Caving 2026 – A van As, D Cumming-Potvin & J Wesseloo (eds)
© 2026 Australian Centre for Geomechanics, Perth, ISBN 978-1-7636842-3-2
doi:10.36487/ACG_repo/2645_74

Machine-vision-based automation of geotechnical core logging: rock quality designation workflow and comparative assessment at Dund Sukhait

Jan Jaroling [a,*], Chris Sarrazin [a], Hannah Mulshaw [a], Adam Jackson [a,b], Gereltsetseg Tumurkhuu [c]

[a] *Tetra Tech, Canada*

[b] *Czech Geological Survey, Czech Republic*

[c] *Altan Shar Tal LLC, Mongolia*

Abstract

Recent advances in artificial intelligence (AI) and machine learning (ML) are creating new opportunities to automate structural geological data collection in mining through machine-vision-based drillcore logging. These technologies offer improved speed, repeatability, and consistency compared to conventional manual logging. However, industry adoption remains limited, partly due to constraints in legacy data capture systems and inconsistencies in historical geological datasets.

This paper presents the application of an AI-based image analysis platform to structurally interpret drillcore from the Dund Sukhait project in Mongolia, a structurally complex deposit characterised by faulted and variably fractured ground conditions. Four drill holes were processed using automated structural logging workflows, and the resulting datasets were compared against manually logged records. The comparison focused on rock quality designation (RQD), fracture frequency (FF), and classification performance metrics.

Results demonstrate variable statistical correlation between AI-derived outputs and manual logging; however, automated interpretations remain geologically coherent and internally consistent at the domain scale. Interval-scale discrepancies occur both in highly fractured (low-RQD) zones and locally within competent rock, reflecting break-type discrimination and interval segmentation effects. Automated processing achieved an average logging rate of approximately one hour per kilometre of drillcore processed, representing a substantial improvement in throughput relative to conventional logging. While the manual datasets reflected a high level of professional expertise, reliance on spreadsheet-based data capture introduces vulnerabilities that limit reproducibility and integration with advanced analytical workflows.

The findings indicate that both interval-scale break discrimination challenges and traditional data management frameworks influence AI evaluation metrics. For projects such as Dund Sukhait, the adoption of structured databases and purpose-built digital logging systems with integrated quality assurance and quality control (QA/QC), combined with continued refinement of break-type classification algorithms, is essential to enable scalable and reliable ML-assisted structural interpretation.

Keywords: *machine learning, structural geology, drillcore logging, geotechnical data, rock quality designation*

1 Introduction

Structural and geotechnical drillcore logging remains a fundamental component of ground characterisation at mining projects. At the Dund Sukhait project in Mongolia, accurate recording of fracture intensity, rock quality designation (RQD), and structural features is essential for understanding ground conditions within

* Corresponding author. Email address: janek.jaroling@tetratech.com

faulted and variably fractured rock masses. Traditionally, this work relies on manual logging, which is time-intensive and inherently influenced by individual interpretation.

Recent developments in artificial intelligence (AI) and machine learning (ML), particularly machine-vision-based image analysis, provide an opportunity to automate elements of structural logging. While these tools offer potential improvements in speed and consistency, their effectiveness depends on the quality and structure of the underlying geological dataset, which in many operations remains constrained by legacy data capture practices. The workflow builds on earlier comparative investigations between human and imagery-based geotechnical logging (Jaroling et al. 2023).

2 Project background

The Dund Sukhait project in Mongolia provides a practical case study for evaluating both the opportunities and limitations of AI-based machine vision core logging. The project contains an established archive of historical drillhole data and is characterised by structurally complex and variably fractured ground, where manual logging of recovered core is time-intensive and subject to interpretive variability. As part of recent geotechnical investigations supporting ongoing project development, fault classifications and key geotechnical parameters, including rock quality designation (RQD) (Deere et al. 2023), were extracted from 7 boreholes completed in 2025. However, only 4 of these boreholes were used in this study based on image and logging completeness suitable for comparative analysis (Table 1).

Table 1 Summary of boreholes completed during the 2023 drilling program

Borehole ID	Borehole coordinates (mine grid)			Borehole orientation		Final length (m-ah [2])	Core recovery (%)
	Easting	Northing	Collar elevation (m-asl [1])	Azimuth (°)	Inclination (°)		
DSGM01	344,599	4,794,862	1,569	220	65	267.50	99.1
DSGM02	344,862	4,794,905	1,572	180	60	469.50	98.9
DSGM03	345,014	4,794,697	1,571	40	70	327.50	99.3
DSGM04	345,477	4,794,809	1,586	220	60	452.00	99.3

[1] m-asl = metres above sea level, [2] m-ah = metres along hole.

This logging program created an opportunity to directly compare automated machine-vision-derived structural datasets with conventional manual logging results. The comparison was undertaken to assess consistency, accuracy, and practical applicability, and to determine whether AI- and ML-based workflows could be reliably integrated into routine geotechnical data acquisition at Dund Sukhait.

3 Site geology

The Dund Sukhait exploration licence is located approximately 35 km southwest of Bayandalai Soum in South Gobi Aimag, Mongolia, along the southern slopes of the eastern Gobi Altai Range. Regionally, the project area lies along the boundary between the Zoolen Terrane to the south and the Gurvansaikhan Terrane to the north. This terrane boundary trends west-northwest–east-southeast and is interpreted as a dextral shear zone, expressed in the western sector by strongly deformed Carboniferous granitic rocks.

The Gurvansaikhan Terrane comprises a structurally complex assemblage of imbricated thrust sheets representing a Lower Palaeozoic juvenile island-arc system. It consists predominantly of mafic volcanic rocks and oceanic sedimentary sequences that have been metamorphosed under greenschist facies conditions. In contrast, the Zoolen Terrane represents a subduction–accretion complex characterised by deep-water sedimentary sequences with ophiolitic fragments forming a tectonic mélange.

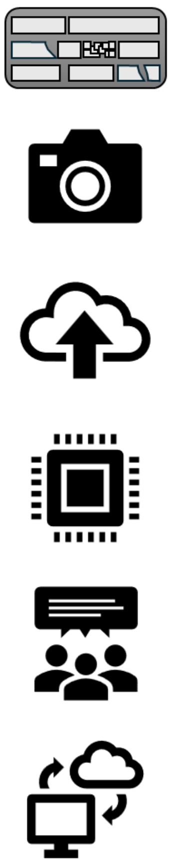
4 Geotechnical background

This structural framework, defined by folded bedding, and intersection fault systems, results in pronounced directional dependence of discontinuity persistence, spacing, and shear strength. Bedding planes on the deposit represent limbs of the open folds; to the south-southwest, steeply dipping planes dominate over relatively flat north-northeast dipping limbs. Cleavage surfaces and axial fold planes commonly form subparallel pervasive discontinuity sets trending east-northeast–west-southwest, while subparallel localised shear zones and intrusive contacts further compartmentalise the rock mass. These structures are segmented by the northwest and northeast trending fault system. These structural orientations are expected to influence block geometry, potential failure mechanisms, and excavation response, particularly where slope or excavation faces are oriented oblique or subparallel to the dominant fabric.

Diamond HQ size core drilling was completed using a YJ 900 drill rig (standard double-tube configuration) to obtain high-quality core (63.5 mm core diameter). The core was logged directly in wooden core boxes on site without orientation. Geotechnical parameters, including RQD, FF, and structural attributes, were recorded to support conventional rock mass assessment. ML geotechnical features, including fracture frequency, RQD estimation, joint validation, feature depth, and colour clustering, were subsequently extracted using automated machine vision analysis.

5 Methodology

For this study, the lithoLens™ machine vision core analysis platform was used for centralised image storage and standardised image preprocessing (including colour correction, white balance adjustment, glare reduction, and noise filtering), followed by machine-learning-based identification and counting of fractures and broken ground. A summary of a generalised workflow is given in Figure 1, with additional details below, which can be adapted by workers towards their specific site needs and unique software requirements or access.



1. Core Recovery
   - Drill core in trays/boxes is aligned.
   - Core block and depth marking QaQc.
2. Core Imagery
   - Photography or core scanning under controlled lighting.
   - Capture images before any manual logging or markup.
3. Image Upload
   - Upload images to a central image management system.
   - Apply standard naming and metadata for consistency.
4. Automated Image Processing
   - Run computer vision and machine learning analysis: e.g., fracture detection and counting
   - Classification of measured features: e.g., fracture types (natural vs induced; weak vs strong brecciation, etc..)
   - Structural orientation data extraction: e.g., alpha angles
5. Review & Validation
   - Geologists review outputs for accuracy.
   - QC checks are applied and corrections made where needed, re-training ML models until outputs match target precision and accuracy (“human-in-the-loop”)
6. Data Integration
   - Export validated data into geological logging systems or databases.
   - Use for interpretation, 3D modelling, reporting, and planning.

Figure 1 Core logging workflow summary

5.1 Core recovery, image capture, and upload

After the core was recovered by drillers, the core was aligned in boxes and inspected by geotechnical staff. This ensures that depth blocks and core box labels reflect accurate details and avoid misrepresentation of extracted logs at later stages. Manual logging included detailed structural descriptors; however, the dataset does not contain a dedicated field explicitly distinguishing natural discontinuities from mechanically induced breaks. For the purposes of the confusion matrix evaluation, natural joints (J) and mechanical joints (MJ) were grouped into a single ‘joint present’ category, and classification comparison was performed against this combined structural class. High-resolution images were captured using conventional digital cameras (e.g. SONY DSC-RX100M3 [by AST], parallel with an Infinix GT 20 Pro mobile phone [by SRK]). Consistent lighting is preferred to reduce image white balance and brightness variation; however, photos from Dund Sukhait were taken under outdoor conditions and sometimes direct sunlight resulted in localised glare and reflection on the core.

Both dry and wet images of the core were taken; the wet photograph was selected for analysis because it better reveals core characteristics. Image capture was done before any manual handling or logging took place, to minimise damage to the core. Once photographed, images were uploaded to the centralised cloud system with metadata, enabling fast, organised access for automated processing. The online platform allows the borehole ID and core depths to be added as the image files were not previously named (Figure 2).

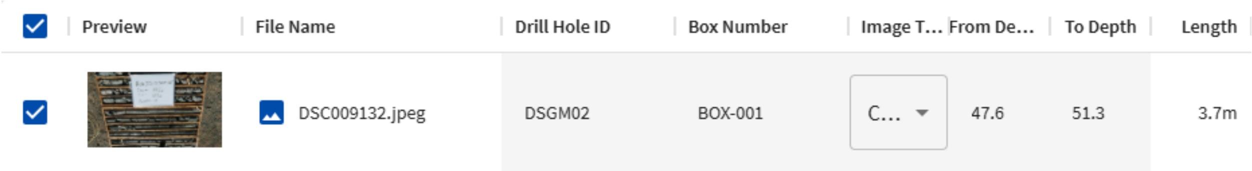


Figure 2 Image upload and manual input of borehole ID and core depth details

Prior to upload to the LithoLens platform, all core images were batch-processed using a Python-based compression workflow to ensure file sizes remained below the 10 MB upload limit. The script reduced image resolution and optimised the JPEG compression parameters while preserving sufficient detail for structural feature detection. This preprocessing step ensured platform compatibility without materially affecting fracture visibility or the quality of RQD interpretation.

5.2 Automated image processing

Prior to training the ML models, a standardised preprocessing and quality assurance procedure was applied to all core images. This process ensured that the training data accurately represented the geological material of interest while minimising artifacts that could degrade model performance or introduce bias.

The next step of data preparation involved ‘cleaning’ the images, and masking them to remove any content that is not the core itself (example given in Figure 3).

Figure 3 Example of a cleaned core image using masking techniques. All masked areas of the image are highlighted in green; only the specified geological material remains

Where core loss exceeded 0.1 m, a standardised placeholder image, known as a ‘dummy’ photo, was inserted to account for core loss and preserve continuity within the depth-registered dataset (Figure 4). The dummy photo represents a rock type with known characteristics (poor rock quality) and provides a form of synthetic image data analogous to the use of ‘rule-of-thumb’ values in traditional manual logging, which are often determined on a site-specific basis.

All images with >0.1 m of core loss were identified, and the depths of core and core loss were recorded. The image was then deleted and re-uploaded alongside the dummy photo the necessary number of times with the new adjusted depth ranges to account for all lost core.

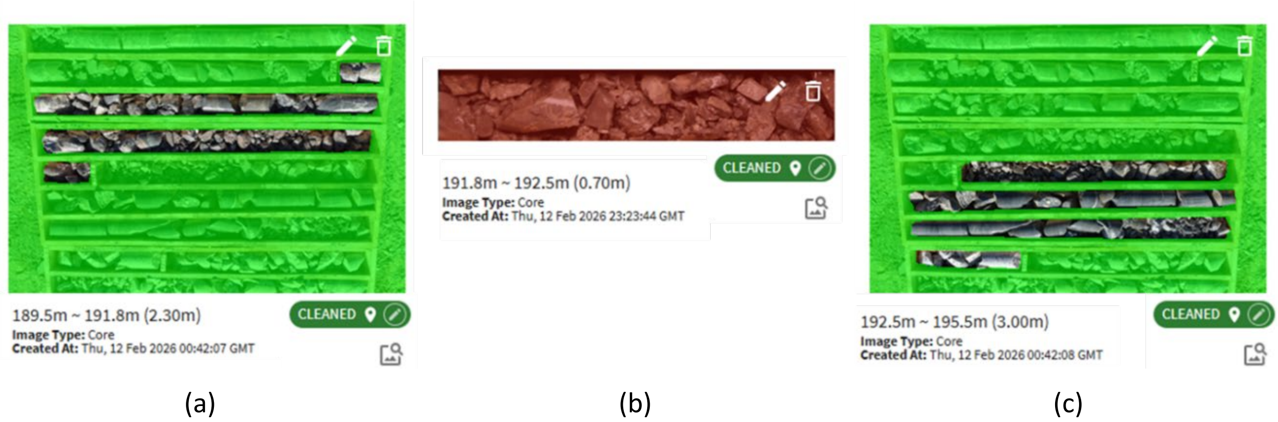


Figure 4 Core loss accounting example using ’dummy’ photos in DSGM01. (a) Core masking and depths edited to account for core loss; (b) ‘Dummy’ photo; (c) Subsequent masked core after core loss

To confirm the ML-derived masking, a manual check is performed to ensure that only relevant geological material was included in the data. False positives may arise from algorithmic misclassification of external objects, such as dark shadows or dark-coloured items near the core box, as core material. This visual inspection confirms that all components have been correctly masked and ensures that all intervals of the core are visible. All unmasked intervals of the core must also align with the depths labelled on the image.

If false positives are observed, corrective actions include manually masking to exclude non-core elements. For example, in the image platform used in this study, a polygon selection tool was used to revise inaccurate portions of automatic masking to correctly delineate the core fragments.

To ensure all quality assurance and quality control (QA/QC) was completed successfully as part of the image preprocessing, the relevant boreholes should be visually inspected. All borehole intervals were examined to inspect for any missing core and to confirm that all dummy photos were inserted correctly.

Following image preprocessing, the machine learning process can be initiated. In this study, ML-generated logging outputs included: colour, qualitative degree of rock quality, fracture frequency, RQD, and semi-quantitative fault classification (Table 2).

Table 2 LithoLens machine learning outputs

Category	Description	Attributes
Standard colour report	Core colour	Core is visualised in blocks of colour
RoQE Net	Rock quality estimation	Rock quality is estimated continuously along the length of the core using machine vision analysis, and is displayed as a bar or line graph Intact: no visible fractures or joints Joint: natural fractures or breaks present Rubble zone: natural broken rock fragments of random size (classified A–D) Mechanical joint: fractures due to mechanical processes (e.g. drilling, handling)
RQD	Rock quality designation	Rock quality at every depth. Displayed as a bar or line graph
Fracture frequency	Number of fractures	Fracture count at every depth. Displayed as a bar or line graph
FC Net	Fault classification	Class 0 = no fault present Class 1 = most significant fault Class 2 = moderate fault Class 3 = least significant fault

The ML processing was selected to proceed in the order above, because some outputs are used to inform others (e.g. fracture counts and lengths of intact rock inform RQD calculations). In our study, a moving window of 0.2 m was used for machine vision algorithms, and RQD used a fixed interval of 3.00 m (approximate run length of drill rods).

5.3 Review and validation

After processing, ML outputs were visualised downhole as a vertical graphic log overlain on the borehole image, for both single holes and multiple holes together. This effort was conducted at multiple scales, as the 0.2 m scale for the image processing can create some local noise (e.g. colour variations as short intervals) which are not significant when examining the overall averages of larger run lengths (e.g. over metres or tens of metres). When a single borehole was examined, multiple ML outputs were visualised together to assess for anticipated correspondence between outputs, e.g. that FF would correspond with the RQD rating. When multiple boreholes were selected, one output was compared at a time across boreholes to assess for geological continuity.

An example of validation is shown in Figure 5. Each variation of a result is shown: a successful, valid log ('true') can be subdivided as accurate in detecting a geological feature ('true positive') or avoiding the generation of a spurious log ('true negative'). Conversely, errors fall into categories of false positive and false negative.

Each validation result is defined as:

- True positive (TP): a natural or mechanical joint is present in the drillcore, and the automated system correctly classifies it as a natural/mechanical joint.
- False positive (FP): no natural or mechanical joint is present, but the system incorrectly classifies a feature (e.g. drilling-induced break or handling fracture) as a natural joint. Or, a natural/mechanical joint is present, but the system incorrectly classifies the feature (e.g. a mechanical break is recognised as a natural joint).
- True negative (TN): no natural/mechanical joint is present, and the system correctly classifies the interval as intact rock.
- False negative (FN): a natural/mechanical joint is present in the core, but the system fails to detect or classify it as a joint.

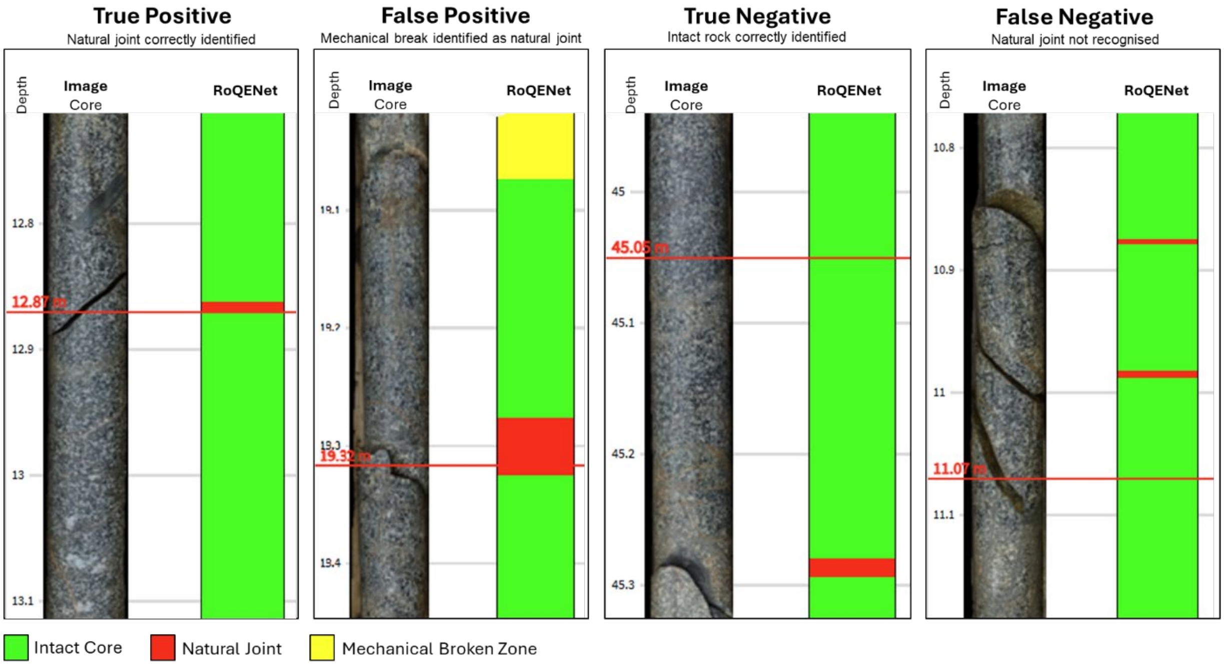


Figure 5 Core images and RoQE Net (rock quality estimation) results in LithoLens, demonstrating the 4 types of validation examples

Finally, the required data files were exported as .csv files. For each output type (e.g. RQD), the data for all the boreholes were exported to a single file for use in modelling and statistical analysis software.

5.4 Performance metrics

Model performance was evaluated using standard binary classification metrics derived from the confusion matrix, including accuracy, precision, recall, F1 score, TP, TN, FP, and FN. Accuracy was calculated as the proportion of correctly classified intervals (TP + TN) relative to the total number of intervals. Precision was defined as TP / (TP + FP), representing the reliability of positive predictions. Recall (sensitivity) was calculated as TP / (TP + FN), indicating the proportion of actual positives correctly identified. The false positive rate (FPR) was calculated as FP / (FP + TN).

6 Results

Automated structural logging was completed for boreholes DSGM01–DSGM04 using the machine vision workflow described in the methodology. Model performance is evaluated at 2 scales: structural classification using confusion matrix metrics, and RQD comparison versus depth. Automated outputs included fracture frequency, RQD, qualitative rock quality estimation, a machine-learning-based rock mass quality classification algorithm from core imagery RoQE Net, and colour cluster analysis. Quantitative results were compared against the client's manually logged dataset to assess classification consistency.

Confusion matrix results Figure 6 show consistent behaviour across bore holes. True negatives represent the largest proportion of the classification (45–55%), indicating reliable identification of intact intervals. False positives remain substantial (~39–44%), reflecting the model's sensitivity to minor discontinuities and mechanical breaks. False negatives are elevated in DSGM03 (13.11%), corresponding to its lower average RQD and higher fracture intensity.

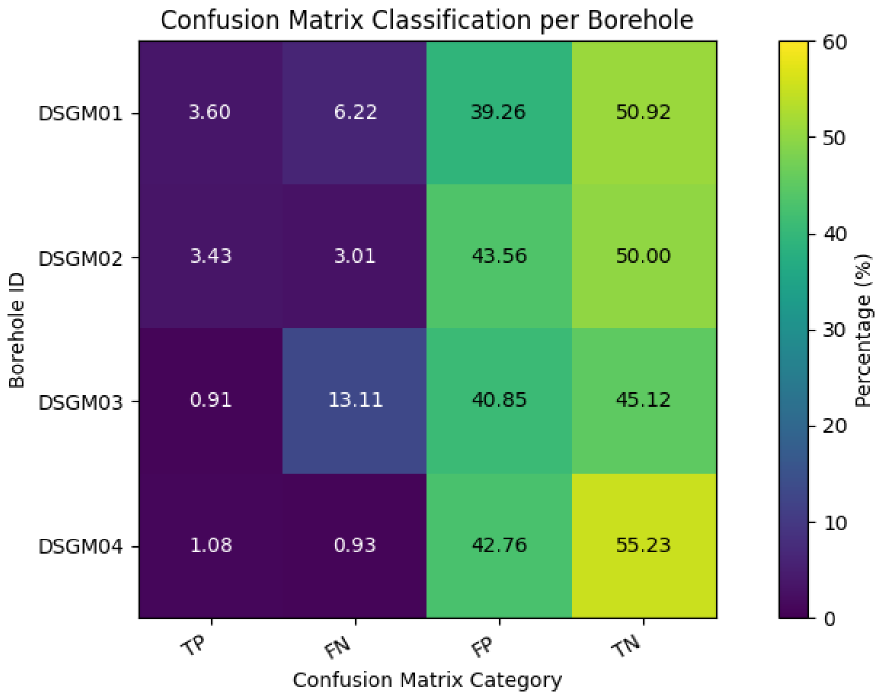


Figure 6 Drillhole-scale confusion matrix (%) showing true positive (TP), false negative (FN), false positive (FP), and true negative (TN) classifications for AI-based structural logging relative to manual logging

When results from all boreholes are combined in Figure 7, true negatives account for 52% of classifications and false positives for 42%, while true positives and false negatives represent smaller proportions (3% and 2%, respectively). These discrepancies reflect small-scale break identification differences, not a failure to recognise major structural domains.

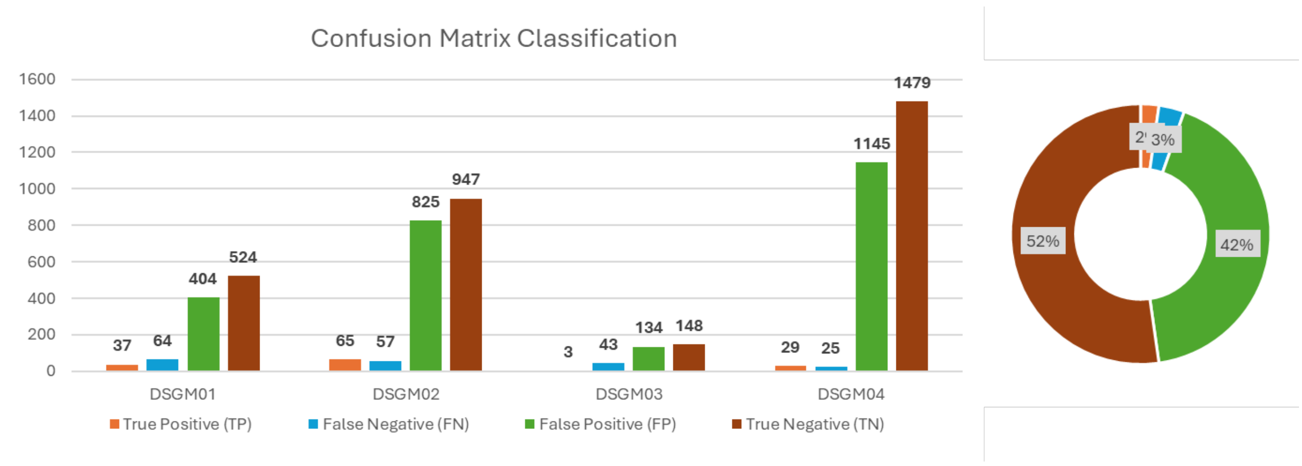


Figure 7 Overall confusion matrix summary (counts and %) for DSGM01–DSGM04

The downhole RQD comparison (Figure 8) demonstrates that RoQE-Net-derived RQD reproduces domain-scale trends observed in the client dataset. Agreement is strongest in low-RQD rubble zones, where both methods consistently identify highly fractured intervals. However, localised divergence is observed in both structurally complex (low-RQD) zones and in competent rock, where LithoLens RQD occasionally exceeds client values by approximately 10–20%. These differences suggest that interval segmentation and break-type filtering influence RQD calculation at small scales.

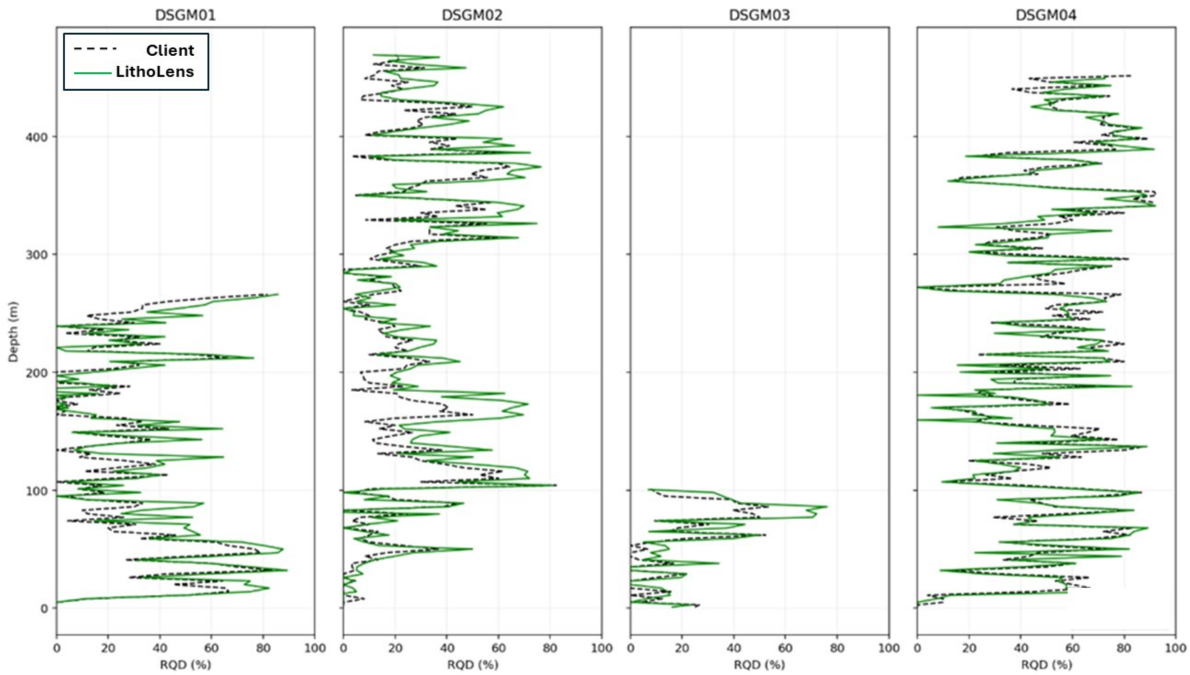


Figure 8 Client and LithoLens RoQE Net rock quality designation comparison for DSGM01–DSGM04

Performance metrics Figure 9 confirm stable overall accuracy (0.46–0.56), with variability in precision and recall between boreholes. DSGM03 exhibits reduced recall consistent with its higher fracture intensity, while DSGM04 demonstrates comparatively stable classification performance.

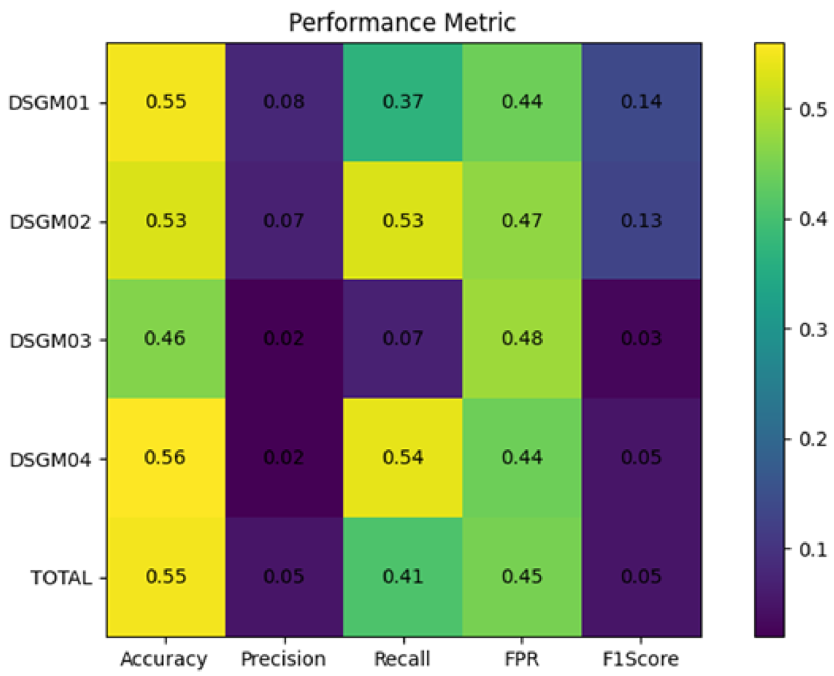


Figure 9 Performance metrics (accuracy, precision, recall, false positive rate (FPR), and F1-score)

7 Discussion

The results indicate that AI-derived structural logging remains stable at the domain scale, consistently identifying intact and highly fractured intervals. However, interval-scale discrepancies occur in both low-RQD and high-RQD zones.

In highly fractured ground, closely spaced discontinuities and complex break geometries increase ambiguity in distinguishing natural from mechanical breaks, contributing to elevated false negative rates. Similar interval-scale ambiguity in image-based structural classification has been reported in previous studies (Seifert et al. 2024; Owen et al. 2024).

In competent rock, localised upward deviations in LithoLens-derived RQD suggest that break filtering and segmentation logic affect the treatment of short intact fragments. These deviations are not indicative of systematic model instability but rather reflect sensitivity to interval-scale classification thresholds.

Geological variability, therefore, influences performance at small scales, while domain-scale structural trends remain coherent between AI and manual datasets. This behaviour is consistent with prior AI-based structural logging investigations (Seifert et al. 2024), where performance variability was shown to correlate with fracture density and dataset structure.

Overall, discrepancies arise from a combination of geological complexity, challenges in break-type discrimination, and interval segmentation effects. The workflow significantly improves logging efficiency while maintaining reproducible structural characterisation at project scale. Continued refinement of break-type identification and segmentation logic represents the primary path to further improvement.

8 Conclusion

Application of the machine learning workflow at Dund Sukhait demonstrates consistent reproduction of domain-scale RQD and structural trends across boreholes. While interval-scale discrepancies occur in both highly fractured and locally competent rock, overall model behaviour remains stable and geologically coherent. Classification differences arise primarily from break-type discrimination and segmentation sensitivity rather than random algorithm instability. These findings support the use of AI-assisted structural logging as a complementary tool to traditional methods, particularly when implemented within standardised classification frameworks and supported by ongoing algorithm refinement.

Acknowledgement

The authors would like to thank SRK Consulting for facilitating coordination with the client and supporting access to the project data used in this study. Appreciation is extended to LAMA Investments Mongolia for permitting the use of drilling and logging datasets, and to the Czech Geological Survey for their technical collaboration and data support. The authors also acknowledge Tetra Tech for enabling the preparation of this manuscript and supporting the continued development of AI-assisted geotechnical workflows. Finally, the authors thank the Australian Centre for Geomechanics for the opportunity to present and discuss this work within the broader geotechnical community.

References

Deere, DU & Deere, DW 1988, 'The rock quality designation (RQD) index in practice', in L Kirkdale (ed.), *Rock Classification Systems for Engineering Purposes*, American Society for Testing and Materials, Philadelphia, pp. 91–101.

Jaroling, J, Shalmani, S & Hood, S 2023, 'Machine versus human – imagery geotechnical data collection', *Proceedings of GeoConvection 2023*, Canadian Energy Geoscience Association, Calgary, pp. 1–2

Owen, G, Longridge, L, Jaroling, J, Shalmani, SM & Kamal, D 2024, 'Automating the identification and classification of faults from drill core imagery using machine learning', in *Proceedings of Slope Stability 2024*, Brazilian Geotechnical Society, Belo Horizonte.

Seifert, N, Owen, G, Longridge, L, Quartey, L, Jaroling, J & Shalmani, S 2024, 'A comparison of machine learning vs manual geotechnical logging for RQD and fracture frequency', in *Proceedings of Slope Stability 2024*, Brazilian Society of Geotechnical Engineering, Belo Horizonte.