SoilNet – Soil Texture Classification using Convolutional Neural Network

By Lam Hoang Phan

Abstract

Soil texture classification is a critical task in agriculture, as it directly impacts water retention, nutrient availability, and crop planning. Traditional approaches rely on manual inspection and laboratory analysis, which are time-consuming and subject to human error. This study introduces SoilNet, a Convolutional Neural Network (CNN)-based model designed to automate soil texture classification using image data. The proposed architecture consists of four convolutional blocks for hierarchical feature extraction, followed by a fully connected layer that outputs class probabilities for three texture categories: coarse, medium, and fine. The model was trained and evaluated on a combined dataset of 3,702 soil images from Roboflow's SOIL (v1) dataset and additional Kaggle samples. After hyperparameter tuning - optimizing epochs, learning rate, kernel size, and activation function - SoilNet achieved an outstanding 99.46% accuracy, 0.0182 test loss, and near-perfect precision, recall, and F1-scores. These results demonstrate the model's robustness and its potential as a reliable tool for rapid, scalable, and objective soil texture classification in precision agriculture.

Keywords: Soil texture, Convolutional Neural Network, Classification

1. Introduction

Soil texture plays a critical role in determining water retention, nutrient dynamics, and overall agricultural productivity. Accurate classification of soil types enables better crop planning and resource management. However, traditional methods for soil texture analysis rely heavily on manual inspection and laboratory testing, which are both time-consuming and labor-intensive.

By automating soil texture classification with a Convolutional Neural Network (CNN), this project aims to reduce manual work, save time, and provide farmers with a fast, consistent, and scalable solution. The resulting model, SoilNet, has the potential to support precision agriculture efforts by making soil analysis more accessible and reliable, ultimately contributing to improved crop yield and resource efficiency.

1.1 Related Works

Traditional methods for determining soil texture, such as hydrometer and sieve analyses, remain accurate but are time-consuming and impractical for large-scale applications. Early computational studies applied machine learning models like Support Vector Machines and Partial Least Squares Regression on spectral reflectance or handcrafted features to automate texture estimation (Viscarra Rossel et al., 2006; Gholizadeh et al., 2013).

Although these approaches provided promising results, their reliance on manual feature extraction limited scalability and adaptability across diverse soil types. Recent research has shifted toward deep learning-based approaches, particularly Convolutional Neural Networks (CNNs), which can automatically learn spatial and spectral patterns from raw soil images. CNNs have achieved superior accuracy in texture classification compared to traditional models (Zhang et al., 2018; Dhakal et al., 2020). Subsequent studies have employed transfer learning (Gonzalez et al., 2021) and data fusion techniques combining multispectral and ground-level imagery (Li et al., 2022) to improve generalization. However, issues such as dataset standardization and model interpretability persist.

Building upon these developments, this study introduces a custom CNN architecture specifically optimized for soil texture classification using ground-level RGB images. Unlike prior works that rely on pretrained models or mixed data modalities, the proposed network is trained from scratch on curated soil imagery to better capture fine-grained textural variations. The model integrates deeper convolutional blocks with batch normalization and dropout layers to improve generalization and reduce overfitting. By focusing on interpretability and robustness across heterogeneous soil conditions, this approach aims to establish a reproducible and efficient framework for soil texture classification under practical agricultural settings.

1.2 Neural Networks

Neural network is a Machine Learning model that adheres to the complex function of human brains. These models consist of layers of interconnected nodes (perceptrons) that enable tasks such as pattern recognition and decision-making. The Neural Network model varies in the architecture - how the model processes and computes data, and each architecture specializes in unique tasks. Different architectures include Feedforward Neural Network, Convolutional Neural Network, Recurrent Neural Network, and Generative Adversarial Network. This paper will focus on the architecture of Convolutional Neural Network.

Convolutional Neural Network (CNN) is a type of neural network architecture designed specifically for analyzing and recognizing spatial or visual patterns in data, such as images. A CNN is composed of three main types of layers that work together to extract and interpret image features. The convolutional layers apply a series of learnable filters to the input images, allowing the network to automatically detect essential features like edges, gradients, and texture patterns. The pooling layers then reduce the spatial dimensions of these extracted features by summarizing regions, which helps retain the most relevant information while improving computational efficiency and robustness. Finally, the fully connected (dense) layers integrate the extracted spatial features to form high-level representations and produce final classifications.

Since soil texture classification fundamentally depends on visual cues – such as grain structure, color variation, and surface patterns – CNNs are particularly well-suited for this task. By progressively learning from low-level textures to high-level patterns, CNNs can accurately distinguish between different soil types, enabling precise and automated analysis of soil properties. Figure 1 illustrates the architecture of a simple CNN model.

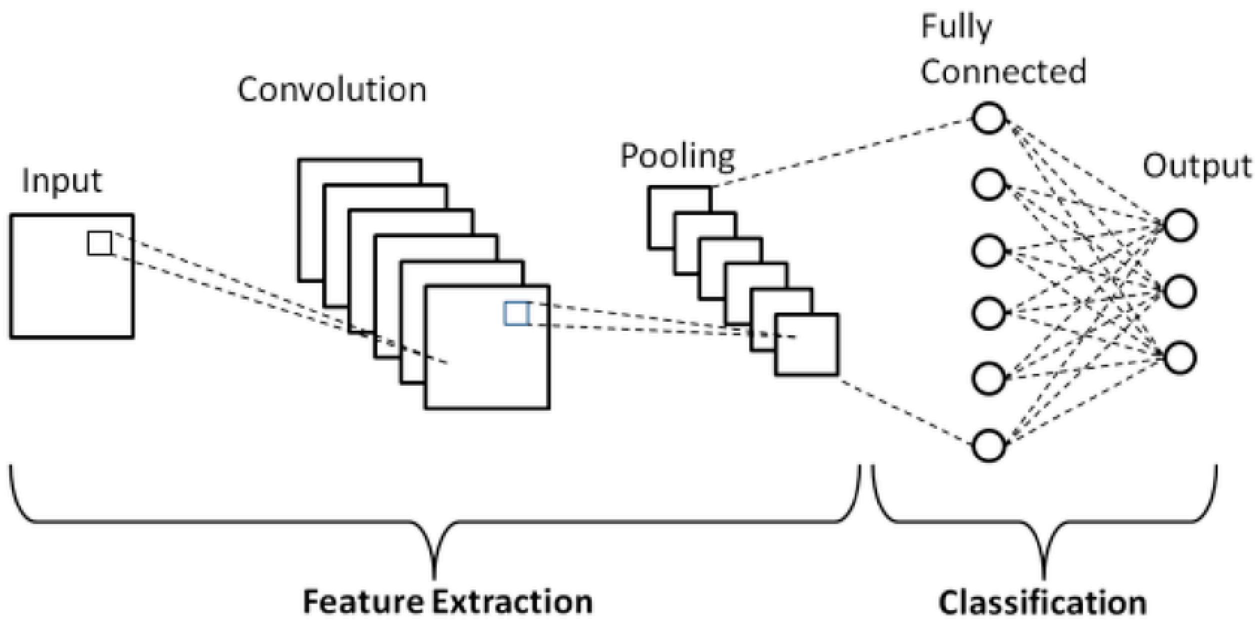

Figure 1. General architecture of Convolutional Neural Network

Throughout the span of the paper, a frequent number of technical terms will be used. Table 1 briefly explains those terms.

Table 1. Related terminology

Terms	Definition
Filter and Kernel	Small matrices applied to input data to extract features by convolution. They slide over the input to produce feature maps
Stride	The step size of the filter as it moves across the input
Padding	Adding extra pixels (usually zeros) around the input to control output size and preserve edge information
Batch Normalization	Normalizes the input of each layer to stabilize and speed up training
Pooling	Reduces spatial dimensions by summarizing regions. This makes the model more efficient and robust
Activation function	Introduces non-linearity into the network. Common examples of activation function are sigmoid, ReLU, Tanh
Epoch	One complete pass of the training dataset through the model
Learning rate	A hyperparameter that controls the step size when updating weights. It balances speed and stability of learning

2. Methodology

SoilNet's code can be referenced and examined through this GitHub repository: **SoilNet**

2.1 Environment Setup

The experimental environment was configured on Google Colab, leveraging an NVIDIA T4 graphical processing unit (GPU) to facilitate efficient model training and inference. All algorithms were implemented in the Python programming language, with PyTorch serving as the primary deep learning framework. Additional libraries were employed to support various tasks, including Scikit-Learn for metrics computation, Matplotlib for data visualization, OpenCV for image preprocessing, NumPy for high-performance numerical computation, and Pandas for structured data manipulation and analysis.

2.2 Dataset Preparation

Generally, the soil system could be categorized to up to 13 types, but this paper reduced the amount of labels down to 3 - coarse, fine, and medium – by grouping several soil types together; these are the 3 classes we are trying to predict.

The primary dataset utilized in this study was the SOIL (v1) dataset obtained from Roboflow, consisting of 3,573 images categorized into three soil texture classes. The dataset initially contained 4 soil types. I relabeled the soil types as follows: loamy sand was categorized as coarse, lal soil as medium, and clay soil as fine. The remaining category, sandy soil, is excluded from the analysis because it can not be easily categorized into one of the three classes. To increase the diversity of soil images, an additional dataset was incorporated from Kaggle (Pondy, 2022), which contained three soil types. For consistency with the classification scheme, these were categorized as follows: alluvial soil as coarse, black soil as fine, and red soil as medium.

Prior to training, preprocessing steps were applied, including loading images with OpenCV, resizing all inputs to a uniform resolution of 256 × 256 pixels, and the labels were converted to the onehot–encoded format: the ground truth of the label "Coarse" is stored as [1.0 , 0 , 0], "Medium" is stored as [0 , 1.0, 0], and "Fine" is stored as [0 , 0 , 1.0]. The combined dataset was then partitioned into 80% for training and 20% for testing, ensuring a balanced evaluation of model performance.

2.3 Model Architecture

The convolutional neural network (CNN) designed for soil texture classification consists of a hierarchical feature extraction module followed by a classification head. The feature extraction module is composed of four convolutional blocks. The number of filters increases from 16 to 32, 64, and finally 128, allowing the model to capture increasingly complex features. After the final layer, the extracted information is flattened into a 128-dimensional feature vector and passed through a fully connected layer with three output nodes representing the soil texture classes: coarse, medium, and fine. A softmax function then produces the final class probabilities.

2.4 Hyperparameters

To establish a baseline configuration, the SoilNet model was first trained using a standard set of hyperparameters. Specifically, the training was conducted for 20 epochs with the Adam optimizer and a learning rate of 0.001. A kernel size of 3×3 was employed across the convolutional layers to effectively capture local spatial features, while the ReLU activation function was applied to introduce non-linearity and mitigate vanishing gradient issues. This initial setup served as the reference point against which subsequent hyperparameter tuning experiments were evaluated. The initial model achieved an overall accuracy of 93.66% with a test loss of 0.1684, indicating strong predictive capability. It obtained a precision of 0.9439, a recall of 0.9366, and an F1-score of 0.9358, reflecting a well-balanced performance between correctly identifying positive samples and minimizing false positives.

Table 2. Hyperparameter tuning results for the SoilNet

Hyperparameter	Values Tested	Best Value	Accuracy	Precisinn	Recall	F1 - Score
Baseline			93,66%	0.9439	0.9366	0.9358
Number of epochs	30, 50, 70	50	99,73%	0.9973	0.9973	0.9973
Learning rate	0.01, 0.001, 0.0001	0.001	93,66%	0.9439	0.9366	0.9358
Kernel size	3 x 3, 5 x 5	5 x 5	95,14%	0.9577	0.9514	0.9513
Activation function	ReLU, LeakyReLU	LeakyReLU	95,82%	0.9609	0.9582	0.9584

Finally, a model with a set of hyperparameters that output the best metrics has been tested. The tuned-model has been trained through 50 epochs with a learning rate of 0.001, kernel size 5 x 5, and activation function LeakyReLU. After combining the best hyperparameters, the model achieved an outstanding performance with a test loss of 0.0182 and an impressive accuracy of 99.46%. The model also reached a precision of 0.9946, recall of 0.9946, and F1-score of 0.9946, reflecting a near-perfect balance between precision and recall. This is the final hyperparameter set of SoilNet.

2.5 Training

The final model was trained on the final dataset using the optimized hyperparameters identified in Section 2.4. Specifically, training was performed for 50 epochs with the Adam optimizer, a learning rate of 0.001, a 5×5 kernel size, and the LeakyReLU activation function. The training process involved monitoring the loss curve at each epoch to ensure convergence and prevent overfitting. Early stopping was not required as the validation loss consistently decreased, demonstrating stable and effective learning throughout the training process.

3. Results

3.1 Training loss

In each epoch, the model monitored the average loss of batches and then visualized.

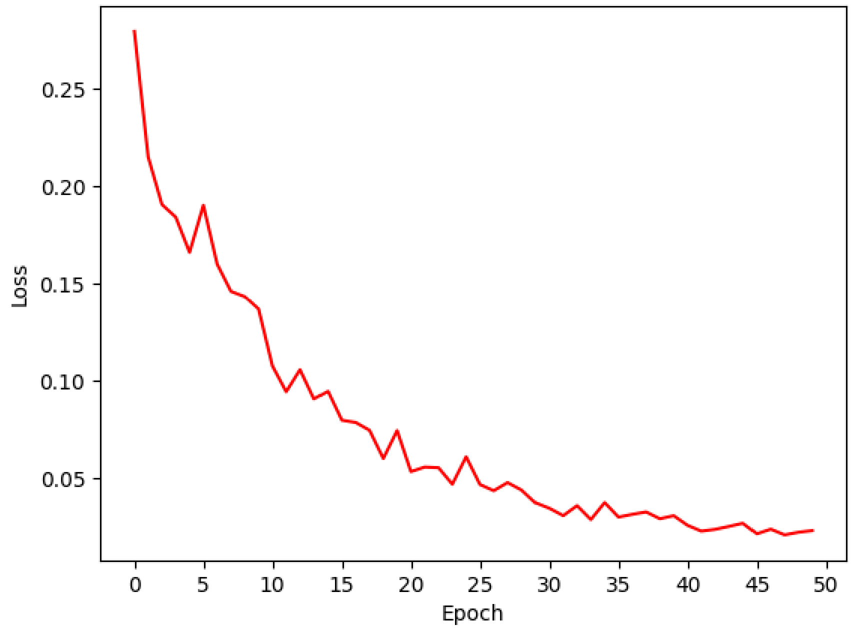

Figure 2. Training loss curve of the CNN model

Figure 2 illustrates the training loss over 50 epochs. The curve shows a steady and smooth decline, indicating that the model is learning effectively and converging without signs of overfitting. The loss stabilizes toward the later epochs, confirming that the chosen number of epochs (50) allows the model to fully learn the data distribution.

3.2 Confusion matrix

Model evaluation was carried out on the held-out test set using the final tuned configuration. Predictions were generated by selecting the class with the highest softmax probability, and performance was assessed using accuracy, precision, recall, and F1-score.

After the fine-tuned model is trained, the model is tested on the test dataset, which accounts for 20% of the dataset. The final model achieved 99.46% accuracy and a test loss of 0.0182, with near-perfect precision (0.9946), recall (0.9946), and F1-score (0.9946). These results indicate that the tuned model generalizes exceptionally well and can reliably classify all three soil types with minimal error.

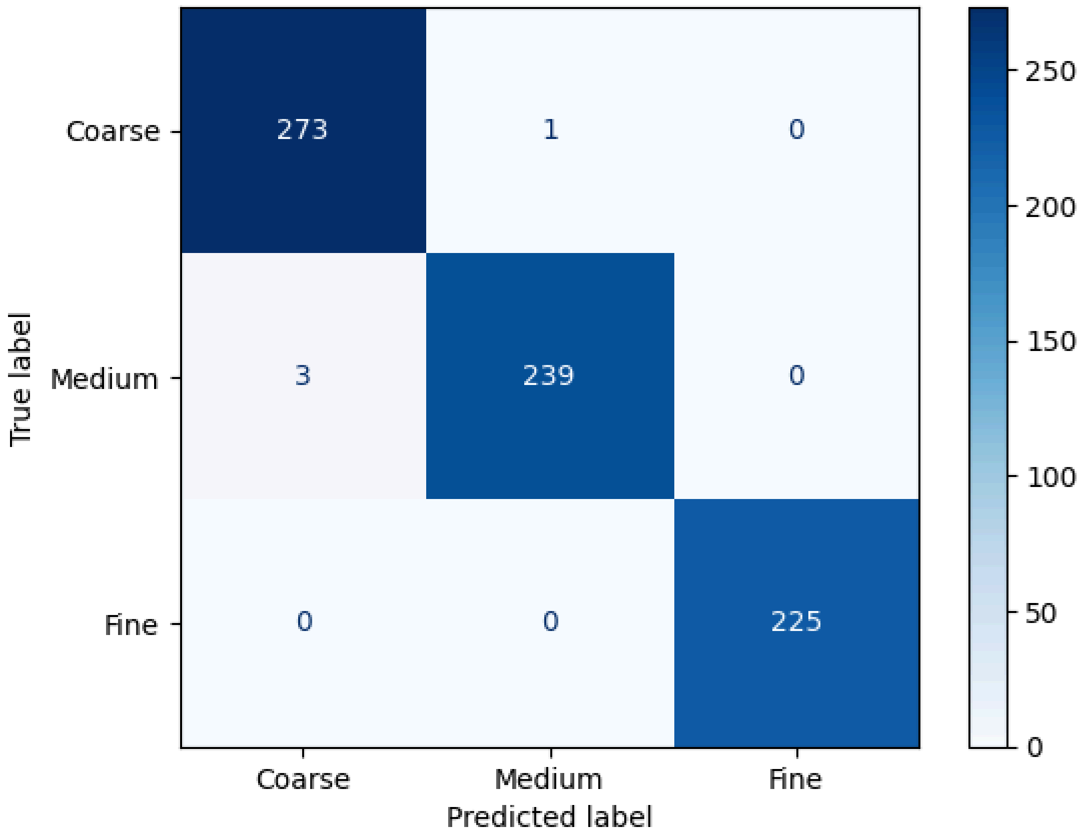

Figure 3. Confusion Matrix for the SoilNet prediction

Figure 3 presents the confusion matrix of the final tuned model evaluated on the test set. The matrix demonstrates near-perfect classification performance across all three soil types. Only four misclassifications occurred in total: one Coarse sample was predicted as Medium, and three Medium samples were predicted as Coarse. Fine soil samples were classified with perfect accuracy. This further supports the reported high precision, recall, and F1-scores, confirming that the model generalizes exceptionally well.

4. Discussion

The experimental results demonstrate that the proposed SoilNet model is highly effective for soil texture classification. The steady decrease in training loss and the absence of overfitting indicate that the chosen hyperparameters – particularly the use of a 5×5 kernel size and the LeakyReLU activation function – were appropriate for capturing discriminative features while maintaining model generalization.

When comparing the baseline configuration to the tuned model, there is a substantial improvement in all performance metrics, with accuracy increasing from 93.66% to 99.46%. This improvement highlights the importance of systematic hyperparameter tuning, as both kernel size and activation function play a critical role in enhancing feature extraction capabilities. The nearly perfect precision, recall, and F1-scores, as well as the minimal misclassifications observed in the confusion matrix, indicate that the model is highly reliable across all soil classes.

Moreover, the results validate the suitability of CNN-based approaches for automating soil texture classification. Traditional methods, which are often manual and time-consuming, can now be complemented or replaced with deep learning models to enable faster and more consistent analysis. However, while the performance is excellent on the current dataset, future work should evaluate the model's robustness on larger and more diverse datasets, including images captured under different lighting conditions, soil moisture levels, and geographic regions. Additionally, techniques such as data augmentation and transfer learning could be explored to further improve generalization and reduce dependence on large, labeled datasets.

References

Albawi, S., Mohammed, T. A., & Al-Zawi, S. (2017, August). *Understanding of a convolutional neural network*. In *2017 International Conference on Engineering and Technology (ICET)* (pp. 1–6). IEEE. https://doi.org/10.1109/ICEngTechnol.2017.8308186

Dhakal, S., Shrestha, S., & Sharma, S. (2020). *Soil texture classification using convolutional neural networks*. Journal of Applied Remote Sensing, 14(3), 1–12.

Gholizadeh, A., Borůvka, L., Saberioon, M., & Kozák, J. (2013). *Visible, near-infrared, and mid-infrared spectroscopy applications for soil assessment*. Applied Spectroscopy, 67(12), 1349–1362.

Gonzalez, A., Lopez, D., & Carrasco, J. (2021). *Transfer learning for soil classification using convolutional neural networks*. Computers and Electronics in Agriculture, 190, 106423.

Li, W., Wang, J., & Zhang, Y. (2022). *Fusion of multispectral and ground-level soil imagery using deep convolutional networks for texture classification*. Remote Sensing, 14(5), 1103.

Md Sohag. (2025). *SOIL (v1) dataset* [Data set]. Roboflow Universe. https://universe.roboflow.com/md-sohag-w9lng/soil-aoweg

Pondy, J. (2022). *Soil Image Dataset* [Data set]. Kaggle. https://www.kaggle.com/datasets/jayaprakashpondy/soil-image-dataset

Viscarra Rossel, R. A., Walvoort, D. J. J., McBratney, A. B., Janik, L. J., & Skjemstad, J. O. (2006). *Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties*. Geoderma, 131(1–2), 59–75.

Zhang, Y., Qin, Q., & Liu, J. (2018). *Soil texture classification from hyperspectral data using convolutional neural networks*. ISPRS Journal of Photogrammetry and Remote Sensing, 146, 183–192.