

CORSIA : Core Regression for Soil Image Analysis

Joanne Boulos^{1,3}[0009–0006–1054–1996], Véronique Eglin¹[0000–0001–8738–2088],
Bertrand Kerautret²[0000–0001–8418–2558], Edith Larue³[0000–0002–1832–8683],
and Jean-Marie Côme³

¹ INSA Lyon, CNRS, Université Claude Bernard Lyon 1, LIRIS, UMR5205, 69621
Villeurbanne, France

² Université Lumière Lyon 2, CNRS, INSA Lyon, Université Claude Bernard Lyon 1,
LIRIS, UMR5205, 69007 Lyon, France

³ Ginger BURGEAP, R&D Departement, 69425 Lyon, France

Abstract. Predicting continuous granulometric distributions from soil core images remains a challenging regression problem due to the stochastic and texture-dominated nature of the data. We propose a lightweight CNN-based framework that directly infers continuous granulometric distributions without explicit particle segmentation. The method combines hierarchical prediction with a texture-aware loss to address structured regression from complex visual patterns and multi-scale texture organization. The key originality of our approach lies in its ability to recover a continuous distribution from textures that are intrinsically stochastic yet statistically homogeneous, thereby going beyond traditional particle-level analysis. The model is trained on representative soil samples and is designed to improve robustness across heterogeneous soil textures and varying acquisition conditions while maintaining generalization performance across datasets consistently. Experimental results show that the proposed method consistently outperforms existing approaches in terms of accuracy, stability, and robustness, while remaining computationally efficient, making it suitable for scalable and near real-time soil core analysis workflows.

Keywords: Soil Core Image Analysis · Petrophysical Properties · Convolutional Neural Networks · Regression Model · Particle Size Distribution Curve.

1 Introduction

Image-based analysis of soil cores (cylindrical soil samples) plays a critical role in geotechnical investigations, providing essential information on subsurface properties for engineering and environmental assessment. Conventional methods rely on extracting soil cores and sending them to the laboratory for analysis, a process that is invasive, time-consuming, and prone to uncertainty due to mechanical disturbances during extraction and transport [1]. These limitations coupled with the rapid advancement of imaging and image processing techniques, motivate the

development of image-based approaches capable of delivering substantial time savings while being readily deployable under field conditions.

However, inferring soil type or particle size distribution from images is not straightforward. For fine-grained soils such as silts and clays, individual particles are often smaller than the image resolution and images mainly exhibit textures rather than distinct objects. The task is therefore to infer particle distributions from texture patterns, which remains challenging due to the non-uniqueness of visually similar textures and the additional variability introduced by minerals, organic matter and water content. Traditional approaches rely on segmentation or classification [6], but segmentation becomes unreliable when individual particles are not fully visible, while classification fails to capture the continuous nature of soil properties. In practice, granulometry is a continuous quantity described by proportions of particle sizes, which must remain non-negative and sum to one to ensure physical consistency. Recent works have explored computer vision [11,7] and machine learning [15,8] techniques for soil image analysis. However, many approaches rely on computationally expensive deep learning models [12], limiting their applicability in real-time applications. In practical scenarios, such as on-site geotechnical investigations, models need to run on resource-constrained edge devices like smartphones or tablets and provide fast and reliable predictions. This requires lightweight models with low computational cost and fast inference.

In this paper, we address soil granulometry as a regression problem where continuous particle size distributions are inferred directly from image textures. We propose a lightweight CNN–MLP architecture specifically designed for texture-to-distribution regression. Unlike standard CNN-based regression pipelines, the model operates under limited spatial resolution and relies on intermediate feature statistics rather than explicit object representations. A convolutional network is used to extract compact texture features from the images, while a multi-layer perceptron maps these features to granulometric distributions. To address the lack of explicit particle representation, we include a texture-aware loss based on Gram-matrix regularization [14], which captures second-order feature correlations and helps preserve statistical texture patterns relevant to the underlying particle distributions.

The main contributions of this work are: (i) the formulation of soil granulometry as a texture-based regression problem, (ii) the design of a lightweight model suitable for real-time deployment on mobile devices, and (iii) the introduction of a texture-aware regularization strategy to improve prediction robustness.

2 Related Work

Early research in optical granulometry focused on particle detection and segmentation, assuming that individual grains are visible in images. Classical segmentation-based approaches [17] present good performance for coarse soils such as sands and gravels, but they fail when particle size approaches or falls below image resolution, which is typical for fine-grained soils such as silts and clays.

In these cases, particle boundaries disappear and segmentation errors lead to unreliable granulometric estimates [9].

To overcome these limitations, recent works have shifted toward regression-based approaches that directly predict soil properties such as particle size distributions (PSD) from raw images. These methods rely on the observation that granulometric information is still encoded in texture statistics, even when individual particles are not visible. CNN-based regression models have therefore become the dominant approach, enabling multi-scale feature extraction without explicit segmentation. A representative contribution is Soranzo et al. [16], who demonstrated that CNNs trained on smartphone-acquired soil images can predict full PSD curves from laboratory measurements, showing the feasibility of end-to-end regression while avoiding segmentation. However, their analysis excludes purely clayey soils, and interpolation-based reconstruction of PSD curves remains less reliable in the distribution tails. Similarly, St-Cyr et al. [13] introduced a large-scale dataset with a hierarchical CNN framework for PSD estimation, improving robustness to heterogeneity but under controlled acquisition conditions, limiting broader generalization. Granitzer et al. [4] further explored deep architectures for soil classification, including vision transformers, highlighting the relevance of modern vision models, although their task was limited to coarse categorical labels rather than continuous PSD regression. Other recent works, including PSDNet [10], confirm the growing interest in learning-based PSD estimation from images, particularly in smartphone-based or field-deployable setups. However, most existing methods remain dependent on relatively controlled imaging conditions and high-quality data, limiting their robustness in real-world scenarios.

Overall, the literature reveals a clear transition from segmentation-based pipelines to regression-based methods. Despite these advances, existing approaches remain limited in their ability to generalize among different soil textures and imaging conditions. Most existing regression models rely on relatively high quality images or are trained under controlled conditions, which limits their robustness in realistic deployment scenario. The present work builds on these insights by further investigating regression-based approaches under varying imaging resolutions, with a focus on predictive robustness and on characterizing their applicability to fine-grained soils.

3 Methodology

3.1 Model

The primary objective of this work is to predict continuous soil particle size distributions directly from soil core images, using texture information without the requirement of individual particle identification. Fine-grained soils such as silts and clays present visual patterns rather than discrete objects, making direct particle segmentation unreliable. Predicting granulometric distributions from such images is inherently challenging: the model must infer structured outputs that satisfy physical constraints, including non-negativity, ordering of

size fractions, and mass conservation, while capturing subtle differences between visually similar textures. Standard CNN regression models often struggle in this scenario because they are not designed to leverage hierarchical relationships between macro- and micro-scale features. To achieve this, our approach focuses on three key principles:

1. Local feature extraction with CNNs: Convolutional neural networks are used to capture sub-resolution textural patterns (at scales below the pixel resolution), which encode information about fine-grained soil particles that are not visually separable.
2. Feature aggregation and final prediction: Local features are aggregated and processed via a multi-layer perceptron (MLP) to estimate the proportions of different soil sub-classes, effectively translating texture information into continuous granulometric distributions.
3. Combined loss function: A hybrid loss is employed, combining a global L1 term with a texture-aware, feature-based regularization term. This design enhances both overall accuracy and local fidelity in the predicted distributions.

Building upon these principles, we propose CORSIA, a lightweight CNN–MLP framework specifically tailored for soil granulometry (Figure 1). The model combines local feature extraction and hierarchical regression to bridge the gap between macro-level proportions and fine-grained subclass distributions, enabling structured prediction of continuous granulometric distributions from texture-dominated inputs. A three-layer CNN extracts compact, high-level texture features from the input images, sufficient to capture local particle aggregates while remaining computationally efficient for real-time deployment. These features are first used to predict macro-level proportions (4D vector) for Clay (A), Silt (L), Sand (S) and Gravel (G), trained with a combination of L1 and texture-based loss. Macro-level predictions provide a contextual prior that guides subclass estimation, which is performed by concatenating the CNN features with macro probabilities and passing them to a final MLP that outputs the proportions of ten subclasses. Freezing the CNN parameters during subclass training stabilizes feature extraction and prevents overfitting, ensuring that subclass predictions remain consistent with macro-level structure. This hierarchical design is motivated by the observation that coarse-grained soil composition provides a strong prior that constrains finer-scale subclass distributions, improving stability and reducing prediction ambiguity. Each soil sample is interpreted according to domain-specific standards (Figure 2). In this study, we focus on the USDA agronomy standard, where Clay (A) and Silt (L) are each subdivided into two subclasses, Sand (S) into five subclasses and Gravel (G) into one subclass, covering particle sizes from 0 to 4 mm. This results in ten regression outputs per sample, enabling the reconstruction of a detailed granulometric curve from each soil image. By combining hierarchical macro-to-subclass learning with lightweight texture-aware CNN features, CORSIA addresses the challenge of predicting continuous distributions from stochastic soil textures while remaining suitable for real-time, field-deployable applications.

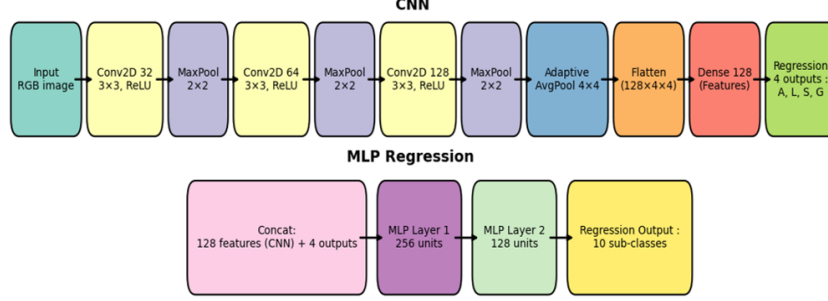


Fig. 1: Proposed model architecture

Agency	USDA (2014)	Clay				Silt				Sand				Gravel
		fine		coarse		fine		coarse		very fine	fine	medium	coarse	fine
	mm	0.0002		0.002		0.002		0.05		0.1	0.25	0.5	1	2
Geotechnics	Doughie (1989)	Clay				Silt				Sand				Gravlier
						fine		coarse		very fine	fine	medium	coarse	fine
	mm					0.004	0.008	0.016	0.031	0.062	0.125	0.25	0.5	1
														2
														4
Sedimentology	Bati et Pye (2012)	Clay				Silt				Sand				Gravlier
		very fine	fine	medium	coarse	very fine	fine	medium	coarse	very coarse	very fine	fine	medium	coarse
	mm	0.00012	0.00025	0.0005	0.001	0.002	0.004	0.008	0.016	0.031	0.062	0.1	0.25	0.5
														1
														2
														4

Fig. 2: Granulometric class divisions (0–2 mm) according to selected standards per application [5]

3.2 Texture Loss Function

To ensure that the CNN model captures both global image fidelity and fine-grained structural cues, we introduce a texture loss function that combines a standard pixel-level reconstruction term with a texture-based regularization component. Classical reconstruction losses, such as the L1 loss, are effective for enforcing overall structural accuracy by penalizing pixel-wise deviations between the predicted output and the ground-truth image. However, they do not capture repetitive patterns, localized variations, and high-frequency structures characteristic of natural textures, often leading to oversmoothed outputs.

Our formulation is inspired by the texture loss proposed by Dong et al. [3], which highlights the importance of inter-channel correlations for capturing fine-grained texture patterns. These correlations are computed via the Gram matrix of intermediate CNN feature maps, which captures texture-related statistical dependencies rather than absolute feature activations. Unlike approaches that explicitly match Gram matrices between predicted and reference features, our formulation does not enforce direct consistency with ground-truth statistics. Instead, it introduces a soft constraint on second-order feature interactions, preventing correlations from becoming excessively large while still allowing meaningful co-activation patterns to emerge. Rather than enforcing explicit alignment with a reference Gram matrix, the loss acts as a soft constraint on inter-channel relationships, encouraging stable and non-dominant correlation structures.

Building on these insights, we design a combined loss function for the proposed regression model. The network predicts a vector of macro-level proportions (A, L, S, G) for each soil sample. We apply a weighted L1 loss directly on the predicted proportion vectors, which is robust to outliers and promotes accurate estimation of class proportions.

$$\mathcal{L}_{L1}(\mathbf{p}, \mathbf{y}) = \frac{1}{C} \sum_{c=1}^C w_c |p_c - y_c|; \quad (1)$$

where:

- $C = 4$ is the number of macro classes,
- p_c is the predicted proportion for class c ,
- y_c is the corresponding ground-truth proportion,
- w_c is a class-specific weight.

To complement global accuracy, we include a texture regularization term applied to intermediate CNN feature maps $\mathbf{F}_{\text{CNN}} \in \mathbb{R}^{B \times C}$.

$$\mathcal{R}_{\text{texture}}(\mathbf{F}_{\text{CNN}}) = \frac{1}{B C^2} \sum_{b=1}^B \|G_b\|_F^2, \quad G_b = \frac{F_b F_b^\top}{H \cdot W}; \quad (2)$$

where:

- B is the batch size,
- C is the number of feature channels,
- $F_b \in \mathbb{R}^{C \times HW}$ is the flattened feature map of sample b ,
- G_b is the Gram matrix of F_b , computed as $G_b = \frac{F_b F_b^\top}{H \cdot W}$,
- $\|\cdot\|_F$ denotes the Frobenius norm.

The Gram matrix encodes inter-channel correlations, providing a compact descriptor of local texture patterns. In our formulation, the Frobenius norm acts as a regularizer on the magnitude of these correlations. Large unbalanced correlations indicate unstable texture patterns. By penalizing excessively large or unstable co-activations between feature channels, this term discourages overly dominant texture patterns while maintaining a balanced representation of feature interactions. As a result, the model is guided to learn stable and well-distributed second-order statistics.

The final combined loss is then expressed as:

$$\mathcal{L} = \alpha \cdot \mathcal{L}_{L1} + \beta \cdot \mathcal{R}_{\text{texture}}; \quad (3)$$

where α and β balance the relative contributions of $\mathcal{L}_{L1}$ and texture regularization. By combining these two terms, the loss balances two complementary objectives: maintaining global structural accuracy via $\mathcal{L}_{L1}$ and constraining local texture statistics via $\mathcal{R}_{\text{texture}}$ to avoid extreme correlations. The weighting coefficients α and β balance global reconstruction quality and texture regularization. For the second component of our model, the MLP, we used the L1 loss, as it promotes stable and accurate estimation at fine-grained levels. Ablation studies in Section 4.2 further quantify the contribution of each component.

4 Experiments

After presenting the methodology and model architecture, this section outlines the dataset, experimental setup, and comparative evaluations, including ablation studies.

4.1 Dataset

A total of 240 samples were prepared to represent all lithological classes in the triangle shown in Figure 3, illustrating the non-unique mapping between particle composition and texture classes according to the United States Department of Agriculture (USDA) [2]. The samples were obtained following a well-defined protocol, using particle sizes ranging from 0 to 4 mm. Coarser particles, such as gravel up to 4 mm, were incorporated to extend the range of properties. Handling pure clay in the laboratory was not feasible, so the clay content in the samples was limited to 15%, which explains why the samples are mainly distributed toward the silt-sand region of the triangle in Figure 4, which illustrate the texture classification of our samples. All soils used in the sample preparation were sent for laboratory analysis.

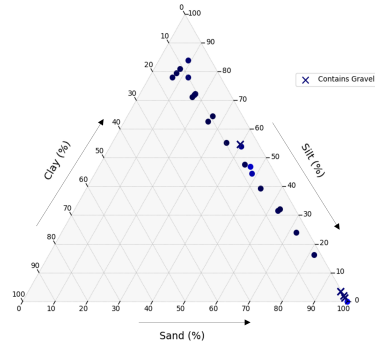
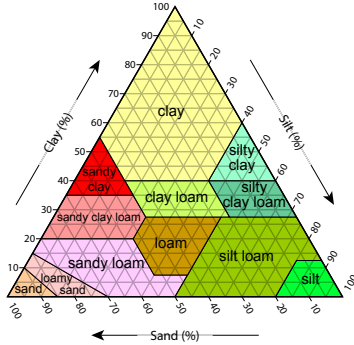


Fig. 3: Texture triangle from USDA [2]. Fig. 4: Texture classification triangle.

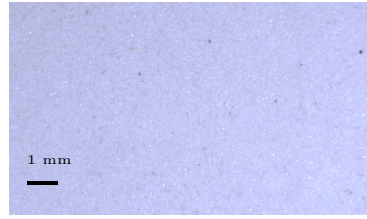
Although the number of physical samples is limited, extensive imaging under controlled variations (granulometry and water content) ensures sufficient diversity for model training and validation. To prevent data leakage, the dataset is split at the physical sample level, so that all images from a given soil sample are strictly assigned to either the training or the test set. These samples were imaged using an optical system, with a camera of a spatial calibration of 2.6 $\mu\text{m}/\text{pixel}$. All images are RGB images of size 500×500 pixels, this size being determined based on a representative elementary volume study. This acquisition process generated a total of 257 428 training images and 64 484 test images. Representative examples of these acquisitions are shown in Fig 5. The large number

of images ensures a robust dataset that can serve as a calibrated reference for the proposed work.

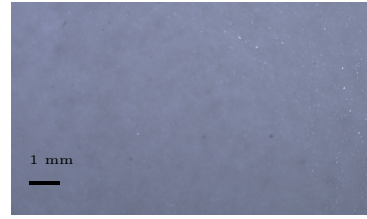
The textures visible in these images are primarily determined by the intrinsic properties of the soils. Clay, silt and sand exhibit distinct visual characteristics, ranging from smooth and cohesive appearances to intermediate and coarse granular patterns, reflecting differences in particle size and organization. Water slightly modifies these visual characteristics by creating a thin layer that diffuses light, enhancing surface smoothness and slightly masking very fine details. However, the overall texture remains dominated by the soil composition itself. These variations are critical: they influence the visual signal captured in images and therefore affect the features used by the classification model. Recognizing how different soil types and minor water effects alter image appearance is essential for understanding the dataset, interpreting model outputs and refining the training process for more accurate predictions.

4.2 Experimental Settings

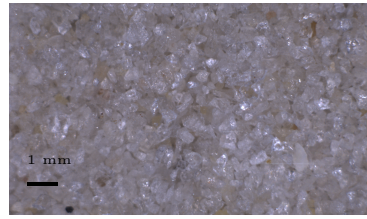
The CNN model is trained from scratch with 80 epochs, using Adam optimizer and a batch size of 64. The learning rate is set to 0.001. Moreover, the parameters α and β in our proposed texture loss function are empirically set to 1 and 0.1, respectively, based on experiments showing stable performance, as this configuration provides a good trade-off between reconstruction accuracy and texture regularization as shown in Table 1. In addition to this partial ablation on the loss function, we conduct a complementary architectural ablation study to evaluate the contribution of the hierarchical design and feature fusion strategy. Specifically, we compare several variants: (i) our proposed full model combining CNN features and macro predictions, the 4D vector (M0), (ii) a CNN-only variant without macro information (M1), (iii) a macro-only variant without image features (M2), (iv) a simplified linear fusion model (M3), and (v) a baseline using a linear fusion and only global CNN features (M4). The results are reported in Table 2. The results show that the full model (M0) consistently achieves the best performance, confirming the complementarity between CNN features and macro-level predictions. In contrast, removing either component (M1, M2) leads to a noticeable performance drop, highlighting the importance of both visual and compositional information. Additionally, the comparison between MLP and linear fusion (M0 vs M3) demonstrates the benefit of non-linear feature interactions. These variants help to show how different parts of the model contribute, but a more detailed ablation study, looking at architecture depth, freezing strategy, and loss design, would better isolate individual effect of each component. Our dataset was split into 80% for training and 20% for testing. For the MLP model, we also used the Adam optimizer, a batch size of 64, and a learning rate of 0.001, training for 40 epochs. Our hardware is NVIDIA Tesla V100. The MLP serves to integrate the local feature representations extracted by the CNN into a global estimate of sub-class proportions. CNN features capture textures and patterns within limited receptive fields, providing local information that alone is insufficient for image-level prediction. By concatenating macroclass labels with



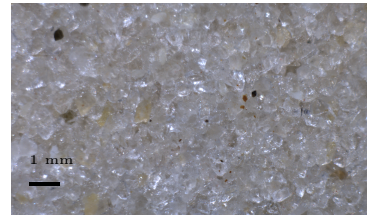
(a) Silt - W0



(b) Silt - W100



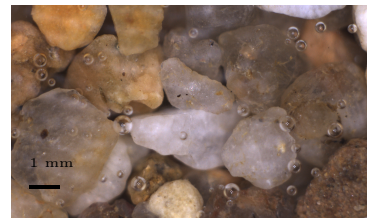
(c) Sand - W10



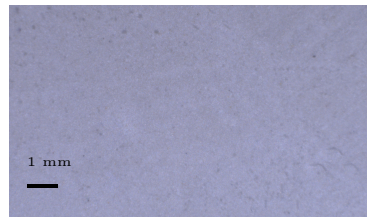
(d) Sand - W100



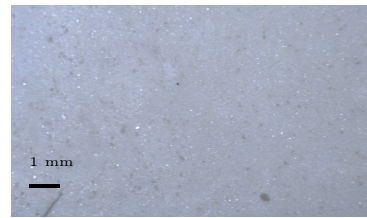
(e) Coarse Sand - W0



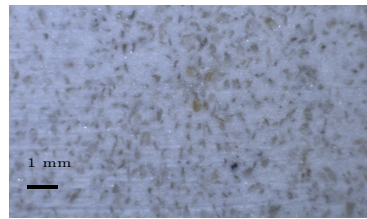
(f) Coarse Sand - W95



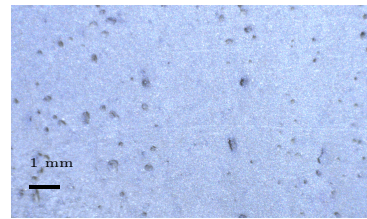
(g) Clay - W10



(h) Clay - W100



(i) Sandy Silt - W0



(j) Sandy Silt - W100

Fig. 5: Representation of the soil lithologies with different water saturation levels (W expressed in %) from our dataset.

these features, the MLP learns to model complex nonlinear interactions and aggregate local signals into a coherent global output.

α	β	R^2	RMSE
0.1	0.1	0.9815	0.00058
0.1	0.5	0.9701	0.00161
0.5	0.1	0.9798	0.00109
0.5	0.5	0.9885	0.00062
1.0	0.1	0.9890	0.00059
1.0	0.5	0.9885	0.00062

Table 1: A detailed comparison of different α and β settings.

Model	Features Used	Fusion Type	Sub $R^2 \uparrow$	Sub RMSE $\downarrow$	Loss Value $\downarrow$
M0	CNN + 4D	MLP	0.9767	0.015426	0.35143
M1	CNN	MLP	0.9699	0.017118	0.43900
M2	4D	MLP	0.6728	0.050958	1.87785
M3	CNN + 4D	Linear	0.9164	0.028186	1.64637
M4	CNN	Linear	0.8851	0.032411	2.27780

Table 2: Ablation study over architectural variants.

4.3 Baseline Models

Only two studies in the literature address the prediction of soil particle size distribution curves, namely Soranzo et al. [16] and St-Cyr et al. [13]. However, neither their trained models nor their complete datasets are publicly available, which prevents any direct comparison with our approach.

So, we evaluate a set of representative deep learning architectures under identical training conditions on our dataset. This provides a controlled and consistent benchmark against widely used state of the art vision backbones adapted to the regression task. Specifically, we consider ResNet50, DenseNet121, EfficientNetB0, and ViT16, in addition to our proposed CNN. For a fair evaluation, all these architectures are adapted to our regression task by appending the same multilayer perceptron (MLP) head used in our model, replacing their original classification task. Each model is evaluated using our proposed texture-based loss function.

While this does not reproduce full pipelines from prior work, it enables a controlled comparison of representation learning capacity under identical data conditions against established deep learning models. The results corresponding to the different model configurations are reported in Table 3.

Model	$R^2 \uparrow$	RMSE $\downarrow$	Time(h) $\downarrow$	Size(MB) $\downarrow$
ResNet50	0.96684	0.0188	141	91.68
DenseNet121	0.97137	0.0172	86	27.53
EfficientNetB0	0.84877	0.0362	262	16.54
ViT16	0.47956	0.0787	112	328.05
CORSIA (ours)	0.97544	0.0158	70	0.55

Table 3: Comparison of regression performance across architectures using the proposed texture-based loss, with training time and model size.

In conclusion, the results presented in Table 3 demonstrate that the proposed CORSIA model outperforms the evaluated pre-trained architectures in terms of regression performance, achieving both higher coefficients of determination (R^2) and lower root mean squared error (RMSE). In addition, CORSIA requires less overall computational time, as reflected by the total execution time reported. While pre-trained models are often considered strong baselines, they do not provide a favorable trade-off in this study, as they incur higher computational costs without yielding superior performance. ResNet50 and DenseNet121 achieve performance levels relatively close to CORSIA in terms of R^2 and RMSE; however, their optimization remains limited, with early stabilization during training suggesting reduced capacity for further improvement. In contrast, ViT16 exhibits less stable training behavior and does not reach comparable performance levels, indicating difficulties in adapting to the considered regression task. Overall, CORSIA combines accuracy and efficiency within a lightweight architecture characterized by a reduced memory footprint. These properties make it particularly well-suited for embedded applications, especially for deployment on tablets in field conditions where computational resources are limited. Consequently, the proposed model represents a practical and effective solution, balancing simplicity, robustness, and competitive performance.

4.4 Superiority of Texture Loss

To show the effectiveness and superiority of our proposed texture Loss function, we test our proposed CNN network architecture by employing three different loss functions including cross-entropy loss (CE), L1 loss (L1) and Mean-Squared Error loss (MSE) to predict our first 4 classes. The Table 4 below reports the classification accuracy rates of the resulting macro-level model with the three loss functions compared to the proposed texture loss function. The performance is measured using Macro R^2 and Macro RMSE to assess both the explained

variance and the prediction error on the four macro classes.

Loss Function	Macro R^2	Macro RMSE
L1 Loss	0.8241	0.09797
CE Loss	0.9719	0.02114
MSE Loss	0.9621	0.05813
Proposed Loss	0.9890	0.00059

Table 4: Performance comparison of the model using four different loss functions in terms of Macro R^2 and Macro RMSE.

The results show that the choice of loss function significantly impacts model performance. While L1 loss leads to weaker results and higher prediction errors, both CE and MSE losses provide strong baseline performance under the conditional setting. In contrast, the proposed loss achieves the best overall performance, combining the highest Macro R^2 with the lowest RMSE. This indicates a better balance between accurately capturing data variability and minimizing prediction error. Although the evaluation is conducted under conditional sampling, which may not fully reflect real-world conditions, these results suggest that the proposed loss offers improved robustness and generalization. Its design allows it to better capture texture-specific patterns, leading to more accurate and stable predictions.

Overall, the proposed loss stands out as the most effective choice among the evaluated methods for tasks requiring reliable and precise classification of the targeted classes.

4.5 Results and Discussions

To assess the performance and robustness of our model, we conducted a 5-fold cross-validation on the dataset. For the 240 tested samples, the model achieved good results. Table 5 summarizes the performance metrics averaged over the 5 folds: both macro-averaged (A, L, S, G) and micro-averaged (over the 10 sub-classes), R^2 and RMSE values are reported.

Across all folds, the model demonstrates a high and consistent predictive accuracy, as reflected by a macro- R^2 score of 0.939 ± 0.002 . The limited variability observed across folds indicates that the model generalizes well and is not sensitive to the particular data partitioning used during training and evaluation. In addition to the overall performance, the results aggregated at the sub-target level further confirm the stability of the model. This suggests that the learned representations capture the underlying relationships in the data in a consistent manner across different prediction tasks. A similar behavior is observed for the error-based metrics. Both macro and sub-target RMSE values remain low across

Metric	Fold1	Fold2	Fold3	Fold4	Fold5	Mean	Std
Macro_R2	0.96030	0.89330	0.90060	0.98900	0.95480	0.93960	0.04112
Macro_RMSE	0.00242	0.00513	0.00401	0.00060	0.00217	0.00287	0.00175
Micro_R2	0.87688	0.81928	0.86168	0.79643	0.85116	0.84109	0.03270
Micro_RMSE	0.03697	0.04093	0.04092	0.04891	0.03778	0.04110	0.00472

Table 5: Model performance metrics across 5-fold cross-validation.

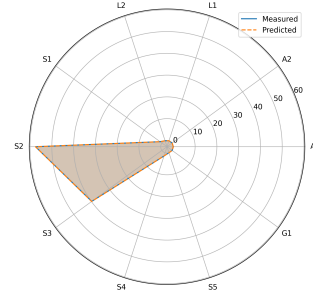
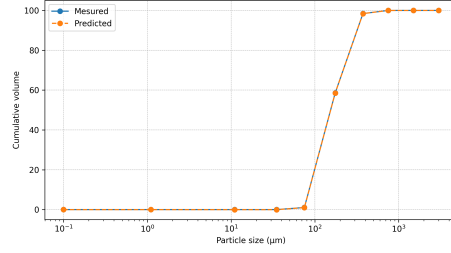
folds, with limited dispersion, indicating reliable error control and reinforcing the robustness of the model predictions. The combination of high coefficient of determination and low prediction error across multiple folds highlights the effectiveness of the proposed approach.

Overall, these results demonstrate that the model consistently delivers accurate predictions across all evaluated metrics, thereby supporting its suitability for deployment in scenarios that require both reliability and reproducibility.

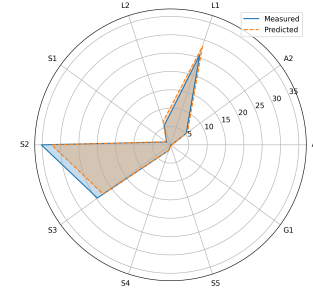
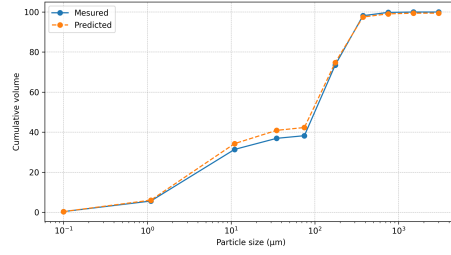
Figure 6 illustrates the predicted vs experimentally measured granulometric curves and radar visualisation for some selected samples. Each duo is given with its individual RMSE to allow a detailed per-sample evaluation of performance. The percentages of S, L, A, and G represent the sample’s soil composition. The granulometric curve obtained on Sample 1 (100%S of Figure 6 (a)) demonstrates excellent alignment across the curve, confirming the ability of the model to predict textures dominated by a single particle size. The results of on Sample 2 (62.9%S, 31.5%L, 5.6%A of Figure 6 (b)) highlights that predictions remain accurate for mixed compositions, although minor deviations are observed in the silt fraction, likely caused by the less pronounced texture patterns. The curve obtained on Sample 3, (5.1% G, 91.2%S, 3.4%L, 0.3%A of Figure 6 (c)) shows that the model captures the grain-size distribution despite coarse dominance and remains robust under intermediate saturation conditions.

5 Conclusion

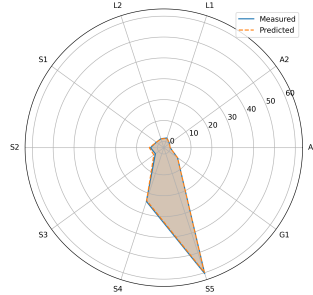
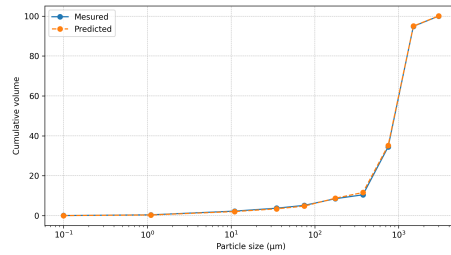
The results presented in this study demonstrate that the proposed convolutional regression framework CORSIA, combined with the texture loss, effectively captures the granulometric structure of soil samples and produces highly accurate and consistent predictions across both macro and subclass levels. By explicitly incorporating texture information, the model outperforms some traditional loss functions, confirming the importance of texture-aware learning for soil granulometry. These findings highlight the potential of the proposed approach for robust and reliable inference from soil core images. For future work, we plan to adapt and validate this framework on real soil image acquisitions, addressing challenges such as heterogeneous lighting, natural soil variability, and more variable water content, to further assess its generalization capability and practical applicability in real-world soil monitoring.



(a) Sample 1 : 100%*S* with 100% water saturation ($RMSE = 0.00024$).



(b) Sample 2 : 62.9%*S*, 31.5%*L*, 5.6%*A* with 0% water saturation ($RMSE = 0.01365$).



(c) Sample 3 : 5.1%*G*, 91.2%*S*, 3.4%*L*, 0.3%*A* with 50% water saturation ($RMSE = 0.00432$).

Fig. 6: Predicted vs experimental granulometric curve and radar visualisation for various compositions of gravel (*G*), sand (*S*), silt (*L*) and clay (*A*) including different water saturations.

Acknowledgements This work was conducted as part of an industrial PhD funded by the French National Association for Research and Technology. The author thanks Ginger BURGEAP and LIRIS for their supervision and scientific guidance. We thank the CNRS/IN2P3 Computing Center for providing computing resources for this work.

References

1. Geoengineer. Geoengineer Link, last accessed 17 Nov 2025
2. Aragon, C.: Soil textures according to the usda (usda)
3. Dong, Y., Lu, N., Li, X.: Dense attention networks for texture classification. *Neurocomputing* **634**, 129833 (2025)
4. Granitzer, A.N., Beck, J., Leo, J., Tschuchnigg, F.: Explainable insight into the vision-based classification of soil core samples from close-range images. In: Third Workshop on the Future of Machine Learning in Geotechnics. pp. 242–247 (2025)
5. Hostyn, G.: Contribution des fractions granulométriques grossières au fonctionnement des sols très anthropisés. *SIReNA* (2022)
6. Hu, Y., Wang, J., Wang, X., Yu, J., Zhang, J.: Efficient virtual-to-real dataset synthesis for amodal instance segmentation of occlusion-aware rockfill material gradation detection. *Expert Systems with Applications* **238**, 122046 (2024)
7. Inazumi, S., Intui, S., Jotisankasa, A., Chaiprakaikeow, S., Kojima, K.: Artificial intelligence system for supporting soil classification. *Results in Engineering* **8**, 100188 (2020)
8. Jia, S., Li, H., Wang, Y., Tong, R., Li, Q.: Hyperspectral imaging analysis for the classification of soil types and the determination of soil total nitrogen. *Sensors* **17**(10), 2252 (2017)
9. Jiang, Z., Leung, A., Liu, J.: Segmentation uncertainty of vegetated porous media propagates during x-ray ct image-based analysis. *Plant and Soil* **511**(1), 969–995 (2025)
10. Manashti, J., Pirnia, P., Manashty, A., Ujan, S., Toews, M., Duhaime, F.: Psd-net: Determination of particle size distributions using synthetic soil images and convolutional neural networks. *arXiv preprint arXiv:2303.04269* (2023)
11. Mengistu, A., Alemayehu, D.: Soil characterization and classification: A hybrid approach of computer vision and sensor network. *International Journal of Electrical & Computer Engineering* (2088-8708) **8**(2) (2018)
12. Pandiri, D.K., Murugan, R., Goel, T.: Smart soil image classification system using lightweight convolutional neural network. *ESA* **238**, 122185 (2024)
13. Plante St-Cyr, T., Duhaime, F., Dubé, J., Grenier, S.: Dataset of soil images with corresponding particle size distributions for photogranulometry (2025), <https://arxiv.org/abs/2506.17469>
14. S., H.: Introduction to Linear Algebra and the Theory of Matrices (1961)
15. Shenbagavalli, R., Ramar, K.: Classification of soil textures based on laws features extracted from preprocessing images on sequential and random windows. *Bonfring International Journal of Advances in Image Processing* **1**, 15 (2011)
16. Soranzo, E., Guardiani, C., Wu, W.: Convolutional neural network prediction of the particle size distribution of soil from close-range images. *Soils and Foundations* **65**(1), 101575 (2025)
17. Yang, J., Yu, W., Fang, H., Huang, X., Chen, S.: Detection of size of manufactured sand particles based on digital image processing. *PLOS ONE* **13** (2018). <https://doi.org/10.1371/journal.pone.0206135>