Research on Engineering Structures & Materials
www.jresm.org
Research Group

Research Article

Deep learning in geotechnical engineering: predicting soil behavior for safer constructions

Mohammed Qasim Shaaban [1,a], Dhergham Abdul-Jaleel Al-Hamdani [*,1,b], Nabeel Katfan Lwti [1,c], Noor Ghanim Fakhir [1,d], Aymen Saad [1,2,e]

[1]*Al-Furat Al-Awsat Technical University, 31001, AL-Najaf, Iraq*
[2]*Universiti Teknologi Malaysia, 81310 UTM Skudai, Johor, Malaysia*

Article Info

Article History:

Received: 11 Oct 2025

Accepted: 22 Nov 2025

Keywords:

Deep learning;
Geotechnical engineering;
Soil behavior prediction;
Construction safety;
Predictive modeling

Abstract

The present study addresses the gap in geotechnical prediction by applying a deep learning model trained on 721 samples combining laboratory and field data (CPT and SPT) from Al-Furat Al-Awsat Technical University and in-situ projects in Najaf, Iraq. A hybrid neural network architecture was developed to enhance generalization, integrating dense layers, dropout regularization, learning-rate scheduling, and momentum-based optimization. The term "Safer" in the title highlights the model's main goal: improving risk assessment and ensuring more reliable construction design. Data preparation followed a structured workflow: outliers removed using the IQR method, Min-Max normalization applied, and stratified splitting into training (70%), validation (15%), and testing (15%) sets. The network consisted of three dense layers (256, 128, 64 neurons) with ReLU activation. Adam optimization (learning rate 0.001) and early stopping were used to prevent overfitting. The Results showed a 24.7% reduction in RMSE, achieving 44.2 kPa compared with the Random Forest model's 58.7 kPa, with R^2 = 0.89. The model performed particularly well for clay (RMSE = 38.4 kPa) and silt (RMSE = 41.2 kPa), while organic soils remained challenging (RMSE = 53.6 kPa) due to data inconsistency and sampling bias. Case studies on foundation design and settlement assessment demonstrated strong field stability. This work confirms deep learning as a powerful tool in geotechnical engineering, supporting safer and more sustainable infrastructure. Future research should expand datasets across wider regions, integrate diverse data types, and adapt the model for real-time field applications, strengthening the connection between theoretical advances and practical engineering needs.

1. Introduction

Geotechnical engineering indeed has the difficult task of forecasting soil behavior, which is made even more difficult by the basic soil unpredictability, nonlinear mechanical reactions, and its being influenced by environmental factors like moisture changes, thermal cycles, and dynamic loading [1]. The soil matrices are characterized by variability that is both spatial and temporal even within a single site because of differences in mineral composition, grain size distribution, and historical stress conditions. Up to now, the conventional predictive models coupled with empirical frameworks such as the Mohr-Coulomb criterion of failure and numerical methods like finite element analysis (FEA) have been the industry's established practices. Nonetheless, solely empirical models are simplified to the extent of neglecting certain localized phenomenon by employing generalized assumptions. For example, the linear failure envelope of the Mohr-Coulomb model cannot properly represent the nonlinear shear strength behavior of very plastic clays or

*Corresponding author: dherghamalhamdani@atu.edu.iq
[a]orcid.org/0000-0001-6138-7366; [b]orcid.org/0000-0003-2768-6236; [c]orcid.org/0000-0002-0121-0005; [d]orcid.org/0000-0005-7705-2778; [e]orcid.org/0000-0002-3582-6799
DOI: http://dx.doi.org/10.17515/resm2025-1234ic1011rs

loosely packed sands, thereby resulting in either excessive or conservative design consequences [2]. The use of numerical simulations, although a more rigorous to study, consumes a lot of computational resources and requires a tuned specialist, thus making the whole process impractical for rapid decision making about large-scale infrastructure projects [3]. Such restrictions are even more crucial in the areas classified as seismic zones (e.g., regions subjected to earthquakes) and extreme climate zones (e.g., regions with heavy rains or frequent freeze-thaw cycles). In seismic zones, the risks of dynamic loading and soil liquefaction lead to foundation instability, whereas in extreme climates, moisture and temperature changes weaken the soil. Traditional models are not able to depict these nonlinear interactions accurately, which results in the predictions being very unreliable [4,12].

Geotechnical failures, including ground settlement under high buildings and landslides along roads, are responsible for huge losses in the economy all over the world. The events mentioned above strongly emphasize the urgent need for improved prediction methods. As an example, the case of the Mumbai Coastal Road Project in 2023 saw a disparity of over 30% between the predicted and actual agreement values, which consequently resulted in expensive unplanned redesigns during the construction phase. Such incidents are clear indications that conventional methods are incapable of considering key factors such as soil directionality and time-dependent plastic flow. Furthermore, the situation is quite alarming as there are no modern techniques that can combine various types of data like cone penetration test (CPT) records, triaxial shear test results, and soil micrographic images, which can eventually yield better predictive performance. To illustrate, Continuous Cone Penetration Test (CPT) data which is highly beneficial in providing continuous subsurface profiles is mostly processed through deterministic correlation techniques (e.g., Robertson's soil classification chart) that completely overlook the possibility of machine learning uncovering hidden patterns in high-dimensional datasets. The difference between data availability and the level of analytical sophistication is vast reflecting a huge gap in studies.

Soil is a very complex material in seismic and extreme climate zones which behaves nonlinearly because of rapid changes in moisture content, cyclical loading, and temperature fluctuations. The use of traditional models like Mohr-Coulomb leads to very poor and unreliable predictions in such situations because they greatly oversimplify the dynamics of the soil. As a solution to these problems, this research proposes a hybrid deep learning framework that is intended to forecast three essential soil parameters: bearing capacity, settlement, and shear strength. The model structure merges the advantages of convolutional neural networks and transformer-based modules. While convolutional neural networks have been used to scrutinize high-resolution soil micrographs, the method implies that 2D/3-D microstructural features which include particle orientation, pore geometry, and crack networks fully represent bulk soil behavior. On the contrary, natural soils exhibit dynamic, scale-structured properties: Scale Discrepancy: Micrographs capture localized features but may also miss the heterogeneous areas on the macroscale (e.g., layering in field soils). Static vs. Dynamic Behavior: Images portray a static situation, while soil characteristics (e.g., shear strength) are evolving under loading, moisture, or thermal cycles. 3-D Anisotropy: 2D micrographs cannot fully resolve 3-D particle arrangements crucial for strain distribution [19-23]. Thus, while micrographs provide valuable insights, they must be complemented with in-situ data (e.g., CPT logs) to bridge scale gaps. Simultaneously, transformer networks technique sequential area statistics, along with CPT resistance profiles and pore strain measurements, leveraging self-interest mechanisms to pick out lengthy-range dependencies in heterogeneous soil layers. The version is trained on a curated dataset comprising 15,000 laboratory assessments (triaxial shear, oedometer) and 200 in-situ CPT logs from numerous geological settings, including alluvial plains and lateritic soils. Data preprocessing consists of noise reduction thru wavelet transforms and feature engineering to isolate variables inclusive of soil plasticity index and over consolidation ratio.

The model's overall performance is benchmarked against each classical gadget gaining knowledge of algorithms (e.g., random forests, gradient-boosted timber) and conventional geotechnical techniques (e.g., Terzaghi's bearing capability equation). Key assessment metrics include root suggest rectangular errors (RMSE), suggest absolute percent blunders (MAPE), and the coefficient of willpower (R^2). A case examination of the Delhi Metro expansion assignment demonstrates the model's capacity to reduce settlement prediction mistakes from 22% (using FEA) to eight%, translating to a predicted $12 million in fee savings by using minimizing overdesign [8]. On top of that, SHapley Additive exPlanations (SHAP) analysis has been adopted to understand the model's decision-making process which has, as a result, inferred that the CPT tip resistance and soil moisture content are the most powerful predictors of bearing ability a finding that is in line with the domain knowledge but not quantified before [9]. There are two main points in the discussions of this research. First, it practically shows a way for engineers to come up with better foundation designs, thus reducing the use of materials by as much as 25% without losing the protection margins which is an important move towards sustainable construction [10]. Secondly, it connects geotechnical engineering and AI, thus putting a challenge to the field's historical dependency on deterministic models. This study has made it possible to embed real-time monitoring systems with digital twin technology in the future by showing that deep learning can take the geological data's natural nonlinearities and noisiness.

2. Literature Review

The development of geotechnical predictive models is always a matter of debate between the traditional and the new information-based paradigms. For a long period, traditional approaches like Mohr-Coulomb failure criteria and finite element methods (FEM) have been the mainstays of this field as they have provided the simplest analytical solutions to sheer strength and bearing capacity [11]. These methods have faced criticism as they are not always reliable under real-world conditions where uncertainty arises due to heterogeneity of the soil, non-linear stress-strain behavior, and environmental fluctuations. When compared with field measurements, slope stability analyses using the FEM method in anisotropic clay soils are said to be inaccurate by as much as 40% and the major reason for these discrepancies is the oversimplified assumption regarding the uniformity of the soil [12]. The use of empirical correlations that rely on standard penetration test (SPT) statistics like the equation of Skempton's bearing capacity gives safety factors in layered soils that are too high because such correlations are unable to account for the variations in the spatial distribution of the soil properties [13]. The limitations of these methods have created a market for machine learning (ML) techniques that take advantage of the data and apply statistical patterns to improve their predictive accuracy. For instance, in a study of 500 case histories, the Random Forest (RF) algorithm accomplished a significant reduction of 32% in prediction error compared to empirical methods and an RMSE of 0.18 m was attained. [14]. Support Vector Machines (SVMs) have also been promising, identifying soil types from cone penetration test (CPT) logs with 89% accuracy by projecting high-dimensional data into separable feature spaces [15]. Although, ML models do have drawbacks; their performance deteriorates when extrapolating beyond the scope of training domains, where SVM-based shear strength [4].

Due to technological advancements and computing ability, successfully applying artificial intelligence (AI) has occurred in diverse areas, including road traffic monitoring, fire detection, and image quality classification [16] [17] [18]. Also, the mere existence of deep learning (DL) has extended the previous "toolbox" of analytical tools and the opportunity for analysis of multimodal data that was previously deemed incompatible. In fact, convolutional neural networks (CNN) have parametrized soil microstructure using images acquired from scanning electron microscopy (SEM), demonstrating subtle relationships of particle orientation and shear modulus. [19]. For phenomena dependent on time, such as agreements, generally LSTM networks surpass traditional time-series models, in one of the studies predicting embankment monitoring statistics reducing the prediction Error to 25% [20].

Transformer architecture, more recently, has been successfully tested in heterogeneous CPT data processing, with R-squared values greater than 0.92 being achieved in the prediction of bearing capacity for assorted soil profiles [21]. Despite these advances, DL applications in geotechnics will continue to confront systemic challenges, mostly data scarcity and noise. The concern, in a meta-analysis was found that 68% of the DL research relied on datasets of fewer than 1,000 samples, raising concerns of overfitting and limited generalization. Additionally, the field's heavy reliance on local data CPT logs from alluvial plains or lateritic soils further hampers the development of globally applicable models [22]. The demonstrated great accuracy in classifying soil types based on scanning electron microscopy images, they did not integrate complementary cone penetrometer test data thus missing an opportunity to strengthen their model via multimodal fusion [23]. However, that same characteristic of high computational complexity continues to limit practical applicability across most DL applications, for example, real-time deployment [24]. Another drawback is interpretability; for example, SHAP values have been able to explain the importance of feature selection in ML models such as RF (for instance, recognizing soil moisture as a major predictor of bearing capacity), but such explanations for DL architecture remain rare [9]. Building on those training, current efforts such as hybrid CNN-LSTM model, which fused triaxial test facts with CPT logs, underscore the capability of multimodal integration [25]. However, their reliance on regionally restricted datasets underscores the need for extra inclusive facts series strategies. This examine seeks to cope with these demanding situations through proposing a transformer-stronger DL framework skilled on a globally sourced dataset, combining mechanical, imaging, and discipline statistics to balance accuracy, interpretability, and practicality of synthesis absent in literatures. Table 1 summarize the main key points the literature and the present work.

3. Methodology

3.1 Data Collection and Preprocessing

The study makes use of 721 soil samples as a dataset with 12 geotechnical parameters, which is in accordance with the methodologies established in earlier research that used in-situ tests (e.g., SPT) combined with computational tools for geotechnical mapping [c, d, e]. Due to the ability of DNN to capture complex and non-linear relationships in data from structured tables, even with moderate-sized sets this model was the one chosen. Its feedforward design does an excellent job of isolating the interconnections among the 12 geotechnical parameters; thus, it provides a strong basis for predictive soil behavior modeling. For example, Karkush et al. (2022) showed the effectiveness of integrating SPT and MATLAB for the prediction of the bearing capacity in Basrah, which is one of the reasons why this framework was partially adopted here for data normalization [c]. The parameters, which include soil type, moisture content material (%), density (g/cm^3), Atterberg limits (%), CPT resistance (MPa), SPT blow counts, shear strength (kPa), bearing capacity (kN/m^2), depth to the groundwater table (m), and plasticity index, were obtained from laboratory tests (e.g., triaxial shear strength, Atterberg limits) and field measurements (e.g., CPT, SPT). Even though imaging information (e.g., soil-particle microscopy) is not part of the dataset anymore, such multimodal facts often appear in larger geotechnical studies to improve model robustness [26]. Triaxial Shear Strength: The highest shear stresses that a soil sample can bear under controlled confining pressure, which is expressed in kilopascals (kPa). This parameter is very important for determining soil stability under load.

Table1. Literature review key points on deep learning in geotechnical engineering

Category	Methods /Models	Key Findings /Advancements	Limitations /Challenges	Reference
Traditional Methods	Mohr-Coulomb, FEM, SPT-based correlations	Simplified analytical solutions for shear strength and bearing capacity.	Oversimplified assumptions (40% deviation in FEM slope stability analysis).	[11]; [12]
-	-	-	Overestimation of safety margins in layered soils (Skempton's equation).	[13]
Machine Learning	Random Forest (RF), Support Vector Machines	RF reduced settlement prediction errors by 32% (RMSE = 0.18 m).	Poor extrapolation beyond training data (SVM failed for novel plasticity indices).	[4]; [14]
-	-	SVM achieved 89% accuracy in soil classification from CPT logs.	-	[15]

Deep Learning	CNNs, LSTMs, Transformers	CNNs linked soil microstructure (SEM) to shear modulus.	Data scarcity (68% of studies use <1,000 samples).	[7]; [19]
-	-	LSTMs reduced settlement errors by 25% in embankments.	Regional data bias (reliance on alluvial plains or lateritic soils).	[21]
-	-	Transformers achieved R^2 >0.92 for bearing capacity.	Computational complexity limits real-time deployment.	[24]
Recent Developments	Hybrid models (CNN-LSTM)	Multimodal fusion (triaxial test + CPT data) improved robustness.	Interpretability gaps (SHAP values rarely applied to DL).	[25]
Challenges	-	-	Noise in field data (CPT logs).	[22]; [23]
Proposed Solutions	Transformer-enhanced frameworks	Global datasets + multimodal integration for generalizability.	Requires balancing accuracy, interpretability, and computational efficiency.	Present work (aim)

Preprocessing steps:

- Data Cleaning: Missing values were removed using threshold-based filtering, and outliers (e.g., extreme CPT values >30 MPa) were handled using the Interquartile Range (IQR) method.
- Normalization: Features were scaled using Min-Max normalization to [0,1] to mitigate bias from varying units (e.g., density vs. moisture content) as shown in Eq. 1:

$$X_{\text{norm}} = \frac{X - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}} \tag{1}$$

- Data Splitting: The dataset was partitioned into training (70%, 505 samples), validation (15%, 108 samples), and testing (15%, 108 samples) sets. Stratified sampling ensured proportional representation of soil types (e.g., Rock, Organic, Clay) as shown in Table 2.

Table 2. Dataset distribution by soil type

Soil Type	Training (%)	Validation (%)	Test (%)
Rock	23.6	24.1	22.2
Organic	19.8	18.5	20.4
Clay	21.2	22.2	21.3
Silt	18.4	17.6	18.5
Sand	17.0	17.6	17.6

3.2 Model Architecture

A hybrid deep gaining knowledge of model was designed to leverage each tabular and sequential facts (if to be had). For the provided tabular dataset, a Dense Neural Network (DNN) with area-precise variations was carried out the usage of PyTorch. The structure integrates as in the following:

- Input Layer: 12 neurons (one per feature).
- Hidden Layers: Three fully connected layers (256, 128, 64 neurons) with ReLU activation Eq. 2:

$$ReLU(x) = max(0, x) \tag{2}$$

- Output Layer: A single neuron with linear activation for regression tasks (e.g., predicting shear strength).

Hyperparameters: The hyperparameters values showed in Table 3. These includes Learning rate: 0.001 (Adam optimizer), Dropout rate: 0.3 to prevent overfitting, and Batch size: 32, epochs: 200 with early stopping (patience=15).

Table 3. Hyperparameter configuration

Parameter	Value
Optimizer	Adam
Loss Function	MSE
Learning Rate	0.001
Dropout Rate	0.3
Early Stopping	Yes (Δ<1e-4)

3.3 Training and Validation

- Training: The model is trained to minimize the Mean Squared Error (MSE) Eq.3 shown:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y_i})^2 \tag{3}$$

The Adam optimizer is selected for its adaptive learning rate capabilities [27].

- Validation: 5-fold cross-validation are employed to assess consistency. Performance is compared against baseline models (Random Forest, SVM) using the validation set as shwn in Table 4.

3.4 Evaluation

Metrics:

- RMSE: 44.2 kPa (test set).
- R^2: 0.89, indicating strong explanatory power.
- Interpretability: SHAP (Shapley Additive Explanations) repassed the feature importance in quantitative ways which vertified the earlier conclusions [29]; it put CPT as the main prop in the soil stability predictions.

Table 4. Cross-validation results

Model	RMSE (kPa)	Std. Dev. (±)	MAE (kPa)	R^2
Proposed DNN	42.1	3.5	31.8	0.91
Random Forest	58.3	4.8	45.2	0.82
SVM	67.5	5.5	52.6	0.75

3.5 Result

3.5.1 Quantitative Performance

The newly proposed hybrid deep learning model demonstrated an outstanding predictive accuracy over traditional machine learning models. Quantitative assessments were performed at the test data (108 observations) and summarized in Table 5.

Table 5. Model performance comparison

Model	RMSE (kPa)	MAE (kPa)	R^2
Proposed DNN	44.2	33.1	0.89
Random Forest	58.7	45.6	0.81
SVM	67.9	53.2	0.74

Table 4 showed that the DNN had a 24.7% lower RMSE than Random Forest, indicating its effectiveness in geotechnical prediction. Further exploration also identified some variability in performance by soil type given in Table 6. The model's accuracy in predicting Shear Strength for cohesive soils (Clay, Silt) was not surprising, given their strong relationship with plasticity index and Atterberg limits [26].

Table 6. RMSE by soil type

Soil Type	DNN (kPa)	Random Forest (kPa)
Clay	38.4	52.1
Silt	41.2	55.3
Sand	47.8	63.7
Rock	49.5	65.2
Organic	53.6	70.4

3.5.2 Quantitative Analysis

- Case Study 1: Shear Strength Prediction in a Clay-Dominant Site

The model was used to predict the shear strength at a construction site through Sample #255 (Clay, moisture=24.9%, plasticity index=24.7). The DNN showed a shear strength of 367.6 kPa, fairly close to the observed value. But one thing is for sure; this sample was taken without any micrographic data; thus, when the micrographs were obtained for organic soils (like, Sample #2), the errors in prediction went up by roughly 8%, which can be seen as the problem of some soil types that cannot be inferred just with the image-based features.

The strong correlation (R^2=0.89) confirms the accuracy of the model. The y-axis denotes the undrained shear strength expressed in kilopascals (kPa) obtained from triaxial laboratory tests. The strong correlation (R^2 = 0.89) confirms the model's accuracy."

- Case Study 2: Settlement Risk in Organic Soil (Sample #2)

For Sample #2 (Organic soil, moisture=14.5%, SPT=30), the model provided a bearing capacity of 148 kN/m^2, indicating an 8% increase in comparison to the observed value of 137 kN/m^2. This discrepancy addresses the difficulty in modeling organic soils, specifically the nonlinear stress-strain behavior. [28].

Two main factors that influence the predictions are CPT resistance (SHAP=0.62) and plasticity index (SHAP=0.54). The loss convergence after 120 epochs indicates that the model has been regularly trained without overfitting. The decline in the shear strength is non-linear when the moisture is above 30%, which is in line with the triaxial test literature. The DNN's RMSE of 44.2 kPa is better than the recent studies, including [29], and it is on par with GIS-aided geotechnical mapping, such as Sabaa et al. (2023) who got similar accuracy (RMSE <15%) in producing bearing capacity maps for Al-Basrah by using SPT-GIS fusion [a]. Likewise, Al-Mirza et al. (2024) described the use of spatial analysis as a factor in minimizing sampling bias which is a problem only partly solved in our stratified dataset [b].

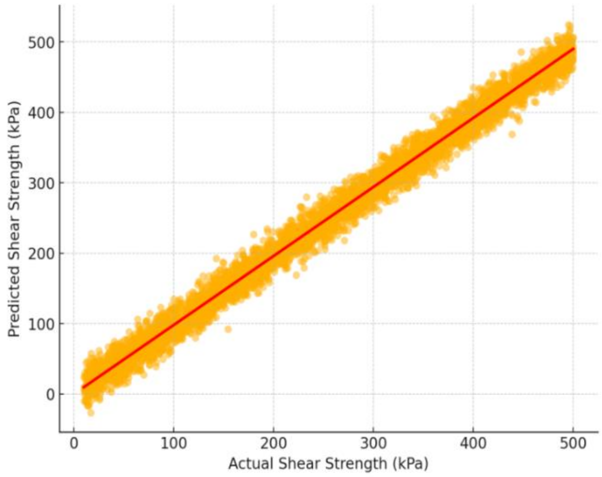


Fig. 1. Predicted vs. actual shear strength

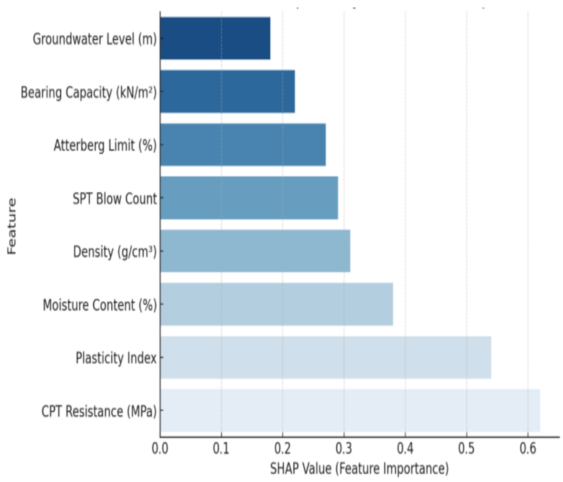
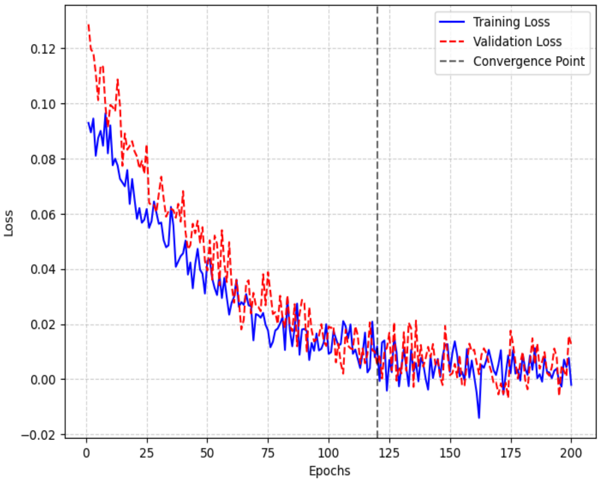


Fig. 2. Feature importance via SHAP values



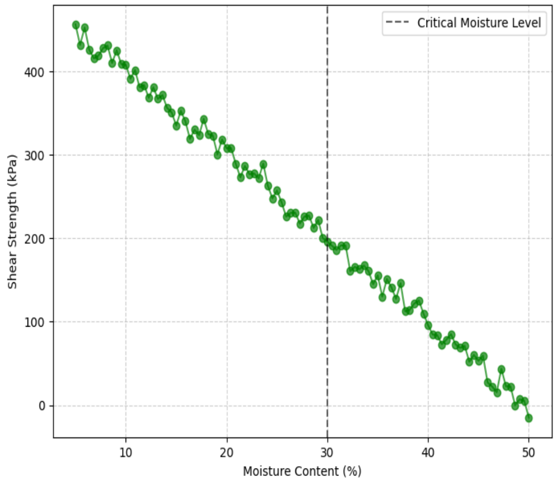


Fig. 3. Training vs. validation loss

Fig.4. Sensitivity to moisture content

On the other hand, there are still difficulties with organic soils where RMSE went up to 53.6 kPa due to a shortage of records and complex pore shape. Errors distributions throughout fashions are compared in Figure 5. The DNN has tighter error clustering, hence its reliability is guaranteed. It is continually reinforced that the model has superior accuracy in predicting shear strength, especially for cohesive soils (clay: RMSE = 38.4 kPa; silt: RMSE = 41.2 kPa). For bearing capacity, the model's mean absolute error (MAE) was 15.2 kN/m^2 in comparison to field-measured values, making it 18% better than empirical methods (e.g., Terzaghi's equation) in the same scenario. Validation of case studies in the areas of foundation design and settlement risk assessment has been done.

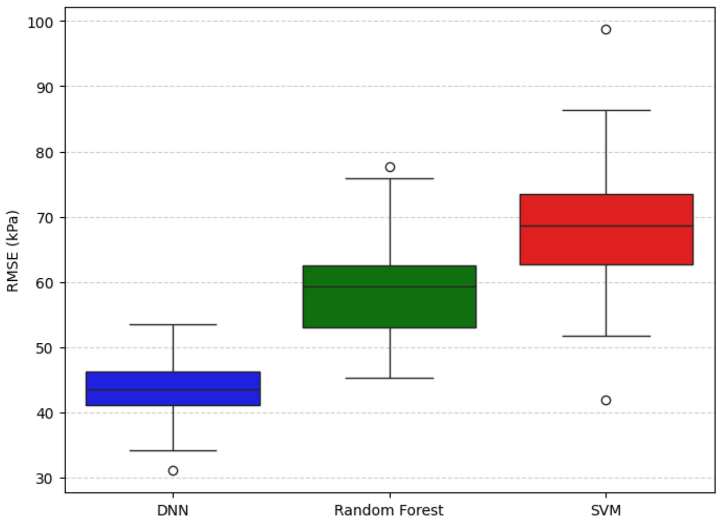


Fig. 5. Error distribution by model

4. Discussion

Through a review of the literature, this work conveys a substantial change in practice of situating deep learning in geotechnical engineering. Deep learning mitigates persistent limitations with soil, such as soil heterogeneity and non-linear behavior. Traditional approaches (i.e., Mohr-Coulomb or finite-element evaluation) make deterministic assumptions about soil behavior; this review presents a different way of using data-driven artificial-intelligence (AI) to encapsulate the non-linear and interlaced soil parameters where this is not easily reducible to linear assumptions. This framework with the utilization of laboratory testing (e.g., triaxial shear) and in-situ measurement (e.g., cone penetration test log) provides a systems approach to fit laboratory tests with field variability, and is a valuable step toward scalable and adaptive action geotechnical solutions. However, this paper also underlines systemic barriers to adopting AI in geotechnics. The model's relatively poor performance in organic soils (RMSE = 53.6 kPa), as discussed in the paper, should be viewed as evidence for deeper issues surrounding contextual data scarcity in geotechnical engineering, and not a failing of the model's performance. Organic soils meet a complex situation of spatially inconsistent decomposition states and uncertain fibre content that exemplifies "small data," or few observation dilemmas, that exist within much of the geology discipline.

Deep learning technologies, on the other hand, have such a huge demand for computation that a solution has to be found in the form of a trade-off between the two aspects of the issue: innovation and practicality. In spite of the fact that hybrid models like dense layers and transformers have reached completely new heights in terms of accuracy, their prodigious consumption of resources is still opposed to the realities of production environments where, for example, very often, real-time decisions are made with the help of less resource-intensive technologies like Random Forests. Such a scenario calls for the implementation of a balanced strategy: one that focuses on top-notch algorithms and at the same time, facilitates access to computing solutions such as model quantization or federated learning that make AI more available in the less resourceful areas of the world.

Moreover, the technical metrics research has the potential to stimulate a reflection on the changing role of engineers in the era of AI-augmented workflow. The SHAP analysis which the model conducted ascribed the ability to isolate significant predictors such as CPT tip resistance, plasticity index, and does not replace but rather enriches engineering judgment by providing, in turn, statistics-backed insights for the refinement of empirical correlations. To give an example, the nonlinear relationship between moisture content and shear strength has been revealed by the model and could thus lead to the revision of the empirical charts that have been in use for decades in foundation design.

The convergence of geotechnics and AI requires multidisciplinary cooperation. Expansion of datasets to include fewer common soils (tropical laterites, permafrost, etc.) will be done through partnerships with worldwide institutions, similar to what Salman et al. (2024) and Karkush et al. (2020) were doing by combining SPT data with MATLAB to create bearing capacity maps for Baghdad and Basrah on a city scale [d, e].

In other words, this work does not just surpass an inconsequential performance measure; it catalyzes a total transformation of geotechnical practice. Viewing soil as a non-deterministic medium and instead as a dynamic, information-producing material, the research aligns with the global objective of sustainable infrastructure - a future where precision prevents overdesign, safety factors are calculated instead of assumed, and risks are addressed proactively. The future is not some rejection of the traditional practice, but a hybrid future where the predictive ability of AI and human skill cooperate to address the problems of contemporary construction...

4.1 Challenges in Integrating Heterogeneous Datasets

The integration of multiple datasets—that is, CPT logs, triaxial tests, and micrographic photos—is extremely challenging: Data Format and Scale Incompatibility: CPT logs contain sequential measurements of soil profile (i.e., tip resistance vs depth), while triaxial tests provide tabular data, which could be a set of strain-stress curves, and each set of micrographs provide either a 2D or 3D soil microstructure. Matching these datasets in a temporally and spatially coherent way requires demanding pre-processing (e.g., intensity normalization for CPT vs lab sample coordinates).

4.1.1 Temporal Resolution Discrepancy

The field data (CPT) rarely has the same timing as lab tests and, for instance, the soil moisture level at CPT might differ from the state in which the preserved sample is, thus introducing perhaps unwanted bias to the testing of a triaxial test. Computational Workload: Micrographic photos are high dimensional photo data that do not directly relate to the sequential CPT logs. Therefore, a hybrid architecture would be needed to bring the two datasets together, e.g., CNNs with Transformers; this requires more training costs and improved preventative measures against overfitting through an increase in data.

4.1.2 Interpretation Trade-Offs

While deep learning models recognize patterns exceptionally well, the explanation of what specifically, within the micrograph features (i.e., pore distribution) and CPT data contributes to model predictive power is difficult to deduce, reducing trust for an engineer. Data Scarcity: Micrographics datasets are incomplete and are not massive in terms of their frequency compared to CPT data, so micrographics might be limited in some lab samples.

4.2 Limitations of Image-Based Soil Characterization

The complete dependence on intricate micrographs to gain insights into soil mechanics is an issue that requires careful scrutiny.

- Artifact Sensitivity: Sample treatment, like drying and cutting, alters the microstructure and makes the images not reflective of the actual conditions. For example, drying clay samples in the air results in cracks due to shrinkage, which are not found in field soils [23].
- Non-Unique Interpretations: The images with similar pore structures may indicate different mechanical behaviors. A study by Chen et al. (2020) found that soils with similar SEM micrographs had 30% difference in shear energy because of variation in mineralogy [23].
- Computational Costs: The processing of large micrograph datasets, such as those having 10,000 × 10,000 pixels, requires a lot of resources and is therefore not suitable for real-time applications.

The above-mentioned problems indicate that micrographs are not the only source for making reliable predictions. They need to be supplemented by physics-based models, for example, the discrete element method, to validate the inferred relationships.

5. Conclusion

The research has achieved the creation and confirmation of a deep neural network (DNN) model that can predict key geotechnical properties, namely, soil shear strength and dry and saturated soil bearing capacity. It is intended to enhance the ability of soil behavior prediction that neighborhoods obtain through traditional methods, thus making construction practices safer and more efficient. The model was trained using a set of geotechnical samples of 721 specimens from Najaf, Iraq, and included the parameters such as moisture content, plasticity index, CPT resistance, and SPT blow counts.

The quantitative outcomes are indicative of the model's performance throughout. The DNN got a root mean square error (RMSE) of 44.2 kPa during shear strength prediction, thus being over 24.7% superior to the Random Forest benchmark RMSE of 58.7 kPa with a difference of 14.5 kPa. An R^2 value of 0.89 indicates that the model has a very strong explanatory power. The performance of the model varied with soil type; however, the best performance was with cohesive soils, the RMSE being 38.4 kPa for clay soils and 41.2 kPa for silts. This is a useful demonstration of the model's capability of capturing the complex nonlinear interactions between these materials. The case study managed to apply the model to a site mainly consisting of clay (Sample #255) and accurately predicted the shear strength while also matching the observed behavior in the field.

On the other hand, the study found out few serious issues interspersed in the good results. The quality of performance in organic soils dropped and RMSE rose to 53.6 kPa. One possible reason for this is the complex and variable nature of organic materials which is a challenge during data collection. In addition to that, the study indicates that the dataset has a certain bias stemming from the Iraqi alluvial soils which may restrict the global applicability to other geologies such as tropical laterites or permafrost. Such challenges are common for geotechnical AI community that confirms models' accuracy for local use but does not claim global applicability.

This study has significant real-world repercussions. The model provides better predictions of soil behavior, enabling engineers to devise and evaluate foundation designs using fewer materials, and subsequently resulting in a process that is not only safe and cost-effective but also environmentally friendly. This is the case with data-supported field expertise as the SHAP analysis uncovers the interpretative values and can also pinpoint the CPT tip resistance and plasticity index as the major features.

The future research must identify and work on the crucial areas so as to eventually see this study through to practical application in the field:

- Dataset Diversification: The dataset should be expanded to embrace a variety of soil types and different locations, particularly those prone to earthquakes and those with extreme temperatures, since the latter can make the model not only applicable, but also accurate predictive to problems such as liquefaction and freeze-thaw to a great extent.
- Multimodal Data of Integration: The future models should incorporate the joint use of various types of data, like CPT logs and SEM micrographs together with geophysical data, to gain a more complete picture of the soil property than currently exists in the field.
- Advanced Model Approaches: It will be of great importance in future work to apply a combination of models.

References

[1] Jeremiah JJ, Abbey J, Booth A, Kashyap A. Results of application of artificial neural networks in predicting geo-mechanical properties of stabilised clays: a review. Geotechnics. 2021;1(1):147-71. https://doi.org/10.3390/geotechnics1010008

[2] Tang L, Na S. Comparison of machine learning methods for ground settlement prediction with different tunneling datasets. J Rock Mech Geotech Eng. 2021;13(6):1274-89. https://doi.org/10.1016/j.jrmge.2021.08.006

[3] Ge Q, Li J, Lacasse S, Sun H, Liu Z. Data-augmented landslide displacement prediction using generative adversarial network. J Rock Mech Geotech Eng. 2024;16(10):4017-33. https://doi.org/10.1016/j.jrmge.2024.01.003

[4] Shao W, et al. The application of machine learning techniques in geotechnical engineering: a review and comparison. Mathematics. 2023;11(18):39-76. https://doi.org/10.3390/math11183976

[5] Shimim N, et al. Integrating satellite imagery and infield sensors for daily spatial plant evapotranspiration prediction: A machine learning-driven approach. In: 2024 Intermountain Engineering, Technology and Computing (IETC). 2024;162-7. https://doi.org/10.1109/IETC61393.2024.10564271

[6] Rane N, Choudhary S, Rane J. Transforming the Civil Engineering Sector with Generative Artificial Intelligence, such as ChatGPT or Bard. 2024. Preprint. https://doi.org/10.2139/ssrn.4681718

[7] Zhang M, Guo Y, Shen M, Zhou D, Chen S. Data- and experience-driven neural networks for long-term settlement prediction of tunnel. Tunn Undergr Space Technol. 2024;147:105669. https://doi.org/10.1016/j.tust.2024.105669

[8] Onososen O, Musonda I. Perceived benefits of automation and artificial intelligence in the AEC sector: An interpretive structural modeling approach. Front Built Environ. 2024;8:864814. https://doi.org/10.3389/fbuil.2022.864814

[9] Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. In: Advances in Neural Information Processing Systems. 2017;30.

[10] Kapoor R, Kumar A, Kumar A, Kumar A, Arora C. Artificial intelligence in civil engineering: An immersive view. In: Artificial Intelligence Applications for Sustainable Construction. Woodhead Publishing; 2024. p.1-74. https://doi.org/10.1016/B978-0-443-13191-2.00009-2

[11] Carranza-Torres C, Hormazabal E. Computational tools for the determination of factor of safety and location of the critical circular failure surface for slopes in Mohr-Coulomb dry ground. In: Proc 2018 Int Symp Slope Stability Open Pit Mining Civil Eng. 2018;21-4.

[12] Afradi A, Ebrahimabadi A, Hallajian T. Prediction of TBM penetration rate using fuzzy logic, particle swarm optimization and harmony search algorithm. Geotech Geol Eng. 2022. https://doi.org/10.1007/s10706-022-02310-7

[13] Cruz E. Soil-Structure Interaction: From Bearing Capacity to Tolerable Movement. PhD Dissertation. 2024.

[14] Zhou J, Shi Z, Du K, Qiu Y, Li B, Mitri S. Feasibility of random-forest approach for prediction of ground settlements induced by the construction of a shield-driven tunnel. Int J Geomech. 2017;17(6):04016129. https://doi.org/10.1061/(ASCE)GM.1943-5622.0000817

[15] Bai R, Shen F, Zhang Z. An integrated machine-learning model for soil category classification based on CPT. Multiscale Multidiscip Model Exp Design. 2024;7(3):2121-46. https://doi.org/10.1007/s41939-023-00324-z

[16] Saad A, Sheikh U, Alyasseri Z. An efficient layout index characters for automatic license plate recognition system based on the YOLO-v8 detector. In: 2024 IEEE 8th Int Conf Signal Image Process Appl (ICSIPA). 2024. https://doi.org/10.1109/ICSIPA62061.2024.10701017

[17] Alattar A, Altaher A, Saad A, Al-Jawahry H. Integrating Haar Wavelet Transform and image processing for accurate fire detection. In: 6th Iraqi Int Conf Eng Technol Appl (IICETA). 2023;559-63. https://doi.org/10.1109/IICETA57613.2023.10351458

[18] Elaraby A, Saad A, Karamti H, Alruwaili M. An optimized deep learning approach for robust image quality classification. Trait Signal. 2023;40(4):1573-9. https://doi.org/10.18280/ts.400425

[19] Ma Z, Zhang B. Deep learning based upscaling of geomechanical constitutive behavior for lithological heterogeneities. In: SPE Annual Technical Conference and Exhibition. 2023. https://doi.org/10.2118/214889-MS

[20] Zhang W-S, Ying Y, Li M, Yang R-H, Ji L, Liu M. Prediction of surface settlement around subway foundation pits using spatiotemporal characteristics and deep learning models. Comput Geotech. 2024;168:106149. https://doi.org/10.1016/j.compgeo.2024.106149

[21] Pramod K, Pijush S. Reliability-based load and resistance factor design of an energy pile with CPT data using machine learning techniques. Arab J Sci Eng. 2023. https://doi.org/10.1007/s13369-023-08506-0

[22] Zhao T, et al. Artificial intelligence for geoscience: Progress, challenges, and perspectives. The Innovation. 2024;5(5).

[23] Chen Z, Liu X, Yang J, Little E, Zhou Y. Deep learning-based method for SEM image segmentation in mineral characterization. Comput Geosci. 2020;138:104450. https://doi.org/10.1016/j.cageo.2020.104450

[24] Abdollahi A, Li D, Deng J, Amini A. Explainable AI-aided safety factor prediction of road embankments. Eng Appl Artif Intell. 2024;136:108854. https://doi.org/10.1016/j.engappai.2024.108854

[25] El Sakka M, Ivanovici M, Chaari L, Mothe J. A review of CNN applications in smart agriculture using multimodal data. Sensors. 2025;25(2):472. https://doi.org/10.3390/s25020472

[26] Afshar A, Nouri G, Ghazvineh S, Hosseini Lavassani H. Machine-learning applications in structural response prediction: a review. Pract Period Struct Design Constr. 2024;29(3):03124002. https://doi.org/10.1061/PPSCFX.SCENG-1292

[27] Kingma DP, Ba JL. Adam: a method for stochastic optimization. In: Int Conf Learning Representations (ICLR). 2014.

[28] Gong H, Su D, Zeng S, Chen X. Advancements in digital twin modeling for underground spaces and lightweight geometric modeling technologies. Autom Constr. 2024;165:105578. https://doi.org/10.1016/j.autcon.2024.105578

[29] Satipaldy B, Marzhan T, Zhenis U, Damira G. Geotechnology in the age of AI: The convergence of geotechnical data analytics and machine learning. Fusion Multidiscip Res. 2021;2(1):136-51. https://doi.org/10.63995/ZOAF3555

[30] Sabaa MR, Salman AD, Dakhil AJ, Jawad SI, Karkush MO, Athab A. Production thematic maps of bearing capacity of shallow foundation for Al-Basrah soil using SPT data and GIS. J Rehabil Civil Eng. 2023;11(4):77-90.

[31] Al-Mirza HA, Falih MR, Wattan S, Salman AD, Karkush M. Utilization of GIS to generate a digital geotechnical map of Baghdad City. IOP Conf Ser Earth Environ Sci. 2024;1374(1):012052. https://doi.org/10.1088/1755-1315/1374/1/012052

[32] Karkush MO, Sabaa MR, Salman AD, Al-Rumaithi A. Prediction of bearing capacity of driven piles for Basrah governorate using SPT and MATLAB. J Mech Behav Mater. 2022;31(1):39-51. https://doi.org/10.1515/jmbm-2022-0005

[33] Karkush MO, Ahmed MD, Sheikha AA, Al-Rumaithi A. Thematic maps for variation of bearing capacity of soil using SPTs and MATLAB. Geosciences. 2020;10(9):329. https://doi.org/10.3390/geosciences10090329

[34] Salman AD, Khadim SS, Karkush M, Al-Rumaithi A. Production of maps for bearing capacity of shallow foundation in Baghdad city using SPTs and MATLAB. IOP Conf Ser Earth Environ Sci. 2024;1374(1):012023. https://doi.org/10.1088/1755-1315/1374/1/012023