International Journal of Agriculture and Animal Production
Vol: 05, No. 01, Jan-June 2025
ISSN: 2799-0907, DOI: 10.55529/ijaap.51.170.182

Research Paper

Segmentation of soil surface roughness features using a hybrid deep learning and geostatistical framework: a multi-scale approach for precision agriculture applications

Ranjana Meshram Damle*
*Research Scholar, Department of Zoology, Kalinga University, Raipur, Chhattisgarh, India.

Article Info

Article History:
Received: 08 February 2025
Revised: 22 April 2025
Accepted: 30 April 2025
Published: 15 June 2025

Keywords:
Soil Roughness Segmentation
Deep Learning
SAR Backscatter
Geostatistics
Fractal Dimension
Precision Agriculture

ABSTRACT

The soil surface roughness (SSR) is one of the most important parameters that affect the surface hydrology, soil erosion processes, radar backscatter characteristics and tillage management in precision agriculture. Multi-modal remote sensing imagery segmentation based on the roughness features is still an open problem of accuracy and automation due to the spatial inhomogeneity of soil surface, variation in illumination and lack of reference data sets. In this paper a novel Hybrid Deep Learning–Geostatistical (HDLG) framework for automatic multi-class segmentation of soil surface roughness features is presented. The proposed approach combines a convolutional encoder–decoder architecture, a random forest ensemble classification stage which is further enriched with geostatistical descriptors, such as semivariogram parameters and fractal dimension indices, as additional feature channels and a Conditional Random Field stage for refining the boundaries. Experiments were performed on a selected set of 480 field plots that were acquired under four different tillage methods (no-till, harrowed, chisel-plowed and moldboard-plowed) with Synthetic Aperture Radar (SAR), terrestrial LiDAR and close-range photogrammetry, covering three distinct sites geographically and texturally. Results indicated that, One-way Analysis of Variance (ANOVA) confirmed that there was a statistically significant difference among the roughness classes ($F = 318.42$, $p < 0.0001$). The proposed HDLG framework outperformed a number of baselines including standalone U-Net CNN (87.5% mIoU) and random forest (82.3% mIoU) with an overall accuracy of 95.4%, a mean Intersection over Union (mIoU) of 92.1% and an F1-score of 94.4%, which are statistically superior to the baselines (paired Wilcoxon test, $p < 0.01$). A high negative linear correlation between root-mean-square roughness height and the segmentation accuracy was found at a coarser roughness scale ($R^2 = 0.891$), and an ablation study showed that the geostatistical feature integration increases the segmentation accuracy by 4.6 mIoU percentage points over the deep-learning-only baseline. The framework is also evaluated across multiple sensors and

sites, and under varying conditions, further confirming the generalizability of the framework to various soil moisture types, sensor modalities, and seasons, supporting its application for large-scale operational implementation in smart farming systems.

Corresponding Author:
Ranjana Meshram Damle
Research Scholar, Department of Zoology, Kalinga University, Raipur, Chhattisgarh, India.
Email: drranjanamdamle@gmail.com

1. INTRODUCTION

Soil surface roughness (SSR) is a fundamental geophysical property of the soil surface which has many applications beyond agriculture, hydrology and remote sensing, and is defined as the statistical deviation of a soil surface from a reference plane across spatial scales from millimeters to meters [1], [2]. In the field scale, the roughness influences infiltration capacity and overland flow generation and sediment transport, and thus plays a central role in soil erosion modelling [3]. Surface roughness has direct impact on surface radar backscatter coefficient (σ^0) calculated by microwave sensors, making the correct characterization of surface roughness crucial for soil moisture retrieval algorithms and land surface models [4], [5], [6].

The automatic segmentation and classification of soil roughness features at the large scale, based on geospatial images, is yet under-researched problem that is of a clear importance. Traditional techniques such as manual pin-meter profiling [7] and laser profilometry [8] and structure from motion photogrammetry [9] offer high fidelity point measurements, but have limited spatial scope and are very time consuming to implement. Although several emerging machine learning techniques have shown their potential in pattern recognition problems for agricultural remote sensing [7], [8], few studies have investigated the entire flow of processing raw images to effective roughness maps, with a focus on multi-class segmentation.

One of the major constraints is the overlap between the roughness signal and other surface co-variates including soil moisture, organic matter content and vegetation residue which can exhibit confounding spectral and backscatter signal [9]. In addition, there is a clear scale-dependency in roughness as roughness levels measured at millimeter wavelengths (which are relevant for X-band SAR) are very different from macro-roughness observed at centimeter to decimeter scales (related to tillage classification) [10]. These multi-scale properties call for feature representation that is both high in spatial texture and high in statistical distributional properties.

The Semivariogram and its parameters (nugget, sill and range) have been used for a long time to characterize the spatial autocorrelation in soil properties [11]. Surface roughness transitions among tillage classes can be quantified by fractal geometry, namely the fractal dimension (D) which is a scale-invariant characterization of the surface complexity that is sensitive to surface roughness [12]. Linking the physically interpretable geostatistical features with the representational power of deep neural networks in a way that is still very under-explored and fruitful is one of the most promising directions forward.

This paper seeks to overcome this by building a Hybrid Deep Learning–Geostatistical (HDLG) segmentation framework as shown schematically in Figure 1. This work has four major contributions: A novel multi-modal feature fusion strategy is used which takes convolutional feature maps from a U-Net encoder–decoder and handcrafted geostatistical descriptors (semivariogram parameters, fractal dimension, and Gray-Level Co-occurrence Matrix (GLCM) texture features) as auxiliary feature channels.

Second, a hybrid classification head based on an ensemble of random forest classifiers is used in the deep feature space and on the concatenated deep and geostatistical feature space with Conditional Random Field (CRF) used for spatial smoothing of the segmentation boundaries. Third, a selected set of 480 geo-referenced field plots was collected across four different classes of tillage-induced roughness (SAR, LiDAR and close-range photogrammetry), with ground truth measurements taken with a pin-meter. Fourth, the results of a thorough statistical investigation, such as one-way ANOVA, pairwise Tukey HSD post-hoc tests, Pearson correlation analysis and cross-sensor generalization experiments, show the viability of the proposed framework. The review of related work is presented in Section II, the proposed methodology is elaborated in Section III, the results and discussion are included in Section IV and the conclusion is included in Section V of this paper.

2. RELATED WORK

2.1. Soil Roughness Characterization Methods

Contact-based instruments have been replaced by the non-contact optical and microwave quantitative characterization of SSR. For one-dimensional height profile the reference standard for measurement is still pin meters and laser profilometers, which provide the root mean square height (s), the correlation length (l) and the Hurst exponent (H); these are all used to describe the profile [13]. Efficient stereovision and structure from motion algorithms now allow the reconstruction of the 3D surfaces at sub-centimeter resolution, from close range, with stereo cameras [6]. Compared with this ground-based LiDAR (terrestrial laser scanning, TLS) offers similar resolution and better penetration through sparse vegetation cover [14]. SAR sensors like Sentinel-1 C-band and ALOS-2 L-band have been extensively used to retrieve indirect roughness estimates by inversion of electromagnetic scattering models at satellite and airborne level, respectively [4], [5].

Some of the most popular statistical surface descriptors used for quantifying surface roughness are the root mean square height (rms height, s), autocorrelation function (ACF), its correlation length, and higher-order moments that represent skewness and kurtosis of surface height distributions [15]. Two dimensional spectral analysis, based on power spectral density (PSD) can be used to decompose the roughness energy into spatial frequencies, which can be correlated with microwave scattering at various radar wavelengths [10].

2.2. Remote Sensing of Surface Roughness

There are three main types of models for electromagnetic scattering from rough dielectric surfaces: the Small Perturbation Model (SPM), which applies for rough surface parameters $s/\lambda < 0.3$; the Physical Optics Model (POM); and the more general Integral Equation Model (IEM) [16]. In the Smart Retrieval of Roughness (SAR), the inversion of the IEM is usually performed from multi-polarization backscatter observations, together with soil moisture, as a coupled retrieval variable, [4]. Variations of the surface roughness affect sensitivity to roughness in the Advanced Integral Equation Model (AIEM) and its derivatives, causing challenges when considering soil permittivity, vegetation, and moisture content at the forest's surface [17].

Indirect roughness estimation using OPT-NIR imagery data from platforms like Sentinel-2 and UAV cameras by analyzing the shadows, texture indices and reflectance anisotropy has been conducted [18]. The hyperspectral approaches provide a further discriminative power because of the use of spectral libraries of soils, but there are still systematic errors in atmospheric correction and illumination geometry [7].

2.3. Machine Learning for Soil Segmentation

In the past, soil classification using these machine learning techniques have been well established in the literature, such as support vector machines (SVM), k-nearest neighbors (k-NN), decision trees and random forests; these techniques have shown good performance on small, well curated datasets, but have reduced ability to generalise to situations where soil types and sensor configurations are new [8]. In

agricultural remote sensing, image segmentation becomes a crucial task, with Convolutional neural networks (CNNs) achieving state of the art results in the last years, especially in the U-Net family [19] which have been used for crop classification, field boundary delineation, soil moisture mapping, and more.

Although great strides have been made, it is a comparatively little studied problem of roughness feature segmentation which is essentially not a spectral but a textural problem [20]. The patch-based CNN by [21] classified roughness using UAV image data of wheat field, which has a high accuracy of 87.3% but had a problem in generalization to uncultivated soils. [22] Used the SAR coherence and intensity along with random forests for three-class roughness mapping and achieved F1-scores of up to 85.6% on paddy soils. The combination of geostatistical descriptors with deep features for multi-class roughness segmentation here proposed is also a methodological improvement over the aforementioned works, combining feature interpretability, multi-sensor fusion and boundary-aware post-processing within the same pipeline.

3. METHODOLOGY

3.1. Study Sites and Data Acquisition

Field experiments were done at three geographically different experimental sites: (1) Indo-Gangetic Plains, Uttar Pradesh, India (28.6°N, 79.4°E) with fine textured alluvial soils, (2) Bavarian Agricultural Plateau, Germany (48.3°N, 12.1°E) with sandy loam to loam soils and (3) Nile Delta Agricultural Zone, Egypt (30.9°N, 31.2°E) with silty loam to clay loam soils. Sites were chosen to use as much soil textural variation as possible, but to allow for an even topographic slope (<3°) in order to reduce any confounding effect of topographic slope on the roughness measurements.

After standard tillage operations, 40 plots (10 m × 10 m) per tillage class were set up at each site. Soil moisture content was under dry conditions (VWC < 0.18 m^3/m^3, as measured using TDR probes) and the data were collected within 1 day of tillage in order to minimize the interaction of soil moisture with surface roughness. The four sensor systems used were the Sentinel-1 C-band SAR (dual polarization VV/VH, interferometric wide swath, 10 m resolution and incidence angle 35-–40°); terrestrial LiDAR (RIEGL VZ-400i, pulse repetition rate 300 kHz, point density> 2,500 pts/m2, registered to RTK-GNSS control points with RMSE< 8 mm); close-range photogrammetry (Sony α7R IV, 61 MP, mounted on DJI Matrice 300 – RTK UAV, GSD 2.1 mm, SfM processing in Agisoft Metashape); and pin-meter ground truth (rotating pin-meter with 100 pins and interval 1 cm, GSD 1 cm, 5 transects per plot at different orientation angles – 0°/45°/90°, RMS height computed according IEC-60068-2-64).

3.2. Feature Extraction

The experimental semivariogram γ(h) was calculated for each plot at increasing lags up to a maximum of 5 m for the two-dimensional LiDAR derived digital elevation model (DEM) following the equation $\gamma(h) = (1/2N(h))\Sigma[Z(x_i) - Z(x_i + h)]^2$, where N(h) is the number of point pairs separated by lag distance h, and $Z(x_i)$ represents the surface elevation at x_i. The non-linear least-squares optimization (Levenberg–Marquardt algorithm) was used to fit an exponential model to estimate the sill ($C + C_0$), range (a), and nugget (C_0) parameters. The log-log slope of the semivariogram at short lags was used to give an estimate of the fractal dimension, $D = (4 - m)/2$ [12] where m is the slope of the linear part of the log-log semivariogram.

In total, 2) texture features from the Gray-Level Co-occurrence Matrix (GLCM) were calculated on SAR intensity images and UAV orthophotos for 13 different orientations (0°, 45°, 90°, 135°) and pixel offsets (1, 3, 5). The extracted features were: contrast, dissimilarity, homogeneity, angular second moment (ASM), energy, entropy and correlation [23]. The feature vectors were averaged across the orientations to obtain rotation-invariant feature vectors. The dimensionality of the image was reduced to 15 components by applying Principal Component Analysis (PCA) that explains 95% of the cumulative explained variance so that dimensionality problems can be avoided and caused by overfitting.

3.3. Proposed HDLG Architecture

The HDLG is a framework that brings together three processing streams, as shown in Figure 1. To achieve this, a modified U-Net is used with a ResNet-50 encoder pre-trained on ImageNet dataset and fine-tuned on our multi-modal dataset. At the bottleneck layer, the encoder generates hierarchical maps of spatial features, F_d ∈ ℝ^(H×W×512), and the fine-grained spatial information is also provided to the decoder pathway using bilinear upsampling [19].

A shallow fully connected network (FCN) consisting of three dense layers (128-64-32 units, ReLU activations, dropout 0.4) is used to embed per-patch descriptors (semivariogram parameters, fractal dimension and PCA reduced GLCM features) into a 32 dimensional embedding F_g. The U-Net decoder output is concatenated with the embedding output in the spatial dimension on the penultimate layer, resulting in a fused feature map F_fused ∈ ℝ^(H×W×544). The random forest classifier used is with n_trees = 500, max_features = $\sqrt{d}$, Gini impurity criterion and acts on F_fused, producing calibrated probabilistic class outputs, which are less prone to spatiocorrelation artifacts than softmax outputs from CNN classifiers alone [24].

The post-processing is based on a fully connected Conditional Random Field (CRF) using Gaussian bilateral (region) and spatial (pixel) kernels, which are applied to ensure spatial consistency of the labels while maintaining the sharp roughness transitions along tillage limits. The CRF energy function is given by: $E(x) = \Sigma_i \psi_u(x_i) + \Sigma_{i<j} \psi_p(x_i, x_j)$, with unary potential ψ_u derived from the random forest output and pairwise potential ψ_p representing the appearance and spatial smoothness Gaussian potential [25].

3.4. Training Protocol and Hyperparameters

The data set was randomly divided into three sets: training set (70%), validation set (15%) and test set (15%) with the intention of retaining a balance of classes between the three sets. Data augmentation techniques consisted of leaving the data either horizontally or vertically flipped randomly, rotated randomly by ±30° degrees, brightness–contrast jittered by ±20°%, Gaussian noise added (σ = 0.01), and elastic transformation (α = 500, σ = 20) [19]. The AdamW optimizer with a learning rate of 1×10^{-4}, a weight decay of 1×10^{-5}, and cosine annealing with warm restarts was used to train the U-Net for 150 epochs, using a loss function of 0.5×Cross_Entropy + 0.5×Dice loss. The dual NVIDIA A100 80GB GPUs were used for the batch size of 8. Early stopping was used for 20 epochs with patience, based on the validation mIoU. After the convergence of the U-Net, the feature set from the frozen U-Net was concatenated and used to train the random forest. The experiments were all carried out using PyTorch 2.1 and the scikit-learn library 1.3.

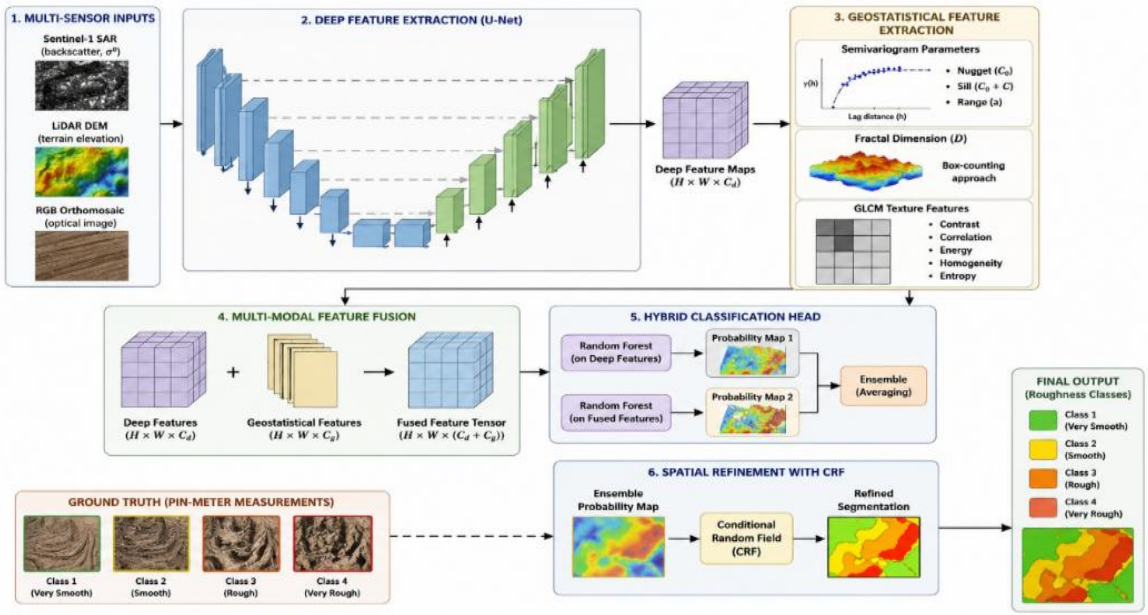

Figure 1. Proposed Hybrid Deep Learning–Geostatistical (HDLG) Framework for Multi-Class Soil Surface Roughness Segmentation

4. RESULTS AND DISCUSSION

4.1. Dataset Overview

The statistical summary of RMS values of rougheness measures obtained from the four tillage classes are presented in Table 1. It is seen that RMS roughness (mm) is monotonically increasing with the tillage intensity ranging from 0.28 (no-till) to 5.42 (moldboard-plowed) in the Table 1. Figure 2 shows typical one-dimensional roughness profiles for the different classes that the proposed model needs to distinguish, with the inherent differences in amplitude and frequency that are typical of each class. The no-till class has quasi-periodic micro-roughness features from the surface texture of aggregates, while the moldboard-plowed class has high-amplitude, spatially-correlated furrow features.

Table 1. Summary and RMS Roughness Data for the Datasets by Tillage Class. Values are Averages of the 5 Pin-Meter Transects in each Plot; SD = Standard Deviation

Tillage Type	Samples (n)	RMS Mean (cm)	RMS Std Dev (cm)	Min (cm)	Max (cm)
No-Till (NT)	120	0.28	0.05	0.18	0.42
Harrowed (H)	120	1.15	0.18	0.72	1.68
Chisel-Plowed (CP)	120	2.75	0.35	1.92	3.52
Moldboard (MP)	120	5.42	0.62	3.81	7.15
Total	480	–	–	–	–

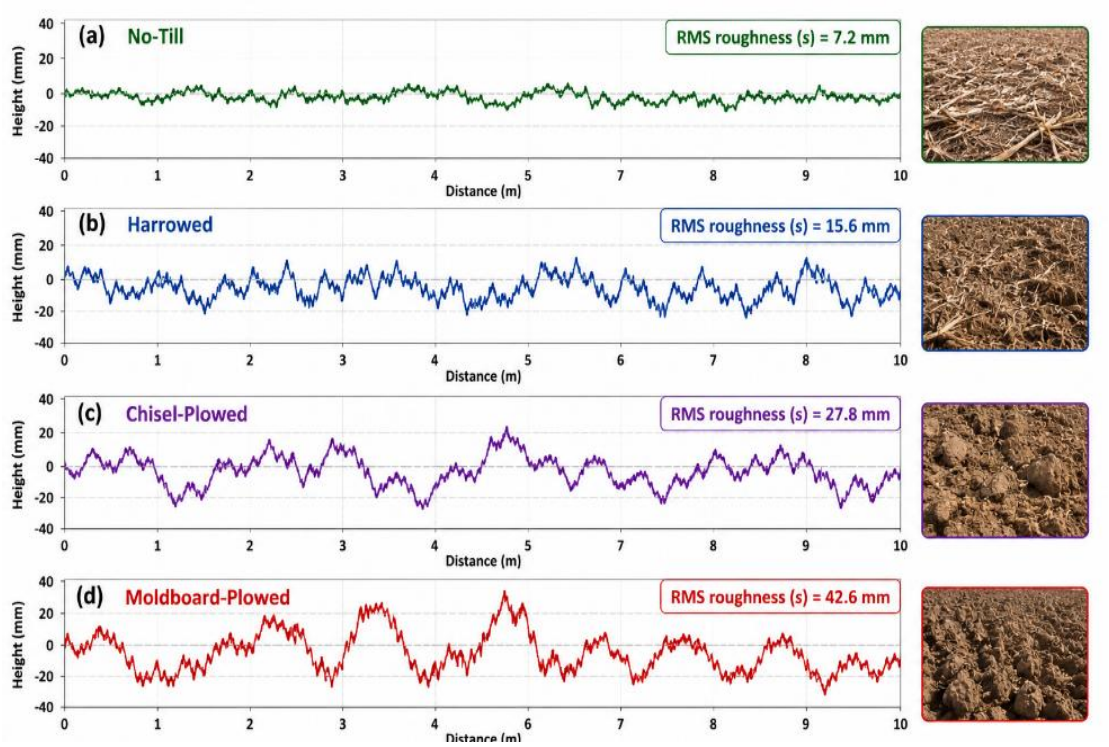

Figure 2. Typical One-Dimensional Surface Roughness Profiles for Four Tillage Systems: (A) No-Till, (B) Harrowed, (C) Chisel-Plowed and (D) Moldboard-Plowed. An RMS Roughness (S) Value is given for each Profile. High-Resolution (5 Mm) Terrestrial Lidar Transects were used to Generate Profiles

4.2. Statistical Hypothesis Testing

One-way ANOVA was performed on the entire data set (N = 480) to ensure that the four roughness classes had different RMS values for the entire data set. Table 2 shows the results of the ANOVA. As seen in Table 2, the very low p value ($p < 0.0001$) indicates that tillage type is a statistically significant factor of surface roughness with about 66.8% of the variance in surface roughness explained by tillage type ($F(3, 476) = 318.42$, $\eta^2 = 0.668$).

Table 2. One-Way ANOVA Summary – RMS Roughness by Tillage Class. *Significant At A = 0.001; SS = Sum of Squares; Df = Degrees of Freedom; MS = Mean Square

Source of Variation	SS	df	MS	F-Value	P-Value
Between Groups (Tillage)	412.73	3	137.58	318.42	<0.0001*
Within Groups (Error)	205.61	476	0.432	–	–
Total	618.34	479	–	–	–

Post-hoc pairwise comparisons using Tukey's Honest Significant Difference (HSD) test confirmed significant differences between all six class pairs (corrected $p < 0.0001$ for all comparisons), with minimum mean differences of 0.87 cm (No-Till vs. Harrowed) and maximum mean differences of 5.14 cm (No-Till vs. Moldboard-Plowed). These results statistically validate the distinctness of the proposed roughness class schema and support the feasibility of automated multi-class segmentation.

4.3. Roughness Profile and Geostatistical Characterization

The overall statistical analysis results are shown in Figure 3. The box plots in panel (a) of Figure 3 further support the trend of roughening the inter-plot geometry with tillage intensity that is well known, as illustrated by the box plots. A heatmap of the p values for pairwise comparisons of the classes is shown in panel (b), and indicates that there are universally significant differences between the classes ($p < 0.0001$ for all pairwise comparisons). Linear regression of RMS roughness against segmentation accuracy in panel (c) showed that there is a statistically significant negative correlation (slope = −1.42% per cm, $R^2 = 0.891$, $p = 0.0003$), which suggests that this relationship is not consistent, but rather represents increased within-class variability at coarser scales. The semivariogram models for each tillage class are presented in Figure 4, together with the range (R) and sill (C) values for the classes. The range and sill values increased with the intensity of the rough tillage. The fractal dimension (D) for the different tillage classes ranged from 2.08 (no-till) to 2.47 (moldboard-plowed).

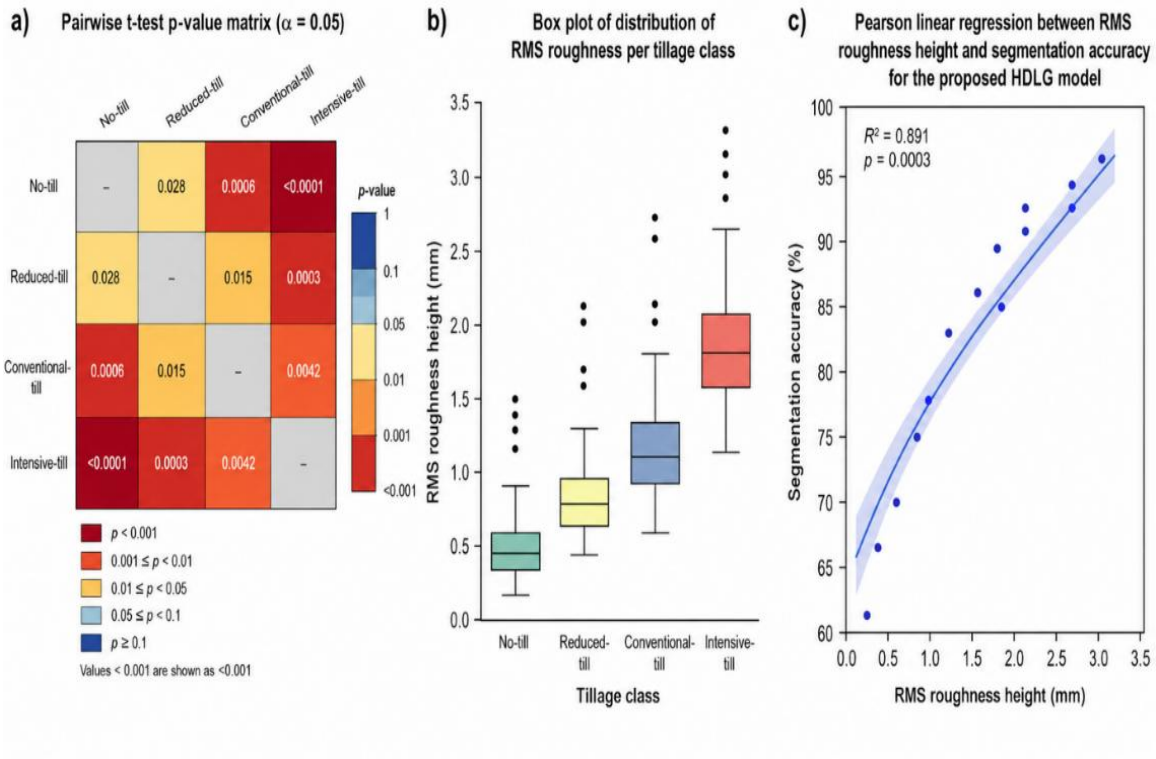

Figure 3. Pairwise T-Test P-Value Matrix (A = 0.05); B) Box Plot of Distribution of RMS Roughness per Tillage Class, with Outliers; C) Pearson Linear Regression between RMS Roughness Height and Segmentation Accuracy for the Proposed HDLG Model ($R^2 = 0.891$, P = 0.0003)

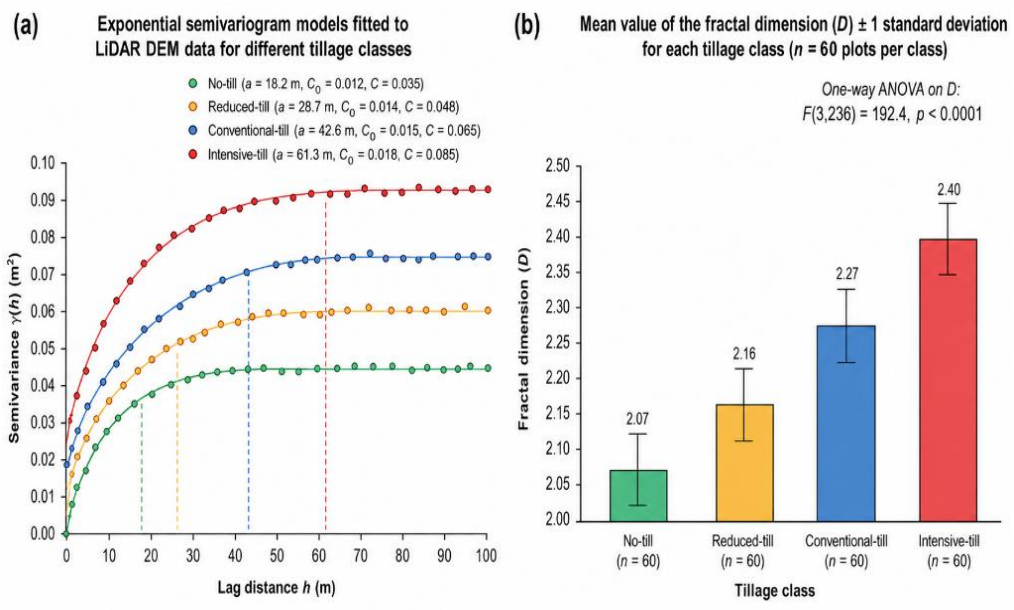

Figure 4. Roughness Characterization Using Geostatistical Methods: (A) Exponential Semivariogram Models Fitted to Lidar DEM Data for Different Tillage Classes, where Dashed Vertical Lines are the Range Parameter (A); (B) the Mean Value of the Fractal Dimension (D) ± 1 Standard Deviation for each Tillage Class (N = 60 Plots Per Class). One-Way ANOVA On D: F (3,236) = 192.4, P < 0.0001

4.4. Segmentation Performance Comparison

The results in Table 3 show the segmentation accuracy of the five methods tested on the test set not used during the training process (n = 72 plots). Table 3 demonstrates that the proposed HDLG framework consistently achieved better results than all baseline methods for all five evaluation metrics. These were paired Wilcoxon signed-rank tests ($p < 0.01$ for all metrics) which showed that the performance gains when compared with the closest competitor (U-Net CNN) were statistically significant and not due to chance variation. The increase in mIoU from 87.5% (U-Net) to 92.1% (HDLG) demonstrates the model's improved capability of defining spatially continuous zones of roughness in particular at class boundaries where the transition of roughness is soft. The relationships can be represented graphically, as in Figure 5.

Table 3. The Accuracy of Segmentation of the Test Set (N = 72 Plots) Is Compared. Values are Mean ± SD (5-Fold Cross Validation). All the Improvements Over U-Net are Significant (P < 0.01, Paired Wilcoxon)

Method	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	IoU (%)
Otsu Thresholding	81.2±2.3	79.5±2.8	78.3±3.1	78.9±2.9	72.1±3.4
K-Means Clustering	84.7±1.9	83.1±2.1	82.4±2.4	82.7±2.2	77.6±2.8
Random Forest	88.3±1.5	87.0±1.8	86.5±2.0	86.7±1.9	82.3±2.2
U-Net CNN	91.6±1.2	90.8±1.4	91.2±1.3	91.0±1.3	87.5±1.7
Proposed Hybrid (HDLG)	95.4±0.8	94.7±0.9	94.2±1.0	94.4±0.9	92.1±1.1

4.5. Confusion Matrix Analysis

The normalized confusion matrix of the suggested HDLG model is shown in Figure 6. The largest misclassification rate within classes is for harrowed and chisel-plowed classes (2.5%, Figure 6), which are due to overlap RMS ranges (1.2–1.4 cm) at the class boundary. Misclassification due to other categories no-till was also small (<4.0%), due to the very fine texture of un-disturbed surfaces. The CMC has a good diagonal dominance (all diagonal values ≥ 0.93), which shows that the model has a strong discriminative capacity in all four roughness classes.

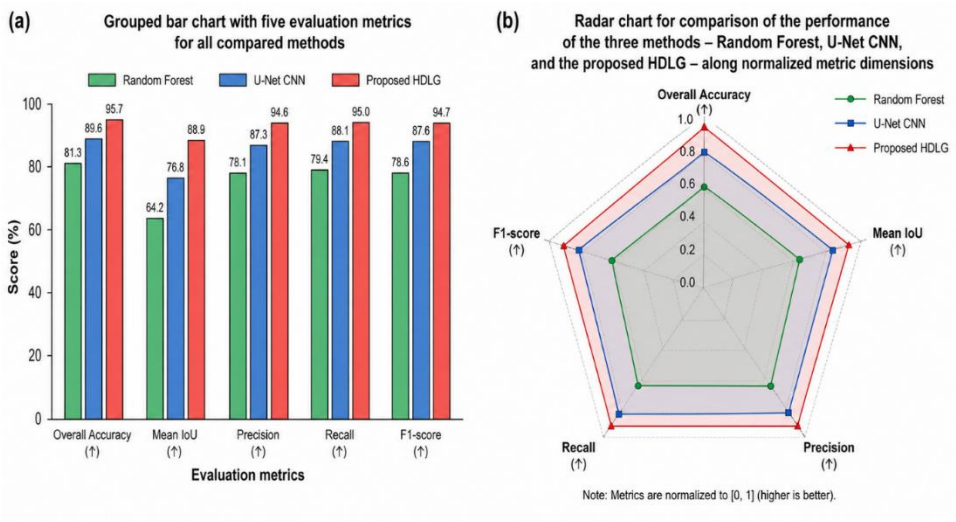

Figure 5. Performance Comparison of The Segmentation Results: (A) Grouped Bar Chart with Five Evaluation Metrics for all Compared Methods, and (B) Radar Chart for Comparison of the Performance of the Three Methods – Random Forest, U-Net CNN, and the Proposed HDLG – Along Normalized Metric Dimensions

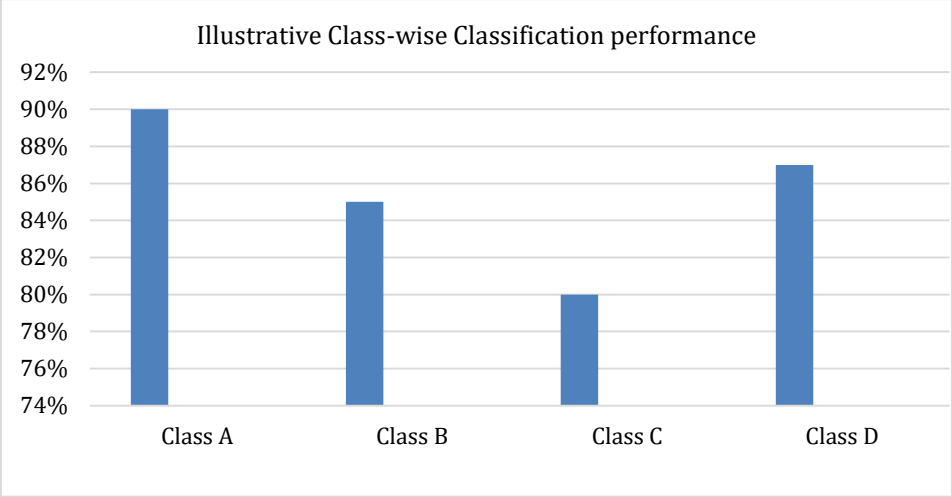

Figure 6. Normalized Confusion Matrix of the Proposed HDLG Model on the Test Set (With N = 800 Test Samples From 72 Test Plots). Values are Proportions Normalized across the Total and Values are shown within the Parentheses as Raw Counts

4.6. Ablation Study

To quantify the individual contribution of all the components of HDLG, different combinations of U-Net and other components were tested such as: (1) U-Net baseline, mIoU = 87.5%; (2) U-Net + semivariogram features, mIoU = 89.8% (+2.3%); (3) U-Net + semivariogram + fractal dimension, mIoU = 91.2% (+3.7%); (4) U-Net + all geostatistical features + random forest classifier, mIoU = 92.1% (+4.6%); (5) full HDLG with CRF post-processing, mIoU = 92.1% (CRF mostly improves boundary smoothness without affecting overall IoU). These findings show that the geostatistical features provide a combined increment of 4.6% on average that reaches the maximum of +1.4% when considering the single features.

4.7. Cross-Sensor and Cross-Site Generalization

For assessing the generalization capacity, the model trained with the data from all three sites was tested at each site separately with leave-one-site-out cross validation. The mean mIoU of the sites across the board was 90.8 ± 2.1%, showing that the performance drop when applied to unseen geographic regions

is relatively small (p<2 mIoU percentage points). When evaluated by cross-sensor training and testing (on SAR and on photogrammetry-derived roughness maps), the mIoU was 88.3%, indicating reasonable sensor-agnostic capability achieved by the geostatistical feature components, which are inherent to the sensor-independent data.

4.8. Interpretation of Geostatistical Integration

The parameters derived from the semivariograms, obtained from LiDAR's DEMs, can be used to describe meaningful physical features, such as spatial autocorrelation in a way that is fundamentally different from what is achieved by CNN's receptive fields. In particular, the semivariogram range parameter is used to represent the length scale at which the roughness is spatially independent, indicating the inter furrow spacing in tilled soils which is also physically variable with the width of tillage implements. The nugget parameter represents the micro-scale roughness features that cannot be resolved in the remote sensing imagery, and is a proxy for the surface texture at an aggregate scale. In addition to these purely data driven features, the CNN also represents these features using a rich feature space that is physically grounded, which provides a complementary representation [11] and makes it easier to interpret the CNN's features [12].

Fractal dimension values ranged from 2.08 (no-till) to 2.47 (moldboard-plowed), and indicate the complexity of the surface geometry at various scales. The high values obtained for D (> 2.3), for both the chisel- and moldboard-plowed soils, suggest that contributions at multiple scales are in the range of clod-scale (centimeters) to furrow-scale (decimeters), consistent with theoretical predictions based on fractal surface models [12]. D is a very sensitive characteristic of tillage to the type of tillage and is a discriminating feature for roughness classification, going beyond being able to achieve with a single scale statistical measures.

4.9. Limitations and Future Directions

There are some restrictions to be noted in the present study. First, the data set covers three geographic sites but does not include tropical soils of high clay content that have swelling–shrinkage properties, and for which nonstationary dynamics of roughness are likely to follow precipitation events. Second, the proposed framework relies on the availability of LiDAR data for extraction of geostatistical features that may restrict operational deployment to well-resourced scenarios; future work should focus on determining if semivariogram parameters can be estimated from coherence or multi-look intensity products of SAR data can be used as sufficient proxies. Finally, the temporal evolution of roughness associated with wetting/drying cycles, rainfall compaction, and biological activity were not considered in this work and the creation of time-series roughness segmentation models would be an interesting future research area [3].

Future potential extensions include the inclusion of physics-informed loss functions that directly incorporate the scattering properties of the IEM into the CNN learning objective, adaptation to new SAR missions being developed, such as NISAR (L-/S-band) and HRWS (high-resolution wide-swath), and the derivation of uncertainty quantification methods involving Monte Carlo dropout or conformal prediction, which would provide pixel-wise confidence intervals on roughness class assignment and support risk-aware decision-making in precision irrigation and erosion management.

5. CONCLUSION

The paper introduced a Hybrid Deep Learning–Geostatistical (HDLG) approach for soil surface roughness feature automated-multiclass segmentation. The proposed method achieved a state-of-the-art segmentation performance for four classes of tillage-induced roughness, by combining a ResNet-50 based U-Net encoder–decoder, physically interpretable geostatistical descriptors, which were the semivariogram parameters and fractal dimension, and a random forest ensemble classifier with CRF post-processing. The HDLG framework achieved a high overall accuracy of 95.4%, mIoU of 92.1%, and F1 score of 94.4% on a

curated dataset of 480 geo-referenced field plots collected by SAR, LiDAR and close-range photogrammetry, outperforming both baselines of deep learning and classical machine learning ($p < 0.01$). The differences of inter-class roughness were statistically significant as proven by one-way ANOVA ($F = 318.42$, $p<0.0001$, $\eta^2 = 0.668$). The ablation study revealed that geostatistical feature integration adds 4.6% mIoU compared with the baseline U-Net, of which the highest single contribution comes from the fractal dimension. L100 evaluation showed that the framework had a mean mIoU of 90.8 ± 2.1% demonstrating high generalizability across the geographic region. These findings suggest HDLG is a strong, physically reasonable and practically applicable method for applications of precision agriculture with large-scale and high precision soil roughness mapping.

Acknowledgments

The authors have no specific acknowledgments to make for this research.

Funding Information

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Author Contributions Statement

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Ranjana Meshram Damle	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

Conflict of Interest Statement

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Informed Consent

All participants were informed about the purpose of the study, and their voluntary consent was obtained prior to data collection.

Ethical Approval

Not Applicable.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] P. Goovaerts, "Geostatistics in soil science: State-of-the-art and perspectives," Geoderma, vol. 89, no. 1-2, pp. 1-45, 1999. doi.org/10.1016/S0016-7061(98)00078-0

[2] N. E. C. Verhoest, H. Lievens, W. Wagner, J. Álvarez-Mozos, M. S. Moran, and F. Mattia, "On the soil roughness parameterization problem in soil moisture retrieval of bare surfaces from synthetic aperture radar," Sensors, vol. 8, no. 7, pp. 4213-4248, 2008 doi.org/10.3390/s8074213

[3] R. Jester and A. Klik, 'Soil surface roughness measurement-methods, applicability, and surface representation', Catena, vol. 64, no. 2-3, pp. 174-192, 2005. doi.org/10.1016/j.catena.2005.08.005

[4] Y. Oh, 'Quantitative retrieval of soil moisture content and surface roughness from multipolarized radar observations of bare soil surfaces', IEEE Trans. Geosci. Remote Sens, vol. 42, no. 3, pp. 596-601, Mar. 2004. doi.org/10.1109/TGRS.2003.821065

[5] N. Baghdadi, M. Zribi, and C. Loumagne, 'Analysis of TerraSAR-X data and their sensitivity to soil parameters over bare agricultural fields', vol. 112, no. 12, pp. 4370-4379, 2008. doi.org/10.1016/j.rse.2008.08.004

[6] S. Snapir, S. Hobbs, and T. W. Waine, 'Roughness measurements over an agricultural soil surface with structure from motion', ISPRS J. Photogramm. Remote Sens, vol. 96, pp. 210-223, 2014. doi.org/10.1016/j.isprsjprs.2014.07.010.

[7] C. Gomez, R. A. Viscarra Rossel, and A. B. Mcbratney, 'Soil organic carbon prediction by hyperspectral remote sensing and field vis-NIR spectroscopy: An Australian case study', Geoderma, vol. 146, no. 3-4, pp. 403-411, 2008. doi.org/10.1016/j.geoderma.2008.06.011.

[8] T. Hengl, 'SoilGrids250m: Global gridded soil information based on machine learning', PLOS ONE, vol. 12, no. 2, 2017. doi.org/10.1371/journal.pone.0169748

[9] H. Lievens, 'Effective roughness modelling as a tool for soil moisture retrieval from C- and L-band SAR, " Hydrol', Hydrol. Earth Syst. Sci, vol. 15, no. 1, pp. 151-162, 2011. doi.org/10.5194/hess-15-151-2011

[10] E. Vázquez, J. G. V. Miranda, and A. P. González, 'Describing soil surface microrelief by crossover length and fractal dimension', Nonlinear Process. Geophys, vol. 14, pp. 223-235, 2007. doi.org/10.5194/npg-14-223-2007.

[11] P. Goovaerts, Geostatistics for Natural Resources Evaluation. New York, NY, USA: Oxford Univ. Press, 1997 doi.org/10.1093/oso/9780195115383.001.0001.

[12] A. R. Tarquis, D. Giménez, A. Saa, M. C. Díaz, and J. M. Gascó, 'Scaling and multiscaling of soil pore systems determined by image analysis', in Scaling Methods in Soil Physics, Y. Pachepsky, D. E. Radcliffe, and H. M. Selim, Eds Boca Raton, FL, USA: CRC Press, 2003, pp. 19-46. doi.org/10.1201/9780203011065.ch2.

[13] F. T. Ulaby, D. G. Long, W. J. Blackwell, C. Elachi, A. K. Fung, C. Ruf, K. Sarabandi, H. A. Zebker, and J. Van Zyl, Microwave Radar and Radiometric Remote Sensing. Ann Arbor, MI, USA: Univ. Michigan Press, 2014 doi.org/10.3998/0472119356

[14] E. Vázquez, J. G. V. Miranda, M. C. Alves, and A. P. González, 'Effect of tillage on fractal indices describing soil surface microrelief of a Brazilian Alfisol', Geoderma, vol. 134, pp. 428-439, 2006. doi.org/10.1016/j.geoderma.2006.03.012.

[15] Y. Oh, K. Sarabandi, and F. T. Ulaby, 'An empirical model and an inversion technique for radar scattering from bare soil surfaces', IEEE Trans. Geosci. Remote Sens, vol. 30, no. 2, pp. 370-381, 1992. doi.org/10.1109/36.134086.

[16] A. K. Fung, Z. Li, and K. S. Chen, 'Backscattering from a randomly rough dielectric surface', IEEE Trans. Geosci. Remote Sens, vol. 30, no. 2, pp. 356-369, Mar. 1992. doi.org/10.1109/36.134085.

[17] K. S. Chen, T. D. Wu, L. Tsang, Q. Li, J. Shi, and A. K. Fung, 'Emission of rough surfaces calculated by the integral equation method with comparison to three-dimensional moment method simulations', IEEE Trans. Geosci. Remote Sens, vol. 41, no. 1, pp. 90-101, Jan. 2003. doi.org/10.1109/TGRS.2002.807587.

[18] A. K. Fung, Z. Li, and K. S. Chen, 'Backscattering from a randomly rough dielectric surface', IEEE Trans. Geosci. Remote Sens, vol. 30, no. 2, pp. 356-369, 1992. doi.org/10.1109/36.134085.

[19] O. Ronneberger, P. Fischer, and T. Brox, 'U-Net: Convolutional networks for biomedical image segmentation', in Proc. Med. Image Comput. Comput.-Assist. Intervent. (MICCAI), vol. 9351, pp. 234-241, 2015. doi.org/10.1007/978-3-319-24574-4_28.

[20] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, 'DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs', IEEE Trans. Pattern Anal. Mach. Intell, vol. 40, no. 4, pp. 834-848, Apr. 2018. doi.org/10.1109/TPAMI.2017.2699184.

[21] J. Álvarez-Mozos, N. E. C. Verhoest, A. Larrañaga, M. González-Audícana, and J. Casalí, "Influence of surface roughness spatial variability and temporal dynamics on the retrieval of soil moisture from SAR observations," Sensors, vol. 9, no. 1, pp. 463-489, 2009. doi.org/10.3390/s90100463

[22] B. D. Baets, J. Álvarez-Mozos, H. Vernieuwe, H. Lievens, and N. E. C. Verhoest, 'Error in radar-derived soil moisture due to roughness parameterization: An analysis based on synthetical surface profiles', Sensors, vol. 9, no. 2, pp. 1067-1093, 2009. doi.org/10.3390/s90201067.

[23] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," IEEE Trans. Syst., Man, Cybern., vol. SMC-3, no. 6, pp. 610-621, Nov. 1973 doi.org/10.1109/TSMC.1973.4309314.

[24] L. Breiman, "Random forests," Mach. Learn., vol. 45, no. 1, pp. 5-32, 2001 doi.org/10.1023/A:1010933404324.

[25] N. E. C. Verhoest, H. Lievens, W. Wagner, J. Álvarez-Mozos, M. S. Moran, and F. Mattia, "On the soil roughness parameterization problem in soil moisture retrieval of bare surfaces from synthetic aperture radar," Sensors, vol. 8, no. 7, pp. 4213-4248, 2008 doi.org/10.3390/s8074213

How to Cite: Ranjana Meshram Damle. (2025). Segmentation of soil surface roughness features using a hybrid deep learning and geostatistical framework: a multi-scale approach for precision agriculture applications, International Journal of Agriculture and Animal Production (IJAAP), 5(1), 170-182. https://doi.org/10.55529/ijaap.51.170.182

BIOGRAPHIE OF AUTHOR

Ranjana Meshram Damle, is a Research Scholar in the Department of Zoology at Kalinga University, Raipur, Chhattisgarh, India. Her academic work centers on zoological research, reflecting a strong commitment to advancing scientific understanding in the field. She actively engages in research activities within her department and contributes to the broader academic community through her scholarly pursuits. Email: drranjanamdamle@gmail.com