19th WORLD CONFERENCE OF THE ASSOCIATED RESEARCH CENTRES FOR THE URBAN UNDERGROUND SPACE
Belgrade, Serbia, November 4-7, 2025.

ACUUS BELGRADE 2025

Paper No: 5.7.76
DOI: 10.46793/ACUUS2025.5.7.76

APPLICATION OF IMPROVED MACHINE LEARNING METHODS TOWARD BETTER ACCURACIES IN PREDICTING TBM PERFORMANCE

Dansheng Yao [1], Mengqi Zhu [2], Hehua Zhu [3]

Abstract: Tunnel Boring Machines (TBMs) play a pivotal role in modern underground construction, especially for tunneling through complex geological environments. Accurate real-time classification of surrounding rock masses is critical for optimizing TBM operation, yet conventional methods relying on engineering labels often suffer from subjectivity and label noise. To address these challenges, we propose a novel semi-supervised teacher-student framework that integrates engineering expertise, class prototype learning, and a Graph Neural Network (GNN)-based self-distillation mechanism. The model leverages soft labels, generated from domain-specific knowledge, to incorporate label uncertainty while improving feature cohesion. By employing self-supervised learning and iterative refinement of class prototypes, the framework effectively minimizes label noise, enhances classification boundaries, and adapts to unseen data. Experimental results on a real-world TBM tunneling dataset demonstrate that the proposed method significantly improves classification performance, offering more reliable and robust predictions in the presence of noisy labels. This work lays the foundation for future advancements in AI-driven geological inference, providing enhanced safety, efficiency, and reliability for TBM tunneling projects.

Keywords: Tunnel Boring Machine, Rock Mass Classification, Label Noise Correction, Self-Distillation, Prototype Learning.

1. INTRODUCTION

Tunnel Boring Machines (TBMs) are critical equipment widely used in modern underground engineering, particularly in tunnel construction. With their efficient and precise excavation capabilities, TBMs significantly enhance operational efficiency(Fu, Qiu, Xue, Shao, and Lan, 2024; Hansen, Erharter, and Marcher, 2024; M. Zhang, Ji, Zhou, Ding, and Wang, 2024, Gangrade et al., 2022). Ensuring the smooth operation of TBMs under complex geological conditions is essential, and this heavily relies on the accurate classification of surrounding rock masses. Accurate rock mass classification not only helps engineers assess the working environment of the TBM but also guides the adjustment of operational parameters, thereby optimizing excavation efficiency, reducing equipment failure, and mitigating risks during tunnel construction(Li, Tao, Du, and Wang, 2024; Qiu et al., 2022).

Traditional rock mass classification methods, such as the Rock Mass Rating (RMR), Barton’s Quality Index (BQ), and the Q-system, are typically based on geological surveys and rely on expert experience and onsite measurements. While these methods have been widely adopted in practical engineering, they often suffer from a high degree of subjectivity(Ambah and Elmo, 2024; Azad et al., 2024; Rehman et al., 2018). Their applicability is limited in complex and unstable geological environments where traditional approaches may fail to accurately reflect the true characteristics of the rock mass, consequently affecting TBM excavation decisions and operational efficiency(Yang, Mitelman, Elmo, and Stead, 2021).

With the development of machine learning and artificial intelligence (AI) technologies, data-driven approaches have been increasingly introduced into rock mass classification, offering promising solutions to address the limitations of traditional methods. AI-based approaches leverage big data analytics and pattern recognition to

[1] MSc, Dansheng, Yao, Intelligent Science and Technology, Shanghai Research Institute for Intelligent Autonomous Systems, Tongji University, Shanghai 200092, China , sorano@tongji.edu.cn
[2] PhD, Mengqi, Zhu, Geotechnical and Underground Engineering, Department of Geotechnical Engineering, College of Civil Engineering, Tongji University, Shanghai 200092, China, mqzhu@tongji.edu.cn
[3] Prof., Hehua Zhu, Geotechnical and Underground Engineering, Department of Geotechnical Engineering, College of Civil Engineering, Tongji University, Shanghai 200092, China, zhuhehua@tongji.edu.cn

efficiently process large-scale geological data, overcoming some of the key challenges of conventional techniques (Mooney, M. A. et al., 2023). However, the application of AI methods in real-world scenarios still faces significant challenges, particularly the issue of label noise, which is especially pronounced when the labels are derived from TBM operational data.

Most current AI-based methods for rock mass classification rely on supervised learning, where models are trained on large quantities of labeled data to minimize the error between predicted and true labels, thereby improving classification accuracy(Adeli, 2001; Jiang and Zhang, 2020; C. Zhang, Tao, Wang, and Fan, 2024). While supervised learning methods have achieved impressive results on standardized datasets, their robustness diminishes when applied to real-world engineering data, particularly in the presence of noisy labels. In practice, rock mass grading labels are often based on geological survey results or subjective judgments made by engineers during TBM operations. These labels are prone to measurement errors, environmental changes, and inconsistencies in engineering expertise, resulting in label noise. Such noise can negatively impact the accuracy of model training, leading to overfitting or underfitting, and ultimately limiting the generalization ability of the model under complex geological conditions.

To address the challenge of label noise, researchers have explored unsupervised and semi-supervised learning approaches. Unsupervised learning methods, which rely on the intrinsic structure of the data for classification, completely eliminate dependence on labels, thus avoiding the negative effects of label noise. For example, several studies(Xue et al., 2023; Q. Zhang, Liu, and Tan, 2019) have demonstrated the potential of clustering algorithms and dimensionality reduction techniques in analyzing TBM data, offering valuable insights into overcoming the limitations of traditional supervised learning approaches. However, unsupervised methods often fail to leverage existing labeled data, which frequently contains substantial domain knowledge and geological guidelines that accurately reflect the fundamental properties of rock masses. As a result, while unsupervised learning excels at reducing label dependence, it cannot fully utilize the engineering value embedded in labeled data.

Semi-supervised learning, on the other hand, combines a small amount of labeled data with a large amount of unlabeled data to enhance model learning through techniques such as pseudo-labeling and consistency regularization. Significant progress has been made in this area. For instance, several studies(Honggan Yu et al., 2021; Hongjie Yu and Mooney, 2023) have successfully applied semi-supervised learning methods to classify complex geological data, achieving notable improvements in classification performance and laying a solid foundation for addressing the label noise problem. However, semi-supervised learning also has significant limitations in real-world engineering scenarios. TBM data often lacks a sufficient quantity of high-quality labeled data that can be considered “completely accurate,” as well as a large-scale unlabeled dataset. Consequently, traditional semi-supervised learning methods struggle to fully exploit the available data in such contexts. Furthermore, semi-supervised learning tends to be highly sensitive to noisy labels, which can result in an underestimation of the value of labeled data and limit classification performance.

To overcome these limitations, it is essential to design an approach that addresses the unique challenges of TBM data. Such an approach should integrate multiple strategies: leveraging domain knowledge to refine and improve label reliability, fully utilizing the information embedded in noisy labels without dismissing their engineering value, and extracting intrinsic data structures to reduce overreliance on imperfect labels. By combining these elements, we aim to maximize the utility of available labeled data, enabling more robust and accurate classification even in the presence of noisy labels.

This paper proposes an innovative approach to tackle these challenges, combining class prototype learning and self-supervised learning. The proposed model classifies data by computing class prototypes, refines label distributions using domain knowledge, and leverages self-supervised learning to extract the intrinsic structure of the data. This approach balances label knowledge with the inherent data distribution, providing a robust solution for noisy label correction and improved rock mass classification in TBM operations.

2. METHODOLOGY

In this section, we first define the problem and provide an overview of our approach. Then, we introduce the workflow of the proposed framework. Next, we describe each component of the framework, including the guidance generator, which incorporates domain-specific prior knowledge as guidance, the teacher-student structure, the feature iterators based on graph neural networks, along with the Semantic Level Graph (SLG) and Class Level Graph (CLG) modules, and finally, we detail our proposed loss function.

2.1. Problem Definition

In practical problems, it is often difficult to obtain samples from the true distribution D of a pair of random variables $(\mathrm{X}, \mathrm{Y}) \in \mathcal{X} \times \{1,2, \dots, K\}$, where X is the feature space, and K is the number of classes. Specifically, we are given a set of noisy labeled data, where the feature space $\mathcal{X}$ consists of time-series samples, and the labels Y are corrupted due to noise.

For time series data, let:

(1) $x_i \in \mathbb{R}^{n_{\text{feature}} \times L}$ represent the input feature vector of size nfeature and length L, where L is the time series length.

(2) $\tilde{\mathbf{y}}_i \in \mathbb{R}^K$ represent the noisy label vector for the i-th sample, where $\tilde{y}_i[c]$ is the probability that the sample x_i belongs to class c.

The noisy data $\{(x_i, \tilde{y}_i)\}_{i=1}^{n}$ are independently drawn from a noisy distribution $\widetilde{D}$ of samples (X, Y), where the labels $\tilde{y}_i$ are corrupted versions of the true labels yi. The objective of this work is to learn a robust model $f: \mathcal{X} \times \mathbb{R}^K \to \mathbb{R}^K$ that can both correct the label noise and make accurate predictions for the test samples. Rather than directly learning from noisy labels, we introduce class prototypes into the model's learning process. These class prototypes serve as the reference points in the feature space, which guide the model towards a more accurate understanding of the underlying class distributions. The model should ideally minimize the following objective, which now incorporates the learning of class prototypes pc for each class c:

$$\min_f \mathbb{E}_{(x_i, \tilde{y}_i) \sim \widetilde{D}} [\mathcal{L}(f(x_i), y_i)] \quad (1)$$

where $\mathcal{L}(\cdot,\cdot)$ is a loss function that quantifies the difference between the predicted label f(xi) and the true label yi.

In this setting, we assume that the model learns to approximate the true class distribution via the class prototypes, pc, which are learned iteratively through both supervised and self-supervised losses. The class prototype pc represents the most typical feature vector for class c, helping the model to refine its understanding of each class by minimizing the distance between the predicted feature representation $f(x_i)$ and the corresponding class prototype.

By learning these class prototypes through soft labels and optimizing the feature representations, we expect the model to improve its robustness to label noise. Furthermore, the iterative self-supervised updates allow the model to fine-tune these prototypes by minimizing the intra-class feature distances and maximizing the inter-class feature distances. This ultimately leads to a model that more accurately reflects the true distribution of the classes in the presence of noisy labels.

2.2. Framework Overview and Workflow

The proposed model addresses noisy labels and enhances generalization through a teacher-student framework, as illustrated in Fig. 1. The key idea is to learn robust class prototypes that guide the model toward accurate predictions, even under label noise.

During training, TBM operation data is fed into the backbone of both teacher and student models to extract feature embeddings and class probabilities. Simultaneously, hard labels from engineering records are input into the Guidance Generator to produce Leader Soft Labels (LSL)—a refined form of supervision that reflects label uncertainty and domain knowledge.

These soft labels help approximate true class distributions, guiding the learning of class prototypes. Feature representations are projected by the Feature Embedding Head, while the Classification Head produces class probabilities. These outputs, along with LSL, are used to iteratively refine class prototypes via a Self-Distillation Block, which aligns features within classes and separates those from different classes.

The teacher model further processes the LSL and feature embeddings to generate LSL Embeddings, serving as high-confidence targets to guide the student model. A composite loss function includes:

- **Leader Loss** for guiding the teacher via LSL,
- **CLG Loss** for supervising the student using teacher outputs,
- **SLG Loss** for encouraging better feature learning through semantic-level self-supervision.

Inspired by Xiao et al. (2024)(Xiao et al., 2024), this framework reduces the distance between features and their class prototypes, improving both robustness and accuracy.

We adopt the U-Time architecture (Perslev, Jensen, Darkner, Jennum, and Igel, 2019) as the backbone for feature extraction from multi-dimensional TBM time-series data. This fully convolutional network efficiently captures temporal patterns and outputs compact representations for downstream modeling.

At inference time, only the student model is used to make predictions based on the learned representations.

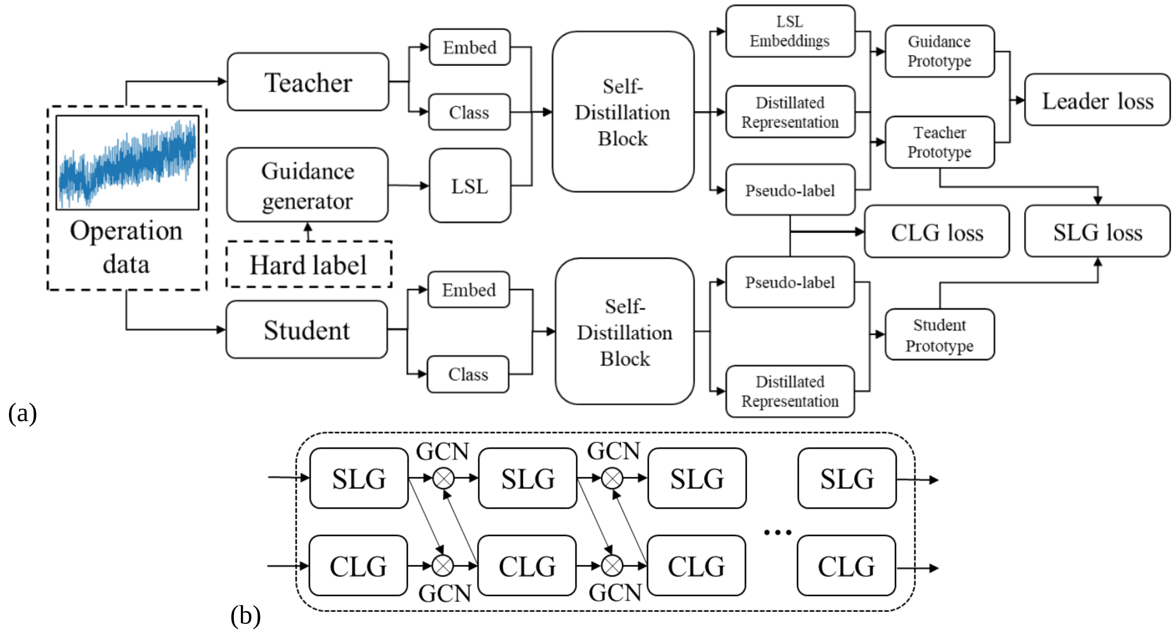

Figure 1. *Overview of the Model. (a) The Overview of the Framework. (b) The Detailed Structure of Self-Distillation Block.*

2.3. Guidance Generator

To mitigate label noise and integrate domain knowledge, we propose a Guidance Generator that assigns adaptive label confidence based on data stability. For stable segments with consistent labels (Figure 2), a high confidence of 0.7 is assigned to the labeled class, while the remaining 0.3 is uniformly distributed across others. For transition segments near label change points, mismatched labels receive a low confidence of 0.1, and matching labels are weighted using a Gaussian decay based on temporal proximity.

These confidence values are not arbitrarily chosen but are theoretically grounded. Specifically: (1) the 0.7 weight in stable regions reinforces consistent patterns while retaining residual uncertainty via uniform smoothing; and (2) in transition zones, the 0.1 mismatch penalty downweights unreliable annotations, while Gaussian weighting emphasizes temporal coherence. From an algorithmic standpoint, these values act as soft priors integrated into the loss, rather than hard constraints. Due to the smooth nature of the weighting mechanism and the regularization effect of the Gaussian decay, our method is empirically robust to minor variations in these values.

$$P(x) = \text{peak} \cdot e^{-\frac{(x - x_{\text{change}})^2}{2\sigma^2}} \tag{2}$$

This ensures minimal confidence at the change point and higher values as distance increases. The approach provides more reliable supervision, enhancing model robustness under uncertain labeling conditions.

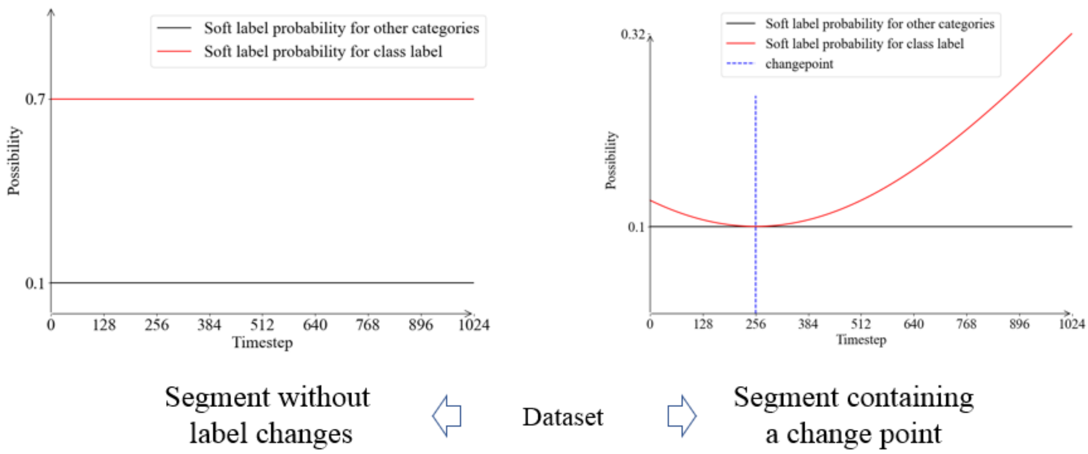

Figure 2. *Guidance Generator Model Overview for TBM Data Segmentation*

2.4. Graph-Based Representation Learning

To enhance feature learning under label noise and incorporate structural dependencies within the data, we construct two complementary graph structures: a Semantic-Level Graph (SLG) and a Class-Level Graph (CLG). The SLG captures local feature similarity between time steps, while the CLG models global consistency across predictions sharing the same class. Both graphs are constructed within each mini-batch and serve to propagate contextual information through neighborhood aggregation.

To further refine these representations, we introduce a mutual self-distillation mechanism between SLG and CLG. At each training iteration, the CLG is updated based on the semantically enriched features from the SLG, while the SLG is refined using class-level affinity derived from the CLG. This bidirectional interaction is modulated by dynamic confidence thresholds, which identify high- and low-confidence regions, allowing the model to balance reliance on confident predictions and contextual cues. After several iterations, we obtain enhanced feature embeddings from the SLG and refined class distributions from the CLG, which are used for pseudo-label generation and downstream supervision.

2.5. Loss Function

The overall training objective integrates three key components:

(1) Leader Loss, which aligns the teacher model's representations with soft labels provided by the Guidance Generator via prototype-level supervision.

(2) CLG Loss, which supervises the student model using confidence-weighted pseudo-labels derived from the CLG.

(3) SLG Loss, which enforces intra-class compactness and inter-class separability in the feature space, both at the batch level and across batches through global prototypes.

The combined loss encourages the model to learn discriminative, noise-resilient representations by jointly leveraging domain-informed soft supervision and structural consistency. This design facilitates robust training under uncertain labeling conditions and enhances generalization to unseen geological patterns.

3. EXPERIMENT

3.1. Project background

We evaluate our method using real-world operational data from the Yinchao water conveyance tunneling project in northeastern Inner Mongolia, China. The project employed an open-type gripper TBM (diameter: 5.2 m, thrust: 11,340 kN), which excavated a 55 km tunnel section over 509 days, with 4–10 cycles per day.

The dataset includes multi-dimensional time-series measurements (e.g., thrust, torque, penetration rate) recorded during excavation, along with corresponding geological labels derived from the Chinese Hydropower Classification (HC) system (Grades I–V). The distribution is dominated by Grade III rockmass (63%), followed by Grades II (19%), IV (14%), and V (4%).

3.2. Data preprocessing

To enhance training stability and label reliability, we applied a structured preprocessing pipeline following related studies (Xue et al., 2023; Zhu et al., 2022). TBM operational data were segmented into excavation cycles using thrust, torque, and net advance rate signals. Stable phases were then extracted from each cycle for subsequent analysis.

Geological records identified 74 rock mass transition points. Around each point, 2,294 samples were generated by symmetrically extracting 31 segments of 1,024 time steps from both sides. To balance the dataset, 2,287 additional samples were proportionally drawn from geologically stable segments across classes. Mixed-face conditions—where multiple lithologies coexisted at the excavation face—were not included, ensuring that each sample segment corresponds to a single dominant rock type.

The dataset was partitioned into training, validation, and testing subsets in a 60:20:20 ratio with consistent class distributions. While alternative splitting ratios were not exhaustively explored, the model design emphasizes robustness through confidence-weighted supervision and temporal context modeling. These properties reduce reliance on specific data partitions, and the theoretical formulation is not tightly coupled to a particular train/test ratio.

3.3. Results analysis

3.3.1. Evaluation and Visualization of Model Performance

Our model is designed to correct label noise rather than mimic it. However, successful correction presupposes the model's ability to extract meaningful geological patterns from noisy data. Therefore, a moderate alignment between model predictions and the original engineering labels is expected: overly high agreement may indicate overfitting to noise, while overly low agreement suggests failure to learn from the data.

To evaluate this balance, we first assess model performance using the original hard labels provided by the engineering team. These labels, although imperfect, serve as a practical supervision source. The model is trained on them and further refined through the proposed self-distillation process. Evaluation metrics—including Intersection over Union(IoU), Dice score, and overall accuracy—are reported in Table 1. As anticipated, the model shows limited alignment for Grades II–IV due to label and data noise, but achieves significantly better agreement for Grade V, which is known to be more reliably labeled and geologically distinct.

To further test generalization, we conduct an extended evaluation across the entire 55 km excavation range. The full dataset is segmented into non-overlapping 1024-step sequences, and a majority-vote scheme is applied within each segment to assign a stable label. This helps reduce prediction variance and improves interpretability.

Figure 3 presents the model's predictions across the full tunnel length. Despite noisy supervision, the model exhibits strong alignment with geological trends, accurately capturing major transitions while maintaining robustness to local fluctuations. This confirms that the model not only avoids overfitting to noise but also effectively learns from weak labels. Class-wise precision, recall, and F1 scores from this extended evaluation are summarized in Table 2.

Table 1. *Comparison between the proposed method and existing methods.*

	Grade II	Grade III	Grade IV	Grade V	Mean
IoU	0.4807	0.4156	0.4401	0.8360	0.5431
Dice	0.6493	0.5871	0.6112	0.9107	0.6896
Accuracy			0.6640		

Table 2. *Comparison between the proposed method and existing methods.*

	Precision	Recall	F1 score
Grade II	0.59	0.72	0.65
Grade III	0.86	0.73	0.79
Grade IV	0.62	0.75	0.67
Grade V	0.54	0.99	0.70

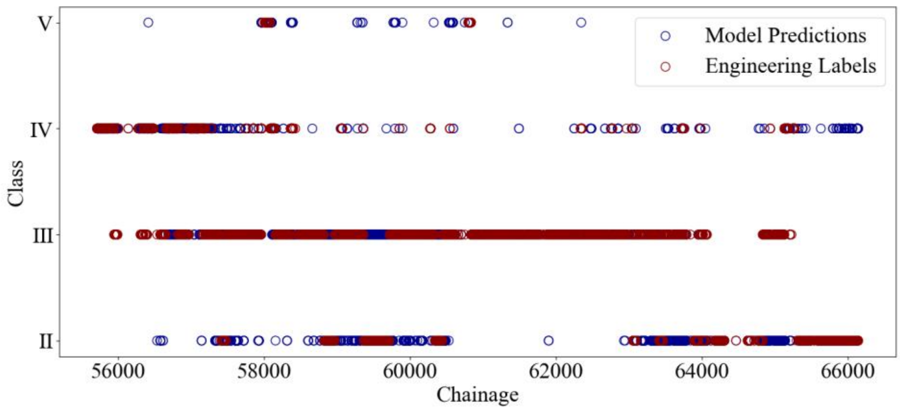

Figure 3. *Comprehensive Classification Evaluation Over Full Excavation Length*

3.3.2. Evaluation of TBM Penetration in Soft Rock Zones and Labeling Discrepancies

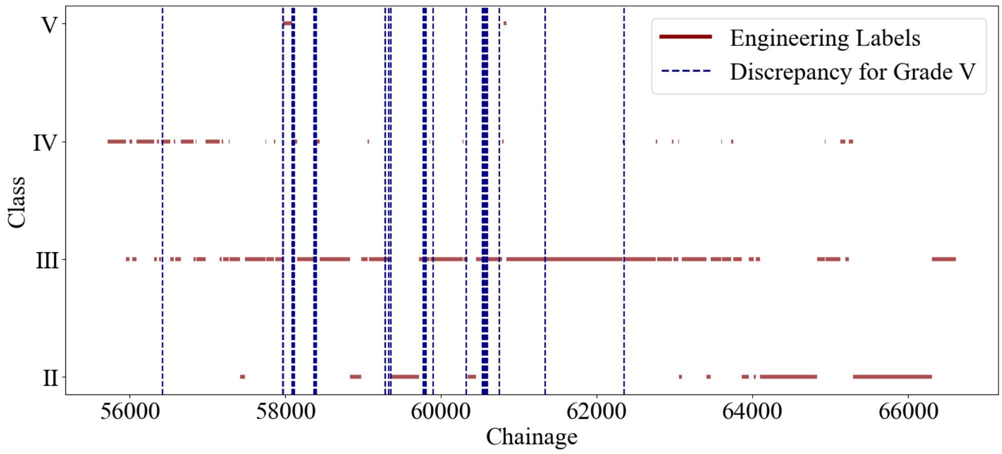

Figure 4. *Distribution of Label Discrepancies for Grade V across the Entire Chainage*

Accurate geological surveys are crucial for the safety, efficiency, and cost management of TBM tunneling, especially in soft rock areas where incorrect rock mass classification can lead to accidents and delays. The significant discrepancies observed between the proposed model and the labels provided by field engineers for Grade V rock masses, as shown in Figure 3, indicate the need for further investigation of the data related to Grade V. This is particularly important because accident-prone areas are often found in soft rock regions, and the classification of weak strata (Grade V) is critical to ensuring safety. To this end, we visualized the spatial locations where the model's predictions differ from the noisy labels for Grade V, as illustrated in Figure 4. These discrepancies are primarily observed at the boundaries between Grade IV and Grade V excavation sections, where excavation markers from the engineering team are subject to changes. This suggests that the model may interpret the feedback from the engineers as either premature or insufficiently timely in reflecting changes in rock mass classification.

Our model, however, has successfully identified all sections that were considered vulnerable by the engineering units, which were previously classified as Grade V. In addition to recognizing the segments that the engineers deemed weak, our model also discovered hidden vulnerable sections—areas that were not immediately identified in the continuous engineering assessments. This ability to detect previously overlooked weak strata significantly contributes to improving overall safety by highlighting additional risks that might have otherwise remained undetected.

Compared to traditional methods, our model demonstrates a higher level of accuracy in processing noisy data, effectively extracting geological features, and minimizing the interference of noise on classification results. By optimizing the identification of weak strata (such as Grade V), the model provides more reliable decision support in real-world engineering applications. In doing so, it not only enhances the safety and efficiency of tunneling operations but also reduces engineering risks, construction delays, and costs.

4. CONCLUSIONS

Accurate classification of surrounding rock masses is critical for optimizing excavation efficiency, ensuring equipment stability, and maintaining construction safety in Tunnel Boring Machine (TBM) operations. Traditional classification methods primarily rely on labels provided by engineering teams, which are often subject to human bias and noise. These issues reduce the robustness and adaptability of such methods under complex geological conditions. Addressing these challenges is essential for enhancing TBM operational efficiency and mitigating risks in underground construction.

This study proposes an innovative framework to address the challenges posed by noisy engineering labels in real-time rock mass classification during TBM operations. By generating domain-knowledge-driven Leader Soft Labels (LSL) and incorporating a Graph Neural Network (GNN)-based self-distillation mechanism, the proposed method achieves significant improvements in both classification accuracy and feature representation. Experimental results demonstrate that, compared to baseline models, the framework exhibits notable advantages in feature clustering, intra-class cohesion, and inter-class separation. These improvements are further validated through t-SNE visualizations. The findings confirm that the proposed model effectively mitigates the impact of label noise and achieves stable and accurate classification under complex geological conditions.

Furthermore, a comparison between the model predictions and noisy engineering labels reveals a delay in updating the classification of the weakest rock masses (Grade V). This observation provides critical insights for real-time geological assessments, highlighting the need to focus on dynamic adjustments for weak rock mass classifications. Accurate classification of weak rock masses is particularly important for ensuring TBM safety and efficiency. To address this issue, it is recommended to adopt the proposed model or similar noise-reduction frameworks when performing geological assessments. Such models can reduce the interference of noisy engineering labels, especially in weak rock mass regions where geological conditions are highly complex and subject to frequent changes. Accurate classification of weak rock masses enables timely adjustment of excavation parameters, such as cutterhead rotational speed, thrust force, and torque, thereby reducing equipment wear and construction risks.

Despite the significant progress achieved in this study, there remain several areas for further improvement. The current framework relies on relatively simple prior knowledge assumptions. Future work could incorporate richer engineering insights and statistical expertise to further optimize the model. For instance, adaptive adjustments to soft labels could be applied to specific spatial regions, such as drill sites or areas where engineers have high confidence in the surrounding rock conditions. Additionally, integrating spatial-temporal characteristics of geological data could enhance the model's robustness and generalizability in practical applications.

5. REFERENCES

[1] Adeli, H. (2001). Neural Networks in Civil Engineering: 1989–2000. Computer-Aided Civil and Infrastructure Engineering, 16(2), 126–142. https://doi.org/10.1111/0885-9507.00219

[2] Ambah, E., and Elmo, D. (2024, June 23). Is there a Universal Rock Mass Classification System? Presented at the 58th U.S. Rock Mechanics/Geomechanics Symposium. OnePetro. https://doi.org/10.56952/ARMA-2024-0784

[3] Azad, M. A., Najeh, T., Raina, A. K., Singh, N., Ansari, A., Ali, M., … Singh, S. K. (2024). Development of correlations between various engineering rockmass classification systems using railway tunnel data in Garhwal Himalaya, India. Scientific Reports, 14(1), 10716. https://doi.org/10.1038/s41598-024-60289-y

[4] Fu, K., Qiu, D., Xue, Y., Shao, T., and Lan, G. (2024). TBM tunneling strata automatic identification and working conditions decision support. Automation in Construction, 163, 105425. https://doi.org/10.1016/j.autcon.2024.105425

[5] Gangrade, R., Grasmick, J., Trainor-Guitton, W., & Mooney, M. (2022). Risk-based methodology to optimize geotechnical site investigations in tunnel projects. *Tunnelling and Underground Space Technology*, *127*, 104589.

[6] Hansen, T. F., Erharter, G. H., and Marcher, T. (2024). Towards reinforcement learning—Driven TBM cutter changing policies. Automation in Construction, 165, 105505. https://doi.org/10.1016/j.autcon.2024.105505

[7] Jiang, S., and Zhang, J. (2020). Real-time crack assessment using deep neural networks with wall-climbing unmanned aerial system. Computer-Aided Civil and Infrastructure Engineering, 35(6), 549–564. https://doi.org/10.1111/mice.12519

[8] Li, Z., Tao, Y., Du, Y., and Wang, X. (2024). Classification and Prediction of Rock Mass Boreability Based on Daily Advancement during TBM Tunneling. Buildings, 14(7), 1893. https://doi.org/10.3390/buildings14071893

[9] Mooney, M. A., Grasmick, J. G., & Gangrade, R. (2023). Methods for local big data integration to reduce geotechnical uncertainty and risk on subsurface infrastructure projects. In *Expanding Underground-Knowledge and Passion to Make a Positive Impact on the World* (pp. 2799-2806). CRC Press.

[10] Perslev, M., Jensen, M. H., Darkner, S., Jennum, P. J., and Igel, C. (2019). U-Time: A fully convolutional network for time series segmentation applied to sleep staging. In Proceedings of the 33rd International Conference on Neural Information Processing Systems (pp. 4415–4426). Red Hook, NY, USA: Curran Associates Inc.

[11] Qiu, D., Fu, K., Xue, Y., Tao, Y., Kong, F., and Bai, C. (2022). TBM Tunnel Surrounding Rock Classification Method and Real-Time Identification Model Based on Tunneling Performance. International Journal of Geomechanics, 22(6), 04022070. https://doi.org/10.1061/(ASCE)GM.1943-5622.0002379

[12] Rehman, H., Ali, W., Naji, A. M., Kim, J., Abdullah, R. A., and Yoo, H. (2018). Review of Rock-Mass Rating and Tunneling Quality Index Systems for Tunnel Design: Development, Refinement, Application and Limitation. Applied Sciences, 8(8), 1250. https://doi.org/10.3390/app8081250

[13] Wang, X., Zhu, H., Zhu, M., Zhang, L., and Ju, J. W. (2021). An integrated parameter prediction framework for intelligent TBM excavation in hard rock. Tunnelling and Underground Space Technology, 118, 104196. https://doi.org/10.1016/j.tust.2021.104196

[14] Xiao, H., Hong, Y., Dong, L., Yan, D., Zhuang, J., Xiong, J., … Peng, C. (2024). Multi-Level Label Correction by Distilling Proximate Patterns for Semi-supervised Semantic Segmentation. IEEE Transactions on Multimedia, 26, 8077–8087. https://doi.org/10.1109/TMM.2024.3374594

[15] Xue, Y.-D., Luo, W., Chen, L., Dong, H.-X., Shu, L.-S., and Zhao, L. (2023). An intelligent method for TBM surrounding rock classification based on time series segmentation of rock-machine interaction data. Tunnelling and Underground Space Technology, 140, 105317. https://doi.org/10.1016/j.tust.2023.105317

[16] Yang, B., Mitelman, A., Elmo, D., and Stead, D. (2021). Why the future of rock mass classification systems requires revisiting their empirical past. Quarterly Journal of Engineering Geology and Hydrogeology, 55(1), qjegh2021-039. https://doi.org/10.1144/qjegh2021-039

[17] Yu, Honggan, Tao, J., Qin, C., Xiao, D., Sun, H., and Liu, C. (2021). Rock mass type prediction for tunnel boring machine using a novel semi-supervised method. Measurement, 179, 109545. https://doi.org/10.1016/j.measurement.2021.109545

[18] Yu, Hongjie, and Mooney, M. (2023). Characterizing the as-encountered ground condition with tunnel boring machine data using semi-supervised learning. Computers and Geotechnics, 154, 105159. https://doi.org/10.1016/j.compgeo.2022.105159

[19] Zhang, C., Tao, M.-X., Wang, C., and Fan, J.-S. (2024). End-to-end generation of structural topology for complex architectural layouts with graph neural networks. Computer-Aided Civil and Infrastructure Engineering, 39(5), 756–775. https://doi.org/10.1111/mice.13098

[20] Zhang, M., Ji, A., Zhou, C., Ding, Y., and Wang, L. (2024). Real-time prediction of TBM penetration rates using a transformer-based ensemble deep learning model. Automation in Construction, 168, 105793. https://doi.org/10.1016/j.autcon.2024.105793

[21] Zhang, Q., Liu, Z., and Tan, J. (2019). Prediction of geological conditions for a tunnel boring machine using big operational data. Automation in Construction, 100, 73–83. https://doi.org/10.1016/j.autcon.2018.12.022

[22] Zhu, M., Zhu, H., Gutierrez, M., Ju, J. W., Zhuang, X., and Wu, W. (2022). Predicting Tunneling-Induced Ground Collapse Based on TBM Operational Data and Geological Data. International Journal of Computational Methods, 19(08), 2142015. https://doi.org/10.1142/S0219876221420159