Petroleum Science 23 (2026) 5624–5635

Contents lists available at ScienceDirect

Petroleum Science

journal homepage: www.keaipublishing.com/en/journals/petroleum-science

Original Paper

A knowledge-guided multi-task learning method for drill cuttings analysis

Li Hou [a,b], Xin Li [a,b], Wei-Zhi Tian [c], Qian-Xiao Su [a,b], Rui-Qi Huang [a,b], Ming-Hui Song [c], Peng Du [c], Yi-Li Ren [a,b,d,*]

[a] Research Institute of Petroleum Exploration & Development, PetroChina, Beijing, 100083, China
[b] Artificial Intelligence Technology R&D Center for Exploration and Development, CNPC, Beijing, 100083, China
[c] Mud Logging Technology Research and Development Center & GWDC Mud Logging Company, CNPC, Panjin, 124010, Liaoning, China
[d] Tsinghua University, Beijing, 100084, China

ARTICLE INFO

Article history:
Received 19 August 2025
Received in revised form
7 February 2026
Accepted 22 May 2026
Available online 27 May 2026

Edited by Jia-Jia Fei

Keywords:
Drill cuttings
Multi-task learning
Prior knowledge
Collaboration classifier

ABSTRACT

Drill cuttings serve as a fundamental source of geological information in oil and gas exploration and development. However, traditional analysis remains largely manual, resulting in low efficiency and high subjectivity. While recent studies have introduced artificial intelligence, most focus on single-attribute tasks and purely on data-driven approaches, often neglecting geological domain knowledge. This highly limits model generalization and interpretability in real-world applications. To address these challenges, we propose a multi-attribute intelligent identification model based on multi-task learning, which integrates geological prior knowledge and introduces a novel collaboration classifier to enhance the comprehensiveness, robustness, and interpretability of drill cuttings analysis. Specifically, the model constructs structured inter-attribute correlations as prior knowledge and employs a collaboration classifier designed based on the law of total probability, enabling collaborative modeling and cross-attribute information exchange. These innovations significantly improve recognition performance across multiple attribute dimensions. Experimental results demonstrate that the proposed model achieves a mean accuracy of 93.90% and a mean F1-score of 93.08% in identifying cuttings attributes including color, shape, and grain type. These results validate the effectiveness of the structured prior-guided collaborative modeling strategy and highlight its potential for complex multi-attribute analysis tasks in petroleum geology.

1. Introduction

Drill cuttings constitute the most accessible and continuous geological materials acquired during drilling, offering low-cost, real-time, and systematic information about subsurface formations (Costa et al., 2023; Kadyrov et al., 2022). The diverse geological properties preserved in cuttings provide comprehensive support for reconstructing stratigraphic, lithologic, and reservoir characteristics, establishing a core foundation for geological interpretation throughout the exploration and development (Yousef and Morozov, 2024; Kadyrov et al., 2022; Sanei et al., 2020; Salim and Lagraba P, 2018). Building on these attributes, cuttings serve as a vital bridge between direct geological observation and downhole measurements in practical operations, thereby enhancing the continuity and reliability of geological evaluation and engineering decision-making (Singh et al., 2023; Reyes et al., 2015; Guilherme et al., 2011; Hogg et al., 1996). As a near-continuous record along the wellbore, cuttings analysis also provides timely feedback for optimizing drilling parameters and ensuring operational safety (Ramba et al., 2021). Collectively, these attributes underscore the comprehensive geological and engineering value of drill cuttings, highlighting their indispensable role in oil and gas exploration and development.

However, despite their indispensable value, traditional approaches to cuttings analysis, including manual visual examination and thin-section petrography, remain constrained by low efficiency, limited resolution, and a strong reliance on expert experience. To address these shortcomings, techniques such as X-

* Corresponding author.
E-mail address: renyili@petrochina.com.cn (Y.-L. Ren).
Peer review under the responsibility of China University of Petroleum (Beijing).

https://doi.org/10.1016/j.petsci.2026.05.037
1995-8226/

ray diffraction (XRD), X-ray fluorescence (XRF), and natural gamma spectroscopy have been employed to improve accuracy and extract more information (Carugo et al., 2013; Marsala et al., 2011; Alixant and Collie, 1999). However, these laboratory-based methods are labor-intensive, rely heavily on operator expertise, and lack real-time capability, which limits their suitability for field operations and rapid decision-making. These constraints highlight the need for intelligent, automated cuttings analysis systems to overcome these limitations and enhance geological interpretation and operational efficiency. With advances in image processing, classical algorithms have been applied to segment particles and extract features from drill cuttings images to assist geological identification (Huo et al., 2021; Guo et al., 2020; Perez et al., 2015; Młynarczuk et al., 2013). While these methods offer certain improvements, they are highly sensitive to hyperparameter settings and prone to over- or under-segmentation in complex geological settings.

The rapid advancement of deep learning aligns well with the growing demand for intelligent drill cuttings analysis. Its strong capabilities in autonomous feature learning and efficient inference have driven the development of intelligent logging systems. Early studies combined deep learning with traditional image processing to improve the accuracy and reliability of cuttings analysis (Xu et al., 2022; Patel and Chatterjee, 2016; Singh et al., 2010), demonstrating its feasibility and practical value. However, these methods still rely heavily on parameter tuning, resulting in inherently unstable performance and poor robustness.

In response, researchers are increasingly exploring advanced methods for drill cuttings analysis. Existing studies employ CNN-based models to identify grain types in cuttings, effectively reducing manual workload and simplifying traditional workflows (Tolstaya et al., 2023; Kathrada and Adillah, 2019). Other work focuses on the automatic extraction of regions of interest, including individual particles and hydrocarbon-bearing areas (Yi et al., 2024; Huo et al., 2024; Zhang et al., 2022; Yamada and Di Santo, 2022). Segmentation networks trained on these target regions enable efficient recognition and quantitative characterization of cuttings features. More recently, large model techniques have been explored to improve robustness and analytical capability under complex and variable operational conditions (Shan et al., 2024).

Although the above studies have substantially advanced drill cuttings analysis, most existing studies remain focused on single-task learning (STL), such as particle segmentation or the identification of a specific attribute. However, drill cuttings analysis is fundamentally multi-dimensional, and its attributes are inherently correlated, requiring joint interpretation for reliable geological understanding (Costa et al., 2023). Consequently, STL models cannot fully capture these dependencies and are limited in their ability to support comprehensive analysis. These challenges highlight the need for multi-task learning (MTL) frameworks capable of modeling multiple attributes simultaneously and exploiting their inherent correlations (hereafter, “task” and “attribute” are used interchangeably depending on context). By sharing representations across tasks, MTL enhances generalization, improves performance, and reduces computational cost (Vandenhende et al., 2022). These advantages stem from the fact that related tasks can act as mutual regularizers, reinforcing one another to improve identification accuracy (Ruder, 2017).

In addition, most existing approaches rely predominantly on data-driven deep learning and lack effective integration of geological prior knowledge as structural constraints or guiding information. This limitation becomes especially pronounced when dealing with multiple, strongly correlated attributes, where purely data-driven models often struggle to capture domain-consistent relationships. Embedding domain-specific geological knowledge as a prior bias during learning therefore offers a promising way to enhance both scientific validity and practical applicability (Xu et al., 2025; Zuo et al., 2024).

To address these limitations, this study proposes a correlation prior and collaborative enhancement (CpCe) model for the intelligent identification of multiple attributes in drill cuttings analysis. CpCe integrates structured geological knowledge with MTL mechanism, introducing a correlation-prior graph that encodes domain-informed attribute relationships and a collaborative inference mechanism that supports cross-attribute reinforcement for joint multi-attribute prediction. This design allows the model to more effectively capture intrinsic geological constraints and inter-attribute interactions that conventional works often overlook. By combining prior-guided representation learning with cross-task interaction, CpCe enhances identification consistency, interpretability, and robustness across multiple attributes. Overall, this framework provides a solution for leveraging geological knowledge and intrinsic attribute dependencies to advance multi-attribute drill-cuttings analysis.

2. Data

2.1. Dataset construction

The data used in this study were collected from representative structural units of the Liaohe Depression in the Bohai Bay Basin, including the Taoyuan and Yulou-Rehetai structural belts in the Eastern Sag, the Niuju Structural Belt in the northern depression, the Fault-Nose Structural Belt on the southern Western Slope of the Western Sag, and the Shuguang low buried-hill belt in the central Western Slope. This cross-structure sampling strategy increases the geological diversity captured by the dataset and helps reduce potential locality-induced sampling bias. After drill cuttings are brought to the surface by drilling mud, they are cleaned, dried, flattened, and photographed under natural light. These images typically contain many particles. While prior studies have focused on segmenting and extracting individual particles (Huo et al., 2024; Shan et al., 2024; Kathrada and Adillah, 2019), our work centers on multi-attribute collaborative identification. Rather than developing a new segmentation method, we use the segment anything model (SAM) (Kirillov et al., 2023), a generalized zero-shot segmentation tool, to simplify preprocessing and improve segmentation efficiency.

In this study, three representative attributes, namely color, shape, and grain type (hereafter referred to as grain), are selected as learning targets for training and evaluation. After removing low-quality particles, the remaining samples are annotated by domain experts to construct a high-quality multi-attribute dataset. Labeling is based on direct visual examination of the particles. For pure minerals like quartz and feldspar, grain labels are assigned directly; others are labeled based on composition, following standard drill cuttings logging practices. Color labels include gray, grayish black, grayish white, grayish green, grayish brown, and red. Shape labels include angular, blocky, and rounded. Grain labels include mudstone, sandstone, basalt, granite, feldspar, dolomite, quartz, and diabase. After curation, 3935 annotated images were obtained. Data distribution is shown in Fig. 1. The dataset sufficiently covers the common and geologically reasonable attribute combinations, while the missing combinations are extremely rare in nature and therefore do not compromise the representational completeness of the dataset or introduce sampling bias.

All images are uniformly expanded to 128×128 pixels for standardized input. The dataset is split into training and test sets

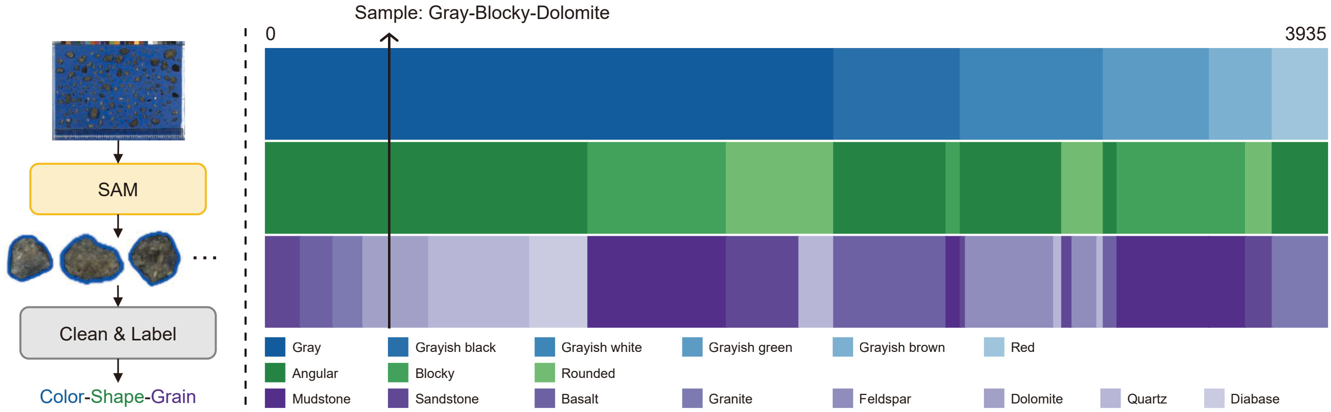

Fig. 1. Dataset construction and distribution for multi-attribute drill cuttings identification. The left illustrates the data processing pipeline; the right presents the distribution of labeled samples.

using an 8:2 stratified strategy based on multi-attribute combinations. For instance, of the 513 "gray-blocky-mudstone" samples, 410 are allocated to training and 103 for testing. This results in 3148 training and 787 test samples. Stratification ensures balanced attribute distribution and enhances model generalization in multi-attribute identification.

2.2. Statistical analysis

The correlation among drill cuttings attributes underpins the multi-attribute collaborative identification in this study. To assess label relationships across color, shape, and grain, we conducted joint frequency and correlation analyses on the dataset. Specifically, we calculated pairwise co-occurrence frequencies and corresponding positive pointwise mutual information (PPMI) to quantify association strength. Fig. 2(a–c) display the marginal distributions of color, shape, and grain, Fig. 2(d) presents the joint co-occurrence frequencies among attribute categories, and Fig. 2(e) shows the PPMI values that highlight significant positive associations.

The joint frequency and PPMI matrices (Fig. 2(d) and (e)) reveal clear inter-attribute label associations. Frequent combinations such as "gray-angular" reflect expected co-occurrence under high marginal probabilities, while less frequent pairs like "grayish green-blocky" and "grayish white-feldspar" exhibit elevated PPMI values, indicating stronger-than-expected dependencies. These patterns suggest that both high- and low-frequency combinations convey meaningful statistical information. The diagonal elements demonstrate sufficient label coverage (Fig. 2(a–c)), with grain relatively balanced and color and shape showing a few dominant classes. Overall, the observed correlations among attributes confirm that categories are not independent, supporting the effectiveness of learning joint representations and motivating the proposed multi-attribute collaborative identification framework.

3. Method

3.1. Overview

The CpCe model proposed in this study enables collaborative identification of multiple attributes in drill cuttings by integrating color, shape, and grain recognition within a unified framework. Although color can be extracted accurately through traditional color-space methods, it is also formulated as a learnable task so that color features can participate in shared representation learning and contribute cross-task regularization, which is not achievable with fixed hand-crafted descriptors. Unlike prior studies focused on single-attribute classification or purely data-driven cross-attribute modeling (Vandenhende et al., 2022; Ruder, 2017), CpCe introduces two key innovations. First, CpCe employs a correlation prior graph to represent domain-informed relationships among attributes and to provide structural constraints for the learning process, which improves model stability and interpretability. Second, it incorporates a collaborative inference mechanism that integrates prior knowledge with visual features and serves as the central component of the MTL framework by enabling explicit information exchange across tasks. These two components work together to combine data-driven learning with geological expertise and improve the robustness and consistency of multi-attribute identification.

Specifically, CpCe consists of three main components: an image encoder, a prior encoder, and a collaborative classifier, as illustrated in Fig. 3. The image encoder extracts high-dimensional feature representations from drill cuttings images using a standard convolutional backbone. The prior encoder uses stacked modified GAT (Veličković et al., 2018) layers to learn structured correlations among attribute labels, and the resulting prior-aware features provide guidance for multi-attribute identification. Finally, the image and prior features are fused through the collaborative classifier, which performs cross-task interaction based on the law of total probability to produce unified multi-attribute inference within the MTL framework.

This architecture adopts a parallel design in which the image encoder and prior encoder operate independently before fusion. Such separation allows visual features and structured prior knowledge to be learned without mutual interference, preserving the interpretability of the prior and preventing it from being implicitly shaped by image content. Meanwhile, deferring fusion to the collaborative classifier enables more stable cross-task interaction within the MTL framework and improves robustness in complex geological scenarios.

3.2. Correlation prior

In this study, a graph-based structure is employed to model correlation priors among attribute labels. The prior graph is constructed using label co-occurrence statistics from the dataset (Zhu et al., 2023; Wang et al., 2020; Chen et al., 2019), a widely adopted way that provides a direct and interpretable measure of inter-attribute correlations.

Formally, the prior graph is defined as $G = (V, E)$, where V denotes the set of nodes corresponding to all attribute categories,

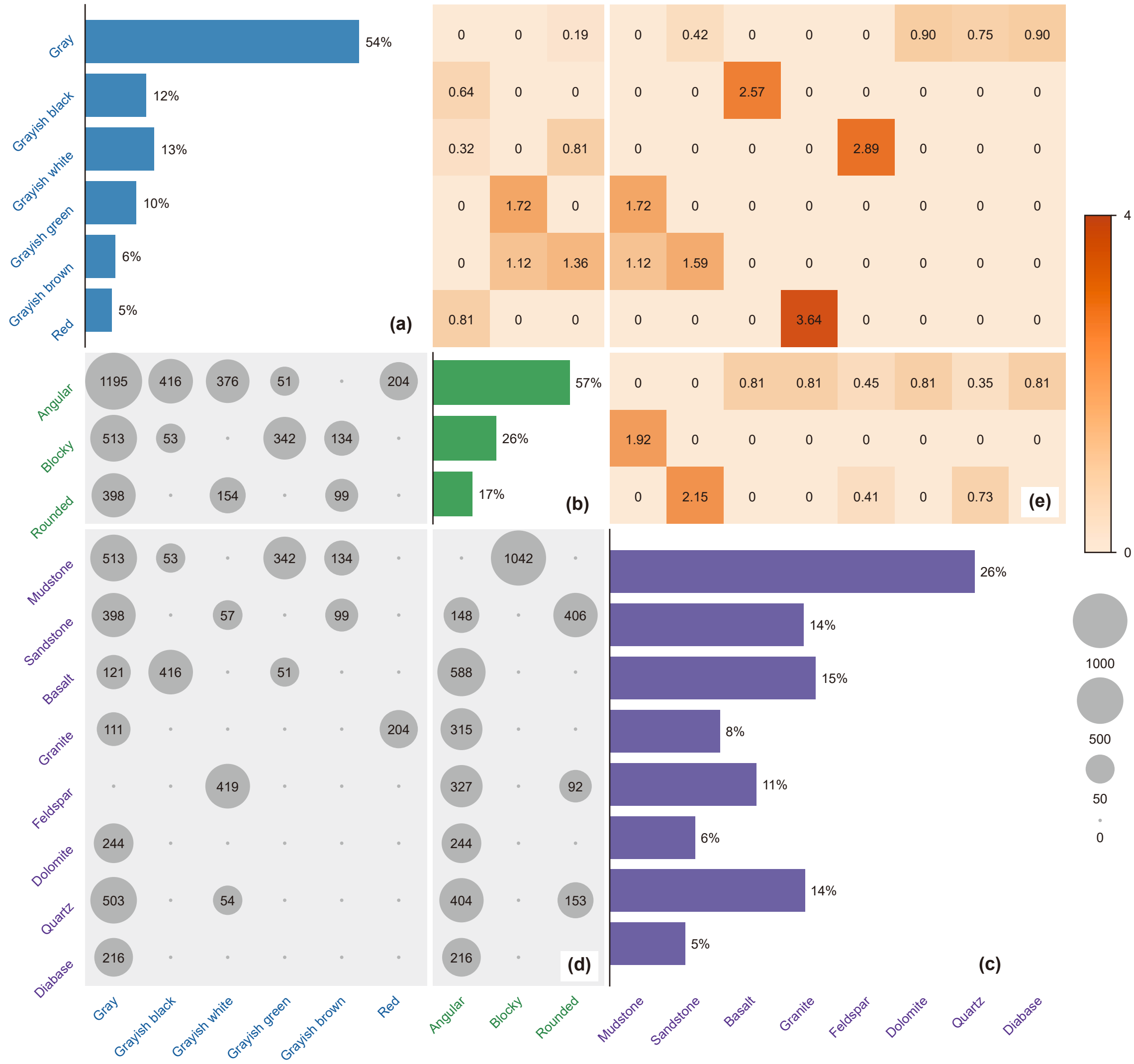

Fig. 2. Statistical analysis of the multi-attribute drill cuttings dataset. (**a**–**c**) show the marginal distributions of color, shape, and grain; (**d**) presents joint frequencies among attribute categories; and (**e**) displays the corresponding PPMI values.

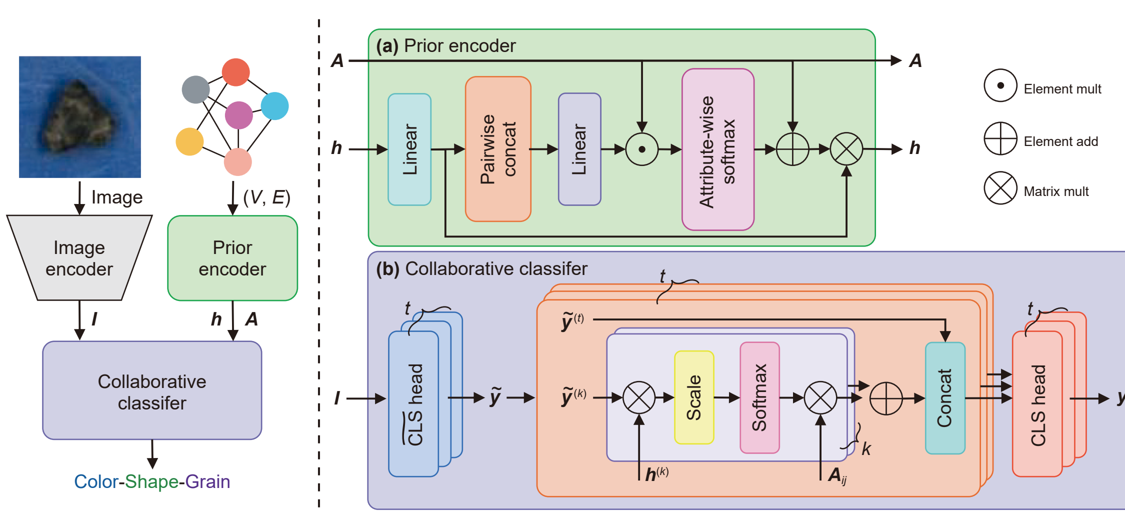

Fig. 3. Overview of the CpCe model. The left illustrates the overall architecture; (**a**) shows the prior encoder design, and (**b**) shows the collaborative classifier design.

and E represents the edges capturing their correlations. Each node is initialized using one-hot encoding, and all node features are stacked to form the node feature matrix $\boldsymbol{h} \in \mathbb{R}^{C \times C}$, where $C = 17$ represents the total number of labels across color, shape, and grain. To construct the edge set E, we compute a correlation matrix $\boldsymbol{A} \in \mathbb{R}^{C \times C}$ using co-occurrence statistics. The co-occurrence matrix $\boldsymbol{M} \in \mathbb{R}^{C \times C}$ and label frequency vector $\boldsymbol{N} \in \mathbb{R}^{C}$ are accumulated across all samples, and the conditional probability matrix $\boldsymbol{P}$ is obtained by normalizing each row of $\boldsymbol{M}$ with the corresponding element in $\boldsymbol{N}$. Then, adding the identity matrix retains self-correlation and produces the final correlation matrix used in our model.

The resulting graph G is an asymmetric directed graph, as shown in Fig. 4, where each edge $\boldsymbol{A}_{ij}$ encodes the conditional probability of observing label j given label i. This structured dependency graph provides inductive bias that helps the model capture co-occurrence among color, shape, and grain, thereby improving both performance and interpretability in multi-attribute identification.

Although the prior graph is constructed from statistical co-occurrence, these correlations often reflect meaningful geological tendencies, as color, shape, and grain of drill cuttings are jointly controlled by mineral composition, sedimentary environment, and diagenetic processes. For example, light-colored particles commonly derive from quartz- or feldspar-rich rocks or mature sandstone, whereas dark or gray-black grains often originate from mafic igneous rocks or organic-rich mudstone formed under reducing conditions. Angular grains are typically associated with crystalline igneous rocks such as basalt and granite, reflecting limited transport, while rounded grains generally indicate more durable minerals or longer transport histories, as observed in mature sandstone. These geologically interpretable tendencies explain why certain label pairs show strong statistical correlations in the correlation matrix $\boldsymbol{A}$. At the same time, because this prior is estimated purely from co-occurrence statistics, it inherits the inherent limitations of statistically derived priors. Statistical correlation does not necessarily imply geological causation, and the estimated dependencies may be influenced by dataset-specific factors such as sampling distribution, stratigraphic coverage, and labeling conventions. Therefore, strong co-occurrence relationships may reflect local or basin-specific settings rather than universally valid geological rules. For this reason, the correlation structure in our framework is explicitly treated as a probabilistic soft prior, providing inductive bias without being enforced as a deterministic constraint.

3.3. Prior encoder

The GAT provides a flexible mechanism for modeling node relationships through learnable, edge-specific attention weights. However, the original GAT does not explicitly address the attribute heterogeneity present in multi-attribute drill cuttings data. It aggregates all neighbors using a single global normalization, which may blur distinctions among attribute groups and introduce cross-attribute interference. To address this issue, we adopt an attribute-grouped attention mechanism that normalizes attention weights within each attribute group, enabling more faithful modeling of structured inter-attribute relationships.

Moreover, the original design computes attention weights solely from node features and does not utilize the varying edge strengths of attribute associations encoded in the prior graph. To incorporate this structural information, we modulate attention coefficients with the correlation matrix $\boldsymbol{A}$, enabling the model to distinguish different levels of inter-attribute correlation.

Finally, the prior graph is estimated from limited co-occurrence statistics. As a result, some statistically strong correlations may be noisy or inconsistent with established geological knowledge. If a fixed $\boldsymbol{A}$ were used directly, such correlations could be over-amplified and lead to spurious dependencies. To mitigate these potential conflicts, we introduce a learnable refinement scheme that updates $\boldsymbol{A}$ during training, allowing the structured prior to gradually adapt toward relations supported by the input evidence and reducing sensitivity to dataset-specific co-occurrence noise.

Specifically, the node feature matrix $\boldsymbol{h} \in \mathbb{R}^{C \times C}$ and the correlation matrix $\boldsymbol{A} \in \mathbb{R}^{C \times C}$, both derived from the structured correlation prior graph, serve as inputs to the prior encoder. The attention similarity matrix $\boldsymbol{e} \in \mathbb{R}^{C \times C}$ is then computed to estimate attention coefficients between node pairs.

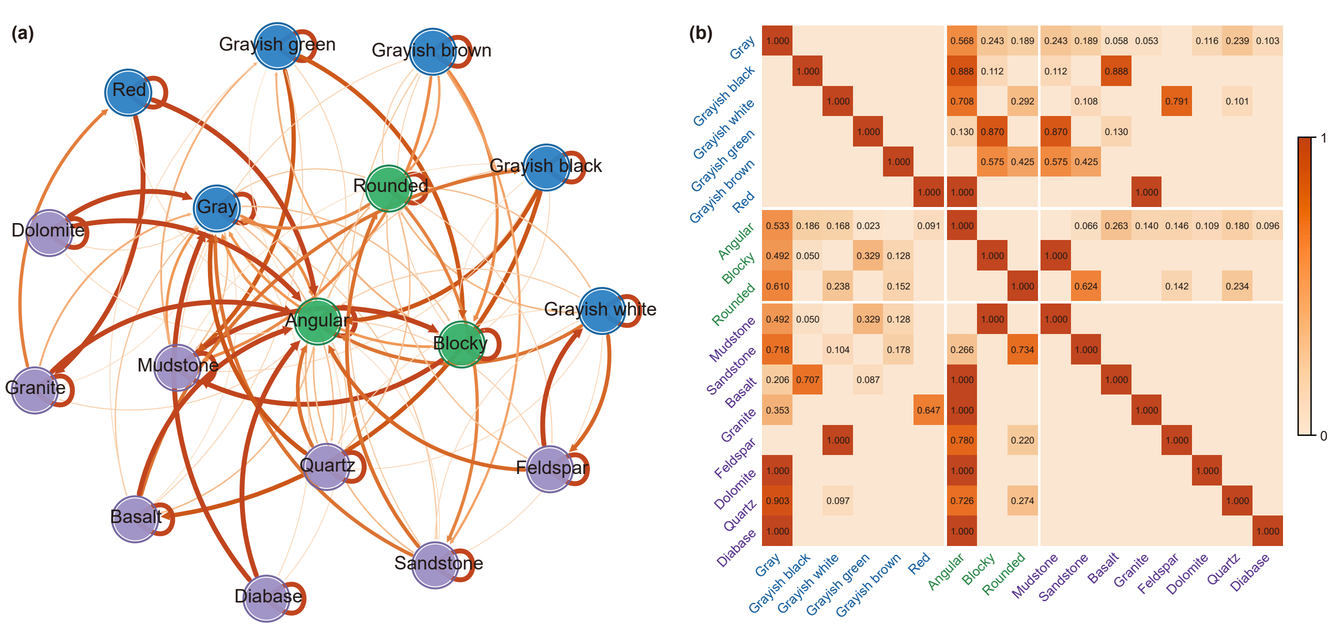

Fig. 4. Visualization of the correlation prior graph G. (**a**) Graph structure view; (**b**) matrix representation, with Y-axis as source and X-axis as target (source → target).

$$\boldsymbol{e}_{ij} = LeakyReLU\left(\boldsymbol{\alpha}^{\mathrm{T}}\left[\mathscr{W}\boldsymbol{h}_i \parallel \mathscr{W}\boldsymbol{h}_j\right]\right)\cdot\boldsymbol{A}_{ij} \tag{1}$$

Here, $\mathscr{W} \in \mathbb{R}^{\mathrm{F}\times\mathrm{F}'}$ is a learnable linear transformation and $\boldsymbol{\alpha} \in \mathbb{R}^{2\mathrm{F}\times 1}$ projects the concatenated features into scalar similarity scores. In the first layer, $F' = C$, corresponding to the total number of attribute labels. The concatenation operator $[\cdot \parallel \cdot]$ fuses the transformed features of node i and its neighbor j. The *LeakyReLU* activation introduces non-linearity, and $\boldsymbol{A}$ modulates $\boldsymbol{e}_{ij}$ to embed prior-defined edge strengths into the attention computation.

Next, we normalize the attention coefficients to obtain the final attention score matrix $\boldsymbol{a} \in \mathbb{R}^{\mathrm{C}\times\mathrm{C}}$, which governs the message aggregation.

$$\boldsymbol{a}_{ij}^{(t)} = \frac{\exp\left(\boldsymbol{e}_{ij}\right)}{\sum_{k\in\mathscr{N}_i^{(t)}} \exp(\boldsymbol{e}_{ik})}, j \in \mathscr{N}_i^{(t)} \tag{2}$$

Here, $\mathscr{N}_i^{(t)}$ denotes the neighbors of node i that belong to attribute type t, under the condition that node i itself does not belong to type t. To preserve attribute semantics, attention normalization is performed independently within each attribute group, producing an attribute-aware block structure that enhances the modeling of inter-attribute dependencies. In addition, since labels within the same attribute are mutually exclusive, intra-attribute attention is restricted to a unit diagonal, preventing invalid information flow and maintaining the structural integrity of the multi-attribute graph.

The normalized attention scores guide the aggregation of node features from their attribute-specific neighbors. The aggregated features are then passed through an *ELU* activation. To further enhance structural awareness, we add a prior-guided residual connection using $\boldsymbol{A}$ to reinforce structural correlations that may not be fully captured by attention alone. The final output $\boldsymbol{h}' \in \mathbb{R}^{\mathrm{C}\times\mathrm{F}}$ thus integrates both learned attention patterns and prior structural information.

$$\boldsymbol{h}' = ELU(\boldsymbol{a}\mathscr{W}\boldsymbol{h} + \boldsymbol{A}\mathscr{W}\boldsymbol{h}) \tag{3}$$

To mitigate bias introduced during the construction of $\boldsymbol{A}$, a learnable parameter β (initialized to 0.1) is used to adaptively update $\boldsymbol{A}$ using the learned attention scores. This refinement allows $\boldsymbol{A}$ to gradually align with the relational patterns discovered during training, thereby improving the robustness of the structured prior.

$$\boldsymbol{A} = (1-\beta)\boldsymbol{A} + \beta\boldsymbol{a} \tag{4}$$

3.4. Collaborative classifier

After processing the drill cuttings image through the image encoder, the resulting high-dimensional feature is globally aggregated via average pooling to obtain the image representation $\boldsymbol{I} \in \mathbb{R}^{\mathrm{F}}$ (e.g., $F = 512$ for ResNet-18 (He et al., 2016)). This image encoder is shared across all tasks in CpCe, forming the hard parameter sharing backbone that provides a unified visual representation for all tasks. The prior encoder and correlation prior are also shared, while each attribute is equipped with task-specific prediction heads built upon this shared representation. The attribute label feature matrix $\boldsymbol{h} \in \mathbb{R}^{\mathrm{C}\times\mathrm{F}}$, together with the attribute correlation matrix $\boldsymbol{A} \in \mathbb{R}^{\mathrm{C}\times\mathrm{C}}$, are then jointly fed into the final decision module. Here, $\boldsymbol{A}$ encodes probabilistic relationships among attribute labels and serves as a structural prior for multi-attribute identification of drill cuttings.

We first define an initial prediction $\tilde{\boldsymbol{y}}^{(t)} \in \mathbb{R}^{\mathrm{F}}$, which serves as guidance for each attribute t. This preliminary prediction is obtained by passing the global image feature $\boldsymbol{I}$ through an attribute-specific pre-classifier, capturing the model's initial semantic understanding of the drill cuttings image.

$$\tilde{\boldsymbol{y}}^{(t)} = \widetilde{CLS}^{(t)}(\boldsymbol{I}) \tag{5}$$

Each pre-classifier $\widetilde{CLS}$ consists of two linear layers, with *LayerNorm* and *ReLU* applied in between.

Next, the initial prediction $\tilde{\boldsymbol{y}}^{(t)}$, together with the label feature matrix $\boldsymbol{h}$ and the correlation matrix $\boldsymbol{A}$, are used for collaborative fusion to enhance final attribute identification. This process allows the tasks to influence one another through probabilistic conditioning, providing a flexible interaction mechanism on top of the shared backbone. Specifically, we adopt a total probability formulation, in which the label distribution $\boldsymbol{P}(t) \in \mathbb{R}^{C_t}$ for attribute t (with C_t denoting the number of categories) is computed as the sum of conditional probabilities given all other related attributes.

$$\boldsymbol{P}(t) = \sum_{k\in\mathscr{N}^{(t)}} \boldsymbol{P}(k)\cdot\boldsymbol{P}(t|k) \tag{6}$$

Here, $\mathscr{N}^{(t)}$ denotes the set of attributes excluding t. For a related attribute k, the prior probability $\boldsymbol{P}(k) \in \mathbb{R}^{C_k}$ is calculated by computing the similarity between the initial prediction $\tilde{\boldsymbol{y}}^{(k)} \in \mathbb{R}^{\mathrm{F}}$ and the label feature matrix $\boldsymbol{h}^{(k)} \in \mathbb{R}^{C_k\times\mathrm{F}}$. The matrix $\boldsymbol{A}_{ij}$ encodes the conditional probability $\boldsymbol{P}(t|k) \in \mathbb{R}^{C_k\times C_t}$, where i and j index the labels of attributes k and t, respectively.

Inspired by the law of total probability, we adopt the attention mechanism (Vaswani et al., 2017) to construct the collaborative classifier. Specifically, for each attribute t, the label distribution $\boldsymbol{P}(t)$ is computed by treating $\tilde{\boldsymbol{y}}^{(k)}$ as the query (Q), $\boldsymbol{h}^{(k)}$ as the key (K), and $\boldsymbol{A}_{ij}$ as the value (V), where k denotes attributes other than t.

$$\boldsymbol{P}(t) = \sum_{k\in\mathscr{N}^{(t)}} softmax\left(\frac{\tilde{\boldsymbol{y}}^{(k)}\boldsymbol{h}^{(k)\mathrm{T}}}{\sqrt{\mathrm{F}}}\right)\boldsymbol{A}_{ij} \tag{7}$$

To improve classification performance, we fuse the image features with the collaborative fusion features. Specifically, for each attribute t, $\boldsymbol{P}(t)$ is concatenated with the initial prediction $\tilde{\boldsymbol{y}}^{(t)}$ and passed to the final classification layer for prediction. These final classifiers are task-specific, allowing each attribute to refine its own prediction while still benefiting from shared representations and inter-attribute information provided by the collaborative module.

$$\boldsymbol{y}^{(t)} = CLS^{(t)}\left(\left[\tilde{\boldsymbol{y}}^{(t)} \parallel \boldsymbol{P}(t)\right]\right) \tag{8}$$

The classifier *CLS* shares the same structure as the pre-classifier $\widetilde{CLS}$, consisting of two linear layers with *LayerNorm* and *ReLU* between them. The pre-classifier is used to generate intermediate representations for guidance, while the final classifier outputs the probability distribution over the classes of each attribute. This separation of functional roles between guidance and classification enhances both the accuracy and robustness of multi-attribute identification. Overall, the collaborative classifier operates on top of a hard parameter sharing backbone and, through probabilistic conditioning, provides flexible task-to-task interactions. Together with the task-specific prediction heads, this establishes a complete MTL framework that integrates shared representations, attribute-specific decision, and prior-aware collaborative fusion.

3.5. Experimental settings

The CpCe model was implemented in PyTorch. ResNet-18 (without the classification head) was used as the image encoder and initialized with official pre-trained weights to accelerate convergence. The prior encoder consisted of a three-layer stack of the modified GAT. During both training and inference, the model takes the drill cuttings image together with the fixed structured prior graph as inputs. The prior graph is constructed once from the training set and remains unchanged as an input, whereas the image varies across samples. Although the correlations associated with the prior graph are refined during training via a learnable mechanism, this refinement is part of the learned model parameters; thus, the pre-constructed prior graph remains an input to the prior encoder during inference.

Training was conducted with a batch size of 128 for 50 epochs using Focal Loss for each attribute, and the overall loss was defined as an unweighted sum of the three task losses. Optimization employed the AdamW optimizer with a weight decay of 1×10^{-4}, momentum of 0.9, an initial learning rate of 5×10^{-4}, and a poly decay schedule. As multi-attribute identification is essentially a classification task, Accuracy (Acc) and F1-score (F1) were used as the primary evaluation metrics.

4. Results and discussion

4.1. Performance of multi-attribute identification of drill cuttings

After training CpCe on the training set, we evaluated its performance on the test set. To comprehensively assess its effectiveness, we compared it against several representative MTL methods widely used in computer vision, including MTAN (Liu et al., 2019), NDDR-CNN (Gao et al., 2019) and cross-stitch (Misra et al., 2016). Additionally, we implemented an MTL-baseline model featuring a shared backbone with separate task-specific prediction heads. For fairness, all models were trained and evaluated using identical training protocols and hyperparameter settings. The multi-attribute identification results on the test set are summarized in Table 1.

The experimental results demonstrate that CpCe achieves the highest performance across all three attribute recognition tasks, substantially outperforming the comparison models in both Acc and F1. Specifically, CpCe attains Acc values of 92.12%, 94.41%, and 95.17% for the color, shape, and grain tasks, respectively, with corresponding F1 of 91.24%, 92.88%, and 95.13%. The overall mean Acc and mean F1 reach 93.90% and 93.08%, respectively, significantly exceeding those of the baseline models. These results indicate that CpCe exhibits strong task-level collaboration and effectively enhances multi-attribute identification performance.

We further evaluated the class-specific F1-scores for each attribute, as shown in Fig. 5. The results reveal clear differences in class-level stability among models. CpCe exhibits the most regular and balanced patterns across all tasks, indicating superior consistency across classes and stronger adaptability. For example, performance of comparison models drops markedly for grayish green, grayish brown, and rounded. The first two are highly sensitive to illumination conditions and clay-rich components, leading to blurred and unstable color boundaries and high intra-class variability. Meanwhile, rounded shows large morphological variation driven by transport distance, reworking intensity, and fragmentation processes. Under these conditions, CpCe maintains more consistent performance than other models, reflecting the effect of the proposed structured prior-guided collaborative modeling strategy. Additionally, although CpCe shows slightly lower performance on quartz, a category characterized by broad variation in transparency, size, and morphology, this localized decline does not affect the overall prediction pattern, as CpCe leverages cross-attribute co-occurrence to mitigate intra-class heterogeneity across tasks.

4.2. Analysis of attribute collaboration

To evaluate the effectiveness of collaborative multi-attribute learning, this study leverages the intrinsic correlations among drill cuttings attributes to perform joint identification, which substantially enhances overall performance compared with STL models. To further verify this advantage, we designed a set of STL experiments, where separate models were independently trained for color, shape, and grain. All STL models adopted the same backbone network (ResNet-18) and training strategy as CpCe, ensuring fairness and consistent comparisons. The corresponding results are reported in the STL section of Table 1.

As shown in Table 1, although the STL model achieves reasonable performance on individual tasks, such as 93.39% Acc and 93.31% F1 in grain classification, its overall effectiveness is limited by the absence of inter-task information sharing. This independent training strategy prevents the model from capturing potential correlations among attributes. The limitation is particularly evident in color and shape classification, where STL attains lower Acc (87.80% and 87.17%) and F1 (85.58% and 83.34%) compared with the MTL-based CpCe. These performance gaps are especially critical in geological analysis, where attributes such as color, shape, and grain exhibit strong intrinsic associations. In contrast, MTL models like CpCe benefit from shared representations and joint optimization across tasks, resulting in consistently higher accuracy and stability. These findings highlight the key advantage of MTL in enhancing generalization and robustness by leveraging inter-attribute dependencies and shared semantic representations (Vandenhende et al., 2022; Ruder, 2017).

To further evaluate model behavior, confusion matrices for both STL and CpCe were visualized in Fig. 6. These results intuitively support the earlier performance comparisons. In Fig. 6, the STL

Table 1
Performance of each model on the test set after training.

	Model	Color		Shape		Grain		Mean	
		Acc	F1	Acc	F1	Acc	F1	Acc	F1
STL	ResNet-18	87.80	85.58	-	-	-	-	89.45	87.41
		-	-	87.17	83.34	-	-		
		-	-	-	-	93.39	93.31		
MTL	Cross-stitch	88.06	86.52	89.45	87.41	94.28	94.14	90.60	89.36
	NDDR-CNN	88.44	87.09	91.49	89.09	93.90	93.90	91.27	90.02
	MTAN	88.95	87.47	92.88	90.98	94.28	94.03	92.04	90.83
	MTL-baseline	90.60	89.60	91.23	89.15	93.65	93.23	91.83	90.66
	CpCe	92.12	91.24	94.41	92.88	95.17	95.13	93.90	93.08

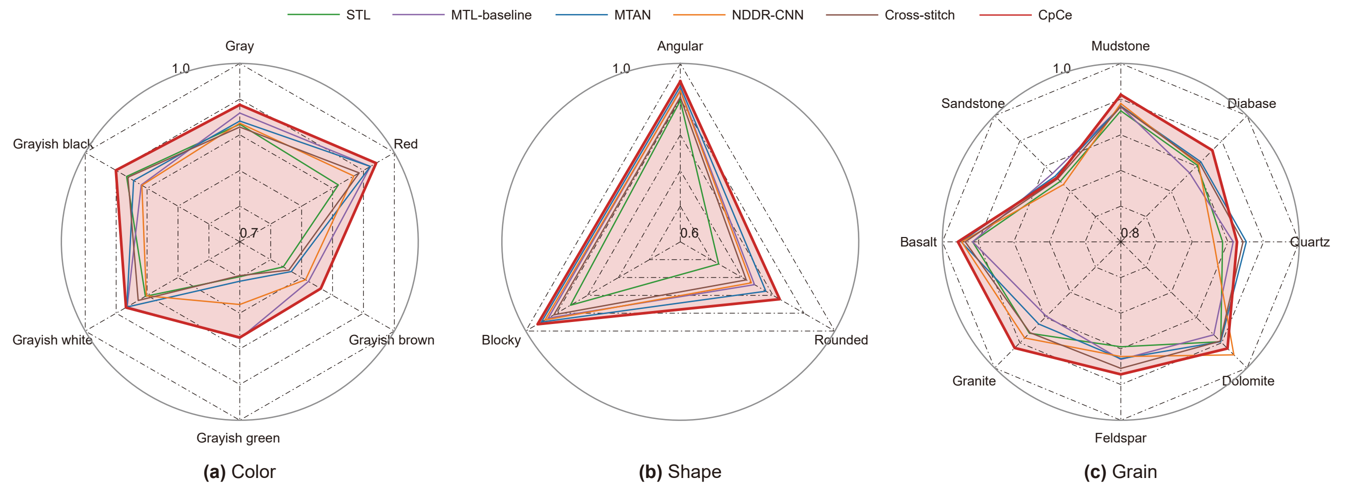

Fig. 5. Class-wise F1-scores for each model on (**a**) color, (**b**) shape, and (**c**) grain.

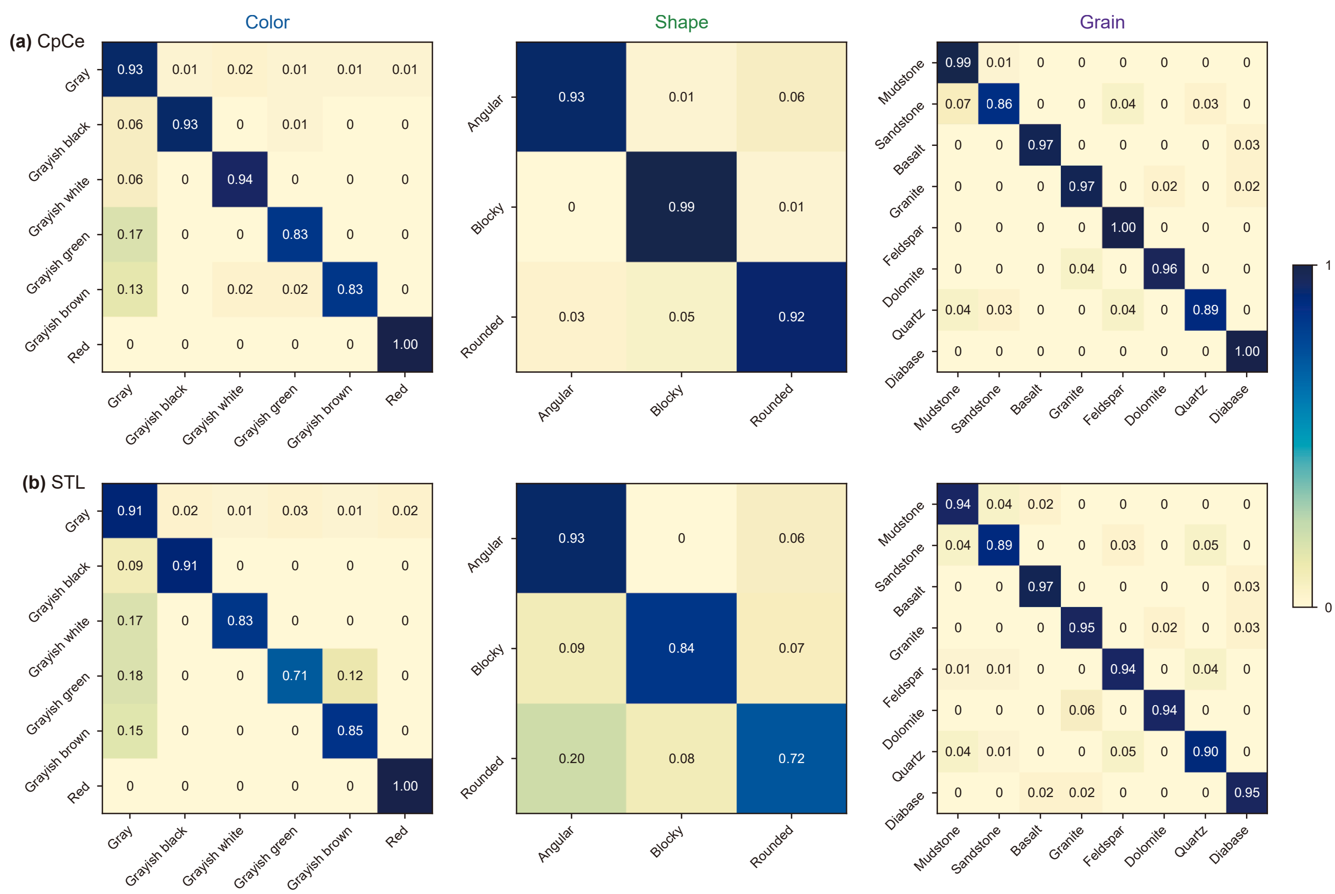

Fig. 6. Confusion matrices of CpCe and STL models for each attribute on the test set.

models show noticeable misclassifications across all tasks, particularly in color and shape. The confusion between grayish white and gray as well as angular and rounded are most prominent. These errors mainly stem from the visual similarity of certain cuttings under drilling-fluid wetting, illumination changes, or blurred fracture edges, together with inherent subjectivity in manual labeling. Moreover, imbalanced class distributions make STL more sensitive to local data biases, reducing its stability on fine-grained distinctions. In contrast, CpCe mitigates these limitations by leveraging cross-attribute correlations. Grayish white typically appears in quartz- or feldspar-rich, or compositionally mature sandstone fragments, whereas gray more often occurs in mudstone or mafic igneous rocks such as basalt and diabase. Similarly, rounded particles are commonly associated with well-transported, texturally mature sandstones, while angular particles frequently originate from near-source, immature fragments including mafic igneous rocks. By integrating such attribute co-occurrence patterns, CpCe can utilize complementary geological cues when single-attribute visual information is ambiguous, effectively reducing the misclassifications observed in STL.

4.3. Analysis of prior encoders

To evaluate the effect of the introduced enhancements in the GAT module of the prior encoder, we conducted ablation experiments targeting three core components: attribute-grouped attention (AttrA), edge structure integration (EdgeI), and correlation matrix refinement (CorR). Each component was independently removed to assess its individual contribution to model performance. The results are summarized in Table 2.

The results indicate that the baseline employing the original GAT achieves a mean Acc of 92.29% and a mean F1 of 90.95% across the three tasks. When AttrA is introduced, the mean F1 increases to 92.13%, demonstrating that partitioning neighbors by attribute type enhances the model's ability to capture inter-attribute dependencies. The integration of EdgeI also yields notable improvements, raising the mean F1 to 91.14%, suggesting that incorporating edge strengths helps encode topological semantics. Furthermore, the addition of CorR leads to marked gains, particularly in the shape and grain tasks, increasing the overall mean F1 to 91.43%. This highlights the limitations of using a fixed correlation matrix and supports the benefit of adaptive refinement during training to improve flexibility and generalization. Ultimately, the full CpCe model achieves the best overall performance, with a mean Acc of 93.90% and a mean F1 of 93.08%, confirming the complementary contributions of all three components to multi-attribute identification accuracy and robustness.

4.4. Analysis of refined correlation matrix

To illustrate the effect of the correlation refinement mechanism, we visualized the updated correlation matrix $\boldsymbol{A}$ after training, as shown in Fig. 7.

The refined correlation matrix reduces the dominance of prior single-cue associations by reallocating weights to more accurately reflect the intrinsic characteristics of drill cuttings, indicating a shift toward the integration of multiple geological signals. Several representative weight changes and their geological implications illustrate this improvement. For example, grayish-black cuttings are no longer strongly linked to basalt. The decrease in “grayish black→basalt” and the increase in “grayish black→mudstone” capture the common sources of dark-colored cuttings, which include the intrinsic dark tone of basalt as a mafic igneous rock and the dark appearance produced by organic matter enrichment and clay-rich matrices in fine-grained mudstones. Similarly, the strengthened “quartz→grayish white” and weakened “quartz→gray” association align with the light gray to whitish appearance typically exhibited by quartz-rich cuttings. For morphology, the increase in “gray→blocky” and the decrease in “gray→angular” indicate that the model reduces its reliance on angular fragments when linking gray cuttings to fragment morphology, while progressively considering blocky fragments as a plausible product of brittle fragmentation in gray rocks. Furthermore, the adjusted roundness patterns for sandstone, where the reduced gap between angular and rounded weights better reflects the natural range of roundness associated with differences in textural maturity, mineral composition, and transport history. Finally, the decline in “grayish white→feldspar” and the redistribution of weights toward other grains show that whitish cuttings may originate not only from feldspar but also from quartz or quartz-rich mature sandstones, thereby mitigating the overly deterministic mappings imposed by the initial prior.

Overall, the refined matrix exhibits a clear depolarization from rigid single-feature associations toward multi-factor geological explanations. Probabilities are redistributed across several geologically plausible pathways rather than concentrated in single dominant links. This shift suggests that the model enhances robustness and generalization in settings characterized by mixed sources and ambiguous feature expression.

4.5. Analysis of model cost

To assess the computational cost of CpCe, we evaluate parameter count, floating point of operations (FLOPs), and inference latency across different models. All inference tests are conducted with single-sample input under identical settings. Since the three STL variants (color/shape/grain) share almost identical computational profiles, we report the grain-based STL model as their representative. Table 3 summarizes the computational cost during inference.

Although CpCe introduces additional modules, namely the prior encoder and the collaborative classifier, its overall computational overhead remains moderate. Compared with the MTL-baseline, CpCe increases the number of parameters and FLOPs only slightly, indicating that both components are lightweight. In terms of inference speed, CpCe runs slower than the STL model and the MTL-baseline, which is expected given its cross-task reasoning and prior-guided fusion mechanism. However, its inference latency is still faster than that of other MTL frameworks, demonstrating that CpCe maintains competitive efficiency among MTL models. Overall, CpCe offers a balanced trade-off between accuracy and efficiency, making it feasible for practical field deployment.

4.6. Cross-well generalization evaluation

To further assess the transferability of CpCe, we collected an additional set of samples from wells located in the Shuangtaizi Structural Belt of the Western Sag and the Jinganpu Structural Belt of the Damintun Sag within the Liaohe Depression of the Bohai Bay Basin. These wells are spatially independent from those used for training and were not included in the original dataset, providing

Table 2
Ablation results for modified graph attention components in the prior encoder.

	AttrA	EdgeI	CorR	Color		Shape		Grain		Mean	
				Acc	F1	Acc	F1	Acc	F1	Acc	F1
CpCe	×	×	×	91.36	90.01	91.49	88.71	94.03	94.13	92.29	90.95
	✓			92.76	91.94	92.88	91.11	93.77	93.34	93.14	92.13
		✓		91.23	89.67	92.88	90.93	93.27	92.83	92.46	91.14
			✓	89.45	87.58	93.90	92.06	94.79	94.65	92.71	91.43
	✓	✓	✓	92.12	91.24	94.41	92.88	95.17	95.13	93.90	93.08

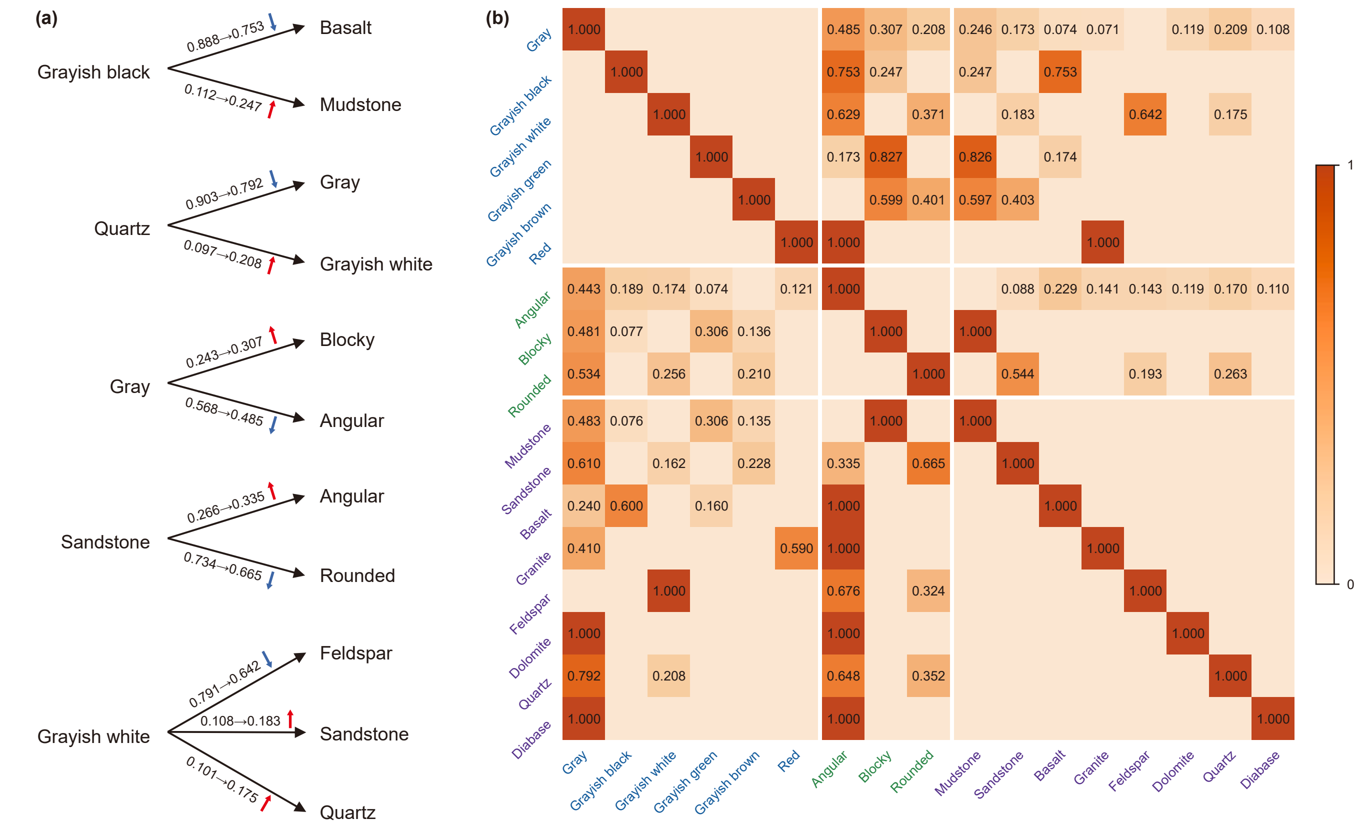

Fig. 7. Refined correlation matrix ***A*** after training. (**a**) Partial node correlations; (**b**) Full matrix, with *Y*-axis as source nodes and *X*-axis as target nodes (source → target).

Table 3
Computational cost of each model during inference.

	Model	Params, M	FLOPs, G	Latency, ms/sample
STL	ResNet-18	11.24	0.60	2.27
MTL	Cross-stitch	33.74	1.79	7.86
	NDDR-CNN	36.91	1.98	8.27
	MTAN	12.67	0.86	8.90
	MTL-baseline	11.38	0.60	2.55
	CpCe	12.31	0.62	5.98

unseen wells for cross-well evaluation under the same basin-scale geological framework. A total of 347 cutting samples were obtained, none of which were used during training. We applied the previously trained models to this new dataset and computed Acc and F1 to assess their cross-well generalization performance. The results for all models are summarized in Table 4.

As shown in Table 4, CpCe consistently outperforms the comparison models on the unseen wells. Although a certain performance drop is observed compared with the in-domain results, CpCe maintains relatively stable Acc and F1 across different attributes. This suggests that the model exhibits effective cross-well generalization within the same basin. The results indicate that, by modeling cross-attribute correlations and incorporating geological prior knowledge, CpCe is less sensitive to well-specific variations and can learn attribute representations that remain effective under unseen wells. Overall, the cross-well evaluation demonstrates the robustness and generalization capability of CpCe in practical engineering scenarios. Extending the proposed approach to a broader range of geological settings is expected to involve additional challenges and will be investigated in future work.

5. Conclusion

This paper presents correlation prior and collaborative enhancement (CpCe), a multi-attribute identification model for

Table 4
Generalization performance of all models on unseen wells.

	Model	Color		Shape		Grain		Mean	
		Acc	F1	Acc	F1	Acc	F1	Acc	F1
STL	ResNet-18	73.49	60.90	-	-	-	-	76.08	68.70
		-	-	75.50	68.49	-	-		
		-	-	-	-	79.25	76.72		
MTL	Cross-stitch	75.50	66.42	78.96	75.91	77.81	76.73	77.43	73.02
	NDDR-CNN	75.79	64.60	76.66	71.99	78.67	78.85	77.04	71.81
	MTAN	78.39	68.21	76.66	72.28	78.67	76.33	77.91	72.28
	MTL-baseline	74.93	65.70	79.83	76.69	79.54	77.26	78.10	73.21
	CpCe	80.40	70.14	83.00	79.78	82.42	82.17	81.94	77.36

drill cuttings that leverages inter-attribute correlations as prior knowledge to enable joint modeling and collaborative classification. CpCe constructs a correlation prior graph based on the co-occurrence patterns of attribute labels and incorporates two core modules: a prior encoder and a collaborative classifier. The prior encoder employs a modified GAT to capture structured attribute dependencies, while the classifier integrates image and graph features within a total probability framework to facilitate cross-attribute information exchange and enhance identification accuracy. A comprehensive statistical analysis of the drill cuttings dataset confirms the existence of meaningful inter-attribute associations, providing support for collaborative learning. Experimental results demonstrate that CpCe consistently outperforms representative MTL and STL models in terms of accuracy, efficiency, and generalization capability. Overall, CpCe provides a robust and generalizable framework for multi-attribute drill cuttings analysis, with potential applicability to other geoscientific tasks involving structured relational knowledge.

CRediT authorship contribution statement

Li Hou: Writing – original draft, Visualization, Validation, Methodology, Investigation, Data curation. **Xin Li:** Writing – review & editing, Supervision. **Wei-Zhi Tian:** Writing – review & editing, Data curation. **Qian-Xiao Su:** Writing – review & editing, Investigation, Formal analysis. **Rui-Qi Huang:** Writing – review & editing. **Ming-Hui Song:** Writing – review & editing. **Peng Du:** Writing – review & editing. **Yi-Li Ren:** Writing – review & editing, Supervision, Conceptualization.

Data availability

The data that has been used is confidential.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was financially supported by the National Natural Science Foundation of China (No. 42372175), National Science and Technology Major Project-Key Equipment and Software for Smart Oil and Gas Fields (No. 2025ZD1401501), China National Petroleum Corporation Technology Project (No. 2023DJ84).

References

Alixant, J., Collie, D.A.L., 1999. Spectral gamma ray measurement on cuttings: An alternative to MWD. In: SPE/IADC Middle East Drilling Technology Conference. https://doi.org/10.2118/57550-MS.

Carugo, C., Malossi, A., Balossino, P., et al., 2013. Advanced cuttings analysis improves reservoir characterisation and reduces operating times in shale gas drilling project. In: International Petroleum Technology Conference. https://doi.org/10.2523/IPTC-17186-MS.

Chen, Z., Wei, X., Wang, P., et al., 2019. Multi-label image recognition with graph convolutional networks. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 5177–5186. https://doi.org/10.1109/CVPR.2019.00532.

Costa, L.C., Carvalho, C.F., Soares, A.S.F., et al., 2023. Physical and chemical characterization of drill cuttings: A review. Mar. Pollut. Bull. 194, 115342. https://doi.org/10.1016/j.marpolbul.2023.115342.

Gao, Y., Ma, J., Zhao, M., et al., 2019. NDDR-CNN: Layerwise feature fusing in multi-task cnns by neural discriminative dimensionality reduction. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. CVPR, pp. 3200–3209. https://doi.org/10.1109/CVPR.2019.00332.

Guilherme, I.R., Marana, A.N., Papa, J.P., et al., 2011. Petroleum well drilling monitoring through cutting image analysis and artificial intelligence techniques. Eng. Appl. Artif. Intell. 24, 201–207. https://doi.org/10.1016/j.engappai.2010.04.002.

Guo, L., Qian, J., Si, J., et al., 2020. Image detection method of drill cuttings. J. Geophys. Eng. 17, 249–257. https://doi.org/10.1093/jge/gxz103.

He, K., Zhang, X., Ren, S., et al., 2016. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778. https://doi.org/10.1109/CVPR.2016.90.

Hogg, A.J.C., Mitchell, A.W., Young, S., 1996. Predicting well productivity from grain size analysis and logging while drilling. Pet. Geosci. 2 (2.1.1), 1–15. https://doi.org/10.1144/petgeo.

Huo, F., Li, A., Zhao, X., et al., 2021. Novel lithology identification method for drilling cuttings under PDC bit condition. J. Pet. Sci. Eng. 205, 108898. https://doi.org/10.1016/j.petrol.2021.108898.

Huo, F., Liu, K., Dong, H., et al., 2024. Research on cuttings image segmentation method based on improved MultiRes-Unet++ with attention mechanism. Signal, Image Video Process 18, 799–808. https://doi.org/10.1007/s11760-024-03192-3.

Kadyrov, R., Nurgaliev, D., Saenger, E.H., et al., 2022. Digital rock physics: Defining the reservoir properties on drill cuttings. J. Pet. Sci. Eng. 210, 110063. https://doi.org/10.1016/j.petrol.2021.110063.

Kathrada, M., Adillah, B.J., 2019. Visual recognition of drill cuttings lithologies using convolutional neural networks to aid reservoir characterisation. In: SPE Reservoir Characterisation and Simulation Conference and Exhibition, pp. 17–19. https://doi.org/10.2118/196675-MS.

Kirillov, A., Mintun, E., Ravi, N., et al., 2023. Segment anything. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 3992–4003. https://doi.org/10.1109/ICCV51070.2023.00371.

Liu, S., Johns, E., Davison, A.J., 2019. End-to-end multi-task learning with attention. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1871–1880. https://doi.org/10.1109/CVPR.2019.00197.

Marsala, A.F., Loermans, T., Shen, S., et al., 2011. Real-time mineralogy, lithology, and chemostratigraphy while drilling using portable energy-dispersive X-ray fluorescence. In: SPE EUROPEC/EAGE Annual Conference and Exhibition. https://doi.org/10.2118/143468-MS.

Misra, I., Shrivastava, A., Gupta, A., et al., 2016. Cross-stitch networks for multi-task learning. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3994–4003. https://doi.org/10.1109/CVPR.2016.433.

Młynarczuk, M., Górszczyk, A., Ślipek, B., 2013. The application of pattern recognition in the automatic classification of microscopic rock images. Comput. Geosci. 60, 126–133. https://doi.org/10.1016/j.cageo.2013.07.015.

Patel, A.K., Chatterjee, S., 2016. Computer vision-based limestone rock-type classification using probabilistic neural network. Geosci. Front. 7, 53–60. https://doi.org/10.1016/j.gsf.2014.10.005.

Perez, C.A., Saravia, J.A., Navarro, C.F., et al., 2015. Rock lithological classification using multi-scale gabor features from sub-images, and voting with rock contour information. Int. J. Miner. Process. 144, 56–64. https://doi.org/10.1016/j.minpro.2015.09.015.

Ramba, V., Selvaraju, S., Subbiah, S., et al., 2021. Optimization of drilling parameters using improved play-back methodology. J. Pet. Sci. Eng. 206, 108991. https://doi.org/10.1016/j.petrol.2021.108991.

Reyes, R., Kyzym, I., Rana, P.S., et al., 2015. Cuttings analysis for rotary drilling penetration mechanisms and performance evaluation. In: 49th U.S. Rock Mechanics/Geomechanics Symposium. ARMA-2015-764.

Ruder, S., 2017. An overview of multi-task learning in deep neural networks. arXiv preprint arXiv:1706.05098. https://doi.org/10.48550/arXiv.1706.05098.

Salim, A., Lagraba P, J.O., 2018. Utilizing drill cuttings to enhance characterization and description of tight carbonate reservoirs. In: SPE Annual Technical Conference and Exhibition. https://doi.org/10.2118/191662-MS.

Sanei, H., Ardakani, O.H., Akai, T., et al., 2020. Core versus cuttings samples for geochemical and petrophysical analysis of unconventional reservoir rocks. Sci. Rep. 10, 7920. https://doi.org/10.1038/s41598-020-64936-y.

Shan, L., Liu, Y., Du, K., et al., 2024. Drilling rock image segmentation and analysis using segment anything model. Adv. Geo-Energy Res. 12, 89–101. https://doi.org/10.46690/ager.2024.05.02.

Singh, H., Li, C., Cheng, P., et al., 2023. Automated real-time formation evaluation from cuttings and drilling data analysis: state of the art. Adv. Geo-Energy Res. 8, 19–36. https://doi.org/10.46690/ager.2023.04.03.

Singh, N., Singh, T.N., Tiwary, A., et al., 2010. Textural identification of basaltic rock mass using image processing and neural network. Comput. Geosci. 14, 301–310. https://doi.org/10.1007/s10596-009-9154-x.

Tolstaya, E., Shakirov, A., Mezghani, M., et al., 2023. Fast reservoir characterization with AI-based lithology prediction using drill cuttings images and noisy labels. Journal of Imaging 9 (7), 126. https://doi.org/10.3390/jimaging9070126.

Vandenhende, S., Georgoulis, S., Van Gansbeke, W., et al., 2022. Multi-task learning for dense prediction tasks: A survey. IEEE Trans. Pattern Anal. Mach. Intell. 44, 3614–3633. https://doi.org/10.1109/TPAMI.2021.3054719.

Vaswani, A., Shazeer, N., Parmar, N., et al., 2017. Attention is all you need. In: Advances in Neural Information Processing Systems.

Veličković, P., Cucurull, G., Casanova, A., et al., 2018. Graph attention networks. In: International Conference on Learning Representations.

Wang, Y., He, D., Li, F., et al., 2020. Multi-label classification with label graph superimposing. Proc. AAAI Conf. Artif. Intell. 34, 12265–12272. https://doi.org/10.1609/aaai.v34i07.6909.

Xu, P., Gan, C., Wang, L., et al., 2022. A multi-feature extraction-based image identification method for rock debris in the drilling process. In: 2022 China Automation Congress. CAC, pp. 6666–6671. https://doi.org/10.1109/CAC57257.2022.10054959.

Xu, Y., Zuo, R., Bai, Y., 2025. Geological knowledge-guided dual-branch deep learning model for identification of geochemical anomalies related to mineralization. J. Geophys. Res.: Machine Learning and Computation 2 (1), e2024JH000468. https://doi.org/10.1029/2024JH000468.

Yamada, T., Di Santo, S., 2022. Instance segmentation of piled rock particles based on mask R-CNN. In: 2022 IEEE International Geoscience and Remote Sensing Symposium, pp. 163–166. https://doi.org/10.1109/IGARSS46834.2022.9883183.

Yi, J., Yang, Z., Hu, Y., et al., 2024. Dc-Segnet: A semantic segmentation network for fluorescent detection of drilling cuttings. SSRN preprint. https://doi.org/10.2139/ssrn.4976028.

Yousef, I., Morozov, V.P., 2024. Exploring the potential of drill cuttings for reservoir characterization: A case study from the volga-ural basin, Russia. Petroleum Research 9, 193–205. https://doi.org/10.1016/j.ptlrs.2024.01.012.

Zhang, Y., Yuan, B., Wang, Y., et al., 2022. A real-time oil content analysis method of cuttings based on deep learning. IEEE Access 10, 132083–132094. https://doi.org/10.1109/ACCESS.2022.3229760.

Zhu, X., Liu, J., Liu, W., et al., 2023. Scene-aware label graph learning for multi-label image classification. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pp. 1473–1482. https://doi.org/10.1109/ICCV51070.2023.00142.

Zuo, R., Yang, F., Cheng, Q., et al., 2024. A novel data-knowledge dual-driven model coupling artificial intelligence with a mineral systems approach for mineral prospectivity mapping. Geology 53 (1), 284–288. https://doi.org/10.1130/G52970.