DOI: https://doi.org/10.68732/2444137dec69a4d9

ISBN: 978-9989-2053-7-8

Reproducible geotechnical metrics from AI-assisted core logging

L. S. L. Tan[1], M. Rinne[1], M. Janiszewski[1] & I. Avramov[2]
[1]*Aalto University, Espoo, Finland*
[2]*Boliden Zinkgruvan, Zinkgruvan, Sweden*
tansunlie8@hotmail.com
0009-0008-1250-6849

Abstract

Geotechnical core logging provides crucial input for rock engineering, including mine design, stability assessment, and rock-mass classification. Yet the process remains prone to human error and lacks reproducibility. Core box photographs, widely collected in the industry, represent a largely underutilised primary data source for recomputing geotechnical metrics. This study investigates the practical benefits of applying artificial intelligence (AI) to such images, focusing on reproducibility and enabling multiple rock-mass characterisation workflows from metric-agnostic outputs.

An AI-based instance segmentation method (Segment Anything Model; SAM), combined with context-specific filtering, was applied to an image dataset from the Boliden Zinkgruvan mine, Sweden.The workflow automatically extracts the intact core pieces in each image, digitally measures their lengths, and produces a fracture-position database. Natural and mechanical fractures are then differentiated using markings produced by a geologist, enabling computation of rock quality designation (RQD). The results are compared with manual measurements.

The workflow achieved a mean absolute difference between AI-computed and manually measured RQD of 7.8 percentage points. Outlier analysis suggested that large discrepancies between manual and computed RQD may be related to human error, indicating that AI-assisted methods can support quality assurance/quality control (QA/QC) of manual logs.

The results demonstrate that recording fracture positions rather than interval-based metrics reported over fixed lengths (e.g., RQD every 1 m) enables transparent recalculation of RQD and supports the derivation of additional metrics such as fracture frequency (P10). This highlights how AI-assisted fracture extraction can increase reproducibility in core logging and produce design-independent data that can be transformed into multiple geotechnical metrics and used across different analysis frameworks (empirical methods, numerical simulations, DFNs).

The findings confirm that AI-assisted core logging enables reproducible downstream metric computation and design-independent rock-mass metrics from standard core photographs, aiding QA/QC of manual logs and use across different analysis frameworks.

Keywords

Geotechnical core logging, Core box images, Rock quality designation (RQD), Artificial intelligence.

1 Introduction

Manual geotechnical core logging provides essential input for mine design, stability assessment, and operational decision-making. Despite its central role, the process is prone to error and lacks reproducibility. Differences in conventions over time, as well as human errors, can lead to values that are difficult to trace once entered into the database and may subsequently contribute to design errors.

Rock quality designation (RQD; Deere et al. 1967) is defined as the percentage of intact core pieces longer than 10 cm over a given interval, considering only natural fractures. RQD is one of the most widely used empirical geotechnical metrics and forms a key input parameter in established rock mass classification systems such as the Q-system(Barton et al. 1974) and the rock mass rating (RMR; Bieniawski 1973) system. Although RQD is fundamentally based on physical measurements of intact core lengths, it is typically recorded as a single percentage value over defined intervals, making subsequent validation difficult. This also makes it impossible to derive other fracture-spacing metrics, such as 1D fracture intensity (P10; Dershowitz and Herda 1992), without relogging the core.

Core box images represent a largely underutilized data source in this context. Such images are routinely captured at most mining operations, primarily for archival purposes, yet they contain detailed geometric information that can be processed digitally.

Instance segmentation is particularly suitable for this application, as it enables pixel-level separation of individual intact core pieces, preserving object-level geometric information. This instance-level representation provides a direct basis for computing fracture-spacing metrics. Previous studies applying instance segmentation methods, primarily based on Mask R-CNN architectures(He et al. 2017), have demonstrated the feasibility of extracting RQD from core box images (Tims et al. 2024; Zhang et al. 2024). However, these approaches have typically relied on laborious datasets to retrain generic models on core box images, raising concerns regarding accessibility, domain adaptation and generalizability. Moreover, the potential of such methods as quality control/quality assurance of RQD, and for broader geotechnical applications has received limited attention.

This study proposes an instance segmentation workflow that operates without retraining. Applied to core box images, the workflow produces metric-agnostic fracture position data from which RQD can be reproducibly recomputed and compared against manually logged values, enabling systematic error analysis.

2 Methodology

The dataset comprises digital photographs of core boxes acquired during routine logging at Boliden Zinkgruvan, an underground base metal mine in Örebro County, Sweden. Images were taken after completion of logging and entry of rock quality designation (RQD) values into the drillhole database, which stores geological and geotechnical data by user-defined depth intervals. In this study, RQD was logged using geotechnical intervals rather than fixed 1 m intervals. Images were acquired using a digital camera mounted in a fixed metal frame to ensure consistent position and viewing angle (Fig.1a).

For the subset of images used in this study, natural fractures were marked directly on the core using blue cross markings on either side of each fracture prior to wetting and photography (Fig.1b). RQD values stored in the drillhole database are referred to as manual RQD and serve as the reference for comparison with the AI-derived computed RQD.

Computed RQD was obtained through a multi-stage workflow comprising 2D image processing followed by 1D reconstruction. In the 2D stage, images were first corrected for perspective. The Segment Anything Model (SAM; Kirillov et al., 2023) was then applied to generate instance segmentation masks, i.e., pixel regions corresponding to visually distinct objects in the core box image. Context-specific filtering retained only masks corresponding to intact core pieces and depth tags while resolving overlaps between candidate masks and removing artefacts. For each accepted mask, centreline length was measured and optical character recognition (OCR) was applied to depth tags to extract numerical values.

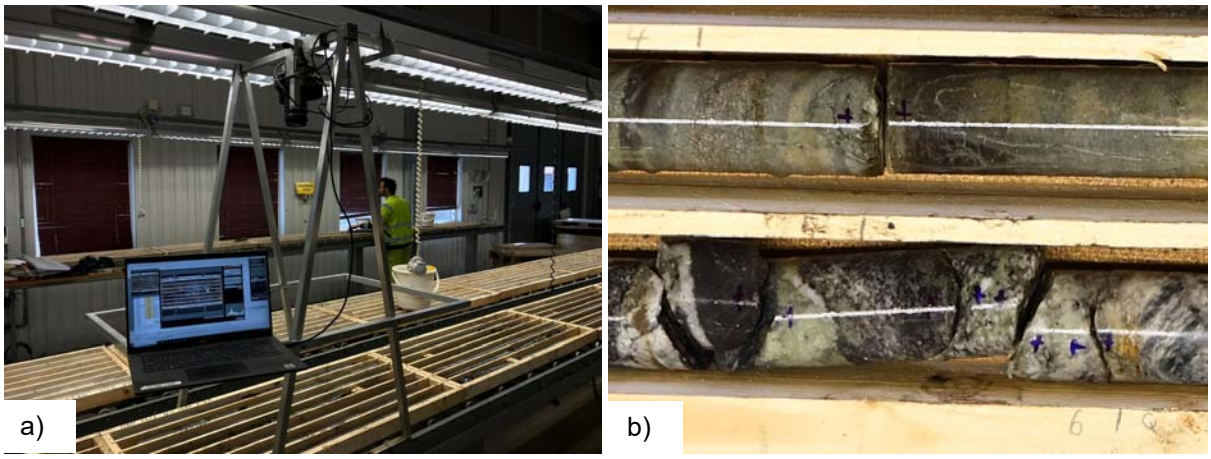


Fig.1 (a) Fixed-frame camera setup ensuring consistent position and viewing angle. (b) Natural fractures marked using blue crosses on either side of each fracture.

The first step of the 2D processing stage was perspective correction. All images correspond to a single core box type with fixed geometry, enabling conversion from pixel-based measurements to physical lengths. Corner points of the box were manually selected in each image and mapped to true proportions using a homography, resulting in images with consistent scale and geometric proportions that contain only the core box region (Fig. 2a). All subsequent image processing steps were performed automatically without additional manual input.

Following geometric correction, instance segmentation was performed using the Segment Anything Model (SAM) with grid-based sampling, employing the ViT-H backbone. As this produces a large number of overlapping candidate masks, context-specific filtering criteria were applied to retain only valid core piece segments and depth tags. The filtering relies on a sequence of heuristic geometric constraints tailored to the template-like structure of core box images, in which cylindrical core pieces are arranged in ordered rows. Filtering is performed in three stages: (i) individual geometric constraints, which enforce basic plausibility of each mask based on continuity and row confinement; (ii) relational geometric constraints, which resolve overlap between masks by identifying near-duplicates, reflections, and masks spanning multiple core pieces; and (iii) targeted removal of artefacts, such as masks of wooden panels, small chips, and background texture. Together, these steps progressively reduce the candidate masks to a non-overlapping set of accepted masks that consistently correspond to physical core pieces and depth tags, forming a reliable basis for subsequent digital measurements (Fig. 2b, c).All subsequent processing steps operate exclusively on the accepted masks.

Depth tags, which appear as cubes, were identified among the accepted masks based on their geometric proportions. Optical character recognition (OCR) was applied to these masks using Google Cloud Vision's document text detection feature (Google LLC 2026), and individual digit detections were algorithmically combined to form a single numerical depth value per tag. The resulting depth values were manually corrected prior to comparison with logged RQD. Corrections were made by visual inspection of the depth tag, with partially obscured digits inferred from visible portions and tag context. For each core piece mask, the centreline length was measured by multiplying the pixel length of the mask centreline by the image scale. The resulting computed length and depth measurements are illustrated in Fig. 2c. These centreline lengths, together with additional geometric attributes, OCR-derived depth information, mask geometries, and unique mask identifiers, were stored in a relational database, hereafter referred to as the segmentation database.

Using the geometric and OCR data stored in the segmentation database, the workflow reconstructs a one-dimensional (1D) representation of the core along the borehole axis. Masks containing OCR values were treated as depth tags and excluded from fracture representation. Core piece masks adjacent to a depth tag were assigned fixed start or end depths, while all masks between successive depth tags were rescaled to span the OCR-derived numerical depth interval. Where the numerical interval exceeded the summed centreline lengths of the core pieces, the remaining length was retained as core loss. This procedure yields a fracture position database in which all detected fractures, irrespective of origin, are assigned a unique depth along the borehole.

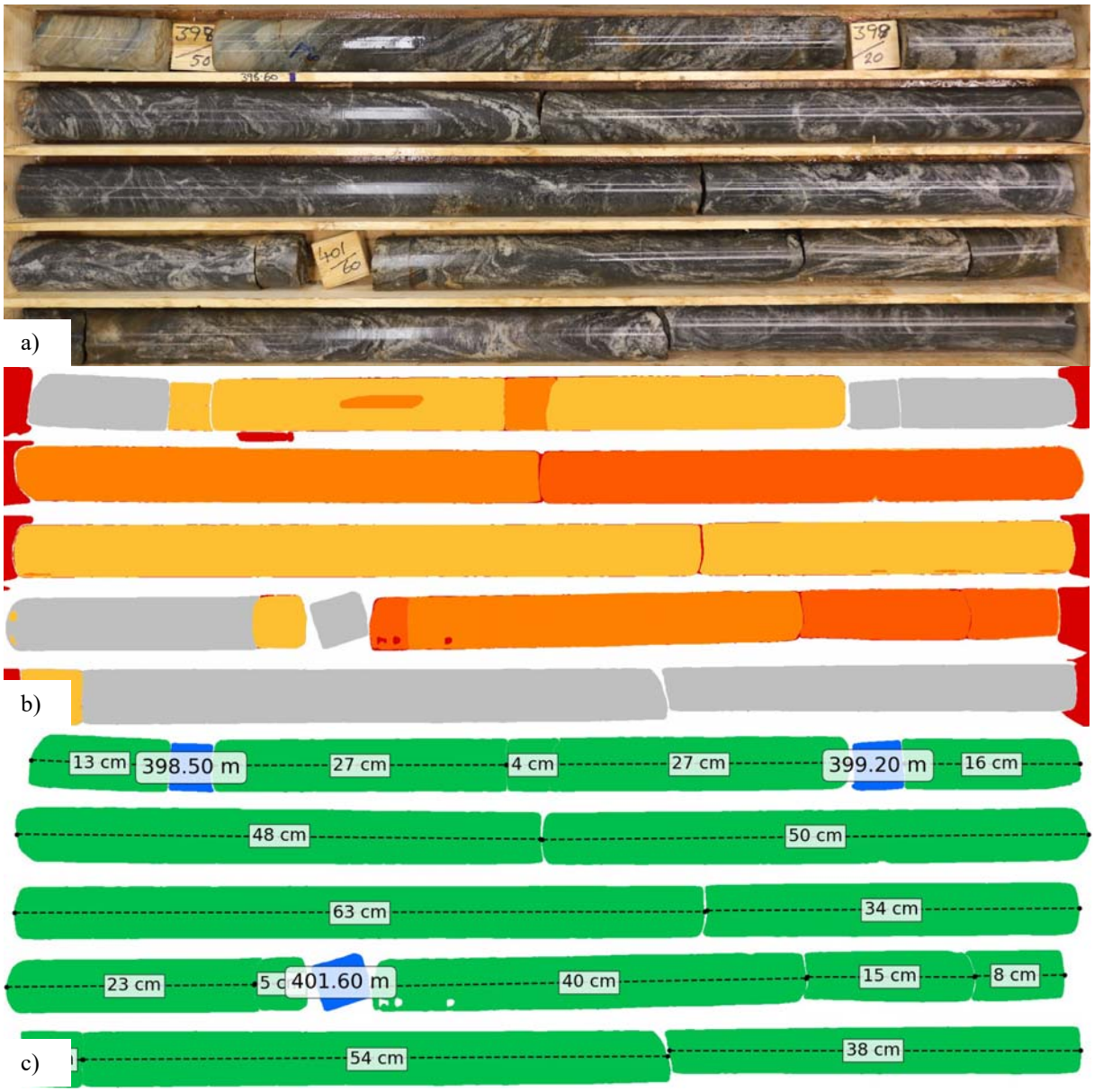


Fig. 2 (a) Perspective corrected core box image. (b) Candidate masks and the effect of filtering, showing removed detections (red), overlapping core piece masks (yellow to dark orange, indicating increasing numbers of overlapping pixels), and non-overlapping masks (grey). (c) Accepted masks showing core pieces (green) with centre lines and lengths (cm), and detected depth tags (blue) with OCR-derived depths (m).

The resulting fracture position database contains both mechanical (drilling- or handling-induced) and natural fractures. For RQD computation, only natural fractures were considered. These were identified using the cross markings in the images. Natural fracture positions were used to compute RQD within the same geotechnical intervals in which RQD was manually logged. During this calculation, RQD values were reduced proportionally to the amount of core loss within the interval.

3 Results

The comparison between manually logged and AI-computed RQD values per geotechnical logging interval (n = 105) is shown in Fig. 3. Interval lengths ranged from 0.2 m to 3.0 m, with an average length of 2.4 m. The model detected 86% of natural fractures. Of the 69 markers, 46 (67%) were correctly recognized, 10 (14%) were undetected, and 13 (19%) were misread.

The AI-computed RQD showed a mean absolute difference of 7.8 percentage points, a root-mean-square deviation of 18.9 percentage points, and a bias of +3.2 percentage points. In total, 73.3% of intervals were within ±5 percentage points of the logged RQD. Additionally, 7.6% of intervals exceeded ±20 percentage points deviation; these intervals averaged 0.4 m in length, compared to 2.6 m for the remaining intervals.

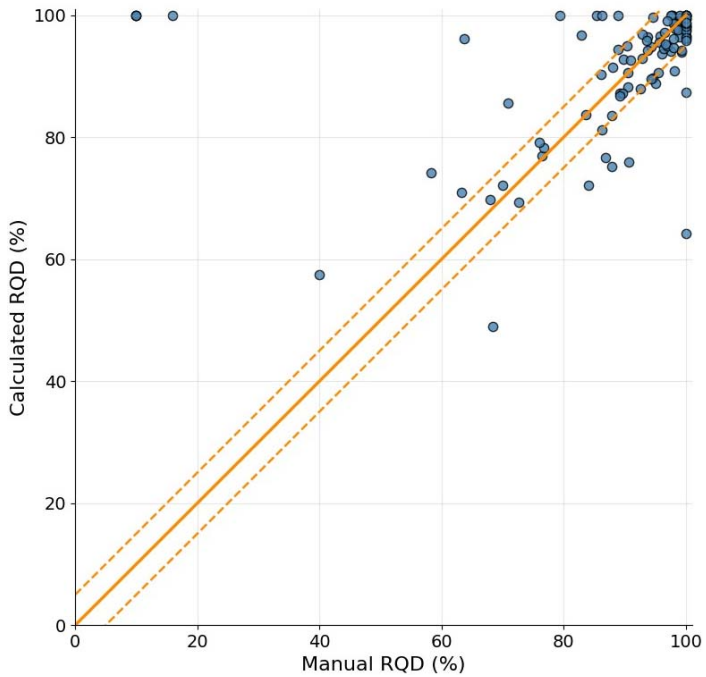


Fig. 3 Comparison of manually logged and AI-computed RQD values per geotechnical interval (n = 105). The solid line denotes the line of agreement, and the dashed lines indicate 5 percentage point deviation from agreement.

4 Discussion

The proposed workflow differs from previous instance segmentation approaches applied to core box images (Tims et al. 2024; Zhang et al. 2024) in that it does not require retraining a model. Instead, it combines zero-shot segmentation using SAM with context-specific filtering tailored to the structure of core box images. This reduces the barrier to implementation at operating mines.Direct quantitative comparison is not possible as Tims et al. (2024)report no quantitative accuracy metrics and Zhang et al. (2024)use a different dataset and methodology with no core loss accounting and unclear definitions regarding natural and mechanical fractures.

A mean absolute difference of 7.8 percentage points indicates overall agreement between computed and manually logged RQD within engineering tolerances. The mean bias of +3.2 percentage points is consistent with the segmentation performance of 86%, as undetected fractures result in higher RQD values. The high root-mean-square deviation reflects the influence of outliers, primarily associated with shorter intervals. The dataset comprised good quality rock, leading to clustering in the upper right of the comparison plot. Manually logged RQD was not treated as ground truth: both manual and AI-derived RQD contained errors. Tolerances (±0.2 m) in depth markings make it impossible to derive exact fracture positions from core boxes, preventing strict validation of either measurement. Comparison against televiewer data could provide a basis for directly comparing accuracy. Outlier analysis showed that the largest deviations were associated with manual errors (presumably data entry errors and misattribution of depth), as well as procedural factors not attributable to either method (e.g., core loss written on the box). This suggests that large RQD discrepancies may indicate intervals that warrant review, which could simplify manual QC and image inspection.

Most OCR errors were caused by poor image quality (incomplete marker visibility, poor lighting, and separation of numbers over multiple lines). These errors could likely be reduced through workflow improvement, and this step may be automated by replacing handwritten tags with standardized text. While the perspective correction relies on the manual selection of box corner points, this step can be automated retrospectively by training a detection algorithm on the rows of the core box, or procedurally by printing fiducial markers on the row corners. The workflow can directly be implemented at sites where mechanical breaks are marked in the images, or where all breaks are logged respective of origin. In the latter case, the reported RQD corresponds to pseudo-RQD, as defined by Tims et al. (2024), since values are also lowered by mechanical breaks. Under these conditions, the method may be deployed either as an ongoing quality assurance tool or retrospectively for quality control.At sites where mechanical fractures are not distinguished by markings on the core

prior to photography, implementation for quality assurance would require adjustments to the logging procedure. Retrospective quality control may be attempted by investigating the relationship between the occurrence of natural and artificial breaks.

As emphasised by Potvin et al. (2012), separating intrinsic data collection from design-dependent analysis is fundamental to geotechnical logging. Storing individual fracture positions rather than fixed-interval metrics preserves the primary geometric information, enabling derivation of RQD, P10, or other metrics at any required interval and ensuring that reported values can be traced back to detected fractures.

Testing across a wider range of rock qualities is required before general deployment. Additionally, engineering acceptance still relies on human judgment, and automation of this single parameter does not necessarily yield time savings.

5 Conclusion

This study demonstrates that zero-shot instance segmentation combined with context-specific filtering can extract fracture positions directly from standard core box photographs, without retraining or annotated datasets. Although the initial classification of natural fractures remains manual, storing fracture positions rather than interval-based RQD values enables fully reproducible recomputation of metrics and supports quality assurance and quality control (QA/QC) of manual logs.AI-computed RQD aligned with manually logged RQD within engineering tolerances for a dataset comprising predominantly good-quality rock, while identifying intervals that warrant review. The workflow is most applicable at sites where mechanical breaks are marked or where all visible breaks are logged. This represents a step toward more transparent, reproducible, and design-independent geotechnical data collection that can support multiple analysis frameworks, from empirical classification methods to numerical modelling.

References

Barton N, Lien R, Lunde J (1974) Engineering classification of rock masses for the design of tunnel support. Rock mechanics 6:189–236. https://doi.org/10.1007/BF01239496

Bieniawski ZT (1973) Engineering classification of jointed rock masses. Civil Engineering 1973:335–343

Deere DU, Hendron AJ, Patton FD, Cording EJ (1967) Design of surface and near-surface construction in rock. In: ARMA US rock mechanics/geomechanics symposium. ARMA, p ARMA-66

Dershowitz WS, Herda HH (1992) Interpretation of fracture spacing and intensity. In: ARMA US rock mechanics/geomechanics symposium. ARMA, p ARMA-92

Google LLC (2026) Optical character recognition (OCR). Google Cloud Vision documentation. https://docs.cloud.google.com/vision/docs/ocr. Accessed 18 Feb 2026

He K, Gkioxari G, Dollár P, Girshick R (2017) Mask R-CNN. In: 2017 IEEE International Conference on Computer Vision (ICCV). pp 2980–2988

Kirillov A, Mintun E, Ravi N, et al (2023) Segment Anything. https://doi.org/10.48550/arXiv.2304.02643

Potvin Y, Dight P, Wesseloo J (2012) Some pitfalls and misuses of rock mass classification systems for mine design. Journal of the South African Institute of Mining and Metallurgy 112:697–702

Tims S, Caté A, LeRiche A, Saunders E (2024) Automated geotechnical logging of core box images using machine learning. CIM Journal 1–10. https://doi.org/10.1080/19236026.2024.2385269

Zhang Y, Ma Y, Li Y, Wen L (2024) A Deep Learning Approach of RQD Analysis for Rock Core Images via Cascade Mask R-CNN-Based Model. Rock Mech Rock Eng. https://doi.org/10.1007/s00603-024-04100-4