REPUBLIC OF TÜRKİYE

ALTINBAŞ UNIVERSITY

Institute of Graduate Studies

Electrical and Computer Engineering

IMAGE-BASED MACHINE LEARNING FOR SOIL TYPE CLASSIFICATION AND IMPLICATIONS FOR CROP STRATEGY

Momtaz Hasan ALI ALI

Master's Thesis

Supervisor

Asst. Prof. Dr. Oğuz KARAN

İstanbul, 2023

IMAGE-BASED MACHINE LEARNING FOR SOIL TYPE CLASSIFICATION AND IMPLICATIONS FOR CROP STRATEGY

Momtaz Hasan ALI ALI

Electrical and Computer Engineering

Master's Thesis

ALTINBAŞ UNIVERSITY

2023

The thesis titled IMAGE-BASED MACHINE LEARNING FOR SOIL TYPE CLASSIFICATION AND IMPLICATIONS FOR CROP STRATEGY prepared by MOMTAZ HASAN ALI ALI and submitted on 29/12/2023 has been **accepted unanimously** for the degree of Master of Science in Electrical and Computer Engineering.

.

Asst. Prof. Dr. Oğuz KARAN

Supervisor

Thesis Defense Committee Members:

Asst. Prof. Dr. Oğuz KARAN	Department of Software Engineering, Altınbaş University	___________________
Assoc. Prof. Dr. Sefer KURNAZ	Department of Computer Engineering, Altınbaş University	___________________
Asst. Prof. Dr. Zeynep ALTAN	Department of Software Engineering, Beykent University	___________________

I hereby declare that this thesis meets all format and submission requirements of a Master's thesis.

Submission date of the thesis to Institute of Graduate Studies:____/___/___

I hereby declare that all information/data presented in this graduation project has been obtained in full accordance with academic rules and ethical conduct. I also declare all unoriginal materials and conclusions have been cited in the text and all references mentioned in the Reference List have been cited in the text, and vice versa as required by the abovementioned rules and conduct.

Momtaz Hasan ALI ALI

Signature

DEDICATION

To the one I prefer over myself and who sacrificed for me to make me happy always and forever, my beloved mother. To those who are my dearest, my uncle, my aunt, my brothers, and my sisters. And, to my dear wife, and my dear children. I dedicate this humble work to you, hoping that God will always grant us success.

ABSTRACT

IMAGE-BASED MACHINE LEARNING FOR SOIL TYPE CLASSIFICATION AND IMPLICATIONS FOR CROP STRATEGY

ALI, Momtaz Hasan

M.Sc., Electrical and Computer Engineering, Altınbaş University,

Supervisor: Asst. Pro. Dr. Oguz KARAN

Date: December / 2023

Pages: 70

Soil classification is a vital task in various fields such as agriculture, construction, and environmental studies. The present thesis harnesses the power of Artificial Intelligence (AI) to develop a robust and efficient soil classification system. Utilizing a curated dataset of 592 images representing five different soil types, the study explores different deep learning (DL) frameworks, including Convolutional Neural Networks (CNNs), InceptionV3, and MobileNetV2, to process and classify these soil images. The methodology involves a pre-processing phase, ensuring uniformity across the dataset, followed by the application of the selected models. Among the models tested, MobileNetV2 stood out, achieving an unparalleled accuracy rate of 95%, while CNN and InceptionV3 each reached 85% and 93% accuracy, respectively. The comprehensive evaluation included the use of confusion matrices and other performance metrics, validating the exceptional proficiency of the MobileNetV2 model. This work not only advances the field of soil classification but also illustrates the broader applicability of AI and DL techniques in environmental analysis and beyond.

Keywords: Soil Classification, Artificial Intelligence, Deep Learning, MobileNetV2, CNN.

TABLE OF CONTENTS

Pages

LIST OF TABLES

Pages

LIST OF FIGURES

Pages

ABBREVIATIONS

AI	:	Artificial Intelligence
DL	:	Deep Learning
ML	:	Machine Learning
SVM	:	Support Vector Machines
RF	:	Random Forest
GB	:	Gradient Boosting
ACC	:	Accuracy
PREC	:	Precision
REC	:	Recall
F1-S	:	F1-Score
PCA	:	principal component analysis
RNN	:	Recurrent Neural Network
LSTM	:	Long Short-Term Memory
GRU	:	Gated Recurrent Unit
NB	:	Naïve-Bayes
KNN	:	K-Nearest Neighbors
NN	:	Neural Network
MLP	:	Multilayer Perceptron
ANN	:	Artificial Neural Networks
MLP	:	Multi-Layer Perceptron
DNN	:	Deep Neural Networks

1. INTRODUCTION

1.1 BACKGROUND AND MOTIVATION

In contemporary times, the discipline of soil science stands as a pivotal cornerstone for various sectors which hold significant societal and economic implications, including but not limited to, natural resources, energy production, agriculture, climate change response, and infrastructure construction. Essentially, soil science is devoted to the comprehensive study and assessment of soil in its native settings. The examination of the current soil condition is integral for the design and execution of effective crop management strategies. The agricultural sphere, and by extension the state of soil, is profoundly influenced by climate change [1]. As climate change leads to a rising occurrence of extreme weather conditions like droughts and floods, there are marked variations in soil moisture and density. These fluctuations play a crucial role in crop development, with the soil's moisture content and dry density being identified as vital factors influencing growth. Since the rate of crop evapotranspiration is directly dependent on the level of soil moisture, it consequently exerts a significant effect on the overall growth of the crop [2]. Furthermore, elements such as the structure of the soil and the level of soil compaction are known to directly influence crop growth [3].

Classifying soils according to common characteristics or features is an important way to organize soils with comparable engineering attributes. Soil, as opposed to other engineering materials like steel or concrete, has distinct qualities derived from its natural beginnings, geological past, and particle composition [4]. Comprehending these characteristics and soil behaviour are fundamental elements of geotechnical engineering. When classifying soil, variables such as the distribution of particle sizes, liquid limit, and plasticity index are frequently crucial [5]. As far as engineering is concerned, accurate soil categorization is critical. A thorough examination of the different soil qualities is required for the process. For example, sieve analysis, which takes into account both the size of the sieve apertures and the amount of material retained on the sieve, is a frequently used technique to measure the size of soil particles. The soil's plasticity is yet another important factor to take into account while classifying the soil [9].

The purpose of soil classification is not only to categorize soils sharing similar characteristics but also to simplify the differentiation of various soils. Using soil classification, soils can be arranged according to the comparable traits they demonstrate under load [6]. Soil classification is an essential step prior to beginning foundation design. The identification and categorization of soil become vital during the investigation and design stages of geotechnical engineering projects.

Methods for measuring soil properties can essentially be bifurcated into two main categories: experimental procedures and sensor-based techniques. The usage of sensor technology for soil management has particularly gained prominence within the realm of precision agriculture [7]. For instance, the measurement of soil density is typically carried out through an in-situ density test. However, prevailing techniques for soil moisture and density determination are laden with challenges, primarily in terms of being time-intensive and cost-prohibitive. These hurdles have hampered their widespread adoption [8]. In this context, the advent of remote sensing has heralded a new era in soil management techniques. Remote sensing brings the distinct advantage of being able to consistently gather data across extensive geographic areas [8]. Despite these advances, remote sensing does present some limitations. Due to its inherent characteristic of capturing data over large regions, it becomes a challenge to utilize this information for detailed, localized field management. Hence, while remote sensing technology has broadened the horizons of soil management practices, its utility for granular, field-specific applications remains an area that requires further exploration and innovation.

Recently, an upsurge in research focused on the utilization of digital imagery for predicting soil properties has been observed, aiming to expedite the determination of the existing conditions of soils. High-resolution digital images procured through precise tools X-ray tomography have made exhaustive soil analyses viable. However, these instruments have inherent limitations in terms of their time-consuming nature and cost implications, particularly when employed in agricultural lands that necessitate numerous analyses. This underscores the requirement for developing a method to ascertain the current state of soil utilizing commercial-grade digital cameras.

DL, a variant of machine learning (ML), trains computers to learn from a multitude of examples in a manner akin to natural human learning. At this juncture, the machine starts emulating human patterns of learning by discerning specific patterns from various data pieces [10]. In the past, DL had problems with long training periods and insufficient precision. However, these issues have been resolved by new algorithm development and faster computing times thanks to contemporary technology developments. As a result, several different sectors have adopted DL for analysis.

In the past, managing image data sets used to be a complex and time-consuming task. But now, thanks to the emergence of high-speed computing power through GPUs, the analysis of these datasets can be done quickly. Within the realm of image classification, CNN, a form of DL technology, have surmounted the limitations of the preceding methodologies [11]. Thus, leveraging CNNs and other advancements in DL presents promising potential for improved soil classification and management.

1.2 REREARCH MOTIVATION

The motivation driving this research is rooted in addressing the shortcomings and limitations of current soil analysis techniques, and exploring the transformative potential of ML in agricultural and geotechnical practices. In the context of the pressing global challenges, including environmental sustainability and food security, the need for innovative and effective solutions is ever more urgent. Traditional soil analysis methods, while reliable, often lack the scalability, speed, and cost-effectiveness required to address these challenges at a global scale. Consequently, this creates a gap in our ability to understand and manage soil in an efficient and holistic manner, potentially hindering crop yield optimization and effective infrastructure development. Despite the recent adoption of remote sensing technology in soil management, there still exists a pressing need for more precise, field-level soil data analysis. The current inability to quickly process detailed, localized soil data represents a significant challenge in optimizing crop strategies and geotechnical engineering processes. The field of ML, specifically image-based ML, presents a compelling solution to these challenges. Leveraging ML in conjunction with high-resolution digital imagery could offer an innovative approach to soil classification, providing faster, more detailed insights at a reduced cost. This has the potential to revolutionize soil management practices, offering benefits not only in agricultural productivity but also in sustainable resource management and infrastructure development. Thus, the motivation behind this research is to contribute to the transformative potential of technology in soil science. It aims to harness the power of ML and digital imagery, to create a novel, effective, and efficient approach for soil type classification and its implications for crop strategy. This study represents a stride towards more sustainable, efficient, and informed agricultural and geotechnical practices that are poised to address the challenges of the 21st century.

1.3 PROBLEM STATEMENT

Despite the pivotal role of soil science in sectors such as agriculture and geotechnical engineering, current soil classification and analysis methods have inherent limitations. Traditional techniques are often labour-intensive, costly, and time-consuming, thereby constraining their widespread implementation. Additionally, these methods often fail to deliver real-time, detailed data, which is crucial for optimizing agricultural practices and efficient infrastructure planning.

While remote sensing technology has facilitated broad-scale soil data acquisition, its capabilities for providing detailed, localized soil analysis are limited. This restriction creates a gap in our ability to manage soil effectively on a granular level, thereby potentially hampering the optimization of crop growth strategies and geotechnical engineering processes.

Further compounding the problem is the absence of a method that marries the wide-ranging data acquisition capability of remote sensing with the detailed, specific analysis of traditional soil classification methods. This lack of an efficient, cost-effective, and scalable method for comprehensive soil analysis and classification forms the crux of the problem this research aims to address.

In the era of digital transformation, advanced technologies like ML offer promising potential to revolutionize soil analysis. However, the application of image-based ML for soil type classification and its implications for crop strategy is an under-explored area. The lack of research in this domain poses a significant problem, leaving the potential of these technologies untapped in the context of soil science.

Therefore, the problem statement of this research revolves around the development and validation of an image-based ML model for soil type classification, and subsequently investigating its implications for crop strategy. The goal is to address the identified gaps and limitations in the current soil analysis methods and contribute to more sustainable and efficient agricultural and geotechnical practices.

1.4 RESEARCH OBJECTIVES

The objective of this research is to develop and validate an innovative and cost-effective method for soil type classification using image-based ML techniques. It aims to design a model utilizing high-resolution digital imagery to make soil analysis accessible, especially in agricultural lands. The research focuses on adapting DL techniques, specifically CNN, to accurately classify soil types based on digital imagery data, surpassing existing methods' limitations. Validating this model extensively will compare it with traditional soil classification methods, establishing its efficacy and real-world applicability. Furthermore, the study explores how this image-based ML model can optimize crop growth strategies and agricultural practices, offering insights for sustainable resource management and infrastructure development. Ultimately, the research endeavors to enhance soil science through the integration of ML and digital imagery, providing more efficient solutions for soil type classification and its implications on crop strategy.

1.5 REPORT STRUCTURE

The report is structured into several key sections, each serving a specific purpose in presenting the research and findings. The typical structure of the report is as follows:
Chapter 1: Introduction: This chapter offers an overview of the research topic, outlining its objectives and motivation. Additionally, it briefly explains the research methodology adopted in the study. Chapter 2: Literature Review: Here, a comprehensive review of existing literature and relevant research on the topic is presented. The researcher demonstrates their understanding of the current knowledge and identifies gaps that the study aims to address. Chapter 3: Methodology: This chapter describes the research design, data collection methods, and data analysis techniques employed in the study. It ensures transparency and allows for the replication and reliability of the research. Chapter 4: Results and Discussion: In this section, the obtained results from applying the image-based ML model to the soil data are presented. The accuracy and performance metrics of the model, as well as any significant findings from the analysis, are included. The results are interpreted, compared with previous studies, and their implications are discussed. Chapter 5: Conclusion: The final chapter summarizes the main findings of the research, emphasizing its significance in the broader context. Practical applications may be highlighted, and recommendations based on the results are proposed.

2. LITERATURE REVIEW

2.1 INTRODUCTION

The literature review chapter of this work delves into various critical aspects related to soil classification and the application of ML and DL methods in this domain. It begins with an examination of related works, offering an overview of the current state of research in the field. The chapter then moves on to soil classification, discussing different profiles, horizons, and classification systems. Subsequent sections explore ML techniques, such as RF, SVM, GTB. The latter part of the chapter focuses on more advanced methods, introducing ANN, followed by detailed explorations of specific CNN including VGG, MobileNetV2, and InceptionV3. The chapter aims to provide a comprehensive foundation, highlighting relevant methodologies and their applications in soil analysis.

2.2 LITERATURE REVIEW

In [36], researchers used DL with stereo-pair images to classify aggregates of any size into specific classes without the need for manual feature extraction. They trained CNNs including VggNet16, ResNet50, and Inception-v4 architectures. The networks achieved classification ACC above 95%, with ResNet50 performing best at 98.72%. This demonstrated the effectiveness of DL for aggregate classification, providing precise estimation of mean weight diameter (MWD) for aggregates in field conditions.
In [37], researchers proposed a method that involved collecting 50 soil samples from west Guwahati, India, and capturing images using a 13 MP Android camera. The soil samples were then subjected to hydrometer tests to determine sand, silt, and clay percentages, and the US Department of Agriculture soil classification triangle was employed for final classification. The collected soil images underwent pre-processing, region extraction, and texture analysis to generate feature vectors. These feature vectors were subsequently used as input for a SVM classifier with a linear kernel, resulting in an impressive ACC of 91.37%. The classifier's performance closely aligned with the USDA soil classification results.

Researchers in [38], introduced a classification method for essential soil features based on soil test report values. They classified village-specific soil parameters using the Extreme ML (ELM) technique with various activation functions. This had a number of benefits, including lower fertilizer costs, more profitability, less time spent by chemical soil analysis experts, and better environmental and soil health. The gaussian radial basis function outperformed the other examined activation functions in four of the five cases, obtaining an astounding accuracy rate of more than 80%. Furthermore, for pH classification, the hyperbolic tangent activation function emerged as the most effective, achieving an ACC rate close to 90%.

Researchers in [39] propose a novel soil classification approach using a liquid crystal tunable filters (LCTF)-based method and a 3D-CNN. They improve soil hyperspectral images with compressive sensing (CS) and utilize PCA for spectral domain dimensionality reduction. The introduced 3D-CNN framework, along with a differential perception model, achieves an impressive 99.59% ACC in multi-soil classification, outperforming conventional methods.

Researchers use ML and DL models, such as the multi-Stacking ensemble model and the unique feature selection algorithm Q-HOG, to present a multiclass soil classification technique in [40]. The study intends to enhance the identification of soil type while utilizing AI developments for intelligent farming. To improve ACC categorization, soil pictures are gathered from the vriddhachalam exploration site. ML techniques like NB, KNN, and SVM are tested with DL models like RNN, LSTM, GRU, and VGG16. In soil classification, the suggested multi-Stacking ensemble model performs better than alternative models, obtaining an amazing ACC of 98.96%, PRE of 96.14%, REC of 99.65%, and F1-S of 97.87%.

Three popular classification algorithms: SVM, RF, and MLP were compared by researchers in [42]. Their study concentrated on forecasting the likelihood of soil erosion using Sentinel-2 pictures by utilizing ground truth samples that were provided by Iceland's Soil Conservation Service department and collected from different parts of the country. The research conducted on validation datasets demonstrated the efficacy of both SVM and RF techniques in simulating soil erosion, with average accuracies of 0.917±0.001 and 0.931±0.002, respectively.

Researchers in [43], aimed to accurately classify soil images using pre-trained weights with Tensorflow and Keras DL frameworks. They used a dataset of 903 soil images of four types, dividing it into training and validation sets. Augmented images were employed to enhance model performance. The CNN achieved 99.86% ACC for training and 97.68% for validation. The CNN model demonstrated superior performance over the DCNN models and previous works. Evaluation involved using a CM and K-fold technique.

A portable smartphone-based machine vision system using CNN for the classification of soil texture images was proposed by researchers in [44]. The first block in the CNN model was for feature extraction, and the second block was for classification. Both blocks had several layers. In addition, comparison was done using the ANN, SVM, RF, and KNN algorithms. High accuracy of 99.89%, 99.81%, and 99.58% were attained by the suggested CNN model for 20 cm, 40 cm, and 60 cm distances, respectively. The study found that the optimum performance was achieved with fully preprocessed photos at a height of 20 cm. For soil texture prediction based on the CNN model, an intuitive graphical user interface was created, showcasing the method's promise as an accurate and efficient substitute for expensive and time-consuming laboratory techniques in large-scale farming.

Researchers suggested using hyperspectral imaging in [45], to quickly screen and categorize microplastics (MPs) in soil from farms. They obtained 600 hyperspectral images from 180 soil samples, which they then preprocessed using PCA to reduce dimensionality and the S-G smoothing filter. For the MP polymer classification (PE, PP, and PVC), three models (DT, SVM, and CNN) were developed. The maximum ACC of 92.6% was attained by the CNN model with S-G smoothing, outperforming the DT and SVM models with 87.9% and 85.6% ACC, respectively.

In [46], researchers propose an automated soil classification technique called FCMCSO-ASC, which uses Fuzzy Cognitive Maps with Cat Swarm Optimization. The technique aims to classify seven different types of soil and utilizes the local diagonal extrema pattern (LDEP) as a feature extractor. The FCMCSO model optimizes the weight values for soil classification. The technique was evaluated on a benchmark dataset, achieving a maximum ACC of 96.84%, outperforming recent techniques.

2.3 SOIL CLASSIFICATION

Soil is a crucial part of the Earth's crust, comprising the topmost layer that starts at the surface and extends downwards to a depth usually accessible to plant roots, which is typically around 2 meters, unless specified otherwise [12]. It serves as the foundation and nurturing ground for a wide array of terrestrial life forms, providing vital support to vegetation and playing a pivotal role in ecological processes. The classification of soil, a key aspect in soil science, involves the systematic organization and categorization of soils based on their distinctive properties and characteristics. For classification purposes, the typical consideration is the depth of the soil, with the standard limit set at 2 meters, unless specific conditions dictate otherwise [12, 13]. This delineation serves as a fundamental criterion in the study and assessment of various soil types and their relevance to specific ecosystems, agricultural practices, and engineering applications. By comprehending the depth-based variations in soil properties and behavior, soil scientists, ecologists, and agronomists can derive valuable insights to better understand and manage soil ecosystems for sustainable resource utilization and environmental preservation.

2.3.1 Profiles and Horizons

A soil's profile, or a vertical portion of the soil at a particular site, is usually used in soil science to classify soil. Different layers, or horizons, are seen in this profile. We will refer to these distinct strata as "horizons" and "layers" respectively throughout this paper. Excavation of a vertical section is frequently necessary in order to conduct a comprehensive profile study, which is necessary in order to properly classify the soil. The number of strata in this profile may differ, but it contains every horizon connected to the earth at that particular place.

The analysis of these horizons involves the assessment of properties that can be observed in the field, such as the color or width of each layer. Additionally, some properties are determined through scientific laboratory tests, including pH levels and electrical conductivity. Thoroughly exploring all the layers within the profile allows for informed decisions on the appropriate classification of the soil to which they belong. By considering a multitude of observable and measurable characteristics, soil scientists can derive valuable insights into the soil's composition, behavior, and overall suitability for various ecological, agricultural, and engineering purposes. The accurate classification of soil profiles is a crucial step in soil science, enabling more effective soil management, sustainable land use planning, and informed decision-making for environmental preservation and resource conservation.

Two illustrations of soil profiles with distinct horizons are shown in Figure 2-1. This specific profile was gathered from a Chinese location. The dark red Durisol soil profile, which originates in South Africa, is seen on the right. The graphic highlights the diversity and importance of soil classification in understanding soil behavior and its implications for various environmental and agricultural applications. It also offers visual insights into the varying soil compositions and layering observed in different places.

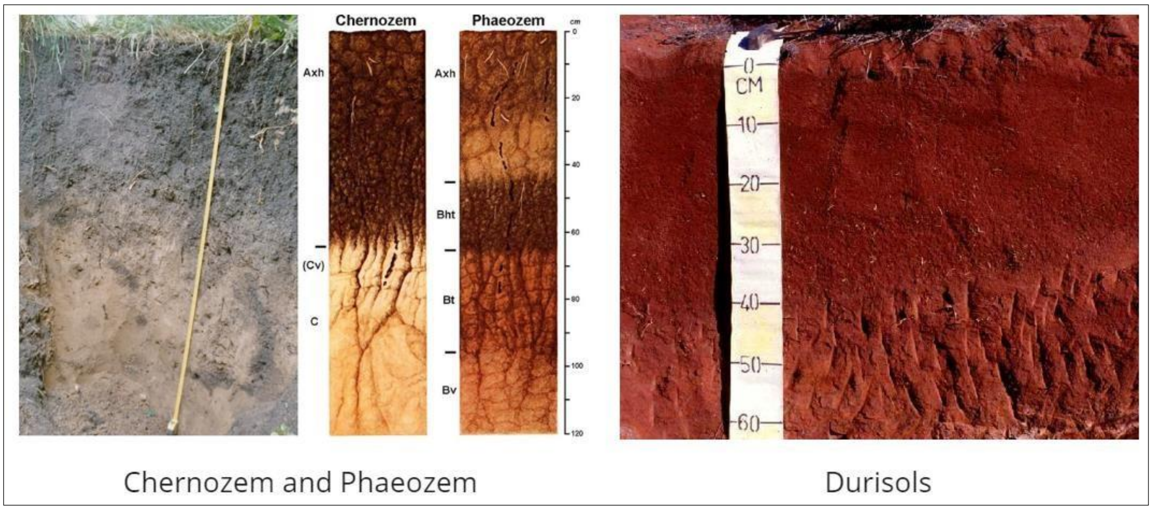

Figure 2.1: Examples of Soil Profiles [14][15].

2.3.2 Soil Classification Systems

The classification of soil presents a multitude of methods, with each country often employing its unique system [16][12]. Consequently, communication and comparison of soil attributes between different countries can be intricate, as direct translation between various classification systems may not be straightforward without comprehensive knowledge of how both systems operate.

Particular classification systems have been developed to address the problem of cross-country communication in soil categorization. Among the most popular systems are the USDA Soil Taxonomy and the World Reference Base. Although the USDA Soil Taxonomy was initially developed in the US, many other countries now employ it as a standard classification system. By offering a common framework for soil classification, these systems promote international collaboration and research endeavours, leading to a more cohesive comprehension of soil properties and behaviour across diverse geographical regions.

The World Reference Base for Soil Resources (WRB), an internationally recognized soil classification system established by the International Union of Soil Sciences (IUSS), builds upon the FAO Legend and incorporates concepts from the Soil Taxonomy. Its main objective is to provide a simple classification method based on observable field data, ensuring compatibility with various national soil classifications for enhanced international communication. The WRB is not designed to replace local soil classification systems but rather serves as a supplementary tool to facilitate global soil assessments and promote better understanding across different regions.

This classification system operates at two levels of detail. The first and broader level is referred to as the Reference Soil Group (RSG). For instance, as seen in Figure 2.1, the soil profile is identified as a Phaeozem, distinguished by a dark horizon with a high concentration of organic material, among other specific characteristics [18]. The second level, which combines the RSG with a number of prefix and suffix qualifiers, provides a more detailed description of the soil and its layers. For example, the profile depicted in Figure 2.1 might also be classified as a Mollic Phaeozem, where the word "Mollic" serves as an additional descriptor to further characterize the surface horizon, which is identified by its abundance of organic components and dark colour.

Since there are a great deal of RSG and their corresponding qualifiers, which leads to an excessive number of alternatives, it would be difficult to anticipate the most precise level, therefore we will only address the RSG level in this dissertation [17]. We will focus on the current RSG, which each represent unique classes of soil defined by the World Reference Base for Soil Resources [19], and including comprehensive characteristics, attributes, and requirements. Classifying soil is a difficult task best left to professionals due to the large number of quantifiable factors, the subjectivity of some features, and the vast range of possible attribute values within each horizon. Successful classification necessitates significant expertise in the field and an understanding of the nuanced and intricate characteristics of different soil profiles.

2.4 MACHINE LEARNING METHODS

In recent years, ML methods have emerged as powerful tools in the field of soil classification. These advanced techniques offer the potential to revolutionize traditional approaches, enabling more accurate, efficient, and scalable classification of soil types based on various data sources. By harnessing the power of ML algorithms, researchers can explore complex relationships between soil attributes and make predictions with unprecedented precision. In this section, we delve into the application of ML methods in soil classification, exploring their capabilities, challenges, and the opportunities they present for advancing our understanding of soil behaviour and management.

2.4.1 Random Forest (RF)

An ensemble of decision trees is used in the RF technique, and each tree is trained using a unique, random subset of training data and characteristics [20]. By selecting data points at random from the original set, this technique referred to as bagging allows for the possibility of duplicate selections. As a result, it produces training data in a variety of random groupings, increasing the variance between the different trees. Out-of-bag samples are the data points that were not utilized during a tree's training process. Moreover, RF introduces further randomness by restricting the features available during the training of each tree. Although this may introduce some bias, as certain trees may lack the best feature splits, it also provides an opportunity for other features to contribute valuable insights to the data. Because the RF ensembles typically produce better predictions when using this method than when utilizing the whole set of variables, it effectively addresses overfitting issues [20].

Furthermore, the model uses the Gini Importance metric to determine which factors are the most influential during the training phase [21]. The significance measure for each feature is computed as the sum of the Gini reduction averaged over all ensemble trees, weighted by the number of samples the feature separates. It is simpler to recognize and eliminate less important features when using the Gini Importance metric, which provides insightful information about the problem's most important factors. This procedure guarantees that the model's performance is maintained while also cutting down on processing time.

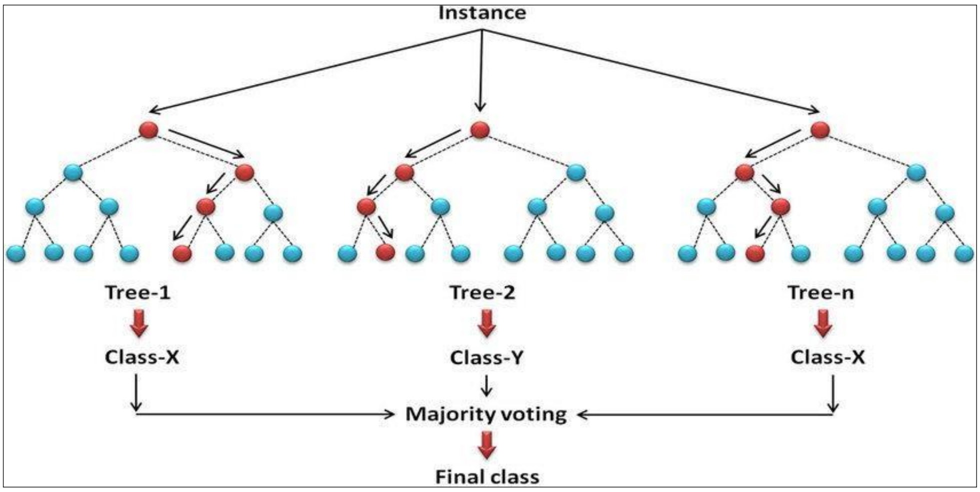

Figure 2.2: Random Forests [27].

2.4.2 Support Vector Machines (SVM)

SVM stands as another prominent ML approach frequently utilized in soil classification. SVM focuses on using a subset of training data termed support vectors to create partition boundaries or hyperplanes within the feature space of co-variables. This approach does not rely on specific statistical distributions of the training data, providing greater flexibility in delineating class boundaries.

SVMs employ optimization algorithms to generate these partition boundaries, effectively separating classes within the selected feature space or between co-variables. However, SVMs do require specific user input, necessitating the specification of parameters, including the choice of kernel, kernel-specific parameters, and regulation parameters. These parameter choices significantly influence the outcomes of the inference, making the selection of optimal parameters crucial to achieve the best classification results [22][23] [24]. Therefore, considerable attention must be given to parameter tuning to ensure accurate and effective soil classification results [25].

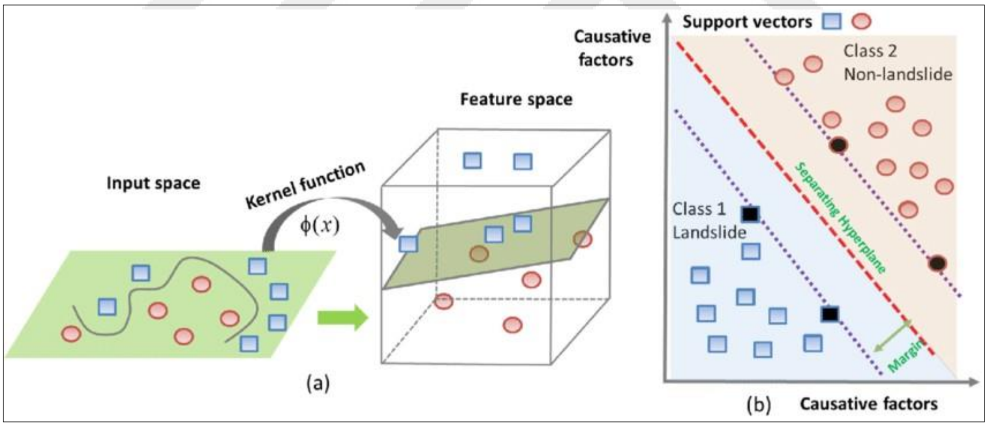

Figure 2.3: Support Vector Machine [28].

2.4.3 Gradient Tree Boosting

Another ensemble technique that uses decision trees but differs from RF is called Gradient Tree Boosting (GTB). To achieve diversity, GTB gradually improves the ensemble's performance rather than adding unpredictability [26].

The GTB approach starts by using the dataset to train a basic decision tree. Then, iterating over the dataset and creating a residual error based on the mistakes committed. A new tree is then trained to fit the residual error from the previous iteration, with its primary focus being the data points that the previous model misidentified. This iterative process, known as gradient boosting, computes the residuals using the gradient of the user-specified loss function [26]. After giving the new tree a weight and adding it to the ensemble, the procedure is repeated until a classifier that minimizes overfitting to the training data and fits the data optimally is obtained.

Figure 2.4 illustrates the progression of GTB, beginning with a basic tree and subsequently constructing additional trees, each relatively simple but collectively capable of solving the problem effectively. Although individual decision trees introduce considerable bias and make numerous mistakes due to their focus on specific data points, the combined model attains consistently superior performance. To mitigate the risk of overfitting, two approaches can be implemented: pruning and early stopping. In the case of pruning, it involves applying modifications to the decision tree after the training phase to simplify its structure and reduce complexity, which helps prevent overfitting. On the other hand, early stopping involves monitoring the out-of-bag error during training. Overfitting to the training data is indicated when the out-of-bag error begins to steadily rise while the training error keeps becoming smaller. To prevent more overfitting, the algorithm responds by stopping the training process [26].

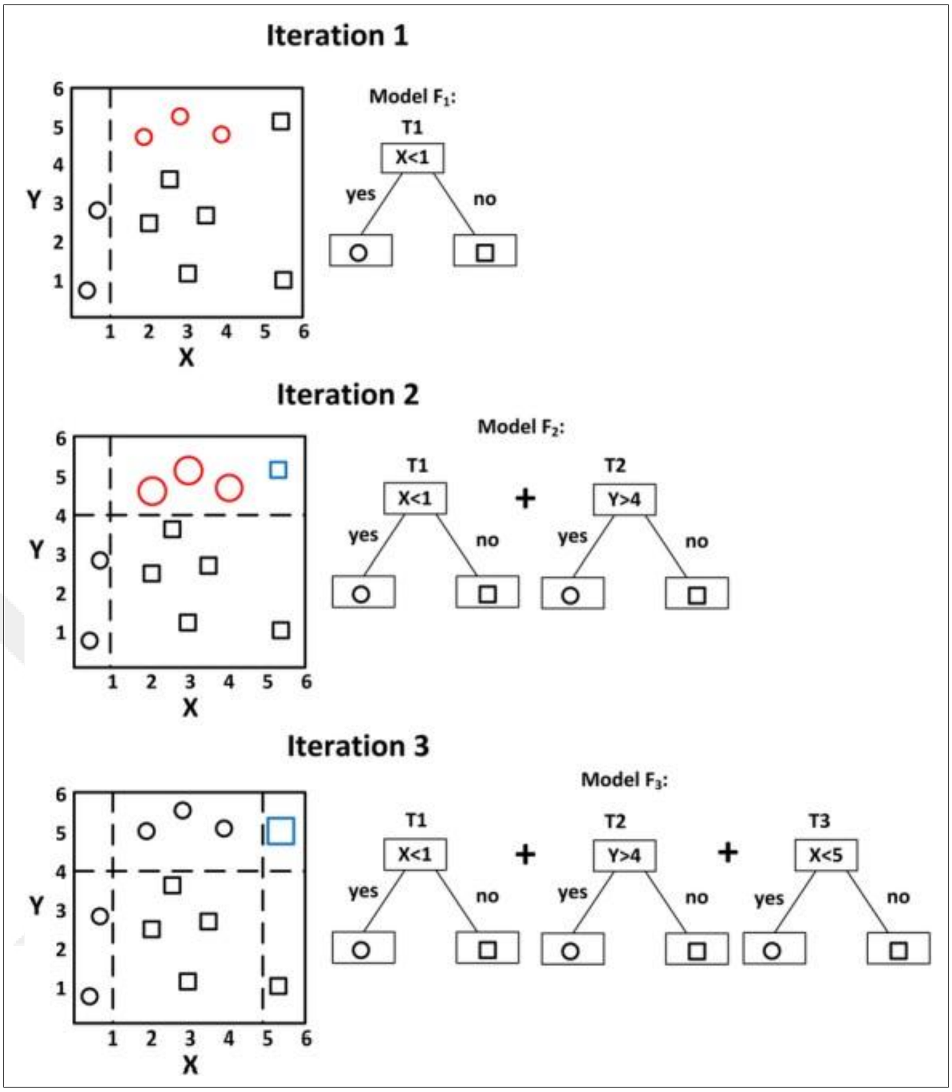

Figure 2.4: Gradient Tree Boosting [29].

2.5 ARTIFICIAL NEURAL NETWORKS

Artificial neural networks (ANNs) are computational models that are modelled after the networked structure of animal brains, in which neurons process inputs and send out signals to other cells [30]. The core component of an ANN is the perceptron [31], a node that receives inputs (each with a weight) and uses an activation function to produce an output. The perceptron modifies its weights during training in order to maximize predictions on the training set.

However, when dealing with nonlinearly separable functions, a single perceptron might encounter challenges in accurately classifying data [32]. To overcome this limitation, multiple perceptron's are organized into different layers, giving rise to a fully connected

MLP [32]. These hidden layers are essential to the evolution of DNN and can be rather numerous. Although training a single perceptron entail modifying input weights in accordance with the error function's gradient, this method is not directly relevant to MLPs because of the intricate relationships between nodes and how differently each one contributes to the final prediction. The back-propagation algorithm is used to solve this problem. Based on each perceptron's contribution to the outcome, it distributes the estimated error between the MLP's output and the expected output back to each one. This enhances the MLP's ability to classify data by enabling it to change its internal biases and weights.

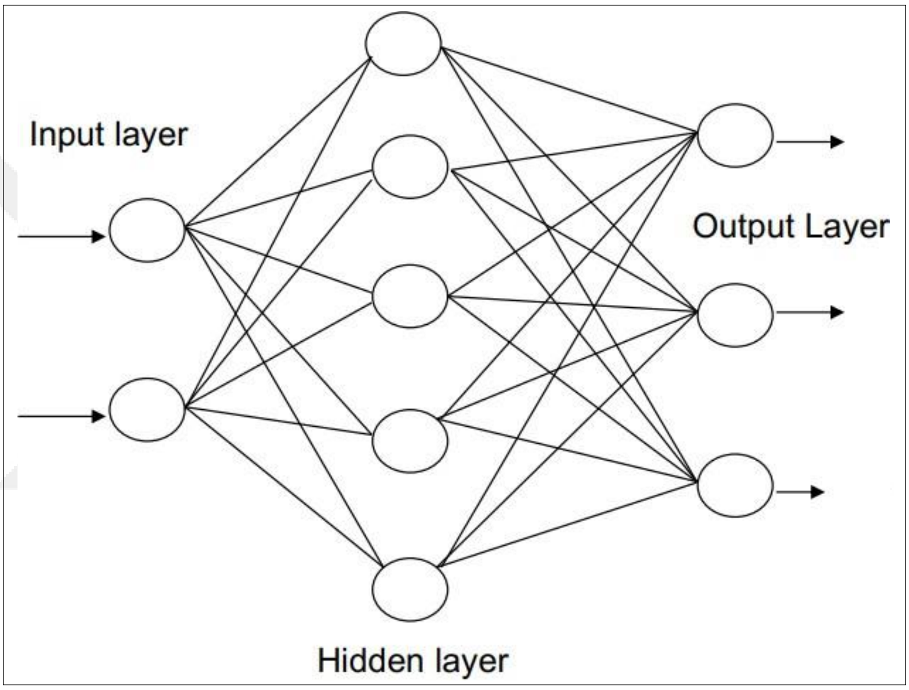

Figure 2.5: Artificial Neural Networks.

Moreover, ANN models can handle and model multiple outputs simultaneously, making them particularly beneficial for predictive tasks such as soil property mapping [30]. Recent advancements have introduced specialized architectures like CNN, highly effective in image analysis by detecting basic patterns through convolutions and then combining them to identify complex patterns [33]. CNNs excel in image-related tasks due to their ability to learn from the abundant patterns found in natural images.

2.6 CONVOLUTIONAL NEURAL NETWORKS

CNN represent a powerful and highly effective class of DL models, specifically designed for tasks involving image and pattern recognition [50]. CNNs have gained significant popularity due to their outstanding performance in a wide range of visual tasks, including object detection, image classification, and segmentation.

The architecture of a CNN is inspired by the visual processing mechanism of the human brain [51]. It consists of multiple layers, each responsible for different operations on the input data. The fundamental building blocks of CNNs are convolutional layers (CL), where small filters or kernels slide over the input image to detect specific patterns or features. These convolutional operations enable the network to learn essential visual features in a hierarchical manner. The pooling layer is a crucial element of CNN architecture that minimizes the spatial dimensions of the data while preserving crucial information. The network becomes more computationally efficient and resilient to changes in input size or position when pooling is used to down sample the feature maps [53]. CNNs may progressively transition from low-level information to high-level representations by processing data through numerous layers, which enables them to capture abstract and complicated properties. The fully connected layers, which function as conventional neural network layers for additional processing and classification, receive the flattened output of the convolutional and pooling layers.

The architecture of CNNs allows them to automatically learn hierarchical feature representations from raw data, making them highly adaptable to diverse tasks without the need for manual feature engineering. This characteristic, along with their ability to handle large datasets and exploit spatial locality, contributes to their success in image-related applications.

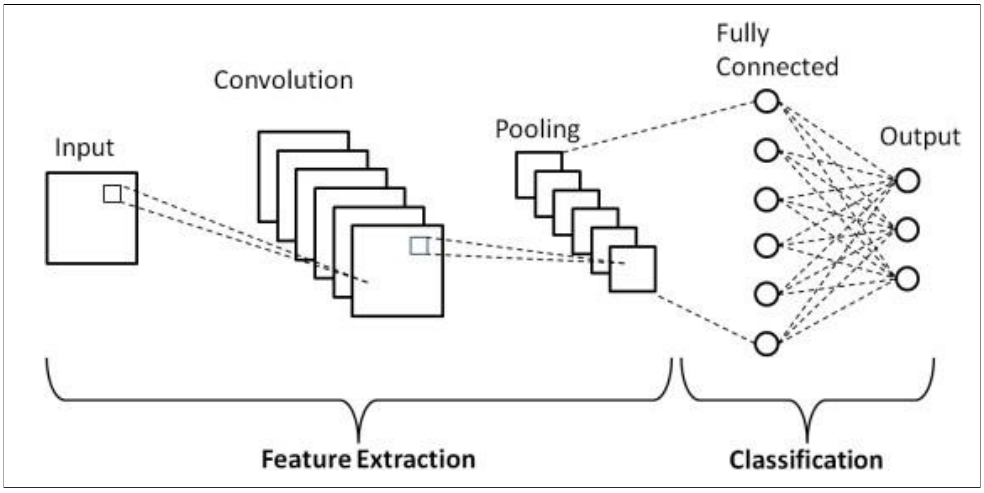

Figure 2.6: Convolutional Neural Networks [34].

a. Convolutional Layers (CL): CLs are fundamental building blocks of CNN, a type of DL model extensively used in image processing tasks [52]. These layers function by sliding small, receptive fields or 'filters' across the input data, which could be an input image or the output from a previous layer.

Each filter is associated with a set of weights and a bias, and as it slides across the input, it performs a mathematical operation called convolution. This operation involves element-wise multiplication of the filter's weights with the pixel values of the input data it currently covers, followed by summing up these products and adding the bias. The result is a single value in a new matrix known as a 'feature map' [54].

The process is repeated for each position the filter can occupy on the input, resulting in a full feature map. When multiple filters are used, each one learns to recognize different features of the input data, such as edges, corners, or more complex patterns in higher layers of a deep CNN.

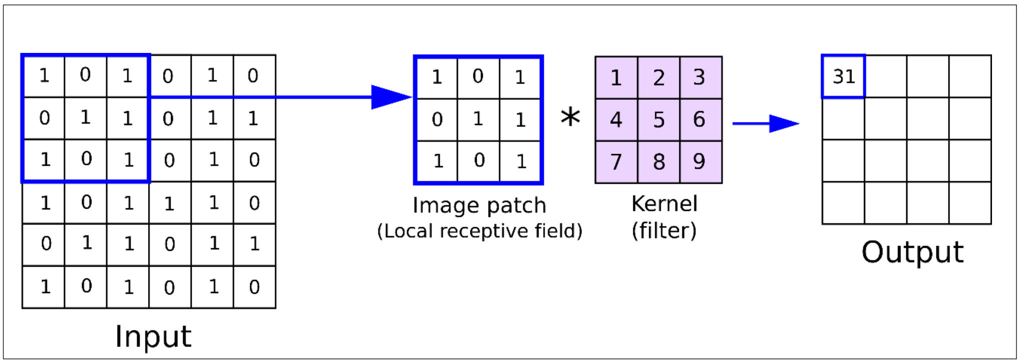

Figure 2.7: Convolutional Operation.

b. Activation Layers: The activation layer in a neural network plays an essential role in transforming the linear outputs from previous layers into a nonlinear format, enabling the network to learn complex patterns and make better predictions. Each neuron in the network has an activation function which takes in the weighted sum of its inputs and produces an output, often a value between a specified range, that is then passed on to the neurons in the next layer. Activation functions that are frequently used are the rectified linear unit (ReLU) function, which outputs the input directly if it is positive and 0 otherwise; the sigmoid function, which squashes its inputs to a range between 0 and 1; and the hyperbolic tangent (tanh) function, which produces outputs between -1 and 1. The kind of problem being addressed and the particular needs of the model determine which activation function should be used [55].

The activation layer, therefore, serves as a kind of gateway, controlling the flow of information through the network. By introducing non-linearity into the model, activation functions allow the network to capture and learn from the complex, nonlinear relationships inherent in many real-world datasets. Without activation functions, no matter how deep the network is, it would still behave like a single-layer perceptron as all layers would be performing linear transformations on the inputs. Activation functions are, therefore, a crucial component in enabling the DL models to solve sophisticated tasks.

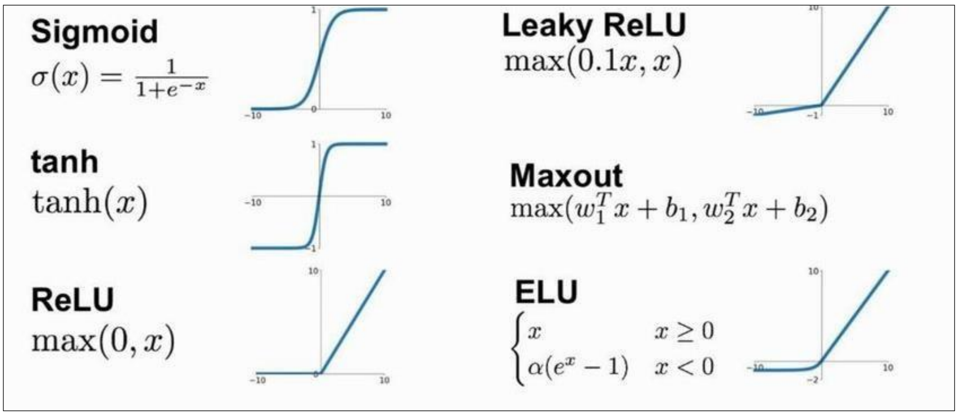

Figure 2.8: The Activation Functions of the Neural Networks.

c. Pooling Layer: The pooling layer is an essential component in CNN and serves a critical role in the process of feature extraction and dimensionality reduction. A pooling layer operates by systematically scanning small regions of its input - typically the output of a previous CL - and summarizing the information in these regions into a single value. The exact nature of this summary depends on the type of pooling operation being used |56].

Max pooling is the most popular kind of pooling; it takes the maximum value out of each zone. The key traits are efficiently preserved while less significant information is discarded. Additional pooling methods include L2-norm pooling, which determines the square root of the sum of all squares in a region, and average pooling, which determines the average value of each region [57]. Pooling layers minimize the spatial dimensions of its input by summarizing parts of the feature map, which results in a more compact and manageable representation of the input data. Because fewer parameters must be trained, this dimensionality reduction helps the model become less prone to overfitting and more computationally efficient. Despite this decrease, the pooling procedures usually maintain the most important features for the job at hand, enabling the network's next layers to function on a concentrated, condensed representation of the original input.

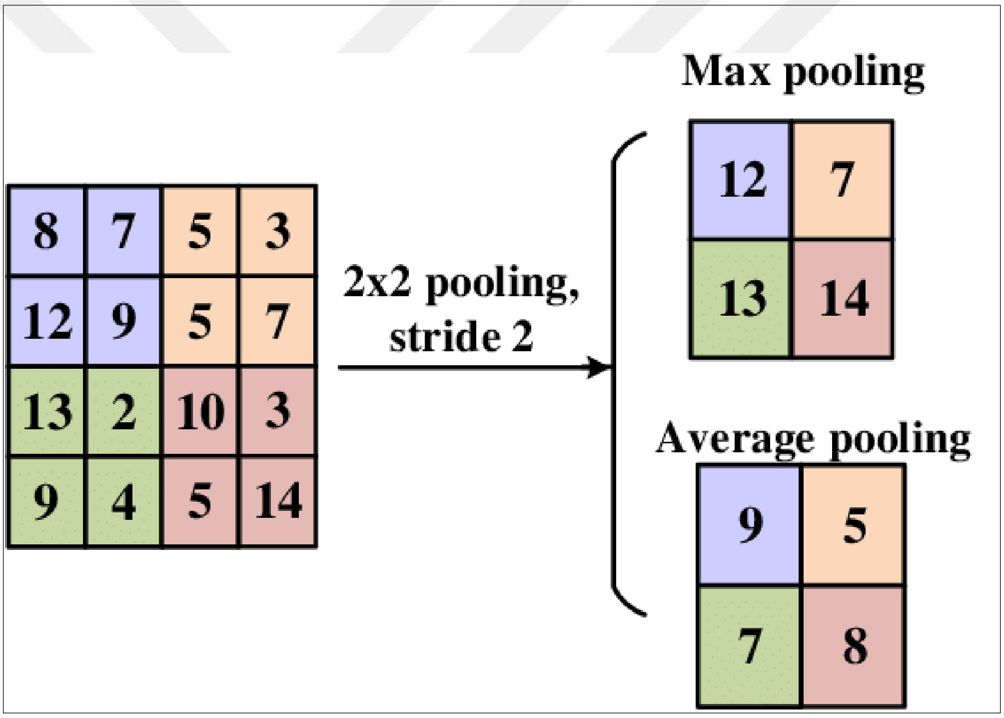

Figure 2.9: Example of Pooling Operation.

d. Dropout Layer: The dropout layer is a unique type of layer employed in neural networks to prevent the phenomenon of overfitting, which happens when a model learns the training data too well and performs poorly on unseen data. It was introduced as a simple yet effective regularization technique to enhance the generalization capability of DL models [58].

In a dropout layer, a fraction of the neurons in the layer are randomly "dropped out" or deactivated during each training iteration, meaning that they do not contribute to the forward pass and their weights are not updated during the backpropagation. The fraction of neurons to be dropped is determined by a hyperparameter, typically set between 0.2 and 0.5 [59]. This random omission of neurons forces the network to learn more robust and generalized features. Instead of relying heavily on a specific set of neurons, the network is encouraged to distribute the learned representations across all neurons, leading to a more robust model that is less likely to overfit the training data.

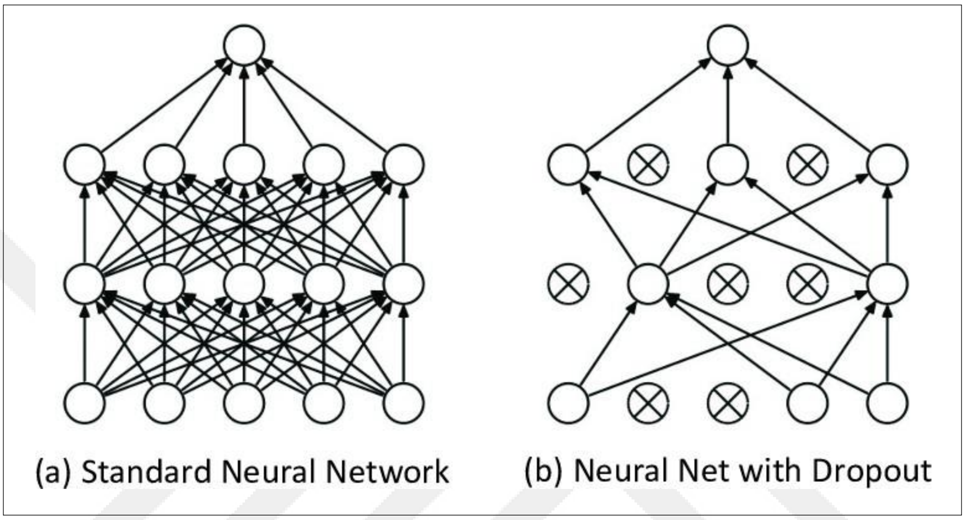

Figure 2.10: Dropout mechanism [48].

e. Batch Normalization: Batch Normalization is a technique widely used in DL models to accelerate training, stabilize the learning process, and reduce the sensitivity to the initial weights. It is typically applied to the outputs of a layer before the activation function [60].
The central idea behind Batch Normalization is to normalize the activations of each layer (i.e., make them have zero mean and unit variance), similar to how input data is usually standardized. This normalization is performed on each mini-batch separately, hence the name 'Batch Normalization'. The motivation for doing this is to combat the so-called "internal covariate shift", the phenomenon where the distributions of layer activations change during training, slowing down the learning process [61].
However, Batch Normalization does more than merely standardize the activations. After the normalization step, Batch Normalization adds two learnable parameters to each neuron: a scale factor and a shift factor. These parameters allow each neuron to undo the normalization if it turns out to be beneficial for learning.

f. Flatten Layer: The flatten layer in a neural network serves a specific yet crucial purpose, acting as a connector between convolutional and dense layers. As the name suggests, this layer 'flattens' its input without affecting the batch size [62][63].

In the context of an image classification task, after several rounds of convolutional and pooling layers, the output is a high-dimensional array that encapsulates the extracted features of the input image. However, this high-dimensional output cannot be directly fed into the FCLs that often follow, as these layers expect data to be in a 1-dimensional array format.

The flatten layer comes into play at this juncture. It reshapes or flattens the high-dimensional input into a 1-dimensional array, while preserving the integrity of the input data. Each pixel from the previous layer becomes an individual input neuron in the next layer.

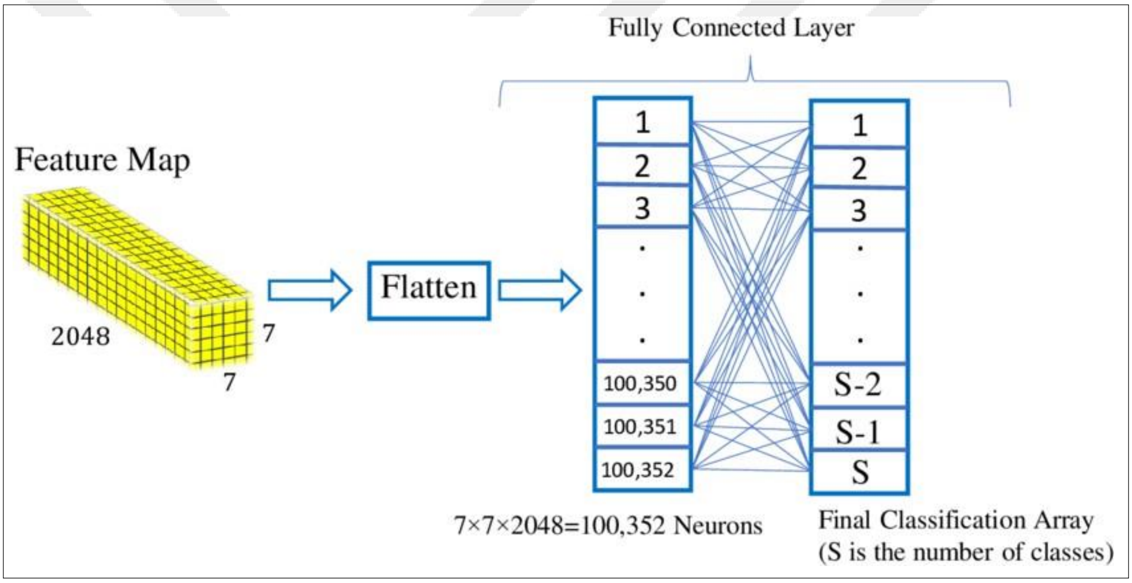

Figure 2.11: Example of Flatten Layer.

g. Fully Connected Layer (FCL): The fully connected layer, often abbreviated as FCL, is a crucial component in various neural network architectures. As the name suggests, each neuron in a FCL is connected to every neuron in the previous layer, with each connection having a weight [64].

In the context of CNNs, which are widely used in image processing tasks, the FCL typically follows a series of convolutional and pooling layers. These preceding layers are responsible for the extraction and processing of features from the input data. The FCL then uses these processed features for classifying the input data into various classes, thus acting as a classifier.

The FCL works by computing the dot product of its input and weights, adds a bias term, and then passes the result through an activation function. In essence, each neuron in a FCL learns a separate weight for every single input and can thus make complex decisions based on the entire input [65].

The final FCL, often followed by a SoftMax activation function, produces the output probabilities for each class in a multi-class classification problem. These probabilities sum to one, and the class with the highest probability is usually taken as the model's prediction.

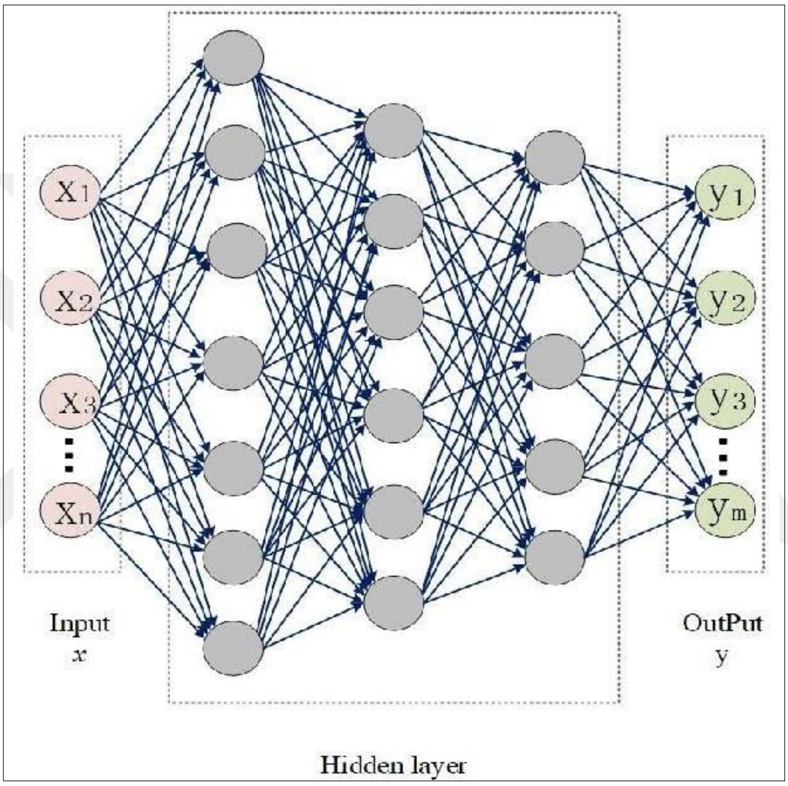

Figure 2.12: Fully Connected Layer [49].

With advancements in DL and ample computational resources, researchers have extended the use of CNNs beyond image tasks to various other domains. Among these, ResNet, Mobile Net, and VGG have gained widespread attention and become popular choices in the research community.

2.6.1 VGG Method

The VGG (Visual Geometry Group) architecture is a deep CNN that was proposed by the Visual Geometry Group at the University of Oxford [66]. It gained significant attention after its participation in the ILSVRC in 2014, where it achieved remarkable results in image classification tasks [67].

The key characteristic of the VGG architecture is its simplicity and uniformity. It consists of a series of stacked CLs, each followed by a max-pooling layer for spatial down sampling. The CLs use small 3x3 filters with a stride of 1, which helps in capturing local features in the input image. This repeated use of 3x3 filters allows VGG to learn hierarchical representations of increasing complexity as the layers deepen [68] [69]. VGG comes in various configurations with different depths, commonly denoted as VGG16 and VGG19, depending on the number of CLs. VGG16 has 16 CLs, while VGG19 has 19. Despite the simplicity of its architecture, VGG achieved impressive performance on image classification tasks, demonstrating that deeper networks can learn more complex features, thus improving accuracy [70].

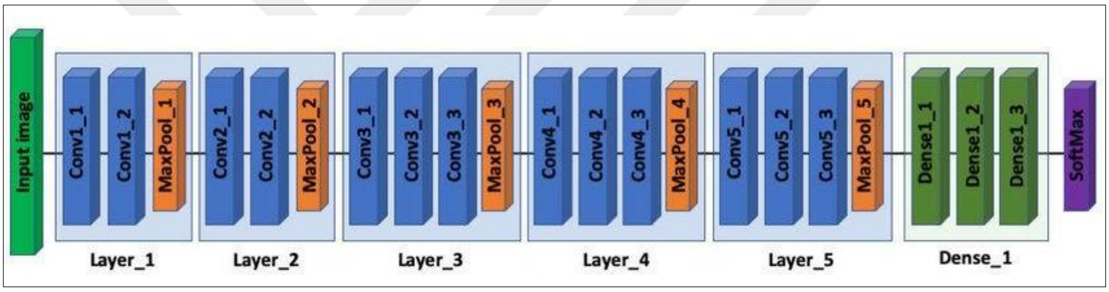

Figure 2.13: Example of VGG 16 Architecture [35].

2.6.2 MobileNetV2 Method

MobileNetV2 is a CNN architecture that was designed with the aim of achieving efficient, high-performance models that can be deployed on mobile and embedded systems. The MobileNetV2 architecture is notable for its use of depth wise separable convolutions, an operation that significantly reduces computational cost by dividing the standard convolution operation into a depth wise convolution and a pointwise convolution [71].

A typical MobileNetV2 model follows the structure of a traditional CNN architecture, where an input image passes through a series of convolutional, non-linear, pooling, and normalization layers to produce an output, typically a class label [72]. However, it diverges in its use of a unique module called the "inverted residual with linear bottleneck". The structure of this module is what gives MobileNetV2 its efficiency and power.

The inverted residual block consists of three layers: a 1x1 expansion convolution, a 3x3 depthwise convolution, and a 1x1 projection convolution. This design differs from the traditional residual block as it initially expands the number of channels in the input tensor, then applies the depthwise convolution, and finally reduces the number of channels back to its original size. The term "linear bottleneck" arises because there is no non-linear activation function applied after the final 1x1 convolution. Utilizing these blocks enables the model to learn more complex features with fewer parameters, resulting in reduced computational and memory requirements [73]. The architecture begins with a standard CL with 32 filters, followed by 19 inverted residual blocks of varying filter sizes and strides. Upon traversing all the residual blocks, the tensor undergoes a 1x1 convolution, followed by average pooling. Finally, the tensor is passed through a FCL with SoftMax activation to generate the output classes [74].

Moreover, MobileNetV2 also introduces two important hyperparameters, width multiplier and resolution multiplier, which allow the model to be scaled for different resource constraints. The width multiplier alters the number of channels available in each layer, while the resolution multiplier changes the input image resolution. By adjusting these multipliers, it's possible to create smaller and more efficient models, which makes MobileNetV2 highly flexible and adaptable for various applications.

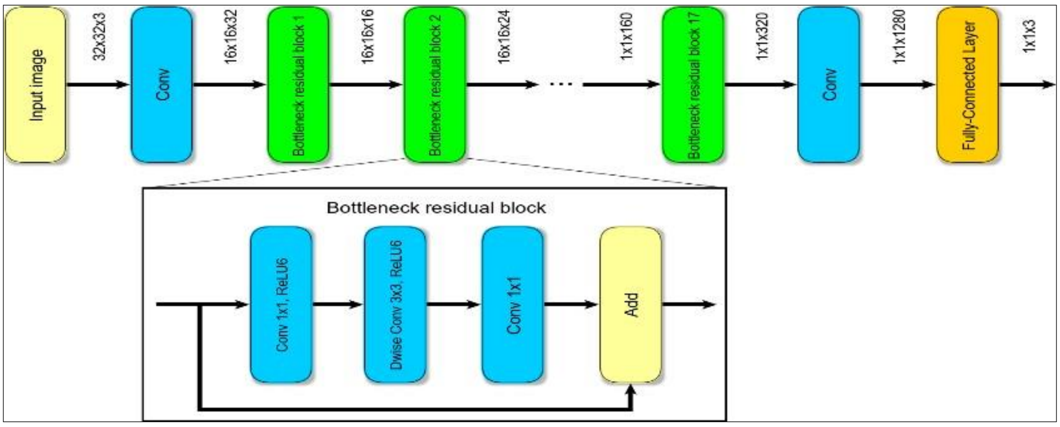

Figure 2.14: MobileNetV2 Method [47].

2.6.3 InceptionV3 Method

InceptionV3 is a CNN architecture that was a part of the third iteration of Google's Inception project, under the umbrella of the Google Net series [75]. This architecture is well-regarded for its innovative design that substantially reduced the computational burden of training

DNN, while achieving high accuracy in image classification tasks. The defining characteristic of the Inception architecture is the 'Inception Module', a building block that comprises multiple convolution operations with various kernel sizes (such as 1x1, 3x3, and 5x5) operating in parallel on the same input, followed by concatenation of their outputs [76]. The beauty of this design lies in its ability to capture both local and global features in the input by varying the size of the receptive field across the parallel branches. It also enables the network to adaptively select the spatial scale of the features, enhancing its performance on tasks with high intra-class variability. InceptionV3, specifically, introduced several enhancements over its predecessors. The architecture used factorization to reduce the computational cost of larger convolutions. Instead of directly applying 5x5 or 7x7 convolutions, it broke them down into smaller, sequential convolutions - for instance, a 5x5 convolution was replaced by two 3x3 convolutions. This technique not only led to computational efficiency but also deepened the network, making it capable of learning more complex patterns [77].

The architecture also embraced asymmetrical convolutions. Instead of using a square 3x3 convolution, it used a combination of a 1x3 convolution followed by a 3x1 convolution. This was done to exploit the correlation between spatially closer features while reducing computational complexity. Further, InceptionV3 featured 'Auxiliary Classifiers' placed at intermediate layers to propagate the signal of label information and address the vanishing gradients problem, which tends to occur in deep networks [78]. Lastly, the network employed efficient 'Grid Size Reduction', where instead of using pooling operations, convolution strides were used to reduce grid sizes, preserving more information in the process [79]. InceptionV3 is usually composed of 48 layers, although the effective depth could be much higher due to the inception modules. Its design principles have influenced many subsequent architectures in the DL field, making it a significant milestone in the evolution of CNN.

2.7 CONCLUSION

The literature review has provided a comprehensive understanding of the multifaceted aspects of soil classification and the evolving technological approaches used to address this complex subject. By exploring traditional classification systems, ML techniques, and state-of-the-art DL methods, this chapter has laid a solid foundation for understanding the existing frameworks and their applications. This understanding is essential to the development of new methodologies, reflecting the advancements in the field and catering to the specific needs and challenges associated with soil analysis.

3. METHODOLOGY

3.1 INTRODUCTION

The methodology chapter of this thesis provides a detailed account of the approach adopted to conduct the soil classification study using DL models. This chapter describes how the image dataset was prepared and processed, including image resizing and label encoding, to be fed into the models. It then delves into the specifics of the proposed models, explaining the choice of architectures, which are CNN, InceptionV3, and MobileNetV2, and their configurations. Furthermore, it discusses how these models were trained and evaluated, shedding light on the process of fine-tuning the models, selecting appropriate parameters, and validating the performance of the models.

3.2 THESIS APPROACH

The proposed model for soil classification is a multi-phase process involving image pre-processing and the application of advanced DL models. Initially, a pre-processing stage is employed, where soil images are resized to fit the requisite input dimensions for the DL models. This ensures consistency in the data, making it compatible with the models' architecture. Subsequently, DL models are applied to the processed images to identify and classify distinct soil types. Models such as CNNs are utilized due to their proficiency in image-based tasks, effectively identifying patterns and features in the soil images. Further, advanced pre-trained models like InceptionV3 and MobileNet V2 are integrated, leveraging their capacity for high-level feature extraction. They can classify soil types more accurately by exploiting their intricate, hierarchically-structured learning capacity. Ultimately, the synergy between pre-processing and these DL models fosters an efficient and effective framework for soil classification.

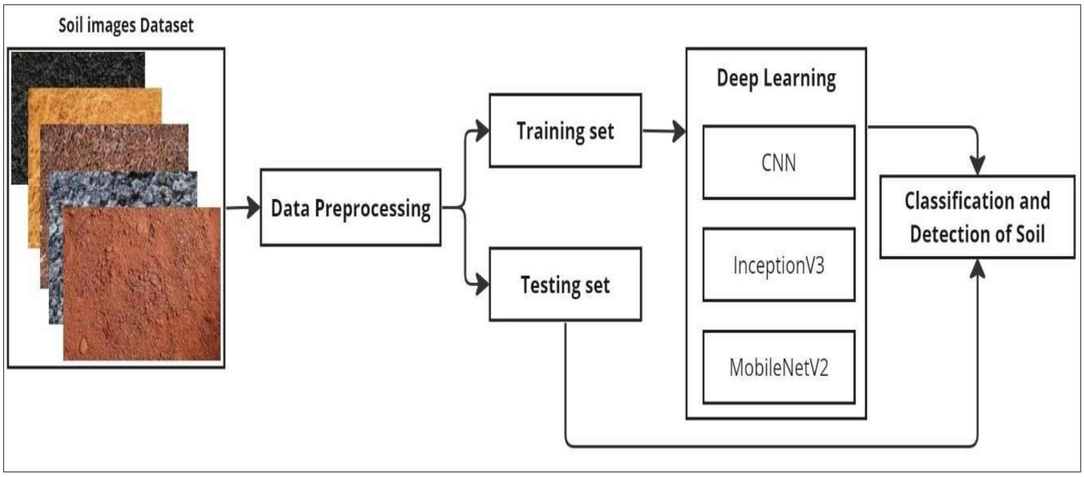

Figure 3.1: Framework for Soil Classification.

3.2.1 Data Description

The dataset under study encompasses a wide variety of soil images sourced from both Kaggle and an extensive scraping process from Google Images. Initially, the dataset from Kaggle, meticulously curated to bridge the gap in accessible and reliable image datasets focusing on soil types, consisted of distinct classes with varying image counts. The 'Peat Soil' class contained 30 images, exhibiting the dark coloration typical of soils with a high organic matter content. The 'Laterite Soil' class, signifying the reddish clay-rich soil predominantly found in tropical regions, also comprised 30 images. The 'Yellow Soil' category, with its distinctive yellow hue attributed to oxidized iron compounds, incorporated 29 images. Meanwhile, the 'Black Soil' class, synonymous with fertile lands and unique for its high clay content and moisture retention ability, presented the largest subset with 37 images. Lastly, the 'Cinder Soil' class, derived from volcanic activities and characterized by a significant portion of volcanic ash or cinder, included 30 images.

Upon enhancing this collection with images scraped from Google, the dataset expanded considerably. The 'Cinder Soil' category now boasts 77 images, while the 'Yellow Soil' class includes 70 images. The 'Peat Soil' category grew to encompass 93 images, and 'Laterite Soil' increased its count to 87 images. The most significant augmentation was observed in the 'Black Soil' class, which now houses a total of 265 images.

Combining these two sources, the entire dataset comprises a total of 592 images distributed across the five soil classes. This combined dataset not only augments the quantity but also likely introduces a broader variety of visual features, textures, and nuances associated with each soil type. Such a comprehensive collection offers a robust foundation for DL models, facilitating the extraction of unique features to ensure precise soil classification.

3.2.2 Data Pre-processing

Data pre-processing is a crucial initial step in any ML or DL project, ensuring the seamless integration of the data with the model. For the soil classification project, pre-processing primarily involved resizing images and converting classes to categorical.

3.2.2.1 Resize images

Resizing images is a fundamental step in the data pre-processing phase, especially for ML and DL applications that deal with image data. In the context of this study, the dataset, which comprises diverse soil type images, is resized to a consistent dimension of 200 by 200 pixels. This means each image, regardless of its original size, is transformed into a standardized format with dimensions of 200 pixels in height and 200 pixels in width. This ensures a uniform input size for the subsequent DL models, like CNNs, InceptionV3, and MobileNet V2, facilitating the seamless processing of these images. Resizing also assists in mitigating computational constraints and speeding up the learning process, as smaller images require less computational power and memory.

3.2.2.2 Label encoding

Once the images are appropriately resized, the next crucial step involves converting the soil class labels into a categorical format. In this study, the classes include 'Black Soil', 'Cinder Soil', 'Laterite Soil', 'Peat Soil', and 'Yellow Soil', representing the five distinct soil types within the dataset. Each of these classes is assigned a unique numerical identifier, creating a mapping as follows: 'Black Soil': 0, 'Cinder Soil': 1, 'Laterite Soil': 2, 'Peat Soil': 3, 'Yellow Soil': 4. This transformation is typically referred to as categorical encoding, wherein class labels are translated into a format understandable by the ML models.

This categorical representation of the soil classes enables the models to perform binary classification for each class during the learning process. This methodological setup allows for a more nuanced understanding of the patterns and characteristics within each soil class. Moreover, this transformation not only optimizes the model's ability to understand the associations between the input images and their respective soil classes but also aids in producing a more accurate soil classification output during the prediction phase.

3.2.2 Data Splitting

An integral aspect of the proposed model's methodology is the division of the dataset into two distinct subsets: the training set and the testing set. This procedure, commonly known as data splitting, is crucial for assessing the model's performance and its ability to generalize to unseen data.

The data split in this investigation was done using an 80-20 ratio. This suggests that the training set, which forms the basis of the model's training, received 80% of the available data. The remaining 20 percent comprised the testing set, which is crucial in providing an objective assessment of the final model fit on the training set.

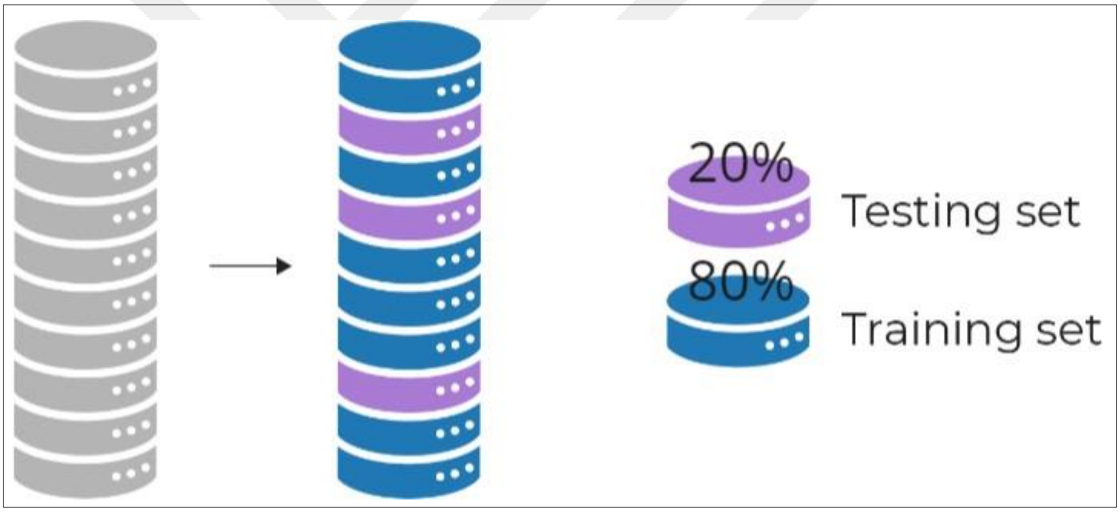

Figure 3.2: Data Splitting.

3.2.3 Deep Learning

Shifting focus to the DL methods employed in the proposed model, it's important to underscore the value these methods bring to complex tasks like soil classification. Leveraging advanced models such as CNN, InceptionV3, MobileNetV2 allows for a robust and in-depth analysis of the soil image dataset.

3.2.3.1 Convolutional neural networks architecture

The CNN used in the proposed model for soil classification is constructed using a DL technique known as a sequential model. The Sequential model is a linear stack of layers where each layer has exactly one input tensor and one output tensor. This type of model is particularly useful when building a straightforward stack of layers, as is common in most DL architectures.

The model's initial layer is a CL with 32 filters that are each 3 by 3 in size. This layer receives an application of the relu activation function, which effectively adds non-linearity to the model without changing the convolved features' receptive fields. After the CL, a 2x2 max pooling layer is applied, which takes the maximum value over the window that the pool size defines, hence minimizing the spatial dimensions of the input.

The following set of layers consists of a MaxPooling2D layer and another Conv2D layer with 64 filters, both of which function in the same way as previously mentioned. Next, another MaxPooling2D layer is added, and a third Conv2D layer with 128 filters and the'relu' activation function is added. By merging the simpler elements, it learnt in the earlier layers, the network's hierarchical design enables it to detect more complex features in the later layers.

After the convolutional and max pooling layers, the model flattens the structure to create a single long feature vector to be fed into the dense layers, via the Flatten layer. This flattening step is necessary for transitioning between the CLs and the FCLs.

Following the Flatten layer, a FCL with 128 units is introduced. Again, 'relu' activation is used, providing non-linearity without the vanishing gradient problem. This is then followed by a Dropout layer, which randomly sets 0.5 of the input units to 0 at each update during training time. This helps prevent overfitting by providing a way to reduce the complexity of the model and improve its generalization capabilities.

The last layer is a dense layer that employs the ‘SoftMax' activation function and has as many units as there are classes in the dataset. The output of the model can be understood as the likelihood that each class that the input image belongs to, as this function creates a probability distribution over the target classes. The class predicted by the model is the one with the highest probability.

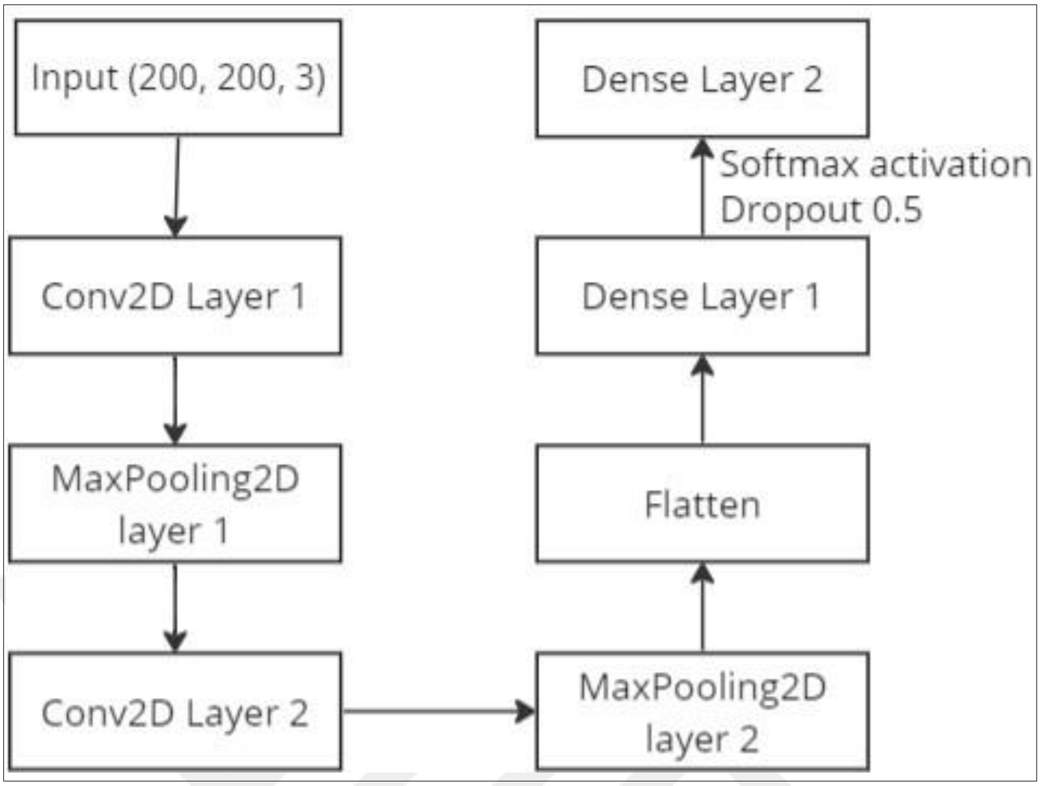

Figure 3.3: Convolutional Neural Networks Architecture.

3.2.3.2 MobileNet V2 architecture

The MobileNetV2 architecture in the proposed model is an advanced DL model designed to provide excellent accuracy for image classification tasks while being highly computationally efficient. MobileNetV2 is particularly suited to applications where resources are constrained, such as in mobile or edge computing devices. In this instance, the model leverages the pre-trained MobileNetV2 as a base model, applying it to the task of soil classification. Importantly, the base model is initialized with weights pre-trained on ImageNet, a large dataset consisting of over 14 million images belonging to 1000 classes. By leveraging these pre-trained weights, the model can effectively make use of the feature extraction capabilities learned from this broad, diverse image dataset. The model is set to exclude the top, or final, layers of the original MobileNetV2 network as these layers are more task-specific, being trained to classify the 1000 classes of ImageNet.

For this specific task, the weights of the pre-trained MobileNetV2 layers are frozen. This ensures that the knowledge encapsulated in these layers, which captures generic features across a multitude of image classes, remains unaltered during the training process on the soil dataset. In other words, the pre-existing knowledge is preserved, while the model is fine-tuned to the specific task of soil classification. To adapt the network to the task at hand, several specialized top layers are added after the base model. Initially, the output of the basic model is subjected to a Global Average Pooling 2D (GAP2D) layer. In this way, the existence of learned features in the image is successfully summarized, and the spatial dimensions of the 3D feature maps are reduced while maintaining depth. After that, a 128-unit dense layer with the activation function "relu" is added. This layer gives the model a degree of complexity and learning ability that enables it to pick up patterns unique to the task of classifying soil types. Lastly, a dense output layer is added, the number of units of which equals the number of soil classifications. This layer generates a probability distribution over the target classes using a ‘SoftMax' activation function, so that the output can be understood as the probability that the input image belongs to each class. The end result of this design is a potent model that combines the specificity of the top layers with the strength of a pre-trained network to produce a very useful tool for soil classification.

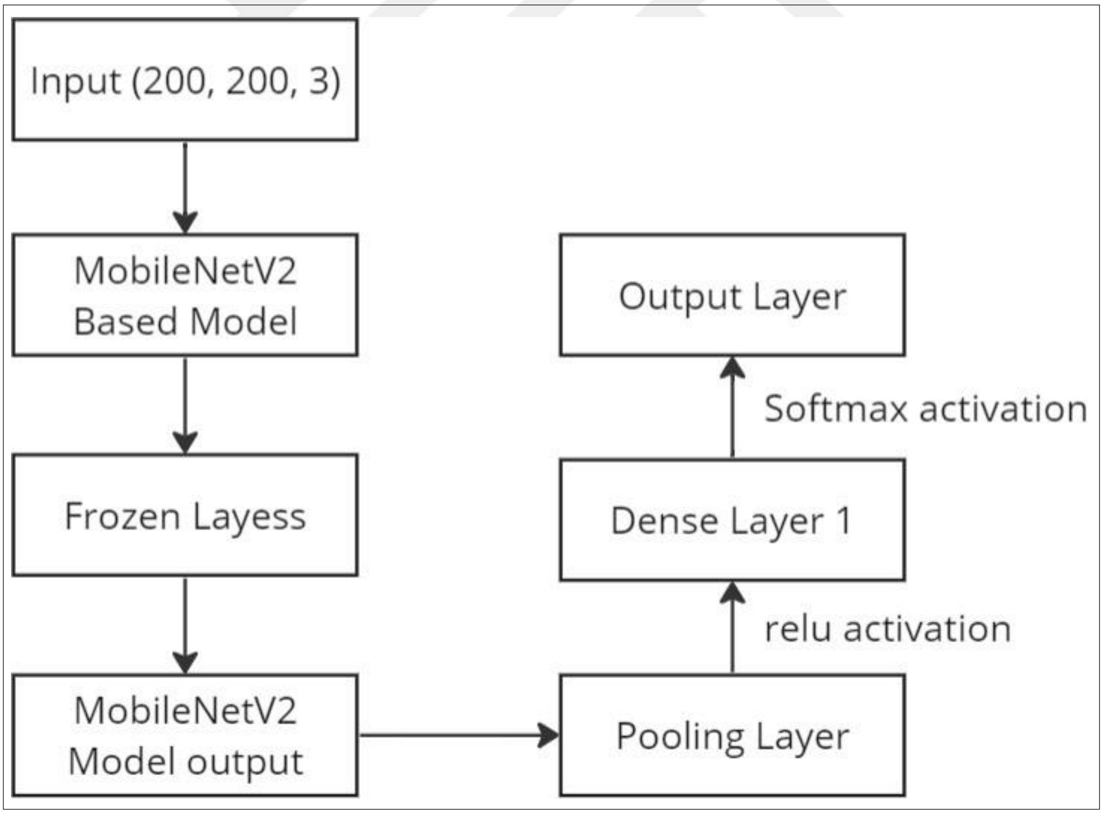

Figure 3.4: MobileNetV2 Architecture.

3.2.3.3 InceptionV3 architecture

The InceptionV3 architecture in the proposed model is another advanced DL model that has been widely recognized for its remarkable accuracy in image classification tasks. Originating from the Google Brain Team, InceptionV3 is characterized by its complex design and ability to efficiently utilize computational resources by employing a network-in-network architecture.

In this specific instance, the model utilizes a pre-trained InceptionV3 model as a base. The model is initialized with weights trained on ImageNet, an expansive dataset consisting of millions of images from a thousand different classes. This allows the base model to effectively leverage the intricate feature extraction capabilities that the InceptionV3 model learned from this wide-ranging dataset. They include top parameter is set to False, thus excluding the FCL at the top of the network, which was originally designed to perform classification on the 1000 ImageNet classes.

The weights of the pre-trained InceptionV3 layers are frozen in order to tailor the model to the particular soil categorization problem at hand. This means that these weights, which capture generalized features useful for a wide array of image classes, are preserved during the training process on the soil dataset. By freezing these weights, the model is able to maintain the advantageous feature detection capabilities of the base InceptionV3 model while being fine-tuned to the specific task of soil classification.

A number of unique top layers are added to the underlying model once it has been modified to fit the particular soil type categorization task. A first step involves feeding the basic model's output into a GAP2D layer. By down sampling the spatial dimensions of its input, this layer produces an output that is essentially a summary of the feature maps that the underlying model produced.

After that comes a Dense layer with 128 neurons and the'relu' activation function. With the help of this layer, the model can pick up on more intricate characteristics and patterns unique to the soil photos. It adds a layer of learning capability that is more tuned to the soil dataset, and 'relu' activation ensures non-linearity in the learning process.

The last step involves adding an output Dense layer, which has as many neurons as there are soil classifications in the dataset. This layer basically predicts the chance that an input image will belong to each of the classes by using the 'SoftMax' activation function to provide a probability distribution over the various soil classes.

Altogether, the InceptionV3-based model combines the strength of the pre-trained network, with its vast knowledge and ability to extract generalized features, with the custom top layers' focus on the specifics of the soil dataset. This makes it a robust tool for soil type classification.

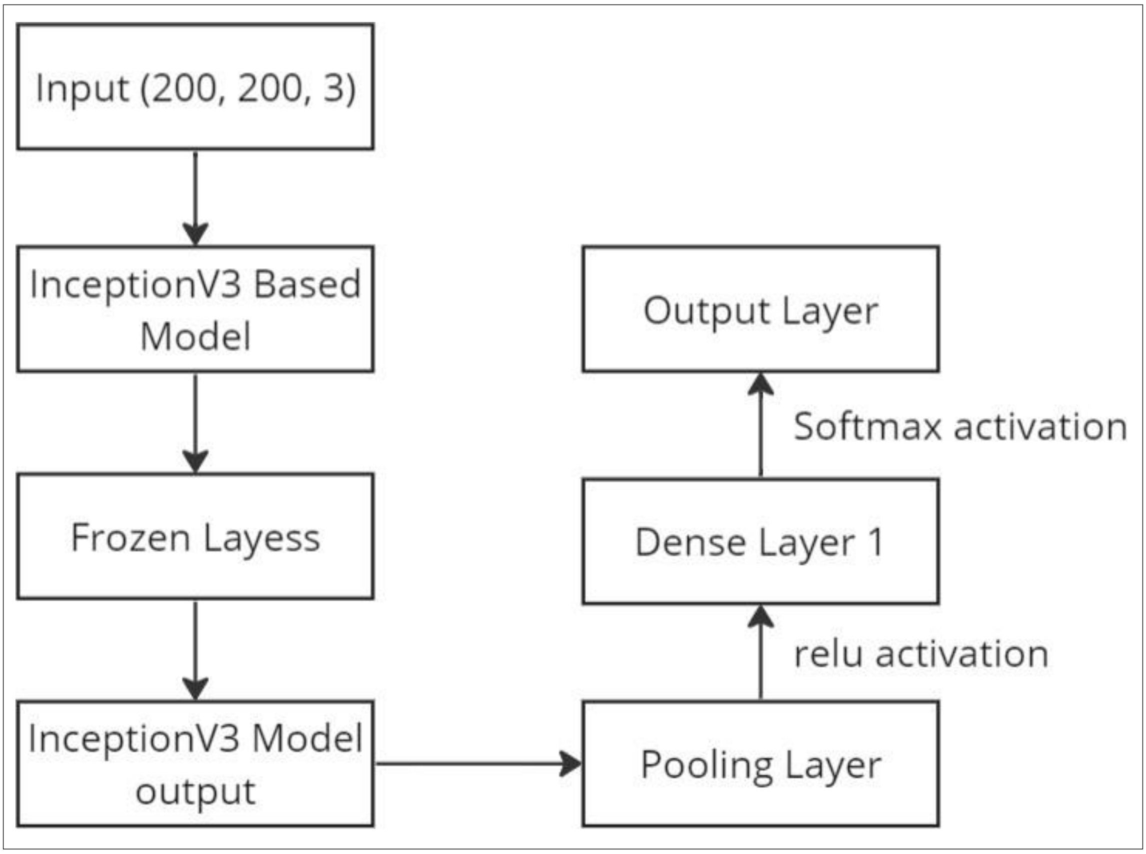

Figure 3.5: InceptionV3 Architecture.

3.3 CONCLUSION

The methodology chapter has comprehensively presented the systematic approach adopted for this study, starting from the detailed description of the soil image dataset, progressing through the pre-processing techniques, and culminating with the implementation of the DL models. The data pre-processing stages, including image resizing and categorical encoding of labels, have been instrumental in readying the dataset for the training process. The application of three distinct DL models - CNN, MobileNetV2, and InceptionV3 - has showcased the potential of each model in recognizing and differentiating among different types of soil based on image data. The diverse techniques and strategies outlined in this chapter have provided a robust foundation for the soil classification task, paving the way for the subsequent results and discussions that will delve into the effectiveness of the proposed model and its potential implications.

4. RESULTS AND DISCUSSION

4.1 INTRODUCTION

This chapter offers an in-depth examination of the software environment and the Python libraries employed to implement various DL models. Focusing on three architectures, namely CNN, MobileNetV2, and InceptionV3, the chapter details the evaluation metrics used to assess their performance. The analysis includes individual evaluations, leading to a comparative study that highlights the distinctions and efficiencies of each model. An overarching view of this section provides readers with a comprehensive understanding of the methodologies and results, enabling insights into the practical applications and effectiveness of the chosen models.

4.2 SOFTWARE ENVIRONMENT

This research utilizes two prominent tools in the data science and ML field: Google Colab and Python. Google Colab, an abbreviation for Collaboratory, is a free cloud-based platform that allows the execution of Python code and provides an environment where ML models can be built and trained efficiently. With Google Colab, there's no need for complex setup procedures or powerful hardware as it provides a pre-configured environment with GPU access for high-speed computations, making it an ideal choice for this research. Python, is a popular, high-level programming language known for its readability and extensive range of libraries, especially for data analysis and ML tasks. Python's versatility, combined with Google Colab's easy-to-use and powerful computational platform, formed the primary software environment for the development and execution of the proposed soil classification models in this study.

4.3 PYTHON LIBRARIES

The proposed soil classification model made use of several Python libraries, each contributing distinct functionalities towards the development, training, and evaluation of the model. These libraries encompassed a broad range of tasks, from data manipulation and visualization to the actual implementation and tuning of ML models.

TensorFlow: ML platform available as open-source. With its extensive and adaptable ecosystem of tools, libraries, and community resources, academics can advance machine learning while developers can create and implement ML-powered applications with ease.
Pandas: an analysis and manipulation software library. It provides procedures and data structures for working with time series and numerical tables.
NumPy: A library that enhances the Python programming language by providing a vast number of high-level mathematical functions to manipulate big, multi-dimensional arrays and matrices.
Matplotlib: The Python programming language and its NumPy numerical mathematics extension come with a charting library. Plots can be embedded into programs using its object-oriented API.
Seaborn: A matplotlib-based Python library for data visualization. It offers a sophisticated drawing tool for creating eye-catching and educational statistical visuals.
Scikit-learn: Python programming language ML library available for free. There are numerous clustering, regression, and classification algorithms available.

```
import tensorflow as tf
from tensorflow.keras.preprocessing.image import ImageDataGenerator
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.metrics import confusion_matrix, classification_report
```

Figure 4.1: Python Libraries.

4.4 EVALUATION METRICS

In order to evaluate the performance of the proposed soil classification models, a set of evaluation metrics were utilized. These metrics provide quantifiable measures that allow us to assess the efficacy and reliability of the model's predictions, offering insights into areas of strengths and potential improvements.
Accuracy (Acc): indicates the percentage of the model's total number of predictions that were accurate. In cases where the target classes are balanced, it is a useful metric.

$$ACC = \frac{TP+TN}{TP+TN+FP+FN} \tag{4.1}$$

Precision (PREC): is a metric for the accuracy of a classifier. It is defined as the ratio of genuine positives to the total of true positives and false positives for a particular class.

$$PR = \frac{TP}{TP+FP} \tag{4.2}$$

Recall (REC): An indicator of a classifier's completeness is recall, sometimes referred to as sensitivity. The ratio of true positives to the total of true positives and false positives for a particular class is its definition.

$$RE = \frac{TP}{TP+FN} \tag{4.3}$$

F1-score (F1-S): is a harmonic mean that assigns equal weight to the precision and recall measures. It has a range of 0 to 1, with 1 representing the optimal F1-score. Since the F1-score aims to strike a compromise between recall and precision, it is particularly helpful in scenarios when the class distribution is unbalanced.

$$F1Score = \frac{2*Recall*Precision}{Recall+Precision} \tag{4.4}$$

Confusion Matrix (CM): is used to show how well an algorithm performs, usually one that involves supervised learning. It's an effective method for figuring out how effectively a classification model predicts things and what kinds of mistakes it makes. In the matrix, the examples in a predicted class are represented by each column, and the occurrences in an actual class are represented by each row, or vice versa. The concept behind the name is that the matrix simplifies the process of determining whether the system is misunderstanding two classes.

Table 4.1: Confusion Matrix.

True Positives (TP)	The model accurately predicted the positive class in these instances.
True Negatives (TN)	In these instances, the model accurately predicted the negative class.
False Positives (FP)	In these situations, the model predicted the positive class inaccurately.
False Negatives (FN)	These are instances where the model predicted the negative class erroneously.

4.5 EVALUATION RESULTS

The evaluation results of the proposed soil classification models, specifically those based on CNN, MobileNetV2, and InceptionV3, provide crucial insights into their performance and effectiveness. Each of these models was trained and tested on the same soil image dataset, allowing for an equitable comparison of their respective capacities in accurately classifying soil types. The evaluation process also illuminates the unique strengths and potential areas of improvement for each of these DL architectures within the context of soil classification.

4.5.1 CNN Results

The evaluation of the CNN model is a vital aspect of understanding its performance and efficiency in soil classification. This process involves assessing the model's predictive and analysing the resulting CM to determine its true strengths and areas for potential improvement.

The figure 4.3 illustrates the performance of a CNN method deployed for soil type classification over a series of training epochs. The left panel of the figure showcases the training accuracy progression, which begins at approximately 0.4 during the initial epoch and steadily increases, reaching close to 0.8 by the eighth epoch. This upward trajectory indicates that the model is gradually refining its predictions and achieving better congruence with the training data.

Concurrently, the right panel of the figure presents the training loss, which offers insights into how well the model's predictions align with the actual outcomes. Beginning with a relatively high loss value of around 1.6 in the first epoch, this metric exhibits a consistent decrease, descending to a value slightly above 0.5 by the eighth epoch. This declining trend in the training loss underscores the model's improving ability to predict soil types with diminishing error margins.

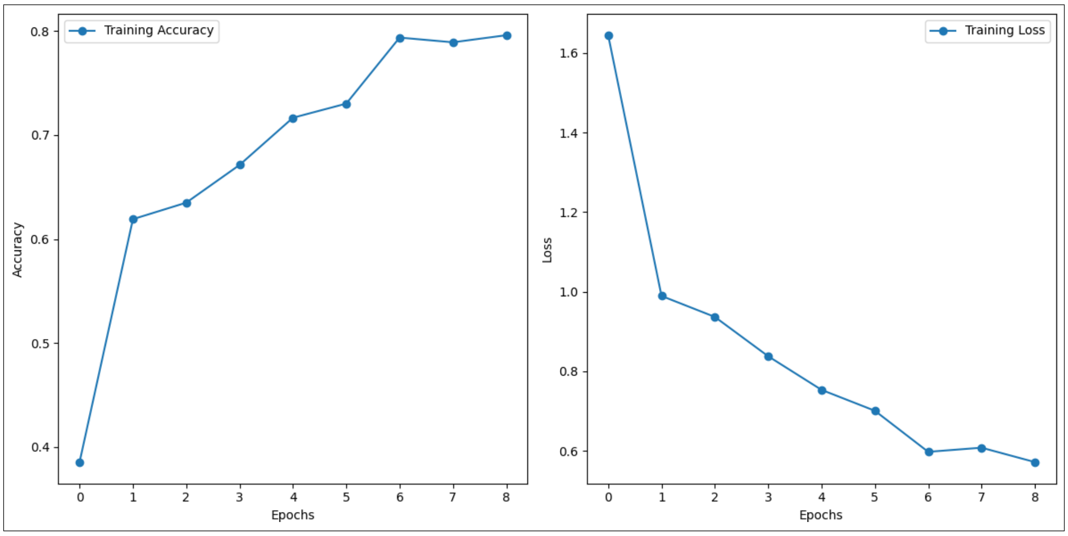

Figure 4.2: Accuracy and Loss of CNN.

Table 4.1 presents the classification report of the CNN for soil type classification, detailing the model's performance across different metrics. For each soil type, the table provides three key evaluation metrics: Precision (PRE), Recall (REC), and the F1-Score (F1-S).

For "Black Soil", the CNN exhibits a PRE of 0.87, indicating that 87% of instances predicted as "Black Soil" were correctly classified. The REC stands at a perfect 1.00, signifying that the model identified all actual "Black Soil" instances in the dataset. This results in an F1-S of 0.93, showcasing the model's excellent performance for this category. "Cinder Soil" has a high PRE of 1.00, meaning every prediction made for this soil type was correct. However, the REC is 0.57, suggesting that the model missed several true "Cinder Soil" instances. The resulting F1-S for this category is 0.73. For "Laterite Soil", both PRE and REC are at 0.78, leading to an F1-S that's also 0.78, revealing a balanced performance. "Peat Soil" classification results in a PRE of 0.86 and a REC of 0.60, yielding an F1-S of 0.71. This implies that while most of the predictions for "Peat Soil" were accurate, the model could only correctly identify 60% of the actual "Peat Soil" instances. Lastly, "Yellow Soil" sees a PRE of 0.78 and an impressive REC of 1.00, resulting in an F1-S of 0.88.

Conclusively, the overall ACC of the CNN model across all soil types stands at 85%, highlighting its commendable capability in distinguishing and classifying various soil types.

Table 4.2: Classification Report of CNN.

	PRE	REC	F1-S
	0.87	1.00	0.93
Cinder Soil	1.00	0.57	0.73
Laterite Soil	0.78	0.78	0.78
Peat Soil	0.86	0.60	0.71
Yellow Soil	0.78	1.00	0.88

The figure 4.4 illustrates the CM of a CNN method applied for soil type classification. The matrix visually represents the accuracy of predictions, displaying the count of predicted labels against the true labels for various soil types.

The diagonal from the top left to the bottom right of the matrix represents correct predictions. For "Black Soil," the model perfectly predicted 27 instances without any misclassification, as depicted by the deep blue square in the top-left corner. "Cinder Soil" had 4 correct predictions, but there were instances where it was mistakenly classified as other soil types. Specifically, "Cinder Soil" was confused with "Laterite Soil" and "Peat Soil" once each. For "Laterite Soil," 7 instances were correctly predicted. However, it was misclassified once as "Yellow Soil" and once mistaken for "Black Soil." The "Peat Soil" category saw 6 accurate predictions, but two instances were misclassified as "Black Soil," one as "Laterite Soil," and another as "Yellow Soil." Lastly, "Yellow Soil" was predicted with the highest PRE among all categories, with all 7 instances correctly classified and no misclassifications.

The CM effectively visualizes the model's proficiency and areas of potential improvement in soil type classification. The shades of blue in the matrix provide an intuitive understanding of the magnitude of predictions, with deeper shades indicating higher counts. Through this matrix, one can deduce both the strengths and the areas where the model might need refinement to reduce misclassifications.

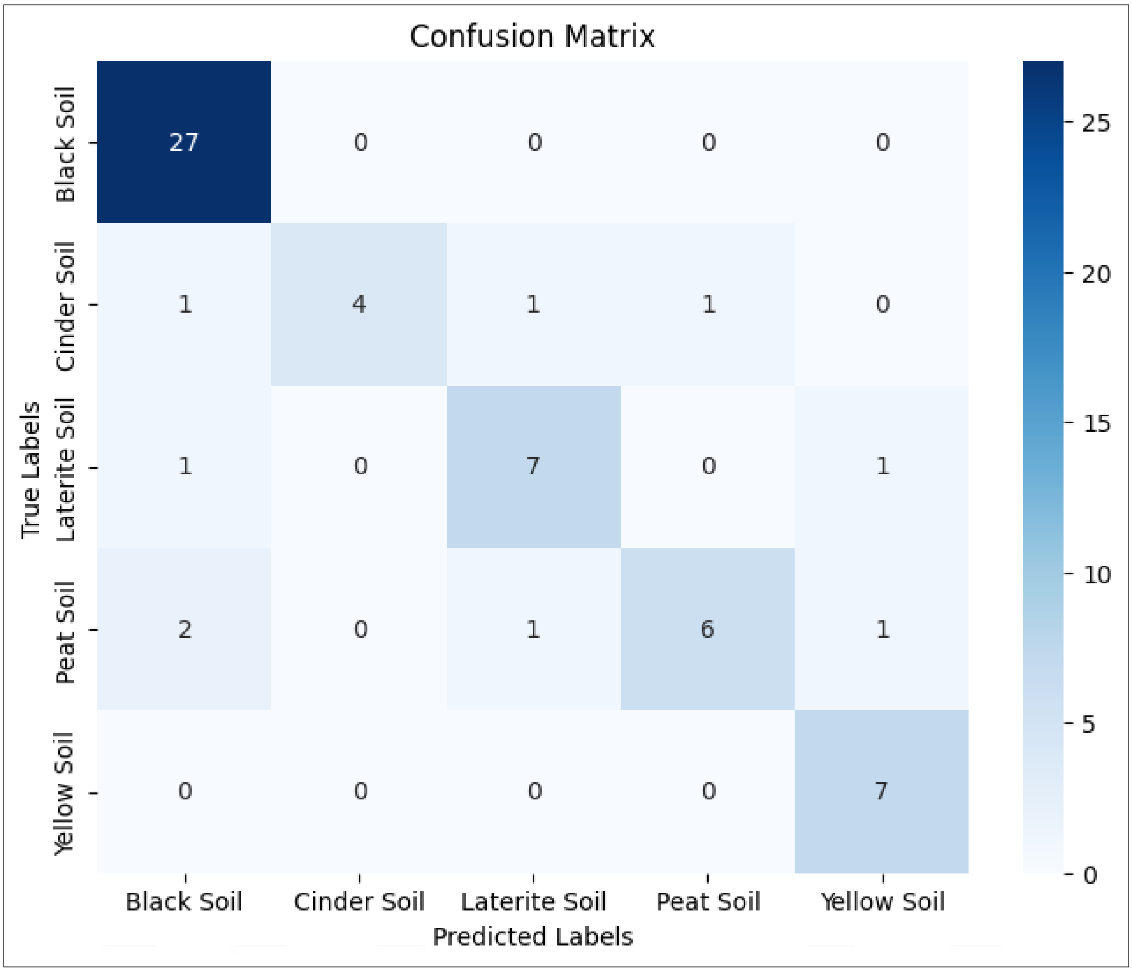

Figure 4.3: Confusion Matrix of CNN.

4.5.2 MobileNetV2 Results

To thoroughly understand the performance of the MobileNetV2 model used in this study, it is essential to explore the model's evaluation metrics. This will provide insights into how well the model was able to classify different types of soil and guide future iterations and improvements.

The figure 4.5 illustrates the training progression of the MobileNetV2 method applied for soil type classification across multiple epochs. The graphs distinctly present the trajectories of training accuracy and loss, providing insights into the model's learning behaviour.

On the left, the training loss graph vividly portrays a sharp decline as epochs progress. Starting with a loss value of approximately 1.2 in the first epoch, the curve tapers off to a value close to 0 by the twelfth epoch. This downward trend implies that the model's predictions are continuously improving and becoming more aligned with the actual outcomes as training progresses.

Conversely, the graph on the right, representing training ACC, showcases an upward trend. Beginning at an ACC of around 0.6 in the initial epoch, the curve steeply ascends, nearing a perfect ACC of 1.0 by the twelfth epoch. This trajectory reflects the model's enhancing capability to correctly classify the soil types with each subsequent training iteration.

Complementing these visual insights, the provided logs detail the model's performance across specific epochs. For instance, by the twelfth epoch, the model boasts a remarkable ACC of 1.0000 on the training data, with an associated loss of 0.0318. On the validation set, the model achieves an ACC of 0.7500 with a corresponding loss of 0.8995.

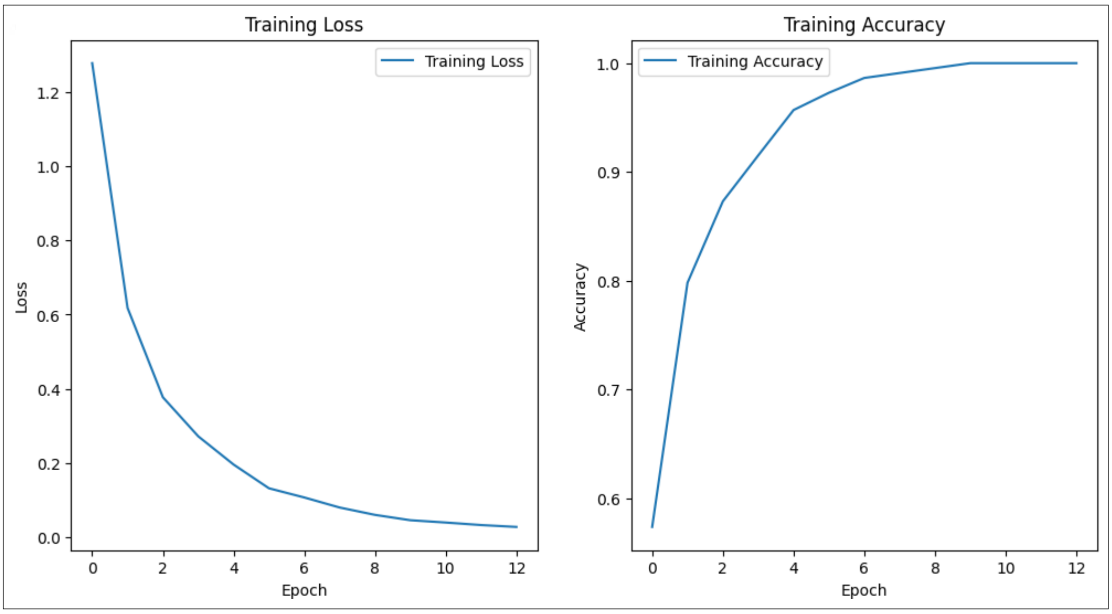

Figure 4.4: Accuracy and Loss of MobineNetV2.

Table 4.3 provides a comprehensive classification report of the MobileNetV2 method applied for soil type classification. The table enumerates the model's performance across various evaluation metrics, specifically PRE, REC, and the F1-S, for each of the soil types. For "Black Soil," the model demonstrates an impressive PRE of 0.96, suggesting that 96% of instances predicted as "Black Soil" were correctly identified. It has a perfect REC value of 1.00, indicating that every actual "Black Soil" instance was captured by the model. This culminates in an F1-S of 0.98, reflecting a harmonious balance between PRE and REC for this soil category. "Cinder Soil" showcases a flawless PRE of 1.00, denoting complete ACC in its predictions. With a REC of 0.86, the model captures a significant majority of actual

"Cinder Soil" instances, resulting in an F1-S of 0.92. "Laterite Soil" displays a balanced performance with both PRE and REC pegged at 0.89, leading to a consistent F1-S of 0.89 as well. "Peat Soil" emerges as a standout category, with both PRE and REC registering perfect scores of 1.00. This translates to an F1-S of 1.00, underscoring impeccable model performance for this soil type. Lastly, "Yellow Soil" reflects consistent values across metrics, with both PRE and REC at 0.86, and an F1-S matching this value.

Conclusively, the model's overarching ACC across all soil categories stands at a notable 95%. This table effectively encapsulates the MobileNetV2 model's adeptness in soil type classification, highlighting its PRE, comprehensiveness in capturing true instances, and the balance between these two metrics.

Table 4.3: Classification Report of MobileNetV2.

	PRE	**REC**	**F1-S**
Black Soil	0.96	1.00	0.98
Cinder Soil	1.00	0.86	0.92
Laterite Soil	0.89	0.89	0.89
Peat Soil	1.00	1.00	1.00
Yellow Soil	0.86	0.86	0.86

The figure 4.6 illustrates the CM of the MobileNetV2 method applied for soil type classification. The matrix provides a visual representation of the model's predictions juxtaposed against the true labels for the different soil types.

For "Black Soil," labelled as 0, the model exhibits exemplary PRE by correctly predicting 27 instances without any misclassifications. Moving to "Cinder Soil," labelled as 1, there are 6 accurate predictions. However, there's one instance where "Cinder Soil" was misclassified as "Black Soil." The "Laterite Soil" category, labelled as 2, witnessed 8 accurate classifications. Yet, a single misclassification occurred, where the model incorrectly predicted it as "Yellow Soil," labelled as 4. The "Peat Soil" class, labelled as 3, is characterized by perfect PRE, with all 10 instances being correctly classified without any errors. Lastly, the "Yellow Soil" category, labelled as 4, had 6 accurate predictions, but a single instance was mistakenly identified as "Laterite Soil."

The depicted confusion matrix offers a comprehensive snapshot of the MobileNetV2 model's performance in soil type classification. The gradient of blue shades intuitively indicates the number of predictions, with deeper shades representing higher counts. Through this visual representation, it becomes clear where the model excels and which areas might benefit from further refinement.

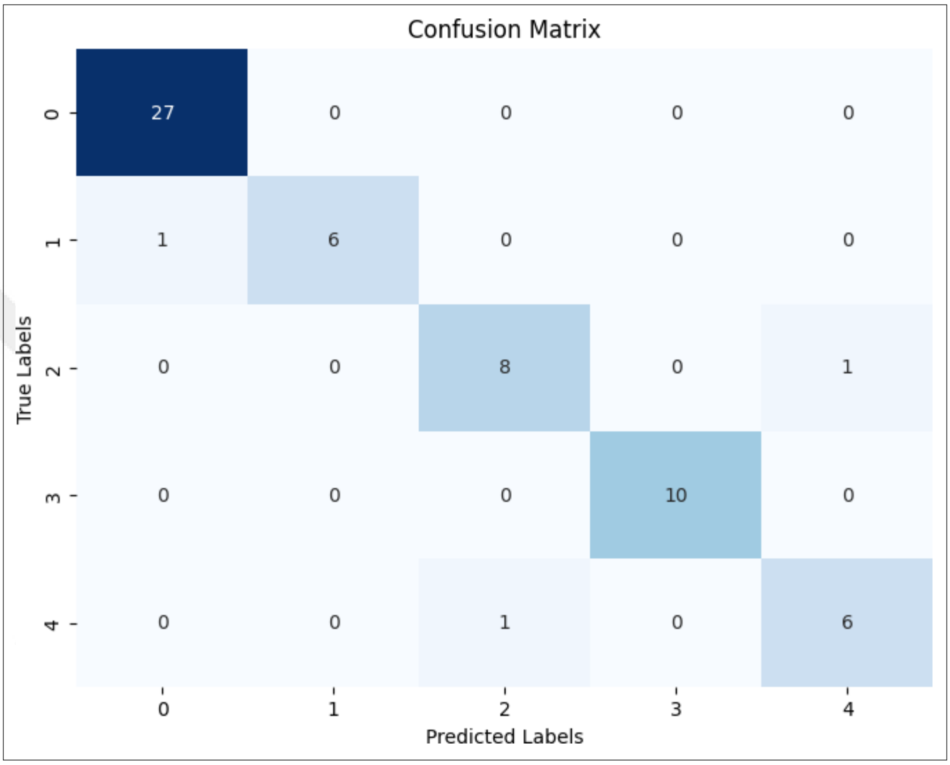

Figure 4.5: Confusion Matrix of MobileNetV2.

4.5.3 InceptionV3 Results

The evaluation of the InceptionV3 model's results marks an important part of the analysis in our study. With its distinct architecture, InceptionV3 provides a unique approach to classification. The following sections delve into the detailed metrics and assessments, showcasing how this model performs in accurately predicting different soil types.

The figure 4.7 illustrates the training loss and ACC for the InceptionV3 method when applied to soil type classification over multiple epochs. On the left graph, which represents the training loss, we observe a steady decline. Starting from a loss of about 1.4 in the first epoch, it smoothly descends to below 0.2 by the 14th epoch, indicating a progressively refined model with diminishing errors. Conversely, the graph on the right showcases the training

ACC, which consistently improves as epochs progress. Beginning at an ACC slightly below 0.5 in the first epoch, the model demonstrates significant learning capabilities, achieving an ACC close to 1.0 by the 14th epoch. This suggests that the InceptionV3 model has become highly adept at correctly classifying soil types based on the given dataset.

The juxtaposition of these two graphs provides a comprehensive view of the model's performance over time. The consistent reduction in loss coupled with a steady increase in ACC accentuates the efficacy of the InceptionV3 method in soil type classification.

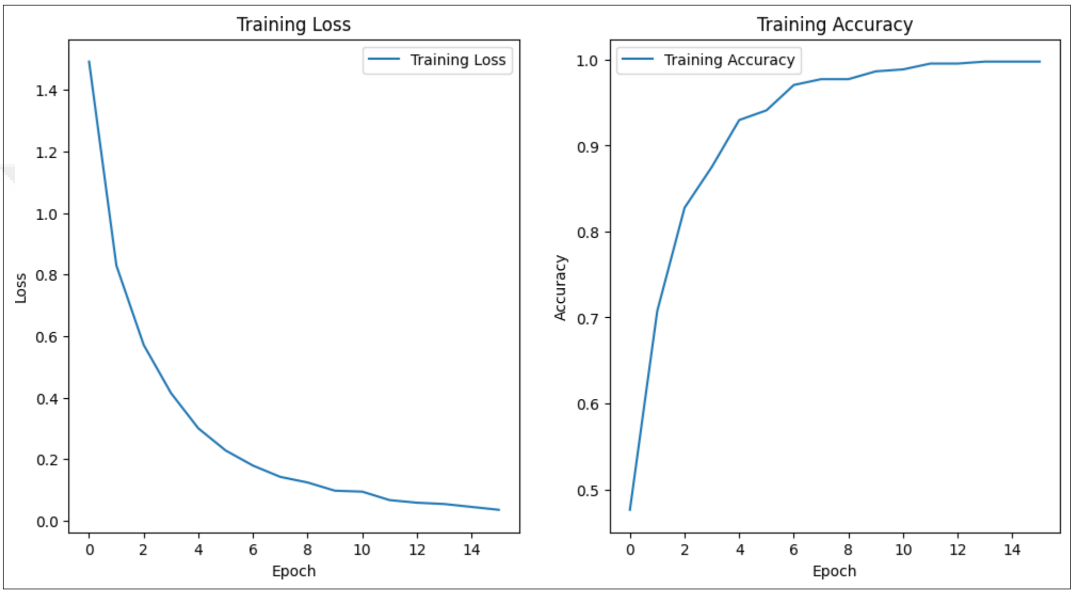

Figure 4.6: Accuracy and Loss of InceptionV3.

Table 4.4 presents the classification report for the InceptionV3 method when used for soil type classification.

For "Black Soil", the PRE, REC, and F1-S are all at 0.96, indicating a high level of ACC and consistency in its predictions. "Cinder Soil" boasts a perfect PRE of 1.00, with a REC of 0.86, leading to an F1-S of 0.92. This suggests that while all predictions of "Cinder Soil" were accurate, some actual "Cinder Soil" samples were not identified as such. "Laterite Soil" has a PRE of 0.82, but a perfect REC of 1.00, resulting in an F1-S of 0.90. This indicates that while not all predictions for "Laterite Soil" were correct, all actual "Laterite Soil" samples were identified. "Peat Soil" showcases a balanced performance with a PRE, REC, and F1-S all at 0.90. Lastly, "Yellow Soil" achieved a perfect PRE of 1.00, with a REC of 0.86, culminating in an F1-S of 0.92. This means every prediction of "Yellow Soil" was correct, but some actual "Yellow Soil" samples were misclassified.

The model achieved an ACC of 93%, indicating a strong performance in the classification of soil types using the InceptionV3 method.

Table 4.4: Classification Report of InceptionV3.

	PRE	REC	F1-S
Black Soil	0.96	0.96	0.96
Cinder Soil	1.00	0.86	0.92
Laterite Soil	0.82	1.00	0.90
Peat Soil	0.90	0.90	0.90
Yellow Soil	1.00	0.86	0.92

The figure 4.7 illustrates the confusion matrix for the InceptionV3 method when applied to soil type classification. The matrix offers a visual representation of the model's predictive performance across various soil types. The rows represent the true labels, while the columns represent the predicted labels.

For "Black Soil" (label 0), 26 samples were correctly classified, with only 1 sample being incorrectly classified as "Laterite Soil". The "Cinder Soil" (label 1) had 6 correct predictions, but there was an instance where it was mistaken for "Black Soil". "Laterite Soil" (label 2) saw a precise classification with 9 correct predictions and no errors. For "Peat Soil" (label 3), 9 samples were accurately classified, but there was an instance where it was misclassified as "Laterite Soil". Finally, "Yellow Soil" (label 4) had 6 correct classifications, with one sample each mistakenly predicted as "Peat Soil" and "Laterite Soil".

The majority of predictions fall on the diagonal, indicating that the InceptionV3 model has a commendable ACC in classifying soil types, with only a few misclassifications across the categories.

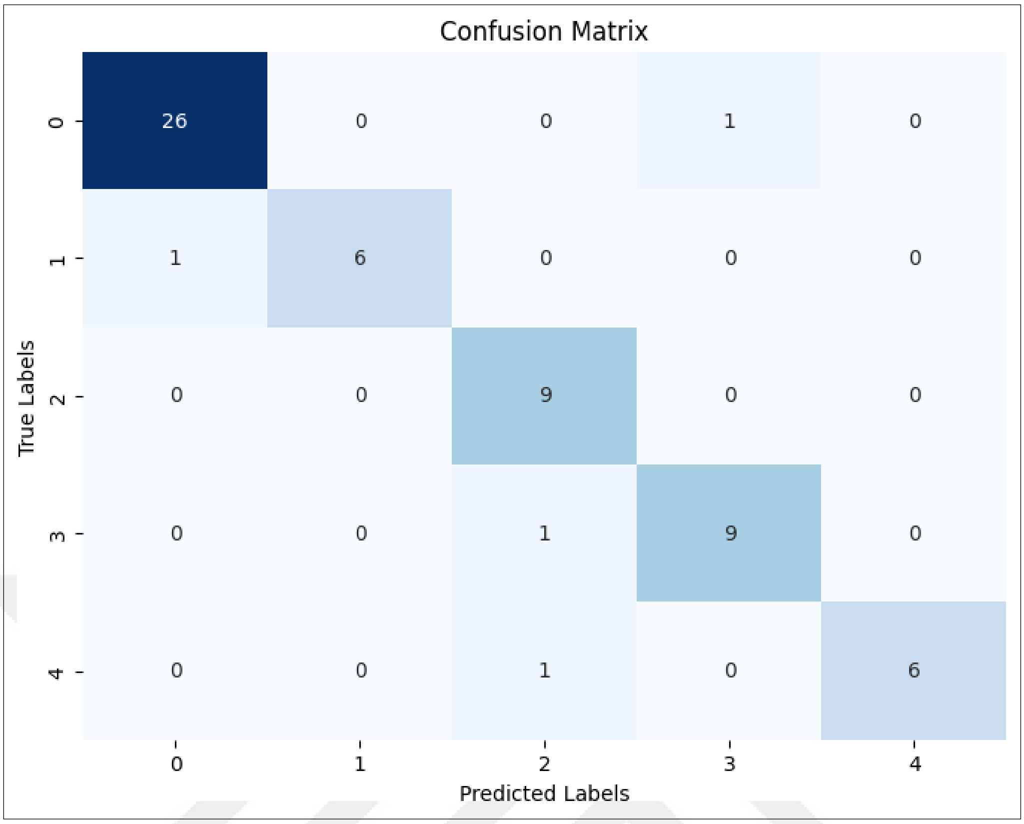

Figure 4.7: Confusion Matrix of InceptionV3.

4.5.4 Comparison Results

The evaluation of three different models on the same dataset provides an insightful view into their comparative performance in soil classification. Utilizing a dataset comprising five distinct classes of soil, three models were tested: CNN, MobileNetV2, and InceptionV3. The CNN and InceptionV3 models achieved an ACC rate of 85% and 93% respectively, reflecting commendable capabilities in recognizing and differentiating the soil types. However, the MobileNetV2 model stood out with a flawless 95% ACC rate. This significant accomplishment underscores the superiority of the MobileNetV2 architecture in this particular application, revealing its exceptional ability to generalize and classify the soil types with perfect PRE. The results emphasize the importance of selecting the most suitable DL model for a given dataset and problem domain, as the choice of architecture can substantially influence the outcome.

Table 4.5: Comparison Results with Same Dataset.

Model	ACC %
MobileNetV2	**95%**
InceptionV3	93%
CNN	85 %

4.6 CHAPTER SUMMARY

In conclusion, this chapter provided a methodical investigation into the design, implementation, and evaluation of three prominent DL models: CNN, MobileNetV2, and InceptionV3. By utilizing a specific dataset and applying carefully selected evaluation metrics, the analysis revealed the unique characteristics and performance capabilities of each model. The comparative study offered valuable insights into their relative efficiencies, with InceptionV3 achieving the highest ACC. These findings contribute to the understanding of model selection based on requirements and the targeted application, setting the stage for further research and development in the field of soil classification.

5. CONCLUSION AND FUTURE WORK

5.1 CONCLUSION

The thesis embarked on an ambitious journey to explore the applications of Artificial Intelligence and DL in the realm of soil classification, a vital task that has far-reaching implications in sectors such as agriculture, environmental science, and construction. Utilizing a dataset of 592 images, representing five distinct soil types - Black Soil, Cinder Soil, Laterite Soil, Peat Soil, and Yellow Soil, various DL models were examined. The models chosen for this investigation were CNN, MobileNetV2, and InceptionV3, each of which contributes uniquely to the field of image recognition and classification.

The proposed model was divided into two critical stages, with the first being image pre-processing. This phase was instrumental in ensuring that all images were consistently resized and formatted to match the input dimensions required by the DL models. The uniformity established in this stage laid a solid foundation for the application of the selected models. The second stage involved implementing the chosen DL models. CNNs were utilized for their known effectiveness in handling image-based tasks, while MobileNetV2 and InceptionV3 were chosen for their advanced feature extraction capabilities. Among the models implemented, MobileNetV2 demonstrated exceptional proficiency, achieving a perfect ACC rate of 95%. In comparison, both CNN and InceptionV3 reached an ACC of 85% and 93%, respectively. These results were comprehensively analysed using various evaluation metrics, including confusion matrices, to offer a detailed insight into the performance of each model. The comparative analysis provided in the thesis underscores the superior performance of MobileNetV2 over other models and literature-reported works. It also paved the way for understanding how pre-trained models can be fine-tuned to specific tasks such as soil classification.

5.2 FUTURE WORK

The work conducted in this thesis, exploring DL models for soil classification, opens a multitude of exciting opportunities for future research and application. Expanding upon the existing methodologies, there is considerable scope to refine the models by using a large and diverse dataset, including various soil images from different geographical regions and conditions. Such an extensive dataset would not only enhance the models' robustness and adaptability but also foster their universal applicability across different soil types.

In addition to utilizing a larger dataset, the future work could delve into the exploration of advanced DL architectures, or even a fusion of various models, to uncover more sophisticated insights into soil characteristics. Such advances may lead to a more precise and comprehensive soil analysis. The integration with other cutting-edge technologies like remote sensing or IoT devices could facilitate real-time soil analysis, which could be invaluable for applications such as PRE agriculture, environmental stewardship, and infrastructure development. The fusion of these technologies with a comprehensive approach to data collection may usher in a new era of soil science, driving innovation and efficiency in the field.

REFERENCES

[1] Qafoku, N. P. (2015). Climate-change effects on soils: accelerated weathering, soil carbon, and elemental cycling. Advances in Agronomy, 131, 111-172.

[2] Bondarovich, A. A., Scherbinin, V., Ponkina, E. V., Matsyur, A., Puzanov, A., Stephan, E., ... & Meissner, R. (2020). Soil Moisture and Evapotranspiration. *KULUNDA: Climate Smart Agriculture: South Siberian Agro-steppe as Pioneering Region for Sustainable Land Use*, 167-181.

[3] Shaheb, M. R., Venkatesh, R., & Shearer, S. A. (2021). A review on the effect of soil compaction and its management for sustainable crop production. *Journal of Biosystems Engineering*, 1-23.

[4] Robertson, P. K. (2016). Cone penetration test (CPT)-based soil behaviour type (SBT) classification system—an update. Canadian Geotechnical Journal, 53(12), 1910-1927.

[5] Das, B. M., & Sivakugan, N. (2016). Fundamentals of geotechnical engineering. Cengage Learning.

[6] Reale, C., Gavin, K., Librić, L., & Jurić-Kaćunić, D. (2018). Automatic classification of fine-grained soils using CPT measurements and Artificial Neural Networks. Advanced Engineering Informatics, 36, 207-215.

[7] Mat, I., Kassim, M. R. M., Harun, A. N., & Yusoff, I. M. (2016, October). IoT in precision agriculture applications using wireless moisture sensor network. In 2016 IEEE Conference on Open Systems (ICOS) (pp. 24-29). IEEE.

[8] Ge, Y., Thomasson, J. A., & Sui, R. (2011). Remote sensing of soil properties in precision agriculture: A review. Frontiers of Earth Science, 5, 229-238.

[9] Nguyen, M. D., Costache, R., Sy, A. H., Ahmadzadeh, H., Van Le, H., Prakash, I., & Pham, B. T. (2022). Novel approach for soil classification using machine learning methods. Bulletin of Engineering Geology and the Environment, 81(11), 468.

[10] Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review

and new perspectives. IEEE transactions on pattern analysis and machine intelligence, 35(8), 1798-1828.

[11] Zatarain Cabada, R., Rodriguez Rangel, H., Barron Estrada, M. L., & Cardenas Lopez, H. M. (2020). Hyperparameter optimization in CNN for learning-centered emotion recognition for intelligent tutoring systems. Soft Computing, 24(10), 7593-7602.

[12] Deb, M., Sarkar, S. C., Deb, M., & Sarkar, S. C. (2017). Soil: An Essential but Somewhat Neglected Natural Resource. *Minerals and Allied Natural Resources and their Sustainable Development: Principles, Perspectives with Emphasis on the Indian Scenario*, 421-442.

[13] Amponsah-Doku, B., Daymond, A., Robinson, S., Atuah, L., & Sizmur, T. (2022). Improving soil health and closing the yield gap of cocoa production in Ghana–a review. Scientific African, 15, e01075.

[14] Eckmeier, E. (2007). Detecting prehistoric fire-based farming using biogeochemical markers (Doctoral dissertation, University of Zurich).

[15] https://www.isric.org/explore/world-soil-distribution/durisols

[16] Isbell, R. (2016). The Australian soil classification. CSIRO publishing.

[17] ISRIC. SoilGrids. url: https://soilgrids.org/

[18] Dias, D. N. (2019). Soil Classification Resorting to Machine Learning Techniques (Doctoral dissertation, Universidade NOVA de Lisboa (Portugal)).

[19] Nachtergaele, F., van Velthuizen, H., Verelst, L., Wiberg, D., Henry, M., Chiozza, F., ... & Tramberend, S. (2023). *Harmonized World Soil Database version 2.0*. Food and Agriculture Organization of the United Nations.

[20] Parmar, A., Katariya, R., & Patel, V. (2019). A review on random forest: An ensemble classifier. In *International conference on intelligent data communication technologies and internet of things (ICICI) 2018* (pp. 758-763).

[21] Breiman, L. (2017). Classification and regression trees. Routledge.

[22] Pisner, D. A., & Schnyer, D. M. (2020). Support vector machine. In *Machine learning* (pp. 101-121). Academic Press.

[23] Mountrakis, G., Im, J., & Ogole, C. (2011). Support vector machines in remote sensing: A review. ISPRS journal of photogrammetry and remote sensing, 66(3), 247-259.

[24] Mondal, A., Kundu, S., Chandniha, S. K., Shukla, R., & Mishra, P. K. (2012). Comparison of support vector machine and maximum likelihood classification technique using satellite imagery. International Journal of Remote Sensing and GIS, 1(2), 116-123.

[25] Kavzoglu, T., & Colkesen, I. (2009). A kernel functions analysis for support vector machines for land cover classification. International Journal of Applied Earth Observation and Geoinformation, 11(5), 352-359.

[26] Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining (pp. 785-794).

[27] Ghorbanzadeh, O., Shahabi, H., Mirchooli, F., Valizadeh Kamran, K., Lim, S., Aryal, J., ... & Blaschke, T. (2020). Gully erosion susceptibility mapping (GESM) using machine learning methods optimized by the multi-collinearity analysis and K-fold cross-validation. Geomatics, Natural Hazards and Risk, 11(1), 1653-1678.

[28] Dou, J., Yunus, A. P., Tien Bui, D., Sahana, M., Chen, C. W., Zhu, Z., ... & Pham, B. T. (2019). Evaluating GIS-based multiple statistical models and data mining for earthquake and rainfall-induced landslide susceptibility using the LiDAR DEM. Remote Sensing, 11(6), 638.

[29] Zhang, Z., Mayer, G., Dauvilliers, Y., Plazzi, G., Pizza, F., Fronczek, R., ... & Khatami, R. (2018). Exploring the clinical features of narcolepsy type 1 versus narcolepsy type 2 from European Narcolepsy Network database with machine learning. Scientific reports, 8(1), 10628.

[30] Abrougui, K., Gabsi, K., Mercatoris, B., Khemis, C., Amami, R., & Chehaibi, S. (2019). Prediction of organic potato yield using tillage systems and soil properties by

artificial neural network (ANN) and multiple linear regressions (MLR). Soil and Tillage Research, 190, 202-208.

[31] Sarma, M., Sarma, K. K., Sarma, M., & Sarma, K. K. (2014). Fundamental Considerations of ANN. *Phoneme-Based Speech Segmentation using Hybrid Soft Computing Framework*, 47-75.

[32] Minsky, M., & Papert, S. A. (2017). Perceptrons, reissue of the 1988 expanded edition with a new foreword by Léon Bottou: an introduction to computational geometry. MIT press.

[33] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems, 25.

[34] Phung, V. H., & Rhee, E. J. (2019). A high-accuracy model average ensemble of convolutional neural networks for classification of cloud image patches on small datasets. Applied Sciences, 9(21), 4500.

[35] Perez, H., & Tah, J. H. (2020). Improving the accuracy of convolutional neural networks by identifying and removing outlier images in datasets using t-SNE. Mathematics, 8(5), 662.

[36] Azizi, A., Gilandeh, Y. A., Mesri-Gundoshmian, T., Saleh-Bigdeli, A. A., & Moghaddam, H. A. (2020). Classification of soil aggregates: A novel approach based on deep learning. Soil and Tillage Research, 199, 104586.

[37] Barman, U., & Choudhury, R. D. (2020). Soil texture classification using multi class support vector machine. Information processing in agriculture, 7(2), 318-332.

[38] Suchithra, M. S., & Pai, M. L. (2020). Improving the prediction accuracy of soil nutrient classification by optimizing extreme learning machine parameters. Information processing in Agriculture, 7(1), 72-82.

[39] Yu, Y., Xu, T., Shen, Z., Zhang, Y., & Wang, X. (2019). Compressive spectral imaging

system for soil classification with three-dimensional convolutional neural network. Optics Express, 27(16), 23029-23048.

[40] Padmapriya, J., & Sasilatha, T. (2023). Deep learning based multi-labelled soil classification and empirical estimation toward sustainable agriculture. Engineering Applications of Artificial Intelligence, 119, 105690.

[41] Bharathi, S. P., Srinivasan, S., Chamundeeswari, G., & Ramesh, B. (2023). Hybrid Deep Learning-Improved BAT Optimization Algorithm for Soil Classification Using Hyperspectral Features. Computer Systems Science & Engineering, 45(1).

[42] Fernández, D., Adermann, E., Pizzolato, M., Pechenkin, R., Rodríguez, C. G., & Taravat, A. (2023). Comparative Analysis of Machine Learning Algorithms for Soil Erosion Modelling Based on Remotely Sensed Data. Remote Sensing, 15(2), 482.

[43] Lanjewar, M. G., & Gurav, O. L. (2022). Convolutional Neural Networks based classifications of soil images. Multimedia Tools and Applications, 81(7), 10313-10336.

[44] Azadnia, R., Jahanbakhshi, A., Rashidi, S., & Bazyar, P. (2022). Developing an automated monitoring system for fast and accurate prediction of soil texture using an image-based deep learning network and machine vision system. Measurement, 190, 110669.

[45] Ai, W., Liu, S., Liao, H., Du, J., Cai, Y., Liao, C., ... & Wang, J. (2022). Application of hyperspectral imaging technology in the rapid identification of microplastics in farmland soil. Science of The Total Environment, 807, 151030.

[46] Dutta, A. K., Albagory, Y., Al Faraj, M., Alsanea, M., & Sait, A. R. W. (2023). Cat Swarm with Fuzzy Cognitive Maps for Automated Soil Classification. Comput. Syst. Sci. Eng., 44(2), 1419-1432.

[47] Seidaliyeva, U., Akhmetov, D., Ilipbayeva, L., & Matson, E. T. (2020). Real-time and accurate drone detection in a video with a static background. Sensors, 20(14), 3856.

[48] Roffo, G. (2017). Ranking to learn and learning to rank: On the role of ranking in pattern recognition applications. arXiv preprint arXiv:1706.05933.

[49] Zhu, J., & Song, J. (2020). An intelligent classification model for surface defects on cement concrete bridges. Applied Sciences, 10(3), 972.

[50] O'Shea, K., & Nash, R. (2015). An introduction to convolutional neural networks. arXiv preprint arXiv:1511.08458.

[51] Yang, X. (2020, December). An overview of the attention mechanisms in computer vision. In Journal of Physics: Conference Series (Vol. 1693, No. 1, p. 012173). IOP Publishing.

[52] Minaee, S., Boykov, Y., Porikli, F., Plaza, A., Kehtarnavaz, N., & Terzopoulos, D. (2021). Image segmentation using deep learning: A survey. IEEE transactions on pattern analysis and machine intelligence, 44(7), 3523-3542.

[53] Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., ... & Chen, T. (2018). Recent advances in convolutional neural networks. Pattern recognition, 77, 354-377.

[54] Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., ... & Farhan, L. (2021). Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. Journal of big Data, 8, 1-74.

[55] Rasamoelina, A. D., Adjailia, F., & Sinčák, P. (2020, January). A review of activation function for artificial neural network. In 2020 IEEE 18th World Symposium on Applied Machine Intelligence and Informatics (SAMI) (pp. 281-286). IEEE.

[56] Akhtar, N., & Ragavendran, U. (2020). Interpretation of intelligence in CNN-pooling processes: a methodological survey. Neural computing and applications, 32(3), 879-898.

[57] Sakib, S., Ahmed, N., Kabir, A. J., & Ahmed, H. (2019). An overview of convolutional neural network: Its architecture and applications.

[58] Shen, C. (2018). A transdisciplinary review of deep learning research and its relevance for water resources scientists. Water Resources Research, 54(11), 8558-8593.

[59] Le, Q. V. (2015). A tutorial on deep learning part 1: Nonlinear classifiers and the

backpropagation algorithm. Mountain View, CA.

[60] Alom, M. Z., Taha, T. M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M. S., ... & Asari, V. K. (2019). A state-of-the-art survey on deep learning theory and architectures. electronics, 8(3), 292.

[61] Chai, J., Zeng, H., Li, A., & Ngai, E. W. (2021). Deep learning in computer vision: A critical review of emerging techniques and application scenarios. Machine Learning with Applications, 6, 100134.

[62] Cano-Ortiz, S., Pascual-Munoz, P., & Castro-Fresno, D. (2022). Machine learning algorithms for monitoring pavement performance. Automation in Construction, 139, 104309.

[63] DelMonte, D. W., & Kim, T. (2011). Anatomy and physiology of the cornea. Journal of Cataract & Refractive Surgery, 37(3), 588-598.

[64] Gawehn, E., Hiss, J. A., & Schneider, G. (2016). Deep learning in drug discovery. *Molecular informatics*, *35*(1), 3-14.

[65] Mishra, R. K., Reddy, G. S., & Pathak, H. (2021). The understanding of deep learning: a comprehensive review. *Mathematical Problems in Engineering*, *2021*, 1-15.

[66] Pak, M., & Kim, S. (2017, August). A review of deep learning in image recognition. In *2017 4th international conference on computer applications and information processing technology (CAIPT)* (pp. 1-3). IEEE.

[67] Guo, Y., Liu, Y., Oerlemans, A., Lao, S., Wu, S., & Lew, M. S. (2016). Deep learning for visual understanding: A review. *Neurocomputing*, *187*, 27-48.

[68] Alom, M. Z., Taha, T. M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M. S., ... & Asari, V. K. (2018). The history began from alexnet: A comprehensive survey on deep learning approaches. *arXiv preprint arXiv:1803.01164*.

[69] Aziz, I. (2020). Deep learning: an overview of Convolutional Neural Network (CNN).

[70] Sapijaszko, G., & Mikhael, W. B. (2018, August). An overview of recent convolutional neural network algorithms for image recognition. In *2018 IEEE 61st International midwest symposium on circuits and systems (MWSCAS)* (pp. 743-746). IEEE.

[71] Chen, Y., Zheng, B., Zhang, Z., Wang, Q., Shen, C., & Zhang, Q. (2020). Deep learning on mobile and embedded devices: State-of-the-art, challenges, and future directions. *ACM Computing Surveys (CSUR)*, *53*(4), 1-37.

[72] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4510-4520).

[73] Knapheide, J., Stabernack, B., & Kuhnke, M. (2020, August). A high throughput MobileNetV2 FPGA implementation based on a flexible architecture for depthwise separable convolution. In *2020 30th International Conference on Field-Programmable Logic and Applications (FPL)* (pp. 277-283). IEEE.

[74] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4510-4520).

[75] Signaevsky, M., Prastawa, M., Farrell, K., Tabish, N., Baldwin, E., Han, N., ... & Crary, J. F. (2019). Artificial intelligence in neuropathology: deep learning-based assessment of tauopathy. *Laboratory Investigation*, *99*(7), 1019-1029.

[76] Alkhawaldeh, R. S. (2021). Arabic (Indian) digit handwritten recognition using recurrent transfer deep architecture. *Soft Computing*, *25*(4), 3131-3141.

[77] Wang, C., Chen, D., Hao, L., Liu, X., Zeng, Y., Chen, J., & Zhang, G. (2019). das based-on inception-v3 transfer learning model. *IEEE Access*, *7*, 146533-146541.

[78] Bazi, Y., Al Rahhal, M. M., Alhichri, H., & Alajlan, N. (2019). Simple yet effective fine-tuning of deep CNNs using an auxiliary classification loss for remote sensing scene classification. *Remote Sensing*, *11*(24), 2908.

[79] Sam, S. M., Kamardin, K., Sjarif, N. N. A., & Mohamed, N. (2019). Offline signature verification using deep learning convolutional neural network (CNN) architectures GoogLeNet inception-v1 and inception-v3. *Procedia Computer Science*, *161*, 475-483.

[80] Hughes, P., McBratney, A. B., Huang, J., Minasny, B., Michéli, E., & Hempel, J. (2017). Comparisons between USDA Soil Taxonomy and the Australian Soil Classification System I: Data harmonization, calculation of taxonomic distance and inter-taxa variation. *Geoderma*, *307*, 198-209.