banerjee)
https://doi.org/10.21833/ijaas.2024.08.021
Corresponding author's ORCID profile:
https://orcid.org/0000-0002-7361-1553

on the farm sector is another consequence of the world's rapidly expanding population. Across the board, from crop selection to product marketing, today's losses are at an all-time high.

Consequently, specific methods that rely on computers to aid farmers are necessary. The most notable advancements in the computer business are deep learning (DL) and machine learning, along with other technologies like blockchain and the Internet of Things. Numerous fields have already used it, including robotics, metrology, economics, biology, and more. One of the many applications of machine learning in agriculture is the ability to forecast crop yields and quality and identify diseases and weeds. Artificial intelligence harvesting bots could one day replace human workers, cutting down on labor costs in the agricultural sector. DL can make it easier to monitor growing areas for factors such as crop humidity, soil composition, and temperature. To help farmers, we will apply machine learning and DL-based apps to solve the problems previously mentioned in the agriculture industry. An essential part of the activity that affects the development of agriculture is the technique of soil forecasting. Limestone reduces to mineral soil particles by weathering, fracturing, and transmission. These processes may progress at different rates depending on the strength of the forces acting upon them, which impacts the result. Soil, on the other hand, is more than just a bunch of different-sized mineral particles; it is a complicated mixture of organic matter, humus, and necessary minerals that have broken down in the presence of water and air, and it is home to all sorts of plant and animal life. Soil is more complex than the mere combination of mineral particles. An extremely long time is required for this mixture to reach the characteristic texture and necessary depth of the thing we have called "soil." Soil prediction makes use of numerous efficient technologies. Experts can also make educated guesses based on visual cues (Mondal and Banerjee, 2021).

DL is the way to go when it comes to soil prediction using images. Soil property prediction makes use of a plethora of time-honored methods. The CNN outperformed conventional techniques, such as the multi-tasking model, PLS regression, and Cubist regression trees. It was especially evident when compared with the outcomes produced by CNN. Using soil organic carbon content prediction as an example, the multi-task CNN outperformed PLS by 87% and Cubist by 62% in accuracy. Compared to the traditional back-propagation neural network, the accuracy of soil moisture predictions improved by 9.1%. Because of this, the theoretical basis for soil moisture prediction is improving.

The soil composition results from the interplay between the natural surroundings and anthropogenic actions. The primary focus of this study pertains to delineating the inherent characteristics of the Earth's surface, which exhibit distinct temporal and spatial attributes. Alterations in land cover can lead to climate and environmental modifications, significantly impacting socio-economic and ecological aspects (Banerjee et al., 2023). The predominant land cover comprises diverse soil types, encompassing arable lands, forests, meadows, and unvegetated terrains. Hence, the prompt and precise classification of various soil types is essential in land cover analysis, soil exploration, and cartography. The initial method of classification involves the utilization of a topographic map of land use, which is acquired through the amalgamation of data obtained from on-site surveys. Remote sensing image classification technology is predominantly employed to classify diverse soil types (Mondal and Banerjee, 2021). The employment of visible and near-infrared spectroscopy gadgets is a rapid and non-invasive technique for obtaining measurements. This substance has been extensively employed in various industries, such as medicine, agriculture, and oil, as evidenced by sources. The method of spectral analysis can obtain valuable information regarding a substance in an indirect manner. The outcome was achieved after a rigorous process of developing a proficient correction model linking the spectrum and information (Sharma et al., 2023). Implementing spectral technology in soil classification involves substituting remote sensing image data with spectral information, facilitating the creation of diverse soil models. The labeling of soil can be efficiently and non-invasively achieved promptly. Research has demonstrated that DL algorithms can accurately classify images to determine the presence or absence of nutrient-rich soil. DL is a computational approach that emulates the neural architecture of the human brain and has demonstrated significant advancements in various domains, including but not limited to image processing (Folorunso et al., 2023). The Support Vector Machine (SVM) is a widely utilized machine learning classification technique founded on statistical learning theory. The concept underlying SVM involves projecting input samples from a low-dimensional feature space to a high-dimensional space using nonlinear mapping.

Sahana and Mollah (2022) DL involved the progressive transformation of the input signal through multiple layers, resulting in a new feature space representation of the original signal. The system acquires a hierarchical feature representation through automatic learning, ultimately leading to the attainment of the classification outcome. CNN is a DL network architecture that has demonstrated efficacy in image classification, leading to widespread adoption across various domains (Waikar et al., 2020). The CNN technique has emerged as a promising and non-invasive approach for the quality assessment of agricultural commodities, encompassing the identification and classification of various produce such as fruits, vegetables, and others. Several studies have reported favorable outcomes from this method (Awais et al., 2023). CNN is commonly employed to develop classification models involving many samples. A CNN classification model was subsequently developed (Lanjewar and Gurav, 2022). The study analyzed the classification outcomes under varying label sample conditions and compared the results obtained through a shallow network SVM (Rani et al., 2023). This study evaluates the viability of using CNN to classify land cover with limited sample sizes. Additionally, this research aims to investigate novel approaches for swift, non-invasive, and precise soil classification based on the CNN methodology.

2. Literature survey

An analysis of CNN algorithms for the diagnosis of plant diseases was carried out by Abade et al. (2021). A total of 121 papers that were published between the years 2010 and 2019 were analyzed by the authors. During this review, it was determined that TensorFlow was the framework utilized the most frequently, while Plant Village was chosen as the dataset used the most frequently. To identify plant diseases through leaf photographs, Dhaka et al. (2021) provided an overview of the fundamental methodologies of CNN models. They also compared the available frameworks, pre-processing techniques, and CNN models. The datasets and performance measures utilized to evaluate the model's effectiveness are also investigated in this work. Van Klompenburg et al. (2020) conducted a systematic literature review on applying Machine Learning in predicting crop yield. The researchers determined that neural networks, particularly CNN, LSTM, and DNN, are primarily used for predicting crop productivity. Additionally, it was indicated that the quantity of characteristics varies depending on the investigation. In certain instances, the accuracy of yield forecast relies on object counting and detection rather than relying solely on tabular data.

The analysis of Dharani et al. (2021) about crop yield prediction via Deep Learning demonstrated that hybrid networks and the RNN-LSTM networks surpass all other networks in terms of performance. The superior performance of RNN and LSTM can be attributed to their efficient storage and feedback loop mechanisms. Their findings indicate that these networks possess greater predictive accuracy due to their ability to handle time-series data on crop yield.

Verma et al. (2022) proposed an improved method for identifying leaf diseases in plants. This method uses a feature computation technique based on Squeeze and Excitation (SE) Networks, which is used before processing the data using the original Capsule networks. SE-Alex-CapsNet achieves a higher accuracy of 92.1% with a picture size of 64X64, while Capsule Network only achieves an accuracy of 85.53%. The recommended methodology can be employed to create a mobile application optimized for low processing capabilities, specifically designed for low-cost cell phones and intended for use by farmers. The classification accuracies of six state-of-the-art CNN models are provided for comparison: AlexNet, SqueezeNet, ResNet50, VGG16, VGG19, and Inception V3.

In computer vision and related disciplines, image classification is vital in processing image depth for various purposes, such as uniform picture classification. It undergoes significant procedures, such as image processing, feature extraction, classifier building, and learning training (Wang et al., 2019). The primary focus of conventional image classification algorithms is to classify images based on fundamental features extracted from the pictures. It can establish the foundation for computers to extract semantic information from photos.

Lee et al. (2019) conducted research utilizing Deep Learning techniques to construct a crop yield platform that can predict itself, focusing specifically on crop diseases. The crop disease diagnostic module demonstrated that the CNN algorithm surpassed the RCNN and YOLO algorithms in performance. Utilizing a Rectified Linear Unit (ReLU) activation function in the artificial neural network yielded the most accurate results for the CYP module. Zhang et al. (2020) conducted a comprehensive analysis of the utilization of Deep Learning in dense agriculture scenarios. The study encompassed many applications such as recognition and classification, detection, counting, and yield estimate. The survey results demonstrated that Deep Learning surpasses other methods in densely populated environments. Extracting features is a time-consuming task to analyze and challenging to apply to image classification. Concurrently, it is not feasible to handle large datasets using traditional machine learning methods; improving the design of features, selecting relevant features, and training the model is also tricky, resulting in unsatisfactory classification outcomes. As a result, various fields of application are affected by picture categorization algorithms that depend on conventional machine learning (Ding et al., 2019). Research has shown the capability to utilize fundamental, elementary characteristics for image classification and identification, including texture, shape, and color. Traditional methods commonly utilize SVM to categorize photos, which supplies either a solitary characteristic or an assemblage of characteristics. Artificial neural network classifiers have significantly progressed in picture classification in recent years. Standardizing the arrangement of fundamental components such as color, shape, and texture helps enhance the precision of picture categorization. DL uses a multilayer network architecture to train on extensive datasets, utilizing a method known as layer-by-layer feature extraction to obtain the image's higher-level characteristics. A DL network model utilizes multiple hidden layers to extract an image's fundamental properties and intricate aspects. DL yields superior and more precise image categorization characteristics than traditional machine learning methods.

The study demonstrated that traditional techniques like Gaussian Mixture Models are more effective than modern approaches such as U-Net, Faster R-CNN, and CNN for yield mapping (Koirala et al., 2019). comprehensively analyzed by applying

Deep Learning techniques in fruit counting and yield estimation. The study demonstrated the efficacy of deep learning techniques in extracting significant characteristics and suggested the utilization of CNN detectors, deep regression, and LSTM to calculate the fruit load accurately.

In their study, Deorankar and Rohankar (2020) proposed using hyperspectral data to classify soil texture. They develop and deploy three 1-dimensional (1D) convolutional neural networks (CNN): the Lucas CNN, the LucasResNet, which incorporates an identity block as a residual network, and the LucasCoordConv with an extra layer for coordinates. In addition, we adapt two pre-existing 1D CNN techniques for the classification job at hand. They assess the performance of the CNN techniques and contrast them with a random forest classifier. Consequently, we depend on the readily available LUCAS topsoil dataset. The CNN technique with the lowest number of layers is the most effective classifier. The LucasCoordConv achieves optimal performance in terms of average accuracy.

This literature review comprehensively analyses how deep learning methods are used to forecast crop yields using remote sensing data. This study aims to conduct a thorough literature analysis to identify the current research gaps in a specific field of deep learning techniques. The review will also guide the study of the influence of vegetation indices and environmental conditions on crop development. We evaluate and incorporate research studies from journals, conferences, and internet databases in the systematic literature review. These studies are then organized and presented in relation to the research topics indicated in our study. The literature study reveals that CNNs are predominantly utilized in image recognition. It encompasses a wide range of areas within the subject of image categorization. Based on the current research, it has been determined that agricultural picture classification remains a difficult task. Based on the analysis of past research, it was also discovered that there is a lack of imagery depicting productive soil in agriculture. Therefore, we have generated a dataset specifically for our study. Our goal is to aid farmers by offering them the necessary help and expertise to improve the quality of their agricultural harvests and their financial situation. This research aims to categorize soil types by analyzing photos captured using mobile cameras in various locations such as Hooghly, Burdwan, Purulia, Bankura, and others. This research utilizes contemporary convolutional neural network (CNN) architecture. Subsequently, the proposed methodology will also be compared with existing studies.

3. Method

3.1. Pre-processing

A low-pass filter removes high-frequency noise and artifacts from the image. Smoothing filters utilize an evolving window operator to modify the value of individual pixels within an image. This adjustment depends on a function of the surrounding local area of pixels. As the operator progresses across the image, all pixels are impacted. The smoothing filter gradually improves image quality by removing imperfections. The term "soil" denotes a mixture of small rock particles and organic humus that develops on the Earth's surface, creating a favorable habitat for plant growth. The importance of soil in promoting economic and social advancement is of utmost significance. Facilitates the sustenance of human, animal, and plant life by supplying food, fodder, and renewable energy. The soil classification in India can be traced back to ancient times, albeit with less comprehensiveness compared to contemporary classifications. In the ancient period, soil classification was determined exclusively by two factors: the soil type and its fertility. The planned research study is divided into individual steps and depicted in a step-by-step style in Fig. 1.

Regions such as West Bengal, Jharkhand, Bihar, Uttar Pradesh, Haryana, and Punjab are commonly characterized by alluvial soil. This soil exhibits notable proportions of silt, sand, and clay. Usually, a blend of potash, phosphoric acid, lime, and organic matter is incorporated into these soils. The land in question has the potential to yield crops such as rice, wheat, jute, and potatoes. Furthermore, it is distinguished by a dark hue. The Deccan plateau in India harbors black soils widely dispersed across multiple states, namely Maharashtra, Madhya Pradesh, parts of Karnataka, Andhra Pradesh, Gujarat, and Tamil Nadu. The soil in question is known as regur soil, carpus soil, or black cotton soil. The soil in question displays a notable presence of titanium oxide. Usually, it combines with lime, iron, magnesium, and potash. This soil type is highly suitable for the cultivation of Carpus cotton. Various crops, including tobacco, wheat, sorghum, and oilseeds, are cultivated. Red soil is a commonly occurring soil type in several regions of India, such as Tamil Nadu, the eastern parts of Andhra Pradesh, the Chotonagpur region, and the Bankura and Birbhum districts of West Bengal. The genesis of these soils ascribes to the occurrence of granite, gneiss, and quartzite rocks dating back to the Archaean era. The distinctive red color of red soil is attributed to its high iron content. This soil has been traditionally regarded as infertile. However, crops such as groundnut, potato, maize, paddy, wheat, dal, oilseed, orange, and vegetables are cultivated through chemical fertilizers. The laterite soil is a commonly occurring soil type in the high-altitude areas of the Western Ghats, Purghats, Rajmahal Hills, Assam, and Meghalaya. These soil types are commonly encountered in regions with high temperatures and significant rainfall. The soil analysis displays the existence of iron oxide, various organic substances, phosphate, and calcium. The soil's water-holding capacity is significantly low due to the plentiful presence of gravel. It is important to note that laterite soils are highly suitable for cultivating cashew nuts. Apart from other agricultural produce, crops such as cotton, tea, coconut, and pulses are also cultivated. The predominant location of desert soils is in the desert regions of Rajasthan, northern Gujarat, western Haryana, and southwestern Punjab. The soil in question is categorized as chernozem soil. The soil in question exhibits hues of either brownish yellow or light yellow. A sandy texture and salinity typically characterize the soils in question. The soils in question are typically considered unsuitable for agricultural use due to their limited ability to sustain crop growth. Bihar, Uttar Pradesh, Haryana, Punjab, the Sundarbans area, and Rann of Kutch are characterized by saline and alkaline soil. Multiple names, including Re, Ushar, Kannar, and Thur, know this specific soil type. The soil analysis displays the occurrence of sodium, potassium, and magnesium. The soil's agricultural viability is considered insufficient. Peat and wetland soils are commonly found in the coastal regions of West Bengal, Odisha, Tamil Nadu, and the coastline of Kerala. These soil types are commonly found in areas with high precipitation and humidity. Organic matter is present in the soil.

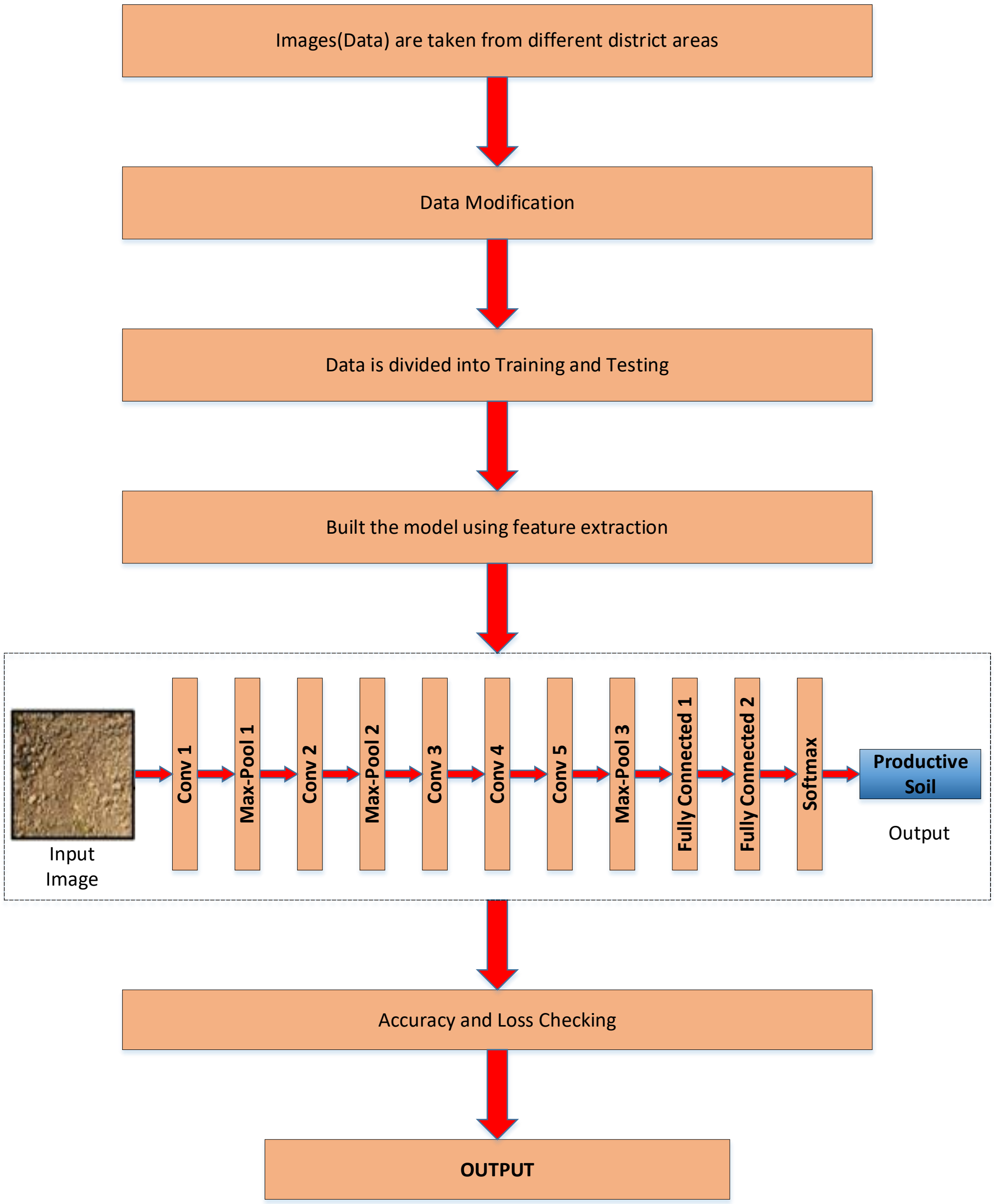

Fig. 1: Step-by-step workflow of the proposed research work

The soil in question is known to be suitable for cultivating paddy and jute crops. Soils in mountainous and forested regions are predominantly situated on inclines that experience elevated precipitation levels. The soils under consideration display a significant level of organic matter concentration. The soil in question must display more potash, phosphorus, and lime. The soil in question possesses inherent infertility. Although the soil may appear infertile, it cultivates crops such as tea, coffee, medicinal plants, and various fruits.

We are in the process of creating a dataset consisting of productive soil samples. To accomplish this, we are gathering datasets of unproductive soil from various sources such as Google, Kaggle, and other relevant platforms. Fig. 2 displays a selection of unproductive soil photographs. The regions of Hooghly, Burdwan, Purulia, and Bankura are renowned for their significant contributions to grain production. Soil images have been collected from the said locations. The databases have been acquired utilizing various devices, including DSLRs and mobile phone cameras. Soil images were acquired using an Asus Zenfone Max M2 smartphone with a 16-megapixel camera. The camera settings remain at their default values when capturing images, including picture quality set to medium, countdown timer turned off, exposure time set to 1/60 S, and anti-banding *set* to auto. The database was generated utilizing a Redmi Note 7 Pro smartphone equipped with a 48-megapixel camera capable of capturing images at a resolution of 4160 by 3120 pixels under visible light conditions. Fig. 3 displays a selection of productive soil photographs. The images were captured at 1 meter from the soil to the camera to mitigate the impact of uncontrolled illumination sources. The training dataset has 31,600 photos, whereas the testing dataset comprises 6,320 images.

Fig. 2: Images of unproductive soil

Fig. 3: Images of productive soil

3.1.1. Feature extraction

The stage of feature extraction is vital in the process. The extracted data includes all the necessary characteristics for soil classification, including texture, color, and intensity. Consequently, a color moments metric is implemented to differentiate images based on their color attributes. The physical qualities of soil encompass various factors such as hydraulic conductivity, soil texture, stoniness, soil structure, bulk density, soil crusting, soil color, soil compaction, soil infiltration, soil retention, soil drainage, and topsoil depth. These qualities impact infiltration, erosion, nutrient cycling, and biological activity. These attributes also influence the appropriateness of soil for many purposes, such as its ability to absorb stormwater, its usefulness as a foundation for roadways, and its strength for construction.

This part offers a comprehensive introduction to soil physical properties, their associated processes, and the impact of human activities on these properties. It also provides a detailed analysis of soil image categorization about stormwater applications and more resources on sampling, testing, and soil productivity evaluations. Fig. 4 depicts a histogram, the predominant graphical representation for illustrating frequency distributions. A histogram displays the frequency distribution of soil image parameter data points for different classes. Histograms are effective for showing the overall distributional characteristics of soil dataset features. One can observe the approximate locations of the highest points in the distribution, determine whether the distribution is skewed or symmetrical, and identify any outliers.

Hydraulic conductivity measures the speed at which water moves through the empty spaces in the soil. It depends on factors such as the texture, size, and arrangement of the grains, density, and overall soil structure. Image texture is the presence of patterns created by differences in contrast and the unevenness of natural surfaces. These patterns result from various attributes such as roughness, depth, lighting, and color. Texture analysis in computer vision focuses on extracting features, segmenting images, classifying objects, synthesizing textures, and inferring shapes based on texture. By combining textural and geometric data, the integrated approach produced more significant segmentations than traditional methods and improved outcomes for the classic watershed technique. The evaluation uses visual inspection and quality metrics for soil pictures with intricate structures.

The bulk density serves as a reliable measure of soil compaction. The calculation involves dividing the weight of soil without moisture by its volume. This volume encompasses the volume occupied by soil particles and the space between the particles. High bulk density is a reliable sign of reduced soil porosity and increased soil compaction. It can impede root growth and hinder air and water circulation in the soil. Compaction can lead to superficial plant root growth and hinder plant development, negatively impacting crop productivity and decreasing the amount of vegetation covering the soil, diminishing its ability to prevent erosion. Compaction can cause a decrease in water infiltration into the soil, which in turn can increase runoff and erosion. It could be more problematic in locations with sloping ground or water-saturated soils. The significance of texture analysis is the segmentation performance achieved by the decomposition of image texture.

Using texture modulation energy enhances the differentiation of soil texture for classification purposes. It effectively utilizes segmented soil regions by utilizing precomputed characteristics. The annotated photos will be analyzed using geometric, statistical, and textural metrics to map the soil's characteristic attributes. This mapping evaluates the bioecological quality of the soil. A smartphone camera obtains photos of the desiccated soil samples in an imaging setup. The collected photos undergo pre-processing techniques such as RGB extraction, V extraction from HSV bins, and adaptive histogram to enhance the texture of the soil images. Light-SoilNet is a new CNN designed to categorize five types of soil samples: sand, clay, loam, loamy sand, and sandy loam.

Compared with lightweight and pre-trained DL networks, which are currently considered the most advanced, the suggested network undergoes testing. Soil crusting significantly impacts soil production and function. It results in a lower infiltration rate, a decrease in soil moisture, an increase in runoff, and a decrease in groundwater recharge. The hydraulic conductivity of a crust can be significantly lower, up to a hundred or even a thousand times, compared to the underlying soil. Crusting hinders or prevents the sprouting of seedlings. "Soil retention measures" refer to any method or building that keeps dirt from washing away or spreading beyond a designated area. Shoring excavated sections with wood, concrete, or steel structures is the most frequent method. It helps to reduce steep slopes. Design engineers can implement structural soil-retaining solutions to prevent or mitigate soil erosion and ensure the safety of excavation workers. The process by which water seeps into the soil is known as infiltration. The infiltration rate is the rate at which water seeps into the ground. The standard units of measurement are inches per hour. Water, whether from rain or irrigation, must soak into the soil to be helpful. The spatial arrangement of the estimated soil properties for pixels representing bare soils emphasizes, indicating potential enhancements to spatial prediction models for these qualities through the application of Digital Soil Mapping techniques.

3.2. AlexNet

The input to AlexNet is an RGB image of size 227 x 227 (Fig. 5). This means all images in the training set and all test images need to be size 227 x 227. If the input image is not 227 x 227, it converts to 227 x 227 before training the network. The smaller dimension resizes to 227, and the resulting image is cropped to obtain a 227 x 227 image. If the input image is grayscale, it converts to an RGB image by replicating the single channel to obtain a 3-channel RGB image. AlexNet was much larger than previous CNNs used for computer vision tasks. It has 60 million parameters and 650,000 neurons, and it took five to six days to train on two GTX 580 3GB GPUs.

AlexNet consists of 5 Convolutional Layers and 3 Fully Connected Layers. Multiple convolutional kernels extract exciting features in an image. In a single convolutional layer, there are usually many kernels of the same size. For example, the first conv layer of AlexNet contains 96 kernels of size 11 x 11 x 3. The width and height of the kernels are usually the same, and the depth is the same as the number of channels. Max Pooling layers are used to downsample the width and height of the tensors, keeping the depth the same. Overlapping Max Pool layers are similar to the Max Pool layers, except the adjacent windows over which the max is computed overlap. ReLU Nonlinearity: The AlexNet architecture incorporates a significant aspect, namely the utilization of ReLU (Rectified Linear Unit) Nonlinearity. Previously, the conventional approach to training a neural network model utilized either the Tanh or sigmoid activation functions. The AlexNet study demonstrated that ReLU nonlinearity in deep CNNs resulted in significantly faster training times than saturating activation functions such as tanh or sigmoid. The data presented in Fig. 5 indicates that ReLUs (solid curve) utilization resulted in AlexNet achieving a training error rate of 25%, six times faster than a comparable network utilizing tanh (dotted curve). The ReLU function is given by $f(x) = \max(0, x)$.

If an image of soil is present in our training set, its mirror image can also be considered a legitimate representation of soil. Please refer to Fig. 6 for an illustrative example. Flipping the image along the vertical axis can increase the size of the training dataset.

3.3. VGG16

The VGG16 model was determined to be the most effective model on the image dataset. Fig. 6 examines the current architecture of this arrangement for our soil image classification. The input for any of the network setups is a 224 by 224 image with three channels - red (R), green (G), and blue (B) - of a predetermined size. The sole pre-processing conducted involves standardizing the RGB values for each pixel. It accomplishes this by removing the average value from each pixel. The image is processed by the initial stack of two convolutional layers with a tiny receptive size of 3 x 3, then activated using the Rectified Linear Unit (ReLU) function. Both layers consist of 64 filters each. Both the convolution stride and the padding initialize to 1 pixel. This arrangement maintains the spatial resolution, and the output activation map has the exact dimensions as the input image. The activation maps are subsequently subjected to spatial max pooling using a 2 x 2-pixel window and a stride of 2 pixels. It reduces the size of the activations by half. Therefore, the activations' dimensions at the initial stack's conclusion are 112 x 112 x 64.

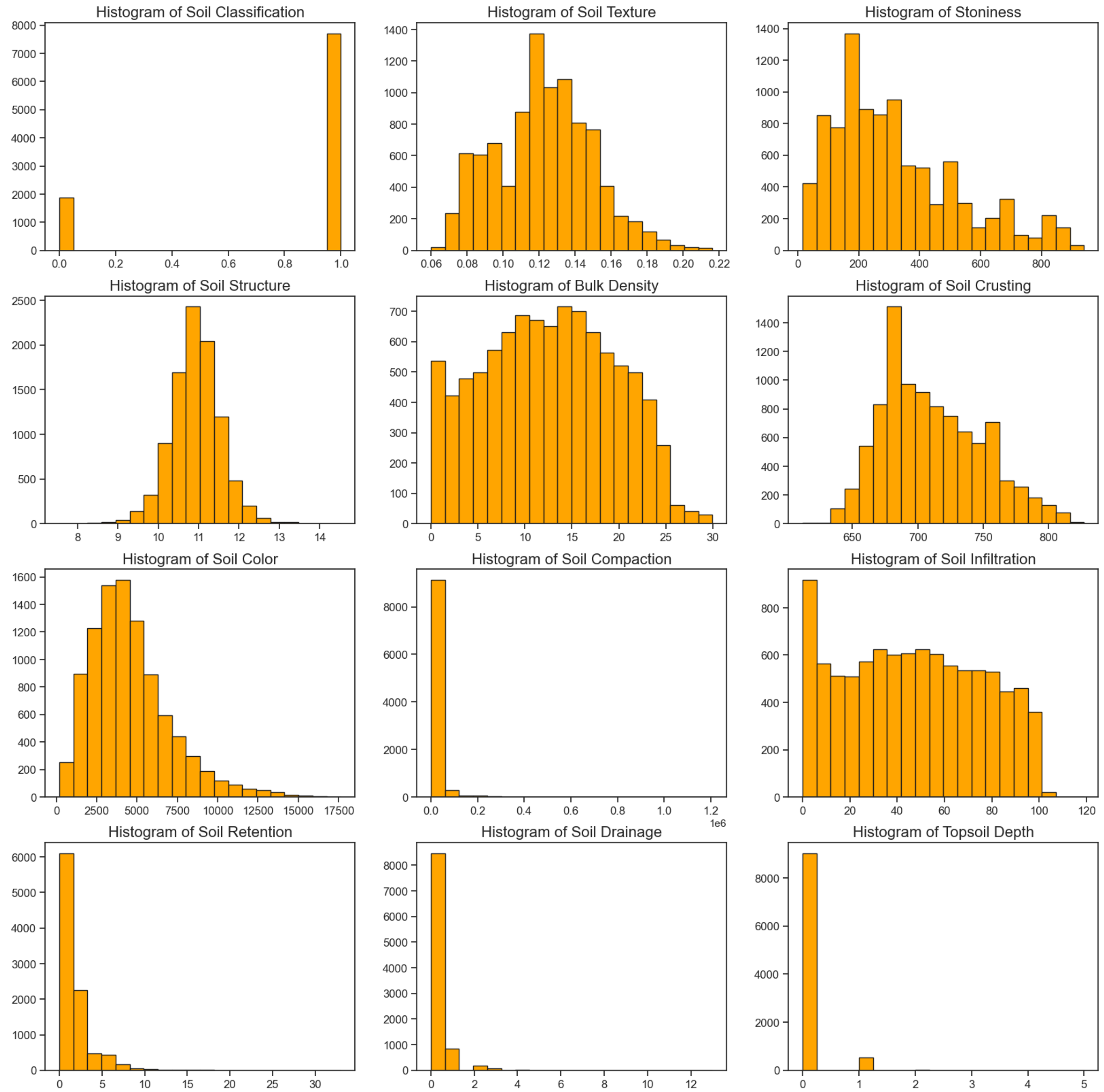

Fig. 4: Histogram of soil parameter

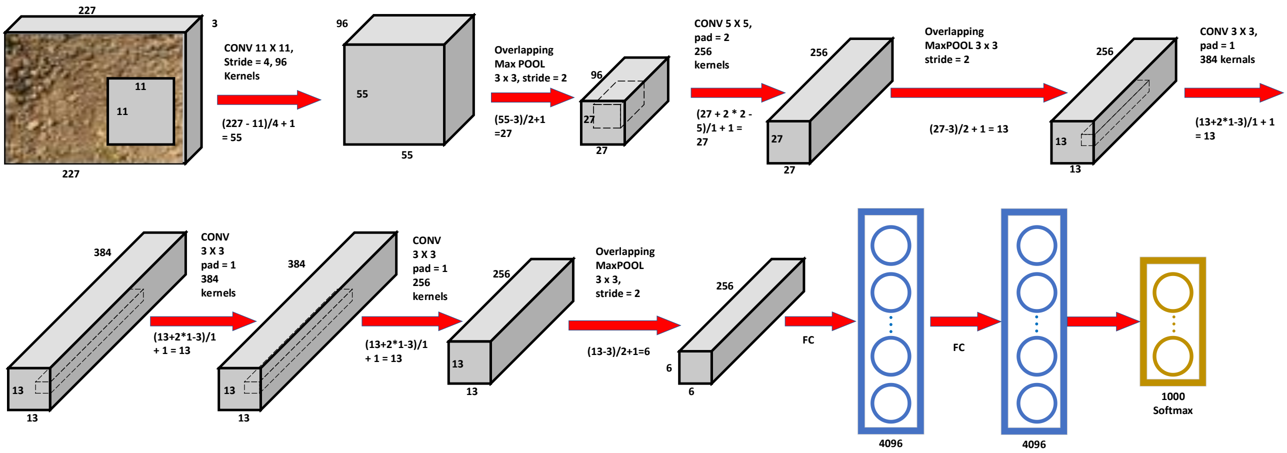

Fig. 5: AlexNet architecture

The activations subsequently pass through a second stack identical in structure but with an increased number of filters. Specifically, the second stack contains 128 filters, compared to the 64 filters in the first stack. Therefore, the dimensions of the stack after the second iteration are 56 x 56 x 128.

Next is a third stack consisting of three convolutional layers and a max pool layer. The number of filters employed in this case is 256, resulting in a stack output size of 28 x 28 x 256—subsequently, two sets of three convolutional layers, each including 512 filters. The result of both stacks will be seven multiplied by seven multiplied by 512. Following the convolutional layers, a flattening layer sits between three fully linked layers. The initial two layers consist of 4,096 neurons each, while the final fully connected layer functions as the output layer and has 1,000 neurons corresponding to the 1,000 potential classes in the dataset. The Softmax activation layer is applied after the output layer to facilitate categorical classification. The proposed model architectures are presented in Table 1.

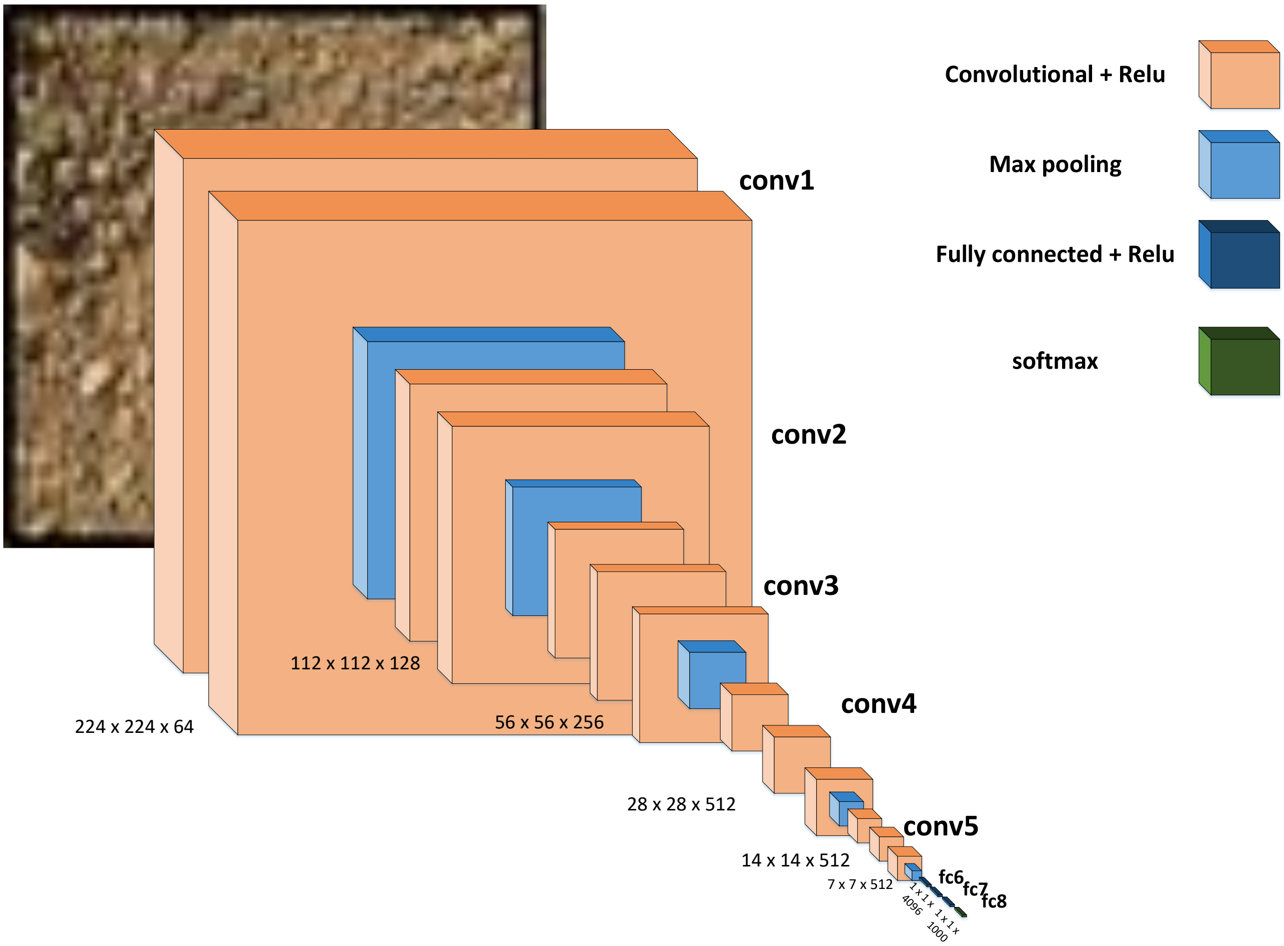

Fig. 6: VGG-16 architecture

4. Experimental setup

This experiment aims to improve price prediction by utilizing two DL approaches. The study guides by compiling a dataset using agricultural data. The experiment uses an Intel CoreTM i5-M520 CPU with a clock speed of 2.40 GHz and 4 GB of RAM. We have selected Python, Keras, TensorFlow, Numpy, and the Pandas library as tools for conducting experiments and achieving practical results. Google Colaboratory, based on the Jupyter Notebook, was utilized for all the training trials in this work. This notebook offers user-friendly libraries, visualization capabilities, and tools for integrating data. This software is free of charge and enables the execution and distribution of Python programs. The primary objective of this platform is to facilitate the advancement of machine learning research and education. This platform is compatible with high-performance hardware such as Parallel Tensor Processing Units (TPUs) and Graphic Processing Units (GPUs). The research utilized the TensorFlow API, which enables the construction of VGG16 and AlexNet models, which are deep neural network architectures. The experiment's dataset consists of 37,923 photos, 80% allocated for training and 20% for the test dataset.

5. Results and discussions

Fig. 7 and Fig. 8 show that the accuracy gradually rises to more than 95 percent, while the loss decreases significantly by the 40th epoch. Around the 25th epoch, there is a peculiar decrease in precision even though the loss gradually and smoothly improves. For forecasting class probabilities, the loss function utilized here is categorical cross-entropy. There is evidence that AlexNet has made progress in terms of loss and accuracy. After one hundred iterations, the accuracy of the model's predictions is 95%, and the loss value is also decreasing. As a result of the relatively small batch size, there were minor variances in the curve that represented the validation procedure. In comparison to the accuracy of the other CNN models (Yadav et al., 2024), the VGG16 model needs to be more accurate. Fig. 9 and Fig. 10 display the loss plots achieved using the AlexNet and VGG16 architectures, respectively.

Table 1: Proposed model architectures

VGG16			AlexNet		
Layer (type)	Output Shape	Param #	Layer (type)	Output Shape	Param #
conv1_1 (Conv2D)	(None, 224, 224, 64)	1792	conv2d (Conv2D)	(None, 55, 55, 96)	34944
conv1_2 (Conv2D)	(None, 224, 224, 64)	36928	max_pooling2d (MaxPooling2D) (	(None, 27, 27, 96)	0
max_pooling2d_1 (MaxPooling2	(None, 112, 112, 64)	0	conv2d_1 (Conv2D)	(None, 27, 27, 256)	614656
conv2_1 (Conv2D)	(None, 112, 112, 128)	73856	max_pooling2d_1 (MaxPooling2	(None, 13, 13, 256)	0
conv2_2 (Conv2D)	(None, 112, 112, 128)	147584	conv2d_2 (Conv2D)	(None, 13, 13, 384)	885120
max_pooling2d_2 (MaxPooling2	(None, 56, 56, 128)	0	conv2d_3 (Conv2D)	(None, 13, 13, 384)	1327488
conv3_1 (Conv2D)	(None, 56, 56, 256)	295168	conv2d_4 (Conv2D)	(None, 13, 13, 256)	884992
conv3_2 (Conv2D)	(None, 56, 56, 256)	590080	max_pooling2d_2 (MaxPooling2)	(None, 6, 6, 256)	0
conv3_3 (Conv2D)	(None, 56, 56, 256)	590080	flatten (Flatten)	(None, 9216)	0
max_pooling2d_3 (MaxPooling2	(None, 28, 28, 256)	0	dense (Dense)	(None, 4096)	37752832
conv4_1 (Conv2D)	(None, 28, 28, 512)	1180160	dropout (Dropout)	(None, 4096)	0
conv4_2 (Conv2D)	(None, 28, 28, 512)	2359808	dense_1 (Dense)	(None, 4096)	16781312
			dropout_1 (Dropout)	(None, 4096)	0
			dense_2 (Dense)	(None, 1000)	4097000

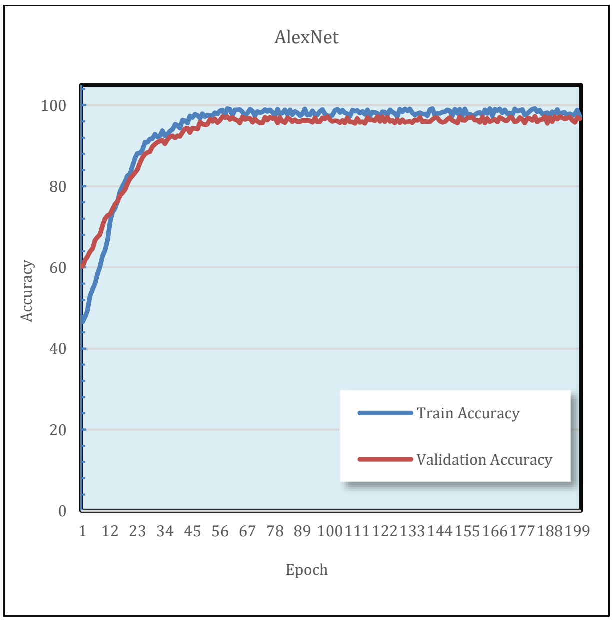

Fig. 7: Accuracy of AlexNet

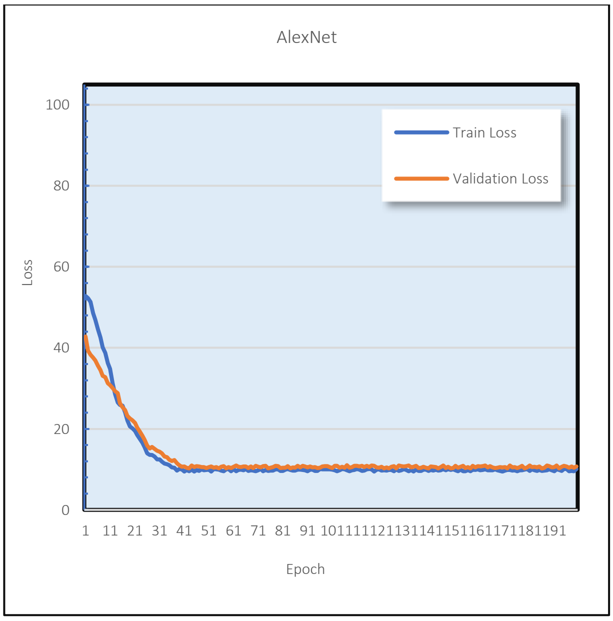

Fig. 9: Loss of AlexNet

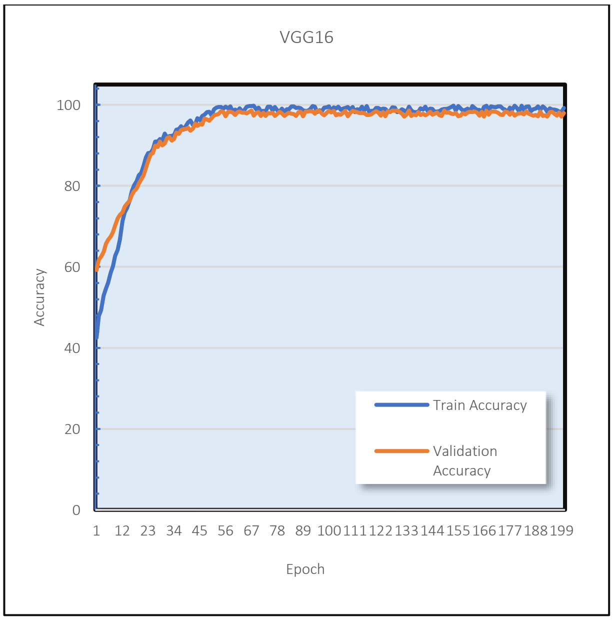

Fig. 8: Accuracy of VGG16

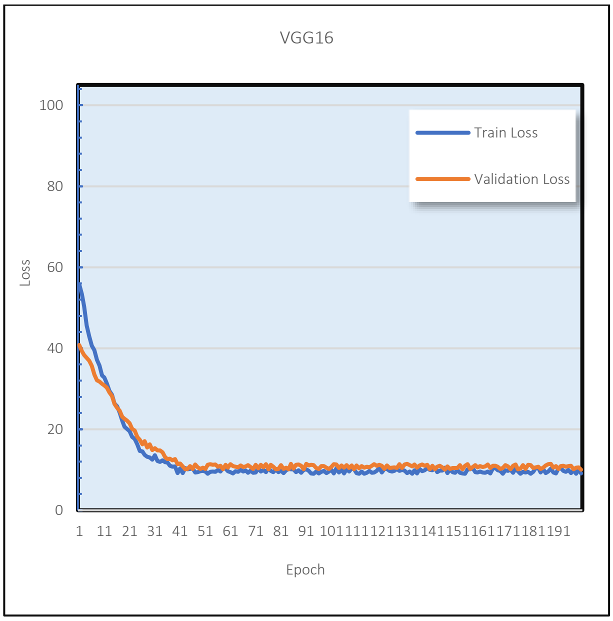

Fig. 10: Loss of VGG16

One way to determine precision is to add all the positive and negative results and divide the total by the number of true positives. This idea is also known as positive predictive value. This ratio of correct guesses to total predictions is known as "recall." Our model's "recall" is the proportion of correctly identified and classified instances. We must use the recall measure to determine the optimum model when the cost of making a false negative exceeds a false positive. An F1 score is required. As we have seen, the fraction of appropriately detected positives and negatives is accuracy.

$$Precision = TP/(TP + FP) \times 100 \quad (1)$$
$$Recall = \frac{True\ Positives}{TP + FN} \times 100 \quad (2)$$
$$F1\ score = 2 * \frac{Precision * Recall}{Precision + Recall} \times 100 \quad (3)$$
$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP} \times 100 \quad (4)$$

What follows is an examination of the "F1 Score," a tool for gauging the dataset's available accuracy. False positive and negative results can damage businesses in terms of money and other intangibles, yet we tend to overlook a lot of different considerations when making decisions. In this case, the F1 score—which aims to balance accuracy and recall—may be the better metric. True Negatives (TN), False Negatives (FN), False Positives (FP), and True Positives (TP) are the terms used here (Banerjee and Mondal, 2023).

This study presents an authentic dataset collected from various areas in West Bengal. Comparing the results with earlier research is challenging because we needed help identifying relevant research articles on soil image classification. Our research surpasses earlier studies in picture classification, demonstrating superior performance. Compared to other models, AlexNet's superior prediction performance stands out. The precision and recall values of AlexNet significantly surpass those of other models, underscoring its effectiveness. The learning rate of other models was found to enhance their accuracy and precision values. According to Table 2, the AlexNet design offers the highest level of training data accuracy, a remarkable 98.97 percent.

In contrast, the VGG-16 design achieved 97.38 percent, and the testing data accuracy offered 97.92 and 96.93 percent, respectively. Similarly, the AlexNet architecture achieves a precision of 98.25%, surpassing the 97.68% achieved by the VGG-16 design and the 80% achieved by the Region-based CNN and You Only Look Once (YOLO) architecture (Yadav et al., 2024). Additionally, an F1-Score of 97.86% was achieved by utilizing AlexNet, in contrast to 97.11% when utilizing VGG-16.

Table 2: Evaluating the performance of various network models

Model	Training data accuracy	Testing data accuracy	Training loss	Testing loss	Recall	Precision	F1-score
AlexNet	98.97	97.92	0.48	0.78	99.31	98.25	97.86
VGG16	97.38	96.93	0.65	0.92	98.27	97.68	97.11
Region-based CNN (Yadav et al., 2024)	80.11	79.12	0.1037	0.1041	74.64	99.08	-
YOLO (Yadav et al., 2024)	79.23	78.56	0.01512	9.1793×10^{-3}	97.23	96.31	-
Single shot detector (Yadav et al., 2024)	78	77.57	0.1951	0.2708	66.93	95.26	-

By examining the chart that displays the accuracy and loss of the model, it becomes evident that the VGG-16 model achieves an accuracy of 97.38% and a loss of 0.65. In comparison, the Region-based CNN model achieves an accuracy of 80% and a loss of 0.1037. The application of the AlexNet model results in an accuracy of 98.97% and a loss of 0.48, which is lower than that of VGG-16. Figs. 7 and 8 comprehensively compare the accuracy and loss experienced by three different models, while Fig. 11 offers an in-depth analysis of the models' performance. These precision values underscore the reliability and confidence in the results obtained from the models. Our suggested model performed satisfactorily compared to past research regarding recall and precision.

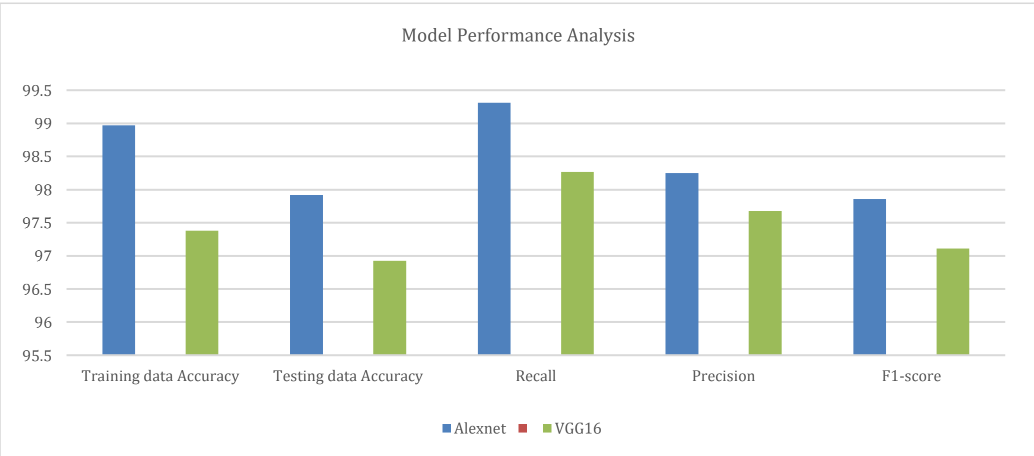

Fig. 11: Model performance

6. Conclusion

When contrasted with more conventional, labor-intensive manual approaches, soil classification using deep learning can accomplish more in less time. Deep learning models can quickly process massive amounts of soil photos and accurately extract complex patterns that could be difficult for human specialists to detect. Soil characteristics and spatial variability can be better understood using deep learning models that accurately classify soil. Soil management approaches, precision agriculture, and land use planning can all benefit from this data, which means more crops per acre and less waste. Soil classification using deep learning can be integrated into educational programs to bring attention to the significance of soil and the need to conserve it. Due to the opportunities it presents to educate farmers and landowners about their soil resources, there will undoubtedly be a flurry of new developments. Soil image detection approaches based on DL streamline the process of detecting plant diseases and pests by combining multiple processes and connections into a single end-to-end feature extraction. This approach offers significant development prospects and has enormous promise. Despite the rapid development of soil image classification technology, its transition from academic research to practical agricultural use is still in progress. Unresolved issues remain. This has yet to reach a mature use stage in real natural environments.

In our future research, we will focus on resolving the issues related to real-time data gathering and creating a sophisticated DL model capable of identifying soil production in a specific area of agricultural land. Moreover, we are actively striving to integrate a mobile application with the trained model derived from this project. The real-time soil fertility monitoring system will assist farmers and benefit the agriculture sector. Future Work's objective is to develop an application that utilizes an image of soil as input and accurately forecasts whether the soil is fertile or infertile. Therefore, we will generate a comprehensive soil image dataset encompassing several districts, with a minimum of 400000 photos.

Compliance with ethical standards

Conflict of interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

Abade A, Ferreira PA, and de Barros Vidal F (2021). Plant diseases recognition on images using convolutional neural networks: A systematic review. Computers and Electronics in Agriculture, 185: 106125. https://doi.org/10.1016/j.compag.2021.106125

Awais M, Naqvi SM, Zhang H, Li L, Zhang W, Awwad FA, Ismail EA, Khan MI, Raghavan V, and Hu J (2023). AI and machine learning for soil analysis: An assessment of sustainable agricultural practices. Bioresources and Bioprocessing, 10(1): 90. https://doi.org/10.1186/s40643-023-00710-y **PMid:38647622 PMCid:PMC10992573**

Banerjee S and Mondal AC (2021). Nutrient food prediction through deep learning. In the Asian Conference on Innovation in Technology, IEEE, Pune, India: 1-5. https://doi.org/10.1109/ASIANCON51346.2021.9545014

Banerjee S, Ckakraboity S, and Mondal AC (2023). Machine learning based crop prediction on region wise weather data. International Journal on Recent and Innovation Trends in Computing and Communication, 11(1): 145-153. https://doi.org/10.17762/ijritcc.v11i1.6084

Deorankar AV and Rohankar AA (2020). An analytical approach for soil and land classification system using image processing. In the 5th International Conference on Communication and Electronics Systems, IEEE, Coimbatore, India: 1416-1420. https://doi.org/10.1109/ICCES48766.2020.9137952

Dhaka VS, Meena SV, Rani G, Sinwar D, Ijaz MF, and Woźniak M (2021). A survey of deep convolutional neural networks applied for prediction of plant leaf diseases. Sensors, 21(14): 4749. https://doi.org/10.3390/s21144749 **PMid:34300489 PMCid:PMC8309553**

Dharani MK, Thamilselvan R, Natesan P, Kalaivaani PC D, and Santhoshkumar S (2021). Review on crop prediction using deep learning techniques. Journal of Physics: Conference Series, 1767(1): 012026. https://doi.org/10.1088/1742-6596/1767/1/012026

Ding M, Sun Z, Wei L, Cao Y, and Yao Y (2019). Infrared target detection and recognition method in airborne photoelectric system. Journal of Aerospace Information Systems, 16(3): 94-106. https://doi.org/10.2514/1.I010655

Folorunso O, Ojo O, Busari M, Adebayo M, Joshua A, Folorunso D, Ugwunna CO, Olabanjo O, and Olabanjo O (2023). Exploring machine learning models for soil nutrient properties prediction: A systematic review. Big Data and Cognitive Computing, 7(2): 113. https://doi.org/10.3390/bdcc7020113

Koirala A, Walsh KB, Wang Z, and McCarthy C (2019). Deep learning–Method overview and review of use for fruit detection and yield estimation. Computers and Electronics in Agriculture, 162: 219-234. https://doi.org/10.1016/j.compag.2019.04.017

Lanjewar MG and Gurav OL (2022). Convolutional neural networks based classifications of soil images. Multimedia Tools and Applications, 81(7): 10313-10336. https://doi.org/10.1007/s11042-022-12200-y

Lee S, Jeong Y, Son S, and Lee B (2019). A self-predictable crop yield platform (SCYP) based on crop diseases using deep learning. Sustainability, 11(13): 3637. https://doi.org/10.3390/su11133637

Mondal A and Banerjee S (2021). Effective crop prediction using deep learning. In the International Conference on Smart Generation Computing, Communication and Networking, IEEE, Pune, India: 1-6. https://doi.org/10.1109/SMARTGENCON51891.2021.9645872

Rani S, Mishra AK, Kataria A, Mallik S, and Qin H (2023). Machine learning-based optimal crop selection system in smart agriculture. Scientific Reports, 13(1): 15997. https://doi.org/10.1038/s41598-023-42356-y **PMid:37749111 PMCid:PMC10520008**

Sahana T and Mollah AF (2022). An effective pipeline for depth image-based hand gesture recognition. In: Das AK, Nayak J, Naik B, Vimal S, and Pelusi D (Eds.), Computational intelligence in pattern recognition. CIPR 2022. Lecture notes in networks and systems, 725: 489-503. Springer, Singapore, Singapore. https://doi.org/10.1007/978-981-99-3734-9_40

Sharma P, Dadheech P, and Senthil ASK (2023). AI-enabled crop recommendation system based on soil and weather patterns. In: Gupta RK, Jain A, Wang J, Bharti SK, and Patel S (Eds.), Artificial intelligence tools and technologies for smart farming and agriculture practices: 184-199. IGI Global, Hershey, USA. https://doi.org/10.4018/978-1-6684-8516-3.ch010

Van Klompenburg T, Kassahun A, and Catal C (2020). Crop yield prediction using machine learning: A systematic literature review. Computers and Electronics in Agriculture, 177: 105709. https://doi.org/10.1016/j.compag.2020.105709

Verma S, Chug A, Singh RP, Singh AP, and Singh D (2022). SE-CapsNet: Automated evaluation of plant disease severity based on feature extraction through squeeze and excitation (SE) networks and capsule networks. Kuwait Journal of Science, 49(1): 1-31. https://doi.org/10.48129/kjs.v49i1.10586

Waikar VC, Thorat SY, Ghute AA, Rajput PP, and Shinde MS (2020). Crop prediction based on soil classification using machine learning with classifier ensembling. International Research Journal of Engineering and Technology, 7(5): 4857-4861.

Wang S, Jiang F, Zhang B, Ma R, and Hao Q (2019). Development of UAV-based target tracking and recognition systems. IEEE Transactions on Intelligent Transportation Systems, 21(8): 3409-3422. https://doi.org/10.1109/TITS.2019.2927838

Yadav SP, Jindal M, Rani P, de Albuquerque VHC, dos Santos Nascimento C, and Kumar M (2024). An improved deep learning-based optimal object detection system from images. Multimedia Tools and Applications, 83(10): 30045-30072. https://doi.org/10.1007/s11042-023-16736-5

Zhang Q, Liu Y, Gong C, Chen Y, and Yu H (2020). Applications of deep learning for dense scenes analysis in agriculture: A review. Sensors, 20(5): 1520. https://doi.org/10.3390/s20051520 **PMid:32164200 PMCid:PMC7085505**