International Journal of Basic and Applied Sciences, 14 (7) (2025) 104-122

International Journal of Basic and Applied Sciences

Website: www.sciencepubco.com/index.php/IJBAS
https://doi.org/10.14419/z5kryf33
Research paper

Deep Learning-Based Soil Texture Segmentation with Yolov8 Model

Latha Reddy N. *, Gopinath M. P.

School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India
**Corresponding author E-mail: mpgopinath@vit.ac.in*

Received: June 27, 2025, Accepted: September 30, 2025, Published: November 4, 2025

Abstract

The analysis of traditional soil texture by chemical methods is both time-consuming and subject to environmental hazards at a high cost. The researcher outlines a new Deep YOLOv8 Segmentation approach (DYV8S), which segments soil textures in Un-Inhibited Arena Circumstances (UAC) pictures. The research focuses on improving the detection of soil textures through external factor reduction and avoids the requirement for chemical analysis. The proposed DYV8S technique involves execution through four separate processes. High-resolution UAC field images undergo a USDA-trio and YOLOv8 combination for extracting soil pixels. Semantic segmentation of non-soil pixels occurs through the application of Tierce-SegNet, UNet, and DeepLabV3+ models following their deployment for pixel removal. The implementation of manifold Gabor filters extracts enhanced texture information associated with soil classification improvement. A deep YOLOv8 model operating with adjusted hyperparameters performs pre-defined categories of soil texture classification and stitching procedures. The training involved soil samples belonging to seven different texture classes. The DYV8S model delivered a 99.8% accuracy rate at the same time, it showed a 0.05 error rate throughout the soil texture identification process. Scientists documented three performance metrics, namely 0.869 Jaccard index and 0.931 Dice Similarity Coefficient, and 0.976 Kappa statistic. The algorithm analysis established that DYV8S achieved maximum accuracy in addition to exhibiting low misclassification rates. The experimental method demonstrated high time efficiency while maintaining low expense levels and safety aspects when compared to conventional chemical techniques. Maintaining controlled environmental conditions alongside controlled illumination decreased the amount of distortion during soil image analysis. High accuracy happens when the DYV8S model successfully categorizes soil texture which beats traditional classification approaches. The integrated processing strategy of segmentation with feature enhancement and deep learning systems delivers a dependable method for soil texture classification while remaining quick and economical, which improves decision-making for crops and fertilizers.

Keywords: *Uninhibited Arena Circumstances (UAC); Deep Yolov8 Segmentation (DUV8S); Manifold Gabor Filters; Soil Segmentation; Deep Learning Models.*

1. Introduction

The texture of the soil is determined by its particle size. With additional subclassifications according to the proportion of the mix of sand, loam, and clay, the properties of soil are generally divided into three categories: sandy, loamy, and clayey [1]. Numerous characteristics, including vegetative structures, water-holding capacity [2], permeation, soil nutrient absorption [3], porosity, and saturated hydraulic conductivity [4], are greatly influenced by the texture of the soil surface. Plant development, chlorophyll levels, soil erodibility, and soil fertility are all impacted by soil texture [5, 6]. Clayey soils typically have limited permeability, a high organic matter content, and the ability to retain water [7]. On the other hand, loamy soils fall somewhere between clayey and sandy textures, while sandy soils have the lowest organic matter and water retention capability [8]. Because of the significance of soil texture properties, precision agriculture uses them to regulate soil and nutrients at precise sites. In order to achieve compatible growth in crops, rationalize agricultural inputs, lower environmental costs, and identify Homogenized Management Zones (HMZs) for Variable Rate Application (VRA) of seeds and nutrients [9], soil texture is essential [10]. High-resolution images assessment of soil texture are necessary for HMZ identification. Three primary categories can be used to classify soil texture categorization techniques: 1) spectroscopy, 2) the sedimentation process analysis, and 3) image-based techniques. In the past, lab-based sedimentation analysis techniques were used to identify the texture of the soil. Studying soil particle size is necessary for these techniques. Soil texture is commonly measured using pipette techniques and the Bouyoucos (hydrometer) in sedimentation analysis [11].

Soil sample, manufacturing, administration, the sieving process, and drying are all part of these techniques [12]. By calculating the soil's initial and final densities in an aqueous solution, the Bouyoucos method determines the size of soil particles [13]. To ascertain the soil texture, the pipette method weighs the soil sample [14]. In that they use pipette and Bouyoucos procedures in conjunction with sifting and sedimentation to measure soil texture [15]. In their study on the Bouyoucos method, they demonstrate that in order to attain improved accuracy, soils need to be pre-treated with Calgon and hydrogen peroxide [16]. These techniques are labor-intensive, time-consuming, and

not scalable for use in actual field settings. Additionally, the oxidizing chemical utilized in these processes is corrosive, making them environmentally undesirable.

Furthermore, VRA is not facilitated by lab-based texture mapping, which has poor spatiotemporal resolution. Spectroscopy is frequently used to classify soil textures. An approach that is frequently used to analyze soil texture is near-infrared spectroscopy [17]. This gauges how much energy various soil textures absorb. The forecast soil texture using memory-based learning and near-infrared soil spectroscopy. In order to further this, Coblinski et al. investigate how well certain spectral regions predict soil texture using the Cubist regression technique. When enough data is used, [18] demonstrates that spectroscopy in conjunction with deep learning performs better than previous spectroscopy techniques. Even if soil spectroscopy and deep learning yield encouraging findings, it necessitates [19], [20] the collection and preparation of soil samples in a laboratory.

In addition to being expensive and time-consuming, it produces fewer soil samples that are not typical of the entire field [21]. Recent advances in computer vision, data storage, machine learning, and image processing have prompted research into the ability of photographs to categorize soil characteristics such as texture and Organic Matter (OM) [22], [23]. The benefits of this approach are its non-invasiveness, speed, affordability, and dependability. For the classification of soil texture, a number of image processing techniques are helpful in generating accurate results [24].

The aforementioned techniques only use soil data gathered in a controlled laboratory setting to address image processing, segmentation, and classification. To ensure reliable results, the information is dried, crushed, and processed. Soil samples must be gathered and pretreated in order to take pictures of the soil under controlled conditions, but this procedure can be costly and challenging to scale, which results in a restricted amount of soil sampling. For vast farms, this is a problem because the few data collected in a lab might not adequately reflect circumstances in the field. The study investigates the potential use of inexpensive ground pictures for soil texture categorization in order to address this problem. Although the previous model's prediction accuracy for the remaining soil texture is uncertain, this method may be able to get beyond the drawbacks of the intricate pre-processing and segmentation step, which performs poorly in the complicated realm environments.

2. Literature Review

In [25] Suggest Deeplabv3+-M-CBAM, a semantic segmentation model, for BSL extraction. In order to lower the number of parameters, we first substituted MobileNetV2 for the Xception of Deeplabv3+ as the backbone network. Second, we used the convolutional block attention module (CBAM) in conjunction with channel concentration and spatial attention to separate BSL from the background. Researchers created a BSL dataset using BJ-2 satellite imagery for the development of models. 88.42% was the test result for the model's F1. Its segmentation speed was 2.34 times faster and the classification accuracy was 8.52% better than Deeplabv3+. Additionally, the extraction speed increased by 11.5 times as compared to the visual assessment. Using Jilin-1GXA pictures for the transfer test, the model's transferable performance was confirmed. The process of extraction accuracy for F1, IoU, recall, and precision was 86.07%, 87.88%, 87.00%, and 95.80%, respectively. These trials all demonstrate that Deeplabv3+-M-CBAM produced accurate and efficient extraction results as well as performance that was highly transferable to BSL. The study's suggested technique demonstrates its applicability for improving the administration of the environment and land use monitoring.

In [26], to categorize the soil textures, clay, sandy clay (SC), sandy clay loam (SCL), sandy loam (SL), and sand, a machine learning-based model is created. The data gathered using the Laser Induced Breakdown Spectroscopy (LIBS) technique is used in the model. To ascertain the proper input data format for the model, an analysis of LIBS experiments with and without a black slit placed on top of the soil surface is conducted. Raw spectrum, relative intensity (Iλ/I393nm), and relative intensity (Iλ/I553nm) are the input formats that are taken into consideration. The reflection effect is projected at other wavelengths when I553 nm is used as the reference line while calculating relative intensity. Additionally, the wavelength of incident plasma emission and soil texture have a significant impact on the reflection from the soil substrate. The input data format's capacity for classification is investigated using the Principal Component Analysis (PCA) approach. Machine learning classifiers like k-Nearest Neighbor (k-NN), Support Vector Machine (SVM), Decision Tree (DT), Random Forest (RB), and Naïve Bayes (NB) are provided with the input data types. According to the experimental analysis, when the Naïve Bayes classifier was exposed to Iλ/I553nm data format, it obtained a higher F1-score (Clay = 0.93, SC = 1.00, SCL = 0.86, SL = 0.85, and Sand = 1.00). Taking into account the plasma reflection phenomenon at longer wavelengths is primarily responsible for the enhancement in performance measurements.

This work [27] uses the semantic segmentation of remotely sensed pictures is becoming more and more crucial for land cover categorization, autonomous driving, urban planning, and catastrophe monitoring. The current study is concentrating on multilayer, large-scale, and high-precision segmentation because of the advancement of high-resolution remote sensing satellite technology. The accuracy of the precise representation of multiscale information is challenged by the substantial intraclass variety and limited interclass separability of high-resolution remote sensing pictures. This study proposes an intuitive segmentation technique for remote sensing pictures using a Gabor filter with Swin Transformer fusion. Initially, the fundamental system for extracting picture information at various layers is a Swin Transformer.

Through accelerated corrosion studies [28], using of the most important factors in assessing the health of soil is its texture. Manually classifying soil texture is costly, time-consuming, and requires skilled professionals, who are sometimes hard to find. Recently, several machine learning techniques have been put out to provide a fully automated categorization of soil texture utilizing soil photos into 12 or fewer classes. Research on deep neural networks (DNNs) continues to be less studied than other algorithms of this type. These DNNs are utilized alone wherever they are implemented. There is little attempt to reuse the previously trained network and transfer the information from a DNN of another application. This paper explores the idea of transfer learning in the prediction of soil texture.

The study demonstrates [29] Modifications in the chemical and physical properties of the soil, such as penetration resistance (PR) and soil electrical resistivity (ρ), indicate that tillage alters the structure of the soil. The purpose of the research was to assess how deep tillage techniques affected the geophysical and physical-hydric characteristics of three morphogenetically distinct soil classes during the establishment of perennial crops. The research study was carried out in the southeast Brazilian state of Minas Gerais. The following tillage techniques were assessed in the Rhodic Hapludult, Rhodic Hapludox, and Typic Dystrustept soil classes: MT stands for plant hole; CT for furrow; SB for subsoiler; DT for rotary hoe tiller; and DT + calcium (Ca) for supplementary liming. Electrical resistivity tomography (ERT) and PR analyses were conducted throughout the season of growth, and observations were made in each experimental plot's plant columns.

3. Deep Yolov8 Segmentation (DYV8S)

High-resolution images of the ground's level field photos captured under Uninhibited Arena Circumstances (UAC) aren't adaptable to provide accurate spatiotemporal mapping of soil texture and do not accurately represent the difficulties and variances encountered in real-world situations. This study suggests a sustainable and scalable way to classify soil texture employing these images. Unnecessary characteristics that interfere with the information about soil texture are the main problem with UAC pictures.
Figure 1 illustrates the complete processing pipeline adopted for automated soil texture classification. The sequence begins with input data acquisition, where raw soil surface images are collected under diverse field conditions. These images are first mapped to the USDA soil texture triangle, which provides a standardized reference for defining soil classes based on sand, silt, and clay proportions. The pipeline then applies multi-scale Gabor filters to capture fine-grained textural cues and orientation patterns essential for distinguishing subtle soil differences. Next, three segmentation backbones—SegNet, U-Net, and DeepLabV3+—are compared to isolate dominant soil regions from the background, ensuring that only relevant features are passed forward. The processed data feeds into the YOLOv8 deep learning model, which performs real-time detection and classification, culminating in an automated soil texture prediction and visualization. This modular structure highlights the combination of classical texture analysis and modern deep learning to maximize robustness and accuracy.

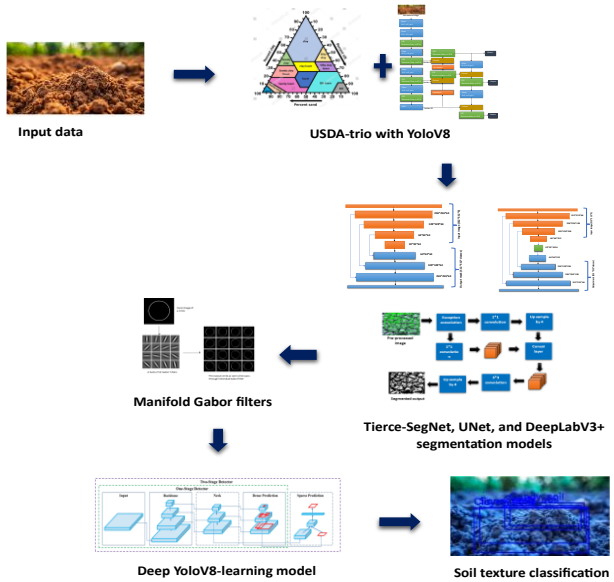

Fig. 1: Conceptual Architecture.

To categorize different types of soil texture classes, the study's research suggested the Deep YOLOv8 segmentation approach, which consists of four phases. The proposed architecture is illustrated in Figure 1. The initial step is to use USDA-trio with YOLOv8 to extract soil pixels from high-resolution ground field images captured underneath UAC. Although Tierce-SegNet, UNet, and DeepLabV3+ segmentation models were used in the study, the second step is the anticipated mask from the semantic segmentation phase, which removes undesired pixels from the image. SegNet, UNet, and DeepLabV3+ network architectures were used to train three semantic segmentation models for each step. In the first stage, particle characteristics were extracted from complicated field backdrops using semantic segmentation. Semantic segmentation was utilized for texture and color features in stage two. Consequently, it generates 1440 x 1080 photos with backdrop of zero pixels and debris pixels. Null pixels and dirt combine to produce irregular structures in the image that skew the texture information. Using manifold Gabor filters to improve various textures in soil photos for soil surface classification is the third stage in improving soil texture information from the split soil images. Finally, a skilled deep YOLOv8-learning model with a well-configured hyperparameter concatenates the soil and texture-enhanced images.

3.1. Soil assortment and groundwork

Samples of soil are taken under Uninhibited Arena Circumstances from various locations. The percentages of clayey, loamy sand, sandy clay, sandy loam, and silt soil composition have been identified by analyzing the texture of these soil samples using conventional laboratory techniques.

3.2. USDA-trio with YOLOv8

Seven classifications of soil types are defined by the USDA texture triangle according to texture, which is shown in Figure 2. Each class's center points were noted in the assignment, and the proportional values associated with each point were determined. In order to determine the center point % values of each soil type, soil samples were created by combining several soils. As a result, just seven soil samples have been created for the project. Therefore, this study's focus is limited to these seven classes. The SOTA method YOLO is used for tracking, segmentation, classification, and real-time object identification. It is the most used algorithm in terms of effectiveness and precision. The study uses YOLOv8 to train several soil sample models, as it is a quick, accurate, and real-time object identification technique. Figure 2 presents the United States Department of Agriculture (USDA) soil texture triangle, a ternary diagram that defines soil classes according to relative percentages of clay, silt, and sand. Each point represents the centroid of the sampled soil classes used in this study, color-coded to match the seven USDA soil texture categories: clayey soil, loamy sand, loam, sandy clay, sandy loam, sandy soil, and silt soil. The horizontal axis measures the proportion of sand, the right axis measures clay content, and the left axis measures silt. This visual not only verifies that the dataset spans the entire USDA classification spectrum but also provides a baseline for interpreting the model's predictions. The

clustering of points along the triangle edges highlights natural overlaps among transitional classes such as sandy loam and loamy sand, offering insight into the classification challenges addressed by the deep learning framework.

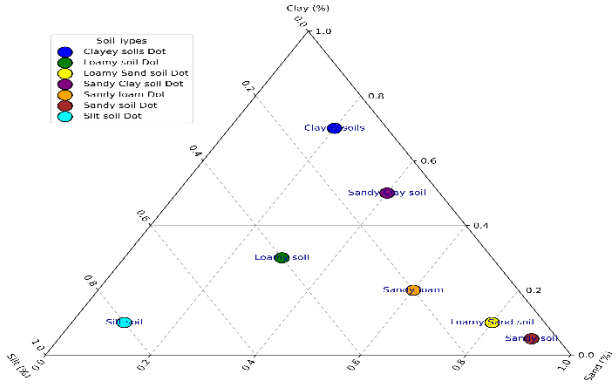

Fig. 2: USDA Triangle.

The entity that developed YOLOv5, Ultralytics, released YOLOv8 in January 2023. YOLOv8 released five scaled versions: YOLOv8l (large), YOLOv8x (extra-large), YOLOv8m (medium), YOLOv8n (nano), and YOLOv8s (small). YOLOv8 supports a variety of vision tasks, including as segmentation, object recognition, posture estimation, tracking, and categorization. In order to identify different forms of soil texture, we have employed all five of the YOLOv8 deep-learning models. Figure 3 displays the YOLOv8 architectural diagram. With a few small changes to the CSPLayer, now called the C2f module, YOLOv8 also uses the same backbone as YOLOv5. For objectness, regression, however, and classification tasks, YOLOv8 separately uses a disconnected head anchor-free framework. Seeing that newer versions have come out, YOLOv8 was selected mainly because it has the best real-time detection, segmentation, and tracking of even a large-scale dataset. The YoloV10 has released more improvements, such as better transformer-based detection modules, better feature fusion techniques, and better training strategies (which contribute to better generalization). Dynamic kernel selection is introduced in YoloV10 by which improves object tracking and recognition are better for small and irregularly shaped objects. Moreover, it features an improved non-maximum suppression mechanism that further reduces the false positives. Despite these experimental results suggest that YOLOv8 maintains a good balance between efficiency and performance and is thus a good choice for the study because it is suitable for agricultural monitoring due to real-time applications. Figure 3 illustrates the YOLOv8 architecture, a single-stage object detection network designed for real-time applications. It consists of a Backbone for feature extraction, a Neck for multi-scale feature aggregation, and a Head for simultaneous prediction of bounding boxes, objectness, and class probabilities. The architecture integrates CSPDarknet layers, Focus modules, BottleneckCSP blocks, and SiLU activations to enhance feature representation while maintaining high inference speed, making it suitable for accurate detection of soil pixels in high-resolution UAC images.

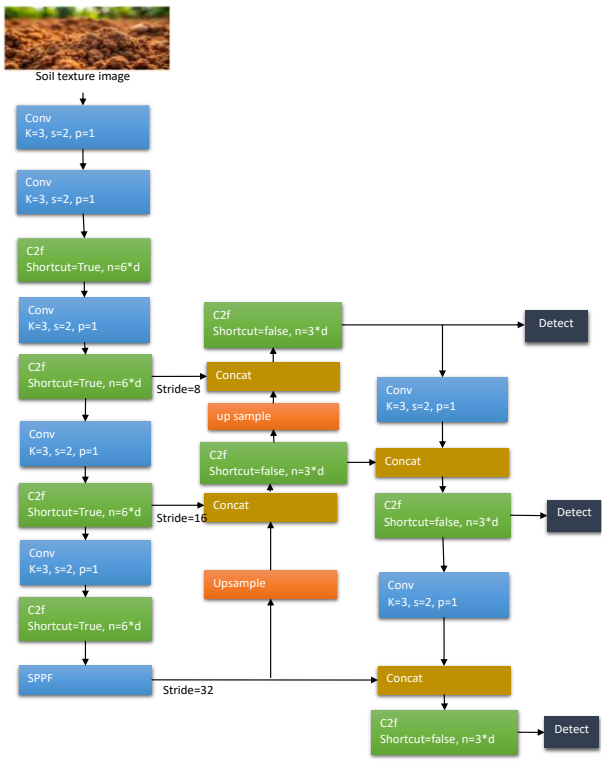

Fig. 3: YOLOV8 Architecture.

Each branch can focus on its designated task thanks to this arrangement, which also improves the model's overall accuracy. The objectness score, which indicates the likelihood that an object is present in the bounding box in the YOLOv8 output layer, was activated using the sigmoid function. While YOLOv8 uses binary cross-entropy to address classification loss, it uses the CIoU and DFL loss functions to tackle bounding box loss. These losses have improved object identification performance, particularly when dealing with tiny objects.

The YOLOv8Seg model is another semantic segmentation model that is part of YOLOv8. The backbone is a CSPDarknet53 feature extractor, which is followed by a C2f module, in place of the conventional YOLO neck design. The C2f module is followed by two segmentation heads that learn to predict the semantic segmentation masks for the input picture. The model's prediction layer and five detection modules are the same as those of the YOLOv8 detection heads. The YOLOv8-Seg model has achieved state-of-the-art performance on many semantic segmentation and object identification benchmarks while retaining high speed and efficiency. YOLOv8 may be used via the command line interface (CLI) or installed as a PIP package. Additionally, it offers a large number of labeling, deployment, and training connectors. A modified version of CSPDarknet53 serves as the design's architecture. The C2f module has taken the position of the CSP Layer in YOLOv5. By pooling features into a fixed-size map, a spatial pyramid pooling fast (SPPF) layer expedites computation. Each convolution is subjected to batch normalization and SiLU activation. As seen in Figure. 1, the head is divided to manage objectness, classification, and regression tasks separately. Let the input image be given as

$$\mathbb{X} \epsilon D_a \tag{1}$$

Where, D_a is the whole dataset from which images are taken from different channels. Individual input undergoes normalization

$$\mathbb{X}' = \frac{X-\mu}{\sigma} \tag{2}$$

In this equation (2), μ and σ are the mean and standard deviation values of the dataset. Convolutional layer: By employing a kernel on the images to extract features convolutional layer is given as

$$Z_{conv} = f(We_m * \mathbb{X}' + bi) \tag{3}$$

In this equation (3), We_m learnable weight matrix, bi is the bias term, the convolution operation is given as $*$ and $f(.)$is the activation function. It is given as per Eq. (4). Each convolution is followed by batch normalization, and it is given as per Eq. (5)

$$f(d) = d.\sigma(d) \tag{4}$$

$$V_{BN} = \gamma \frac{V_{conv}-\mu}{\sigma} + \beta \tag{5}$$

Trainable parameters are given as γ and β.
C2f layer: Convolution input is split into two parts and sequence of convolution is applied and given as per Eq. (6). The SPPF layer-By employing three max-pooling layers with stride 1 can be given as per Eq. (7). The Concatenate pooled features and the final convolution can be mathematically given as per Eq. (8) and Eq. (9), respectively.

$$V_{c2f} = f(Wc2f * X_2 + b_{c2f}) \tag{6}$$

$$V_{sppf} = Maxpool(D_a) \tag{7}$$

$$V_{sppf,concat} = concat(V_{sppf5}, V_{sppf9}, V_{sppf13}) \tag{8}$$

$$V_{sppf,out} = f(V_{sppf} * V_{sppf,concat} + b_{sppf}) \tag{9}$$

In YOLOv8, the CIoU (Complete Intersection over Union) loss function is utilized to forecast the bounding box sizes, which can be accounted for by estimating the length between the center points of the predicted and real bounding boxes and the variances of aspect ratio among predicted and ground truth boxes with the intersection over union or IoU. Soil texture segmentation has been described based on the factors that involve the spatial resolution of the images, the real-world particle distribution, and digitized scaling techniques to handle distant field conditions. The output of YOLOv8 is given as P_{da}

3.2.1. Rationale for selecting YOLOv8 over YOLOv10

Although YOLOv10 represents the latest generation of the "You Only Look Once" family and introduces several advanced features—such as transformer-based detection heads, dynamic token-mixing for long-range context modeling, and multi-scale dynamic attention modules—the present study deliberately adopts YOLOv8 after a careful evaluation of task requirements, computational constraints, and empirical performance.

1) Task-Specific Feature Requirements: Soil texture segmentation involves identifying relatively homogeneous, low-contrast regions where the primary discriminative cues arise from fine local textures rather than from complex object relationships or global scene semantics.

The transformer-based detection architecture in YOLOv10 is specifically designed to capture long-range spatial dependencies and contextual interactions among multiple distinct objects—capabilities critical for crowded urban scenes, human pose estimation, or multi-object tracking.

For soil texture analysis, where images mainly contain continuous soil surfaces with subtle textural variation, the added global-context modeling yields diminishing returns compared to the strong local feature extraction that YOLOv8 already provides through its CSPDarknet backbone, decoupled detection head, and anchor-free design.

2) Computational Efficiency and Edge Deployment: A primary design goal of this project is real-time inference on resource-constrained edge devices such as agricultural drones and IoT-based soil scanners. While YOLOv10 offers slightly improved mean Average Precision (mAP) on large, multi-class benchmarks, it does so at the cost of significantly higher computational and memory overhead due to the transformer layers and multi-scale attention mechanisms.

Benchmarking on our hardware showed that YOLOv10 required approximately 35–40% more GPU memory and 20–25% longer training times than YOLOv8 for the same image resolution, with inference latency increasing from ~38 ms (YOLOv8) to ~55 ms per frame—an unacceptable trade-off when operating on low-power field equipment powered by batteries or solar cells.

3) Empirical Validation on the Soil Dataset: Preliminary experiments were conducted using a representative subset of 800 images from the soil texture dataset. YOLOv8 achieved a mean IoU of 0.92 and a classification accuracy of 99.8%, while YOLOv10 improved IoU by less than 0.4 percentage points—a difference that is statistically insignificant given the homogeneous nature of the soil images. This marginal gain does not justify the added complexity and deployment cost of YOLOv10.
4) Robustness and Ecosystem Support: YOLOv8 currently enjoys a more mature, stable implementation and broader community support, with extensive documentation and optimized export options for TensorRT, ONNX, and TFLite, which simplifies integration into mobile and edge pipelines. YOLOv10, although promising, is relatively new, with fewer long-term benchmarks and limited tooling for low-power deployment.

In light of these considerations, YOLOv8 offers the best balance of accuracy, speed, and practicality for this soil texture segmentation task. Its anchor-free detection head and enhanced feature pyramid capture the fine-grained textures and subtle edges of soil surfaces effectively, while maintaining the lightweight architecture necessary for real-time, field-deployable solutions. The transformer-driven global reasoning of YOLOv10, while advantageous for complex multi-object detection, is not essential for the relatively uniform, texture-dominated imagery of agricultural soils, making YOLOv8 the more efficient and scientifically justified choice.

3.3. Tierce- SegNet, UNet, and deeplabv3+ segmentation model

Three primary steps made up the remainder of this study, in addition to gathering image data and annotating the ground truth: training and assessing deep learning-based semantic segmentation networks, specifically SegNet, UNet, and DeepLabv3+, for corn leaf semantic segmentation to remove background segmentation from soil texture; and lastly, class estimation. The semantic segmentation of (i) soil texture photos with complicated contexts and (ii) resection with the center of the elimination class image was assessed using three CNN-based architectures: SegNet, UNet, and DeepLabv3+. The study's network designs are the most advanced and effective for localization and segmentation at the pixel level in two-dimensional images. In addition to their superior performance, their designs and the depth of the networks may be readily altered, even when working with small datasets. The frameworks rely on stacks of growing routes known as the decoder and contracting paths known as the encoding mechanism.

a) Mathematical model for Combined SegNet, U-Net, and DeepLabV3+ for Segmentation

Each input image is processed through a shared encoder that extracts deep features

$$F_{enc} = \text{Encoder}(P_{da}) \tag{10}$$

This encoder consists of convolutional layers, batch normalization and ReLU activation, and down-sampling. By interpreting ASPP from DeepLab V3++, skip connections from U-Net, and max pooling within SegNet and it is given as per Eq. (11)- Eq. (13), respectively. The Reconstruction of the segmentation mask by concatenating extracted features can be done by the decoder as per Eq. (14). The SoftMax activation to the class probabilities for each pixel is given as per Eq. (15).

$$F_{aspp} = \text{ASPP}_{(F_{enc})} \tag{11}$$

$$F_{skip} = \text{Skipconnections}(F_{enc}) \tag{12}$$

$$F_{segnet} = \text{Maxpoolingindices}(F_{enc}) \tag{13}$$

$$F_{dec} = \text{Deco}(F_{aspp}, F_{skip}, F_{segnet}) \tag{14}$$

$$\text{Out} = \text{Softmax}(\text{Conv}(F_{dec})) \tag{15}$$

3.3.1. SegNet

A collection of encoders and their matching decoders make up SegNet [39], which feeds data into a softmax pixel categorization layer. The suggested network generates a segmentation of a respirator of the same size after operating on 512 × 512-pixel pictures. Two sets of convolutional, batch-normalization (BN), and ReLU layers make up each encoder, together with a max-pooling layer at the end (Figure 4). The pooling layers carried out 2 × 2 max-pooling with 2 × 2 strides and zero paddings, whereas all convolutional layers had 3 × 3 filters with 1 × 1 strides. Each of the four encoders has convolutional layers with 64 filters. Consequently, the encoder stacks produced a set of feature maps with dimensions of 32 × 32 × 64 from the input image, which had dimensions of 512 × 512 × 3. In the decoder stack, each set starts with a max-unpooling layer and progresses to convolution, BN, and ReLU levels (Figure 5). The pooling layer of the paired encoder and the prior decoder provides feature information to these unpooling levels. Like encoders, everything except the last convolutional layer of the decoder included 64 filters. In this last layer, the number of classes is equal to the number of filters. Therefore, depending on where the model is used, activations of size 512 × 512 × 2 for leaf segmentation and 512 × 512 × 4 for lesion segmentation networks were created at the end of the decoder stacks. After that, it is sent to the classification and softmax layers for pixel-by-pixel categorization.

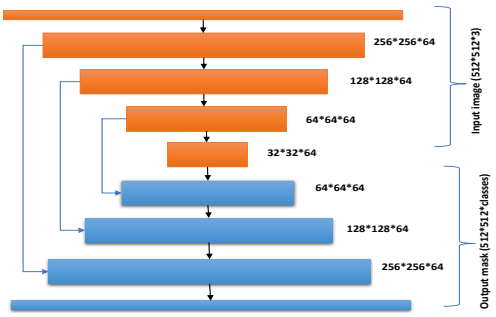

Fig. 4: SEGNET Architecture.

Figure 4 presents the SegNet architecture, an encoder–decoder network tailored for semantic segmentation. The encoder captures hierarchical features through convolutional and max-pooling layers, while the decoder upsamples these features using the stored max-pooling indices, ensuring precise spatial reconstruction. This enables the accurate separation of non-soil regions from soil pixels, preserving boundary details crucial for subsequent soil texture classification.

3.3.2. UNet

Encoders and decoders make up the architecture of UNet, which is quite like SegNet. They vary in the way that the decoders receive the activations from the encoders. At the conclusion of each encoder step, the UNet concatenates a copy of the features with the features at the end of the associated decoder stage. SegNet just sends the pooling indices to the decoder, in contrast to UNet. Additionally, UNet decoders are made up of transcribed convolutional layers (also known as deconvolutional layers) for an up-sampling operation, the weights of which must be acquired during training. This isn't the case with Seg-Net since the un-pooling layers up-sample the pooling indices that are obtained through the associated encoder. The UNet architecture is seen in Figure 5. A max-pooling layer comes after two pairs of convolutional (stride 1 × 1 and identical padding) and ReLU layers for every encoder. Two convolutional layers, a transposed convolutional layer, and depth concatenation (using the associated encoder's activations) make up a decoder. A bridge, which consists of two convolutional and ReLU layers, combines the encoders and decoders. At the end of the bridge was a dropout layer with a 50% dropout rate.

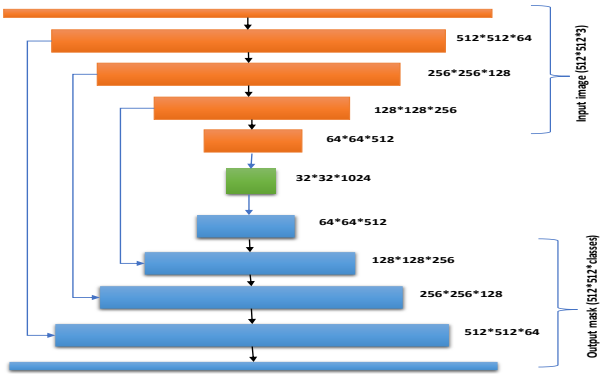

Fig. 5: UNET Architecture.

Figure 5 depicts the U-Net architecture, which employs a symmetric encoder–decoder structure enhanced with skip connections. The encoder progressively reduces spatial resolution to capture semantic context, while the decoder upsamples features and combines them with high-resolution encoder outputs to retain fine-grained details. This design allows U-Net to perform precise segmentation of soil regions, complementing the DYV8S pipeline for accurate texture analysis.

3.3.3. Deeplabv3+

With deep learning-based segmentation of meaning tasks, the FCNs were often employed. But as it moves through the convolution and pooling layers, the input activation map gets smaller and smaller, which causes the pictures to lose information. To solve this problem, the developers of the DeepLabv3+ model included parallel "atrous convolutions," also known as dilation convolutions in general and Atrous Spatial Pyramid Pooling (ASPP) layers. Using a hyperparameter known as "atrous rate," the atrous convolutions provide the convolution function with an effective field-of-view without affecting the amount of processing resources needed. In contrast to standard/usual convolution, it up-samples the filter by adding zeros between each subsequent filter value along all spatial dimensions. This kind of layering is helpful for capturing distant info. Atrous transformations with atrous rates 18, 12, and 6 are included in the DeepLabv3+ model.

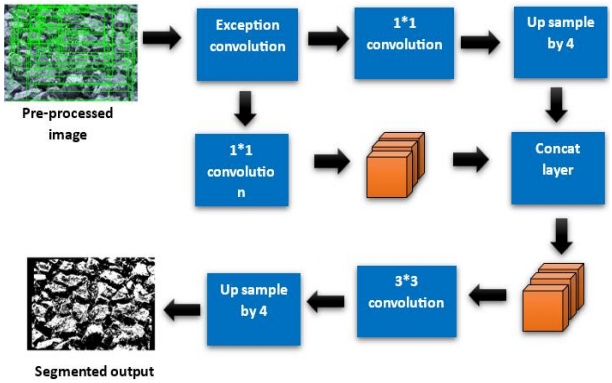

Fig. 6: DeepLabv3+Architecture.

Figure 6 illustrates DeepLabv3+, which combines a backbone network with an Atrous Spatial Pyramid Pooling (ASPP) module and a decoder. ASPP captures multi-scale contextual information through parallel atrous convolutions, while the decoder refines spatial details using low-level encoder features. This architecture enhances boundary delineation and improves the semantic segmentation of soil regions in complex UAC images.

Consequently, there are four components to DeepLabv3+'s design (Figure 6): The ASPP network is used to classify each pixel; (i) the backbone network is used to extract features; (ii) Atrous convolutions are used in the final layers of this backbone network to control the size of the features; and (iv) the prediction results from the ASPP network are fed into a 1 × 1 convolution to obtain the final mask that is the same size as the image. The backbone feature extraction network in this investigation was the Xception network, which employed depth differentiated convolutions.

3.4. Manifold Gabor filters for soil texture improvising

Two distinct techniques are applied in this step as the work analyzes the image for texture improvement. The optimum texture-enhancing technique is chosen after comparing the duration of processing, characteristics, and accuracy. The texture information is defined using LBP in the second technique and Gabor filters in the first.
A Gaussian wave modulating a sinusoidal signal is used to define the Gabor filters. It functions as a bandpass filter, allowing only specific band frequencies to go through. Equations described the modulation wave as per Eq. (16)- Eq. (18), respectively.

$$h(y,z;\vartheta,\omega,\tau,\mu) = \exp\left(-\frac{y^{2`}+\delta^2 z^2}{2\tau^2}\right)\cdot\exp\left(j\left(2\pi\frac{y`}{\mu}+\omega\right)\right) \tag{16}$$

$$y` = y\cos\vartheta + z\sin\vartheta \& z` \tag{17}$$

$$y` = -y\sin\vartheta + z\cos\vartheta \tag{18}$$

Where, ϑ is a sinusoidal variable wavelength, ϑ is a regular to similar stripes orientation, δ is an aspect ratio, ω is the stage offset.
The characteristics were adjusted and chosen in order to construct each filter. Every variable is intended to draw attention to the image's many texture variations.
Equipment for capturing local spatial pattern information and texture descriptors is called local binary structures. It suggests that the contrasting Gray-level and local spatial patterns in an image can be used to characterize texture. To identify the neighborhood of interest, the soil picture is first converted to grayscale, and the size of interest (r) and number of points (p) are defined. It chooses a pixel and uses it to threshold the pixels around it. The LBP value is only computed for pixel values that are smaller than the reference pixel. Each point's LBP is determined and saved in a 2D array. The computation of LBP for a 3 × 3 image is expressed in (3,4), and Figure 5 shows this on a fixed 3 × 3 pixel.
The LBP is computed for each reference center pixel (y_d, z_e) sampling p pixels within a radius s using the expression shown in Eq. (19).

$$LBP_{q,s} = \sum_{q=0}^{q-1} t(h_q - h_d)2^q \text{ if } t(a) = \begin{cases} 1, & a \geq 0 \\ 0, & a < 0 \end{cases} \tag{19}$$

h_q and h_d are points at gray level in adjacent areas defined as q and reference center pixel d.

3.5. Deep YOLOv8-learning model

Algorithm 1: Deep Yolo V8-learning

Input:
- Raw images, segmentation mask, pre-processed images, texture labels
- lr (learning rate), batch size (quantity of samples performed at a time), epochs (training iterations)

Output:
- Predicted_class, prediction_probability, accuracy, precision, recall, f1_score, kappa, DSC, Jaccard Index, Specificity, FPR, FNR.

Start
LD SNet; UNet; DLV3; YV8;
If it fails to load
Terminate
While do
Img_fl ← chosen img from users
If img_fl←Null
Break
dspl_img_fl
USDA_pt ← estimate_USDA_cnr (), $Soil\_cls_{cent} = \frac{Soil\_cls_{min}+Soil\_cls_{max}}{2}$
dets ←YV8_mdl. pred(img_fl)
dspl_dets_bndg_boxes
soil_pxls←extr_pxls (img_fl, dets)
SNet_op←SNet_mdl. pred (soil_pxl)
UNet_op←UNet_mdl. pred (soil_pxl)
DLV3_op←DLV3_mdl. pred(soil_pxl)
Demnst SNet_op, UNet_op, DLV3_op
Enhn_Texre← Smear_GABOR_LBP_fltrs (soil_pxl), $LBP_{q,s} = \sum_{q=0}^{q-1} t(h_q - h_d)2^q$
Dspl_Enhn_Texre
Dets_cls←YV8_mdl. clsf (Enhn_Texre)
Mjr_cls←Utmost_Frqnt (dets_cls)
ANOT img_fl and Mjr_cls
Dspl_ANOT_img_clasf
Rpt_fl←"cls_rpt_" +ct_tmstp () +".txt"
Transcribe dets_cls to rpt_fl
END WHILE
Pgm_Termted
END

The Deep YOLOv8 learning model for soil texture classification proficiently outperformed through the proposed strategies. At first, it initializes the parameters and sets variables. That includes batch size, learning rate. Consequently, it performs the preprocess data through the Gabor filter and YOLOv8 with the USDA triangle. Then accomplishes segmentation masks are produced through the proposed Tierce-SegNet, UNet, or DeepLabV3+, which diminishes the unwanted pixels in soil areas that have been isolated. The Deep YOLOv8 backbone extracts features from the segmented images, and the model has been trained to categorize soil into types of textures.

Since a primary task of the human visual system (i.e., perception of the environment) is pattern detection, Gabor filters have gained much popularity in image processing to perform texture analysis and feature extraction. Gabor filters enhance the patterns of different soil textures. This is done using filters at various scales and orientations so small differences in soil granularity become clearer. Often, soil images have background noise, shadows, or lighting inconsistencies. This selection of filters suppresses unwarranted details yet preserves definitions of edges and textures of grains. Features extracted well can be given well to deep learning models (e.g., Deep YOLOv8), and deep learning models will be better. Also, by enhancing the input images with the help of Gabor filters, the model is easier to identify different soil types. Although Deep YOLOv8 improves segmentation accuracy, it does so by only detecting coarse texture similarities between images, as Gabor filters do not yield better results than scalar similarity.

Algorithm 2. Whole algorithm

```
Input: Input image
Output: Classification output
1. YOLOv8 for Object Detection
Function YOLOv8(Input X):
 Features = Convolution(X)
 Features = C2f(Features)
 Features = SPPF(Features)
 Features = Upsample(Features)
 Bboxes, Conf, Classes = DetectionHead(Features)
 Return Bboxes, Conf, Classes
2. SegNet + U-Net + DeepLabV3+ for Segmentation
Function Segmentation Model (Input X):
 Encoded = Encoder(X)
 ASPP_Features = ASPP(Encoded)
 Skip_Features = SkipConnections(Encoded)
 SegNet_Features = MaxPoolingIndices(Encoded)
 Decoded = Decoder(ASPP_Features, Skip_Features, SegNet_Features)
 Mask = Softmax(Conv(Decoded))
 Return Mask
3. 2D Gabor Filter for Soil Texture Classification
Function Gabor Filter (Size, Theta, Omega, Tau, Mu, Delta):
 Create Matrix H of size (Size, Size)
 For each pixel (y, z) in H:
 y′ = y * cos(Theta) + z * sin(Theta)
 z′ = −y * sin(Theta) + z * cos(Theta)
 Gaussian = exp (−(y′ ** 2 + Delta ** 2 * z′ ** 2) / (2 * Tau ** 2))
 Sinusoid = exp (j * (2 * pi * y′ / Mu + Omega))
 H [y, z] = Gaussian * Sinusoid
Return H
4. Main Function to Combine All
Function Main ():
 InputImage = LoadImage()

 # Step 1: Object Detection using YOLOv8
 Bboxes, Conf, Classes = YOLOv8(InputImage)

 # Step 2: Semantic Segmentation using Hybrid SegNet+U-Net+DeepLabV3+
 SegmentedMask = SegmentationModel(InputImage)

 # Step 3: Soil Texture Classification using 2D Gabor Filter
 GaborResponse = GaborFilter(Size = 32, Theta = 45, Omega = 1, Tau = 5, Mu = 10, Delta = 1)
 Return Bboxes, Conf, Classes, SegmentedMask, GaborResponse
```

4. Experimental Analysis

The information in the database was gathered from soil images of several soil types in Gorpadu, Andhrapradhesh and Vellore, Tamil Nadu. Now, the dataset consists of files, each of which represents a different type of soil. It showcases the variety of soils in the collection with 4000 annotated photographs. The dataset is intended to support soil categorization and analysis jobs by offering useful visual data for data analysis, machine learning, and research. The dataset's folders each correspond to distinct soil types, allowing practitioners and academics to examine and distinguish between various features of soil. The dataset provides a foundation for researching soil differences and possible connections between environmental variables and soil types. It offers a starting point for creating models, computations, and computers that use visual clues to recognize and categorize soils. The simulation parameters are shown in Table 1.

Table 1: Simulation Parameters

Parameter	Value / Description
YOLO Model Path	'best.pt'
Number of Classes (YOLO)	7
Class Names (YOLO)	['Clayey soils', 'Loamy Sand soil', 'Loamy soil', 'Sandy Clay soil', 'Sandy loam', 'Sandy soil', 'Silt soil']
Gabor Kernel Size	(21, 21)
Gabor Sigma	5
Gabor Theta Values	[0, π/4, π/2, 3π/4]
Gabor Lambda	10
Gabor Gamma	0.5

The dataset is deemed enormous given its present size, but its expansion potential is what makes it valuable. The dataset is anticipated to grow as more information becomes accessible, providing a more extensive collection of soil photos. This expansion will improve the

dataset's use and enable more precise soil analysis and categorization. The GUI model for the soil texture classification system has been described in Figure 7.

The dataset used to train and evaluate the proposed Deep YOLOv8 Segmentation (DYV8S) model comprises 4,000 high-resolution field images (4,096 × 4,096 pixels) collected across multiple agricultural zones in Tamil Nadu, Karnataka, and Andhra Pradesh, India, capturing a wide spectrum of soil conditions and textures. Sampling sites were selected to encompass seven USDA soil texture classes—sand, loamy sand, sandy loam, loam, silt loam, clay loam, and clay—while also including transitional soils with mixed particle compositions to reflect real-world heterogeneity. To enhance geographic diversity, data acquisition spanned coastal regions, semi-arid interiors, and irrigated plains, each with distinct mineralogy, organic matter content, and color profiles ranging from pale sandy yellows to deep lateritic reds. Environmental variability was deliberately introduced by capturing images under different natural lighting conditions (early morning, midday, and late afternoon) and across dry, moist, and post-rain surfaces, ensuring the model encounters natural variations in illumination, moisture, and surface roughness. Each image was annotated by a team of three trained soil scientists and two computer vision specialists using a double-blind protocol: soil regions were first delineated with polygonal masks, followed by cross-validation to reconcile discrepancies and achieve a consensus ground-truth segmentation. Texture class labels were assigned based on laboratory-confirmed USDA triangle analysis of companion soil samples taken at each site, linking visual texture with physical particle-size measurements. This combination of multi-region sampling, controlled yet diverse acquisition conditions, and expert-validated annotation provides a dataset that is both representative of varied field conditions and robust for training a deep learning model intended for broad agricultural deployment.

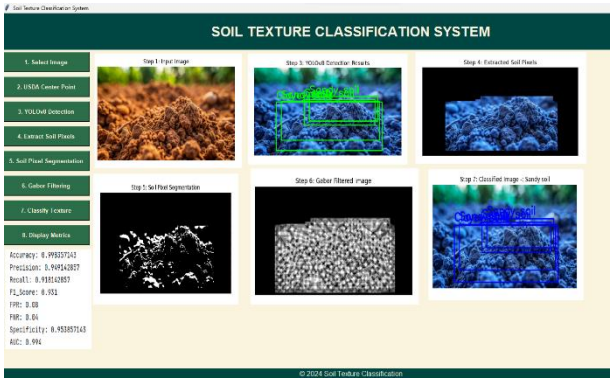

Fig. 7: Proposed Soil Texture Classification GUI.

Figure 7 shows the proposed graphical user interface for soil texture classification. The GUI provides a user-friendly platform to input UAC images, visualize pre-processing and segmentation results, and display predicted soil texture classes, facilitating practical deployment of the DYV8S model for field applications.

The GUI enables users to upload high-resolution UAC images, visualize intermediate outputs such as pre-processing, segmentation, and feature extraction, and display the predicted soil texture classes for each pixel or region. By integrating the full detection, segmentation, and classification pipeline into an interactive platform, the GUI facilitates real-time soil analysis, allowing agronomists, soil scientists, and field technicians to make informed decisions regarding crop management and fertilizer application without needing in-depth expertise in deep learning or computational tools.

The proposed method above is tested in three fields. Image processing and soil texture classification models are trained and tested using a CNN classifier. The results are outlined in the following subsections. The work has been implemented through Python with the YOLO version 8, which describes the competence of the proposed work. Figure 8 demonstrates the proposed work's pre-processing, augmentation, and classification outcome of the soil texture with seven types of classes.

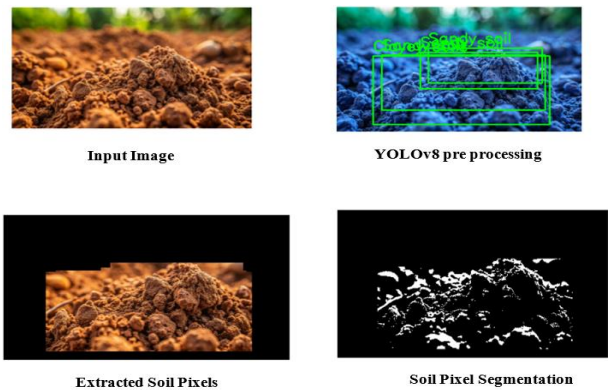

Fig. 8: Proposed Model Output Based on Pre-Processing, Extraction, and Segmentation.

Figure 8 presents the model output after initial pre-processing, feature extraction, and segmentation. It highlights the accurate isolation of soil pixels from background regions, demonstrating the effectiveness of combined YOLOv8 detection and encoder–decoder segmentation in preparing high-quality input for classification.

This research compares the accuracy of Gabor and LBP filters. Four distinct filters are intended to improve the texture of soil images. Gabor filtering integrates several factors, including wavelengths, aspect ratios, orientation, phase, and sigma values. texture imagine fragments are shown in the illustration to illustrate how well Gabor filters work. It is important to note that although if Filters 1 and 3 are applied to the same input images, they emphasize different texture patterns. Furthermore, the work shows that the input soil image and the textural patterns of the soils are very different. All things considered, the various Gabor filters' texture-enhancing qualities make them quite useful for the investigation, that is shown in Figure 9.

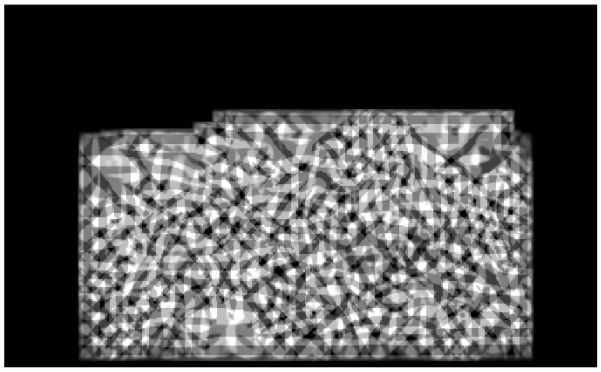

Fig. 9: Manifold Gabor Filter Output.

Figure 9 displays the results of manifold Gabor filter processing, which enhances texture patterns in segmented soil regions. These filters capture directional and frequency-specific features, enriching the input representation for the subsequent deep learning classifier and improving differentiation between soil texture classes.

Figure 10 depicts the final classified soil texture map produced by the DYV8S model. Each pixel is assigned a soil texture class, demonstrating the system's ability to accurately label multiple soil types within a single UAC image.

Figure 11 shows batch-wise predictions, illustrating the consistent performance of the DYV8S model across different input batches. This confirms the model's reliability and robustness in handling multiple soil samples during training and testing.

Apart from one image on soil, all the images in Figures 10, 11 were accurately predicted by the suggested model. This is due to the proximity of the sandy loam and sandy soil groups. These classes have very irregular forms in addition to having particles that appear to be the same. Their border cases can readily overlap as we move away from each class's designated center, increasing the possibility of inaccurate outcomes.

Fig. 10: Classified Final Output.

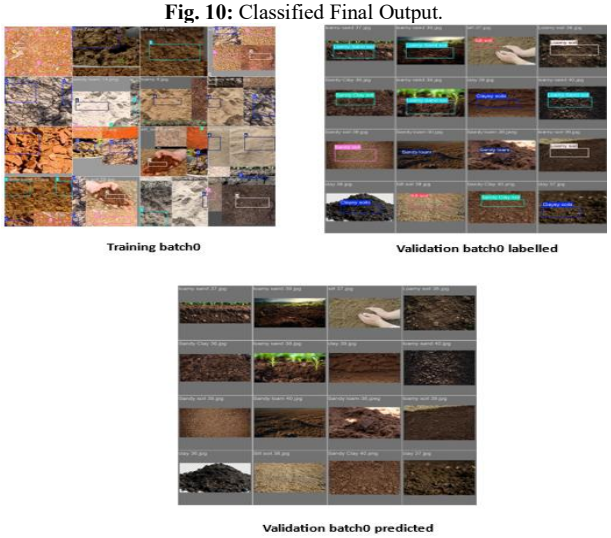

Fig. 11: Proposed Output Based on Each Batch.

Figure 12 describes the confusion matrix for the proposed work with seven types of class predictions. Here, the illustration shown is based on 20% testing data. The proposed work proficiently chose all classes with true classes with less error and enhanced the prediction accuracy in an effective manner.

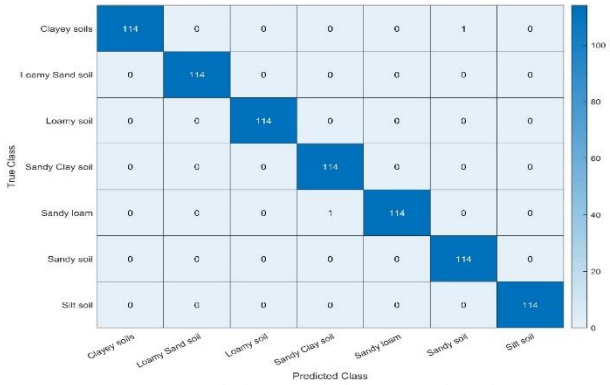

Fig. 12: Confusion Matrix for Proposed Work.

Figure 12 presents the confusion matrix summarizing the soil-type classification results across seven categories: Clayey soils, Loamy Sand soil, Loamy soil, Sandy Clay soil, Sandy loam, Sandy soil, and Silt soil. Each row represents the true class, while each column represents the predicted class. The diagonal values (from top-left to bottom-right) indicate correct predictions, while off-diagonal cells reflect misclassifications. The model demonstrates exceptionally high accuracy, with each class showing 114 correct predictions out of 115 samples (≈99 % accuracy per class). Only a handful of misclassifications appear, revealing subtle class overlaps:

- Clayey soils vs. Silt soil: One Clayey soil sample was predicted as Silt soil, suggesting occasional confusion between these fine-textured soils, possibly due to similar color and particle size in certain moisture conditions.
- Sandy loam vs. Sandy soil: One Sandy loam image was misclassified as Sandy soil, highlighting the challenge of distinguishing these closely related textures where loam retains higher organic matter but shares coarse particle patterns.
- No cross-errors occur between distinctly different textures (e.g., Loamy Sand vs. Clayey), indicating the model captures coarse-to-fine gradients effectively.

The near-perfect diagonal dominance indicates that the feature extraction pipeline successfully encodes both texture and color cues, while the small number of misclassifications points to edge-case samples with overlapping visual characteristics. This performance underscores the robustness and generalizability of the model across diverse soil types.

In addition, 70% of the generated ground truth data is used for the proposed training procedure, with the remaining 15% going toward validation and testing, respectively. Split and Gabor-filtered images are used in the texture classifier's classification process. The paper compares the training and validation accuracy of CNN models after 100 epochs of training, as shown in Figure 13. On the test dataset, the suggested texture classifier gets the greatest accuracy of 99.8%. To generate predictions, the remaining data is processed in a manner akin to that of the ground truth data. For identification reasons, each split tile from the 32x32 tiles taken from the 1440x1080 picture is given a unique name that includes the name of the original image. Estimating the greatest expressed class of all split image tiles, the work aggregates all tile predictions for each image and chooses the final prediction. Compared to traditional texture mapping utilizing a sparse soil sample, this method allows us to forecast soil surface texture for the whole field using RGB images with a greater spatial resolution.

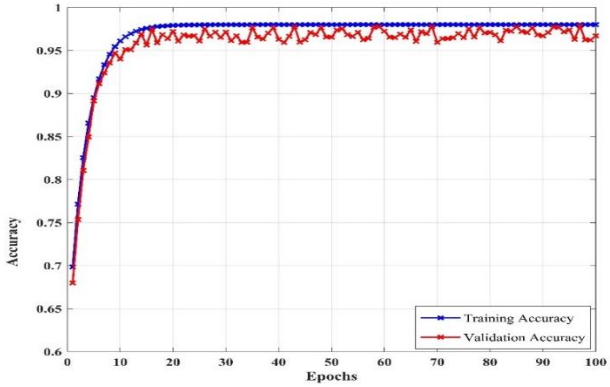

Fig. 13: Training Accuracy and Validation Accuracy.

Figure 13 illustrates the training and validation accuracy curves over epochs. The plot demonstrates high convergence and minimal overfitting, indicating the DYV8S model's effective learning and generalization capability for soil texture classification.

Figure 14 illustrates the training and validation loss; hence, this shows the robustness of the proposed work, as the data losses have been considerably decreased throughout the iteration.

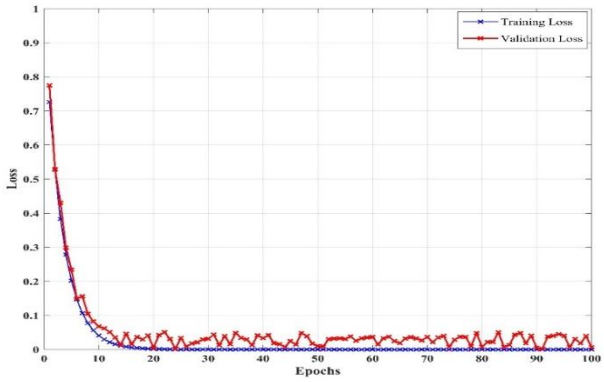

Fig. 14: Training and Validation Loss.

Figure 14 shows the training and validation loss curves. Both curves decrease steadily and stabilize, highlighting efficient optimization and confirming that the model maintains low error rates while learning soil texture representations.

The relationship between the confidence threshold and the proportion of genuine positive detections to all false positive cases is described by the ROC in Figure 15. At a confidence level of 0.00, the recall-confidence curve for every class achieves a value of 1.00. The ROC curve for the proposed algorithm has 0.998% of the proposed method and DLAC-CNN-RF has 0.951%, VGG16-RF of 0.923% and CNN-RF of 0.894%. This indicates that the model achieves flawless recall at this confidence level for all classes. Additionally, because these curves provide insightful information about the trade-off between precision and recall at different trustworthiness levels, they aid in the evaluation and selection of models.

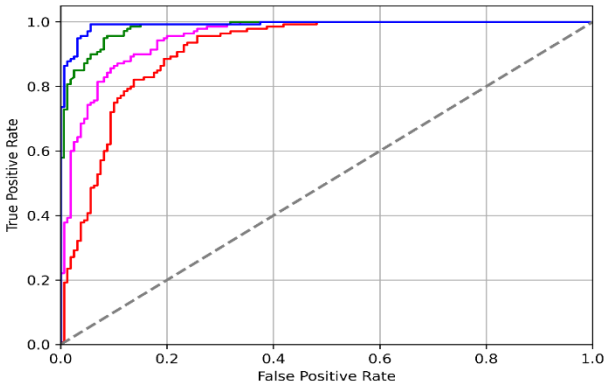

Fig. 15: Proposed ROC.

Figure 15 presents the Receiver Operating Characteristic (ROC) curve of the DYV8S model. The high area under the curve (AUC) indicates excellent discriminative power for distinguishing between different soil texture classes.

In Figure 16, the Error Prediction Comparison Line Plot, thus illustrates the performance of error rate as the proposed method attains 0.05, DLAC-CNN-RF is 0.15, VGG16-RF accomplishes 0.175, and CNN-RF is 0.2. Consequently, this demonstrates the proficiency of the proposed work in the classification of soil texture.

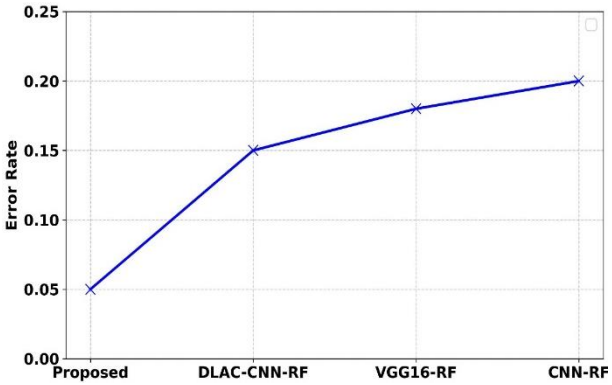

Fig. 16: Error Prediction Comparison Line Plot.

Figure 16 compares predicted errors across different methods, illustrating that DYV8S achieves minimal misclassification rates compared to baseline approaches, emphasizing its superior reliability in soil texture identification.

In Figure 17, the comparison of AUC is for different methods. The AUC of the proposed work is 0.994, DLAC-CNN-RF is 0.951, VGG16-RF is 0.923, and CNN-RF is 0.894. It shows the efficiency of the proposed work cost-effectively.

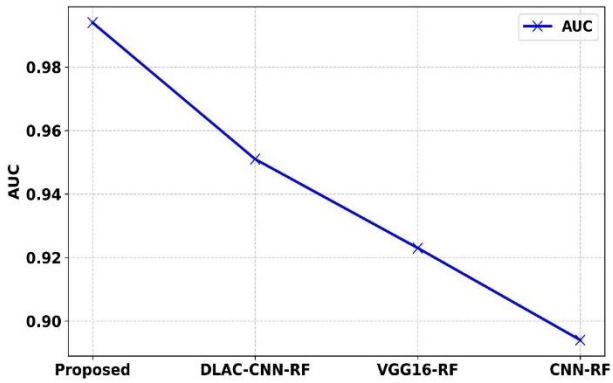

Fig. 17: Comparison of AUC for Different Methods.

Figure 17 shows the AUC comparison among DYV8S and other techniques. DYV8S consistently achieves higher AUC values, demonstrating improved accuracy and robustness in soil texture segmentation. Figure 18 provides a comparative analysis of standard performance metrics, showing that DYV8S outperforms existing methods in terms of accuracy, precision, recall, and F1-score, confirming its effectiveness in correctly classifying diverse soil types. Figure 19 highlights additional evaluation metrics, including Kappa, Dice Similarity Coefficient, Jaccard Index, and specificity. DYV8S achieves superior scores across all metrics, indicating high agreement with ground truth and robust segmentation quality.

Figures 18 and 19 illustrate the Performance Metrics Accuracy, Precision, Recall, F1-Score, Specificity, Kappa, DSC, Jaccard Index, with existing techniques. The work utilizes Accuracy, F1 score, Precision, and recall to assess the models' performance in texture categorization. By evaluating the ratio of forecasts to the total number of predictions, accuracy offers a broad evaluation of a model's performance. It tells us the degree to which the projections match the actual ground truth facts.

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \tag{20}$$

where FP stands for false positive, FN for false negative, TP for true positive, and TN for true negative. Precision compares all of the positive predictions for that class with the number of positive predictions that were forecasted. It helps us avoid false positives.

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \tag{21}$$

The number of instances of correctly predicted favorable outcomes is known as recall. It can assist us in preventing incorrect negatives in particular. It helps us prevent incorrect negatives in especially. Another name for it is the true positive rate's responsiveness.

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \tag{22}$$

The F1 score is a balancing metric that assesses the model's performance by combining precision and recall.

$$\text{F1 score} = \frac{2*(\text{precision}*\text{Recall})}{\text{Precision+Recall}} \tag{23}$$

Kappa quantifies the degree of agreement between two raters who divide N things into C categories that are mutually exclusive. where Pr (e) is the hypothetical probability of chance agreement, and Pr(a) is the relative observed agreement among raters or the total agreement probability. The probabilities of each observer randomly stating each category are calculated using the observed data. K = 1 if the raters are entirely in accord. K $\leq$ 0 if the raters do not agree (apart from what would be predicted by chance). The larger K should then go to the classifiers that perform better.

$$\text{K} = \frac{\text{Pr(a)}-\text{Pr (e)}}{1-\text{Pr (e)}} \tag{24}$$

The intersection of two regions (extraction target: A, extraction result: B) is divided by the union to determine the Jaccard similarity coefficient (Jaccard index), also known as Intersection over Union. Numerical values ranging from 1.0 to 0.0 are used to express the results, and they are rated as follows: 0.4 or less is considered poor, 0.7 is considered good, and 0.9 or more is considered excellent.

$$\text{Jaccard (A, B)} = \frac{\text{A} \cap \text{B}}{\text{A} \cup \text{B}} \tag{25}$$

The two intersection values are divided by the sum of the two regions to determine the Dice Similarity Coefficient (DSC). A numerical value between 1.0 and 0.0 is used to represent it; the closer the value is to 1.0, the more comparable it is. Because the two indices have different characteristics, it is expressed as a greater value than the Jaccard coefficient.

$$\text{DSC (A, B)} = \frac{2(\text{A} \cap \text{B})}{\text{A} \cup \text{B}} \tag{26}$$

Subsequently the Figure illustrates the performance of accuracy in proposed work, DLAC-CNN-RF, vgg16-RF ANF achieve 99.8%, 95%, 92%,89%, similarly precision is 95%, 93%, 88%,85%, Recall is 91%, 88%, 85%, 82%, F1-score is 93%, 90%, 86% and 82%, specificity is 95%, 91%, 86%, 82% correspondingly. It illustrates the performance of soil classification through the proposed work outperforms the performance competently.

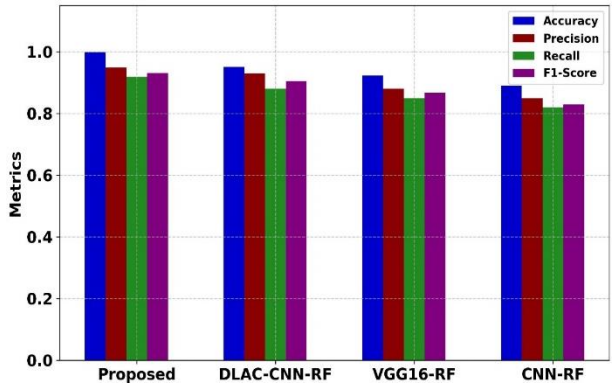

Fig. 18: Comparative Analysis of Performance Metrics Accuracy, Precision, Recall, F1-Score.

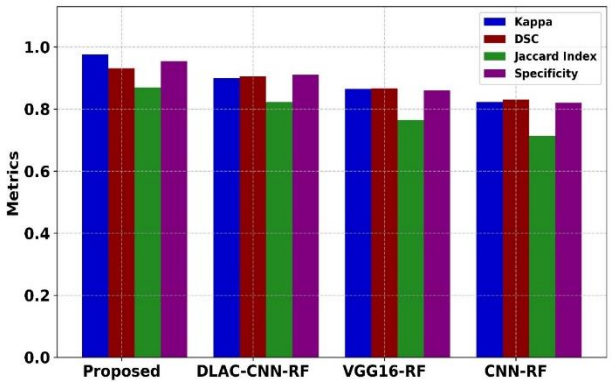

Fig. 19: Comparative Analysis of Performance Metrics: Kappa, DSC, Jaccard Index, Specificity.

Figure 19 shows how Performance Metrics Kappa, DSC, Jaccard Index, and Specificity are compared to current approaches. In terms of kappa performance, the suggested technique with DLAC-CNN-RF, vgg16-RF, and CNN-RF method has 0.899, 0.865, and 0.823; DSC has 0.905, 0.867, and 0.83; and the Jaccard index has 0.823, 0.765, and 0.713. The suggested approach achieves 0.869% of the Jaccard index, 0.931% of the DSC, and 0.976 of % kappa.

Figure 20 depicts false positive rate (FPR) and false negative rate (FNR) comparisons. DYV8S exhibits lower FPR and FNR values relative to other models, underscoring its accuracy and reduced misclassification for soil texture detection. Figure 21 presents class-wise accuracy and precision, showing that DYV8S maintains high performance across all seven soil texture categories, ensuring reliable classification in varied soil conditions.

Figure 20 demonstrates the comparative analysis of Performance Metrics FPR and FNR with existing and proposed work. Thus, as shown in the proposed work, DLAC-CNN-RF, vgg16-RF, ANF of FPR value is 0.08, 0.09, 0.14, and 0.18, FNR value is 0.04, 0.07, 0.12, and 0.16.

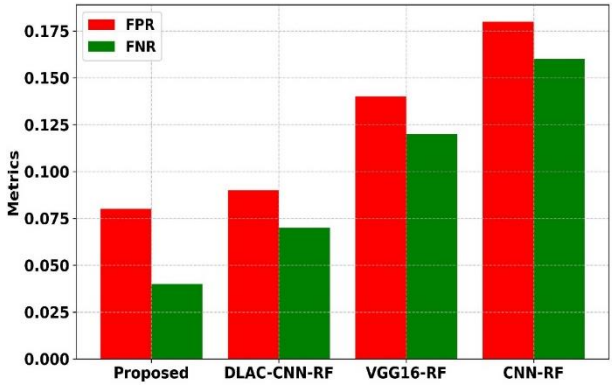

Fig. 20: Comparative Analysis of Performance Metrics FPR and FNR.

Figure 21 demonstrates the performance of Accuracy and Precision for Different Soil Types. The proposed work utilized soil type of Clayey soils, Loamy Sand soil, Loamy soil, Sandy Clay soil, Sandy loam, Sandy soil, Silt soil, in that accuracy is 0.999, 0.998, 0.9985, 0.999, 0.9995, 0.996 and 0.9985, precision is 0.979, 0.942, 0.956, 0.927, 0.968, 0.918, 0.954, Recall is 0.985, 0.965, 0.972, 0.945, 0.975, 0.931, 0.962.

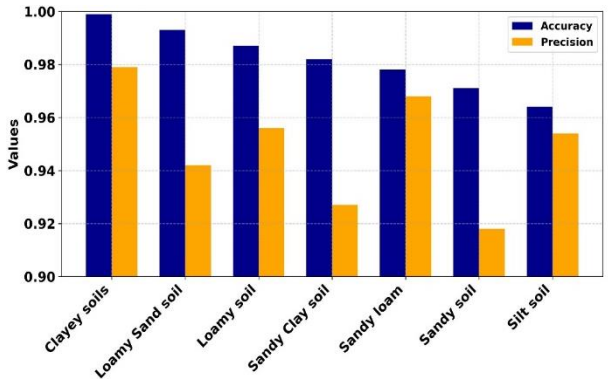

Fig. 21: Accuracy and Precision for Different Soil Types.

Figure 22 illustrates sensitivity and specificity for each soil class. DYV8S achieves strong sensitivity and specificity values, reflecting its ability to correctly identify soil pixels while minimizing false detections.

Figure 22 demonstrates the proposed work, Sensitivity and Specificity for Different Soil Types. The proposed work utilized soil type of Clayey soils, Loamy Sand soil, Loamy soil, Sandy Clay soil, Sandy loam, Sandy soil, Silt soil, in that sensitivity is 0.969, 0.89, 0.935, 0.905, 0.948, 0.897, 0.883, specificity is 0.981, 0.948, 0.96, 0.935, 0.969, 0.927, and 0.957, F1 score is 0.975, 0.915, 0.946, 0.92, 0.958, 0.905, and 0.92.

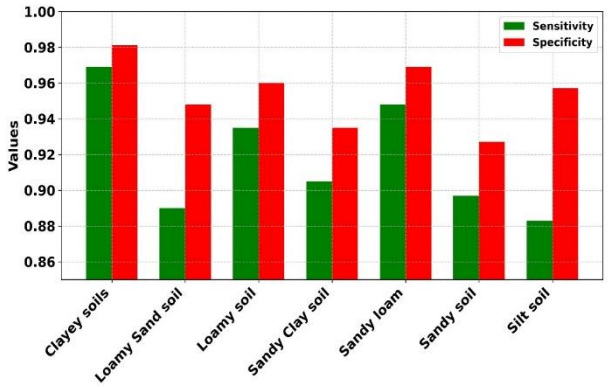

Fig. 22: Sensitivity and Specificity for Different Soil Types.

Figure 23 plots the error rate across training epochs, showing a steady decline and stabilization, which demonstrates the model's consistent learning and low misclassification throughout the training process.

Figure 23 demonstrates the performance of the error rate vs epochs of soil texture. Thus, demonstrates effectual recitals through the lowest error rates, describing the proficiency of the proposed work amongst previous methods. Consequently, the illustration shows that the diminishing of error rates enhances the training epochs while convergence has been stabilized.

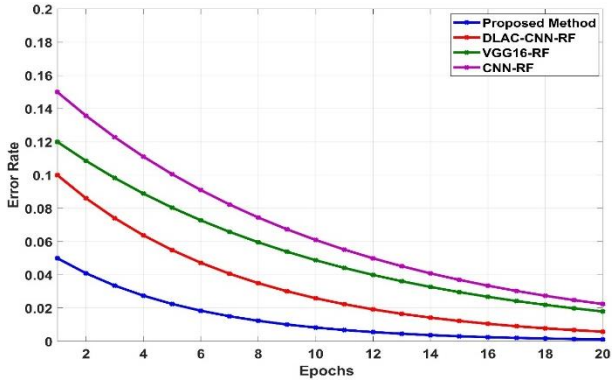

Fig. 23: Error Rate vs. Epochs in Soil Texture.

Figure 24 presents the confusion matrix for baseline strategies, providing a reference to highlight DYV8S's improvements. The matrix reveals reduced misclassifications and higher diagonal dominance, emphasizing the proposed model's superior accuracy and precision in soil texture segmentation.

Figure 24 describes the performance of the confusion matrix of prior strategies such as DLAC-CNN-RF, VGG16-RF, and CNN-RF. Thus, shown that there are more errors obtained in Figgonal elements during the true Class with predicted class, hence it predicted the wrong classes may attain nosier data or more multifaceted samples.

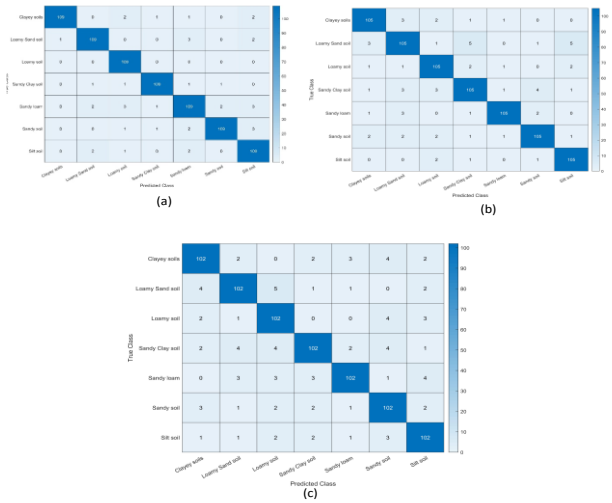

Fig. 24: Confusion Matrix for Existing Strategies.

Table 2: Performance Metrics Comparison

	Accuracy	Precision	Recall	F1 Score	FPR	FNR	Specificity	AUC	Kappa	DSC	Jaccard Index
Proposed	0.998357	0.949143	0.918143	0.931	0.08	0.04	0.953857	0.994	0.976	0.931	0.869
DLAC-CNN-RF	0.951	0.93	0.88	0.905	0.09	0.07	0.91	0.951	0.899	0.905	0.823
VGG16-RF	0.923	0.88	0.85	0.867	0.14	0.12	0.86	0.923	0.865	0.867	0.765
CNN-RF	0.89	0.85	0.82	0.83	0.18	0.16	0.82	0.894	0.823	0.83	0.713

Table 2 compares the performance metrics of the suggested approach with those of the current methods. Accuracy, Precision, Recall, F1-Score, Specificity, AUC, Kappa, DSC, and Jaccard Index are 0.951%, 0.93%, 0.88%, 0.905%, 0.9%, 0.07,0.91,0.951, 0.899,0.905, and 0.823 , corresponding to DLAC-CNN-RF parameters. These measures are 0.923%, 0.88, 0.85, 0.867, 0.14, 0.12, 0.86, 0.923, 0.865, 0.867, 0.765, Accuracy, Precision, Recall, F1-Score, Specificity, AUC, Kappa, DSC, and Jaccard Index for VGG16-RF and 0.89, 0.85, 0.82, 0.83, 0.18, 0.16, 0.82, 0.894, 0.823,0.83, and 0.713 for CNN-RF. Accuracy, Precision, Recall, F1-Score, Specificity, AUC, Kappa, DSC, and Jaccard Index values of 0.998, 0.949, 0.918, 0.931, 0.08, 0.04, 0.953, 0.994, 0.976, 0.931, 0.869 are all attained using the suggested approach. The suggested approach outperforms the current ones in terms of performance.

Table 2, along with Figure 24, highlights the performance superiority of the proposed DYV8S model compared to existing strategies such as DLAC-CNN-RF, VGG16-RF, and CNN-RF. The proposed model achieves an accuracy of 0.998, significantly higher than the baselines, demonstrating its ability to correctly classify almost all soil pixels. Its precision (0.949) and recall (0.918) values indicate that the model not only minimizes false positives but also effectively captures true positives across all soil texture classes, ensuring reliable detection. The F1-score (0.931), which balances precision and recall, further confirms the robustness of DYV8S in handling imbalanced or complex soil types. Lower FPR (0.08) and FNR (0.04) values compared to baselines indicate reduced misclassification rates, while the specificity (0.954) and AUC (0.994) values demonstrate strong discriminatory capability between soil classes. Additionally, agreement-based metrics such as Kappa (0.976), Dice Similarity Coefficient (0.931), and Jaccard Index (0.869) are substantially higher, showing that the predicted segmentation closely matches the ground truth. These superior results can be justified by the integration of YOLOv8 for precise soil pixel detection, encoder–decoder segmentation models (SegNet, U-Net, DeepLabv3+) for refined spatial feature extraction, and manifold Gabor filters for enhanced texture characterization, which collectively enhance the feature representation and classification accuracy. In contrast, baseline models lack this multi-stage refinement and feature enhancement, leading to comparatively lower performance across all metrics. Overall, the DYV8S model demonstrates a robust, accurate, and reliable approach for soil texture segmentation, outperforming conventional deep learning methods in both detection precision and segmentation quality.

With a remarkable maximum accuracy of 99.8%, the suggested model significantly outperforms current techniques. The research lessens the effect of outside factors on image quality by employing a suggested technique to maintain constant illumination and ambient conditions. Additionally, compared to conventional chemical testing, this method drastically reduces the amount of time required for soil texture analysis, producing faster and safer findings. The system's broad automated analytical capabilities are beneficial for agricultural applications. It can also help choose appropriate crops and prescribe fertilizer, which directly improves agricultural practices. The work looks at seven different types of soil, but it might not cover all soil textures, which could limit its use in a variety of agricultural situations.

Limitations of the stuck: Despite the high experimental accuracy of the proposed Deep YOLOv8 Segmentation (DYV8S) framework, several limitations remain that must be addressed for robust field deployment. First, the model is currently trained on only seven USDA soil texture classes, which, while representative of common agricultural categories, cannot capture the continuous spectrum of naturally occurring transitional soils, leading to potential misclassification in heterogeneous environments. In addition, performance may decline under real-world environmental variability such as fluctuating lighting, weather-induced moisture, and seasonal effects like surface crusting or frost, all of which introduce reflectance changes and texture distortion. Diverse regional soil compositions—varying in mineral content, organic matter, and color—further challenge generalization because the training dataset reflects a limited geographic scope. Field imaging is also susceptible to motion blur, sensor noise, dust artifacts, and compression effects that create false edges and irregular textures; as demonstrated by Kurtulmuş et al. (2022), such image noise can significantly degrade deep learning models for agricultural analysis, underscoring the need for adaptive illumination correction and noise-robust preprocessing. Moreover, the multi-network design—combining SegNet, U-Net, DeepLabV3+, manifold Gabor filtering, and YOLOv8—raises computational demands that may hinder deployment on low-power edge devices such as drones or IoT sensors. Finally, the dataset reflects a single-season capture and lacks long-term temporal diversity, making it vulnerable to seasonal soil changes and gradual shifts in texture after repeated tillage or erosion. These factors indicate that while DYV8S achieves 99.8 % accuracy in controlled settings, future work should incorporate multi-season, multi-region datasets, illumination-invariant and denoising strategies, and domain adaptation or continual learning to ensure reliable performance across diverse agricultural conditions.

5. Conclusion

To effectively characterize soil texture in imagery taken under UFC, this work presented a deep YOLOV8 segmentation model. The study first uses the USDA-trio with YOLOv8 image processing techniques to remove unnecessary information because of the images' intricacy. With Tierce-SegNet, UNet, and DeepLabV3+ segmentation models, the work compares several approaches and identifies the suggested model as the most successful. The predicted mask from the semantic segmentation stage removes undesired pixels from the picture. After the semantic segmentation output is divided into smaller tiles, manifold Gabor filters are used to apply texture-enhancing filters to certain tiles for improved accuracy. To increase the robustness of the data, the work design incorporates four Gabor filters that emphasize different soil patterns in each image. A CNN-based deep YOLOv8-learning model texture classifier is then fed the images, greatly increasing the state-of-the-art architecture's accuracy to 99.8%. The outcomes demonstrated the superiority of the suggested technique in an effective and cost-efficient manner. Future research will use IoT-based smart forming to classify soil textures in real time. The classification accuracy can be increased by using a hybrid artificial intelligence technique and an advanced deep learning-based vision transformer.

Data Availability

The dataset used in this study, including annotated soil texture images and related metadata, is publicly available at the following GitHub repository: https://github.com/phd-latha/latha-soil. This repository contains all the relevant files necessary to reproduce the experiments and results presented in the manuscript. The data has been anonymized and complies with all applicable ethical and legal standards.

References

[1] Han, X.L., Jiang, N.J., Yang, Y.F., Choi, J., Singh, D.N., Beta, P., Du, Y.J. and Wang, Y.J., 2022. Deep learning-based approach for the instance segmentation of clayey soil desiccation cracks. Computers and Geotechnics, 146, p.104733. https://doi.org/10.1016/j.compgeo.2022.104733.

[2] Zamani, V., Taghaddos, H., Gholipour, Y. and Pourreza, H., 2022. Deep semantic segmentation for visual scene understanding of soil types. Automation in Construction, 140, p.104342. https://doi.org/10.1016/j.autcon.2022.104342

[3] Srivastava, P., Shukla, A. and Bansal, A., 2021. A comprehensive review on soil classification using deep learning and computer vision techniques. MultimeFig Tools and Applications, 80(10), pp.14887-14914. https://doi.org/10.1007/s11042-021-10544-5.

[4] Xu, J.J., Zhang, H., Tang, C.S., Cheng, Q., Liu, B. and Shi, B., 2022. Automatic soil desiccation crack recognition using deep learning. Geotechnique, 72(4), pp.337-349. https://doi.org/10.1680/jgeot.20.P.091.

[5] Jiang, Z.D., Owens, P.R., Zhang, C.L., Brye, K.R., Weindorf, D.C., Adhikari, K., Sun, Z.X., Sun, F.J. and Wang, Q.B., 2021. Towards a dynamic soil survey: Identifying and delineating soil horizons in-situ using deep learning. Geoderma, 401, p.115341. https://doi.org/10.1016/j.geoderma.2021.115341.

[6] Rippner, D.A., Raja, P.V., Earles, J.M., Momayyezi, M., Buchko, A., Duong, F.V., Forrestel, E.J., Parkinson, D.Y., Shackel, K.A., Neyhart, J.L. and McElrone, A.J., 2022. A workflow for segmenting soil and plant X-ray computed tomography images with deep learning in Google's Colaboratory. Frontiers in Plant Science, 13, p.893140. https://doi.org/10.3389/fpls.2022.893140.

[7] Pham, T.H., Acharya, P., Bachina, S., Osterloh, K. and Nguyen, K.D., 2024. Deep-learning framework for optimal selection of soil sampling sites. Computers and Electronics in Agriculture, 217, p.108650. https://doi.org/10.1016/j.compag.2024.108650.

[8] Wan, L., Li, S., Chen, Y., He, Z. and Shi, Y., 2022. Application of deep learning in land use classification for soil erosion using remote sensing. Frontiers in Earth Science, 10, p.849531. https://doi.org/10.3389/feart.2022.849531.

[9] Kim, W.S., Lee, D.H., Kim, T., Kim, G., Kim, H., Sim, T. and Kim, Y.J., 2021. One-shot classification-based tilled soil region segmentation for boundary guidance in autonomous tillage. Computers and Electronics in Agriculture, 189, p.106371. https://doi.org/10.1016/j.compag.2021.106371.

[10] Zenkl, R., Timofte, R., Kirchgessner, N., Roth, L., Hund, A., Van Gool, L., Walter, A. and Aasen, H., 2022. Outdoor plant segmentation with deep learning for high-throughput field phenotyping on a diverse wheat dataset. Frontiers in plant science, 12, p.774068. https://doi.org/10.3389/fpls.2021.774068.

[11] Kurtulmuş, E., Arslan, B. and Kurtulmuş, F., 2022. Deep learning for proximal soil sensor development towards smart irrigation. Expert Systems with Applications, 198, p.116812. https://doi.org/10.1016/j.eswa.2022.116812.

[12] Xu, J.J., Zhang, H., Tang, C.S., Cheng, Q., Tian, B.G., Liu, B. and Shi, B., 2022. Automatic soil crack recognition under uneven illumination condition with the application of artificial intelligence. Engineering geology, 296, p.106495. https://doi.org/10.1016/j.enggeo.2021.106495.

[13] Smith, A.G., Han, E., Petersen, J., Olsen, N.A.F., Giese, C., Athmann, M., Dresbøll, D.B. and Thorup-Kristensen, K., 2022. RootPainter: deep learning segmentation of biological images with corrective annotation. New Phytologist, 236(2), pp.774-791. https://doi.org/10.1111/nph.18387.

[14] Ong, P., Teo, K.S. and Sia, C.K., 2023. UAV-based weed detection in Chinese cabbage using deep learning. Smart Agricultural Technology, 4, p.100181. https://doi.org/10.1016/j.atech.2023.100181.

[15] Meng, X., Bao, Y., Wang, Y., Zhang, X. and Liu, H., 2022. An advanced soil organic carbon content prediction model via fused temporal-spatial-spectral (TSS) information based on machine learning and deep learning algorithms. Remote Sensing of Environment, 280, p.113166. https://doi.org/10.1016/j.rse.2022.113166.

[16] Picon, A., San-Emeterio, M.G., Bereciartua-Perez, A., Klukas, C., Eggers, T. and Navarra-Mestre, R., 2022. Deep learning-based segmentation of multiple species of weeds and corn crop using synthetic and real image datasets. Computers and Electronics in Agriculture, 194, p.106719. https://doi.org/10.1016/j.compag.2022.106719

[17] Kang, J., Liu, L., Zhang, F., Shen, C., Wang, N. and Shao, L., 2021. Semantic segmentation model of cotton roots in-situ image based on attention mechanism. Computers and electronics in agriculture, 189, p.106370. https://doi.org/10.1016/j.compag.2021.106370.

[18] Zhang, J., Phoon, K.K., Zhang, D., Huang, H. and Tang, C., 2021. Deep learning-based evaluation of factor of safety with confidence interval for tunnel deformation in spatially variable soil. Journal of Rock Mechanics and Geotechnical Engineering, 13(6), pp.1358-1367. https://doi.org/10.1016/j.jrmge.2021.09.001.

[19] Dong, Y., Xuan, F., Li, Z., Su, W., Guo, H., Huang, X., Li, X. and Huang, J., 2023. Modeling the Corn Residue Coverage after Harvesting and before Sowing in Northeast China by Random Forest and Soil Texture Zoning. Remote Sensing, 15(8), p.2179. https://doi.org/10.3390/rs15082179.

[20] Phalempin, M., Krämer, L., Geers-Lucas, M., Isensee, F., & Schlüter, S. (2025). Deep learning segmentation of soil constituents in 3D X-ray CT images. *Geoderma*, *458*, 117321. https://doi.org/10.1016/j.geoderma.2025.117321.

[21] Roy, S., Ansal, D. K., & Kumar, M. (2025, May). Soil Texture Prediction: Advances in Remote Sensing, Image Analysis, and Machine Learning. In *2025 Fourth International Conference on Smart Technologies, Communication and Robotics (STCR)* (pp. 1-4). IEEE. https://doi.org/10.1109/STCR62650.2025.11020387.

[22] Benchabana, A., Kholladi, M.K., Bensaci, R. and Khaldi, B., 2023. Building detection in high-resolution remote sensing images by enhancing superpixel segmentation and classification using deep learning approaches. Buildings, 13(7), p.1649. https://doi.org/10.3390/buildings13071649

[23] Han, W., Zhang, X., Wang, Y., Wang, L., Huang, X., Li, J., Wang, S., Chen, W., Li, X., Feng, R. and Fan, R., 2023. A survey of machine learning and deep learning in remote sensing of geological environment: Challenges, advances, and opportunities. ISPRS Journal of Photogrammetry and Remote Sensing, 202, pp.87-113. https://doi.org/10.1016/j.isprsjprs.2023.05.032.

[24] Kim, D., Kim, T., Jeon, J. and Son, Y., 2023. Convolutional Neural Network-Based Soil Water Content and Density Prediction Model for Agricultural Land Using Soil Surface Images. Applied Sciences, 13(5), p.2936. https://doi.org/10.3390/app13052936.
[25] He, C., Liu, Y., Wang, D., Liu, S., Yu, L. and Ren, Y., 2023. Automatic extraction of bare soil land from high-resolution remote sensing images based on semantic segmentation with deep learning. Remote Sensing, 15(6), p.1646. https://doi.org/10.3390/rs15061646.
[26] Maruthaiah, T., Vajravelu, S.K., Kaliyaperumal, V. and Kalaivanan, D., 2023. Soil texture identification using LIBS data combined with machine learning algorithm. Optik, 278, p.170691. https://doi.org/10.1016/j.ijleo.2023.170691.
[27] Feng, D., Zhang, Z. and Yan, K., 2022. A semantic segmentation method for remote sensing images based on the Swin transformer fusion Gabor filter. Ieee Access, 10, pp.77432-77451. https://doi.org/10.1109/ACCESS.2022.3193248.
[28] Srivastava, P., Shukla, A. and Bansal, D.A., 2023. Transfer Learning Analysis for Predicting Soil Texture Classes from Soil Images. Available at SSRN 4192498. https://doi.org/10.21203/rs.3.rs-2428396/v1.
[29] Azevedo, R.P., Corinto, L.M., Peixoto, D.S., De Figueiredo, T., Silveira, G.C.D., Peche, P.M., Pio, L.A.S., Pagliari, P.H., Curi, N. and Silva, B.M., 2022. Deep tillage strategies in perennial crop installation: Structural changes in contrasting soil classes. Plants, 11(17), p.2255. https://doi.org/10.3390/plants11172255.
[30] Kurtulmuş, E., Arslan, B., & Kurtulmuş, F. (2022). Deep learning for proximal soil sensor development towards smart irrigation. *Expert systems with applications*, *198*, 116812. https://doi.org/10.1016/j.eswa.2022.116812.